JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

P450 Fingerprinting Method for Rapid Discovery of Terpene Hydroxylating P450 Catalysts with Diversified Regioselectivity

Kaidong Zhang, Shady El Damaty, and Rudi Fasan*

Department of Chemistry, University of Rochester, Rochester, New York 14627, United States

Supporting Information

ABSTRACT: Engineered P450 enzymes constitute attractive catalysts for the selective oxidation of unactivated C−H bonds in complex molecules. A current bottleneck in the use of P450 catalysis for chemical synthesis is the time and effort required to identify the P450 variant(s) with the desired level of activity and selectivity. In this report, we describe a method to map the active site configuration of engineered P450 variants in high throughput using a set of semisynthetic chromogenic probes. Through analysis of the resulting 'fingerprints', reliable predictions can be made regarding the reactivity of these enzymes toward complex substrates structurally related to the fingerprint probes. In addition, fingerprint analysis offers a convenient and time-effective means to assess the regioselectivity properties of the fingerprinted P450s. The described approach can represent a valuable tool to expedite the discovery of P450 oxidation catalysts for the functionalization of relevant natural products such as members of the terpene family.

Methods for selective oxidation of aliphatic (sp^3) C−H bonds are of huge synthetic relevance, in particular when viable for the late-stage oxidation of complex molecules and natural products.[1] For this purpose, the use of cytochrome P450 monooxygenases offers the advantage that the regio- and stereoselectivity of these catalysts can be modulated by protein engineering and potentially directed also toward energetically and/or stereoelectronically 'unbiased' sp^3 C−H sites. Cytochrome P450 enzymes have been isolated and engineered for a variety of relevant applications.[2−4] Recently, the systematic utilization of P450 variants with diversified substrate profile and regioselectivity has constituted a powerful strategy toward the late-stage transformation of single and multiple unactivated sp^3 C−H bonds in small-molecule substrates through P450-mediated chemoenzymatic synthesis.[5] The lack of time-effective methods to gain access to P450 catalysts with varying regio- and stereoselectivity properties, however, currently limits the scope of this approach. While various methods are available for high-throughput screening of P450 activity,[6] these do not provide information regarding the regio/stereoselectivity of the screened P450s, which has to be established on a case-by-case basis through laborious and time-consuming HPLC or GC analyses. In this report, we introduce a method which can address this limitation and streamline the discovery of P450s with the desired regioselectivity toward complex terpene substrates.

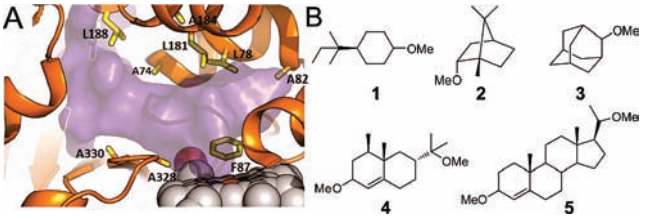

Figure 1. (A) Surface representation of the active site shape and volume of $P450_{BM3}$ (Reprinted with permission from ref 8. Copyright 2008 Elsevier. PDB 1BVY). (B) Chromogenic probes used to map the active site geometry of the engineered P450 variants.

Enzyme activity profiling across multiple substrates has been used in the context of proteases, lipases, and kinases to generate 'fingerprints' for enzyme identification and classification.[7] Our goal was to develop a fingerprinting method for P450 enzymes which could provide an indirect map of the size and geometry of their active site (Figure 1A). We envisioned the resulting fingerprint could (a) relay information regarding the accessibility of the active site in the enzyme and thus its substrate scope, and (b) enable a qualitative assessment of its regio- and stereoselectivity properties, as these are in large part dictated by the configuration of the enzyme active site.[8,9] To this end, we prepared a set of semisynthetic probes based on hydrocarbon scaffolds with marked differences in structure, size, and bulkiness (**1**−**5**, Figure 1B). 'Reporter' methoxy groups (−OCH_3) were installed on these structures to enable rapid profiling of P450 function on the probes using a Purpald-based colorimetric assay, which detects the product (formaldehyde) of P450-dependent demethylation of this functional group.[4,5] These designs were also intended to couple the screen readout with the C−H oxidation activity of the P450, as this has proven important toward the isolation of P450 catalysts which exhibit more coupled catalytic cycles and can therefore support higher substrate turnover numbers.[2,4]

