

Mechanism-Based Strategy for Optimizing HaloTag Protein Labeling

Sérgio M. Marques^{a,b†}, Michaela Slanska^{a†}, Klaudia Chmelova^{a,b}, Radka Chaloupkova^{a,c}, Martin Marek^{a,b}, Spencer Clark^d, Jiri Damborsky^{a,b}, Eric T. Kool^{d*}, David Bednar^{a*}, Zbynek Prokop^{a,b*}

^a*Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Brno, Czech Republic*

^b*International Clinical Research Center, St. Anne's University Hospital, Brno, Czech Republic*

^c*Enantis Ltd., Biotechnology Incubator INBIT, Brno, Czech Republic*

^d*Department of Chemistry, Stanford University, Stanford, California, USA*

[†]These authors contributed equally to this study.

*Corresponding authors: Eric T. Kool eric.kool@stanford.edu; David Bednar, 222755@mail.muni.cz; Zbynek Prokop, zbynek@chemi.muni.cz

Highlights

- A comprehensive kinetic and computational study provided detailed insights into the mechanism of HaloTag ligand incorporation.
- Both the ligand binding and the subsequent chemical step can be decisive for the efficiency of covalent ligand incorporation.
- Natural dehalogenases provide high efficiency in ligand incorporation with no need for intensive optimization using directed evolution.
- A simple strategy is proposed to identify the optimal tag from a wide range of natural dehalogenases.

Abstract

HaloTag labeling technology has introduced unrivaled potential in protein chemistry, molecular and cellular biology. A wide variety of ligands have been developed to meet the specific needs of diverse applications, but only a single protein tag, DhaAHT, is routinely used for their incorporation. Following a systematic kinetic and computational analysis of different reporters, tetramethylrhodamine and three 4-stilbazolium-based fluorescent ligands, we showed that the mechanism of incorporating different ligands depends both on the binding step and the efficiency of the chemical reaction. By studying the different haloalkane dehalogenases DhaA, LinB, and DmmA, we found that the architecture of the access tunnels is critical for the kinetics of both steps and the ligand specificity. We show that highly efficient labelling with specific ligands is achievable with natural dehalogenases. We propose a simple protocol for selecting the optimal protein tag for a specific ligand from a wide pool of available enzymes with diverse access tunnel architectures. The application of this protocol eliminates a need for expensive and laborious protein engineering.

Keywords

HaloTag, enzyme kinetics, molecular modeling, reaction mechanism, ligand binding, nucleophilic substitution, protein engineering, access tunnel, numerical integration.

Abbreviations

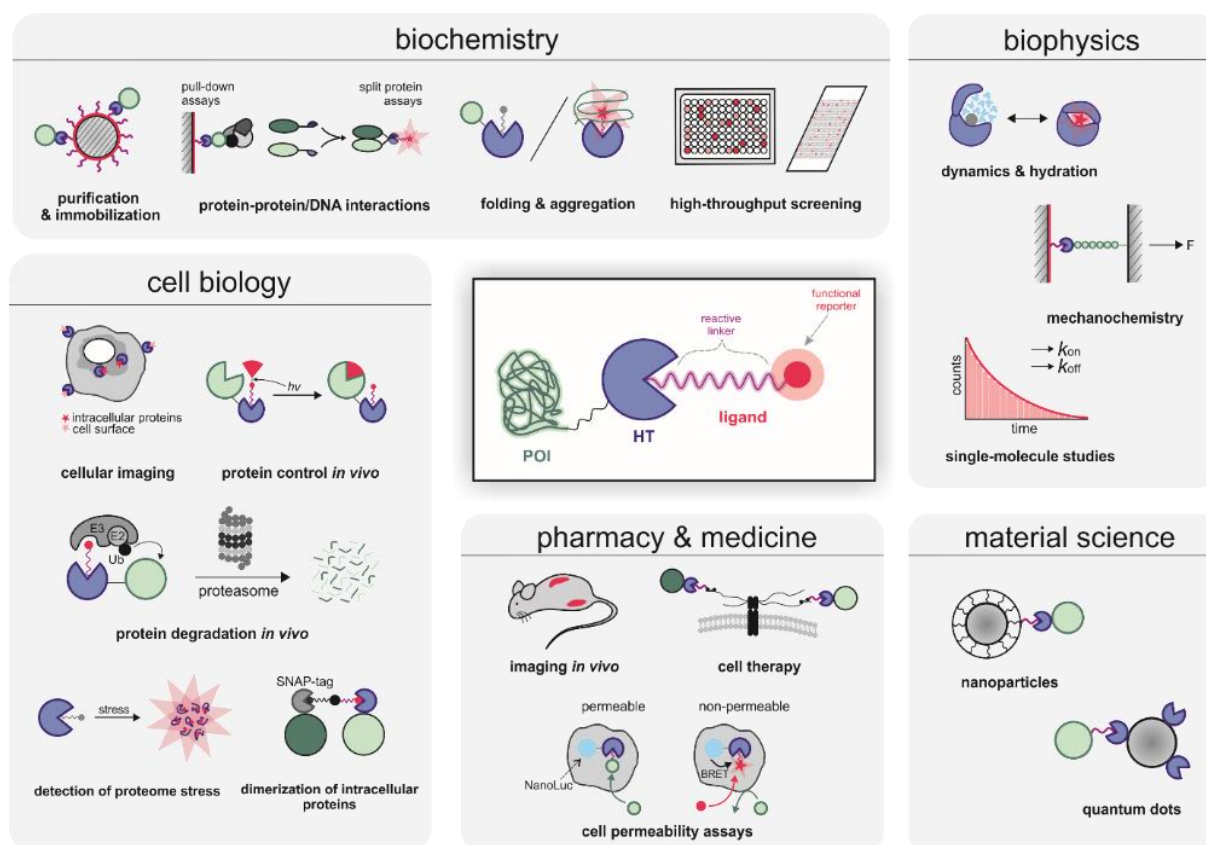
HLD, haloalkane dehalogenase; TMR, tetramethylrhodamine; PBS, phosphate-buffer saline; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; MSM, Markov state model; NAC, near-attack conformation; QM/MM, quantum mechanics/molecular mechanics.

Introduction

Genetically encoded protein labeling methods are widely employed in protein chemistry, molecular and cellular biology. More recently, self-labeling protein tags designed for covalent conjugation to small-molecule ligands functionalized with biorthogonal linkers have gained widespread attention. One of the most popular self-labeling methods, HaloTag®, uses engineered haloalkane dehalogenase (HLD) genetically fused to the proteins of interest, which covalently binds synthetic ligands bearing various functionalities, such as a strong light-up fluorescence response. The original concept of bifunctional linkers developed by Dick Janssen and co-workers¹ for covalent capturing and ribosomal/phage display of HLDs was translated to *in vivo* and *in vitro* analysis of mammalian proteins by Wood and co-workers^{2–4}.

Since its development and commercialization by Promega, HaloTag® has become a valuable research tool for a broad range of applications (**Figure 1A**) including protein purification⁵ and immobilization⁶, enhancement of the soluble expression of recombinant proteins⁷, cellular protein imaging^{8,9}, imaging *in vivo*¹⁰, and single-molecule studies^{11–13}. The technology is applicable for analyses of protein-protein and protein-nucleic acid interactions^{14,15}, proteome stress^{16,17}, protein folding and aggregation^{18,19}, dynamics and hydration^{20–22}, or cell-permeability²³. HaloTag fusions enable protein control *in vivo*^{24–26}, including degradation^{27,28} or dimerization²⁹ of proteins of interest. Further applications include high-throughput screening methods, microarrays, and chip technology^{30–32}, intracellular detection of pH³³ or biologically important ions^{34,35}, mechanochemistry^{36,37}, functionalization of nanoparticles³⁸, and quantum dots³⁹. Recently, the potential of the HaloTag technology in cell therapy was discovered, as it has been used for cell surface modification permitting angiogenesis, increased motility, and immune shielding⁴⁰.

A



B

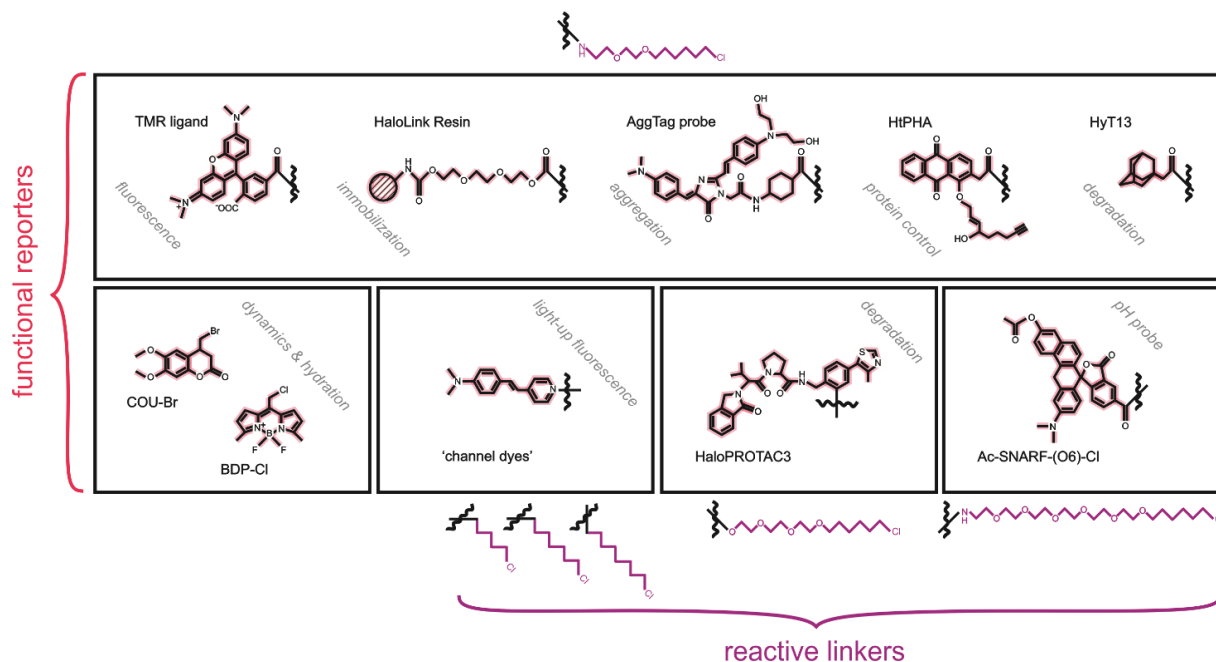


Figure 1. HaloTag® technology in chemistry, biology, biochemistry and biophysics. (A) Widespread applications of the HaloTag® technology. (B) The HaloTag ligands contain two crucial components (i) a reactive linker that initiates the formation of a covalent bond with the HaloTag protein, and (ii) a functional reporter.

