JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Discovery and Characterization of Heme Enzymes from Unsequenced Bacteria: Application to Microbial Lignin Degradation

Margaret E. Brown,[†] Mark C. Walker,[‡] Toshiki G. Nakashige,[†] Anthony T. Iavarone,[§] and Michelle C. Y. Chang*[,†,‡]

[†]Departments of Chemistry and [‡]Molecular and Cell Biology, University of California, Berkeley, California 94720-1460, United States
[§]The QB3/College of Chemistry Mass Spectrometry Facility, Berkeley, California 94720-3220, United States

Supporting Information

ABSTRACT: Bacteria and other living organisms offer a potentially unlimited resource for the discovery of new chemical catalysts, but many interesting reaction phenotypes observed at the whole organism level remain difficult to elucidate down to the molecular level. A key challenge in the discovery process is the identification of discrete molecular players involved in complex biological transformations because multiple cryptic genetic components often work in concert to elicit an overall chemical phenotype. We now report a rapid pipeline for the discovery of new enzymes of interest from unsequenced bacterial hosts based on laboratory-scale methods for the de novo assembly of bacterial genome sequences using short reads. We have applied this approach to the biomass-degrading soil bacterium *Amycolatopsis* sp. 75iv2 ATCC 39116 (formerly *Streptomyces setonii* and *S. griseus* 75vi2) to discover and biochemically characterize two new heme proteins comprising the most abundant members of the extracellular oxidative system under lignin-reactive growth conditions.

Living organisms have used the power of evolution to innovate new and remarkable chemical transformations that not only allow them to compete successfully with rivals for environmental resources but can also address synthetic challenges of interest to human society. Indeed, the activation of small molecules such as carbon dioxide, oxygen, or nitrogen for downstream multistep reactions takes advantage of the high availability of gaseous substrates to fix C, N, and O for cell growth and is also a longstanding focus of development of homogeneous and heterogeneous chemical catalysts.[1−3] One of the most important natural carbon-based feedstocks is plant biomass, which represents a large renewable resource for industrial production of fuels and chemicals with 10−50 billion tons produced globally per year.[4] However, the majority of the carbon in plant material remains trapped in structural biopolymers comprised of cellulose, hemicellulose, and lignin rather than easily accessed sugars and starches (Figure 1A).[5] Lignin is the most chemically difficult of these feedstocks to approach because of its aromatic monomer units and diverse set of C−C and C−O cross-links.[6] Its high structural complexity consequently requires multiple families of enzymes to achieve its breakdown.[7,8] The current model for lignin breakdown is derived from studies of fungal metalloenzymes and is based on an oxidative combustion of lignin through

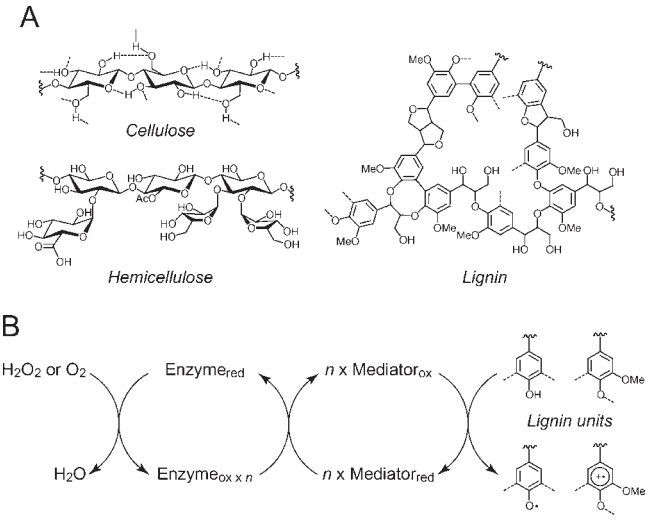
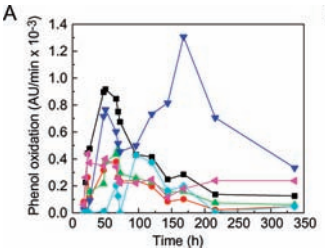