To test the viability of the fingerprinting method, we tested 10 previously described variants of the fatty acid hydroxylase $P450_{BM3}$ from *Bacillus megaterium*.[10] These P450 variants exhibit varying activity and selectivity in the oxidation of non-native small-molecule substrates.[5] Parallel reactions with **1**−**5** were carried out using these P450s in purified form and a NADPH cofactor regeneration system containing a thermostable phosphite dehydrogenase (PTDH).[11] Notably, each variant was found to be associated to a unique fingerprint, supporting the

Received: October 25, 2010
Published: February 22, 2011

dx.doi.org/10.1021/ja109590h | *J. Am. Chem. Soc.* 2011, 133, 3242–3245

ability of these profiles to capture the functional differences among these P450s (Figure S1). In addition, the relative activity of each P450 variant on **1**, **2**, and **3** varied considerably in spite of the comparable size of these molecules (±20 Da), indicating that these probes can effectively report on the different geometric constraints within the active site of these enzymes.

Encouraged by these results, we investigated the utility of this method for high-throughput analysis of engineered $P450_{BM3}$ libraries in search of variants with diversified reactivity and selectivity properties toward terpene substrates. To this end, triple and quadruple mutant libraries were constructed by site-saturation mutagenesis of positions V78, S81, V82, A87, L181, and V184 in $P450_{BM3}$ variant FL#62, which exhibited high activity on all the probes (Figure S1, Table S1). Targeting these sites was expected to be most effective in altering the active site configuration of FL#62 variant as the corresponding amino acid residues project their side chains toward the heme pocket and substrate channel of the enzyme based on the available crystal structure of $P450_{BM3}$[12] (Figure 1A). The P450 libraries were expressed in 96-well plates (DH5α cells) and screened against probes **1**−**5** in parallel to acquire a fingerprint for each functional P450 variant occurring in these libraries. Reactions with the probes (1 mM, 60 min) were carried out using cell lysate in the presence of the phosphite/PTDH cofactor regeneration system, and probe activity was quantified based on absorbance at 550 nm after 60-min incubation with Purpald (30 mM). To enable comparisons among the fingerprints, the probe activities were normalized to that of a reference P450, $P450_{BM3}$(F87A), which has minimal yet detectable activity on all five probes (Figure S1). Wild-type $P450_{BM3}$ was unsuitable for this purpose as it shows no activity on three out of the five fingerprint probes (Figure S1). From the screening of ∼10 000 recombinants, a total of 1220 catalytically active P450 variants were identified (threshold: >20% parental activity on at least one probe). Comparative analysis of the normalized P450 fingerprints revealed the occurrence of 261 variants (21%) featuring a unique profile, a representative sample (25) of which is provided in Figure 2. Since the variation coefficient of the assay is 15%, a variation larger than ±20% in at least one of the five fingerprint components served as criterion to define two fingerprints as distinct. The 261 variants with unique profile were pooled to form a collection of fingerprinted P450 catalysts for the subsequent proof-of-principle studies.

Probes **1**, **2**, **4**, and **5** incorporate carbon skeletons (cyclohexane, bicyclic norbornane, bicyclic decalin, and steroid) occurring in numerous terpenes with relevant pharmacological activity[13] and of practical value for asymmetric catalysis.[14] By reflecting the geometric compatibility between the P450 active site and these molecular scaffolds, we expected the fingerprints to be useful for predicting the reactivity of the P450s toward substrates sharing the same core structure. This hypothesis was based on the observation that in most P450s substrate recognition is primarily mediated by hydrophobic and van der Waals interactions rather than directional H-bonds or ionic interactions.[15,16] In particular, the recent engineering of $P450_{BM3}$ into a propane monooxygenase illustrates how substrate-active site complementarity is sufficient to grant catalytic proficiency on a substrate (propane) which cannot be recognized through H-bonding or electrostatic interactions.[8]