A wide range of diverse HaloTag ligands have been designed and synthesized offering a variety of properties (**Figure 1B**), e.g., improved photostability and brightness¹², high biocompatibility or fluorogenicity allowing “no-wash” labeling protocols^{41–43} or providing specific affinity handles⁴. Despite the great diversity of ligands used, the majority of their applications always utilize the same tag protein DhaAHT, without considering the choice of another protein partner for better recognition of a specific ligand. The labeling reaction proceeds via a two-step kinetic pathway, the binding of the ligand, and the following chemical conversion, leading to the stable covalent alkyl-enzyme complex. The latter unimolecular step cannot be easily optimized by modifying the labeling protocol and it depends solely on the optimal reactive orientation of the bound ligand. The architecture and dynamics of the protein access tunnel and active site thus plays a key role. The 10,000-fold improvement in binding efficiency of the DhaAHT tag required for the successful protein imaging was achieved using a focused directed evolution on the access tunnel residues⁴⁴. This study has already shown that the efficiency obtained by molecular evolution differs significantly among individual ligands, despite sharing the same reactive linker. Similar effects were observed in our recent studies focused on the engineering of access tunnels in HLDs^{45,46}. The binding efficiency of HaloTag ligands varied across seven orders of magnitude for HLDs with different architectures of their access tunnels. Interestingly, the change in the functional reporter strongly affected the labeling efficiency even for the ligands with the same reactive linker^{45,46}. The results collectively suggest that the broadly used DhaAHT tag may not be the optimal tag for the incorporation of various available ligands. Since DhaAHT was introduced, the portfolio of available dehalogenases has significantly expanded^{47,48} and currently offers an interesting range of variants (**Figure 2, Table 1**).

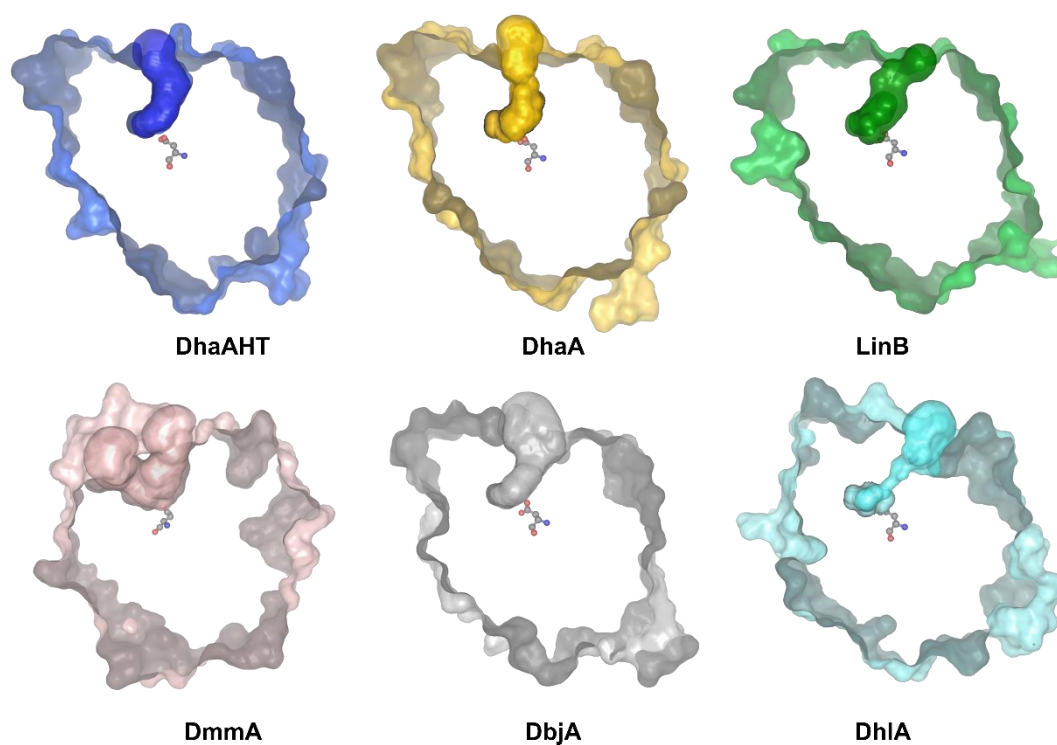


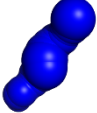

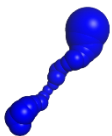




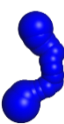



Figure 2. Tunnel diversity among the haloalkane dehalogenase family. The tag-optimized DhaAHT and five natural HLDs with different access tunnels and active site architectures. The proteins are represented by a cross-section of their surface, the active site is illustrated by the catalytic nucleophile, shown in ball and sticks, and the main tunnels connecting the active site to the surface are shown as a full surface. The tunnels and images were generated with Caver Analyst 2⁴⁹.

Table 1. Structural characteristics of the main tunnel in several natural HLDs and their variants.^a

Enzyme	Mutations ^b	Tunnel shape ^c	Geometry	PDB ID
DhaA	NA		Length: 16.1 Å Bottleneck radius: 1.45 Å Curvature: 1.19	4E46
LinB	NA		Length: 14.2 Å Bottleneck radius: 1.24 Å Curvature: 1.10	1MJ5
DmmaA	NA		Length: 14.4 Å Bottleneck radius: 1.58 Å Curvature: 1.19	3U1T

DbjA	NA		Length: 13.9 Å Bottleneck radius: 1.51 Å Curvature: 1.23	3A2M
DhIA	NA		Length: 17.6 Å Bottleneck radius: 0.65 Å Curvature: 1.18	2YXP
DbeA	NA		Length: 12.2 Å Bottleneck radius: 1.76 Å Curvature: 1.21	6XY9
DhaAHT	DhaA + K175M* + C176G* + V197I + H272F + Y273L + A292G		Length: 14.0 Å Bottleneck radius: 1.68 Å Curvature: 1.24	Model based on 4E46
DhaA31	DhaA + I135F + 176Y* + V245F* + L246I* + Y273F		Length: 14.0 Å Bottleneck radius: 0.97 Å Curvature: 1.28	3RK4
DhaA115	DhaA + E20S + F80R + C128F + T148L* + A155P + A172I* + C176F* + D198W + V219W + C262L + D266F		Length: 16.6 Å Bottleneck radius: 0.87 Å Curvature: 1.36	6SP5
LinB32	LinB + L177W*		Length: 12.7 Å Bottleneck radius: 1.31 Å Curvature: 1.36	4WDQ
DhIA-W175Y	DhIA + W175Y*		Length: 15.6 Å Bottleneck radius: 1.00 Å Curvature: 1.19	1BEE

^aThe main tunnel calculated with CAVER 3.02⁵⁰ in the respective structures after adding hydrogens atoms using PyMOL 2.3.2⁵¹; the tunnel origin is defined at the carboxylic-O atoms of Asp106; a probe radius is 0.6 Å; shell depth 4 Å; and shell radius 4 Å; ^bthe mutations affecting the p1 tunnel are marked with asterisk; NA means not applicable; ^call tunnels are shown from the same viewpoint, being the active site located at the lower end of the tunnels and the protein surface at the top.

In this study, we present a comprehensive kinetic and computational study of a mechanism of the HaloTag ligand incorporation. We investigated the effect of two different functional reporters, tetramethylrhodamine (**TMR**) and 4-stilbazolium, and different lengths of the reactive linker. We compared DhaAHT optimized by the directed evolution with three natural dehalogenases DhaA, LinB, and DmmA. The results have shown striking differences in the kinetics of both of the involved reaction steps, the ligand binding and the consecutive chemical step. The efficiency of each step is different for individual protein-ligand pairs. Strikingly, the most efficient reaction was not determined for DhaAHT and **TMR**, albeit systematically optimized by directed evolution. The wild-type enzymes LinB and DmmA showed the highest efficiency of the probe incorporation with the 4-stilbazolium probes. Our study proposes a new concept of selecting the protein tag matching the specific ligands.

Results

We performed comprehensive kinetic and theoretical study for the incorporation of two different HaloTag ligands, commercial tetramethylrhodamine (**TMR**) and three 4-stilbazolium-based ligands (**1B**, **1D**, and **1E**) with different lengths of the reactive linker (**Figure 3**). The 4-stilbazolium-based dyes have shown a stronger fluorogenic response upon labeling and easier synthetic routes than any of the previous HaloTag labels, which could be highly beneficial for their applicability⁴². We compared DhaAHT optimized by directed evolution with analogs of three natural dehalogenases DhaA, LinB, and DmmA. DhaAH272F, LinBH272F, and DmmAH315F, correspond to the natural enzymes with a single additional mutation in the catalytic base (H272F for DhaA and LinB, H315F in DmmA). This mutation in the catalytic histidine leads to the interruption of the catalytic cycle by preventing the hydrolytic step, and to the formation of the covalent alkyl-enzyme intermediate as the final complex.

Protein expression and purification

The haloalkane dehalogenases genes *linBH272F*, *dhaAH272F*, *dhaAHT*, and *dmmAH315F* were cloned into pAQN, pET21b or pET24a vectors and transformed into *Escherichia coli* BL21 (**Supplementary Table S1**). The enzymes were overexpressed and purified by metal-affinity chromatography. The purity of proteins was analyzed by SDS-PAGE (**Supplementary Figure S1**). A detailed description of the gene cloning, expression, and purification are provided in the **Supplementary Information (Section I)**.

