ACS Catalysis

pubs.acs.org/acscatalysis Research Article

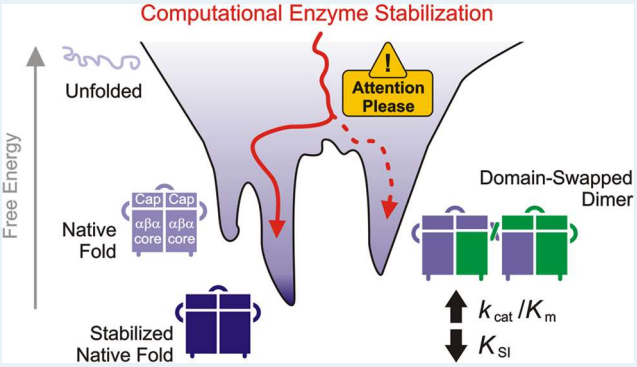
# Computational Enzyme Stabilization Can Affect Folding Energy Landscapes and Lead to Catalytically Enhanced Domain-Swapped Dimers

Klara Markova, Antonin Kunka, Klaudia Chmelova, Martin Havlasek, Petra Babkova, Sérgio M. Marques, Michal Vasina, Joan Planas-Iglesias, Radka Chaloupkova, David Bednar, Zbynek Prokop, Jiri Damborsky,* and Martin Marek*

Cite This: *ACS Catal.* 2021, 11, 12864−12885 Read Online



**ABSTRACT:** The functionality of an enzyme depends on its unique three-dimensional structure, which is a result of the folding process when the nascent polypeptide follows a funnel-like energy landscape to reach a global energy minimum. Computer-encoded algorithms are increasingly employed to stabilize native proteins for use in research and biotechnology applications. Here, we reveal a unique example where the computational stabilization of a monomeric $\alpha/\beta$-hydrolase enzyme ($T_m$ = 73.5 °C; $\Delta T_m$ > 23 °C) affected the protein folding energy landscape. The introduction of eleven single-point stabilizing mutations based on force field calculations and evolutionary analysis yielded soluble domain-swapped intermediates trapped in local energy minima. Crystallographic structures revealed that these stabilizing mutations might (i) activate cryptic hinge-loop regions and (ii) establish secondary interfaces, where they make extensive noncovalent interactions between the intertwined protomers. The existence of domain-swapped dimers in a solution is further confirmed experimentally by data obtained from small-angle X-ray scattering (SAXS) and cross-linking mass spectrometry. Unfolding experiments showed that the domain-swapped dimers can be irreversibly converted into native-like monomers, suggesting that the domain swapping occurs exclusively *in vivo*. Crucially, the swapped-dimers exhibited advantageous catalytic properties such as an increased catalytic rate and elimination of substrate inhibition. These findings provide additional enzyme engineering avenues for next-generation biocatalysts.




**KEYWORDS:** *protein folding, protein design, α/β-hydrolase, haloalkane dehalogenase, domain swapping, energy landscape, oligomerization, catalytic efficiency, substrate inhibition*



## ■ INTRODUCTION

Protein folding is a process in which a polypeptide chain folds into its native state, a well-defined three-dimensional structure, which is a prerequisite for its proper biological function. During this process, a polypeptide chain undergoes many conformational changes as it navigates through the partially folded states down the energy landscape to the free energy minimum—a native state.[1] The folding energy landscape of a protein is encoded by its amino acid sequence. Consequently, any modification of the protein's sequence may alter its folding pathway. Energy landscapes of natural proteins are shaped by the evolution toward functional structures. Protein engineering aims to modulate the protein sequence so that the native state is able to perform its function even under unnatural conditions. Computationally designed substitutions are typically introduced to create favorable stabilizing interactions in the native state.[2−4] Despite numerous, and often remarkably successful, attempts at stabilization,[3,5,6] the associated changes to the folding energy landscape have not been investigated, to the best of our knowledge.

Domain swapping[7,8] is a specific form of protein oligomerization in which small elements or domains are exchanged among identical polypeptide chains.[9,10] Domain swapping can result in dimers, closed oligomers, or filament-like oligomers.[8] The mechanism of domain swapping requires either nearly complete unfolding[11] or opening of the protein into a partially unfolded state, allowing the exchange of the structural elements between two or more identical chains.[12] Most interactions that are crucial for stabilizing a monomeric

Received: July 26, 2021
Revised: September 22, 2021

   https://doi.org/10.1021/acscatal.1c03343
*ACS Catal.* 2021, 11, 12864−12885

form are also preserved in the domain-swapped structure. Additionally, new interactions can be formed in the domain-swapped structures and create a secondary interface.[13] A hinge region is a protein segment that adopts different backbone conformations in the monomer and domain-swapped oligomer, connecting the swapped and the nonswapped domains of the single polypeptide chain.[14] The hinge region acts as a conformational switch, with its length and sequence composition being critical for the domain-swapping propensity.[10] Engineering of putative hinge loops is an attractive strategy for *de novo* domain-swapping design.[15,16]

Domain swapping can be biologically relevant as a regulatory mechanism, for example, in receptor binding proteins[17] or DNA-binding proteins.[18,19] It has been reported as an effective strategy for protein function modification *in vitro*,[20] and an analogous mechanism was utilized in the successful design of the ligand-triggered functional switch of staphylococcal nuclease and ribose binding protein.[21] Domain swapping was also proposed as one of the evolutionary mechanisms of protein oligomerization.[22] A possible link between domain swapping and protein deposition diseases has been implied.[23] Crystal structures of the human prion protein and cystatin C, involved in the prion and amyloid diseases, respectively, revealed the formation of dimers through a domain-swapping mechanism.[24,25] Structural and functional studies of domain swapping are essential for our understanding of this phenomenon.

In this study, we show that the computationally driven stabilization of a monomeric haloalkane dehalogenase DhaA115 ($T_m^{app}$ = 73.5 °C; $\Delta T_m^{app}$ > 23 °C)[5,26] coincidentally altered the protein folding landscape, resulting in the formation of stable domain-swapped intermediates. Our structural findings reveal that the intended stabilizing mutations were frequently found in the cryptic hinge regions and introduced secondary interfaces where they made new noncovalent interactions between the misfolded, intertwined polypeptide chains. We further demonstrate that the domain-swapped dimers can be irreversibly converted into native-like monomers using thermal or chemical denaturation, suggesting that the domain swapping occurs exclusively *in vivo*. Collectively, our results highlight an unprecedented example when computationally guided enzyme stabilization is unexpectedly accompanied by the in-cell formation of catalytically enhanced domain-swapped intermediates.

## ■ RESULTS

***De Novo* Oligomerization Due to Computer-Aided Stabilization.** We have already reported that our previous computer-aided engineering of a haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* was unexpectedly accompanied by *de novo* formation of enzyme oligomers.[26,27] While the DhaA wild-type enzyme ($T_m^{app}$ = ∼50.4 °C) exists solely in a monomeric form, the most stabilized 11-point mutant DhaA115 ($T_m^{app}$ = ∼73.4 °C) exists in monomeric (∼75.7%), dimeric (∼20.3%), and higher oligomeric (∼4.0%) forms (Figure 1 and Table S1). A single band of ∼35 kDa can be detected using electrophoresis in denaturing conditions for both wild-type DhaA and DhaA115, while native non-denaturing electrophoretic analysis revealed multiple distinct oligomeric states of hyperstable DhaA115 (Figure 1A). Notably, the gel-separated fractions corresponding to the putative dimeric form of DhaA115 repeatedly migrated as a specific double band, suggesting the existence of different conformers. Analytical size-exclusion chromatography (SEC) and analytical ultracentrifugation (AUC) of DhaA115 confirmed the presence of a monomeric, dimeric, and a small amount of nonspecified oligomers of higher molecular weights (Figure 1B,C). Further SEC analysis at increased protein concentrations showed that the oligomeric content was independent of protein concentration (Table S1).

Next, we carried out circular dichroism (CD) spectroscopy on SEC-purified protein fractions to check the effect of DhaA115 oligomerization on protein folding. The shapes of the CD spectra of all measured species showed characteristics typical of an $\alpha/\beta$-hydrolase fold, specifically one positive peak at 195−197 nm and two negative minima at approximately 209 and 224 nm. CD spectra of DhaA115 monomeric and dimeric forms were highly similar, indicating that the secondary structure of the DhaA115 monomer was preserved in its dimeric form (Figure 1D). Finally, differential scanning fluorimetry experiments identified similar melting transitions with the midpoint at 74 °C, preceded by an additional minor transition with a midpoint around 53.5 °C observed in the case of the DhaA115 dimer (Figure 1E).

**DhaA115 Oligomerizes through the Domain-Swapping Mechanism.** To gain structural insights into the mechanism of DhaA115 oligomerization, we attempted crystallization of its oligomeric forms. Despite intensive efforts, we were not able to obtain well-diffracting crystals of the Dha115 higher oligomeric species. In contrast, the crystallization of the DhaA115 dimeric form was successful and yielded two types of well-diffracting crystals. The first belonged to the space group $P2_12_12_1$, while the second crystal form belonged to the $P12_11$ space group (Table S2). The structures were solved by molecular replacement, and the initial models were further refined with several cycles of manual building and automatic refinement, yielding structural models with good deviations from ideal geometry (Table S2). Most of the residues could be built in the electron density maps, except for a few residues on N-terminal and C-terminal ends.

Surprisingly, careful inspection of the electron density maps unambiguously revealed two intertwined polypeptide chains (protomers), which formed dimers through a domain-swapping mechanism (Figure 2). Importantly, the DhaA115 domain-swapped dimer structure solved in the space group $P2_12_12_1$, hereafter referred to as DhaA115-DSD1, structurally differed from that found in the second type of crystals processed in space group $P12_11$, which is further referred to as DhaA115-DSD2. Both DhaA115 domain-swapped dimers differ from each other in the position of the hinge loop, the only region of the protein that adopts a different structure in the monomeric and domain-swapped dimeric structures.

The dimeric structure of DhaA115-DSD1 is mediated by the altered conformation of the L9 loop, which connects the $\beta6$ strand with the cap domain-forming $\alpha4$ helix (Figure 2A). Specifically, the $W_{141}$PEFA element of the L9 loop functions as the hinge region in DhaA115-DSD1. As a consequence, the $\alpha4$ helix is slightly deflected (by an angle of ∼10°) when compared to its monomeric counterpart. The structure of DhaA115-DSD1 further reveals a new secondary interface, a site where the two intertwined protomers extensively interact with each other and thus stabilize the domain-swapped dimer (Figure 2A). The secondary interface has a predominantly hydrophobic and aromatic character (F144, L148, I172, and F176), although polar contacts and hydrogen bonding are also involved (K175, V177, P249, and A250). The secondary


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

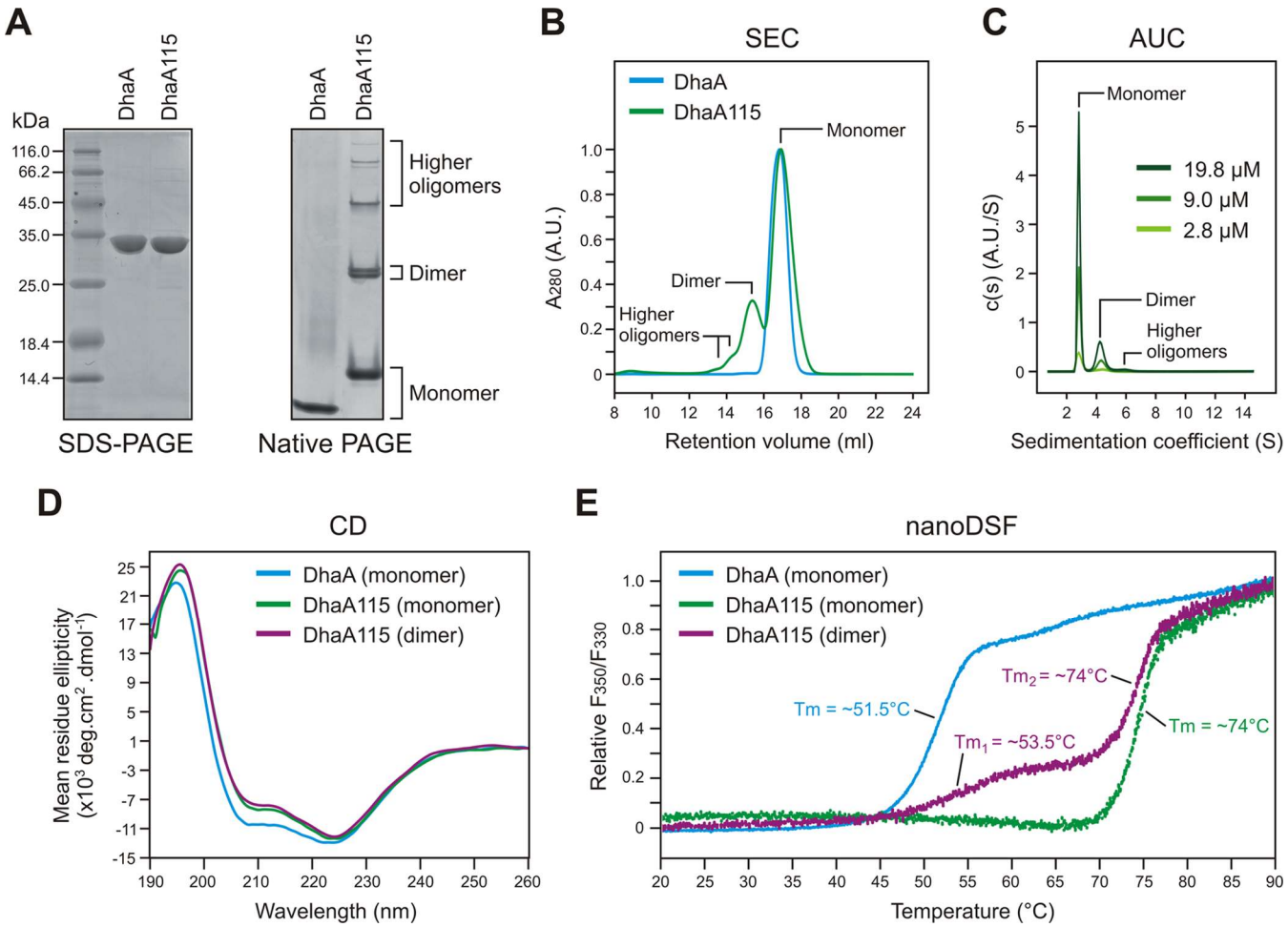



**Figure 1.** Biophysical and biochemical characterization of the hyperstable engineered DhaA115. (A) Electrophoretic separation of DhaA and DhaA115 proteins by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (left panel) and native nondenaturing PAGE (right panel). (B) Analysis of DhaA and DhaA115 by analytical size-exclusion chromatography (SEC). (C) Analysis of DhaA115 by analytical ultracentrifugation (AUC). (D) Circular dichroism (CD) spectra of DhaA monomers, DhaA115 monomers, and DhaA115 dimers. (E) Differential scanning fluorimetry (nanoDSF) of DhaA monomers, DhaA115 monomers, and DhaA115 dimers. Note that the melting of DhaA115 dimers is accompanied by two major melting points: $T_{m1}$ = 53.5 °C and $T_{m2}$ = 74 °C.

interface shapes a pore with a diameter of ∼4 to 5 Å between the intertwined protomers, where an unambiguous linear electron density occupying this pore is present (Figure 2A). We interpreted this electron density as a poly(ethylene glycol) molecule, which was one of the components of the crystallization buffer. Interestingly, three residues introduced into DhaA115 by the computational protein design as stabilizing mutations (L148, I172, and F176)[5] are involved in the formation of this secondary interface.

A structurally distinct domain-swapping topology was observed in the crystal structure of DhaA115-DSD2, where the asymmetric unit consisted of two domain-swapped dimers. Domain swapping in DhaA115-DSD2 is mediated by the conformational change of the L13 loop (residues $L_{194}$KPVW). The L13 loop connects $\alpha$6 and $\alpha$7 helices, which are part of the cap domain (Figure 2B). Interestingly, the DhaA115-DSD2 dimer does not contain an extensive stabilizing secondary interface, as observed in the case of DhaA115-DSD1. At the DhaA115-DSD2 secondary interface, two opposite arginine residues, R21, from the individual protomers form an arginine pair, whose repulsive nature is stabilized by hydrogen bonding with the negatively charged D73 and bound water molecules (Figure 2B). The serine residue S20 designed by computational stabilization[5] is present in the vicinity of the DhaA115-DSD2 secondary interface, although it does not seem to be directly participating in its formation.