Figure 1. (A) Representative structural units of the biopolymers found in lignocellulosic biomass, which is made up of cellulose (∼35−60% of biomass), hemicellulose (∼20−35% of biomass), and lignin (∼10−25% of biomass). (B) Oxidative reactions initiating the radical-dependent breakdown of lignin. Small-molecule mediators, including veratryl alcohol, Mn(II), and oxalate, are oxidized directly by fungal peroxidases and laccases, which utilize H_2O_2 or O_2, respectively. The oxidizing equivalent is then transferred from the mediator to the lignin substrate to generate either a phenoxyl or phenyl radical equivalent, leading to bond scission events at the site of oxidation and eventual lignin depolymerization.

various radical-mediated paths that ultimately generate phenoxyl and phenyl radicals on the substrate followed by depolymerization (Figure 1B).[7−9] While this overall oxidative strategy is most likely conserved, it is also known that bacteria exhibit quite different reactivity toward lignin compared to fungi based on their altered product distribution, but these enzymes have yet to be biochemically characterized.[10−15] In order to explore new chemical strategies for lignin degradation, we have initiated a program aimed at discovery and characterization of oxidative and accessory enzymes in lignin-reactive soil bacteria, which exhibit particularly rich activity toward the depolymerization and utilization of biomass-derived carbon sources.[16] In this report, we present a pipeline for rapid discovery and functional identification

Received: April 29, 2011
Published: June 14, 2011

dx.doi.org/10.1021/ja203972q | *J. Am. Chem. Soc.* 2011, 133, 18006–18009



Bacterium	Miscanthus APPL (mg per 10 mL)
Amycolatopsis sp. 75iv2	2.6
Streptomyces viridosporus	0.6
Streptomyces griseus	1.6
Streptomyces badius	1.3
Rhodococcus opacus	1.5
Arthrobacter chlorophenolicus	0.5

Figure 2. Characterization of lignin reactivity in several species of soil bacteria. (A) Monitoring extracellular peroxidase activity from bacterial cultures grown on *M. giganteus* lignocellulose as a substrate [*A.* sp. 75iv2, (blue ▼); *S. viridosporus*, (■); *S. griseus*, (green ▲); *S. badius*, (red ●); *R. opacus*, (turquoise ◆); *A. chlorophenolicus*, (pink, tilted solid triangle)]. Peroxidase activity was measured spectrophotometrically by trapping of the 2,4-dichlorophenoxyl radical with 4-aminoantipyrene as a colored adduct. (B) Mass of APPL produced after 14 d of cell culture.

of new enzymes from unsequenced bacteria under lignin-reactive growth conditions, an approach that can be applied to any culturable microbe of interest.

To identify bacteria with robust lignin reactivity, we selected several species of soil bacteria that have been reported to exhibit lignin-degrading or lignolytic-like ability—*Amycolatopsis* sp. 75iv2 (formerly named *Streptomyces setonii* and *S. griseus* 75vi2),[17,18] *S. viridosporus*,[13,14,17,19] *S. griseus*,[20] *S. badius*,[13,19] *Arthrobacter cholorophenolicus*,[21] and *Rhodococcus opacus*.[15,22]—for culture in the presence of a lignocellulosic carbon source from a perennial grass (*Miscanthus giganteus*). As phenol oxidation is a key reaction in the cleavage of lignin cross-links, we tracked the peroxidase-mediated phenoxyl radical generation capacity in the extracellular protein fraction (Figure 2A) in addition to the production of acid-precipitable polymeric lignin (APPL), which is a bacteria-specific lignin degradation product (Figure 2B).[12,13] From these studies, we determined that *Amycolatopsis* sp. 75iv2 appeared to be the most active strain by both standards and turned our attention to developing an efficient lab-scale pipeline for identifying candidate ligninase and accessory enzymes from an unsequenced microbial host using a combination of next-generation sequencing and functional proteomics.

Using massively parallel sequencing based on the Illumina platform, we assembled an inexpensive de novo draft genome sequence for *A.* sp. 75iv2 from paired-end 84-base reads (Supporting Information [SI]). Assembly of the short reads using Velvet[23] and Image[24] produced a genome that spans 168 contigs with an N_{50} of 209 kB and whose size falls within the estimated range for closely related actinomycetes (8–10 MB)[20,25−28] (SI). Its recent reclassification from *S. setonii* or *S. griseus* 75vi2 into the *Amycolatopsis* genus is consistent with the observed genetic cell-wall markers comprising the absence of genes for the incorporation of glycine cross-linkers in the peptidoglycan layer and biosynthesis of mycolic acids as well as *murE* phylogeny (SI).[28] *A.* sp. 75iv2 also maintains a rich collection of genes for biomass deconstruction with many predicted secreted enzymes within the glycosyl hydrolase superfamily and a considerable number of clusters for aromatic degradation through various meta- and ortho-ring cleavage pathways (SI).[29] With regard to lignin reactivity, *A.* sp. 75iv2 possesses several potential oxidative systems based on the presence of predicted secreted laccases, peroxidases, and peroxide-generating enzymes (SI).

