AFOSR MURI: Fundamentals and Bioengineering of Enzymatic Fuel Cells
FA9550-06-1-0264

Final Report: Part I

Bioengineering of Enzymes as Electroctalysts

Thrust Lead:

Scot Banta, PhD, Associate Professor
Columbia University, New York, NY

in collaboration with:

Scot Calabrese Barton, PhD, Associate Professor
Michigan State University, East Lancing, MI

Shelley Minteer, PhD, Professor
St. Louis University, St. Loius, MI,
Now at the University of Utah, Salt Lake City, Utah

and

Plamen Atanassov, PhD, Professor
University of New Mexico, Albuquerque, NM

20120918192

Table of Contents

Chapter 1 Overview

The activities at Columbia University were mainly devoted towards Thrust Area I, Bioengineering of Electrocatalysis. Contributions were made to other Thrust Areas as well. The contributions can be divided into 5 areas and these contributions are summarized below. The following chapters contain the manuscripts that were produced in each area of contribution.

The first area of contribution was the development of bi-functional enzymatic hydrogels for use in bioelectrode modification. Bioelectrocatalytic enzymes are often immobilized on electrodes using polymers or other materials. As these systems are developed, the enzymes will be engineered for improved performance. And, there has been a good deal of recent interest in creating hydrogels using proteins and peptides. So, we developed the idea of engineering redox proteins to self-assemble into hydrogels that could be used as bioelectrode modifications. Our first effort in this area is described in Chapter 2 where we engineered fluorescent proteins to self-assemble into hydrogels. This was accomplished by adding helical appendages to the proteins and the helical appendages will dimerize and tetramerize into coiled-coil bundles. We were able to demonstrate the formation of robust hydrogels with the monomeric green fluorescent protein and by using the tetrameric red fluorescent protein we were able to make hydrogels with cross-linking mediated by both the helical appendages and through protein-protein interactions. Erosion experiments demonstrated that the addition of the globular protein caused the hydrogels to erode more slowly as compared to peptide hydrogels that do not contain a protein. And, FRET experiments using mixed fluorescent protein hydrogels show that the hydrogels do not phase-separate which has been a limitation in other mixed enzyme systems. Once this groundwork was laid, we then went on to incorporate redox enzymes into the hydrogels. In Chapter 3 we describe our efforts to engineer the small laccase (SLAC) enzyme from *Streptomyces coelicolor*. This enzyme can reduce oxygen to water with high current densities, and is one of the most active laccase enzymes at neutral pH. SLAC is a multimer and so only one helical appendage was needed to form a hydrogel. The addition of the appendage caused a 2-order of magnitude loss in activity of the enzyme, but it was still active. In order to create electrical contact with redox enzymes, mediators are often employed and we had previously shown that SLAC performs well in an osmium-modified redox polymer. Therefore, we engineered a self-assembling protein hydrogel to bind osmium moieties and we engineered the SLAC enzyme with the self-assembling helical appendages. When the SLAC and osmium hydrogels were combined and applied to an electrode, we were able to demonstrate that the hydrogel protein engineering approach could produce highly active bioelectrodes. We further developed this approach by engineering enzymes for use on the biofuel cell anode. First we worked with a thermostable alcohol dehydrogenase (AdhD) from *Pyrococcus furiosus* and this is described in Chapter 4. The addition of the helical appendages to AdhD did not impact the kinetic activity of the enzyme, and the enzyme formed robust hydrogels. The hydrogels retained the thermostability of the enzyme and the mechanical properties of the hydrogel changed at high temperatures presumably due to melting of the coiled-coil bundles and not the AdhD enzyme. Enzymatic activity in the bulk hydrogels was monitored by following the intrinsic fluorescence of the nicotinamide cofactor. In order to expand the types of enzymes that we could engineer to be bifunctional, we added the helical appendages to an organophosphate hydrolase enzyme and this is

described in Chapter 5. This was interesting as we were able to form hydrogels from the protein expressed in the *E. coli* extract without purifying it. And, we found that the enzyme, which has a metal center, made more robust hydrogels when a polyhistidine purification tag was included. This suggests that additional cross-linking was occurring between the histidine tag and the active site, and these hydrogels were incredibly robust. In fact, we were able to make a hydrogel and show that it remained intact and active after 5 months submersed in buffer in the refrigerator. Continuing on our dehydrogenase work, we engineered 3 dehydrogenase enzymes to self-assemble and form a metabolic pathway and this is described in Chapter 6. The 3 enzymes are able to oxidize methanol to carbon dioxide. Each of the enzymes produces NADH which can be oxidized at a bioanode. The mixed dehydrogenase hydrogel was applied to a novel biofuel cell design and high current and power densities were achieved. Finally in Chapter 7 a review of protein hydrogel technology is presented and the work performed in this MURI program is highlighted.

The next area of contribution focused on the engineering of a dehydrogenase enzyme for use in biofuel cell applications. One approach for creating biofuel cells is to find native enzymes that have a desired activity. This approach is often limited by the features of the enzyme such as thermostability and cofactor requirements, and this is difficult to change for each enzyme that is utilized. We decided to take a different approach and to identify a good enzyme with some desirable features and then to engineer it for different activities that will be useful in biofuel cell development. As described in Chapter 4 we began working with the AdhD dehydrogenase from *Pyrococcus furiosus*. This enzyme is monomeric, extremely theromstable, and uses the cofactor NAD(H) to oxidize or reduce longer chain secondary alcohols and aldehydes. Many features of this enzyme are desirable, however the substrate specificity is not ideal since secondary alcohols are generally not usefuel fuels in biofuel cells. As described above, in Chapter 4 we demonstrate the engineering of the AdhD enzyme to be bifunctional so that it self-assembles into enzymatic hydrogels. In Chapter 8 we describe our efforts to rationally engineer the cofactors specificity of the enzyme. The enzyme uses NAD(H) as a cofactor and other groups have demonstrated the mutations needed to broaden the cofactor specificity in related enzymes. We made two point mutations to the cofactor binding pocket of the AdhD enzyme which made it more active with NAD(H) and it broadened the enzymatic activity so that the enzyme could use NADP(H) as a cofactor. We also performed pre-steady state kinetics on the enzyme and the various mutants and our results showed an unexpected richness of behavior that demonstrates the enzyme undergoes conformational changes during the catalytic cycle which had not been previously explored. In Chapter 9 we describe our efforts to change the substrate specificity of the AdhD enzyme. AdhD is a member of the aldo keto reductase superfamily of enzymes, and these enzymes have a variety of substrate specificities which is mediated by 3 flexible surface loops on the top of the enzyme. The surface loops from the mesostable human aldose reductase, which has activity with glucose, were transferred to AdhD and we showed that this introduced glucose activity into AdhD and this also introduced a change in cofactor specificity as well. Interestingly when the loops were swapped in the double mutant enzyme from Chapter 8, the enzyme was inactive which means we discovered two routes to changing cofactor specificity but the two methods are not additive. Once we demonstrated that we could engineer the cofactor and substrate specificity of AdhD, we wanted to see if we could use it in a biofuel cell device. The native cofactors NAD(H) and NADP(H) are expensive, marginally stable, and are large molecules. The use of biomimetic cofactors would allow for more flexibility in the design and operation of these devices, and we found that our double mutant AdhD

4

enzyme described in Chapter 8 has low activity with the biomimetic cofactor nicotinamide mononucleotide (NMN). The use of this cofactor also allowed the enzyme to use various substrates including arabinose. We measured the kinetics of the mutant enzyme with NMN and arabinose and we collaborated with the Minteer group to produce a biofuel cell using the biomimetic cofactor. This work is described in Chpater 10 and the results show that even though the enzyme is less active with NMN, the cofactor is smaller than NADH and thus the use of NMN results in a biofuel cell with improved performance presumably due to improved mass transport.

In addition to exploring dehydrogenases for use in the biofuel cell anode, our third area of contribution is the engineering of glucose oxidase for anode operation. There has been a good deal of interest in glucose oxidase as it can oxidize glucose without the need for a nicotinamide cofactor. However, synthetic mediators are often used for electron transfer, as the enzyme does not perform well for direct electron transfer to an electrode since its cofactor is buried deeply inside the enzyme. We took a new approach to address this problem and we engineered a surface accessible thiol group near the active site of the enzyme. We then mixed the mutant enzymes with modified gold nanoparticles. The results of these experiments are in Chapter 11 where we show that the modification does negatively impact the kinetic activity of the enzyme, but the enzyme/nanoparticle conjugates are capable of making direct electron transfer to the electrode through the attached gold nanoparticle.

Our fourth area of contribution is in the area of using computational protein design to attempt to create new laccase enzymes for the biofuel cell cathode. Some of the fungal laccases are favored for cathode applications since they have high redox potentials, but they are difficult to express, they are glycosylated, they are most active at low pH and only few have been expressed recombinantly. The bacterial laccases address many of these limitations but they tend to have lower redox potentials. We set out to use computational design to form a protein that adopts the same structure and active site as a fungal laccase but could be expressed in *E. coli*. The protein did not fold into the desired conformation and this suggests that the native proteins may require chaperones or other factors that facilitate folding. This work, described in Chapter 12, was one of the largest protein redesign projects ever attempted and it helped highlight the limitations of our understanding of laccase protein folding.

The final area of contribution is in the modeling of enzymatic biofuel cells. Metabolic control analysis was developed to understand how control is distributed over metabolic networks and pathways. Often it is found that native pathways have distributed control which eliminates a single rate-limiting step and allows a dynamic response to perturbations. We applied this analysis to a simple glucose/oxygen biofuel cell and determined when control would shift form the anode to the cathode. This work is described in Chapter 13 and this framework will be even more valuable when multi-enzyme anodes are explored.

Going forward it will be critical to continue engineering enzymes for bioelectrocatalytic systems such as enzymatic biofuel cells. One of the next areas to explore will be the engineering of enzymes and nanomaterials simultaneously so that enzymatic nanoarchitectures can be created that are designed to interface into devices. This way features such as redox potential, activity, and transport limitations can be addressed at the design stage, and this will continue to lead to new designs and devices with improved performance.

Chapter 2

Wheeldon, I.R., Calabrese Barton, S., and **Banta, S.** (2007) "Bioactive Proteinaceous
Hydrogels from Designed Bi-functional Building Blocks" *Biomacromolecules* **8**(10)
2990-2994.

Bioactive Proteinaceous Hydrogels from Designed Bi-Functional Building

Blocks

Ian R. Wheeldon, Scott Calabrese Barton, Scott Banta*

Department of Chemical Engineering, Columbia University in the City of New York, New

York, NY 10027, USA,

sbanta@cheme.columbia.edu

†Correspondence should be address to S. Banta, email: sbanta@cheme.columbia.edu. *

Present address: Chemical Engineering and Materials Science, Michigan State University,

East Lansing, MI 48824-1226, USA

Stimulus-responsive, or 'smart', protein-based hydrogels are of interest for many bioengineering applications, but have yet to include biological activity independent of structural functionality. We have genetically engineered bi-functional building blocks incorporating fluorescent proteins that self-assemble into robust and active hydrogels. Gellation occurs when protein building blocks are cross-linked through native protein-protein interactions and the aggregation of α-helical hydrogel-forming appendages. Building blocks constructed from different fluorescent proteins can be mixed to enable tuning of fluorescence loading and hydrogel strength with a high degree of independence. FRET experiments suggest a macro-homogenous structure and that intra-gel and inter-protein reactions can be engineered. This design approach will enable the facile construction of complex hydrogels with broad applicability.

Introduction

Concerted effort has been aimed at the development of hydrogels for biomedical applications such as drug delivery, and tissue engineering[1-3]. Peptide motifs and protein domains are frequently incorporated because they can be readily engineered to form reversible physical cross-links in supramolecular assemblies and hydrogel networks. Hydrogels can be further engineered to undergo structural changes in response to environmental stimuli such as ionic strength, pH, temperature or analyte concentration[4]. Many examples exist including engineered α-helical domains used to non-covalently cross-link protein[5-8] and hybrid[9] hydrogels, amphiphilic β-sheets that assemble into responsive supramolecular networks of oligopeptides[10-13], and binding proteins that form physical cross-links in synthetic polymer networks that respond to changes in analyte concentrations[14,15]. Engineered polyvalent binding proteins have also been used to form structurally responsive planar networks[16].

We foresee a next generation of protein hydrogels with broader application enabled by the design of bi-functional building blocks with both bioactivity and self-assembly functionalities. Enzymatic, optically active, analyte-binding and signaling protein building blocks will self-assemble into ordered supramolecular structures with tunable physical properties and bioactivities. With the development of a series of compatible building blocks, each with distinct function, the composite material as a whole will gain new structure-function relationships. Just as protein secondary and tertiary structure dictates the functionality of the building blocks, the composition of the supramolecular assemblies will dictate hydrogel functionality.

We have begun to explore this new class of hydrogels by starting with recombinant proteins and adding hydrogel functionality through the addition of physical cross-linking appendages. Here we report the development of three natively-folded and functional proteins that form hydrogels with strong elastic character and minimal erosion rates. Bioactivity is derived from a fluorescent protein and hydrogel formation is mediated through native protein-protein interactions and aggregation of α-helical (H) domains. We show that neat and mixed hydrogels are macrohomogeneous and that fluorescence loading is tunable.

Experimental Section

Expression plasmid constructs. Plasmid expressing the hydrogel forming protein AC10Acys[6], termed here H-S-H, was a kind gift from David Tirrell (CalTech). Green fluorescent protein (GFP), enhanced cyan fluorescent protein (ECFP) and red fluorescent

7

protein (DSRED) were excised from pRSETb S65T[29], pCMVECFP and pCMVDSREDexpress (Clonetech), respectively, with the addition of upstream and downstream restriction sites. GFP S65T and ECFP without a terminal stop codon were ligated into pQE9AC10Acys at the unique SphI site downstream of the polyelectrolyte C10 (or S) region resulting in pQE9HSGFPH and pQE9HSECFPH. GFP S65T with terminal stop codon was ligated into pQE9AC10Acys at the same location as above resulting in pQE9HSGFP. DSRED with terminal stop codon was ligated into pQE9AC10Acys at the unique SphI and SpeI sites resulting in pQE9HSDSRED. Plasmids were propagated in 5α *E.coli* (New England Biolabs) and expression was carried out in SG13009 *E.coli* expression cell line (Qiagen).

Protein Expression and Purification. Expression and purification of all protein hydrogels were preformed in an identical manner. One to four liters of TB media, supplemented with 200 mg mL^{-1} of ampicillin (Sigma) and 50 mg mL^{-1} of kanamycin (Sigma), were inoculated with mature culture of SG13009 containing the appropriate plasmid and induced with 0.5mM IPTG (Promega) at an OD_{600} = 0.85 - 1.05. Growth prior to induction occurred at 37 °C. Temperature was reduced to 20 °C during overexpression. Cells were harvested by centrifugation at 15,000 g for 15 minutes after 18-20 hours. Cell pellets were re-suspended in nickel affinity chromatography binding buffer as recommended for His-Trap purification (GE Healthcare). Cells were disrupted by sonication after a freeze thaw cycle to -80 °C. The resulting lysate was clarified by centrifugation at 15,000 g for 30 minutes. Proteins were purified from the crude lysate via Fast low Pressure Liquid Chromatography (ATKA 900 FPLC, GE Healthcare) with a nickel affinity column (His-Trap Crude, 5mL, GE Healthcare). Fractions containing the desired proteins were pooled and buffer exchanged over a 30 kDa cellulose filter (Milipore). Buffer exchange to 3.4 mM monobasic sodium phosphate and 16.6 mM dibasic sodium phosphate over a 30 kDa filter served two purposes; to remove the high levels of salts resultant from the purification and to remove proteins with a molecular weight less than ~30 kDa. Purified protein was further concentrated by centrifugation over a 10 kDa cellulose filter (Centriplus, Millipore) to 15-35 mg mL^{-1}. Aliquoted samples were frozen to -20 °C and lyophilized (FastFreeze Bench Top, Labconco) for 24 hours. Solutions at higher concentrations (> ~25 mg mL^{-1}) were incubated at -80 °C for one hour prior to lyophilization to ensure complete freezing.

Hydrogel Formation. The reversibility of coiled coil formation due to the loss of helical secondary structure at high pH was exploited by Shen et al.[6] during hydrogel formation. Hydrogels were formed by transferring solutions of hydrogel forming concentration at high pH to the desired location, drying in-place and re-hydrating at neutral pH. This method of manipulating the structure of the H-domains does not lend itself to the manipulation of proteins with (integral) tertiary structure due to possible denaturation at high pH. Attempts at forming fluorescent hydrogels through a method as described by Shen et al.[6] results in protein denaturation and precipitation at high pH, and incomplete re-folding of the upon re-hydration at neutral pH. As such, a new, constant pH route to hydrogel formation was conceived. Hydrogel samples were formed by concentrating buffered protein samples to a concentration of 0.2X of the desire final concentration (~15-35 mg mL^{-1}). Samples were lyophilized and re-hydrated with one fifth the initial volume of water. Vortexing and mechanical mixing aided dissolution of the lyophilized protein. Air bubbles were removed from the hydrogel samples by centrifugation for 5 minute at 10,000 g. Samples prepared in microtiter plates were not vortexed or mixed but were centrifuged at 2,000 rpm for 5 minutes after re-hydration. 1 M NaOH and 1 M HCl were used to adjust pH when necessary.

8

Hydrogel pH was measured by two different fine range pH indication papers (Whatman and pHydrion).

Hydrogel erosion. Hydrogels re-hydrated with 10 μl of aqueous solution were prepared in 96 well microtiter plates and covered with 250 μl of 100 mM Na$_i$P, pH 7.0±0.2, 25 °C. Hydrogel erosion rate was determined by monitoring the protein concentration of a sample of open buffer solution over time (H-S-H, ε_{280} = 0.331, H-S-DSRED, ε_{280} = 0.971 and H-S-GFP-H, ε_{280} = 0.549 mg ml^{-1}). Release of GFP from H-S-H hydrogels was monitored by measuring the total fluorescence of the open buffer solution over time.

Hydrogel Rheology. Rheologicial oscillatory shear experiments were carried out using a TA Instruments AR1200 constant stress rheometer. An 8 mm stainless steel parallel plate geometery with 500 mm gap was used in small amplitude oscillatory shear experiments (1 to 100 to 1 rad s^{-1}) at a constant strain of 1% at a constant temperature of 22 °C. Evaporation of water during the course of the experiments was prevented by applying a small bead of mineral oil to the edge of the sample. Samples were tested within the linear viscoelastic regime.

Confocal microscopy. An Olympus IX-70 confocal microscope was used to examine mixed hydrogels of H-S-GFP-H and H-S-DSRED in H-S-H. XY plane images were captured under 10, 20 and 40X magnification. Green field images were excited with 488 nm Ar laser and emissions were filtered with a 510 nm long pass and a 539 nm short pass filters. Red field images were excited with a 568 nm Kr laser and emissions were filtered with a 605 nm band pass filter. Red and green images were taken separately and are shown with false color. Hydrogels were prepared, as described above, in micro-centrifuge tubes. Aliquots of the hydrogels were placed between two microscope slide cover slips (Fisher) for imaging.

Fluorescence measurements and determining FRET efficiency. Ten-μL mixed hydrogel samples of H-S-ECFP-H and H-S-DSRED and twenty-μL samples of ECFP and DSRED with 1:1 ratio of cyan to red fluorophores were prepared in black 384 well microtiter plates (Corning). Donor, cyan fluorophore, intensity control-samples of H-S-ECFP-H and H-S-H and ECFP in solution were also made. Fluorescence emission at 480 nm was measured using a SpectraMax M2 microplate reader (Molecular Devices) with excitation at 410 nm. FRET efficiency, D_{FRET}, was determined by donor emission quenching (Em. 480 nm) in the FRET sample in comparison to the cyan only control sample, $D_{FRET} = 1 - I_{DA}/I_D$, where I_{DA} is the fluorescence intensity of the donor in the presence of the acceptor and I_D is the fluorescence intensity of the donor only sample. Fluorophore concentrations were limited to those within a linear concentration-fluorescence intensity region.

Statistical analysis. Two-way ANOVA statistics were accomplished with the commercial statistical package contained in Microsoft Excel. Six separate data sets were measured for both the unmodified and hydrogel samples (n = 6). Statistical significance was achieved for parameters with a P value < 0.05.

Results and Discussion

Four new fluorescent protein constructs were produced: GFP and ECFP with N- and C-terminal H-domains (H-S-GFP-H and H-S-ECFP-H), and DSRED and GFP with a single N-terminal H-domain (H-S-DSRED and H-S-GFP). The N-terminal fusions in all the constructs include a randomly coiled, highly soluble, polyelectrolyte domain (S) located between the H-domain and the fluorescent protein. The H- and S-domains, described in reference 6 were selected as they are well characterized[6,7,17,18] and readily adaptable to our design. Both GFP

9

and ECFP are monomeric while DSRED forms tetramers[19]. Schematic representations and selected sequences of the fusion proteins studied herein are presented in Figure 1. Protein expression and hydrogel formation by lyophilization and re-hydration while maintaining protein structure are described above in the Experimental Section.

We found that upon re-hydration both functionalities were maintained; concentrated samples (>2 mM) of H-S-GFP-H, H-S-ECFP-H and H-S-DSRED form strong hydrogels with vivid color. Our interpretations of the physically cross-linked networks of H-S-GFP-H (and H-S-ECFP-H) and H-S-DSRED are shown in Figures 2a and 2b. Helical domains are illustrated as fully formed bundles, but it is possible that some helices, particularly those C-terminal in H-S-GFP-H, could be sterically prevented from completely overlapping with all the members of a given bundle. Small amplitude oscillatory shear experiments confirm the formation of hydrogels for 3.4 mM samples of H-S-GFP-H and H-S-DSRED as evidenced by a plateaued G' substantially greater than G'' over a wide range of frequencies[20] (Figures 2c and 2d). H-S-ECFP-H exhibits similar behavior (Supporting Information, SI). The significance of the physical cross-links formed by specific protein-protein interactions in the tetramers of DSRED is evident in the lack of hydrogel formation of H-S-GFP. While GFP modified with a single H-domain does exhibit some elastic character ($G' \sim 100$), likely due to its tendency to form weak dimers at high concentrations[21], it has approximately equal viscous character and does not form a stable hydrogel (SI).

The plateau G' of a 3.4 mM (7.5 wt%) sample of the engineered H-domains, H-S-H[5,6], was measured at a value of 560 ± 130 Pa. With two H-domains asymmetrically fused to the termini of GFP, H-S-GFP-H, a 3.4 mM (18 wt%) sample exhibits a plateau G' of 2500 ± 300 Pa (Figure 2c). H-S-DSRED of the same molar concentration (14 wt%), but half the concentration of H-domains, reaches a plateau G' of 1900 ± 200 Pa (Figure 2d). Measurements were made on 3 separate samples and errors are standard deviations. These results show that incorporating a folded globular protein into the hydrogel structure can contribute to the physical properties of the network in a beneficial way.

One proposed mechanism of protein release from the H-S-H hydrogel surface is that intramolecular and small-number intermolecular closed loops form at the surface of the structure and diffuse into solution within the time scale of H-domain exchange between coiled coils[6]. In our hands a 3.4 mM (7.5 wt%) hydrogel of H-S-H is more than 60% eroded after 2.5 hours in quiescent open buffer solution (Figure 3a) (previously published reports show near complete dissolution of a 7wt% sample in 150 minutes[17]). In the same time span only 8% and 14% of 3.4 mM samples of H-S-DSRED and H-S-GFP-H, respectively, erode (Figures 3a and 3b). The presence of a fluorescent protein alone is not sufficient to suppress the erosion rate, as GFP entrapped within a hydrogel of H-S-H quickly diffuses in to solution (Figure 3b).

It has been shown that the erosion rate of H-S-H can be suppressed by tailoring the number of H-domains per coiled coil, thereby impeding closed loop formation[17]. We hypothesize that the asymmetry of the H-S-GFP-H monomer likely results in steric hindrance, impeding facile contact between H-domains of one, or even two, monomers thus reducing the number of closed loops. Additionally, weak dimerization between GFP molecules could increase network connectivity, further preventing erosion. In the case of the H-S-DSRED hydrogels, the dual modalities of connectivity combined with the reduced formation of closed loops stemming from the tetrameric nature of DSRED are the likely causes of the observed suppressed erosion.

Building blocks with compatible hydrogel forming domains can be mixed and fluorescence loading can be modulated by varying the ratio of fluorescent to non-active monomers (Figure 4a). Mixed color hydrogels can also be made from mixtures of two or three different building blocks (Figures 4a-c). Mixed samples do form stable hydrogels but with rheological properties intermediate to those of the individual components (SI). Hydrogels of a single monomer are necessarily homogeneous, but importantly, mixed hydrogels appear to be macro-homogeneous. Confocal microscopy imaging shows an even distribution of both red and green fluorophores with minimal aggregation at the micrometer scale in a hydrogel comprising 250 μM each of H-S-GFP-H and H-S-DSRED and 3 mM H-S-H (Figures 4d, 4e and SI)

Fluorescent proteins were employed in part so that we could non-invasively probe the hydrogel structure using Förster Resonance Energy Transfer (FRET). With a suitable FRET pair separated by less than ~10 nm, energy from an excited donor (ECFP) can be non-radiatively transferred to an acceptor (DSRED)[22]. FRET efficiency drops with the inverse sixth power of the distance between flurophores as $E = 1/\left[1 + (r/R_0)^6\right]$ where R_0, the Förster radius, is the distance at which FRET efficiency is 50% (R_0 = 4.0 nm, see SI). Here we compare the FRET efficiency in mixed hydrogels of H-S-ECFP-H, H-S-DSRED and H-S-H and unmodified ECFP and DSRED in solution to show the macro-homogeneity of mixed hydrogels and to demonstrate that compatible monomers can be used to control a multi-component and distance-dependent process (Figure 4f). If substantial segregation of building block components of the mixed hydrogels occurs upon re-hydration, one would expect minimal FRET efficiency. This is not the case, as a measurable FRET signal is observed in hydrogels with greater than 50 μM total fluorophore (1:1 ECFP to DSRED). This result, along with the confocal imagery (Figures 4d and 4e), suggests that mixed hydrogels are macro-homogeneous.

The significant reduction in FRET efficiency, as quantified by two-way ANOVA, seen in the hydrogel samples compared to solutions of equal concentration (P = 0.028, n = 6) is likely due in part to a reduction in rotational freedom of the fluorophore. Restricting fluorophore rotation can result in a decrease in R_0[23]. Additionally, the decrease in FRET efficiency could be due to an increase in fluorophore separation, r, due to slight swelling of the hydrogel structure. A similar decrease in FRET efficiency has recently been shown with alginate hydrogels modified with synthetic fluorophores[24]

Conclusions

The FRET experiments and hydrogel rheology results demonstrate that we can control an intra-gel, interactive, process such that hydrogel architecture and bioactivity can be independently modulated. This level of control may advance many potential applications including active tissue engineering scaffolds[3], heterogeneous catalysis, and electrode and surface modifications for biocatalysis[25,26]. It may be possible to incorporate an entire metabolic pathway into a single stable hydrogel. The design approach is not specific to fluorescent proteins and H-domains but is, in principal, applicable to any bioactive protein and to other peptide motifs and protein domains that form physical cross-links. The hydrogel-forming motifs can be varied to include response to external stimuli[27,28] and can be tailored to control connectivity within the network[18]. More sophisticated hydrogels can be readily created by engineering additional bi-functional building blocks.

Supporting Information Available. Results of small amplitude oscillatory shear experiments of H-S-ECFP-H, H-S-GFP and mixed hydrogels. Confocal microscopy images of mixed hydrogels. Calculation of the Förster radius, R_0. This material is available free of charge via the Internet at http://pubs.acs.org

12

Protein constructs
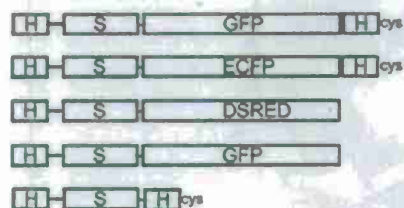
Protein sequences

H = SGDLENE VAQLENE VRSLEDE AAELEQK
        VSRLKNE IEDLKAE
S = (AGAGPEG)$_{10}$
GFP = green fluorescent protein (S65T)
ECFP = enhance cyan fluorescent protein
DSRED = DSRED express

Figure 1. Fusion protein constructs and partial sequences. 'H' signifies an α-helical domain and S, a randomly coiled polyelectrolyte domain[6]. DSRED and GFP(S65T) are described elsewhere[19, 29]
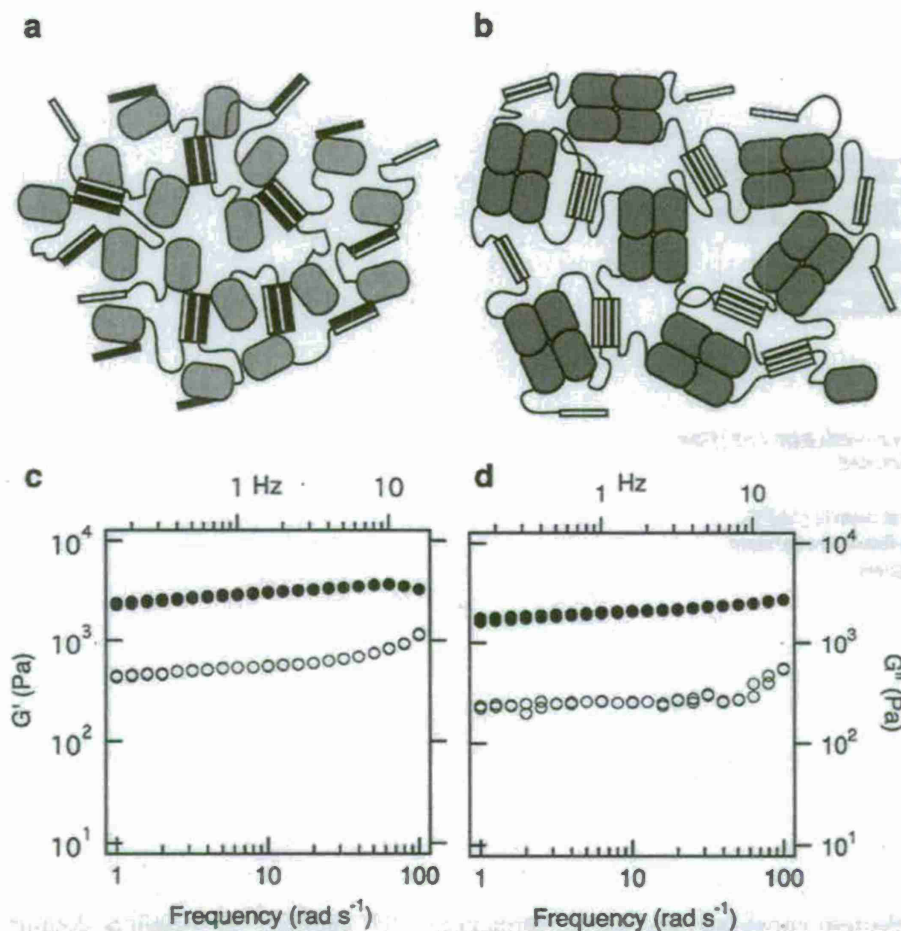
13

Figure 2. Hydrogels from bi-functional fluorescent building blocks. (a,b) Schematic representations of the physically cross-linked fluorescent protein hydrogels, ovals are fluorescent proteins, bars are H-domains, (a) H-S-GFP-H and (b) H-S-DSRED. Cross-links between H-S-GFP-H monomers are made by the formation of tetrameric coiled coil bundles of the H-domains and through disulfide bonds between two terminal cystines[6]. Cross-links between H-S-DSRED monomers are made by the formation of DSRED tetramers[19] and also by the formation of coiled coil bundles of the N-terminal H-domains. (c,d) Examples of small amplitude oscillatory shear experiments of 3.4 mM samples of (c) H-S-GFP-H and (d) H-S-DSRED, G′,● and G″, ○, 100 mM Na$_i$P and pH 7.0±0.2. Molar concentration is stated as the number of moles of protein per liter of aqueous solution used for re-hydration.
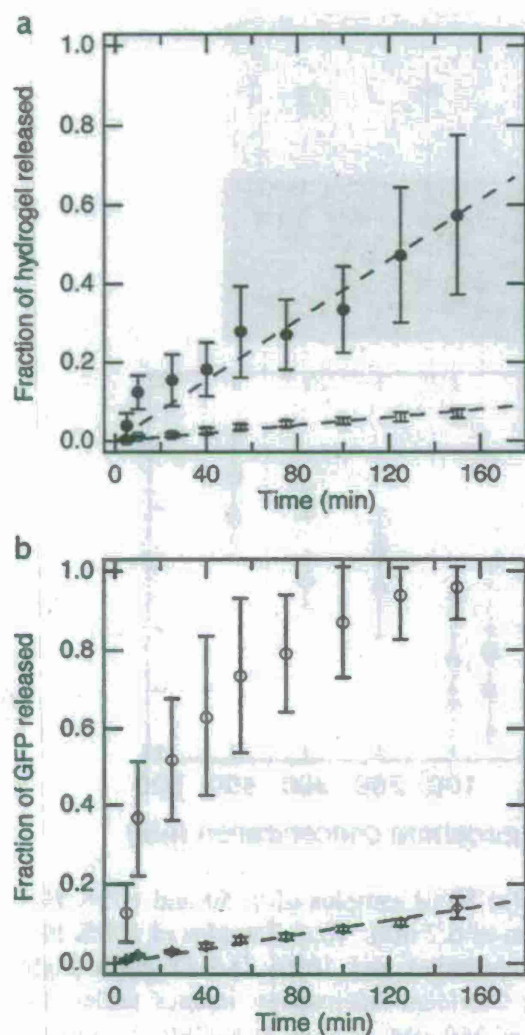
Figure 3 Hydrogel erosion. (a,b) Fractional release of 3.4 mM samples of (a) H-S-H, ●, and H-S-DSRED, ■, and (b) fractional release of 3.4 mM GFP from a mixed hydrogel of 3.4 mM H-S-H with 3.4 mM of entrapped GFP, □, and 3.4 mM H-S-GFP-H, ○. Error bar are the standard deviation of at least 6 replicates. Linear fits to the H-S-H, H-S-DSRED and H-S-GFP-H data sets result in slopes of 400, 53 and 93 pmoles $min^{-1}cm^{-2}$ (MW_{H-S-H}, 22470, $MW_{H-S-DSRED}$, 41875 , $MW_{H-S-GFP-H}$, 52710).
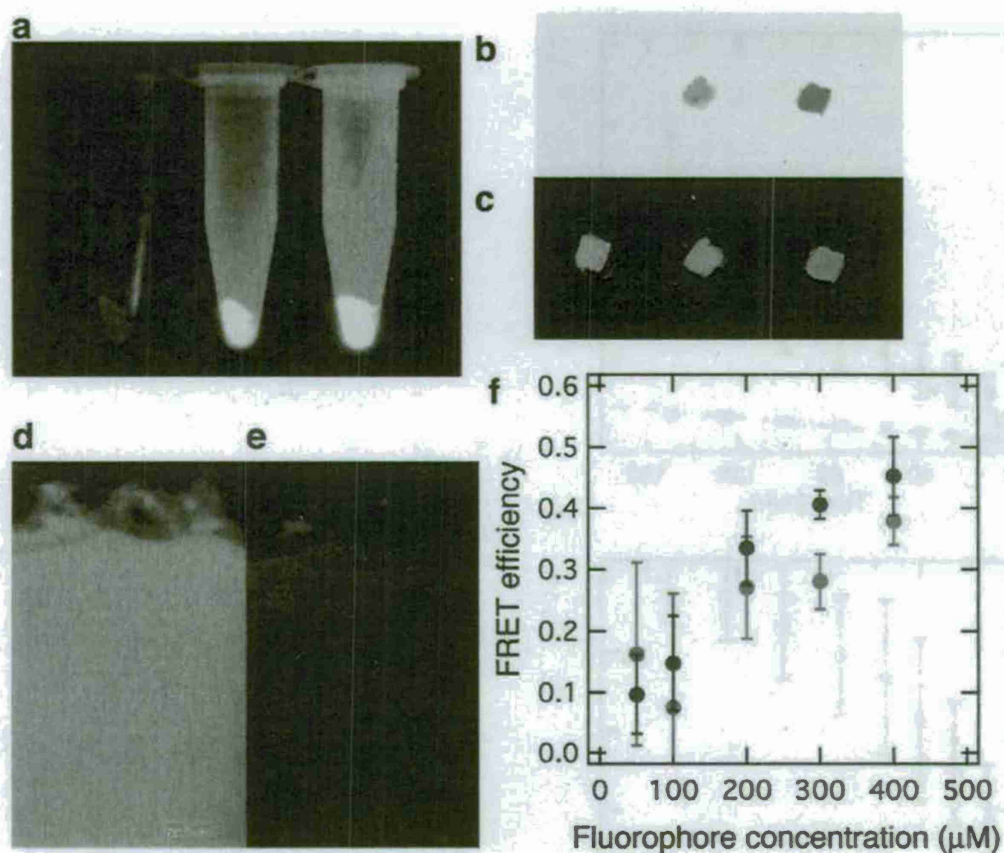
Figure 4. Mixed flourescent protein hydrogels. (a) 50 μl samples of 0, 50 and 100% H-S-GFP-H in H-S-H, constant H-domain concentration of 6.7 mM. (b,c) Samples of 100% H-S-GFP-H (left), 50/50% H-S-GFP-H/H-S-DSRED (middle) and 100% H-S-DSRED (right) under (b) visible light and (c) UV light. (d,e) Confocal microscopy images under 20X magnification of the edge of a mixed hydrogel of 250 μM of each H-S-GFP-H and H-S-DSRED and 3 mM H-S-H, (d) Ex(488 nm), 505 nm long pass and 545 nm short pass emission filters and, (e) Ex(568 nm), 705 nm short pass emission filter. (f) Donor intensity FRET efficiency of unmodified ECFP and DSRED (1:1, ECFP:DSRED) in solution,● and hydrogels of H-S-ECFP-H, H-S-DSRED (1:1, HSECFPH:HSDSRED) and H-S-H with constant H-domain concentration of 6.7 mM, ●, pH 7.

References and Notes

1.   Cushing, M. C.; Anseth, K. S., MATERIALS SCIENCE: Hydrogel Cell Cultures. *Science* 2007, 316, (5828), 1133-1134.

2.   Peppas, N. A.; Hilt, J. Z.; Khademhosseini, A.; Langer, R., Hydrogels in biology and medicine: From molecular principles to bionanotechnology. *Advanced Materials* 2006, 18, (11), 1345-1360.

3.   Lutolf, M. P.; Hubbell, J. A., Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nature Biotechnology* 2005, 23, (1), 47-55.

4.   Ulijn, R. V.; Bibi, N.; Jayawarna, V.; Thornton, P. D.; Todd, S. J.; Mart, R. J.; Smith, A. M.; Gough, J. E., Bioresponsive hydrogels. *Materials Today* 2007, 10, (4), 40-48.

5.   Petka, W. A.; Harden, J. L.; McGrath, K. P.; Wirtz, D.; Tirrell, D. A., Reversible hydrogels from self-assembling artificial proteins. *Science* 1998, 281, (5375), 389-392.

6.   Shen, W.; Lammertink, R. G. H.; Sakata, J. K.; Kornfield, J. A.; Tirrell, D. A., Assembly of an artificial protein hydrogel through leucine zipper aggregation and disulfide bond formation. *Macromolecules* 2005, 38, (9), 3909-3916.

7.   Xu, C. Y.; Breedveld, V.; Kopecek, J., Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules* 2005, 6, (3), 1739-1749.

8.   Mi, L. X.; Fischer, S.; Chung, B.; Sundelacruz, S.; Harden, J. L., Self-assembling protein hydrogels with modular integrin binding domains. *Biomacromolecules* 2006, 7, (1), 38-47.

9.   Wang, C.; Stewart, R. J.; Kopecek, J., Hybrid hydrogels assembled from synthetic polymers and coiled-coil\ protein domains. *Nature* 1999, 397, (6718), 417-420.

10.   Zhang, S. G.; Holmes, T.; Lockshin, C.; Rich, A., Spontaneous Assembly of a Self-Complementary Oligopeptide to Form a Stable Macroscopic Membrane. *Proceedings of the National Academy of Sciences of the United States of America* 1993, 90, (8), 3334-3338.

11.   Schneider, J. P.; Pochan, D. J.; Ozbas, B.; Rajagopal, K.; Pakstis, L.; Kretsinger, J., Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide. *Journal of the American Chemical Society* 2002, 124, (50), 15030-15037.

12.   Ozbas, B.; Rajagopal, K.; Schneider, J. P.; Pochan, D. J., Semiflexible chain networks formed via self-assembly of beta-hairpin molecules. *Physical Review Letters* 2004, 93, (26).

13.   Yokoi, H.; Kinoshita, T.; Zhang, S. G., Dynamic reassembly of peptide RADA16 nanofiber scaffold. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102, (24), 8414-8419.

17

14.   Ehrick, J. D.; Deo, S. K.; Browning, T. W.; Bachas, L. G.; Madou, M. J.; Daunert, S., Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. *Nature Materials* 2005, 4, (4), 298-302.

15.   Murphy, W. L.; Dillmore, W. S.; Modica, J.; Mrksich, M., Dynamic Hydrogels: Translating a Protein Conformational Change into Macroscopic Motion. *Angewandte Chemie International Edition* 2007, 46, (17), 3066-3069.

16.   Ringler, P.; Schulz, G. E., Self-assembly of proteins into designed networks. *Science* 2003, 302, (5642), 106-109.

17.   Shen, W.; Zhang, K. C.; Kornfield, J. A.; Tirrell, D. A., Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nature Materials* 2006, 5, (2), 153-158.

18.   Oshea, E. K.; Rutkowski, R.; Stafford, W. F.; Kim, P. S., Preferential Heterodimer Formation by Isolated Leucine Zippers from Fos and Jun. *Science* 1989, 245, (4918), 646-648.

19.   Bevis, B. J.; Glick, B. S., Rapidly maturing variants of the Discosoma red fluorescent protein (DsRed) (vol 20, pg 83, 2002). *Nature Biotechnology* 2002, 20, (11), 1159-1159.

20.   Kavanagh, G. M.; Ross-Murphy, S. B., Rheological characterisation of polymer gels. *Progress in Polymer Science* 1998, 23, (3), 533-562.

21.   Ohashi, T.; Galiacy, S. D.; Briscoe, G.; Erickson, H. P., An experimental study of GFP-based FRET, with application to intrinsically unstructured proteins. *Protein Sci* 2007, 16, (7), 1429-1438.

22.   Wallrabe, H.; Periasamy, A., Imaging protein molecules using FRET and FLIM microscopy. *Current Opinion in Biotechnology* 2005, 16, (1), 19-27.

23.   Gryczynski, Z.; Gryczynski, I.; Lakowicz, J. R., Basics of Fluorescence and FRET. In *Molecular Imaging: FRET Micrscopy and Spectroscopy;* Periasamy, A., Day, R.N., Eds; The American Physiological Society Methods in Physiology Series; Oxford University Press: New York, NY, 2005; p 21-56.

24.   Kong, H. J.; Kim, C. J.; Huebsch, N.; Weitz, D.; Mooney, D. J., Noninvasive probing of the spatial organization of polymer chains in hydrogels using fluorescence resonance energy transfer (FRET). *Journal of the American Chemical Society* 2007, 129, (15), 4518

25.   Heller, A., Electron-conducting redox hydrogels: design, characteristics and synthesis. *Current Opinion in Chemical Biology* 2006, 10, (6), 664-672.

26.   Calabrese Barton, S.; Gallaway, J.; Atanassov, P., Enzymatic biofuel cells for Implantable and microscale devices. *Chemical Reviews* 2004, 104, (10), 4867-4886.

27.   Chockalingam, K.; Blenner, M.; Banta, S., Design and application of stimulus-responsive peptide systems. *Protein Engineering, Design and Selection* 2007, 20, (4), 155-161.

28.   Banta, S.; Megeed, Z.; Casali, M.; Rege, K.; Yarmush, M. L., Engineering protein and peptide building blocks for nanotechnology. *Journal of Nanoscience and Nanotechnology* 2007, 7, (2), 387-401.

29.   Heim, R.; Cubitt, A. B.; Tsien, R. Y., Improved Green Fluorescence. *Nature* 1995, 373, (6516), 663-664.

Chapter 3

Wheeldon, I.R., Gallaway, J.W., Calabrese Barton, S., and **Banta, S.** (2008) "Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks" *Proceedings of the National Academy of Sciences of the United States of America* **105**(40) 15275-15280.

Bioelectrocatalytic hydrogels from electron-conducting metallo-polypeptides co-assembled with bifunctional enzymatic building blocks

Ian R. Wheeldon[*], Joshua W. Gallaway[*], Scott Calabrese Barton[†], and Scott Banta[*]

[*]Department of Chemical Engineering, Columbia University, 500 West 120th Street, New York, NY, USA, 10027

[†]Department of Chemical and Materials Engineering, Michigan State University, 1262 Engineering Building, East Lansing, MI, USA 48824-1226

Corresponding Author:

Scott Banta

500 West 120th Street, 826 Mudd

New York, NY, USA

10027

Phone: (212) 854-7531

Fax: (212) 854-3054

Email: sbanta@cheme.columbia.edu

Abstract

Here we present two bifunctional protein building blocks that co-assemble to form a bioelectrocatalytic hydrogel that catalyzes the reduction of dioxygen to water. One building block, a metallo-polypeptide based on a previously designed triblock poly-peptide, is electron-conducting. A second building block is a chimera of artificial α-helical leucine zipper and random coil domains fused to a poly-phenol oxidase, small laccase (SLAC). The metallo-polypeptide has a helix-random-helix secondary structure and forms a hydrogel via tetrameric coiled-coils. The helical and random domains are identical to those fused to the poly-phenol oxidase. Electron-conducting functionality is derived from the divalent attachment of an osmium bis-bipyrdine complex to histidine residues within the peptide. Attachment of the osmium moiety is demonstrated by mass spectroscopy (MS-MALDI-TOF) and cyclic voltammetry. The structure and function of the α-helical domains are confirmed by circular dichroism spectroscopy and by rheological measurements. The metallo-polypeptide shows the ability to make electrical contact to a solid-state electrode and to the redox centers of modified SLAC. Neat samples of the modified SLAC form hydrogels, indicating that the fused α-helical domain function as a physical cross-linker. The fusion does not disrupt dimer formation, a necessity for catalytic activity. Mixtures of the two building blocks co-assemble to form a continuous supramolecular hydrogel that, when polarized, generates a catalytic current in the presence of oxygen. The specific application of the system is a biofuel cell cathode, but this protein engineering approach to advanced functional hydrogel design is general and broadly applicable to biocatalytic, biosensing and tissue engineering applications.

Introduction

Protein engineering provides the toolset to design and produce peptides and proteins that form the building blocks of new bio-inspired materials. The toolset allows for the manipulation of natural and artificial DNA sequences encoding the peptides or proteins of interest, and the subsequent biological production of the translated products. The methodology is powerful in that it allows for exact control over the identity and sequence of each residue and, consequently, the structural folding patterns of the resultant peptides or proteins (1). And, just as function stems from structure, it is also controlled within the protein engineering scheme of materials design. There are a number of successful examples of hydrogels designed for tissue engineering and drug delivery applications that use functional protein domains to obtain structural responsiveness to environmental cues (2,3,4). These examples include responsiveness to pH (5), temperature (6), shear stress (7) and ligand binding (8, 9) among others (2, 3 and reference therein, 10). A wider range of applications, such as bioelectrocatalysis and biosensing, will benefit from the advantages of protein-based materials design provided that a second functionality, like catalytic activity or electron conduction, can be engineered into the biologically-inspired building blocks. It is here that our current interests lie; the development of multi-functional proteins that self-assemble into supramolecular structures and have bioelectrocatalytic functionally.

In the past decade, important advancements have been made in the development of bioelectrocatalysis for biosensing and biofuel cell applications (11, 12). In particular, much success has been achieved with systems that couple biological recognition with an electrochemical signal by electrically connecting the redox active centers of oxidoreductases with a solid-state electrode (11, 13, 14). The most advanced systems are synthetic redox polymer-enzyme hydrogels (15, 16) . The redox polymer is often poly(vinyl)pyridine or poly(vinyl)imidazole (PVI) with pendent osmium moieties complexed to a fraction of the nitrogen-containing aromatic rings (14-17). Osmium complexes are used as electron mediators as the redox potential can be tailored to a desired value through ligand substitution, and the complexes are more photostable than other transition metal complexes such as those based on ruthenium (17). Osmium metal and most compounds are generally considered nontoxic (17), although the tetroxide (not used in this work) is highly toxic in cases of acute exposure. The toxicological properties of the salts used here have not been evaluated.

While much work has been focused on optimizing electron diffusion within the redox polymer-enzyme hydrogels and on improving catalytic performance, little work has been done in evaluating the nano- and micro-scale structure of these hybrid materials. In general, enzymes are non-specifically entrapped within the hydrogels by electrostatic interactions, and high loadings of redox moieties are required to ensure electrical contact to the immobilized enzymes (14, 15). There is also evidence that other similar synthetic polymer-enzyme hydrogels are inherently heterogeneous in structure and activity (18,19,20).

Here we propose a bioelectrocatalytic hydrogel constructed from bifunctional protein building blocks. Mixtures of electron conducting and catalytic building blocks will self-assemble into a bioelectrocatalytic supramolecular hydrogel. The physical properties and bulk function of the hydrogel can be independently tuned, as those properties are dependent on the identity and amount of each building block. The building blocks are based on a previously designed triblock polypeptide, here termed HSH, comprised of two α-helical leucine zipper domains (H-domain) separated by a randomly coiled domain (S-domain) (5,

22

21). The helical domains assemble into tetramic coiled-coils thus forming an ordered supramolecular hydrogel. Catalytic functionality is derived from an oxidoreductase to which H-domains, identical to those of the tri-block polypeptide (HSH), have been genetically fused. Electron conduction functionality is achieved through modification of the HSH building block with a redox moiety.

We explore this class of multi-functional hydrogels by creating a bioeletrocatalytic system that catalyzes the reduction of dioxygen to water at neutral pH, which has proven difficult with synthetic redox-polymer systems. A poly-phenol oxidase from *Streptomyces coelicolor* catalyzes the oxidation of a wide range of substrates along with the concomitant reduction of dioxygen to water (22), and it is here used as the basis of the enzymatic building block. The scheme of attaching osmium redox moieties to PVI is an opportune starting point to begin adding electron-conducting functionality to the system though attachment to the imidazole side chain of histidine residues within the hydrogel forming polypeptide. When co-assembled, the two protein building blocks produce a hydrogel that can be used as an electrode surface modification for a cathode in a biofuel cell or as an oxygen biosensor. Thus we demonstrate that the design and the co-assembly of bifunctional protein building blocks is a general protein engineering method for the creation of multi-functional biomaterials.

Results and Discussion

The bioelectrocatalytic hydrogel is made from two compatible building blocks, an engineered bifunctional metallo-polypetide that self assembles into a hydrogel and is electron conducting, and a poly-phenol oxidase, modified with an N-terminal α-helical domain (H-domain), that is enzymatic and self-assembles into a supramolecular network. Both the metallo-polypeptide and modified poly-phenol oxidase are bifunctional in that they exhibit two distinct functionalities. Physical cross-linking functionality of the metallo-polypeptide is derived from the previously developed tri-block polypeptide, HSH, that forms a supramolecular hydrogel due to coiled coil formation of α-helical leucine zipper domains (5). A second functionality, electron conduction, is derived from the divalent attachment of osmium bis-bipyridine via ligand exchange (14-17) with suitable amino acid side chains. The modified poly-phenol oxidase, a small laccase (SLAC) (22), was genetically engineered to exhibit physical cross-linking functionality through the addition of the H- and S-domain encoding sequences to the N-terminus of the gene. The protein constructs, relevant amino acid sequences, and the osmium bis-bipyridine complex are shown in Figure 1. Of specific importance is the location of histidine residues within HSH as the imidazole side chains are the putative location of osmium attachment. N-terminal to the first H-domain is a 6 × histidine tag. There are also two single histidines, one within each of the linker regions, N- and C-terminal to the S-domain. The full amino acid sequence is provided in the supporting information (SI).

The attachment procedure of osmium bis-bipyridine to HSH is uncomplicated. A 20:1 molar ratio of osmium to purified HSH was dissolved in deionized water and heated to 70 °C under argon, while refluxing. Two metallo-polypeptide products were synthesized, OsHSH-1 and OsHSH-2, from reactions lasting 6 and 18 hours, respectively. Both products are purple in color. Reaction for 6 hours yields a polypeptide product that migrates on a polyacrylamide gel under denaturing conditions to a distance equivalent to that of HSH (Figure 2a). The product also contains polypeptides that migrate to twice the molecular

23

weight (MW) of a single HSH monomer as well as a smearing of material between 90 and 120 kDa. The diffuse bands are dimers, trimers and tetramers of HSH cross-linked with osmium complex that have undergone two ligand substitutions.

The MW of HSH as measured by mass spectrometry (MS MALDI-TOF) is 22,050 Da but it is retarded in the electrophoretic gel to an apparent MW of 32 kDa. The same electrophoretic behavior is seen with monomers of OsHSH-1 as MS MALDI-TOF analysis reveals a species with a MW of 22,060 with two additional broad peaks at 22,570 and 23,080 Da (SI). The MW of an attached osmium bis-bypridine complex is 502.6, 538.1 or 573.5, depending on the number of chloro ligands and ionic chlorides associated with the complex, suggesting that OsHSH-1 is a mixed population of unmodified HSH, HSH with a single osmium complex, and HSH to which two complexes have been bound.

After reaction for 18 hours, the majority of the polypeptide product OsHSH-2 appears to be highly polymerized, based on limited migration of product into the gel (Figure 2a). A smear of protein appears at a MW of 120 kDa and larger, as well as smaller diffuse bands at approximately 90 kDa and 60 kDa and a faint band at 32 kDa. Similar to OsHSH-1, the higher MW species are multimers of osmium modified HSH. The high degree of multimerization seen in OsHSH-2 suggests that it no longer functions as a monomeric (or dimeric, or trimeric) building block in a supramolecular hydrogel, but more likely, will function as a chemically cross-linked high MW hydrogel. It is included here as evidence of osmium bis-bipyridine ligand substitution and as it is a distinct, viable product in and of itself.

The standard redox potential, $E°$, of the osmium complex is dependent on the identity of each ligand and, as such, can be tailored to a desired value (17). Ligand substitution can also be confirmed by measurement of a shift in redox potential. Substitution of a chloride ligand of $[Os(bpy)_2Cl_2]Cl$ by imidazole or poly(N-vinylimidazole) shifts $E°$ from 190 mV vs standard hydrogel electrode (SHE) to 400–470 mV; substitution of the second chloride results in a further shift in $E°$ to 710 – 850 mV(15). The major voltammetric peaks of OsHSH-1 and OsHSH-2 in dilute solution are centered at potentials of 470 and 800 mV, respectively (Figure 2b). Minor peaks are also observed at 800 mV for OsHSH-1 and 470 mV for OsHSH-2, but in light of the *a priori* estimates of potential shifts resulting from ligand exchange, reaction for 6 hours results predominately in single ligand substitution of chloride by an imidazole side chain of a histidine residue. At longer reaction times, a second chloride ligand is exchanged. The assertion that $[Os(bpy)_2Cl_2]Cl$ undergoes a second ligand exchange is supported by the apparent multimerization of OsHSH-2; the complex acts as a cross-link between hydrogel monomers at one residue per polypeptide resulting in a chemically cross-linked network of high MW.

MS MALDI-TOF analysis of protease-digested samples of OsHSH-1 also supports the proposed ligand exchange. Five fragments seen in trypsinized samples of OsHSH-1 are not observed in similarly treated HSH and correspond to the MW of an expected histidine containing peptide fragment plus an osmium bis-bipyridine complex with zero, one or two chlorides (SI). The same is true of four fragments seen in AspN digested samples (SI).

Physical cross-linking of HSH into a supramolecular hydrogel is mediated by the agglomeration of the H–domains into coiled-coil bundles, thus secondary structure is

24

essential to hydrogel functionality. Circular dichroism (CD) spectroscopy of OsHSH-1 and OsHSH-2 indicates that the secondary structure is not compromised as evidenced by the spectral minima at 208 and 222 nm, characteristics of α-helices (23) (Figure 1c). Deconvolution of the spectra using the Johnson and Henessey method (24) shows that there is no statistical difference between the helical content of OsHSH-1 and HSH (36.3 ±2.2 % and 36.2 ±2.4 %, respectively; n=5, p=0.979) and only a slight, but statistically relevant, decrease in OsHSH-2 (32.2 ±2.0%; n=5, p=0.0199). Additionally, both osmium-modified polypeptides retain the designed pH dependent structure (5) as seen in the loss of helical content with increasing pH (Figure 2c inset).

Rheological experiments show that upon rehydration in aqueous solution buffered to pH 7.0, 7.5 wt% samples of OsHSH-1 and OsHSH-2 form hydrogels. In small amplitude oscillatory shear experiments both OsHSH-1 and OsHSH-2 exhibit storage modulus, G′, plateau values consistently greater than the elastic modulus, G″, a characteristic indicative of hydrogel formation (25). A 7.5 wt% sample of OsHSH-1 at pH 7.0 and 22°C exhibits a G′ plateau value of 380±70 Pa, and a similar sample of OsHSH-2 has a G′ of 640±40 Pa. In comparison, a 7.5 wt% sample of HSH at pH 7.0 has a G′ of 560±130 Pa (26). The G″ values of all samples reach plateaus below 100 Pa. The lower phase transition limit was not rigorously studied, but hydrogel formation of OsHSH-1 and OsHSH-2 was visually confirmed at concentrations as low as 6.0 wt%; the samples did not appear to flow or deform under gravity on a time scale of tens of hours.

The function of the H–domains is also confirmed by the formation of hydrogels of OsHSH-1 through a change in pH from highly basic to near-neutral. At pH 12, secondary structure is lost and OsHSH-1 does not form a hydrogel, but at pH 7 the metallo-polypeptide regains secondary structure and physical cross-linking functionality. Five-µL samples of 7.5 wt% OsHSH-1 at pH 12 were aliquoted on to glassy carbon electrodes and dried overnight. Hydrogels were formed upon re-hydration at neutral pH. OsHSH-2 does not exhibit the same functional pH dependence, as 7.5 wt% samples at pH 12 do form hydrogels regardless of the loss of secondary structure. Examples of a 7.5 wt% OsHSH-1 hydrogel at pH 7 and solution of the same concentration at pH 12 are shown in Figure 3a.

Hydrogel strength, in terms of G′, can be increased with the addition of building blocks as this increases the helical cross-linking concentration. A 15 wt% sample of OsHSH-1 shows an increase in G′ to 950±170 Pa (Figure 3b). We have previously shown that building blocks with compatible helices form macrohomogeneous mixed hydrogels (26). Therefore by mixing HSH with OsHSH-1 it is possible to independently tune hydrogel strength and osmium concentration by varying the amounts and ratio of HSH to OsHSH-1.

The high degree of multimerization seen in OsHSH-2 contributes substantially to the overall cross-linking of the network, but physical cross-linking due to coiled-coil formation is still thought to occur. A decrease in the ratio of CD signal at 222 and 208 nm from 0.84 to 0.78, under native conditions compared with samples in 50% trifluoroethanol (TFE), a compound that is known to disrupt coiled-coil formation, suggests that OsHSH-2 forms coiled-coils (27). A similar decrease in the signal ratio is observed with HSH. The typical ratios of elipticity at 222 to 208 nm for coiled coils, 1.0, and single helices, 0.8 (27), are not strictly applicable in this case as the polypeptides are only 32-36% helical. The random coil

25

domain (S-domain) separating the helices contributes to the negative dichroic signal between 210 and 195 nm (23) decreasing the expected signal ratio of 222 to 208 nm. We point to the change in signal ratios of both HSH and OsHSH-2 as support for the change in tertiary structure under native and coil-disruptive conditions.

