JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Computational Design of Catalytic Dyads and Oxyanion Holes for Ester Hydrolysis

Florian Richter,[†,‡,×] Rebecca Blomberg,[⊥,×,+] Sagar D. Khare,[†] Gert Kiss,[∥,△] Alexandre P. Kuzin,[#] Adam J. T. Smith,[∥,◆] Jasmine Gallaher,[†] Zbigniew Pianowski,[⊥] Roger C. Helgeson,[∥] Alexej Grjasnow,[▽] Rong Xiao,[¶] Jayaraman Seetharaman,[#] Min Su,[#] Sergey Vorobiev,[#] Scott Lew,[#] Farhad Forouhar,[#] Gregory J. Kornhaber,[¶] John F. Hunt,[#] Gaetano T. Montelione,[¶,⊗] Liang Tong,[#] K. N. Houk,[∥] Donald Hilvert,*[,⊥] and David Baker*[,†,‡,§]

[†]Department of Biochemistry, [‡]Interdisciplinary Program in Biomolecular Structure and Design, and [§]Howard Hughes Medical Institute, University of Washington, Seattle, Washington 98195, United States

[⊥]Laboratory of Organic Chemistry, ETH Zurich, 8093 Zurich, Switzerland

[∥]Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095, United States

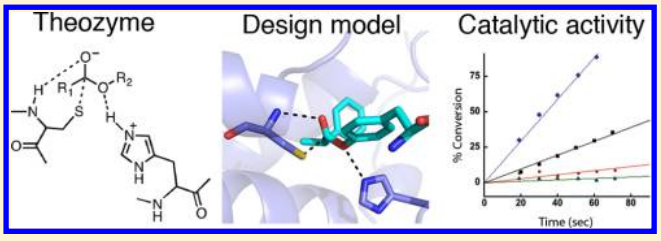
Northeast Structural Genomics Consortium, [#]Department of Biological Sciences, Columbia University, New York, New York 10027, and [¶]Center for Advanced Biotechnology and Medicine, Department of Molecular Biology and Biochemistry, Rutgers, The State University of New Jersey, Piscataway, New Jersey 08854, United States

[⊗]Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey, Piscataway, New Jersey 08854, United States

[▽]Department of Biochemistry, Ludwig-Maximilians-Universität München, 81377 München, Germany

Supporting Information

ABSTRACT: Nucleophilic catalysis is a general strategy for accelerating ester and amide hydrolysis. In natural active sites, nucleophilic elements such as catalytic dyads and triads are usually paired with oxyanion holes for substrate activation, but it is difficult to parse out the independent contributions of these elements or to understand how they emerged in the course of evolution. Here we explore the minimal requirements for esterase activity by computationally designing artificial catalysts using catalytic dyads and oxyanion holes. We found much higher success rates using designed oxyanion holes formed by backbone NH groups rather than by side chains or bridging water molecules and obtained four active designs in different scaffolds by combining this motif with a Cys-His dyad. Following active site optimization, the most active of the variants exhibited a catalytic efficiency (k_{cat}/K_M) of 400 M^{-1} s^{-1} for the cleavage of a *p*-nitrophenyl ester. Kinetic experiments indicate that the active site cysteines are rapidly acylated as programmed by design, but the subsequent slow hydrolysis of the acyl-enzyme intermediate limits overall catalytic efficiency. Moreover, the Cys-His dyads are not properly formed in crystal structures of the designed enzymes. These results highlight the challenges that computational design must overcome to achieve high levels of activity.



INTRODUCTION

Hydrolytic enzymes play an important role in numerous physiological and pathological processes such as inflammation,[1] angiogenesis,[2] cancer,[3,4] and diabetes.[5] These enzymes are also established tools for the industrial synthesis of fine chemicals. For example, lipases and hydrolases are often used for the production of optically pure molecules[6] and the modification of complex natural products such as antibiotics,[7] steroids,[8] and the anti-cancer drug Taxol.[9] Hydrolytic enzymes have been mimicked by cyclodextrins,[10] cyclophanes,[11] and other synthetic molecules.[12] These organic compounds provide hydrophobic binding sites for their substrates, and, like natural lipases and serine proteases, employ alcohols as nucleophiles to effect ester cleavage.[10] However, in these systems only the first step of ester hydrolysis is typically accelerated, namely the nucleophilic attack of a hydroxyl group on the ester substrates to give a covalent acyl intermediate. In contrast, true turnover catalysis has been achieved by catalytic antibodies.[13] Structural studies show that these antibodies usually operate by stabilizing the negatively charged intermediate of the hydrolysis reaction rather than by nucleophilic catalysis.[14] Phosphonate haptens generally fail to program for more elaborate arrays of catalytic functionality, although a nucleophilic histidine was elicited in at least one case.[15,16] A nucleophilic histidine was also used in a designed thioredoxin with hydrolytic activity.[17]

Received: April 18, 2012
Published: August 7, 2012

   dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197−16206

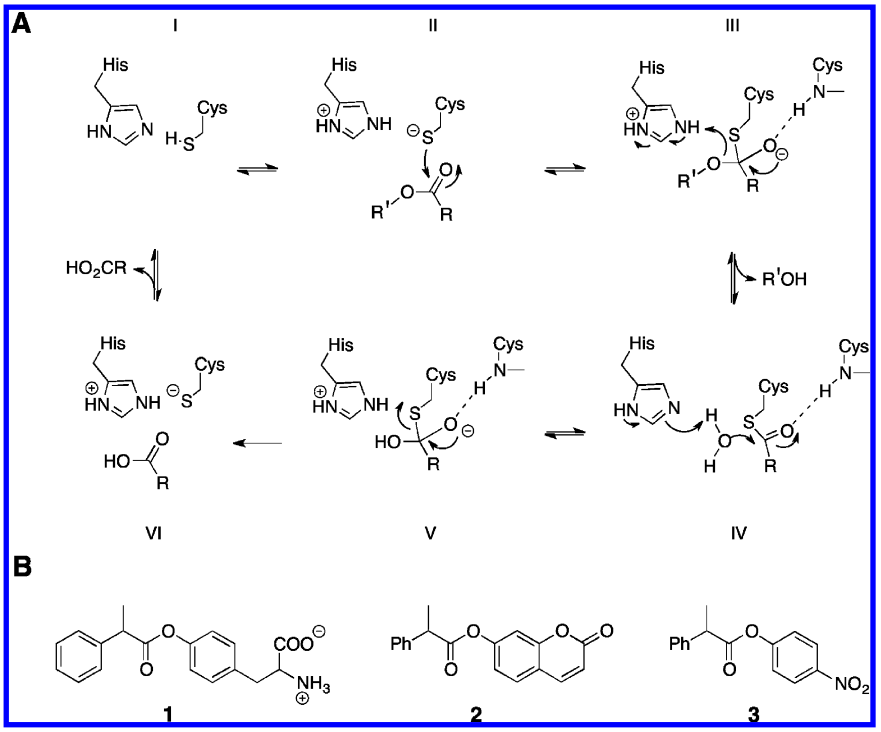


Figure 1. Programmed mechanism and model substrates of the *de novo* designed esterases. (A) A Cys-His dyad, in combination with an oxyanion binder, were used to hydrolyze activated esters via covalent catalysis. (B) The tyrosyl ester **1** served as the target substrate for computational design; the fluorogenic coumarin ester **2** and the chromogenic *p*-nitrophenyl ester **3** were used for screening purposes.