Given the larger size of the substrate compared to cell dimensions, most enzymes involved in biomass deconstruction

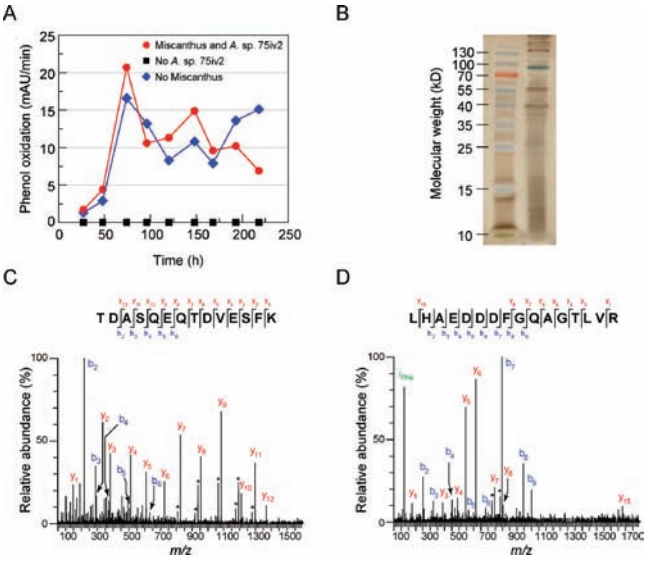
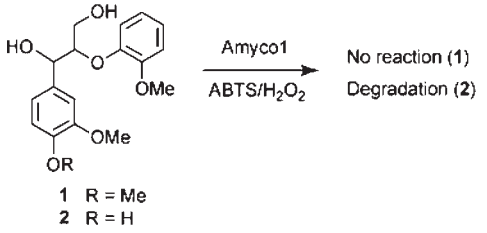


Figure 3. (A) Time course for *A.* sp. 75iv2 growth monitoring extracellular peroxidase activity by oxidation of 2,4-dichlorophenol (DCP). (B) Overlaid heme- and silver-stained SDS-PAGE gel showing the extracellular protein fraction at 3 d. Amyco1 and Amyco2 appear as a heme-stained doublet. (C) Representative tandem mass spectrum of a tryptic peptide derived from Amyco1, a heme peroxidase identified in the secretome of *A.* sp. 75iv2. (D) Representative tandem mass spectrum of a tryptic peptide derived from the Amyco2 catalase.

are thought to reside in the extracellular environment. We therefore carried out a functional secretome analysis of *A.* sp. 75iv2 in order to identify the enzymes that form the extracellular oxidative system. The extracellular peroxidase activity was tracked over 10 d in order to find the peak production time points in cultures grown in the presence of *M. giganteus* lignocellulose (Figure 3A). A combination of silver- and heme-staining indicated that the secreted protein profile was relatively simple and that one-dimensional separation was sufficient to isolate individual proteins (Figure 3B). Peptide sequences from in-gel tryptic digests of the heme-stained protein bands from different time points and growths repeatedly confirmed the presence of two major extracellular hemoproteins, which were named Amyco1 and Amyco2 (Figure 3CD). A full proteomic analysis of the secretome also indicated that Amyco1 and Amyco2 were the two most likely extracellular heme protein candidates on the basis of their relatively high abundance and that laccases or dye-decolorizing peroxidases (DyPs), which are known for their ability to degrade aromatics, were not detected in the extracellular protein fraction under these growth conditions (SI). The genes encoding Amyco1 and Amyco2 were cloned from *A.* sp. 75iv2 genomic DNA, heterologously expressed in *Escherichia coli*, and purified to homogeneity. Characterization by UV−visible spectroscopy and inductively coupled plasma-optical emission spectrometry (ICP-OES) showed that both proteins contain an iron-porphyrin cofactor (SI). The functions of Amyco1 and Amyco2 were assigned respectively as bifunctional catalase−peroxidase and catalase based on sequence homology and were validated using in vitro biochemical assays (SI). As a catalase, Amyco2 may be involved in protection against oxidative damage to the host organism. Indeed, the white-rot fungus, *Phanerochaete chrysosporium*, demonstrates among the highest activities for peroxide-dependent lignin degradation and continually secretes catalases during cell growth.[30]