To test this hypothesis, we investigated a panel of target compounds structurally related (i.e., core structure related) to the cyclohexane-based probe **1** (pentylcyclohexanol (**6**), menthol (**7**)), the norbornane-based probe **2** (borneol (**8**), camphorsultam (**9**)), and the decaline-based probe **4** (11,12-dihydronootkatone (**10**), sclareolide (**11**)) (Figure 3). The P450s in the 261-member collection were ranked according to their predicted activity on **6**−**11** using single fingerprint component analysis, where probe **1**-activity was used as predictor for **6**- and **7**-activity, probe **2**-activity as predictor for **8**- and **9**-activity, and probe **4**-activity as predictor for **10**- and **11**-activity. The 40 top-ranking P450s for each substrate pair were extracted from the collection and tested for hydroxylation activity on the target substrates in reactions in 96-well plates (∼0.1 mol % P450, KPi pH 8.0, 16 h) followed by GC analysis. For **6**, **8**, **10**, which are most closely related to the probes, activity predictions were confirmed in 87−97% of the cases (activity threshold value: 100 total turnovers, TTN), as summarized in Figure 3. Importantly, excellent rates of correct activity predictions (90−100%) were found also in the context of **7**, **9**, and **11**, which are more distantly related to the probe structures. On average, 78% of the identified P450 variants were found to support more than 400 total turnovers (57%, >750 TTN; 30%, >1000 TTN), indicating that the large majority of the identified P450 catalysts could be already useful for synthesis at preparative scale.[5] The fingerprint-based activity predictions were further tested by characterizing 10 bottom-ranking P450 variants for each pair of target substrates. These enzymes were

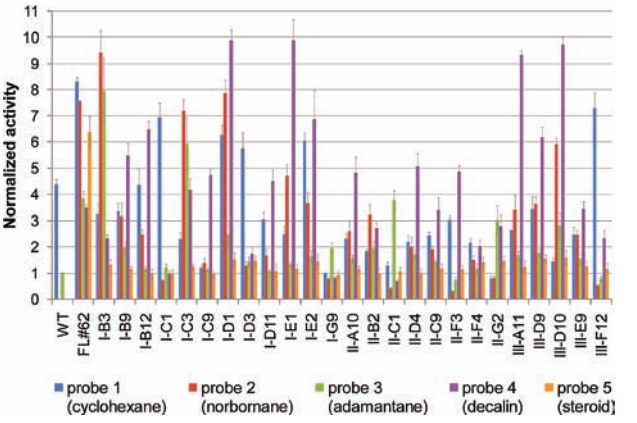

Figure 2. Fingerprints of $P450_{BM3}$ (= WT), FL#62 and 25 FL#62-derived variants from the collection of engineered P450 catalysts. Probe activities are normalized to those of the reference enzyme $P450_{BM3}$-(F87A). Mean values and standard deviations were calculated from three replicates.

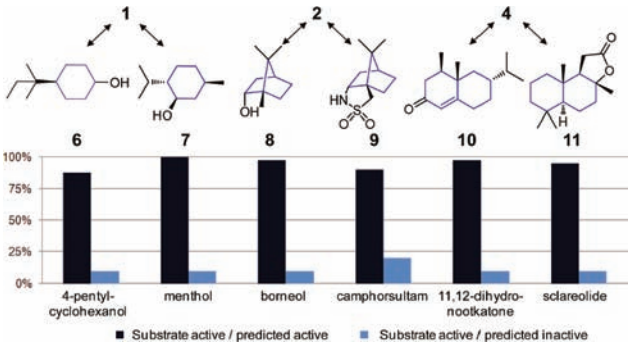

Figure 3. Fraction of catalytically active P450 variants (TTN ≥ 100) among the predicted active (dark blue) and predicted inactive (light blue) members of the P450 collection based on single fingerprint component analysis. The core structure shared by the probes and the target substrates is highlighted in blue.