Natural hydrolytic enzymes often utilize a serine or cysteine as a nucleophile, which is deprotonated by a hydrogen-bonded histidine.[18] Precisely positioned hydrogen bond donors, so-called "oxyanion holes",[19] stabilize the oxyanion intermediate. The importance of these elements has been demonstrated by mutagenesis experiments in which the removal of any of these functional groups leads to drastic losses in activity. However, to our knowledge there have to date been no complementary efforts to build esterase catalysts in a bottom-up approach based on this catalytic machinery. With the approaches enumerated in the previous paragraph it is challenging to program an appropriately positioned nucleophile and oxyanion stabilization in the same catalyst, and hence difficult to systematically assess the extent to which functional esterase active sites can be constructed using combinations of these catalytic elements.

Here, we explore the fundamentals of esterase catalytic machinery using computational enzyme design.[20−22] The Rosetta3[23] *de novo* enzyme design protocol[24] is used to embed catalytic dyads and appropriately positioned oxyanion holes into catalytically inert protein scaffolds. We find that a minimalist catalytic schema consisting of a cysteine nucleophile, a nearby histidine together with a backbone NH group to stabilize the oxyanion intermediate is sufficient to generate primitive esterases. The relatively high success rate in generating esterases with this strategy suggests that a similar mechanism could have been employed by the nascent ancestors of modern enzymes.

RESULTS

Computational Design. We set out to explore the extent to which active esterases could be generated using the catalytic elements found in natural hydrolytic enzymes. Esterases often employ serines[25] and cysteines[26] as nucleophiles. We focused on cysteine as the reactive group as it is more nucleophilic than serine, has a lower pK_a, and is hence a better leaving group. In natural cysteine hydrolases, a histidine residue, usually oriented and activated by another hydrogen bond acceptor such as Asn/Gln or backbone oxygen, acts as a general acid/base to deprotonate the nucleophilic cysteine in the first step and the water in the second, and to protonate the leaving group of the tetrahedral intermediates.[27] Two or three hydrogen bond donors stabilize the oxyanion reaction intermediate. A backbone amide group often forms at least one of these oxyanion contacts. The general hydrolysis mechanism catalyzed by this active site arrangement is depicted in Figure 1A.

Three esters (**1**, **2**, and **3** in Figure 1) were chosen as model substrates. These three compounds have identical acyl groups, but differ in the degree of activation of their aromatic leaving groups. Tyrosyl ester **1** was used for the computational design process. Although it is the least activated substrate, cleavage of this compound would lead to the production of tyrosine, and thus allow for the development of a high-throughput growth selection assay based on complementation of an auxotrophic bacterial strain unable to biosynthesize this essential amino acid.[28] Hydrolysis of umbelliferyl ester **2**, which exhibits intermediate reactivity, can be monitored by a sensitive fluorescence assay. The *p*-nitrophenyl leaving group of ester **3**, the most activated substrate, is identical to that of *p*-nitrophenyl acetate, a model substrate often used in other studies of ester hydrolysis.

In a set of 214 scaffold proteins,[21] we used RosettaMatch[29] to search for constellations of protein backbones that could accommodate these functional groups (Cys, His, Asn/Gln, and two backbone NH oxyanion hole contacts) in cysteine hydrolase-like geometries. Initial calculations showed that no placements could be found for this five-residue arrangement (data not shown).

dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197−16206

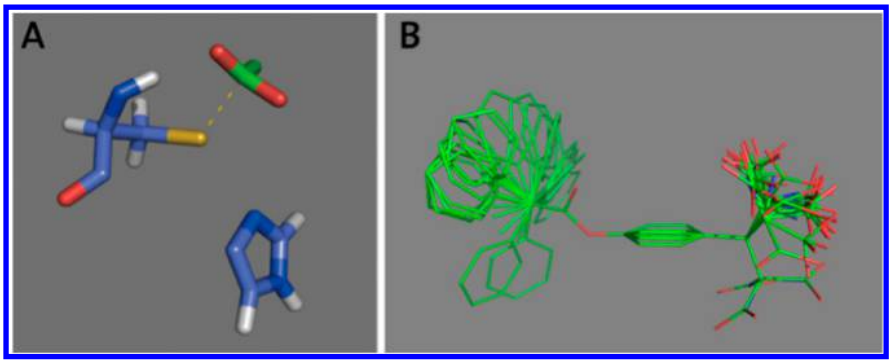


Figure 2. Snapshots of the computational design process. (A) Representation of the calculated theozyme of the ester substrate framed by the catalytic dyad (Cys-His) and the backbone NH-oxyanion contact. To increase the number of matches, both the histidine sidechain and the substrate could rotate with respect to the cysteine (not shown). Note that in this case, the backbone NH contact is made by the cysteine itself. (B) Image of the theoretical conformer ensemble of tyrosyl ester **1**.

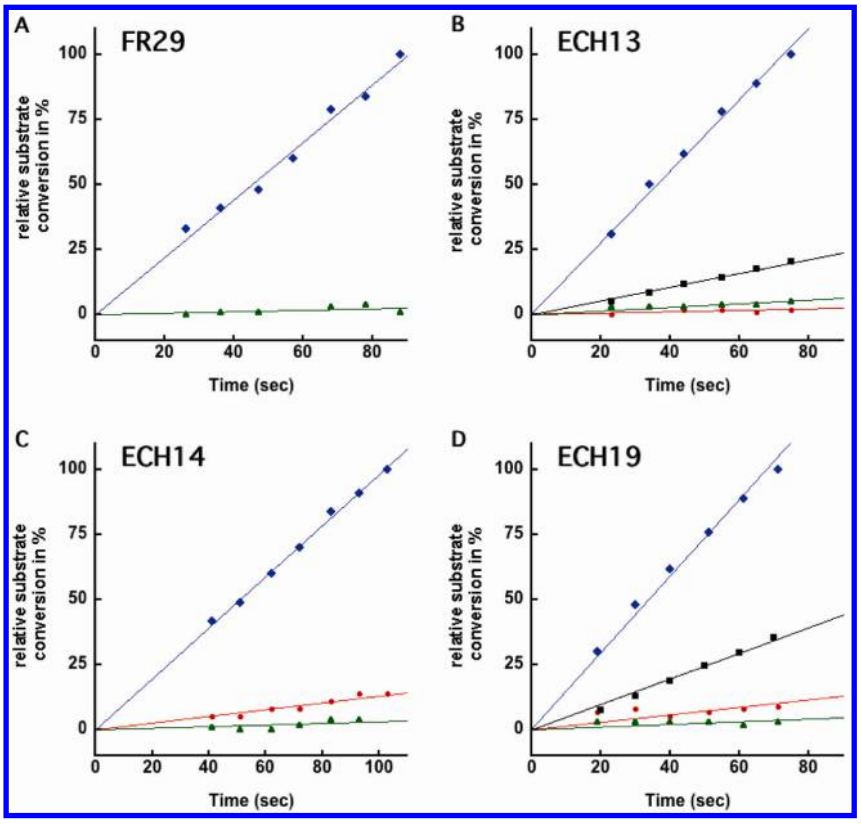


Figure 3. Experimental characterization of the active designs and their respective knockout variants. The progress curves of the parental designs are depicted in blue, the traces of the single knockout variants (cysteine) are shown red, traces of the double knockout variants (cysteine and histidine) are illustrated in green, and the histidine knockout variants for ECH13 and ECH19 are shown in black: (A) FR29, (B) ECH13, (C) ECH14, and (D) ECH19. The enzymes (5 μM) were tested with the coumarin ester **1** (FR29 design, 20 μM; ECH designs, 50 μM) and the reaction progress was monitored by measuring the appearance of the fluorescent coumarin product (excitation wavelength, 340 nm; emission wavelength, 452 nm). For each graph, the amount of substrate that was converted by the active designs after 70−100 s was set to 100% and then used to normalize the entire data set. The background was subtracted in all cases and the linear fits were extrapolated to zero substrate conversion.