The other molecular interactions between the two protomers in DhaA115-DSD2 are facilitated by the hinge region, which allows intrinsic conformational flexibility of the dimeric structure. This conformational freedom is apparent from the superimposition of the two noncrystallographic DhaA115-DSD2 dimers found in the asymmetric unit (Figure S1). Importantly, the hinge region in DhaA115-DSD2 is composed of the sequence element $L_{194}$KPVW, where the pyrrolidine rings of the two opposing prolines, P196, from each protomer are arranged around the noncrystallographic two-fold axis (Figure 2B). These hinge regions interact with each other through multiple nonpolar and hydrophobic contacts, thus forming the two-loop bundle that bridges the two pseudo-$\alpha/\beta$-hydrolase folds. Importantly, one of the stabilizing mutations, D198W, is located in the DhaA115-DSD2 hinge region.[5]

**DhaA115 Domain-Swapped Dimers Exist in Solution.** We speculated as to whether the domain-swapped DhaA115 dimers could be a crystallographic artifact. We initially probed SEC-separated dimeric fractions of the DhaA115 enzyme using


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

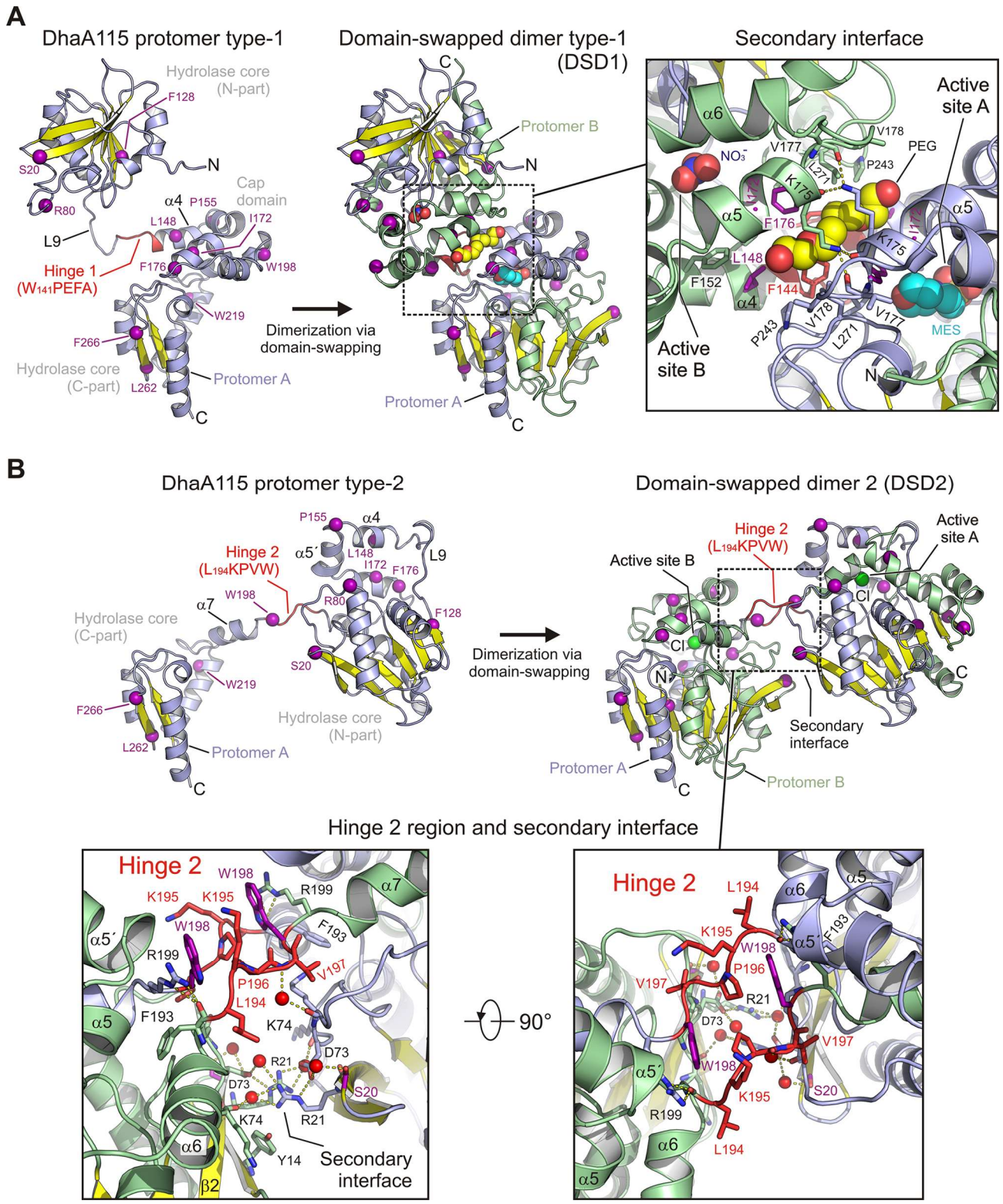
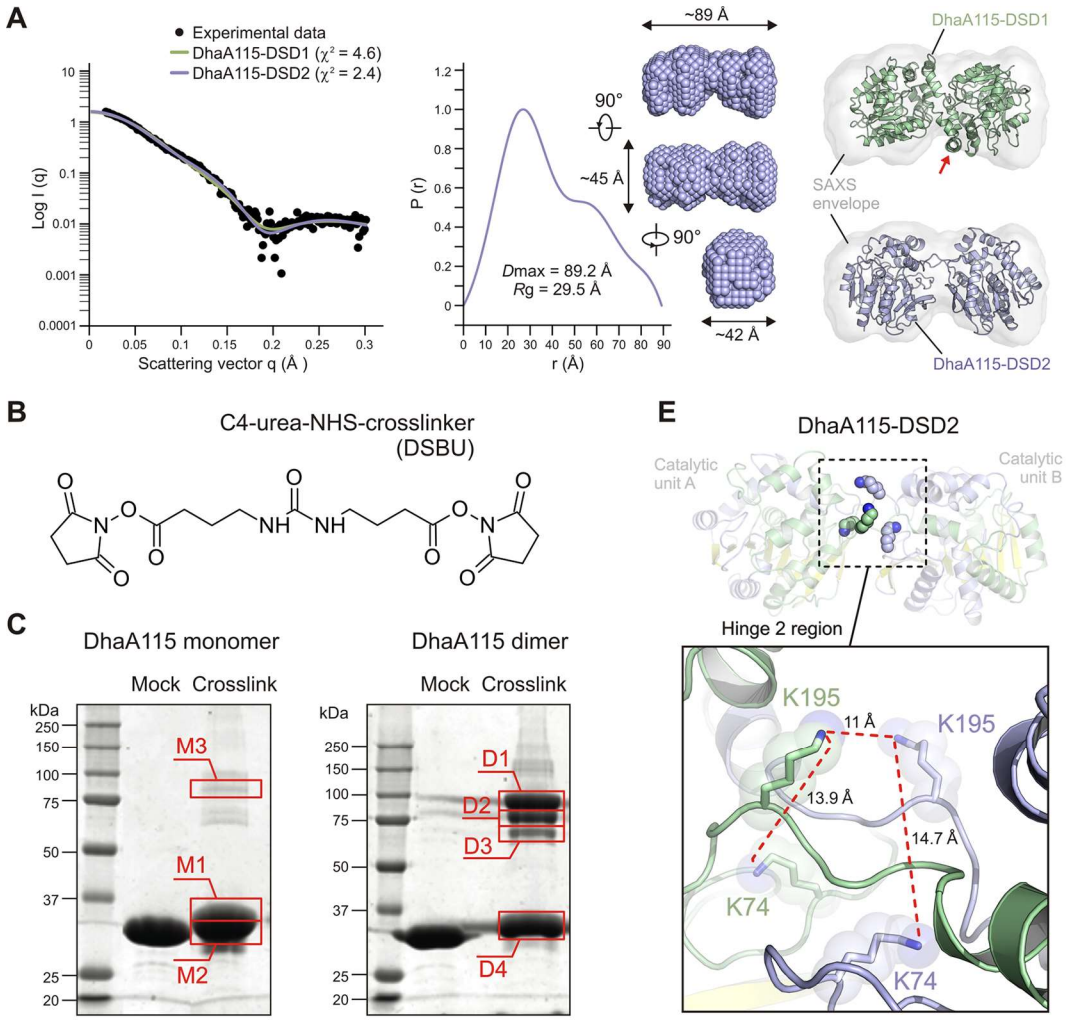



**Figure 2.** X-ray structures of domain-swapped dimers DhaA115-DSD1 and DhaA115-DSD2. (A) Cartoon representations of DhaA115 protomer type-1 (left panel), the DhaA115-DSD1 dimer (middle panel), and close-up view of the new secondary interface (right panel). The central eight-stranded $\beta$-sheet (yellow), the helices (light blue), and the hinge region 1 ($W_{141}$PEFA) sequence (red). The designed stabilizing mutations are shown as purple spheres; the poly(ethylene glycol) (PEG) molecule bound between the two catalytic units is shown as yellow spheres; the 2-(*N*-morpholino)ethanesulfonic acid (MES) bound in the active site A is shown as cyan spheres; and the nitrate ($NO^{3-}$) anion bound in the active site B is shown as blue spheres. The key residues involved in the secondary interface are shown as sticks; hydrogen bonds are shown as yellow dashed lines. (B) Cartoon representations of DhaA115 protomer type-2 (left panel), the DhaA115-DSD2 dimer (right panel), and close-up views of the hinge 2 region and the new secondary interface (bottom panels). Color-coding is the same as in A panels. The hinge 2 region ($K_{194}$PVW) sequence (red); and the chloride ($Cl^-$) anions bound in the active sites are shown as green spheres.


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885




D

| Sample | Merox score | Site 1 | Peptide 1 | Site 2 | Peptide 2 | Probability (%) | No. matching spectra |
|---|---|---|---|---|---|---|---|
| M1 | – | – | – | – | – | – | – |
| M2 | 171 | 195 | EPFL**K**PVWR | 74 | SD**K**PDLDYR | 80 | 4 |
| M3 | – | – | – | – | – | – | – |
| **D1** | 218 | 195 | EPFL**K**PVWR | 195 | EPFL**K**PVWR | 100 | 7 |
| **D2** | 206 | 195 | EPFL**K**PVWR | 195 | EPFL**K**PVWR | 100 | 7 |
| | 187 | 195 | EPFL**K**PVWR | 74 | SD**K**PDLDYR/BIAPDLI GMGKSD**K**PDLDYR | 76 | 13 |
| D3 | 207 | 195 | EPFL**K**PVWR | 195 | EPFL**K**PVWR | 100 | 3 |
| | 115 | 195 | EPFL**K**PVWR | 74 | SD**K**PDLDYR | 77 | 2 |
| D4 | – | – | – | – | – | – | – |

**Figure 3.** Characterization of the DhaA115 dimers by small-angle X-ray scattering (SAXS) and cross-linking coupled mass spectrometry (XL-MS). (A) Solution structure of DhaA115 dimers determined by SAXS. The experimental SAXS scattering curve for the DhaA115 dimer (black dots) is shown against the calculated scattering curves derived from the DhaA115-DSD1 (green line) and DhaA115-DSD2 (blue line) crystal structures (left panel). The distance distribution function of the DhaA115 dimer was computed from the X-ray scattering pattern using the GNOM program (middle panel). The *Ab initio* molecular envelope generated from SAXS data is shown in semitransparent gray color and superposed on the DhaA115-DSD1 (green) and DhaA115-DSD2 (blue) crystal structures (right panels). (B) Structural formula of the C4-urea-NHS-cross-linker (DSBU). (C) SDS-PAGE separation of cross-linker-treated DhaA115 monomers (left panel) and DhaA115 dimers (right panel). Red frames (M1-3 and D1-4) depict protein specimens that were excised, processed, and analyzed by liquid chromatography with tandem mass spectrometry (LC-MS/MS). (D) Table of cross-linked peptides identified in separated fractions (M1-3 and D1-4) by LC-MS/MS analysis. MeroX software detected homodipeptides (EPFLK$_{195}$PVWR) only in the samples D1 and D2 that originated from DhaA115 dimers. The lysine residues that were covalently cross-linked are in bold. (E) Cartoon representation of the DhaA115-DSD2 structure with shown lysine residues (K74 and K195) present in the hinge region and its vicinity. Lysine-to-lysine covalent links mediated by the DSBU reagent are shown as red dashed lines.


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

a small-angle X-ray scattering (SAXS) technique. The SAXS profile (Figure 3A) fits well to the scattering profile calculated using both crystallographic dimers, although the fit for the DhaA115-DSD2 structure ($\chi^2$ = 2.40) is slightly better than that for DhaA115-DSD1 ($\chi^2$ = 4.62). The radius of gyration ($R_g$) obtained from the SAXS data was ∼29.5 Å. The profile of the pair distance distribution function has a double bell-like shape with the main peak at ∼27.2 Å, and trails off to the maximum dimension ($D_{max}$) of ∼89.2 Å. The *ab initio* model reconstructed from the SAXS data perfectly matches the DhaA115-DSD2 dimer, while there are observable discrepancies between the SAXS envelope and the DhaA115-DSD1 crystal structure (Figure 3A). We also calculated the $R_g$ values for the molecular dynamics (MD) simulations: $R_g$ = 17.83 ± 0.04 Å for the monomeric DhaA115, $R_g$ = 26.89 ± 0.11 Å for DhaA115-DSD1, and $R_g$ = 28.61 ± 0.16 Å for DhaA115-DSD2. Notably, the latter value for DhaA115-DSD2 is very close to the SAXS-determined value for this dimer (∼29.5 Å).

We further employed chemical cross-linking coupled with mass spectrometry (XL-MS) to capture domain-swapped dimers in a solution. The cross-linking experiment for the SEC-purified DhaA115 monomers with a cleavable 11-atom cross-linker (4,4′-ureylenedibutyric acid bis-NHS ester) produced complexes that contained intramolecular heterodipeptides cross-linked at lysines K74 and K195 (Figure 3B−D). In contrast, two major complexes were detected in the same experiment for the dimeric fraction of DhaA115. The first one was the same intramolecular heterodipeptide between the lysine residues K74 and K195 as observed in the DhaA115 monomeric fraction (Figure 3B−D). The second covalently cross-linked complex, exclusively detected in the dimeric fraction of DhaA115, contained intermolecular homodipeptides formed between lysines K195 of the two individual polypeptide chains. The distance between the lysines K74 and K195 in the crystal structures of both monomeric and dimeric forms of DhaA115 was 12−14 Å, which corresponds to the length of the cross-linker (12.5 Å) used in the experiment. This explains why the intramolecular heterodipeptide created between these lysines was detected in both monomeric and dimeric fractions when cross-linked.

By contrast, the intermolecular homodipeptides linked through lysine K195 were detected solely in the dimeric fraction (Figure 3B−D). Careful inspection of the crystal structures revealed that the cross-linked homodipeptides detected most likely originated from the DhaA115-DSD2 since lysine K195 is located in the hinge region of this conformer. As a result, lysines K195 from the individual protomers move close to this region and are ∼11 Å apart (Figure 3E). In contrast, the distance between these identical residues in the two protomers of DhaA115-DSD1 is 55 Å, too far for a successful cross-linking with the 11-atom cross-linker. Collectively, the results of SAXS and XL-MS experiments prove that DhaA115 domain-swapped dimers exist in solution and imply that the DhaA115-DSD2 is the dominant species.

**MD Simulations Imply Higher Flexibility of the Domain-Swapped Dimers.** Next, we used MD simulations to study the conformational dynamics of the domain-swapped dimers in the solution. Multiple MD simulations of 100 ns length were carried out with the two dimeric structures, and their flexibility was analyzed using root-mean-square deviation (RMSD) of the atomic positions. These simulations were reasonably stable after 100 ns, as assessed by the plateaus observed in the respective RMSD plots near the end of each MD (Figure S2). Interestingly, the two domain-swapped dimers deviated from their respective crystal structures considerably with RMSD fluctuations above 2−3 Å, in contrast to the relatively rigid structures of monomeric DhaA115 and wild-type DhaA (RMSD below 1 Å).[26]

To investigate the main reason for such large deviations, we clustered those trajectories by the RMSD and analyzed the centroid structures (Figures S3 and S4). We found that the structural changes mainly consisted of the repositioning of the two globular units with respect to each other, while each catalytic unit remained very stable and superimposable with itself (Figures S3C and S4C). The large topological changes observed during the MD simulations are caused by the conformational freedom of the hinge regions of the domain-swapped dimers. Domain-swapped dimers of DhaA115 can adopt spatially different conformations in the crystal state and in solution, which is most likely influenced by (i) the length and flexibility of the hinge region, (ii) nature of the secondary interface, and possibly (iii) buffer composition. For instance, in the MD simulations of DhaA115-DSD1, the most populated state corresponds to the cluster with the largest deviation (cluster 1, RMSD 2.75 ± 0.20 Å), which was also the dominant conformation by the end of both MD simulations. Conversely, the most populated conformation during MD simulation of DhaA115-DSD2 was the one closest to the crystal structure (cluster 1, RMSD 1.24 ± 0.20 Å), and the most deviated one was populated the least (cluster 4, RMSD 1.82 ± 0.22 Å).

The B-factors of the protein backbone were calculated as a measure of the respective residue flexibility. The B-factors of the two dimers are higher than for the monomer and greater for DhaA115-DSD1 than for DhaA115-DSD2 (Figure S5A). This is in agreement with the RMSD results described above. However, we noted that most of these fluctuations were due to the rocking and tilting movements of the two catalytic units with respect to one another. Therefore, we carried out a stepwise analysis of the MDs by aligning the systems to each catalytic unit at the time, calculating the respective B-factors, and then combining them. These new results (Figures S5B and S6) showed that DhaA115-DSD1 is still the most flexible of the DhaA115 variants tested here. DhaA115-DSD2 showed B-factors similar to those observed for the monomeric DhaA115 when considering the individual catalytic units and even slightly lower for some residues.