It has been shown that the predominate mechanism of erosion of HSH hydrogels is diffusion of small-number closed loop complexes formed via strand exchanged between coiled-coils (28). When a close loop forms at the surface of the hydrogel it can diffuse into open buffer solution on a time scale faster than reattachment to the greater network via a subsequent strand exchange. With the additional cross-linking due to the osmium modification, erosion rate is suppressed as closed loop formation is impeded, or, at least, the number of monomers required for closed loop formation is increased. HSH is completely eroded after 180 minutes (28). After 5.5 hours OsHSH-1 is more than 95% eroded and OsHSH-2 is only$17\pm4\%$ eroded (Figure 3c). Erosion rate can be further suppressed by treating re-hydrated hydrogels with glutaraldehyde thus further preventing closed loop formation. Chemical cross-linking between amine groups results in near complete suppression of erosion after 5 minutes of treatment with 1 % glutaraldehyde solution (SI). Effects of glutaraldehyde cross-linking on the nano- and micro-structure of the supramolecular network have not yet been studied. Additionally, non-specific chemical cross-linking could have an effect on the electron diffusion coefficient within OsHSH-1 hydrogels and on the catalytic properties of enzymes within an OsHSH-1 hydrogel. Other strategies for controlling the erosion rate of coiled coil cross-links have been studied elsewhere (26, 28).

Electron conducting functionality of the metallo-polypeptide hydrogels is best demonstrated through the intended application, the bioelectrocatalyic reduction of dioxygen to water. To do so, a second building block, one that co-assembles with OsHSH-1 into supramolecular hydrogel and exhibits catalytic activity towards the reduction of dioxygen to water, is required. We have recently shown that the H- and S-domains of HSH can be appended to the termini of fluorescent proteins without disrupting native structure and function, suggesting that a catalytic building block could be made via a similar design (26). As SLAC has previously been shown to be active at neutral pH in the presence of osmium bis-bipyridine (29) it is an ideal candidate for the catalytic building block. The sequences encoding the H- and S-domains were ligated to the N-terminus of the SLAC gene and the chimeric protein construct, HS-SLAC, was expressed.

In the initial work characterizing SLAC, Machczynski et al. demonstrates that dimer formation is required for catalysis (22), this is also the case with HS-SLAC. Eluate from nickel affinity chromatography purified HS-SLAC contains putative monomers and dimers, but only the higher molecular weight protein (the putative dimers) exhibit catalytic activity in a native in-gel activity assay (Figure 4a). The active band in the native gel corresponds to a protein that runs on a denaturing gel to an apparent MW between 120 and 160 kDa (SI). The calculated MW of the HS-SLAC dimer is 97,398 Da, but in light of the electrophoretic behavior of HSH and OsHSH-1, the high apparent MW is not unexpected. Unmodified SLAC dimers were used as a positive control, and as it also exhibits in-gel activity with 30 mM dimethoxyphenol (DMP) in air-saturated solution at pH 7.0.

26

Further characterization of HS-SLAC dimers with dilute solution activity assays reveals an apparent Mechalis-Menton constant, K_M, towards DMP in air saturated solution of 4.2 ± 0.5 mM and a k_{cat} of 2 ± 0.1 min^{-1} at pH 7.0 (Figure 4b). Comparison to unmodified SLAC reveals that, while the N-terminal addition of the H- and S-domains did not prevent dimer formation, catalytic activity is impaired. Under similar conditions (pH 7.2, air-saturated solution and with DMP as co-substrate) unmodified SLAC exhibits a k_{cat} of 350 min^{-1} and an apparent K_M towards DMP of 4 mM (22).

A bifunctional catalytic building block requires that the structure and function of the cross-linking domains also be intact. CD spectroscopy shows that the appended H-domain does have helical secondary structure as evidenced by the introduction of spectra minima at 208 and 222 nm, characteristics of α-helices (Figure 4c). Deconvolution of the SLAC and HS-SLAC spectra with the CDSSTR alogorithm (30) suggests an increase in total helical content from 0 to $7.0\pm1.0\%$. Cross-linking functionality of the H-domain was confirmed by hydrogel formation of a 32 wt% sample of HS-SLAC at pH 7.0. Hydrogel formation was visually confirmed in the same manner as the lower phase transition limit of OsHSH-1 and -2, the sample did not flow or deform under gravity. The exceedingly high concentration of HS-SLAC (320 mg mL^{-1}) required to attain an equal helical concentration as 7.5 wt% OsHSH-1 was a major impediment to further rheological characterization. Visual confirmation of hydrogel formation, as compared to the non-gel forming unmodified SLAC at the same concentration and under the same conditions, provides sufficient evidence for physical cross-linking functionality.

Figure 5a is a schematic representation of a mixed hydrogel of OsHSH-1 and HS-SLAC. The active dimers of HS-SLAC are incorporated within the network of OsHSH-1 by means of physical cross-linking at the coiled-coil bundles. Cross-linking within the network occurs by three modes: coiled-coil formation; dimer formation of HS-SLAC; and osmium-mediated cross-linking within OsHSH-1. Cross-linking due to coiled-coil formation in OsHSH-1 and HS-SLAC was demonstrated by hydrogel formation of neat samples of both constructs at pH 7 (Figure 3a and text above). The helical domain fused to the N-terminus of HS-SLAC is identical to those of OsHSH-1, and will form coiled-coils within the metallo-polypeptide network. SLAC is enzymatically active only as a dimer, therefore a catalytically active hydrogel necessarily includes cross-linking due to protein-protein interactions between HS-SLAC monomers (Figure 4a). Thirdly, as seen in the electrophoretic analysis of OsHSH-1 and OsHSH-2 (Figure 2a) the modification of HSH with osmium bis-bipyridine results in the formation of dimers, trimers, and tetratmers, that in the supramolecular hydrogel contribute, in part, to the total cross-linking of the network. While it is experimentally difficult to quantify the cross-linking density of each mode, there is evidence demonstrating the existence of each. In the case of neat OsHSH-1 and HS-SLAC hydrogels it is clear (based on evidence presented above) that coiled-coil formation is the dominant mode, this holds true for mixed hydrogels of OsHSH-1 and HS-SLAC as well.

A mixed hydrogel of 7.5 wt% OsHSH-1 and 2.5 wt% HS-SLAC on glassy carbon in aqueous solution buffered to pH 7.0 with 100 mM Na$_i$P at 25 °C produces a catalytic current of 12.3 ± 5.2 μA cm^{-2} 15 minutes after rehydration (Figure 5b). No catalytic current is observable under N$_2$ sparging, and an osmium-free control of 7.5 wt% HSH and 2.5 wt% HS-SLAC shows no redox activity. Bioelectrocatalysis requires electron transfer within the electrode surface modification though self-exchange between redox moieties and via

heterogeneous electron transfer from the redox moieties to the redox centers of immobilized enzymes. The generation of a catalytic current from a mixed OsHSH-1 and HS-SLAC hydrogel demonstrates the ability of OsHSH-1 to fulfill the redox mediator functions. In addition, the catalytic current demonstrates that HS-SLAC maintains enzymatic function while bound within the hydrogel network.

Synthetic redox polymer hydrogels benefit from electrostatic forces to immobilize enzymes within the film, limiting the choice of enzymes to those that are both active and polyanionic at operating conditions. SLAC is active at pH 7.0, but with a pI of 8.4 (22) it is polycationic and therefore not matched to synthetic redox hydrogel systems operating at neutral pH (29); however, a PVI-Os redox polymer-SLAC hydrogel can initially achieve 1.5 mA cm^{-2} at pH 7.0 and 40 °C (29). A similar system with a fungal laccase from *Trametes versicolor* reaches 0.2 mA cm^{-2} under similar conditions (29). Milli-ampere per cm^2 current densities have also been demonstrated with bilirubin oxidase from *Trachyderma tsunodae* immobilized in a PVI-Os redox polymer on high surface area carbon cloth (31).

While the protein-based system presented here does not attain the catalytic current densities of the state-of-the-art hybrid systems, it achieves bioelectrocatalysis with much less osmium redox mediator. The upper limit of osmium concentration in a hydrogel can be estimated from the MS MALDI-TOF results of OsHSH-1 presented above. OsHSH-1 is a mixed population of unmodified HSH, HSH with a single osmium complex and HSH to which two complexes have been bound. If we assume that the OsHSH-1 population is comprised entirely of HSH with two complexes, the osmium concentration of a 7.5 wt% hydrogel is 6.5 mM assuming the hydrogels do not swell beyond the volume used to rehydrate the sample (26). A reasonable assumption of the lower limit is 0.5 osmium moieties per HSH or 1.7 mM. The upper limit of osmium concentration is 1.5 to 2 orders of magnitude less than the 100-500 mM typically founded in synthetic redox-polymer systems (14). That a catalytic current is produced with such a low osmium concentration suggests that the electron diffusion (via electron transfer between osmium moieties) is efficient and that the apparent electron diffusion coefficient is large. Additionally, as demonstrated above (Figure 3b) the concentration of osmium moieties is easily controlled.

The major redox couple of OsHSH-2 lies at a potential more oxidizing than that of HS-SLAC, indicating that the osmium moieties of OsHSH-2 would not serve as a reducing substrate in this system. Therefore this configuration was not tested under bioelectrocatalytic conditions. Neat hydrogels of OsHSH-2 do, however, show redox activity in a cyclic voltammogramm (SI).

Conclusions

Here we have shown that an electrode surface modification made from a metallo-polypeptide and an engineered chimeric poly-phenol oxidase can support the bioelectrocatalytic reduction of dioxygen to water. The protein building blocks are bifunctional in that the exhibit physical cross-linking functionality in addition to electron transport or enzymatic functionality. The bioelectrocatalytic system present here is, to the best of our knowledge, the first example of an all protein-based, self-assembling, electrode surface modification. The bifunctionality of the protein building blocks allows for the functional components of the bioelectrocatalytic system (i.e. the redox mediator and

oxidoreductase) to also act as a physical structure of the hydrogel. The modularity of the system allows for the independent tuning of the physical properties and bulk functionality. The system presented here can function as a cathode in a biofuel cell or biosensor as is, but importantly, the design is general and can be applied to other advanced hydrogel applications such as tissue engineering and drug delivery.

Experimental Section

Polypeptide expression and purification. Plasmid encoding the hydrogel forming polypeptide AC10A(trp), here termed HSH, was a kind gift from D. Tirrell (CalTech, USA). Expression was carried out as previously described (26). Briefly, 750 mL of Terrific Broth (TB) media (Sigma) with 200 μg mL^{-1} of ampicillin (Sigma) and 50 μg mL^{-1} of kanamycin (Sigma) was inoculated with 10 mL of mature SG13009 *Escherichia coli* (Qiagen) containing pQE9AC10Atrp. The culture was induced with 1.5 mM β-D-1-thiogalactopyranodise (IPTG; Promega) at an OD of 0.85 – 1.05. Expression continued at 37 °C for 5 hours. HSH was purified by nickel affinity chromatography (His-Trap Crude, GE Healthcare) to greater than 95% purity as judged by SDS-PAGE. The purified polypeptide was lyophilized after dialysis against water for 3 days.

Construction of pQE9HSslac and expression and purification of HS-SLAC. The plasmid pSLAC was a kind gift from G. Canters (Leiden University, The Netherlands). The SLAC gene was extracted from pSLAC by polymerase chain reaction (SI). The upstream primer adds a unique SphI restriction site and the downstream primer a unique HindIII site. The SLAC gene was ligated into pQE9AC10Acys (also a gift from D. Tirrell) at the unique SphI and HindIII sites. The resulting plasmid, pQE9HSslac, was propagated in a 5α *Escherichia coli* cell line (New England Biolabs, NEB), extracted and inserted into the SG13009 expression cell line.

Expression of HS-SLAC was done in 750 mL batches of TB media with 200 μg mL^{-1} and 50 μg mL^{-1} kanamycin inoculated with 10 mL of overnight culture of SG13009 containing pQE9HSslac. Cultures were grown to an OD=1.4–1.5 at 30°C prior to induction with 0.5 mM IPTG. Expression continued for 18–20 hours at 24°C. Cells were harvested by centrifugation and purified by nickel-affinity gel filtration chromatography. Detailed procedures are given in SI.

Synthesis of Os-HSH. The osmium complex [Os(bpy)$_2$Cl$_2$]Cl was synthesized from K$_2$OsCl$_6$ and bipyridine (Fisher Chemical) as previously described (15). Twenty mg mL^{-1} of purified HSH in DI water was mixed with 10 mg mL^{-1} of [Os(bpy)$_2$Cl$_2$]Cl (1:20 molar ratio) and the reaction vessel was immersed in heating oil held to 70 °C. The reaction was allowed to proceed for 6 to 18 hours under argon atmosphere and with constant stirring. The synthesis product was filtered over a 10 kDa cellulose filter with 500 times the reaction volume (1 L) each of DI water and PBS (50 mM Na$_i$P, 500 mM NaCl). The product was lyophilized after 3 days of dialysis against DI water.

Hydrogel preparation. Neat hydrogels of OsHSH-1 and OsHSH-2 were prepared by one of two methods: 1) rehydrating lyophilized protein with aqueous solution buffered to pH 7.0 with 100 mM Na$_i$P; 2) aliquots of OsHSH-1 solution at pH 12 were dried in-place and re-hydrated in aqueous solution buffered to pH 7.0 with 100 mM Na$_i$P. Preparation of mixed

hydrogels on glassy carbon electrode surface was accomplished by drying approximately 3 μL of wet hydrogel on the electrode surface and rehydrating prior to analysis. The rate of hydrogel erosion was measured as previously described (26).

Electrochemical measurements. All electrochemical measurements were done in a 3-electrode experiment with a platinum wire as the counter electrode, a Ag/AgCl reference electrode, and glassy carbon working electrode. Bioelectrocatalytic and dilute solution experiments were done in 100 mM Na$_i$P, pH 7.0. All measurements were taken using a μAutoLab potentiostat. Rotating disk electrode (RDE) experiments were conducted with a polished 3 mm glassy carbon RDE. Potentials are stated versus the standard hydrogel electrode (SHE).

Circular Dichroism. Experiments were conducted with a Jasco J-815 spectrometer. Samples at a concentration of approximately 5 μM were prepared from lyophilized polypeptide with 10 mM Na$_i$P buffer of appropriate pH.

Mass Spectroscopy. MS MALDI-TOF analysis of protease digested and un-digested samples were done at the Protein Core Facility at Columbia University using an Applied Biosystems Voyager DE Pro. Modified and un-modified HSH were digested with Trypsin and AspN (New England Biolabs) for 16 hours at 37 °C.

Rheology. Small amplitude oscillatory shear experiments were preformed with a TA Instruments AR1200 constant stress rheometer equipped with an 8 mm steel parallel plate with the gap of 500 μm, a constant strain of 1% and at 22 °C. A bead of mineral oil around the edge of the sample was used to prevent dehydration of the hydrogels.

Activity assays. Activity of HS-SLAC in dilute solution was measured with dimethoxyphenol (DMP) and air saturated aqueous solution with 10 mM Na$_i$P, pH 7.0. Oxidation of DMP was monitored at 468 nm, $\varepsilon = 14,800$ M^{-1}cm^{-1} as described in ref (22). In-gel activity assays were done in 16% Tricene native gels (Invitrogen). Prior to fixing and Coomassie-blue staining, the gel was washed with 100 mM Na$_i$P, pH 7.0, for 30 minutes and incubated with a fresh buffer solution contain 30 mM DMP for an additional 30 minutes. Activity was confirmed by observing a purple precipitate of oxidized DMP.

References

1. van Hest J C M, Tirrell D A (2001) Protein-based materials, toward a new level of structural control. *Chem Comm* 19: 1897-1904.

2. Kopecek J (2007) Hydrogel biomaterials: A smart future? *Biomaterials* 28(34): 5185-5192.

3. Ulijn R V, Bibi N, Jayawarna V, Thornton P D, Todd S J, Mart R J, Smith A M, Gough, J E (2007) Bioresponsive hydrogels. *Mater Today* 10: 40-48.

4. Chockalingam K, Blenner M, Banta S (2007) Design and application of stimulus-responsive peptide systems. Prot Eng Des Sel 20(4):155-161.

5. Petka W A, Harden J L, McGrath K P, Wirtz D, Tirrell D A (1998) Reversible hydrogels from self-assembling artificial proteins. Science 281(5375): 389-392.

6. Wang C, Stewart R J, Kopecek J (1999) Hybrid hydrogels assembled from synthetic polymers and coiled-coil\ protein domains. Nature 397(6718): 417-420.

7. Haines-Butterick L, Rajagopal K, Branco M, Salick D, Rughani R, Pilarz M, Lamm M S, Pochan D J, Schneider J P (2007) Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells. Pro Nat Acad Sci Unit States USA 104(19): 7791-7796.

8. Ehrick J D, Deo S K, Browning T W, Bachas L G, Madou M J, Daunert S (2005) Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. Nat Mater 4:298-302.

9. Murphy W L, Dillmore W S, Modica J, Mrksich M (2007) Dynamic Hydrogels: Translating a Protein Conformational Change into Macroscopic Motion. Angew Chem Int Ed 46:3066-3069.

10. Chau Y, Luo Y, Cheung A C Y, Nagai Y, Zhang S G, Kobler J B, Zeitels S M, Langer R (2008) Incorporation of a matrix metalloproteinase-sensitive substrate into self-assembling peptides - A model for biofunctional scaffolds. Biomaterials 29:1713-1719.

11. Newman J D, Setford S J (2006) Enzymatic biosensors. Mol Biotechnol 32(3): 249-268.

12. Bullen R A, Arnot T C, Lakeman J B, Walsh F C (2006) Biofuel cells and their development. Biosens Bioelectron 21(11):2015-2045.

13. Ikeda T, Kano K. (2003) Bioelectrocatalysis-based application of quinoproteins and quinoprotein-containing bacterial cells in biosensors and biofuel cells. Biochim Biophy Acta Proteins Proteomics 1647(1-2): 121-126.

14. Heller, A. (2006) Electron-conducting redox hydrogels: design, characteristics and synthesis. Curr Opin Chem Biol 10(6): 664-672.

15. Gallaway J W, Calabrese Barton S (2008) Kinetics of redox polymer-mediated enzyme electrodes. J Am Chem Soc 130(26): 8527-8536

16. Mano N, Soukharev V, Heller A (2006) A laccase-wiring redox hydrogel for efficient catalysis of O2 electroreduction. J Phys Chem B 110(23): 11180-11187.

17. Calabrese Barton S, Gallaway J, Atanassov P (2004) Enzymatic biofuel cells for Implantable and microscale devices. Chem Rev 104(10): 4867-4886.

18. Bartlett P N, Cooper J M (1993) A Review of the Immobilization of Enzymes in Electropolymerized Films. J Electroanal Chem 362(1-2):1-12.

19.   Maciejewska M, Schafer D, Schuhmann W (2006) SECM imaging of spatial variability in biosensor architectures. *Electrochem Comm* 8(7):1119-1124.

20.   Iwasaki Y, Tobita T, Kurihara K, Horiuchi T, Suzuki K, Niwa O (2002) Imaging of electrochemical enzyme sensor on gold electrode using surface plasmon resonance. *Biosens Bioelectron* 17(9):783-788.

21.   Shen W, Lammertink R G H, Sakata J K, Kornfield J A, Tirrell D A (2005) Assembly of an artificial protein hydrogel through leucine zipper aggregation and disulfide bond formation. *Macromolecules* 38(9):3909-3916.

22.   Machczynski M C, Vijgenboom E, Samyn B, Canters G W (2004) Characterization of SLAC: a small laccase from Streptomyces coelicolor with unprecedented activity. *Protein Sci* 13(9): 2388-2397.

23.   Johnson W C (1990) Protein Secondary Structure and Circular-Dichroism- a Practical Guide. *Protein Struct Funct Genet* 7(3):205-214.

24.   Hennessey J P, Johnson W C (1981) Information-Content in the Circular-Dichroism of Proteins. *Biochemistry* 20(5):1085-1094.

25.   Kavanagh G M, Ross-Murphy S B (1998) Rheological characterisation of polymer gels. *Progr Polymer Sci* 23(3):533-562.

26.   Wheeldon I R, Barton S C, Banta S (2007) Bioactive proteinaceous hydrogels from designed bifunctional building blocks. *Biomacromolecules* 8(10):2990-2994.

27.   Lau, S. Y. M., Taneja, A. K. & Hodges, R. S. (1984) Synthesis of a Model Protein of Defined Secondary and Quaternary Structure - Effect of Chain-Length on the Stabilization and Formation of 2-Stranded Alpha-Helical Coiled-Coils. *J Biol Chem* 259(21):3253-3261.

28.   Shen W, Zhang K C, Kornfield J A, Tirrell D A (2006) Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nat Mater* 5(2):153-158.

29.   Gallaway J, Wheeldon I, Rincon R, Atanassov P, Banta S, Barton S C (2008) Oxygen-reducing enzyme cathodes produced from SLAC, a small laccase from Streptomyces coelicolor. *Biosens Bioelectron* 23(8):1229-1235.

30.   Sreerama N, Woody R W (2000) Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287(2):252-260.

31.   Mano N, Kim H H, Heller A (2002) On the relationship between the characteristics of bilirubin oxidases and O-2 cathodes based on their "wiring". *J Phys Chem B* 106(34):8842-8848.
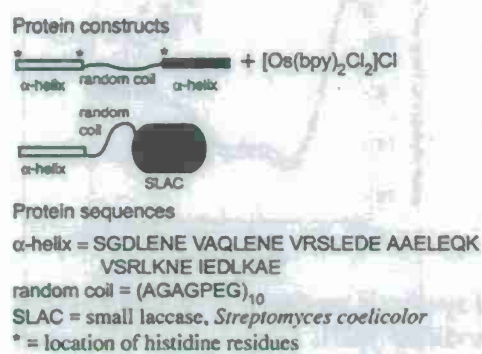
Protein constructs

α-helix   random coil   α-helix   + [Os(bpy)₂Cl₂]Cl

random coil

α-helix

SLAC

Protein sequences

α-helix = SGDLENE VAQLENE VRSLEDE AAELEQK
        VSRLKNE IEDLKAE
random coil = (AGAGPEG)$_{10}$
SLAC = small laccase, *Streptomyces coelicolor*
* = location of histidine residues

Figure 1. Protein constructs with partial sequences, and the osmium redox species used in this study.
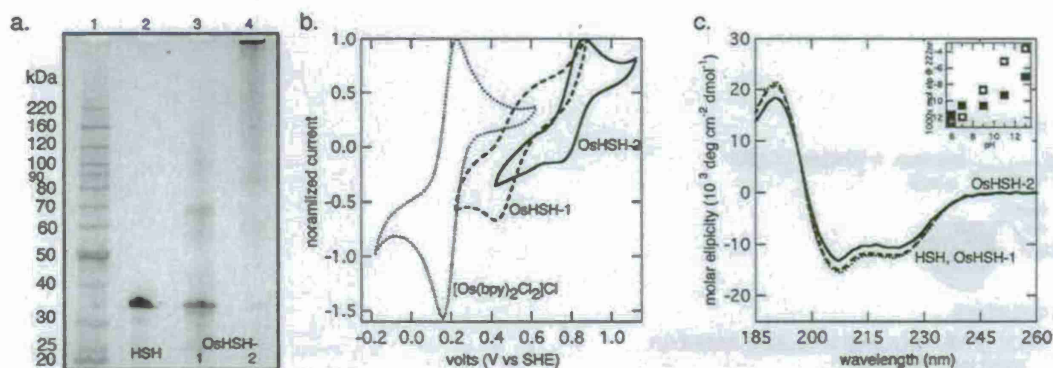
Figure 2. Characterization of $[Os(bpy)_2Cl_2]Cl$ + HSH synthesis products. a. 4-12% Bis-Tris SDS-PAGE; Lane 1, protein standards, Lane 2, unmodified HSH, Lane 3, Os-HSH-1, and Lane 4, Os-HSH-2. b. Cyclic voltammograms of 10 mg mL^{-1} $[Os(bpy)_2Cl_2]Cl$, 35 mg mL^{-1} Os-HSH-1, and 35 mg mL^{-1} Os-HSH-2 in quiescent solution, 100 mVs^{-1}, 100 mM Na$_i$P, pH 7.0. The voltammograms are normalized to 1.0 based on the maximum current density, i_{max}. In µA cm^{-2}, i_{max} values are $[Os(bpy)_2Cl_2]Cl$: 5.3, OsHSH-1: 1.3, OsHSH-2: 62. c. CD spectra of 5 mM samples in 10 mM Na$_i$P, pH 7, at 22 °C. inset Mean residue ellipticity of Os-HSH-1 (open) and OsHSH-2 (closed) at pH 6 – 13. b. and c. Os-HSH-1 (dashed) and Os-HSH-2 (solid). b. $[Os(bpy)_2Cl_2]Cl$ (dots), c. HSH (dots).
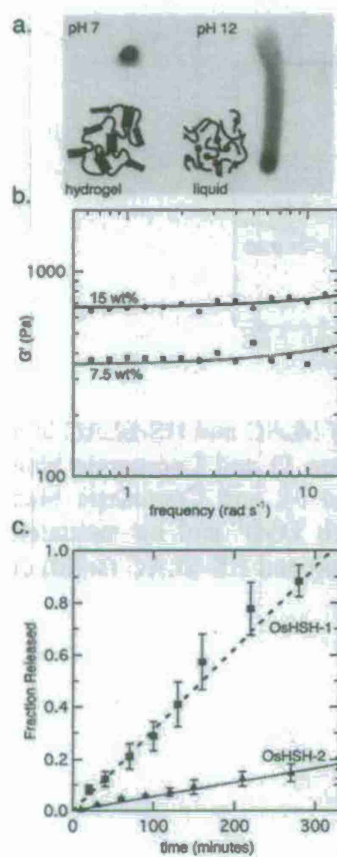
Figure 3. Characterization of OsHSH hydrogels. a. Picture of 7.5 wt% samples of OsHSH-1 on a vertical glass slide at pH 7.0 (left) and pH 12 (right). The picture was taken immediately after standing the glass side on edge. b. Small amplitude oscillatory shear experiments of 7.5 (squares) and 15 wt% (circles) sample of OsHSH-1, 100 mM NaiP, pH 7, 22°C. c. Fractional release of 7.5 wt% samples of OsHSH-1 (squares) and OsHSH-2 (triangles) into quiescent open buffer solution, 100 mM Na$_i$P, pH 7.0 25°C.
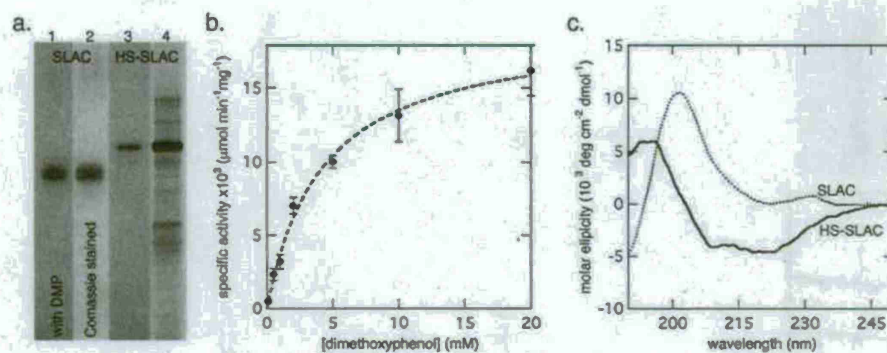
Figure 4. Characterization of HS-SLAC. a. In-gel activity assay of SLAC and HS-SLAC in a 16% tricene native gel. SLAC washed with 30 mM DMP (Lane 1) and Coomassie blue stained (Lane 2). HS-SLAC washed with 30 mM DMP (Lane 3), and Coomassie blue stained (Lane 4). b. Dilute solution activity of HS-SLAC with DMP and air saturated solution, 10 mM Na_iP, pH 7.0, 25°C. c. CD spectra of SLAC (dots) and HS-SLAC (solid) at pH 7.0.
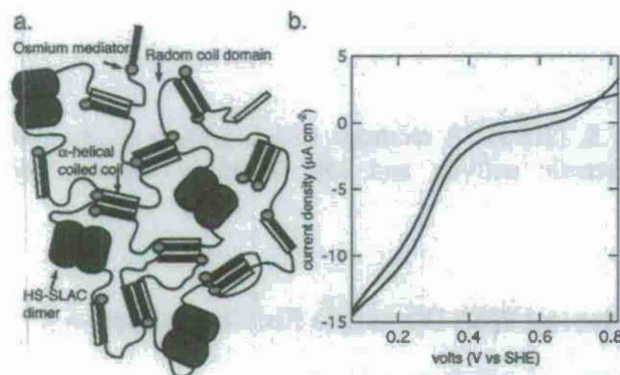
Figure 5. Bioelectrocatalytic hydrogel. a. Schematic diagram of OsHSH-1 and HS-SLAC supramolecular hydrogel. d. Bioelectrocatalysis of a mixed hydrogel of 7.5 wt% OsHSH-1 and 2.5 wt% HS-SLAC on a 3 mm glassy carbon rotating disk electrode. Buffered to pH 7.0 with 100 mM Na$_i$P, 900 rpm, 25 °C, 1 mM O$_2$.

Chapter 4

A Chimeric Fusion Protein Engineered With Disparate Functionalities—

Enzymatic Activity and Self Assembly

Ian R. Wheeldon, Elliot Campbell, and Scott Banta*

Department of Chemical Engineering, Columbia University in the City of New York, New

York, NY 10027, USA.

*Correspondence should be addressed to S. Banta, email: sbanta@cheme.columbia.edu.

Abstract

The fusion of protein domains is an important mechanism in molecular evolution, and a valuable strategy for protein engineering. We are interested in creating fusion proteins containing both globular and structural domains so that the final chimeric protein can be utilized to create novel bioactive biomaterials. Interactions between fused domains can be desirable in some fusion protein applications, but in this case the optimal configuration will enable the bioactivity to be unaffected by the structural cross-linking. To explore this concept, we have created a fusion consisting of a thermostable aldo-keto reductase, two α-helical leucine zipper domains, and a randomly coiled domain. The resulting protein is bifunctional in that: (1) it can self-assemble into a hydrogel material as the terminal leucine zipper domains form inter-protein coiled coil cross-links; and (2) it expresses alcohol dehydrogenase and aldo-keto reductase activity native to AdhD from *Pyrococcus furiosus*. The kinetic parameters of the enzyme are minimally affected by the addition of the helical appendages, and rheological studies demonstrate that a supramolecular assembly of the bifunctional protein building blocks forms a hydrogel. An active hydrogel is produced at temperatures up to 60 °C, and we demonstrate the functionality of the biomaterial by monitoring the oxidation and reduction of the native substrates by the gel. The design of chimeric fusion proteins with both globular and structural domains is an important

advancement for the creation of bioactive biomaterials for biotechnology applications such as tissue engineering, bioelectrocatalysis, and biosensing and for the study of native assembled enzyme structures and clustered enzyme systems such as metabolons.

Introduction

Molecular evolution relies on diversity in protein structures, and one way this is accomplished is through the fusion of native protein domains. There has also been much success in the design and engineering of novel chimeric fusion proteins for new applications. Examples of globular protein fusions include enzymatic switches[1], light activated DNA binders[2], drug and gene delivery systems[3,4,5] single chain antibodies[6], tethered enzymes[7,8], and many others. There has also been interest in creating chimeric structural fusion proteins, such as silk-elastin fusions[9,10], and leucine zipper-elastin fusions[11]. We are interested in creating chimeric fusions with both globular and structural functionalities such that both domains, and the functions of those domains, contribute independently to the final protein construct.

Examples of fusion proteins that combine structural functionality and chemical or bioactive functionality are relatively uncommon. The literature describes fusions of organophosphate hydrolase[12] (OPH) and the protein G Fc binding domain[13] to elastin-like peptides for immobilization to hydrophobic surfaces. Also described are OPH[14] and horseradish peroxide[15] fusions to a cellulose binding domain, for immoblization on the cellulose surfaces and calmodulin fusions to OPH and β-lactamase for reversible immobilization on to appropriately modified surfaces[16]. Perhaps the most common use of protein fusions is in biotechnologies for heterologous expression of recombinant proteins and in the purification of such products.[17,18,19] In expression and purification technologies, the fusion is often temporary as cleavage of the fusion generally occurs en route to the final product.

We have previously created bifunctional fusion proteins that self-assemble into bioactive and enzymatic biomaterials. So far, we have demonstrated the addition of structural self-assembly domains to green fluorescent protein (GFP), the tetrameric *Discosoma* red fluorescent protein (DSRed) and a polyphenol oxidase (SLAC) from *Streptomyces coelicolor*.[20,21] The bifunctional proteins are fusions of the globular mesostable protein-of-interest (GFP, DSRED and SLAC) and the hydrogel forming triblock polypeptide[22], AC10A (here termed HSH). The triblock polypeptide has a helix-random coil-helix structure; the terminal α-helical leucine zipper domains form tetrameric coiled coils leading to the formation of a supramolecular hydrogel structure. Our bifunctional proteins are fusions of the domains of the triblock polypeptide to the termini of the protein-of-interest (or insertions of the protein-of-interest in triblock polypeptide). The addition of self-assembly functionality to globular proteins is highly useful in that with compatible assembly domains one can produce mixed supramolecular structures from more than one type of bifunctional fusion protein.[20,21]

In this study, we describe a new bifunctional enzyme that self-assembles to form a thermostable, 3-dimensional supramolecular hydrogel that has aldo-keto reductase (AKR) activity. This is again accomplished through N- and C-terminal fusions of α-helical leucine zipper cross-linking domains[22] to the α/β barrel structure of an alcohol dehydrogenase with

AKR activity, AdhD from *Pyrococcus furiosus*[23]. The monomers are able to self-assemble into a bioactive enzymatic hydrogel that is stable to temperatures in excess of 60 °C.

AdhD is a member of the AKR superfamily that catalyzes the oxidation of secondary alcohols under basic conditions (optimum pH 8.8) and reduction of ketones under slightly acidic conditions (optimum pH 6.1) with a strong preference towards NAD(H) as a co-factor. Activity increases up to 100 °C and AdhD exhibits latent activity towards primary alcohols, xylose, glucose, arabinose and glyceraldehydes, among others.[23] A thermostable bifunctional AdhD is a platform from which one, with additional protein engineering work to modify substrate specificity, could develop mimics to cellular metabolic pathways that require co-localization or multi-enzyme complexes.[7] The complete oxidation of an alcohol to carbon dioxide for electrical power production[24,25], and biosensing of transient intermediates within a metabolic pathway are applications that would benefit from such a hydrogel system.

The primary concern is that the addition of a second functionality by genetic fusion will eliminate or drastically inhibit the first functionality.[21] We show that the fused α–helical domains maintain helical secondary structure and that the α/β barrel remains highly thermostable. Additionally, we demonstrate that the kinetic parameters measured in dilute solution for diol oxidation and ketone reduction are minimally affected by the fusions to AdhD. With rheological characterization and erosion studies we show supramolecular assemblies of the bifunctional enzyme to be robust and thermostable hydrogels. Finally, we demonstrate the functionality of the system as a hydrogel made from bifunctional AdhD building blocks can catalyze the oxidation of a secondary alcohol and the reduction of a ketone.

Results

Design, expression and purification of a bifunctional AKR

Of primary concern in the design of the bifunctional AdhD protein is that the desired functions, self-assembly and enzymatic activity, are retained in the final construct and that neither is significantly altered. To aid in the design, a homology model of AdhD from *Pyrococcus furiosus* (**accession no. Q8TZM9**) was produced in order to evaluate the potential impact the bifunctional construct may have on catalytic activity. The homology model (Figure 1a.) was generated using ESyPred3D and MODELLER[26] with primary template prostaglandin F synthase from *Trypanosoma brucei* (1VBJ, 31.1% identities). Structures were analyzed using MolProbity[27] and verified against other members of the aldo-keto reductase superfamily. The conserved residues in the catalytic active site of the AKR superfamily (D58, Y63, K89, H121, AdhD numbering) lie at the top of the α/β barrel, at the same end but on the opposite side, as the C-terminal helix (H2, as per AKR nomeclature[28]). The N-terminus lacks a defined structure, is not buried in the α/β barrel structure, and is spatially located at the opposite end of the barrel structure.

Naïvely, inspection of the homology model suggests that fusions to the N-terminus should be innocuous, but that substantial modification to the C-terminus could impair catalytic activity. Concern about the latter modification is supported by evidence from other authors that have shown that single amino acid mutation to the C-terminus of different AKRs can significantly alter catalytic activity.[29] We have previously shown that the fusions of an

α-helical domain (H) and randomly coiled (S) domain in series to the N-terminus and a single H–domain to the C-terminus of GFP results in robust hydrogel with beneficial properties.[20] We hypothesized that the asymmetric order of the domain fusions leads to a reduction in the rate of hydrogel erosion in open buffer solution due to the suppression of closed loop formation. Therefore we employ a similar fusion strategy in this case. A schematic representation of AdhD with both N- and C-terminal modifications is shown in Figure 1B.

AdhD is active to temperatures in excess of 90 °C and has a half-life greater than 2 hours at 100 °C.[23] An enzymatically active, thermostable hydrogel made from self-assembling HS-Adh-H building blocks requires that physical cross-linking through α-helical coiled coil formation also occur at elevated temperatures. The structure of the appended H-domains, if correctly folded, should confer thermostability to the coiled coil bundles to temperatures upwards of 80 °C at pH 6 and to 50 °C at pH 9.[22] Optimization of the coiled coil melting temperature at different pH values is possible[30]; however, the helices used here are amenable to our design goals as they will allow for the investigation of both the reductive and oxidation reaction at high temperatures.

A new bifunctional protein that exhibits AKR activity and self-assembles to form a supramolecular hydrogel was produced. A gene encoding the protein, HS-Adh-H, was constructed by genetically fusing an H- and S-domain to the N-terminus of AdhD and a single H-domain to the C-terminus. The appended H-domains are identical in sequence. A control construct, HS-Adh, with only the N-terminal modification was also constructed. The full amino acid and genetic sequences of HS-Adh-H, HS-Adh, and AdhD are presented in the supporting information (Supporting information, text). The H-and S-domains are blocks of an engineered tri-block polypeptide, HSH.[22,31]

The protein constructs were expressed in *Escherichia coli* and purified by cell lysis at 80 °C followed by strong anion exchange and size exclusion chromatography (SEC). The large amounts of protein required for hydrogel formation prevented extensive use of SEC, therefore hydrogel samples were purified by anion exchange purification only. The two-step purification resulted in samples of approximately 95% purity and three-step purification resulted in a slight increase in purity as judged by SDS-PAGE. Yields, prior to size exclusion chromatography, of 15 to 25 mg per liter of culture were achieved. Protein yields were not substantially reduced after SEC. HS-Adh-H elutes from SEC (20 mM TRIS, 500 mM NaCl, pH 7.8) in three broad peaks between protein standards of 44,000 and 75,000 kD; however, samples of each fraction appear identical with SDS-PAGE analysis.

Structure and stability of a thermostable α/β barrel with cross-linking appendages

Circular dichroism (CD) spectroscopy confirms that the H-domains of the bifunctional fusion protein HS-Adh-H do form α-helices. The CD spectra of HS-Adh-H and AdhD were recorded at intervals of one-half pH units from pH 6 to 9 each at a temperature of 22, 45, 60 and 90 °C. Spectra of HS-Adh-H and AdhD at pH 6 and 9, at 22 and 90 °C, are shown in Figures 2a. and b. No aggregation was observed upon heating and cooling of dilute solutions. The spectra suggest that the wild-type AdhD has no perceptible structural change with an increase in pH from 6 to 9 and only a small change with an increase in temperature

41

from 22 to 90 °C. Conversely, both pH and temperature have an effect on the structure of HS-Adh-H.

Deconvolutions of the spectra support these claims. The α-helical and β-sheet content of HS-Adh-H and AdhD at pH 6 (top) and 9 (bottom) at temperature intervals of 22, 45, 60 and 90 °C are shown in Figure 2c. Wild-type AdhD shows a small exchange of α-helical to β-sheet secondary structure with increasing temperature. The random, or unstructured, content remains constant. In comparison to AdhD, the appended HS- and H-domains add to the α-helical content of HS-Adh-H at low temperatures, and the total β-sheet content is reduced by a similar amount. At 90 °C and pH 6, α-helical content of HS-Adh-H drops below that of AdhD (the β-sheet content increases concomitantly). The decrease in HS-Adh-H α-helical content occurs at a lower temperature when buffered to pH 9, with the %-helices reaching a value slightly lower than that of AdhD at 45 °C. The decrease in α-helical content with increasing temperature at pH 9 is not as large as the decrease observed at pH 6, as at 22 °C and pH 9 the α-helical content of HS-Adh-H is only slightly higher than AdhD.

The deconvolution and spectral data strongly suggests that the HS- and H-domain fusions to the N- and C-terminus of AdhD, respectively, result in the addition of α-helical secondary structure and that the α-helical content of HS-Adh-H decreases with increasing temperature and increasing pH. The effect that the fusions have on the ability of the H-domains to form physical cross-links through coiled coil formation is demonstrated in the rheological characterization of concentrated samples presented below.

A concern is that the appended domains will dramatically limit the stability of the highly thermostable α/β barrel structure of AdhD. The data presented in Figure 3a (and Figure 2c.) shows that HS-Adh-H undergoes some thermal denaturation at temperatures below 90 °C, while wild type AdhD does not. We hypothesize that the initial change of HS-Adh-H from the native state is due to the loss of α-helical structure of the appended domains and not unfolding of the core of the α/β barrel. This partially unfolded structure (denatured helical appendages with an intact, or nearly intact, α/β barrel core) forms a stable intermediate prior to complete denaturation. The folded state of the α/β barrel core in HS-Adh-H at elevated temperature is supported by evidence of enzymatic activity at 90 °C presented below. AdhD in 6 M guanidine hydrochloride (GdHCl) buffered to pH 8 denatures with a single unfolding transition with a T_m of 76 °C (Figure 3b). Without denaturant, HS-Adh-H undergoes a structural change with a T_m of 40 °C at pH 8 (Figure 3a). With denaturant, HS-Adh-H undergoes a second unfolding transition with a T_m of 72 °C (Figure 3b). The observed unfolding transitions in HS-Adh-H are thermodynamically separated to such a degree that we were unable to identify a denaturant concentration that allows for both transitions to occur within a single temperature scan. The stable α/β barrel core of HS-Adh-H in denaturant differ in molar ellipticity per residue in comparison to AdhD (Figure 3b) as the HS-Adh-H construct contains an additional 261 residues in random or un-structured conformation in 6 M GdHCl.

The data presented in Figure 3 supports the existence of the hypothesized intermediate of HS-Adh-H as unfolding proceeds through two thermally induced transitions: the first transition is reversible, and the second is not. Thermodynamic analysis of the

42

folded-unfolded transition is not possible as complete unfolding of both AdhD and HS-Adh-H in denaturant are irreversible upon cooling. Aggregation of both AdhD and HS-Adh-H was observed upon cooling of denaturant solutions but not in solutions without denaturant. Comparison of the melting temperatures of AdhD and HS-Adh-H in 6 M GdHCl does suggest that the α/β barrel structure of HS-Adh-H is slightly destabilized by the appended domains.

The melting temperatures of the first transient of HS-Adh-H decrease with increasing pH ($T_M > 85$ °C at pH 6, 59 °C at pH 7, 40 °C at pH 8, 38 °C at pH 9, see Supplementary material SFigure. 1). The T_M of the triblock polypeptide HSH is similar to the T_M of the first unfolding transition of HS-Adh-H at neutral pH and under slightly acidic conditions (T_M of HSH at pH 6 > 80 °C, and at pH 7 ~55 °C).[22] Under basic conditions, pH 9.5, T_M of HSH is approximately 15 °C greater than the T_M of HS-Adh-H at pH 9.0. Decrease in the stability of the coiled coil domains with increasing pH is expected as the negative charge of deprotonated glutamic acid side chains at positions *e* and *g* of the leucine zipper heptad repeat *abcdefg* destabilize the structure.[32]

The first functionality: alcohol dehydrogenase and aldo-keto reductase activities

The N- and C-terminal fusion of HS- and H-domains to AdhD, respectively, do not eliminate catalytic activity. In dilute solution assays (specifically, 90 nM or 5×10^{-4} wt% enzyme) and under near saturating concentrations of substrate and co-factor there is no significant difference between the oxidative activity of HS-Adh-H and AdhD with 2,3-butanediol and NAD^+ at pH 8.8. Turnover number with 2 mM NAD^+ and 100 mM 2,3-butanediol increases from less than 0.1 s^{-1} at 25 °C to nearly 40 s^{-1} at 90 °C (Figure 4). Under conditions favoring diol oxidation, the CD analysis (presented above) provides evidence of thermal denaturation at temperatures above 38 °C. Catalytic activity at temperature greater than 45 °C confirms the existence of a stable, partially unfolded intermediate.

A similar trend of increasing activity with increasing temperature is observed in dilute solution kinetic assays with ketone reduction. There is no significant difference between HS-Adh-H and wild type AdhD turnover number with 250 μM NADH and 100 mM 3-hydroxy-2-butanone measured at 25, 45 and 60 °C (Figure 4). Activity increases from 0.1 s^{-1} at 25 °C to 1.5 s^{-1} at 60 °C. At 90 °C reductive activity of HS-Adh-H increases to 12 ±2 s^{-1}, and to 7 ±2 s^{-1} for the unmodified AdhD. The significant difference in reductive activities at 90 °C is unexpected as there is no significant difference in activity at all other temperatures for both reactions.

Characteristic to the AKR superfamily is an ordered bi bi reaction mechanism requiring the sequential binding of NAD(P)(H) co-factor followed by substrate binding.[29,33,34] Consistent with this mechanism are the trends observed in double reciprocal plots (1/activity vs 1/[substrate]) of both AdhD and HS-Adh-H.[35] The intersection of linear fits to a set of data series of inverse activity as a function of inverse co-factor concentration occurs below the x-axis (1/[NAD^+]) for the substrate oxidation reaction and above the x-axis (1/[NADH]) for the substrate reduction reaction (see Supplementary material SFigure.2).

The steady state kinetic parameter $k_{i,NAD(H)}$ in the ordered bi bi mechanism (Equation 1, Materials and Methods) is equivalent to the equilibrium dissociation constant for the co-factors, $K_{D,NAD(H)}$.[36] A conveniently located tryptophan residue in the co-factor binding pocket (W92, AdhD numbering) allows for accurate measurement of co-factor binding, as tryptophan fluorescence is quenched upon co-factor binding.[37] Fluorescence titrations reveal a significant difference in dissociation constants, K_{D,NAD^+} and $K_{D,NADH}$, of wild type AdhD and HS-Adh-H. The terminal fusions slightly increase affinity for NAD^+ (HS-Adh-H, $K_{D,NAD^+} = 106$ μM; AdhD, $K_{D,NAD^+} = 110$ μM), while decreasing the affinity towards NADH (HS-Adh-H, $K_{D,NADH} = 47$ μM; AdhD, $K_{D,NADH} = 38$ μM) (Table 1). The Michaelis-Menten constants K_{M,NAD^+} and $K_{M,NADH}$ also reflect the change in co-factor binding, as a statistically significant reduction in K_{M,NAD^+} (HS-Adh-H, 40 μM, AdhD, 57 μM) and a statistically significant increase in $K_{M,NADH}$ (HS-Adh-H, 225 μM, AdhD, 145 μM) are observed. A significant change in Michaelis-Menten constants for the substrate ($K_{M,S}$) is not observed in either the oxidation reaction (HS-Adh-H, 22 mM, AdhD, 21 mM) or the reduction reaction (HS-Adh-H 0.24 mM, AdhD, 0.67 mM).

The change in co-factor binding does not result in a change in maximum rate of kinetic turnover, as there is no statistical difference in k_{cat} for either the oxidation reaction (HS-Adh-H, 3.0 s^{-1}, AdhD, 2.7 s^{-1}) or reduction reaction (HS-Adh-H, 0.9 s^{-1}, AdhD, 0.8 s^{-1}). It is possible that a difference in k_{cat} arises at higher temperatures, as is seen in Figure 4. Additionally, the steady state kinetic parameters of the ordered bi bi mechanism presented in Table 1 capture, within experimental error, the turnover number measured at near saturating conditions of both reactions for AdhD and HS-Adh-H measured at 45 °C (Figure 4) (e.g. the predicted turnover number for HS-Adh-H with 2000 μM NAD^+ and 100 mM 2,3-butandiol at pH 8.8 and 45 °C is 2.4 ±1.0, and the experimental value was measured to be 1.4 ±0.2 [p = 0.11, n = 5,4]). The AdhD results are consistent with previously published reports[23].

The second functionality: supramolecular assembly and hydrogel formation

At 10 wt%, HS-Adh-H forms a supramolecular hydrogel via physical cross-linking of monomers through coiled-coil formation of two or more H-domains (it has previously been shown that the H-domains used in this work tend to form tetrameric coiled-coils[38]). It is also possible that cross-linking between monomers occurs due to dimer formation between the α/β barrel cores as is seen in some AKR family members.[39] Three separate negative controls, 8 wt% HS-Adh-H, 10 wt% HS-Adh and 20 wt% AdhD confirm that two H-domains per monomer are required to form a sufficiently cross-linked structure, and that a minimum of 10 wt% HS-Adh-H is required to from a stable hydrogel structure. Evidence of hydrogel formation, i.e. a shear storage modulus (G′) that is greater than the shear loss modulus (G″) over a range of oscillation frequencies[40], along with the G′ values of 10 and 14 wt% samples of HS-Adh-H and negative controls of 10 wt% HS-Adh and 8 wt% HS-Adh-H are shown in Figure 5a.

A minimum concentration of 10 wt% HS-Adh-H is common to hydrogels at pH 6.3, 7.0, 8.0 and 9.0 (all hydrogel pH values ±0.2). The shear storage modulus of 10 wt% samples of HS-Adh-H at all studied pH values is between 100 and 200 Pa with a loss modulus no greater than 50 Pa (Figure 5b and STable 1). As expected, at low pH (pH < 4)

we observed protein precipitation and at high pH (pH >12) hydrogel structure is lost due to a loss in secondary structure of the appended H-domains.[20,22]

Previous works have demonstrated that G′ increases with wt% protein used to make the hydrogel material.[20,21,41] As applications such as enzymatic surface modifications and bioelectrocatalysis generally require more rigid hydrogels we investigated the temperature dependence of 14 and 18 wt% hydrogels at pH 7 and 9, respectively. At 14 wt%, 22 °C and pH 7, HS-Adh-H forms a hydrogel with a G′ of 960 ±140 Pa, a value similar to a previously reported monomeric fluorescence bifunctional protein HS-ECFP-H, which attains a G′ value of 1000 Pa at 15 wt%. With additional monomers within the structure (18 wt%), G′ increases to 3000 ±540 Pa, 22 °C and pH 9. At An 18 wt% hydrogel sample at pH 9 demonstrates that the increase in hydrogel strength is not limited to neutral pH and that G' can be increased to at least 3000 kPa. In both cases (14 wt%, pH 7 and 18 wt%, pH 9) there is no meaningful change in G′ or G″ up to 60 °C (Figure 5c.). At temperatures above 60 °C, G′ decreases and G″ increases, but a hydrogel persists (as G′ > G″) to the end of the temperature ramp at 75 °C. Ten-wt% samples of HS-Adh-H also persist at high temperatures. At pH 6.3, 7.0, 8.0, and 9.0, 10 wt% samples of HS-Adh-H maintain hydrogel characteristics at temperatures up to 65 °C (Figure 5c and Supplementary material SFigure 3).

The coiled coil tertiary structure is transient in that strand exchange occurs between coiled coils.[31,42] That is to say that an individual H-domain within a coiled coil can exchange places with another H-domain of a different coiled coil. Strand exchange results in a small number of monomers forming a separate unit (closed loop) that is not attached to the larger hydrogel structure. Surface erosion, the loss of protein multimers (or monomers) from the surface of the hydrogel, occurs when closed loops at the surface of the hydrogel diffuse away into open buffer solution. We have previously shown that the asymmetrical structures of HS-GFP-H and HS-DSRED result in suppression of closed loop formation leading to an increase in hydrogel longevity.[20] Erosion rate can also be suppressed by creating a mismatch of aggregation number between N- and C-terminal H-domains.[42] We have also shown that covalent cross-linking after hydrogel formation results in a near complete suppression of erosion.[21]

As expected, HS-Adh-H does erode in quiescent buffer solution (Figure 6). At 25 °C and pH 7, a 12 wt% (2.2 mM) HS-Adh-H hydrogel erodes at a rate of 120 ±10 pmol min^{-1}cm^{-2} (a 10 μL gel is nearly 30% eroded after 2.5 hrs), a value comparable to that of HS-GFP-H (18 wt% or 3.4 mM, 93 pmol min^{-1}cm^{-2})[20]. Under similar conditions a 7.5 wt% (3.4 mM) sample of triblock polypeptide, HSH, readily forms closed loops and completely erodes within 150 minutes[42]. At 45 °C, the erosion rate of HS-Adh-H increases to 390 ±30 pmol min^{-1}cm^{-2}, likely due to an increase in the rate of strand exchange at the elevated temperature.

Hydrogel formation does not prevent enzymatic activity. In Figure 7 we demonstrate the bifunctionality of HS-Adh-H: enzymatic reaction within a self-assembled hydrogel of HS-Adh-H building blocks. In hydrogel samples re-hydrated with buffer containing NAD$^+$ co-factor we monitor the production and consumption of NADH by fluorescence detection upon initiating the oxidative and reductive reactions at 45 °C with 2,3-butanediol and 3-hydroxy-2-butanone, respectively. After the addition of the substrate solutions the final HS-Adh-H concentration was 10 wt%, a concentration sufficient to form a hydrogel structure.

Importantly, we observe that liquid solutions added to HS-Adh-H hydrogel samples are rapidly absorbed into the hydrogel. Reaction and hydrogel conditions were optimized so that we could observe a second change in the redox state of 'in-gel' co-factor while moving towards a new equilibrium point after addition of 3-hydroxy-2-butanone. An optimized set of conditions were found at pH 7 were the reaction rate of the oxidation reaction is sub-optimial.[23]

Upon addition of 3.8 µL of 100 mM 2,3-butanediol to a 14.2 µL sample of HS-Adh-H containing NAD^+, the reaction proceeds towards an equilibrium state favoring NADH due to the limiting concentration of co-factor and high enzyme loading (Figure 7a). Upon addition of 2 µL of 20 mM 3-hydroxy-2-butanone, the reduction reaction proceeds towards an equilibrium that favors NAD^+ and diol. Enzymatic activity is due to HS-Adh-H monomers that are incorporated within the hydrogel as there is no open buffer solution at the surface of the hydrogel for erosion to occur. The reaction rate during the first minute after addition of oxidation of 2,3-butanediol in figure 7a is greater than the reaction rate due to eroded monomers alone, if erosion had been able to occur during that time (Supporting Information, SFigure 4).

Similar reaction profiles with the initiation of the oxidation and reduction reactions are observed at 60 °C (see Supporting Material SFigure 5). Additionally, control of the redox state of 'in-gel' co-factor is possible through changes in hydrogel pH (Figure 7b). A basic shift from slightly acidic to basic conditions induces the concomitant oxidation of 2,3-butanediol and reduction of NAD^+ while the system moves towards a new equilibrium that favors reduced co-factor.

DISCUSSION

Here we demonstrate that a fusion protein of α-helical leucine zipper domains to the termini of the thermostable AdhD from *Pyrococcus furiosus* results in a bifunctional protein building block that self-assembles into a thermostable enzymatic hydrogel. The bifunctional protein building block, HS-Adh-H, expresses the disparate functions of its constituent parts. The α/β core of AdhD is catalytically active and the α-helical leucine zipper domains form coiled coil cross-links in a supramolecular hydrogel structure.

In contrast to some members of the AKR superfamily[29,33], mutation to the C-terminus of AdhD, specifically addition to the C-terminus, does not dramatically affect substrate binding or catalytic functionality. The N- and C-terminal fusions to AdhD do alter co-factor binding, but in such a way as to not inhibit turnover under saturating conditions. AdhD lacks substrate binding loops common among some members of the AKR superfamily, and one or more of these loops is often at the C-terminus[33] (see Supplementary material SFigure 6). The relatively benign nature of the C-terminal fusion to AdhD in HS-Adh-H is evidence for the lack of importance of the C-terminal domain in substrate binding and enzymatic activity of AdhD. The terminal fusions to AdhD from *Pyrococcus furiosus* do not eliminate enzymatic activity. In fact, catalytic turnover at saturating conditions is unaffected. Co-factor binding is affected, but modification to the C-terminus results in a less than one fold difference in $k_{cat}/K_{M,NAD^+}$ for the oxidation of 2,3-butanediol (AdhD, 50 $s^{-1}mM^{-1}$, HS-Adh-H, 75 $s^{-1}mM^{-1}$) and $k_{cat}/K_{M,NADH}$ for the reduction of 3-hydroxy-2-butanone is relatively unchanged (AdhD, 6 $s^{-1}mM^{-1}$, HS-Adh-H, 4 $s^{-1}mM^{-1}$).

46

The minimal disruption of AdhD enzymatic activity in the HS-Adh-H fusion is noteworthy as it is not always the case that fusions are benign. A bifunctional protein of similar design to HS-Adh-H, HS-SLAC (SLAC is a dimeric polyphenol oxidase) results in more than two orders of magnitude decrease in k_{cat}/K_M.[21] Conversely, fusion of an elastin-like peptide domain to organphosphate hydrolase (OPH) results in only a 16% decrease in k_{cat}/K_M.[12] OPH activity is minimally affect in a cellulose binding domain fusion (CBD)[14], but enzymatic activity is decreased by one order of magnitude in a calmodulin-OPH fusion[16]. Additionally, there is no measurable difference in horseradish peroxide (HRP) activity between the wild-type and a CBD-HRP fusion.[15] Cross-comparison of the different fusions does not provide specific insight into the different protein engineering problems, but does highlight the success of the HS-Adh-H fusion and leads to the simple observation that each case is unique.

The demonstration of enzymatic activity within a hydrogel sample of HS-Adh-H presented in Figure 7 does not provide data for evaluation of the specific activity of active sites within the hydrogel construct. It is possible that immobilization AdhD within the supramolecular structure results in a decrease in enzymatic capacity either through structural change to the active site or due to substrate and co-factor diffusion limitations within the hydrogel. The enzymatic activity shown in Figure 7 and SFigure 5 is due to monomers that are crosslinked within the hydrogel as there is no open buffer solution for erosion to occur as added substrate solution is quickly absorbed into the hydrogel and a hydrogel structure is maintain throughout the course of the experiment. It should also be noted that in the experiments where reversibility is demonstrated (Figure 7a, SFigure 5), the reactions were performed at pH 7.0 which is not optimal for either the oxidative or reductive reactions, and therefore the kinetics of the enzymes will be slower than what is reported in Table 1 at the more optimal pH values.

Circular dichroism analysis suggests that HS-Adh-H passes through an enzymatically active but partially unfolded intermediate in which the appended helical domains change conformation while the AdhD α/β barrel core remains intact. The pH dependent T_M of the first unfolding transition is greater than 85 °C at pH 6 and decreases to 38 °C at pH 9. The destabilization of the coiled-coil structure with increasing pH is due to the increase in negative charge with progressive deprotonation of glutamic acid residues at positions e and g.[32] The apparent decrease in stability of the H-domains of HS-Adh-H with respect to HSH could possibly be a result of some interference of helical coil formation by the AdhD protein core. Since there is no linker region on the C-terminal side, it is possible that this H-domain is sterically inhibited from complete formation of coiled coils.

