Protein Design

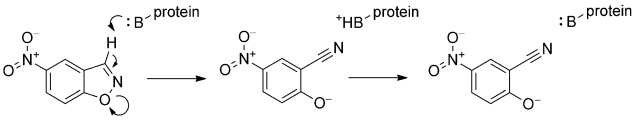
DOI: 10.1002/anie.201302339

A Single Mutation in a Regulatory Protein Produces Evolvable Allosterically Regulated Catalyst of Nonnatural Reaction**

*Olesia V. Moroz, Yurii S. Moroz, Yibing Wu, Alissa B. Olsen, Hong Cheng, Korrie L. Mack, Jaclyn M. McLaughlin, Elizabeth A. Raymond, Krystyna Zhezherya, Heinrich Roder, and Ivan V. Korendovych**

In memory of Robin M. Hochstrasser

Despite the progress in creating novel catalysts for nonnatural reactions and reengineering existing enzymes to adopt new properties, the problem of creating a catalyst for a particular chemical transformation is far from solved. Current successful approaches to develop enzymes for nonnatural reactions combine sophisticated computational algorithms with directed evolution methods.[1] Most available algorithms focus on introducing multiple mutations into existing protein scaffolds and optimizing interactions to satisfy a potential function developed to predict the thermodynamic stability of a protein.[2] However, catalysis often requires destabilizing specific groups to enhance their chemical reactivity and thus a purely thermodynamic approach may not sample the whole range of available scaffolds. Could one introduce a highly reactive functional group in a protein *without* extensive engineering of the protein scaffold followed by optimization by directed evolution to achieve function? Using a very basic minimalist design strategy we have shown that a single mutation in a non-catalytic calmodulin scaffold produced a small (74 amino acid long) allosterically regulated enzyme, AlleyCat,[3] that catalyzes Kemp elimination (Scheme 1), a benchmark reaction for testing various catalyst design methodologies.



Scheme 1. Kemp elimination.

While catalytic efficiency of the original AlleyCat catalyst is modest, here we show that our design approach produced a catalyst that could be optimized using standard directed evolution techniques. The activity of the evolved catalyst is on par with those designed by the best computational algorithms and catalytic antibodies.

AlleyCat contains only the C-terminal domain of calmodulin. In our directed evolution studies we reintroduced the N-terminal domain to obtain an additional way of purifying the protein and to improve expression yield in *E. coli*. The N-terminal domain did not significantly influence the catalytic activity. The resulting protein CaM F92E has essentially the same enzymatic efficiency and pK_a values as AlleyCat (Table 1). Next, we identified eight positions near the active site likely to influence the reactivity of the catalyst and performed saturation mutagenesis to generate libraries with complete coverage of all possible single mutants in these positions. Several mutants with substantially improved activity were identified; the best (M144R) showed more than 6-fold increase in enzymatic efficiency. In the second round of directed evolution, we performed gene shuffling of the mutants identified in the first round to achieve an overall 14-fold improvement in activity with only two mutations. Two subsequent rounds of error-prone PCR identified mutations outside the active-site region, to produce an overall 65-fold increase in k_{cat}/K_M. After one more round of gene shuffling and an additional round of error-prone PCR, we turned back to saturation mutagenesis. Reetz et al. showed[4] that extensive use of saturation mutagenesis often provides a more optimal path to finding active mutants. In agreement with this finding we observed that saturation mutagenesis in a position found to have no effect on activity in round 1, yielded a mutant (A88Q) with an additional 2-fold activity increase. Subsequent removal of the N-terminal domain resulted in a further improved 74-residue catalyst AlleyCat7. Detailed kinetic parameters for the evolved proteins are summarized in Table 1 and the Supporting Information. Overall, with only

[*] Dr. O. V. Moroz, Dr. Y. S. Moroz, A. B. Olsen, K. L. Mack, J. M. McLaughlin, E. A. Raymond, K. Zhezherya, Prof. Dr. I. V. Korendovych
Department of Chemistry, Syracuse University
111 College Place, Syracuse, NY 13244 (USA)
E-mail: ikorendo@syr.edu
Dr. Y. Wu
Department of Pharmaceutical Chemistry
University of California, San Francisco
555 Mission Bay Blvd. South, San Francisco, CA 94158 (USA)
Dr. Y. S. Moroz
Department of Chemistry, Taras Shevchenko National University of Kiev, 64 Volodymyrska St., Kiev, 01601 (Ukraine)
Dr. H. Cheng, Prof. Dr. H. Roder
Fox Chase Cancer Center
333 Cottman Ave., Philadelphia, PA 19111 (USA)

[**] We thank Prof. Robert P. Doyle for providing access to CD instrument. The Spectroscopy Support Facility at FCCC provided access to a 600 MHz NMR spectrometer.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201302339.

Table 1: Kinetic parameters of the evolved mutants.

