ChemPubSoc Europe

DOI: 10.1002/cbic.201300085

Decoupled Roles for the Atypical, Bifurcated Binding Pocket of the ybfF Hydrolase

Elizabeth E. Ellis,[a] Chinessa T. Adkins,[b] Natalie M. Galovska,[a] Luke D. Lavis,[b] and R. Jeremy Johnson*[a]

Serine hydrolases have diverse intracellular substrates, biological functions, and structural plasticity, and are thus important for biocatalyst design. Amongst serine hydrolases, the recently described ybfF enzyme family are promising novel biocatalysts with an unusual bifurcated substrate-binding cleft and the ability to recognize commercially relevant substrates. We characterized in detail the substrate selectivity of a novel ybfF enzyme from *Vibrio cholerae* (Vc-ybfF) by using a 21-member library of fluorogenic ester substrates. We assigned the roles of the two substrate-binding clefts in controlling the substrate selectivity and folded stability of Vc-ybfF by comprehensive substitution analysis. The overall substrate preference of Vc-ybfF was for short polar chains, but it retained significant activity with a range of cyclic and extended esters. This broad substrate specificity combined with the substitutional analysis demonstrates that the larger binding cleft controls the substrate specificity of Vc-ybfF. Key selectivity residues (Tyr116, Arg120, Tyr209) are also located at the larger binding pocket and control the substrate specificity profile. In the structure of ybfF the narrower binding cleft contains water molecules prepositioned for hydrolysis, but based on substitution this cleft showed only minimal contribution to catalysis. Instead, the residues surrounding the narrow binding cleft and at the entrance to the binding pocket contributed significantly to the folded stability of Vc-ybfF. The relative contributions of each cleft of the binding pocket to the catalytic activity and folded stability of Vc-ybfF provide a valuable map for designing future biocatalysts based on the ybfF scaffold.

Introduction

Serine hydrolases (EC: 3.1) constitute a large class of ubiquitous and highly conserved cellular enzymes with broad substrate specificity and broad biological function.[1] The diverse members of the serine hydrolase family are grouped together based on their conserved protein fold (α/β hydrolase protein fold) and conserved catalytic machinery (a catalytic triad of serine, histidine, and an acidic residue).[2] This conserved catalytic triad can, however, catalyze a range of chemical transformations with a diverse group of biological substrates, depending on the structural orientation of the active site and binding pocket.[3] This plasticity in the reactivity of serine hydrolases has been exploited by protein engineers to design novel and commercially applicable reactivity on serine hydrolase scaffolds.[3a,b,4] Bioengineered serine hydrolases have found applications in variety of processes, including biofuel production, pharmaceutical synthesis, and commercial resolution of chiral esters.[3a,b,5] Expanded understanding of the serine hydrolase structure and comprehensive databases of engineered hydrolases have greatly aided the development of novel serine hydrolase biocatalysts.[1b,3c,6] New structural and enzymatic families of serine hydrolases, however, continue to be discovered, thus presenting new scaffolds and reactivities for the continued development of serine hydrolase biocatalysts.[3c]

Amongst the serine hydrolase family, the serine hydrolase ybfF from *Escherichia coli* (Ec-ybfF) was shown to be a promising target for biocatalyst development, as it catalyzes important chemical transformations and contains an unusual bifurcated binding pocket that can accommodate a range of substrates.[7] The ybfF protein is also highly conserved across bacteria, with homologues in a variety of pathogenic bacteria, including *Vibrio cholerae*, but it has low sequence similarity ($<20\%$) with other serine hydrolases.[7a] The crystal structure of Ec-ybfF confirmed that it is a serine hydrolases with the canonical α/β hydrolase fold and a GxSxG catalytic motif, where the central serine is the catalytic nucleophile required for the hydrolysis reaction.[7a]

The structure also revealed an unusual, buried substrate-binding pocket that bifurcates into two distinct binding clefts (Figure 1). A charged malonate molecule was bound tightly into the broader cleft of the substrate-binding pocket, in close proximity to the catalytic serine, thus leading to the classification of ybfF as a malonyl-CoA thioesterase (Figure 1).[7a] The narrower cleft of the substrate-binding pocket contained three highly ordered water molecules, but was measured to contain only space sufficient to accommodate an unbranched four-atom chain.[7a] Additionally, the structure of Ec-ybfF showed that the enzyme contains a separate helical cap domain atop

[a] E. E. Ellis, N. M. Galovska, Dr. R. J. Johnson
Department of Chemistry, Butler University
4600 Sunset Ave, Indianapolis, IN 46208-3443 (USA)
E-mail: rjjohns1@butler.edu

[b] Dr. C. T. Adkins, Dr. L. D. Lavis
Janelia Farm Research Campus, The Howard Hughes Medical Institute
19700 Helix Dr, Ashburn, VA 20147-2408 (USA)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300085.

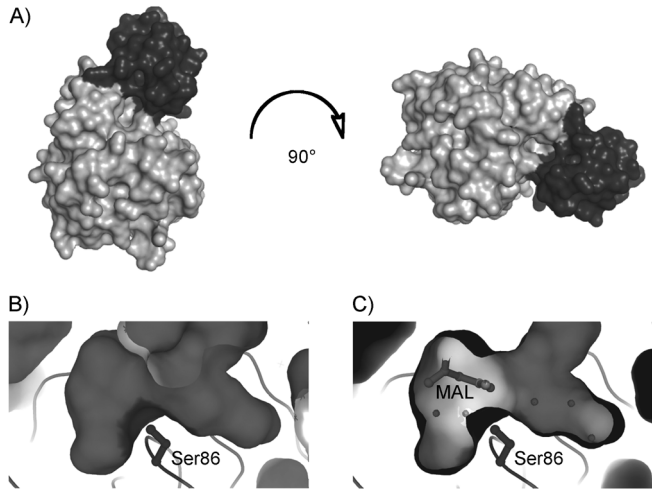


Figure 1. Structure of ybfF and its binding pocket. A) General architecture of *E. coli* ybfF (PDB ID: 3BF8) showing α/β hydrolase domain (light grey) and helical cap domain (dark grey). B) Dual-cleft substrate binding pocket. The exterior surface has been removed to show the nucleophilic serine (Ser86) and the bifurcated binding pocket of ybfF: broader binding cleft (left) and narrower binding cleft (right). C) Interior view of the binding pocket with bound malonate (MAL, ball and stick) and water molecules (spheres). All figures were made with PyMOL.

the α/β hydrolase domain and likely serves to recognize coenzyme A.[7a] A similar structural architecture (hydrolase and cap domains) was also observed in Rv0045c, a ybfF homologue from *Mycobacterium tuberculosis*. In the structure of Rv0045c (PDB ID: 3P2M) the majority of the cap domain was, however, disordered, thus making comparisons of the substrate-binding pocket architecture between the bacterial homologues difficult.[8] The unique nature of the ybfF binding pocket was also modeled to facilitate the rapid, stereospecific hydrolysis by ybfF of 1,2-*O*-isopropyideneglycerol (IPG) esters, a potential pharmaceutical starting material in the production of β-blockers, prostaglandins, and leukotrienes.[7b] Focused substitutions around the entrance to the binding cleft endowed Ec-ybfF with the highest known enantioselectivity for *R*-IPG butyrate and caprylate.[9]

Here, we characterized the role of the dual substrate-binding pocket of a novel ybfF enzyme from V. cholera (Vc-ybfF) to determine the enzymatic activity and substrate specificity of ybfF proteins. Using a library of 21 fluorogenic ester substrates, we measured the substrate specificity of Vc-ybfF, by varying chain length, steric hindrance, polarity, and degree of saturation. Against the optimal substrates, the dual substrate-binding surface of Vc-ybfF was then mapped by using alanine-scanning mutagenesis to assess the relative contributions of different residues to the enzymatic activity and structural stability of ybfF. Focused substitutions of key residues of the substrate-binding pocket were then used to identify residues essential to the catalytic activity of Vc-ybfF. The global picture of the substrate-binding pocket of ybfF shows that residues key to catalytic activity and thermal stability of ybfF are located at the interface between the α/β hydrolase domain and helical cap domain. Each of the dual lobes of the substrate-binding pocket provides distinct contributions to the overall activity and stability of ybfF, thus suggesting the ability to decouple the thermal stability and substrate selectivity of ybfF in future biocatalyst designs.

