Directed evolution of the peroxidase activity of a *de novo*-designed protein

Shona C. Patel[1,3] and Michael H. Hecht[2,4]

[1]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544, USA, [2]Department of Chemistry, Princeton University, Princeton, NJ 08544, USA and [3]Present address: Merck, Summit, NJ 07901, USA

[4]To whom correspondence should be addressed.
E-mail: hecht@princeton.edu

Collections of *de novo*-designed proteins provide a unique opportunity to probe the functional potential of sequences that are stably folded, but were neither explicitly designed nor evolutionarily selected to perform any particular type of activity. A combinatorial library of folded proteins was designed previously using a strategy that exploits the binary patterning of polar and non-polar amino acids to drive sequences to fold into four-helix bundles. Although these novel four-helix bundles were not explicitly designed for function, previous characterization of several hundred arbitrarily chosen sequences showed that many of them bound the heme cofactor, and several of these novel heme proteins catalyzed peroxidase activity at levels substantially above background. Here, we show that these moderately active *de novo* heme proteins can serve as non-natural starting points for laboratory-based evolution: Random mutagenesis followed by color-based screening of a relatively small number (hundreds or thousands) of variants yielded novel sequences with improved peroxidase activity. Biochemical characterization of the purified proteins showed that the evolved variants were nearly 3-fold more active than the parental sequence. These results demonstrate that *de novo*-designed proteins can be utilized as a novel feedstock for the evolution of enzyme activity.

Keywords: binary code/directed evolution/peroxidase/protein design/synthetic biology

Introduction

Designing proteins *de novo* provides a unique opportunity to explore sequences, structures and functions that are chemically reasonable, but which have not been sampled by biological evolution. In the initial steps toward such explorations, we designed combinatorial libraries of α-helical sequences, determined the three-dimensional structures of several proteins from these libraries, and probed their potential for binding and catalytic function (Kamtekar *et al.*, 1993; Rojas *et al.*, 1997; Moffet *et al.*, 2000; Wei *et al.*, 2003a,b; Hecht *et al.*, 2004; Go *et al.*, 2008; Patel *et al.*, 2009; Fisher *et al.*, 2011; Arai *et al.*, 2012; Cherny *et al.*, 2012).

Combinatorial libraries of α-helical proteins were constructed using the binary code strategy for protein design. A canonical alpha helix has 3.6 residues per turn. Accordingly, a non-polar amino acid is placed every 3 or 4 positions in the primary sequence, generating the following pattern: ○ ● ○○ ● ● ○○ ● ○○ ● ● ○, where ○ and ● represent arbitrarily chosen polar and non-polar residues, respectively (Kamtekar *et al.*, 1993; Hecht *et al.*, 2004). This pattern encodes an amphipathic helix, where all non-polar residues align on one face of the helix and all polar residues align on the opposite face. When four such helices are linked, the hydrophobic effect drives them to pack against one another, thereby forming a four-helix bundle with non-polar residues pointing toward the protein core, and polar residues exposed to solvent. Several proteins from previously constructed libraries have been shown to form stable well-ordered structures (Wei *et al.*, 2003a). Furthermore, determination of the structures of two proteins (S-824 and S-836) provided evidence that these binary-patterned proteins folded into four-helix bundles as designed (Wei *et al.*, 2003b; Go *et al.*, 2008) (Fig. 1).

Although the proteins in our binary-patterned libraries were not explicitly designed for any particular activities, characterization of several hundred arbitrarily chosen sequences showed that many of them bound the heme cofactor, and several of these novel heme proteins catalyzed peroxidase activity at levels substantially above background (Patel *et al.*, 2009).

The observation of cofactor binding and catalytic activity in proteins from a library of α-helical bundles that had neither been designed nor selected (*in vivo* or *in vitro*) for function suggested that this superfamily of *de novo* proteins could be viewed as a novel 'feedstock' for evolution. Thus, by starting with a binary-patterned protein with relatively low-level activity, one can mimic natural selection by introducing random mutations and screening for variants with improved activity. This process is reminiscent of numerous directed evolution experiments that have been conducted on various natural proteins (Joo *et al.*, 1999; Farinas *et al.*, 2001; Glieder *et al.*, 2002; Gould and Tawfik, 2005). However, there is a significant difference. Previous experiments in laboratory-based evolution started with sequences that were themselves products of billions of years of natural selection. In contrast, laboratory evolution of our binary-patterned sequences enables the exploration of evolutionary trajectories emanating from naive sequences that are not biased by natural selection for biological function in living organisms.