Mutation of the catalytic triad residue asparagine/glutamine to alanine does not abolish catalytic activity in natural cysteine hydrolases,[30] suggesting that a Cys-His dyad should be sufficient for activity. Furthermore, the oxyanion intermediate can, in principle, be stabilized by any appropriately positioned hydrogen bond donors including water molecules (which presumably perform this function in the uncatalyzed reaction). Therefore, we generated theozymes[31,32] as described in the Methods section consisting of the central cysteine residue programmed to carry out nucleophilic attack, a histidine residue to assist with the various proton shuffling steps occurring during the reaction, and three possible oxyanion-stabilization schemes (Figure S10). In the first set of designs, a backbone NH group serves as the oxyanion stabilizer, and in the second and third theozymes, explicit water molecules or side-chain functional groups were used. An example of the first theozyme is shown in Figure 2A. RosettaMatch[29] was then used to scan for matches to these theozymes in the set of 214 protein scaffolds. For the three-residue theozyme I (featuring one backbone NH), a total of 207 unique matches was identified in

dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197−16206

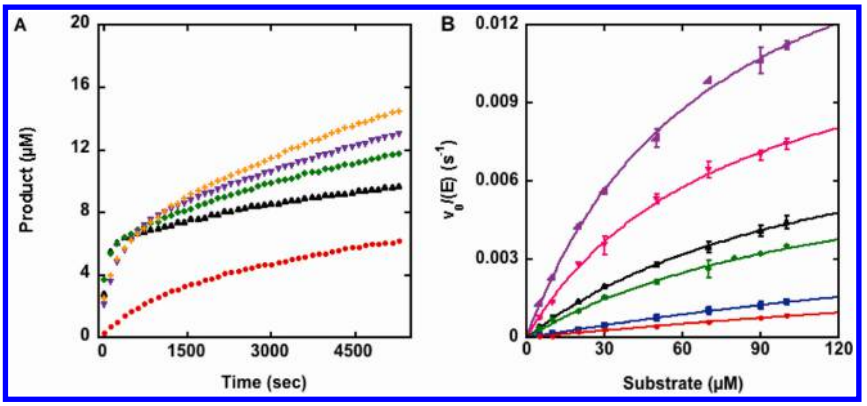


Figure 4. Kinetic analysis of the designed esterases: (A) Two-phase progress curves of selected *de novo*-designed ester hydrolases. The conversion of 100 μM coumarin ester **1** by 10 μM of FR29 (red), ECH13 (black), ECH19 (green), FR29 T112L (purple) and FR29 A44S/V151L (orange) consists of an initial fast phase followed by a second, slow phase. (B) Michaelis–Menten plots of the hydrolysis of coumarin ester **2** and by the *in silico*-designed ester hydrolases and their best evolved variants (red, FR29; blue, ECH14; green, ECH19; black, ECH13; pink, FR29 A44S T112L V151L; purple, ECH19 K354P P364W). Only the slopes of the fast phases were considered for the determination of k_2 and K_M.

81 distinct scaffolds. In 178 of the matches, the amide backbone of the cysteine itself provided a hydrogen bonding contact to the oxyanion of the tetrahedral intermediate, as often observed in natural cysteine and serine hydrolases. Every match was designed 100 times and the resulting designs were filtered and ranked as described in the Methods section. We selected 31 theozyme I designs, 12 theozyme II designs, and 12 theozyme III designs for experimental testing (Table 1S).

Initial Activity Screen. Genes encoding the 55 designs were cloned into the pET29b+ vector (Novagen) and the proteins were expressed and purified by Ni-NTA affinity chromatography. Purified soluble protein was obtained for 19 theozyme I designs, 5 theozyme II designs, and 8 theozyme III designs; the remaining designs did not yield soluble protein. Hydrolytic activity was evaluated by measuring the increase in fluorescence due to the hydrolysis of coumarin ester **2**. Four of the theozyme I designs showed measurable activity, but none of the theozyme II or theozyme III designs were active, suggesting that backbone NH groups might make more robust interactions with the tetrahedral oxyanion than polar side chains or localized water molecules. For each of the four active designs, substitution of the catalytic cysteine and histidine residues with alanine, as well as the single knockout of the catalytic cysteine, either abolished activity completely (FR29, ECH13, ECH19) or decreased it considerably (ECH14), indicating that the source of the observed activity is in fact the designed active site (Figure 3).

Optimization of the Active Designs. Although the computational design process afforded novel active sites capable of cleaving the activated coumarin ester, the rate accelerations over background provided by the esterases are modest. Guided by visual inspection and by evaluation of active site dynamics, we explored the effects of mutations on catalytic activity for each of the designs. The optimized variants were generated as described in the SI. For FR29 we constructed and analyzed 46 single mutants and for the ECH designs we screened 5–8 mutants each. Mutations increasing activity more than 1.5-fold were combined into second-generation variants and the hydrolytic activity was re-tested. For design FR29 this step was carried out twice to generate third-generation variants, which contained up to seven mutations compared to the parental design (SI, Figures 1S and 2S). A set of three mutations (A44S/T112L/V151L) was found to increase the overall catalytic efficiency 17-fold relative to the original design. ECH13 showed a low tolerance to additional mutations, and most variants exhibited a dramatic decrease in expression yield (SI, Figure 1S-A). ECH14 was more tolerant of substitutions, but since mutation of the active site cysteine does not completely eliminate activity we did not explore these further.

ECH19 is based on a periplasmic binding protein (PBP)[33] scaffold that contains two large domains connected by a flexible hinge region, allowing the two domains to open and close around a cavity. The crystal structure of the closed form was used in the design, but most PBPs are in the open conformation in the absence of ligand. MD simulations[34] of ECH19 show an irreversible transition from closed to open conformation, both with and without substrate (Figure S7). Single-mutant $\Delta\Delta G$ calculations[35] on the open and closed forms of the ECH19 scaffold suggested that incorporation of a proline at position 354, which is located in the hinge region, would increase the stability of the closed form. A variant with two mutations, K354P and P364W, enhanced the esterase activity 4-fold compared to the original design (SI, Figure 1S-C). The P364W mutation was intended to enhance the binding of the esters by optimizing packing around the acyl moiety of the substrate.

Biophysical and Kinetic Characterization of the Ester Hydrolases. For more detailed biophysical and biochemical characterization, the designed hydrolases were purified by an additional anion exchange step following standard Ni-NTA affinity chromatography. Variant FR29 was further purified by either gel filtration or GST-affinity chromatography. The specific activity was found to be independent of the purification method, arguing against contamination by endogenous esterases. Consistent with this conclusion, active site alanine mutants of the designed esterases, which were purified in an identical manner to the active biocatalysts, did not convert the ester substrates above the buffer background rate. Mass spectrometric analysis confirmed the identity of the individual variants (SI, Table 2S), and circular dichroism (CD) measurements verified that they were folded and exhibited similar stabilities to the parental designs (SI, Figure 3S and Table 3S).