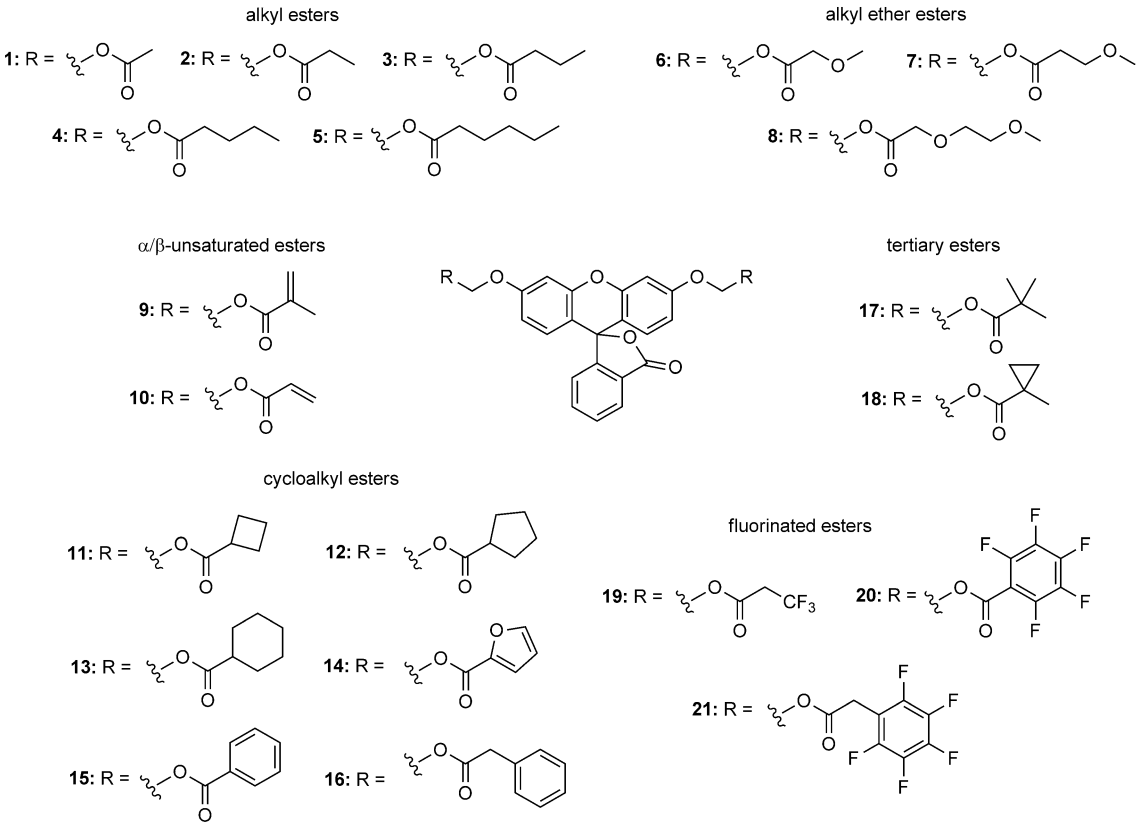
Results

Serine hydrolase substrate library

The binding pocket and active site of ybfF is located at the interface between the α/β hydrolase domain and the cap domain (Figure 1). The entrance to the bifurcated binding pocket contains the nucleophilic serine (Ser86), which is positioned to initiate the classic catalytic mechanism of serine hydrolases (Figure 1 B and C). The larger binding cleft (Figure 1 B and C, left) contains a well-positioned malonate in the Ec-ybfF structure, which is held in position by hydrogen bonds to Arg120 and Tyr209.[7a] The larger cleft was proposed to accommodate the incoming carboxylic acid portion of the ester substrate, with space to accommodate substrates of up to six carbons.[7a] The larger cleft was also modeled to fit IPG esters of four and eight carbons with stereoselectivity for the *R* isomer.[7b] In comparison, the narrower binding cleft (right side in Figure 1 B and C) has sufficient space to accommodate only straight chains of approximately four atoms and contains three well-positioned water molecules in the Ec-ybfF structure.[7a] Along the exposed interface between the α/β hydrolase domain and the cap domain, a hydrophobic cleft is present that can accommodate longer chains, including palmitoyl-CoA or the CoA moiety of malonyl-CoA. The distinct binding surfaces and diversity of currently identified substrates for ybfF complicate the assignment of the function of ybfF and complicate manipulation of the ybfF scaffold in biocatalyst designs.

Given the current list of substrates hydrolyzed by ybfF, we set out to clarify the substrate specificity of Vc-ybfF by using a library of diverse fluorogenic esters (Scheme 1). Each of the fluorogenic ester substrates contains an invariant fluorescein core with various R-groups attached to the phenolic oxygens as acyloxymethyl ethers. The acyloxymethyl ether attachment provides lower background fluorescence and increased sensitivity over traditional esterase substrates, thus permitting rapid, precise measurement of full enzyme kinetics in microplate format.[10] A library of 21 acyloxymethyl ether substrates (Scheme 1) was synthesized by a previously published procedure,[10a,b] and correct synthesis was confirmed by NMR and LC-MS (see the Experimental Section).

The ester functionalities of the fluorogenic substrates were chosen to mimic distinct biological classes of ester-based substrates, including alkyl esters, polar esters (alkyl ether esters), unsaturated esters, aryl esters (cycloalkyl esters), tertiary esters, and fluoroesters (Scheme 1).[1a] The library also focused on basic building blocks for each of the classes, by favoring simple ester structures over more complex, branched structures. Certain classes in the library (alkyl esters and cycloalkyl esters) were more highly populated than other classes, based on the relative abundance of similar substrates amongst natural esters and their potential for clarifying the binding pocket structure of ybfF. A limitation of the current library is the inabil-



Scheme 1. Library of fluorogenic esterase substrates. Each substrate is composed of diacyloxymethyl ether fluorescein (center) and R groups. The differing R groups are organized into classes based on chemical functionality. All substrates were synthesized as described previously[10b,c] (see the Experimental Section).

ity to classify serine hydrolases into the subcategories carboxyl esterase and lipase, as solubility limitations restricted the maximum length of the alkyl chain to six carbons, although seven atoms are possible after addition of ethers within the chain (substrates **6**–**8**). To confirm the specificity of Vc-ybfF with substrates of longer carbon-chain length, a commercial series of *p*-nitrophenyl substrates of varying lengths was also tested for reactivity. The substrate-specificity profile of Vc-ybfF was then defined by mapping its precise enzymatic activity toward the entire 21-member fluorogenic library and the four *p*-nitrophenyl substrates.

Substrate specificity of ybfF enzymes

We characterized the activity of a homologue of Ec-ybfF to expand the limited knowledge about this unusual enzyme family. We chose to study ybfF from *V. cholerae* (Vc-ybfF) because of the global health epidemic of cholera and the sequence conservation (45% identity) between Vc-ybfF and Ec-ybfF. Vc-ybfF was cloned into a bacterial expression vector, expressed in *E. coli*, and purified to homogeneity by Ni-metal affinity chromatography (Figure S1 in the Supporting Information). The wild-type Vc-ybfF enzyme had an unfolding curve (T_m = 48 °C; Table 1) and size exclusion chromatogram indicative of a well-folded enzyme (Figure S3). Vc-ybfF appeared to be an extended dimer in solution; this is consistent with the crystallographic state of Ec-ybfF (Figure S3). Vc-ybfF was highly active against a variety of ester-based substrates, with a broad pH tolerance and classic Michaelis–Menten kinetics (Figure S2). For each of the 21 fluorogenic substrates, the k_{cat}, K_m, and k_{cat}/K_m values were determined by microplate analysis of the purified enzyme (Table S1). Enzymatic efficiency (k_{cat}/K_m) was used for comparisons between the different substrates (Figure 2).

The global landscape of Vc-ybfF substrate specificity showed high selectivity toward alkyl ether esters (substrates **6**–**8**; Figure 2). Vc-ybfF cleaved the alkyl ether esters with a catalytic efficiency more than ten times higher than for the other fluorogenic substrates (Table S1). Within the alkyl ether esters, Vc-ybfF had the highest activity with **6**, which contains four atoms, consistent with the predicted size selectivity of ybfF (four atom units; Figure 2C).[7a] The selectivity of Vc-ybfF for polar esters matched the significant activity of Ec-ybfF toward IPG esters.[7b]

Although the activity was significantly lower with the remaining library members, distinct patterns of substrate selectivity emerged when comparing the enzymatic efficiency values within different subgroups (Figure 2). Within alkyl esters (**1**–**5**), Vc-ybfF shows bell-shaped selectivity with maximal activity against C4 (**3**; Figure 2B), again confirming the preference of ybfF for substrates of four atoms.[7a] This level of detail about substrate selectivity was not available with the *p*-nitrophenyl substrates, but the global picture again confirms the selectivity of ybfF as highest with four-carbon substrates (*p*-nitrophenyl butyrate; Figure 2E). Similar size preferences were

Table 1. Biochemical characterization of Vc-ybfF and variants.

ybfF variant	Substrate **3** K_m [μM][a]	Substrate **3** k_{cat}/K_m [$M^{-1}s^{-1}$]	Substrate **6** K_m [μM]	Substrate **6** k_{cat}/K_m [$M^{-1}s^{-1}$]	T_m [°C][c]
WT	6.6±0.8	210±40	25±2	6500±500	48±0
H20A	3.3±0.5	2.7±0.4	2.2±0.2	1500±200	42.5±0.9
N28A	3.6±0.5	120±20	16±4	3300±900	46.5±0
S86A	n.d.[b]	n.d.	21±9	0.29±0.13	49.5±0
M87A	0.18±0.08	94±42	1.5±0.3	5200±1100	44±2
K90A	n.d.	n.d.	1.0±0.1	23±2	41.5±0.9
D110A	5±1	0.43±0.10	2.8±0.5	57±10	43.5±0
Y116A	5±1	8.8±1.9	4.5±0.4	3300±300	43.5±0
R120A	3.2±0.3	58±6	23±3	1600±200	45.5±0.9
H121A	7.3±0.9	84±12	16±4	2200±600	42±0
F158A	6.2±0.4	140±10	13±0.8	5300±300	45±0
Y209A	1.9±0.1	0.50±0.03	3.9±0.9	88±21	45.5±0.9
H235A	19±4	0.074±0.020	0.5±0.2	58±23	44±0.9
W236A	27±77	0.085±0.30	5±1	12±2	39±0
R120K	2.1±0.1	130±10	12±0.7	4500±300	45±0
Y209F	1.8±0.2	170±20	6.9±0.8	5000±600	45±0

[a] Kinetic constants for substrates **3** and **6** were determined by measuring the increase in fluorogenic enzyme substrate fluorescence over time.[10a] Data were fitted to a standard Michaelis–Menten equation to determine values for k_{cat}, K_M, and k_{cat}/K_M. Kinetic measurements for each substrate were repeated three times; values are given as the mean±SEM, [b] n.d.: measurement below the detection limit. [c] Values for T_m were determined by following the change in SYPRO Orange fluorescence with increasing temperature. Melting curve experiments were repeated three times for each variant, and the T_m values are reported as the mean±SEM.