A second thermal unfolding transition in HS-Adh-H is observed with 6 M GdHCl at a T_M 4 °C less than the single thermal unfolding transition of AdhD in 6 M GdHCl (Figure 3b). The two state unfolding of HS-Adh-H is confirmed by the catalytic activity of the HS-Adh-H fusion at pH 6.1 and 8.8 at 90 °C (Figure 4). The minimal difference in kinetic parameters of HS-Adh-H with respect to the wild type enzyme and the measurement of enzymatic activity at temperatures above the first unfolding transition imply that the α/β barrel core of HS-Adh-H remains folded and the observed unfolding transition at low temperature is the loss or change of structure of the appended helical domains. Combined the CD analysis and dilute solution kinetic assays demonstrate two aspects of the structure of HS-Adh-H: 1) the terminal

47

fusions add α-helical structure to the protein independent of the α/β barrel core; and 2) the fusions do not substantially reduce the highly thermostable nature of the AdhD core.

Secondary structure analysis of the CD measurements suggests that at pH > 7 α-helical formation, and consequently hydrogel formation, is limited to temperatures of approximately 40 °C. At pH 8 and 9, 10 wt% samples of HS-Adh-H at pH from 6.3 to 9.0 are stable to temperatures up to 65 °C (Figure 5c and Supporting Material SFigure 5 and STable 1). The T_M of the H-domains at pH 8 and 9, as determined under dilute solution conditions required for the CD analysis, are 40 °C and 38 °C respectively. It is likely that α-helical secondary structure is concentration dependent and is stabilized by the formation of coiled coil bundles; consequently, hydrogel stability (through physical cross-linking between monomers by coiled-coil formation and potentially by protein-protein interactions) is limited not to the temperature dependence observed in the dilute solution experiments. The T_M data does not predict the temperature-dependent behavior of the secondary structure of HS-Adh-H samples at hydrogel forming concentrations (i.e. concentrations that are more than three orders of magnitude higher than the CD analysis). The discrepancy in thermal stability of the α-helical secondary structure in dilute solution and HS-Adh-H hydrogels is also observed at neutral pH.

Five-wt% samples of HSH[22] and 10 wt% samples of a triblock polypeptide with HSH structure but with different H-domain sequences[43] show an increase in liquid-like character as temperature is increased from 23 to 55 °C; however, direct comparison of an HS-Adh-H supramolecular structure to one of HSH is inappropriate as the structures are distinctly different. In the case presented here, the C-terminal cross-links are immobilized by a thermostable α/β barrel potentially adding to the stability of the motif. Additionally, an increase in cross-linking density could results from specific and non-specific protein-protein interactions between α/β cores and suppression of closed loop formation increases connectivity of the network. While the differences in systems are substantial, a comparison of the two systems reveals that the inclusion of the AdhD protein with the supramolecular structure has significant implications on the mechanisms of connectivity within the hydrogel.

The N-terminal fusion of a randomly coiled domain is included in the design as it is highly soluble[22] and it allows for physical separation between protein cores within the hydrogel. It has been shown that it is not essential to hydrogel formation[44] provided that the construct is sufficiently soluble. Also, we envision control over hydrogel porosity by controlling the length and placement of the S-domain.

Conclusions

The chimeric fusion protein, HS-Adh-H, self-assembles to form a thermostable enzymatic hydrogel. The protein is bifunctional in that it forms the physical structure of a hydrogel while retaining the enzymatic activity of the enzyme. The appended α-helical leucine zipper domains are responsible for the formation of a physically cross-linked hydrogel at a minimum concentration of 10 wt% protein. The N- and C-terminal fusions to AdhD minimally affect native enzymatic activity. Enzymatic activity of the bifunctional protein increases with temperature, and hydrogel formation is lost at high temperatures; we produce a rigid hydrogel with enzymatic activity at 60 °C. Our design will have use in a broad range of biotechnology applications such as enzymatic hydrogels for heterogeneous

48

catalysis, electrode modifications for bioelectrocatalysis, enzymatically active surface coatings for biosensors, tissue engineering scaffolds, and the development of artificial metabolons. We also present it as an example of a general design for functional, and multifunctional hydrogels.

Materials and Methods

Chemicals and Reagents

Mono and dibasic sodium phosphate, glycine, sodium chloride, sodium hydroxide, 3-hydroxy-2-butanone, 2,3-butanediol, guanidine HCl, Trizma HCl and Base (TRIS), hydrochloric acid, β-Nicotinamide adenine dinucleotide reduced disodium salt (NADH) and β-Nicotinamide adenine dinucleotide (NAD$^+$) were purchased from Sigma-Aldrich and used without modification. Isopropyl β-D-1-thiogalactopyranoside (IPTG; Promega) was also used without modification. Ampicillin, kanamycin, spectinomycin, and Terrific Broth media were also purchased from Sigma-Aldrich. Premade sodium dodecyl sulfate polyacrylamide gels for electrophoresis (SDS-PAGE) were purchased from Invitrogen. All protein concentrations were determined by Bradford assay with bovine serum albumin standards (Pierce).

Construction of pQE9HSadhH and pQE9HSadh

The plasmid, pWUR85, with adhD from *Pyrococcus furiosus* was a kind gift from John van der Oost (Wageningen University, The Netherlands). The expression plasmid and the tRNA helper plasmid, pSJS1244, are described in ref(23). The adhD gene was extracted from pWUR85 by polymerase chain reaction with forward and reverse primers that include the addition of a SphI site both upstream and downstream of the gene (adhD(SphI)-F ATATAA<u>GCATGC</u>ATGGAATGGCAAAAAGGGTAAATG, the forward primer with unique SphI site (underlined) and adhD(SphI)-R, AATATA<u>GCATGC</u>CCACACACCTCCTTGCCAT, the reverse primer with unique SphI site (underlined)). The resulting fragment was ligated into pQE9AC10Acys[31] (a kind gift from David Tirrell, California Institute of Technology) at the unique SphI site between the C10 and Acys domain encoding regions. Successful transformants were propagated in 5α *Escherichia coli* cell line (NEB). The resulting expression plasmid, pQE9HSadhH, was also transformed into SG13009 *Escherichia coli* (Qiagen) harboring the repressor plasmid pREP4 and pSJS1244. Successful expression of HS-Adh-H from SG13009 cell line requires ampicillin (pQE9HSadhH), kanamycin (pREP4) and spectinomycin (pSJS1244). The plasmid encoding HS-Adh was constructed in an identical manner as described above with one exception, the adhD gene was extracted from pWUR85 with a downstream primer adding a unique SpeI site (adhD(SpeI)-R, CGTATA<u>ACTAGT</u>TCACACACACCTCCT-TGC with unique SpeI site (underlined)).

Expression and purification of AdhD, HS-Adh-H and HS-Adh

Expression of AdhD was done in a similar manner as previously described[23]. Expression of HS-Adh-H and HS-Adh was done in 750 mL batches of Terrific Broth media supplemented with 200 μg mL^{-1} ampicillin, 50 μg mL^{-1} kanamycin and 50 μg mL^{-1} spectinomycin inoculated with 10 mL of mature SG13009 *E. coli* harboring pQE9HSadhH (or pQE9HSadh), pREP4 and pSJS1244. Expression was induced with 0.5 mM of Isopropyl β-D-1-thiogalactopyranoside upon reaching an OD$_{600}$ of 0.8-0.9. Expression was allowed to

continue for 15-16 hours at 27 °C prior to harvesting. Growth prior to induction occurred at 37 °C. Cells were harvested by centrifugation at 10,000g for 10 minutes and resuspended in 100 mL of 20 mM TRIS pH 7.5 per 750 mL culture. Cells were lysed by heating to 80 °C for 1 hour and clarified by centrifugation for 30 minutes at 10,000g. HS-Adh-H (or HS-Adh) was purified from the heat stable lysate by Fast Protein Liquid Chromatography (ÄKTA FPLC, GE HealthCare) using a strong anion exchange column (Q-FF, GE HealthCare). After injection of the lysate, the column was washed with 20 mM TRIS pH 7.8 with 200 mM NaCl. The protein of interest was eluted from the column with a linear gradient of NaCl in 20 mM TRIS pH 7.8 from 200 mM NaCl to 500 mM NaCl. Ninety to ninety-five percent pure HS-Adh-H (as judged by SDS-PAGE) elutes in a broad peak from 300 mM to 450 mM NaCl. Fractions containing HS-Adh-H (or HS-Adh) were pooled and concentrated over a 30 kDa cellulose filter (Millipore) while exchanging the buffer to 10 mM dibasic sodium phosphate. The resulting concentrated samples of HS-Adh-H and HS-Adh were approximately 95% pure. Samples used in kinetic assays and circular dichroism experiments were further purified by size exclusion chromatography (SEC; HiLoad 16/20, Superdex 200, GE HealthCare) with 20 mM TRIS pH 7.8, 500 mM NaCl. Excess salt was removed from the size exclusion eluate by buffer exchange over 30 kDa cellulose filters (Amicon, Millipore). The SEC results were compared to low molecular weight calibrations standards (Gel Filtration LMW Calibration Kit, GE HealthCare).

Hydrogel formation

Hydrogel samples ranging from 10 to 18 wt% (100-180 mg mL^{-1}) were prepared from lyophilized HS-Adh-H (or HS-Adh). Protein was lyophilized from anion exchange purified samples after buffer exchange to 10 mM dibasic sodium phosphate adjusted with 1 M NaOH to pH 9 and after concentration to 15-25 mg mL^{-1} (approximately 1/10 final hydrogel concentration). Aliquots of the samples were frozen to -80 °C and lyophilized to dryness. Hydrogels were formed by re-hydrating the dried samples to the desired weight percent and buffer concentration while accounting for the initial sample buffer. Hydrogel pH was adjusted by adding 1 M NaOH or 1 M HCl in place of equal volumes of water required for re-hydration. Final hydrogel pH was measured by fine range pH paper (Whatman and pHydrion).

Hydrogel Rheology

Small amplitude oscillatory shear experiments were preformed with a TA Instruments AR 1200 constant stress rheometer equipped with an 8 or 20 mm steel parallel plate with the gap of 500 μm, a constant strain of 1% at 22 – 75 °C (Peltier plate temperature control). A bead of mineral oil around the edge of the sample was used to prevent dehydration of the hydrogels during testing.

Hydrogel Erosion

Hydrogels re-hydrated with 10 μl of aqueous solution were prepared in 96 well microtiter plates and covered with 250 μL of 100 mM sodium phosphate, pH 7.5. Percent erosion was determined by monitoring the absorbance at 280 nm of a sample of open buffer solution over time.

Circular Dichroism

Experiments were conducted with a Jasco J-815 CD spectrometer with Peltier junction temperature control. Five-μM samples of HS-Adh-H (purified by SEC) and AdhD in 10 mM sodium phosphate buffer were analyzed in a 1 mm quartz cuvette. Solution pH was adjusted with 1 M NaOH and 1 M HCl as required. Spectral deconvolution was accomplished with CDPro software[45]. Each spectrum was deconvoluted with each of SELCON3, CONTILL, and CDSSTR, in each case with 4 protein reference sets. The secondary structure composition is given as an average of the 12 deconvolutions with the associated standard deviation. Alpha-helical and β-sheet contents are stated as the sum of the ordered and disordered helical and sheet deconvolution results.

Protein denaturation studies

The extent of folding was determined by monitoring the circular dichroic absorbance at 222 nm while increasing temperature at a rate of 1 °C per minute. Samples were prepared as described for all circular dichroism analysis. Six-molar guanidine hydrochloride was used in place of phosphate buffer when required. Melting temperature, T_M, taken as the midpoint parameter of sigmoidal fits to temperature scan data at 222 nm.

Activity assays and determining the steady state kinetic parameters

Oxidative activity of AdhD and HS-Adh-H (purified by SEC) was measured with 2,3-butanediol and NAD^+ co-factor in 50 mM glycine buffer, pH 8.8. Reductive activity was measured with 3-hydroxy-2-butanone and NADH co-factor in 100 mM sodium phosphate buffer, pH 6.1. The steady state kinetic parameters of the ordered bi bi reaction mechanism, k_{cat}, $K_{M,S}$ and $K_{M,NADH(H)}$ were determined by fitting initial rate data to equation 1.

$$Rate = \frac{[E]^{total} k_{cat} [NAD(H)][S]}{k_{iNAD(H)} K_{M,S} + K_{M,S}[NAD(H)] + K_{M,NAD(H)}[S] + [NAD(H)][S]} \quad (1)$$

where S is 2,3-butanediol in the oxidative reaction and 3-hydroxy-2-butanone in the reductive reaction. The constant $k_{iNAD(H)}$ was determined by fluorescence titration as described below. The initial rates were determined by following the absorbance of produced, or consumed, NADH at 340 nm ($\varepsilon = 6.22$ mM^{-1}cm^{-1}) with a SpectraMax M2 microplate reader (Molecular Devices). Assays were performed in 96-well assay plates with an enzyme concentration of $1 - 5$ μg mL^{-1} at 45 °C. Reaction conditions for the oxidative reaction were combinations of 5, 10, 25, 50, 100, 250, 500, and 1000 mM NAD^+ co-factor and 5, 10, 20, 40, 60, 80, and 100 mM 2,3-butanediol. Reaction conditions for the reductive reaction were combinations of 1, 5, 10, 25, 50, 100, 175, and 250 mM NADH and 0.1, 0.5, 1, 2.5, 5, 10, 20, and 40 mM 3-hydroxy-2-butanone. Each data set was repeated in at least quadruplicate.

The steady state kinetic parameter $k_{iNAD(H)}$ is equivalent to the equilibrium binding constant, K_D, of the co-factor in the active site of the enzyme[36]. A conveniently located tryptophan residue at position 92 in the wild type and position 272 in HS-Adh-H allows for accurate determination of the binding constant as tryptophan fluorescence is quenched upon binding (Ex. 280 nm, Em. 330-340 nm)[37]. For both the wild type and the bifunctional enzyme, the NAD^+ and NADH equilibrium binding constants were determined by equation 2.

$$\Phi = \frac{k_{D,NAD(H)}^{-1}[NAD(H)]}{1 + k_{D,NAD(H)}^{-1}[NAD(H)]} \qquad (2)$$

where Φ is the percent of NAD(H) bound as determined by fluorescent titration. Experiments were done in a 1 cm quartz cuvette with 2 mL of 2 μM enzyme, to which 2 μL aliquots of concentrated co-factor was added. At each point the fluorescence emission at 330 nm for NAD^+ titration or at 450 nm for NADH titrations, with excitation at 280 nm, was recorded. No more than 10 aliquots of concentrated NAD(H) (3, 5 or 10 mM) were added during each titration to ensure a negligible change in enzyme concentration. At least three titration were performed to determine each of K_{D,NAD^+} and $K_{D,NADH}$ for both the wild-type and bifunctional enzyme. Titrations were performed at the same conditions as described above for the reduction and oxidation reactions.

The effect of temperature on enzyme activity was determined by initiating a buffered enzyme-substrate solution, equilibrated to the desired temperature, with co-factor. Assays were performed in a 1 cm quartz cuvette with a 1.5 mL reaction volume. The temperature dependent rate of degradation of NADH was determined from control assays without enzyme and subtracted when appropriate.

All data fits were done using IGOR Pro software with a 95 % confidence interval. Statistical significance is reached with $p \leq 0.05$ with Student's t-test.

Hydrogel enzymatic activity

Enzymatic activity of hydrogel samples was determined by monitoring NADH fluorescence (Ex. 340 nm, Em. 450 nm). Hydrogel samples were re-hydrated with buffer containing NAD^+ co-factor and reaction was initiated by substrate addition. Assays were performed in 384 well black assay plates. Conversion of 'in-gel' NAD^+ co-factor to NADH and again to NAD^+ as accomplished in 20 μL hydrogel of 10 wt% HS-Adh-H rehydrated with 2 mM NAD^+ (final volume, wt% and concentration after addition of 2,3-butanediol), heated to 45 or 60 °C, and buffered to pH 7 with 100 mM sodium phosphate. 3.8 μL of 100 mM 2,3 butanediol added at t=0 to initiate the reduction of NAD^+ to NADH. Twenty-mM 3-hydroxy-2-butanone added at t=10 minutes to initiate oxidation of 'in-gel' NADH. Sample pH and concentrations of substrates and co-factor were selected so that the equilibriums would favor near complete conversion of the limiting concentration of NAD^+ to NADH upon addition of diol, and the oxidation of NADH to NAD^+ upon ketone addition.
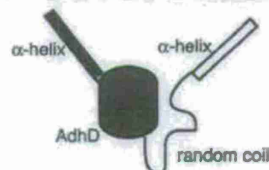
Acknowledgment

52

a. AdhD wt homology model

C-Terminus

N-Terminus
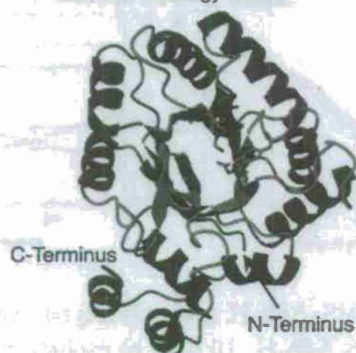
b. HS-Adh-H bifunctional building block

α-helix   α-helix

AdhD   random coil

Protein sequences

α-helix = SGDLENE VAQLENE VRSLEDE
             AAELEQK VSRLKNE IEDLKAE
random coil = $(AGAGAGPEG)_{10}$
AdhD = alcohol dehydrogenase, AKR
                  Pyrococcus furiosus

Figure 1. Structure of AdhD and bifunctional HS-Adh-H. **a.** Top view of the homology model of AdhD from *Pyrococcus furiosus* with N- and C-termini indicated and active site residue side chains shown. C-terminus is at the top of the α/β barrel, out of the page. N-terminus is at the bottom of the α/β barrel, into the page. **b.** Schematic of HS-Adh-H. AdhD with α-helical (H) and randomly coiled (S) domains fused to the N-terminus and an α-helical (H) domain fused to the C-terminus. Partial protein sequences are stated below, full sequence is provided in the Supporting Material.
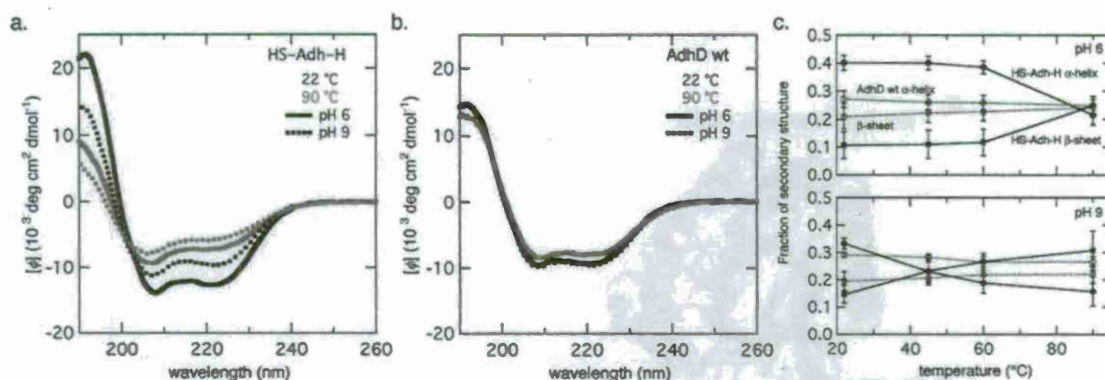
Figure 2. Circular dichroism (CD) analysis of HS-Adh-H and wild type AdhD. **a.** CD spectra in molar ellipticity per residue of 5 μM HS-Adh-H, 10 mM phosphate buffer. **b.** CD spectra of AdhD, conditions same as a. **c.** α-helix and β-sheet content, determined by spectral deconvolution, of spectra of HS-Adh-H and AdhD at pH 6 and 9 at temperatures of 22, 45, 60 and 90 °C.
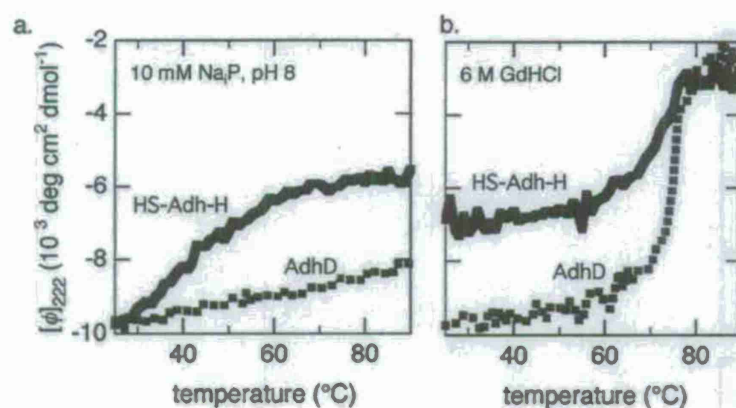
Figure 3. Thermal denaturing of HS-Adh-H and AdhD. Molar ellipticity per residue at 222 nm, $[\phi]_{222}$, of HS-Adh-H (solid lines) and AdhD (dots) from 25 to 90 °C. Samples prepared in (a) 10 mM sodium phosphate buffer, pH 8 and (b) 6 M guanidine hydrochloride.

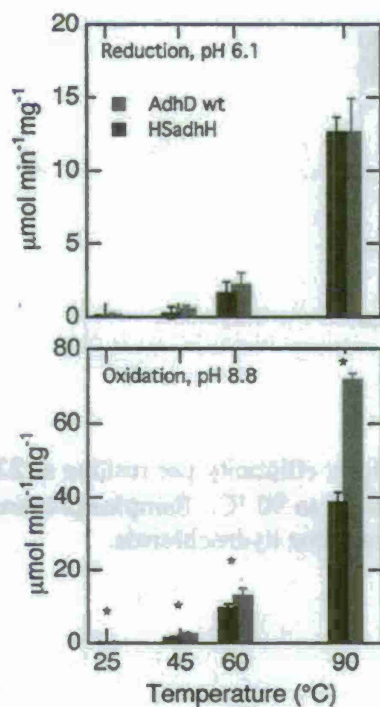**Figure 4.** Turnover number of AdhD (grey) and HS-Adh-H (black) with saturating substrate concentrations at 25, 45, 60 and 90 °C. Oxidation of 100 mM 2,3-butanediol (top) with 2000 μM NAD$^+$, buffered to pH 8.8 with 100 mM sodium phosphate. Reduction of 100 mM 3-hydroxy-2-butanone (bottom) with 250 μM NADH, buffered to pH 6.1 with 100 mM sodium phosphate. Statistically significant difference indicated by * ($p<0.05$). Error bars are standard deviations.
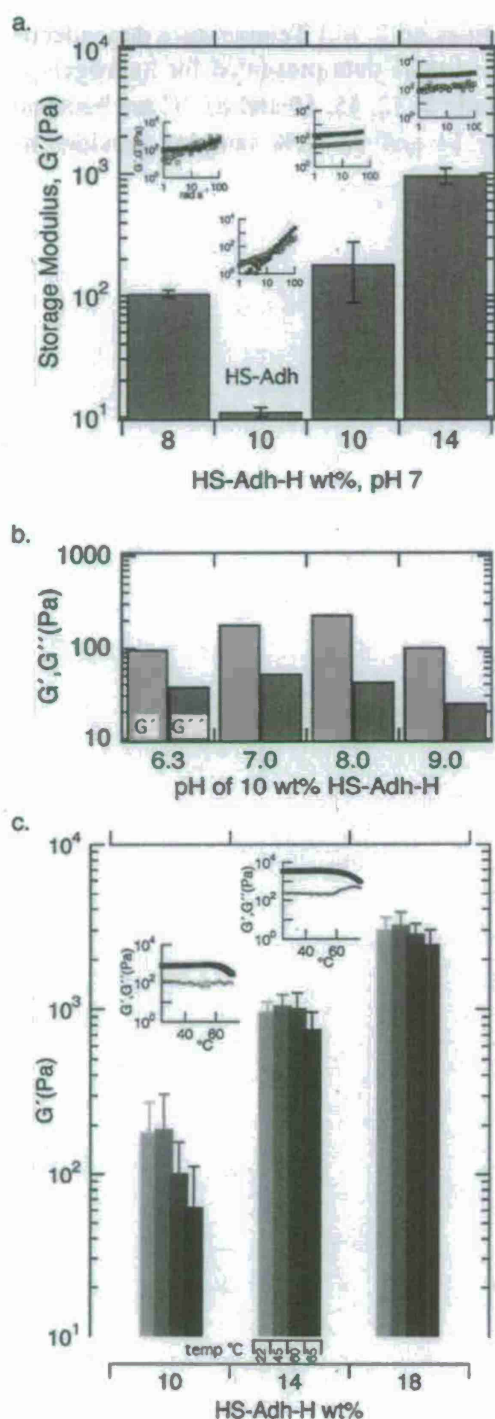
Figure 5. HS-Adh-H hydrogel rheology. (a.) Storage modulus, G′, of 8 wt% HS-Adh-H and 10 wt% HS-Adh negative controls and 10 and 14 wt% HS-Adh-H hydrogels at 22 °C, pH 7. Small amplitude oscillatory shear frequency sweeps depicting G′ (closed) and G″ (open) of each sample is depicted as an inset above each data bar. G′ (b.) G′ and G″ of 10 wt%

57

samples of HS-Adh-H at pH 6, 7, 8 and 9. All pH values ±0.2. (c.) Temperature dependence of 10, 14 and 18 wt% HS-Adh-H hydrogels. 10 and 14 wt% data presented for hydrogels at pH 7, 18 wt% data taken a pH 9. G′ of each wt% sample at 22, 45, 60 and 65 °C are bars and temperature scans of G′ (closed) and G″ (open) for 14 and 18 wt% samples provided as insets.

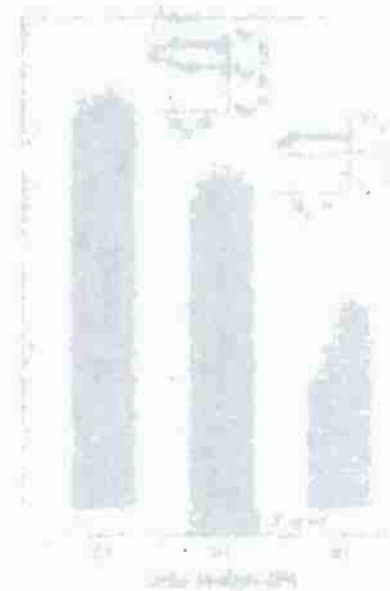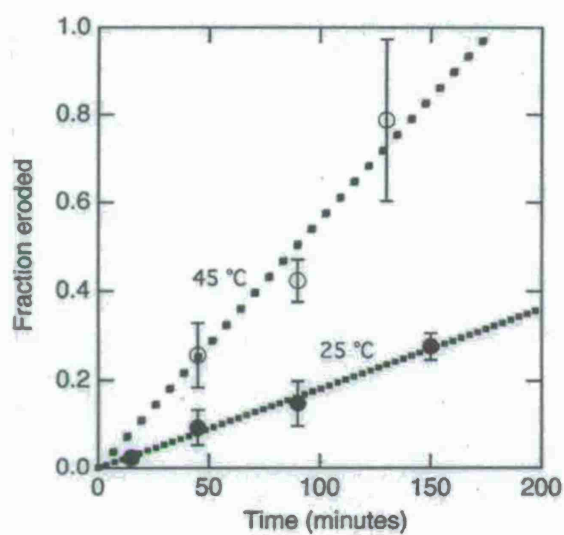**Figure 6**. Fraction eroded of a 12 wt% (2.2 mM) hydrogel of HS-Adh-H, pH 7, 25 °C (closed) and 45 °C (open circles). Ten-µL hydrogel samples in 25X quiescent buffer solution, 100 mM sodium phosphate, pH 7. Error bars are standard deviations, n≥5.
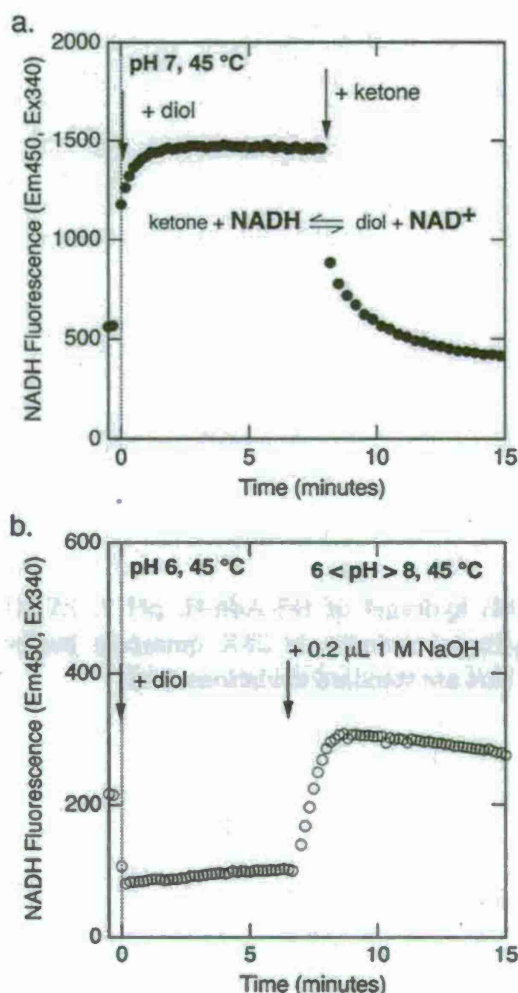
Figure 7. Enzymatic hydrogel activity. (a.) Conversion of in-gel NAD^+ to NADH by initiating the oxidation of 2,3-butanediol and the conversion of produced NADH to NAD^+ by initiating the reduction of 3-hydroxy-2-butanone at pH 7 and 45 °C in a 10 wt% hydrogel of HS-Adh-Ḥ (final wt% after the additions of 2,3-butanediol at t=0 and 3-hydroxy-2-butanone at t=8 minutes). Hydrogel re-hydrated with 2 mM NAD^+ (final concentration). Reduction of NAD^+ to NADH initiated with 21 mM 2,3-butanediol (diol). Perturbation to a new equilibrium initiated with 20 mM 3-hydroxy-2-butanone (ketone). (b.) Initiation of the oxidation of 2,3 butanediol and the conversion of in-gel NAD^+ to NADH by a basic shift in pH. Times of substrate additions indicated by arrows.

Table 1. Kinetic constants for HS-Adh-H and wild type AdhD for the oxidation of 2,3-butanediol at pH 8.8 and 45 °C and the reduction of 3-hydroxy-2-butanone at pH 6.1 and 45 °C. Statistically significant difference indicated by * (p<0.05).

		k_{cat} (s^{-1})	$K_{M,S}$ (mM)	$K_{M,NAD(H)}$ (μM)	$K_{D,NAD(H)}$ (μM)
oxidation reaction, pH 8.8	HS-Adh-H	3.0 ±0.7	22 ±2.4	40 ±6.2*	106 ±1.8*
	AdhD	2.7 ±0.6	21 ±1.7	57 ±5.7*	110 ±1.5*
reduction reaction, pH 6.1	HS-Adh-H	0.9 ±0.2	0.24 ±0.04	225 ±35*	47 ±2*
	AdhD	0.8 ±0.2	0.67 ±0.13	145 ±26*	38 ±1*

61

REFERENCES

1.   Guntas, G., Mansell, T. J., Kim, J. R. & Ostermeier, M. (2005). Directed evolution of protein switches and their application to the creation of ligand-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 11224-11229.

2.   Strickland, D., Moffat, K. & Sosnick, T. R. (2008). Light-activated DNA binding in a designed allosteric protein. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 10709-10714.

3.   Pardridge, W. M. (2008). Re-engineering biopharmaceuticals for delivery to brain with molecular Trojan horses. *Bioconjugate Chemistry* **19**, 1327-1338.

4.   Dietz, G. P. H. & Bahr, M. (2004). Delivery of bioactive molecules into the cell: The Trojan horse approach. *Molecular and Cellular Neuroscience* **27**, 85-131.

5.   Gao, S., Simon, M. J., Morrison Iii, B. & Banta, S. (2009). Bifunctional chimeric fusion proteins engineered for DNA delivery: Optimization of the protein to DNA ratio. *Biochimica et Biophysica Acta (BBA) - General Subjects* **1790**, 198-207.

6.   Blenner, M. A. & Banta, S. (2008). Characterization of the 4D5Flu single-chain antibody with a stimulus-responsive elastin-like peptide linker: A potential reporter of peptide linker conformation. *Protein Science* **17**, 527-536.

7.   Conrado, R. J., Varner, J. D. & DeLisa, M. P. (2008). Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Current Opinion in Biotechnology* **19**, 492-499.

8.   Bulow, L. & Mosbach, K. (1991). Multienzyme Systems Obtained by Gene Fusion. *Trends in Biotechnology* **9**, 226-231.

9.   Kluge, J. A., Rabotyagova, O., Leisk, G. G. & Kaplan, D. L. (2008). Spider silks and their applications. *Trends in Biotechnology* **26**, 244-251.

10.   Haider, M., Megeed, Z. & Ghandehari, H. (2004). Genetically engineered polymers: status and prospects for controlled release. *Journal of Controlled Release* **95**, 1-26.

11.   Zhang, K., Diehl, M. R. & Tirrell, D. A. (2005). Artificial polypeptide scaffold for protein immobilization. *Journal of the American Chemical Society* **127**, 10136-10137.

12.   Shimazu, M., Mulchandani, A. & Chen, W. (2003). Thermally triggered purification and immobilization of elastin-OPH fusions. *Biotechnology and Bioengineering* **81**, 74-79.

13.   Tanaka, G., Funabashi, H., Mie, M. & Kobatake, E. (2006). Fabrication of an antibody microwell array with self-adhering antibody binding protein. *Analytical Biochemistry* **350**, 298-303.

14.   Richins, R. D., Mulchandani, A. & Chen, W. (2000). Expression, immobilization, and enzymatic characterization of cellulose-binding domain-organophosphorus hydrolase fusion enzymes. *Biotechnology and Bioengineering* **69**, 591-596.

15.   Fishman, A., Levy, I., Cogan, U. & Shoseyov, O. (2002). Stabilization of horseradish peroxidase in aqueous-organic media by immobilization onto cellulose using a cellulose-binding-domain. *Journal of Molecular Catalysis B: Enzymatic* **18**, 121-131.

16.   Daunert, S., Bachas, L. G., Schauer-Vukasinovic, V., Gregory, K. J., Schrift, G. & Deo, S. (2007). Calmodulin-mediated reversible immobilization of enzymes. *Colloids and Surfaces B: Biointerfaces* **58**, 20-27.

17. Di Guan, C., Li, P., Riggs, P. D. & Inouye, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. *Gene* **67**, 21-30.

18. Lavallie, E. R., Diblasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F. & McCoy, J. M. (1993). A Thioredoxin Gene Fusion Expression System That Circumvents Inclusion Body Formation in the Escherichia-Coli Cytoplasm. *Bio-Technology* **11**, 187-193.

19. Banki, M. R., Feng, L. A. & Wood, D. W. (2005). Simple bioseparations using self-cleaving elastin-like polypeptide tags. *Nature Methods* **2**, 659-661.

20. Wheeldon, I. R., Barton, S. C. & Banta, S. (2007). Bioactive proteinaceous hydrogels from designed bifunctional building blocks. *Biomacromolecules* **8**, 2990-2994.

21. Wheeldon, I. R., Gallaway, J. W., Barton, S. C. & Banta, S. (2008). Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks. *Proceedings of the National Academy of Sciences* **105**, 15275-15280.

22. Petka, W. A., Harden, J. L., McGrath, K. P., Wirtz, D. & Tirrell, D. A. (1998). Reversible hydrogels from self-assembling artificial proteins. *Science* **281**, 389-392.

23. Machielsen, R., Uria, A. R., Kengen, S. W. M. & van der Oost, J. (2006). Production and characterization of a thermostable alcohol dehydrogenase that belongs to the aldo-keto reductase superfamily. *Applied and Environmental Microbiology* **72**, 233-238.

24. Palmore, G. T. R., Bertschy, H., Bergens, S. H. & Whitesides, G. M. (1998). A methanol/dioxygen biofuel cell that uses NAD(+)-dependent dehydrogenases as catalysts: application of an electro-enzymatic method to regenerate nicotinamide adenine dinucleotide at low overpotentials. *Journal of Electroanalytical Chemistry* **443**, 155-161.

25. Calabrese Barton, S., Gallaway, J. & Atanassov, P. (2004). Enzymatic biofuel cells for Implantable and microscale devices. *Chemical Reviews* **104**, 4867-4886.

26. Lambert, C., Leonard, N., De Bolle, X. & Depiereux, E. (2002). ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics* **18**, 1250-1256.

27. Lovell, S. C., Davis, I. W., Adrendall, W. B., de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). Structure validation by C alpha geometry: phi,psi and C beta deviation. *Proteins-Structure Function and Genetics* **50**, 437-450.

28. Sanli, G., Dudley, J. I. & Blaber, M. (2003). Structural biology of the aldo-keto reductase family of enzymes - Catalysis and cofactor binding. *Cell Biochemistry and Biophysics* **38**, 79-101.

29. Banta, S., Swanson, B. A., Wu, S., Jarnagin, A. & Anderson, S. (2002). Optimizing an artificial metabolic pathway: Engineering the cofactor specificity of Corynebacterium 2,5-diketo-D-gluconic acid reductase for use in vitamin C biosynthesis. *Biochemistry* **41**, 6226-6236.

30. Xu, C. Y., Joss, L., Wang, C., Pechar, M. & Kopecek, J. (2002). The influence of fusion sequences on the thermal stabilities of coiled-coil proteins. *Macromolecular Bioscience* **2**, 395-401.

31. Shen, W., Lammertink, R. G. H., Sakata, J. K., Kornfield, J. A. & Tirrell, D. A. (2005). Assembly of an artificial protein hydrogel through leucine zipper aggregation and disulfide bond formation. *Macromolecules* **38**, 3909-3916.

32. Oshea, E. K., Lumb, K. J. & Kim, P. S. (1993). Peptide Velcro - Design of a Heterodimeric Coiled-Coil. *Current Biology* **3**, 658-667.

33. Sanli, G., Banta, S., Anderson, S. & Blaber, M. (2004). Structural alteration of cofactor specificity in Corynebacterium 2,5-diketo-D-gluconic acid reductase. *Protein Science* **13**, 504-512.

34. Jez, J. M., Bennett, M. J., Schlegel, B. P., Lewis, M. & Penning, T. M. (1997). Comparative anatomy of the aldo-keto reductase superfamily. *Biochemical Journal* **326**, 625-636.

35. Leskovac, V. (2003). *Comprehensive Enzyme Kinetics*. 1 edit, Kluwer Academic/Plenum Publishers, New York.

36. Segel, I. H. (1993). *Enzyme Kinetics*. Wiley Classic Library edit, Wiley-Interscience Publication.

37. Ratnam, K., Ma, H. & Penning, T. M. (1999). The arginine 276 anchor for NADP(H) dictates fluorescence kinetic transients in 3 alpha-hydroxysteroid dehydrogenase, a representative aldo-keto reductase. *Biochemistry* **38**, 7856-7864.

38. Kennedy, S. B., deAzevedo, E. R., Petka, W. A., Russell, T. P., Tirrell, D. A. & Hong, M. (2001). Dynamic structure of a protein hydrogel: A solid-state NMR study. *Macromolecules* **34**, 8675-8685.

39. Kozma, E., Brown, E., Ellis, E. M. & Lapthorn, A. J. (2002). The crystal structure of rat liver AKR7A1 - A dimeric member of the aldo-keto reductase superfamily. *Journal of Biological Chemistry* **277**, 16285-16293.

40. Kavanagh, G. M. & Ross-Murphy, S. B. (1998). Rheological characterisation of polymer gels. *Progress in Polymer Science* **23**, 533-562.

41. Shen, W., Kornfield, J. A. & Tirrell, D. A. (2007). Structure and mechanical properties of artificial protein hydrogels assembled through aggregation of leucine zipper peptide domains. *Soft Matter* **3**, 99-107.

42. Shen, W., Zhang, K. C., Kornfield, J. A. & Tirrell, D. A. (2006). Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nature Materials* **5**, 153-158.

43. Xu, C. Y., Breedveld, V. & Kopecek, J. (2005). Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules* **6**, 1739-1749.

44. Cao, Y. & Li, H. B. (2008). Engineering tandem modular protein based reversible hydrogels. *Chemical Communications*, 4144-4146.

45. Sreerama, N. & Woody, R. W. (2000). Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical Biochemistry* **287**, 252-260.

Chapter 5

Catalytic biomaterials: Engineering organophosphate hydrolase to form self-assembling enzymatic hydrogels

Hoang D. Lu[1], Ian R. Wheeldon[1,2], and Scott Banta[1,3]

[1]Department of Chemical Engineering, Columbia University, New York, NY 10027, USA.

[2]Current Address: Department of Medicine, Center for Biomedical Engineering, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA and Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[3]To whom correspondence should be addressed.
Email: sbanta@columbia.edu

Running Title: Organophosphate Hydrolase Hydrogels

Abstract

Organophosphate (OP) neurotoxins have contaminated the environment, contributed to millions of poisonings annually, and have been used as chemical weapons. Biomaterials incorporating the native activity of the organophosphate hydrolase (OPH) enzyme are of interest for applications including OP sensing, environmental bioremediation, and prophylactic decontamination. We have engineered and characterized four novel hydrogel-forming OPH variants by genetically fusing the OPH enzyme with α-helical leucine zipper domains (H), unstructured soluble linker domains (S), and polyhistidine purification tags. The appended H domains form physical cross links between the enzymes and enables self-assembly of the enzymes into hydrogels. The addition of the H and S fusions significantly increased the expression levels of soluble protein. OPH constructs with biterminal H domains form hydrogels at lower protein weight percents and exhibit higher enzymatic activity than those variants modified with a single H domain fusion. Polyhistidine tags were not useful for purification but they were not benign, as the addition of the 6His tags increased the hydrogel forming abilities of the proteins with a concomitant reduction in both the k_{cat} and K_M values. Active enzymatic hydrogels could be made from concentrated unpurified crude protein lysates, significantly simplifying the processing and utilization of the biomaterials. And, a simple proteinaceous bioactive surface coating exhibiting OPH activity is demonstrated. The hydrogels were stable over long term storage, as activity was retained after cold storage in buffer after 5 months. These new protein constructs further show the use of rational protein design to create novel, bifunctional, self-assembling units for the formation of catalytic biomaterials.

Keywords: Protein hydrogel / Leucine zipper / Organophosphate hydrolase/ Bifunctional protein/ Surface modification

Introduction

Organophosphates (OPs) can deactivate serine proteases and inhibit vital metabolic functions upon exposure, which can lead to subsequent neurological failure and death (Chambers and Oppenheimer, 2004). OP compounds have been developed and exploited for their neurological toxicity. For example, parathion has been widely used a pesticide and sarin has been used as a chemical nerve agent. The widespread use of OP-based pesticides has resulted in extensive environmental contamination, which contributes to millions of cases of OP poisonings annually (Karalliedde and Senanayake, 1989, Mulbry *et al.*, 1996, Singh and Walker, 2006). The use and possession of OP-based weapons has been banned among party members of the Chemical Weapons Convention (Chauhan *et al.*, 2008). For these reasons, there is a broad interest in the development of methods to both detect and to degrade OP compounds.

Organophosphate hydrolase (OPH) from *Flavobacterium sp.* (Mulbry and Karns, 1989) is a dimeric metalloenzyme that catalyzes the hydrolysis of the P–O, P–F, or P–S phosphoric acid ester bonds of OP compounds. OPH consists of two 36 kDa monomers and is active over a broad pH range (pH 6.5 to 12), and it exhibits Michaelis-Menten kinetics (Chen-Goodspeed *et al.*, 2001, Efremenko and Sergeeva, 2001). Many research groups have reported the modification and use of OPH in biosensors and decontamination systems. For example, OPH-polymer complexes and glutaraldehyde cross-linked OPH hydrogels have been applied in electrochemical and spectroscopic OP sensors (Mulchandani *et al.*, 2001, Trojanowicz, 2002). OPH-based sensors have been engineered with OP detection limits as low as parts per trillion (Lei *et al.*, 2007, Luckarift *et al.*, 2007, White and Harmon, 2005). OPH-based materials have also been used in bioreactors for the industrial-scale destruction of OP weapon stockpiles, and OPH-displaying microorganisms have been engineered for the bioremediation of OP contaminated ecosystems (Richins *et al.*, 2000, Singh and Walker, 2006). OPH has also been added to composite materials, including firefighting foams, sponges, paint coatings, cotton, and a variety of polymers for use in personal OP decontamination (Russell *et al.*, 2003).

We have been developing a protein engineering approach to make enzymes bifunctional, such that they retain their native activity while gaining the ability to self-assemble into hydrogel biomaterials. This is accomplished by fusing the enzymes with previously designed α-helical leucine zipper domains (H), which reversibly assemble and form non-covalent cross-links (Shen *et al.*, 2005). We have demonstrated this approach for creating fluorescent protein hydrogels (Wheeldon *et al.*, 2007), bioelectrocatalytic protein hydrogels (Wheeldon *et al.*, 2008), and hydrogels made from a thermostable dehydrogenase enzyme (Wheeldon *et al.*, 2009). In the present work, we have applied this protein engineering strategy to the OPH enzyme to enable it to self-assemble and form a catalytic biomaterial that can degrade OP compounds. Cartoon representations of the newly created bifunctional OPH constructs are shown in Figure 1. This new technology could be used in the development of new decontaminating or bioremediating surfaces, as well as surface modifications for OP biosensors.

Materials and Methods

Materials

β-D-1-thiogalactopyranoside (IPTG; Promega) and complete protease inhibitor cocktail (Roche) were used without modification. Sodium dodecyl sulfate polyacrylamide

67

electrophoresis gels (SDS-PAGE) and running buffers were purchased from Invitrogen. Restriction endonucleases were purchased from New England Biolabs and used as directed. Amicon centrifugal filter devices (Millipore) and Bradford protein concentration assay kit (Pierce) were also used as directed. Terrific Broth powder was purchased from Sigma-Aldrich. All other chemicals were purchased from Sigma-Aldrich and used without modification.

Plasmid Constructs

The plasmid pQE9AC10Acys, expressing the hydrogel forming triblock polypeptide AC10Acys (Shen *et al.*, 2005) here termed HSH, was a kind gift from David Tirrell (California Institute of Technology). The plasmid encoding the OPH gene from *Flavobacterium sp.* ATCC27551, pE1OPD (Barnard *et al.*, 2005) was a kind gift from David Wood (Princeton University).

The OPH gene was isolated from the pE1OPD plasmid using the forward primer 5'-ATATATGGATCCATGTCTATCGGT-3' and the reverse primer 5-ATATATAAGCTTTTATGAGCGCCG-3' by overlap-extension PCR. These primers introduced an upstream *Bam*HI restriction site, and a downstream *Hind*III restriction site. The doubly digested product was ligated into pQE9AC10Acys at the first vector upstream *Bam*HI site and the unique *Hind*III site, resulting in pQE9OPH, coding for the OPH protein with an N-terminal hexa-histidine tag (6His-OPH).

The OPH was also isolated from pE1OPD using the forward primer 5'-ATATAAGCATGCATGGAATGTCTATCGGTAC-3' and the reverse primer 5'-GCTGTAACTAGTTCATGACGCCCGC-3' by overlap extension PCR. These primers introduced an upstream *Sph*I restriction site and a downstream *Spe*I restriction site. The doubly digested fragment was ligated into pQE9AC10Acys at the unique *Sph*I and *Spe*I sites, resulting in pQE9HS-OPH, coding for the protein 6His-HS-OPH. This was repeated with a different reverse primer, 5'-AATATAGCATGCCTGACGCCCGCAAG-3', which incorporated a downstream *Sph*I restriction site. The *Sph*I-digested fragment was ligated into pQE9AC10Acys at the unique *Sph*I site, resulting in pQE9HS-OPH-H, coding for the protein 6His-HS-OPH-H.

Complimentary oligonucleotides, 5-AGATCTGGATCCCATAGTTAATTTCTCCTCTT TAATGAATTCAGTTTC-3' and 5'-GAAACTGAATTCATTAAAGAGGAGAAATTAACTA TGGGATCCAGATCT-3' were hybridized, double digested with *Eco*RI and *Bam*HI and ligated into pQE9AC10Acys following double digestion with the same enzymes. The resultant vector, no longer containing a 6His tag, was named pQE9-del.

The HS-OPH and HS-OPH-H genes were excised from their respective plasmids using unique *Eco*RI and *Bam*HI sites, and these were ligated into the doubly-digested pQE9-del vector. This resulted in the plasmids pQE9-delHS-OPH and pQE9-delHS-OPH-H which code for the proteins HS-OPH and HS-OPH-H respectively.

The five plasmid constructs were propagated into *E. coli* strain SG13009, which contains the pREP4 repressor plasmid. The fidelity of all constructed plasmids was confirmed by DNA sequencing. The DNA and amino acid sequences of the four new bifunctional OPH constructs can be found in the Supplementary Materials.

Protein Expression and purification

All protein constructs were expressed identically in Terrific Broth growth medium with one half the prescribed glycerol content (5 g/L glycerol), as per previous OP expression protocols (Omburo *et al.*, 1992). One liter cultures of growth media supplemented with 200

68

$\Box$g/mL ampicillin and 50 $\Box$g/mL kanamycin were inoculated with 1 mL of saturated overnight culture. Upon reaching $OD_{600} \approx 0.5$, IPTG was added to a final concentration of the 0.5 mM and $CoCl_2$ was also added to a final concentration of 1 mM. Expression was allowed to continue for 15-16 hours at 24 °C, and cells were harvested by centrifugation at 15,000g for 10 minutes. Cell pellets (from 1/3 L of culture) were resuspended in 20 mL of Buffer A (50 mM HEPES, 100 µM $CoCl_2$, pH 8.5) and pelleted by centrifugation. The pellets were again resuspended in 35 mL of Buffer A containing protease inhibitor, and sonicated (Misonix 3000) on ice for six minutes. The lysate was clarified by centrifugation at 15,000g for 30 minutes. The supernatant was subjected to fractionation by ammonium sulfate precipitation over ice by slowly adding concentrated aqueous ammonium sulfate in Buffer A to 40% saturation (28 mg ammonium sulfate/100 mL of Buffer A) while mixing. A precipitated pellet was collected by centrifugation at 10,000g for 10 minutes. The salt precipitated pellet was spun for an additional minute and the remaining supernatant was aspirated. The pellet was resuspended in 10 mL of Buffer A, and thoroughly desalted by diafiltration over a 3 kDa centrifugal filter (Amicon Ultra 15). When required, the desalted solution was subsequently loaded onto a gel filtration column (HiLoad 16/60 Superdex 200pg, GE Healthcare) equilibrated with Buffer A and eluted at 0.5 mL min^{-1} at 4 °C. Fractions containing purest recombinant OPH, as identified by SDS-PAGE, were pooled. Protein was either used directly, or it was further concentrated by ultrafiltration and stored at -80 °C.

Concentrations of pure protein samples were determined by measuring A_{280} values and using theoretical extinction coefficients [ε_{HS-OPH}=.69; $\varepsilon_{T-HS-OPH}$=.70; $\varepsilon_{HS-OP-H}$=.61; $\varepsilon_{T-HS-OPH-H}$=.62 mg^{-1} mL] (Gill and von Hippel, 1989). Protein concentrations were verified using the Bradford method. Expression and purity were monitored throughout the purification process by SDS-PAGE under denaturing conditions using 4-12% Bis-Tris polyacrylamide gels.

Protein Activity Assays

The activity of freshly purified protein constructs in dilute solution was measured by monitoring the accumulation of p-nitrophenolate, a paraoxon, parathion, and methyl parathion hydrolysis product, colorimetrically (λ_{max} = 405 nm) (Votchitseva et al., 2006) at 25 °C using a SpectraMax M2 spectrophotometer. Activity at various pH values was measured using 0.50 mM parathion substrate in 50 mM carbonate or 50 mM Tris buffer with an enzyme concentration of approximately 10^{-8} M. Activity measurements under saturating conditions were also recorded with 1.0 mM paraoxon, 0.50 mM parathion, or 1.0 mM methyl parathion as substrates in 50 mM carbonate buffer (pH 10.5). Other researchers have reported OPH activity on a mass basis of the enzyme. Here we define a unit (U) as moles of substrate hydrolyzed per mol enzyme per second at a specified substrate concentration since our constructs vary significantly in molecular weight.

The kinetic parameters k_{cat} and K_M were determined by non-linear regressions of reaction rates in 50 mM carbonate buffer (pH 10.5) with varying parathion concentrations. Enzymatic activity of the protein constructs at hydrogel-forming concentrations was verified by visually confirming the production yellow p-nitrophenolate after 10uL of 0.4 M parathion in ethanol was added to approximately 100 uL of the protein hydrogels.

Hydrogel Preparation

Hydrogels were formed by either concentrating protein solutions directly or dissolving lyophilized protein in buffer. For hydrogel formation by concentration, protein was concentrated by diafiltration (3 kDa cellulose filter; Amicon Ultra 15). For controlled

hydrogel formation, protein aliquots with known concentrations and volumes were prepared in glass vials and frozen at -80 °C. The samples were lyophilized and rehydrated with 100 µL distilled water. Mechanical mixing with a pipette tip and low speed centrifugation aided in dissolution and homogenization of the rehydrated protein. Hydrogel formation was determined by observing the rehydrated protein's extent of adherence to the top of a glass vial container upon inversion at room temperature.

Hydrogel films were prepared by spreading small amounts of hydrogel onto a surface of a glass microscope slide. Greater than 20 µL of hydrogel was spotted on the surface of one slide and a second slide was placed firmly on top of the first, sandwiching the hydrogel into a thin film. Excess hydrogel was wiped away, and the second slide was removed with a gentle sliding motion.

Statistical analysis. All measurements were performed in at least triplicate, and reported errors are standard deviations. In the cases were two-way ANOVA statistics were performed, statistical significance was achieved for parameters with a p value < 0.05.

Results

Protein Expression and Purification. Four new bifunctional protein constructs based on the organophosphate hydrolase enzyme were designed and produced: HS-OPH and HS-OPH-H with and without N-terminal hexa-histidine tags (Figure 1A-D). Physically cross-linked networks of the bifunctional proteins are schematically shown in Figure 1E and F. 6His-OPH (without helical appendages) was also engineered to serve as a control construct. SDS-PAGE of whole cell lysate of *E. coli* expressing 6His-OPH show a distinct band of the anticipated monomeric molecular weight (~36kDa), but the majority of this protein is in the insoluble fraction (Figure 2A) which is similar to what has been observed by other researchers working with overexpressed OPH modified with a polyhistidine tag (Cha *et al.*, 2000, Wu *et al.*, 2000). No further work was done with 6His-OPH since it was mostly expressed in an insoluble form under these conditions. SDS-PAGE of clarified lysate of *E. coli* expressing the four proteins with the helical appendages shows that the addition of the HS domains to the OPH protein significantly increased the expression of soluble protein. Ammonium sulfate fractionation followed by size exclusion chromatography resulted in samples of greater than >95% purity as judged by SDS-PAGE (Figure 2B).

Purification of the constructs with both the helical appendages and the hexa-histidine tags by Ni-NTA affinity chromatography unexpectedly resulted in low yields (data not shown). Since the hexa-histidine tags appeared to have an impact on the kinetics and the gel forming abilities of the constructs, the 6His containing constructs were characterized in this work, although the 6His tags were not used for purification. All of the new OPH constructs used here were purified in the same way, when necessary, by ammonium salt precipitation and size exclusion chromatography.

Protein Activity. Activity assays with saturating substrate concentrations were performed at varying pH values (Figure 3). The four fusion constructs had similar pH optimas (pH 10.5 - 11.0), and had similar pH-dependent activity profiles between pH 7.5 and 12.0; however, each construct exhibited a different specific activity. The most active construct, HS-OPH-H had a maximum activity of 1.3 ± 0.1 U at pH 11.0, and the least active construct, 6His-HS-OPH, had a maximum activity of 0.42 ± 0.01 U at pH 10.5. For all pH values tested, the

70

specific activities at saturating substrate values were ordered according to: HS-OPH-H > 6His-HS-OPH-H > HS-OPH > 6His-HS-OPH.

Activity assays using different OP substrates reveal the same trend, with the activity of HS-OPH-H > 6His-HS-OPH-H > HS-OPH > 6His-HS-OPH for paraoxon, parathion, and methyl parathion substrates (Table 1). The rate of hydrolysis was highest with the paraoxon substrate and slowest with methyl parathion substrate for all four variants.

Michaelis-Menten kinetic parameters were determined for the four protein constructs with parathion as a substrate (Figure 4, Table2). As expected the k_{cat} values were consistent with the saturation activity measurements, with HS-OPH-H > 6His-HS-OPH-H > HS-OPH > 6His-HS-OPH; the highest k_{cat} was 1.50 ± 0.05 s^{-1} and the lowest k_{cat} was 0.48 ± 0.03 s^{-1}. The addition of a 6His fusion significantly lowered both the k_{cat} and K_M values for the constructs, but when combined, the addition of the 6His tags had no significant impact on the k_{cat}/K_M values. In contrast, the addition of the C-terminal H domain caused a significant increase in the k_{cat} values for the constructs, which also lead to a significant improvement in the k_{cat}/K_M values.

Hydrogel Formation. The constructs ability to self-assemble into hydrogels was determined by rehydrating lyophilized protein with small amounts of distilled water. Mechanical mixing and low-speed centrifugation were required to ensure the hydrogels were homogenous and free of bubbles. Hydrogel formation was judged by inspection as well as by adhesion of the rehydrated protein to the top of a glass vial upon inversion, similar to the methods used by others (Cao and Li, 2008, Das *et al.*, 2006).

All of the protein constructs were capable of forming hydrogels, but the minimum weight percent of protein required to form a hydrogel varied among each construct. Protein samples with varying weight percent values were prepared to identify a minimum weight percent range where constructs would form hydrogels. Images of protein hydrogels above and below the critical weight percent are shown (Figure 5A). HS-OPH formed a hydrogel at 18 wt% (4.3 mM) and not at 14 wt% (3.2 mM). HS-OPH-H formed a hydrogel at 14 wt% (2.8 mM) but not at 11 wt% (2.1 mM), 6His-HS-OPH formed a hydrogel at 11 wt% (2.4 mM) but not 8 wt% (1.7 mM), and 6His-HS-OPH-H formed a hydrogel at 8 wt% (1.5 mM) but not 4 wt% (0.70 mM) (Table 2, Supplementary Table 1).

Hydrogels were also formed from protein collected after different steps of the purification process (Figure 5B). First, to simplify the processing procedure, the samples were concentrated by diafiltration instead of lyophization. Samples were obtained during the purification of the 6His-HS-OPH-H construct, and hydrogels were made from clarified crude cell lysate, protein purified by ammonium sulfate precipitation, and protein purified by size exclusion chromatography (Figure 5B i, ii, and iii, respectively). All of these samples formed hydrogels, and all of the hydrogels retained OPH catalytic activity as they visibly showed the production of p-nitrophenolate from parathion degradation. The kinetic activities of the proteins in the bulk hydrogels were not quantified because the presence of the hydrogel complicates absorbance measurements, and because the OP hydrolysis reaction occurred too rapidly to be accurately measured at the high protein concentrations required for hydrogel formation.

The ability of the 6His-HS-OPH to function as a catalytic film coating was investigated (Figure 5C). Half of a glass slide was coated with a film of the 6His-HS-OPH hydrogel. Drops of parathion substrate placed on the hydrogel covered slide turned from clear

71

to yellow, indicating hydrolysis of the substrate. Drops of buffer alone on hydrogel covered slide remained clear. Moreover, substrate drops on bare glass also remained clear. The results show that the protein film was catalytically active and could readily degrade OP samples.