For all four designs, cleavage of ester **2** exhibits a biphasic time course with an initial fast phase followed by a second slow phase (Figure 4A). The slow phase is roughly 2–3-fold above the spontaneous hydrolysis reaction, whereas the fast phase

dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197–16206

varies considerably depending on the analyzed variant. Similar behavior was observed for the cleavage of ester **3**. The burst phase was studied as a function of substrate concentration to determine steady-state parameters (Figure 4B, Table 1). The

Table 1. Kinetic Parameters of the *in Silico*-Designed Ester Hydrolases and Their Evolved Variants[a]

enzyme	substrate	$k_2 \times 10^3$ (s^{-1})	K_M (μM)	k_2/K_M (M^{-1} s^{-1})
FR29	**2**	5.1 ± 0.7	500 ± 180	10
	3	4.1 ± 0.6	120 ± 20	34
FR29 A44S/T112L/ V151L	**2**	13.3 ± 0.6	78 ± 7	170
	3	15.4 ± 1.1	38 ± 5	405
ECH13	**2**	9.6 ± 0.4	120 ± 10	80
	3	17.6 ± 1.9	57 ± 10	309
ECH14	**2**	6.3 ± 0.9	360 ± 60	17
	3	8.2 ± 1.9	130 ± 40	63
ECH19	**2**	7.7 ± 0.7	125 ± 20	62
	3	10.0 ± 0.7	44 ± 5	227
ECH19 K354P/ P364W	**2**	19.5 ± 0.7	73 ± 5	267
	3	n.d.	n.d.	n.d.

[a]Measurements were performed in buffer (25 mM HEPES, 100 mM NaCl, 5% acetonitrile) at pH 7.5 and 29 °C. Only the rates of the initial phase were considered for the determination of k_2 and K_M. n.d. = not determined

computationally designed ester hydrolases exhibit Michaelis–Menten kinetics. Apparent bimolecular rate constants (k_{cat}/K_M) for the acylation step range between 10 M^{-1} s^{-1} (FR29) and 80 M^{-1} s^{-1} (ECH13) for the conversion of coumarin ester **2**. *p*-Nitrophenyl ester **3**, which is also a much better mimic of tyrosyl ester **1**, toward which the computational designs were generated, is hydrolyzed with up to 4-fold higher catalytic efficiency. In this case, the values for k_{cat}/K_M range between 30 M^{-1} s^{-1} (FR29) and 320 M^{-1} s^{-1} (ECH13). The slightly higher turnover numbers can be attributed to the greater reactivity of substrate **3** compared to ester **2**. However, cleavage of ester **3** is also characterized by a 2- to 3-fold lower Michaelis constant, consistent with a better fit to the active site.

As summarized in Table 1, the catalytic efficiency of the acylation step could be successfully improved up to 17-fold by introducing point mutations into the parental designs. The best third-generation ester hydrolase variant, FR29 A44S/T112L/V151L, exhibits k_{cat}/K_M values of approximately 400 M^{-1} s^{-1}. This value is within the range achieved by typical catalytic antibodies (Table 2). However, the best hydrolytic antibody has been found to display a 1000-fold higher catalytic efficiency[15,16] and the natural cysteine protease papain cleaves *p*-nitrophenyl hippurate with a k_{cat}/K_M of 1.8×10^5 M^{-1} s^{-1}.[36]

To assay the cleavage of tyrosyl ester **1**, which was the designed substrate but less reactive than esters **2** and **3**, we monitored formation of the product tyrosine by HPLC. Although no cleavage of this substrate was detected in the presence of the designed proteins, this compound inhibits the hydrolysis of ester **2** indicating binding to the active site. IC_{50} values were determined for the most active hydrolases FR29 A44S/T112L/V151L and ECH19 K354P/P364W by recording

Table 2. Kinetic Parameters for Ester Cleavage Catalyzed by Artificial Biocatalysts

enzyme	$k_{cat} \times 10^{-3}$ (s^{-1})	K_M (μM)	k_{cat}/K_M (M^{-1} s^{-1})
FR29 (best variant)	15.4 ± 1.1	38 ± 5	405
CNJ206[52] [a]	7 ± 0.8	80 ± 10	87
48G7[53] [b]	35	113	310
17E8[54] [c]	817 ± 28	215 ± 33	3700
43C9[55] [d]	460 ± 50	470 ± 160	979

[a]30 mM TBS, pH 8.0. [b]10 mM Tris-HCl, 50 mM NaCl, pH 8.2, 37 °C. [c]50 mM Tris-HCl, 150 mM NaCl, pH 8.7. [d]100 mM ACES, 50 mM Tris-HCl, 50 mM CAPS, pH 8.5, 25 °C.

esterase activity in the presence of increasing concentrations of inhibitor. Tyrosyl ester **1** inhibits the FR29 and ECH19 variants with an IC_{50} of 37 and 45 μM corresponding to K_i values of 23 and 27 μM, respectively (Figure 5).[37]

Burst kinetics is expected for a two-step reaction mechanism in which the substrate rapidly reacts with the enzyme to form stoichiometric amounts of a stable enzyme-bound intermediate that subsequently breaks down slowly (eq 1). The observation

$$E + S \rightleftharpoons ES \underset{P_1}{\overset{k_1}{\rightarrow}} EI \xrightarrow{k_2} E + P_2 \quad (1)$$

of the burst phase is thus evidence that the four novel hydrolases function through the designed mechanism; the fact that biphasic kinetics is not observed for the corresponding cysteine-knockout variants supports this conclusion. For the ECH designs and the optimized FR29 variants the second phase is much slower than the first phase ($k_2 \ll k_1$) and the observed bursts approximately correspond to the concentration of enzyme employed, and hence to a single turnover. This correlation provides additional evidence that the observed activities are due to the designs and not to a highly active esterase contaminant.

Typically, the acyl-enzyme intermediate that is formed in serine and cysteine proteases undergoes hydrolysis, but can also undergo aminolysis when in the presence of amine nucleophiles.[38] Since aminolysis of thioesters is known to be significantly faster than for esters,[39] we attempted to facilitate liberation of the acyl-enzyme complex by adding amine nucleophiles spanning a range of reactivity to the assay mixtures (SI, Figure 6S) but no increase in activity was observed.

Characterization of the Acyl-Enzyme Intermediate by Mass Spectrometry. In analogy to natural cysteine hydrolases, the most likely intermediate formed during reaction of the designed proteins with substrate is an acyl-enzyme adduct. To test this hypothesis, we attempted to isolate this species and characterize it by mass spectrometry. For FR29, ECH13, and ECH19, a shift of 133 Da was observed for the major mass peak of the protein after incubation with substrate **2**. This increase corresponds to the acyl fragment of the substrate, as expected for acylated proteins. Consistent with this interpretation, the identically treated variants of FR29 and ECH19, in which the active site cysteine and histidine were mutated to alanine, did not show an increase in mass following incubation with substrate. In the case of ECH14, the protein was modified up to seven times and mutation of the active site cysteine did not alter the acylation pattern (SI, Figure 5S and Table 4S). In contrast to the results obtained with activated ester **2**, no acyl-enzyme intermediate was observed with tyrosyl derivative **1**,

dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197–16206

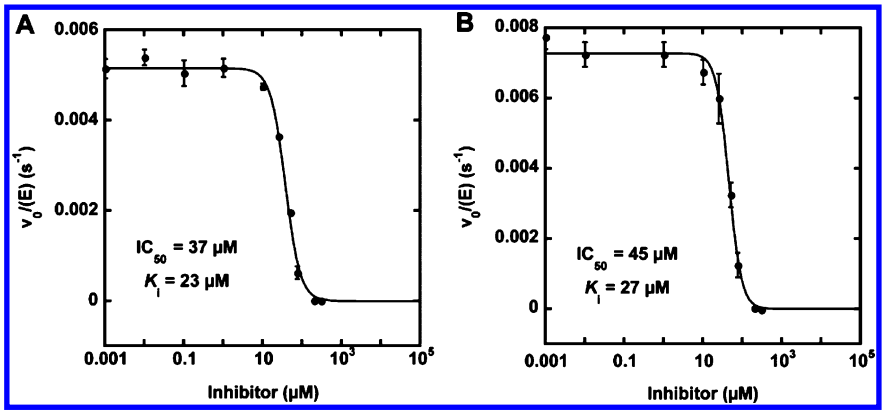


Figure 5. Inhibition of a FR29 and a ECH19 variant by tyrosyl ester **1**. The IC_{50} of tyrosyl ester **1** was determined by incubating (A) FR29 A44S/T112L/V151L ([E] = 2 μM) and (B) ECH19 K354P/P364W ([E] = 2 μM) with increasing amounts of the inhibitor ([c] = 0.001−300 μM) before 50 μM of coumarin ester **2** was added and the release of coumarin was monitored. The IC_{50} was determined by curve fitting and converted into the corresponding K_i using the Cheng−Prusoff equation.[37]

which inhibits the enzymes but is not cleaved and thus does not covalently modify the proteins (Table 5S).

