Article

New Tricks for Old Proteins: Single Mutations in a Non-enzymatic Protein Give Rise to Various Enzymatic Activities

Yurii S. Moroz, Tiffany T. Dunston, Olga V. Makhlynets, Olesia V. Moroz, Yibing Wu, Jennifer H. Yoon, Alissa B. Olsen, Jaclyn M. McLaughlin, Korrie L. Mack, Pallavi M. Gosavi, Nico A.J. Van Nuland, and Ivan V Korendovych

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.5b07812 • Publication Date (Web): 10 Nov 2015

New Tricks for Old Proteins: Single Mutations in a Non-enzymatic Protein Give Rise to Various Enzymatic Activities

Yurii S. Moroz,[1,‡,†] Tiffany T. Dunston,[1,‡] Olga V. Makhlynets,[1,‡] Olesia V. Moroz,[1,‡,†] Yibing Wu,[2] Jennifer H. Yoon,[1] Alissa B. Olsen,[1] Jaclyn M. McLaughlin,[1] Korrie L. Mack,[1] Pallavi M. Gosavi,[1] Nico A. J. van Nuland[3] and Ivan V. Korendovych[1,*]

[1] Department of Chemistry, Syracuse University, 111 College Place, Syracuse NY, 13244

[2] Department of Pharmaceutical Chemistry University of California – San Francisco 555 Mission Bay Blvd. South, San Francisco, CA 94158, USA

[3] Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

ABSTRACT: Design of a new catalytic function in proteins, apart from its inherent practical value, is important for fundamental understanding of enzymatic activity. Using a computationally inexpensive, minimalistic approach that focuses on introducing a single highly reactive residue into proteins to achieve catalysis we converted a 74-residue-long C-terminal domain of calmodulin into an efficient esterase. The catalytic efficiency of the resulting stereoselective, allosterically regulated catalyst, nicknamed AlleyCatE, is higher than that of any previously reported *de novo* designed esterases. The simplicity of our design protocol should complement and expand the capabilities of current state-of-art approaches to protein design. These results show that even a small non-enzymatic protein can efficiently attain catalytic activities in various reactions (Kemp elimination, ester hydrolysis, retroaldol reaction) as a result of a single mutation. In other words, proteins can be just one mutation away from becoming entry points for subsequent evolution.

Introduction. Nature has developed the means to facilitate an enormous variety of chemical transformations with exceptional regio- and stereo-selectivity starting with simple reactants in water. Moreover, the repertoire of enzymes for new reactions continues to grow. Just during past decades new proteins emerged to degrade synthetic chemicals introduced into the environment.[1-6] Much effort has been dedicated to understand how proteins evolve to adapt new functions and the consensus emerged that new function originates from latent, promiscuous enzymes or broad specificity enzymes that serve as starting points for further evolution.[2,7-9] Indeed, redesigning existing proteins to perform new reactions starting with chemically similar substrates has been very fruitful.[10-13] Over the past years principles and methods of directed evolution have been perfected to allow for substantial improvement of very modest initial activity by screening relatively small focused combinatorial libraries.[14,15] While optimization of existing enzymes has been successful, it does require measurable initial activity and is often limited by substrate scope, enzyme size, preparation issues, etc. Achieving reasonable initial activity is especially difficult for reactions with no obvious natural analogs. Computational design lifts this fundamental limitation by providing such starting points.[16,17] Combination of sophisticated computational algorithms with directed evolution has been shown to successfully create catalysts for a wide range of chemical transformations.[18-22] Importantly, initial configuration of the computationally introduced active site can be far from optimal; Hilvert *et al.* showed that directed evolution may even completely reshape the catalytic site during the optimization of a computationally designed retroaldolase.[23]

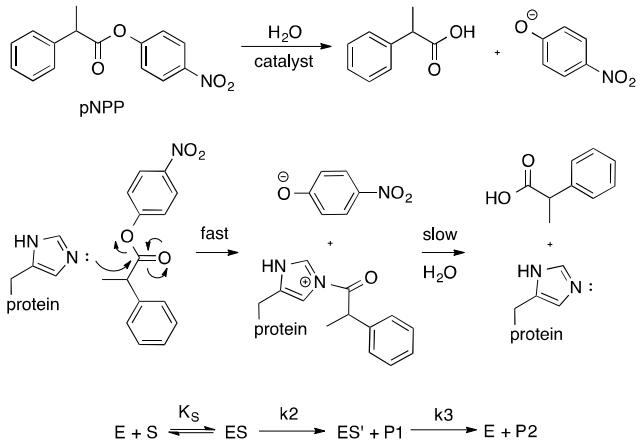
Current methods for computational enzyme design rely on precise positioning of multiple functional groups deemed to be necessary for catalysis into the rigid protein scaffold to minimize a scoring function that predicts thermodynamic stability of a protein. Despite successes of this approach to create good starting points for subsequent directed evolution, computationally designed enzymes still have fairly modest activity.[24,25] Moreover, simultaneous introduction of multiple (often more than 10) mutations is often so detrimental to the protein stability that additional rounds of mutagenesis may be required to ensure that the designed catalysts are properly folded.[26]

While we are still far away from being able to create efficient catalysts *de novo*, lessons from successes and failures of current state of the art methods of enzyme design[27] present us with an opportunity. Indeed, if directed evolution is so powerful, can we focus on introducing a single highly reactive functional group into a protein without extensive engineering of the protein scaffold to achieve function? Such a minimalist strategy offers high-throughput and low computational cost.