The long term stability of hydrogels was also partially investigated. A sample of 6His-HS-OPH-H was formed into hydrogel and was subsequently stored for over five months in Buffer A at 4°C. The sample retained its three dimensional shape (Figure 5D) and was still active with paraoxon as a substrate (Supplementary Figure 2). A sample of lyophilized 6His-HS-OPH-H protein was stored for over six months at -20°C and upon rehydration, the sample was capable of hydrogel formation and paraoxon degradation (not shown).

Discussion

Protein engineering provides a powerful tool set to design and produce macromolecules with novel functionalities and new materials with responsive action and catalytic functions (Banta *et al.*, 2010, Chockalingam *et al.*, 2007, van Hest and Tirrell, 2001). We have developed a method for endowing enzymes the ability to self-assemble into enzymatic biomaterials by fusing ☐-helical leucine zipper domains (H domains) to the protein termini (Wheeldon et al., 2007). In the present work, we created four variants of the OP-hydrolyzing enzyme, OPH, and we have demonstrated that hydrogels made from these proteins retain the ability to degrade OP compounds (Figure 5). These new catalytic materials will be useful in the development of new OP detection, bioremediation, and decontamination platforms.

The unmodified OPH protein has previously been shown to be prone to inclusion body formation upon high yield production in *E. coli* (Cha *et al.*, 2000) which has limited its widespread use. Similar to other reports (Cha *et al.*, 2000), in our hands the expression of the OPH gene with a hexa-histidine tag resulted in protein that was predominantly in the insoluble fraction; however, the addition of the H and S domains to OPH significantly increased the soluble overexpression of the proteins (Figure 2). The increase in soluble expression is likely due to the presence of the S domain (polypeptide $(AGAGAGPEG)_{10}$) which has previously been shown to improve protein solubility (Richins *et al.*, 2000).

All of the new hydrogel forming protein constructs retained native OPH activity and were able to hydrolyze various OP compounds including paraoxon, parathion, and methyl parathion (Table 1, Figure 3 and 4). The activity-pH relationship of all hydrogel forming constructs were found to be similar to that of the wild type OPH (Votchitseva *et al.*, 2006). The wild type OPH enzyme has previously been characterized by several research groups, and there has been a range of values reported for the specific activity of the enzyme, which depend on the expression systems and constructs used. Omburo et al. reported one of the highest specific activities (8120 µmoles of paraoxon degraded per minute per mg OPH, or approximately 4,870 U) with paraoxon as a substrate, and this was obtained with OPH produced in a low-level, constitutive expression system (Omburo *et al.*, 1992). Cha et al has reported a much lower specific activity (0.014 µmoles paraoxon degraded per minute per µgram 6His-OPH, or approximately 8.5 U) with paraoxon substrate, but this was obtained using a hexa-histidine modified OPH overexpressed with a strong promoter system (Cha *et al.*, 2000). The constructs made in the present work are more consistent with the latter report, as the activities for the bifunctional constructs with paraoxon were close to 1 U. Leader peptides fused to the N-terminus of OPH have been previously been observed to decrease the activity of OPH (Mulbry and Karns, 1989), and this effect may have occurred in our case as

72

well, as all N-terminally modified constructs exhibit lower enzymatic activity than what has been reported for the wild type. But importantly, the activity is retained by the mutant enzymes, and when self-assembled into high-density hydrogels, the high concentration of enzymes will dominate the overall turnover rate of the biomaterial.

Comparisons of the kinetic properties of the four constructs revealed unanticipated trends. Two-way ANOVA shows that the addition of the C-terminal H domain increased the k_{cat} values for the constructs while having no significant impact on the K_M values. The addition of a 6His tag to the constructs resulted in decreased k_{cat} and K_M values, leading to no effect on the k_{cat}/K_M values (Table 2). Changes in the K_M values have been previously observed when polyhistidine tags have been fused to other metalloenzymes including OPH, and it has been proposed that the histidine residues may interact with the divalent metal ion in the active site (Efremenko et al., 2007, Lai et al., 1994). Votchitseva and Efremenko report a K_M of OPH (60 μM) (Votchitseva et al., 2006) which is similar to the values exhibited by the mutants in this work. Omburo et al. reported the K_M of wild type OPH to be dependent on the divalent metal ions present in the culture media used for protein expression, ranging from 40 μM with high Ni^{2+} content, 130 μM with high Co^{2+} content, and 400 μM with high Cd^{2+} content (Omburo et al., 1992).

All of the protein constructs were able to self-assemble to form hydrogels as evidenced by visual inspection and through evaluation using a vial inversion test (Figure 5A). However, there was a difference in the minimum protein concentration required for hydrogel formation. A single N-terminal H and S fusion to OPH was sufficient for hydrogel assembly, demonstrating that hydrogel formation of the HS-OPH construct is mediated both by cross-linking of the appended H domains as well as protein-protein interactions that occur upon OPH dimerization. This agrees with previous observations that have demonstrated that dimeric or multimeric proteins require only a single H fusion to form hydrogels, while monomeric proteins require two terminal H fusions to form hydrogels (Wheeldon et al., 2007, Wheeldon et al., 2008). Protein with an additional C-terminal H domain forms hydrogels at a lower weight percent, and this reflects the increased cross-linking abilities that occur from the added H-domains.

The addition of an N-terminal 6His tag to either HS-OPH or HS-OPH-H further decreased the protein weight required for hydrogel formation. The construct with the lowest concentration required for hydrogel formation was the 6His-HS-OPH-H mutant. It is possible that the 6His domain interacts with the divalent metal in OPH, adding an additional physical cross-link. This assertion is consistent with the suggestions that polyhistidine tag protein fusions can mediate structural assembly and oligomerization (Efremenko et al., 2007, Salgado et al., 2008). It could also be possible that two histidine tags bind a single metal ion to form a physical cross-link between protein constructs.

Comparison of the protein concentrations used in the inversion tests (Figure 5A) show that the inversion tests are not simply reporting protein concentration effects. The protein with the lowest concentration for hydrogel formation (6His-HS-OPH-H) forms a hydrogel material at 1.5 mM, while the other three constructs at higher protein concentrations (ranging from 1.7 mM to 3.2 mM) do not form hydrogels, and appear instead to be viscous liquids upon inversion.

A sample of the 6His-HS-OPH-H hydrogel was subjected to 5 months of cold storage in a buffered solution, and the sample remarkably retained its three dimensional shape and did not erode into the buffer. This result is in stark contrast to the relatively fast erosion rate of the HSH peptide (Shen et al., 2006) and the slower, but measurable, erosion rate of the

73

hydrogels formed using fluorescent proteins (Wheeldon *et al.*, 2007). The superior stability of the 6His-HS-OPH-H hydrogel may possibly be attributed to the additional cross-linking provided by the addition of 6His tag.

Importantly, all of the OPH protein constructs with terminal H domains (both N-terminal only, and N- and C-terminal) were catalytically active in hydrogel form (Figure 5A and B). A change in color corresponding to the production of p-nitrophenolate, the hydrolysis product of parathion, is visible after reaction in bulk hydrogel samples.

Additionally, we used 6His-HS-OPH to demonstrate the utility of the constructs as a surface modification. A film of hydrogel coated on a glass slide produces a catalytically active surface coating (Figure 5C) that could potentially be used as an OP decontaminating or protective coating.

In order to further explore the potential utility of the new constructs, we investigated alternative methods for hydrogel formation. To create hydrogels for rigorous investigation, multiple purifications steps are used, and hydrogels were formed by lyophization and rehydration, a procedure that is both time consuming and expensive. For larger scale applications, hydrogel formation at various stages of purification was explored. Protein purification was proven to be unnecessary for hydrogel formation, as hydrogels were formed with protein obtained from clarified crude cell lysates, concentrated via simple membrane filtration (Figure 5B). The ability to form hydrogels without the need for purification or modification will significantly lower the cost of utilizing of this technology.

The stability of the engineered proteins will be critical for certain applications. A hydrogel sample that was subjected to cold storage in a buffer solution remained in a hydrogel state and retained its catalytic activity. A lyophoylized protein sample was stored for at least six months at -20°C and upon rehydration it was able form a hydrogel capable catalyzing of OP degradation. The stability of the OP-based hydrogels is certainly partially due to the inherent stability of the OPH enzyme, which has been reported to exhibit a Gibbs free energy change of over 40kcal/mol upon unfolding (Grimsley *et al.*, 1997). The stability of enzyme with the helical appendages is greater than what has been reported for the wild type OPH stored in similar conditions (Yair *et al.*, 2008), but it is consistent with a report that demonstrated that cross-linked OPH retains catalytic activity after three months of storage (Laothanachareon *et al.*, 2008). Remarkably, the hydrogel left in a buffered solution did not erode away, which suggests that the OPH-based hydrogels are more robust than any of the bifunctional proteins that we previously reported (Wheeldon *et al.* 2007, Wheeldon *et al.*, 2008, Wheeldon *et al.*, 2009).

OPH has great potential in many applications and technologies, but the high costs of preparation and processing is a major barrier to its widespread adaptation (Yair *et al.*, 2008). OPH is prone to form inclusion bodies in high-level *E. coli* expression systems (Cha *et al.*, 2000), and efforts to immobilize OPH onto surfaces and into different materials further increase costs. Here we present a simple protein engineering strategy that results in the creation of self-assembling OPH-active hydrogels which are highly expressed in the soluble fraction and are readily purified. Hydrogel formation via self-assembly can be controlled by adjusting protein concentration, and does not require the addition of chemical cross-linking agents. This may be especially advantageous in applications involving gelation on surfaces where contact with chemical cross-linking reagents should be avoided, such as application to the skin. For some large scale applications, the catalytic biomaterials can be created by simple concentration of crude cell extract, without the need for additional purification of the proteins. The protein required for spontaneous hydrogel formation is high (ranging from 8%

74

to 18% gel weight), but this may be reduced by increasing the number of helical appendages to the system, such as by including the HSH peptide to create mixed hydrogels (Wheeldon *et al.*, 2007). The self-assembling protein-based OPH hydrogels have the potential to create new protective gel coatings, act as catalysts in decontamination reactors, and to be used as simple surface modifications for OP biosensors.
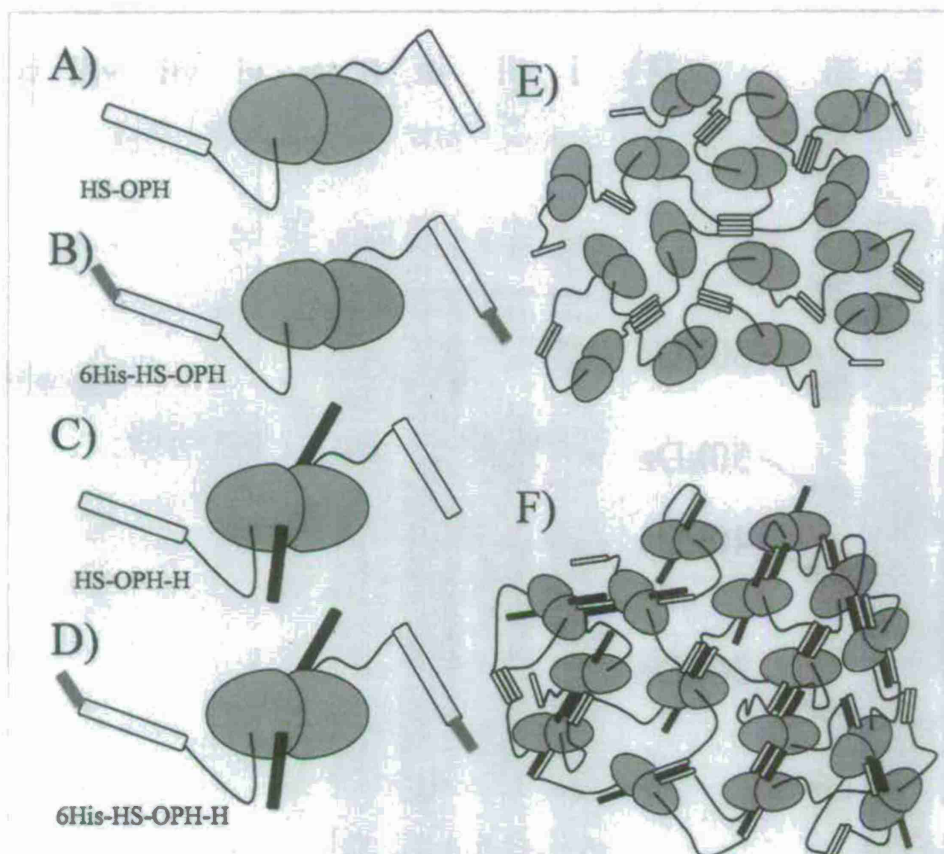
Acknowledgements

References

Banta S., Wheeldon I.R. and Blenner M.A. (2010) *Annu Rev Biomed Eng*, **In Press**.

Barnard G.C., McCool J.D., Wood D.W. and Gerngross T.U. (2005) *Appl Environ Microbiol*, **71**, 5735-5742.

Cao Y. and Li H. (2008) *Chem Commun (Camb)*, 4144-4146.

Cha H.J., Wu C.F., Valdes J.J., Rao G. and Bentley W.E. (2000) *Biotechnol Bioeng*, **67**, 565-574.

Chambers J. and Oppenheimer S.F. (2004) *Toxicol Sci*, **77**, 185-187.

Chauhan S., Chauhan S., D'Cruz R., Faruqi S., Singh K.K., Varma S., Singh M. and Karthik V. (2008) *Environ Toxicol Pharmacol*, **26**, 113-122.

Chen-Goodspeed M., Sogorb M.A., Wu F.Y., Hong S.B. and Raushel F.M. (2001) *Biochemistry*, **40**, 1325-1331.

Chockalingam K., Blenner M. and Banta S. (2007) *Protein Engineering Design & Selection*, **20**, 155-161.

Das D., Dasgupta A., Roy S., Mitra R.N., Debnath S. and Das P.K. (2006) *Chemistry*, **12**, 5068-5074.

Efremenko E., Lyagin I., Votchitseva Y., Sirotkina M. and Varfolomeyev S. (2007) *Biocatalysis and Biotransformation*, **25**, 103-108.

Efremenko E.N. and Sergeeva V.S. (2001) *Russ Chem Bull*, **50**, 1826-1832.

Gill S.C. and von Hippel P.H. (1989) *Anal Biochem*, **182**, 319-326.

Grimsley J.K., Scholtz J.M., Pace C.N. and Wild J.R. (1997) *Biochemistry*, **36**, 14366-14374.

Karalliedde L. and Senanayake N. (1989) *Br J Anaesth*, **63**, 736-750.

Lai K., Dave K.I. and Wild J.R. (1994) *J Biol Chem*, **269**, 16579-16584.

Laothanachareon T., Champreda V., Sritongkham P., Somasundrum M. and Surareungchai W. (2008) *World Journal of Microbiology and Biotechnology*, **24**, 3049-3055.

Lei C., Valenta M.M., Saripalli K.P. and Ackerman E.J. (2007) *J Environ Qual*, **36**, 233-238.

Luckarift H.R., Greenwald R., Bergin M.H., Spain J.C. and Johnson G.R. (2007) *Biosens Bioelectron*, **23**, 400-406.

Mulbry W.W., DelValle P.L. and Karns J.S. (1996) *Pesticide Science*, **48**, 149-155.

Mulbry W.W. and Karns J.S. (1989) *J Bacteriol*, **171**, 6740-6746.

Mulchandani A., Chen W., Mulchandani P., Wang J. and Rogers K.R. (2001) *Biosens Bioelectron,* **16,** 225-230.

Omburo G.A., Kuo J.M., Mullins L.S. and Raushel F.M. (1992) *J Biol Chem,* **267,** 13278-13283.

Richins R.D., Mulchandani A. and Chen W. (2000) *Biotechnol Bioeng,* **69,** 591-596.

Russell A.J., Berberich J.A., Drevon G.F. and Koepsel R.R. (2003) *Annu Rev Biomed Eng,* **5,** 1-27.

Salgado E.N., Lewis R.A., Faraone-Mennella J. and Tezcan F.A. (2008) *J Am Chem Soc,* **130,** 6082-6084.

Shen W., Lammertink R.G.H., Sakata J.K., Kornfield J.A. and Tirrell D.A. (2005) *Macromolecules,* **38,** 3909-3916.

Shen W., Zhang K., Kornfield J.A. and Tirrell D.A. (2006) *Nat Mater,* **5,** 153-158.

Singh B.K. and Walker A. (2006) *FEMS Microbiol Rev,* **30,** 428-471.

Trojanowicz M. (2002) *Electroanalysis,* **14,** 1311-1328.

van Hest J.C.M. and Tirrell D.A. (2001) *Chem Commun,* 1897-1904.

Votchitseva Y.A., Efremenko E.N., Aliev T.K. and Varfolomeyev S.D. (2006) *Biochemistry (Mosc),* **71,** 167-172.

Wheeldon I.R., Barton S.C. and Banta S. (2007) *Biomacromolecules,* **8,** 2990-2994.

Wheeldon I.R., Campbell E. and Banta S. (2009) *J Mol Biol,* **392,** 129-142.

Wheeldon I.R., Gallaway J.W., Barton S.C. and Banta S. (2008) *Proc Natl Acad Sci U S A,* **105,** 15275-15280.

White B.J. and Harmon H.J. (2005) *Biosens Bioelectron,* **20,** 1977-1983.

Wu C.F., Cha H.J., Rao G., Valdes J.J. and Bentley W.E. (2000) *Appl Microbiol Biotechnol,* **54,** 78-83.

Yair S., Ofer B., Arik E., Shai S., Yossi R., Tzvika D. and Amir K. (2008) *Crit Rev Biotechnol,* **28,** 265-275.

Protein Domains and Amino Acid Sequence

/ = H   α-helical leucine zipper domain
SGDLENEVAQLEREVRSLEDEAAELEQKVSRLKNEIEDLKAE

ↄ = S   Random coil, soluble domain
(AGAGAGPEG)$_{10}$

● = OPH   Organophosphate hydrolase monomer
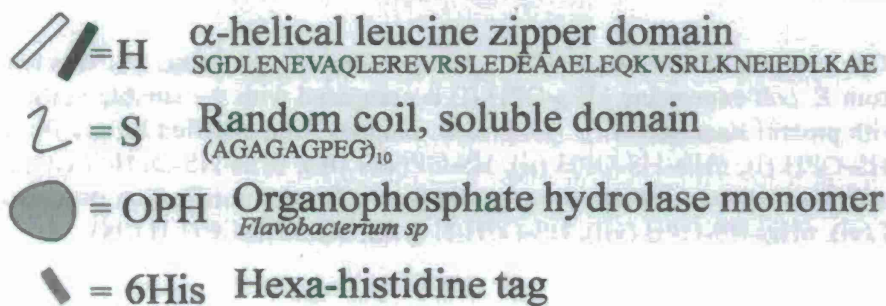*Flavobacterium sp*

▌ = 6His   Hexa-histidine tag

Fig. 1. OPH hydrogel forming bifunctional proteins. Cartoon representations the four new bifunctional OPH constructs (A-D) and the predicted cross linking abilities of the hydrogel forming enzymes HS-OPH and HS-OPH-H (E,F).
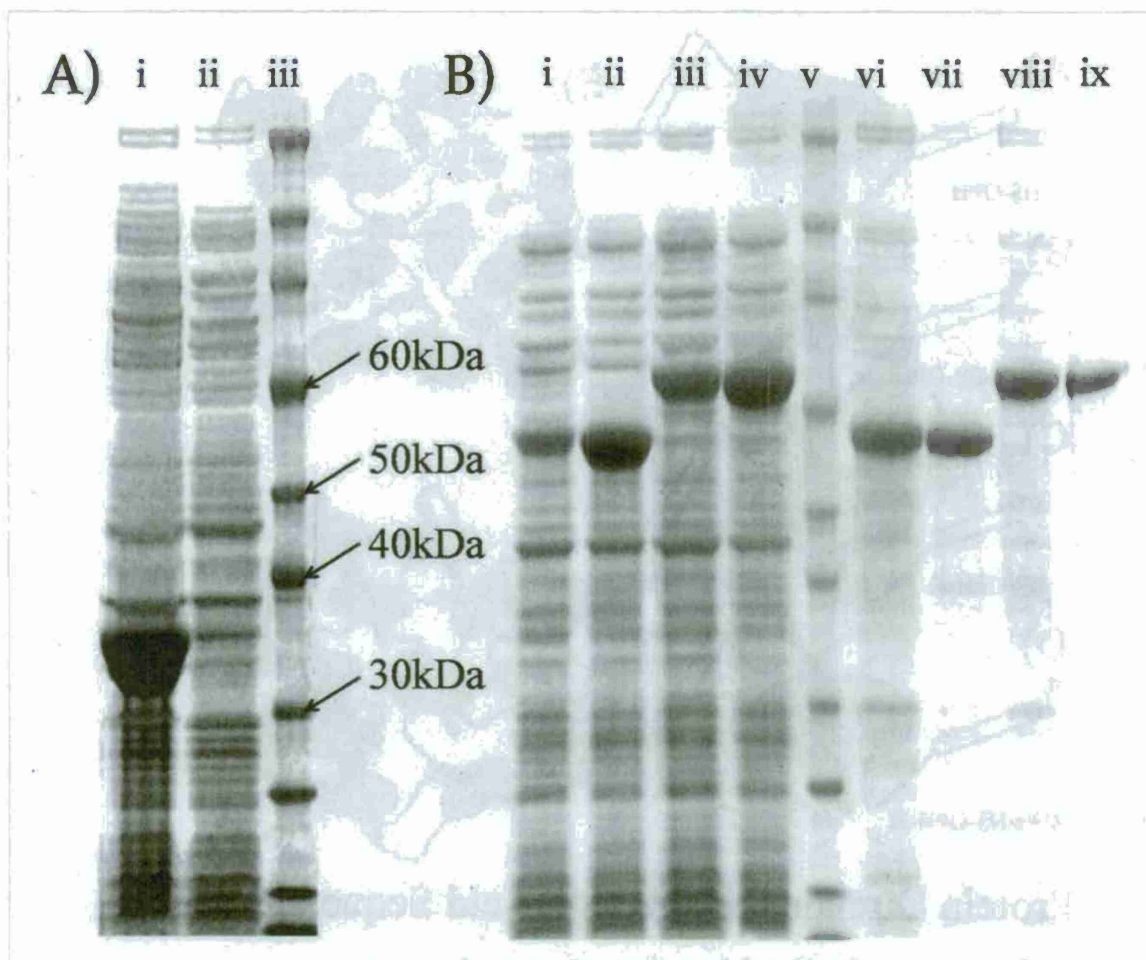
Fig. 2. SDS-PAGE analysis of OPH and OPH hydrogel forming proteins. (A) Whole cell protein extract from *E. coli* expressing 6His-OPH (i) is compared with the soluble portion of the sample (ii) with protein standards (iii). (B) Protein samples from clarified lysates from *E. coli* expressing HS-OPH (i), 6His-HS-OPH (ii), HS-OPH-H (iii), 6His-HS-OPH-H (iv), and protein standards (V) are shown. Protein samples obtained after the purification process are shown, HS-OPH (vi), 6His-HS-OPH (vii), HS-OPH-H (viii), 6His-HS-OPH-H (iix).
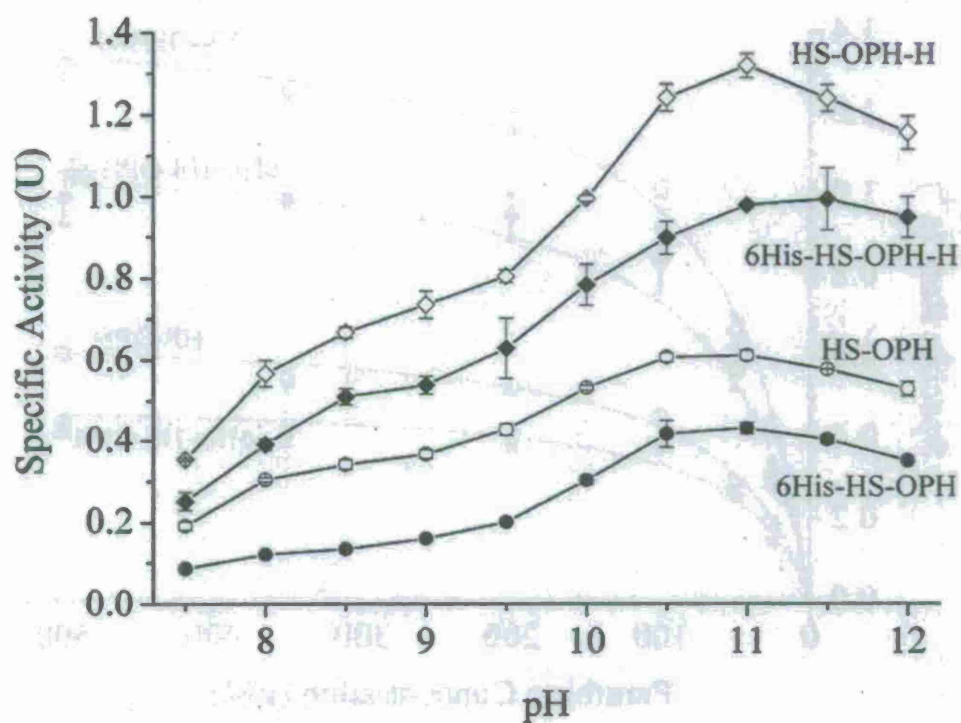
Fig. 3. Effect of pH on the specific activities of the OPH hydrogel forming proteins. The specific activity of each protein construct in dilute solution at 25 °C with 0.5 mM parathion substrate is shown at various pH values. Assays between pH 7.5 and pH 9.0 were conducted in solutions buffered with 50 mM Tris, and assays between pH 9.5 and pH 12.0 were carried out in solutions buffered with 50 mM carbonate. All data was collected in triplicate and error bars represent standard deviations.
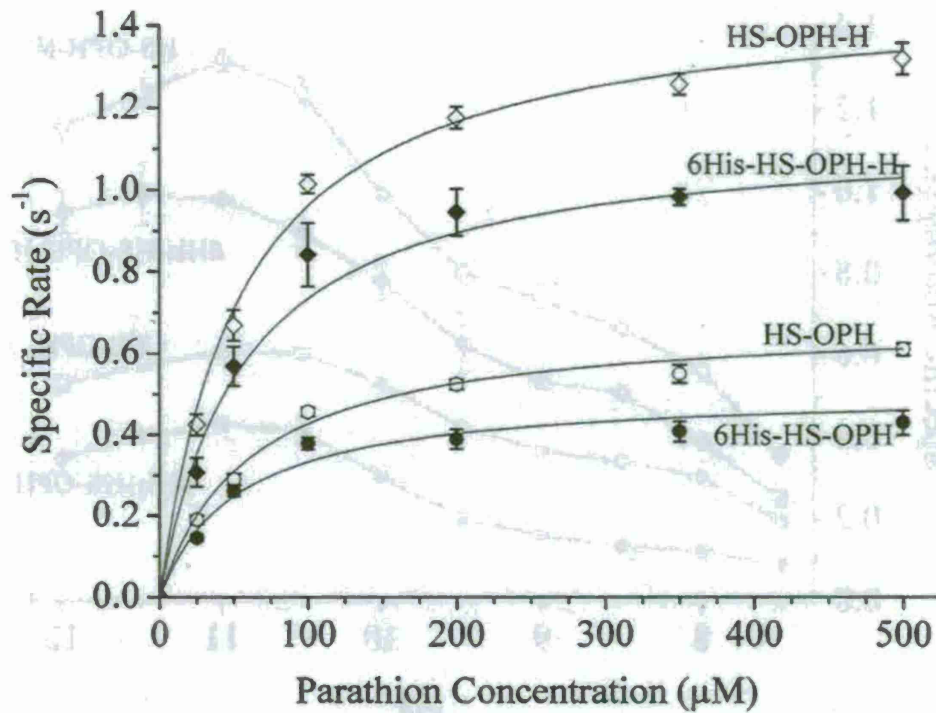
Fig. 4. Kinetic analysis of OPH hydrogel forming proteins. The activity of each protein construct in 50 mM carbonate buffer at pH 10.5 and 25 °C is shown at various concentrations of the parathion substrate. The data were fit to the Michaelis-Menten equation using non-linear regression, and the best fit lines are also shown. The resulting k_{cat} and K_M values can be found in Table 2. All data was collected in triplicate and error bars represent standard deviations.

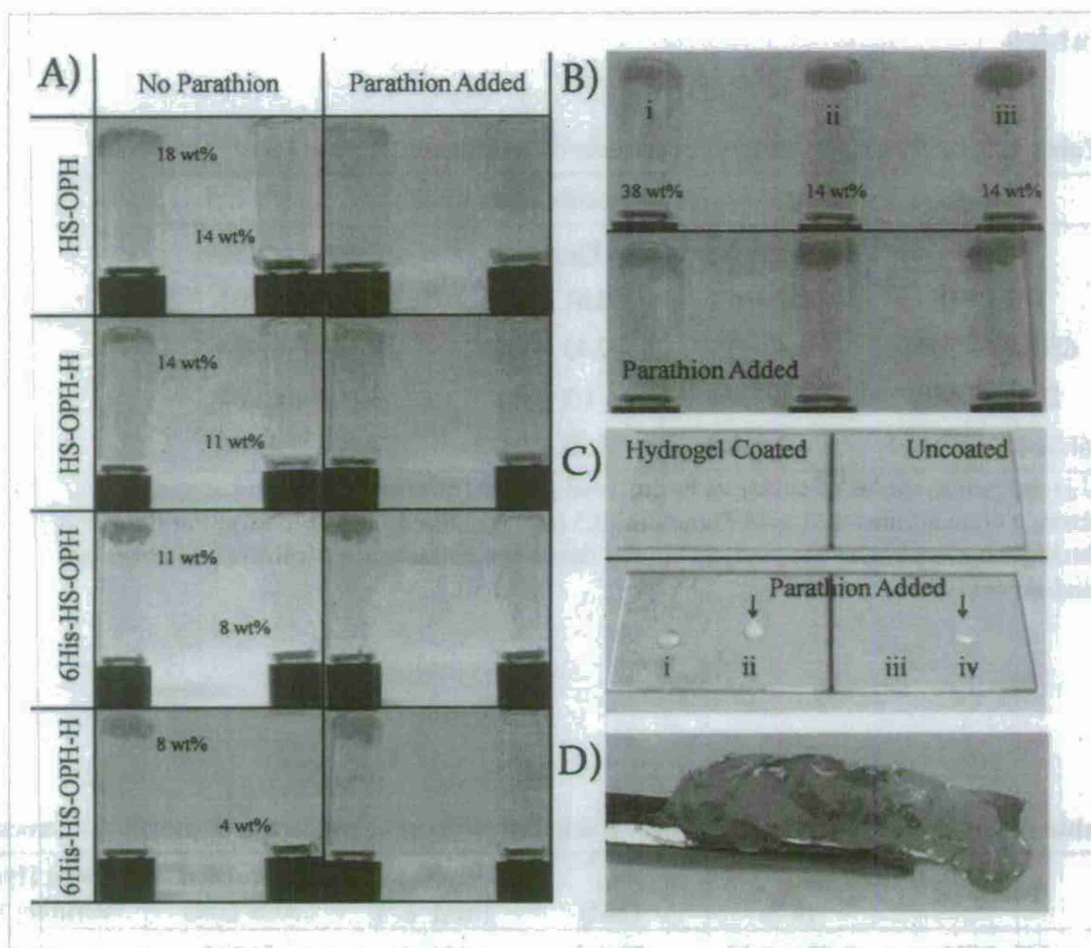**Fig. 5. Hydrogel formation by OPH hydrogel forming proteins.** (A) Concentrated protein samples were lyophilized and rehydrated with 100μL of distilled water to various weight percent protein mixtures in glass vials at room temperature. The protein weight percent of each sample is labeled adjacent to the sample. All experiments were done in duplicate. (Left) Rehydration with water only, no reactant added. (Right) Rehydration with water and 10 μL of 0.4 M parathion. All vials were inverted 10 mins after rehydration. (B) 200 μL of the 6His-HS-OPH-H hydrogels made from protein collected from clarified cell lysate (i), protein pooled from the salt precipitation purification step (ii), and protein collected from the size exclusion chromatography step (iii) (top). The same hydrogel samples after 10 uL of 0.4 M parathion in ethanol was added to each vial before inversion (bottom). (C) A glass slide covered by a thin film of 6His-HS-OPH on the left side of the blue divider is shown. 5 μL of 50 mM carbonate buffer pH 10.5 without parathion (i) and with 0.50 mM parathion (ii) is placed on top of the glass slide covered with protein film, and the same buffer without parathion (iii) and with parathion (iv) is placed on top of the uncoated section of the glass slide. The yellow color indicates the degradation of the parathion substrate to produce p-nitrophenolate. (D) Image of a 6His-HS-OPH-H hydrogel on a spatula that has been stored for over five months in 50 mM HEPES buffer pH 8.5 at 4°C.

81

Tables

Table 1. Specific activity of OPH constructs with different OP substrates

Name	Specific Activity (U) [a]		
	Paraoxon	Parathion	Methyl Parathion
HS-OPH	2.2 ± 0.1	0.61 ± 0.01	0.015 ± 0.001
6His-HS-OPH	1.6 ± 0.1	0.43 ± 0.03	0.011 ± 0.003
HS-OPH-H	5.0 ± 0.1	1.3 ± 0.1	0.036 ± 0.003
6His-HS-OPH-H	4.1 ± 0.2	0.99 ± 0.07	0.025 ± 0.008

[a] U is defined as moles of substrate hydrolyzed per mol enzyme per second at specific substrate concentrations (1 mM Paraoxon, 0.5 mM Parathion, 1 mM Methyl Parathion) in 50mM carbonate buffer pH 10.5, 25°C. All data were collected in triplicate and errors are standard deviations.

Table II: Summary of kinetic parameters and hydrogel forming properties of protein constructs.

Name	k_{cat} [a] (s^{-1})	K_M [b] (μM)	k_{cat}/K_M [c] $(mM^{-1}s^{-1})$	Calculated MW (Da)	Lowest Hydrogel forming wt%
HS-OPH	0.67 ± 0.02	59 ± 3	11 ± 1	51437	18
6His-HS-OPH	0.47 ± 0.03	42 ± 4	11 ± 1	52561	11
HS-OPH-H	1.5 ± 0.1	56 ± 1	26 ± 1	58117	14
6His-HS-OPH-H	1.1 ± 0.1	50 ± 5	23 ± 2	59240	8

Parathion substrate was used for kinetics experiments.

All data were collected in triplicate. Errors are standard deviations.

[a] The effect of the addition of the 6His tag and the C-terminal H domain on the k_{cat} value is statistically significant (2-way ANOVA, $p<0.05$).

[b] The effect of the addition of the 6His tag on the K_M value is statistically significant (2-way ANOVA, $p<0.05$).

[c] The effect of the addition of the C-terminal H domain on the k_{cat}/K_M value is statistically significant (2-way ANOVA, $p<0.05$).

Chapter 6

A self-assembling hydrogel created from three modified dehydrogenases enables the complete oxidation of methanol in an enzymatic biofuel cell

Yang Hee Kim and Scott Banta*

[*]   Dr. Y.H. Kim, Prof. S. Banta
Department of Chemical Engineering
Columbia University
500 West 120th Street, New York, NY 10027 (USA)
Fax: (+1) 212-854-3054
E-mail: sbanta@columbia.edu
Homepage: http://www.columbia.edu/~sb2373

Protein engineering involves the manipulation of amino acids to bring about desired improvements in proteins. Breakthroughs are still being reported in the design and improvement of enzymes while at the same time the field has expanded to include efforts to improve structural proteins that are having an impact in the biomaterials arena. Various protein and peptide domains have been engineered to create new functional materials for a variety of applications.[1] Here we report an advancement of this approach where we create a new catalytic biomaterial via the engineering of three dehydrogenase enzymes for self-assembly. When combined, the resulting new catalytic biomaterial is able to fully oxidize methanol to carbon dioxide using NAD(H) as a redox mediator, and we demonstrate the application of this material as an anode modification in an enzymatic biofuel cell.[2]

Hydrogels can be created from proteins and peptides by outfitting them with cross-linking domains. Pioneering work by Tirrell and colleagues demonstrated that alpha-helical leucine zipper domains could be used to create peptides that self-assemble into hydrogels via coiled-coil interactions[3], and we have expanded on this line of research by demonstrating that these domains can be appended to globular proteins[4]. The hydrogels are cross-linked by coiled-coil motifs formed via the leucine zipper domains and protein/protein interactions can also contribute to the cross-linking of the gels depending on the quaternary structure of the proteins. So far, we have described the addition of helical appendages to a variety of different globular proteins including: fluorescent proteins[4a], a thermostable alcohol dehydrogenase[4b], an organophosphate hydrolase enzyme[4c], and a small laccase enzyme[4d]. When the latter enzyme was combined with osmium-modified peptides, a bioelectrocatalytic hydrogel was formed that could reduce oxygen to water and could function as a cathode modification for an enzymatic biofuel cell.[4d] In almost every case the addition of the helical appendages has had a minimal impact on the catalytic activity of the enzymes, and robust hydrogels have been demonstrated.

Here we extend this approach to create an enzymatic hydrogel that supports a functional synthetic metabolic pathway. Three NAD(H)-dependent dehydrogenase enzymes from different sources were modified for self-assembly. The first enzyme was a tetrameric alcohol dehydrogenase (ADH) from *Bacillus stearothermophilus* which is able to oxidize methanol to formaldehyde.[5] The second enzyme was a tetrameric human aldehyde dehydrogenase (ALDH2) which can oxidize formaldehyde to formate.[6] The final enzyme, a dimeric formate dehydrogenase (FDH1) from *Saccharomyces cerevisiae,* is able to oxidize formate to CO_2.[7] When combined these enzymes produce a synthetic metabolic pathway capable of the complete oxidation of methanol.[8] A schematic diagram of this reaction is as shown in Figure 1(a).

An alpha-helical leucine zipper domain (H) and randomly structured soluble peptide domain (S) were genetically appended to the N-termini of each of the three dehydrogenase genes. The three new bifunctional enzyme constructs (HSADH, HSALDH2, and HSFDH1) were overexpressed in *E. coli* and purified as described in the Supporting Information. HSADH and HSFDH1 were readily expressed and purified, while the HSALDH2 enzyme required the addition of the maltose binding protein (MBP) to enable functional expression. An intein domain was added between the MBP and HSALDH2 such that it spontaneously cleaved after expression within the cells and thus the HSALDH2 protein could be purified as though no fusion protein had been included in the construct.[9]

The kinetics of the purified bifunctionalized enzymes, all of which follow the ordered bi bi kinetic mechanism, were measured in dilute solution to determine the impacts of modifications on the kinetic parameters (Table 1, Figure S3 in the Supporting Information). The kinetic parameters of the HSALDH2 enzyme were largely similar to what has been reported in the literature, while unexpectedly the kinetic parameters of the both the HSADH and HSFDH1 enzymes were both found to be improved by the addition of the helical appendages. Both modified enzymes showed significant increases in catalytic efficiency (k_{cat}/K_m) as compared to the published values for the wild type enzymes. The Michaelis constant (K_m) for the substrate of the HSADH enzyme was three orders of magnitude smaller than what has been reported for the unmodified enzyme while the k_{cat} value was found to be increased by 120-fold. As a result, the catalytic efficiency (k_{cat}/K_m) of HSADH was increased six-orders of magnitude. The catalytic efficiency of HSFDH1 was found to be increased by two orders of magnitude compared to literature values. The change in K_m for the substrate

84

was not significantly different but the k_{cat} value was found to be two orders of magnitude higher than what has been reported for the unmodified enzyme. We have previously observed that the addition of the helical appendages to both ends of a different alcohol dehydrogenase enzyme resulted in almost no impact on the kinetic parameters.[4b] We have also observed that the addition of an N-terminal helical appendage can dramatically improve the functional expression of an organophosphate hydrolase enzyme.[4c] A similar increase in specific activity could partially explain the improvements in the k_{cat} values observed here, as this modification could result in an increase in the amount of active enzymes that are purified however further experiments will be needed to prove this hypothesis.

Each of the newly bifunctionalized enzymes was purified and concentrated (Figure S5 in the Supporting Information) and each was able to form a hydrogel (3.4 mM protein in 100 mM phosphate buffer pH 7.0) as assessed by an inversion test as well as by the ability of the material to retain its shape on the end of a spatula (Figure S5 in the Supporting Information). The hydrogels formed by the three enzymes as well as a hydrogels formed by a mixture of the three enzymes were very similar to the hydrogels formed by all of the other globular proteins we have previously investigated. Since all three enzymes are multimers, the hydrogels are cross-linked via protein/protein interactions in addition to the coiled-coil interactions added by the helical appendages (Figure 1b).

Hydrogels were formed from a mixture of the three enzymes so that activity of each enzyme was normalized to be identical (0.9 U each, where $1 U = 1 \mu mol\ min^{-1}$), and these mixed gels were used to create an anode for an enzymatic biofuel cell. The hydrogel was formed by dissolving the lyophilized enzymes with 100 mM Tris buffer (pH 8.5) along with NAD^+ (3 mM) methylene blue (0.5 mM), and Na_2SO_4 (1 M) and this material was spread on a stainless steel mesh anode. The methylene blue lowers the overpotential required for the electrochemical oxidation of NADH to NAD^+ at the electrode (Figure 1a).[10] To complete the biofuel cell, a commercially prepared air-breathing cathode was used consisting of a gas permeable ELAT electrode with $0.5\ mg\ cm^{-2}$ platinum on carbon which was assembled with a Nafion112 proton exchange membrane (Figure 1c). The biofuel cell was operated by adding drops of the fuels dissolved in buffer onto the anodic hydrogel. Since the hydrogels were operating in air, where they could potentially dry during operation, and the amount of fuel added was small, all measurements were made as soon as a stable potential was obtained.

The theoretical maximum cell voltage of a methanol/oxygen fuel cell is 1.2 V (Figure 2). However, the theoretical maximum cell voltage of this system is 0.81 V as the maximum cell voltage is decreased by the addition the methylene blue redox mediator which has a relatively high standard potential. Figure 2 shows the polarization curves of the anode as well as the whole fuel cell when methanol (100 mM) is applied to the system. The cell voltage of the complete methanol/oxygen enzymatic biofuel cell was found to be 0.69 ± 0.05 V. As would be observed with a general fuel cell, the polarization curves show an activation loss in the low current density area and as the current density increases, the ohmic loss becomes proportionally more important. The activation and ohmic loss together result in a significant voltage drop, while voltage drop due to mass transfer limitations was apparently small in the complete biofuel cell system.

The three enzyme metabolic pathway was designed to fully oxidize methanol to CO_2 via formaldehyde and formate intermediates with a concomitant generation of 6 electrons per molecule of methanol oxidized. By operating the complete biofuel cell with the two intermediates as fuels, the functionality of all three enzymes in the pathway can be investigated. The fact that the power curves increase as the reduction state of the fuel is increased from formate up to methanol indicates that all three enzymes are functioning in the anodic hydrogel (Figure 3, Table S1 in the Supporting Information). The maximum power density (obtained at the apex of the power curves in Figure 3b) was proportional to the number of enzymes involved in the reaction as each enzyme step contributes to the reduction of NAD^+ and reduced NADH can then be electrochemically reoxidized at the electrode to produce current. As expected, the highest power density ($3.52\pm0.16\ mW\ cm^{-2}$) and current density ($26.4\pm1.8\ mA\ cm^{-2}$) were observed when methanol was used in the biofuel cell.

The current efficiency can be evaluated by comparing the theoretical maximum current to the actual current measured in the enzymatic biofuel cell.[11] The theoretical maximum

current density can be calculated using the measured kinetic parameters of the enzymes (Table 1) along with the assumptions that all of the enzymes in the hydrogel are well mixed and active, the NAD^+ regeneration rate is very fast so that the majority of the cofactor is in the oxidized state, and mass transfer effects can be neglected (Supporting Information). The calculated current efficiencies were 40.8 % for methanol, 14.6 % for formaldehyde, and 7.6 % for formate. The variation in the efficiencies may be explained by mass transfer limitations, and this is consistent with the polarization curves obtained using the different fuels (Figure 3a) where mass transfer limitations are most apparent when formate and formaldehyde are fed to the anode. Another factor that may contribute to the variation in the efficiencies is the differences in the kinetic behavior of the three enzymes. Although the gels were created so that the activity of each enzyme was normalized (0.9 U each), the kinetic behavior of the enzymes are very different so their activities will not be the same when operating away from the conditions used to define a Unit of activity (1 mM NAD^+ and 10 mM substrate) (Figure S4 in the Supporting Information). Also, the difference of temperatures between the operation of the enzymatic biofuel cell (ambient temperature) and temperature of the enzyme kinetics study (37 °C) also likely lowered the efficiency. Under the operating conditions with methanol as a fuel, the HSADH enzyme was the rate limiting step, and this resulted in a lower reaction rate compared to when the intermediate fuels were used. Thus, when methanol is used, the system is kinetically limited, not mass transfer limited, which results in a high efficiency and a lack of mass transfer limitations in the polarization curve (Figure 3a).

In this work we have used a protein engineering approach to make a new catalytic biomaterial that supports a synthetic metabolic pathway and can be used as an anode modification for a biofuel cell. The protein engineering strategy enabled the enzymes to self-assemble while either retaining or enhancing the catalytic performance of the enzymes. Although the biofuel cell created in this work was not optimized, it still resulted in higher power densities that have been previously reported in the literature for similar systems. Palmore and coworkers reported a methanol/O_2 enzymatic biofuel cell using diaphorase and methyl viologen for the oxidation of NADH with a maximum power density of 0.68 mW cm^{-2}.[8a] In later work, Akers and coworkers constructed a methanol/oxygen enzymatic biofuel cell and Addo and coworkers made methanol/oxygen (air) enzymatic biofuel cells using salt-extracted tetrabutylammonium bromide/Nafion membranes to immobilize enzymes. The maximum power densities of these biofuel cells were 1.55 mW cm^{-2}[8b] and 261±7.6 µW cm^{-2}[8c] respectively. Higher power can certainly be obtained by further optimizing this system and by stacking multiple cells together.

On the large-scale, enzymes can be produced inexpensively, and these hydrogels may be a competitive alternative to other chemical catalysts for methanol oxidation. This general approach can also be applied to other enzymes so that new biocatalytic biomaterials could be used for additional bioelectrocatalysis applications as well as other technologies involving heterogenous biocatalysis.

[1] a) S. Banta, I. R. Wheeldon, M. A. Blenner, *Annu. Rev. Biomed. Eng.* **2010**, *12*, 167-186; b) S. Banta, Z. Megeed, M. Casali, K. Rege, M. L. Yarmush, *J. Nanosci. Nanotechnol.* **2007**, *7*, 387-401; c) D. N. Woolfson, M. G. Ryadnov, *Curr. Opin. Chem. Biol.* **2006**, *10*, 559-567; d) W. F. Daamen, J. H. Veerkamp, J. C. van Hest, T. H. van Kuppevelt, *Biomaterials* **2007**, *28*, 4378-4398; e) S. A. Maskarinec, D. A. Tirrell, *Curr. Opin. Biotechnol.* **2005**, *16*, 422-426; f) F. G. Omenetto, D. L. Kaplan, *Science* **2010**, *329*, 528-531.
[2] a) S. Calabrese Barton, J. Gallaway, P. Atanassov, *Chem. Rev.* **2004**, *104*, 4867-4886; b) S. D. Minteer, B. Y. Liaw, M. J. Cooney, *Curr. Opin. Biotechnol.* **2007**, *18*, 228-234; c) P. Atanassov, C. Apblett, S. Banta, S. Brozik, S. Calabrese Barton, M. Cooney, B. Y. Liaw, S. Mukerjee, S. D. Minteer, *Interface* **2007**, *Summer 2007*, 28-31.
[3] a) W. A. Petka, J. L. Harden, K. P. McGrath, D. Wirtz, D. A. Tirrell, *Science* **1998**, *281*, 389-392; b) W. Shen, R. G. H. Lammertink, J. K. Sakata, J. A. Kornfield, D. A. Tirrell,

86

Macromolecules **2005**, *38*, 3909-3916; c) W. Shen, K. Zhang, J. A. Kornfield, D. A. Tirrell, *Nat. Mater.* **2006**, *5*, 153-158; d) H. R. Marsden, A. Kros, *Angew. Chem. Int. Ed.* **2010**, *49*, 2988-3005.

[4] a) I. R. Wheeldon, S. C. Barton, S. Banta, *Biomacromolecules* **2007**, *8*, 2990-2994; b) I. R. Wheeldon, E. Campbell, S. Banta, *J. Mol. Biol.* **2009**, *392*, 129-142; c) H. D. Lu, I. R. Wheeldon, S. Banta, *Protein Eng. Des. Sel.* **2010**, *23*, 559-566; d) I. R. Wheeldon, J. W. Gallaway, S. C. Barton, S. Banta, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 15275-15280.

[5] M. C. Sheehan, C. J. Bailey, B. C. Dowds, D. J. McConnell, *Biochem. J.* **1988**, *252*, 661-666.

[6] a) A. A. Klyosov, *Biochemistry* **1996**, *35*, 4457-4467; b)H. Glatt, K. Rost, H. Frank, A. Seidel, R. Kollock, *Arch. Biochem. Biophys.* **2008**, *477*, 196-205.

[7] A. E. Serov, A. S. Popova, V. V. Fedorchuk, V. I. Tishkov, *Biochem. J.* **2002**, *367*, 841-847.

[8] a) G. T. R. Palmore, H. Bertschy, S. H. Bergens, G. M. Whitesides, *J. Electroanal. Chem.* **1998**, *443*, 155-161; b) N. L. Akers, C. M. Moore, S. D. Minteer, *Electrochim. Acta* **2005**, *50*, 2521-2525; c) P. K. Addo, R. L. Arechederra, S. D. Minteer, *Electroanalysis* **2010**, 22, 807-812.

[9] a) K. V. Mills, H. Paulus in *Homing Endonucleases and Inteins, Vol. 16* (Eds.: M. Belfort, D. W. Wood, B. L. Stoddard, V. Derbyshire), Springer Berlin Heidelberg, **2005**, pp. 233-255; b) Chong, M.-Q. Xu in *Homing Endonucleases and Inteins, Vol. 16* (Eds.: M. Belfort, D. W. Wood, B. L. Stoddard, V. Derbyshire), Springer Berlin Heidelberg, **2005**, pp. 273-292; c) K. Shingledecker, S.-q. Jiang, H. Paulus, *Archives of Biochemistry and Biophysics* **2000**, *375*, 138-144; d) W.-Y. Wu, C. Mee, F. Califano, R. Banki, D. W. Wood, *Nat. Protocols* **2006**, *1*, 2257-2262.

[10] A. A. Karyakin, E. E. Karyakina, W. Schuhmann, H.-L. Schmidt, S. D. Varfolomeyev, *Electroanalysis* **1994**, *6*, 821-829.

[11] W. Gellett, J. Schumacher, M. Kesmez, D. Le, S. D. Minteer, *J. Electrochem. Soc.* **2010**, *157*, B557-B562.
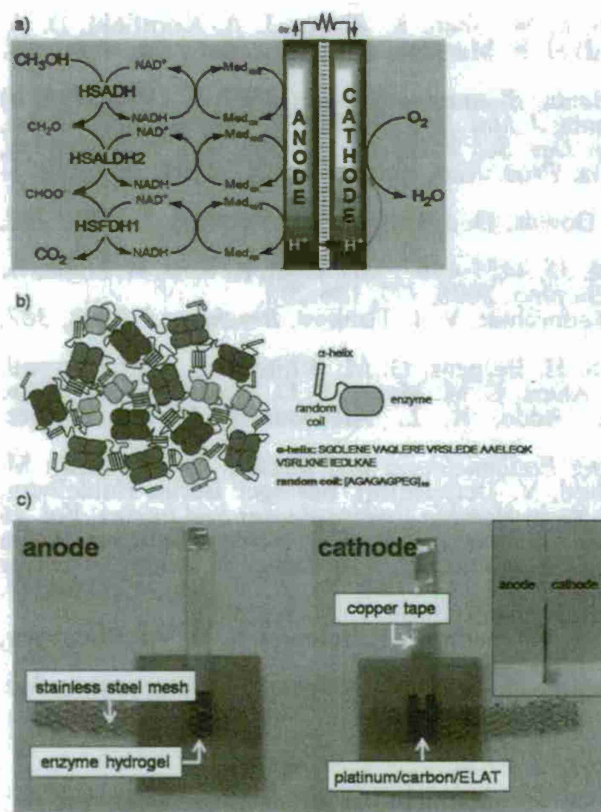
Figure 1. a) Schematic diagram of methanol/O_2 enzymatic biofuel cell. b) Left: Schematic representation of the physically crosslinked enzyme hydrogel (red: ADH, purple: ALDH2, and green: FDH1), Right: Schematic of protein modification where an alpha helix and random coil are added to the N-terminus of each monomer. c) Picture of methanol/O_2 enzymatic biofuel cell using the self-assembled protein hydrogel on the anode.
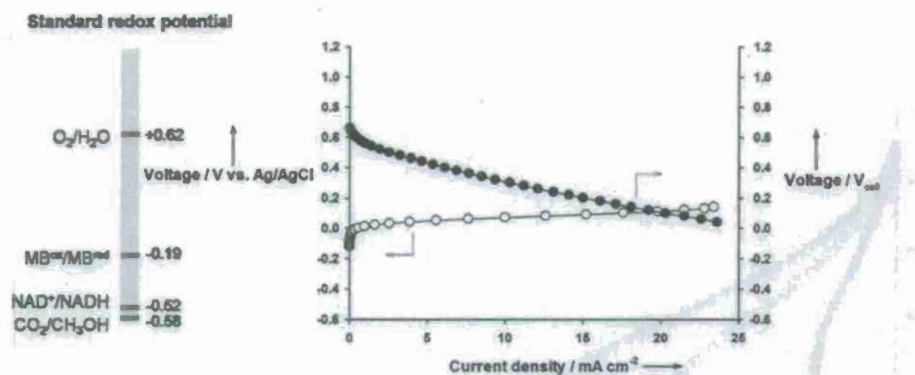
Figure 2. Representative polarization curves for the hydrogel modified anode (O) and the complete enzymatic biofuel cell (●).
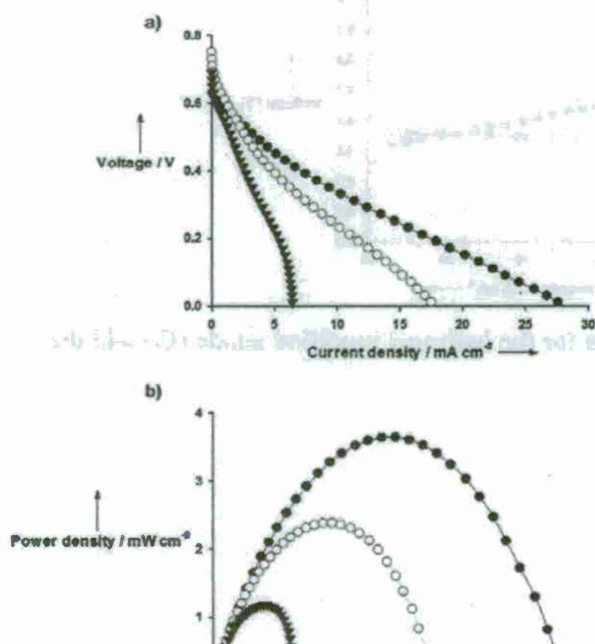
Figure 3. Performance of the enzymatic fuel cell. (a) polarization and (b) power curves when different fuels are added to the anode; methanol (●), formaldehyde (○), and formate (▼).

Table 1. Kinetic parameters of hydrogel forming enzymes for the oxidation of methanol (HSADH), formaldehyde (HSALDH), and formate (HSFDH).

	k_{cat}	$K_{m,S}$	K_{m,NAD^+}	k_{i,NAD^+}
	$[s^{-1}]$	[mM]	[mM]	[mM]
HSADH[a]	7.2±0.53	$(12.0±9)×10^{-3}$	$(7.6±2.6)×10^{-3}$	$(69.5±72.8)×10^{-3}$
ADH[b]	0.06	20	-	-
HSALDH2[a]	58.3±19.8	21.7±12.1	0.14±0.18	0.21±0.14
ALDH2[c]	67.5±8.33	0.32±0.08	-	-
ALDH2[d]	21.8	0.42±0.08	-	-
HSFDH1[a]	1670±80	7.2±1.5	0.8±0.1	1.7±0.6
FDH1[e]	6.5±0.4	5.5±0.3	0.036±0.005	

[a] T=37 °C, pH 8.5; [b] T=37 °C, pH 7.0[3]; [c] T=25 °C, pH 9.5[6a]; [d] T=ambient temperature, pH 9.5[6b]; [e] T=30 °C, pH 7.0[7].

Chapter 7

Title: Protein engineering in the development of functional hydrogels

Authors: Scott Banta, Ian R. Wheeldon[a] and Mark Blenner[b]

Department of Chemical Engineering, Columbia University in the City of New York, New York, NY, 10027

Current addresses:

a. Department of Medicine, Center for Biomedical Engineering, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

b. Immune Disease Institute, Department of Pathology, Harvard Medical School, Boston, MA 02115, USA

Email addresses:

Scott Banta: sbanta@cheme.columbia.edu

Ian Wheeldon: iwheeldon@rics.bwh.harvard.edu

Mark Blenner: blenner@idi.harvard.edu

Corresponding author: Scott Banta, Associate Professor, 820 Mudd, Mail Code: 4721

Phone: +1 212-854-7531   Fax: +1 212-854-3054

Abstract Proteins, which are natural heteropolymers, have evolved to exhibit a staggering array of functions and capabilities. As scientists and engineers strive to tackle important challenges in medicine, novel biomaterials continue to be devised, designed, and implemented to help to address critical needs. This review aims to cover the present advances in the use of protein engineering to create new protein and peptide domains that enable the formation of advanced functional hydrogels. Three types of domains are covered in this review, (*a*) the leucine zipper coiled-coil domains, (*b*) the EF-hand domains, and (*c*) the elastin-like polypeptides. In each case, the functionality of these domains is discussed as well as recent advancements in the use of these domains to create novel hydrogel-based biomaterials. As protein engineering is used to both create and improve protein domains, these advances will lead to exciting new biomaterials for use in a variety of applications.

1. INTRODUCTION

In nature we find myriad examples of consensus amino acid sequences that reliably fold in to three-dimensional structures with distinct functions, called protein domains. Researchers have employed a number of different domains to find solutions to challenges in biomedical engineering materials such as scaffolds for tissue engineering, active or reactive drug delivery systems, and in vivo and in vitro biosensors (1--9). These applications often benefit from bioactive, stimuli-responsive, or "smart" **hydrogel** biomaterials, i.e., the biomaterial undergoes a change in response to the local environment. The event (the action, reaction, or phenomenon) that occurs in response to the environmental change or cue is a result of the structure, and consequent function, of the protein domain or motif incorporated into the biomaterial. For example, β-sheets can form ordered supramolecular structures in response to temperature changes (6, 10, 11); α-helical coiled coils form pH-dependent physical crosslinks in hydrogel structures (12--14); the calcium-binding protein camodulin undergoes a conformation change upon ligand binding causing macroscopic changes in hydrogel structures (15, 16); elastin-like polypeptides undergo a reversible inverse phase transition to form insoluble β-spiral aggregates (17, 18); and controlled degradation of a hydrogel matrix is possible with the incorporation of protease cleavage sequences (19, 20).