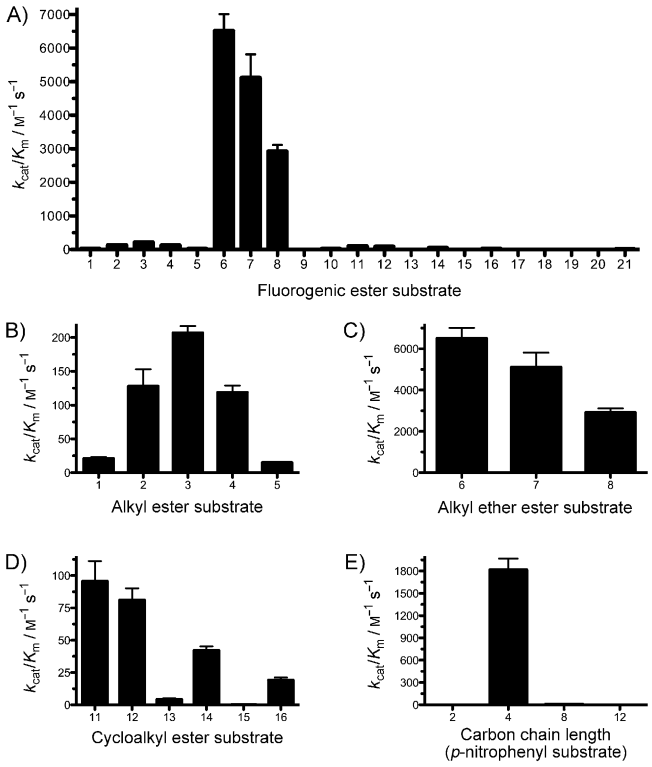


Figure 2. Kinetic characterization of Vc-ybfF with the library of fluorogenic ester substrates. A) Catalytic specificity (k_{cat}/K_M) of Vc-ybfF with each of the 21 substrates (Scheme 1). To illustrate patterns within classes of substrates, important subclasses are compared: B) alkyl esters, C) alkyl ether ester, D) cycloalkyl ester, and E) *p*-nitrophenyl esters. Detailed kinetic results for each substrate are provided in Table S1.

found for the cycloalkyl esters: substrates with rings of four or five carbons (cyclobutyl (**11**), cyclopentyl (**12**), and furanyl (**14**)) showed significantly higher activity than substrates with six-membered rings (Figure 2D). The selectivity for smaller cycloalkyls was independent of aromaticity or flexibility in ring attachment, although adding one extra carbon between the ester and phenyl ring (**15** versus **16**) did increase the activity over 50-fold (Table S1).

For a variety of substrates, including α/β unsaturated, tertiary, and fluorinated esters, Vc-ybfF had low but measurable activity (Figure 2). This low activity is common to the majority of bacterial hydrolases, as only a small class of enzymes readily catalyzes the hydrolysis of fluorinated esters,[11] and tertiary esters are often used as challenging substrates in directed evolution studies.[12] Given the unusual binding pocket structure of ybfF and its known ability to catalyze commercially relevant ester hydrolysis reactions,[7b] these low-activity substrates could be used in future experiments to test the adaptability of the ybfF binding pocket to new substrates and reactions through directed-evolution studies.

Structural determinants of ybfF enzymatic activity

The global picture of the substrate specificity of ybfF shows that it prefers short, polar esters. Given the sizes of each cleft of the ybfF binding pocket, this substrate-specificity profile does not reveal which cleft of the ybfF binding pocket is interacting with the substrates. To differentiate the roles of each cleft of the Vc-ybfF binding pocket, comprehensive site-directed mutagenesis was performed on each of cleft. As an initial assessment of their contributions to the substrate selectivity, 13 amino acid residues surrounding the binding pocket were substituted with alanine to scan the binding surface for groupings of residues that make significant contributions to the enzymatic activity of ybfF (Figure 3A).

These 13 residues were chosen based upon their proximity to the binding pocket, their potential roles in catalysis, their distribution across both clefts of the binding pocket, and their conservation across species (Figure 3A and B). The majority of the 13 residues are highly conserved across ybfF homologues from bacteria to fungi to humans, with only highly conservative substitutions across homologues (Figure 3B). Amongst the 13 residues, the least-conserved binding-pocket residues are proposed to be involved in substrate recognition (Tyr116, Arg120, and His121) where diverse residues are present at these positions in mammalian homologues of ybfF (Figure 3B and Table S3).

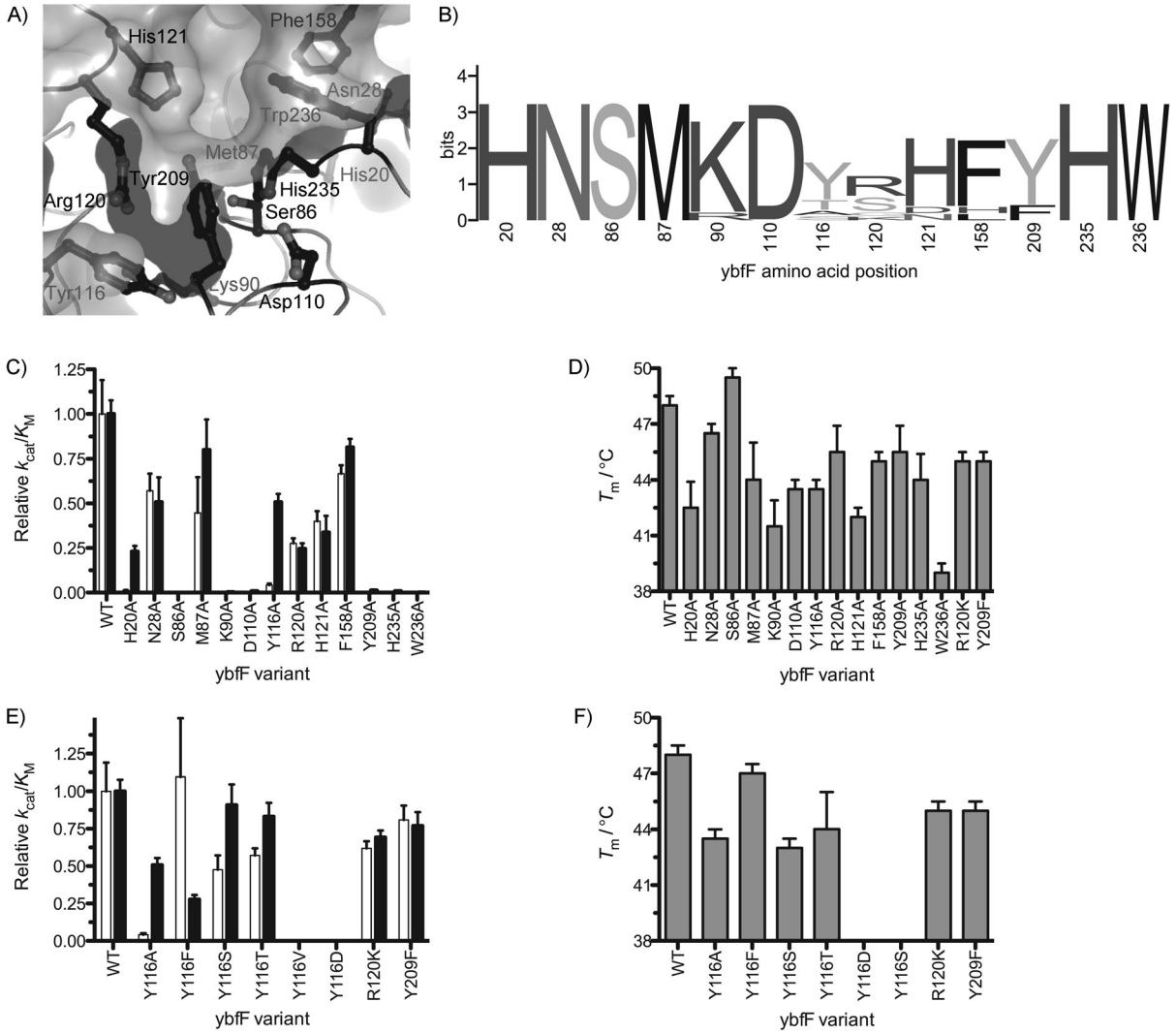


Figure 3. Substrate-binding pocket residues important to the catalytic activity and thermal stability of Vc-ybfF. A) Residues (ball and stick) were substituted individually with alanine, and the effects on catalytic activity and thermal stability were determined. B) Conservation of substrate binding residues across bacterial and mammalian homologues of ybfF. Sequences were aligned with ClustalW,[25b] and relative weightings were evaluated with Weblogo.[25a] (Detailed sequence analysis in Table S3 and Figure S4.) C) Relative catalytic activity of Vc-ybfF variants with substrates **3** and **6.** D) Thermal stability of Vc-ybfF variants, determined by measuring the increase in SYPRO Orange fluorescence in response to increasing temperature. Results are shown with standard error values. (Detailed kinetic and thermal stability analysis in Table 1.) E) Relative catalytic activity and F) thermal stability of Y116X variants. Experiments conducted as for (C) and (D). (Detailed kinetic and thermal stability analysis in Table S2.)

Vc-ybfF variants were then constructed with single alanine substitutions of each of the 13 amino acids by site-directed mutagenesis, and purified to homogeneity (Figure S1 and Table S4). Each of the 13 Vc-ybfF variants was tested for enzymatic activity with two of the highly active fluorogenic substrates (**3** and **6**) from the general screen (Figure 3C and Table 1). With one important exception (Tyr116), the relative activity differences between the two substrates for the 13 Vc-ybfF variants were minor (within experimental error; Figure 3C). Given the similarity in activity with the two substrates, comparisons of relative enzymatic activity between the binding pocket variants were conducted from the results with **6**.

Many of the residues significantly affecting the catalytic activity were predicted based on the structure of Ec-ybfF, including complete loss of activity upon substitution of the nucleophilic serine (Ser86) and the general acid-base (His235; Figure 3C). The third amino acid in the catalytic triad (Asp110), which forms hydrogen bonds with His235, had a lesser role in catalysis. The decreased importance of the acidic amino acid in the catalytic triad has been observed in other hydrolases, and was predicted for Ec-ybfF based on the shifted orientation of Asp110 in relation to His235.[7a,13] Serine hydrolases also utilize an oxyanion hole (in addition to the catalytic triad) to stabilize the negative charge on the tetrahedral transition state.[2,13b] The oxyanion hole in Ec-ybfF was proposed to be formed by the backbone nitrogens of Leu22 and Met87 (Vc-ybfF numbering), and this was supported by modeling of bound IPG substrates.[7a,9] Although Leu22 was not tested in the present study, substitution of Met87 with alanine did significantly reduce the overall catalytic efficiency of ybfF, thus confirming its important role in catalysis.