**Domain-Swapped Dimer Opens and Collapses to Monomers during Thermal Unfolding.** Temperature denaturation of domain-swapped dimers of DhaA115 was monitored by several biophysical techniques and analyzed globally to resolve the unfolding mechanism. The unfolding of wild-type monomeric DhaA proceeds through one or more intermediates (unpublished data). Refolding of DhaA from a heat-denatured state leads to a mixture of a native-like state and an ensemble of aggregated states of various sizes. We hypothesized that the domain swapping of DhaA115 proceeds with a similar intermediate formation during refolding, while the introduced mutations stabilize local minima on the protein folding landscape. To test this hypothesis, we carried out unfolding and refolding experiments with monomeric and dimeric fractions of DhaA115 using circular dichroism (CD) and fluorescence spectroscopy in combination with differential scanning calorimetry (DSC). The monomer unfolding shows a single transition with the midpoint at ∼73 °C when monitored with spectroscopic techniques (Figure S7) and can be fitted


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

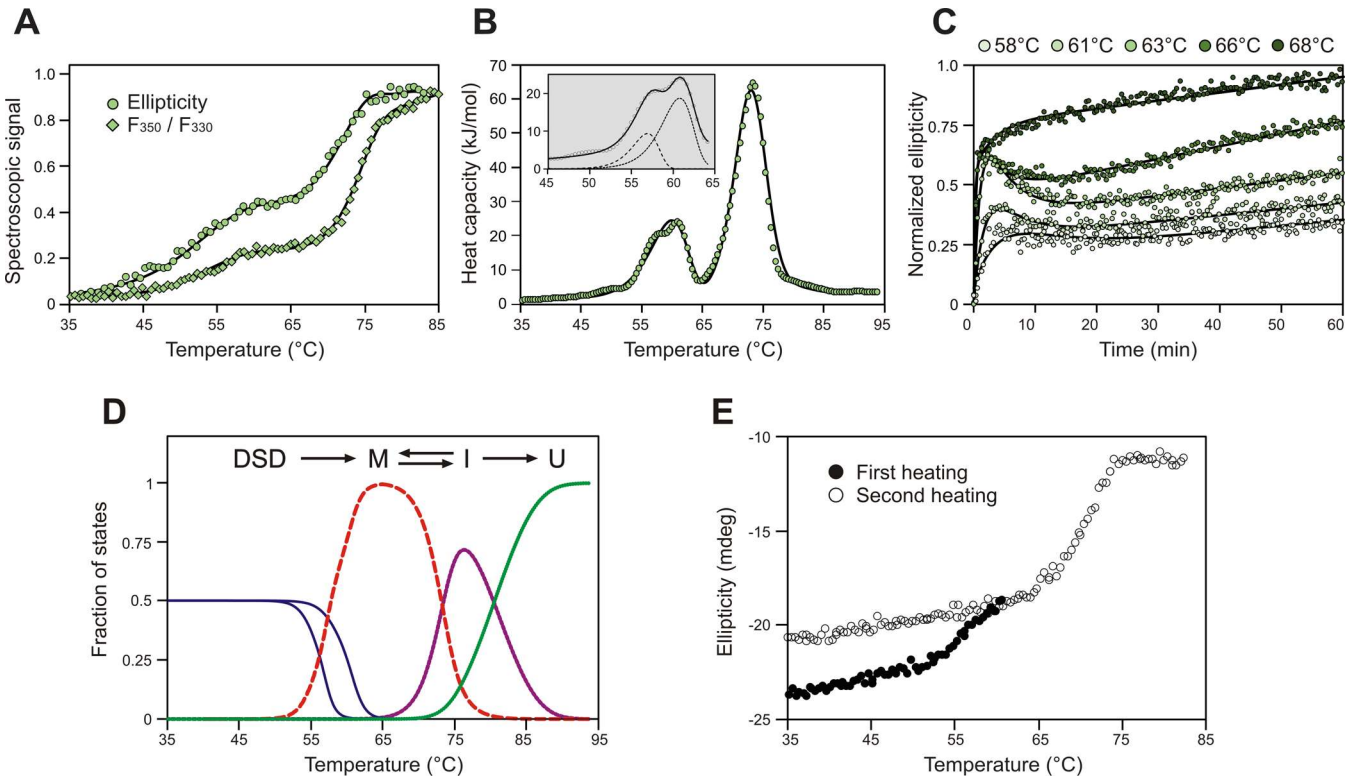



**Figure 4.** Global analysis of DhaA115 dimer unfolding. (A) Temperature scanning experiments were carried out at a 1 °C·min$^{-1}$ scan rate followed by CD (circles) and fluorescence (diamonds) spectroscopies. Ellipticity at 224 nm and the ratio of fluorescence intensities at 350 and 330 nm are plotted against the temperature. (B) Differential scanning calorimetry thermograph. The inset presents deconvolution of the individual dimer dissociation. (C) Unfolding kinetics measured by monitoring changes of ellipticity at 224 nm at different temperatures. (D) Fraction of states calculated from the global fitting (black lines in all graphs) as a function of temperature: DSD, domain-swapped dimers (blue); M, monomer (red); I, intermediate (purple); and U, unfolded state (green). (E) Reheating experiment of DhaA115 dimers measured by monitoring changes in ellipticity at 224 nm. The sample was first heated to 60 °C (black circles), then cooled down to 25 °C, and finally heated to 80 °C (open circles) at a 1 °C·min$^{-1}$ scan rate.

together globally with calorimetric and unfolding kinetics data to the three-state partially reversible mechanism (Lumry–Eyring model).[28,29] Refolding was carried out by heating the monomer at various concentrations to different temperatures, to and above the unfolding transition range, and cooling it down at the same rate (1 °C·min$^{-1}$). The resulting oligomeric states of the sample were then analyzed using SEC and native PAGE. As seen with the wild-type, refolding of DhaA115 yielded a mixture of monomeric protein and aggregates (Figure S8A). The fraction of the aggregated states was concentration-dependent and increased with the temperature to which the sample was heated. Variations in resident time for which the protein was kept at an elevated temperature, in protein concentration or in the speed of heating and cooling did not induce dimer formation (Figure S9).

The denaturation curves for the domain-swapped dimers show two major transitions (Figure 4A,B). The first unfolding transition occurs between 45 and 65 °C and is associated with a minor loss of helicity and partial exposure of hydrophobic residues as indicated by a red shift in the fluorescence spectrum. The corresponding heat capacity peak begins near the midpoint of the spectral denaturation curves at 52 °C and ends at 65 °C, where spectroscopic signals plateau. Interestingly, this peak comprises two transitions, most likely corresponding to the melting of two different dimers in the same temperature range. The second major transition follows the first one and coincides with the single unfolding transition of the monomer. Unfolding occurs between 65 and 78 °C and is associated with further loss of ellipticity, extensive red shift, and a single heat capacity peak. These results suggest that, with increasing temperature, the dimer undergoes structure "opening" followed by a collapse to a monomer-like structure, which then proceeds to complete unfolding along the same pathway as a native monomer. This view is further supported by the analysis of oligomeric states after heating of the dimer to temperatures above the first transition point and subsequent cooling. This procedure leads to the formation of functional monomers (Figure S8B). Kinetics of the dimer unfolding measured using CD spectroscopy in the 58–70 °C range showed triphasic behavior: (i) fast decrease of ellipticity, (ii) a slightly slower phase with the opposite direction of amplitude, and (iii) a very slow phase with small amplitude in the direction of the first phase (Figure 4C). The second phase is lost at temperatures above 67 °C, and the curves can be fitted with a double exponential curve. At this range, rates of the second phase correspond with the unfolding rates of the monomer (Figure S7C). Kinetic data are highly consistent with our interpretation of the temperature scanning experiments.

Furthermore, we fitted all data globally to a four-state model, a partially reversible unfolding mechanism, using Calfitter v1.3 (Figure 4D).[29,30] The model provides an acceptable fit to the data and enables a robust estimation of the energy barriers of each unfolding transition (Figure S10). The largest deviation


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864–12885

from the data was observed for the first DSC peak since it includes transitions of the two dimers, which occur at similar temperatures. Since the dimers are in approx. 1:1 ratio in solution, based on native PAGE analysis, their unfolding transitions can be resolved by fitting the first transition to the unfolding model with two parallel pathways (Figure 4B). The resulting average of $\Delta H_{cal}$ values for each dimer, 180 and 88 kJ/mol (average ∼134 kJ/mol), corresponds well with the $\Delta H_{cal}$ value of this transition from the global fit (∼133 kJ/mol). The second transition can be approximated with the Lumry−Eyring model, as with the monomer. We speculate that as the monomeric unit unfolds at high temperatures, the refolding rate becomes significantly slower compared to the rate of irreversible denaturation (I → U), which results in an overestimation of the energy barrier of this step and poor resolution of the last unfolding step. The overall $\Delta H_{cal}$ for the whole transition (∼630 kJ/mol) corresponds well with the value obtained by global fitting of the monomer unfolding data (∼650 kJ/mol). We conclude that the collected data can be explained with this minimal unfolding model.

**Refolded Monomers Cannot Form Domain-Swapped Dimers.** Similar results were obtained from the chemical denaturation of DhaA115. The unfolding of DhaA115 using urea is more potent and leads to a completely unfolded polypeptide chain, in contrast to the heat-denatured state, which retains a residual structure (Figure S11). Denaturation curves of both monomer and dimers plotted as ellipticity at 224 nm and an average emission wavelength of fluorescence spectra, each show two transitions and can be fitted using a three-state unfolding model (Figure S12). The unfolding curves of both oligomeric states nearly perfectly overlap with the midpoints around 4 and 6 M urea. This indicates that the dimer first splits into monomer-like structures at low urea concentration and then unfolds along the same pathway as a native monomer. Analysis of oligomeric states in the presence of 0−3 M urea using the native PAGE confirmed this interpretation, showing a gradual loss of bands corresponding to dimers and a simultaneous increase in the monomeric fraction with its midpoint around 2.2 M urea (Figures S13 and S12A). Refolding of the DhaA115 was achieved by dilution or dialysis of the partially or completely denatured sample from 5 and 8 M urea, respectively. In all cases, refolding yielded a mixture of native-like monomers and misfolded aggregates at all protein concentrations tested (Figure S14). Precipitation occurred at a protein concentration above 1 $mg \cdot mL^{-1}$. The ratio between the aggregates and monomers was concentration-dependent. However, no dimer formation was observed during any of these unfolding/refolding experiments.

**Structure-Based Mutagenesis Reveals Importance of Mutations in the Hinge.** Next, we questioned whether the domain swapping is induced primarily by the mutations in the hinge regions or those at the secondary interfaces. Based on the structure of the DhaA115-DSD1 dimer, we first introduced three mutations into wild-type DhaA to create DhaA176 (T148L + A172I + C176F). The effect of mutations based on DhaA115-DSD2 was tested in two-point mutant DhaA178 (E20S + D198W). In the second round of mutagenesis, the stabilizing mutation A155P was inserted into both of these mutants, yielding the four-point mutant DhaA177 (T148L + A155P + A172I + C176F) and the three-point mutant DhaA179 (E20S + A155P + D198W). All newly constructed mutants were expressed and purified as soluble proteins, and their proper folding was confirmed with CD spectroscopy (Figure S15). Analysis of the oligomeric species content confirmed that all four variants form dimeric forms (Figure S15B), although much less abundantly than DhaA115. Despite the intensive effort, we could only obtain well-diffracting crystals for the dimeric fraction of DhaA177, which is the variant with the highest fraction of oligomeric species among the four additional constructs.

The DhaA177 crystals belong to the $P2_12_12_1$ space group, and its structure was solved by molecular replacement at 2.55 Å resolution (Table S2). The asymmetric unit is formed by four polypeptide chains that are assembled into two non-crystallographic domain-swapped dimers. To our surprise, these domain-swapped dimers, hereafter referred to as DhaA177-DSD3, revealed a third, completely novel mode of domain swapping. In this conformation, the short L10 loop ($P_{155}D$), connecting the $\alpha$4 and $\alpha$5′ cap helices, functions as a hinge (Figure 5A,B). The residue P155, introduced by the computational design as a stabilizing mutation,[5] plays a central role in the hinge region of DhaA177-DSD3. No apparent stabilizing secondary interface is present in the structure of DhaA177-DSD3, and all physical contacts between the two catalytic units are entirely mediated through the hinge region (Figure 5C).

Inspection of crystallographic packing revealed that the DhaA177 domain-swapped dimers are tightly packed into a Z-shaped tetramer through an extensive dimer-to-dimer interface (Figure 5D,E). This interface is built around the side chains of two F144 residues from different dimers that are packed against each other. Amino acid composition of the dimer-to-dimer interface in DhaA177-DSD3 resembles the secondary interface observed in the DhaA115-DSD1 structure (Figure 2A). Expansion along the crystallographic two-fold screw axis reveals that the DhaA177 tetramers, which are composed of two domain-swapped dimers, are repeated building blocks of crystallographic open-ended filament-like structures in the corresponding crystal lattice (Figure 5F).

The mutagenesis experiments showed that four stabilizing mutations are sufficient to induce domain swapping in the DhaA enzyme. A topological comparison of all three types of observed DhaA domain-swapped dimers is shown in Figure 6. In all three types of DhaA domain-swapped dimers, the swapping is mediated by a hinge loop located in the cap domain, as visualized in a morphing movie (Movie S1). Structurally, a malleable cap domain shapes the morphology of internal enzyme access tunnels, which are known determinants of catalytic properties for this enzyme family.[31,32] Therefore, we hypothesized that the domain-swapped dimers could exhibit novel catalytic properties.

**Abundance of Swapped-Dimers Increases with the Time of Protein Overexpression.** The collected data indicated that the domain swapping must occur *in vivo* during recombinant protein overexpression. Therefore, we carried out time-course protein expression experiments coupled with analytical size-exclusion chromatography to detect various oligomeric species of the DhaA115 enzyme during its overexpression in *Escherichia coli* (Figure S16). The only monomeric form of DhaA115 was detected 2 h after the induction. Dimers and larger oligomers, presumably tetramers and hexamers, started to accumulate to a measurable concentration (∼3 to 5% of total protein) after 4 and 6 h after the induction, respectively. The distribution of the oligomeric states following the overnight expression (12−16 h) remained relatively constant at ∼69% of the monomer, ∼24%


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

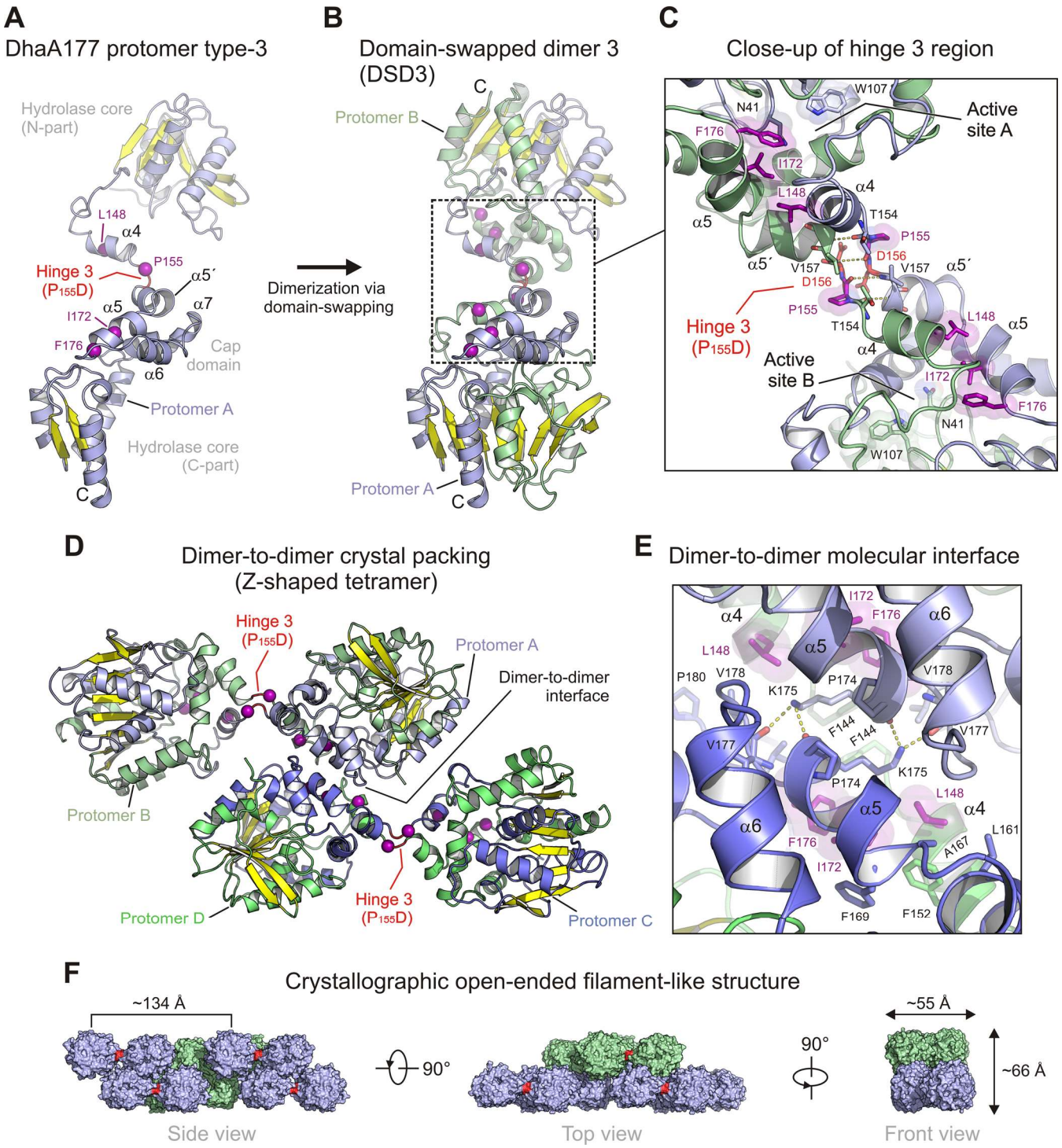



**Figure 5.** X-ray structure of domain-swapped dimer DhaA177-DSD3. Cartoon representations of (A) DhaA177 protomer type-3, (B) DhaA177-DSD3 dimer, (C) close-up view of the hinge 3 region, (D) dimer-to-dimer crystal packing, (E) close-up view of the dimer-to-dimer molecular interface, and (F) surface representation of crystallographic open-ended filament-like DhaA177 structure. The color-coding is the same as in Figure 2. The designed stabilizing mutations are depicted as purple spheres; the hinge 3 region ($P_{155}D$) is colored in red. The key residues stabilizing the intertwined protomers are shown as sticks; hydrogen bonds are shown as yellow dashed lines.

of the dimer, and ∼7% of the higher-order oligomers (Figure S16). These experiments suggest that a critical concentration of the recombinant protein must accumulate in producer cells before domain swapping can occur.

