doi:10.2533/chimia.2009.309 Chimia 63 (2009) 309–312

Catalysts on Demand: Selective Oxidations by Laboratory-Evolved Cytochrome P450 BM3

Jared C. Lewis and Frances H. Arnold*

Abstract: Efficient catalysts for selective oxidation of C–H bonds using atmospheric oxygen are highly desirable to decrease the economic and environmental costs associated with conventional oxidation processes. We have used methods of directed evolution to generate variants of bacterial cytochrome P450 BM3 that catalyze hydroxylation and epoxidation of a wide range of nonnative substrates. This fatty acid hydroxylase was converted to a propane monooxygenase (PMO) capable of hydroxylating propane at rates comparable to that of BM3 on its natural substrates. Variants along the PMO evolutionary lineage showed broadened substrate scope; these became the starting points for evolution of a wide array of enzymes that can hydroxylate and derivatize organic scaffolds. This work demonstrates how a single member of enzyme family is readily converted by evolution into a whole family of catalysts for organic synthesis.

Keywords: Biocatalysis · Evolution · Drug metabolite · Oxidation · P450

Simple reduced hydrocarbons obtained from petroleum feedstocks are used to produce over 90% of the world's synthetic organic chemicals, including pharmaceuticals, agrochemicals, and bulk commodity products.[1] This process begins with the oxidation of hydrocarbons to install functional groups that enable subsequent synthetic manipulations. While a variety of oxidants are currently used in these transformations on a massive scale, many provide only moderate yields or selectivities and produce environmentally hazardous by-products. Furthermore, these methods are not applicable for the synthesis of fine chemicals and advanced intermediates with more complex structures and sensitive functional groups; extensive effort is required to manipulate the oxidation state of these molecules throughout their preparation.[2] There is great interest therefore in developing catalysts capable of selective hydrocarbon oxidation under mild conditions using oxygen as the terminal oxidant.[3]

*Correspondence: Prof. F. H. Arnold
California Institute of Technology
Division of Chemistry and Chemical Engineering, 210-41
Pasadena, CA 91125 USA
Tel.: +1 626 395 4162
Fax: +1 626 568 8743
E-mail: frances@cheme.caltech.edu

The advantages of such reactions have not been overlooked in nature, and diverse enzymatic solutions to this challenge enable organisms to utilize hydrocarbons and other hydrophobic substrates in metabolic processes.[4] The cytochrome P450 monooxygenase superfamily is a remarkable example of how nature can generate a whole spectrum of catalysts from one common framework: more than 7,000 P450 sequences identified from all kingdoms of life catalyze the oxidation of a huge variety of organic compounds. Each of these enzymes harbors a cysteine-bound heme cofactor responsible for this activity,[5] but the cytochrome P450 BM3 from *Bacillus megaterium* (BM3) possesses a number of features that make it particularly attractive for applications in chemical synthesis.[6] BM3 is one of only a handful of known P450s in which the heme domain and the diflavin reductase domains (FMN and FAD) required for generation of the active oxidant are fused in a single polypeptide chain.[6] Thus, BM3 is soluble, readily overexpressed in a variety of heterologous hosts, and requires only atmospheric oxygen and a supply of nicotinamide adenine dinucleotide phosphate (NADPH) for hydroxylation activity.[7] The substrates of BM3, however, are largely limited to C12–C18 fatty acids, which it hydroxylates at subterminal positions at rates of thousands of turnovers per minute, making it one of the most active hydroxylases known.[8]

Given the engineering-friendly properties of BM3 and the potential utility of C–H bond hydroxylation, we set out about ten years ago to explore and expand the substrate scope of this enzyme with the goal of creating efficient enzyme and whole-cell catalysts for a variety of oxidative transformations. Extensive effort has been devoted by the research community to rational mutagenesis of P450 enzymes to probe the roles of various residues on substrate binding and catalysis.[9] Guided by numerous crystal structures of substrate-bound and unbound P450s, these studies demonstrated that modification of properties was possible but time consuming, complicated by complex interactions between structure and catalysis, and limited to modest modifications of properties. Inspired by the engineering approach that generated the impressive array of P450s in nature, we have instead used directed evolution, iterative rounds of mutagenesis and functional screening, to generate new BM3-based catalysts.[10]