Kinetic analysis

Fluorescence intensity and anisotropy were used to systematically analyze the concentration dependence of DhaAHT, DhaAH272F, LinBH272F, and DmmAH315F reactions with **TMR** and **1B**, **1D**, and **1E**. The conventional fitting and numerical integration methods were applied to obtain detailed information about the individual rate and equilibrium constants related to the two-step model of the HaloTag reaction (**Figure 3**).

Initially, the kinetic data were analyzed by conventional exponential fitting using nonlinear regression. To compare the consistency of the data with earlier published results, the apparent second-order rate constants were calculated following the procedure used originally by Los and co-workers⁴. The value of the apparent rate constant obtained for the incorporation of **TMR** into DhaAHT, $2.3 \times 10^6 \text{ M}^{-1}\cdot\text{s}^{-1}$, corresponds well with the value reported by Los and co-workers of $2.7 \times 10^6 \text{ M}^{-1}\cdot\text{s}^{-1}$.⁴ Next, the concentration dependence of the kinetic data (**Figure 3A**) was explored to provide detailed information about the kinetic pathway and the estimates of the true rate and equilibrium constants. Although the single-exponential fit of DhaAHT traces obtained with **TMR** provided satisfying statistics ($\chi^2/\text{DoF} = 2.29$; p-value = 0.28), the use of a double-exponential function showed significantly improved goodness of fit ($\chi^2/\text{DoF} = 1.21$; p-value = 0.43) and distinguished two separate phases, well consistent with expected two-step kinetic model for the reaction (**Figure 3C**).

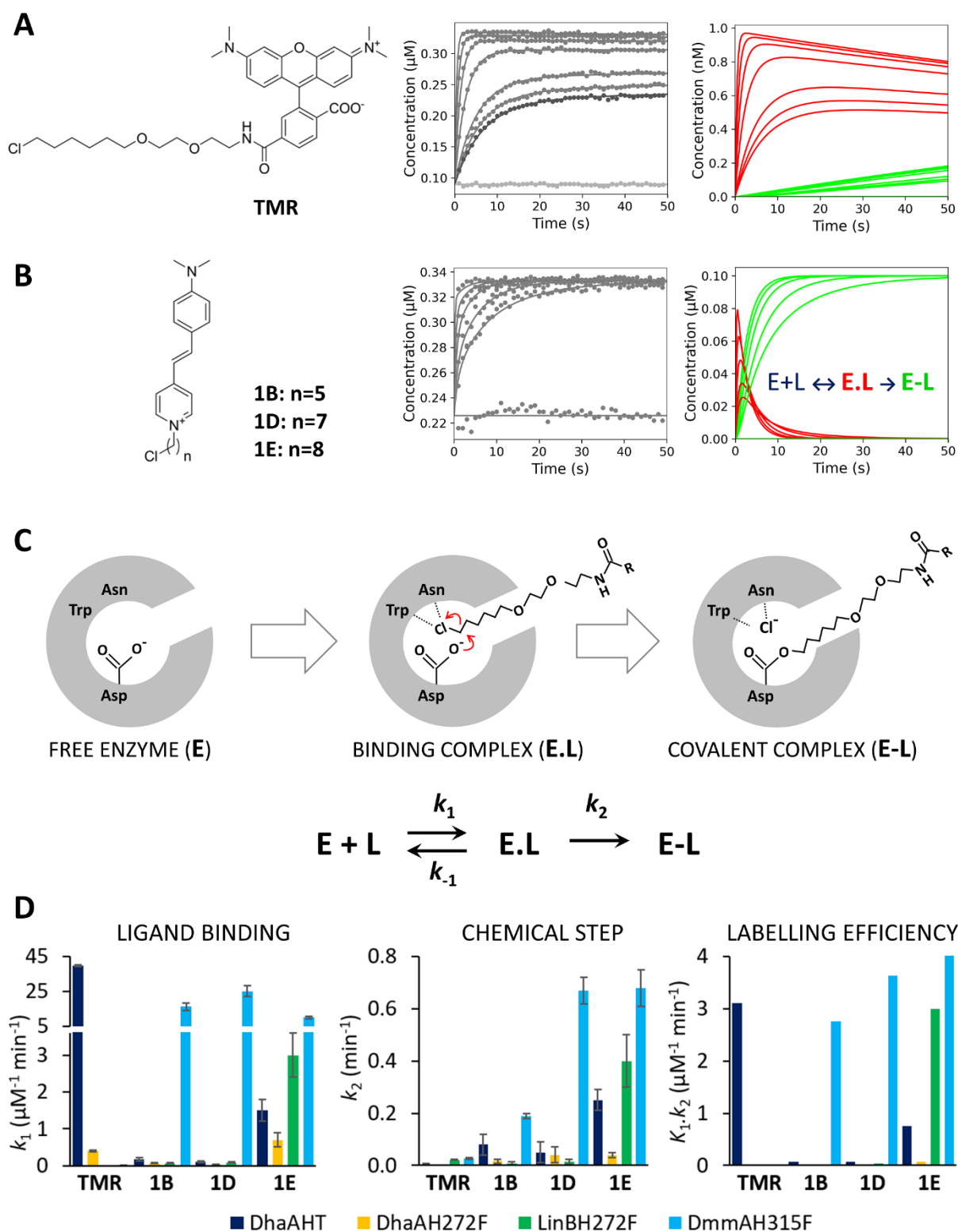


Figure 3. Kinetic mechanism of tetramethylrhodamine (TMR) and 4-stilbazolium ligand incorporation. (A) Chemical structure of **TMR** ligand (left) and anisotropy kinetic traces obtained upon mixing 0.001 μM **TMR** with 0 - 0.064 μM DhaAHT (middle). (B) Chemical structure of 4-stilbazolium ligands (left) and anisotropy kinetic traces obtained upon mixing 0.1 μM **1E** with 0 - 2 μM LinBH272F (middle). The anisotropy experiments were performed at 30 °C in PBS with 0.01 % (w/v) CHAPS and pH 7.4. The solid lines represent the best global fit to the kinetic data. The right figures show the time course of concentration of binding complex (E.L) and covalent alkyl-enzyme complex (E-L) obtained by numerical simulation. (C) Scheme of the HLD reaction with a halogenated ligand. The chemical

mechanism is adopted from Verschueren *et al.*⁵² The kinetic model of HaloTag reaction: E is the enzyme, L is the ligand, E.L is enzyme-ligand complex, E-L is the covalent alkyl-enzyme complex, k_1 and k_{-1} is the rate of association and dissociation of enzyme-ligand complex, respectively, and k_2 represents the rate of chemical step (nucleophilic substitution S_N2). (D) The kinetic parameters obtained by numerical analysis of anisotropy data. Error bars represents the standard error of the fitted parameters. The rigorous confidence contour analysis of variance of fitted parameters is presented in Supplementary Information (Supplementary Table S5). The kinetic experiments were performed in two to three independent replicates.

The concentration dependence of the obtained rates was analyzed analytically by a secondary fitting to the approximate rate equations derived for the two-step model (**Supplementary Equations S4** and **S5**). The analysis provided initial estimates of the rate constant for association ($k_1 = 41 \pm 4 \mu\text{M}^{-1}\cdot\text{min}^{-1}$) and dissociation ($k_{-1} = 0.08 \pm 0.04 \text{ min}^{-1}$) for **TMR**. DhaAHT bound complex and the rate constant for subsequent chemical step resulted in the formation of the covalent alkyl-enzyme complex ($k_2 = 0.06 \pm 0.04 \text{ min}^{-1}$). In the case of DhaAHT reaction with 4-stilbazolium-based ligands (**Figure 3B**), the binding phase gradually disappears in the dead-time of the measurement with increasing concentration of the enzyme, and only single-exponential fit provided reasonable estimates of the rates and amplitudes. The concentration dependence of the observed rate (**Supplementary Equation S6**) allowed to define the initial estimates of the equilibrium dissociation constant for the enzyme-ligand bound complex ($K_D = k_{-1}/k_1$) ranging from 0.6 to 1.3 μM and the rate of consecutive chemical step (k_2) ranging from 0.03 to 0.33 min^{-1} for individual 4-stilbazolium ligands.

The conventional analysis of the kinetics of DhaAHT indicated substantial differences in the reaction with the two types of tested ligands. The **TMR** probe shows a rapid binding to a tight enzyme-ligand complex followed by relatively slow chemical conversion leading to the final covalently bound complex. The accumulation of the reversible enzyme-ligand bound complex (E.L) is well visible from the anisotropic data (**Figure 3A middle**), which shows a significant concentration dependence of the equilibrium signal with amplitudes defined by the equilibrium dissociation constant of enzyme-ligand complex. The numerical simulation of the fraction of individual reaction species (**Figure 3A right**) illustrates the course of the reaction involving rapid binding of **TMR** into DhaAHT associated with the accumulation of enzyme-ligand bound complex (E.L, red), which is slowly transformed into the final covalent alkyl-enzyme complex (E-L, green) (**Figure 3A right**). In contrast, the kinetics of the 4-stilbazolium ligand reaction is dominated by the chemical step leading to the dominant accumulation of the final covalent complex (E-L). Anisotropic traces thus reach the same level of the signal in equilibrium, which is defined by the total ligand concentration (**Figure 3B**).

Although the conventional approach is currently the most widely used method of kinetic data analysis, it has several limitations, such as the loss of an important relationship between velocity and amplitude, or the accumulation of errors associated with successive calculation of a large number of temporary parameters to estimate a small number of relevant kinetic constants. To overcome these limitations, we performed a global data analysis based on

numerical methods. The parameter estimates obtained by conventional analysis (**Supplementary Figures S2 and Table S2**) were used as initial starting values for the numerical fitting. A detailed description of the conventional and numerical analysis of the kinetics data including rigorous statistical assessment is provided in the **Supplementary Information (Section II)**.