**Domain-Swapped Dimers Exhibit Enhanced Catalytic Efficiency.** Careful inspection of crystallographic structures revealed that the catalytic pentad of domain-swapped dimers is not provided by the single polypeptide chain, as is the case in the DhaA115 monomer (Figure 7A). Specifically, the catalytic histidine (H272) in the swapped dimers is provided by a second intertwined polypeptide chain (protomer B), which further supported the hypothesis that the swapped dimers might behave differently and exhibit unusual catalytic properties.


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

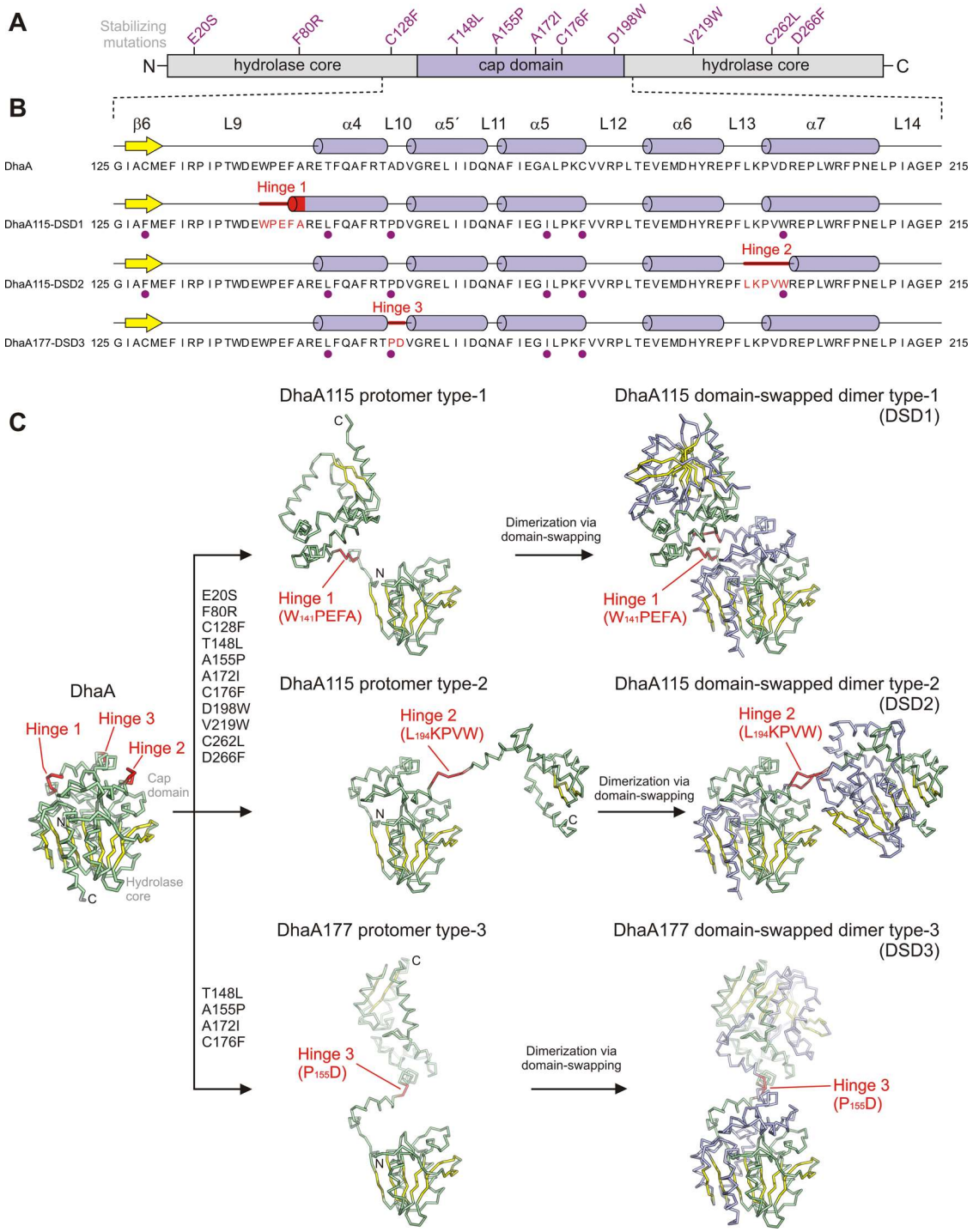



**Figure 6.** Structural comparison of three different domain-swapping topologies. (A) Schematic representation of the protein sequence showing the domain topology of DhaA115 and the positions of the stabilizing mutations. (B) Partial structure-based alignment of DhaA, DhaA115, and DhaA177 sequences. The stabilizing mutations are highlighted with violet dots. The secondary structure elements are shown above every sequence. The hinge regions are colored red. (C) Ribbon representations of the DhaA monomeric structure along with three different domain-swapped dimers DhaA115-DSD1, DhaA115-DSD2, and DhaA177. The color-coding is the same as in Figure 2.

To find out how domain swapping affects catalysis, the dehalogenase activities of the monomeric and dimeric DhaA115 forms were assayed with five halogenated substrates at 37 °C. The activity of DhaA115 dimeric form increased 3.5−8 times with all tested substrates compared to the monomeric counterpart (Figure 7B). To determine whether the several-fold improvement in dehalogenase activity was not merely the result of an optimum temperature shift, we analyzed temperature profiles of monomer and dimer activity with 1,2-dibromoethane (Figure 7C, left). Indeed, the shift of temperature maxima of the dimer was observed compared to the monomeric counterpart. The catalytic activity at the temperature maxima showed that the domain-swapped dimer of DhaA115 is ∼5.5 times more active (Figure 7C, left). Moreover, the temperature profile of the dimer activity has two peaks at temperatures that correspond well with the onset


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

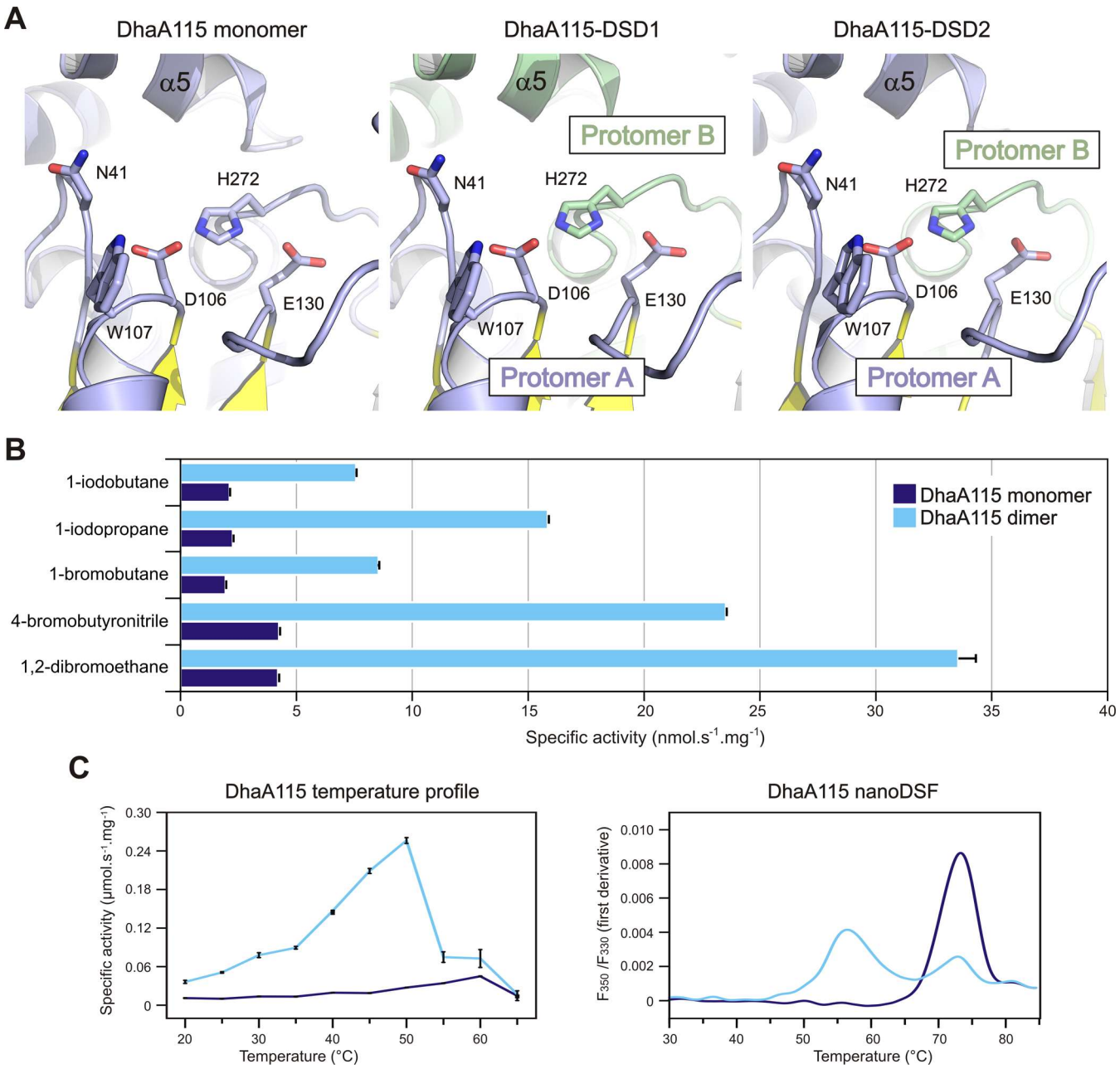
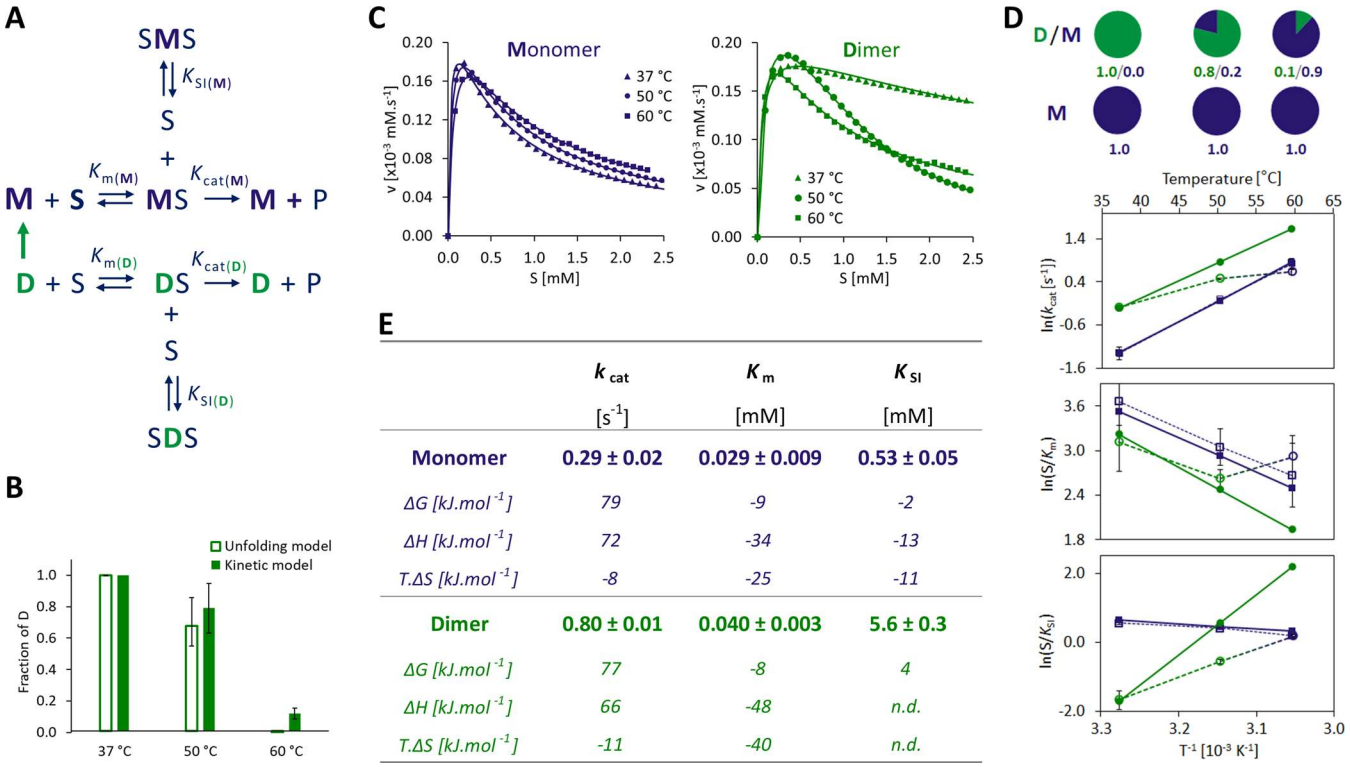



**Figure 7.** Comparison of the catalytic properties of monomers and domain-swapped dimers. (A) Positioning of active site residues (catalytic pentad) in DhaA115 monomer (left panel), DhaA115-DSD1 (middle panel), and DhaA115-DSD2 (right panel) structures. Note that the catalytic histidine base (H272) in the swapped dimers is provided by a second polypeptide chain (protomer B), which distinguishes them from the monomeric counterpart. (B) Dehalogenase activity of DhaA115 monomeric and dimeric fractions toward five halogenated substrates. The activity was determined in glycine buffer at 37 °C and pH 8.6. Presented are averages from the three independent experiments, and the error bars represent standard deviation. (C) Activity-temperature profiles of the monomeric and dimeric fraction of DhaA115 with 1,2-dibromoethane (left panel) and temperature stability measurements using nanoDSF (right panel).

temperatures of the dimer dissociation and subsequent unfolding transitions observed in the temperature-induced unfolding (Figure 7C, right). The increased catalytic activity of the dimer gradually decreased to the level of monomer activity at temperatures above the first transition, where we assume the dissociation to monomer occurs.

These observations are further supported by the steady-state kinetic measurements with isolated monomeric and dimeric fractions of DhaA115 performed at different temperatures (Figure 8). The complex steady-state kinetic data collected with a monomer and a dimer at different temperatures were fitted globally. The data fitting used numerical integration of rate equations derived from an input model incorporating the specific kinetic pathway for the monomer and the dimer as well as the dynamic transition of the dimer to the monomer due to dimer dissociation expected at elevated experimental temperatures (Figure 8A). The specific fractions of the dimeric form (Figure 8B) were calculated from the global unfolding model (Figure 4) for each experimental temperature. The fractions served as an initial parameter for the kinetic modeling (Figure 8C,D). The global numerical model and simultaneous analysis of these complex data provided estimates of individual kinetic


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885




E

| | $k_{cat}$ [$s^{-1}$] | $K_m$ [mM] | $K_{SI}$ [mM] |
|---|---|---|---|
| **Monomer** | **0.29 ± 0.02** | **0.029 ± 0.009** | **0.53 ± 0.05** |
| *ΔG [kJ.mol⁻¹]* | *79* | *-9* | *-2* |
| *ΔH [kJ.mol⁻¹]* | *72* | *-34* | *-13* |
| *T.ΔS [kJ.mol⁻¹]* | *-8* | *-25* | *-11* |
| **Dimer** | **0.80 ± 0.01** | **0.040 ± 0.003** | **5.6 ± 0.3** |
| *ΔG [kJ.mol⁻¹]* | *77* | *-8* | *4* |
| *ΔH [kJ.mol⁻¹]* | *66* | *-48* | *n.d.* |
| *T.ΔS [kJ.mol⁻¹]* | *-11* | *-40* | *n.d.* |

**Figure 8.** Steady-state kinetics and thermodynamics of DhaA115. (A) Kinetic model for the reaction of a monomer (M, blue) and a dimer (D, green), including the dynamic transition of the dimer to the monomer due to dimer dissociation at elevated temperatures. (B) Fraction of the dimeric form is calculated from the global unfolding (empty columns) and global kinetic (filled columns) models as a function of experimental temperature. (C) Steady-state kinetic data was obtained for monomeric and dimeric fractions with 1,2-dibromoethane in glycine buffer at pH 8.6 at 37, 50, and 60 °C. The solid lines represent the best fit of the global kinetic model. (D) Actual fraction of the monomer and the dimer at individual kinetic experiments under different experimental temperatures. The dimer fraction is stable at 37 °C, while it is dissociated into the monomer at higher temperatures. The Arrhenius plot of the temperature dependence of $k_{cat}$ (upper graph), the Van't Hoff plot of the temperature dependence of the equilibrium constants $S/K_m$ (middle graph), and $S/K_{SI}$ (lower graph), where [S] = 1 mM (reference experimental conditions). The kinetic parameters were obtained for the reaction with the monomeric (circles) and dimeric (squares) enzymes. The empty data points represent the average from kinetic parameters obtained in four independent experiments analyzed analytically by nonlinear fitting (eq 3, Table S3). The error bars indicate a 95% confidence interval. In the case of the dimer, the change of the color from green to blue indicates the dynamic transition of the dimer to the monomer due to dimer dissociation at elevated temperatures. The filled points and solid lines show the best fit to the complex global kinetic model representing the behavior of exclusive monomers and dimers. (E) The kinetic and thermodynamic parameters of 1,2-dibromoethane conversion catalyzed by DhaA115 monomeric and dimeric form. The values and standard errors were obtained by global fitting complete kinetic dataset, including analysis of the monomer and the dimer at three experimental temperatures. The thermodynamic parameters represent the free energy of the transient state for $k_{cat}$ and the free energy of the ground state for the equilibrium constants $K_m$ and $K_{SI}$. The entropy and enthalpy were not determined (n.d.) for the weak unfavorable $K_{SI}$ in the case of the dimer.

and thermodynamic parameters for both monomeric and dimeric forms (Figure 8E).