Here we took on a task of *de novo* designing a catalyst for ester hydrolysis, a reaction that finds many practical applications ranging from fine chemical synthesis to biofuel production. We chose *p*-nitrophenyl esters as substrates due to the abundance of data for catalysts developed by various approaches including computational design,[28,29] combinatorial screening[30-35] and catalytic antibody selection.[36-40] While low

rates of *p*-nitrophenyl ester hydrolysis are relatively easy to achieve,[41] it is profoundly humbling that the best artificial esterase that doesn't show significant product inhibition is only modestly active with a k_{cat} value of just 3.3 min^{-1} (Table S1, Supporting Information).[36]

Scheme 1. Overall reaction (top) and the mechanism of histidine catalyzed ester hydrolysis.



It is well known that common folds are quite frequent in proteins catalyzing a variety of different reactions (e.g. TIM barrel) and successful computationally designed catalysts take advantage of existing enzymatic folds to achieve new function. But does that mean that non-enzymatic proteins lack such ability? Calmodulin (CaM), an allosterically regulated protein, uses metal binding to achieve high fold stability. Stability has been long postulated to play a critical role for designability,[42] thus we set out to explore whether calmodulin could assume esterase activity. We chose chromogenic p-nitrophenyl-(2-phenyl)-propanoate (pNPP, Scheme 1) as a representative ester substrate to compare our studies with previous design work.[29]

Naturally occurring esterases often rely on multiple amino acid residues (arranged in dyads and triads) aided by transition state stabilization provided by oxyanion holes to achieve efficient catalysis. However, it has been demonstrated that a histidine residue in a protein scaffold may catalyze ester hydrolysis, albeit with modest activity.[28] Therefore we decided to engineer a single histidine residue into the calmodulin scaffold to hydrolyze pNPP.

Results and Discussion. The design process is schematically outlined in Fig. 1. First, we ensured that the hydrophobic cavity of the C-terminal domain of CaM (c-CaM) could accommodate the substrate. Docking both enantiomers of pNPP into c-CaM using AutoDock Vina[43] shows that the substrate fits well in the hydrophobic cavity of the protein (Fig. 1, Step 1). Next, we identified all the residues facing the substrate (shown in orange in Fig. 1, Step 2) and computationally evaluated the effect of mutating these residues to histidines using Rosetta. The protein backbone was fixed and the obtained models were sorted by score (Fig. 1, Step 3). Most of the mutants were able to accommodate histidine's side chain without clashes; one of the mutants (A128H) that showed relatively high Rosetta score was removed from further consideration. The second step of optimization was to determine the possibility of Michaelis complex formation. Both enantiomers of pNPP were docked into the computational models obtained in the previous step. The poses found from the docking studies were then evaluated to establish the feasibility of transition state geometry as described in detail in the Experimental Section. Only a single mutant (c-CaM M144H) satisfied all parameters described above, importantly, the R enantiomer of pNPP provided a better fit in the hydrophobic cavity of c-CaM (Fig. 1, Step 4). By analogy with our previous work[44,45] we named c-CaM M144H AlleyCatE, where E stands for "esterase".

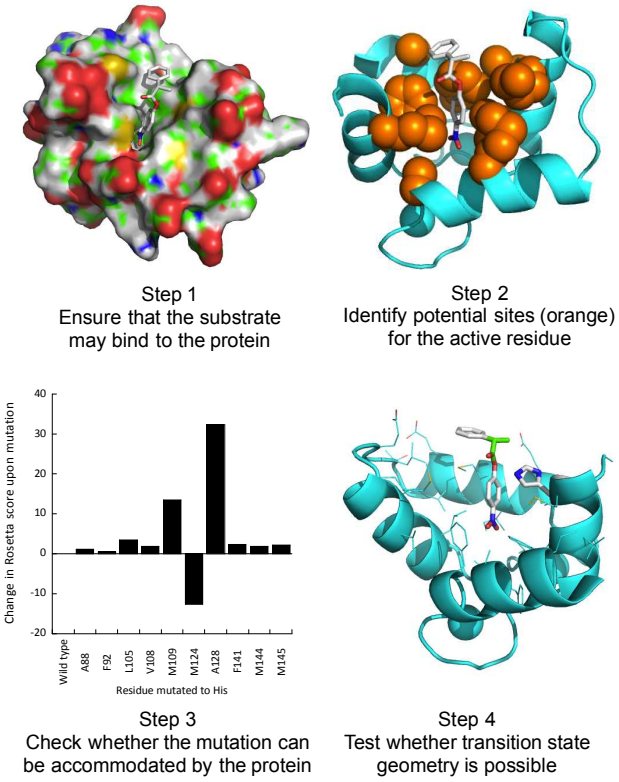


Figure 1. Overview of the design methodology.

AlleyCatE was then recombinantly expressed in *Escherichia coli* (*E. coli*) and purified by affinity chromatography. AlleyCatE shows the classic "burst phase" dependence of product formation on time (Fig. 2), consistent with the acyl intermediate formation described in Scheme 1. The acyl intermediate was directly observed by MALDI-TOF mass spectrometry (Figs. S1-S4, Supporting Information). The MALDI spectrum for unmutated c-CaM shows no acylation in the presence of pNPP (Fig. S5, Supporting Information) confirming that AlleyCatE is acylated at the newly introduced H144. However, it should be noted that after several turnovers MALDI-TOF spectra of AlleyCatE in the presence of substrate show multiple acylation events (up to 4 total) suggesting that the initially formed acyl-H144 intermediate can acylate AlleyCatE (Fig. S6, Supporting Information). AlleyCatE showed significant esterase activity with a k_2/K_s of 70,000 ± 17,000 $M^{-1}min^{-1}$ and an apparent k_{cat}/K_M of 4,800 ± 300 $M^{-1}min^{-1}$ for hydrolysis of R-pNPP (Fig. 2, Table 1). AlleyCatE undergoes at least 12 turnovers and shows no signs of inhibition by p-nitrophenol or 2-phenylpropionic acid (Figs. S7, S8, Supporting Information). In agreement with the design, AlleyCatE is enantioselective: R-pNPP is hydrolyzed ~3-fold faster than its enantiomer (Table 1). This level of activity, achieved with a single mutation in a non-enzymatic scaffold, is to our knowledge higher than that of any previously reported *de novo* designed