The global data fitting used numerical integration of the rate equations from an input model (**Figure 3C**) searching a unique set of kinetic parameters (**Supplementary Figure S3 and Table S3**) that explain the original raw data and produce a minimum χ^2 value⁵³. The observable anisotropy signal was defined as the sum of the contributions of each species to the total signal with scaling factors for each species (**Supplementary Table S4**). In addition to monitoring the standard errors and residuals, the global fitting of kinetic data allowed to perform a rigorous analysis of variance referred to as a confidence contour analysis⁵⁴. This analysis confirmed the high quality of the global fit, with all obtained kinetic parameters being well constrained by the experimental data (**Supplementary Table S5**). In the same way, the complex kinetic analysis was performed systematically comparing DhaAHT with three non-optimized natural variants DhaAH272F, LinBH272F, and DmmAH315F in the reaction with **TMR** and **1B**, **1D**, and **1E** (**Supplementary Figure S3 and Table S3**). The specific rate constants defining velocity of individual reaction steps, ligand binding (k_1) and chemical conversion (k_2), as well as the overall labeling efficiency defined by $K_1.k_2$, the product of the equilibrium constant for the ground-state binding $K_1 = 1/K_D = k_1/k_{-1}$ and the rate of the consecutive chemical step k_2 , are summarized for each enzyme variant in **Figure 3D**.

Unlike **TMR**, which provides only the possibility of instrumentally more complex anisotropy/polarization measurements, the 4-stilbazolium-based ligands provide the additional advantage in the tracking of fluorescence intensity signal (**Supplementary Figure S4**), commonly available in most laboratories. The increase in fluorescence intensity observed upon the incorporation of ligands into the enzymes was 5-, 2- and 10-fold for **1B**, **1D**, and **1E** ligand, respectively. Such increases provide sufficient signal for *in vitro* enzymology studies. Additionally, the strong signal change observed especially for the **1E** ligand provides a promising alternative to **TMR** in no-wash applications for cell labeling experiments.

Computational studies

DhaAHT, DhaAH272F, LinBH272F, and DmmAH315F were modeled from the crystal structures by *in silico* mutagenesis with Rosetta⁵⁵, and the respective access tunnels from the active site to the surface were calculated using CAVER 3.02⁵⁰. The tunnels found in DhaAHT, DhaAH272F, LinBH272F, and DmmAH315F showed considerably different geometric properties, especially the main tunnel, p1 (**Figure 4A** and **Supplementary Figure S5**). P1 is considerably wider in DhaAHT (bottleneck radius of 1.68 Å) than in DhaAH272F (1.29 Å) or LinBH272F (1.35 Å), and slightly wider than in DmmAH315F (1.62 Å). This fact suggests a higher accessibility of the DhaAHT active site as compared to DhaAH272F and LinBH272F, which provides a first explanation for the generally higher binding rates of the probes to DhaAHT and DmmAH315F than to the other two proteins. Moreover, the orientation of p1 is very different in LinBH272F and DmmAH315F compared to that of the DhaA variants. This suggests that LinBH272F and DmmAH315F may have different chemical and geometric preferences for the ligands that they can bind, compared to DhaAHT and DhaAH272F.

To understand the large differences found in the experimental kinetic measurements, we selected two representative probes (**TMR** and **1E**) and two proteins (DhaAHT and LinBH272F) to study their molecular binding in more detail. These systems were chosen because of the high binding specificity found among two of the corresponding pairs, i.e., DhaAHT with **TMR** and LinBH272F with **1E**. The **TMR** and **1E** ligands were modeled and then refined with the Density Function Theory, which provided the energy-minimized structures and the partial atomic charges (see **Supporting Information, Section III**, for details). The binding of both probes to DhaAHT and LinBH272F was studied by molecular dynamics (MD), using the adaptive sampling approach⁵⁶. The simulations started with the probes located in the bulk solvent, and consecutive rounds (epochs) of multiple MD simulations were performed. According to the adaptive sampling method, the starting points for the new MDs in each epoch were chosen from the previously sampled states based on the distance between the reacting groups in the probe and the enzyme. Each system was simulated for a total time of 20 μ s. Markov state model (MSM) analysis was performed to obtain the relevant kinetic ensembles describing the binding of the molecular probes to the active sites of the proteins. Four Markov states could describe well the binding process, consisting of one fully bound state, two intermediates, and a fully unbound state (**Figure 4B** and **Supplementary Figures S9** and **S10**). The kinetic parameters were calculated for the transitions between the most unbound state and the fully bound state. The results showed that the estimated binding rates (DhaAHT: $k_1 = 2.99 \pm 0.45 \times 10^8 \text{ M}^{-1}\cdot\text{s}^{-1}$ for **TMR**, $1.09 \pm 0.23 \times 10^8 \text{ M}^{-1}\cdot\text{s}^{-1}$ for **1E**; LinBH272F: $k_1 = 7.6 \pm 2.5 \times 10^7 \text{ M}^{-1}\cdot\text{s}^{-1}$ for **TMR**, $1.22 \pm 0.21 \times 10^8 \text{ M}^{-1}\cdot\text{s}^{-1}$ for **1E**; see **Supplementary Table S7**) followed the exact same order as the experimental ones, and k_1 was highest for DhaAHT with **TMR**, and lowest for LinBH272F with **TMR**. The computational and experimental rates, however, differed by several orders of magnitude, being higher for the theoretical values. Such discrepancy is not unprecedented^{56,57} and will be discussed below. Regarding the unbinding rates (k_{-1} ; **Supplementary Table S7**), they were all slower than the binding, which is consistent with the

majority of the experimental results obtained here, although the order was not strictly observed. The slowest unbinding was predicted for **TMR** with DhaAHT, while the experiments showed the lowest unbinding rate for **TMR** with LinBH272F. The predicted binding affinity, given by $K_1 = 1/K_D$, also partially followed the experimental trends, where DhaAHT and **TMR** were correctly predicted with the highest affinity. However, the predicted order among the other pairs was incorrect, where LinBH272F and **1E** showed experimentally the second strongest affinity. Interestingly, the highest probability of the bound state (P_{bound}) was obtained for **TMR** with DhaAHT, with 0.260 ± 0.043 , and the lowest was found for **1E** with LinBH272F, with 0.047 ± 0.008 (**Supplementary Table S7**).

Inspecting visually the bound states, we found that, in every system, the probes used exclusively the p1 tunnel (**Figure 4B** and **Supplementary Figure S10**). We also observed that the 4-stilbazolium aromatic system of **1E** was partially inserted in the tunnel, both in DhaAHT and LinBH272F. Conversely, due to its longer linker, **TMR** presented its aromatic moiety completely outside of the protein. Moreover, the bound conformations of **TMR** followed the natural orientation of the p1 tunnel in DhaAHT, while **1E** followed the tilted orientation of the p1 tunnel in LinBH272F (**Figure 4**). This interesting finding is in line with the fact that **1E** is a better binder with LinBH272F and DmmAH315F than **TMR**. It also suggests a higher complementarity of the 4-stilbazolium-based probes with the LinB and DmmA variants in comparison with **TMR**. Analyzing in more detail the interactions found in the LinBH272F-**1E** complex, we found that in the bound state the aromatic system formed close hydrophobic contacts with L150, V173, L177, and L248, located in the tunnel, and with L179, P245, A247, A271, located on the extension of the tunnel mouth. Interestingly, the negatively charged D147 residue also formed electrostatic interactions with the 4-stilbazolium system due to its delocalized positive charge. We expect that the many interactions and constraints of the aromatic system of **1B**, **1D**, and **1E** within the access tunnel contribute to strong fluorescence effects upon binding, as Clark and co-workers⁴² have previously suggested.