The turnover number ($k_{cat}$) and degree of substrate inhibition ($K_{SI}$) were improved significantly for the dimeric form, while no significant differences were observed in the level and temperature dependence of $K_m$. An increase in the catalytic rate of the dimeric form is associated with a decrease in the activation enthalpy ($\Delta H^{\neq}_{M} - \Delta H^{\neq}_{D} = 9.7$ kJ·mol$^{-1}$). The favorable enthalpic effect is partially compensated by the unfavorable drop of activation entropy between the monomer and the dimer ($T\Delta S^{\neq}_{M} - T\Delta S^{\neq}_{D} = 7.0$ kJ·mol$^{-1}$). Thus, for the dimer, either the transition state for the subsequent chemical conversion is more ordered or the ground state of the enzyme−substrate complex is stabilized as a consequence of the structural arrangement of the polypeptide chains in the domain-swapped conformation. Interestingly, the domain-swapped conformation fundamentally changed the thermodynamics of the substrate inhibition. Unlike the monomeric form, whose reaction is substantially affected by the exergonic formation of the substrate inhibitory complex ($K_{SI}$ = 0.53 ± 0.05 mM; $\Delta G^0 = -2$ kJ·mol$^{-1}$), for the dimer, the formation of the inhibited complex has become an endergonic process unfavorable at the experimental conditions ($K_{SI}$ = 5.6 ± 0.3 mM; $\Delta G^0 = 4$ kJ·mol$^{-1}$). In summary, the elevated activity of the dimer is due to increased catalytic turnover and also reduced substrate inhibition under given reaction conditions.

**Domain Swapping Changed the Access Tunnels.** In an attempt to rationalize the differences in the catalytic activities of the monomeric and dimeric forms, we calculated the access tunnels connecting the active sites to the bulk solvent. This was performed for the crystal structures and the MD simulations using CAVER 3.02,[33] and it allowed us to assess how the tunnels' geometry differed among the different variants. Moreover, it helped in evaluating how those tunnels varied in the dynamic systems in comparison to the respective static crystallographic structures. The analysis of the crystal structures showed that both the main tunnel (p1; bottleneck radius of 0.68 Å) and the slot tunnel (p2; bottleneck radius of 0.59 Å) were significantly narrowed in DhaA115 compared to the wild-type DhaA (Table S4; Figures S17 and S18). The p1


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

tunnel also became second in priority, while the p2 became the major route to the active site, as we discussed previously.[26] Interestingly, the crystal structure of dimer DhaA115-DSD1 shows that the domain swapping slightly altered the positioning of the $\alpha$4 helix and compensated the extent of the tunnel blockage to the first active site (site A) to partially recover the main and slot tunnels to bottlenecks of radii 1.19 and 1.29 Å, respectively. The second active site of DhaA115-DSD1 (site B), however, remains occluded. The dimensions of the p2 tunnel in DhaA177-DSD3 (bottleneck radii of 0.85 and 0.93 Å for sites A and B, respectively) are slightly wider than in the DhaA115 monomer (bottleneck radius of 0.59 Å) and DhaA115-DSD2 (bottleneck radii of 0.80 and 0.85 Å for sites A and B, respectively). The difference in the bottleneck size of these p2 tunnels is due to the different orientations of the R133 side chain.

Next, we calculated the access tunnels in all of the snapshots from MD trajectories and observed similar trends. All of the DhaA115 variants showed the slot tunnel (p2) is more relevant than the main tunnel (p1), with larger rates of occurrence and higher average bottleneck radii. The average bottleneck radii in the simulations were generally larger than in the crystal structures. This is a result of the commonly observed tunnel fluctuations due to the protein dynamics.[34] Such fluctuation resulted in some snapshots displaying significantly open tunnels, with the bottleneck radii reaching maximum values between 1.4 and 1.5 Å (Table S4), which are sufficiently wide to allow the access of water molecules. One exception was site A of DhaA115-DSD1, which remarkably decreased the size of its tunnel during the MD simulations in comparison to the crystal structure. This made site A a lot more similar to site B and also similar to all of the other sites in DhaA115 variants in terms of their accessibility. It also demonstrates that the crystallographic configuration of DhaA115-DSD1 might be influenced by the co-crystallized ligand located in the active site.

Another interesting finding is that the access tunnels were more open in the dimeric structures than in the monomers. The tunnels in the dimers showed higher occurrence rates than in the monomer. Altogether, the computational analysis of both crystal structures and MD simulations suggest that the active sites in the swapped dimers are slightly more accessible than in the monomeric counterpart. In combination with the higher flexibility identified in MDs of DhaA115-DSD1, it may explain their experimentally proven improved catalytic rate and elimination of substrate inhibition.

## ■ DISCUSSION

In this work, we provided direct experimental evidence that contemporary computer algorithms used for protein stabilization may affect protein folding. We previously designed hyperstable DhaA115[5,27] using force field calculations and evolutionary analysis with automated software FireProt.[5,35] A recent structural study carried out on the monomer of DhaA115 revealed an intricate network of molecular interactions that reinforce the engineered $\alpha\beta\alpha$-sandwich architecture.[26] Mutations to bulky aromatic amino acids at the protein surface trigger long-distance backbone changes through multiple cooperative interactions. These interactions produce an unprecedented double-lock system that closes the molecular gates to the active site and reduces the volumes of the access tunnels.[26] Surprisingly, we have observed that this computer-aided stabilization of DhaA was accompanied by partial *de novo* oligomerization.[27]

Here, we showed that while the wild-type DhaA exists solely as a monomer, the engineered DhaA115 variant can additionally form dimers and higher-order oligomeric forms when overexpressed in *E. coli*. Crystallographic analysis of the DhaA115 dimeric form revealed two different dimer topologies, both formed by a so-called three-dimensional domain-swapping mechanism. The crystallographic structures of DhaA variants determined in this study are the first domain-swapped structures of haloalkane dehalogenases to be seen. Several complementary experiments ruled out the possibility of the structure as a crystallization artifact.[36] The SAXS and cross-linking MS experiments provided direct experimental evidence that the domain-swapped dimers exist in the solution. Crucially, these structural observations raised a fundamental question: why does domain swapping occur during computer-aided stabilization of a protein?

Our findings demonstrate that the domain swapping in DhaA occurs through solvent-exposed loops—cryptic hinge regions—which are parts of the cap domain and represent regions, which can serve as hinges for domain swapping upon introduction of mutations (Figures S19 and S20). Interestingly, stabilizing mutations are frequently found in these cryptic hinge loops, in their vicinity and/or in the secondary interfaces, where they contribute to the noncovalent interactions between the intertwined polypeptide chains. In the DhaA115-DSD1 structure, no engineered mutation is present in the hinge loop L9 ($W_{141}$PEFA). However, the three designed mutations (T148L, A172I, and C176F) take part in a novel hydrophobic patch that stabilizes the domain-swapped dimer (DSD1) topology through the newly formed secondary interface. In DhaA115-DSD2, the stabilizing mutation D198W is localized in the hinge loop L13 ($L_{194}$KPVW). Additionally, mutation E20S participates in the extended network of protein−solvent interactions that stabilize the secondary interface. In the DhaA177-DSD3 dimer, mutation A155P is a part of the hinge region ($P_{155}$D). It has been previously demonstrated that prolines are frequently found in the hinge loops.[37] Native or engineered proline residues are found in all three respective hinges of our domain-swapped dimers. The structural analyses of these dimers show that the putative stabilizing mutations designed by force field calculations not only stabilized the monomeric form[26] but, at the same time, coincidentally increased domain-swapping propensity. Multi-point mutations to bulky hydrophobic and/or aromatic residues created a new hydrophobic interface, which is responsible for the co-translational misfolding via intermolecular interactions. We conclude that the domain-swapping propensity can be affected when protein engineering efforts simultaneously: (i) manipulate the sequence of cryptic hinge regions and (ii) introduce new interaction interfaces that increase the domain-swapping propensity by lowering the free energy of the incorrectly folded (i.e., domain-swapped) intermediates. Recently, it has been shown that the activity and selectivity of a cumene dioxygenase could be impressively modulated via the introduction of proline- and glycine-rich motifs into active-site loops.[38] Concerning the results reported here, we propose that such loop-engineering might change folding energy pathways under certain circumstances.

Our study illustrates that the DhaA protein was evolutionarily optimized toward folding into the monomeric protein. A few mutations are sufficient to modify the energy landscape


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

and lead to domain-swapped intermediates. Domain-swapped dimers are thermodynamically less stable than their monomeric counterparts, and functional native-like monomers can be restored by thermal dissociation at elevated temperatures. MD simulations suggest that the dimers are less stable due to a significant increase in their structural fluctuation. There is cumulative evidence that domain swapping played a key role in the emergence of oligomeric proteins during evolution.[39] There are also many examples describing how domain swapping is a detrimental process leading to protein misfolding and aggregation, associated with human pathologies.[23,40]

Domain swapping of the DhaA115 dimer cannot be induced *in vitro* by refolding from denatured states because folding to the monomeric state and/or aggregation are energetically more favorable and preferred pathways. Arguably, the formation of increasingly larger aggregates observed during *in vitro* refolding of the wild-type DhaA and DhaA115 could be caused by consecutive domain swapping, but additional experiments need to be carried out to validate this hypothesis. Undoubtedly, the domain swapping must therefore occur *in vivo* during recombinant protein expression. We monitored the time course of protein expression in *E. coli* and were able to detect a fraction (∼5%) of dimer at 4 h after expression induction. The fraction of dimer triples in the course of the next 2 h and then gradually plateaus at ∼24% of the total DhaA115 after overnight expression. This suggests that the critical concentration of the recombinant protein must accumulate before domain swapping can occur and that the crowding effect might play a role in this process. However, other effects such as the formation of an intermediate during co-translational folding, interaction with chaperones, or other effects might be responsible for the domain swapping.

Surprisingly, the *de novo* oligomerization of DhaA115 significantly increased its catalytic efficiency and at the same time eliminated the substrate inhibition. Previously, Fraser and co-workers experienced that the stabilization of $\alpha$E7 carboxylesterase through directed evolution also yielded new protein oligomers, which, however, were not formed via domain swapping.[41] Unlike our DhaA115 domain-swapped dimers, their dimers and tetramers displayed significantly lower catalytic activities when compared to the monomeric counterpart. To understand the extent of domain swapping on biocatalysis, we did a comparative analysis of published intermolecular protein swapping topologies (Table S5). We could find altered enzymatic activities in two reported instances—RNase A[42] and metallo-$\beta$-lactamase VIM2.[43] The domain-swapped dimers of native RNase A exhibited the increased specific activity than its monomeric form.[42] Baier and co-workers obtained the domain-swapped dimer of metallo-$\beta$-lactamase VIM2 upon a directed evolution on an originally monomeric protein template.[43] The swapped-dimer of the engineered VIM2-R10 variant expressed at higher levels but exhibited ∼8.5-fold decrease in catalytic activity when compared to the monomer. Our results together with the literature survey suggest that the protein domain swapping can be a rare but not unexpected outcome of protein engineering efforts, regardless of the structural fold. Nonetheless, such engineered domain-swapped proteins might have their stability increased (Table S5). Recently, Koga and colleagues revealed that computational design can also yield proteins with swapped secondary structure elements within one polypeptide chain.[44] We note that both the directed evolution and the computational design can end up in the misfolded proteins, indicating that this phenomenon can be much more frequent than we expect.

The advantageous catalytic properties of DhaA115 swapped-dimers can be attributed to their unique inherent properties. The MD simulations and access tunnel calculations provided some explanations. First, the higher flexibility of the swapped dimers suggests that they can be more malleable than the monomer and thus more prone to accommodate the substrates or products traveling to and from the active site. Second, the slightly larger tunnels found in the swapped dimers, both in the crystal structures as in the MD simulations, suggest the same possibility. Although those calculated tunnels are still too narrow for the practical transport of ligands, in the presence of those molecules, they could open significantly in an induced-fit manner, as we observed previously with other DhaA variants.[34] Finally, a more robust explanation and dissection of those catalytic differences would require more thorough computational and experimental studies, which are beyond the scope of this manuscript.

## ■ CONCLUSIONS

We have elucidated the structural basis of the domain-swapping mechanism of computationally stabilized DhaA by comprehensive biophysical, biochemical, and computational analyses. The results provide a new view of the structure–function relationships of haloalkane dehalogenases and their folding energy landscapes. Our comprehensive study revealed hidden consequences for protein folding through computational protein stabilization, which need to be taken into account when applying rational stabilization to biomolecules of biological and pharmaceutical interest. In addition, our discovery highlights the importance of biophysical characterization techniques, for instance, native nondenaturing PAGE, SEC, AUC, and SAXS, which should always be applied during the control and quality checks of computationally designed proteins. Knowledge derived from this study can be exploited in future projects for the rational design of stable and catalytically enhanced enzymes and the study of folding, aggregation, and stability of proteins.

## ■ EXPERIMENTAL SECTION

**Protein Production and Purification.** His-tagged DhaA variants were overexpressed from pET21b recombinant plasmids in *E. coli* BL21(DE3). The expression was induced using 0.5 mM IPTG at 20 °C for 16 h. The cells were harvested using centrifugation at 11 806*g* at 4 °C for 10 min. The pellet was resuspended in purification buffer A (500 mM NaCl, 10 mM imidazole, 20 mM potassium phosphate buffer pH 7.5) and sonicated using Sonic Dismembrator Model 705 (Fisher Scientific) in 3 cycles, each of 2 min (5 s pulse/5 s pause) with 50% amplitude. Disrupted cells were centrifuged at 21 000*g* at 4 °C for 1 h. His-tagged proteins were purified on a Ni-chelating column (Ni-NTA Superflow cartridge) equilibrated with the purification buffer A. The column-bound enzymes were eluted with a gradient of purification buffer A supplemented with 300 mM imidazole. The eluted proteins were further purified with size-exclusion chromatography using a HiLoad 16/600 Superdex 200 gel filtration column (GE Healthcare). The separated peak fractions were pooled and concentrated with an Amicon Ultra centrifugal filter unit (Merck Millipore Ltd), and protein concentrations were


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

measured with a DeNovix DS-11 Spectrophotometer (DeNovix Inc.).

**Crystallization, X-ray Data Collection, and Structure Determination.** Diffraction-quality crystals of DhaA115-DSD1 were obtained using the sitting-drop vapor diffusion method. Briefly, the DhaA115 protein was dialyzed into 50 mM Tris−HCl buffer (pH 7.5) and concentrated to ∼11.5 mg·mL$^{-1}$. Crystallization was set up using the sitting-drop vapor diffusion method with Cryschem 24-well crystallization plates (Hampton Research) at 19 °C. After 4 to 8 days, crystals appeared in a mixture (1:1) of protein and crystallization buffer consisting of 0.1 M MES/imidazole pH 6.5, 10% (w/v) PEG 20 K, 20% (v/v) PEG MME 550, and 0.09 M NPS buffer system (0.03 M $NaNO_3$, 0.03 M $Na_2HPO_4$, and 0.03 M $(NH_4)_2SO_4$). The crystals obtained were directly flash-frozen in liquid nitrogen with no additional cryo-protection. X-ray frames of DhaA115-DSD1 were collected using the MX beamline I03 (Diamond Light Source, U.K.) at a wavelength of 0.97625 Å with a Pilatus 6M-F detector. Diffracting crystals of DhaA115-DSD2 were analogously obtained using a vapor diffusion technique in a crystallization buffer containing 0.1 M Bis-Tris (pH 5.5), 0.2 M $NH_4NO_3$, and 16% PEG 3350. After 3 to 6 days at 19 °C, the crystals were harvested and flash-frozen in a reservoir solution supplemented with 20% glycerol. X-ray data for DhaA115-DSD2 were collected from the ESRF ID23-1 beamline[45] (Grenoble, France) at a wavelength of 0.861 Å using a Pilatus 6 M detector. Crystals of DhaA177-DSD3 were also obtained using a vapor diffusion technique. Diffracting crystals of DhaA177 were obtained from a crystallization buffer consisting of 0.1 M Tris (pH 8.5), 0.2 M $LiSO_4$, and 1.26 M $(NH_4)_2SO_4$. After 3 to 9 days at 19 °C, the crystals were harvested and flash-frozen in a reservoir solution supplemented with 20% glycerol. X-ray frames of DhaA177-DSD3 were collected from the PX3 beamline at SLS (Paul Scherrer Institute, Switzerland) at a wavelength of 0.99987 Å using a Pilatus 2M-F detector. The crystallographic data were processed using XDS[46] and Aimless.[47] Initial phases were solved by molecular replacement using Phaser[48] software implemented in the Phenix package.[49] The X-ray structure of wild-type DhaA (PDB: 4HZG[50]) was used as a search model for molecular replacement during structure determination. The structure refinements were carried out over several cycles of automated refinement by the phenix.refine[51] program, with manual model building, carried out in Coot.[52] The final models were validated using tools provided by Coot[52] and Molprobity.[53] Graphical visualizations of structural data were created using PyMOL.[54] Structural superposition was carried out using the secondary structure matching (SSM) superimpose tool in the Coot.[55] Atomic coordinates and structure factors of the DhaA115-DSD1, DhA115-DSD2, and DhaA177 domain-swapped enzyme variants were saved in the Protein Data Bank (www.wwpdb.org)[56] under the PDB codes 6TY7, 6XT8, and 6XTC.