In addition to substitution of the expected catalytic triad residues, substitution of four additional residues (His20, Lys90, Tyr209, and Trp236) in Vc-ybfF led to complete loss of catalytic activity with both fluorogenic substrates. Two of these residues (Lys90 and Tyr209) interact directly with the bound malonate in Ec-ybfF, and this suggests direct roles for Lys90 and Tyr209 in substrate positioning.[7a] In spite of the interaction with malonate in Ec-ybfF, substitutions of Arg120 and His121 with alanine produced nearly identical (three- to fourfold) changes in catalytic efficiency (Figure 3C). More conservative substitutions (Arg120 and Tyr209 with lysine and phenylalanine, respectively) resulted in near wild-type activity, thus suggesting that size and charge are more important than hydrogen bonding for these amino acids in Vc-ybfF (Figure 3E). Of the other two residues essential for catalytic activity, Trp236 was predicted to play an important role in ybfF based on its position at the interface of the helical cap and hydrolase domain; this was confirmed by the complete loss of activity in W236A Vc-ybfF.[7a] Substitution of the analogous tryptophan residue in Ec-ybfF also led to the greatest increase in enantioselectivity toward *R*-IPG esters.[9] The His20 residue, although completely conserved, was not however previously predicted to play an important role in catalysis by ybfF.

Amongst the 13 alanine mutants, the Y116A variant showed the greatest difference in activity between substrates **3** and **6**. With the alkyl substrate (**3**), catalytic activity was significantly decreased, whereas with **6**, it showed only a twofold decrease in catalytic activity (Figure 3C). The difference in the activity of Y116A Vc-ybfF between **6** and **3** was also ten times greater than the difference measured for wild-type ybfF (Figure 2). Intriguingly, the residue in eukaryotic organisms homologous to Tyr116 is either serine or threonine (Figure 3B), thus suggesting that the residue at position 116 might act as a substrate-selectivity filter for different ybfF homologues.

To confirm this hypothesis, Tyr116 was substituted with various residues (phenylalanine, serine, threonine, valine, and aspartate), and each of the resulting Vc-ybfF variants was characterized for activity with **3** and **6** (Figure 3D). Even with multiple attempts at growth and purification by a variety of expression strategies, two of the Tyr116 variants (Y116V and Y116D) produced only limited soluble, well-folded protein, thus kinetic characterization was impossible. Phenylalanine substitution, however, showed the opposite pattern to Y116A: the activity toward **3** was near wild-type levels, whereas the activity toward **6** was significantly lower. The enzymatic efficiency values for the serine and threonine variants were somewhere in between, with slightly decreased activity toward **3,** and wild-type activity with **6**. Thus, single substitutions of Tyr116 with serine or threonine likely provide eukaryotic-type homologues, with greater selectivity toward polar substrates that are not evident in bacterial homologues. The selectivity in mammalian homologues with serine or threonine at position 116 is not, however, as large as is possible with an alanine substitution, which interestingly was not observed in any of the homologues (Figure 3B). Full elucidation of the role of Tyr116 in controlling the substrate specificity of ybfF will require more detailed investigation, as eukaryotic homologues of ybfF also have differences in the adjacent amino acids (Arg120 and His121), and concerted substitution of these residues might be required to fully shift the substrate specificity of the bacterial and human homologues.

Structural determinants of ybfF stability

Definitive explanation of the role of each binding pocket cleft of Vc-ybfF requires consideration of its contribution to the catalytic activity and structural stability; large structural changes upon substitution would shift the catalytic activity of the enzyme. Thermal stability measurements assess the role of each binding-pocket residue in maintaining the folded structure of ybfF. The thermal denaturation of each Vc-ybfF variant was measured, and T_m values for the mid-point of the two state transition are shown in Figure 3D. A large shift in T_m between wild-type and a particular Vc-ybfF variant suggests an important role for the substituted residue in maintaining the folded structure of Vc-ybfF.

Significant changes in folded stability ($>5\,^{\circ}C$) were observed for the H20A, K90A, H121A, and W236A variants (Figure 3D). Apart from H121A, each showed near-zero catalytic activity, thus confirming the connection between folded stability and catalytic efficiency (Figure 3C). The change in stability for W236A ($9\,^{\circ}C$) was especially dramatic and indicates an essential role for Trp236 in maintaining stability (Figure 3D). Based on the connected nature of the stability and catalytic activity of this group of substitutions, a "stability threshold" can be identified: residue changes that result in a T_m less than $42\,^{\circ}C$ drastically affect the stability and hence catalytic activity of ybfF.

Substitution of the catalytic serine (S86A) is the only Vc-ybfF variant that showed increased protein stability. This is likely due to the observed distorted geometry adopted by Ser86.[2] As ybfF is a member of the hydrolase-fold family, it contains a conserved GxSxG nucleophilic elbow, which adopts a tight turn and thus strained bond angles.[2] Removal of this strain upon alanine substitution led to elevated thermal stability. This phenomena is, however, unique to Ser86, as substitution of the neighboring amino acid (Met87) resulted in a protein that was slightly less stable than wild-type Vc-ybfF (Figure 3D).

For Tyr116, thermal stability differences do not explain the observed shifts in catalytic activity (Figure 3E and F). Alanine substitution decreased the thermal stability ($T_m = 43.5\,^{\circ}C$), but similar changes in thermal stability were seen for substitution with serine ($T_m = 43\,^{\circ}C$) and threonine ($T_m = 44\,^{\circ}C$). The phenylalanine variant exhibited near-wild-type stability ($T_m = 47\,^{\circ}C$); none of the Tyr116 variants had stabilities below the $42\,^{\circ}C$ threshold (that defines the instability mark for Vc-ybfF). The thermal stability curves for the Y116D and Y116V variants also failed to show a folding transition, thereby confirming improper folding of these Tyr116 variants. With the lack of major thermal stability changes, the shifts in catalytic activity for the Tyr116 variants can be directly related to differences in substrate interaction. This further supports the argument that Tyr116 is essential to the substrate selectivity profile of Vc-ybfF.

Discussion

Serine hydrolases constitute a diverse enzyme family, and catalyze a range of biological functions by using the same basic structural fold.[1a,2,14] Large variations in the structure of the binding pocket and the addition of secondary domains (lipase lids and helical caps) within serine hydrolases allow family members to recognize diverse intracellular substrates.[1a,2] This diversity of substrate selectivity and general adaptability of the α/β hydrolase fold has made serine hydrolases common targets for directed evolution and now for inhibitor design.[3b,c,14b] Within the diverse serine hydrolase superfamily, the ybfF protein family shows an unusual architecture of its substrate-binding pocket and a unique helical cap (Figure 1).[7a] With only limited sequence similarity to other hydrolases, limited enzymatic characterization, and the known ability to catalyze commercially applicable reactions, the ybfF protein family provides an interesting opportunity to understand the plasticity of substrate binding and reactivity in serine hydrolases.

The overall substrate specificity landscape of ybfF places it in the carboxylesterase family, with strict substrate selectivity for smaller, only slightly branched substrates (Figure 2). The preference for polarity close to the ester bond reinforces the proposed biological assignment as malonyl-CoA or succinyl-CoA esterase.[7a] Yet like many bacterial esterases where limited numbers of hydrolases are present in the proteome,[15] ybfF has the ability to accept a range of substrates, especially a variety of cycloalkyl substrates, as ambiguous substrates in cases of evolutionary necessity.[16] To refine the structural factors determining the catalytic activity and substrate specificity of ybfF, we wanted to assign a role to each cleft of the unusual bifurcated binding pocket of ybfF in its activity. The ability of Vc-ybfF to accommodate chains larger than four atoms and to catalyze reactions with bulky aromatic groups indicates that the larger binding cleft is the major substrate-binding cleft in ybfF, as the smaller cleft does not have sufficient space to accommodate these larger substrates (Scheme 1 and Figure 4). The mutagenesis results support this conclusion, as the majority of residues essential (or highly important) to the catalytic activity of Vc-ybfF (with the exception of Tyr236) are located around the larger binding cleft (Figure 4A–C). Additionally, the proposed selectivity residue (Tyr116), which differs between bacterial and mammalian homologues of ybfF, is at the bottom of the larger binding cleft, thus suggesting full immersion of the substrates into the large binding cleft (Figures 3A and 4B, C). The identification of the larger binding cleft as the major substrate-recognition region also reinforces the modeled binding of IPG esters to ybfF.[7b,9]

The role of the smaller binding lobe in the ybfF binding pocket is less well-defined, as the presence of clearly positioned water molecules in the structure of Ec-ybfF led to the hypothesis that this cleft positions the water molecules for hydrolysis of the acyl–enzyme intermediate (Figure 1).[7a] Substitution of residues around the narrow cleft did not, however, lead to significant changes in catalytic activity, as would be predicted if the cleft played this key role in catalysis (Figure 3 and 4).