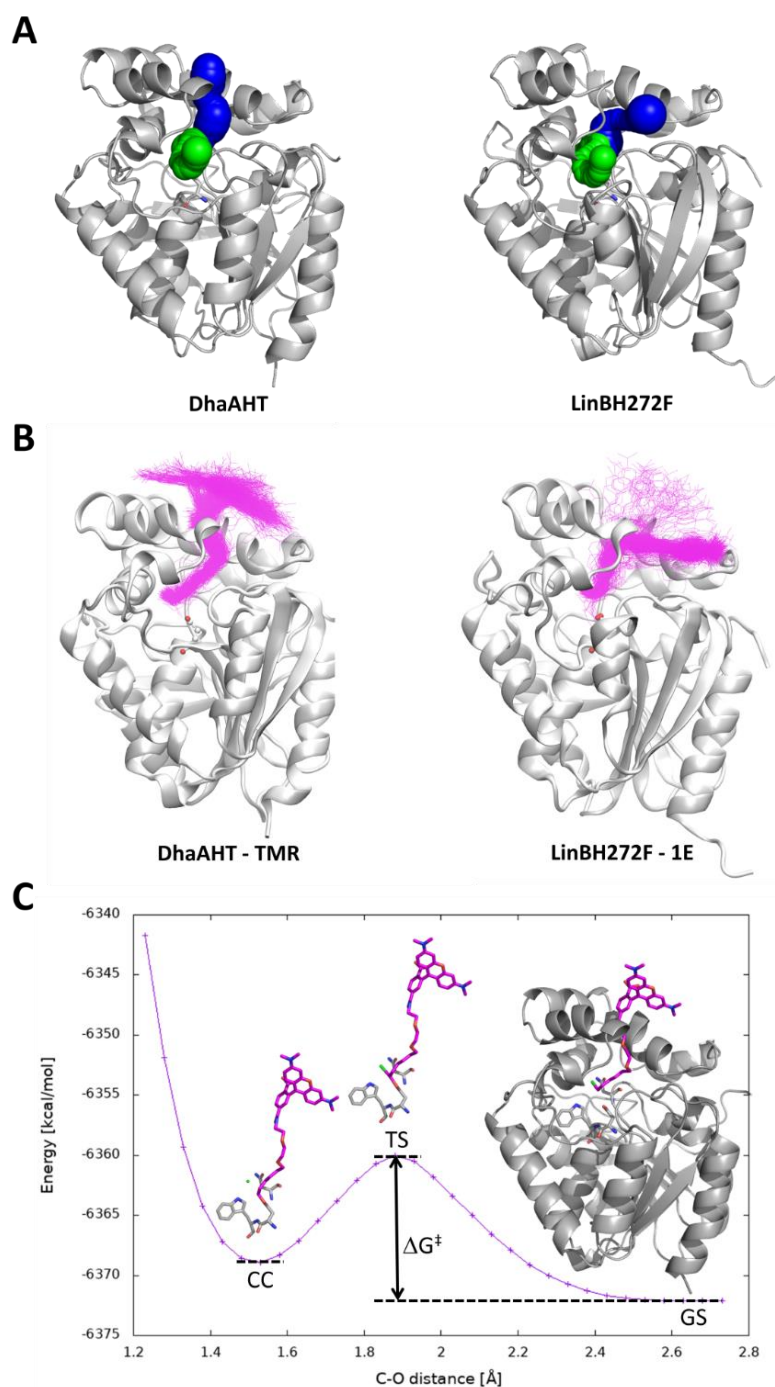


Figure 4. Molecular modeling of DhaAHT and LinBH272F and their binding to TMR and 1E. (A) The main access tunnel p1 (blue) and the slot tunnel p2 (green). (B) Structures of the complexes in the bound state obtained from Markov state analysis of the molecular dynamics simulations with a superimposition of the respective probes (magenta). (C) Potential energy surface of the S_N2 reaction between DhaAHT and **TMR**, obtained from an adiabatic mapping of the distance between the reacting atoms of the protein ($D106-COQ^-$) and the TMR probe (CH_2Cl). ΔG^\ddagger is the activation barrier of the reaction, where the ground state (GS), transition state (TS), and ligand-enzyme covalent complex (CC) are depicted. **TMR** is shown as the magenta sticks, the chloride ion as the green ball, the nucleophile D106 and the halide-stabilizing residues N41 and W107 as grey sticks.

Next, we predicted the reactivity of the **TMR** and **1E** probes towards DhaAHT and LinBH272F and compared them with the experimental results. We started by analyzing the pre-reactive complexes found during the respective MDs, hereafter termed as *near-attack conformation* (*NAC*). We estimated the constant of formation of this pre-reactive complex, K_{NAC} , based on the total number of *NAC*s found and the probability of the bound state, P_{bound} . Surprisingly, the highest K_{NAC} was obtained for **1E** with LinBH272F, and the lowest for **TMR** with DhaAHT (**Supplementary Table S7**), with a difference of nearly two orders of magnitude. This suggests that, despite the binding of **TMR** with DhaAHT being extremely fast, the probability of the system adopting a potentially reactive conformation is very low. In contrast, once **1E** and LinBH272F reach the bound state, the reactive conformation is achieved much faster than for any of the other systems. This perfectly follows the trends of the experimental kinetic rates obtained of the chemical step (k_2 in **Figure 3** and **Supplementary Table S7**). We then applied a hybrid QM/MM adiabatic mapping to estimate the energy barriers of the S_N2 reaction, ΔG^\ddagger_{SN2} (**Figure 4C**). The predicted ΔG^\ddagger_{SN2} values (**Supplementary Table S7**) showed the lowest activation barrier for LinBH272F-**TMR** (12.1 ± 1.9 kcal·mol⁻¹), followed by LinBH272F-**1E** (13.8 ± 1.8 kcal·mol⁻¹), and the highest barrier for DhaAHT-**TMR** (15.5 ± 1.3 kcal·mol⁻¹). This indicates that once the *NAC* has been achieved, LinBH272F provides a better environment for performing the S_N2 step with both of the probes than DhaAHT.

Finally, the K_{NAC} and ΔG^\ddagger_{SN2} were combined (**Supplementary Information, Section III**) to estimate the overall activation energy of the second kinetic step, ΔG^\ddagger_2 . This step 2 (**Figure 3C**) is a direct measure of reactivity, and the estimated and experimental values of ΔG^\ddagger_2 can be directly compared. As a result, the highest calculated ΔG^\ddagger_2 value was obtained for DhaAHT with **TMR**, 2.5 kcal·mol⁻¹ above the second highest energy barrier of DhaAHT-**1E** and 4.4 kcal·mol⁻¹ above the ΔG^\ddagger_2 value for LinBH272F-**1E** (**Supplementary Table S7**). This is in reasonably good agreement with the experimental values, where DhaAHT with **TMR** also showed the highest ΔG^\ddagger_2 value, 2.7 kcal·mol⁻¹ above that of LinBH272F with **1E**.

Discussion

We conducted a detailed kinetic and computational study of the reaction of a tetramethylrhodamine-based (**TMR**) and three 4-stilbazolium-based ligands (**1B**, **1D**, and **1E**) with reacting linkers of different lengths, and the haloalkane dehalogenase optimized by directed evolution (DhaAHT) and the analogs of three natural dehalogenases (DhaAH272F, LinBH272F, and DmmAH315F). The kinetic study showed substantial differences in the reaction kinetics between the individual enzymes as well as among the different ligands. The **TMR** probe showed a very high rate of binding towards the engineered DhaAHT ($k_1 = 39.7 \pm 0.6$ $\mu\text{M}^{-1}\cdot\text{min}^{-1}$), which is in high contrast with the other studied enzymes, which showed rates of binding ranging between 0.001 and 0.4 $\mu\text{M}^{-1}\cdot\text{min}^{-1}$. In comparison to the non-optimized DhaAH272F, the engineering of the access tunnel of DhaA⁴⁴ led to a significant improvement of the ligand binding, while it did not compromise the catalytic step. The engineered DhaAHT showed even a slightly decreased activation barrier of the chemical step ($\Delta\Delta G^\ddagger_2 = -0.6$ kcal·mol⁻¹

¹), although the ground state energy ($\Delta\Delta G_0$) of the enzyme-ligand complex was significantly lower ($-3.3 \text{ kcal}\cdot\text{mol}^{-1}$) in comparison to DhaAH272F (**Supplementary Table S6**). Interestingly, the non-optimized LinBH272F showed a slow binding towards **TMR**, but it displayed the highest rate of the consecutive chemical conversion leading to the formation of the covalent enzyme-**TMR** complex.

The importance of the chemical step for the efficiency of the HaloTag labeling reactions is more pronounced with the 4-stilbazolium-based ligands. Even though the binding of **1E** into LinBH272F is orders of magnitude slower and weaker, the elevated velocity of the following chemical step ensures the fully comparable labeling efficiency of **1E** with LinBH272F ($K_1.k_2 = 3.0 \text{ }\mu\text{M}^{-1}\cdot\text{min}^{-1}$) to that observed for the reaction of **TMR** with the engineered DhaAHT ($K_1.k_2 = 3.1 \text{ }\mu\text{M}^{-1}\cdot\text{min}^{-1}$). The weaker binding of **1E** to LinBH272F ($\Delta\Delta G_0 = 2.8 \text{ kcal}\cdot\text{mol}^{-1}$) is compensated by a lower activation energy, resulting in a fast consecutive chemical conversion ($\Delta\Delta G^\ddagger = -2.8 \text{ kcal}\cdot\text{mol}^{-1}$). The reaction of **1E** with LinBH272F illustrates that the desired labeling efficiency can be achieved not only by an improved ligand binding, but also by an acceleration of the chemical reaction. The reaction mechanism observed for LinBH272F can be explained by the specific architecture of its access tunnel. The narrower tunnel bottleneck compromises to some extent the ligand transport, but at the same time reduces the active site solvation and makes the productive binding more probable. Moreover, the tunnel lining residues lower the initial entropy and promote the contact of the reacting atoms, possibly through specific interactions with the ligand. All these effects may have a positive effect on increasing the rate of the carbon-halogen bond cleavage (S_N2), and have been described in previous studies focused on the engineering of access tunnels, for both DhaA and LinB⁵⁸⁻⁶¹. However, the narrow architecture of the access tunnel in LinB makes this enzyme more sensitive to the length of the reactive linker. This is important to allow the formation of favorable interactions between the aromatic system of the probes and the residues lining the tunnel. The eight-carbon linker of **1E** was the only one providing optimal length for LinB labeling since shorter linkers have not been able to achieve even remotely the efficiency of **1E**.

Wider and more accessible access tunnels seem to be more universal, as evidenced by the reaction of the non-optimized DmmAH315F with 4-stilbazolium ligands. DmmA has the most open main tunnel ensuring easy access of the substrates to its active site. The reaction of DmmAH315F with all 4-stilbazolium-based ligands showed rapid binding, but also rapid chemical steps. The resulting labeling efficiency thus surpasses the commercial reaction of **TMR** with DhaAHT. The reaction of LinBH272F with **1E** and DmmAH315F with all the 4-stilbazolium ligands showed that natural variants can provide high efficiency useful for HaloTag applications without time-demanding protein optimization by directed evolution. It is also interesting that, just as DhaAHT is highly specific for the reaction with **TMR**, DmmAH315F showed high efficiency only in reaction with the 4-stilbazolium-based ligands, but not with **TMR**. Clearly, it is important to select an appropriate protein for the binding of a specific ligand.

Using several computational methods, we simulated and predicted the kinetics and thermodynamics of the two-step binding process of four representative fluorescent probe/protein systems, namely for the **TMR** and **1E** probes with DhaAHT and LinBH272F. We found disparities in the absolute values of the calculated kinetic rates. Such differences have been reported previously^{56,57}, and can be attributed to the bias intrinsic to the simulation method (adaptive sampling) or the conditions used in our MD simulations, namely the force field and the solvent model. The ligand transport in proteins is highly influenced by the solvent and its respective bulk properties, such as diffusivity. In spite of being one of the most widely used water models in molecular simulations, the TIP3P model has a higher diffusivity than pure water. It is also known to overestimate the diffusion properties of amino acids⁶², and we can presume that that same holds for many other solvated molecules. Importantly, our results showed significant correlations with some of the experimental parameters and revealed important clues for different aspects of the molecular binding on focus here. We could qualitatively replicate the order in the k_1 binding rates, with DhaAHT-**TMR** showing the highest k_1 value, followed by LinBH272F-**1E**, and partially the order of affinities. The unbinding rates, however, were less consistent with the experimental results. The simulation of the binding process revealed how the probes interacted differently with the proteins. Both used solely p1 tunnel to bind the proteins, but the preferred orientation of **TMR** was more compatible with the geometry of the p1 tunnel in the DhaA variants, while **1E** adopted an orientation more similar to the p1 tunnel found in the native LinBH272F and DmmAH315F. This reveals a complementarity intrinsic to those two pairs that seem to explain the labeling efficiencies described above. The large number of interactions formed between **1E** and residues lining the tunnel of LinBH272F support this hypothesis.