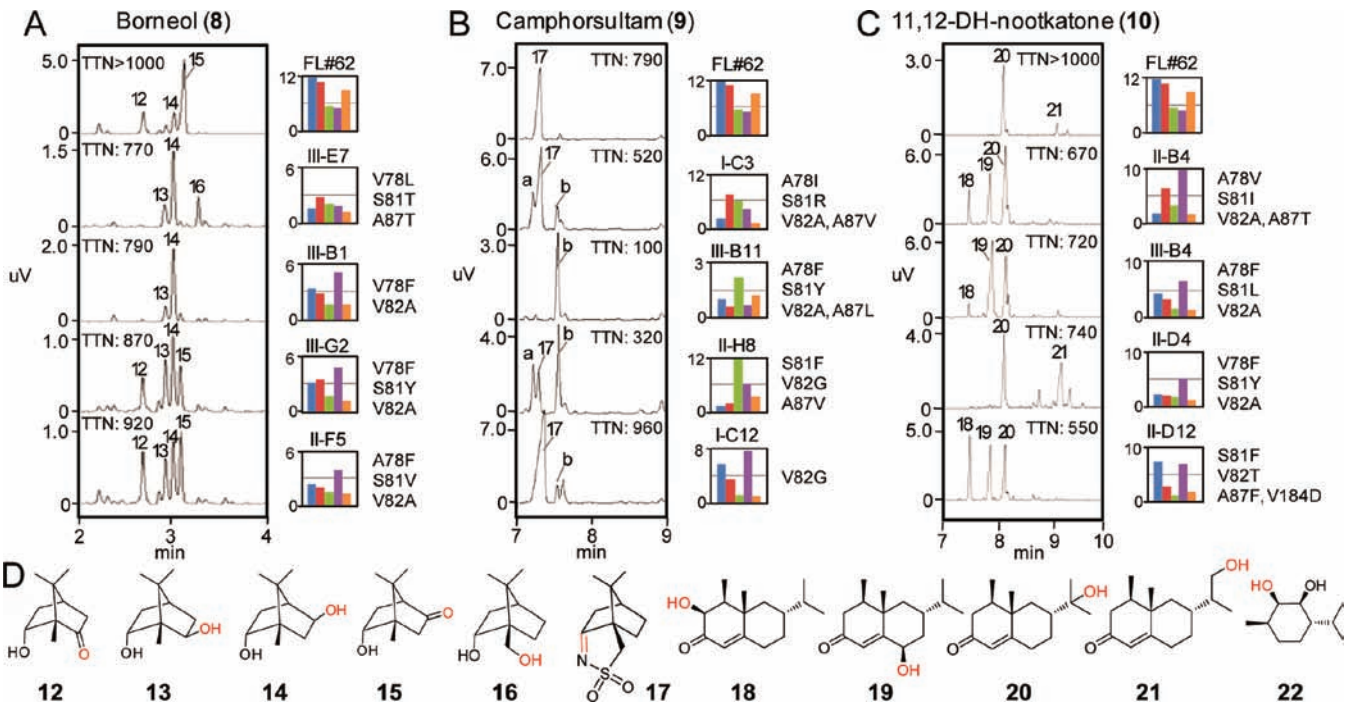

Figure 4. Product distribution (GC), total turnovers (TTN), fingerprints, and amino acid mutations (vs parent enzyme FL#62) of **8**-active (A), **9**-active (B), and **10**-active (C) variants extracted from the P450 collection. GC peak integrations are provided in Tables S2−S4. Amino acid mutations in FL#62 (vs $P450_{BM3}$) are provided in Table S1. Oxidation products **a** and **b** were not characterized. (D) Hydroxylation products isolated from the reactions with **8**, **9**, **10**, and **7**.

predicted to be inactive as their fingerprints indicated no activity on **1**, **2** or **4**, respectively. On average, 88% of these variants showed no detectable oxidation activity on the target substrate (Figure 3), further supporting the reliability of the method and indicating that the rate of 'missed predictions' is, on average, less than 15%. Altogether, these studies demonstrated the efficiency of the fingerprint analysis strategy in accelerating the discovery of synthetically useful P450 catalysts for a variety of target compounds (**6**−**11**) through simple inspection of their fingerprints.