Starting with the previously characterized binary-patterned proteins, S-824 and S-836, we demonstrate that mutagenesis of a *de novo*-designed protein, followed by screening of a relatively small number (hundreds or thousands) of
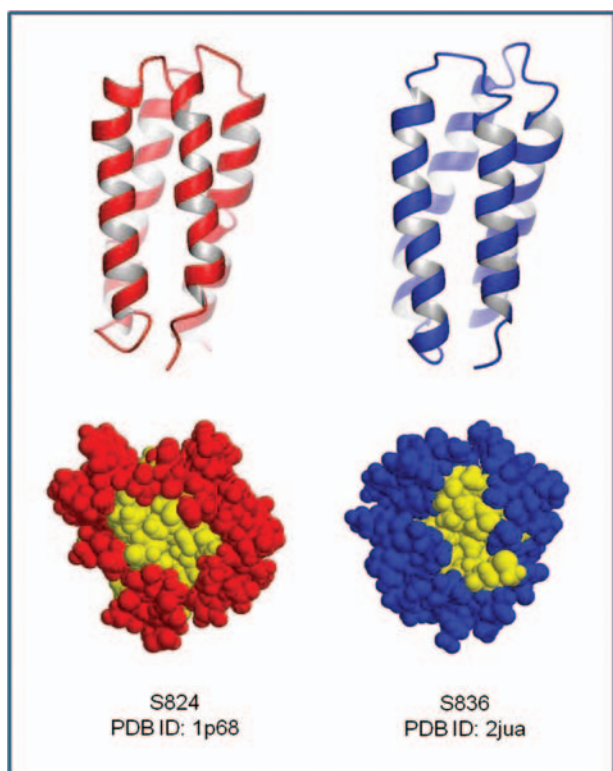
Fig. 1. Nuclear magnetic resonance spectroscopy solution structures of proteins S-824 and S-836. Top: Ribbon diagrams of four-helix bundle proteins. Bottom: Space-filling models of corresponding proteins with non-polar amino acids sequestered in the core (yellow) and polar amino acids on the exterior (red or blue). Figure adapted from Wei *et al.* (2003b) and Go *et al.* (2008).

variants can yield novel sequences with improved peroxidase activity.

Materials and methods

Random mutagenesis and construction of libraries of mutants

Error-prone polymerase chain reaction (PCR) using nucleotide analogs was used to introduce random mutations into genes encoding the novel proteins (Zaccolo *et al.*, 1996). The nucleotide analogs d-PTP (2′-deoxy-P-nucleoside-5′-triphosphate) and 8-oxo-dGTP (8-oxo-2′-deoxyguanosine-5′-triphosphate) were obtained from TriLink BioTechnologies. The products were gel purified (ZymoClean) and amplified using error-free PCR. The final PCR products were purified, concentrated (ZymoClean) and double-digested with NdeI and BamHI. Digestion mixtures were run on a low-melting 2% agarose gel, and the band at ∼300 bp was excised and purified. The mutation rate was three to four base mutations per sequence of 306 bp. Mutagenized gene sequences were cloned into the pET-3a vector. Previously, a stuffer fragment of lambda DNA had been cloned into the vector to ensure that only double-cut vector would be isolated following double digestion and gel purification. The purified double-digested vector was ligated with the library insert at 16°C for 12 h, desalted and electroporated into ElectroMAX competent cells (200 μl) at 2.3 kV (Transporator Plus) using a 2-mm gap cuvette. Transformation efficiency was ∼ 10^8 colony forming

units/μg DNA. After overnight growth, the colonies were scraped and pooled, and libraries of plasmid DNA were isolated for subsequent experiments.

Screen for peroxidase activity in cell lysates assayed in 96-well plates

BL21(DE3) *Escherichia coli* cells were transformed via electroporation with plasmid library DNA, and the cells were grown overnight at 37°C on LB/amp (100 μg/ml) plates. The colonies were picked and placed in 96-deep well plates containing auto-induction media (600 μl) and grown overnight at 37°C on a plate shaker. Auto-induction media was prepared by combining glycerol (0.4% v/v), glucose (0.05% w/v) and α-lactose (0.2% w/v) in LB/amp (Studier, 2005). Auto-induction media facilitates overexpression of the target protein without the need to monitor cell density or add isopropyl β-D-1-thiogalactopyranoside (IPTG). Following protein expression, aliquots of 100 μl were saved and the remaining cells were harvested by centrifugation. The cells were disrupted by addition of BugBuster lysis solution (200 μl; Novagen) and activity buffer (300 μl, 50 mM Tris–HCl, pH 8). The lysed cells were centrifuged and the supernatant was assayed for peroxidase activity. Samples for the assay were prepared by mixing protein supernatant (50 μl), hemin chloride (5 μM final concentration; Sigma), ABTS (1 mg/ml final concentration; Sigma), H_2O_2 (0.006% final concentration) in activity buffer (50 mM Tris–HCl, pH 8) for a final volume of 200 μl. After addition of H_2O_2, product formation was detected at 650 nm using a plate reader.