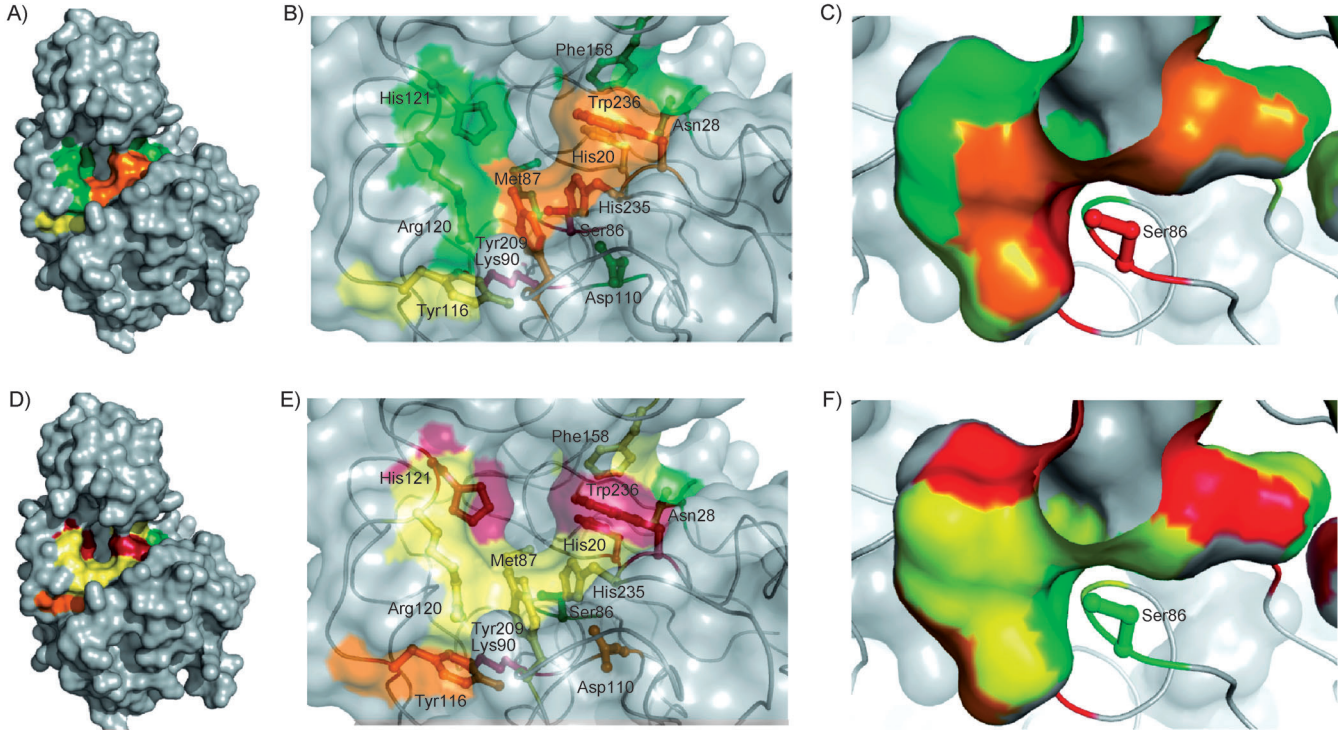


Figure 4. Substrate-binding residue relevance to Vc-ybfF catalytic activity and thermal stability. Surface representations of A) and B) exterior and C) interior of the binding pocket showing relative contributions of residues to catalytic activity with substrate **6**. Colors represent relative changes in catalytic activity (k_{cat}/K_M) upon substitution to alanine: red (>1000-fold); orange (>400-fold); yellow (>20-fold); green (<20-fold). D–F) Relative contributions or residues to thermal stability (T_m): red (>5.5 °C), orange (>4.5 °C), yellow (>2.5 °C), and green (<2.5 °C).

Instead, residues surrounding the narrow cleft seemed more directly associated with thermal stability and proper folding of the Vc-ybfF structure, as this region is enriched in residues important for the stability of Vc-ybfF, including His20, Phe158, His235, and Trp236 (Figure 4D–F). Our results cannot, however, discount the modeled role for the narrow binding cleft in fitting short alkyl chains bound to IPG esters, as the difunctionalized fluorescein is likely too large to fit within the larger binding pocket of ybfF, as has been proposed for the IPG groups.[7b,9] Our results do, however, suggest that this role for the narrow binding pocket is likely minor in standard substrate-binding and catalytic activity.

In addition to surrounding the narrow binding cleft, key residues for the stability of ybfF (His121 and Trp236) are at the entrance to the binding pocket and at the interface of the helical cap (Figure 4D–F). Two potential explanations for the positioning of these key stability residues at this position are: 1) rigid structural positioning is required to maintain an open entrance to the tunnel or 2) these residues are required for movement of the entrance tunnel to accommodate larger substrates. Movement of lid domains in lipases is required for their full catalytic activity and to provide additional regulation.[17] Although ybfF is not a lipase, with its specificity for shorter ester chains (Figure 2), the predicted location of the coenzyme A molecule along the interface could provide a similar mechanism for substrate-induced activation in ybfF.[7a] Movement between the helical cap and hydrolase domain in ybfF could also explain the necessity for a unique helical cap and the placement of the buried binding pocket at the interface of these domains. Similar flexibility was observed in the loop connecting the lid and hydrolase domains in the structure of the ybfF homologue, Rv0045c.[8b] Confirmation of this proposed movement will require further substitution and kinetics analysis with longer substrates.

The broad substrate selectivity of ybfF and the high enantioselectivity of ybfF with IPG esters make ybfF a potential new commercial biocatalyst.[7b,9] An initial study on activity enhancement of Ec-ybfF toward IPG esters focused on four amino acids (Leu22, His121, Ile186, and Trp236; Vc-ybfF numbering) that were in direct contact with the modeled IPG esters.[9] Only substitutions at Ile186 and Trp236 made significant improvements to the activity of Ec-ybfF, with nonpolar substitutions at Tr236 providing the largest increase in enantioselectivity toward *R*-IPG esters.[9] These results confirm the potential for improving the biocatalyst properties of ybfF, but further development of ybfF as a biocatalyst will require broader directed-evolution studies. For these studies, the assignment of key residues involved in activity, stability, and substrate selectivity can be used to decrease the size of the random library and to increase the probability of finding a successful variant.[3c,18] Our results have determined both potential selectivity residues (Tyr116, Arg120, Tyr209) and residues to avoid in future substitution studies (His20, His121, His235, Trp236) because of their essential roles in stability or activity. The map of each residues role in the catalytic activity and stability of ybfF (Figure 4) also highlights hot spots for residues not varied in this study and second sphere residues positioned on the periphery of the binding pocket. The divergent roles for the large and narrow binding cleft will also allow future substitution analysis to decouple the effect of a substitution on the catalytic activity and on thermal stability of ybfF. Future mutagenesis studies will now focus on concerted saturation mutagenesis of selectivity residues and seek to determine if concerted substitutions at these residues can redirect the substrate selectivity of Vc-ybfF.

In addition to their potential roles as novel biocatalysts, members of the ybfF enzyme family have also been implicated as potential therapeutic targets. Abhydrolase 11 (ABHD11), the human homologue of ybfF, was recently identified as a biomarker for aggressive lung adenocarcinoma and was previously linked to the development of Williams syndrome, a complex neurodevelopmental disorder caused by a large deletion in chromosome 21.[14a,19] Although no explicit studies of the activity or function of ABHD11 have been conducted to date, two current inhibitors of ABHD11 have been identified in activity-based protein-profiling screens with triazole and carbamate core structures.[14b,20] We note that these inhibitor structures parallel the substrate specificity profile of Vc-ybfF, as each inhibitor contains a slightly bulky cyclic side chain connected to a polar alkyl chain of less than four carbons.[14b] This dual functionality of the ABHD11 inhibitors not only matches the substrate specificity profile of Vc-ybfF, but also the structure of the substrate-binding pocket of ybfF (Figure 4). Given the structural similarity between ABHD11 inhibitors and the substrate-selectivity profile of ybfF (Figure 2), the current ABHD11 inhibitors could be used as primary leads in designing ybfF-specific inhibitors or could be used as in cellulo modulators of ybfF activity to better assign the biological function of ybfF.

Conclusions

The ybfF protein family exhibits unusual structural, functional, and kinetic characteristics, and presents potential avenues for biocatalyst development. The structural architecture of the bifurcated binding pocket of ybfF allows it to recognize a range of short, bulky, and slightly polar ester substrates. This substrate selectivity is mainly provided by the larger substrate binding cleft and by key substrate-selectivity residues (Tyr116, Arg120, and Tyr209). The narrow binding pocket, which lacks a defined role in the catalytic activity of Vc-ybfF but is essential to the stability, presents an intriguing location for future biocatalyst design. The distinct roles of the large and narrow binding clefts in catalytic activity and thermal stability can now be exploited to decouple these two important functions when designing future biocatalysts. The combined contributions of each cleft of this unusual binding pocket to the catalytic activity and substrate selectivity of ybfF also provides a new mechanism for serine hydrolases to differentiate between their diverse biological substrates.

Experimental Section

Synthesis of fluorogenic ester substrates:

General chemical synthesis: Compounds **1**–**5**, **9**, **11**, **17**, and **18** were synthesized as described previously.[10a,b,c] All reagents were of the highest grade available and obtained from Sigma–Aldrich or Fisher Scientific. Preactivated, powdered 4 Å molecular sieves (Sigma–Aldrich) were used as received. Anhydrous solvents were drawn from Sigma–Aldrich Sure-Seal bottles. Thin-layer chromatography was performed with aluminum-backed plates coated with silica gel containing F_{254} phosphor (Sigma–Aldrich) and visualized by UV illumination, charring, or staining with I_2, ceric ammonium molybdate, or phosphomolybdic acid.

Flash chromatography was performed on an Isolera 4 system with SNAP columns (Biotage, Uppsala, Sweden). Because of closely eluting byproducts in the alkylation reaction, several chromatography runs were sometimes required to obtain pure material, thus resulting in low isolated yields. The term "concentrated in vacuo" (below) refers to removal of solvents and other volatile materials by using a rotary evaporator or speedvac system at variable pressure (controlled diaphragm pump; $\geq$0.5 mm Hg) while maintaining the water-bath or chamber temperature below 40 °C. NMR spectra were obtained with a 400 MHz Avance II$^+$ spectrometer (Bruker) at the Janelia Farm Research Campus (Ashburn, VA).