While the individual fingerprint components were useful to predict substrate acceptance in the P450 variants, we envisioned that inspection of the whole fingerprint could provide a means to anticipate differences in the regioselectivity properties of these enzymes. In P450s, the regio/stereoselectivity of the oxidation reaction depends upon the orientation of the substrate above the distal side of the heme prior to oxidative attack by the oxo-ferryl porphyrin π-cation radical species.[16,17] This is influenced by the active site configuration of the P450 enzyme, which can be mapped through the described fingerprinting approach. To test this idea, we isolated P450 variants featuring divergent fingerprints and compared the product distribution after reaction with **6**−**11**. With the majority (5/6) of these compounds, P450s with different fingerprints exhibited also important differences in regioselectivity, as illustrated by the examples in Figure 4 (see also Tables S2−S4). This occurred with a frequency of 43% (**6**), 32% (**7**), 42% (**8**), 41% (**9**), and 45% (**10**), indicating that a large fraction (~40% on average) of the active site changes captured by fingerprinting affected the binding mode of these substrates during catalysis. With **11**, a sole oxidation product ((*S*)-3-hydroxy sclareolide) was obtained with various P450 variants, suggesting that the ability of this compound to adopt alternate orientations within the enzyme active site may be limited. To shed light on the sites targeted by hydroxylation in **6**−**10**, the major oxidation products were isolated from larger scale reactions (50 mg substrate, 0.2 mol % P450, 24 h) and their identity elucidated by 1D and 2D-NMR (^{1}H−^{1}H COSY, HMBC, HSQC, NOESY). The data corresponding to **8** (borneol) and **10** (11,12-dihydronootkatone) are of particular interest because they highlight two key aspects of the overall approach. First, the isolated P450 variants were found to target, collectively, 40% of the sp^3 C−H sites occurring in these compounds (3/7 and 4/10, respectively), including tertiary, secondary, and even primary positions as indicated in Figure 4D. While the selectivity of these variants for some of these positions is moderate (<40%, Tables S2−S4), these results are remarkable considering the small number of variants tested (40) and their straightforward identification through fingerprinting and fingerprint analysis. Another important finding was that P450-catalyzed hydroxylation in these substrates occurred also at positions which are remote with respect to the reporter functional group in the corresponding probe (e.g., products **13**, **14**, **19** in Figure 4). Altogether, these studies demonstrate the ability of the described fingerprint-based method not only to expedite the search of P450 oxidation catalysts with diversified regioselectivity but also to enable the discovery of P450s useful for targeting aliphatic positions across the whole carbon skeleton of terpenes structurally related to the fingerprint probes.

Given the observed relationship between fingerprint and regioselectivity, we anticipated that two variants sharing a similar fingerprint would display similar selectivity in substrate oxidation. To test this hypothesis, we isolated three variants from the 78/81/87 library (5-G9, 5-C4, 5-C2) possessing an identical fingerprint. Characterization of these variants revealed that these enzymes exhibit remarkably similar product profiles across multiple substrates (Figure 5, Table S5). While these P450s differ from each other by up

dx.doi.org/10.1021/ja109590h |*J. Am. Chem. Soc.* 2011, 133, 3242–3245

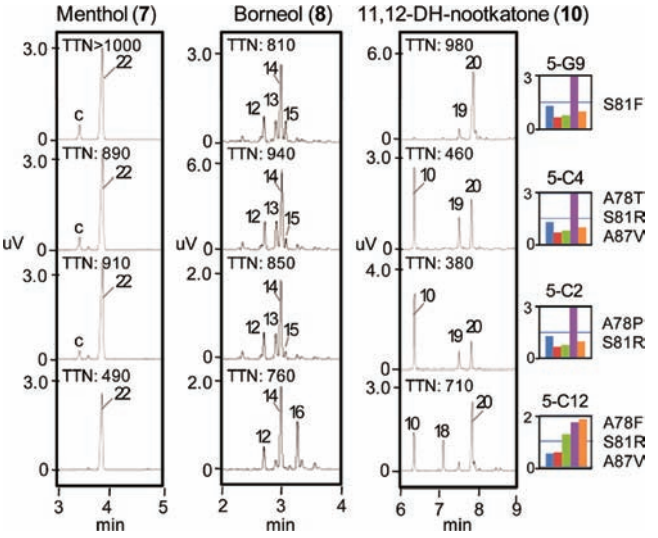

Figure 5. Product distribution (GC), TTNs, fingerprints, and amino acid mutations (vs parent enzyme FL#62) of three P450 variants sharing an identical fingerprint and one having a different fingerprint from the same library. GC peak integrations are provided in Table S5. Oxidation product **c** was not characterized.