To reveal the location of the modification, the acylated proteins were analyzed by matrix-assisted laser desorption/ionization tandem mass spectrometry (MALDI-MS/MS). For variants FR29 A44S/V151L/M133Y (a precursor of the most active FR29 variant) and ECH13 the active site cysteine was found to be acylated. In the case of ECH13, an additional modification site was identified, namely a cysteine at the C-terminus of the protein. Unfortunately, low peptide intensities precluded conclusive MALDI-MS/MS analysis of ECH14 and ECH19.

The MS data thus show that the designed cysteine functions as the catalytic nucleophile in three of the four designs (ECH13, ECH19, FR29), while in ECH14 the results are less clear. These findings fit well with the previously reported observation that ECH13 and ECH19, but not the cysteine-knockout variants, efficiently react with cysteine-hydrolase specific probes.[40]

Structural Characterization. Crystal structures of the apo forms of ECH13 (at 1.6 Å resolution, PDB code 3u13), ECH14 (at 3.2 Å, 3uak), ECH19 P364W (at 2.55 Å, 3u1o), and FR29 A44S/T112L/V151L (at 2.8 Å, 3u1v) were determined (Figure 6, SI Table 10S). The overall backbone structures of the designed proteins were similar to the design models. However, in each case either the designed histidine residue or the nucleophilic cysteine adopt a conformation different from the design model, and the dyad is not formed as desired; this was also observed in MD simulations (Figures S7, S8, S9). Furthermore, in each case, either the histidine or cysteine is in regions of the protein with relatively high flexibility. For two designs that were based on the ligand-bound, holo conformation of scaffolds which undergo global conformational changes upon ligand binding (ECH19, based on a periplasmic binding protein[33] and FR29, based on a tryptophanyl-tRNA synthetase[41]), crystal conformations resemble the unliganded, apo conformation of the scaffold. MD simulations also suggest that the unbound conformation is thermodynamically more favored. Incorporation of backbone flexibility and avoidance of alternative, non-catalytic states can consequently be envisioned as avenues for improvement of the design methodology.

Even though dyads are not formed in the apo-crystal structures of ECH13 and ECH19, substituting the catalytic histidine with alanine leads to diminished catalytic activity (Figure 3). This observation indicates that the histidine does take part in the reaction and that the energy gap between the designed (holo) and observed (apo) conformations may be overcome upon ligand binding. Although preliminary experiments to crystallize the enzymes with compound **1** and **2** have been unsuccessful, structures with bound ligand will be important for verifying this hypothesis. The catalytic activities would presumably be significantly higher if the desired theozyme geometry had been fully realized in the designs.

■ DISCUSSION

We have described the construction of primitive esterases by combining Cys-His dyads with appropriately positioned oxyanion holes. Comparison of the structures and activities of the designs to those of native hydrolases provides insight into the contributions of these functional elements to catalysis and where improvements can be made in the design process.

The first insight relates to the relative ease with which the developing negative charge on the carbonyl oxygen can be stabilized by backbone NH groups. We experimented with different strategies for oxyanion stabilization using oxyanion holes formed by backbone NH groups, side-chain NH groups, or water molecules. Notably, all of the active designs utilized a backbone NH group for oxyanion stabilization. Four out of 19 (~20%) soluble designs with backbone NH based oxyanion holes had activity, while none of the designs using side-chain or discrete water-mediated oxyanion stabilization were active. Although direct evidence for an oxyanion binding site is lacking, and there may be other reasons why the theozyme II and III designs did not work, this anecdotal correlation is suggestive. In all active designs except ECH19, the backbone amide of the active site cysteine residue itself provides oxyanion stabilization, and in all active designs except FR29, the catalytic Cys residue and the oxyanion donor are at the N-terminus of an α-helix which may provide additional oxyanion stabilization due to its helix dipole. This feature might represent an intrinsic advantage of theozyme I. Backbone oxyanion holes are found in almost all proteases and many esterases and have apparently evolved independently multiple times. Because the protein backbone is on average more rigid than side chains, it is possible that the preferential use of backbone amides for oxyanion hole stabilization in designed (and natural) hydrolases reflects the advantages of pre-organized active sites that are poised to stabilize the oxyanion intermediate, thereby aiding nucleophilic

dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197−16206

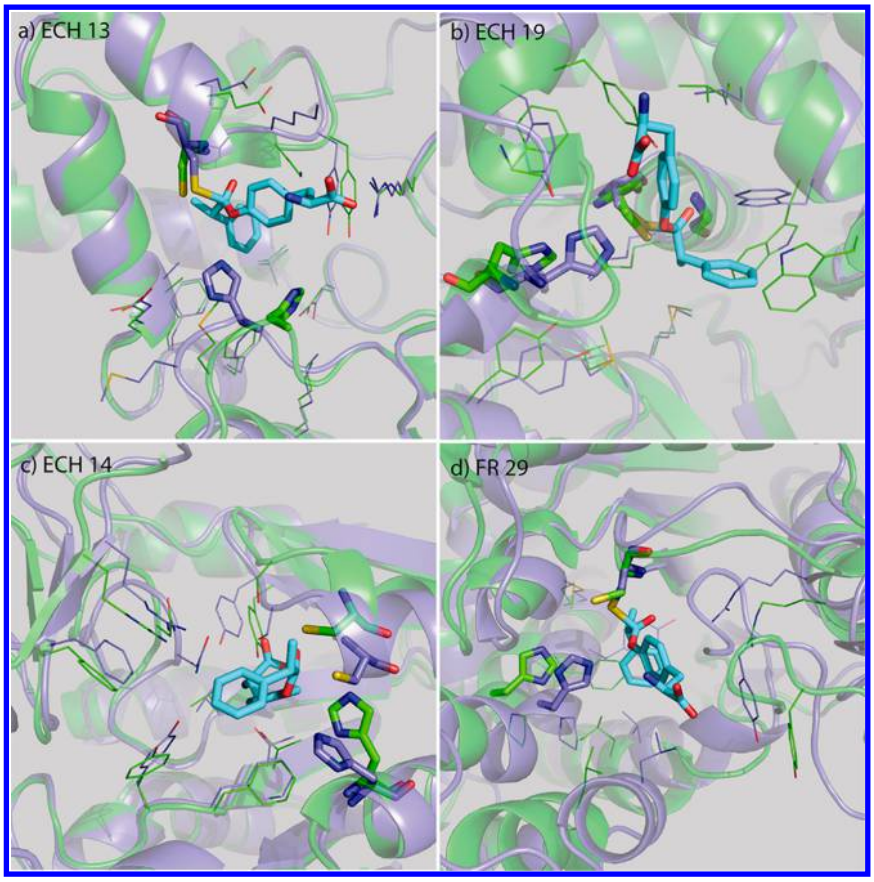


Figure 6. Crystal structures of the four active designs. In each case, the design model is shown in purple with the ligand in cyan, and the crystal structure in green. The theozyme residues and the ligand are shown in stick representation, and selected other active site residues in line representation. (A) ECH13: The Cα rmsd between design model and crystal structure is 0.97 Å over the 15 active site residues. The catalytic histidine, His100, is in a rotameric conformation different from the design model, and instead of pointing toward the ligand and Cys45, it makes a hydrogen bond with Asp10. This alternative conformation is facilitated by a small backbone shift between residues Pro99 and Phe103, and the observed close interaction between His100 and Asp10 would not be possible with the scaffold backbone conformation that served as the template for the design. (B) ECH19 P364W: The design was based on the closed conformation of a periplasmic binding protein, but the apo protein crystallized in the open form, with an rmsd of 4.1 Å to the design model and 1.6 Å to the open form of the scaffold protein (PDB 2uvg). The designed active site is mostly located in one of the scaffold's two domains, close to the interdomain cleft. When superimposing design model and crystal structure based solely on the active-site containing domain, the resulting rmsd is 1.5 Å. However, the catalytic His226 does not interact with Cys161 as designed, but adopts a different rotameric conformation to interact with the side chain-hydroxyl of Tyr250 and the backbone oxygen of Phe221. The segment from Tyr218 to Lys230 that contains His226 has high relative B-factors, suggesting that it is fairly flexible. (C) ECH14: the crystal structure has an rmsd of 1.4 Å to the design. The catalytic dyad is not formed, as the Cys132 containing loop-helix stretch between residues 127 and 140 moves upward away from the active site and His104 reorients around χ^2. This unexpected movement may result from the W130S mutation, since W130 stacks against the PLP cofactor of the wild-type scaffold and thus locks this backbone segment into the conformation used as the design template. (D) FR29 A44S/T112L/V151L: The apo-structure of FR29 is more similar to the unliganded, more open conformation of the scaffold (PDB 1D2R, 0.86 Å rmsd) than to the ligand-bound structure (PDB 1mau, 2.7 Å rmsd) which was used as the template for the design. The catalytic dyad is not formed, since in the apo form of the scaffold the helix-turn-helix motif between residues 106 and 132 that contains the catalytic His125 moves outward relative to the catalytic Cys9, leading to shift in His125 Cα–Cys9 Cα distance from 10.9 to 12.4 Å. As the backbone of most of the designed active site residues shifts between the liganded and the apo structure, the active site is generally more open than in the design model.

attack. In natural hydrolases, the full oxyanion hole is often formed by two backbone NH groups or a backbone NH plus a sidechain amide, and the removal of only one of these interactions can reduce catalytic efficiency 10^2–10^3-fold.[42] Adding a second interaction to the partial oxyanion hole of the computational designs might therefore improve catalytic efficacy. This suggestion is supported by comparison with the antibody esterases, which have more completely formed oxyanion holes and significant activity even without a catalytic dyad or triad.[14]

The second insight concerns the relationship between the reactivity of the catalytic cysteine and the effectiveness of the active site for catalysis. The active site cysteines in ECH13 and ECH19 react as strongly with cysteine protease-specific probes as the nucleophiles in natural cysteine proteases.[41] Nevertheless, the acylation efficiency of the designed esterases is more than 3 orders of magnitude lower than that observed for the natural hydrolase papain.[43] Thus, activating the cysteine is not difficult even for primitive designs. The contrast between high nucleophilicity of the active site cysteine yet low catalytic activity may be due to relatively poor oxyanion stabilization, or as suggested in studies with thiosubtilisin,[44] a failure to protonate the substrate leaving group. Nevertheless, simply placing a cysteine in a protein binding pocket is not sufficient to generate an active hydrolase, as evidenced by the lack of activity observed for the other soluble designs.

dx.doi.org/10.1021/ja3037367 | J. Am. Chem. Soc. 2012, 134, 16197–16206

The third insight concerns the effectiveness of the combination of the catalytic dyad and minimalist oxyanion holes utilized in the designs for acylation versus deacylation. The most proficient of the designs accelerates acylation, the first step of the reaction sequence, with an apparent bimolecular rate constant of about 400 M^{-1} s^{-1} and a rate acceleration of approximately 3.7 × 10^3-fold over background. However, in contrast to naturally occurring enzymes and antibody catalysts, the computationally designed proteins are not efficient multiple turnover catalysts. Kinetic and mass spectrometric studies establish that the nucleophilic cysteine attacks the substrate and becomes acylated as programmed by design, but the resulting acyl-enzyme intermediate is hydrolyzed only slowly. Deacylation evidently places more demands on the catalytic machinery than the acylation step. This is due, at least in part, to the fact that the designs work on activated esters. It may also reflect the fact that the theozyme does not capture the full complexity of this multi-step transformation. While theozyme I may be a good representation for the initial transthioesterification step, it does not explicitly model general base delivery of water to the acyl-enzyme intermediate or possible contributions of conformational dynamics and preorganization toward catalysis.

The fourth insight relates the contribution of the histidine of the engineered Cys-His dyad to the actual geometry of the active site. The crystal structures and the MD simulations show that the engineered Cys-His dyad is usually not in the designed conformation, yet mutation of the histidine to alanine reduces the rate of catalysis. Evidently histidine can promote weak ester cleavage by a nearby cysteine in the absence of a stable hydrogen bond between the two, which suggests that the familiar catalytic triad could have evolved in a stepwise fashion, with a histidine in the vicinity of a cysteine promoting catalysis and only later fixed in place by the third amino acid in the triad. By analogy, increasing the catalytic efficacy of the designed esterases will likely require expansion of the dyad motif into a full catalytic triad to stabilize the histidine in the desired conformation and promote its alternating roles as general base and acid in the overall catalytic cycle. The hydrogen bonds between thiol donors and nitrogen acceptors are not as strong as those between oxygen and nitrogen donors,[45] and hence the interaction energy between the cysteine and the histidine might not suffice to hold the histidine in position. Backing up the histidine with a hydrogen-bond acceptor would also favor the correct tautomeric state necessary for the histidine to deprotonate the cysteine in the acylation step and water in the deacylation step.[46] The extent to which such a designed triad is exposed to bulk-solvent is also likely to be important. Mutation of naturally occurring enzymes illustrates the importance of these effects; for example the turnover number (k_{cat}) of papain drops by 2 orders of magnitude when the catalytic asparagine of the Cys-His-Asn triad is replaced with an alanine residue[30] and larger effects have been observed in other hydrolases.[42]

There are a number of avenues for improving the primitive esterases described in this manuscript. Foremost among these is converting the poorly formed catalytic dyad into a full catalytic triad to hold the histidine in place for cysteine activation, leaving group protonation, and water activation in the deacylation step. Achievement of this goal would be expected to speed up the rate-limiting deacylation step as well as acylation of the enzymes by less activated ester and amide substrates. Supplementing the single backbone NH group in the primitive oxyanion holes with additional backbone or side-chain hydrogen bond donors for better tetrahedral intermediate stabilization is also likely to be important. Creating these more sophisticated active sites will require going beyond the fixed-backbone approach utilized in this work since, as found in the RosettaMatch calculations described in the Results section, the five required functional groups cannot be placed in the proper relative orientations without modification of the backbone of the scaffolds.