Round, method[a]	Mutation[b]	$(k_{cat}/K_M)_{max}$[c] [s^{-1} M^{-1}]	k_{cat} [s^{-1}][d]	K_M [mM][d]	pK_a[c]
AlleyCat	F92E	5.8 ± 0.3			6.9 ± 0.1
CaM F92E	F92E	6.3 ± 0.2			6.98 ± 0.05
1, SM	M144R	47 ± 2	0.10 ± 0.02	4.1 ± 1.0	7.21 ± 0.06
2, GS	H107I	93 ± 2[e]	0.17 ± 0.02[e]	2.6 ± 0.4[e]	6.73 ± 0.03[e]
3, epPCR	L112R	200 ± 3	0.37 ± 0.04	3.1 ± 0.4	6.70 ± 0.03
4, epPCR	I85L	296 ± 5	0.67 ± 0.09	2.9 ± 0.5	6.93 ± 0.03
5, GS	A128T	414 ± 4	0.54 ± 0.02	2.0 ± 0.1	6.60 ± 0.02
		524 ± 9[e]	0.78 ± 0.07[e]	1.9 ± 0.2[e]	6.47 ± 0.03[e]
6, epPCR	M124L	543 ± 4	0.65 ± 0.06	2.2 ± 0.3	6.67 ± 0.01
7, SM	A88Q	1069 ± 14[e]	1.86 ± 0.10[e]	2.1 ± 0.2[e]	6.43 ± 0.02[e]
AlleyCat7		1283 ± 13[e]	1.87 ± 0.14[e]	2.0 ± 0.2[e]	6.39 ± 0.02[e]
			3.2 ± 0.2[f]	2.4 ± 0.2[f]	

[a] SM: saturation mutagenesis, GS: gene shuffling, epPCR: error-prone PCR. [b] In addition to the ones in the preceding round. The numbering scheme refers to the positions of residues in parent chicken calmodulin.[7] [c] $(k_{cat}/K_M)_{max}$ and pK_a values were obtained from the pH activity profile as described in detail in the Supporting Information. Each individual k_{cat}/K_M value was obtained from the linear portion of the Michaelis–Menten plot. [d] Individual k_{cat} and K_M values were obtained from fitting the Michaelis–Menten equation to the data at pH 7.0. Note that due to limited substrate solubility these values could be subject to higher uncertainty. [e] With affinity tag removed. [f] At pH 8.0.

7 rounds of directed evolution, having screened less than 6000 mutants, we achieved a 220-fold improvement in k_{cat}/K_M. This value compares well to the results of extensive directed evolution of Kemp eliminases reported by Tawfik and Baker: the catalytic efficiencies of KE07 and KE59, that, as in AlleyCat, use a single Glu as a catalytic base, were improved by 210-fold and 180 to 380-fold, respectively.[2b,5] The k_{cat} value for the evolved protein AlleyCat7 is only ca. 3-fold lower than that of the best evolved KE variant (KE59 R13-3/11H) and ca. 5-fold higher than that of a catalytic antibody for Kemp elimination.[6] The k_{cat}/k_{uncat} value for AlleyCat is more than 10^6. The improvement in enzymatic efficiency of AlleyCat over 7 rounds of directed evolution is achieved almost exclusively due to increase in k_{cat}. Unlike the KE family designs, improvement in k_{cat} is not a result of increased basicity of the active residue as the apparent pK_a of the catalytic Glu92 residue generally decreases in the evolved variants. This is consistent with our overall approach to introduce a highly reactive glutamate residue without complete pre-optimization of the active site.

Mutation of Glu92 to Gln reduces activity by at least 2200-fold, confirming that the original reaction mechanism is preserved. Evolution of catalytic activity of AlleyCat did not significantly influence its affinity for calcium; hence, the allosteric regulation of activity is still intact in the evolved protein as the protein is completely inactive in the absence of Ca^{2+}. Moreover, the protein can be cycled through the on and off states without any loss of activity or product inhibition (Figure 1).

Most of the newly identified mutations (M144R, L112R, A128T and A88Q) introduced additional polar groups in the hydrophobic region of the protein. Two mutations, I85L and M124L, could be improving packing around the active site. H107I increases steric bulk around the catalytic residue and could be locking the active-site carboxylate in the productive conformation. Interestingly, the backbone amide proton resonance of H107 in AlleyCat is quite broad, but the peak of I107 in CaM F92E 5R HSQC spectrum is sharp and well defined, consistent with improved stability of the helix formed by residues 103–111.