Scheme 1. Reactivity of Amyco1 with Lignin Model Dimers 1 and 2



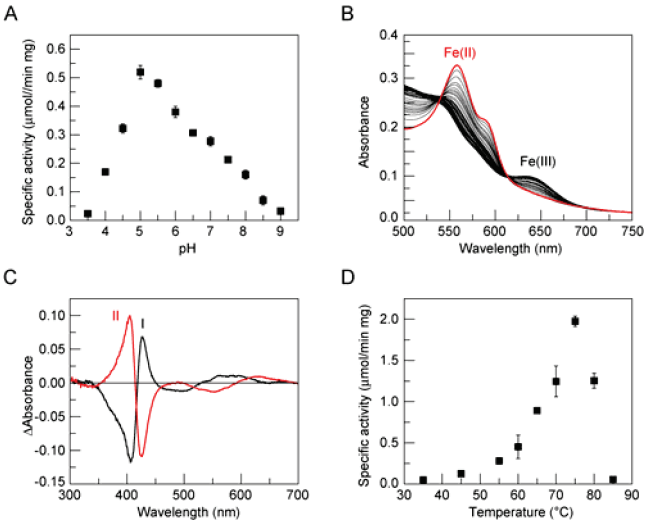


Figure 4. Biochemical characterization of Amyco1. (A) pH−rate profile for Amyco1 for DCP oxidation. Data are mean ± s.d. (n = 3). (B) Changes in the electronic absorption spectrum indicating stepwise reduction of the ferric state to the ferrous state during redox titration. (C) Difference spectra indicating formation of species assigned to compound I- and compound II-like states formed upon rapid mixing with H_2O_2 at 5 °C. (D) Temperature−rate profile for an Amyco1 orthologue from *Geobacillus kaustophilus* monitoring DCP oxidation. Data are mean ± s.d. (n = 3).

On the basis of the phenol oxidation reactivity of Amyco1, we carried out additional experiments for its characterization. Reaction of purified Amyco1 with synthetic lignin model dimers **1** and **2** (Scheme 1) indicate that Amyco1 is competent to carry out degradation of lignin at unprotected phenolic sites but not fully protected methoxylated sites (SI). Further biochemical characterization indicates that Amyco1 exists as a homodimer based on size-exclusion chromatography, which measures an apparent molecular mass of 194 kD (SI). Additional studies show that the pH−rate profile for Amyco1 is centered at pH 5, which is consistent with the lower pH of soil in the presence of organic acids produced by decaying plant matter and the low pH optimums for fungal lignin peroxidases (Figure 4A).[31] The midpoint potential for the Fe(III)−Fe(II) couple of Amyco1 (−171 mV) was measured by potentiometric redox titration and found to fall within the range of other heme peroxidases (Figure 4B).[32] Rapid mixing of heme-reconstituted Amyco1 with excess H_2O_2 captured the formation of two species by stopped-flow UV−visible spectroscopy whose spectral characteristics[33] led to their assignment as the two- and one-electron oxidized compound I- and compound II-like states, respectively (Figure 4C), which supports a model in which each Amyco1 active site is competent to oxidize two equivalents of phenol per reaction cycle. A homology search through sequenced metagenomes and genomes revealed that closely related Amyco1 orthologues are widely distributed through bacteria and archaea. A gene encoding a thermophilic Amyco1 orthologue from *Geobacillus kaustophilus* was then synthesized and its peroxide-dependent phenol oxidation activity was confirmed by biochemical characterization of the enzyme, which showed a temperature optimum near the growth temperature of the host organism (Figure 4D).

To close, we have described a general pipeline to rapidly discover enzyme candidates involved in multistep transformations of complex substrates in whole organisms. We have applied this strategy to studies of lignin reactivity and identified and characterized two new extracellular heme proteins from *A.* sp. 75iv2, an unsequenced soil bacterium valued for its ability to degrade lignocellulose and produce APPL. Amyco1 is competent for phenol rather than aromatic ring oxidation, but analysis of the genome of *A.* sp. 75iv2 shows that it contains potential demethylases for lignin-related compounds,[34] which could act synergistically for the degradation of biomass by uncapping new phenolic sites. In addition to functionally identified candidates, we can now also begin to explore the function and reactivity of the full oxidative system of *A.* sp 75iv2, including laccases and DyPs, and their synergy with other enzyme families.