Fluorescein bis((2-methoxy)acetoxymethyl ether) (6). The following protocol is representative. Methoxyacetic acid (730 mg, 8.08 mmol, 1.0 equiv) was added to a solution of tetrabutylammonium hydrogensulfate (274 mg, 0.808 mmol, 0.1 equiv) and potassium carbonate (4.47 g, 32.3 mmol, 4.0 equiv) in H_2O (10 mL) and CH_2Cl_2 (10 mL). A solution of chloromethyl chlorosulfate (1.23 mL, 12.12 mmol, 1.5 equiv) in CH_2Cl_2 (5 mL) was then added dropwise, and the reaction was stirred vigorously (~600 rpm) at room temperature under N_2 for 4 h. The mixture was then diluted in H_2O (30 mL) and CH_2Cl_2 (50 mL), and the layers were separated. The aqueous layer was extracted with CH_2Cl_2 (2× 50 mL), the organic layers were combined and dried over $MgSO_4$, filtered, and concentrated in vacuo (***note:*** volatile). The residue was filtered through a plug of silica gel, washed with CH_2Cl_2 (100 mL), and carefully concentrated in vacuo (***note:*** volatile), to afford chloromethyl 2-methoxyacetate as a colorless liquid (0.22 g, 20%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.77 (s, 2H), 4.12 (s, 2H), 3.48 ppm (s, 3H).

Chloromethyl 2-methoxyacetate (220 mg, 1.59 mmol, 4.0 equiv) was dissolved in HPLC-grade acetone (4.0 mL). NaI (241 mg, 1.59 mmol, 4.0 equiv) was added, and the mixture was stirred for 24 h at room temperature. The reaction mixture was adsorbed onto celite and purified by column chromatography (silica gel, CH_2Cl_2 (4→34% *v/v* in hexanes). Removal of solvent gave a colorless oil (***note:*** volatile), which was taken up in anhydrous CH_3CN (7.0 mL) under N_2. Fluorescein (0.13 g, 0.40 mmol, 1.0 equiv), powdered 4 Å molecular sieves (100 mg), and anhydrous Ag_2O (0.23 g, 1.0 mmol, 2.5 equiv) were added, then the mixture was stirred for 72 h at room temperature, diluted in CH_2Cl_2 (100 mL), and filtered through a pad of celite. The resulting solution was concentrated in vacuo to give a yellow-brown oil, and purified by column chromatography (silica gel, EtOAc (10→80% *v/v*) in hexanes containing CH_2Cl_2 (30% *v/v*) as cosolvent) to afford **6** as a white solid (56.5 mg, 27%). ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.03 (d, *J* = 7.2 Hz, 1H), 7.68 (ddd, *J* = 7.4, 7.4, 1.1 Hz, 1H), 7.63 (ddd, *J* = 7.4, 7.1, 0.6 Hz, 1H), 7.15 (d, *J* = 7.6 Hz, 1H), 6.96 (m, 2H), 6.74 (m, 4H), 5.87 (s, 4H), 4.10 (s, 4H), 3.46 ppm (s, 6H).

Fluorescein bis(((3-methoxy)propionyloxy)methyl ether) (7). Data for chloromethyl 3-methoxypropanoate: 0.40 g, 32%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.72 (s, 2H), 3.68 (t, *J* = 6.2 Hz, 2H), 3.36 (s, 3H), 2.66 ppm (t, *J* = 6.2 Hz, 2H). Data for **7**: 199 mg, 13%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.03 (d, *J* = 7.3 Hz, 1H), 7.68 (ddd, *J* = 7.5, 7.4, 1.4 Hz, 1H), 7.63 (dd, *J* = 7.2, 7.2 Hz, 1H), 7.15 (d, *J* = 7.4 Hz, 1H), 6.99–6.93 (m, 2H), 6.76–6.71 (m, 4H), 5.81 (s, 4H), 3.67 (t, *J* = 6.2 Hz, 4H), 3.32 (s, 6H), 2.65 ppm (t, *J* = 6.2 Hz, 4H).

Fluorescein bis((2-(2-methoxyethoxy)acetoxy)methyl ether) (8). Data for chloromethyl 2-(2-methoxyethoxy)acetate: 0.57 g, 39%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.76 (s, 2H), 4.24 (s, 2H), 3.78–3.72 (m, 2H), 3.63–3.56 (m, 1H), 3.39 ppm (s, 3H). Data for **8**: 32 mg, 7%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.03 (d, *J* = 7.3 Hz, 1H), 7.68 (ddd, *J* = 7.5, 7.3, 1.2 Hz, 1H), 7.64 (ddd, *J* = 7.4, 7.4, 1.2 Hz, 1H), 7.15 (d, *J* = 7.5 Hz, 1H), 6.97–6.93 (m, 2H), 6.77–6.69 (m, 4H), 5.86 (s, 4H), 4.22 (s, 4H), 3.76–3.69 (m, 4H), 3.59–3.54 ppm (m, 4H), 3.36 (s, 6H).

Fluorescein bis((acryloyloxy)methyl ether) (10). Data for chloromethyl acrylate: 0.30 g, 31%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 6.54 (dd, *J* = 17.3, 1.2 Hz, 1H), 6.16 (dd, *J* = 17.3, 10.5 Hz, 1H), 5.98 (dd, *J* = 10.5, 1.2 Hz, 1H), 5.80 ppm (s, 2H). Data for **10**: 110 mg, 35%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.03 (ddd, *J* = 7.3, 1.4, 0.8 Hz, 1H), 7.67 (ddd, *J* = 7.4, 7.4, 1.4 Hz, 1H), 7.63 (ddd, *J* = 7.4, 7.4, 1.2 Hz, 1H), 7.18–7.13 (m, 1H), 7.00–6.96 (m, 2H), 6.78–6.71 (m, 4H), 6.51 (dd, *J* = 17.3, 1.3 Hz, 2H), 6.16 (dd, *J* = 17.3, 10.5 Hz, 2H), 5.94 (dd, *J* = 10.5, 1.3 Hz, 2H), 5.88 ppm (s, 4H).

Fluorescein bis(((cyclopentanecarbonyl)oxy)methyl ether) (12). Data for chloromethyl cyclopentanecarboxylate: 0.99 g, 75%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.71 (s, 2H), 2.80 (tt, *J* = 8.6, 7.3 Hz, 1H), 1.99–1.53 ppm (m, 8H). Data for **12**: 354 mg, 40%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.06–8.00 (m, 1H), 7.68 (ddd, *J* = 7.4, 7.4, 1.4 Hz, 1H), 7.63 (ddd, *J* = 7.4, 7.4, 1.2 Hz, 1H), 7.17 (dd, *J* = 7.6, 1.0 Hz, 1H), 7.01–6.92 (m, 2H), 6.79–6.68 (m, 4H), 5.78 (s, 5H), 2.80 (tt, *J* = 8.6, 7.4 Hz, 2H), 2.03–1.44 ppm (m, 16H).

Fluorescein bis(((cyclohexanecarbonyl)oxy)methyl ether) (13). Data for chloromethyl cyclohexanecarboxylate: 1.29 g, 90%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.71 (s, 2H), 2.37 (tt, *J* = 11.2, 3.7 Hz, 1H), 2.02–1.84 (m, 4H), 1.83–1.15 ppm (m, 6H). Data for **13**: 211 mg, 19%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.03 (ddd, *J* = 7.3, 1.2, 1.0 Hz, 1H), 7.68 (ddd, *J* = 7.4, 7.4, 1.4 Hz, 1H), 7.63 (ddd, *J* = 7.4, 7.4, 1.2 Hz, 1H), 7.17 (ddd, *J* = 7.6, 1.3, 1.0 Hz, 1H), 6.98–6.92 (m, 2H), 6.76–6.70 (m, 4H), 5.78 (s, 4H), 2.37 (tt, *J* = 11.3, 3.7 Hz, 2H), 1.97–1.86 (m, 4H), 1.83–1.12 ppm (m, 16H).

Fluorescein bis((furan-2-carboxy)methyl ether) (14). Data for chloromethyl furan-2-carboxylate: 1.02 g, 79%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.65 (dd, *J* = 1.7, 0.8 Hz, 1H), 7.31 (dd, *J* = 3.5, 0.9 Hz, 1H), 6.56 (dd, *J* = 3.5, 1.7 Hz, 1H), 5.92 ppm (s, 2H). Data for **14**: 309 mg, 33%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.02 (ddd, *J* = 7.4, 1.4, 0.9 Hz, 1H), 7.67 (ddd, *J* = 7.4, 7.4, 1.5 Hz, 1H), 7.64–7.59 (m, 3H), 7.28 (dd, *J* = 3.5, 0.9 Hz, 2H), 7.15 (ddd, *J* = 7.4, 1.4, 0.9 Hz, 1H), 7.04 (d, *J* = 2.4 Hz, 2H), 6.80 (dd, *J* = 8.8, 2.5 Hz, 2H), 6.75 (d, *J* = 8.8 Hz, 2H), 6.53 (dd, *J* = 3.6, 1.7 Hz, 2H), 6.08–5.91 ppm (m, 4H).

Fluorescein bis((benzoyloxy)methyl ether) (15). Data for chloromethyl benzoate: 1.06 g, 77%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.13–8.04 (m, 2H), 7.68–7.55 (m, 1H), 7.52–7.43 (m, 2H), 5.96 ppm (s, 2H). Data for **15**: 68 mg, 7%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.13–8.05 (m, 4H), 8.02 (ddd, *J* = 7.3, 1.5, 0.8 Hz, 1H), 7.66 (ddd, *J* = 7.4, 7.4, 1.5 Hz, 2H), 7.63–7.54 (m, 3H), 7.52–7.38 (m, 4H), 7.20–7.11 (m, 1H), 7.05 (d, *J* = 2.4 Hz, 2H), 6.81 (dd, *J* = 8.8, 2.4 Hz, 2H), 6.75 (d, *J* = 8.8 Hz, 2H), 6.28–5.79 ppm (m, 4H).