esterases (Table S1, Supporting Information). The observed k_{cat} values (up to 0.24 min^{-1}) for AlleyCatE are comparable to the values obtained for the less active catalytic antibodies (0.33-0.4 min^{-1})[38,39] for this reaction and are two orders of magnitude lower than the k_{cat} values shown by best catalytic antibodies for *p*-nitrophenyl ester hydrolysis under the same conditions.[36,37,46,47] While the observed degree of enantioselectivity is low when compared to promiscuous esterase activity found in natural enzymes,[48] it does represent a starting point for subsequent improvement by directed evolution.

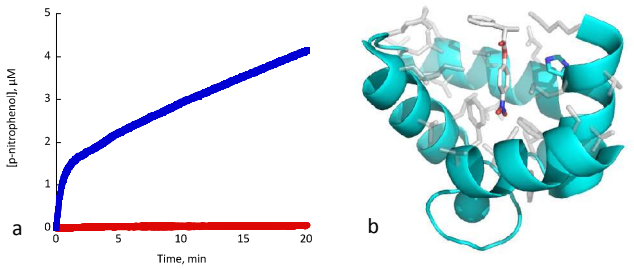


Figure 2. a) Hydrolysis of R-pNPP (50 μM) catalyzed by AlleyCatE (2 μM) in the presence (blue) and in the absence (red) of 10 mM $CaCl_2$. b) NMR structure of the AlleyCatE. The substrate was docked into the structure using AutoDock Vina. The side chain for the active residue (M144H) is shown in color. The side chains forming the hydrophobic cavity are shown in grey, the rest of the side chains are omitted for clarity. PDB code 4bya.

Proper positioning of the histidine residue in the cavity is critical, introducing histidine in any of the other eight positions facing the hydrophobic pocket resulted in no significant esterase activity.

One of the key features of calmodulin is its allosteric regulation. Binding of calcium ions to the apo c-CaM results in a drastic conformational change that exposes protein's hydrophobic region to solvent. AlleyCatE is active only in the presence of Ca^{2+} confirming that the allosteric regulation is still preserved in this mutant (Fig. 2a). AlleyCatE's activity drops by at least 100-fold in the absence of Ca^{2+}.

Introducing a histidine residue into the hydrophobic region of CaM resulted in a small loss of the free energy of folding, well tolerated by the highly stable calmodulin fold as shown by chemical denaturation studies (Fig. S9, Table S2, Supporting Information). The NMR structure of AlleyCatE in the calcium-bound state is very similar (backbone rmsd 0.78 Å) to that of designed model; importantly, the observed rotameric state of the active histidine residue is the same as predicted (Fig. 2b).

The activity of AlleyCatE to hydrolyze pNPP is well above the background of natural esterases in *E. coli* crude lysate (Fig. S10, Supporting Information) making directed evolution to further improve catalytic efficiency simple and practical. Reetz *et al.* showed that saturation mutagenesis of residues around the active site offers the highest probability of success,[49] therefore we employed this approach to show the feasibility of AlleyCatE's further improvement. Saturation mutagenesis performed on nine residues around the active residue produced several proteins with higher activity. The best variant (c-CaM M144H A128T, named AlleyCatE2) was characterized in detail. The k_2/K_s value of AlleyCatE2 is 2-fold higher and the overall apparent k_{cat}/K_M is nearly 40% higher than those of AlleyCatE (Table 1). The allosteric regulation of the activity is fully preserved in AllyCatE2 (Fig. S11, Supporting Information). These results obtained from a very limited library of mutants show that AlleyCatE is evolvable and additional rounds of directed evolution may improve the reactivity even further. Previously we have shown that catalytic efficiency of AlleyCat, a calmodulin-based Kemp eliminase, can be easily improved by 200-fold.[50]

Calmodulin provides a relatively featureless hydrophobic cavity for catalyst design, thus we have probed the degree to which such protein fold can provide substrate selectivity. We employed *p*-nitrophenyl acetate (pNPA), a common substrate for measuring esterase activity of various enzymes, as our reference point.

Table 1. Kinetic parameters for pNPP AlleyCatE and AlleyCatE2.

Protein	Substrate	k_2/K_S, $M^{-1}\,min^{-1}$	$k_2 \times 10^{-2}$, min^{-1}	K_S, μM	k_{cat}/K_M, $M^{-1}\,min^{-1}$	$k_{cat} \times 10^{-2}$, min^{-1}	K_M (app), μM
AlleyCatE	R-pNPP	70,000±17,000	140±45	22±6	4,800 ± 300	8.8±0.7	18±1
	S-pNPP	15,000±200	140±40	86±26	1,700 ± 200	3.7±0.4	21±3
	pNPA	1,950±170	75±13	400±70	330 ± 45	24±10	700 ±300
AlleyCatE2	R-pNPP	120,000±27,000	160±30	13±3	6,600 ± 600	28 ± 1	42 ±4
	S-pNPP	26,400±1,000	220±30	83 ± 11	2,400 ± 100	6.5±0.3	27 ± 2
	pNPA	4,600±500	90±9	200±24	360 ± 30	7.0±0.5	200 ±15

The enzymatic efficiencies of pNPA hydrolysis catalyzed by c-CaM mutants are substantially lower (more than 35-fold) that those for pNPP despite similar k_{uncat} values for the two substrates (Table 1). This change in the k_{cat}/K_M value originates mostly from the increase in K_M, suggesting that even the simplest of design principles can provide a certain degree of substrate recognition.