Because the protein domain is the center of the responsive action or "smart" phenomena, manipulation of the domain can be used to control the material properties of the hydrogel. The protein engineering tool set allows for the manipulation of natural and artificial DNA sequences encoding the peptides or proteins of interest, as well as the subsequent biological production of the translated products (21). This methodology is powerful in that it allows for exact control over the identity and sequence of each residue and, consequently, the structural folding patterns of the resultant protein or peptide. Of the generally accessible biological expression systems (yeast, mammalian, and bacterial), heterologous expression in *Escherichia coli* is the most common. Furthermore, the availability of commercial prepackaged kits for traditional molecular biological experiments and techniques has made this methodology accessible to the

nonexpert molecular biologist. Functional peptides of ~100 amino acids can also be produced by solid-phase polypeptide synthesis (22). A schematic of the **protein engineering methodology of responsive materials design** is presented in **Figure 1**.

The past ten or more years have produced a number of successful examples that use variations of this methodology. As a number of reviews have presented excellent surveys of the broad range of protein-based functional materials (2, 3, 5), we review only the examples that best present the case for protein engineering design of responsive or smart hydrogels. A summary of these examples is given in Table 1. We focus on the aspects of the protein domains that bring about the desired action and the chemical and physical phenomena that drive the action. In doing so we review not only those works that present responsive materials but also many protein engineering and chemical biology efforts that have led to the understanding of the mechanisms of protein domains and motifs used in the responsive materials. We also highlight protein domains with unrealized potential for hydrogel materials applications.

One of the most widely exploited protein domains is the α-helical coiled-coil or leucine zipper domain. Coiled coils have been extensively studied (23--29) and have been successfully used as physical crosslinks in hybrid hydrogels and supramolecular hydrogel structures (12--14). These engineering efforts have led to the development of hydrogels that structurally respond to pH, temperature, and ionic strength (30--35). Control over the number of strands per bundle, strand identity, and bundle strength is also possible (23, 31, 36, 37). Beta-sheet, β-turn, and elastin-like peptides can also form reversible physical crosslinks that support hydrogel structures (6, 17, 38). Biomaterials that are structurally responsive to light, shear stress, pH, temperature, and ionic strength have been engineered (11, 39--42). In addition to physical crosslinking, the conformational change of elastin-like domains has been exploited to manipulate hydrogel swelling and degradation rate (17, 18). Finally, the ligand and calcium-binding domains of calmodulin have also been extensively studied (43--47) and have proven successful in altering macroscopic properties of hybrid hydrogels in response to binding events (15, 48--50)

Nature is rich with diversity: There are many protein domains and peptide motifs with interesting ligand binding properties, with structures that undergo conformation changes, and that have complex functions yet to be used in responsive materials. Thus far, the protein domains and peptide motifs used in responsive materials have been quite successful in imparting specific and controllable action. Advancements in the three-dimensional architecture of hydrogel structures for tissue engineering applications, advancements in active and passive control of drug delivery, and the development of novel biosensors have been made. We foresee the extension of this work toward the development multifunctional hydrogels that combine the characteristics of a number of protein domains to create more complex systems that will address biomedical engineering problems yet unresolved.

2. LEUCINE ZIPPER COILED COILS

2.1. Hydrogel Applications

The leucine zipper domains, or simply leucine zippers, are a structural motif commonly found in transcription factors. Much of the pioneering work in elucidating the structure

94

of the motif was focused on the fos and jun oncogenes and the GCN4 transcription factor (23, 24, 36, 37, 51). The helices are characterized by a heptad repeat of the form *abcdefg* where *a* and *d* are leucine, or nonpolar, residues and *e* and *g* are charged residues. The side chains of the nonpolar leucine residues lie in a plane along the length of a helix; the hydrophobic nature of the plane leads to the formation of multistranded coiled coils. The motif's name reflects the predominance of leucine residues at the *a* and *d* positions.

The aggregation or assembly of two or more domains into a coiled-coil bundle is reversible with changes in pH, temperature, and ionic strength. The conditions under which assembly occurs are dependent on the primary sequence. The reversible assembly makes the leucine zipper domain ideal to serve as responsive physical crosslinks in hydrogel structures. Pekta et al. first demonstrated that a supramolecular hydrogel structure assembled from protein subunits could be made responsive to temperature and pH (12). A triblock polypeptide with terminal leucine zipper domains and a central randomly coiled domain self-assembles at neutral pH and temperatures below 40°C. A physically crosslinked hydrogel network forms as the leucine zippers aggregate into tetrameric coiled coils. Wang et al. first described the use of leucine zipper domains in a hybrid synthetic polymer-protein material (13). The hybrid material undergoes a volume change in response to temperature change as leucine zipper coiled coils dissociate at high temperature. Since the publication of these two works, a number of studies have shown that the temperature, ionic strength, and pH at which hybrid and protein materials undergo volume change and hydrogel formation can be engineered at the primary sequence level (28--32, 34, 35).

Physical crosslinking with leucine zipper coiled coils has also been used to create cell binding scaffolds. Triblock polypeptides with leucine zipper domains and a central domain containing the cell adhesion ligand, arginine-glycine-aspartic acid (RGD), have been shown to bind cells in a reversible manner (52--54). Triblock polypeptides with antibody binding ability have also been demonstrated (55). In our lab we have used leucine zipper fusions to create enzymatic protein building blocks that self-assemble into enzymatic hydrogels (56, 57). These works combine the self-assembly function of leucine zippers with the catalytic function of an enzyme creating bifunctional building blocks. We have also shown that different building blocks with compatible leucine zipper domains can assemble into multifunctional hydrogel structures (14). The association of α-helices in coiled-coil bundles can be transient (27, 33, 34), i.e., a helix from one bundle can exchange places with a helix from a different bundle. This phenomenon has been used to tailor the rate of erosion of supramolecular triblock polypeptide structures (33).

A number of recent publications have shown the utility of α-helices in fiber formation (58--60). These exciting works have recently been reviewed (61) and are therefore not covered here. In the cases presented in this review, the leucine zipper domain has been used as a physical crosslink, both as the only source of crosslinking in supramolecular protein structures (**Figure 1**, top left) and as a minor source of crosslink density (**Figure 1**, top right). The domain has been used to impart responsive and reversible behavior in hydrogels. The utility of the domain lies in the robustness of the structure and in the long history of investigation into the mechanisms of coiled-coil formation.

2.2. Coiled-Coil Formation Oligomerization Number

Leucine zipper domains form supercoils of two or more α-helical strands with a characteristic heptad repeat of *abcdefg*. The natural coiled-coil shape is a left-handed supercoil of two or more right-handed α-helices. The structure of the helices results in 3.5 residues per turn, as opposed to the 3.6 residues per turn common to the α-helical secondary structure. The leucine residues, or isoleucine or valine residues, at positions *a* and *d* form a hydrophobic plane along the length of each helix. Charged residues occupy positions *e* and *g* and charged or neutral residues are commonly found at positions *b, c,* and *f* (23, 36). Formation of the coiled-coil structure is driven by the hydrophobic effect as the hydrophobic plane of the *a* and *d* positions buries within the core of the motif. Using the dimer structure shown in **Figure 2** (top, left) as an example, the motif can be envisioned as a ladder with the backbones of the α-helices as the sides, and the sets of *a* and *g* and *d* and *e* residues (along with the *a'* and *g'* and *d'* and *e'*) as the rungs.

The X-ray crystal structures shown in **Figure 2** are a set of examples of coiled coils demonstrating the diversity of natural and engineered structures. The series is also historical, in that the discovery or design is from earliest (left) to most recent (right). The dimer structure (left) is that of the GCN4 leucine zipper domain (the final 33 C-terminal residues of the full GCN4 protein) and was first solved by Kim et al. (51). The heptamer structure (right) is an engineered seven-helix bundle designed by Liu and coworkers (62). Parallel (**Figure 2**, top, third from the left) and antiparallel (**Figure 2**, top, second from the right) orientations are also possible (63).

Alber et al. demonstrate that dimeric, trimeric, and tetrameric coiled-coil configurations are possible by altering the packing characteristics of the hydrophobic residues at positions *a* and *d*. The identity of the *a* and *d* residues drives the formation of different quaternary structures owing to the destabilizing effects of isoleucine and valine in specific packing arrangements (37). As can be seen in the end view of each example, the center of the core forms a channel. At 1. 0 to 1. 3 Å in diameter, the channel in tetra-, tri-, and dimers is too small to allow water to enter the hydrophobic core. Coiled coils with five (64) and seven (62) (top, right) strands have also been demonstrated.

The homodimer forms a shallow left-handed superhelix with parallel strands. The structure shows the typical core packing as well as the presence of salt bridges between the charged side chains at positions *e* and *g'*, and *g* and *e'*. The asparagine side chain visible in the core of the structure forms hydrogen bonds with its opposing partner. Despite a destabilization of helical and coiled-coil structure, asparagines are commonly found in native coiled-coil structures to promote specific dimerization (24, 51).

2.3. Engineering Coiled-Coil Stability and Specificity

Substantial effort has been devoted to understanding the specificity of homo- and heterodimerization (25, 27, 65). A seminal work by O'Shea et al. demonstrates that homo- and heterodimerization is driven not by stabilizing salt bridges of the *e* and *g* residues but by the destabilizing effect of homodimers with similar charge (24). In a detailed study of the length of the charged side chains at *e* and *g*, Ryan & Kennan demonstrate that strand exchange can occur in a dimer pair where a more stabilizing strand replaces a less stable strand (27). They do not conclude that the salt bridges have an overall stabilizing effect, but that preference is made toward the more stable dimer pair. In fact, it is has been shown that some salt bridges are destabilizing.

A number of works have explored leucine zipper coiled-coil temperature stability as a function of domain length. Sui et al. present evidence of a three-heptad repeat minimum for dimer formation, and also demonstrate that temperature stability increases with increasing length (66). Kopecek and coworkers show that the correlation between temperature stability and domain length is also observed with triblock polypeptides with terminal leucine zipper domains (30); however, coiled-coil stability does not necessarily translate into hydrogel stability as the coiled-coil melting temperature, as measured by circular dichroism spectroscopy, is greater than the sol-gel transition temperature of protein hydrogels assembled from triblock polypeptides (12, 32).

2.4. Underexplored Helical Structures—Heme-Binding Helical Bundles

There are many helical motifs and domains that have yet to be used or that have only recently been used in responsive materials design. The assembly of long fibers from the association of designed α-helices (58--61) is such an example. One possibility with great potential yet to be explored is the incorporation of heme-binding helical bundles into responsive protein and hybrid hydrogels. An example of a heme-binding helical bundle is shown in figure 3. Dutton and coworkers have designed four-helix bundles with controlled heme binding (67, 68), and have recently demonstrated the design and engineering of a dioxygen transport protein (69). Shu et al. have successfully constructed peptide-polymer conjugates with PEG (polyethylene glycol) polymer chains selectively bound to the exterior side chains of a trimeric heme-binding domain (70). These works, in combination with the development of self-assembled and responsive hydrogels could lead to new and very exciting avenues of protein-based functional materials design and engineering.

3. CAMODULIN AND THE EF-HAND DOMAIN

3.1. Responsive Hydrogel Applications

The EF-hand motif is a calcium-binding domain common to many calcium-binding proteins and calcium sensor proteins that are essential for intracellular signal transduction. The domain is characterized by a helix-loop-helix structure that binds Ca^{2+} in a cooperative manner (44, 47). The prototypical calcium sensor protein is calmodulin (CaM). When the four EF-hand domains of CaM bind cystolic Ca^{2+} the protein undergoes a conformational change to reveal a ligand binding site. CaM can bind many different ligands, but the target is often an enzyme that becomes activated. In this way CaM facilitates the transduction of a cystolic Ca^{2+} signal into a biochemical action (46).

The structure of CaM is such that it undergoes two distinct conformational changes. Upon calcium binding the unstructured-globular CaM takes on an extended dumbbell shape (Figure 2, center). The extended state is an activated state, as a ligand binding site is revealed. A second conformational change occurs upon ligand binding as the extended dumbbell shape collapses around a peptide ligand (Figure 2, center right). Structures of all three states are shown in Figure 2 (center). The conformational change of CaM upon Ca^{2+} and ligand binding are substantial. The distance between the globular ends of CaM in the extended state is approximately 50 Å (71). The distance is reduced to approximately 15 Å upon ligand binding (72). The mechanical action in response to biological cues is ideal for engineering responsive hydrogels. Two excellent examples

97

of the protein engineering methodology of responsive materials design are those from Murphy and coworkers (16, 49, 50, 73) and from Ehrick et al. (15, 48).

Ehrick et al. first describe an acrylamide hydrogel with CaM and a phenothiazine peptide ligand covalently bound to the hydrogel network (15). In the presence of Ca^{2+}, CaM binds the immobilized phenothiazine ligand, thus adding to the crosslink density of the hydrogel. Chelation of Ca^{2+} from the EF hand-domains of CaM inhibits ligand binding and results in an increase in hydrogel volume as the physical crosslinks formed by CaM-phenothiazine interactions are disrupted. Swelling is reversible with the subsequent addition of calcium. This mechanism is schematically represented in **Figure 1** (top, right). An increase in hydrogel volume also occurs in response to soluble CaM-binding ligand cues. In the presence of a soluble ligand that binds with higher affinity than the immobilized phenothiazine, CaM-phenothiazine interactions are replaced with CaM-soluble ligand interactions, resulting in an increase in volume.

The Murphy group has also produced CaM-based responsive hydrogels. In their approach, CaM is modified with diacrylated poly(ethylene glycol) (**PEGDA**) to produce a triblock PEGDA-CaM-PEGDA prepolymer solution. Hydrogels responsive to Ca^{2+} and CaM-ligand binding are made upon UV crosslinking of PEGDA-CaM-PEGDA monomers in the presence of photoinitiator. The mechanical action of the bound CaM in response to binding events can be modulated by tuning the PEGDA chain length. With PEGDA of MW 575, Sui et al. produce a hydrogel that undergoes an 80% change in volume in response to binding the antipsychotic drug trifluoperazine in the presence of Ca^{2+} (49). The volume change is reversible, losing only minimal dynamic range after 10 swelling/deswelling cycles.

One key to incorporating the CaM domain into both the acrylate and photocurable PEGDA hydrogels is the mutation of two threionine residues, one at each end of the extended dumbbell conformation. Mutation of the threonines to cysteines (T34C and T110C) does not disrupt Ca^{2+} or ligand binding and allows for selective modification of the cysteine residues. The thiol side chains of the cysteine residues are available for modification by Michael-type addition to acrylate groups (16), by reaction with N-succinimidylacrylate (NHS) chemistries (15) or by other thiol-specific chemistries.

The mechanical action and conformational change of the calmodulin domain are functionalities that are much different in character than the functionalities of the leucine zipper coiled coils. These functionalities, mechanical action and conformational change, have been shown to be adaptable to hydrogels to produce a high degree of swelling and volume control, characteristics that make these systems amenable to applications such as active drug release (50), biosensing (49), microlenses, (48) and microfluidic gates (15).

3.2. Conformational Change, and Calcium and Ligand Binding

The EF-hand domain takes its name from the E and F helices first described in the crystal structure of parvalbumin, a calcium-binding protein of the CaM superfamily (74). Much has been written about the structure and function of the EF-hand domain because of the importance of calcium signaling in intracellular processes. Here we review the mechanism of calcium and ligand binding to provide sufficient background for a detailed understanding of the responsive behavior of the systems described above. A more complete description of the structure and function of the EF-hand domain can be found in a recent review by Gifford et al. (75).

98

Typically, as is the case for CaM, EF-hand domains are found in pairs. Each pair interacts to form globular domains, each with two EF-hand domains, that cooperatively bind Ca^{2+}. In CaM, a flexible linker joins two pairs of EF-hand domains to make a 16.7-kDa protein with a pI of 4.6. Apo-CaM is globular and the loops adjoining the helices of each EF-hand domain are exposed (**Figure 2**, center, left). Each loop provides the bulk of the Ca^{2+} ligands, as five of the seven ligands per ion are donated from the residues within a flexible loop. Tight Ca^{2+} binding is a result of the high number of acidic residues that donate oxygen ligands to the bound ions. Binding affinity can be altered through mutations to the 12-residue consensus loop adjoining the helical domains of the motif (47, 76). Ca^{2+} binding causes a change in orientation of the helices in each EF-hand domain; the repositioning of the helices in turn cause a conformational change in each pair of EF-hand domains. Detailed aspects of the cooperative nature of Ca^{2+} binding and the spatial repositioning of the helix-loop-helix structure are reviewed in References 44, 46, and 75.

Cellular concentrations of Ca^{2+} range from 10^{-7}--10^{-8} M in an inactive state and up to 10^{-5}--10^{-6} M in an active state. The binding K_d of EF-hand proteins ranges from 10^{-4} to 10^{-9} M (47), but at low concentrations (~10^{-7} M), the EF-hand domains of CaM are inactive and do not bind Ca^{2+}. CaM is active when the Ca^{2+} concentration reaches ~10^{-6} M. The conformation change that occurs from apo-CaM to CaM with the binding of four Ca^{2+} ions results in a structural change from globular to extended states (**Figure 2**, center, middle). A hydrophobic patch is revealed in the extended, or dumbbell, state as four previously buried methionine residues within each pair of EF-hand domains are exposed. The hydrophobic patch is the site of ligand binding. The linker joining the globular ends of the extended conformation is highly flexible. Upon ligand binding a second conformational change occurs, as the flexible linker joining the globular ends of the extended conformation collapses around the ligand (45, 77) (**Figure 2**, middle, right).

3.3. Calmodulin Ligands

Calcium-binding proteins with EF-hand domains such as calbindin D_{9K} and parvalbumin are calcium buffers, i.e., the EF-hand domains of calcium buffer proteins act to maintain the intracellular calcium concentration and bind calcium without transmitting a signal in response to a change in calcium concentration. The exposure of the ligand binding site in the extended conformation of CaM results in the eventual transmission of the calcium-binding event to a ligand binding event. The structure of the flexible linker that contains the ligand binding site is such that it allows for the binding of many different ligands. Although binding is promiscuous, it is strong (K_d of 10^{-7} to 10^{-11} M^{-1}) (44, 46). The diversity of ligands that bind to CaM speaks to the broad importance of the protein in intracellular signaling (46).

Despite little sequence homology in CaM ligands, the structures are often similar. Generally, CaM-ligand sequences are predicted to fold into an amphiphilic helical structure (44, 46), and the amino acid sequence of peptide ligands is highly regular.[**AU: OK now, or is something missing?**ok now] CaM ligands are classified by the position of the hydrophobic residues. The 1--10 type has hydrophobic residues at positions 1 and 10, the 1--14 type has hydrophobic residues at positions 1 and 14, and so on for 1--16, and the 1--10 subclass 1-5-10 (44). An active and growing list of CaM

ligands can be found at http://calcium. uhnres. utoronto. ca/ctdb. Combined, the regularity of sequence pattern and the diversity of potential ligand provide a powerful advantage to CaM-based biosensors.

3.4. Underexplored Conformational Changing Domains---☐☐ Rolls

The beta roll domain is a calcium-binding motif (Figure 3a) common to proteins secreted through the bacterial Type 1 Secretion System (T1SS) (78). Two parallel beta sheets connected by calcium-binding turns form a parallel beta helix (79). The beta roll sequence is comprised of tandem repeats of GGXGXDXUX, where X can be any residue and U is a hydrophobic amino acid (80). GGXGXD forms the calcium-binding turn and XUX forms a short ☐strand, completing one half of a beta helix turn. Calcium is coordinated by aspartate amino acids (D) and the glycine amino acids (G) allow turns to form above and below the calcium ions. NMR and biophysical data suggest the beta roll is largely unstructured in the absence of calcium (81). The beta roll likely acts as a calcium-induced structural switch preventing premature folding of proteins utilizing the T1SS (82). Once in the extracellular environment, beta roll formation prevents diffusion back through the outer membrane pore.

The beta roll from *Serratia marcescens* serralysin was used to modulate the distance between 6-phospho-☐galactosidase (PGAL) proteins to demonstrate the possibility of using responsive proteins to tune biomaterial pore size. Electron micrograph data show that chelating calcium with EDTA induced 3-13-nm changes in PGAL separation (83). Theoretically, the repeating structure of the beta roll should make possible engineering efforts to control protein separation by simply adding more beta roll repeats to the desired constructs.

Studies of the calcium-responsiveness of the isolated beta roll have proven difficult, requiring either folding aids, such PEG (84), or the inclusion of natural nonbeta roll flanking sequences (81). Recent work in our lab has demonstrated that nonnative globular proteins are also capable of enabling calcium-responsive behavior (85, 86). This finding suggests that beta rolls may find utility as calcium-responsive elements in protein-based hydrogel systems.

4. ELASTIN AND ELASTIN-LIKE PEPTIDES

4.1. Hydrogel and Responsive Hydrogel Applications

Engineering of elastin-like peptides (ELPs) has matured greatly over the past 30 years and is one of the best characterized proteins used for tissue engineering and drug delivery applications. ELPs are peptides 5 to 1500 amino acids in length, generally made from the pentamer consensus sequence, VPGXG, where X is any amino acid. ELPs exist in disordered conformations below their transition temperatures and form more ordered ☐turns above their transition temperatures (**Figure 2**, bottom). Above the transition temperature, ELPs form reversible proteinaceous networks comprised of repetitive ☐ turns forming a ☐spiral. The combination of ELPs' inherent elasticity and the inverse phase transition has led to the development of different materials applications including synthetic-ELP hybrid hydrogels with reversible swelling, inverse phase transition hybrid materials, and other switchable interfaces.

ELP-cell interactions have been well studied. Some ELP technology may be considered mature in the tissue engineering and responsive biomaterials fields. ELPs are generally incorporated into polymer-based tissue engineering scaffolds for their (a) desirable mechanical properties. Other tissue engineering applications for ELPs involve (b) genetic manipulation of ELPs for favorable crosslinking chemistry, and (c) using the inverse phase transition of ELPs to create a reversible physically crosslinked network. In each application, bacterial expression of ELPs enables exquisite control of ELP mechanical, chemical, and materials properties. Furthermore, bacterial expression offers ELP monodispersity that is otherwise unattainable with synthetic approaches.

ELPs are most often employed for their mechanical properties. ELPs with lysine residues periodically used as the guest residue were chemically crosslinked to form hydrogels with reversible swelling properties (87), which are able to form without detriment to implanted cell viability (88). Conjunctival epithelial cells can be proliferated on ELP substrates (89). By creating ELP blocks of lysine-containing hydrophobic ELPs, and blocks containing aliphatic and hydrophilic ELPs, the mechanical properties of the hydrogel can be tuned (87). Chemical crosslinking takes advantage of the lysine residues, and hydrogel crosslink density can be controlled via the lysine molar fraction. Because ELPs can be genetically encoded, other responsive domains can be incorporated into the hydrogel using standard molecular biology (17, 18, 90). ELP was injected into osteochondral defects and crosslinked in situ, showing increased infiltration over three months (91). Using an ELP hydrogel with urokinase plasminogen activator cleavage sites and RGD cell adhesion sites, PC-12 cells were able to grow, differentiate, and exhibited extensive neurite outgrowth (92). The ELP inverse phase transition was used to create a cell sheet---and after confluence, the temperature was dropped below the transition temperature and the sheet was detached (93). In drug delivery applications, ELPs are used for their acervation properties, where upon inverse phase transition, soluble ELPs form an aggregate. Using local hyperthermia, the ELP phase transition was used to target ELPs to solid tumors (94). ELPs have also been used in creating switchable interfaces (95). ELP was fused to Interleukin 1 Receptor antagonist (IL1-Ra) to create a persistent "depot" of IL1-Ra. Proteolytic cleavage released soluble IL1-Ra. ELPs have been used to control the attachment of cells in cell-based biochips (96). Hyperthermophilic targeting of an anticancer lactoferrin L12 domain to a Tat-ELP caused a 30-fold increase in cytotoxicity of pancreatic adenocarcinoma cells in vitro (97, 98). In situ depot formation of ELP aggregates encapsulating antibiotics has prolonged release properties with first-order time constants ranging from days to weeks, opening up the possibility of using ELPs to prolong drug delivery (99).

4.2. Conformational Change, Inverse Phase Transition

There are several excellent reviews on ELP conformational change and inverse phase transition (90, 100--102). We cover these topics only in sufficient detail to understand how ELP inverse phase transition can be used in tissue engineering and drug delivery applications, and general trends observed in the engineering of ELP properties. ELPs are frequently composed of repeating pentamers with the sequence VPGXG, where X is known as the guest residue. The guest residue can be any amino acid except Pro. The prototypical ELP contains many repeats of the VPGVG unit, denoted $(VPGVG)_n$. At lower temperatures, this biopolymer is soluble in aqueous environments and forms a

101

crosslinked network above its inverse transition temperature (103). Structural information for ELPs is limited because the crosslinked ELP is insoluble. Various NMR, X-ray, CD, and FTIR studies have shown that in its soluble form, ELPs are mostly disordered, whereas crosslinked ELPs are higher in β-turns and form β-spiral structures (104–106). The inverse temperature transition (T_t) is named as such because typically proteins lose structure upon increasing temperature. On the contrary, ELPs become more structured. ELPs are believed to be hydrated below T_t, and increasing temperature causes dehydration and β-turn formation within each ELP repeat (101, 107, 108). Entropy increases associated with the liberation of structured water of hydrophobic hydration drive ELP structure formation. The dehydrated, structured ELP then aggregates in order to bury more hydrophobic regions and minimize hydration, forming a physically crosslinked network. This inverse temperature transition is reversible (109). Here, increasing temperature drives ELP conformational change that results in an aggregate reversibly crosslinked network. In addition to thermal responsiveness, crosslinking can occur via any mechanism by which the hydrophobicity of the ELP can be increased (110). The prominent means to enhance hydrophobicity are increasing chain length, ionic strength, and pH. In each of these cases, the enhanced hydrophobicity acts to effectively lower the T_t, thereby allowing isothermal phase transitioning (111).

4.3. Responsiveness Tuning: Temperature, Ionic Strength, pH, Chain Length, Fusion

ELPs are highly desired biomaterials because they can be genetically encoded through recursive directional ligation, allowing precise control of chemical composition as well as a degree of monodispersity not possible using synthetic methods (112). Furthermore, the conditions under which ELPs can undergo inverse phase transition are predictable and therefore can be engineered. Depending on the application, a variety of methods are available for tuning responsiveness. The most obvious method is to change the identity and molar fraction of the guest residue. Methods to predict the effect of the guest residue and guest residue molar fraction are extensively covered elsewhere (113). The effect of chain length and ELP concentration on phase transition has been described previously (114). Increasing concentration or increasing chain length both decreases T_t. Longer ELPs have a weaker dependence on concentration. Increasing ionic strength has the effect of enhancing hydrophobicity and decreasing T_t (101). Changing pH likewise changes the ionization of certain amino acid side chains, causing a change in hydrophobicity and T_t (110).

Fusing ELPs to other proteins allows the ELP to make the protein responsive to stimulus (115, 116). However, the surface hydrophobicity of the fused protein likewise affects the inverse phase transition of the ELP (117). Since surface hydrophobicity can tune the inverse phase transition, fusion to allosteric calmodulin, whose surface hydrophobicity changes upon calcium and ligand binding, has been used to modulate inverse phase transition of a calmodulin-ELP fusion (118). This concept should be able to extend to ELP fusions with other responsive protein domains. Should protein domains and protein-ligand pairs exist, be discovered, or designed that change surface hydrophobicity upon binding a target biomolecule of interest, then fusing an ELP to this protein domain may make possible specific target-driven hydrogel formation.

5. CONCLUSION

102

Increasingly, environmentally responsive or smart hydrogels are finding use in biomedical applications such as active drug delivery, biosensing, and tissue engineering. The technologies that create a change in bulk hydrogel properties in response to the environmental cue are often protein domains or peptide motifs. Here we present evidence in support of the argument for the design of responsive hydrogel materials by a protein engineering methodology. The structure and function of proteins, protein domains, and peptide motifs are encoded at the genetic level. The codon sequence within a gene encodes a primary amino acid sequence. The secondary and tertiary protein structures are dictated by the primary sequence. Function is a consequence of structure. Mutation of the DNA sequence leads to changes in the function(s) of the expressed protein, protein domain, or peptide motif. Assembled, or self-assembled, supramolecular protein structures and polymer-protein conjugates that derive responsive behavior from the protein units in the hydrogel structure are susceptible to the same control---from the genetic level to responsive function. This methodology has been put forth as a means of macromolecule design (21); responsive materials design is an extension of the broader concept of engineering the biological production of useful macromolecules.

The examples presented here, leucine zipper coiled-coil domains, the calcium-binding protein calmodulin, and elastin-like peptides, are excellent examples of the protein engineering methodology of responsive materials design. These are not the only examples, as amphiphilic peptides (119), β-sheets peptides (120), and stimulus-responsive peptides (121, 122) have also been used to create protein-based hydrogels.

The leucine zipper domain is, for materials design purposes, a reversible self-assembly domain. Hydrophobic forces drive the assembly of coiled-coil bundles as the hydrophobic planes along the length of the α-helices are buried. The mutations required to engineer responsive behavior have been mapped by the extensive literature that collectively describes the workings of the leucine zipper coiled coil. Leucine zipper coiled-coil assembly, in terms of aggregation number (37, 62), strand orientation (63), strand specificity (24, 27), and temperature stability (66), can be tuned.

The mechanical action of the CaM protein is unique to the protein domains and motifs that have thus far been used in responsive materials. Ligand binding in the presence of Ca^{2+} causes a collapse of the extended state and reduces the end-to-end distance by approximately 300%. In capturing the effect of the conformational change, hybrid synthetic polymer-protein hydrogels have been demonstrated with exquisite volume control (48, 49).

The elastin-like pentamer consensus sequence, VPGXG, has proven to be a highly useful motif. Engineering of biotechnologies using ELP repeats has matured beyond that of using either leucine zippers or calmodulin. The inverse phase transition of ELP-based hydrogels and the temperature-induced conformational change have been used to engineer responsive hydrogels for application in drug delivery (94, 95, 97, 98) and in tissue engineering (91--93). ELP technologies are also at the forefront of multifunctionailty. (118)

The responsive hydrogel works presented here are, in our view, those that best exemplify the protein engineering methodology of responsive materials design. As protein-based materials design is multidisciplinary in nature, we review many protein engineering and biological chemistry works in describing the relevant protein structures and functions. For protein-based functional materials to move forward in a significant

and meaningful way, a deep understanding of the structure and function of the relevant domains is essential. In our view, there exist many underexplored designs of materials based on the proteins, protein domains, and peptide motifs discussed above. New designs and new materials incorporating domains yet unexplored will provide new solutions to complex biomedical engineering problems.

SUMMARY POINTS

1. The protein engineering methodology of responsive materials design uses the central dogma of molecular biology to produce proteins, protein domains, and peptide motifs that are used to create responsive protein or protein-polymer conjugate hydrogels.

2. Leucine zipper domains can be used as reversible physical crosslinks in self-assembled supramolecular protein hydrogels and in hybrid polymer-protein hydrogels. The mutations required for tuning the oligomerization number, strand orientation, temperature stability, and oligomerization specificity are known and have been mapped during the course of investigating the leucine zipper coiled coils.

3. The conformational change that the calcium sensor protein calmodulin undergoes upon ligand binding can be harnessed to generate a finely tuned and reversible change in hydrogel volume. Hybrid polymer-calmodulin hydrogels sensitive to calmodulin-ligand binding can be used as simple and effective biosensors.

4. Elastin-like peptides undergo an inverse phase transition that allows triggered supramolecular assembly of biocompatible elastomeric structures for tissue engineering and drug delivery applications. Increased temperature, ionic strength, and hydrophobicity can trigger soluble unstructured ELPs to undergo a conformational change to form dehydrated beta turns and reversible aggregation. The genetic manipulation of ELPs allows one to tune the inverse phase transition, the crosslinking chemistry, and to create chimeric fusions with enhanced materials properties.

5. The structure-function relationships of the leucine zipper domain, calmodulin protein, and elastin-like peptides have been well studied. The mutations leading to changes in function have been mapped throughout the course of investigating the mechanisms of oligomerization, ligand binding, and phase transitions.

6. Compatible physical crosslinking domains such as leucine zippers can be used to create supramolecular structures from a variety of differently functional building blocks, resulting in multifunctional hydrogels.

7. Many domains with interesting and useful functions have yet to be incorporated into responsive hydrogel materials. For example, the calcium-binding β-roll domain might prove useful as a conformational change unit that is responsive to ion binding; and heme-binding helix bundles might prove useful in engineering responsive hydrogels with oxygen binding properties.

FUTURE ISSUES

1. The multidisciplinary field of protein-based responsive hydrogels is still in its infancy. There exist many possible protein engineering-based solutions to complex biomedical engineering problems such as active drug delivery, tissue engineering, and biosensing.

2. Natural systems such as biochemical pathways and organelles are highly multifunctional. Mimicking such systems may require merging many

different protein domains and functional proteins into multifunctional hydrogel structures.

3. There are many potential applications for hydrogel materials that can elicit a bulk change in response to a surface phenomenon or surface binding event. We foresee highly interesting and useful research into systems that can translate a specific biological surface event into a change in bulk mechanical property.

4. Many medical problems require in vivo solutions. We anticipate that the next evolution of responsive hydrogels will begin to address in vivo biocompatibility.

ACKNOWLEDGMENTS

Figure 1 The protein engineering methodology of responsive materials design. Control of the primary amino acid sequence at the DNA level leads to control of the three-dimensional structure and function of the responsive protein domain or peptide motif. Responsive or "smart" materials can be made in one of two manners: first, through self-assembly of protein building block into supramolecular structures; second, by incorporating the responsive protein domain into a hybrid synthetic polymer-protein hydrogel material. In each case the materials are made responsive to an external cue due to the structure and function of the protein domain(s) used in the material. Fine-tuning of the structure and consequent function of the protein domain can be achieved at the DNA level, as mutation to the protein encoding sequence leads to specific changes in the primary amino acid sequence, and the secondary, tertiary, and quaternary structures. The environmental cues are generally pH, temperature, ionic strength, and ligand binding. Calmodulin (1DMO) is shown as an example protein product; calmodulin in extended dumbbell conformation (3CLN) is shown as an example of a ligand binding protein; a two-stranded leucine zipper domain is shown as an example self-assembly domain; and an elastin-like peptide is shown as an example of a conformational change domain.

Figure 2 The structure of protein domains and peptide motifs most commonly used in protein and hybrid synthetic polymer-protein--responsive hydrogels. *Top*: Leucine zipper, coiled coils from left to right: parallel two (2ZTA), three (1GCM), four (1GCL), antiparallel four (2B1F) and parallel seven (2HY6) member coiled coils. Protein data bank files are stated in parenthesis. The parallel trimer, tetramer, and heptamer coiled coils are mutants of the GCN4 dimer structure. The series of leucine zipper coiled coils is shown to demonstrate the variety of structures that are possible through mutation to the primary amino acid sequence of the heptad repeat *abcdefg* of the leucine zipper domain. *Middle*: The conformational states of the calcium sensor protein calmodulin from left to right: unstructured-globular, nonligand binding (1DMO); extended "dumbbell", bound Ca^{2+} with revealed ligand binding sites (3CLN); globular, bound ligand (2BBM). *Bottom*: At low temperatures (T), elastin-like peptide is shown as unstructured and water soluble (*left*). Above the inverse phase transition temperature, the elastin-like peptide undergoes a conformational change, taking on a beta-turn structure (*right*).

Figure 3 New domains for responsive hydrogels. These protein domains have not yet been used to create responsive hydrogels. The structures are shown here to demonstrate the potential of these domains in engineering new responsive materials. (*a*) Beta-roll domain (2Z8X); (*b*) a heme binding helix bundle (3EH5).

LITERATURE CITED

1. Langer R, Tirrell DA. 2004. Designing materials for biology and medicine. *Nature* 428:487--92
2. Ulijn RV, Bibi N, Jayawarna V, Thornton PD, Todd SJ, et al. 2007. Bioresponsive hydrogels. *Mater. Today* 10:40--48

3. Ulijn RV, Smith AM. 2008. Designing peptide based nanomaterials. *Chem. Soc. Rev.* 37:664--75

4. Kopecek J. 2007. Hydrogel biomaterials: a smart future? *Biomaterials* 28:5185--92

5. Chow D, Nunalee ML, Lim DW, Simnick AJ, Chilkoti A. 2008. Peptide-based biopolymers in biomedicine and biotechnology. *Mater. Sci. Eng. R-Rep.* 62:125--55

6. Zhang SG. 2003. Fabrication of novel biomaterials through molecular self-assembly. *Nat. Biotechnol.* 21:1171--78

7. Peppas NA, Hilt JZ, Khademhosseini A, Langer R. 2006. Hydrogels in biology and medicine: from molecular principles to bionanotechnology. *Adv. Mater.* 18:1345--60

8. Furth ME, Atala A, Van Dyke ME. 2007. Smart biomaterials design for tissue engineering and regenerative medicine. *Biomaterials* 28:5068--73

9. Lutolf MP, Hubbell JA. 2005. Synthetic biomaterials as instructive extracellular microenvironments for morphogenesis in tissue engineering. *Nat. Biotechnol.* 23:47--55

10. Pochan DJ, Schneider JP, Kretsinger J, Ozbas B, Rajagopal K, Haines L. 2003. Thermally reversible hydrogels via intramolecular folding and consequent self-assembly of a de novo designed peptide. *J. Am. Chem. Soc.* 125:11802--3

11. Haines-Butterick L, Rajagopal K, Branco M, Salick D, Rughani R, et al. 2007. Controlling hydrogelation kinetics by peptide design for three-dimensional encapsulation and injectable delivery of cells. *Proc. Natl. Acad. Sci. USA* 104:7791--96

12. Petka WA, Harden JL, McGrath KP, Wirtz D, Tirrell DA. 1998. Reversible hydrogels from self-assembling artificial proteins. *Science* 281:389--92

13. Wang C, Stewart RJ, Kopecek J. 1999. Hybrid hydrogels assembled from synthetic polymers and coiled-coil protein domains. *Nature* 397:417--20

14. Wheeldon IR, Barton SC, Banta S. 2007. Bioactive proteinaceous hydrogels from designed bifunctional building blocks. *Biomacromolecules* 8:2990--94

15. Ehrick JD, Deo SK, Browning TW, Bachas LG, Madou MJ, Daunert S. 2005. Genetically engineered protein in hydrogels tailors stimuli-responsive characteristics. *Nat. Mater.* 4:298--302

16. Murphy WL, Dillmore WS, Modica J, Mrksich M. 2007. Dynamic hydrogels: translating a protein conformational change into macroscopic motion. *Angew. Chem. Int. Ed.* 46:3066--69

17. Chilkoti A, Dreher MR, Meyer DE, Raucher D. 2002. Targeted drug delivery by thermally responsive polymers. *Adv. Drug Deliv. Rev.* 54:613--30

18. Chilkoti A, Dreher MR, Meyer DE. 2002. Design of thermally responsive, recombinant polypeptide carriers for targeted drug delivery. *Adv. Drug Deliv. Rev.* 54:1093--111

19. Lutolf MP, Lauer-Fields JL, Schmoekel HG, Metters AT, Weber FE, et al. 2003. Synthetic matrix metalloproteinase-sensitive hydrogels for the conduction of tissue regeneration: engineering cell-invasion characteristics. *Proc. Natl. Acad. Sci. USA* 100:5413--18

20. Chau Y, Luo Y, Cheung ACY, Nagai Y, Zhang SG, et al. 2008. Incorporation of a matrix metalloproteinase-sensitive substrate into self-assembling peptides---a model for biofunctional scaffolds. *Biomaterials* 29:1713--19

21. van Hest JCM, Tirrell DA. 2001. Protein-based materials, toward a new level of structural control. *Chem. Commun.* 19:1897--904

22. Nilsson BL, Soellner MB, Raines RT. 2005. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* 34:91--118

23. Oshea EK, Rutkowski R, Stafford WF, Kim PS. 1989. Preferential heterodimer formation by isolated leucine zippers from fos and jun. *Science* 245:646--48

24. Oshea EK, Lumb KJ, Kim PS. 1993. Peptide velcro---design of a heterodimeric coiled-coil. *Curr. Biol.* 3:658--67

25. Litowski JR, Hodges RS. 2002. Designing heterodimeric two-stranded alpha-helical coiled-coils--- effects of hydrophobicity and alpha-helical propensity on protein folding, stability, and specificity. *J. Biol. Chem.* 277:37272--79

26. Schnarr NA, Kennan AJ. 2001. Coiled-coil formation governed by unnatural hydrophobic core side chains. *J. Am. Chem. Soc.* 123:11081--82

27. Ryan SJ, Kennan AJ. 2007. Variable stability heterodimeric coiled-coils from manipulation of electrostatic interface residue chain length. *J. Am. Chem. Soc.* 129:10255--60

28. Kennedy SB, Littrell K, Thiyagarajan P, Tirrell DA, Russell TP. 2005. Controlled structure in artificial protein hydrogels. *Macromolecules* 38:7470--75

29. Kennedy SB, deAzevedo ER, Petka WA, Russell TP, Tirrell DA, Hong M. 2001. Dynamic structure of a protein hydrogel: a solid-state NMR study. *Macromolecules* 34:8675--85

30. Xu CY, Kopecek J. 2008. Genetically engineered block copolymers: influence of the length and structure of the coiled-coil blocks on hydrogel self-assembly. *Pharm. Res.* 25:674--82

31. Xu CY, Joss L, Wang C, Pechar M, Kopecek J. 2002. The influence of fusion sequences on the thermal stabilities of coiled-coil proteins. *Macromol. Biosci.* 2:395--401

32. Xu CY, Breedveld V, Kopecek J. 2005. Reversible hydrogels from self-assembling genetically engineered protein block copolymers. *Biomacromolecules* 6:1739--49

33. Shen W, Zhang KC, Kornfield JA, Tirrell DA. 2006. Tuning the erosion rate of artificial protein hydrogels through control of network topology. *Nat. Mater.* 5:153—58

33. This work exemplifies the protein engineering methodology of materials design

34. Shen W, Lammertink RGH, Sakata JK, Kornfield JA, Tirrell DA. 2005. Assembly of an artificial protein hydrogel through leucine zipper aggregation and disulfide bond formation. *Macromolecules* 38:3909--16

35. Wang C, Kopecek J, Stewart RJ. 2001. Hybrid hydrogels cross-linked by genetically engineered coiled-coil block proteins. *Biomacromolecules* 2:912--20

36. Lumb KJ, Kim PS. 1995. Measurement of interhelical electrostatic interactions in the Gcn4 leucine-zipper. *Science* 268:436--39

37. Harbury PB, Zhang T, Kim PS, Alber T. 1993. A switch between 2-stranded, 3-stranded and 4-stranded coiled coils in gcn4 leucine-zipper mutants. *Science* 262:1401--07

38. Schneider JP, Pochan DJ, Ozbas B, Rajagopal K, Pakstis L, Kretsinger J. 2002. Responsive hydrogels from the intramolecular folding and self-assembly of a designed peptide. *J. Am. Chem. Soc.* 124:15030--37

39. Haines LA, Rajagopal K, Ozbas B, Salick DA, Pochan DJ, Schneider JP. 2005. Light-activated hydrogel formation via the triggered folding and self-assembly of a designed peptide. *J. Am. Chem. Soc.* 127:17025--29

40. Kayser V, Turton DA, Aggeli A, Beevers A, Reid GD, Beddard GS. 2004. Energy migration in novel pH-triggered self-assembled beta-sheet ribbons. *J. Am. Chem. Soc.* 126:336--43

41. Aggeli A, Bell M, Boden N, Keen JN, Knowles PF, et al. 1997. Responsive gels formed by the spontaneous self-assembly of peptides into polymeric beta-sheet tapes. *Nature* 386:259--62

42. Aggeli A, Bell M, Boden N, Carrick LM, Strong AE. 2003. Self-assembling peptide polyelectrolyte beta-sheet complexes form nematic hydrogels. *Angew. Chem. Int. Ed.* 42:5603--6

43. Chin D, Means AR. 2000. Calmodulin: a prototypical calcium sensor. *Trends Cell Biol.* 10:322--28

44. Yamniuk AP, Vogel HJ. 2004. Calmodulin's flexibility allows for promiscuity in its interactions with target proteins and peptides. *Mol. Biotechnol.* 27:33--57

45. Meador WE, Means AR, Quiocho FA. 1993. Modulation of calmodulin plasticity in molecular recognition on the basis of X-ray structures. *Science* 262:1718--21

46. Ikura M, Ames JB. 2006. Genetic polymorphism and protein conformational plasticity in the calmodulin superfamily: two ways to promote multifunctionality. *Proc. Natl. Acad. Sci. USA* 103:1159--64

47. Ikura M. 1996. Calcium binding and conformational response in EF-hand proteins. *Trends Biochem. Sci.* 21:14--17

48. Ehrick JD, Stokes S, Bachas-Daunert S, Moschou EA, Deo SK, et al. 2007. Chemically tunable lensing of stimuli-responsive hydrogel microdomes. *Adv. Mater.* 19:4024—27

48. This work is an excellent example of the fine control afforded by capturing the mechanical action of the calmodulin protein.

49. Sui ZJ, King WJ, Murphy WL. 2008. Protein-based hydrogels with tunable dynamic responses. *Adv. Funct. Mater.* 18:1824—31

49. The engineering of a photocurable hydrogel with conjugated calmodulin could potentially be useful for many different applications.

50. King WJ, Mohammed JS, Murphy WL. 2009. Modulating growth factor release from hydrogels via a protein conformational change. *Soft Matter* 5:2399--406

51. Oshea EK, Klemm JD, Kim PS, Alber T. 1991. X-Ray structure of the Gcn4 leucine zipper, a 2-stranded, parallel coiled coil. *Science* 254:539--44

52. Mi LX, Fischer S, Chung B, Sundelacruz S, Harden JL. 2006. Self-assembling protein hydrogels with modular integrin binding domains. *Biomacromolecules* 7:38--47

53. Fischer SE, Mi LX, Mao HQ, Harden JL. 2009. Biofunctional coatings via targeted covalent cross-linking of associating triblock proteins. *Biomacromolecules* 10:2408--17

54. Fischer SE, Liu XY, Mao HQ, Harden JL. 2007. Controlling cell adhesion to surfaces via associating bioactive triblock proteins. *Biomaterials* 28:3325--37

55. Cao Y, Li HB. 2008. Engineering tandem modular protein based reversible hydrogels. *Chem. Commun.* 2008:4144--46

56. Wheeldon IR, Gallaway JW, Barton SC, Banta S. 2008. Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks. *Proc. Natl. Acad. Sci. USA* 105:15275—80

56. This is one of the first examples of the assembly of a multifunctional proteinaceous hydrogel

57. Wheeldon IR, Campbell E, Banta S. 2009. A chimeric fusion protein engineered with disparate functionalities-enzymatic activity and self-assembly. *J. Mol. Biol.* 392:129--42

58. Woolfson DN, Ryadnov MG. 2006. Peptide-based fibrous biomaterials: some things old, new and borrowed. *Curr. Opin. Chem. Biol.* 10:559--67
59. Woolfson DN. 2005. The design of coiled-coil structures and assemblies. In *Fibrous Proteins: Coiled-Coils, Collagen and Elastomers,* ed. D Parry, J Squire, pp. 80--106 [**AU: Please list the editor(s) of this volume as well as the page range of the Woolfson article**]. San Diego: Elsevier.
60. Banwell EF, Abelardo ES, Adams DJ, Birchall MA, Corrigan A, et al. 2009. Rational design and application of responsive alpha-helical peptide hydrogels. *Nat. Mater.* 8:596--600
61. Gunasekar SK, Haghpanah JS, Montclare JK. 2008. Assembly of bioinspired helical protein fibers. *Polymers Adv. Technol.* 19:454--68
62. Liu J, Zheng Q, Deng YQ, Cheng CS, Kallenbach NR, Lu M. 2006. A seven-helix coiled coil. *Proc. Natl. Acad. Sci. USA* 103:15457--62
63. Deng YQ, Liu J, Zheng Q, Eliezer D, Kallenbach NR, Lu M. 2006. Antiparallel four-stranded coiled coil specified by a 3-3-1 hydrophobic heptad repeat. *Structure* 14:247--55
64. Malashkevich VN, Kammerer RA, Efimov VP, Schulthess T, Engel J. 1996. The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel? *Science* 274:761--65
65. Naik RR, Kirkpatrick SM, Stone MO. 2001. The thermostability of an alpha-helical coiled-coil protein and its potential use in sensor applications. *Biosens. Bioelectron.* 16:1051--57
66. Su JY, Hodges RS, Kay CM. 1994. Effect of chain-length on the formation and stability of synthetic alpha-helical coiled coils. *Biochemistry* 33:15501--10
67. Huang SS, Koder RL, Lewis M, Wand AJ, Dutton PL. 2004. The HP-1 maquette: from an apoprotein structure to a structured hemoprotein designed to promote redox-coupled proton exchange. *Proc. Natl. Acad. Sci. USA* 101:5536--41
68. Koder RL, Valentine KG, Cerda J, Noy D, Smith KM, et al. 2006. Nativelike structure in designed four alpha-helix bundles driven by buried polar interactions. *J. Am. Chem. Soc.* 128:14450--51
69. Koder RL, Anderson JLR, Solomon LA, Reddy KS, Moser CC, Dutton PL. 2009. Design and engineering of an O_2 transport protein. *Nature* 458:305--9
70. Shu JY, Tan C, DeGrado WF, Xu T. 2008. New design of helix bundle peptide-polymer conjugates. *Biomacromolecules* 9:2111--17
71. Zhang MJ, Yuan T. 1998. Molecular mechanisms of calmodulin's functional versatility. *Biochem. Cell Biol. Biochim. Biol. Cell.* 76:313--23
72. Tan RY, Mabuchi Y, Grabarek Z. 1996. Blocking the Ca^{2+}-induced conformational transitions in calmodulin with disulfide bonds. *J. Biol. Chem.* 271:7479--83
73. Sui Z, King WJ, Murphy WL. 2007. Dynamic materials based on a protein conformational change. *Adv. Mater.* 19:3377--80
74. Kretsing Rh, Nockolds CE. 1973. Carp muscle calcium-binding protein. 2. Structure determination and general description. *J. Biol. Chem.* 248:3313--26
75. Gifford JL, Walsh MP, Vogel HJ. 2007. Structures and metal-ion-binding properties of the Ca^{2+}-binding helix-loop-helix EF-hand motifs. *Biochem. J.* 405:199--221
76. McPhalen CA, Strynadka NCJ, James MNG. 1991. Calcium-binding sites in proteins---a structural perspective. *Adv. Protein Chem.* 42:77--144
77. Ikura M, Clore GM, Gronenborn AM, Zhu G, Klee CB, Bax A. 1992. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* 256:632--38

78. Davidson AL, Dassa E, Orelle C, Chen J. 2008. Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* 72:317--64

79. Baumann U, Wu S, Flaherty KM, McKay DB. 1993. Three-dimensional structure of the alkaline protease of *Pseudomonas aeruginosa*: a two-domain protein with a calcium binding parallel beta roll motif. *EMBO J.* 12:3357--64

80. Baumann U. 1994. Crystal structure of the 50 kDa metallo protease from *Serratia marcescens*. *J. Mol. Biol.* 242:244--51

81. Bauche C, Chenal A, Knapp O, Bodenreider C, Benz R, et al. 2006. Structural and functional characterization of an essential RTX subdomain of *Bordetella pertussis* adenylate cyclase toxin. *J. Biol. Chem.* 281:16914--26

82. Chenal A, Guijarro JI, Raynal B, Delepierre M, Ladant D. 2009. RTX calcium binding motifs are intrinsically disordered in the absence of calcium: implication for protein secretion. *J. Biol. Chem.* 284:1781--89

83. Ringler P, Schulz GE. 2003. Self-assembly of proteins into designed networks. *Science* 302:106--09

84. Lilie H, Haehnel W, Rudolph R, Baumann U. 2000. Folding of a synthetic parallel beta-roll protein. *FEBS Lett.* 470:173--77

85. Blenner M, Shur O, Szilvay G, Cropek D, Banta S. . 2010. Calcium-induced folding of a repeat in toxin (RTX)-motif via C-terminal entropic stabilization. Submitted

86. Szilvay GR, Blenner MA, Shur O, Cropek DM, Banta S. 2009. A FRET-based method for probing the conformational behavior of an intrinsically disordered repeat domain from *Bordetella pertussis* adenylate cyclase. *Biochemistry* 48:11273--82

87. Trabbic-Carlson K, Setton LA, Chilkoti A. 2003. Swelling and mechanical behaviors of chemically cross-linked hydrogels of elastin-like polypeptides. *Biomacromolecules* 4:572--80

88. Lim DW, Nettles DL, Setton LA, Chilkoti A. 2008. In situ cross-linkinig of elastin-like polypeptide block copolymers for tissue repair. *Biomacromolecules* 9:222--30

89. Martinez-Osorio H, Juarez-Campo M, Diebold Y, Girotti A, Alonso M, et al. 2009. Genetically engineered elastin-like polymer as a substratum to culture cells from the ocular surface. *Curr. Eye Res.* 34:48--56

90. Chilkoti A, Christensen T, MacKay JA. 2006. Stimulus responsive elastin biopolymers: applications in medicine and biotechnology. *Curr. Opin. Chem. Biol.* 10:652--57

91. Nettles DL, Kitaoka K, Hanson NA, Flahiff CM, Mata BA, et al. 2008. In situ crosslinking elastin-like polypeptide gels for application to articular cartilage repair in a goat osteochondral defect model. *Tissue Eng.* (*Part A*)14:1133--40

92. Straley KS, Heilshorn SC. 2009. Independent tuning of multiple biomaterial properties using protein engineering. *Soft Matter* 5:114--24

93. Isenberg BC, Tsuda Y, Williams C, Shimizu T, Yamato M, et al. 2008. A thermoresponsive, microtextured substrate for cell sheet engineering with defined structural organization. *Biomaterials* 29:2565--72

94. Meyer DE, Shin BC, Kong GA, Dewhirst MW, Chilkoti A. 2000. *Drug targeting using thermally responsive polymers and local hyperthermia*. Presented at Int. Symp. Tumor Targeted Delivery Syst. Bethesda, MD

95. Hyun J, Lee WK, Nath N, Chilkoti A, Zauscher S. 2004. Capture and release of proteins on the nanoscale by stimuli-responsive elastin-like polypeptide "switches". *J. Am. Chem. Soc.* 126:7330--35

111

96. Shamji MF, Whitlatch L, Friedman AH, Richardson WJ, Chilkoti A, Setton LA. 2008. An injectable and in situ-gelling biopolymer for sustained drug release following perineural administration. *Spine* 33:748--54

97. Massodi I, Thomas E, Raucher D. 2009. Application of thermally responsive elastin-like polypeptide fused to a lactoferrin-derived peptide for treatment of pancreatic cancer. *Molecules* 14:1999--2015

98. Massodi I, Bidwell GL, Davis A, Tausend A, Credit K, et al. 2009. Inhibition of ovarian cancer cell metastasis by a fusion polypeptide Tat-ELP. *Clin. Exp. Metastasis* 26:251--60

99. Adams SB, Shamji MF, Nettles DL, Hwang P, Setton LA. 2009. Sustained release of antibiotics from injectable and thermally responsive polypeptide depots. *J. Biomed. Mater. Res. Part B. Appl. Biomater.* 90B:67—74

99. Easily injectable in-situ cross-linked ELPs formed antibiotic depots that demonstrated long release times, on the order days.

100. Urry DW. 1999. Elastic molecular machines in metabolism and soft-tissue restoration. *Trends Biotechnol.* 17:249--57

101. Urry DW, Parker TM. 2002. Mechanics of elastin: molecular mechanism of biological elasticity and its relationship to contraction. *J. Muscle Res. Cell Motil.* 23:543--59

102. Urry DW, Hugel T, Seitz M, Gaub HE, Sheiba L, et al. 2002. Elastin: a representative ideal protein elastomer. *Philos. Trans. R. Soc. London Ser. B* 357:169--84

103. Urry DW, Trapane TL, Prasad KU. 1985. Phase-structure transitions of the elastin polypentapeptide-water system within the framework of composition-temperature studies. *Biopolymers* 24:2345--56

104. Chang DK, Venkatachalam CM, Prasad KU, Urry DW. 1989. Nuclear Overhauser effect and computational characterization of the beta-spiral of the polypentapeptide of elastin. *J. Biomol. Struct. Dyn.* 6:851--58

105. Urry DW, Trapane TL, Iqbal M, Venkatachalam CM, Prasad KU. 1985. Carbon^{-13} NMR relaxation studies demonstrate an inverse temperature transition in the elastin polypentapeptide. *Biochemistry* 24:5182--89

106. Schmidt P, Dybal J, Rodriguez-Cabello JC, Reboto V. 2005. Role of water in structural changes of poly(AVGVP) and poly(GVGVP) Studied by FTIR and Raman spectroscopy and ab initio calculations. *Biomacromolecules* 6:697--706

107. Glaves R, Baer M, Schreiner E, Stoll R, Marx D. 2008. Conformational dynamics of minimal elastin-like polypeptides: the role of proline revealed by molecular dynamics and nuclear magnetic resonance. *Chemphyschem* 9:2759--65

108. Li B, Alonso DO, Daggett V. 2001. The molecular basis for the inverse temperature transition of elastin. *J. Mol. Biol.* 305:581--92

109. Urry DW, Gowda DC, Parker TM, Luan CH, Reid MC, et al. 1992. Hydrophobicity scale for proteins based on inverse temperature transitions. *Biopolymers* 32:1243--50

110. Urry DW, Peng SQ, Parker TM. 1992. Hydrophobicity-induced pK shifts in elastin protein-based polymers. *Biopolymers* 32:373--79

111. Urry DW. 1992. Free energy transduction in polypeptides and proteins based on inverse temperature transitions. *Prog. Biophys. Mol. Biol.* 57:23--57

112. Meyer DE, Chilkoti A. 2002. Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules* 3:357--67

113. Urry DW, Luan CH, Parker TM, Gowda DC, Prasad KU, et al. 1991. Temperature of polypeptide inverse temperature transition depends on mean residue hydrophobicity. *J. Am. Chem. Soc.* 113:4346--48

114. Meyer DE, Chilkoti A. 2004. Quantification of the effects of chain length and concentration on the thermal behavior of elastin-like polypeptides. *Biomacromolecules* 5:846--51

115. Blenner MA, Banta S. 2008. Characterization of the 4D5Flu single-chain antibody with a stimulus-responsive elastin-like peptide linker: a potential reporter of peptide linker conformation. *Protein Sci.* 17:527--36

116. Megeed Z, Winters RM, Yarmush ML. 2006. Modulation of single-chain antibody affinity with temperature-responsive elastin-like polypeptide linkers. *Biomacromolecules* 7:999--1004

117. Trabbic-Carlson K, Meyer DE, Liu L, Piervincenzi R, Nath N, et al. 2004. Effect of protein fusion on the transition temperature of an environmentally responsive elastin-like polypeptide: a role for surface hydrophobicity? *Protein Eng. Des. Sel.* 17:57--66

118. Kim B, Chilkoti A. 2008. Allosteric actuation of inverse phase transition of a stimulus-responsive fusion polypeptide by ligand binding. *J. Am. Chem. Soc.* 130:17867—73

118. Demonstrates for the first time, allosteric regulation of ELP inverse phase transitioning.

119. Zhang SG, Marini DM, Hwang W, Santoso S. 2002. Design of nanostructured biological materials through self-assembly of peptides and proteins. *Curr. Opin. Chem. Biol.* 6:865--71

120. Banta S, Megeed Z, Casali M, Rege K, Yarmush ML. 2007. Engineering protein and peptide building blocks for nanotechnology. *J. Nanosci. Nanotechnol.* 7:387--401

121. Chockalingam K, Blenner M, Banta S. 2007. Design and application of stimulus-responsive peptide systems. *Protein Eng. Des. Sel.* 20:155--61

122. Casali M, Banta S, Zambonelli C, Megeed Z, Yarmush ML. 2008. Site-directed mutagenesis of the hinge peptide from the hemagglutinin protein: enhancement of the pH-responsive conformational change. *Protein Eng. Des. Sel.* 21:395--404

Table 1 Engineered characteristics and features of the protein domains in responsive hydrogels

	Function	Reversibility and responsiveness	Engineering features	Applications and associated references
Leucine zipper	Self-assembly Physical crosslinking	pH, temperature, ionic strength	Hetero- and homo-complexes Oligomerization number Oligomerization orientation Thermal stability Polymer conjugation	Reversible hydrogels (80, 88, 118) Tissue engineering scaffolds (28, 29, 71) Drug delivery (89) Enzymatic hydrogels (114, 115)
Calmodulin	Ligand binding Conformational change	Analyte binding Ligand binding	Ligand binding Analyte binding Polymer conjugation	Drug delivery and biosensor (47, 93) Microfluidic gates and microlenses (26, 27)
Elastin-like peptides	Conformational change Phase change	pH Temperature Ionic strength Hydrophobicity	Disorder to b-turn transition Tunable Inverse Phase Transition Guest residue chemistry	Compliant hydrogels (54, 62, 74) Reversible hydrogels (1, 38, 70) Drug delivery (19, 20, 63, 64)

Chapter 8

Campbell, E., Wheeldon, I.R., and **Banta, S.** (2010) "Broadening the cofactor specificity of a thermostable alcohol dehydrogenase using rational protein design introduces novel kinetic transient behavior" *Biotechnology and Bioengineering* **107**(5) 763-774.