The overall chemical step was dissected into the pre-organization of the bound state to form the pre-reactive complex (NAC) and the S_N2 reaction. We estimated these two partial steps from our computational approach and calculated the total activation energy of the second kinetic step (ΔG^\ddagger_2), which can be compared with the parameter determined experimentally. We found that DhaAHT-**TMR** displayed the lowest overall reactivity (with the highest ΔG^\ddagger_2 value), which is in good agreement with the experimental data. It showed not only the worst efficiency in achieving a productive binding mode (lowest K_{NAC}) but also presented the highest activation barrier to the S_N2 reaction. Conversely, LinBH272F-**1E** was the most efficient system in adopting the pre-reactive conformation after the binding (highest K_{NAC}). A low ΔG^\ddagger_{SN2} also resulted in LinBH272F-**1E** having a rather low activation barrier to the overall chemical step, with an estimated ΔG^\ddagger_2 value below that of DhaAHT-**TMR** by 4.4 kcal·mol⁻¹. Some of the discrepancies between the theoretical and experimental values were likely due to a poor sampling of the fully bound states, which may have not been sufficient to provide an accurate ensemble distribution of the pre-reactive state. However, our results provided sufficient clues to explain why, although DhaAHT-**TMR** presented the highest binding rate, its reactivity is very far from ideal. In contrast, although the LinBH272F-**1E** system had poorer binding rates, it is much more efficient on the chemical step. Overall, the binding of the **1E** probe to the non-optimized LinBH272F protein revealed a reasonable

binding/reactivity trade-off, which resulted in a labeling efficiency very close to that of DhaAHT-**TMR**.

Some of the effects discussed above can be extrapolated to the DmmAH315F-**1E** system, which presented the best binding efficiency among all the tested pairs. Hence, we hypothesize that the binding of **1E** to DmmAH315F is fast due to the combination of a sufficiently wide access tunnel and a good complementarity of its architecture with the **1E** probe, which lead to a high number of favorable interactions. Secondly, strong probe-enzyme interactions can contribute to a stable and highly reactive DmmAH315F-**1E** complex, thus leading to a fast chemical step. The combination of a fast binding rate and a fast chemical step resulted in a system with the highest labeling efficiency.

Conclusions

Here we have demonstrated that not only the ligand accessibility is important for the binding of diverse probes to HaloTag proteins, but also the subsequent chemical step can significantly affect the ligand specificity and labeling efficiency. We have identified substantial differences in the kinetics of binding and chemical reaction between individual enzymes with different ligands. The **TMR** probe showed a rapid binding to DhaAHT, which was followed by a slow chemical conversion to the alkyl-enzyme complex. In contrast, the binding of the 4-stilbazolium-based ligands to DhaAHT and other tag proteins was much slower than with **TMR**, but the chemical step was greatly improved in most cases. Interestingly, we found that the best efficiencies for the incorporation of several 4-stilbazolium-based probes (namely **1D** and **1E**) were achieved with the analogs of natural non-optimized dehalogenases, LinBH272F and DmmAH315F, which provided high kinetic rates for both binding and chemistry. This demonstrates that different natural proteins can be effective for the incorporation of specific probes without the need for demanding protein engineering procedures. Moreover, the 4-stilbazolium-based ligands, due to a better light-up response upon binding, may provide better detection limits and thus could be preferable to the traditional probes, e.g. for simple fluorescence assays, analysis of binding interactions, or microscopy imaging.

We propose that, before conducting laborious optimization rounds by directed evolution, a rapid screening of the available natural dehalogenases could lead to the identification of potential candidates for optimal tag proteins. Thus, one could benefit from the very diverse pool of tunnel architectures already available among the known haloalkane dehalogenases. Calculation of the respective access tunnels with CAVER and molecular docking could provide good first filters for this selection. The subsequent utilization of more robust computational methods, like molecular dynamics and quantum mechanics, can help identifying the ideal enzyme-probe pairs.

Materials and Methods

All materials and methods are described in detail in **Supplementary Information**.

Acknowledgements

This work has been supported by the Czech Ministry of Education (grants INBIO CZ.02.1.01/0.0/0.0/16_026/0008451, ESFRI RECETOX LM2018121 and ESFRI ELIXIR LM2018131), and the European Union (SinFonia 814418 and TEAMING 857560). The computational resources were supplied by the project “e-Infrastruktura CZ” (e-INFRA LM2018140) provided within the program Projects of Large Research, Development and Innovations Infrastructures. MM acknowledges support from the internal Grant Agency of the Masaryk University (MUNI/H/1561/2018). ETK acknowledges support from the U.S. National Cancer Institute (CA217809).

References

- (1) Pieters, R. J.; Fennema, M.; Kellogg, R. M.; Janssen, D. B. Design and Synthesis of Reagents for Phage Display Screening of Dehalogenases. *Bioorg. Med. Chem. Lett.* **1999**, *9* (2), 161–166. [https://doi.org/10.1016/S0960-894X\(98\)00697-0](https://doi.org/10.1016/S0960-894X(98)00697-0).
- (2) Los, G. V.; Darzins, A.; Karassina, N.; Zimprich, C.; Learish, R.; McDougall, M. G.; Encell, L. P.; Friedman-Ohana, R.; Wood, M.; Vidugiris, G.; Zimmerman, K.; Otto, P.; Klaubert, D. H.; Wood, K. V. HaloTagTM Interchangeable Labeling Technology for Cell Imaging and Protein Capture. *Cell Notes* **2005**, *11*, 2–6.
- (3) Los, G. V.; Wood, K. The HaloTagTM: A Novel Technology for Cell Imaging and Protein Analysis. In *High Content Screening: A Powerful Approach to Systems Cell Biology and Drug Discovery*; Taylor, D. L., Haskins, J. R., Giuliano, K. A., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, 2006; pp 195–208. <https://doi.org/10.1385/1-59745-217-3:195>.
- (4) Los, G. V.; Encell, L. P.; McDougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3* (6), 373–382. <https://doi.org/10.1021/cb800025k>.
- (5) Mishra, V. Affinity Tags for Protein Purification. *Curr. Protein Pept. Sci.* **2020**, *21* (8), 821–830. <https://doi.org/10.2174/1389203721666200606220109>.
- (6) Döbber, J.; Pohl, M. HaloTagTM: Evaluation of a Covalent One-Step Immobilization for Biocatalysis. *J. Biotechnol.* **2017**, *241*, 170–174. <https://doi.org/10.1016/j.jbiotec.2016.12.004>.
- (7) Sun, C.; Li, Y.; Taylor, S. E.; Mao, X.; Wilkinson, M. C.; Fernig, D. G. HaloTag Is an Effective Expression and Solubilisation Fusion Partner for a Range of Fibroblast Growth Factors. *PeerJ* **2015**, *3*, e1060. <https://doi.org/10.7717/peerj.1060>.
- (8) Singh, V.; Wang, S.; Chan, K. M.; Clark, S. A.; Kool, E. T. Genetically Encoded Multispectral Labeling of Proteins with Polyfluorophores on a DNA Backbone. *J. Am. Chem. Soc.* **2013**, *135* (16), 6184–6191. <https://doi.org/10.1021/ja4004393>.