Evolution of protein function is often limited by the ability of the protein scaffold to accept additional mutations. We used circular dichroism (CD) spectroscopy to characterize the secondary structure and stability of the mutants identified in our studies. All evolved proteins showed well-defined helical structure, nearly identical to that of holo calmodulin. Chemical denaturation experiments with guanidinium chloride were performed to quantify the stability of the protein fold. In line with previous directed evolution studies,[5a,8] mutations that greatly improved activity in the initial rounds had a profound detrimental effect on scaffold stability. However, additional mutations did not substantially disrupt the fold and after round 5 an overall increase in free energy of unfolding was observed (Figure S25, Table S3, Supporting Information). In all cases the energy provided by calcium binding to EF-hands is sufficient to preserve the proper protein fold.

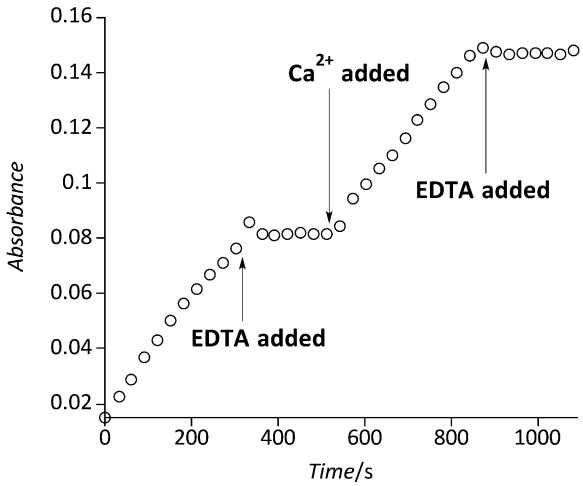


Figure 1. Allosteric regulation of AlleyCat7. Experimental conditions: initial concentrations: 130 nM protein, 100 mM NaCl, 20 mM HEPES buffer, pH 7.0, 0.1 mM $CaCl_2$, 0.1 mM substrate. At 300 s EDTA (ethylenediaminetetraacetic acid) was added to the final concentration of 0.2 mM, followed by addition of $CaCl_2$ at 540 s to the final concentration of 0.3 mM and, again, EDTA at 840 s to the final concentration of 0.5 mM.

The round 5 protein, which we found to be the least stable by chemical denaturation, was structurally characterized by NMR spectroscopy (Figure 2). The structure is very similar to that of the original AlleyCat design (rmsd 1.9 Å), but shows somewhat tighter packing around the active site. As in the original design, the side chain of the active site glutamate is pointing into the middle of the hydrophobic pocket. The position of the side chain of arginine 144 on the other side of

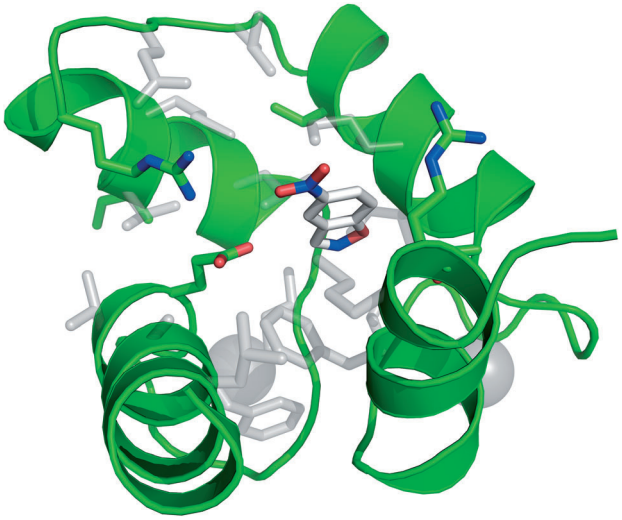


Figure 2. NMR structure of the CaM F92E 5R mutant (only the C-terminal domain is shown). The side chains for mutations introduced through the design (F92E) and directed evolution are shown in color. The side chains for the residues facing the active sites and the calcium atoms (spheres) are shown in gray. The rest of the side chains are omitted for clarity. PDB code 2m3s. The substrate was docked into the structure using Autodock 3.0.

the substrate-binding pocket is consistent with proton-donation to the oxygen of the benzisoxazole ring, shown to be beneficial for the reaction.[9] Interestingly, the L112R mutation brings the arginine side chain within a hydrogen-bonding distance to the carboxylate of the active-site glutamate. This interaction, which does not greatly perturb the pK_a of Glu92, can lock the position of the Glu side chain in the productive conformation and/or stabilize the charged carboxylate in the hydrophobic environment.

AlleyCat7 is effectively ($K_I = 18$ μM) inhibited by 5-nitrobenzotriazole (5-NBT), a transition-state analog. Analysis of the backbone amide NH chemical shift perturbation data for the round 5 variant shows that 5-NBT binds to the C-terminal domain in the pose consistent with our previous modeling studies (Figure S35, Supporting Information).[3] Analysis of the backbone NH residual dipolar couplings shows that the most active round 7 variant has nearly identical fold to that of the round 5 variant (Figure S36, Supporting Information).