One of our earliest goals was to engineer BM3 to hydroxylate linear alkanes, including the short chain gaseous alkanes found in natural gas. Such a catalyst could enable bioconversion of natural gas to more convenient liquid alcohols for transportation or direct application as fuel, which has been a longstanding goal in catalysis with significant economic and environmental implications.[1] While wild type BM3 has only weak activity on long-chain alkanes (and no measurable activity on short alkanes),[8] we hypothesized that mutants of this enzyme displaying enhanced alkane activity might acquire measurable ability to hydroxylate slightly shorter alkanes. Further mutagenesis of active variants, with screening on progressively shorter-chained substrates, could ultimately generate enzymes active on the gaseous alkanes. Breaking down an

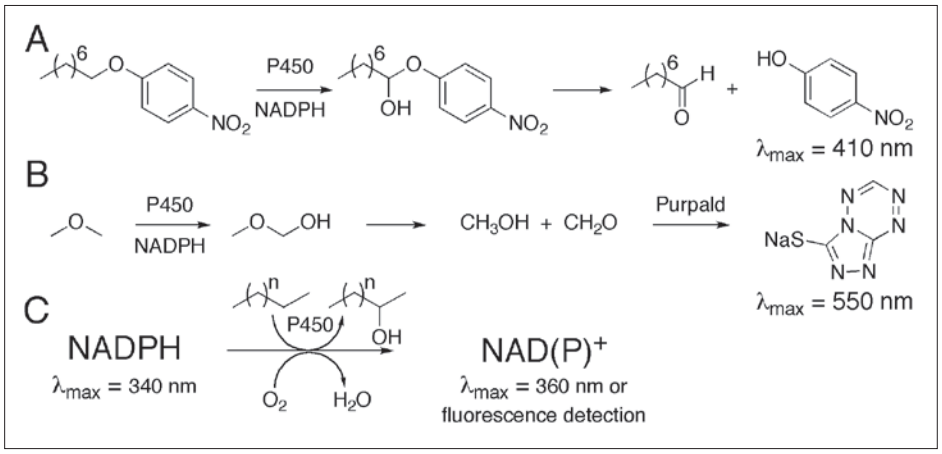
apparently difficult problem (obtaining activity on very small substrates) into a series of smaller problems lowers the bar for each generation, and can even allow new activities to be acquired one mutation at a time. This dramatically decreases the screening requirement, such that beneficial mutations can be found upon screening only a few thousand clones per generation versus the many millions required to find even two simultaneous mutations.

We initially investigated BM3 libraries generated by random mutagenesis of the heme domain.[11] Error-prone PCR reaction mixtures were plated, single colonies were picked, and variants were expressed in 96-well plates. The cells were then lysed and BM3 mutants in the crude lysate were used directly for hydroxylation, which enabled rapid analysis of thousands of P450 sequences using various high throughput assays for alkane hydroxylation. The octane surrogate 1-nitro-4-(octyloxy) benzene (NOB) was used in early experiments (Scheme 1A).[12] Hydroxylation of this substrate at the oxygen-bound methylene generates an unstable hemi-acetal intermediate that decomposes to octanal and *p*-nitrophenol, which can be detected using a plate reader. Similarly, hydroxylation of the propane surrogate dimethylether (DME) generates formaldehyde, which can be detected following reaction with Purpald (Scheme 1B).[13] Consumption of NADPH in the presence of various substrates (octane, propane, *etc.*) was also used in order to ensure that activity on surrogate substrates translated to the desired substrate (Scheme 1C).[14] Ultimately, activity on alkanes was confirmed by GC analysis of the alcohol products.

Five generations of random mutagenesis, recombination of beneficial mutations, and screening for activity on NOB or octane led to the creation of BM3 variant 139-3.[15] This enzyme contains eleven amino acid substitutions relative to BM3, only one of which is in the active site. 139-3 was able to hydroxylate C8–C3 *n*-alkanes. Further rounds of random mutagenesis, recombination, and more focused active site engineering further optimized the activity of 139-3 for hydroxylation of short-chain alkanes, using DME and propane assays to assess the mutants.[16] Variant 35E11, with 17 mutations relative to BM3, displayed high activity on propane, and even provided modest conversion of ethane to ethanol.[16b]