Colony screen for peroxidase activity

BL21(DE3) cells were transformed and grown overnight on LB/amp (100 μg/ml) plates as described above. The colonies were replica plated using nitrocellulose filter paper (Millipore). The original cells were grown at 37°C until the colonies reformed. Duplicate colonies on the nitrocellulose filter paper were placed on LB/amp (100 μg/ml) IPTG (200 μg/ml) plates for 4 h at room temperature to induce protein expression. The cells were then lysed by placing the nitrocellulose paper on top of filter paper soaked in BugBuster lysis solution (Novagen). The nitrocellulose paper was then moved to filter paper soaked in hemin chloride solution (5 μM; Sigma) to allow proteins in the lysate to bind heme. Finally, the nitrocellulose paper was moved to a Petri dish containing ABTS (2 mg/ml, Sigma) and H_2O_2 (0.06%) in activity buffer (50 mM sodium phosphate, pH 6). Colonies with peroxidase activity turned green and were marked. Replica colonies from the original plate were used to reconfirm activity, to make cell stocks and for plasmid preparations.

Protein purification

A single colony was used to inoculate 10 ml of LB/amp (100 mg/l final concentration) which was grown overnight at 37°C. This liquid culture was used to inoculate 1 l of LB/amp (100 mg/l final concentration) at a 1 : 1000 ratio. The 1 l culture was grown at 37°C until reaching an OD600 of ∼0.6 and IPTG (200 mg/l final concentration) was added to induce expression. Protein was released from cells using the freeze-thaw method (Johnson and Hecht, 1994). The protein was resuspended in $MgCl_2$ (100 mM; 10 ml per 1 l of culture) to extract the protein from the perforated cells, and

cellular debris was removed by centrifugation. The supernatant was then acidified (1 M sodium acetate buffer, pH 4; 1 ml per 1 l of culture) and cellular contaminants were removed by acid precipitation and centrifugation at $6000 \times g$ for 10 min. The resulting supernatant was loaded onto a Poros HS cation exchange column and purified using a gradient of NaCl. The protein was concentrated and exchanged into activity buffer using Centricon Plus-20 filters (Millipore).

Peroxidase kinetics and rate constants

Samples for the peroxidase assay were prepared by mixing protein (20 μM) with hemin chloride (5 μM) and ABTS (1 mg/ml) in activity buffer (50 mM Tris–HCl, pH 8). A hemin chloride stock solution (1 mM) was prepared in dimethyl sulfoxide, and an ABTS stock solution (20 mg/ml) was prepared in activity buffer. As the heme concentration is much lower than protein concentration, we assume the maximal concentration of heme–protein complex is 5 μM. On addition of H_2O_2 (0.086 μM – 11 μM), time points were recorded for 30 min at 650 nm. Kinetic data were recorded using the Varioskan Flash Spectral Scanning Multimode Reader and SkanIt software (Thermo Fisher Scientific, Waltham, MA). Peroxidase rate constants were analyzed using Michaelis–Menten kinetics: $1/V = K_M/(k_{cat}[E_0][S]) + 1/(k_{cat}[E_0])$.

Results

Development of two screens for peroxidase activity

The *de novo*-designed four-helix bundles, S-824 and S-836, were targeted for directed evolution. These sequences were chosen as test cases because both proteins had been characterized extensively, and their structures were solved previously by Nuclear magnetic resonance spectroscopy, as shown in Fig. 1 (Wei *et al.*, 2003a,b; Go *et al.*, 2008). Furthermore, as both proteins are very stable, it seemed likely that their scaffolds would tolerate mutations. Mutations in both sequences were created by error-prone PCR using nucleotide analogs (Zaccolo *et al.*, 1996). Mutant libraries of both sequences were initially screened for peroxidase activity in 96-well plates. These screens provide clear results, but are difficult to scale up without robotic instrumentation. An example of a 96-well plate assay for peroxidase activity in cell lysates for a library of S-836 mutants is shown in Fig. 2. The range in the intensity of color indicates a range of activity in the mutant library. Not surprisingly, many mutant sequences display lower levels of activity than the parental S-836 sequence. However, several mutants appear to be more active than S-836. The sequences responsible for the darkest color were chosen for further rounds of mutagenesis and screening.

To facilitate sampling of larger libraries, we also developed a colony-based screen. For this screen, the cells were grown, expressed, lysed and assayed for activity in colonies growing on filter paper above agar Petri dishes. When the colonies are exposed to substrate, those that change color first are picked for further analysis. In our hands, the 96-well format was convenient for screening several hundred sequences, whereas the colony-based assay allowed screening of thousands of clones. Figure 3 shows a control experiment for a colony screen. Each plate contains colonies expressing
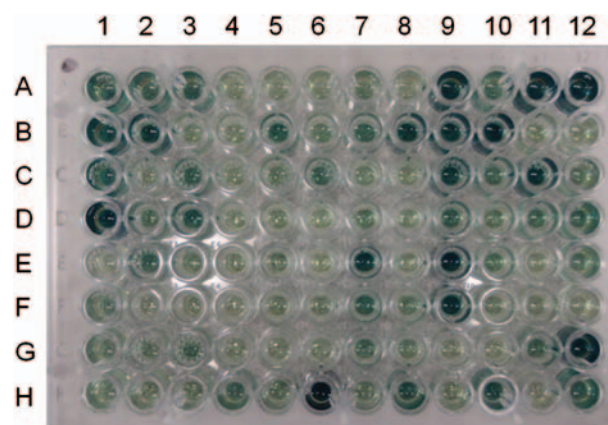