to three amino acid substitutions in their active sites, these mutations have apparently resulted in equivalent active site geometries. Notably, this feature could be captured through inspection of their fingerprints. Thus, we conclude that fingerprint analysis could also offer a convenient strategy to identify, within a library of engineered P450s, those that possess a desired type of regioselectivity, once the corresponding fingerprint is known. The total probe activity would provide then a way to recognize, among these variants, the one(s) which are more catalytically efficient as this parameter is related to the number of turnovers supported by the enzyme. For comparison, another P450 variant was extracted from the same library (5-C12) and found to differ from 5-C4 by a single amino acid (V78 vs F78). Interestingly, this single amino acid substitution causes a remarkable change in the active site configuration of the enzyme as evinced from the difference in the product distribution with **7**, **8**, and **10** and as it could be anticipated from comparison of the respective fingerprints (Figure 5).

In summary, we have reported an efficient and time-effective method to gain insights regarding the reactivity and regioselectivity properties of a P450 enzyme. This approach is amenable to the high-throughput screening of engineered P450 libraries and offers the unprecedented capability to allow for the rapid identification of P450 variants with diversified regioselectivity as well as variants with a specific type of regioselectivity (via fingerprint comparison) within such libraries. We expect this method to prove valuable in the development of P450 catalysts for the functionalization of unactivated C−H bonds in complex terpenes via P450-mediated synthesis.

■ ASSOCIATED CONTENT

Supporting Information. Experimental details and characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
fasan@chem.rochester.edu

■ ACKNOWLEDGMENT

We thank Frances H. Arnold for kindly providing the $P450_{BM3}$ variants used in the preliminary experiments (Figure S1) and Huimin Zhao for kindly providing vectors for the expression of the thermostable PTDH. This work was supported by startup funds from the University of Rochester.

■ REFERENCES

(1) (a) Cook, B. R.; Reinert, T. J.; Suslick, K. S. *J. Am. Chem. Soc.* **1986**, *108*, 7281. (b) Das, S.; Incarvito, C. D.; Crabtree, R. H.; Brudvig, G. W. *Science* **2006**, *312*, 1941. (c) Yang, J.; Gabriele, B.; Belvedere, S.; Huang, Y.; Breslow, R. *J. Org. Chem.* **2002**, *67*, 5057. (d) Wender, P. A.; Hilinski, M. K.; Mayweg, A. V. *Org. Lett.* **2005**, *7*, 79. Chen, M. S.; White, M. C. *Science* **2007**, *318*, 783. (e) Lee, S.; Fuchs, P. L. *J. Am. Chem. Soc.* **2002**, *124*, 13978. (f) Brodsky, B. H.; Du Bois, J. *J. Am. Chem. Soc.* **2005**, *127*, 15391. (g) Chen, M. S.; White, M. C. *Science* **2010**, *327*, 566.

(2) Fasan, R.; Chen, M. M.; Crook, N. C.; Arnold, F. H. *Angew. Chem., Int. Ed.* **2007**, *46*, 8414.