**Semianalytical Size-Exclusion Chromatography.** Protein samples eluted with 300 mM imidazole (60% gradient) using metal-affinity chromatography were loaded on the FPLC system ÄKTA Purifier (GE Healthcare) equipped with a $UV_{280}$ detection system and a Superdex 200 Increase 10/300 GL gel filtration column, equilibrated with 50 mM phosphate buffer (pH 7.5). Elution was carried out using the same buffer at a constant flow rate of 0.8 mL·min$^{-1}$. The contents of oligomeric fractions were evaluated by the peak integration area.

**Analytical Size-Exclusion Chromatography.** The quaternary structure was analyzed using an analytical size-exclusion chromatography (SEC) system equipped with static light scattering, refractive index, ultraviolet, and differential viscometer detectors. The system was calibrated using bovine serum albumin as a protein standard. The Viscotec 305 TDA instrument (Malvern, U.K.) and the column Zenix-C 300 (Sepax Technologies) were equilibrated by 50 mM phosphate buffer pH 7.5. The protein, eluted with 300 mM imidazole during affinity chromatography, was dialyzed overnight in 50 mM phosphate buffer, pH 7.5, concentrated to 2.24, 4.80, and 7.67 mg·mL$^{-1}$, injected into the column, and separated at a constant flow rate of 0.3 mL·min$^{-1}$ of the elution buffer. Retention volumes, molecular weights, hydrodynamic radius, and intrinsic viscosities were evaluated using OmniSec software (Malvern, U.K.).

**Analytical Ultracentrifugation.** Analytical ultracentrifugation (AUC) experiments were carried out using a ProteomeLab XL-I analytical ultracentrifuge (Beckman Coulter, Indianapolis, IN) equipped with an An-60 Ti rotor. Sedimentation velocity experiments were carried out in titanium double-sector centerpiece cells (Nanolytics Instruments, Potsdam, Germany) loaded with 390 $\mu$L of both protein sample and reference solution (100 mM NaCl, 20 mM potassium phosphate buffer pH = 7.5). DhaA115 samples, preincubated with different concentrations (0.096, 0.31, and 0.68 mg·mL$^{-1}$), were analyzed as part of the sedimentation velocity experiment. Data were collected using absorbance and interference optics at 20 °C at a rotor speed of 45 000 rpm. Scans were collected at 280 nm at 5 min intervals and 0.003 cm spatial resolution in a continuous scan mode, with an interference laser duration of 0.1°. The partial specific volume of the protein and the solvent density and viscosity were calculated from the amino acid sequence and buffer composition, respectively, using the software Sednterp (http://bitcwiki.sr.unh.edu). The data were analyzed with the continuous $c(s)$ distribution model implemented in the program Sedfit 15.01b.[57] For the regularization procedure, a confidence level of 0.95 was used. The plots of $c(s)$ distributions were created in GUSSI 1.3.1.[58]

**Native Polyacrylamide Gel Electrophoresis.** The protein sample (∼1 mg·mL$^{-1}$) was mixed with loading buffer (35% glycerol, 0.25 M Tris−HCl pH 6.8, 0.04% Bromophenol Blue) in a ratio of 1:3 and 13 $\mu$L of the mix was loaded into the 12.5% native gel with a 4% stacking gel layer. Electrophoresis was carried out in a Tris-glycine electrophoretic buffer pH 8.3 (25 mM Tris, 192 mM Glycine) at 110V and 4 °C. Protein bands of polyacrylamide gels were stained with InstantBlue Protein Stain (Sigma), following the supplier's protocol, and checked by a GS-800 Calibrated Densitometer.

**Circular Dichroism (CD).** The spectra were recorded at room temperature using a Chirascan spectrometer (Applied Photophysics, U.K.). Some 300−400 $\mu$L of a protein sample at a concentration of ∼0.3 mg·mL$^{-1}$, dialyzed against 50 mM phosphate buffer (pH 7.5), was placed in a 0.1 cm quartz cuvette, and data were collected from 185 to 260 nm, at 100 nm·min$^{-1}$, 1 s response time and 2 nm bandwidth. The final spectrum was an average of three individual scans that had been corrected for baseline noise.

**Differential Scanning Fluorimetry (DSF).** The thermal stability of enzyme variants was determined with label-free nanoDSF using a Prometheus NT.48 instrument (NanoTemper Technologies, Germany), which uses inherent


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

tryptophan fluorescence to monitor protein unfolding. Capillaries were filled with protein samples (~1 mg·mL$^{-1}$) in 50 mM potassium phosphate buffer (pH 7.5), and fluorescence was monitored in the temperature range of 20−90 °C with a 1 °C·min$^{-1}$ heating rate. Melting temperatures were determined from the ratio of tryptophan emission at 330 and 350 nm using PR-ThermControl (NanoTemper Technologies, Germany).

**Dehalogenase Activity Measurements.** Dehalogenase activity was measured using the colorimetric method described by Iwasaki et al.[59] The dehalogenase reaction was tested on five halogenated substrates, namely, 1,2-dibromoethane, 4-bromobutyronitrile, 1-bromobutane, 1-iodopropane, and 1-iodobutane, in 25 mL Reacti Flasks closed with Mininert Valves. Reaction mixtures were composed of 10 mL of glycine buffer (pH 8.6) and 10 μL of a substrate. The mixtures were incubated at 37 °C in a shaking bath for 20 minutes. The reaction was initiated by the addition of 150−600 μL of the enzyme at a concentration of 0.4−1.2 mg·mL$^{-1}$. The reactions were monitored by withdrawing 1 mL of aliquots at regular intervals and stopped by the addition of 100 μL of 35% nitric acid. Halide ions released by the dehalogenase reaction were measured spectrophotometrically at 460 nm after adding 100 μL of Iwasaki solution I (28.4 mM $Hg(SCN)_2$ in ethanol) and 200 μL of Iwasaki solution II (0.56 M $FeNH_4(SO_4)_2{\cdot}12H_2O$, 21% $HNO_3$). Enzyme activity was quantified from the slope of the relationship between the product concentration and the time.

**Temperature Profile Measurements.** The specific activities of individual enzyme variants with 1,2-dibromoethane were assayed from 20 °C to 65 °C in 5° increments. The temperature profiles were measured using a capillary-based droplet microfluidic platform,[60] which enables the characterization of enzymatic activity in droplets for multiple enzymes in one run. Briefly, the droplets were generated using Mitos Dropix (Dolomite, U.K.). A custom sequence of droplets (150 nL aqueous phase, 300 nL oil spacing) was generated using negative pressure (microfluidic pump), and the droplets were guided through a polythene tubing to the incubation chamber. Within the incubation chamber, the halogenated substrate was delivered to the droplets with a combination of microdialysis and partitioning between the oil (FC 40) and the aqueous phase. The reaction solution consisted of a weak buffer (1 mM HEPES, 20 mM $Na_2SO_4$, pH 8.0) and a complementary fluorescent indicator 8-hydroxypyrene-1,3,6-trisulfonic acid (50 μM HPTS). The fluorescence signal was detected using an optical setup with an excitation laser (450 nm), a dichroic mirror with a cut-off at 490 nm filtering the excitation light, and a silicon detector. Using a pH-based fluorescence assay, small changes in the pH could be detected, which enabled monitoring of the enzymatic activity. Reaction progress was analyzed as an end-point measurement recorded after passing through the incubation chamber. The reaction time was determined precisely for each data collection activity and was approx. 4 min. The raw signal was processed with a droplet detection script written in MATLAB 2017b (Mathworks).

**Steady-State Kinetics.** The steady-state kinetics of 1,2-dibromoethane conversion catalyzed by the DhaA115 monomer and dimer were measured using a VP-ITC isothermal titration calorimeter (MicroCal). A reaction vessel was filled with 1.4 mL of the enzyme at a concentration of 0.087−0.87 μM (0.003−0.03 mg·mL$^{-1}$). The substrate solution was prepared by dissolving 10 μL of 1,2-dibromoethane in 4 mL of the 100 mM glycine buffer (pH 8.6); its final concentration was verified with gas chromatography (Agilent) for each measurement (approx. 12−13 mM). The substrate was titrated in 10 μL injections with 150 s intervals into the reaction vessel. Each injection increased the substrate concentration, leading to a change in the reaction rate (a change of heat generated) until the enzymatic reaction was saturated. A total of 28 injections were made during titration. The reaction rates reached after each injection (in units of thermal power) were converted into enzyme turnover values using the apparent molar enthalpy ($\Delta H_{app}$), as shown in eq 1, where [P] is the molar concentration of product generated and $Q$ is the enzyme-generated thermal power.

$$\text{rate} = \frac{d[P]}{dt} = \frac{1}{v.\ \Delta H_{app}} \cdot \frac{dQ}{dt} \tag{1}$$

$$\Delta H_{app} = \frac{1}{[S]_{tot}.\ v} \int_{t=0}^{t=\infty} \frac{dQ(t)}{dt} dt \tag{2}$$

Apparent molar enthalpy ($\Delta H_{app}\Delta H$) was determined using eq 2, where [S] is the molar concentration of a substrate converted by an enzyme in a separate experiment where the reaction was allowed to proceed to completion. The final value was obtained as the average of the integrated area from 4 individual total conversions. The dependence of the calculated rate on the concentration of the substrate [S] was then fitted by nonlinear regression using Origin 8.0 (OriginLab, Northampton, MA) and eq 3, where $k_{cat}$ is a turnover number, $E_0$ is a concentration of enzyme active sites, $K_m$ is Michaelis constant, and $K_{SI}$ is a substrate inhibition constant.

$$\text{rate} = \frac{k_{cat}.\ [S]{\cdot}[E_0]}{K_m + [S]{\cdot}\left(1 + \frac{[S]}{K_{SI}}\right)} \tag{3}$$

All kinetic data collected with the monomer and the dimer at different temperatures were fitted globally using the KinTek Explorer (KinTek Corporation). Data fitting used numerical integration of rate equations from an input model (Figure 8A), searching a set of parameters that produce a minimum $\chi^2$ value using nonlinear regression based on the Levenberg−Marquardt method.[61] Residuals were normalized by sigma value for each data point. The standard error (S.E.) was calculated from the covariance matrix during nonlinear regression. The resulting global kinetic model was used to simulate the temperature dependence of all obtained kinetic parameters. The temperature dependence of the turnover number was analyzed using the Eyring equation, $\ln(k_{cat}/T) = -\Delta H^{\ddagger}/(RT) + \ln(k_B/h) + \Delta S^{\ddagger}/R$, to estimate the enthalpy ($\Delta H^{\ddagger}$) and entropy of activation ($\Delta S^{\ddagger}$), where $R$ is the universal gas constant, $k_B$ is the Boltzmann constant, and $h$ is Planck's constant. The temperature dependence of the equilibrium constants $S/K_m$ and $S/K_{SI}$, where [S] = 1 mM (reference reaction condition), was analyzed using the Van't Hoff equation $\ln(K_{eq}) = -\Delta H^0/(RT) + \Delta S^0/R$, to estimate the ground state enthalpy ($\Delta H^0$) and entropy ($\Delta S^0$). The Gibbs free energy ($\Delta G$) was defined as $\Delta G = \Delta H - T{\cdot}\Delta S$ at a reference temperature of 310.15 K (37 °C).

**Small-Angle X-ray Scattering (SAXS).** The SAXS datasets were collected using the BioSAXS-1000, Rigaku at CEITEC (Brno, Czech Republic). Data were collected at 293 K with a focused (confocal OptiSAXS optic, Rigaku) Cu Kα X-ray (1.54 Å). The sample to detector (PILATUS 100 K, Dectris) distance was 0.48 m, covering a scattering vector ($q$ =


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

$4\pi \sin(\theta)/\lambda$) range from 0.008 to 0.6 Å$^{-1}$. A size-exclusion buffer (50 mM potassium phosphate buffer pH 7.5) was used for the blank measurement. The dimeric form of the DhaA115 protein sample was measured at ∼8.7, 6.3, 4.4, and 2.2 mg·mL$^{-1}$ concentrations. Six separate two-dimensional images were collected for buffer and sample (5 min exposure per image, 30 min total exposure). Radial averaging, data reduction, and buffer subtractions were carried out using SAXSLab3.0.0r1, Rigaku. Six individual scattering curves (5 min exposures) were compared to check the radiation damage, then averaged. Integral structural parameters (Table S1) were determined using PRIMUS/qt ATSAS v.2.8.4.[62] Data points before the Guinier region were truncated. Individual scattering curves from the concentration series were manually merged and truncated to a maximum of q = 0.4 Å$^{-1}$ for further analysis. Refined *ab initio* models were produced by averaging the 10 individual *ab initio* models produced by DAMMIF. The fixed core identified by DAMAVER was used as an input for DAMMIN modeling, where the computation mode was set to "slow", and all other parameters were kept at their defaults. Evaluation of solution scattering and fitting to experimental scattering curves was carried out using CRYSOL, where automatic constant subtraction was allowed; other parameters were kept at their defaults. Superimposition of the atomic and *ab initio* models was carried out using SUPCOMB. Small-angle scattering datasets, experiment details, atomic models, and fits have been saved to the Small-Angle Scattering Biological Data Bank (www.sasbdb.org)[63] as entry SASDHQ7.

**Cross-linking Experiments.** The SEC-separated monomeric and dimeric fractions of the DhaA115 enzyme were cross-linked separately in 50 mM potassium phosphate buffer (pH 7.5) using the MS cleavable cross-linker, DSBU (CF Plus Chemicals, Czech Republic). The cross-linking reaction was carried out at 4 °C for 1 h using a 1:100 protein:cross-linker ratio. Following cross-linking, complexes were either taken directly for in-solution trypsin digestion (sequencing grade, Promega) or subjected to SDS-PAGE (10% gels) to separate cross-linked complexes. SDS-PAGE gels were stained with Biosafe Coomassie (Biorad). Respective bands with cross-linked proteins were excised and further processed for overnight trypsin digestion. The resulting peptides and dipeptides were extracted into LC-MS vials using 2.5% formic acid (FA) in 50% acetonitrile (ACN) and 100% ACN with the addition of poly(ethylene glycol) (0.001%),[64] then concentrated in a SpeedVac concentrator (Thermo Fisher Scientific).

**LC-MS/MS Analysis.** LC-MS/MS analyses of all peptide mixtures were carried out using the RSLCnano system connected to an Orbitrap Fusion Lumos Tribrid mass spectrometer (Thermo Fisher Scientific). Before LC separation, tryptic digests were concentrated online and desalted using a trapping column (100 μm × 30 mm, column compartment temperature of 40 °C) filled with a 3.5 μm X-Bridge BEH 130 C18 sorbent (Waters). After washing of the trapping column with 0.1% FA, the peptides were eluted (flow 300 nL·min$^{-1}$) from the trapping column into an analytical column (Acclaim Pepmap100 C18, 3 μm particles, 75 μm × 500 mm; Thermo Fisher Scientific, column compartment temperature of 40 °C) using a 50 min nonlinear gradient program (1−56% of mobile phase B; mobile phase A: 0.1% FA in water; mobile phase B: 0.1% FA in 80% ACN). Equilibration of the trapping column and the analytical column was completed before sample injection into the sample loop. The analytical column outlet was connected online to the Digital PicoView 550 (New Objective) ion source with a sheath gas option and SilicaTip emitter active (New Objective; FS360-20-15-N-20-C12). ABIRD (Active Background Ion Reduction Device, ESI Source Solutions) was installed. MS data were acquired using a data-dependent strategy with a cycle time of 3 s and with a survey scan (300−1600 $m/z$). The resolution of the survey scan was 120 000 (200 $m/z$) with a maximum injection time of 50 ms. Stepped HCD collision energies of 21, 27, and 33 were used for fragmentation of all precursors. MS/MS were recorded at 30 000 resolution with a maximum injection time of 150 ms and an isolation width of 1.4 $m/z$.[65] Singly and doubly charged ions were excluded from fragmentation as cross-linked peptides are usually +3 or above.[66]

**MS Data Analysis.** Raw files were used to create MGF files using Proteome Discoverer software (Thermo Fisher Scientific; version 1.4) with an in-house Mascot (Matrixscience, version 2.6). During MGF file generation, raw data were filtered out for cRAP proteins (based on http://www.thegpm.org/crap/; 112 sequences in total) to exclude common protein contaminants. MGF files were further processed in MeroX software for the identification of cross-linked peptides. For MeroX (Version: 2.0.1.3) searches, the following settings were used: cross-linker fragments: BuUr (+111.032028 Da) and Bu (+85.05276381 Da); specificity for site 1: K and N-terminus; specificity for site 2: K,S,T,Y and N-terminus. Additionally, the RISE Mode was activated to compensate for 1 missing reporter doublet ion; MS1 accuracy: 10 ppm; MS2 accuracy: 25 ppm; enzyme used: trypsin; max. missed cleavages: arginine 3, Lysine 3; minimum peptide length: 5; max. modifications: 2; peptide mass: 200−6000 Da; static modifications: carbamidomethylation (cysteine, +57.021 Da); dynamic modifications: oxidation (methionine), deamidation (asparagine and glutamine). The false discovery rate (FDR) was set to 5%. Decoy analysis was carried out by shuffling the FASTA database while keeping the amino acids of protease sites in place.