BROADENING THE COFACTOR SPECIFICITY OF A THERMOSTABLE ALCOHOL

DEHYDROGENASE USING RATIONAL PROTEIN DESIGN INTRODUCES NOVEL

KINETIC TRANSIENT BEHAVIOR

Elliot Campbell, Ian R. Wheeldon[‡], and Scott Banta*

Department of Chemical Engineering, Columbia University,

500 West 120th Street, New York, New York 10027

‡ Current Address:
Department of Medicine, Center for Biomedical Engineering, Brigham and Women's Hospital,
Harvard Medical School, Boston, MA 02115, USA and Harvard-MIT Division of Health
Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*To whom correspondence should be addressed:
Email: sbanta@columbia.edu
Telephone: (212) 854-7531
Fax: (212) 854-3054

Running title: Introduction of novel kinetic transient behavior in AdhD

ABSTRACT:

Cofactor specificity in the aldo-keto reductase (AKR) superfamily has been well-studied, and several groups have reported the rational alteration of cofactor specificity in these enzymes. Although most efforts have focused on mesostable AKRs, several putative AKRs have recently been identified from hyperthermophiles. The few that have been characterized exhibit a strong preference for NAD(H) as a cofactor, in contrast to the NADP(H) preference of the mesophilic AKRs. Using the design rules elucidated from mesostable AKRs, we introduced two site-directed mutations in the cofactor binding pocket to investigate cofactor specificity in a thermostable AKR, AdhD, which is an alcohol dehydrogenase from *Pyrococcus furiosus*. The resulting double mutant exhibited significantly improved activity and broadened cofactor specificity as compared to the wild-type. Results of previous pre-steady state kinetic experiments suggest that the high affinity of the mesostable AKRs for NADP(H) stems from a conformational change upon cofactor binding which is mediated by interactions between a canonical arginine and the 2'-phosphate of the cofactor. Pre-steady state kinetics with AdhD and the new mutants show a rich conformational behavior that is independent of the canonical arginine or the 2'-phosphate. Additionally, experiments with the highly active double mutant using NADPH as a cofactor demonstrate an unprecedented transient behavior where the binding mechanism appears to be dependent on cofactor concentration. These results suggest that the structural features involved in cofactor specificity in the AKRs are conserved within the superfamily, but the dynamic interactions of the enzyme with cofactors are unexpectedly complex.

Keywords: Cofactor specificity, aldo-keto reductase, pre-steady state kinetics, enzyme catalysis, alcohol dehydrogenase, site-directed mutagenesis, thermostable.

116

Introduction

The aldo-keto reductases (AKRs) are a family of oxidoreductases with a common $(\alpha/\beta)_8$-barrel structure. They are found in almost every living system and catalyze a wide range of redox reactions (Jez et al. 1997; Jez and Penning 2001). Characteristic to this superfamily is a highly conserved cofactor binding pocket that binds a nicotinamide cofactor in the extended conformation without a Rossmann fold motif (Sanli et al. 2003). Most members of the superfamily that have been studied exhibit a strong preference for NADP(H), suggesting a physiological role as reductases (Veech et al. 1969).

Understanding the determinants of cofactor specificity of dehydrogenases has significant importance from an engineering perspective, as the native cofactor specificity of these enzymes is often not ideal for use in synthetic metabolic pathways and other industrial applications. Altering cofactor specificity of an enzyme in an artificial metabolic pathway can potentially correct a redox imbalance in a process or improve overall product yield, and therefore cofactor engineering is important in applications ranging from cofactor regeneration to bioelectrocatalysis (Banta et al. 2002a; Banta et al. 2002c; Chin et al. 2009; Johannes et al. 2005; Khoury et al. 2009; Minteer et al. 2007; Zhang 2010). We are particularly interested in engineering these dehydrogenase enzymes for use in enzymatic biofuel cells, where the choice of the cofactor (acting as the electron mediator between the enzyme and the electrode) is of critical importance (Glykys and Banta 2009; Moore et al. 2004; Moore et al. 2005).

Several groups have used site-directed mutagenesis to study the structural determinants of cofactor specificity in the AKRs (Bohren et al. 1991; Grimshaw et al. 1995; Kratzer et al. 2006; Kubiseski and Flynn 1995; Kubiseski et al. 1992; Leitgeb et al. 2005; Ma et al. 2000; Ratnam et al. 1999; Tarle et al. 1993; Yamaoka et al. 1992) and there have been a few reports of the broadening of the cofactor specificity to increase the activity of these enzymes with NAD(H) (Banta and Anderson 2002; Banta et al. 2002b; Khoury et al. 2009; Liang et al. 2007; Sanli et al. 2004). Through these efforts, several hot spots for mutagenesis have been identified. The first is a lysine residue that appears partially buried under the bound cofactor and interacts with the pyrophosphate backbone, adenine ribose, and 2'-phosphate of NADP(H). This residue has been conservatively mutated in human aldose reductase (Bohren et al. 1991; Yamaoka et al. 1992) and these studies suggest that interactions with the lysine are important for properly orienting the cofactor within the binding pocket and for positioning the nicotinamide head group for hydride transfer. Later, in an effort to improve the activity of an AKR with NADH, a lysine → glycine mutant was identified with improved kinetic properties (Banta et al. 2002b).

A highly conserved arginine residue has also been shown to form important interactions with the adenosine 2'-phosphate. Studies mutating this canonical arginine have demonstrated a significant impact on activity with NADP(H), while changes with NAD(H) were minor (Kubiseski and Flynn 1995; Ratnam et al. 1999). The mechanism of cofactor binding in a model AKR, rat 3α-hydroxysteroid dehydrogenase (3α-HSD), has been extensively studied and demonstrates a multi-step binding mechanism for the NADP(H) cofactor (Ratnam et al. 1999). A comparison of the crystal structures for the apo enzyme and the enzyme-NADPH binary complex suggests a conformational change takes place upon cofactor binding, similar to that observed in other AKRs (Cooper et al. 2007; Sanli and Blaber 2001; Sanli et al. 2003). Using an arginine→methionine mutant, it was demonstrated that the conformational change was due to the formation of a salt bridge between the arginine and 2'-phosphate of NADP(H), which could be observed as a fluorescence kinetic transient. No transient was observed in the arginine→methionine mutant or when NAD(H) was used as a cofactor, suggesting this transient

117

and corresponding conformational change were dependent upon interactions between the arginine and adenosine 2'-phosphate group (Ratnam et al. 1999). The stopped-flow fluorescence data was consistent with a two-step binding mechanism, where an initial rapid bi-molecular association is followed by a slow isomerization to a tightly bound complex. This serves to greatly increase the affinity of the enzyme for the cofactor, and locks the enzyme in a primed state ready to immediately act upon a substrate (Kubiseski et al. 1992). In mutagenesis work performed with *Cornybacterium* 2,5-diketo-D-gluconic acid reductase (2,5-DKGR), it was demonstrated that an arginine→histidine mutant at this position increased activity with NADH while retaining activity with NADP(H) (Banta et al. 2002b). The solved crystal structure of this mutant shows the histidine side-chain forms a π-stacking interaction with the indole ring of the cofactor, and a kinetic analysis demonstrated an improvement in the free energy of cofactor binding, consistent with the introduction of this stabilizing interaction.

Newly available genome sequences from a variety of hyperthermophiles has led to the identification of several putative thermostable AKRs. Although few have been characterized, sequence alignments indicate that these thermostable AKRs contain a histidine residue in the cofactor binding pocket in place of the highly conserved arginine residue found at this position in mesophilic AKRs (Table I). One such AKR, an alcohol dehydrogenase identified from the hyperthermophilic archaeon *Pyrococcus furiosus* (AdhD), exhibits a strong preference towards NAD(H) as a cofactor (Machielsen et al. 2006). The hyperthermophile sequence data and experimental evidence of the preference of AdhD towards NAD(H) combined with the arginine to histidine mutation identified in NADP(H)-biased AKRs seems to suggest that hyperthermophilic AKRs may preferentially utilize NAD(H).

In the present work, we have rationally mutated the cofactor binding pocket in the thermostable AKR, AdhD, from *Pyrococcus furiosus* guided by the design rules elucidated in the mesostable AKRs. A K249G/H255R double mutant exhibited the greatest improvement in activity with NADP(H), and also had superior activity with NAD(H) compared to the wild-type and the other enzyme variants tested. All enzyme forms also exhibited varying degrees of kinetic transients upon cofactor binding, in contrast to the previous results obtained with a mesostable AKR (Ratnam et al. 1999). Most interestingly, the highly active double mutant exhibited bi-exponential kinetic transients with NADPH where the direction of the fast transient was concentration dependent. Taken together, these results suggest that amino acids identified in the mesostable AKRs can be used to modify the cofactor specificity of AdhD, and the observed kinetic transients are independent of the formation of a guanido-phosphate salt bridge.

Materials and Methods
Chemicals and plasmids: Oligonucleotides were from Integrated DNA Technologies. The QuikChange Site Directed Mutagenesis kit was from Stratagene. Isopropyl-β-D-1-thiogalactopyranoside (IPTG) was from Promega. *E. coli* BL21(DE3) competent cells were from New England Biolabs. Precast sodium dodecyl sulfate-polyacrylamide gels, NuPAGE MOPS running buffer, and broad-range molecular weight marker were from Invitrogen. All other chemicals were from Sigma-Aldrich and used without modification. The *Pyrococcus furiosus* AdhD expression plasmid pWUR85 and tRNA helper plasmid pSJS1244 were a kind gift from Dr. John van der Oost (Wageningen University, The Netherlands) and are described in (Machielsen et al. 2006).

118

Mutant Construction: Single mutants K249G and H255R and double mutant K249G/H255R were created using the QuikChange Site-directed Mutagenesis Kit (see SI). All mutations were verified by DNA sequencing.

AdhD Expression and Purification: Expression and purification of AdhD followed a previously established protocol with minor modifications (Machielsen et al. 2006). After expression, cells were harvested by centrifugation, and pellets were resuspended in 1/10th volume 20 mM Tris-HCl (pH 7.5) before being incubated at 80°C for 1 h. Endogenous proteins and cell debris were then removed by centrifugation for 20 min at 10,000 x g. The supernatant was retained as the heat-stable cell-free extract (HSCFE). Samples were concentrated over a centrifugal filter (30 kDa MWCO) before being applied to a gel filtration column (Superdex 16/200, GE Healthcare) equilibrated in 20 mM Tris HCl (pH 7.8), 100 mM NaCl. Fractions containing active enzyme were pooled and concentrated. Enzyme stocks were diluted to working concentration in 20 mM Tris HCl (pH 7.8) before use. Expression and purification of AdhD mutants followed the same protocol. All enzyme concentrations were determined from A_{280} measurements with a calculated molar extinction coefficient of $\varepsilon_{280} = 52495$ M^{-1} cm^{-1}.

SDS-PAGE: Protein composition was analyzed using NuPAGE 4-12% Bis-Tris Gels with a Novex Mini-Cell system. Samples were prepared as described previously (Machielsen et al. 2006). A broad-range protein marker was used for molecular weight estimation.

Homology Modeling: A homology model of AdhD was generated using ESyPred3D (Lambert et al. 2002) and MODELLER with primary template prostaglandin F synthase from *Trypanosoma brucei* (1VBJ, 31.1% identities). Structures were analyzed using MolProbity (Davis et al. 2007) and verified against other members of the aldo-keto reductase superfamily. Cofactors were inserted into the binding pocket by aligning the backbone of the homology model with 2,5-DKGR from *Corynebacterium* (1A80 with bound NADPH, 1M9H with bound NADH) (Sanli et al. 2004). Figures were generated using YASARA.

Activity Assays: The activity of each mutant was first examined at fixed substrate concentrations above the previously reported K_{MS} for AdhD. Reaction mixtures containing 50 mM glycine (pH 8.8), 100 mM 2,3-butanediol (oxidation reaction) or 100 mM sodium phosphate (pH 6.1), 80 mM 3-hydroxy-2-butanone (reduction reaction) and enzyme were incubated in a 96-well UV-transparent microplate at 45°C in a SpectraMax M2 plate reader (Molecular Devices). Reactions were initiated by the addition of 1-1000 µM cofactor. Initial rates were determined by following the production or depletion of NAD(P)H at 340 nm ($\varepsilon = 6.22$ mM^{-1} cm^{-1}). Data were collected in triplicate, and experiments were repeated three times with fresh solutions. All points were fit simultaneously to (Eq. 1) using non-linear least-squares regression (Igor Pro, Wavemetrics, Inc.) to obtain estimates for the apparent k_{cat} and Michaelis constant for each cofactor (Segel 1993). Reported errors are standard deviations. Statistically significant differences from wild-type AdhD were determined by Student's t-test.

$$V = \frac{E_t k_{cat}^{app} A}{K_A^{app} + A} \qquad \text{Eq. 1}$$

Fluorescence Titrations: Dissociation constants for the enzyme-cofactor complexes were determined by fluorescence titration (Jackman et al. 1992; Ratnam et al. 1999; Stone and Le Bonniec 1997). Briefly, 2 µM enzyme in 50mM glycine (pH 8.8) (for NAD(P)$^+$) or 10 mM potassium phosphate (pH 7.0) (for NAD(P)H) was stirred in a 1 cm quartz cuvette placed in a J-

815 spectrometer (Jasco Inc.) equipped with a Peltier junction temperature control. Samples were excited at 280 nm, and the fluorescence change upon cofactor binding was monitored at 330 nm $(NAD(P)^+)$ or 450 nm (NAD(P)H). The total volume of cofactor added was less than 1% of the reaction volume to limit dilution effects. Experiments were repeated in at least triplicate, and data were fit to a saturation adsorption isotherm.

Steady-state Kinetics: The full kinetic parameters for the wild-type and double mutant AdhD were determined for both the oxidation and reduction reactions with NAD(P)(H). Initial rates at 45°C were measured using a SpectraMax M2 plate reader by following the production or depletion of NAD(P)H at 340 nm ($\varepsilon = 6.22$ mM^{-1} cm^{-1}). Oxidation reactions contained 50 mM glycine (pH 8.8), 1-100 mM 2,3-butanediol, and the appropriate amount of enzyme and were initiated with 1-1000 µM NAD(P)$^+$. Reduction reactions contained 100 mM sodium phosphate (pH 6.1), 1-100 mM 3-hydroxy-2-butanone, and enzyme, and were initiated by the addition of 1-500 µM NAD(P)H. Some cofactor inhibition was observed at concentrations in excess of 1 mM (data not shown). Reactions were initiated with cofactor to limit cofactor degradation during incubation at elevated temperatures, however control experiments indicated that cofactor degradation was not significant over the time the reaction was monitored. Data were collected in at least triplicate, and were fit simultaneously to the ordered bi-bi rate equation ((Segel 1993), Eq. 2) using non-linear least-squares regression. This reaction mechanism was previously verified (Wheeldon et al. 2009). Reported errors are standard deviations.

$$V = \frac{E_t k_{cat} AB}{K_{ia}K_B + K_A B + K_B A + AB} \qquad \text{Eq. 2}$$

Determination of Protein Stability: Unfolding was assessed by following the CD signal at 222 nm in a J-815 CD Spectrophotometer equipped with a Peltier junction temperature control. Scans were made over a range of guanidine hydrochloride (GdnHCl) concentrations with a 1°C min^{-1} temperature ramp from 25°C to 90°C. Prior to analysis, enzyme samples were allowed to equilibrate overnight at room temperature in the appropriate concentration of GdnHCl. The midpoint of a sigmoidal fit to the data at 80 °C was taken as the denaturation midpoint.

Kinetics of Cofactor Binding: The kinetics of cofactor binding were investigated using a SFM-20 stopped flow system (BioLogic Inc.) equipped with a 20 µl fluorescence cuvette (dead time ≈ 13 ms) attached to a J-815 CD Spectrophotometer. All experiments were performed at 25°C. Samples of enzyme (0.75 µM) and cofactor (0.5 – 30 µM) were mixed, and the quenching of intrinsic protein fluorescence (for NAD(P)$^+$, 320 nm cutoff) or the energy transfer between the protein and cofactor (for NAD(P)H, 430 nm cutoff) was monitored upon exciting at 280 nm. Each fluorescence trace is the average of 3-5 shots, and each experiment was repeated three times with fresh solutions. Traces were fit to a mono-exponential or bi-exponential function where applicable, and the resulting rate constants were plotted versus cofactor concentration. These plots were used to obtain estimates of the rate constants for cofactor binding as described previously (Fierke and Hammes 1995). All concentrations are given as the final concentration in the cuvette.

Results

Expression and Purification of AdhD: Site-directed mutagenesis was used to create three new mutant AdhD enzymes, K249G, H255R, and K249G/H255R. The wild-type and new mutant AdhDs expressed in high yields in *E. coli* and were readily purified due to their extreme thermostabilities. A simple and rapid purification scheme consisting of heating the re-suspended cell pellets to both lyse the cells and denature endogenous proteins, followed by centrifugation,

concentration, and size exclusion chromatography yielded homogenous samples as judged by SDS-PAGE (SFigure 1).

Homology Modeling: Previous work and structural insights guided the decision to create the two single mutants and the double mutant of the thermostable AKR AdhD. In order to visualize the potential impact of these mutations on the cofactor binding pocket, a homology model was created. The highly conserved structure of the AKR superfamily enabled the addition of cofactors into the homology model by alignment with crystal structures of a similar AKR (2,5-DKGR) containing bound cofactor. The structural alignment had a RMSD of 1.0Å over 232 aligned residues, and allowed us to identify amino acids that could potentially interact with the cofactor (Figure 1). It seems likely that His255 is in position to form a stacking interaction with the adenine ring of the cofactor, and potentially an ionic interaction with the 2'-phosphate of NADP(H) as well. Additionally, replacement of Lys249 with glycine should increase the volume of the cofactor binding pocket and allow for increased conformational flexibility.

Fluorescence Titrations: The new mutations were made to impact cofactor binding and thereby cofactor specificity. Conveniently located tryptophan residues in the cofactor binding pocket allow for the determination of cofactor dissociation constants for the different mutants by fluorescence titration (STable I). Comparison of dissociation constants between the wild-type enzyme and mutants can then be used to calculate the changes in ground state cofactor binding energies. At 25°C, the H255R mutant lost 0.2 kcal/mol of binding energy with NAD^+ but gained 2.6 kcal/mol with $NADP^+$. The change in binding energy with $NADP^+$ is less than was observed in an arginine$\rightarrow$methionine mutant of 3α-HSD (Ratnam et al. 1999), but is comparable in magnitude to the gain of an electrostatic interaction. The K249G single mutant gained 2.2 and 2.4 kcal/mol of binding energy with NAD^+ and $NADP^+$, respectively, while the K249G/H255R double mutant exhibited a slight gain of 0.5 kcal/mol for NAD^+ and a larger gain of 2.1 kcal/mol with $NADP^+$. Small gains in binding energies were observed with the reduced cofactors in every case, ranging from 0.14 kcal/mol with NADPH for the K249G mutant to 0.57 kcal/mol with NADPH for the H255R mutant. The double mutant gained 0.31 kcal/mol with NADH and 0.25 kcal/mol with NADPH.

Fluorescence titrations were also performed at 45°C with wild-type AdhD and the K249G/H255R mutant to allow for comparison with the K_{ia} term of the ordered bi-bi rate equation obtained from steady-state kinetic experiments as described below (Table II).
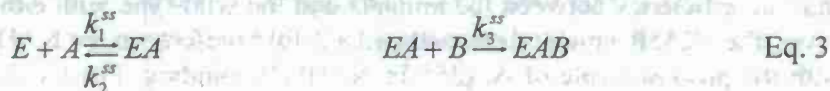
Steady-State Kinetic Analysis: A simplified kinetic analysis of the oxidation and reduction reactions for the wild-type enzyme and the three mutants was performed at a fixed substrate concentration (STable II) In order to estimate the effect of the mutations on cofactor specificity, the apparent catalytic efficiency (k_{cat}/K_A) was compared (Figure 2). In the oxidation reaction, the wild-type enzyme exhibits a similar preference for both NAD^+ and $NADP^+$. No significant difference in catalytic efficiency between the mutants and the wild-type with either cofactor was observed. However, the H255R mutant demonstrated a 2-fold preference for $NADP^+$ over NAD^+, which agrees with the proposed role of Arg255 in NADP(H) binding. For the double mutant in the two oxidation reactions, the Michaelis constants for the substrate (K_B) were later calculated to be larger than the substrate concentration utilized, and therefore the assumption of saturating substrate in these cases is invalid which would lead to an underestimation of the k_{cat}^{app}.

In the reduction direction, larger changes in the apparent kinetic parameters were observed. The wild-type enzyme exhibited significant specificity for NADH over NADPH as evidenced by an order of magnitude difference in the catalytic efficiency. For the H255R mutant, the catalytic efficiency doubled with NADH as compared to the wild-type and it increased by an order of

magnitude with NADPH so that the mutant had no significant specificity between the cofactors. For the K249G mutant, the catalytic efficiency increased 5-fold for NADH while the efficiency with NADPH increased more than 30-fold in comparison to the wild-type. For the double mutant, the catalytic efficiency with NADH increased 4-fold while the efficiency with NADPH increased more than 16-fold.

Since the double mutant exhibited the largest increase in the apparent k_{cat} with both cofactors, the full steady-state kinetic experiments were performed for this mutant and compared to the values for the wild-type enzyme (Table II, SFigure 2). In the oxidation reaction, the k_{cat} with NAD^+ improved by 15-fold for the double mutant over the wild-type enzyme. The impact on activity with $NADP^+$ was even greater, as the double mutant had a k_{cat} nearly two orders of magnitude larger than the wild-type. However, the Michaelis constant for the cofactor also increased significantly in both cases, from 63 μM to 460 μM for NAD^+ and from 5.1 μM to 78 μM for $NADP^+$ in the wild-type and double mutant, respectively. The Michaelis constant for the substrate also increased significantly, from 29 mM for the wild-type to 690 mM for the double mutant with NAD^+, and from 1.3 mM to 200 mM with $NADP^+$. In the reduction reaction, the double mutant has a k_{cat} 3-fold greater with NADH and 6-fold greater with NADPH compared to the wild-type. While the Michaelis constants for the cofactor and substrate increased for the double mutant in the oxidation reaction, they mostly decreased in the reduction reaction. For the cofactor, K_A with NADH decreased from 190 μM to 50 μM, and with NADPH the value decreased from 280 μM to 33 μM. The K_B value increased from 0.9 mM to 13 mM when NADH was the cofactor, but decreased from 6.7 mM to 5.0 mM when NADPH was the cofactor. In the case of the reduction reaction with NADPH, the K_{ia} values were unable to be fit by the model and so the K_D values obtained by fluorescence titration at 45°C were used instead.

In order to simplify the comparison of the impact of the mutations on the steady-state kinetics, the parameters were used to estimate the microscopic rate constants for the simplified reaction mechanism described in Equation 3 (Table III). Generally, the on-rate of the cofactor (k_1^{ss}) increased by 2 to 10-fold in the double mutant compared to the wild-type. In most cases the off-rate (k_2^{ss}) was found to decrease, except in the case of the double mutant with NADPH where the off-rate increased. The net on-rate of the substrate (k_3^{ss}) in the oxidation reaction was only slightly impacted by the mutations, whereas a much stronger effect was observed in the reduction reaction. The on-rate of the substrate with NADH decreased almost 5-fold in the double mutant versus the wild-type, but increased 8-fold with NADPH. The ratio of $k_1^{ss}k_3^{ss}/k_2^{ss}$ is a convenient single parameter for examining the catalytic performance of the mutants (Banta et al. 2002c). When judged by this composite rate constant, the double mutant enzyme is shown to be substantially improved with $NADP^+$ in the oxidative direction and with both NADH and NADPH in the reductive reaction compared to the wild-type (Figure 3).

$$E + A \underset{k_2^{ss}}{\overset{k_1^{ss}}{\rightleftharpoons}} EA \qquad\qquad EA + B \overset{k_3^{ss}}{\longrightarrow} EAB \qquad\qquad \text{Eq. 3}$$

Determination of Enzyme Stability: Guanidine denaturation curves were generated for wild-type AdhD and the K249G/H255R double mutant in order to assess the effect of the mutations on enzyme stability. Both enzymes appeared stable in up to 6M GdnHCl at room temperature, and temperatures greater than 70°C were required to observe an unfolding transition. At 80°C, the denaturation midpoints of both enzymes were comparable (4.8M for the wild-type, K249G and H255R mutants and 4.9M for the K249G/H255R mutant), suggesting that the mutations had little effect on stability (SFigure 3). The unfolding did not appear to be reversible, however, as

little CD signal was regained upon cooling. Thus this data could not be used to calculate ΔG values.

Kinetics of Cofactor Binding: Stopped-flow fluorescence spectroscopy was used to further investigate the mode of cofactor binding in the wild-type and mutant enzymes (Grimshaw et al. 1995; Jackman et al. 1992; Tew and Bottomley 2001). AdhD contains six tryptophan residues, two of which are located near the active site. These residues act as distal reporters of cofactor binding, as the quenching of intrinsic protein fluorescence or energy transfer with the reduced cofactor can be followed (Figure 4). The signal voltage is inversely proportional to the fluorescence intensity, such that negative amplitude corresponds to an increase in fluorescence and vice versa. In this study, all four enzymes tested displayed observable fluorescence transients with NADP(H), and all except H255R displayed observable transients with NAD(H) (SFigure 4). In some cases, transients may have occurred mostly within the dead-time of the stopped-flow (such as H255R with NADH), and in these cases no rate data were obtained. The existence of transients in these cases was confirmed by control stopped-flow experiments diluting the enzyme into buffer and by comparison with the steady-state fluorescence titration data. Plots of the observed rate constant versus cofactor concentration displayed saturation kinetics and were well fit by a hyperbolic function (Table IV, SFigure 5). These are consistent with a two-step binding mechanism (Eq. 4), in which a rapid bimolecular association step is followed by a slow isomerization step (Fierke and Hammes 1995; Jackman et al. 1992). Note that the cofactor binding constants k_1^{ss} and k_2^{ss} obtained from analysis of the steady-state kinetics data include the isomerization step and thus are different than k_1 and k_2 obtained from the transient kinetic data.

$$E + A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E^* \bullet A \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E \bullet A \qquad \text{Eq. 4}$$
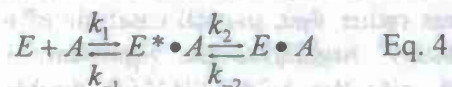
Unprecedented Transients with the Double Mutant and NADPH: Fluorescence traces of NADPH binding with the double mutant exhibited cofactor concentration dependent amplitudes and were best fit with a bi-exponential function. Three regimes were identified based on the amplitude and direction of the fast transient, as the rate of the slow transient remained relatively constant (Figure 5). At low NADPH concentrations (< 2 μM), the fast transient had a negative amplitude, consistent with an increase in FRET efficiency as the nicotinamide head group binds near the active site. At slightly higher concentrations, the amplitude of this fast transient was too small to be reliably fit with a rate constant. Above 5 μM NADPH the amplitude of the fast transient became positive, indicating an initial rapid decrease in fluorescence. Interestingly, a hyperbolic fit to the rates of the fast transient in the first regime extrapolates roughly to the measured rates of the fast transients in the third regime (Figure 5).

Discussion

Knowledge of the cofactor binding mechanism and determinants of cofactor specificity obtained with mesostable AKRs allowed us to readily broaden the cofactor specificity in a thermostable AKR, AdhD. As an arginine→histidine mutation has been previously demonstrated to increase activity with NAD(H), we reasoned the reverse would hold and that a histidine→arginine mutation would increase activity with NADP(H). Additionally, a lysine→glycine mutation was investigated as it was previously found to improve overall activity (Banta et al. 2002b; Banta et al. 2002c). Combining both mutations in 2,5-DKGR yielded a double mutant with significantly improved kinetic properties (Banta et al. 2002c).

In the present work with AdhD, the H255R single mutant exhibited an increased binding affinity toward $NADP^+$ and a concomitant reduction in affinity for NAD^+. A similar trend was

123

observed using a simplified kinetic analysis, as the apparent k_{cat} for H255R was only about 60% of that of the wild-type with NAD^+, but was six-fold higher than the wild-type with $NADP^+$. These results support the idea that an arginine at position 255 is important for recognizing NADP(H), but is not the sole determinant of cofactor specificity.

Kinetics with the K249G single mutant demonstrated a significant increase in k_{cat} compared to both the wild-type and the H255R single mutant with $NAD(P)^+$ and NADH. Previous kinetic and structural studies suggest this residue is important in properly orienting the cofactor in the active site (Banta et al. 2002b; Banta et al. 2002c; Bohren et al. 1991; Yamaoka et al. 1992). As the natural substrates for 2,5-DKGR and AdhD are not known, it is possible this mutation better positions the cofactor for turnover with the non-natural substrates and would impair wild-type functionality. These mutations do not seem to have an additive effect on cofactor binding energy in AdhD, however, as the double mutant only exhibits a slight increase in affinity for NAD^+ and a moderate increase in affinity for $NADP^+$, which is less than would be expected given the changes in binding affinities observed in the single mutants. Regardless, the K249G/H255R double mutant was significantly more active than the wild-type and single mutant enzymes with both NAD(H) and NADP(H) both at moderate (STable 2) and high temperatures (STable 3), and these mutations slightly improved the thermostability of the enzyme (SFigure 3). This impressive result confirmed the design rules established for relaxing cofactor specificity in AKRs and prompted us to further investigate the basis for this change and whether mesostable and thermostable AKRs share a conserved cofactor binding mechanism.

Given the increase in k_{cat} observed in the K249G/H255R double mutant, it is useful to compare individual rate constants rather than overall catalytic efficiencies (Table III, Figure 3). The composite parameter $(k_1k_3/k_2)^{ss}$ highlights the significant improvement in activity and broadened specificity observed with the K249G/H255R double mutant over the wild-type enzyme. Further, the double mutant exhibits an order of magnitude improvement in this parameter when $NADP^+$ is used as the cofactor in place of NAD^+. This is largely a result of a decrease in the off-rate of $NADP^+$, likely due to anchoring of the 2'-phosphate by Arg255 as has been previously proposed.

Stopped-flow fluorescence spectroscopy has been used to probe the difference in binding mechanism between NADP(H) and NAD(H) in rat liver 3α-HSD and suggests the canonical arginine residue forms an electrostatic linkage with the 2'-phosphate of NADP(H), which is observed as a fluorescence kinetic transient (Ratnam et al. 1999). The interaction is accompanied by a conformational change in the cofactor binding pocket which increases the affinity of the enzyme for the cofactor. A kinetic transient was not observed in an arginine→methionine mutant or when NAD(H) was used as a cofactor, suggesting that the transient (and associated conformational change) was both arginine and 2'-phosphate dependent.

Introduction of the canonical arginine residue into AdhD allowed us to examine whether the cofactor binding mechanism established in the 3α-HSD enzyme applies to the thermostable AdhD. Unexpectedly, the reported arginine and 2'-phosphate dependent fluorescence transient observed upon cofactor binding in the mesostable AKRs does not seem to hold for AdhD. In the present work, we demonstrate that a kinetic transient exists in wild-type AdhD which contains a histidine at this position, and also when NAD(H) is used as a cofactor. Similar behavior was observed in the three cofactor binding pocket mutants used to further investigate cofactor specificity, suggesting this conformational behavior is less sensitive to the presence of the arginine residue and 2'-phosphate of NADP(H) than previously suggested. Additionally, when protonated, the histidine mutation is relatively conservative compared to the previously described

124

methionine mutants. This could explain the existence of fluorescence kinetic transients with both the wild-type and H255R enzymes upon NADP(H) binding, as the histidine may be able to form an electrostatic linkage with the negatively charged 2'-phosphate of the cofactor similarly to the canonical arginine. Further exploration of these mutations in mesophilic AKRs, especially the transient behavior of His255 and Gly249 mutants, would lead to a better understanding of these differences. Also, it will be interesting to see whether other thermostable AKRs demonstrate a similar cofactor binding mechanism.

Unprecedented transient behavior was observed in the K249G/H255R double mutant when NADPH was used as a cofactor. Fluorescence traces appeared to be at least bi-exponential, and the amplitude of the fast transient was surprisingly dependent on the concentration of the cofactor. At low NADPH concentrations (< 2 µM), a fast increase in fluorescence intensity was followed by a slow decay to the steady-state value. As the fluorescence signal is due to energy transfer between the enzyme and cofactor, this suggests an initial rapid binding step that brings the nicotinamide head group close to the active site, followed by a slow isomerization moving the head group away to an equilibrium position. As the cofactor concentration increased, the amplitude of the fast transient decreased to the point where the signal was dominated by the slow transient. Above 5 µM NADPH, a fast transient was again observed, but with an amplitude opposite of that at lower cofactor concentrations. A plot of the fast transient observed in regime I versus NADPH concentration is best fit by a hyperbola, suggesting at least a three step reaction mechanism (Fierke and Hammes 1995; Stone and Le Bonniec 1997). To the best of our knowledge, this behavior has not been previously reported in the literature.

The dynamics of NADPH binding suggest the cofactor samples several configurations before reaching an equilibrium position. The increased volume of the cofactor binding pocket afforded by the lysine→glycine mutation seems likely to contribute to the increased conformational flexibility of the cofactor. Multiple cofactor molecules competing for the same binding site could also explain the inverse amplitude observed above 5 µM NADPH, but this does not agree with the steady-state kinetics where cofactor inhibition was only observed at cofactor concentrations several orders of magnitude higher (data not shown).

Estimates of cofactor dissociation constants were obtained through three orthogonal methods: fluorescence titrations, steady-state kinetics, and transient-state kinetics. Fits to the ordered bi-bi rate equation (Eq. 2) were used to determine the full steady-state kinetic parameters for the wild-type and K249G/H255R double mutant. The fit parameter K_{ia} is equivalent to the dissociation constant of the enzyme-cofactor complex (Segel 1993), and was generally in good agreement with the dissociation constant as measured by fluorescence titrations (Table II). The transient-state kinetics investigated by stopped flow fluorescence spectroscopy can also be used to calculate the microscopic rate constants corresponding to each step in the cofactor binding mechanism (Fierke and Hammes 1995). The overall dissociation constant can then be calculated from the microscopic rate constants, and compared to that obtained by fluorescence titrations (Jackman et al. 1992). These results are summarized in Table IV. Almost universally, the dissociation constants calculated from the microscopic rate constants significantly underestimate those obtained by fluorescence titration and steady-state kinetics. Although the source of this disparity is unknown, some difficulty in reconciling stopped-flow fluorescence data with that measured at steady-state has been reported by others (Cooper et al. 2007; Jin and Penning 2006; Ma et al. 2000; Ratnam et al. 1999). Control experiments were performed to rule out artifacts due to mixing effects, non-specific binding, or photobleaching with the fluorescence methods, and the introduction of additional steps in the cofactor binding mechanism could only further

decrease the calculated dissociation constants. Further experiments using T-jump spectroscopy or ITC may be necessary to reconcile these observations and fully elucidate the cofactor binding mechanism.

Broadening cofactor specificity in the AKR superfamily has become almost formulaic, although the mechanism of cofactor binding does not yet seem to be fully elucidated. The ability to change or relax the cofactor specificity of AKRs will be useful in industrial applications, as NAD(H) is more stable and less expensive than NADP(H) (Banta et al. 2002a), and the use of AKRs in specialized applications will benefit from knowledge obtained during cofactor specificity engineering exercises, as it may be advantageous to increase activity with non-natural cofactors that are optimized for the final application (Ryan et al. 2008).

References

Banta S, Anderson S. 2002. Verification of a Novel NADH-Binding Motif: Combinatorial Mutagenesis of Three Amino Acids in the Cofactor-Binding Pocket of Corynebacterium 2,5-Diketo-D-Gluconic Acid Reductase. Journal of Molecular Evolution 55(6):623-631.

Banta S, Boston M, Jarnagin A, Anderson S. 2002a. Mathematical Modeling of in vitro Enzymatic Production of 2-Keto-L-gulonic Acid Using NAD(H) or NADP(H) as Cofactors. Metabolic Engineering 4(4):273-284.

Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S. 2002b. Alteration of the specificity of the cofactor-binding pocket of Corynebacterium 2,5-diketo-D-gluconic acid reductase A. Protein Eng. 15(2):131-140.

Banta S, Swanson BA, Wu S, Jarnagin A, Anderson S. 2002c. Optimizing an Artificial Metabolic Pathway: Engineering the Cofactor Specificity of Corynebacterium 2,5-Diketo-D-gluconic Acid Reductase for Use in Vitamin C Biosynthesis. Biochemistry 41(20):6226-6236.

Bohren K, Page J, Shankar R, Henry S, Gabbay K. 1991. Expression of human aldose and aldehyde reductases. Site-directed mutagenesis of a critical lysine 262. J. Biol. Chem. 266(35):24031-24037.

Chin JW, Khankal R, Monroe CA, Maranas CD, Cirino PC. 2009. Analysis of NADPH supply during xylitol production by engineered *Escherichia coli*. Biotechnology and Bioengineering 102(1):209-220.

Cooper WC, Jin Y, Penning TM. 2007. Elucidation of a Complete Kinetic Mechanism for a Mammalian Hydroxysteroid Dehydrogenase (HSD) and Identification of All Enzyme Forms on the Reaction Coordinate: The Example of Rat Liver 3α-HSD (AKR1C9). J. Biol. Chem. 282(46):33484-33493.

Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB, Snoeyink J, Richardson JS and others. 2007. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res 35(Web Server issue):W375-83.

Fierke CA, Hammes GG. 1995. Transient kinetic approaches to enzyme mechanisms. Methods Enzymol 249:3-37.

Glykys DJ, Banta S. 2009. Metabolic control analysis of an enzymatic biofuel cell. Biotechnol Bioeng 102(6):1624-35.

Grimshaw CE, Bohren KM, Lai CJ, Gabbay KH. 1995. Human aldose reductase: subtle effects revealed by rapid kinetic studies of the C298A mutant enzyme. Biochemistry 34(44):14366-73.

Jackman M, Parry M, Hofsteenge J, Stone S. 1992. Intrinsic fluorescence changes and rapid kinetics of the reaction of thrombin with hirudin. J. Biol. Chem. 267(22):15375-15383.

Jez JM, Bennett MJ, Schlegel BP, Lewis M, Penning TM. 1997. Comparative anatomy of the aldo-keto reductase superfamily. Biochemical Journal 326(3):625-636.

Jez JM, Penning TM. 2001. The aldo-keto reductase (AKR) superfamily: an update. Chemico-Biological Interactions 130-132:499-525.

Jin Y, Penning TM. 2006. Multiple Steps Determine the Overall Rate of the Reduction of 5a-Dihydrotestosterone Catalyzed by Human Type 3 3a-Hydroxysteroid Dehydrogenase: Implications for the Elimination of Androgens. Biochemistry 45(43):13054-13063.

Johannes TW, Woodyer RD, Zhao H. 2005. Directed Evolution of a Thermostable Phosphite Dehydrogenase for NAD(P)H Regeneration. Appl. Environ. Microbiol. 71(10):5728-5734.

Khoury GA, Fazelinia H, Chin JW, Pantazes RJ, Cirino PC, Maranas CD. 2009. Computational design of *Candida boidinii* xylose reductase for altered cofactor specificity. Protein Science 18(10):2125-2138.

Kratzer R, Wilson DK, Nidetzky B. 2006. Catalytic mechanism and substrate selectivity of aldo-keto reductases: Insights from structure-function studies of Candida tenuis xylose reductase. IUBMB Life 58(9):499 - 507.

Kubiseski TJ, Flynn TG. 1995. Studies on Human Aldose Reductase. J. Biol. Chem. 270(28):16911-16917.

Kubiseski TJ, Hyndman DJ, Morjana NA, Flynn TG. 1992. Studies on pig muscle aldose reductase. Kinetic mechanism and evidence for a slow conformational change upon coenzyme binding. J. Biol. Chem. 267(10):6510-6517.

Lambert C, Leonard N, De Bolle X, Depiereux E. 2002. ESyPred3D: Prediction of proteins 3D structures. Bioinformatics 18(9):1250-1256.

Leitgeb S, Petschacher B, Wilson DK, Nidetzky B. 2005. Fine tuning of coenzyme specificity in family 2 aldo-keto reductases revealed by crystal structures of the Lys-274 -> Arg mutant of Candida tenuis xylose reductase (AKR2B5) bound to NAD^+ and $NADP^+$. FEBS Letters 579(3):763-767.

Liang L, Zhang J, Lin Z. 2007. Altering coenzyme specificity of Pichia stipitis xylose reductase by the semi-rational approach CASTing. Microbial Cell Factories 6(1):36.

Ma H, Ratnam K, Penning TM. 2000. Mutation of Nicotinamide Pocket Residues in Rat Liver 3α-Hydroxysteroid Dehydrogenase Reveals Different Modes of Cofactor Binding. Biochemistry 39(1):102-109.

Machielsen R, Uria AR, Kengen SWM, van der Oost J. 2006. Production and Characterization of a Thermostable Alcohol Dehydrogenase That Belongs to the Aldo-Keto Reductase Superfamily. Appl. Environ. Microbiol. 72(1):233-238.

Minteer SD, Liaw BY, Cooney MJ. 2007. Enzyme-based biofuel cells. Current Opinion in Biotechnology 18(3):228-234.

Moore CM, Akers NL, Hill AD, Johnson ZC, Minteer SD. 2004. Improving the Environment for Immobilized Dehydrogenase Enzymes by Modifying Nafion with Tetraalkylammonium Bromides. Biomacromolecules 5(4):1241-1247.

Moore CM, Minteer SD, Martin RS. 2005. Microchip-based ethanol/oxygen biofuel cell. Lab on a Chip 5(2):218-225.

Ratnam K, Ma H, Penning TM. 1999. The Arginine 276 Anchor for NADP(H) Dictates Fluorescence Kinetic Transients in 3α-Hydroxysteroid Dehydrogenase, a Representative Aldo-Keto Reductase. Biochemistry 38(24):7856-7864.

Ryan JD, Fish RH, Clark DS. 2008. Engineering Cytochrome P450 Enzymes for Improved Activity towards Biomimetic 1,4-NADH Cofactors. ChemBioChem 9(16):2579-2582.

Sanli G, Banta S, Anderson S, Blaber M. 2004. Structural alteration of cofactor specificity in Corynebacterium 2,5-diketo-D-gluconic acid reductase. Protein Sci 13(2):504-512.

Sanli G, Blaber M. 2001. Structural assembly of the active site in an aldo-keto reductase by NADPH cofactor. Journal of Molecular Biology 309(5):1209-1218.

Sanli G, Dudley JI, Blaber M. 2003. Structural Biology of the Aldo-Keto Reductase Family of Enzymes: Catalysis and Cofactor Binding. Cell Biochemistry and Biophysics 38:79-101.

Segel IH. 1993. Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems. New York: Wiley. xxii, 957 p.

Stone SR, Le Bonniec BF. 1997. Inhibitory mechanism of serpins. Identification of steps involving the active-site serine residue of the protease. Journal of Molecular Biology 265(3):344-362.

Tarle I, Borhani DW, Wilson DK, Quiocho FA, Petrash JM. 1993. Probing the active site of human aldose reductase. Site-directed mutagenesis of Asp-43, Tyr-48, Lys-77, and His-110. J. Biol. Chem. 268(34):25687-25693.

Tew DJ, Bottomley SP. 2001. Intrinsic fluorescence changes and rapid kinetics of proteinase deformation during serpin inhibition. FEBS Lett 494(1-2):30-3.

Veech RL, Eggleston LV, Krebs HA. 1969. The redox state of free nicotinamide-adenine dinucleotide phosphate in the cytoplasm of rat liver. Biochem J 115(4):609-19.

Wheeldon IR, Campbell E, Banta S. 2009. A Chimeric Fusion Protein Engineered with Disparate Functionalities--Enzymatic Activity and Self-assembly. Journal of Molecular Biology 392(1):129-142.

Yamaoka T, Matsuura Y, Yamashita K, Tanimoto T, Nishimura C. 1992. Site-directed mutagenesis of His-42, His-188 and Lys-263 of human aldose reductase. Biochemical and Biophysical Research Communications 183(1):327-333.

Zhang Y-HP. 2010. Production of biocommodities and bioelectricity by cell-free synthetic enzymatic pathway biotransformations: Challenges and opportunities. Biotechnology and Bioengineering 105(4):663-677.

Table I: Multiple sequence alignment of cofactor binding pocket residues of selected mesostable and thermostable AKRs.

	Protein	Origin	Residue...	24	50	166	167	190	216	219	221	270	271	272	276	279	280
1	aldehyde reductase	Human		W	D	S	N	Q	Y	L	S	K	S	I	R	Q	N
2	3α-HSD	Rat		T	D	S	N	Q	Y	L	S	R	S	F	R	E	L
3	2,5-DKGR	Mesostable bacterium		F	D	S	N	Q	W	L	Q	K	S	V	R	E	N
4	AdhD	Thermostable archaeon		W	D	S	N	Q	Y	L	K	K	A	S	H	E	N
5	Putative AKR	Thermostable archaeon		W	D	S	N	Q	Y	L	K	K	A	I	H	E	N
6	Putative AKR	Thermostable archaeon		Y	D	S	N	Q	W	L	H	R	A	S	H	E	N
7	Putative AKR	Thermostable bacterium		Y	D	A	T	Q	A	L	V	G	M	S	H	E	N

Rat liver 3α-HSD numbering, the shaded positions correspond to positions 249 and 255 as mutated in this study. 1. Human aldehyde reductase (Accession #P14550), 2. Rat liver 3α-hydroxysteroid dehydrogenase (Accession #P23457), 3. *Corynebacterium* 2,5-diketo-D-gluconic acid reductase A (Accession #P06632), 4. *Pyrococcus furiosus* alcohol dehydrogenase D (Accession #NP_579689), 5. Putative AKR from *Thermococcus barophilus* (Accession #EDY40262), 6. Putative AKR from *Thermococcus volcanium* (Accession #NP_111671), 7. Putative AKR from *Aquifex aeolicus* (Accession #NP_213220).

130

Table II: Full steady state kinetic parameters for wild type AdhD and the K249G/H255R double mutant

		Oxidation						Reduction			
		K_D (µM)	K_{ia} (µM)	k_{cat} (s^{-1})	K_A (µM)	K_B (mM)		K_D (µM)	K_{ia} (µM)	k_{cat} (s^{-1})	K_A (µM)	K_B (mM)
wt AdhD	NAD$^+$	59 ±1	37 ±2	1.0 ±0.1	63 ±2	29 ±1	NADH	38 ±1	350 ±200	0.36 ±0.01	190 ±10	0.90 ±0.50
K249G/H255R		45 ±2	11 ±1	15 ±2	460 ±60	690 ±80		20 ±2	14 ±2	1.1 ±0.1	50 ±3	13 ±2
wt AdhD	NADP$^+$	5.7 ±1.0	20 ±8	0.030 ±0.020	5.1 ±0.2	1.3 ±0.1	NADPH	43 ±2	ND	0.20 ±0.02	280 ±40	6.7 ±1.1
K249G/H255R		0.66 ±0.10	1.2 ±0.2	4.7 ±0.2	78 ±5	200 ±9		35 ±2	ND	1.2 ±0.1	33 ±4	5.0 ±0.5

Oxidation reactions were performed at 45°C in 50 mM glycine (pH 8.8) with 2,3-butanediol substrate. Reduction reactions were performed at 45°C in 100 mM sodium phosphate (pH 6.1) with 3-hydroxy-2-butanone substrate. K_D is the cofactor dissociation constant as determined by fluorescence titration under the same conditions. K_A and K_B are the Michaelis constants for the cofactor and substrate, respectively (Equation 2). Reactions were performed in at least triplicate, and errors are standard deviations. ND: The K_{ia} term was unable to be fit by the model, and was instead set equal to the measured K_D (Segel 1993).

Table III: Microscopic rate constants calculated from steady-state kinetic parameters.

		Oxidation					Reduction				
		k_1^{ss} (uM^{-1} s^{-1})	k_2^{ss} (s^{-1})	k_3^{ss} (uM^{-1} s^{-1})	$(k_1k_3/k_2)^{ss}$ (uM^{-2} s^{-1})	$(k_1k_3/k_2)/(k_1k_3/k_2)_{wt}$		k_1^{ss} (uM^{-1} s^{-1})	k_2^{ss} (s^{-1})	k_3^{ss} (uM^{-1} s^{-1})	$(k_1k_3/k_2)^{ss}$ (uM^{-2} s^{-1})	$(k_1k_3/k_2)/(k_1k_3/k_2)_{wt}$
wt AdhD	NAD	0.016	0.59	34	0.93	1.0	NADH	0.0019	0.68	400	1.1	1.0
K249G/H255R		0.033	0.36	22	2.0	2.1		0.022	0.31	85	6.0	5.3
wt AdhD	NADP	0.0059	0.12	23	1.2	1.0	NADPH	0.00072	0.031[a]	30	0.69	1.0
K249G/H255R		0.060	0.07	24	20	17		0.036	1.3[a]	240	6.9	9.9

Rate constants calculated from the relationships: $k_1^{ss} = k_{cat}/K_A$, $k_2^{ss} = k_{cat}K_{ia}/K_A$, and $k_3^{ss} = k_{cat}/K_b$, for the mechanism described in Equation 3.

Table IV: Comparison of cofactor dissociation constants measured by fluorescence titrations and stopped-flow fluorescence spectroscopy

Enzyme	Cofactor	K_D (µM) Measured	K_D (µM) Calculated	Microscopic Rate Constants			Note
				K_1 (µM)	k_2 (s^{-1})	k_{-2} (s^{-1})	
wt AdhD	NAD$^+$	65 ±2	0.65	5.5	3.8	0.51	
	NADH	42 ±1	ND	-	-	-	Small transients
	NADP$^+$	25 ±1	0.01	3.2	9.6	0.019	
	NADPH	29 ±1	0.11	0.70	4.6	0.83	
K249G	NAD$^+$	1.7 ±0.1	NA	1.5	7.9	0	
	NADH	20 ±1	0.66	3.3	26	6.4	
	NADP$^+$	0.4 ±0.1	NA	3.2	12	0	
	NADPH	23 ±1	ND	-	-	-	Small transients
H255R	NAD$^+$	91 ±2	ND	-	-	-	Small transients
	NADH	29 ±1	ND	-	-	-	Small transients
	NADP$^+$	0.3 ±0.1	0.13	14	25	0.23	
	NADPH	11 ±1	0.10	3.0	15	0.52	
K249G/H255R	NAD$^+$	29 ±1	11	78	28	4.5	
	NADH	25 ±1	ND	-	-	-	Small transients
	NADP$^+$	0.7 ±0.1	1.5	18	43	4.0	
	NADPH	19 ±1	0.31	1.7	38	8.2	See discussion

Dissociation constants were calculated as previously described from hyperbolic fits to the observed rate constants of the kinetic transients versus the cofactor concentrations (Figure 5C, SFigure 5). Measured dissociation constants were obtained by fluorescence titrations performed under the same conditions as the stopped flow experiments. NA: Not applicable as no estimate was able to be obtained for k_{-2}.
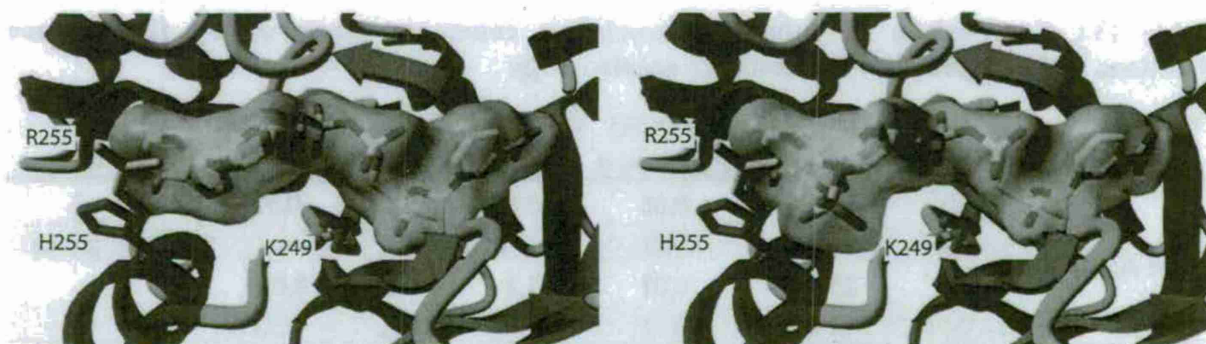
Figure 1: Homology model of the AdhD cofactor binding pocket with bound cofactors
Homology model of AdhD with bound NAD(H) (left) and NADP(H) (right). Side chains of Lys249 and His255 of the wild-type AdhD are shown in purple, and Arg255 of the double mutant is shown in yellow. His255 is in position to form a stacking interaction with the adenine moiety of the cofactor, while Arg255 can form an electrostatic interaction with the 2'-phosphate in NADP(H). Lys249 extends beneath the pyrophosphate group and also forms an electrostatic interaction with the 2'-phosphate in NADP(H).
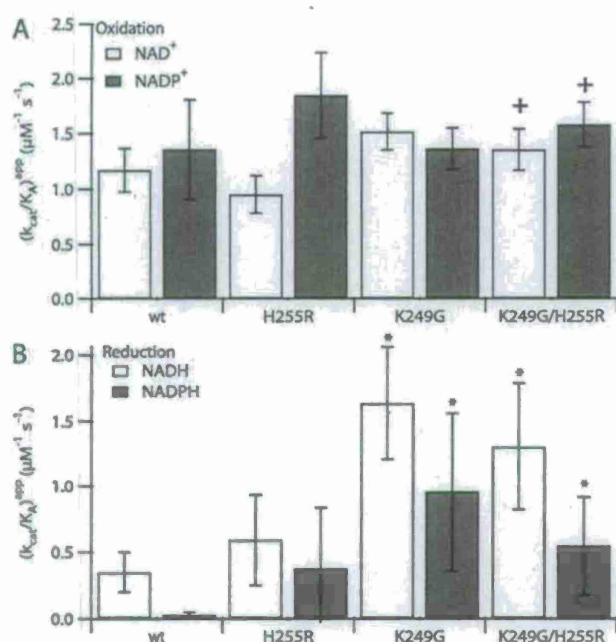
Figure 2: Apparent catalytic efficiencies (k_{cat} / K_M)^{app} of wild-type AdhD and mutants in the oxidation and reduction reactions.

Apparent catalytic efficiencies determined under fixed substrate conditions were calculated using Equation 1. Reaction mixtures contained 50 mM glycine (pH 8.8), 100 mM 2,3-butanediol, 1-1000 μM NAD(P)$^+$, and enzyme (oxidation reaction, A) or 100 mM sodium phosphate (pH 6.1), 80 mM 3-hydroxy-2-butanone, 1-500 μM NAD(P)H, and enzyme (reduction reaction, B) at 45°C. Measurements were performed in triplicate, and experiments were repeated three times with fresh solutions. Error bars are standard deviations. Asterisks indicate statistically significant difference from wild-type AdhD at $p < 0.05$. +: fits to Equation 3 suggest substrate concentration is not saturating, thus the actual k_{cat} is likely higher.
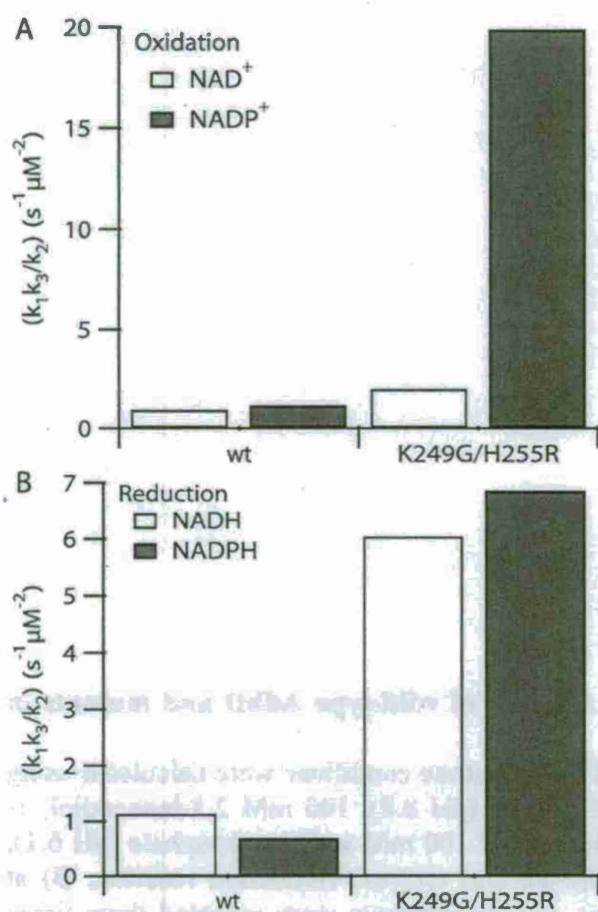
Figure 3: Activity of wt AdhD and K249G/H255R with each cofactor
Microscopic rate constants calculated from the steady-state kinetic parameters (Table III). Comparing the value of $(k_1 k_3 / k_2)^{ss}$ demonstrates the significant improvement in activity and broadened specificity of the double mutant enzyme.
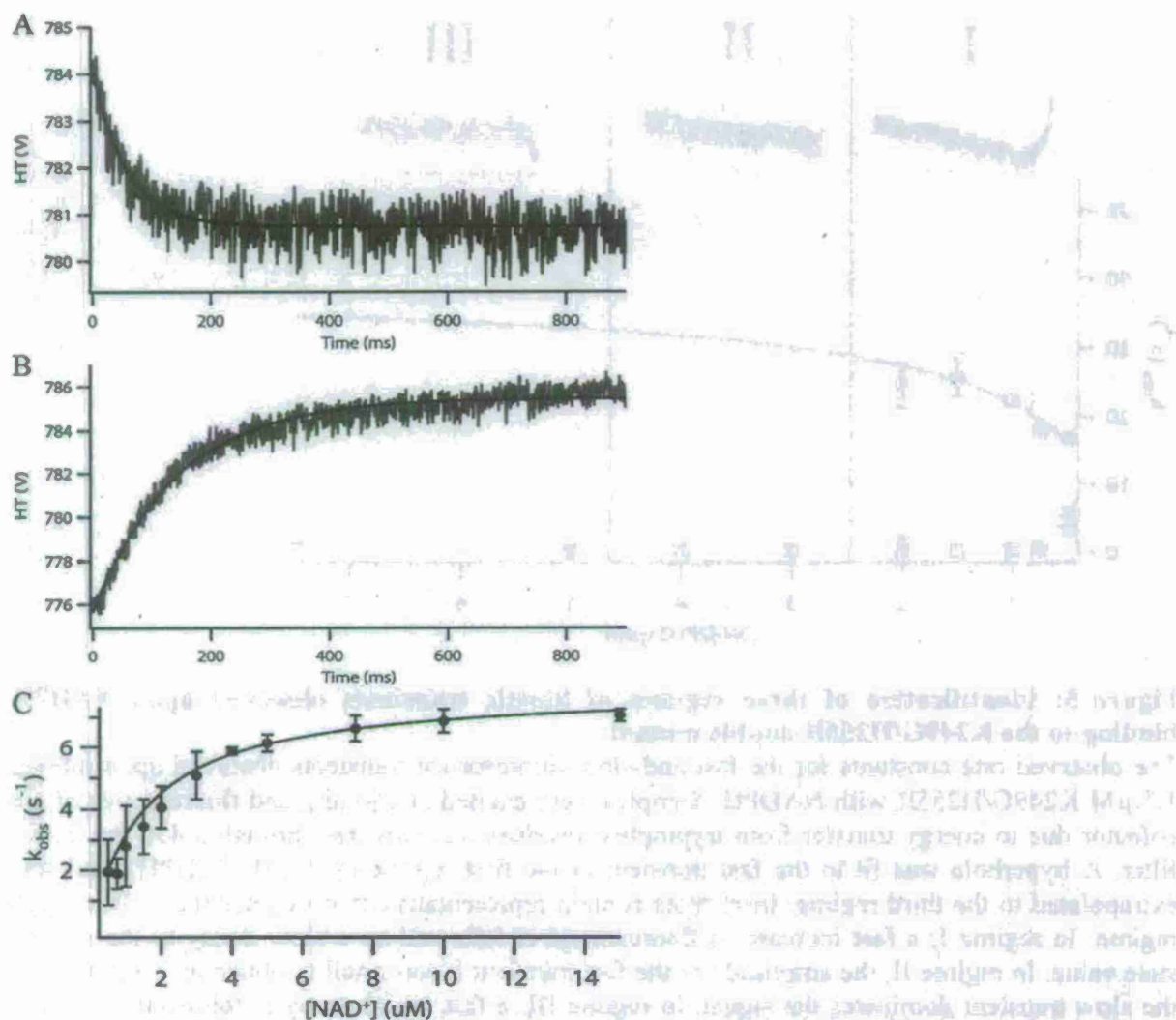
Figure 4: Fluorescence kinetic transients observed upon cofactor binding for the K249G AdhD mutant

Representative fluorescence traces of K249G AdhD mutant (0.75 µM) with (A) 7.5 µM NADH and (B) 7.5 µM NAD^+ fit to a mono-exponential function. (C) Plot of the observed rate constant as a function of cofactor concentration for K249G AdhD with NAD^+. Error bars are standard deviations of at least three independent measurements. The data was fit with a hyperbola to obtain estimates of the microscopic rate constants (Table IV).