■ CONCLUSIONS

Our "inside out" de novo design results complement extensive "top down" studies on the effects of removing catalytic triad residues from naturally occurring proteases and esterases. In nature, this special constellation of amino acids possibly evolved by the stepwise addition of the individual catalytic groups to an ancestral binding pocket. Our results suggest that nascent catalysts could have utilized a cysteine as a nucleophile with a histidine nearby and a backbone hydrogen-bond to stabilize the oxyanion. How subsequent evolution matured such primitive sites into the extremely proficient modern day catalysts is unclear; but improved design methods and laboratory evolution should not only afford progressively more active variants but also further insights into the origins of the strong inter-residue synergies that distinguish highly evolved hydrolytic enzymes.

■ METHODS

Computational Design. Quantum-mechanical methods were used to perform theozyme calculations. The optimal arrangement was computed for a model catalytic triad and oxyanion hole contacts along the reaction path of ester hydrolysis as reported by Smith et al.[18] Reactant, intermediates, and transition states were obtained by systematically stepping along the reaction coordinate, followed by optimizations toward the respective stationary point. Of particular interest for the purpose of enzyme design in general are the transition state (TS) geometries along the reaction steps. Here we focused on the first TS of the acylation step in which the ester substrate undergoes nucleophilic attack by a reactive thiol. The QM theozyme consists of a Cys-His-Glu/Asp triad and two oxyanion hole contacts and was computed at the B3LYP/6-31G(d) level of theory using Gaussian 03 (SI references). Currently, the computational expense of matching more than three groups into a protein scaffold presents a bottleneck in the design protocol, as a consequence of which a stripped down version of the theozyme was utilized (Figure 2A). The conformations of the catalytic Cys/His dyad and of an oxyanion hole are similar to the active site of human cathepsin K which was investigated in a previous QM/MM study.[27] An ensemble of ligand conformers was then generated employing OpenEye's Omega software[48] (Figure 2 B). The final theozyme thus consisted of a conformer library of the substrate in the transition state, the Cys-His catalytic dyad, plus one of the three oxyanion hole possibilities shown in Figure S10. Attachment sites for the theozyme in the scaffold set were found with an improved version[24] of the RosettaMatch algorithm.[28] For theozyme I, the 207 unique matches that were identified were subjected to 100 iterations of the standard Rosetta3 enzyme design protocol.[24] Briefly, all scaffold residues (except the matched catalytic residues) containing either a Cα within 6 Å of a matched ligand atom or both Cα and a Cβ atom within 8 Å of a matched ligand atom were considered design shell residues and mutated to alanine. The resulting structure was then subjected to a gradient-based minimization to optimize the three catalytic interactions. During this minimization, restraints were added to the energy function to enforce the desired theozyme geometry. Three rounds of sequence design with subsequent gradient-based minimization were carried out for the shell residues. Scaffold residues with a Cα within 13 Å of a ligand atom but not part of the design shell were allowed to change their rotameric state. Finally, the designed structures were

repacked without the catalytic restraints. From the resulting 20700 design models, 1071 were selected based on the following criteria: (1) a ligand binding score less than −10.0 Rosetta energy units (REU), (2) no more than two unsatisfied, buried polar ligand atoms, (3) more than 66% of ligand surface area buried by the protein, (4) fewer than three buried unsatisfied polar atoms on the catalytic histidine, (5) fewer than two overall unsatisfied polar atoms compared to the respective wild-type scaffold, (6) packing statistics[49] comparable to the wild-type scaffold, and (7) greater than −2 nonlocal contacts compared to the wild-type scaffold (residues with an interaction score less than −1.0 REU and separated in sequence by at least eight residues were considered to be a nonlocal contact). The cutoff of −10.0 REU for the ligand binding score was determined according to the MASC method[50] by docking substrate **1** to a set of 68 random proteins with the Rosetta ligand docking protocol.[51]

Protein Production, Initial Activity Screening, and Biochemical Characterization. All design constructs were cloned into pET29b and expressed in *E. coli* BL21 cells. Proteins providing crystal structures included NESG targets OR49 (ECH19 P364W), OR51 (ECH13), OR52 (FR29 A44S/T112L/V151L), and OR54 (ECH14). pET expression vectors for these proteins and the mutants listed in Table 6S, 7S, 8S, 9S have been deposited in the PSI Materials Repository (http://psimr.asu.edu/). Purification was carried out as described in the SI. Initial activity determination, CD spectroscopy, melting curves and mass spectrometric analysis were carried out as described in the SI.

Kinetic Measurements. The substrates were synthesized as described in the SI. Reactions were initiated by adding different amounts of coumarin ester **2** (5−100 μM final concentration) or *p*-nitrophenyl ester **3** (5−50 μM final concentration) in acetonitrile to 2 μM of protein (or no protein for the background reaction) in 25 mM HEPES buffer (pH 7.5), containing 100 mM NaCl and 5% acetonitrile. Initial reaction rates were determined as described in the SI. The initial rates divided by the catalyst concentration were plotted against substrate concentration, and k_{cat} and K_M were determined by fitting the data to the Michaelis−Menten equation $v/[\text{catalyst}] = k_2[\text{substrate}]/(K_M + [\text{substrate}])$ using Kaleidagraph software (Synergy Software). The aminolysis and KI determination measurements were performed as described in the SI.

Mass Spectrometry. For standard mass determination, the protein samples were desalted using Illustra Nap-5 columns (GE Healthcare, Glattbrugg, Switzerland) and measured in 0.1% acetic acid (pH 2.0) by ESI-MS on a Daltonics maXis ESI-Q-TOF mass spectrometer (Bruker). The mass spectra of the proteins were deconvoluted using MaxEnt1 software. Protein samples were prepared as described in the SI. Then the samples were acidified, desalted using C_{18} ZipTips and analyzed by MALDI-MS in the positive-ion mode. Acylated peptides were identified by comparison of the mass spectra of treated proteins with the mass spectra of the corresponding negative controls. The identified peptides were additionally fragmented and the obtained fragments were again compared to those of the corresponding non-modified peptides. Additional details can be found in the SI.

Structure Determination and Molecular Dynamics Simulations. The structures were determined as a collaborative project of the Community Target Nomination program of the NIH PSI Northeast Structural Genomics Consortium (www.nesg.org). Details of the experimental procedure, as well as the Molecular Dynamics setup, can be found in the SI.

■ ASSOCIATED CONTENT

Supporting Information

Complete experimental procedures, plus tables and figures illustrating the design and characterization of the artificial esterases. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

hilvert@org.chem.ethz.ch; dabaker@u.washington.edu

Present Addresses

[+]Corporate R&D Division, Firmenich SA, CH-1211 Geneva, Switzerland

[△]Department of Chemistry, Stanford University, Stanford, CA 94305

[◆]Department of Radiation Medicine, University of Kentucky, Lexington, KY 40536

Author Contributions

[×]F.R. and R.B. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to T. Acton, C. Ciccosanti, S. Sahdev, D. Patel, and L. Mayo for experimental support with sample preparation and structure determination. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA), by the Defense Threat Reduction Agency (DTRA) grant HDtRA-11-1-0041 and by the ETH Zürich, and by a grant from the National Institute of General Medical Sciences Protein Structure Initiative (PSI), U54-GM094597 (to G.T.M., J.H., and L.T.). Fellowship support from the Fonds des Verbandes der chemischen Industrie and the Studienstiftung des deutschen Volkes (to R.B.) and the Marie Curie Action within the FP7-PEOPLE program (Z.P.) is gratefully acknowledged.