BM3-catalyzed hydroxylation requires finely tuned conformational rearrangements within and between the heme and reductase domains in order to efficiently couple hydroxylation to NADPH consumption.[17] These processes are disrupted when mutations are introduced or nonnative substrates are utilized. Random mutagenesis also decreases the thermostability of enzymes, which shortens catalyst lifetimes, but more importantly limits further evolution.[18] These issues were addressed using a domain-based approach, in which the heme, FMN, and FAD domains of 35E11 were evolved individually in the context of the holoenzyme and combined in a final step.[19] Optimization of the heme domain involved the introduction of thermostabilizing mutations, followed by random mutagenesis, saturation mutagenesis of active site residues, and recombination of active variants. Random mutagenesis of the FMN and FAD domains revealed a number of beneficial mutations, and saturation mutagenesis of these sites provided further improvements. Finally, the optimized domains were fused to generate $P450_{PMO}$R2 (PMO), which provided more than 45,000 turnovers on propane and produced 2- and 1-propanol in a 9:1 ratio (Fig. 1A). This enzyme displayed activity on propane comparable to that of BM3 on its natural substrates and 98.2% coupling of NADPH consumption to product hydroxylation.

Investigation of the substrate profiles for variants along the lineage for the evolution of PMO revealed that the specificity of BM3 for hydroxylation of long chain alkanes was relaxed in intermediate variants (*e.g.* 139-3, 35E11) and ultimately respecified to propane in the case of PMO (Fig. 1B).[20] This result was notable because only positive selection had been



Scheme 1. Screens used in evolution of BM3. (A) Hydroxylation of nitro-4-(octyloxy)benzene (NOB) generates a colored phenol,[12] (B) hydroxylation of dimethylether (DME) generates formaldehyde, which reacts with Purpald to generate a colored heterocycle,[13] and (C) substrate hydroxylation consumes NADPH.[14] Each of these can be monitored by visible or fluorescence spectroscopy using a plate reader.

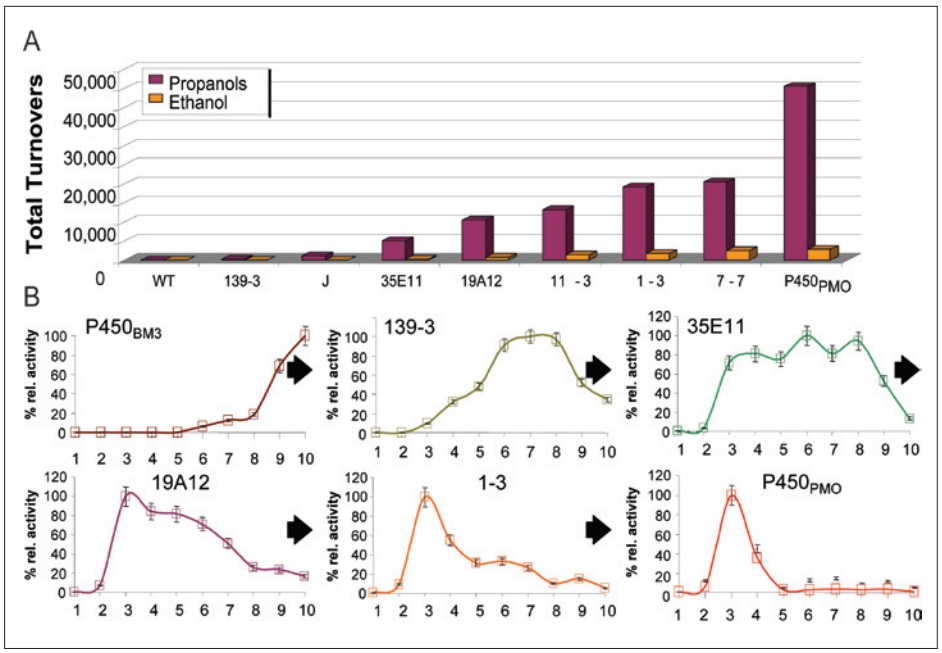
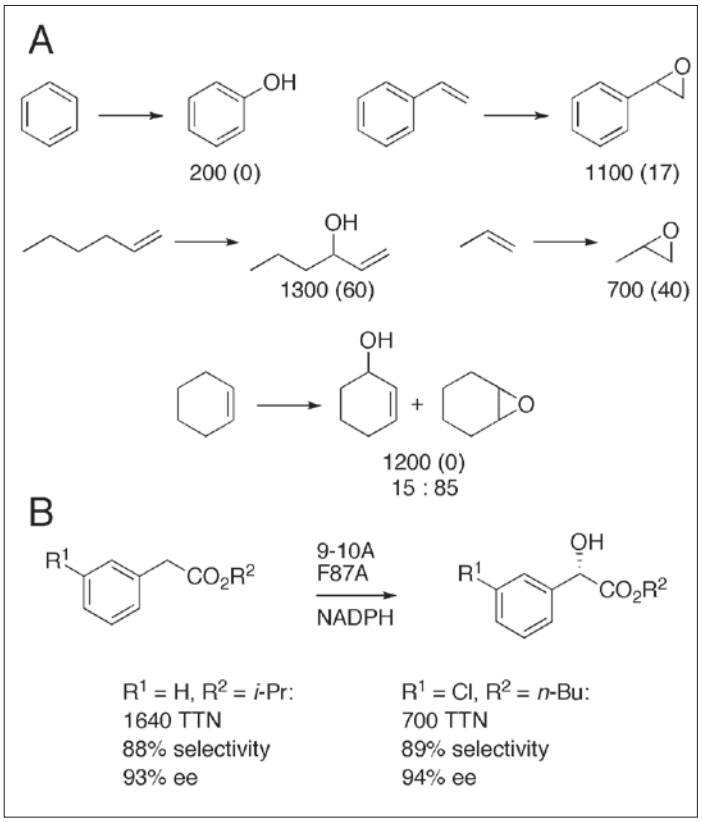