Fluorescein bis((2-phenylacetoxy)methyl ether) (16). Data for chloromethyl 2-phenylacetate: 0.803 g, 54%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.56–7.06 (m, 5H), 5.70 (s, 2H), 3.70 ppm (s, 2H). Data for **16**: 25.4 mg, 4%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.07–8.02 (m, 1H), 7.69 (ddd, J = 7.5, 7.4, 1.4 Hz, 2H), 7.64 (ddd, J = 7.4, 7.4, 1.2 Hz, 1H), 7.36–7.20 (m, 10H), 7.18–7.12 (m, 1H), 6.88 (d, J = 2.3 Hz, 2H), 6.70 (d, J = 8.7 Hz, 4H), 6.67 (dd, J = 8.8, 2.3 Hz, 2H), 5.79 (s, 4H), 3.69 ppm (s, 4H).

Fluorescein bis(((3,3,3-trifluoropropanoyl)oxy)methyl ether) (19). Data for chloromethyl 3,3,3-trifluoropropanoate: 0.37 g, 26%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.75 (s, 2H), 3.28 ppm (q, J = 9.8 Hz, 2H). Data for **19**: 35.2 mg, 11%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 7.95 (ddd, J = 7.3, 1.3, 0.8 Hz, 1H), 7.60 (ddd, J = 7.4, 7.4, 1.4 Hz, 1H), 7.56 (ddd, J = 7.4, 7.4, 1.2 Hz, 1H), 7.08 (ddd, J = 7.6, 1.4, 0.8 Hz, 1H), 6.87 (dd, J = 2.2, 0.6 Hz, 2H), 6.68 (dd, J = 8.8, 0.6 Hz, 2H), 6.65 (dd, J = 8.8, 2.2 Hz, 2H), 5.81–5.74 (m, 4H), 3.18 ppm (q, J = 9.9 Hz, 4H).

Fluorescein bis((perfluorobenzoyloxy)methyl ether) (20). Data for chloromethyl perfluorobenzoate: 1.66 g, 79%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.93 ppm (s, 2H). Data for **20**: 62.2 mg, 5%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.03 (ddd, J = 7.4, 1.5, 0.8 Hz, 1H), 7.68 (ddd, J = 7.5, 7.4, 1.5 Hz, 2H), 7.64 (ddd, J = 7.4, 7.4, 1.2 Hz, 1H), 7.16 (dt, J = 7.5, 1.2, 0.8 Hz, 1H), 7.05–6.98 (m, 2H), 6.82–6.74 (m, 4H), 6.08–5.98 ppm (m, 4H).

Fluorescein bis((2-(pentafluorophenyl)acetoxy)methyl ether) (21). Data for chloromethyl 2-(perfluorophenyl)acetate: 1.58 g, 71%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 5.73 (s, 2H), 3.84–3.79 ppm (m, 2H). Data for **21**: 29.8 mg, 3%; ^{1}H NMR (400 MHz, $CDCl_3$): δ = 8.05 (ddd, J = 7.4, 1.1, 0.8 Hz, 1H), 7.70 (ddd, J = 7.4, 7.4, 1.3 Hz, 1H), 7.65 (ddd, J = 7.4, 7.4, 1.1 Hz, 1H), 7.18 (ddd, J = 7.4, 1.3, 0.9 Hz, 1H), 6.89 (d, J = 2.3 Hz, 2H), 6.76 (d, J = 8.8 Hz, 2H), 6.71 (dd, J = 8.8, 2.4 Hz, 2H), 5.88–5.78 (m, 4H), 3.84–3.78 ppm (m, 4H).

Purification of Vc-ybfF from *V. cholera*e: A bacterial expression plasmid (pANT-GST) containing the VC2097 gene from *V. cholerae* O1 biovar El Tor strain N16961 (GenBank ID: NP_231729.1) was obtained from the Harvard Plasmid Repository (Clone ID: VcCD00035533). The VC2097 gene was subcloned into a pET-22b (+) plasmid (Merck Millipore) between the NdeI and XhoI restriction sites and with an in-frame C-terminal His_6 tag, by using primers VC2097 Forward and VC2097 Reverse (Table S4). Correct sequence insertion was confirmed by DNA sequencing (Genewiz Inc, South Plainfield, NJ).

This resulting plasmid (pET-22b-Vc2097) was transformed into *E. coli* BL21 (DE3) RIPL cells (Agilent). A saturated overnight culture in lysogeny broth (LB) containing ampicillin (100 μg mL^{-1}) and chloramphenicol (30 μg mL^{-1}) was used to inoculate LB (250 mL) containing ampicillin (100 μg mL^{-1}) and chloramphenicol (30 μg mL^{-1}), and the bacterial culture was grown with shaking (225 rpm) at 37 °C. When OD_{600} reached 0.6–0.8, the temperature of the culture was decreased to 16 °C, and isopropyl β-D-1-thiogalactopyranoside (IPTG) was added (final concentration 0.5 mM) for protein induction. After 16–20 h at 16 °C, cells were collected by centrifugation (6000 *g*, 10 min, 4 °C), and the pellet was resuspended in PBS (2 mL) and stored at −20 °C.

To disrupt the bacterial cell wall, lysozyme (50 mg) was added to the thawed cell pellet, and the mixture was set on ice for 30 min. Bug Buster solution (250 μL, 10×; EMD Millipore) was added, and cell lysis proceeded on an orbital shaker for 1 h at 4 °C. To remove insoluble cell material, lysed cells were centrifuged (18500 *g*, 10 min, 4 °C). Ni-NTA agarose (750 μL; Qiagen, Valencia, CA) was added to the soluble fraction, and the mixture was incubated (4 °C, 15 min). The resin was washed with ice-cold PBS (3×1 mL) containing increasing concentrations of imidazole (10, 25, and 50 mM), and collected by centrifugation (2000 *g*, 2 min, 25 °C). Vc-ybfF was eluted in PBS (1000 μL) containing imidazole (250 mM), and dialyzed against PBS overnight at 4 °C with constant stirring (10 K MWCO; Pierce/Thermo Scientific).

The purity of Vc-ybfF was confirmed by SDS-PAGE on a 4–20% gradient gel (purity >95%, Figure S1). The concentration of Vc-ybfF was determined by measuring the absorbance at 280 nm and by calculating the extinction coefficient (ε_{280} = 23950 M^{-1} s^{-1} with all free cysteines) by using ProtParam (http://web.expasy.org/protparam/)

Site-directed mutagenesis and purification: Variants of Vc-ybfF were created by site-directed mutagenesis by using Quikchange II (Agilent, Santa Clara, CA) according to the manufacturer's suggested procedure (mutagenesis primers in Table S4). Correct mutations in Vc-ybfF DNA were confirmed by DNA sequencing (Genewiz). Plasmids with Vc-ybfF variants were transformed into *E. coli* BL21 (DE3) RIPL, and variants of Vc-ybfF were purified by the procedure for wild-type Vc-ybfF (Figure S1). For variants of Vc-ybfF with tyrosine and tryptophan substitutions, specific extinction coefficients were obtained from ProtParam (Tyr ε_{280} = 22460 M^{-1} s$^{-1;}$ Trp ε_{280} = 18450 M^{-1} s^{-1}).

Kinetic measurements with fluorogenic ester substrates: The enzymatic activities of Vc-ybfF and variants were measured with the fluorogenic ester substrates (Scheme 1) in a 96-well microplate assay.[10c] Fluorogenic substrates were prepared as stock solutions in DMSO (10 mM) and were diluted in PBS containing acetylated BSA (PBS–BSA; 0.1 mg mL^{-1}) depending on the K_m of Vc-ybfF for the substrate (10 μM (**1**–**5**, **9**–**18**), 100 μM (**6**–**8**), 1000 μM (**19**–**21**)). Eight serial dilutions in PBS/BSA (1:2; 180 μL each) were made for each substrate. Fluorogenic substrate dilutions (95 μL) were then transferred to wells of a black 96-well microplate (Corning, Lowell, MA).

Wild-type Vc-ybfF (5 μL final concentration 7.5 μg mL^{-1}) or variant (5 μL, final concentration 15 μg mL^{-1}) was added to the diluted fluorogenic substrates in the 96-well microplate (100 μL final volume), and the fluorescence (λ_{ex} = 485 nm, λ_{em} = 528 nm) was measured over 4 min at 25 °C on a Synergy 2 fluorescence plate reader (Biotek Instruments; Winooski, VT). Fluorescence values were measured simultaneously and converted to molar concentration by using a fluorescein standard curve (30 nM–0.23 nM (**1**–**5**, **9**–**18**), 300 nM–2.3 nM (**6**–**8, 19**–**21**)). The initial rates of the reactions were measured in triplicate and plotted against fluorogenic substrate concentration. The saturation enzyme kinetic traces were fitted to a standard Michaelis–Menten equation with Origin 6.1 (OriginLab Corp., Northhampton, MA), and values for k_{cat}, K_M, and k_{cat}/K_M were calculated. The background fluorescence was determined by measuring the kinetics of fluorogenic substrate hydrolysis by the S86A variant of Vc-ybfF (Table 1).

Kinetics measurements for *p*-nitrophenyl substrates with Vc-ybfF: Stock solutions (*p*-nitrophenyl acetate (2 M), *p*-nitrophenyl butyrate (2 M), *p*-nitrophenyl octanoate (200 mM), *p*-nitrophenyl laurate (200 mM)) in acetonitrile were diluted in PBS–BSA (0.1 mg mL^{-1}) to starting concentrations of 20 mM (acetate) or 2 mM (butyrate, octanoate, and laurate). Eight serial dilutions (1:1, total volume 220 μL; 20 mM–156 μM (acetate) and 2 mM–15.6 μM (butyrate, octanoate, and laurate)) were made in PBS–BSA containing acetonitrile (1%). Substrate dilutions (95 μL) were transferred to a clear 96-well microplate, and Vc-ybfF (5 μL; final concentration

15 μg mL^{-1}) was added to start the reaction. The absorbance at 410 nm was measured in a Synergy 2 plate reader (Biotek) over 4 min at 25 °C. The absorbance was converted to molar concentrations by using the extinction coefficient of *p*-nitrophenol (ε_{410} 1.034 mM^{-1} cm^{-1}).[22] Initial rates of the reactions were measured in triplicate, and kinetics parameters were determined as for the fluorogenic assay.