■ ASSOCIATED CONTENT

Supporting Information. Experimental materials and methods. Data that include genome assembly statistics, genome analysis, and full secretome analysis for *A.* sp. 75iv2 as well as additional biochemical characterization data for Amyco1 and Amyco2. Complete references for 25 and 28. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
mcchang@berkeley.edu

■ ACKNOWLEDGMENT

We thank the Energy Biosciences Institute (#OO0J06) for funding this research. M.E.B. acknowledges the John Stauffer Fund for a predoctoral fellowship. M.C.W. and T.G.N. thank the National Institutes of Health NRSA Training Grant (1 T32 GMO66698) and the College of Chemistry Summer Undergraduate Research Program, respectively, for support. We also thank Michael Winter and Emily Derbyshire in the Marletta group for help with redox titrations and stopped-flow kinetic experiments, Stefan Bauer (EBI) for lignocellulose samples as well as John Cronan (UIUC), John Gerlt (UIUC), Charles Nicolet (UC Davis Genome Center), and Will Beeson (UCB) for helpful discussions.

■ REFERENCES

(1) (a) Benson, E. E.; Kubiak, C. P.; Sathrum, A. J.; Smieja, J. M. *Chem. Soc. Rev.* **2009**, *38*, 89–99. (b) Sakakura, T.; Choi, J.-C.; Yasuda, H. *Chem. Rev.* **2007**, *107*, 2365–2387.