Fig. 1. (A) Total turnovers for variants along the PMO lineage on propane and ethane. (B) Activity of variants along the PMO lineage on C_n (n = 1–10) alkanes demonstrated how directed evolution passed through less-specific intermediate catalysts before re-specializing in on propane.[20]

used throughout this process and indicated that negative selection pressure is not necessarily required to obtain high specificity for a new substrate. In addition, the broad substrate scope of the intermediate variants suggested that hydroxylation of additional substrates might be possible,[21] and we began to investigate the activity of these intermediate enzymes on a variety of small molecules.

Initial experiments revealed that 139-3 catalyzed the hydroxylation of a variety of aliphatic and aromatic substrates and the epoxidation of olefins at rates and conversions far greater than BM3 (Scheme 2A).[22] The hydroxylation of 2-aryl acetic acid derivatives by alkane intermediate variant 9-10A was also investigated (Scheme 2B).[23] These substrates present the enzymes a number of different reactive C–H bonds (aromatic, benzyl, and alkyl). Due to differences in the mechanisms of aromatic and alkyl hydroxylation and differences in bond strength between primary, secondary, and benzyl alkyl C–H bonds, regioselectivity in these reactions is difficult to predict.[24] However, 9-10A proved to be an efficient catalyst in this respect, and hydroxylation proceeded with high regio- and enantioselectivity at the benzyl position of certain substrates.



Scheme 2. (A) Oxidation of various small hydrocarbons using 139-3 in the presence of NADPH. Maximum initial rates (nmol NADPH consumed/min/nmol P450) for 139-3 and (BM3) are shown.[22] (B) Hydroxylation of phenyl acetic acid derivatives using 9-10A in the presence of NADPH.[23]

To further explore the reactivities of enzymes lying along the PMO lineage as well as BM3 variants from recombination experiments,[25] we compiled a collection of ~120 variants to screen for activity on a variety of substrates of interest. In a collaboration with scientists from Eli Lilly, we were able to use the enzymes from this small collection, to prepare nearly all of the human metabolites and a number of novel hydroxylated derivatives for the drugs verapamil and astemizole (Fig. 2).[26] The ability to rapidly determine potential metabolites of drug candidates greatly facilitates pharmacological and toxicological evaluations of these molecules in preclinical studies,[27] so the broad scope demonstrated by these BM3 variants should be particularly useful. This panel of enzymes could also be used further upstream in drug development as a means to generate derivatives of candidate structures.