Analysis of the properties of artificial esterases designed by various approaches (Table S1, Supporting Information) is quite instructive. It is clear that by properly arranging a histidine nucleophile in a protein or a peptide assembly using any of the described approaches it is possible to achieve rapid formation of the acyl intermediate. In the case of peptide assemblies the subsequent hydrolysis of the acyl intermediate is fast thus k_2 is the apparent k_{cat}. The major advantage of computa-

tional designs is their ability to utilize existing protein scaffolds for substrate recognition and/or allosteric regulation.[51,52] At the same time the increased hydrophobicity of the substrate binding pocket slows down the subsequent hydrolysis and affects the overall turnover.

Conclusion. In this paper we demonstrate how easily a minimalistic approach to protein design can be applied to different types of chemical reactions: introducing a *single* histidine residue into calmodulin confers esterase activity onto this non-enzymatic protein. The obtained kinetically enantioselective, allosterically controlled esterase is active enough to compete against the background of natural enzymes in the *E. coli* crude lysate, which allowed for further activity improvement using directed evolution. We show that enzyme promiscuity, considered to be a major driving force in evolution of new functions, has profoundly wider meaning. Even a small 74-residue non-enzymatic protein can efficiently attain catalytic activities in various reactions (Kemp elimination,[44] ester hydrolysis, retro-aldol reaction[53]) as a result of just a single mutation. In other words, proteins can be just one mutation away from becoming entry points for subsequent evolution. It remains to be established whether other proteins can be as easily modified to catalyze novel chemical transformations, but there are no reasons to *a priori* assume that this *latent* enzymatic promiscuity is limited to calmodulin.

In addition to providing insight into understanding how proteins evolve to adopt new functionality this work has practical implications for protein design. It is reassuring that the minimalist approach can accurately predict productive mutations. The simplicity of the docking procedure allows for very quick screening of potential catalyst candidates and to identify opportunities that may be discarded by more stringent requirements for the transition state geometry. The advent of tools to shape proteins for binding new substrates[54] will enable this approach to be applied to more complex reactions and protein scaffolds. As various iterative methodologies of design and directed evolution have been highly successful,[55,56] subsequent stepwise improvement of the catalyst candidates identified by minimalist approach is likely to produce efficient catalysts using simple tools at low computational cost following a path that is similar to the one observed in nature.

Experimental section.

General. Reagents and buffers were purchased from BioBasic, Inc. and Santa Cruz Biotechnology, Inc. DNA oligonucleotides were purchased from Integrated DNA Technologies. Enzymes for cloning were purchased from Promega and New England Biolabs. Racemic 4-nitrophenyl 2-phenylpropanoate was prepared from 2-phenyl-propionic acid and 4-nitrophenol using previously reported procedure.[29] The enantiomerically pure (R)- and (S)-4-nitrophenyl-(2-phenyl)-propanoates were prepared from the corresponding commercially available enantiomers of 2-phenyl-propionic acid using the same protocol.

Computational studies. A high resolution crystal structure of calmodulin (PDB entry 1CLL)[57] served as a starting point of the design. The side chain placement was computationally optimized with fixed backbone (fixbb protocol) using Rosetta.[58] Mutations in the resulting model were computationally introduced by changing the amino acid identity in positions A88, F92, L105, V108, M109, M124, A128, F141, M144, M145 to a histidine and side chain placement was again computationally optimized with fixed backbone (fixbb protocol) using Rosetta. The Rosetta score for the mutants were compared to that of the original structure and reported in Fig. 1. The models of R and S isomers of 4-nitrophenyl 2-phenylpropanoate were created with MarvinSketch (ChemAxon) and optimized with UCSF Chimera[59] and ADT 1.5.6.[60] These models were then non-covalently docked into the computationally derived models of CaM mutants using Autodock Vina software.[61] At least 20 docking for each mutant were generated. Then the distances between the δ and ε atoms (both nitrogen and carbon atoms were considered to account for a possibility of a 180 degree rotation around the Cβ-Cγ bond) in the histidine ring and the carbonyl carbon were computed. The poses for the substrates were considered capable of supporting catalysis only if a) the distance between carbonyl carbon of the ester and the nitrogen atom of imidazole ring (either Nε or Nδ) was less than 3.5 Å AND b) the $O{=}C_{carbonyl}\text{-}N_{imidazole}$ angle was between 100 and 140 degrees. The sequences for all proteins discussed in this paper are given in Table S3 (Supporting Information).

Cloning and mutagenesis. All genes were cloned into pMCSG49 expression vector using ligase independent cloning (LIC).[62] This vector contains an N-terminal His_6-tag and a tobacco etch virus (TEV) protease recognition site (ENLYFQ/S). Site directed mutagenesis was performed using Pfu Turbo polymerase (Promega) using manufacturer's standard protocols. The plasmids encoding AlleyCatE and AlleyCatE2 have been deposited in AddGene.