Fig. 2. A 96-well plate shows a screen for peroxidase activity in cell lysates of a library of mutants of *de novo* protein S-836. The darker color corresponds to higher activity and the lighter color corresponds to lower activity. Wells A1–A3 in the top left corner contain lysates of cells expressing the parent sequence, S-836. Wells A4–A6 contain buffer incubated with heme as a negative control. The remaining wells contain lysates of cells expressing mutants of S-836.
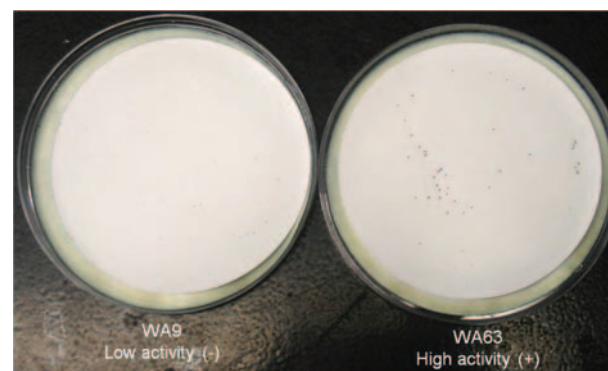

Fig. 3. A colony screen for peroxidase activity. Filters on each agar plate contain colonies expressing a single *de novo* protein, and the number of colonies is approximately the same on both plates. Protein WA9 does not have significant peroxidase activity, and the colonies exhibit little to no color (left). In contrast, protein WA63 yields colonies with a darker blue color, indicative of peroxidase activity (right).

a single known sequence. The negative control (left) shows colorless colonies expressing sequence WA9, which has no peroxidase activity. In contrast, the positive control (right) shows colored colonies for cells expressing the active sequence, WA63. Thus, colonies expressing sequences with different levels of peroxidase activity can be distinguished readily.

Isolation of mutants with increased peroxidase activity

Initially, we used 96-well plates to screen lysates for enhanced peroxidase activity amidst libraries of mutants of proteins S-824 and S-836. Figure 4a shows the results for a single 96-well plate of mutants of S-824. The activity of each mutant is plotted relative to the parental sequence. Not surprisingly, most mutations are deleterious, and most sequences are less active than the parental sequence (ratios <1 in Fig. 4a). However, in this initial screen, four proteins (∼4% of the library) showed activity above the parent by 3 standard deviations (a value of 1.3). Next, we screened the S-836 library in four 96-well plates (Fig. 4b). In this case,
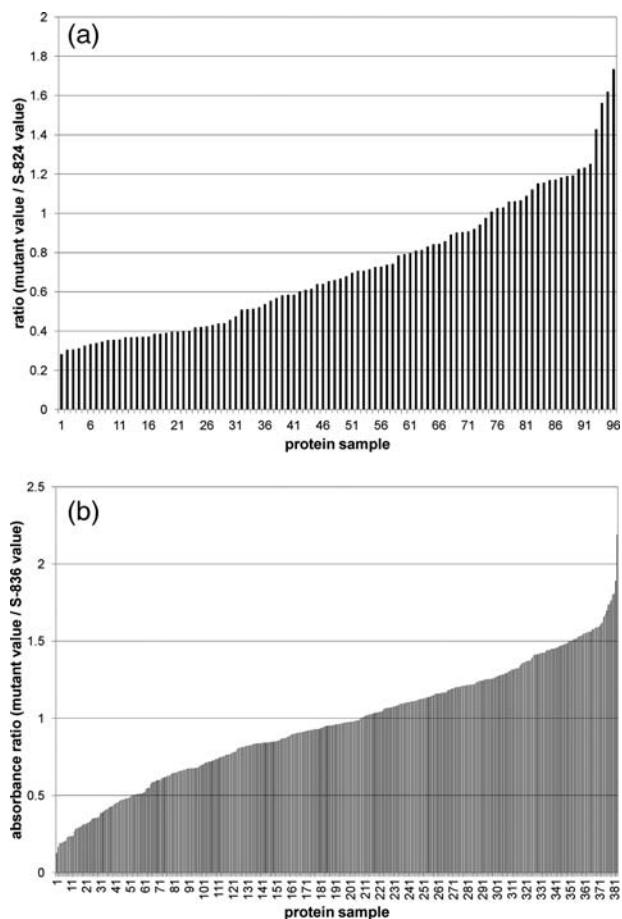
Fig. 4. Screens for enhanced peroxidase activity in 96-well plates. Each bar represents a single protein sequence assayed in the context of a cell lysate, and samples are ranked from lowest to highest absorbance value. The *y*-axis shows the absorbance ratio of each mutant relative to its parent sequence. The standard deviation for the parent is ~10%. Samples with an absorbance ratio >1.3 (i.e. >3 standard deviations above the parent) were considered significant. (a) Peroxidase screen of first generation S-824 mutant library. (b) Peroxidase screen of first generation S-836 mutant library.