Figure 5: Identification of three regimes of kinetic transients observed upon NADPH binding to the K249G/H255R double mutant

The observed rate constants for the fast and slow fluorescence transients observed upon mixing 1.5 µM K249G/H255R with NADPH. Samples were excited at 280 nm, and fluorescence of the cofactor due to energy transfer from tryptophan residues was detected through a 430 nm cutoff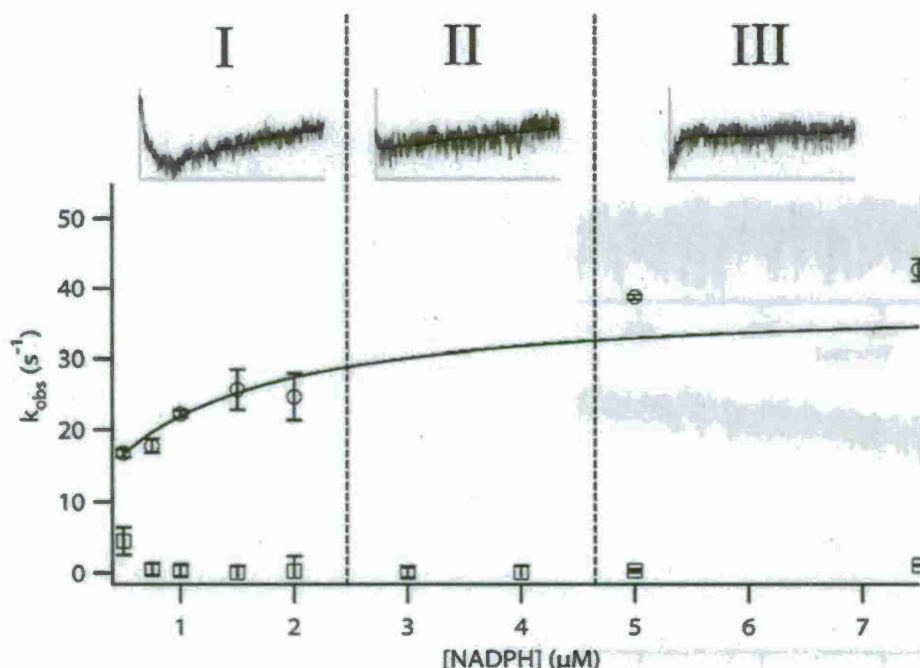 filter. A hyperbola was fit to the fast transient in the first regime (≤ 2 µM NADPH), and was extrapolated to the third regime. Inset plots contain representative fluorescence traces from each regime. In regime I, a fast increase in fluorescence is followed by a slow decay to the steady-state value. In regime II, the amplitude of the fast transient is too small to obtain an estimate, and the slow transient dominates the signal. In regime III, a fast initial decay is followed by a slow decay to the steady-state value. These results are consistent with a concentration-dependent reversal in the direction of the reorientation that occurs in the cofactor binding pocket during the transient conformational change.

138

Chapter 9

Modular exchange of substrate binding loops alters both substrate and cofactor specificity in a member of the aldo-keto reductase superfamily

*Elliot Campbell, Sara Chuang, and Scott Banta**

Department of Chemical Engineering, Columbia University in the City of New York, New York, NY 10027, USA

*To whom correspondence should be addressed. Phone: (212) 854-7531. Fax: (212) 854-3054. E-mail: sbanta@columbia.edu.

Abstract: Substrate specificity in the aldo-keto reductase (AKR) superfamily is determined by three mobile loops positioned at the top of the canonical (α-barrel) structure. These loops have previously been demonstrated to be modular in a well studied class of AKRs, in that exchanging loops between two similar hydroxysteroid dehydrogenases resulted in a complete alteration of substrate specificity (Ma, H., and Penning, T. M. (1999) *PNAS 96*, 11161-11166). Here, we further examine the modularity of these loops by grafting those from human aldose reductase (hAR) into the hyperthermostable AKR, AdhD (alcohol dehydrogenase D), from *Pyrococcus furiosus*. Replacement of Loops A and B was sufficient to impart hAR activity into AdhD, and the resulting chimera retained the thermostability of the parent enzyme. However, no active chimeras were observed when the hAR loops were grafted into a previously engineered cofactor specificity mutant of AdhD, which displayed similar kinetics to hAR with the model substrate DL-glyceraldehyde. The non-additivity of these mutations suggests that efficient turnover is more dependent on the relative positioning of the cofactor and substrate in the active site than on binding of the individual species. The ability to impart the substrate specificities of mesostable AKRs into a thermostable scaffold will be useful in a variety of applications including immobilized enzyme systems for bioelectrocatalysis and fine chemical synthesis.

Aldo-keto reductases (AKRs) comprise a large, diverse family of oxidoreductase enzymes and are found in nearly every species.[1] They share a common $(\alpha/\beta)_8$-barrel structure and catalytic mechanism, but some members of the superfamily share less than 30% sequence homology. These enzymes bind a nicotinamide cofactor in an extended conformation along a cleft that runs through the C-terminal face of the barrel, in contrast to the Rossman-fold motif common in other dehydrogenases. Three mobile loops on the same face form the substrate binding pocket. The physiological role of many of these enzymes is unknown, but they are generally thought to fall into one of three classes. The most studied members of this family are mammalian AKRs involved in steroid and prostaglandin metabolism.[2] These enzymes often have long substrate binding loops and are highly specific for their substrates. Another well studied class of AKRs, the aldose reductases, are involved in the interconversion of glucose to sorbitol and have been investigated as drug targets to prevent complications from diabetes.[3] A third class of AKRs, which have been identified in a wide range of species, have no known function.[4] The substrate binding loops in these enzymes are often truncated, imparting them with fairly broad substrate specificity. Additionally, they have been shown to be upregulated in response to stress, leading to the hypothesis that their physiological role is of general detoxification, metabolizing various aldehydes and ketones to less toxic species.[5]

The advent of high-throughput sequencing has allowed the complete genomes of several species to be elucidated, and the sequence data has yielded several putative members of the aldo-keto reductase superfamily.[6] One of these enzymes, alcohol dehydrogenase D (AdhD), was identified in the hyperthermophilic archaea *Pyrococcus furiosus*, and has been characterized by our group and others. The enzyme has a strong preference for NAD(H) as a cofactor, and oxidizes a range of sugars and alcohols.[6b, 7] The substrate binding loops in this enzyme are significantly truncated compared to AKRs identified from other organisms, with the C-terminal loop (Loop C), completely absent. This likely contributes to the broad substrate specificity and extreme thermostability of the enzyme.

The importance of the mobile loops in substrate binding and specificity was elegantly demonstrated through the creation of several chimeric hydroxysteroid dehydrogenases (HSD) where the substrate binding loops from a 20α-HSD were grafted into a 3α-HSD enzyme scaffold.[8] Replacement of only Loop A resulted in an enzyme with novel 17 HSD activity, while swapping all three substrate binding loops resulted in a complete alteration of substrate specificity, with an increase in catalytic efficiency for the 20α-HSD reaction of 10^{11} compared to the wild-type 3α-HSD enzyme.

Based on this impressive work, we decided to investigate a similar strategy to rationally alter the substrate specificity of AdhD. In an attempt to improve the activity of AdhD with sugars, we created several loop chimeras inserting the substrate binding loops from human aldose reductase, which has activity with glucose.[3a, 9] These loop chimeras are also compared to and combined with a cofactor specificity double mutant of AdhD (K249G/H255R) that exhibits broadened cofactor specificity and improved activity compared to the wild-type enzyme.[7] A summary of loop chimera constructs appears in Table 1.

Whereas the previous work exchanged substrate binding loops between similar hydroxysteroid dehydrogenases, the present work investigates exchanging loops between two distinct AKRs which share less than 30% sequence homology. While AdhD is an extremely thermostable archaeal enzyme with broad substrate specificity, human aldose reductase (hAR) is a mesostable mammalian AKR with a specialized function. Also, AdhD has a strong preference for NAD(H)

141

as a cofactor, while hAR has a strong preference for NADP(H). Thus, in addition to the change in substrate specificity expected due to changing the substrate binding loops, it will also be interesting to observe the effects on cofactor specificity and thermostability.

Materials and Methods

DL-glyceraldehyde, 2,3-butanediol (mixture of DL and meso), 3-hydroxy-2-butanone, all cofactors, media, and buffer components were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise noted and used without modification.

In order to design the AdhD/hAR loop chimeras, sequence and structural alignments were performed. Sequence alignments of AdhD (GenBank 1469842) and hAR (GenBank AAA51713) were performed using the CLUSTALW tool, and structural alignments of AdhD and hAR (PDB 2ACQ) were performed with Yasara. DNA oligos corresponding to the hAR substrate binding loops were obtained from IDT DNA, Inc. (Coralville, IA) and assembled into the AdhD gene using overlap extension PCR (see SI). PCR fragments were doubly digested with NcoI and HindIII and cloned into a similarly digested pET-24d vector. All constructs were verified by DNA sequencing.

AdhD/hAR Loop chimeras were initially expressed in 50ml cultures and purified by heating of the cell extracts, as described previously.[7] Two constructs, A (wt AdhD with Loop A) and D (DM AdhD with Loops AB), were found to be poorly expressed despite efforts to optimize the expression and purification of these samples. Thus only enough enzyme was produced to run the initial screening assay with these constructs. Relatively pure protein (estimated > 90% pure by SDS-PAGE) was obtained in the heat-stable extract, and was used without further purification for initial studies.

Large scale expression and purification of wild-type AdhD and mutants followed a previously described protocol.[7] Typical yields were on the order of 300 - 1200 mg L^{-1} of culture, and samples were estimated to be >98% pure by SDS-PAGE.

The hAR gene was amplified from human placenta QUICK-clone cDNA (Clontech, Mountain View, CA) using forward primer 5'-GGTCTGGGGAGCGCAGCAGC-3' and reverse primer 5'-TTCGAAGCTTTCAAAACTCTTCATGGAAGGGGTAATCCTT-3'. The reverse primer inserted a unique HindIII restriction site (underlined). The purified PCR fragment was doubly digested with NcoI and HindIII and ligated into a similarly digested pET-24d vector containing an N-terminal RGSHis tag for purification. Ligated plasmids were transformed into electrocompetent BLR *E. coli* (Novagen, Gibbstown, NJ) and plated on LB-Kan selection plates. Individual colonies were picked and grown overnight in LB medium supplemented with 50 µg ml^{-1} kanamycin and stored as glycerol stocks. Proper insertion of the hAR gene was verified by DNA sequencing.

Expression and purification of hAR followed a different protocol, as the enzyme is not highly thermostable. One liter expression cultures of Terrific Broth containing 50 µg ml^{-1} kanamycin were inoculated from an overnight culture, and expression was induced at $OD_{600} \approx 0.6$ by the addition of IPTG to 0.2mM. Expression continued for 16h at 37°C with agitation. Cells were harvested by centrifugation, and resuspended in 1/10th volume Binding Buffer (20mM Tris-HCl, 150mM NaCl, 40mM imidazole, pH 7.5) supplemented with HALT Protease Inhibitor (Fisher Scientific). Cells were lysed by sonication on ice for a total of 8 minutes, following cycles of 5 seconds on and 5 seconds off. Cell debris was removed by centrifugation for 20 mins at 10000g. Samples were loaded onto a HisTrap column (GE Healthcare, Piscataway, NJ) equilibrated in Binding Buffer. After rinsing with 10 column volumes of Binding Buffer, His-tagged hAR was

142

eluted with a gradient of 0-100% Elution Buffer (20mM Tris-HCl, 150mM NaCl, 500mM imidazole, pH 7.5) over 20 column volumes. hAR eluted in a single peak at an imidazole concentration of ~150mM. Fractions containing hAR were pooled and concentrated over a 30kDa centrifugal filter and applied to a Superdex 16/200 gel filtration column (GE Healthcare, Piscataway, NJ) equilibrated in 20mM Tris-HCl (pH 7.5) containing 150mM NaCl. Fractions containing enzyme were pooled and concentrated over a 30kDa filter, before being diluted to the desired working concentration in 20mM Tris-HCl (pH 7.5). Typical yields were on the order of 30 mg L^{-1} of culture, and samples were estimated to be >98% pure by SDS-PAGE.

All loop mutants were initially screened in a 96-well plate assay. To test for activity in the oxidation reaction, 10μl of partially purified enzyme (~5 mg/ml) was added to 290μl 50mM glycine (pH 8.8) containing 1mM NAD^+ or $NADP^+$ and 10mM of the indicated substrate in a 96-well UV-transparent microplate. For the reduction reaction, 10μl of partially purified enzyme was added to 290μl 100mM sodium phosphate (pH 6.1) containing 500μM NADH or NADPH and 10mM of the indicated substrate. Plates were incubated at 37°C and imaged under UV light at various time points to monitor the production or depletion of reduced cofactor.

Full kinetic assays were performed on the active enzymes identified by the plate assay. For the oxidation reaction, 10μl of the appropriate concentration of purified enzyme was added to 290μl 50mM glycine (pH 8.8) containing 5-2000μM NAD^+ or $NADP^+$ and 1-100mM of the indicated substrate. For the reduction reaction, 10μl of the appropriate concentration of purified enzyme was added to 290μl 100mM sodium phosphate (pH 6.1) containing 1-500μM NADH or NADPH and 1-100mM of the indicated substrate. Plates were incubated (at 25°C for the reduction reaction, 37°C for the oxidation reaction) in a SpectraMax M2 spectrophotometer (Molecular Devices, Sunnyvale, CA) and the absorbance at 340nm was followed to monitor the production or depletion of NAD(P)H (ε_{340nm} = 6220 M^{-1} cm^{-1}). Experiments were performed in at least triplicate, and the background rate of NAD(P)H degradation was corrected for using blank reactions run in parallel. Kinetic data were fit to the ordered bi-bi rate equation using a non-linear regression program (Igor Pro, Wavemetrics, Inc.). In order to obtain an estimate for the catalytic efficiency (k_{cat}/K_A) of wt AdhD with NADPH and DL-glyceraldehyde for the binding energy calculations, kinetics were run with a high concentration of enzyme and substrate, and the inverse of the slope of a Lineweaver-Burk plot to the data was used.

Cofactor dissociation constants were measured by fluorescence titration, following a previously described protocol[10] Proper folding of the loop chimeras and determination of thermal stability were investigated by CD spectroscopy as described previously.[11]

Cofactor binding energies in the ground-state (ΔG) and transition-state ($\Delta G^{\ddagger}$) were obtained from Eqs. 1 and 2,[12] utilizing the steady-state kinetic parameters.

$$\Delta G = -RT\ln[(K_{ia})_{construct}/(K_{ia})_{wt\,AdhD}] \qquad \text{Eq. 1}$$
$$\Delta G^{\ddagger} = RT\ln[(k_{cat}/K_A)_{construct}/(k_{cat}/K_A)_{wt\,AdhD}] \qquad \text{Eq. 2}$$

Results

A sequence and structural alignment of hAR (PDB ID 2ACQ) with a previously generated homology model of AdhD[7] guided the insertion of the hAR binding loops into AdhD (Figure 1). The structural alignment also identified a short loop in AdhD (corresponding to residues 182-184) not present in hAR, which could potentially sterically interfere with Loop C and prevent it from properly folding over the top of the barrel. Thus additional mutants were generated with

143

this short loop removed (denoted ⧠Δ842to increase the likelihood of Loop C adopting its native conformation.

Loops were inserted at the genetic level through a series of oligonucleotide primers, which were used to PCR fragments of the gene containing the desired loops (see SI). These fragments were then reassembled into a full length gene using overlap-extension PCR, and cloned into a vector for expression.

A concern when grafting in the large substrate binding loops from hAR was a decrease in the thermostability of the AdhD scaffold. However, a thermal purification step was still possible with the mutant enzymes, and they were further characterized by CD spectroscopy and thermal denaturation experiments. Surprisingly, the impact on enzyme stability was minimal, as no change in CD signal was observed from 25°C to 90°C (data not shown).

Loop mutants were initially screened in a plate assay. Mutants were tested for their ability to reduce DL-glyceraldehyde with NADPH, the model substrates for hAR, and in the oxidation and reduction of 2,3-butanediol and 3-hydroxy-2-butanone respectively, the model substrates for AdhD, using both NAD(H) and NADP(H) cofactors. Plates were illuminated by UV to visualize the reduced cofactor and photographed at regular intervals. A representative image of the plate after 45 minutes of incubation is shown in Figure 2. At this time point, only hAR showed appreciable activity with DL-glyceraldehyde, and clearly had a preference for NADPH over NADH. The AdhD double mutant showed the highest activity with 2,3-butanediol and 3-hydroxy-2-butanone, with little difference apparent between the NAD(H) and NADP(H) cofactors. The activity of wt AdhD was lower with these substrates, and a marked preference for NAD+ was observed with 2,3-butanediol. hAR also demonstrated activity with both 2,3-butanediol and 3-hydroxy-2-butanone, and had a slight preference for its preferred cofactor NADP(H) with these model AdhD substrates. Interestingly, constructs C and E also retained significant activity with these substrates, but only when NADP(H) was the cofactor. At longer time points, NADPH/DL-glyceraldehyde activity was also observed in these loop chimeras.

Enzymes that were identified as active in the plate assay were grown in large scale expression cultures and purified to homogeneity as described. A full kinetic analysis was performed with these samples to allow for fitting to the ordered bi-bi rate equation. Kinetic parameters are summarized in Table 2. While the wt AdhD exhibited very little detectable activity with NADPH and DL-glyceraldehyde, the cofactor specificity double mutant (DM AdhD) had a turnover rate more than 50% faster than hAR (46 s^{-1} vs 30 s^{-1}). This was offset by one to two order of magnitude increases in the dissociation constant and Michaelis constant for NADPH, however, leading to a lower catalytic efficiency. The two loop chimeras, C and E, identified as active in the plate assay demonstrated reasonable turnover with NADPH and DL-glyceraldehyde (10-20% of hAR), but again the Michaelis constants were two to three orders of magnitude larger than those for hAR or the DM AdhD. With the model AdhD substrate 2,3-butanediol, an interesting effect is observed with the loop chimeras. Both constructs C and E demonstrated increased activity with this substrate, but had a strict requirement for $NADP^+$ as a cofactor, in contrast to the NAD^+ preference exhibited by wt AdhD. hAR, surprisingly, was found to have the highest catalytic rates with this substrate with both NAD^+ and $NADP^+$ (k_{cat} of 76 s^{-1} and 26 s^{-1}, respectively), however the catalytic efficiency was much higher with its preferred cofactor $NADP^+$.

The combination of relatively low turnover numbers and high Michaelis constants impeded the accurate determination of the full kinetic parameters for the two loop chimeras. In order to enable fitting to the ordered bi-bi rate equation, the K_{ia} term was set equal to the dissociation

144

constant measured by fluorescence titration.[13] As proper saturating conditions were not achieved with these mutants, the resulting kinetic parameters are given as apparent parameters.

Discussion

The modular nature of the aldo-keto reductase substrate binding loops has been confirmed in this work, as the AdhD scaffold was successfully imparted with hAR activity through a loop grafting approach. Whereas a complete reversal of substrate specificity was previously shown to require the exchange of all three substrate binding loops, here it appears that only two loops are necessary for activity. Additionally, the chimeric mutants studied here maintained the high thermostability of the parent enzyme, suggesting that this technique can be used to rapidly stabilize other mesophilic AKRs.

Comparisons of the catalytic efficiencies of the various enzyme constructs are difficult to interpret due to the large difference in Michaelis constants between enzymes. Thus, activities were examined under saturating conditions for the wild-type enzyme by looking at the turnover rate, k_{cat}. The catalytic rate of wt AdhD is much lower than that of hAR with NADP(H), both in the oxidation of the model AdhD substrate 2,3-butanediol (0.03 s^{-1} vs 26 s^{-1}) and the reduction of the model hAR substrate DL-glyceraldehyde (<0.01 s^{-1} vs 30 s^{-1}). The active loop chimeras fall in between, with construct C having similar activities with both substrates (k_{cat} of 4.5 s^{-1} with 2,3-butanediol vs 5.2 s^{-1} with DL-glyceraldehyde) while construct E has a much faster turnover rate with 2,3-butanediol (16 s^{-1} vs 2.8 s^{-1} with DL-glyceraldehyde). Interestingly, the DM AdhD has a much higher turnover rate with DL-glyceraldehyde than even hAR, but is comparable to the construct C with 2,3-butanediol (Figure 3A). The steady-state kinetic parameters can also be used to calculate changes in the cofactor binding energies relative to the wt AdhD enzyme,[12, 14] which may shed light onto the kinetic results. In the ground state, the loop chimeras destabilized the binding of $NADP^+$ by 1-2 kcal/mol, while hAR and the previously engineered DM AdhD had ~2 kcal/mol more favorable binding energies. All constructs except E also demonstrated a decreased free energy of binding with NADPH relative to the wt AdhD, which partially explains the improvement in activity observed in these constructs (Figure 3B). The effect of the loop chimeras is most apparent when comparing the transition-state binding energies. Here, both constructs C and E have a significantly lower transition-state binding energy with DL-glyceraldehyde/NADPH compared to the wt AdhD (by ~ 3-4 kcal/mol), while those with 2,3-butanediol/$NADP^+$ remain relatively unaffected. (Figure 3C) Thus the loop chimeras increase the affinity of the enzyme for the NADPH/DL-glyceraldehyde transition state, but without loss of affinity for 2,3-butanediol.

In contrast to the previous work, the active loop chimeras retained activity with their native substrate. However, the strict reversal in cofactor specificity from NAD^+ to $NADP^+$ was unexpected. As hAR has been shown to prefer NADP(H), this suggests the substrate binding loops can also impact cofactor specificity. In fact, Loop B of some AKRs has been demonstrated to take part in cofactor binding through electrostatic interactions. The crystal structure of hAR indicates that residue Asp216 on Loop B forms a salt-bridge with Lys262 to form the canonical AKR "seat-belt" over the pyrophosphate backbone of the cofactor, thereby locking it into the binding pocket. This motif is likely absent in the wt AdhD, as Loop B is significantly truncated and lacks the charged residue required to form an electrostatic interaction. Grafting Loop B from hAR into AdhD may therefore reconstitute the "seat-belt", and promote binding and proper orientation of the cofactor in the binding pocket. Interestingly, construct C, with both Loops A and B has the second lowest dissociation constant for NADPH, behind only hAR.

145

Given the success in grafting the hAR loops into the AdhD scaffold, it is interesting that the same loops grafted into the double mutant AdhD scaffold, which itself possesses hAR-like activity, results in an inactive enzyme. As the wt AdhD requires at least Loop B of hAR for activity with DL-glyceraldehyde, this supports the hypothesis that cofactor binding and orientation provided by the "seat-belt" mechanism is important for catalysis. One of the mutations in the double mutant, K249G, removes the lysine residue that interacts with the aspartic acid of Loop B. Thus the formation of a "seat-belt" in the DM AdhD loop constructs is unlikely, and if this has a detrimental effect on cofactor binding, could explain the lack of activity observed in these chimeras.

Construct	Enzyme / Loops
A	wt Loop A
B	DM Loop A
C	wt Loops AB
D	DM Loops AB
E	wt Loops ABC
F	DM Loops ABC
G	wt Loops ABC + 1842
H	DM Loops ABC + 1842

Table 1 – AdhD / hAR Loop Chimera Constructs

Enzyme	Cofactor	Substrate	k_{cat} (s^{-1})	K_{ia} (μM)	K_A (μM)	K_B (mM)
wt AdhD	NADPH	DL-glyceraldehyde	< 0.01	43 ± 2[a]	-	-
DM AdhD	NADPH	DL-glyceraldehyde	46 ± 6	35 ± 2[a]	180 ± 40	11 ± 4
wt Loops AB	NADPH	DL-glyceraldehyde	5.2 ± 1.3	4.9 ± 0.7	460 ± 120	450 ± 130
wt Loops ABC	NADPH	DL-glyceraldehyde	2.8 ± 0.6	280 ± 30	1200 ± 300	250 ± 70
hAR	NADPH	DL-glyceraldehyde	30 ± 1	0.45 ± 0.08	20 ± 2	24 ± 1
wt AdhD[b]	NAD^+	2,3-butanediol	1.0 ± 0.1	37 ± 2	63 ± 2	29 ± 1
wt AdhD[b]	$NADP^+$	2,3-butanediol	0.03 ± 0.02	20 ± 8	5.1 ± 0.2	1.3 ± 0.1
DM AdhD[b]	NAD^+	2,3-butanediol	15 ± 2	11 ± 1	460 ± 60	690 ± 80
DM AdhD[b]	$NADP^+$	2,3-butanediol	4.7 ± 2	1.2 ± 0.2	78 ± 5	200 ± 9
wt Loops AB	NAD^+	2,3-butanediol	ND	-	-	-
wt Loops AB	$NADP^+$	2,3-butanediol	4.5 ± 0.8*	161*[a]	750^*	$160 \pm 40^*$
wt Loops ABC	NAD^+	2,3-butanediol	ND	-	-	-
wt Loops ABC	$NADP^+$	2,3-butanediol	$16 \pm 5^*$	$450 \pm 25^*$	$3800 \pm 1200^*$	$470 \pm 160^*$
hAR	NAD^+	2,3-butanediol	76 ± 30	350 ± 40	750 ± 40	350 ± 200
hAR	$NADP^+$	2,3-butanediol	26 ± 2	0.63 ± 0.2	5.4 ± 1.6	180 ± 20

Table 2 – Full Kinetic Parameters with the Model Substrates for hAR and AdhD

*apparent parameters, [a]Measured by fluorescence titration. [b]Data from ref [7].

K_{ia} is the dissociation constant of the enzyme cofactor complex. K_A and K_B are the Michaelis constants of the cofactor and substrate, respectively.

A

B   Δ182-184   Loop A

Loop B

Loop C

C

```
2ACQ  -------ASRLLLNNGAKMPILGLGTWKS---------PPGQVTEAVKVAIDVGYRHIDC  44
AdhD  MAKRVNAFNDLKRIGDDKVTAIGMGTWGIGGRETPDYSRDKESIEAIRYGLELGMNLIDT  60
         .  *   .. *:. :*:***              :  **:: .::::*  .  **

2ACQ  AHVYQN---ENEVGVAIQEKLREQVVKREELFIVSKLWCTYHEKGLVKGACQKTLSDLKL  101
AdhD  AEFYGAGHAEEIVGEAIKEFE------REDIFIVSKVWPTHFGYEEAKKAARASAKRLG-  113
      *..*      *: ** **:*      **::*****:* *:.    .* *.:: .  . *
                      Loop A
2ACQ  DYLDLYLIHWPTGFKPGKEFFPLDESGNVVPSDTNILDTWAAMEELVDEGLVKAIGISNF  161
AdhD  TYIDLYLLHWP-----------------VDDFKKIEETLHALEDLVDEGVIRYIGVSNF  155
       *:****:***                 *  . .:* :*  *:*:*****::: **:***
                      Δ182-184
2ACQ  NHLQVEMILNKPGLKYKPAVNQIECHP---YLTQEKLIQYCQSKGIVVTAYSPLGSPDRP  218
AdhD  N-LELLQRSQEVMRKYEIVANQVKYSVKDRWPETTGLLDYMKREGIALMAYTPLEKG---  211
      * *::    ::    **: ..**::       *::*  : :**.: **:**  .
      Loop B
2ACQ  WAKPEDPSLLEDPRIKAIAAKHNKTTAQVLIRFPMQRN-LVVIPKSVTPERIAENFKVFD  277
AdhD  -------TLARNECLAKIGEKYGKTAAQVALNYLIWEENVVAIPKASNKEHLKENFGAMG  264
             :* .:  :  *. *::**:*** :.: : .: :*.***: . *:: *** .:.
                      Loop C
2ACQ  FELSSQDMTTLLSYNRNWRVCALLSCTSHKDYPFHEEF  315
AdhD  WRLSEEDREMARRCV----------------------  279
      :.**.:*
```
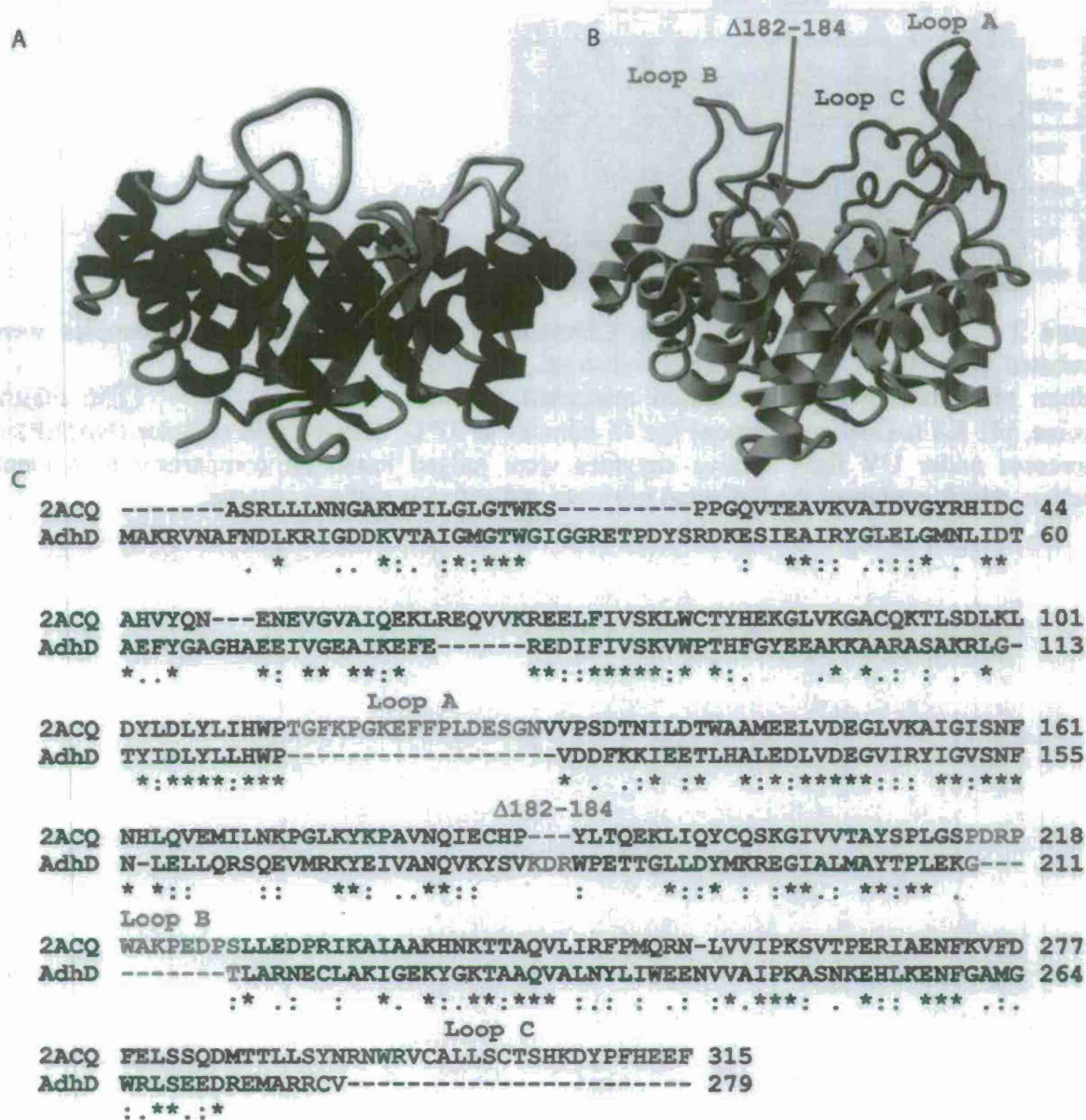
Figure 1 – (A) Homology model of AdhD and (B) crystal structure of hAR (PDB ID 2ACQ) with substrate binding loops indicated. (C) Sequence alignment of hAR and AdhD showing the location of the substrate binding loops and their absence in AdhD.
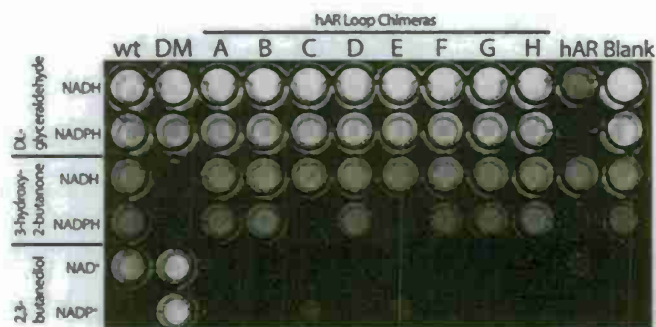
Figure 2 – Plate Assay of hAR Loop Chimeras. Partially purified enzyme samples were incubated with 10mM of the indicated substrate, and 500 ☐M cofactor sodium phosphate, pH 6.1 (reduction reactions) or 500 ☐M cofactor (NAD(P) glycine, pH 8.8 (oxidation reaction) for 45 minutes at 37°C. The reduced cofactor (NAD(P)H) fluoresces under UV light. Active enzymes were judged based on comparison to a blank reaction, which contained the indicated substrate and cofactor without enzyme.
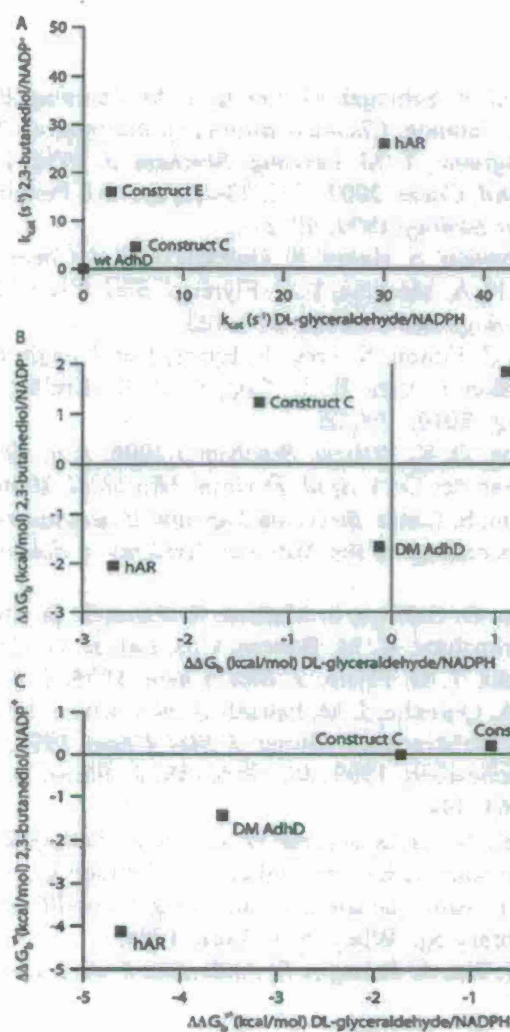
Figure 3 – Effect of Loop Insertions on Activity and Binding Energy. (A) Comparison of turnover rate for AdhD, hAR, and active loop chimeras with DL-glyceraldehyde/NADPH and 2,3-butanediol/NADP$^+$ as substrates. The value of k_{cat} was obtained by fits of kinetic data to the ordered bi-bi rate equation. (B) Change in the ground-state cofactor binding energy[a] of constructs with NADPH and NADP$^+$ relative to wt AdhD. (C) Change in the transition-state binding energies[b] of constructs with DL-glyceraldehyde/NADPH and 2,3-butanediol/NADP$^+$ relative to wt AdhD.

[a] $\Delta\Delta G = -RT\ln[(K_{ia})_{construct}/(K_{ia})_{wt\ AdhD}]$

[b] $\Delta\Delta G^{\ddagger} = RT\ln[(k_{cat}/K_A)_{construct}/(k_{cat}/K_A)_{wt\ AdhD}]$

References

[1]   a)J. M. Jez, M. J. Bennett, B. P. Schlegel, M. Lewis, T. M. Penning, *Biochemical Journal* **1997**, *326*, 625; b)J. M. Jez, T. M. Penning, *Chemico-Biological Interactions* **2001**, *130-132*, 499.

[2]   a)L. J. Askonas, J. W. Ricigliano, T. M. Penning, *Biochem. J.* **1991**, *278*, 835; b)W. C. Cooper, Y. Jin, T. M. Penning, *J. Biol. Chem.* **2007**, *282*, 33484; c)T. M. Penning, *The Journal of Steroid Biochemistry and Molecular Biology* **1999**, *69*, 211.

[3]   a)K. Bohren, J. Page, R. Shankar, S. Henry, K. Gabbay, *J. Biol. Chem.* **1991**, *266*, 24031; b)T. J. Kubiseski, D. J. Hyndman, N. A. Morjana, T. G. Flynn, *J. Biol. Chem.* **1992**, *267*, 6510.

[4]   E. M. Ellis, *FEMS Microbiology Letters* **2002**, *216*, 123.

[5]   J. Scoble, A. D. McAlister, Z. Fulton, S. Troy, E. Byres, J. P. Vivian, R. Brammananth, M. C. J. Wilce, J. Le Nours, L. Zaker-Tabrizi, R. L. Coppel, P. K. Crellin, J. Rossjohn, T. Beddoe, *Journal of Molecular Biology* **2010**, *398*, 26.

[6]   a)E. Di Luccio, R. A. Elling, D. K. Wilson, *Biochem J* **2006**, *400*, 105; b)R. Machielsen, A. R. Uria, S. W. M. Kengen, J. van der Oost, *Appl. Environ. Microbiol.* **2006**, *72*, 233.

[7]   E. Campbell, I. R. Wheeldon, S. Banta, *Biotechnology and Bioengineering* **2010**, *107*, 763.

[8]   H. Ma, T. M. Penning, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 11161.

[9]   a)B. Crosas, D. J. Hyndman, O. Gallego, S. Martras, X. Pares, T. G. Flynn, J. Farres, *Biochem. J.* **2003**, *373*, 973; b)C. E. Grimshaw, K. M. Bohren, C. J. Lai, K. H. Gabbay, *Biochemistry* **1995**, *34*, 14366; c)T. J. Kubiseski, T. G. Flynn, *J. Biol. Chem.* **1995**, *270*, 16911; d)I. Tarle, D. W. Borhani, D. K. Wilson, F. A. Quiocho, J. M. Petrash, *J. Biol. Chem.* **1993**, *268*, 25687.

[10]   a)M. Jackman, M. Parry, J. Hofsteenge, S. Stone, *J. Biol. Chem.* **1992**, *267*, 15375; b)K. Ratnam, H. Ma, T. M. Penning, *Biochemistry* **1999**, *38*, 7856; c)S. R. Stone, B. F. Le Bonniec, *Journal of Molecular Biology* **1997**, *265*, 344.

[11]   I. R. Wheeldon, E. Campbell, S. Banta, *Journal of Molecular Biology* **2009**, *392*, 129.

[12]   A. Fersht, *Enzyme structure and mechanism*, 2nd ed., W.H. Freeman, New York, **1985**.

[13]   I. H. Segel, *Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems*, Wiley Classics Library ed., Wiley, New York, **1993**.

[14]   S. Banta, B. A. Swanson, S. Wu, A. Jarnagin, S. Anderson, *Protein Eng.* **2002**, *15*, 131.

Chapter 10

Campbell, E., Meredith, M., Minteer, S.D., and **Banta, S.,** (2012) "An enzymatic biofuel cell utilizing a biomimetic cofactor" *Chemical Communications* **48**(13) 1898-1900.

Enzymatic biofuel cells utilizing a biomimetic cofactor†

Elliot Campbell,[a] Matthew Meredith,[b] Shelley D. Minteer,[b] and Scott Banta*[a]

[a] *Department of Chemical Engineering, Columbia University in the City of New York, New York, NY 10027, USA.*
E-mail: sbanta@columbia.edu; Fax: +1 212 854 3054;
Tel: +1 212 854 7531;
[b] *Department of Chemistry and Materials Science and Engineering, 315 S 1400 E Room 2020, Salt Lake City, UT 84112, USA.*
E-mail: minteer@chem.utah.edu

The performance of immobilized enzyme systems is often limited by cofactor diffusion and regeneration. Here, we demonstrate an engineered enzyme capable of utilizing the minimal cofactor nicotinamide mononucleotide (NMN^+) to address these limitations. Significant gains in performance are observed with NMN^+ in immobilized systems, despite a decreased turnover rate with the minimal cofactor.

Immobilization of enzymes in polymer films at electrode surfaces has been used extensively in biofuel cell and biosensor applications.[1] Compared to soluble enzymes, the use of immobilized enzymes reduces the amount of protein required and greatly increases the stability and lifetime of the system.[2] Nafion®, a perfluorosulfonated ion-exchange polymer, has been widely used, and modifying the Nafion® membrane with quaternary ammonium salts has been shown to increase mass transport through the film.[1, 3] However, cofactor diffusion has still been suggested to be the rate-limiting step in these systems.[4] In order to address this significant limitation, we have examined the use of an engineered enzyme and minimal cofactor analogs to improve performance through an increased diffusion rate.

Although non-natural biomimetic cofactors may possess some superior properties, they are generally poor substrates for wild type enzymes. For instance, the biomimetic cofactor NMN^+ is the electroactive half of the natural nicotinamide cofactors, yet there have been few reports of the use this truncated cofactor in place of $NAD(P)^+$ for catalysis (Fig. 1a and b).[5] The most notable work in this area involve a series of *N*-benzylnicotinamide derivatives and nicotinamide-5'-ribose methyl phosphate.[6] However, the observed activities with horse liver alcohol dehydrogenase were extremely low (≈ 28 d^{-1}), and the cofactor analogues were sensitive to oxidation. Later experiments with cytochrome P450s and 2-hydroxybiphenyl 3-monooxygenase (HbpA) yielded higher activities, and a few cofactor specificity mutants utilized the nicotinamide derivatives with better than wild-type activity.[7, 8]

Previously, we engineered a thermostable NAD(H)-dependent alcohol dehydrogenase from *Pyrococcus furiosus* (AdhD) for broadened cofactor specificity and improved activity by making two mutations in the cofactor binding pocket.[9] One, a histidine to arginine mutation (H255R), is positioned in a cleft distal to the active site where the adenine indol of the natural cofactor (and the 2'-phosphate of NADP(H)) binds, and is important in determining cofactor specificity.[10] The other, a lysine to glycine mutation (K249G), is located in the bottom of the cofactor binding pocket along the pyrophosphate backbone of the cofactor.[11] Elimination of this bulky side chain likely increases the conformational flexibility of the cofactor in the binding pocket, allowing for an increased turnover rate and broadened specificity.[10-12] Analysis of the two mutants supports this hypothesis, with the H255R mutation increasing activity with NADP(H), while the K249G mutation improves activity with both cofactors.[9] In this work, we subsequently discovered that the K249G/H255R double mutant was able to utilize the minimal cofactor nicotinamide mononucleotide (NMN(H)) for catalysis.

We hypothesize that the increased volume of the binding pocket of AdhD afforded by the K249G mutation allows the truncated cofactor to adopt a conformation that favors catalysis. This is supported by the fact that both the wild-type enzyme and the H255R single mutant possess low activity with NMN^+, while the K249G mutant exhibits over an order of magnitude increase in activity (Table S1, ESI†). Surprisingly, the K249G/H255R double mutant exhibits a higher affinity and a further 2-fold increase in activity with NMN^+. However, the activity of the double mutant with NMN^+ is still one to two orders of magnitude lower than the wild-type enzyme with its natural cofactor (Table S2, ESI†).

The affinity AdhD mutants for the truncated cofactor NMN^+ is clearly lower than for the

natural cofactors (Table S1, ESI†). Nicotinamide-dependent enzymes typically have a high specificity for either NAD(H) or NADP(H), which allows different enzymes to perform both reductions (using NADP(H)) and oxidations (using NAD(H)) simultaneously without spatial separation[13]. As NMN^+ lacks the specificity determining half of the molecule, cofactor affinity is significantly decreased. The reduced affinity is unlikely to have a large impact in immobilized applications, however, due to the high enzyme loadings and an increased local concentration of cofactor and substrate in the polymeric films.

The model substrate for AdhD (2,3-butanediol) was not compatible with the Nafion®/MG anode, so an alternate substrate (D-arabinose) that also displays rapid kinetics with the enzyme was used in the biofuel cell. A full kinetic analysis of the double mutant AdhD enzyme was performed using both substrates and both the native (NAD^+) and truncated (NMN^+) cofactors (Table S2, ESI†). In dilute solution, the enzyme exhibits a k_{cat} two to three orders of magnitude greater with NAD^+ than NMN^+ (15 s^{-1} with NAD^+ vs. 0.018 s^{-1} with NMN^+ for 2,3-butanediol; 65 s^{-1} with NAD^+ vs. 0.55 s^{-1} with NMN^+ for D-arabinose). The impact of the cofactor on the Michaelis constants varies unexpectedly with the substrate, however. When 2,3-butanediol is the substrate, the Michaelis constants for both the cofactor (K_A) and substrate (K_B) decrease (K_A from 460 μM with NAD^+ ... and K_B from 690mM with NAD^+ to 17mM with NMN^+), whereas the Michaelis constants increase when D-arabinose is the substrate (K_A from 480 μM with NAD^+ ... and K_B from 72mM with NAD^+ to 130mM with NMN^+). Thus, enzyme performance is expected to be much higher with NAD^+ as a cofactor, owing both to the two order of magnitude increase in k_{cat} and the significant decreases in the Michaelis constants.

The diffusion coefficients of NAD^+ and NMN^+ through modified Nafion® and their extraction coefficients into the polymer films were determined by cyclic voltammetry (CV) and rotating disc voltammetry (RDV). The biofuel cells described later in this study utilize poly(methylene green) (MG) as an electrocatalyst to oxidize the NADH or NMNH produced by the enzymes during operation. However, MG cannot catalyze the reverse reaction to reduce NAD^+, so a different electrocatalyst was needed to measure the transport properties of NAD^+ and NMN^+ through the films. Poly(neutral red) (PNR) has been shown to be an effective electrocatalyst for the two-electron reduction of NAD^+,[14, 15] and was used in this study to determine the rate at which NAD^+ and NMN^+ diffused through the film to the electrode surface. The diffusion coefficients and extraction coefficients of NAD^+ and NMN^+ are given in Table S3 (ESI†), as determined by Saveant[16] analysis of the RDE data (Figure S1, ESI†) as well as analysis of the variable scan rate CV experiments (Fig. 2a and b). The diffusion coefficient of NMN^+ through the modified Nafion® is faster than NAD^+ by an order of magnitude, likely due to the smaller size of NMN^+ relative to NAD^+ (Fig. 1a and b). NAD^+ exhibited a higher extraction coefficient into the films, possibly due to the more hydrophobic nature of NAD^+ relative to NMN^+. Overall, the flux of NMN^+ through modified Nafion® is higher than NAD^+, as expected for a smaller molecule.

To further study the effect of the minimal cofactor, a biofuel cell was constructed as described in the ESI.† A schematic of the anode is shown in Fig. 1c. Methylene green (MG), an electrocatalyst for cofactor oxidation, is polymerized onto the carbon paper electrode. The MG lowers the overpotential for cofactor oxidation by ~500 mV and produces a greater anodic current as compared to an unmodified electrode.[17, 18] Even though the double mutant AdhD enzyme had a much lower turnover rate with NMN^+ in dilute solution, biofuel cells using NMN(H) as a cofactor performed similarly to ones using NAD(H) (Fig. 2c, Table S4, ESI†). The

open circuit potential (OCP) for the NAD(H) biofuel cells was higher than those using NMN(H) (0.642 vs 0.593 V), while the maximum power densities of the biofuel cells using each cofactor were not statistically different (1.52 ± 0.27 vs 1.37 ± 0.24 respectively). Interestingly, the use of NMN(H) resulted in a 40% increase in maximum current density, which suggests an improvement in mass transfer for the truncated cofactor.

Analysis of the polarization curves with each cofactor provides insight into the processes affecting biofuel cell performance. The OCP of the fuel cell depends on both the formal potential of the cofactor at the electrocatalyst and the rate of accumulation of the reduced cofactor at the electrode surface. The formal potential differences combined with the increased turnover rate of the enzyme with the natural cofactor leads to a higher OCP in the NAD^+ cell. Conversely, the limiting current is predominantly a function of mass transfer within the system. Here, the order of magnitude increase in the diffusion coefficient of NMN^+ compared to NAD^+ results in a greater than 40% increase in the limiting current. As current is proportional to the number of electrons transferred to the electrode, it follows that the faster diffusing NMN(H) can shuttle more electrons between the enzyme and electrode surface per time than NAD(H). Lastly, the maximum power density is dependent on both the kinetic rates of cofactor reduction by the enzyme and cofactor oxidation by MG on the electrode, as well as mass transfer effects and ohmic losses within the system. Surprisingly, the maximum power densities observed with NAD(H) and NMN(H) were not statistically different. This suggests that the enzymatic turnover rate is not limiting in this system; rather performance is dominated by mass transfer effects or by the rate of oxidation of the reduced cofactors by MG.

The use of PNR to measure the diffusion coefficients of the oxidized cofactors through the modified-Nafion® may also provide insight into the relative rates of oxidation of the cofactors at MG. The rate of NMN^+ reduction at the PNR modified electrode was found to be much faster than for NAD^+ (Fig. 2a and b), as shown by the decreased ΔEp from 354 mV for NAD^+ to 229 mV for NMN^+. Given the structural similarities between MG and PNR, it is possible the rate of NMNH oxidation by MG may be higher than that of NADH. While the reasons for this are not clear, the presence of the adenine moiety in NAD(H) may sterically hinder the nicotinamide group from reaching the surface, and may also cause cofactor adsorption through interactions with MG. As this half of the cofactor is absent in NMN(H), the nicotinamide group may be able to more freely interact with the MG, thereby promoting charge transfer.

While there currently exists no framework or general rules for engineering enzymes to use non-natural cofactors, the changing of cofactor specificity between the two natural cofactors NAD(H) and NADP(H) has been extensively researched over the past two decades.[19, 20] Interestingly, the mutations identified in some enzymes to broaden or reverse cofactor specificity seem to improve activity with non-natural cofactors. This effect has been observed both in the cytochrome P450 studied by Ryan et al.[8] and the dehydrogenase (AdhD) examined in this study. Whereas these wild-type enzymes show little to no activity with the non-natural cofactors, some of the cofactor specificity mutants are able to use these truncated cofactors with activities approaching those of the wild-type enzymes with the natural cofactor. It is not unexpected that relaxing the cofactor specificity allows the enzymes to become more permissive in accepting non-natural cofactors, which depending on the rate-limiting step of the reaction, can lead to high levels of activity.

To our knowledge, this is the first reported use of a biomimetic cofactor for bioelectrocatalysis. Despite a significantly decreased enzyme turnover rate with the truncated cofactor, biofuel cell systems constructed with NMN^+ exhibited similar power densities and increased current

densities compared to the natural cofactor NAD^+. This suggests that both mass transport and cofactor oxidation at the electrode are much more important than the turnover rate of the enzyme. The present work opens new avenues of research involving electron relay systems, and will have important applications in many biocatalysis applications. Currently, the critical bottleneck in using alternative cofactors is the lack of enzymes engineered for altered cofactor specificity. In addition, further improvements to the biomimetic cofactors are also likely to improve system performance including stability, redox potential, turnover at the electrode surface, and cost.

Notes and references

1. N. L. Akers, C. M. Moore and S. D. Minteer, *Electrochimica Acta*, 2005, **50**, 2521-2525.
2. C. M. Moore, N. L. Akers, A. D. Hill, Z. C. Johnson and S. D. Minteer, *Biomacromolecules*, 2004, **5**, 1241-1247.
3. T. Klotzbach, M. Watt, Y. Ansari and S. D. Minteer, *Journal of Membrane Science*, 2006, **282**, 276-283.
4. C. M. Moore, S. D. Minteer and R. S. Martin, *Lab on a Chip*, 2005, **5**, 218-225.
5. R. L. Koder and A.-F. Miller, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1998, **1387**, 395-405.
6. H. C. Lo and R. H. Fish, *Angewandte Chemie International Edition*, 2002, **41**, 478-481.
7. J. Lutz, F. Hollmann, T. V. Ho, A. Schnyder, R. H. Fish and A. Schmid, *Journal of Organometallic Chemistry*, 2004, **689**, 4783-4790.
8. J. D. Ryan, R. H. Fish and D. S. Clark, *ChemBioChem*, 2008, **9**, 2579-2582.
9. E. Campbell, I. R. Wheeldon and S. Banta, *Biotechnology and Bioengineering*, 2010, **107**, 763-774.
10. S. Banta, B. A. Swanson, S. Wu, A. Jarnagin and S. Anderson, *Protein Eng.*, 2002, **15**, 131-140.
11. K. Bohren, J. Page, R. Shankar, S. Henry and K. Gabbay, *J. Biol. Chem.*, 1991, **266**, 24031-24037.
12. G. Sanli, S. Banta, S. Anderson and M. Blaber, *Protein Sci*, 2004, **13**, 504-512.
13. R. L. Veech, L. V. Eggleston and H. A. Krebs, *Biochem J*, 1969, **115**, 609-619.
14. A. A. Karyakin, O. A. Bobrova and E. E. Karyakina, *J. Electroanal. Chem.*, 1995, **399**, 179-184.
15. M. N. Arechederra, P. K. Addo and S. D. Minteer, *Electrochim. Acta*, 2011, **56**, 1585-1590.
16. J. Leddy, A. J. Bard, J. Maloy and J. Saveant, *J. Electroanal. Chem.*, 1985, **187**, 205-227.
17. Z. Dong-mei, F. Hui-Qun, C. Hong-yuan, J. Huang-xian and W. Yun, *Analytica Chimica Acta*, 1996, **329**, 41-48.
18. R. A. Rincón, K. Artyushkova, M. Mojica, M. N. Germain, S. D. Minteer and P. Atanassov, *Electroanalysis*, 2010, **22**, 799-806.
19. N. S. Scrutton, A. Berry and R. N. Perham, *Nature*, 1990, **343**, 38-43.
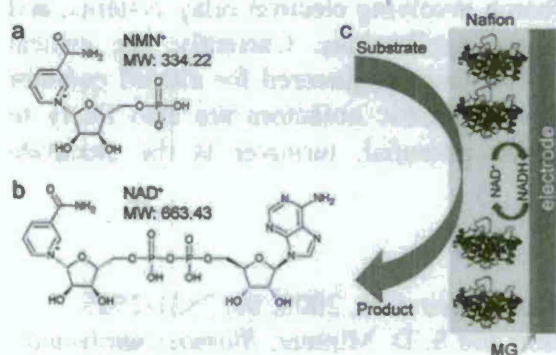20. T. M. Penning and J. M. Jez, *Chemical Reviews*, 2001, **101**, 3027-3046.

Fig. 1 (a) Structure of the biomimetic cofactor NMN$^+$ and (b) natural cofactor NAD$^+$. (c) Schematic of the bioanode. Methylene green mediator is polymerized on carbon paper electrode. Enzyme and cofactor are immobilized in TBAB-modified Nafion® on electrode surface. An air-breathing platinum cathode completes the fuel cell.
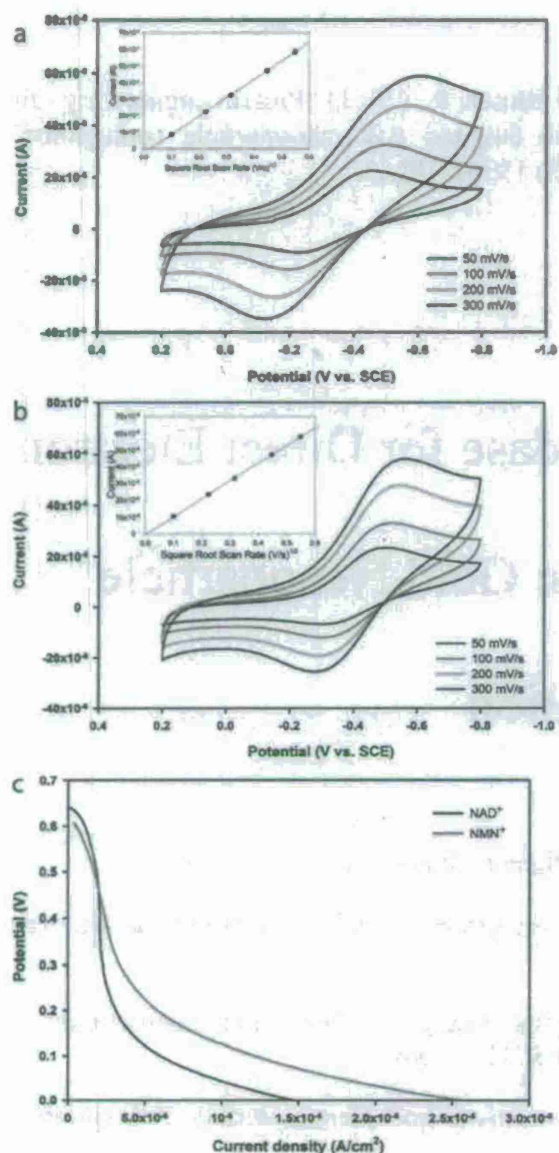
Fig. 2 Representative cyclic voltammograms of (a) NAD^+ and (b) NMN^+ at a PNR-modified GC electrode coated with modified Nafion® at a variety of scan rates. Conditions: quiescent solution, room temperature, 10 mM Tris-HCl, 10 mM KCl, 10 mM NAD^+/NMN^+, pH 7.0. Insets: Plot showing linear relationship between current and the square root of the scan rate. (c) Representative polarization curves of biofuel cells using either NAD^+ or NMN^+ as cofactors. Conditions: Quiescent solution, room temperature, 100 mM sodium phosphate, 100 mM $NaNO_3$, 50 mM arabinose, pH 8.0.