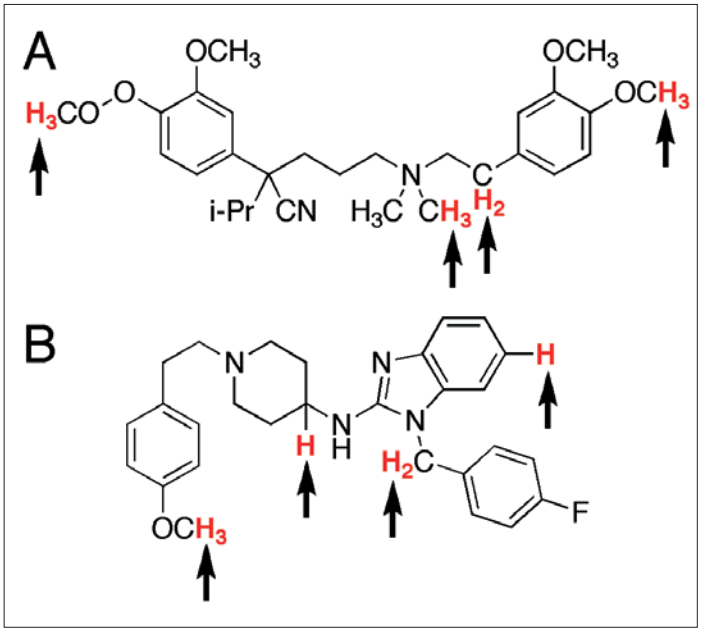
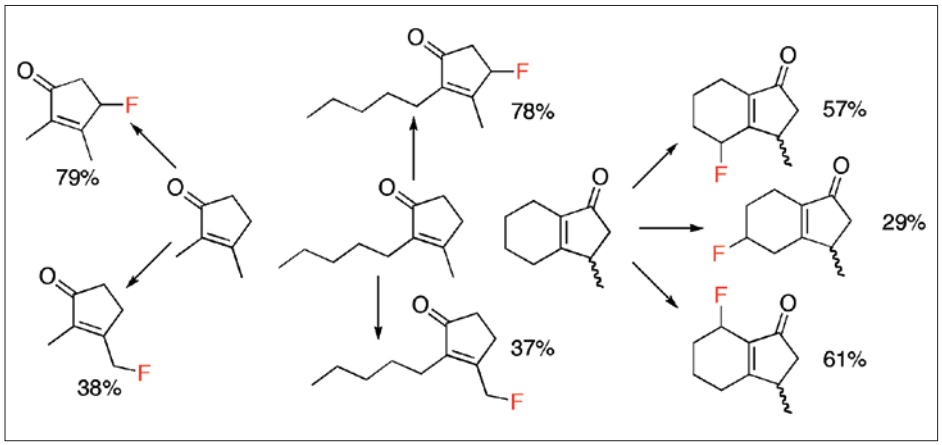


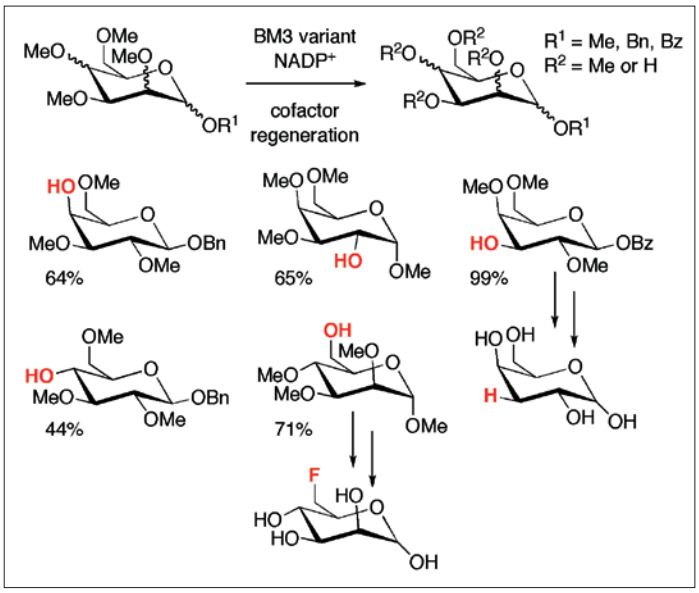
Fig. 2. Structures of (A) verapamil and (B) astemizole. Sites hydroxylated by BM3 variants are designated with arrows.[26] Hydroxylation of heteroatom-bonded alkyl groups resulted in heteroatom dealkylation.