- (9) Hoelzel, C. A.; Zhang, X. Visualizing and Manipulating Biological Processes by Using HaloTag and SNAP-Tag Technologies. *Chembiochem Eur. J. Chem. Biol.* **2020**, *21* (14), 1935–1946. <https://doi.org/10.1002/cbic.202000037>.
- (10) Kosaka, N.; Ogawa, M.; Choyke, P. L.; Karassina, N.; Corona, C.; McDougall, M.; Lynch, D. T.; Hoyt, C. C.; Levenson, R. M.; Los, G. V.; Kobayashi, H. In Vivo Stable Tumor-Specific Painting in Various Colors Using Dehalogenase-Based Protein-Tag Fluorescent Ligands. *Bioconjug. Chem.* **2009**, *20* (7), 1367–1374. <https://doi.org/10.1021/bc9001344>.
- (11) Chen, J.; Zhang, Z.; Li, L.; Chen, B.-C.; Revyakin, A.; Hajj, B.; Legant, W.; Dahan, M.; Lionnet, T.; Betzig, E.; Tjian, R.; Liu, Z. Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. *Cell* **2014**, *156* (6), 1274–1285. <https://doi.org/10.1016/j.cell.2014.01.062>.
- (12) Grimm, J. B.; English, B. P.; Chen, J.; Slaughter, J. P.; Zhang, Z.; Revyakin, A.; Patel, R.; Macklin, J. J.; Normanno, D.; Singer, R. H.; Lionnet, T.; Lavis, L. D. A General Method to Improve Fluorophores for Live-Cell and Single-Molecule Microscopy. *Nat. Methods* **2015**, *12* (3), 244–250, 3 p following 250. <https://doi.org/10.1038/nmeth.3256>.
- (13) Rivas-Pardo, J. A.; Li, Y.; Mártonfalvi, Z.; Tapia-Rojo, R.; Unger, A.; Fernández-Trasancos, Á.; Herrero-Galán, E.; Velázquez-Carreras, D.; Fernández, J. M.; Linke, W. A.; Alegre-Cebollada, J. A HaloTag-TEV Genetic Cassette for Mechanical Phenotyping of Proteins from Tissues. *Nat. Commun.* **2020**, *11* (1), 2060. <https://doi.org/10.1038/s41467-020-15465-9>.
- (14) Urh, M.; Hartzell, D.; Mendez, J.; Klaubert, D. H.; Wood, K. Methods for Detection of Protein-Protein and Protein-DNA Interactions Using HaloTag. *Methods Mol. Biol. Clifton NJ* **2008**, *421*, 191–209. https://doi.org/10.1007/978-1-59745-582-4_13.
- (15) Minner-Meinen, R.; Weber, J.-N.; Albrecht, A.; Matis, R.; Behnecke, M.; Tietge, C.; Frank, S.; Schulze, J.; Buschmann, H.; Walla, P. J.; Mendel, R.-R.; Hänsch, R.; Kaufholdt, D. Split-HaloTag Imaging Assay for Sophisticated Microscopy of Protein-Protein Interactions in Planta. *Plant Commun.* **2021**, *2* (5), 100212. <https://doi.org/10.1016/j.xplc.2021.100212>.
- (16) Liu, Y.; Fares, M.; Dunham, N. P.; Gao, Z.; Miao, K.; Jiang, X.; Bollinger, S. S.; Boal, A. K.; Zhang, X. AgHalo: A Facile Fluorogenic Sensor to Detect Drug-Induced Proteome Stress. *Angew. Chem. Int. Ed Engl.* **2017**, *56* (30), 8672–8676. <https://doi.org/10.1002/anie.201702417>.
- (17) Fares, M.; Li, Y.; Liu, Y.; Miao, K.; Gao, Z.; Zhai, Y.; Zhang, X. A Molecular Rotor-Based Halo-Tag Ligand Enables a Fluorogenic Proteome Stress Sensor to Detect Protein Misfolding in Mildly Stressed Proteome. *Bioconjug. Chem.* **2018**, *29* (1), 215–224. <https://doi.org/10.1021/acs.bioconjchem.7b00763>.
- (18) Wolstenholme, C. H.; Hu, H.; Ye, S.; Funk, B. E.; Jain, D.; Hsiung, C.-H.; Ning, G.; Liu, Y.; Li, X.; Zhang, X. AggFluor: Fluorogenic Toolbox Enables Direct Visualization of the Multi-Step Protein Aggregation Process in Live Cells. *J. Am. Chem. Soc.* **2020**, *142* (41), 17515–17523. <https://doi.org/10.1021/jacs.0c07245>.
- (19) Jung, K. H.; Kim, S. F.; Liu, Y.; Zhang, X. A Fluorogenic AggTag Method Based on Halo- and SNAP-Tags to Simultaneously Detect Aggregation of Two Proteins in Live Cells. *Chembiochem Eur. J. Chem. Biol.* **2019**, *20* (8), 1078–1087. <https://doi.org/10.1002/cbic.201800782>.
- (20) Sykora, J.; Brezovsky, J.; Koudelakova, T.; Lahoda, M.; Fortova, A.; Chernovets, T.; Chaloupkova, R.; Stepankova, V.; Prokop, Z.; Smatanova, I. K.; Hof, M.; Damborsky, J.

- Dynamics and Hydration Explain Failed Functional Transformation in Dehalogenase Design. *Nat. Chem. Biol.* **2014**, *10* (6), 428–430.
<https://doi.org/10.1038/nchembio.1502>.
- (21) Amaro, M.; Brezovský, J.; Kováčová, S.; Sýkora, J.; Bednář, D.; Němec, V.; Lišková, V.; Kurumbang, N. P.; Beerens, K.; Chaloupková, R.; Paruch, K.; Hof, M.; Damborský, J. Site-Specific Analysis of Protein Hydration Based on Unnatural Amino Acid Fluorescence. *J. Am. Chem. Soc.* **2015**, *137* (15), 4988–4992.
<https://doi.org/10.1021/jacs.5b01681>.
- (22) Popa, I.; Rivas-Pardo, J. A.; Eckels, E. C.; Echelman, D. J.; Badilla, C. L.; Valle-Orero, J.; Fernández, J. M. A HaloTag Anchored Ruler for Week-Long Studies of Protein Dynamics. *J. Am. Chem. Soc.* **2016**, *138* (33), 10546–10553.
<https://doi.org/10.1021/jacs.6b05429>.
- (23) Peier, A.; Ge, L.; Boyer, N.; Frost, J.; Duggal, R.; Biswas, K.; Edmondson, S.; Hermes, J. D.; Yan, L.; Zimprich, C.; Sadruddin, A.; Kristal Kaan, H. Y.; Chandramohan, A.; Brown, C. J.; Thean, D.; Lee, X. E.; Yuen, T. Y.; Ferrer-Gago, F. J.; Johannes, C. W.; Lane, D. P.; Sherborne, B.; Corona, C.; Robers, M. B.; Sawyer, T. K.; Partridge, A. W. NanoClick: A High Throughput, Target-Agnostic Peptide Cell Permeability Assay. *ACS Chem. Biol.* **2021**, *16* (2), 293–309. <https://doi.org/10.1021/acscchembio.0c00804>.
- (24) Parvez, S.; Fu, Y.; Li, J.; Long, M. J. C.; Lin, H.-Y.; Lee, D. K.; Hu, G. S.; Aye, Y. Substoichiometric Hydroxynonenylation of a Single Protein Recapitulates Whole-Cell-Stimulated Antioxidant Response. *J. Am. Chem. Soc.* **2015**, *137* (1), 10–13.
<https://doi.org/10.1021/ja5084249>.
- (25) Parvez, S.; Long, M. J. C.; Lin, H.-Y.; Zhao, Y.; Haegele, J. A.; Pham, V. N.; Lee, D. K.; Aye, Y. T-REX on-Demand Redox Targeting in Live Cells. *Nat. Protoc.* **2016**, *11* (12), 2328–2356. <https://doi.org/10.1038/nprot.2016.114>.
- (26) Fang, X.; Fu, Y.; Long, M. J. C.; Haegele, J. A.; Ge, E. J.; Parvez, S.; Aye, Y. Temporally Controlled Targeting of 4-Hydroxynonenal to Specific Proteins in Living Cells. *J. Am. Chem. Soc.* **2013**, *135* (39), 14496–14499. <https://doi.org/10.1021/ja405400k>.
- (27) Neklesa, T. K.; Tae, H. S.; Schneekloth, A. R.; Stulberg, M. J.; Corson, T. W.; Sundberg, T. B.; Raina, K.; Holley, S. A.; Crews, C. M. Small-Molecule Hydrophobic Tagging-Induced Degradation of HaloTag Fusion Proteins. *Nat. Chem. Biol.* **2011**, *7* (8), 538–543. <https://doi.org/10.1038/nchembio.597>.
- (28) Buckley, D. L.; Raina, K.; Darricarrere, N.; Hines, J.; Gustafson, J. L.; Smith, I. E.; Miah, A. H.; Harling, J. D.; Crews, C. M. HaloPROTACS: Use of Small Molecule PROTACS to Induce Degradation of HaloTag Fusion Proteins. *ACS Chem. Biol.* **2015**, *10* (8), 1831–1837. <https://doi.org/10.1021/acscchembio.5b00442>.
- (29) Erhart, D.; Zimmermann, M.; Jacques, O.; Wittwer, M. B.; Ernst, B.; Constable, E.; Zvelebil, M.; Beaufils, F.; Wymann, M. P. Chemical Development of Intracellular Protein Heterodimerizers. *Chem. Biol.* **2013**, *20* (4), 549–557.
<https://doi.org/10.1016/j.chembiol.2013.03.010>.
- (30) Gu, L.; Li, C.; Aach, J.; Hill, D. E.; Vidal, M.; Church, G. M. Multiplex Single-Molecule Interaction Profiling of DNA-Barcoded Proteins. *Nature* **2014**, *515* (7528), 554–557.
<https://doi.org/10.1038/nature13761>.
- (31) Noblin, D. J.; Page, C. M.; Tae, H. S.; Gareiss, P. C.; Schneekloth, J. S.; Crews, C. M. A HaloTag-Based Small Molecule Microarray Screening Methodology with Increased Sensitivity and Multiplex Capabilities. *ACS Chem. Biol.* **2012**, *7* (12), 2055–2063.
<https://doi.org/10.1021/cb300453k>.