Protein expression and purification. The vectors containing the genes of interest were transformed into *E. coli* BL21(DE3) pLysS cells and expressed overnight at 37 °C using ZYM-5052 auto-induction medium supplemented with ampicillin (100 ug/mL) and chloramphenicol (34 mg/L).[63] The cells were lysed in a buffer containing 25 mM Tris (pH 8.0), 20 mM imidazole, 10 mM $CaCl_2$, and 300 mM NaCl on ice with protease inhibitor (phenylmethylsulfonyl fluoride) added. The lysate was then centrifuged at 20,000 g for 1 hour at 4 °C to remove the cell debris and the supernatant was applied onto a Ni-NTA column (Clontech). The column was washed multiple times with the lysis buffer and the protein was eluted with buffer containing 25 mM TRIS, 300 mM imidazole, 10 mM $CaCl_2$, and 300 mM NaCl (pH 8.0). To remove a polyhistidine tag, the proteins were subjected to subsequent TEV digestion. The protein buffer was changed to TEV digestion buffer (50 mM TRIS pH 8.0, 75 mM NaCl) using a BioRad 10 DG desalting column and then dithiothreitol (DTT) and ethylenediaminetetraacetic acid (EDTA) were added to a final concentration 1 mM and 0.5 mM respectively. TEV protease was added to the protein, using 1:10 ratio of OD_{280}(TEV) to OD_{280}(protein). The solution was mixed, sterilized using a 0.22 μm filter, and incubated overnight at 34°C. Following the digest, DTT and EDTA were removed with a desalting column. The obtained protein solution in 20 mM HEPES, 10 mM $CaCl_2$, 100 mM NaCl (pH 7.5) was passed through a Ni-NTA column to remove the undigested protein and the cleaved His-tag. The SDS-PAGE gel of the pure proteins is presented in Fig. S12 (Supporting Information).

Expression and purification of isotopically labeled proteins. U-^{15}N,^{13}C-labeled AlleyCatE was expressed with double antibiotic resistance following the previously described procedure.[64] The plasmid encoding the protein sequence with a N-terminal His_6-tag followed by the TEV protease recognition site (ENLYFQ/S) was transformed into BL21(DE3) pLysS

cells. The next day, a single colony was inoculated into a 50 mL Luria Bertani (LB) culture and was grown overnight at 30 °C. The next morning the overnight culture was added to 1 L of 2x LB and incubated at 37 °C and 230 rpm until OD_{600} reached 0.6 – 0.7. The cell were harvested by centrifugation for 8 minutes at 1,000 g and 30 °C, washed first with unlabeled M9 minimal medium then with 25 mL of M9 minimal medium that contains $^{15}NH_4Cl$ and uniformly ^{13}C-labeled glucose. The cell pellets were resuspended in 300 mL of M9 minimal medium solution containing $^{15}NH_4Cl$ as a nitrogen source, ^{13}C-labeled glucose as a carbon source and incubated at 37 °C for 2.0 hours to allow for the recovery of growth. Protein expression was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. After induction for 2 hours the cells were harvested by centrifugation and proteins were purified using Ni-NTA affinity chromatography as described above.

Crude cell lysate screening and directed evolution. To test the activity of various proteins to catalyze ester hydrolysis in crude cell lysates, the genes for c-CaM, AlleyCatE, CaM and CaM M144H (full length protein) mutant were transformed into *E. coli* BL21 (DE3) pLysS cells. Individual colonies were cultured into 200 µL of LB containing 100 µg/mL of ampicillin in 96-well plates and grown for 6-7 hours at 37 °C. 10 µL of the resulting culture was used to inoculate ZYM-5052 autoinduction media containing 100 µg/mL of ampicillin and grown overnight at 37 °C. The following day, cells were harvested and lysed with cell lysis buffer containing 20 mM HEPES, 10 mM $CaCl_2$, 100 mM NaCl, 0.2% Triton X-100, pH 8.0. Hydrolysis of 25 µM 4-nitrophenyl 2-phenylpropanoate (pNPP) by the crude cell lysates was monitored spectrophotometrically at 405 nm. For directed evolution studies, the screening was done with the full length CaM to improve expression levels. The presence of the N-terminal domain does not affect catalytic activity. Saturation mutagenesis to generate all possible single mutants in positions D80, L105, H107, V108, L112, A128, A88, F89, and V91 was done by following a standard mutagenesis protocol with *Pfu Turbo* DNA polymerase (Stratagene) using an NNK codon and AlleyCatE as a template. The resulting library of about 900 mutants was screened for hydrolytic activity and the mutant that showed the best steady-state activity in the crude lysate (AlleyCatE2) was chosen for further characterization.

UV-vis spectroscopy kinetic assays. Kinetic assays were performed in a 1 cm path length quartz cuvette on the Agilent Cary 60 UV-Vis spectrophotometer monitoring absorbance of the product at 405 nm. The extinction coefficient for p-nitrophenol (12,700 $M^{-1}cm^{-1}$) under these conditions was taken from the literature data.[65] Proteins were diluted to a final concentration of 2.0 µM or 6.0 µM for hydrolysis of pNPP (R and S isomers) and pNPA respectively. Substrate stock solutions of pNPP (1 mM) and pNPA (100 mM) were prepared in acetonitrile. The reaction was monitored in triplicate for 900 seconds (0.100 s averaging time) in 200 µL of buffer containing 20 mM HEPES (pH 7.5), 10 mM $CaCl_2$, 100 mM NaCl. 150 µL of substrate solution in buffer was added first to the cuvette and subsequently 50 µL of the protein solution was added. The final content of acetonitrile was 1.5% in reaction mixtures containing pNPA and 6% in the reaction mixtures containing pNPP. The obtained kinetic traces were analyzed using the method of Kezdy and Bender.[66] Briefly, the kinetic traces were fit to

$$[p-nitrophenolate] = At + B(1 - e^{-bt}) \quad (1)$$

where t is time. Plotting reciprocal values of A and b vs. $1/[S]_0$ at various initial substrate concentrations (Figs. S13-S18, Supporting Information) allows for straightforward determination of k_{cat}, $K_{M(app)}$, k_2 and K_s.[66]