(2) (a) Laplaza, C. E.; Cummins, C. C. *Science* **1995**, *268*, 861–863. (b) Yandulov, D. V.; Schrock, R. R. *Science* **2003**, *301*, 76–78.
(3) (a) Feig, A. L.; Lippard, S. J. *Chem. Rev.* **1994**, *94*, 759–805. (b) Groves, J. T. *J. Inorg. Biochem.* **2006**, *100*, 434–447. (c) Que, L.; Tolman, W. B. *Nature* **2008**, *455*, 333–340.
(4) Greene, N. *Growing energy: How biofuels can help end America's oil dependence*; Natural Resources Defense Council: Washington, DC, 2004; pp 1−86.
(5) Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. *Science* **2007**, *315*, 804–807.
(6) Ralph, J.; Lundquist, K.; Brunow, G.; Lu, F.; Kim, H.; Schatz, P. F.; Marita, J. M.; Hatfield, R. D.; Ralph, S. A.; Christensen, J. H.; Boerjan, W. *Phytochem. Rev.* **2004**, *3*, 29–60.
(7) Kirk, T. K.; Farrell, R. L. *Annu. Rev. Microbiol.* **1987**, *41*, 465–505.
(8) ten Have, R.; Teunissen, P. J. *Chem. Rev.* **2001**, *101*, 3397–413.
(9) (a) Khindaria, A.; Grover, T. A.; Aust, S. D. *Biochemistry* **1995**, *34*, 6020–6025. (b) Khindaria, A.; Barr, D. P.; Aust, S. D. *Biochemistry* **1995**, *34*, 7773–7779.
(10) Kirby, R. Actinomycetes and lignin degradation. In *Advances in Applied Microbiology*; Academic Press: London, UK, 2005.
(11) Crawford, D. L. *Appl. Environ. Microbiol.* **1978**, *35*, 1041–1045.
(12) Crawford, D. L.; Pometto, A. L., III; Crawford, R. L. *Appl. Environ. Microbiol.* **1983**, *45*, 898–904.
(13) Borgmeyer, J. R.; Crawford, D. L. *Appl. Environ. Microbiol.* **1985**, *49*, 273–278.
(14) Thomas, L.; Crawford, D. L. *Can. J. Microbiol.* **1998**, *44*, 364–372.
(15) Ahmad, M.; Taylor, C. R.; Pink, D.; Burton, K.; Eastwood, D.; Bending, G. D.; Bugg, T. D. H. *Mol. BioSyst.* **2010**, *6*, 815–821.
(16) (a) Balser, T. C.; Kinzig, A. P.; Firestone, M. K. Linking soil microbial communities and ecosystem functioning. In *The functional consequences of biodiversity: empirical progress and theoretical extensions*; Princeton University Press: Princeton, NJ, 2002. (b) Torsvik, V.; Øvreås, L. *Curr. Opin. Microbiol.* **2002**, *5*, 240–245. (c) Hättenschwiler, S.; Tiunov, A. V.; Scheu, S. *Annu. Rev. Ecol. Evol. Syst.* **2005**, *36*, 191–218.
(17) Antai, S. P.; Crawford, D. L. *Appl. Environ. Microbiol.* **1981**, *42*, 378–380.
(18) Strandberg, G. W.; Lewis, S. N. *J. Ind. Microbiol. Biotechnol.* **1987**, *1*, 371–375.
(19) Giroux, H.; Vidal, P.; Bouchard, J.; Lamy, F. *Appl. Environ. Microbiol.* **1988**, *54*, 3064–3070.
(20) Ohnishi, Y.; Ishikawa, J.; Hara, H.; Suzuki, H.; Ikenoya, M.; Ikeda, H.; Yamashita, A.; Hattori, M.; Horinouchi, S. *J. Bacteriol.* **2008**, *190*, 4050–4060.
(21) (a) Kerr, T. J.; Kerr, R. D.; Benner, R. *Appl. Environ. Microbiol.* **1983**, *43*, 1201–1206. (b) Westerberg, K.; Elvang, A. M.; Stackebrandt, E.; Jansson, K. *Int. J. Syst. Evol. Microbiol.* **2000**, *50*, 2083–2092.
(22) Trojanowski, J.; Haider, K.; Sundman, V. *Arch. Microbiol.* **1977**, *114*, 149–153.
(23) Zerbino, D. R.; Birney, E. *Genome Res.* **2008**, *18*, 821–829.
(24) Li, H.; Durbin, R. *Bioinformatics* **2009**, *25*, 1754–1760.
(25) Bentley, S. D.; et al. *Nature* **2002**, *417*, 141–147.
(26) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. *Nat. Biotechnol.* **2003**, *21*, 526–531.
(27) Ishikawa, J.; Yamashita, A.; Mikami, Y.; Hoshino, Y.; Kurita, H.; Hotta, K.; Shiba, T.; Hattori, M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14925–14930.
(28) Zhao, W.; et al. *Cell Res.* **2010**, *20*, 1096–1108.
(29) (a) Franklin, F. C.; Bagdasarian, M.; Bagdasarian, M. M.; Timmis, K. N. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 7458–7462. (b) Masai, E.; Katayama, Y.; Nishikawa, S.; Fukuda, M. *J. Ind. Microbiol. Biotechnol.* **1999**, *23*, 364–373. (c) Jiménez, J. I.; Minambres, B.; Garcia, J. L.; Diaz, E. *Environ. Microbiol.* **2002**, *4*, 824–841.
(30) Kwon, S.-I.; Anderson, A. J. *Curr. Microbiol.* **2001**, *42*, 8–11.
(31) Tien, M.; Kirk, T. K.; Bull, C.; Fee, J. A. *J. Biol. Chem.* **1986**, *261*, 1687–1693.
(32) (a) Millis, C. D.; Cai, D.; Stankovich, M. T.; Tien, M. *Biochemistry* **1989**, *28*, 8484–8489.(b) Dunford, B. H. *Heme Peroxidases*; Wiley-VCH: Weinheim, Germany, 1999.
(33) (a) Chouchane, S.; Lippai, I.; Magliozzo, R. S. *Biochemistry* **2000**, *39*, 9975–9983. (b) Yu, S.; Girotto, S.; Zhao, X.; Magliozzo, R. S. *J. Biol. Chem.* **2003**, *278*, 44121–44127. (c) Ghiladi, R. A.; Knudsen, G. M.; Medzihradszky, K. F.; Ortiz de Montellano, P. R. *J. Biol. Chem.* **2005**, *280*, 22651–63.
(34) Sutherland, J. B. *Appl. Environ. Microbiol.* **1986**, *52*, 98–100.

NOTE ADDED AFTER ASAP PUBLICATION

In the version of this Communication published ASAP June 14, 2011, Figure 3D showed a tandem mass spectrum of a second tryptic peptide derived from Amyco1 rather than one from Amyco2. A new version with the correct figure was posted August 22, 2011.