In a similar fashion, highly active variants were identified for the hydroxylation of a variety of small molecules (Scheme 3).[20,28] While these simple, predominantly hydrocarbon substrates are often commercially available, functionalized derivatives are difficult to obtain. Thus the ability to install functional handles in these structures is highly useful for organic synthesis. Furthermore, these hydroxylated products could be subsequently fluorinated in a chemo-enzymatic process that greatly simplified the installation of fluorine into these structures at specific positions.[28] Because the regioselectivity of fluorination is imparted by the P450 enzymes, this process could be particularly useful for improving the metabolic stability of drug candidates, since human P450s would be expected to hydroxylate at least some of these same positions *in vivo*.[29]

As a final example of the utility of BM3 variants in chemical synthesis, we have engineered BM3 to regioselectively deprotect permethylated monosaccharides in a chemoenzymatic method for synthesis of monosaccharide derivatives commonly found in pharmaceuticals and other biologically active molecules (Scheme 4).[30] Synthesis of these valuable compounds is highly challenging using conventional synthetic methods, and efficient methods for their preparation are sorely needed.[31] Screening the BM3 collections revealed a number of variants capable of selectively removing single methyl groups from uniformly protected sugar derivatives, and directed evolution was used to improve both the regioselectivity and activity of these variants. The deprotected monosaccharides were then converted to deoxy-sugars, fluorinated sugars, and even disaccharides (J. Lewis, unpublished results).



Scheme 3. Chemoenzymatic fluorination of common small molecules. Yields for the two-step BM3 hydroxylation/DAST mediated deoxyfluorination sequence are shown.[28]



Scheme 4. Regioselective demethylation of methylated monosaccharide derivatives. Isolated yields of mono-demethylated products are shown, and the conversion of select products to monosaccharide derivatives is shown.[30]

Together, these results demonstrate the ability of directed evolution to generate useful catalysts for oxidation of a wide range of substrates. Creation of the PMO highlights the stunning levels of efficiency possible using this process. In addition, the broad substrate scope of variants along the PMO lineage as well as those obtained by recombination has enabled rapid hydroxylation and derivatization of common organic scaffolds that would be difficult to accomplish using standard chemical means. Specific transformations were easily improved by further directed evolution of variants possessing broad substrate scope. These examples demonstrate the remarkable ability of cytochrome P450 to adapt to new challenges, an ability we expect to continue to exploit in the production of catalysts useful in chemical synthesis.

Acknowledgements

This work was supported by the U.S. Department of Agriculture, grant 2006-35505-16660, and by the U.S. Department of Energy, grant DE-FG02-06ER15762, to F.H.A. J.C.L is supported by the U.S. National Institutes of Health, fellowship 1F32GM079932-01.