CD spectroscopy. Circular dichroism experiments were performed on a JASCO J-715 CD spectrometer in a step scan mode (4 s averaging time) averaging over three runs using a quartz cuvette with a 1 mm path length. Samples were prepared in a buffer containing 4 mM HEPES (pH 7.5), 2 mM $CaCl_2$ and 30 mM NaCl with final protein concentration maintained at 25.0 µM. The sample absorbance never exceeded 1.5 at all wavelengths in order to obtain reliable MRE values. The ellipticity of proteins was monitored at 222 nm for chemical denaturation experiments in the presence of varying concentrations of guanidinium hydrochloride (0 – 6 M). The data collected for c-CaM and AlleyCatE were fitted to the equation below.

$$MRE = \frac{MRE_f + (MRE_u + y_u[D])e^{-\frac{\Delta G - m[D]}{RT}}}{1 + e^{-\frac{\Delta G - m[D]}{RT}}} \quad (2)$$

where *MRE* is the observed mean residue ellipticity; MRE_f, MRE_u are mean residue ellipticities representing the folded and unfolded states for the C-terminal domain, respectively; ΔG is the free energy of unfolding; [D] is the concentration of the denaturant and y_u is the slope for the unfolded state. CD spectrum of AlleyCatE is shown in Fig. S19 (Supporting Information). The thermodynamic parameters yielded by the fit of chemical denaturation curve are summarized in Table S2 (Supporting Information).

NMR Spectroscopy and Structure Calculations. The NMR sample of U-^{15}N,^{13}C -labeled protein were prepared at a concentration 1.1 mM in a buffer containing 5 % D_2O, 0.05 % DSS, 0.01 % NaN_3. All experiments were conducted at 30°C on a Varian 600 MHz spectrometer equipped with cryogenic probe. Chemical shift assignments were made using ^{15}N- and ^{13}C-HSQC, HNCO, HNCACB, CBCA(CO)HN, HBHA(CO)HN, 3D (H)CCH-TOCSY. The assignments were further confirmed by 3D ^{15}N- and ^{13}C-edited NOESY spectra with the mixing time of 100 ms. NMR data were processed, analyzed and visualized using NMRPipe[67] and CARA. ^1H chemical shifts were referenced to internal DSS and ^{13}C and ^{15}N chemical shifts were referenced indirectly via gyromagnetic ratios. Backbone dihedral angles constraints were derived from chemical shifts using TALOS+.[68] ^1H-^1H upper distance constraints for structure calculations were obtained from the analysis of the 3D ^{13}C- and ^{15}N-edited NOESY spectra. Automated NOE assignment and structure calculation was done with CYANA 3.0[69] together with dihedral angle constraints and metal ions constraints, which were taken from the crystal structure of CaM (PDB code 1CLL).[57] The resulting assignments were verified and corrected by interactive spectral analysis. Subsequently the structure calculation was performed iteratively with CYANA, with iterations used to verify and complete resonance assignments, refine NOESY peak lists and optimize the distance calibration constants. The first set of refined structures was obtained using XPLOR-NIH.[70] The final set of structures was further refined by restrained molecular dynamics in explicit water[71] using the program CNS 1.2[72] with the OPLSX force field. Structural statistics and

global quality factors were computed with PSVS 1.5.[73-77] Chemical shifts, NOESY peak lists, and raw time-domain data for AlleyCatE were deposited in the BioMagResBank (BMRB ID 19376); the coordinates were deposited in the Protein Data Bank (PDB code 4BYA). The structure statistics are summarized in Table S4 (Supporting Information).

Mass Spectrometry. MALDI-TOF spectrometry was done on a Bruker Autoflex III instrument. Prior to each experiment the instrument was calibrated with standard samples provided by the manufacturer. Dihydroxybenzoic acid (DHB) or sinapinic acid was used as a matrix. For acylated samples the substrate was added to the proteins and the samples were immediately dried and analyzed.

ASSOCIATED CONTENT

Supporting Information. Details of NMR structural characterization. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

* ikorendo@syr.edu

Present Addresses

†Department of Chemistry, Taras Shevchenko National University of Kiev, 64 Volodymyrska St., Kiev, 01601, Ukraine.

Author Contributions

The manuscript was written through contributions of all authors. ‡These authors contributed equally.

Funding Sources

No competing financial interests have been declared. This work was supported in part by an NSF IGERT fellowship to T.T.D., grant number 1332349 from NSF-EFRI, ORAU Ralph E. Powe Junior Faculty Enhancement award and a Humboldt Fellowship to I.V.K.

ACKNOWLEDGMENT

This paper is dedicated to the memory of Prof. Anatoly I. Brusilovets. We thank Dr. Alexander N. Volkov for insightful discussions and Prof. Robert P. Doyle for providing access to CD instrument.

ABBREVIATIONS

CaM, Calmodulin; pNPA, *p*-nitrophenyl acetate; pNPP, 4-nitrophenyl-(2-phenyl)-propanoate.

REFERENCES

(1) Hartley, C. J.; Newcomb, R. D.; Russell, R. J.; Yong, C. G.; Stevens, J. R.; Yeates, D. K.; La Salle, J.; Oakeshott, J. G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8757.

(2) Khersonsky, O.; Tawfik, D. S. *Annu. Rev. Biochem.* **2010**, *79*, 471.

(3) Janssen, D. B.; Dinkla, I. J.; Poelarends, G. J.; Terpstra, P. *Environ. Microbiol.* **2005**, *7*, 1868.