(3) (a) Lewis, J. C.; Bastian, S.; Bennett, C. S.; Fu, Y.; Mitsuda, Y.; Chen, M. M.; Greenberg, W. A.; Wong, C. H.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 16550. (b) Li, S.; Chaulagain, M. R.; Knauff, A. R.; Podust, L. M.; Montgomery, J.; Sherman, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 18463. (c) Whitehouse, C. J.; Bell, S. G.; Tufton, H. G.; Kenny, R. J.; Ogilvie, L. C.; Wong, L. L. *Chem. Commun. (Cambridge, U.K.)* **2008**, 966. (d) Zehentgruber, D.; Hannemann, F.; Bleif, S.; Bernhardt, R.; Lutz, S. *ChemBioChem* **2010**, *11*, 713. Sun, L.; Chen, C. S.; Waxman, D. J.; Liu, H.; Halpert, J. R.; Kumar, S. *Arch. Biochem. Biophys.* **2007**, *458*, 167. (e) Liu, L.; Schmid, R. D.; Urlacher, V. B. *Biotechnol. Lett.* **2010**, *32*, 841. (f) Bottner, B.; Schrauber, H.; Bernhardt, R. *J. Biol. Chem.* **1996**, *271*, 8028. (g) Tang, W. L.; Li, Z.; Zhao, H. *Chem. Commun. (Cambridge, U.K.)* **2010**, *46*, 5461.

(4) Peters, M. W.; Meinhold, P.; Glieder, A.; Arnold, F. H. *J. Am. Chem. Soc.* **2003**, *125*, 13442.

(5) Rentmeister, A.; Arnold, F. H.; Fasan, R. *Nat. Chem. Biol.* **2009**, *5*, 26.

(6) Rabe, K. S.; Gandubert, V. J.; Spengler, M.; Erkelenz, M.; Niemeyer, C. M. *Anal. Bioanal. Chem.* **2008**, *392*, 1059.

(7) (a) Goddard, J. P.; Reymond, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 11116. (b) Basile, F.; Ferrer, I.; Furlong, E. T.; Voorhees, K. J. *Anal. Chem.* **2002**, *74*, 4290. (c) Maly, D. J.; Huang, L.; Ellman, J. A. *ChemBioChem* **2002**, *3*, 16.

(8) Fasan, R.; Meharenna, Y. T.; Snow, C. D.; Poulos, T. L.; Arnold, F. H. *J. Mol. Biol.* **2008**, *383*, 1069.

(9) (a) Oliver, C. F.; Modi, S.; Primrose, W. U.; Lian, L.; Roberts, G. C. K. *Biochem. J.* **1997**, *327*, 537. (b) Branco, R. J.; Seifert, A.; Budde, M.; Urlacher, V. B.; Ramos, M. J.; Pleiss, J. *Proteins* **2008**, *73*, 597.

(10) Narhi, L. O.; Fulco, A. J. *J. Biol. Chem.* **1987**, *262*, 6683.

(11) McLachlan, M. J.; Johannes, T. W.; Zhao, H. *Biotechnol. Bioeng.* **2008**, *99*, 268.

(12) Sevrioukova, I. F.; Li, H.; Zhang, H.; Peterson, J. A.; Poulos, T. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1863.

(13) Fraga, B. M. *Nat. Prod. Rep.* **2003**, *20*, 392.

(14) (a) Gayet, A.; Bolea, C.; Andersson, P. G. *Org. Biomol. Chem.* **2004**, *2*, 1887. (b) Garcia, C.; LaRochelle, L. K.; Walsh, P. J. *J. Am. Chem. Soc.* **2002**, *124*, 10970. (c) Kumara, S.; Sobhiab, M. E.; Ramachandran, U. *Tetrahedron: Asymmetry* **2005**, *16*, 2599.

(15) (a) Meharenna, Y. T.; Li, H.; Hawkes, D. B.; Pearson, A. G.; De Voss, J.; Poulos, T. L. *Biochemistry* **2004**, *43*, 9487. (b) Yao, H.; McCullough, C. R.; Costache, A. D.; Pullela, P. K.; Sem, D. S. *Proteins* **2007**, *69*, 125. (c) Xu, L. H.; Fushinobu, S.; Takamatsu, S.; Wakagi, T.; Ikeda, H.; Shoun, H. *J. Biol. Chem.* **2010**, *285*, 16844. (d) Denisov, I. G.; Makris, T. M.; Sligar, S. G.; Schlichting, I. *Chem. Rev.* **2005**, *105*, 2253.

(16) Pylypenko, O.; Schlichting, I. *Annu. Rev. Biochem.* **2004**, *73*, 991.

(17) Sono, M.; Roach, M. P.; Coulter, E. D.; Dawson, J. H. *Chem. Rev.* **1996**, *96*, 2841.