■ REFERENCES

(1) Clark, J. D.; Schievella, A. R.; Nalefski, E. A.; Lin, L. L. *J. Lipid Mediat. Cell. Signal.* **1995**, *12*, 83−117.
(2) Mignatti, P.; Rifkin, D. B. *Enzyme Protein* **1996**, *49*, 117−137.
(3) DeClerck, Y. A.; Imren, S.; Montgomery, A. M.; Mueller, B. M.; Reisfeld, R. A.; Laug, W. E. *Adv. Exp. Med. Biol.* **1997**, *425*, 89−97.
(4) MacDonald, T.; DeClerck, Y.; Laug, W. *Thromb. Haemostasis* **1997**, S1541−S1541.
(5) Gorrell, M. D. *Clin. Sci. (Lond)* **2005**, *108*, 277−292.
(6) Panke, S.; Held, M.; Wubbolts, M. *Curr. Opin. Biotechnol.* **2004**, *15*, 272−279.
(7) Sio, C. F.; Quax, W. J. *Curr. Opin. Biotechnol.* **2004**, *15*, 349−355.
(8) Riva, S.; Koskinen, A. M. P.; Klibanov, A. M. *Enzymatic Reactions in Organic Media*; Springer: Berlin, 1995; p 140.
(9) Patel, R. N. *Annu. Rev. Microbiol.* **1998**, *52*, 361−395.
(10) Breslow, R.; Dong, S. D. *Chem. Rev.* **1998**, *98*, 1997−2012.
(11) Murakami, Y. *Top. Curr. Chem.* **1983**, *115*, 107−155.
(12) Cram, D. J. *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 1009−1020.
(13) Tanaka, F. *Chem. Rev.* **2002**, *102*, 4885−4906.
(14) MacBeath, G.; Hilvert, D. *Chem. Biol.* **1996**, *3*, 433−445.
(15) Janda, K. D.; Schloeder, D.; Benkovic, S. J.; Lerner, R. A. *Science* **1988**, *241*, 1188−1191.
(16) Stewart, J. D.; Krebs, J. F.; Siuzdak, G.; Berdis, A. J.; Smithrud, D. B.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7404−7409.
(17) Bolon, D N; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14274−14279.
(18) Smith, A. J. T.; Muller, R.; Toscano, M. D.; Kast, P.; Hellinga, H. W.; Hilvert, D; Houk, K. N. *J. Am. Chem. Soc.* **2008**, *130*, 15361−15373.
(19) (a) Simón, L; Goodman, J. M. *J. Org. Chem.* **2010**, *75* (6), 1831−1840. (b) Kamerlin, S. C. L.; Zhen, T. C.; Warshel, A. *J. Org. Chem.* **2010**, *75*, 6391−6401.
(20) Röthlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.;

Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190−195.
(21) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387−1391.
(22) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St Clair, J. L.; Gallaher, J. L.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. *Science* **2010**, *329*, 309−313.
(23) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y. E.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popovic, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. *Methods Enzymol.* **2011**, *487*, 545−574.
(24) Richter, F.; Leaver-Fay, A.; Khare, S. D.; Bjelic, S.; Baker, D. *PLoS ONE* **2011**, *6* (5), e19230.
(25) Hedstrom, L. *Chem. Rev.* **2002**, *102*, 4501−4523.
(26) Otto, H. H.; Schirmeister, T. *Chem. Rev.* **1997**, *97*, 133−171.
(27) Ma, S.; Devi-Kesavan, L. S.; Gao, J. *J. Am. Chem. Soc.* **2007**, *129*, 13633−13645.
(28) Kast, P.; Asif-Ullah, M.; Jiang, N.; Hilvert, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5043−5048.
(29) Zanghellini, A.; Jiang, L.; Wollacott, A. M.; Cheng, G.; Meiler, J.; Althoff, E. A.; Rothlisberger, D.; Baker, D. *Protein Sci.* **2006**, *15*, 2785−2794.
(30) Vernet, T.; Tessier, D. C.; Chatellier, J.; Plouffe, C.; Lee, T. S.; Thomas, D. Y.; Storer, A. C.; Menard, R. *J. Biol. Chem.* **1995**, *270*, 16645−16652.
(31) Tantillo, D. J.; Chen, J. G.; Houk, K. N. *Curr. Opin. Chem. Biol.* **1998**, *2*, 743−750.
(32) Zhang, X.; DeChancie, J.; Gunaydin, H.; Chowdry, A. B.; Clemente, F. R.; Smith, A. J. T.; Handel, T. M.; Houk, K. N. *J. Org. Chem.* **2008**, *73*, 889−899.
(33) Abbott, D. W.; Boraston, A. B. *J. Mol. Biol.* **2007**, *369*, 759−770.
(34) Kiss, G.; Rothlisberger, D.; Baker, D.; Houk, K. N. *Protein Sci.* **2010**, *19*, 1760−1773.
(35) Kellogg, E. H.; Leaver-Fay, A.; Baker, D. *Proteins* **2011**, *79*, 830−838.
(36) Lowe, G.; Williams, A. *Biochem. J.* **1965**, *96*, 199.
(37) Cheng, Y.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099−3108.
(38) Carter, P.; Abrahmsen, L.; Wells, J. A. *Biochemistry* **1991**, *30*, 6142−6148.
(39) Chu, S. H.; Mautner, H. G. *J. Org. Chem.* **1966**, *31*, 308−&.
(40) Weerapana, E.; Wang, C.; Simon, G. M.; Richter, F.; Khare, S.; Dillon, M. B. D.; Bachovchin, D. A.; Mowen, K.; Baker, D.; Cravatt, B. F. *Nature* **2010**, *468*, 790.
(41) Retailleau, P.; Huang, X.; Yin, Y.; Hu, M.; Weinreb, V.; Vachette, P.; Vonrhein, C.; Bricogne, G.; Roversi, P.; Ilyin, V.; Carter, C. W., Jr. *J. Mol. Biol.* **2003**, *325*, 39−63.
(42) Bryan, P.; Pantoliano, M. W.; Quill, S. G.; Hsiao, H. Y.; Poulos, T. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 3743−3745.
(43) Hinkle, P. M.; Kirsch, J. F. *Biochemistry* **1971**, *10*, 2717.
(44) Brocklehurst, K.; Malthouse, J. P. *Biochem. J.* **1981**, *193*, 819−823.
(45) Patai, S., Ed. *The Chemistry of the Thiol Group*; John Wiley & Sons, Ltd: New York, 1974; Chapter 8
(46) Craik, C. S.; Roczniak, S.; Largman, C.; Rutter, W. J. *Science* **1987**, *237*, 909−913.
(47) Carter, P.; Wells, J. A. *Nature* **1988**, *332*, 564−568.
(48) Bostrom, J.; Greenwood, J. R.; Gottfries, J. *J. Mol. Graph. Model.* **2003**, *21*, 449−462.
(49) Sheffler, W.; Baker, D. *Protein Sci.* **2009**, *18*, 229−239.
(50) Vigers, G. P. A.; Rizzi, J. P. *J. Med. Chem.* **2004**, *47*, 80−89.
(51) Davis, I. W.; Baker, D. *J. Mol. Biol.* **2009**, *385*, 381−392.
(52) Charbonnier, J. B.; Carpenter, E.; Gigant, B.; Golinelli-Pimpaneau, B.; Eshhar, Z.; Green, B. S.; Knossow, M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11721−11725.
(53) Lesley, S. A.; Patten, P. A.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 1160−1165.
(54) Guo, J.; Huang, W.; Scanlan, T. S. *J. Am. Chem. Soc.* **1994**, *116*, 6062−6069.
(55) Stewart, J. D.; Roberts, V. A.; Thomas, N. R.; Getzoff, E. D.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 1994−2003.