(4) Raushel, F. M.; Holden, H. M. *Adv. Enzymol. Relat. Areas Mol. Biol.* **2000**, *74*, 51.

(5) Hall, B. G. *Nat. Rev. Microbiol.* **2004**, *2*, 430.

(6) Weinreich, D. M.; Delaney, N. F.; Depristo, M. A.; Hartl, D. L. *Science* **2006**, *312*, 111.

(7) Jensen, R. A. *Annu. Rev. Microbiol.* **1974**, *30*, 409.

(8) Glasner, M. E.; Gerlt, J. A.; Babbitt, P. C. *Curr. Opin. Chem. Biol.* **2006**, *10*, 492.

(9) O'Brien, P. J.; Herschlag, D. *Chem. Biol.* **1999**, *6*, R91.

(10) Brundiek, H. B.; Evitt, A. S.; Kourist, R.; Bornscheuer, U. T. *Angew. Chem. Int. Ed.* **2012**, *51*, 412.

(11) Fernández-Álvaro, E.; Snajdrova, R.; Jochens, H.; Davids, T.; Böttcher, D.; Bornscheuer, U. T. *Angew. Chem. Int. Ed.* **2011**, *50*, 8584.

(12) Toscano, M. D.; Woycechowsky, K. J.; Hilvert, D. *Angew. Chem. Int. Ed.* **2007**, *46*, 3212.

(13) Lutz, S.; Patrick, W. M. *Curr. Opin. Biotech.* **2004**, *15*, 291.

(14) Jochens, H.; Bornscheuer, U. T. *ChemBioChem* **2010**, *11*, 1861.

(15) Gumulya, Y.; Sanchis, J.; Reetz, M. T. *ChemBioChem* **2012**, *13*, 1060.

(16) Brustad, E. M.; Arnold, F. H. *Curr. Opin. Chem. Biol* **2011**, *15*, 201.

(17) Lutz, S. *Curr. Opin. Biotech.* **2010**, *21*, 734.

(18) Bornscheuer, U. T.; Huisman, G.; Kazlauskas, R. J.; Lutz, S.; Moore, J.; Robins, K. *Nature* **2012**, *485*, 185.

(19) Höhne, M.; Schätzle, S.; Jochens, H.; Robins, K.; Bornscheuer, U. T. *Nat. Chem. Biol.* **2010**, *6*, 807.

(20) Rothlisberger, 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.

(21) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387.

(22) Siegel, J. B.; Zanghellini, A.; Lovick, H. M.; Kiss, G.; Lambert, A. R.; St. Clair, J. L.; Gallaher, J.; Hilvert, D.; Gelb, M. H.; Stoddard, B. L.; Houk, K. N.; Michael, F. E.; Baker, D. *Science* **2010**, *329*, 309.

(23) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. *Nat. Chem. Biol.* **2013**, *9*, 494.

(24) Wijma, H. J.; Janssen, D. B. *FEBS J.* **2013**, *280*, 2948.

(25) Barrozo, A.; Borstnar, R.; Marloie, G.; Kamerlin, S. C. L. *Int. J. Mol. Sci.* **2012**, *13*, 12428.

(26) Khersonsky, O.; Kiss, G.; Rothlisberger, D.; Dym, O.; Albeck, S.; Houk, K. N.; Baker, D.; Tawfik, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 10358.

(27) Korendovych, I. V.; DeGrado, W. F. *Curr. Opin. Struct. Biol.* **2014**, *27*, 113.

(28) Bolon, D. N.; Mayo, S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 14272.

(29) Richter, F.; Blomberg, R.; Khare, S. D.; Kiss, G.; Kuzin, A. P.; Smith, A. J. T.; Gallaher, J.; Pianowski, Z.; Helgeson, R. C.; Grjasnow, A.; Xiao, R.; Seetharaman, J.; Su, M.; Vorobiev, S.; Lew, S.; Forouhar, F.; Kornhaber, G. J.; Hunt, J. F.; Montelione, G. T.; Tong, L.; Houk, K. N.; Hilvert, D.; Baker, D. *J. Am. Chem. Soc.* **2012**, *134*, 16197.

(30) Broo, K. S.; Brive, L.; Ahlberg, P.; Baltzer, L. *J. Am. Chem. Soc.* **1997**, *119*, 11362.

(31) Broo, K. S.; Nilsson, H.; Nilsson, J.; Flodberg, A.; Baltzer, L. *J. Am. Chem. Soc.* **1998**, *120*, 4063.

(32) Wei, Y.; Hecht, M. H. *Protein Eng. Des. Sel.* **2004**, *17*, 67.

(33) Yamauchi, A.; Nakashima, T.; Tokuriki, N.; Hosogawa, M.; Nogami, H.; Arioka, S.; Urabe, I.; Yomo, T. *Protein Eng.* **2002**, *15*, 619.

(34) Patel, S. C.; Bradley, L. H.; Jinadasa, S. P.; Hecht, M. H. *Protein Sci.* **2009**, *18*, 1388.

(35) Maeda, Y.; Javid, N.; Duncan, K.; Birchall, L.; Gibson, K. F.; Cannon, D.; Kanetsuki, Y.; Knapp, C.; Tuttle, T.; Ulijn, R. V.; Matsui, H. *J. Am. Chem. Soc.* **2014**, *136*, 15893.

(36) Tawfik, D. S.; Zemel, R. R.; Arad-Yellin, R.; Green, B. S.; Eshhar, Z. *Biochemistry* **1990**, *29*, 9916.

(37) Tramontano, A.; Janda, K. D.; Lerner, R. A. *Science* **1986**, *234*, 1566.

(38) Pollack, S. J.; Jacobs, J. W.; Schultz, P. G. *Science* **1986**, *234*, 1570.