- (32) Yazaki, J.; Galli, M.; Kim, A. Y.; Nito, K.; Aleman, F.; Chang, K. N.; Carvunis, A.-R.; Quan, R.; Nguyen, H.; Song, L.; Alvarez, J. M.; Huang, S.-S. C.; Chen, H.; Ramachandran, N.; Altmann, S.; Gutiérrez, R. A.; Hill, D. E.; Schroeder, J. I.; Chory, J.; LaBaer, J.; Vidal, M.; Braun, P.; Ecker, J. R. Mapping Transcription Factor Interactome Networks Using HaloTag Protein Arrays. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113* (29), E4238-4247. <https://doi.org/10.1073/pnas.1603229113>.
- (33) Benink, H.; McDougall, M.; Klaubert, D.; Los, G. Direct PH Measurements by Using Subcellular Targeting of 5(and 6-) Carboxysemaphthorhodafluor in Mammalian Cells. *BioTechniques* **2009**, *47* (3), 769–774. <https://doi.org/10.2144/000113220>.
- (34) Taguchi, R.; Terai, T.; Ueno, T.; Komatsu, T.; Hanaoka, K.; Urano, Y. A Protein-Coupled Fluorescent Probe for Organelle-Specific Imaging of Na⁺. *Sens. Actuators B Chem.* **2018**, *265*, 575–581. <https://doi.org/10.1016/j.snb.2018.03.090>.
- (35) Li, D.; Liu, L.; Li, W.-H. Genetic Targeting of a Small Fluorescent Zinc Indicator to Cell Surface for Monitoring Zinc Secretion. *ACS Chem. Biol.* **2015**, *10* (4), 1054–1063. <https://doi.org/10.1021/cb5007536>.
- (36) Popa, I.; Berkovich, R.; Alegre-Cebollada, J.; Badilla, C. L.; Rivas-Pardo, J. A.; Taniguchi, Y.; Kawakami, M.; Fernandez, J. M. Nanomechanics of HaloTag Tethers. *J. Am. Chem. Soc.* **2013**, *135* (34), 12762–12771. <https://doi.org/10.1021/ja4056382>.
- (37) Alonso-Caballero, A.; Echelman, D. J.; Tapia-Rojas, R.; Haldar, S.; Eckels, E. C.; Fernandez, J. M. Protein Folding Modulates the Chemical Reactivity of a Gram-Positive Adhesin. *Nat. Chem.* **2021**, *13* (2), 172–181. <https://doi.org/10.1038/s41557-020-00586-x>.
- (38) Garbujo, S.; Galbiati, E.; Salvioni, L.; Mazzucchelli, M.; Frascotti, G.; Sun, X.; Megahed, S.; Feliu, N.; Prosperi, D.; Parak, W. J.; Colombo, M. Functionalization of Colloidal Nanoparticles with a Discrete Number of Ligands Based on a “HALO-Bioclick” Reaction. *Chem. Commun. Camb. Engl.* **2020**, *56* (77), 11398–11401. <https://doi.org/10.1039/d0cc04355a>.
- (39) Liu, D. S.; Phipps, W. S.; Loh, K. H.; Howarth, M.; Ting, A. Y. Quantum Dot Targeting with Lipoic Acid Ligase and HaloTag for Single-Molecule Imaging on Living Cells. *ACS Nano* **2012**, *6* (12), 11080–11087. <https://doi.org/10.1021/nn304793z>.
- (40) Hakariya, H.; Takashima, I.; Takemoto, M.; Noda, N.; Sato, S.-I.; Uesugi, M. Non-Genetic Cell-Surface Modification with a Self-Assembling Molecular Glue. *Chem. Commun. Camb. Engl.* **2021**, *57* (12), 1470–1473. <https://doi.org/10.1039/d0cc07171d>.
- (41) Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigsmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa, I. R.; Luo, Z.-G.; Schultz, C.; Lemke, E. A.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K. A Near-Infrared Fluorophore for Live-Cell Super-Resolution Microscopy of Cellular Proteins. *Nat. Chem.* **2013**, *5* (2), 132–139. <https://doi.org/10.1038/nchem.1546>.
- (42) Clark, S. A.; Singh, V.; Vega Mendoza, D.; Margolin, W.; Kool, E. T. Light-Up “Channel Dyes” for Haloalkane-Based Protein Labeling in Vitro and in Bacterial Cells. *Bioconjug. Chem.* **2016**, *27* (12), 2839–2843. <https://doi.org/10.1021/acs.bioconjchem.6b00613>.
- (43) Dockalova, V.; Sanchez-Carnerero, E. M.; Dunajova, Z.; Palao, E.; Slanska, M.; Buryska, T.; Damborsky, J.; Klán, P.; Prokop, Z. Fluorescent Substrates for Haloalkane Dehalogenases: Novel Probes for Mechanistic Studies and Protein Labeling. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 922–932. <https://doi.org/10.1016/j.csbj.2020.03.029>.

- (44) Encell, L. P.; Friedman Ohana, R.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Wood, M. G.; Los, G. V.; McDougall, M. G.; Zimprich, C.; Karassina, N.; Learish, R. D.; Hurst, R.; Hartnett, J.; Wheeler, S.; Stecha, P.; English, J.; Zhao, K.; Mendez, J.; Benink, H. A.; Murphy, N.; Daniels, D. L.; Slater, M. R.; Urh, M.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. Development of a Dehalogenase-Based Protein Fusion Tag Capable of Rapid, Selective and Covalent Attachment to Customizable Ligands. *Curr. Chem. Genomics* **2012**, *6*, 55–71. <https://doi.org/10.2174/1875397301206010055>.
- (45) Kaushik, S.; Prokop, Z.; Damborsky, J.; Chaloupkova, R. Kinetics of Binding of Fluorescent Ligands to Enzymes with Engineered Access Tunnels. *FEBS J.* **2017**, *284* (1), 134–148. <https://doi.org/10.1111/febs.13957>.
- (46) Kaushik, S.; Marques, S. M.; Khirsariya, P.; Paruch, K.; Libichova, L.; Brezovsky, J.; Prokop, Z.; Chaloupkova, R.; Damborsky, J. Impact of the Access Tunnel Engineering on Catalysis Is Strictly Ligand-Specific. *FEBS J.* **2018**, *285* (8), 1456–1476. <https://doi.org/10.1111/febs.14418>.
- (47) Vanacek, P.; Sebestova, E.; Babkova, P.; Bidmanova, S.; Daniel, L.; Dvorak, P.; Stepankova, V.; Chaloupkova, R.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Exploration of Enzyme Diversity by Integrating Bioinformatics with Expression Analysis and Biochemical Characterization. *ACS Catal.* **2018**, *8* (3), 2402–2412. <https://doi.org/10.1021/acscatal.7b03523>.
- (48) Vasina, M.; Vanacek, P.; Hon, J.; Kovar, D.; Faldynova, H.; Kunka, A.; Badenhorst, C.; Buryska, T.; Mazurenko, S.; Bednar, D.; Stavros, S.; Bornscheuer, U.; deMello, A.; Damborsky, J.; Prokop, Z. Functional and Mechanistic Characterization of an Enzyme Family Combining Bioinformatics and High-Throughput Microfluidics. *Nature Portfolio*. In review November 15, 2021. <https://doi.org/10.21203/rs.3.rs-1027271/v1>.
- (49) Jurcik, A.; Bednar, D.; Byska, J.; Marques, S. M.; Furmanova, K.; Daniel, L.; Kokkonen, P.; Brezovsky, J.; Strnad, O.; Stourac, J.; Pavelka, A.; Manak, M.; Damborsky, J.; Kozlikova, B. CAVER Analyst 2.0: Analysis and Visualization of Channels and Tunnels in Protein Structures and Molecular Dynamics Trajectories. *Bioinformatics* **2018**, *34* (20), 3586–3588. <https://doi.org/10.1093/bioinformatics/bty386>.
- (50) Chovancová, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8* (10), e1002708. <https://doi.org/10.1371/journal.pcbi.1002708>.
- (51) *The PyMOL Molecular Graphics System, Version 2.3.2*; Schrödinger, LLC, 2019.
- (52) Verschueren, K. H.; Seljée, F.; Rozeboom, H. J.; Kalk, K. H.; Dijkstra, B. W. Crystallographic Analysis of the Catalytic Mechanism of Haloalkane Dehalogenase. *Nature* **1993**, *363* (6431), 693–698. <https://doi.org/10.1038/363693a0>.
- (53) Johnson, K. A.; Simpson, Z. B.; Blom, T. Global Kinetic Explorer: A New Computer Program for Dynamic Simulation and Fitting of Kinetic Data. *Anal. Biochem.* **2009**, *387* (1), 20–29. <https://doi.org/10.1016/j.ab.2008.12.024>.
- (54) Johnson, K. A.; Simpson, Z. B.; Blom, T. FitSpace Explorer: An Algorithm to Evaluate Multidimensional Parameter Space in Fitting Kinetic Data. *Anal. Biochem.* **2009**, *387* (1), 30–41. <https://doi.org/10.1016/j.ab.2008.12.025>.
- (55) Kellogg, E. H.; Leaver-Fay, A.; Baker, D. Role of Conformational Sampling in Computing Mutation-Induced Changes in Protein Structure and Stability. *Proteins* **2011**, *79* (3), 830–838. <https://doi.org/10.1002/prot.22921>.

- (56) Marques, S. M.; Bednar, D.; Damborsky, J. Computational Study of Protein-Ligand Unbinding for Enzyme Engineering. *Front. Chem.* **2019**, *6*, 650. <https://doi.org/10.3389/fchem.2018.00650>.
- (57) Bruce, N. J.; Ganotra, G. K.; Kokh, D. B.; Sadiq, S. K.; Wade, R. C. New Approaches for Computing Ligand-Receptor Binding Kinetics. *Curr. Opin. Struct. Biol.* **2018**, *49*, 1–10. <https://doi.org/10.1016/j.sbi.2017.10.001>.
- (58) Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kutý, M.; Kuta-Smatanova, I.; Dohnalek, J.; Kulhanek, P.; Wade, R. C.; Damborsky, J. Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations. *J. Mol. Biol.* **2009**, *392* (5), 1339–1356. <https://doi.org/10.1016/j.jmb.2009.06.076>.
- (59) Brezovsky, J.; Babkova, P.; Degtjarik, O.; Fortova, A.; Gora, A.; Iermak, I.; Rezacova, P.; Dvorak, P.; Smatanova, I. K.; Prokop, Z.; Chaloupkova, R.; Damborsky, J. Engineering a de Novo Transport Tunnel. *ACS Catal.* **2016**, *6* (11), 7597–7610. <https://doi.org/10.1021/acscatal.6b02081>.
- (60) Kokkonen, P.; Sykora, J.; Prokop, Z.; Ghose, A.; Bednar, D.; Amaro, M.; Beerens, K.; Bidmanova, S.; Slanska, M.; Brezovsky, J.; Damborsky, J.; Hof, M. Molecular Gating of an Engineered Enzyme Captured in Real Time. *J. Am. Chem. Soc.* **2018**, *140* (51), 17999–18008. <https://doi.org/10.1021/jacs.8b09848>.
- (61) Marques, S. M.; Dunajova, Z.; Prokop, Z.; Chaloupkova, R.; Brezovsky, J.; Damborsky, J. Catalytic Cycle of Haloalkane Dehalogenases Toward Unnatural Substrates Explored by Computational Modeling. *J. Chem. Inf. Model.* **2017**, *57* (8), 1970–1989. <https://doi.org/10.1021/acs.jcim.7b00070>.
- (62) Zhang, H.; Yin, C.; Jiang, Y.; van der Spoel, D. Force Field Benchmark of Amino Acids: I. Hydration and Diffusion in Different Water Models. *J. Chem. Inf. Model.* **2018**, *58* (5), 1037–1052. <https://doi.org/10.1021/acs.jcim.8b00026>.