**Thermal Unfolding and Refolding.** The DhaA115 monomer and dimer fractions were separated with SEC using a HiLoad S200 16/60 column (Cytiva); their purity was verified using native PAGE. For DSC experiments, the samples were further dialyzed against 50 mM potassium phosphate buffer, pH 7.5, overnight. All unfolding and refolding experiments were carried out using this buffer. Concentrations of the samples were 0.2 and 1 mg·mL$^{-1}$ for spectroscopic and DSC measurements, respectively. Unfolding transitions were found to be concentration-independent in the 0.1−2 mg·mL$^{-1}$ range tested. Temperature scanning experiments were carried out at a constant scan rate of 1 °C·min$^{-1}$ from 20 to 85 °C, or with higher temperatures when possible. Heat capacity, ellipticity at 224 nm, and the ratio of fluorescence intensities at 350 and 330 nm were recorded using a VP-Capillary DSC (Malvern Panalytical), Chirascan V100 spectrometer (Applied Photophysics), and NT.48 Prometheus (NanoTemper Technologies), respectively. The buffer−buffer scan baseline was subtracted from the protein-buffer heat capacity data before their concentration normalization. Kinetics of unfolding was monitored by recording changes in ellipticity at 224 nm at different constant temperatures. A sample at room temperature was diluted 10 times with a preheated buffer to a final concentration of 0.2 mg·mL$^{-1}$ and immediately transferred to a preheated cuvette placed in the instrument measurement chamber. The whole process took approx. 5 s, which can be considered as the dead time of the measurement. The


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

temperature of the solution was determined using a thermocouple inserted into the cuvette. Analysis of the oligomeric state of the samples after the kinetics monitoring was carried out with native PAGE. Refolding was carried out in several ways. First, samples (0.2 and 1 mg·mL$^{-1}$) were heated and cooled down to different temperatures (60−79 °C) at a 1 °C·min$^{-1}$ scan rate. Their oligomeric state and activity were determined by native PAGE and Iwasaki assay, respectively. Next, the dimer (1 mg·mL$^{-1}$) was kept at 62 °C in the Eppendorf tube and two aliquots were withdrawn at different times (10, 20, 40, 60, 120, and 240 min). One aliquot was transferred to a chilled tube on ice, while the second was allowed to equilibrate to room temperature on a benchtop. The oligomeric state of each aliquot was determined using native PAGE. Finally, the same experiment was repeated but with different concentrations of the dimer (1, 2, 3, 4, 5, 6, and 6.6 mg·mL$^{-1}$) and 10 min incubation time.

**Unfolding Data Analysis.** Data from DSC, CD, fluorescence temperature scanning experiments, and unfolding kinetics of the DhaA115 dimeric fraction were fitted globally to the four-state partially reversible model using MatLab version of CalFitter[30] software. The model involves the irreversible transition of the domain-swapped dimers to monomer subunits, followed by their reversible transition to the intermediate with subsequent irreversible denaturation. The dissection of the total ΔCp for each step was not feasible based on the data. Therefore, several dissections of ΔCp were made during fitting, for example, fixing it for one or more steps at 0 or different values. This resulted in changes in the slope and shape of the $\Delta G^{\ddagger}(T)$ curve, especially at low temperatures. However, the estimation of energy barriers between 50 and 90 °C where the transitions occur is very robust and relatively insensitive to the changes and dissection of ΔCp. Deconvolution of the first DSC peak was achieved using the unfolding model with two parallel pathways included in the online version of CalFitter (N → I1 → D, N → I2 → D).[30] The parameters for energy barriers of the first steps were set to zero and fixed, simulating the situation where two different species in solution irreversibly change to the same final state. The data were cut at 65 °C, and other parameters were allowed to vary. This model is perfectly valid since the dimers in the solution are in an approximate 50/50 ratio. Data analysis of the DhaA115 monomer was carried out in the CalFitter using the three-state partially reversible model (Lumry−Eyring).[30]

**Chemical Denaturation Experiments.** Denaturing buffer was prepared by dissolving urea in 50 mM PB pH 7.5 to a final concentration of 9 M. The precise concentration was determined using refractometry.[67] Next, a dilution series of urea was prepared between 0 and 9 M. The DhaA115 monomer and dimer were added to each aliquot ($c_{final}$ = 0.2 mg·mL$^{-1}$) and were allowed to equilibrate at 25 °C (approx. 24 h). The fluorescence spectrum of each sample was collected between 300 and 400 nm after excitation at 266 nm using an UNcle instrument (Unchained labs). CD spectra were collected in the range of 210−260 nm as an average of 5 measurements with 1 nm bandwidth and 0.5 s integration time. Denaturation curves were plotted as the average emission wavelength of the fluorescence spectra and ellipticity at 224 nm versus urea concentration. Data were fitted to the three-state reversible model using CDpal software.[68] For the dimer/monomer transition, the samples were equilibrated in 0−3 M urea and, after equilibration for 24 h, they were mixed in a 3:1 ratio with the loading dye and analyzed for their oligomeric state using native PAGE. The fraction of monomer in each aliquot was calculated by densitometric analysis of the bands. Refolding was carried out in several ways. First, the monomer was partially or fully denatured in 5 or 9 M urea, respectively, at various initial protein concentrations and subsequently diluted 10 times to native conditions. The oligomeric state was determined using native PAGE. Next, the kinetics of monomer refolding on dilution from denaturing to native conditions was probed by measuring changes in ellipticity at 224 nm over time. Finally, the partially or fully denatured monomer was dialyzed to native conditions for 24 h and then analysis of the oligomeric state was carried out using native PAGE. Both unfolding and refolding were also repeated under reductive conditions by adding 1 mM TCEP to the samples.

**Time-Course Protein Overproduction Assay.** *E. coli* BL21(DE3) cells were transformed with pET21-*DhaA*115 plasmid, plated on agar plates with ampicillin (100 μg·mL$^{-1}$), and grown overnight. The obtained colonies were used for the inoculation of 10 mL of Luria-Bertani (LB) medium with ampicillin (100 μg·mL$^{-1}$), and the cells were grown overnight at 37 °C at 200 rpm. The overnight culture (2 mL) was then used to inoculate 250 mL of LB medium supplemented with ampicillin (100 μg·mL$^{-1}$). The culture was incubated at 37 °C and 115 rpm until an $OD_{600}$ of ∼0.5 was reached, and then expression was induced by the addition of IPTG to a final concentration of 100 μM. The overexpression of proteins was stopped at 2, 4, 6, and 16 (overnight) h post-induction, and cell cultures were harvested by centrifugation (4000*g*, 10 min, 4 °C). The cells were disrupted by sonication within 4 cycles of 2 min processing time at 50% amplitude (Fisher Scientific). The cell lysate was clarified by centrifugation (21 000*g*, 1 h, 4 °C). The crude extract was further purified on a HiTrap IMAC HP 1 mL column charged with $Ni^{2+}$ ions (GE Healthcare, U.K.) using an Äkta purifier (GE Healthcare, U.K.) and an FPLC BioLogic DuoFlow purifier (Bio-Rad). To determine oligomeric states, a size-exclusion chromatography instrument equipped with static light scattering, refractive index, ultraviolet, and differential viscometer detectors was used to analyze the quaternary structure of DhaA115 using a Viscotec 305 TDA instrument (Malvern, U.K.) and Zenix-C SEC-300 column (Sepax Technologies). The instrument was cleaned with 10% methanol and ultrapure water and equilibrated with 50 mM phosphate buffer pH 7.5. The column was washed with ultrapure water and equilibrated with 50 mM phosphate on Äkta FPLC (GE Healthcare, Sweden). The system was calibrated using bovine serum albumin (BSA) as a protein standard. Protein samples were prepared in a concentration > 1 mg·mL$^{-1}$, injected on the column, and separated at a constant flow rate of 0.3 mL·min$^{-1}$. The results were evaluated using OmniSec software (Malvern, U.K.).

**Molecular Dynamics (MD) Simulations.** The crystallographic structures of the DhaA115 dimers were used. The solvent and crystallization molecules were removed, and the double side chains were corrected to keep only the most populated conformation using the *pdb4amber* module of AmberTools 14.[69] The hydrogen atoms were added using the H++ server,[70] calculated in the implicit solvent at pH 7.5, 0.1 M salinity, an internal dielectric constant of 10, and external of 80. The original crystallization solvent was added, and the *tLEaP* program of AmberTools 14 was used to prepare the topology and coordinate files. For this, the force field ff14SB[71] was defined, $Na^+$ and $Cl^-$ ions were added to neutralize the system and achieve 0.10 M concentration of


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885

NaCl salt, and an octagonal box of TIP3P[72] water molecules with the edges at least 10 Å away from the protein atoms was added. MD simulations were carried out using the PMEMD.CUDA[73,74] module of AMBER 14.[69] In total, five minimization steps and twelve steps of equilibration dynamics were carried out before the production MD. The first four minimization steps, composed of 2500 cycles of steepest descent followed by 7500 cycles of the conjugate gradient, were carried out as follows: (i) in the first step, all of the atoms of the protein and ligand were restrained with a 500 kcal·mol$^{-1}$·Å$^2$ harmonic force constant; (ii) in the remaining steps, only the backbone atoms of the protein and heavy atoms of the ligand were restrained, respectively, with 500, 125, and 25 kcal·mol$^{-1}$·Å$^2$ force constants. A fifth minimization step, composed of 5000 cycles of steepest descent and 15 000 cycles of the conjugate gradient, was carried out without any restraints. The subsequent MD simulations used periodic boundary conditions, the particle mesh Ewald method for the treatment of the long-range interactions beyond the 10 Å cutoff,[75] the SHAKE algorithm[76] to constrain the bonds involving the hydrogen atoms, a Langevin thermostat with a collision frequency of 1.0 ps$^{-1}$, and a time step of 2 fs. Equilibration dynamics were carried out in twelve steps: (i) 20 ps of gradual heating from 0 to 310 K, under constant volume, restraining the protein atoms and ligand with a 200 kcal·mol$^{-1}$·Å$^2$ harmonic force constant; (ii) ten MDs of 400 ps each, at constant pressure (1 bar) and constant temperature (310 K), with gradually decreasing restraints on the backbone atoms of the protein and heavy atoms of the ligand with harmonic force constants of 150, 100, 75, 50, 25, 15, 10, 5, 1, and 0.5 kcal·mol$^{-1}$·Å$^2$; (iii) 400 ps of unrestrained MD at the same conditions as the previous restrained MDs. The energy and coordinates were saved every 10 ps. The production MDs were run for 100 ns using the same settings employed in the last equilibration step and carried out in duplicate for each system. The trajectories were analyzed using the *cpptraj*[77] module of AmberTools 14 and visualized using PyMOL 1.7.4[54] and VMD 1.9.1.[78] The simulations of each type were combined in a single one using *cpptraj,*[77] aligned to the respective crystal structures by minimizing the root-mean-square deviation (RMSD) of the backbone atoms (C, N, O, $C_\alpha$), excluding the very flexible terminal residues of each chain (5−7 terminal residues). The B-factors were calculated for the backbone atoms of each system by stepwise fitting each catalytic unit as follows: (i) for the DhaA115 monomer, calculated for residues 9−289 after aligning the simulation by fitting the same atoms; (ii) for DhaA115-DSD1, calculated for residues A:10−141 + B:133−290 (the entire catalytic unit from site A + hinge) by aligning A:10−133 + B:143−290 (catalytic unit from site A), and for residues A:142−290 + B:10−132 (catalytic unit from site B) by aligning the same residues; (iii) for DhaA115-DSD2, calculated for residues A:10−199 + B:133−290 (catalytic unit from site A + hinge) by aligning A:10−192 + B:200−289 (catalytic unit from site A), and for residues A:200−290 + B:10−192 (catalytic unit from site B) by aligning the same residues. This was done to exclude the fluctuations due to the tilting and rocking movements of the units with respect to each other and thus obtain normalized B-factors more comparable to the monomeric unit. The MD trajectories of the dimers were clustered using *cpptraj,*[77] with distance-based metrics of the mass-weighted RMSD for all of the heavy atoms, excluding the highly flexible terminal residues (residues 9−289 were used). The hierarchical agglomerative clustering algorithm was used with average linkage, epsilon 1.5 and sieve 5, and a minimum of 3 clusters. The centroid structures of the different clusters were used for visualization.

**Access Tunnel Calculations.** CAVER 3.02[33] was used to calculate and cluster the tunnels in the crystal structures, the aggregated MD simulations, and the previously reported analogous simulations of DhaA.[27] The tunnels were calculated for every 10 ps-spaced snapshots using a probe radius of 0.7 Å (0.5 Å for the crystal structures), a shell radius of 3 Å, and shell depth of 4 Å. The starting point for the tunnel calculation was defined by the geometric center of the carboxylic oxygen atoms of the catalytic D106. The clustering was performed by the average-link hierarchical Murtagh algorithm, with a weighting coefficient of 1 and a clustering threshold of 3.5 Å. The approximate clustering was allowed only when the total number of tunnels was higher than 20 000, and it was performed using 20 training clusters.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscatal.1c03343.

Supporting information file provides detailed data from crystallographic analysis, molecular dynamics simulations, kinetic experiments, expression analysis, and biochemical characterization (PDF)

Putative conformational interconversions between the DhaA swapped-dimers (Movie S1) (MP4)

## ■ AUTHOR INFORMATION


### Corresponding Authors

**Jiri Damborsky** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic;* orcid.org/0000-0002-7848-8216; Email: jiri@chemi.muni.cz

**Martin Marek** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic;* orcid.org/0000-0001-7220-5644; Email: martin.marek@recetox.muni.cz

### Authors

**Klara Markova** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic*

**Antonin Kunka** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic*

**Klaudia Chmelova** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic*

**Martin Havlasek** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic*

**Petra Babkova** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic*

**Sérgio M. Marques** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic;* orcid.org/0000-0002-6281-7505

**Michal Vasina** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic*

**Joan Planas-Iglesias** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic*

**Radka Chaloupkova** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; Enantis Ltd., 625 00 Brno, Czech Republic*

**David Bednar** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic*

**Zbynek Prokop** − *Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic; International Clinical Research Center, St. Anne's University Hospital Brno, 656 91 Brno, Czech Republic;* orcid.org/0000-0001-9358-4081

Complete contact information is available at: https://pubs.acs.org/10.1021/acscatal.1c03343

### Author Contributions

K.M. and A.K. contributed equally to this study. K.M. and K.C. prepared the protein samples for crystallization, carried out the initial crystallization screenings, and optimized crystallization hits. K.M. prepared the protein samples for SAXS. K.M., K.C., R.C., and M.M. collected diffraction data and solved the protein crystal structures. A.K., M.H., and Z.P. carried out thermal and chemical denaturation experiments. M.V. measured temperature profiles. S.M.M. and D.B. carried out MD analyses. P.B., K.C., R.C., and Z.P. carried out enzyme kinetics experiments. J.P.-I helped with the literature searches. M.M., Z.P., and J.D. designed the project, supervised research, and interpreted data. K.M., A.K., and M.M. wrote the manuscript with contributions from all authors. All authors have approved the final version of the manuscript.

### Notes

The authors declare no competing financial interest.

Atomic coordinates and structural factors have been saved in the Protein Data Bank (www.wwpdb.org)[56] under PDB accession codes: 6TY7, 6XT8, and 6XTC. SAXS datasets, experimental details, atomic model, and fits have been saved in the Small-Angle Scattering Biological Data Bank (www.sasbdb.org)[63] as entry SASDHQ7. The authors will release the atomic coordinates and experimental data upon article publication.

## ACKNOWLEDGMENTS

The authors would like to express their thanks to the Czech Ministry of Education (CZ.02.1.01/0.0/0.0/17_043/0009632, CZ.02.1.01/0.0/0.0/16_026/0008451, LM2018121, LM2015047). This project has received funding from the European Union's Horizon 2020 research and innovation program (Nos. 857560 and 814418). This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No. 792772. M.M. acknowledges financial support from GAMU of Masaryk University (MUNI/H/1561/2018) and M.V. by Brno Ph.D. Talent. The computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA LM2018140) provided within the program Projects of Large Research, Development and Innovation Infrastructures. The CIISB research infrastructure project (LM2018127) is acknowledged for financial support of the measurements at the CF BIC, CF X-ray and CF Prot core facilities. We thank Tomas Klumpler (CEITEC-MU, Brno, Czech Republic) for his assistance during SAXS data collection and processing. The crystallographic experiments were carried out using beamline ID23-1 at the European Synchrotron Facility (ESRF) in Grenoble (France), using beamline I03 at the Diamond Light Source in Didcot (United Kingdom), and using the PXIII beamline at the Swiss Light Source (SLS) in Villigen (Switzerland). We are grateful to the members of the synchrotron facilities for the use of their beamlines and help during data collection.

## REFERENCES

(1) Kazlauskas, R. Engineering More Stable Proteins. *Chem. Soc. Rev.* **2018**, *47*, 9026−9045.

(2) Bommarius, A. S.; Paye, M. F. Stabilizing Biocatalysts. *Chem. Soc. Rev.* **2013**, *42*, 6534.

(3) Goldenzweig, A.; Fleishman, S. J. Principles of Protein Stability and Their Application in Computational Design. *Annu. Rev. Biochem.* **2018**, *87*, 105−129.

(4) Musil, M.; Konegger, H.; Hon, J.; Bednar, D.; Damborsky, J. Computational Design of Stable and Soluble Biocatalysts. *ACS Catal.* **2019**, *9*, 1033−1054.

(5) Bednar, D.; Beerens, K.; Sebestova, E.; Bendl, J.; Khare, S.; Chaloupkova, R.; Prokop, Z.; Brezovsky, J.; Baker, D.; Damborsky, J. FireProt: Energy- and Evolution-Based Computational Design of Thermostable Multiple-Point Mutants. *PLoS Comput. Biol.* **2015**, *11*, No. e1004556.