(39) Wirsching, P.; Ashley, J. A.; Benkovic, S. J.; Janda, K. D.; Lerner, R. A. *Science* **1991**, *252*, 680.

(40) Tanaka, F. *Chem. Rev.* **2002**, *102*, 4885.
(41) Menger, F. M.; Ladika, M. *J. Am. Chem. Soc.* **1987**, *109*, 3145.
(42) Bloom, J. D.; Labthavikul, S. T.; Otey, C. R.; Arnold, F. H. *Proc. Natl. Acad. Sci.* **2006**, *103*, 5869.
(43) Morris, G. M.; Huey, R.; Olson, A. J. *Curr. Protoc. Bioinformatics* **2008**, *Chapter 8 Unit 8.14*.
(44) Korendovych, I. V.; Kulp, D. W.; Wu, Y.; Cheng, H.; Roder, H.; DeGrado, W. F. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 6823.
(45) Mack, K. L.; Moroz, O. V.; Moroz, Y. S.; Olsen, A. B.; McLaughlin, J. M.; Korendovych, I. V. *J. Biol. Inorg. Chem.* **2013**, *18*, 411.
(46) van Vranken, D. L.; Panomitros, D.; Schultz, P. G. *Tetrahedron Lett.* **1994**, *35*, 3873.
(47) Suga, H.; Ersoy, O.; Tsumuraya, T.; Lee, J.; Sinskey, A. J.; Masamune, S. *J. Am. Chem. Soc.* **1994**, *116*, 487.
(48) Taglieber, A.; Höbenreich, H.; Carballeira, J.; Mondière, R. J. G.; Reetz, M. T. *Angew. Chem. Int. Ed.* **2007**, *46*, 8597.
(49) Reetz, M. T. *J. Am. Chem. Soc.* **2013**, *135*, 12480.
(50) Moroz, O. V.; Moroz, Y. S.; Wu, Y.; Olsen, A. B.; Cheng, H.; Mack, K. L.; McLaughlin, J. M.; Raymond, E. A.; Zhezherya, K.; Roder, H.; Korendovych, I. V. *Angew. Chem. Int. Ed.* **2013**, *52*, 6246.
(51) Makhlynets, O. V.; Korendovych, I. V. *Biomolecules* **2014**, *4*, 402.
(52) Makhlynets, O. V.; Raymond, E. A.; Korendovych, I. V. *Biochemistry* **2015**, *54*, 1444.
(53) Raymond, E. A.; Mack, K. L.; Yoon, J. H.; Moroz, O. V.; Moroz, Y. S.; Korendovych, I. V. *Protein Sci.* **2015**, *24*, 561.
(54) Tinberg, C. E.; Khare, S. D.; Dou, J.; Doyle, L.; Nelson, J. W.; Schena, A.; Jankowski, W.; Kalodimos, C. G.; Johnsson, K.; Stoddard, B. L.; Baker, D. *Nature* **2013**, *501*, 212.
(55) Reetz, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5716.
(56) Privett, H. K.; Kiss, G.; Lee, T. M.; Blomberg, R.; Chica, R. A.; Thomas, L. M.; Hilvert, D.; Houk, K. N.; Mayo, S. L. *Proc. Natl. Acad. Sci. U.S.A* **2012**, *109*, 3790.
(57) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quiocho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177.
(58) 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.; Popović, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. *Meth. Enzymol.* **2011**, *487*, 545.
(59) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.
(60) Sanner, M. F. *J. Mol. Graph. Model.* **1999**, *17*, 57.
(61) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455.
(62) Eschenfeldt, W. H.; Stols, L.; Millard, C. S.; Joachimiak, A.; Donnelly, M. I. *Methods Mol. Biol.* **2009**, *498*, 105.
(63) Studier, F. W. *Prot. Expres. and Purif.* **2005**, *41*, 207.
(64) Marley, J.; Lu, M.; Bracken, C. *J. Biomol. NMR* **2001**, *20*, 71.
(65) Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stohr, J.; Smith, T. A.; Hu, X.; DeGrado, W. F.; Korendovych, I. V. *Nat. Chem.* **2014**, *6*, 303.
(66) Bender, M. L.; Kezdy, F. J.; Wedler, F. C. *J. Chem. Ed.* **1967**, *44*, 84.
(67) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277.
(68) Shen, Y.; Delaglio, F.; Cornilescu, G.; Bax, A. *J. Biomol. NMR* **2009**, *44*, 213.
(69) Güntert, P. *Meth. Mol. Biol.* **2004**, *278*, 353.
(70) Schwieters, C. D.; Kuszewski, J. J.; Tjandra, N.; Clore, G. M. *J Magn Reson* **2003**, *160*, 65.
(71) Linge, J. P.; Williams, M. A.; Spronk, C. A.; Bonvin, A. M.; Nilges, M. *Proteins* **2003**, *50*, 496.
(72) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Cryst. Sect. D.* **1998**, *54*, 905.
(73) Luthy, R.; Bowie, J. U.; Eisenberg, D. *Nature* **1992**, *356*, 83.
(74) Laskowski, R. A.; Rullmann, J. A.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477.
(75) Lovell, S. C.; Davis, I. W.; Arendall, W. B.; De Bakker, P. I. W.; Word, J. M.; Prisant, M. G.; Richardson, J. S.; Richardson, D. C. *Proteins* **2003**, *50*, 437.
(76) Sippl, M. J. *Proteins* **1993**, *17*, 355.
(77) Bhattacharya, A.; Tejero, R.; Montelione, G. T. *Proteins* **2007**, *66*, 778.

TOC graphic