~20% of the library showed activity at least 3 standard deviations above the parent. As a larger percentage of the S-836 library (compared with the S-824 library) showed improved activity, we chose this library for further studies.

To examine a larger collection of sequences, we screened the library of S-836 mutants using the colony screen described above. Approximately 2×10^3 colonies were screened. Those that turned green first were picked and reassayed in liquid media. The results of these assays are shown in Fig. 5. All of the picked clones showed greater activity than the parent.

Next, we took the top hits from the 96-well plate screen (clones 1-2B11, 1-2G6) and from the colony screen (clones 1-C3, 1-C8) and mutagenized these to create a second-generation library of mutants. This library was produced in the same way as the first-generation library, using error-prone PCR and nucleotide analogs. The second-generation library of mutants was then screened in four 96-well plates, and the results are shown in Fig. 6. Two percent of the mutants showed activity significantly above the parent S-836.

The top hits from the first-generation 96-well screen (1-2B11, 1-2G6), the first-generation colony screen (1-C3, 1-C8) and the second-generation 96-well screen (2-2H12,
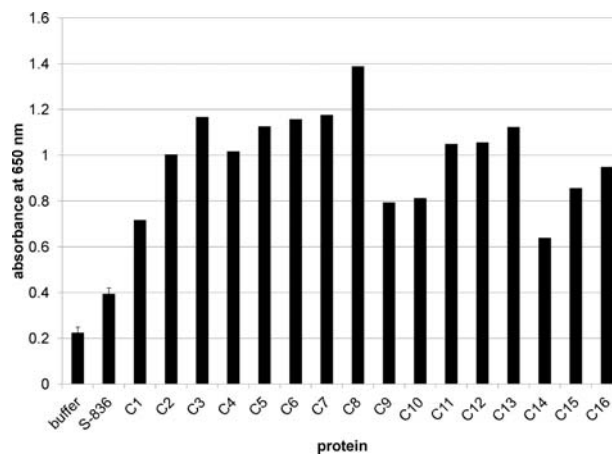

Fig. 5. Assay of peroxidase activity for putative hits isolated from the colony screen of the first generation S-836 mutant library. Sixteen colonies were chosen, and the peroxidase activity of each lysate was compared with the negative control (buffer + heme, first bar) and the parent sequence, S-836 (second bar). The negative control and S-836 samples were analyzed in triplicate and the error bars represent a single standard deviation from three samples. Mutant samples were analyzed once.
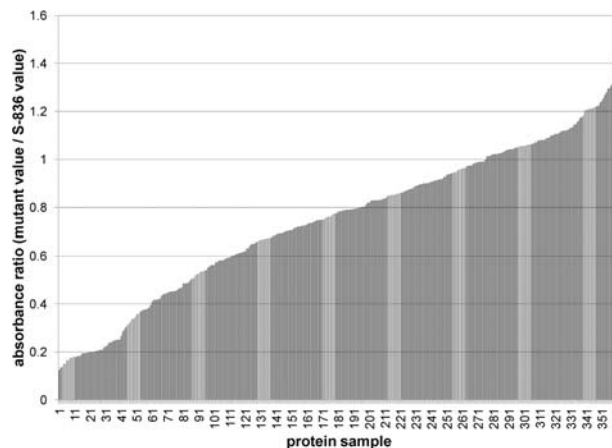

Fig. 6. Peroxidase screen of a second-generation library of mutants of protein S-836. Screens were performed in 96-well plates. Each bar represents a single mutant sequence assayed in the context of a cell lysate, and samples are ranked from lowest to highest absorbance value. The *y*-axis shows the absorbance ratio of each mutant relative to the parent S-836.

2-3F1, 2-4C1) were then reassayed for peroxidase activity in cell lysates. This assay confirmed that all of the putative hits showed activity above the parental S-836 protein. These results are shown in Fig. 7 and the corresponding sequences are shown in Fig. 8. As the activity observed in a cell lysate depends on both the intrinsic activity of a protein and the level of its expression, we estimated expression levels by gel electrophoresis. As shown in Fig. 9, all sequences expressed at similar levels.

Biochemical characterization of purified proteins

To verify that the mutant sequences are indeed more active than the parental S-836 protein, we chose two proteins for purification and biochemical characterization. Sequences 1-2B11 and 2-2H12 were chosen because they displayed relatively high levels of activity in lysates (Fig. 7), and were straightforward to purify. Protein 1-2B11 had been isolated in the 96-well screen of the first-generation library, and

protein 2-H12 was the top hit in the 96-well screen of the second-generation library. Sequence 2-H12 is a variant of 1-2B11 (Fig. 8).