(6) Wijma, H. J.; Floor, R. J.; Jekel, P. A.; Baker, D.; Marrink, S. J.; Janssen, D. B. Computationally Designed Libraries for Rapid Enzyme Stabilization. *Protein Eng. Des. Sel.* **2014**, *27*, 49−58.

(7) Bennett, M. J.; Choe, S.; Eisenberg, D. Domain Swapping: Entangling Alliances between Proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 3127−3131.

(8) Bennett, M. J.; Schlunegger, M. P.; Eisenberg, D. 3D Domain Swapping: A Mechanism for Oligomer Assembly. *Protein Sci.* **1995**, *4*, 2455−2468.

(9) Gronenborn, A. M. Protein Acrobatics in Pairs — Dimerization via Domain Swapping. *Curr. Opin. Struc. Biol.* **2009**, *19*, 39−49.

(10) Rousseau, F.; Schymkowitz, J. W. H.; Itzhaki, L. S. The Unfolding Story of Three-Dimensional Domain Swapping. *Structure* **2003**, *11*, 243−251.

(11) Liu, L.; Byeon, I.-J. L.; Bahar, I.; Gronenborn, A. M. Domain Swapping Proceeds via Complete Unfolding: A $^{19}$F- and $^{1}$H-NMR Study of the Cyanovirin-N Protein. *J. Am. Chem. Soc.* **2012**, *134*, 4229−4235.

(12) Liu, Z.; Huang, Y. Evidences for the Unfolding Mechanism of Three-Dimensional Domain Swapping: The Unfolding Mechanism of 3D Domain Swapping. *Protein Sci.* **2013**, *22*, 280−286.

(13) Mascarenhas, N. M.; Gosavi, S. Understanding Protein Domain-Swapping Using Structure-Based Models of Protein Folding. *Prog. Biophys. Mol. Biol.* **2017**, *128*, 113−120.
(14) Liu, Y.; Eisenberg, D. 3D Domain Swapping: As Domains Continue to Swap. *Protein Sci.* **2002**, *11*, 1285−1299.
(15) Shiga, S.; Yamanaka, M.; Fujiwara, W.; Hirota, S.; Goda, S.; Makabe, K. Domain-Swapping Design by Polyproline Rod Insertion. *ChemBioChem* **2019**, *20*, 2454−2457.
(16) Nandwani, N.; Surana, P.; Negi, H.; Mascarenhas, N. M.; Udgaonkar, J. B.; Das, R.; Gosavi, S. A Five-Residue Motif for the Design of Domain Swapping in Proteins. *Nat. Commun.* **2019**, *10*, No. 452.
(17) Mizuno, H.; Fujimoto, Z.; Koizumi, M.; Kano, H.; Atoda, H.; Morita, T. Structure of Coagulation Factors IX/X-Binding Protein, a Heterodimer of C-Type Lectin Domains. *Nat. Struct. Biol.* **1997**, *4*, 438−441.
(18) Chen, L.-Y.; Huang, Y.-C.; Huang, S.-T.; Hsieh, Y.-C.; Guan, H.-H.; Chen, N.-C.; Chuankhayan, P.; Yoshimura, M.; Tai, M.-H.; Chen, C.-J. Domain Swapping and SMYD1 Interactions with the PWWP Domain of Human Hepatoma-Derived Growth Factor. *Sci. Rep.* **2018**, *8*, No. 287.
(19) Park, C. K.; Joshi, H. K.; Agrawal, A.; Ghare, M. I.; Little, E. J.; Dunten, P. W.; Bitinaite, J.; Horton, N. C. Domain Swapping in Allosteric Modulation of DNA Specificity. *PLoS Biol.* **2010**, *8*, No. e1000554.
(20) Zhou, X.; Wang, H.; Zhang, Y.; Gao, L.; Feng, Y. Alteration of Substrate Specificities of Thermophilic $\alpha/\beta$ Hydrolases through Domain Swapping and Domain Interface Optimization. *Acta Biochim. Biophys. Sin.* **2012**, *44*, 965−973.
(21) Karchin, J. M.; Ha, J.-H.; Namitz, K. E.; Cosgrove, M. S.; Loh, S. N. Small Molecule-Induced Domain Swapping as a Mechanism for Controlling Protein Function and Assembly. *Sci. Rep.* **2017**, *7*, No. 44388.
(22) Nussinov, R.; Xu, D.; Tsai, C.-J. Mechanism and Evolution of Protein Dimerization: Protein Dimerization. *Protein Sci.* **1998**, *7*, 533−544.
(23) Bennett, M. J.; Sawaya, M. R.; Eisenberg, D. Deposition Diseases and 3D Domain Swapping. *Structure* **2006**, *14*, 811−824.
(24) Janowski, R.; Kozak, M.; Abrahamson, M.; Grubb, A.; Jaskolski, M. 3D Domain-Swapped Human Cystatin C with Amyloidlike Intermolecular $\beta$-Sheets. *Proteins* **2005**, *61*, 570−578.
(25) Knaus, K. J.; Morillas, M.; Swietnicki, W.; Malone, M.; Surewicz, W. K.; Yee, V. C. Crystal Structure of the Human Prion Protein Reveals a Mechanism for Oligomerization. *Nat. Struct Biol.* **2001**, *8*, 770−774.
(26) Markova, K.; Chmelova, K.; Marques, S. M.; Carpentier, P.; Bednar, D.; Damborsky, J.; Marek, M. Decoding the Intricate Network of Molecular Interactions of a Hyperstable Engineered Biocatalyst. *Chem. Sci.* **2020**, *11*, 11162−11178.
(27) Beerens, K.; Mazurenko, S.; Kunka, A.; Marques, S. M.; Hansen, N.; Musil, M.; Chaloupkova, R.; Waterman, J.; Brezovsky, J.; Bednar, D.; Prokop, Z.; Damborsky, J. Evolutionary Analysis as a Powerful Complement to Energy Calculations for Protein Stabilization. *ACS Catal.* **2018**, *8*, 9420−9428.
(28) Sanchez-Ruiz, J. M. Theoretical Analysis of Lumry-Eyring Models in Differential Scanning Calorimetry. *Biophys. J.* **1992**, *61*, 921−935.
(29) Mazurenko, S.; Kunka, A.; Beerens, K.; Johnson, C. M.; Damborsky, J.; Prokop, Z. Exploration of Protein Unfolding by Modelling Calorimetry Data from Reheating. *Sci. Rep.* **2017**, *7*, No. 16321.
(30) Mazurenko, S.; Stourac, J.; Kunka, A.; Nedeljković, S.; Bednar, D.; Prokop, Z.; Damborsky, J. CalFitter: A Web Server for Analysis of Protein Thermal Denaturation Data. *Nucleic Acids Res.* **2018**, *46*, W344−W349.
(31) Chaloupková, R.; Sykorova, J.; Prokop, Z.; Jesenska, A.; Monincova, M.; Pavlova, M.; Tsuda, M.; Nagata, Y.; Damborsky, J. Modification of Activity and Specificity of Haloalkane Dehalogenase from Sphingomonas Paucimobilis UT26 by Engineering of Its Entrance Tunnel. *J. Biol. Chem.* **2003**, *278*, 52622−52628.
(32) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. Redesigning Dehalogenase Access Tunnels as a Strategy for Degrading an Anthropogenic Substrate. *Nat. Chem. Biol.* **2009**, *5*, 727−733.
(33) Chovancova, 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*, No. e1002708.
(34) 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*, 1970−1989.
(35) Musil, M.; Stourac, J.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Zendulka, J.; Martinek, T.; Bednar, D.; Damborsky, J. FireProt: Web Server for Automated Design of Thermostable Proteins. *Nucleic Acids Res.* **2017**, *45*, W393−W399.
(36) Zegers, I.; Deswarte, J.; Wyns, L. Trimeric Domain-Swapped Barnase. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 818−822.
(37) Bergdoll, M.; Remy, M.-H.; Cagnon, C.; Masson, J.-M.; Dumas, P. Proline-Dependent Oligomerization with Arm Exchange. *Structure* **1997**, *5*, 391−401.
(38) Heinemann, P. M.; Armbruster, D.; Hauer, B. Active-Site Loop Variations Adjust Activity and Selectivity of the Cumene Dioxygenase. *Nat. Commun.* **2021**, *12*, No. 1095.
(39) Bennett, M. J.; Eisenberg, D. The Evolving Role of 3D Domain Swapping in Proteins. *Structure* **2004**, *12*, 1339−1341.
(40) Jaskólski, M. 3D Domain Swapping, Protein Oligomerization, and Amyloid Formation. *Acta Biochim. Pol.* **2001**, *48*, 807−827.
(41) Fraser, N. J.; Liu, J.-W.; Mabbitt, P. D.; Correy, G. J.; Coppin, C. W.; Lethier, M.; Perugini, M. A.; Murphy, J. M.; Oakeshott, J. G.; Weik, M.; Jackson, C. J. Evolution of Protein Quaternary Structure in Response to Selective Pressure for Increased Thermostability. *J. Mol. Biol.* **2016**, *428*, 2359−2371.
(42) Liu, Y.; Gotte, G.; Libonati, M.; Eisenberg, D. A Domain-Swapped RNase A Dimer with Implications for Amyloid Formation. *Nat. Struct. Biol.* **2001**, *8*, 211−214.
(43) Baier, F.; Hong, N.; Yang, G.; Pabis, A.; Miton, C. M.; Barrozo, A.; Carr, P. D.; Kamerlin, S. C.; Jackson, C. J.; Tokuriki, N. Cryptic Genetic Variation Shapes the Adaptive Evolutionary Potential of Enzymes. *eLife* **2019**, *8*, No. e40789.
(44) Koga, N.; Koga, R.; Liu, G.; Castellanos, J.; Montelione, G. T.; Baker, D. Role of Backbone Strain in de Novo Design of Complex $\alpha/\beta$ Protein Structures. *Nat. Commun.* **2021**, *12*, No. 3921.
(45) Nurizzo, D.; Mairs, T.; Guijarro, M.; Rey, V.; Meyer, J.; Fajardo, P.; Chavanne, J.; Biasci, J.-C.; McSweeney, S.; Mitchell, E. The ID23-1 Structural Biology Beamline at the ESRF. *J. Synchrotron. Rad.* **2006**, *13*, 227−238.
(46) Kabsch, W. XDS. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−132.
(47) Evans, P. R.; Murshudov, G. N. How Good Are My Data and What Is the Resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013**, *69*, 1204−1214.
(48) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(49) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. PHENIX: A Comprehensive Python-Based System for Macromolecular Structure Solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.
(50) Lahoda, M.; Mesters, J. R.; Stsiapanava, A.; Chaloupkova, R.; Kuty, M.; Damborsky, J.; Kuta Smatanova, I. Crystallographic Analysis of 1,2,3-Trichloropropane Biodegradation by the Haloalkane

Dehalogenase DhaA31. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2014**, *70*, 209−217.
(51) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. Towards Automated Crystallographic Structure Refinement with Phenix.Refine. *Acta Crystallogr. D: Biol. Crystallogr.* **2012**, *68*, 352−367.
(52) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126−2132.
(53) Williams, C. J.; Headd, J. J.; Moriarty, N. W.; Prisant, M. G.; Videau, L. L.; Deis, L. N.; Verma, V.; Keedy, D. A.; Hintze, B. J.; Chen, V. B.; Jain, S.; Lewis, S. M.; Arendall, W. B.; Snoeyink, J.; Adams, P. D.; Lovell, S. C.; Richardson, J. S.; Richardson, D. C. MolProbity: More and Better Reference Data for Improved All-Atom Structure Validation. *Protein Sci.* **2018**, *27*, 293−315.
(54) *The PyMOL Molecular Graphics System*; Schrödinger, LLC.
(55) Krissinel, E.; Henrick, K. Secondary-Structure Matching (SSM), a New Tool for Fast Protein Structure Alignment in Three Dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2256−2268.
(56) wwPDB consortium; Burley, S. K.; Berman, H. M.; Bhikadiya, C.; Bi, C.; Chen, L.; Costanzo, L. D.; Christie, C.; Duarte, J. M.; Dutta, S.; Feng, Z.; Ghosh, S.; Goodsell, D. S.; Green, R. K.; Guranovic, V.; Guzenko, D.; Hudson, B. P.; Liang, Y.; Lowe, R.; Peisach, E.; Periskova, I.; Randle, C.; Rose, A.; Sekharan, M.; Shao, C.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J.; Young, J.; Zardecki, C.; Zhuravleva, M.; Kurisu, G.; Nakamura, H.; Kengaku, Y.; Cho, H.; Sato, J.; Kim, J. Y.; Ikegawa, Y.; Nakagawa, A.; Yamashita, R.; Kudou, T.; Bekker, G.-J.; Suzuki, H.; Iwata, T.; Yokochi, M.; Kobayashi, N.; Fujiwara, T.; Velankar, S.; Kleywegt, G. J.; Anyango, S.; Armstrong, D. R.; Berrisford, J. M.; Conroy, M. J.; Dana, J. M.; Deshpande, M.; Gane, P.; Gáborová, R.; Gupta, D.; Gutmanas, A.; Koča, J.; Mak, L.; Mir, S.; Mukhopadhyay, A.; Nadzirin, N.; Nair, S.; Patwardhan, A.; Paysan-Lafosse, T.; Pravda, L.; Salih, O.; Sehnal, D.; Varadi, M.; Vařeková, R.; Markley, J. L.; Hoch, J. C.; Romero, P. R.; Baskaran, K.; Maziuk, D.; Ulrich, E. L.; Wedell, J. R.; Yao, H.; Livny, M.; Ioannidis, Y. E. Protein Data Bank: The Single Global Archive for 3D Macromolecular Structure Data. *Nucleic Acids Res.* **2019**, *47*, D520−D528.
(57) Schuck, P. Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **2000**, *78*, 1606−1619.
(58) Brautigam, C. A. Calculations and Publication-Quality Illustrations for Analytical Ultracentrifugation Data. In *Methods in Enzymology*; Elsevier, 2015; Vol. *562*, pp 109−133.
(59) Iwasaki, I.; Utsumi, S.; Ozawa, T. New Colorimetric Determination of Chloride Using Mercuric Thiocyanate and Ferric Ion. *Bull. Chem. Soc. Jpn* **1952**, *25*, 226.
(60) Buryska, T.; Vasina, M.; Gielen, F.; Vanacek, P.; van Vliet, L.; Jezek, J.; Pilat, Z.; Zemanek, P.; Damborsky, J.; Hollfelder, F.; Prokop, Z. Controlled Oil/Water Partitioning of Hydrophobic Substrates Extending the Bioanalytical Applications of Droplet-Based Microfluidics. *Anal. Chem.* **2019**, *91*, 10008−10015.
(61) 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*, 20−29.
(62) Franke, D.; Petoukhov, M. V.; Konarev, P. V.; Panjkovich, A.; Tuukkanen, A.; Mertens, H. D. T.; Kikhney, A. G.; Hajizadeh, N. R.; Franklin, J. M.; Jeffries, C. M.; Svergun, D. I. ATSAS 2.8: A Comprehensive Data Analysis Suite for Small-Angle Scattering from Macromolecular Solutions. *J. Appl. Crystallogr.* **2017**, *50*, 1212−1225.
(63) Valentini, E.; Kikhney, A. G.; Previtali, G.; Jeffries, C. M.; Svergun, D. I. SASBDB, a Repository for Biological Small-Angle Scattering Data. *Nucleic Acids Res.* **2015**, *43*, D357−D363.
(64) Stejskal, K.; Potesil, D.; Zdrahal, Z. Suppression of Peptide Sample Losses in Autosampler Vials. *J. Proteome Res.* **2013**, *12*, 3057−3062.
(65) Stieger, C. E.; Doppler, P.; Mechtler, K. Optimized Fragmentation Improves the Identification of Peptides Cross-Linked by MS-Cleavable Reagents. *J. Proteome Res.* **2019**, *18*, 1363−1370.
(66) Giese, S. H.; Fischer, L.; Rappsilber, J. A Study into the Collision-Induced Dissociation (CID) Behavior of Cross-Linked Peptides. *Mol. Cell. Proteomics* **2016**, *15*, 1094−1104.
(67) Pace, C. N. Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. In *Methods in Enzymology*; Elsevier, 1986; Vol. *131*, pp 266−280.
(68) Niklasson, M.; Andresen, C.; Helander, S.; Roth, M. G. L.; Zimdahl Kahlin, A.; Lindqvist Appell, M.; Mårtensson, L.-G.; Lundström, P. Robust and Convenient Analysis of Protein Thermal and Chemical Stability: Robust and Convenient Analysis of Protein Stability. *Protein Sci.* **2015**, *24*, 2055−2062.
(69) Case, D. A.; Berryman, J. T.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H.; Goetz, A. W.; Gusarov, S.; Homeyer, N.; Janowski, P.; Kaus, J.; Kolossváry, I.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Luchko, T.; Luo, R.; Madej, B.; Merz, K. M.; Paesani, F.; Roe, D. R.; Roitberg, A.; Sagui, C.; Salomon-Ferrer, R.; Seabra, G.; Simmerling, C. L.; Smith, W.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Kollman, P. A. *AMBER 14*; University of California: San Francisco, 2014.
(70) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating PKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368−W371.
(71) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(72) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(73) Götz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **2012**, *8*, 1542−1555.
(74) Le Grand, S.; Götz, A. W.; Walker, R. C. SPFP: Speed without Compromise—A Mixed Precision Model for GPU Accelerated Molecular Dynamics Simulations. *Comput. Phys. Commun.* **2013**, *184*, 374−380.
(75) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·log(N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(76) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(77) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(78) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33−38.


https://doi.org/10.1021/acscatal.1c03343
ACS Catal. 2021, 11, 12864−12885