Our findings underscore the importance of two factors for determining whether a particular designed protein could be evolvable. First, in agreement with Arnold et al.[10] we find that stability of the protein scaffold is extremely important for functional innovation. While directed evolution approaches that utilize enzymes originating from thermophiles or introduce consensus mutations have been well established, we show that metal-binding domains can also provide the necessary structural stability to the designed protein. The highly stable calmodulin fold tolerated multiple mutations that would be detrimental in the absence of calcium. Additionally, metal-binding loops provide allosteric control of the function and could be independently evolved to respond to different metals.[11] The resulting Kemp eliminases can serve as catalytically amplified colorimetric sensors. Second, evolvability of proteins is linked to their conformational dynamism.[12] While calmodulin has a fairly rigid scaffold, it is capable of tight binding to a great variety of diverse binding partners.[13] This unique property is made possible by a number of flexible, predominantly hydrophobic side chains lining the recognition site, which was used in AlleyCat's design. The large number of possible conformations available for these residues likely improved evolvability.

The current consensus is that new function emerges from latent, promiscuous enzymes that can take on multiple substrates giving a starting point for further evolution.[8c] This hypothesis was practically proven by multiple examples of redesigning proteins to perform novel nonnatural reactions starting from enzymes that have similar function and use chemically similar substrates.[14] What about reactions that have no obvious natural analogs? Here we have shown that our approach, combining simple physico-chemical principles combined with basic computational techniques can give rise to catalysts with enough starting activity for subsequent directed evolution. Moreover, given enough stability, even a small 74-residue protein can be successfully evolved to match activity of catalytic antibodies for a particular nonnatural reaction. We expect that our minimalistic approach based on introducing highly reactive residues into protein scaffolds will greatly enhance future design work, allowing more sophisticated computations to reach their full potential.

Received: March 19, 2013
Published online: April 29, 2013

Keywords: enzyme catalysis · metalloproteins · protein design

[1] a) E. M. Brustad, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2011**, *15*, 201 – 210; b) S. Lutz, *Curr. Opin. Biotechnol.* **2010**, *21*, 734 – 743.
[2] a) J. K. Lassila, H. K. Privett, B. D. Allen, S. L. Mayo, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16710 – 16715; b) D. Röthlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik, D. Baker, *Nature* **2008**, *453*, 190 – 195.
[3] I. V. Korendovych, D. W. Kulp, Y. Wu, H. Cheng, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6823 – 6827.
[4] a) X. Feng, J. Sanchis, M. T. Reetz, H. Rabitz, *Chem. Eur. J.* **2012**, *18*, 5646 – 5654; b) Y. Gumulya, J. Sanchis, M. T. Reetz, *ChemBioChem* **2012**, *13*, 1060 – 1066.
[5] a) O. Khersonsky, G. Kiss, D. Röthlisberger, O. Dym, S. Albeck, K. N. Houk, D. Baker, D. S. Tawfik, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10358 – 10363; b) original KE59 design is unstable and its enzymatic efficieny cannot be accurately determined.
[6] S. N. Thorn, R. G. Daniels, M.-T. Auditor, D. Hilvert, *Nature* **1995**, *373*, 228 – 230.
[7] J. L. Urbauer, J. H. Short, L. K. Dow, A. J. Wand, *Biochemistry* **1995**, *34*, 8099 – 8109.
[8] a) O. Khersonsky, D. Röthlisberger, O. Dym, S. Albeck, C. J. Jackson, D. Baker, D. S. Tawfik, *J. Mol. Biol.* **2010**, *396*, 1025 – 1042; b) O. Khersonsky, D. Röthlisberger, A. M. Wollacott, P. Murphy, O. Dym, S. Albeck, G. Kiss, K. N. Houk, D. Baker, D. S. Tawfik, *J. Mol. Biol.* **2011**, *407*, 391 – 412; c) O. Khersonsky, D. S. Tawfik, *Annu. Rev. Biochem.* **2010**, *79*, 471 – 505.
[9] E. W. Debler, R. Muller, D. Hilvert, I. A. Wilson, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18539 – 18544.

[10] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 5869 – 5874.
[11] K. L. Mack, O. V. Moroz, Y. S. Moroz, A. B. Olsen, J. M. McLaughlin, I. V. Korendovych, *J. Biol. Inorg. Chem.* **2013**, *18*, 411 – 418.
[12] N. Tokuriki, D. S. Tawfik, *Science* **2009**, *324*, 203 – 207.
[13] K. T. O'Neil, W. F. DeGrado, *Trends Biochem. Sci.* **1990**, *15*, 59 – 64.
[14] a) S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291 – 297; b) M. D. Toscano, K. J. Woycechowsky, D. Hilvert, *Angew. Chem.* **2007**, *119*, 3274 – 3300; *Angew. Chem. Int. Ed.* **2007**, *46*, 3212 – 3236.