The kinetic profiles of the purified mutant proteins are compared with S-836 in Fig. 10, and the k_{cat} and K_M values are summarized in Table I. Both mutants show greater activity than S-836. Comparison of the k_{cat}/K_M values shows the progeny are nearly 3-fold more active than the parent.

As expected from the assays in lysates (Fig. 7), the activities of the purified proteins 2-2H12 and 1-2B11 were nearly the same. We consider several factors that might account for the lack of improvement between the first-generation and second-generation proteins: (i) The screens were implemented on relatively small collections of sequences, compared with the extremely large number of mutants possible for a 102-amino acid protein. For example, we screened ~10^3 sequences to isolate the quintuple mutant, 2-2H12. This represents a very small fraction of the 2×10^{16} possible quintuple mutants [$(19 \times 102) \times (19 \times 101) \times (19 \times 100) \times (19 \times 99) \times (19 \times 98) = 2 \times 10^{16}$]. (ii) Most random mutations are likely to be deleterious—either for folding, heme binding or enzymatic activity (Bell, 1997). Therefore, multiple mutations with additive favorable effects would occur rarely, and it would be difficult to find them amid small collections of mutants. (iii) The four-helix bundle structure of protein S-836 was designed for structural stability, not for cofactor binding or catalytic activity, and may not be optimal
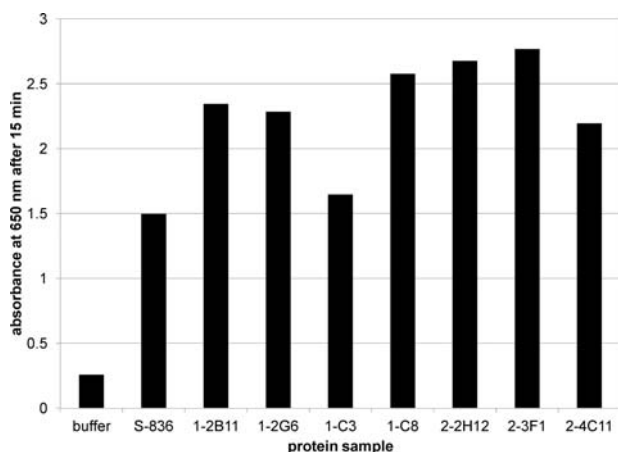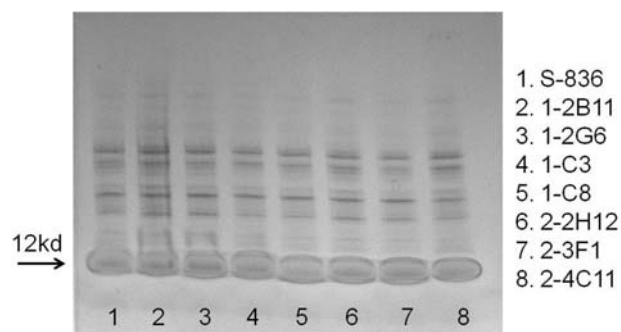

Fig. 9. Expression of protein S-836 and mutants. Expression is detected by the presence of a strong band at the bottom of each gel (MW 12 kD), as indicated by the arrow.
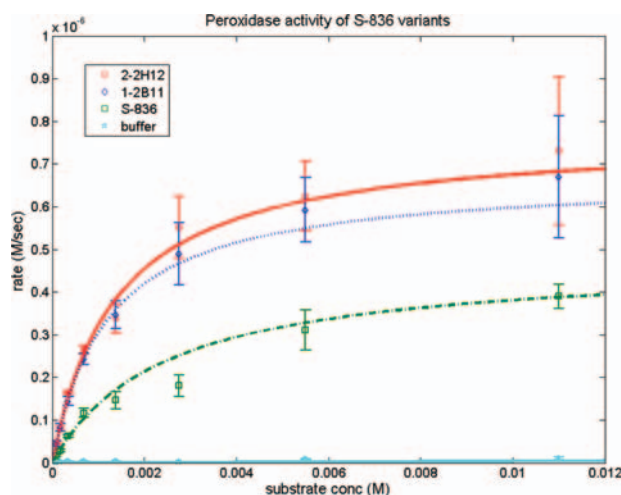

Fig. 10. Kinetic profile for peroxidase activity of protein S-836 (open square, dashed line) and its variants, 1-2B11 (open diamond, dashed line) and 2-2H12 (open circle, solid line). The negligible background of buffer + heme alone (star, solid line) is shown for comparison.

Fig. 7. Confirmation of hits from screens of S-836 mutant libraries. Each mutant is compared with the parental S-836 sequence and to the buffer + heme contol. The number '1' in front of a variant indicates a sequence from the first-generation library and the number '2' in front of a variant indicates a sequence from the second-generation library.