Thermal stability measurement: Similarly to previously published methods,[10c,23] the thermal stabilities of Vc-ybfF and variants were determined by differential scanning fluorimetry (DSF). Wild-type Vc-ybfF or variant (0.30 mg mL^{-1}) was diluted in triplicate in PBS containing SYPRO Orange (1:250 dilution; Invitrogen/Life Technologies). The samples was heated (20–80 °C, 1.5 °C min^{-1}) in a C1000 thermocycler with a CFX96 optical module (Bio-Rad), and SYPRO Orange fluorescence was followed over time (λ_{ex} = 450–490 nm, λ_{em} = 610–650 nm). The melting temperature (T_m) was determined by plotting the first derivative of fluorescence versus temperature and finding the temperature at the midpoint of the transition. As in a previous analysis,[24] all graphs were normalized (minimum fluorescence = 0, maximum fluorescence = 1).

Acknowledgements

We thank Nathan Clarke, Erica Couch, Grace Douglas, Laurel Heckman, Lin Liu, Andrew Lutkewitte, Samantha Minnette, Nathan Tavenor, and Ben Trefilek for assistance with preliminary enzyme purification and enzymatic analysis, Courtney Walker and Ann Yoon for help with size-exclusion chromatography, Paula LeBlanc and Michael Slack for help with pH activity measurements, and Kelly McKenna for help with cloning Vc-ybfF. E.E.E. was supported by a Butler Summer Institute Scholarship. R.J.J. was supported by start-up funds and a Holcomb grant from Butler University, and a Senior Research Grant from the Indiana Academy of Sciences. C.T.A. and L.D.L. were supported by the Howard Hughes Medical Institute.

Keywords: biocatalysts · fluorogenic enzyme substrates · hydrolases · substrate specificity · *Vibrio cholerae*

[1] a) M. Holmquist, *Curr. Protein Pept. Sci.* **2000**, *1*, 209–235; b) R. Kourist, H. Jochens, S. Bartsch, R. Kuipers, S. K. Padhi, M. Gall, D. Bottcher, H. J. Joosten, U. T. Bornscheuer, *ChemBioChem* **2010**, *11*, 1635–1643.
[2] P. D. Carr, D. L. Ollis, *Protein Pept. Lett.* **2009**, *16*, 1137–1148.
[3] a) E. Busto, V. Gotor-Fernández, V. Gotor, *Chem. Soc. Rev.* **2010**, *39*, 4504–4523; b) U. T. Bornscheuer, *FEMS Microbiol. Rev.* **2002**, *26*, 73–81; c) H. Jochens, M. Hesseler, K. Stiba, S. K. Padhi, R. J. Kazlauskas, U. T. Bornscheuer, *ChemBioChem* **2011**, *12*, 1508–1517.
[4] U. T. Bornscheuer, R. J. Kazlauskas, *Angew. Chem.* **2004**, *116*, 6156–6165; *Angew. Chem. Int. Ed.* **2004**, *43*, 6032–6040.
[5] C. J. Yeoman, Y. Han, D. Dodd, C. M. Schroeder, R. I. Mackie, I. K. Cann, *Adv. Appl. Microbiol.* **2010**, *70*, 1–55.
[6] N. Lenfant, T. Hotelier, Y. Bourne, P. Marchot, A. Chatonnet, *Chem. Biol. Interact.* **2012**, *203*, 266–268.
[7] a) S.-Y. Park, S.-H. Lee, J. Lee, K. Nishi, Y.-S. Kim, C.-H. Jung, J.-S. Kim, *J. Mol. Biol.* **2008**, *376*, 1426–1437; b) L. F. Godinho, C. R. Reis, P. G. Tepper, G. J. Poelarends, W. J. Quax, *Appl. Environ. Microbiol.* **2011**, *77*, 6094–6099; c) D. J. Baumler, R. G. Peplinski, J. L. Reed, J. D. Glasner, N. T. Perna, *BMC Syst. Biol.* **2011**, *5*, 182; d) E. Kuznetsova, M. Proudfoot, S. A. Sanders, J. Reinking, A. Savchenko, C. H. Arrowsmith, A. M. Edwards, A. F. Yakunin, *FEMS Microbiol. Rev.* **2005**, *29*, 263–279.
[8] a) J. Guo, X. Zheng, L. Xu, Z. Liu, K. Xu, S. Li, T. Wen, S. Liu, H. Pang, *PLoS One* **2010**, *5*, e13143; b) X. Zheng, J. Guo, L. Xu, H. Li, D. Zhang, K. Zhang, F. Sun, T. Wen, S. Liu, H. Pang, *PLoS One* **2011**, *6*, e20506.
[9] L. F. Godinho, C. R. Reis, R. van Merkerk, G. J. Poelarends, W. J. Quax, *Adv. Synth. Catal.* **2012**, *354*, 3009–3015.
[10] a) L. D. Lavis, T.-Y. Chao, R. T. Raines, *Chem. Sci.* **2011**, *2*, 521–530; b) L. Tian, Y. Yang, L. M. Wysocki, A. C. Arnold, A. Hu, B. Ravichandran, S. M. Sternson, L. L. Looger, L. D. Lavis, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 4756–4761; c) M. K. Hedge, A. M. Gehring, C. T. Adkins, L. A. Weston, L. D. Lavis, R. J. Johnson, *Biochim. Biophys. Acta Gen. Subj.* **2012**, *1824*, 1024–1030.
[11] A. M. Weeks, S. M. Coyle, M. Jinek, J. A. Doudna, M. C. Y. Chang, *Biochemistry* **2010**, *49*, 9269–9279.
[12] a) S. Herter, G. S. Nguyen, M. L. Thompson, F. Steffen-Munsberg, F. Schauer, U. T. Bornscheuer, R. Kourist, *Appl. Microbiol. Biotechnol.* **2011**, *90*, 929–939; b) R. Kourist, U. T. Bornscheuer, *Appl. Microbiol. Biotechnol.* **2011**, *91*, 505–517.
[13] a) C. S. Craik, S. Roczniak, C. Largman, W. J. Rutter, *Science* **1987**, *237*, 909–913; b) L. Polgár, *Cell. Mol. Life Sci.* **2005**, *62*, 2161–2172.
[14] a) J. Z. Long, B. F. Cravatt, *Chem. Rev.* **2011**, *111*, 6022–6063; b) D. A. Bachovchin, B. F. Cravatt, *Nat. Rev. Drug Discovery* **2012**, *11*, 52–68.
[15] a) K. K. Kim, H. K. Song, D. H. Shin, K. Y. Hwang, S. Choe, O. J. Yoo, S. W. Suh, *Structure* **1997**, *5*, 1571–1584; b) A. Pesaresi, G. Devescovi, D. Lamba, V. Venturi, G. Degrassi, *Curr. Microbiol.* **2005**, *50*, 102–109.
[16] B. G. Miller, R. T. Raines, *Biochemistry* **2004**, *43*, 6387–6392.
[17] a) H. van Tilbeurgh, M. P. Egloff, C. Martinez, N. Rugani, R. Verger, C. Cambillau, *Nature* **1993**, *362*, 814–820; b) J. D. Schrag, Y. Li, M. Cygler, D. Lang, T. Burgdorf, H.-J. Hecht, R. Schmid, D. Schomburg, T. J. Rydel, J. D. Oliver, L. C. Strickland, C. M. Dunaway, S. B. Larson, J. Day, A. McPherson, *Structure* **1997**, *5*, 187–202; c) S. Rehm, P. Trodler, J. Pleiss, *Protein Sci.* **2010**, *19*, 2122–2130.
[18] M. T. Reetz, S. Prasad, J. D. Carballeira, Y. Gumulya, M. Bocola, *J. Am. Chem. Soc.* **2010**, *132*, 9144–9152.
[19] a) T. Wiedl, S. Arni, B. Roschitzki, J. Grossmann, S. Collaud, A. Soltermann, S. Hillinger, R. Aebersold, W. Weder, *J. Proteomics* **2011**, *74*, 1884–1894; b) G. Merla, C. Ucla, M. Guipponi, A. Reymond, *Hum. Genet.* **2002**, *110*, 429–438.
[20] a) D. A. Bachovchin, T. Ji, W. Li, G. M. Simon, J. L. Blankman, A. Adibekian, H. Hoover, S. Niessen, B. F. Cravatt, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20941–20946; b) A. Adibekian, B. R. Martin, C. Wang, K. L. Hsu, D. A. Bachovchin, S. Niessen, H. Hoover, B. F. Cravatt, *Nat. Chem. Biol.* **2011**, *7*, 469–478.
[21] a) C. N. Pace, F. Vajdos, L. Fee, G. Grimsley, T. Gray, *Protein Sci.* **1995**, *4*, 2411–2423; b) H. Edelhoch, *Biochemistry* **1967**, *6*, 1948–1954; c) S. C. Gill, P. H. von Hippel, *Anal. Biochem.* **1989**, *182*, 319–326.
[22] S. Kim, S. B. Lee, *Biosci. Biotechnol. Biochem.* **2004**, *68*, 2289–2298.
[23] F. H. Niesen, H. Berglund, M. Vedadi, *Nat. Protoc.* **2007**, *2*, 2212–2221.
[24] D. S. Froese, S. Healy, M. McDonald, G. Kochan, U. Oppermann, F. H. Niesen, R. A. Gravel, *Mol. Genet. Metab.* **2010**, *100*, 29–36.
[25] a) G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, *Genome Res.* **2004**, *14*, 1188–1190; b) R. Chenna, H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, J. D. Thompson, *Nucleic Acids Res.* **2003**, *31*, 3497–3500.

Received: February 19, 2013
Published online on May 13, 2013