Received: January 28, 2009

[1] K. Weissermel, H.-J. Arpe, 'Industrial Organic Chemistry', 4th ed., Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, **2003**.
[2] M. S. Chen, M. C. White, *Science* **2007**, *318*, 783.
[3] a) 'Oxygenases and Model Systems', Ed. T. Funabiki, Kluwer, Boston, 1997; b) S. S. Stahl, *Angew. Chem. Int. Ed.* **2004**, *43*, 3400; c) A. E. Shilov, G. B. Shul'pin, *Chem. Rev.* **1997**, *97*, 2879.
[4] a) E. I. Solomon, U. M. Sundaram, T. E. Machonkin, *Chem. Rev.* **1996**, *96*, 2563; b) L. Que, R. Y. N. Ho, *Chem. Rev.* **1996**, *96*, 2607; c) B. J. Wallar, J. D. Lipscomb, *Chem. Rev.* **1996**, *96*, 2625; d) T. J. Kappock, J. P. Caradonna, *Chem. Rev.* **1996**, *96*, 2659; e) M. Sono, M. P. Roach, E. D. Coulter, J. H. Dawson, *Chem. Rev.* **1996**, *96*, 2841.
[5] a) 'Cytochrome P450: Structure, Mechanism, and Biochemistry', Ed. P. R. Ortiz de Montellano, Kluver Academic/Plenum Publishers, New York, **2005**; b) I. G. Denisov, T. M. Makris, S. G. Sligar, *Chem. Rev.* **2005**, *105*, 2253.
[6] A. J. Warman, O. Roitel, R. Neeli, H. M. Girvan, H. E. Seward, S. A. Murray, K. J. McLean, M. G. Joyce, H. Toogood, R. A. Holt, D. Leys, N. S. Scrutton, A. W. Munro, *Biochem. Soc. Trans.* **2005**, *33*, 747.
[7] R. T. Ruettinger, L.-P. Wen, A. J. Fulco, *J. Biol. Chem.* **1989**, *264*, 10987.
[8] a) Y. Miura, A. J. Fulco, *Biochim. Biophys. Acta* **1975**, *388*, 305; b) S. S. Boddupalli, R. W. Estabrook, J. A. Peterson, *J. Biol. Chem.* **1990**, *265*, 4233.
[9] Y. Lu, S. M. Berry, T. D. Pfister, *Chem. Rev.* **2001**, *101*, 3047.
[10] F. H. Arnold, *Acc. Chem. Res.* **1998**, *31*, 125.
[11] E. T. Farinas, U. Schwaneberg, A. Glieder, F. H. Arnold, *Adv. Synth. Catal.* **2001**, *343*, 601.
[12] U. Schwaneberg, C. Otey, P. C. Cirino, E. Farinas, F. H. Arnold, *J. Biomol. Screen* **2001**, *6*, 111.
[13] H. B. Hopps, *Aldrichimica Acta* **2000**, *33*, 28.
[14] R. S. Matson, R. S. Hare, A. J. Fulco, *Biochem. Biophys. Acta* **1977**, *487*, 487.
[15] A. Glieder, E. T. Farinas, F. H. Arnold, *Nat. Biotechnol.* **2002**, *20*, 1135.
[16] a) M. W. Peters, P. Meinhold, A. Glieder, F. H. Arnold, **2003**, *125*, 13442; b) P. Meinhold, M. W. Peters, M. Y. Chen, K. Takahashi, F. H. Arnold, *ChemBioChem* **2005**, *6*, 1765.
[17] a) A. W. Munro, D. G. Leys, K. J. McLean, K. R. Marshall, T. W. Ost, S. Daff, C. S. Miles, S. K. Chapman, D. A. Lysek, C. C. Moser, C. C. Page, P. L. Dutton, *Trends Biochem. Sci.* **2002**, *27*, 250; b) D. C. Haines, D. R. Tomchick, M. Machius, J. A. Peterson, *Biochem.* **2001**, *40*, 13456; c) M. B. Murataliev, R. Feyereisen, **1996**, *35*, 15029.
[18] J. D. Bloom, S. T. Labthavikul, C. R. Otey, F. H. Arnold, *Proc. Natl. Acad. Sci.* **2006**, *103*, 5869.
[19] R. Fasan, M. M. Chen, N. C. Crook, F. H. Arnold, *Angew. Chem. Int. Ed.* **2007**, *46*, 8414.
[20] R. Fasan, Y. T. Meharenna, C. D. Snow, T. L. Poulos, F. H. Arnold, *J. Mol. Biol.* **2008**, *383*, 1069.
[21] a) R. A. Jensen, *Annu. Rev. Microbiol.* **1974**, *30*, 409; b) A. Aharoni, L. Gaidukov, O. Khersonsky, Q. G. S. Mc, C. Roodveldt, D. S. Tawfik, *Nat. Genet.* **2005**, *37*, 73.
[22] E. T. Farinas, M. Alcalde, F. H. Arnold, *Tetrahedron* **2004**, *60*, 525.
[23] M. Landwehr, L. Hochrein, C. R. Otey, A. Kasrayan, J.-E. Bäckvall, F. H. Arnold, *J. Am. Chem. Soc.* **2006**, *128*, 6058.
[24] F. P. Guengerich, *Chem. Res. Toxicol.* **2001**, *14*, 611.
[25] M. Landwehr, M. Carbone, C. R. Otey, Y. Li, F. H. Arnold, *Chem. Biol.* **2007**, *14*, 269.
[26] A. M. Sawayama, M. M. Chen, P. Kulanthaivel, M.-S. Kuo, H. Hemmerle, F. H. Arnold, *Chem. Eur. J.* **2008**, accepted for publication.
[27] M. D. Johnson, H. Zuo, K. H. Lee, J. P. Trebley, J. M. Rae, R. V. Weatherman, Z. Desta, D. A. Flockhart, T. C. Skaar, *Breast Cancer Res. Treat.* **2004**, *85*, 151.
[28] A. Rentmeister, F. H. Arnold, R. Fasan, *Nat. Chem. Biol.* **2008**, *5*, 26.
[29] K. Muller, C. Faeh, F. Diederich, *Science* **2007**, *317*, 1881.
[30] J. C. Lewis, S. Bastian, C. S. Bennet, F. Yu, Y. Mitsuda, M. M. Chen, W. A. Greenberg, C.-H. Wong, F. H. Arnold, unpublished results.
[31] T. Tsuchiya, in 'Glycoscience: Chemistry and Chemical Biology', Vol. 1; Eds. B. Fraser-Reid, K. Tatsuta, J. Thiem, Springer: Berlin, 2001, pp. 119–194.