Table I. Kinetic rate constants for the peroxidase activity of S-836 and two variants, 1-2B11 and 2-2H12

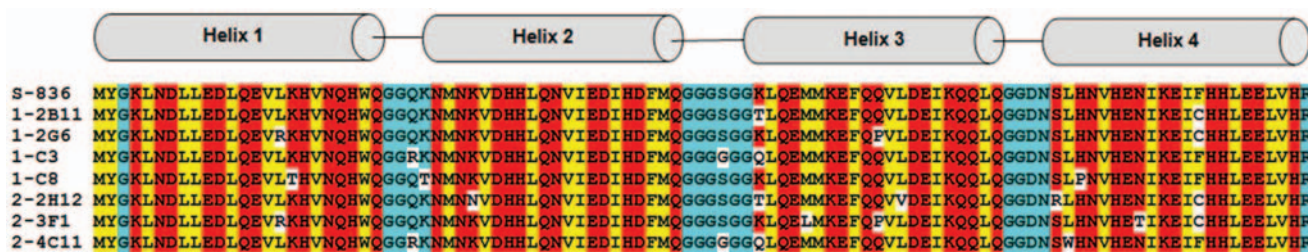| Protein	k_{cat} (s^{-1})	K_M (M)	k_{cat}/K_M (s^{-1} M^{-1})
S-836	0.095 ± 0.042	$2.6 \pm 1.5 \times 10^{-3}$	39 ± 7.5
1-2B11	0.13 ± 0.011	$1.2 \pm 0.31 \times 10^{-3}$	112 ± 17
2-2H12	0.15 ± 0.026	$1.4 \pm 0.28 \times 10^{-3}$	111 ± 4.3

Fig. 8. Sequences of protein S-836 and mutants. Each mutant has two to five amino acid substitutions. Amino acids are highlighted as follows: Binary-patterned polar and non-polar residues are in red and yellow, respectively. Designed turns are in blue, and mutations are white.

for peroxidase activity. Thus, further enhancements in activity may be difficult to achieve.

Discussion

Various natural proteins have been subjected to directed evolution for enhanced activity. These include several heme proteins that were evolved *in vitro* to increase peroxidase activity and/or to alter substrate specificity. Heme-containing peroxidases in nature catalyze the oxidation of hydrogen peroxide with the aid of a reducing agent. One naturally occurring peroxidase is horseradish peroxidase (HRP), which is used as a reporter for various diagnostic assays, as it has broad substrate specificity and high turnover. Morawski *et al.* (2001) implemented directed evolution by utilizing a colony screen to detect HRP mutants with improved thermostability, yielding a variant that was still functional at high temperatures. This opens up the possibility of using HRP in commercial applications such as bioremediation, polymer production or in sensor applications. Another natural peroxidase, cytochrome *c* peroxidase (CCP), which catalyzes the oxidation of ferrocytochrome *c* by hydrogen peroxide, was evolved to have greater specificity toward different small molecule substrates (Iffland *et al.*, 2000; Iffland *et al.*, 2001). Specifically, mutant libraries were created by random mutagenesis and DNA shuffling, and after three rounds of screening $\sim 10^4$ clones per round, the variants were isolated with 300-fold greater values of k_{obs} (apparent first-order rate constant) for the small molecule guaiacol (Iffland *et al.*, 2000). Subsequently, CCP was also evolved toward the substrate ABTS, and mutants with 20-fold increased activity were isolated (Iffland *et al.*, 2001).

There are also natural heme proteins, whose primary function is not peroxidase activity, yet are still able to catalyze this reaction. For example, myoglobin, although best known as an oxygen storage protein, also has peroxidase activity, and this enzymatic activity can be increased by evolution *in vitro*. In one study, Wan *et al.* (1998) used directed evolution to improve the peroxidase activity of myoglobin by creating randomly mutated libraries of the horse heart myoglobin gene and implementing a color-based colony screen toward H_2O_2 and the ABTS reducing agent. After four rounds of mutagenesis and screening, a myoglobin quadruple variant with improved peroxidase kinetics was discovered. The mutant had a k_1 value ~ 25-fold higher than wild type (representing the rate of H_2O_2 oxidation) and a k_3 value 1.6-fold that of wild type (representing ABTS oxidation).

These examples illustrate applications of directed evolution of proteins that are themselves products of billions of years of natural selection. In the current work, we used similar approaches, and like these earlier studies, we screened thousand of sequences to isolate mutant proteins that typically contained one to five amino acid substitutions. However, our starting material was dramatically different: instead of starting with natural sequences that are themselves products of evolution, we applied laboratory evolution techniques to unevolved sequences that were designed *de novo*. Because our proteins are not derived from natural sequences, they are not biased by any requirement to provide life-sustaining functions. Moreover, as they are not the progeny of ancestral sequences, they do not carry the 'evolutionary baggage' associated with the biological and environmental factors that influenced the selection of ancestral sequences through millions of years of evolutionary history. Indeed, the binary code proteins S-824 and S-836 were designed *de novo* for structural stability without any attempt *a priori* to specify heme binding or catalytic activity. Nevertheless, our earlier work showed that these naive proteins bound heme with reasonable affinities, and catalyzed peroxidase chemistry at levels substantially above background (Rojas *et al.*, 1997; Moffet *et al.*, 2000; Das *et al.*, 2006; Das and Hecht, 2007; Patel *et al.*, 2009). Here, we show that these *de novo* proteins can also be used as a novel feedstock to select progeny sequences with improved activities.

Future work will focus on using *de novo* proteins to provide essential biological functions *in vivo*. Initial work in this realm has already demonstrated that proteins from a binary-patterned library of four-helix bundles can function in *E. coli* and sustain cell growth by rescuing the deletion of conditionally essential genes encoding the following four enzymes: phosphoserine phosphatase, citrate synthase, threonine deaminase and enterobactin esterase (Fisher *et al.*, 2011). Those results demonstrated that although the binary-patterned proteins were not derived from natural biological sequences, they nonetheless can provide essential biological functions. Because these functions are required for cell growth, they can be evolved for increased levels of activity, thereby laying the foundation for the eventual construction of novel proteomes that provide numerous functions essential for life.

Funding

References

Arai,R., Kobayashi,N., Kimura,A., Sato,T., Matsuo,K., Wang,A.F., Platt,J.M., Bradley,L.H. and Hecht,M.H. (2012) *J. Phys. Chem. B*, doi:10.1021/jp212438h.

Bell,G. (1997) *The Basics of Selection*. Chapman & Hall, New York.

Cherny,I., Korolev,M., Koehler,A.N. and Hecht,M.H. (2012) *ACS Synth. Biol.*, **1**, 130–138. doi:10.1021/sb200018e.

Das,A. and Hecht,M.H. (2007) *J. Inorgan. Biochem.*, **101**, 1820–1826.

Das,A., Trammell,S.A. and Hecht,M.H. (2006) *Biophy. Chem.*, **123**, 102–112.

Farinas,E.T., Schwaneberg,U., Glieder,A. and Arnold,F.H. (2001) *Adv. Synth. Catal.*, **343**, 601–606.

Fisher,M.A., McKinley,K.L., Bradley,L.H., Viola,S.R. and Hecht,M.H. (2011) *PLoS ONE*, **6**, e15364. doi:10.1371/journal.pone.0015364.

Glieder,A., Farinas,E.T. and Arnold,F.H. (2002) *Nat. Biotechnol.*, **20**, 1135–1139.

Go,A., Kim,S., Baum,J. and Hecht,M.H. (2008) *Protein Sci.*, **17**, 821–832.

Gould,S.McQ. and Tawfik,D.S. (2005) *Biochemistry*, **44**, 5444–5452.

Hecht,M.H., Das,A., Go,A., Bradley,L.H. and Wei,Y. (2004) *Protein Sci*, **13**, 1711–1723.

Iffland,A., Tafelmeyer,P., Saudan,C. and Johnsson,K. (2000) *Biochemistry*, **39**, 10790–10798.

Iffland,A., Gendreizig,G., Tafelmeyer,P. and Johnsson,K. (2001) *Biochem. Biophys. Res. Commun.*, **286**, 126–132.

Johnson,B.H. and Hecht,M.H. (1994) *Biotechnology*, **12**, 1357–1360.

Joo,H., Lin,Z. and Arnold,F.H. (1999) *Nature*, **399**, 670–673.

Kamtekar,S., Schiffer,J.M., Xiong,H., Babik,J.M. and Hecht,M.H. (1993) *Science*, **262**, 1680–1685.

Moffet,D.A., Certain,L.K., Smith,A.J., Kessel,A.J., Beckwith,K.A. and Hecht,M.H. (2000) *J. Am. Chem. Soc.*, **122**, 7612–7613.

Morawski,B., Quan,S. and Arnold,F.H. (2001) *Biotechnol. Bioeng.*, **76**, 99–107.

Patel,S.C., Bradley,L.H., Jinadasa,S.P. and Hecht,M.H. (2009) *Protein Sci.*, **18**, 1388–1400.

Rojas,N.R., Kamtekar,S., Simons,C.T., McLean,J.E., Vogel,K.M., Spiro,T.G., Farid,R.S. and Hecht,M.H. (1997) *Protein Sci.*, **6**, 2512–2524.

Studier,F.W. (2005) *Protein Expr. Purif.*, **41**, 207–234.

Wan,L., Twitchett,M.B., Eltis,L.D., Mauk,A.G. and Smith,M. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 12825–12831.

Wei,Y., Liu,T., Sazinsky,S.L., Moffet,D.A., Pelczer,I. and Hecht,M.H. (2003a) *Protein Sci.*, **12**, 92–112.

Wei,Y., Kim,S., Fela,D., Baum,J. and Hecht,M.H. (2003b) *Proc. Natl. Acad. Sci. USA*, **100**, 13270–13273.

Zaccolo,M., Williams,D.M., Brown,D.M. and Gherardi,E. (1996) *J. Mol. Biol.*, **255**, 589–603.