159

Chapter 11

Engineering of Glucose Oxidase for Direct Electron Transfer via Site-Specific Gold Nanoparticle Conjugation

J. Todd Holland[†,§,#], Carolin Lau[‡,#], Susan Brozik[§], Plamen Atanassov[‡*], Scott Banta[†*]

† Columbia University in the City of New York, Department of Chemical Engineering, New York, NY 10027, USA

‡ University of New Mexico, Center for Emerging Energy Technologies, Department of Chemical & Nuclear Engineering, Albuquerque, NM 87131, USA

§ Sandia National Laboratories, Department of Biosensors and Nanomaterials, Albuquerque, NM 87185, USA

* Corresponding Authors E-mail address: plamen@unm.edu, sbanta@columbia.edu
#These authors contributed equally to this publication

ABSTRACT: Optimizing the electrical communication between enzymes and electrodes is critical in the development of biosensors, enzymatic biofuel cells, and other bioelectrocatalytic applications. One approach to address this limitation is the attachment of redox mediators or relays to the enzymes. Here we report a simple genetic modification of a glucose oxidase enzyme to display a free thiol group near its active site. This facilitates the site-specific attachment of a maleimide-modified gold nanoparticle to the enzyme which enables direct electrical communication between the conjugated enzyme and an electrode. Glucose oxidase is of particular interest in biofuel cell and biosensor applications, and the approach of "pre-wiring" enzyme conjugates in a site-specific manor will be valuable in the continued development of these systems.

Electron transfer in enzymes generally occurs through metal centers or tunneling events that are largely insulated by the surrounding globular protein structure. There is great interest in developing devices and advancing applications that include an enzyme/electrode interface, but efficient electrical communciation between enzymes and electrodes is often hampered by this insulating effect.

Enzymatic biofuel cells are one area where efficient electrical contacts between electrodes and enzymes is critical[1]. Since glucose is ubiquitous and abundant in most living organisms, much of the research in enzymatic biofuel cells has focused on the use of glucose as the fuel source. The glucose-based enzymatic fuel cells developed to date have used either glucose oxidase $(GOx)^{2,3}$ or glucose dehydrogenase $(GDH)^4$ at the anode, while laccase, bilirubin oxidase, and other three copper oxidases have been utilized to reduce oxygen at the cathode.

Achieving direct electron transfer (DET) between a redox enzyme and an electrode is advantageous as it allows one to avoid the problems associated with the use of redox mediators, such as high cost, potential toxicity, and limited stability. Achieving DET depends significantly on the distance between the redox active cofactor and the electrode surface. Several recent publications have reviewed enzymatic DET processess and approaches.[5-11] There have been many reports, with varying degrees of success, to create or modify electrode materials that promote DET with GOx.[12-23] The challenge in the development of this approach is overcoming the long electron tunneling distance. In GOx, as with many redox proteins, the redox-active cofactor, flavin-adenine-dinucleotide (FAD),[24] is buried deeply within the protein core, rendering it inaccessible for direct communication with electrode surfaces. A typical way to overcome this is to add small, mobile redox mediators, which can diffuse into and out of the enzyme active site, ferrying reducing or oxidizing equivalents with them.[25] This approach can have several disadvantages including high cost, potential toxicity, and untethered mediators can diffuse away.

Attachment of such mediators to the enzymatic surface,[12,27-29] or to a surrounding redox polymer hydrogel,[30-32] can potentially solve the diffusion problem. The attachment of conductive nanoparticles (NPs) to the enzyme cofactor has also proven to be a successful strategy, with the resultant enzyme-NP complex reported to have a higher catalytic turnover rate than the unmodified system.[33] However, this approach involved complex chemical synthesis, and reconstitution of native enzymes around modifed co-factors, and none of the above approaches have proven to be amenable to scale-up and practical use.

We sought to bypass such difficulties by taking a protein engineering approach where we genetically modified the enzyme to make it more amenable to the simple and site-specific attachment of gold NPs. Instead of having to make extensive chemical modifications to the protein or its cofactor, or rely upon random placement of possible attachment points for the mediators, why not modify the gene, and let cheap and efficient biological expression systems produce "wire-ready" proteins? Here we report the production of a GOx mutant with a single free sulfhydryl group (cysteine) engineered onto its surface (Figure 1), and the attachment of a single maleimide-labeled gold nanoparticle to it as a redox relay (Figure 1, right).

The parent protein for this effort was a double mutant (T56V/T132S) of the *A. niger* GOx with improved catalytic properties created in the laboratory of Susan Brozik at Sandia Labs[34]. This protein has three native cysteines, where two are involved in a disulfide bond and the third is a free cysteine (Cys 521 in Fig 1). To prevent the attachment of gold nanoparticles to the native free thiol, this was mutated to valine, making the triple mutant (T56V/T132S/C521V) which was the starting point for this project. The C521V mutation did not result in a noticable decrease in

enzymatic activity as compared to the double mutant protein using a horseradish peroxidase (HRP) and ABTS activity assay which measures H_2O_2 production[35,36] (Supporting Table S1).

Site-directed mutagenesis was then used to create additional single mutations in the protein such that cysteine side chains are added at strategic locations on the surface of the protein near the FAD molecule. Five mutations were made at distances from the FAD cofactor ranging from 13.8 to 28.5 Å based on the crystal structure of the native GOx enzyme (Fig. 1, Table 1). The actual distances could be different due to differences in glycosylation of the recombinant enzyme. The activities of the mutant proteins were relatively unaffected by the mutations except for the A449C mutation which exhibited diminished activity. (Supporting Table S1).

Maleimide-modified gold nanoparticles (1.4 nm) were added to the purified mutant proteins (1:1 ratio) to react with the free thiol groups. The attachement of the gold nanoparticles to the purified proteins led to an immediate and dramatic decrease in activity for all of the mutants investigated, as measured by the HRP/ABTS assay. The activity of the A449C mutant was too low to be measured. The remaining 4 mutants exhibited residual activity ranging from 25 to 40% of the original activity and this was stable for a 16 hour time period (supporting information Figure S1).

These GOx-NP conjugates were then tested for DET capability by immobilizing them onto gold electrodes, and control experiments were performed with non-conjugated enzymes (both the mutants and a commercially available wild type GOx). By immobilizing only 5 μL of approximately 1 nM GOx-NP solution, a monolayer coverage of the gold electrode was assumed, which should eliminate electron transfer between excess conjugates in the system. The H447C mutant was estimated to have the shortest distance (Table 1) between the newly introduced cysteine group and the FAD center, making it the most promising candiate for DET. Indeed, only the H447C mutant exhibited DET activity. Therefore only this mutant was investigated further.

The apparent steady state kinetic parameters for the H447C mutant with and without gold NP conjugation were measured using the HRP/ABTS assay with saturating oxygen concentrations and these values were compared to those obtained using the commercially available GOx, also with and without gold NP conjugation (Table 2). The increased activity imparted by the parent double mutations in the H447C quadruple mutant is obvious when compared to the commercially available enzyme. The addition of the gold NPs decreased the overall apparent saturating enzymatic activity ($k_{cat,app}$) of both enzymes (as measued by H_2O_2 production). Interestingly, the addition of the gold NPs decreased the apparent K_M values in each case. Other researchers have observed unexpected decreases in K_M values when enzymes are conjugated with nanoparticles[37,38] and this could be due to subtle structural changes induced by the attachment of the particles. Further experiments were performed to verify this effect on the K_M values.

Hydrogen peroxide is the co-product of glucose oxidation in solution, and as an alternative to the ABTS assay, its concentration can be determined by measuring the current resulting from its direct electrochemical oxidization at 600mV (vs. Ag/AgCl) on gold electrodes. Thus the non-DET activity of the enzyme can also be measured electrochemically. The resulting amperometric response of hydrogen peroxide as co-product is shown in Figure 2. The apparent Michaelis constants ($K_{M,app}$) derived from these measurements were found to follow a similar trend to the solution experiments ($K_{M,app}$ for H447C = 155 ± 14 mM, and $K_{M,app}$ for H447C-AuNP = 75 ± 10 mM). Although the kinetic mechanisms of the ping pong enzymes can be

complex[39], these results suggest that the addition of the gold NPs increases the apparent affinity of the enzyme for the glucose substrate.

The GOx enzyme catalyzes the oxidation of glucose to gluconolactone via reduction of the FAD co-factor to $FADH_2$. The re-oxidation of $FADH_2$ in the ping pong mechanism is normally achieved using oxygen as the electron acceptor. Therefore the competing direct oxidation of $FADH_2$ on an electrode either has to be performed under oxygen free conditions, or the rate of re-oxidation of $FADH_2$ by the electrode (the actual Direct Electron Transfer (DET) reaction) has to be higher than that of the competing oxygen reduction to peroxide reaction (the native electron acceptor process for GOx). All of the GOx mutants and their Au-NP modified analogs listed in Table 1 were tested via cyclic voltammetry (CV) in nitrogen saturated buffer for DET. Only the H447C-AuNP conjugate, exhibited a significant oxidative current starting at -400 mV upon addition of glucose (Figure 3). The steep increase in the oxidative current is due to enzymatically catalyzed glucose oxidation. The starting potential of -400 mV indicates a direct contact between the electrode and the FAD center, which has a formal potential of -460 mV (vs. Ag/AgCl) at pH 7.9. In control experiments (Figure 3) the same mutant without the Au-NP modification showed no apparent DET upon glucose addition. The observed onset of a reductive process for the unmodified enzyme may have several interpretations such as the onset of an oxygen reduction reaction at the electrode or a hydrogen peroxide reduction reaction. Since only the modified enzyme exhibits a DET signal, this leads to the conclusion that electrical contact between the electrode and FAD center is made through the conjugated gold nanoparticle.

A very similar approach of bridging the FAD center wth an electrode via gold NPs was reported by Willner et al and this was discussed as an electron-mediating 2-electron "relay"[40,41]. The major difference in the approach presented here is that the gold NPs are designed to be attached to the surface of the protein near the FAD center, which should decrease the electron transfer distance. Overall we observed similar results; a catalytic glucose oxidation current without the appearnace of the FAD/FADH2 redox wave. Even though the enzymatic oxidation begins at ~-400mV, its shape is tilted towards more anodic potentials. Both effects can be explained by an IR-drop caused by the Au-NP bridge between the enzyme and electrode which is able to serve as and "electron relay" while adding additional resistance to the system.

The fact that the H447C mutant was the only one identified in this study to successfully promote DET underlines the importance of the electron transfer distance. Of the six residues targeted for site-directed replacement with cysteine in this work, H447C is the closest to the FAD in crystal structures (Table 2). The modern Marcus theory[42,43] has been applied to proteins and correctly predicts the exponential decrease in electron transfer rate, k_{ET}, with the distance of the electron transfer, d: $k_{ET} = k_0 \exp[-\beta(d-d_0)]$ (with k_0 the electron transfer rate constant at the distance of closest contact d_0 and the pre-exponential factor β typically in the range 8.5-11.5 nm^{-1})[44]. From the location of the mutation and from the size of the gold NPs used in this study, we can hypothesize that the distances inferred above have been met as evidenced by the voltammetric response to glucose (Figure. 3).

In previous reports, CVs presenting DET of GOx typically show a pair of reversible redox peaks corresponding to the surface bound prosthetic group FAD.[12,14,16-19,21-23] The lack of those peaks in the present work can be hypothesized as a Au-NP bridged contact of the redox center without blocking the active center from substrate access.

The CVs in Figure 3 show an additional reduction process starting at about - 400 mV. Most likely, small amounts of hydrogen peroxide are being produced by the enzyme due to the

presence of some oxygen in the system, and the peroxide is being directly oxidized on the gold electrode. The existence of both catalytic reactions is also supported by potentiometric data. Figure 4 shows the dependence of the open circuit potential (OCP) as a function of increasing glucose concentrations. It clearly shows a negative shift in the redox potential in presence of glucose, reaching a steady state potential of -370 mV at a glucose concentration of about 1 M, which is close to the theoretical redox potential of FAD. According to the Nernst equation, a pure catalysis by $FAD/FADH_2$ should show OCPs without glucose starting at negative potentials, decreasing 60 mV per concentration decade and stabilizing at the formal potential of -450 mV for $FAD/ FADH_2$. Since the decrease in OCP with glucose concentration does not follow perfect Nernstian behavior, the existence of a competing hydrogen peroxide reaction is feasible (Supporting information Figure S2). It should also be noted that there is no direct evidence that all enzymes in the samples are modified with a single Au-NP, so the presense of some unmodified enzymes can also explain the observation of such behavour. However, the closeness of the OCP to the redox potential of $FAD/FADH_2$ is, on the other hand, a testimony to the fact that the majority of the sample has been modified with Au-NPs.

The GOx enzyme has been extensively investigated for use in biosensors and enzymatic biofuel cells. A potentiostatically obtained polarization curve (Figure 5) of the H447C-AuNP conjugate GOx electrode shows the benefits of the genetically engineered GOx wired via gold nanoparticles for use in enzymatic biofuel cell applications.

This study presents the successful site-specific modifictaion of the GOx enzyme with AuNP via protein engineering. The attachement of the NP reduces the apparent catalytic activity of the enzyme while also decreasing the apparent Michaelis constant of the enzyme for its substrate. When immobilized on an electrode, attachment of the AuNP enables direct electrical communciation across the enzyme/electrode interface. This work demonstrates the achievements that can be made in the engineering of proteins for improvments in the interface between biotechnology and nanotechnology. The new mutant enzyme described here holds great promise for use in 3rd generation amperometric biosensors (based on DET) or at the anode of micro-bio-fuel cells due to the tight electrical contact formed between it and the electrode surface after conjugating it with off-the-shelf gold nanoparticles.

Supporting Information. Detailed experimental description of GOx protein engineering with kinetic evaluation of the different mutants is included in the supporting material as well as an evaluation of the potentiometrically obtained data according to the Nernst equation. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

(1)   Calabrese Barton, S.; Gallaway, J.; Atanassov, P. *Chem. Rev.* **2004**, *104*, 4867-4886.

(2)   Yan, Y. M.; Su, L.; Mao, L. Q. *Journal of Nanoscience and Nanotechnology* **2007**, *7*, 1625-1630.

(3)   Barriere, F.; Kavanagh, P.; Leech, D. *Electrochimica Acta* **2006**, *51*, 5187-5192.

(4)   Okuda-Shimazaki, J.; Kakehi, N.; Yamazaki, T.; Tomiyama, M.; Sode, K. *Biotechnology Letters* **2008**, *30*, 1753-1758.

(5)   Ghindilis, A. L.; Atanasov, P.; Wilkins, E. *Electroanal.* **1997**, *9*, 661-674.

(6)   Varfolomeev, S. D.; Kurochkin, I. N.; Yaropolov, A. I. *Biosens. Bioelectron.* **1996**, *11*, 863-871.

(7)   Ikeda, T. In *Frontiers in Biosensorics I; Fundamental Aspects*; Scheller, F. W., Schubert, F., Fedrowitz, J., Eds.; Birkhaeuser: Basel/ Switzerland, 1997; Vol. 80, p 243-266.

(8)   Armstrong, F. A.; Hill, H. A. O.; Walton, N. J. *Acc. Chem. Res* **1988**, *21*, 407-413.

(9)   Hill, H. A. O.; Hunt, N. I. In *Methods in Enzymology*; Academic Press: 1993; Vol. 227, p 501-522.

(10)   Hill, H. A. O.; Higgins, I. J. *Phil. Trans. R. Soc. A* **1981**, *302*, 267-73.

(11)   Cooney, M. J.; Lau, C.; Windmeisser, M.; Liaw, B. Y.; Klotzbach, T.; Minteer, S. D. *J. Mat. Chem.* **2008**, *18*, 667-674.

(12)   De Taxis Du Poet, P.; Miyamoto, S.; Murakami, T.; Kimura, J.; Karube, I. *Anal. Chim. Acta* **1990**, *235*, 255-263.

(13)   Koopal, C. G. J.; de Ruiter, B.; Nolte, R. J. M. *J Chem Soc Chem Comm* **1991**, *23*, 1691 - 1692.

(14)   Chi, Q.; Zhang, J.; Dong, S.; Wang, E. *Electrochim. Acta* **1994**, *39*, 2431-2438.

(15)   Jiang, L.; Mcneil, C. J.; Cooper, J. M. *J Chem Soc Chem Comm* **1995**, *12*, 1293-1295.

(16)   Guiseppi-Elie, A.; Lei, C.; Baughman, R., H. *Nanotechnology* **2002**, *13*, 559-564.

(17)   Liang, W.; Zhuobin, Y. *Sensors* **2003**, *3*, 544-554.

(18)   Cai, C.; Chen, J. *Anal. Biochem.* **2004**, *332*, 75-83.

(19)   Liu, J.; Chou, A.; Rahmat, W.; Paddon-Row, M. N.; Gooding, J. J. *Electroanal.* **2005**, *17*, 38-46.

(20) Liu, Y.; Wang, M.; Zhao, F.; Xu, Z.; Dong, S. *Biosens. Bioelectron.* **2005**, *21*, 984-988.

(21) Ivnitski, D.; Branch, B.; Atanassov, P.; Apblett, C. *Electrochem. Comm.* **2006**, *8*, 1204-1210.

(22) Zhang, J.; Feng, M.; Tachikawa, H. *Biosens. Bioelectron.* **2007**, *22*, 3036-3041.

(23) Wu, P.; Shao, Q.; Hu, Y.; Jin, J.; Yin, Y.; Zhang, H.; Cai, C. *Electrochimica Acta* **2010**, *55*, 8606-8614.

(24) Wilson, R.; Turner, A. P. F. *Biosensors and Bioelectronics* **1992**, *7*, 165-185.

(25) Frew, J. E.; Hill, H. A. O. *E. J. Biochem.* **1988**, *172*, 261-269.

(26) Jmol; an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/.

(27) Degani, Y.; Heller, A. *Journal of Physical Chemistry* **1987**, *91*, 1285-1289.

(28) Battaglini, F.; Bartlett, P. N.; Wang, J. H. *Anal. Chem.* **2000**, *72*, 502-509.

(29) Schuhmann, W. *Biosensors & Bioelectronics* **1995**, *10*, 181-193.

(30) Habermuller, K.; Mosbach, M.; Schuhmann, W. *Fres. J. Anal. Chem.* **2000**, *366*, 560-8.

(31) Mao, F.; Mano, N.; Heller, A. *JACS* **2003**, *125*, 4951-4957.

(32) Barlett, P. N.; Cooper, J. M. *Journal of Electroanalytical Chemistry* **1993**, *362*, 1-12.

(33) Xiao, Y.; Patolsky, F.; Katz, E.; Hainfeld, J. F.; Willner, I. *Science* **2003**, *299*, 1877.

(34) Holland, J. T.; Harper, J. C.; Dolan, P. L.; Manginell, M. M.; Arango, D. C.; Rawlings, J. A.; Apblett, C. A.; Brozik, S. M. *PLoS ONE*, (Submitted).

(35) Baron, A. J.; Stevens, C.; Wilmot, C.; Seneviratne, K. D.; Blakeley, V.; Dooley, D. M.; Phillips, S. E. V.; Knowles, P. F.; Mcpherson, M. J. *Journal of Biological Chemistry* **1994**, *269*, 25095-25105.

(36) Betancor, L.; Fuentes, M.; Dellamora- Ortiz, G.; Lopez-Gallego, F.; Hidalgo, A.; Alonso-Morales, N.; Mateo, C.; Guisan, J. M.; Fernandez-Lafuente, R. *Journal of Molecular Catalysis B-Enzymatic* **2005**, *32*, 97-101.

(37) Pandey, P.; Singh, S. P.; Arya, S. K.; Gupta, V.; Datta, M.; Singh, S.; Malhotra, B. D. *Langmuir* **2007**, *23*, 3333-7.

(38) Keighron, J. D.; Keating, C. D. *Langmuir* **2010**, *26*, 18992-9000.

(39) Glykys, D. J.; Banta, S. *Biotechnol Bioeng* **2009**, *102*, 1624-35.

(40)   Katz, E.; Sheeney-Haj-Ichia, L.; Willner, I. *Angew Chem Int Ed Engl* **2004**, *43*, 3292-300.

(41)   Xiao, Y.; Patolsky, F.; Katz, E.; Hainfeld, J. F.; Willner, I. *Science* **2003**, *299*, 1877-81.

(42)   Marcus, R. A.; Sutin, N. *Biochim. Biophys. Acta* **1985**, *811*, 265-322.

(43)   Marcus, R. A.; Sutin, N. *Inorg. Chem.* **1975**, *14*, 213-16.

(44)   Bartlett, P. N. *Bioelectrochemistry : fundamentals, experimental techniques and applications*; John Wiley & Sons: Chichester, England; Hoboken, NJ, 2008.
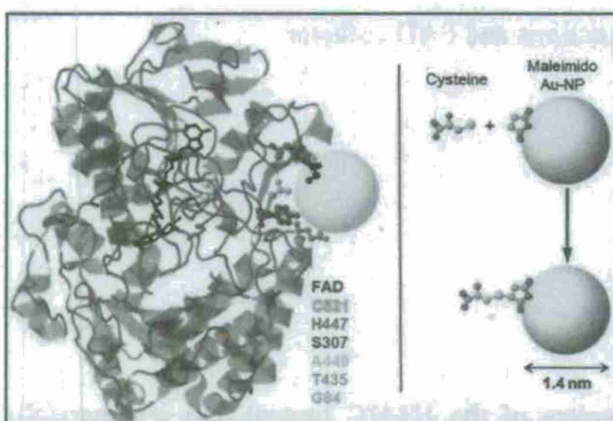
Figure 1: (left) A ribbon diagram of GOx monomer (from *A. niger*)[26], with the FAD molecule shown in blue. The amino acid residues targeted for mutagenesis are highlighted as space-filling models: cysteine (yellow), histidine (red), serine (purple), alanine (orange), tyrosine (pink) and glutamine (light blue). The yellow sphere represents an idealized gold nanoparticle on the same scale as GOx. **(right)** Schematic drawing of the covalent binding chemistry of cysteine to a maleimido modified gold nanoparticle. The molecules are displayed as ball-and-stick, with carbon (grey), oxygen (red), nitrogen (blue) and sulfur (yellow).

Table 1: Estimated distance between cysteine mutations and FAD cofactor

Surface Cysteine	Distance between original amino acid
H447C	13.8 Å
E84C	15.3 Å
A449C	18.6 Å
Y435C	22.2 Å
S307C	28.5 Å

Table 2: Apparent steady state kinetic parameters of the H447C mutant and commercially available GOx before and after covalent attachment of gold NPs.

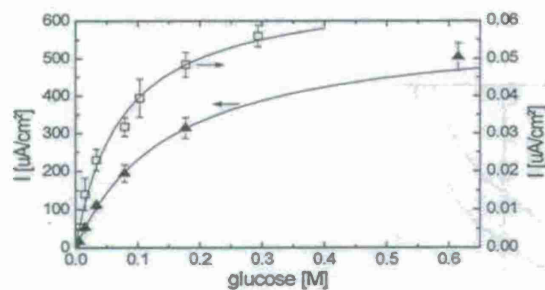|	$k_{cat, app}$ (s⁻¹)	$K_{M,app}$ (mM)	
H447C	425 ± 50	15.0	±
GOx	152 ± 11	96.4	±
H447C-Au	55.3 ± 8.2		±
GOx-Au	22.8 ± 6.3		±

Figure 2: Amperometric response of (▲ **left axis**) H447C mutant and (□ **right axis**) H447C-AuNP mutant with attached Au nanoparticles to glucose oxidation under O_2 saturation, applied potential +0.6V.
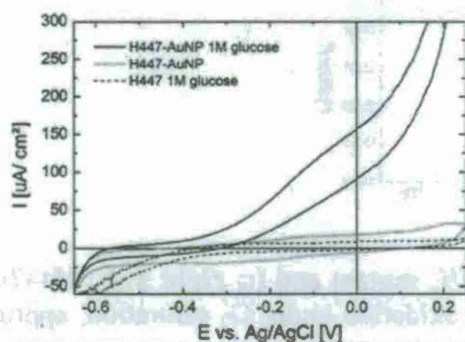
Figure 3: Cyclic voltammogram of H447C-Au NP conjugate gold electrodes in the presence (black line) and absence (grey line) of 1M glucose in N_2 saturated buffer. Unconjugated H44C is shown as a dotted line. The CV with the conjugated enzymes in the presence of glucose (black line) demonstrates the enzymatic glucose oxidation, starting at app. -400mV. 10mV/s, N_2 saturated 0.1M PB, pH 7.
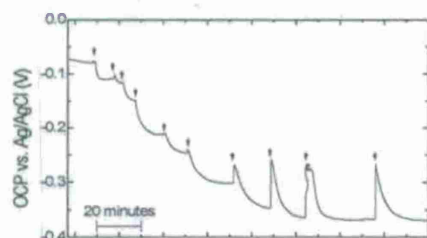
Figure 4: Potentiostatic response of the open circuit potential over time of the H447C-Au NP conjugate under addition of glucose (arrows), in N_2 saturated phosphate buffer (E vs. log concentration can be found in supporting Figure S3).
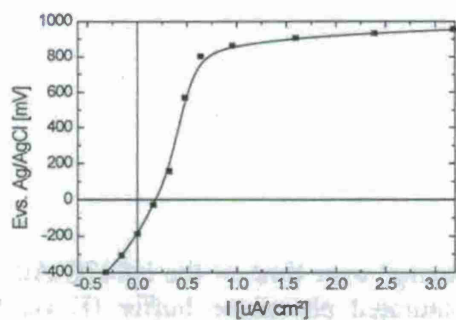
Figure 5: Potentiostatic polarization curve of H447C-Au NP conjugates on a gold electrode with 1M glucose, N$_2$ saturated phosphate buffer. The current density was calculated by taking the geometric surface area of the gold electrode.

Chapter 1

Glykys, D.J., Szilvay, G.R., Tortosa, P., Suarez Diez, M., Jaramillo, A., and **Banta, S.** (2011) "Pushing the limits of automatic computational protein design: computational design, expression, and characterization of a large synthetic protein based on a fungal laccase scaffold" *Systems and Synthetic Biology Systems and Synthetic Biology* 5(1-2) 45-58.

Pushing the Limits of Automatic Computational Protein Design: Design, Expression, and Characterization of a Large Synthetic Protein based on a Fungal Laccase Scaffold

Doris J. Glykys[1], Géza R. Szilvay[1], Pablo Tortosa[2], María Suárez Diez[2], Alfonso Jaramillo[2] and Scott Banta[1]

[1]Department of Chemical Engineering, Columbia University in the City of New York, 500 W 120th St, New York, NY 10027

[2] ISSB, Genopole-Université d'Évry Val d'Essonne-CNRS UPS3201, Batiment Geneavenir 6. 5, rue Henri Desbruères. 91030 Evry Cedex, France

Correspondence to: Scott Banta

Department of Chemical Engineering

Columbia University in the City of New York

820 S. W. Mudd, MC 4721

500 W 120th Street

New York, NY 10027

Telephone: (212) 854-7531

Fax: (212) 854-3054

Email: sbanta@columbia.edu

ABSTRACT

The *de novo* engineering of new proteins will allow the design of complex systems in synthetic biology. But the design of large proteins is very challenging due to the large combinatorial sequence space to be explored and the lack of a suitable selection system to guide the evolution and optimization. One way to approach this challenge is to use computational design methods based on the current crystallographic data and on molecular mechanics. We have used a laccase protein fold as a scaffold to design a new protein sequence that would adopt a 3D conformation in solution similar to a wild-type protein, the *Trametes versicolor* (TvL) fungal laccase. Laccases are multi-copper oxidases that find utility in a variety of industrial applications. The laccases with highest activity and redox potential are generally secreted fungal glycoproteins. Prokaryotic laccases have been identified with some desirable features, but they often exhibit low redox potentials. The designed sequence (DLac) shares a 50% sequence identity to the original TvL protein. The new DLac gene was overexpressed in *E. coli* and the majority of the protein was found in inclusion bodies. Both soluble protein and refolded insoluble protein were purified, and their identity was verified by mass spectrometry. Neither protein exhibited the characteristic T1 copper absorbance, neither bound copper by atomic absorption, and neither was active using a variety of laccase substrates over a range of pH values. Circular dichroism spectroscopy studies suggest that the DLac protein adopts a molten globule structure that is similar to the denatured and refolded native fungal TvL protein, which is significantly different from the natively secreted fungal protein. Taken together, these results indicate that the computationally designed DLac expressed in *E. coli* is unable to utilize the same folding pathway that is used in the expression of the parent TvL protein or the prokaryotic laccases. This sequence can be used going forward to help elucidate the sequence requirements needed for prokaryotic multi-copper oxidase expression.

Keywords: laccase, multi-copper oxidase, computational protein design, protein folding, molten globule

INTRODUCTION

One of the most difficult challenges in synthetic biology is the *de novo* design of proteins with targeted structures and functionalities. This requires the production of a novel amino acid sequence, preventing in most cases the use of experimental methods such as directed evolution to find such sequence. This is due to the fact that the combinatorial space is often too large and efficient evolution strategies require some fitness to start with. We have explored the possibility of using computational methods for this purpose instead. This enables the exploration of a large combinatorial space thanks to smart pair-wise algorithms that allow more efficient searching of the sequence space and this is facilitate by the use of distributed computing. In this manuscript we report this approach with a much larger system than has previously been reported in the literature, the laccase fold.

Laccases are multi-copper oxidase enzymes that catalyze the oxidation of a wide range of substrates including phenolics, polyphenolics, aromatic amines, and other compounds with the concomitant reduction of dioxygen to water (Kunamneni et al. 2008; Mayer and Staples 2002; Nakamura and Go 2005; Riva 2006; Sakurai and Kataoka 2007; Solomon et al. 1996). They have many important industrial applications including: textile and fabric bleaching, lignin degradation for paper production, pollution detoxification, wine clarification, organic chemical synthesis, biosensing and the cathodic reaction in enzymatic biofuel cells.

The most commonly used laccase enzymes for biofuel cells and other industrial applications are the lignolytic glycoproteins secreted by white-rot fungi such as *Trametes versicolor* (Piontek et al. 2002). The *Trametes versicolor* laccase (TvL) is a glycosylated monomer with 499 amino acids and a molecular mass of ~70 kDa. Like other laccases, the protein contains two copper-containing active centers. The first contains a T1, or blue copper center, where two histidines and a cysteine serve as ligands for the copper and this active site is where the oxidation of the substrate takes place. The second copper center is a trinuclear T2/T3 site where eight histidine side chains serve as ligands for a T2 copper and a pair of T3 coppers and in this site the reduction of dioxygen to water occurs. Interestingly, the T2/T3 site is buried deeply in the protein, and is accessible via a tunnel that opens on both sides of the enzyme. Besides the cysteine side chain that coordinates the T1 copper, there are four other cysteines in the sequence that form two disulfide bonds. The enzyme follows the bi-bi ping-pong kinetic mechanism, and it is most active at acidic pH values (Xu 1997; Xu et al. 1996).

TvL and other fungal laccase enzymes tend to have high redox potentials which allow them to oxidize a wide range of substrates, and in a biofuel cell, the high redox potentials are desirable as they lead to high operating voltages (Barton et al. 2004; Glykys and Banta 2009; Shleev et al. 2005). Some of the fungal enzymes have been expressed recombinantly in yeast, and there are a few reports where protein engineering has been performed to attempt to improve their expression, redox potential, and activity (Gelo-Pujic et al. 1999; Madzak et al. 2006; Rodgers et al. 2010; Xu et al. 1998; Xu et al. 1999). But, in published applications involving laccases, the native enzymes from the fungal hosts are most often utilized.

In addition to the fungal enzymes, there are many prokaryotic laccase and laccase-like proteins, and the native functions of these proteins are less well-understood (Alexandre and Zhulin 2000; Claus 2003). There has been a good deal of success in cloning and expressing these enzymes recombinantly in *E. coli*, which means that copper-containing active sites can be

formed and properly folded in prokaryotic hosts. These enzymes could have significant advantages over the fungal enzymes as they could be readily produced economically on a large scale, they are not glycosylated, and they often are active over a range of pH values including neutral pH. In addition, bacterial expression can simplify protein engineering techniques such as directed evolution (Brissos et al. 2009; Bulter et al. 2003; Festa et al. 2008; Gupta and Farinas 2009).

Unfortunately, one of the hallmarks of the prokaryotic laccases is that they tend to exhibit low redox potentials (Endo et al. 2003; Li et al. 2007; Martins et al. 2002). This limits their use in industrial applications, and can significantly reduce their utility in the creation of enzymatic biofuel cells as this leads to a reduction in voltage and therefore, power. For example, a bacterially-derived laccase enzyme (SLAC from *Streptomyces coelicolor*) (Machczynski et al. 2004; Skalova et al. 2009) expressed in *E. coli* has recently been incorporated into an osmium-mediated biofuel cell cathode (Gallaway et al. 2008; Wheeldon et al. 2008). Under acidic conditions its performance was reduced compared to the same cathode made with the TvL enzyme. But, at neutral pH, the highest current density for an enzyme mediated cathode was reported (Gallaway et al. 2008). In order to address this limitation of the prokaryotic laccases, other researchers are working to use protein engineering to attempt to increase the redox potential of the prokaryotic laccase-like proteins to improve their performance (Brissos et al. 2009).

It is likely that other researchers have tried to express the TvL glycoprotein in *E. coli*, but there has yet to be a report of the successful expression of this enzyme in a prokaryotic host (Kunamneni et al. 2008). There may be many reasons for this lack of success, including problems with stability, folding, or solubility of the protein in the prokaryotic host. Such expression requires, among other things, to ensure a proper stability, folding and solubility. One way to achieve such specifications is by using computational design, where automatic methods are used to explore the sequence space to find the optimal solution given an objective function. An objective function ensuring stability and solubility is given by the folding free energy. This requires using a structure template for the folded structure, which could be accomplished by choosing a tertiary structure with the appropriate fold. By picking an existing structure we increase the probabilities of maintaining the folding pathways, although there is currently no general methodology to account for it. However, since the fungal laccase structures are very similar to the prokaryotic laccases, we reasoned that a protein with an identical structure to the TvL protein should be expressible in *E. coli*, in an active form, if properly designed. One way to explore such specifications is by using computational design, where automatic methods are used to explore the sequence space to find the optimal solution given an objective function. An objective function ensuring stability and solubility is given by the folding free energy. This requires using a structure template for the folded structure, which could be accomplished by choosing a tertiary structure with the appropriate fold. By picking an existing structure we increase the probabilities of maintaining the folding pathways, although there is currently no general methodology to account for this.

Here we report the use of a computational approach to attempt to redesign the TvL sequence for increased stability and for expression in a prokaryotic host. The computational approach begins with the TvL protein backbone structure, and new amino acid side chains are identified with the goal of stabilizing this scaffold, thus forming a new protein with the same 3-D

structure as the glycosylated TvL. This methodology has been demonstrated through the redesign of 45 proteins which were successfully compared with their corresponding natural sequences.(Jaramillo et al. 2002) The new TvL-based sequence constitutes, to our knowledge, the largest computational design projects ever attempted with a sequence space of 10^{341} possible states. This size is too large to only use the current pair-wise optimization procedures and it required the use of extensive heuristics.

The newly designed laccase (DLac) gene was synthesized and the gene product was overexpressed in *E. coli* and the majority of the protein was found in inclusion bodies (IBs). Both solubly expressed and refolded proteins were purified and characterized, and the protein appears to be trapped in a molten globule structure that does not bind copper and is inactive under the experimental conditions tested. This molten globule structure seems to be very similar to what is observed when native TvL is chemically denatured and refolded, which suggests that proper folding of the TvL (and DLac) backbone requires a folding pathway that is not enabled by the sequence, or that may not be available in *E. coli*. This new gene will be a valuable starting point for unraveling the sequence requirements for successful laccase folding in prokaryotic hosts.

MATERIALS AND METHODS

Computational Design Considerations

Computational protein design was used to create a new primary sequence that would adopt the same structure as the secreted TvL protein. Ideally this would result in a protein product with the kinetic activity and redox potential of the fungal parent enzyme with the ability to be expressed in a highly stable and active form in a prokaryotic host. TvL was chosen as the starting structure due to its high redox potential (+780 mV vs NHE) and the availability of a 1.90 Å resolution crystal structure (PDB code 1GYC) (Piontek et al. 2002). The structure of the oxidized form of the laccase contains the full complement of coppers needed for its catalytic activity.

Copper atoms are critical for the catalytic function of the enzyme, and the native TvL enzyme has a tunnel through the protein to allow the transport of oxygen and water to and from the trinuclear T2/T3 site active site. Since these features are crucial for the function of the enzyme, the amino acids at the 145 positions within a 12 Å radius of the four copper atoms were fixed to have the same amino acid identity as in the parent sequence. Since proline and cysteine residues can have large impacts on the overall structure of the protein, those residues were also left unavailable for mutation. To add further constraints, the 51 amino acids within 8 Å from the cooper atoms and those that formed tight interactions on the surface of the structure (37 positions) were left unmutated and were also forced to have the same structural conformation as observed in the crystal structure. The 8 Å radius was chosen since the Charmm energy force field has its parameters optimized to use a cutoff of 7 Å for the non-bonded interactions and a 1.4 Å solvation probe. Therefore, 8 Å is the minimum cut-off around the copper atoms that would not change the active site energetic. Outside of these overlapping regions, all of the non-proline and non-cysteine residues were allowed to be substituted to new amino acids with variable conformations. From the original 499 amino acids in the crystal structure, there were 57 positions where the identity was retained but conformational changes were allowed, and a total

of 339 positions (designed positions) where mutations were allowed which leads to a combinatorial space of $20^{339} \approx 10^{341}$ possible sequences for the protein design (Fig. 1).

Computational Protein Design Procedure

The computational protein design software, DESIGNER, has previously been described and validated (Jaramillo et al. 2002; Ogata et al. 2003; Wernisch et al. 2000). The DESIGNER algorithm approaches computational protein design through the inverse folding problem. It begins with a high resolution protein target structure, and the available sequence space is searched to find new amino acid combinations that have optimized intramolecular interactions and will thus cause the protein to adopt the same fold as the desired targeted structure. For each analyzed sequence, atomic models of the folded and unfolded states are constructed using a molecular mechanics force field, CHARMM22, and a rotamer library. The different computed structures are compared to the final desired structures, and they are scored by their folding free energy which is evaluated as the difference between the free energy in the unfolded and folded states.

The free energy of each state contains contributions arising from electrostatic and van der Waals interactions. The electrostatic interaction is represented by a Coulomb term with dielectric constant $\varepsilon=8$. The standard Lennard-Jones potential is used to describe van der Waals interactions. An additional term is included that is linear with the solvent accessible area to account for the solvation energy. Energy of a given state is thus computed as the sum of atomic pair-wise contributions (Eq. 1).

$$G = \sum_{i \neq j} \frac{a_{ij}}{r_{ij}^{12}} - \frac{b_{ij}}{r_{ij}^{6}} + \sum_{i \neq j} \frac{q_i q_j}{\varepsilon r_{ij}} + \sum_i \sigma_i ASA_i$$

Eq. 1

The atomic coefficients are either those of the CHARMM22 force field (a_{ij}, b_{ij}, q_i) or obtained from experimentally measured hydration coefficients for small molecules (σ_i) (Ooi et al. 1987). The unfolded free energy is computed by evaluating the free energy of an ensemble of di-peptides with a composition given by the primary sequence. To rank the sequences, contributions common to all the sequences, like the energetic terms due to the fixed side chains or from the backbone, are neither considered nor computed.

Rotamer Library Refinement

The side chain conformations in the folded state were initially described using all rotamers with a probability higher than 90% from Dunbrack's backbone dependent library (Dunbrack and Karplus 1993). To enlarge this library, it was further expanded to include rotamers with higher interaction energies. For each pair of rotamers on the surface, a local minimization of their structure in a low-dielectric environment ($\varepsilon=1$) was performed, so that the search is biased towards H-bond formation (backbone charges are set to zero to further encourage interaction among the rotamers) (Tortosa and Jaramillo 2006). Only pairs with interaction energies lower than -5 kcal/mol were kept. After a new local minimization in the $\varepsilon=8$ environment, a new refined conformation for each residue of the pair was obtained. To reduce

180

the overwhelming number of rotamers obtained, a final coarse-grained clustering of the conformations was performed and only rotamers differing by more than $25°$ were stored for further use in the energy computation. A total of 48,150 rotamers were finally obtained.

Combinatorial Optimization of Protein Sequence

A standard Monte Carlo simulated annealing (MCSA) optimization algorithm was utilized and Metropolis was employed as a criterion to accept or reject solutions. The temperature followed an exponential cooling scheme, so that at each of the N iterations decreased from the original value $T_0/R = 1$ kcal/mol to the final value $T_f/R = 0.01$ kcal/mol, according to $T \to \alpha T_0$. Due to the size of the combinatorial space, a high number of iterations were chosen for the MCSA (7×10^6) and 120 independent optimizations were run.

Laboratory Materials

The synthetic gene for the designed protein was purchased from DNA2.0 (Menlo Park, CA) as an insert in a standard cloning plasmid (pJ201:10935). Enzymes for DNA cloning and manipulation were from New England Biolabs (Ipswich, MA). *E. coli* strains BL21(DE3) and DH5α as well as the QuikChange Site-directed mutagenesis kit were from Stratagene (La Jolla, CA). Oligonucleotides were obtained from Integrated DNA Technologies (Coralville, IA). The pET-20b(+) expression vector was from Novagen (Darmstadt, Germany). The Gel extraction kit and Miniprep kit were purchased from Qiagen (Valencia, CA). SDS-PAGE gels and dithiothreitol (DTT) were from Invitrogen (Carlsbad, CA). Centricon and Amicon centrifugal filter units were from Millipore (Billerica, MA). Slide-A-Lyzer dialysis cassettes were purchase from Pierce (Rockford, IL). Complete Mini EDTA free protease inhibitor cocktail tablets were obtained from Roche Applied Science (Mannheim, Germany). Isopropyl β-D-1-thiogalactopyranoside (IPTG) from Promega (Madison, Wisconsin). Chromatography columns were from GE Healthcare (Piscataway, NJ). Ampicillin, 2-mercaptoethanol (2-MCE) and phenylmethanesulphonylfluoride (PMSF) from Sigma-Aldrich (St. Louis, MO). All other chemicals were reagent grade or of higher purity and purchased from Sigma-Aldrich (St. Louis, MO) or Fisher Scientific (Fair Lawn, NJ).

Design and Cloning of DLac Gene

The final computationally designed amino acid sequence (DLac) was modified to include an initiation Met codon. The corresponding DNA sequence was created and optimized for expression in *E. coli* using the Gene Designer software program (Villalobos et al. 2006) and several unique restriction endonuclease sites were included within the sequence, including a C-terminal *Xho*I site (Supplementary Material Fig. 1). The DLac gene was digested from the pJ210:10935 plasmid using unique *Nde*I and *Xho*I sites and the fragment was ligated into the pET-20b(+) vector (which contains a C-terminal hexahistidine Tag) using the same restriction sites. This resulted in the addition of 8 amino acids to the C-terminus of the designed DLac sequence (LEHHHHHH). The resulting plasmid (pDLacI) was transformed into *E. coli* strains BL21(DE3) and DH5α and the fidelity of the insertions was verified by DNA sequencing.

Site-Directed Mutagenesis

Residues 453 – 455 in the DLac gene that were designed to be involved in T1 and T2/T3 copper binding (analogous to residues 452-454 in TvL) were mutated using site-directed mutagenesis in order to create an inactive negative control protein. The oligonucleotide primer 5'-CGGGCCCGTGGTTCCTGGTCAGCGTCATCGACTTCCACCTGG-3' and its complementary sequence were used to mutate the amino acid residues 453 – 455 (HCH) to VSV (encoded by the underlined DNA sequence) in pDLacI. The resulting plasmid pDLac-453HCH/VSV was then transformed into BL21(DE3) cells, and the mutations were verified by DNA sequencing.

Soluble Protein Expression and Purification

Single colonies of *E. coli* harboring either pDLacI or pDLacI-453HCH/VSV were used to inoculate 5 mL LB cultures supplemented with 100 μg/mL ampicillin and after 8 hours these were used to inoculate 50 mL cultures which were grown overnight at 37°C to saturation. One liter LB cultures with 100 μg/mL ampicillin were inoculated with 10 mL from the overnight cultures and were grown at 37°C in the presence of 1 mM $CuCl_2$ to an OD_{600} of ~0.9. The temperature was lowered to 25°C and IPTG was added to a final concentration of 1 mM. After 24 hrs of growth, the cells were pelleted by centrifugation at 8,300 x g for 10 minutes and the cell pellets were frozen at -20°C. The thawed pellets were resuspended in 75 mL of ice-cold Resuspension Buffer (20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole, 1 mM DTT, Complete Mini EDTA free protease inhibitor cocktail) and sonicated on ice. The sonicated cells were centrifuged at 25,600 x g for 30 min. The lysate was loaded onto a 5 mL FF HisTrap immobilized metal affinity chromatography (IMAC) columns using an FPLC apparatus. In a typical run, 150 mL of the lysate (obtained from 2 liters of culture) was injected, washed with 5 column volumes (CV) of 90 mM of imidazole (in 20 mM Tris pH 8.0, 500 mM NaCl) and eluted with 5 CV of 500 mM of imidazole in the same buffer. Fractions containing the DLac protein (or mutant), as determined by SDS-PAGE, were pooled and concentrated to 1 – 2 mL using an Amicon Ultra-15 Centrifugal Filter Unit (30 kDa NMWCO). The concentrated samples were loaded onto a Sephadex 200 PG column for size exclusion chromatography (SEC) equilibrated in an SEC Buffer (50 mM Tris pH 8.0, 500 mM NaCl). Fractions containing DLac (or the mutant) were collected and pooled.

Protein Expression and Purification from Inclusion Bodies

One liter cultures with 100 μg/mL ampicillin, inoculated as described above, were grown at 37°C in the presence of 1 mM $CuCl_2$ to an OD_{600} of ~0.6 and IPTG was added to a final concentration of 1 mM. After 4 hrs of growth at 37 °C, the cells were pelleted by centrifugation and frozen as described above. The thawed pellets were sonicated and the lysate was clarified as described above. The supernatant was discarded and the insoluble fractions containing the IBs were resuspended in 30 mL of ice-cold buffer containing 20 mM Tris pH 8.0, 500 mM NaCl, 2 M urea and 2% Triton X-100. The suspension was sonicated on ice. The insoluble material was twice separated by centrifugation at 25,600 x g for 10 min. The resulting washed IB pellets were resuspended in 50 mL Binding Buffer (20 mM Tris pH 8.0, 500 mM NaCl, 6 M guanidine HCl, 5 mM imidazole and 1 mM 2-MCE) and stirred for 30-60 minutes at room temperature. The solution was centrifuged at 25,600 x g for 10 min to remove unsolubilized material.

The solubilized IB proteins were loaded onto IMAC columns for simultaneous purification and refolding of the protein. In a typical run, 100 mL of IB solution was injected to

the column and washed with 10 CV of Binding Buffer followed by 10 CV of Washing Buffer (20 mM Tris pH 8.0, 500 mM NaCl, 6 M urea, 20 mM imidazole and 1 mM 2-MCE). Refolding was completed by changing the Washing Buffer to a Refolding Buffer (20 mM Tris pH 8.0, 500 mM NaCl and 20 mM imidazole) using a linear gradient over 30 CV. Experiments were performed with and without the addition of 1 mM $CuCl_2$ to the Refolding Buffer. After washing with 5 CV of Refolding Buffer, the protein was eluted with 5 CV of Elution Buffer (20 mM Tris pH 8.0, 500 mM NaCl and 500 mM imidazole). The fractions with refolded DLac were pooled. The samples were concentrated and further purified by SEC as described above.

Purification of TvL and SLAC Laccases

TvL was purchased as a crude protein extract. It was further purified by ion exchange chromatography using a HiPrep16/60 DEAE FF ion exchange chromatography (IEC) column in 10 mM sodium phosphate buffer pH 6.2, and the protein was eluted with a linear NaCl gradient from 0 to 1 M, as previously described (Hudak and Barton 2005). Further purification by SEC was performed as described above in a phosphate buffer (50 mM sodium phosphate, 150 mM NaCl, pH 7.2). The TvL fractions were pooled and concentrated (Amicon Ultra-15 Centrifugal Filter Unit, 30 kDa NMWCO).

The SLAC protein was expressed and purified using IEC and SEC as previously described (Gallaway et al. 2008; Machczynski et al. 2004).

PAGE and Mass Spectrometry

The purity and proper size of all of the purified protein samples (soluble DLac, refolded DLac, DLac 453HCH/VSV, TvL, SLAC) was verified using SDS-PAGE under reducing conditions. Samples of the soluble DLac and refolded DLac were excised from the gels and mass spectrometry following tryptic digestion was performed at the Columbia University Medical

UV-Vis Spectroscopy

Absorption scans over a wavelength range from 280 to 800 nm were performed using a SpectraMax M2^e microplate reader (Molecular Devices). Scans of TvL, SLAC, soluble DLac, and refolded DLac samples were compared to blank samples in SEC buffer. All measurements were made in triplicate.

Kinetic Activity Measurements

A series of experiments were performed to measure the catalytic activity of the DLac proteins following just IMAC purification. A variety of substrates with varying redox potentials were explored over a wide range of pH values. All measurements were performed using a SpectraMax M2^e microplate reader. The buffers, substrates, and other conditions are listed in Supplementary Information Table 1 and Table 2. The enzymes (DLac and DLac 453HCH/VSV) were compared to positive controls (TvL and SLAC) and a no-enzyme negative control. In each case, ~7 µg/mL of enzyme was used, and the absorbance at the appropriate wavelength was monitored at 40 $^\circ$C for several hours. Experiments with some of the substrates were repeated with the addition of 1 mM copper chloride. All measurements were made in at least duplicate.

Atomic Absorption Spectroscopy

Samples of the soluble DLac and refolded DLac protein in SEC buffer as well as TvL were sent to an outside laboratory (Galbraith Laboratories Inc., Knoxville, TN) for graphite furnace atomic absorption spectroscopy (GFAAS) analysis using a Perkin Elmer AAnalyst 800 GFAA/FLAA Spectrophotometer.

Circular Dichroism (CD) Spectroscopy

The far-UV CD spectra of SEC purified DLac and TvL samples were measured with a Jasco J-815 CD spectrometer using 0.1 and 0.01 cm optical path length cuvettes. A Peltier-equipped cell holder was used to control the sample temperature at 25°C or 90°C. The CD spectrum of TvL was determined in two buffers (10 mM sodium phosphate buffer at pH 7.2 and in 50 mM Tris buffer containing 500 mM sodium chloride at pH 8.0). The CD spectra of DLac were measured only in the high salt Tris buffer solution due the propensity of the protein to aggregate at lower ionic strengths. Spectrum scans were performed in continuous mode, with a 50 nm/min scan speed, 2 second response time and by accumulating five scans. Mean residue ellipticity values were calculated from ellipticity values using the protein concentrations determined by absorbance at 280 nm and calculated extinction coefficients.

The CD spectrum of TvL in sodium phosphate buffer at pH 8.0 was measured both at 25°C and 90°C using the 0.1 cm path length cuvette. The high absorbance of the 50 mM Tris buffer at pH 8.0 containing 500 mM NaCl (which was needed for DLac to stay soluble) necessitated the use of a 0.01 cm path length cuvette due to the high absorbance of the buffer. However, CD measurement at 90°C was not possible in this cuvette due to rapid evaporation of the 25 µL sample. Heat treated DLac and TvL samples were thus prepared by heating sample solutions in microcentrifuge tubes to 90°C for 10 min and cooled to room temperature followed by CD measurement at 25°C.

Chemical denaturation of TvL and DLac was performed by adding urea to a final concentration of 0 – 8 M in 50 mM Tris buffer, 500 mM NaCl at pH 8.0. The CD spectra of the resulting samples were then measured as described above and then smoothed with a 13 nm window Savitzky–Golay filter using the Spectra Manager (Jasco) software program. The mean residue ellipticity value at 220 nm, the wavelength where denaturation caused the largest change, was used to plot the urea denaturation curves. For the denaturation studies of refolded TvL, the enzyme was prepared by first denaturing the native protein in SEC buffer containing 8 M urea and 1 mM DTT overnight and subsequently dialyzing the sample in a step-wise manner against SEC buffer with decreasing urea concentrations (6M (for 2 hrs), 5M, 4M, 3M (for 1 hr each), 2M (for 2 hrs) and 0 M (overnight)). The resulting refolded TvL was then subjected to denaturation studies using urea as described above. A two-state model was fit to the unfolding data (Greenfield 2006).

RESULTS

Designed DLac Sequence

The computational design and optimization process introduced mutations in 253 of the 339 positions that were available for mutation (Fig. 1), and therefore the final DLac primary sequence shares less than 50% sequence homology with the parent TvL sequence (Fig. 2) (PDB

file of modeled DLac protein is included in Supplementary Material). Amino acids in the core of the enzyme (within 12 Å radius of the four copper) were left unmutated, and amino acids within 8 Å from the cooper atoms were also forced to have the same structural conformation as observed in the crystal structure. This was done in order to attempt to retain the native copper binding sites and their local environment as well as to retain a tunnel through the protein to the active site of the protein that enables the flux of oxygen and water (Figure 1).

Comparing the designed and parent TvL sequences, the number of buried positions (considered to be those where the relative accessible solvent area is less than 20%) slightly increased in the DLac sequence from 139 to 146. Pairs of residues were considered to be tightly interacting whenever their combined van der Waals and electrostatic interaction energies were below -5 kcal/mol and they were considered to have a clash whenever their interaction energy was higher than 2.5 kcal/mol. In the published TvL crystal structure, there are 10 tight interactions. Since the library refinement procedure favored rotamers belonging to highly interacting pairs, the DLac sequence introduced 68 of these pairs, which should produce a more extensive H-bond and salt bridge network on the proposed surface of the protein (Fig. 1). The ratio of polar to non-polar amino acids on the surface of the protein changed from 2.4 to 4.5. The wild type protein had 200 amino acids on the surface and among them, 141 were polar, whereas the DLac protein has 164 polar amino acids on its surface. The optimization process also resulted in a lower number of predicted clashes in the proposed protein structure, as the wild type protein has 8 clashes whereas the DLac sequence should only have three. The minimum folding energy of the final DLac protein was estimated to be -2,436.7 kcal/mol compared to -2,115.1 kcal/mol for the wild type TvL, using the same scoring function, indicating that the final folded structure should have increased stability.

Protein Expression and Purification

The protein was readily overexpressed in *E. coli*, but the majority of the protein was found as insoluble aggregates in IBs. Several parameters were varied in order to improve soluble expression including: temperature, IPTG concentration, copper concentrations, and induction time. The best results were obtained when cultures were moved to 25°C at the time of induction. The expression was not fully optimized, but sufficient quantities of soluble protein were obtained for characterization experiments. Fig. 3 shows the expression levels obtained, and the ratio of protein in the soluble and insoluble fractions. Small quantities of soluble protein were obtained using IMAC and SEC purification. However, the soluble protein precipitated after overnight cold storage unless a large salt concentration (500 mM NaCl) was included in the buffer.

Since the majority of the protein was found in IBs, several refolding protocols were explored. The DLac protein was readily soluble in either 6 M guanidine HCl or 8 M urea, and DTT was added to ensure any disulfides were reduced. The largest amount of protein was obtained when the DLac was refolded on an IMAC column. Copper chloride (1 mM) was added to the Refolding Buffer and/or to the refolded proteins overnight, and following concentration, SEC purification was performed. There were multiple peaks in the SEC chromatograms and a large proportion of the protein was found in the void volume (suggesting aggregation). There was a common peak between the refolded and soluble DLac injections at 73 mL (Fig. 4A and 4B), and it was this fraction which was used for all subsequent experiments. It was observed that if the time between IMAC refolding and SEC injection was minimized, the largest peak of interest was obtained, which suggests that aggregation increased over time.

185

Native TvL did not share the same retention time in the SEC chromatograms as either the soluble or refolded DLac protein. However, the denatured and refolded TvL did share the same retention time as the DLac proteins (Fig. 4C and 4D).

Mass and Copper Content

Trypsin digested samples from an SDS-PAGE gel band showed that both the soluble and refolded DLac protein peaks give rise to the same digestion pattern, which is consistent with the theoretical digestion pattern of the computationally designed protein (Supplementary Material Table 3).

The blue T1 copper, if present in the protein, is characterized by an absorption peak at ~600 nm. For the TvL and SLAC proteins, a wavelength scan reveals the correct distinctive peak. Neither the solubly expressed DLac protein nor the refolded DLac protein showed a peak in this region (Fig. 5). This was independent of whether the proteins were incubated with 1 mM copper chloride before subjecting them to SEC.

GFAAS analysis was unable to detect copper ions in either the soluble or refolded DLac proteins. Copper was present in the expected amount of 4 copper ions per molecule for the TvL sample.

Kinetic Activity Measurements

Although the DLac protein did not appear to bind a significant amount of copper, kinetic assays were still performed in case small amounts of the enzyme had folded properly. In order to maximize the chances of measuring kinetic activity, proteins purified by IMAC (and not SEC) were used. The IMAC samples would include the protein correctly identified by mass spectrometry, as well as any other forms of the protein that might behave differently (and thus separated out) by the SEC column step. Since some of the substrates can autooxidize, the DLac 453HCH/VSV mutant, purified identically, was used as a negative control.

The TvL protein was active at a pH range from 4 to 8, oxidizing all substrates in Supplementary Material Table 2 as observed in colorimetric assays. SLAC was able to oxidize DMP, guaiacol, MBA and SYG at pH values ranging from 7 to 10, with higher activity usually at pH 8 to 9. SLAC also was able to oxidize ABTS at a pH range of 4 to 7, similar to that of TvL. SLAC did not appear to be able to oxidize DMBA.

At pH values ranging of 2 to 11 for all of the substrates in Supplementary Material Table2, the DLac protein (either solubly expressed or refolded) and the triple mutant DLac 453HCH/VSV exhibited identical colorimetric results, indicating a lack of catalytic activity under these experimental conditions.

DLac and TvL conformation

CD spectroscopy was used to study the protein conformation of DLac while comparing it to that of TvL (Fig. 6). The CD spectra of both the refolded DLac and the DLac from the soluble fraction had the same shape with a minimum at about 208 nm (Fig. 6A and 6B). The spectra were deconvoluted and secondary structure compositions of 9% alpha helix, 33% beta sheet, 19% turns, and 35% unordered were calculated for the solubly expressed form of DLac, and 9%,

186

35%, 18% and 39% respectively for the refolded DLac. The CD spectrum of TvL had a maximum at 196 nm and a minimum at around 216 nm resembling a structure that is rich in beta sheets (Fig. 6C). Estimation of secondary structure contents by deconvolution yields 5% helix, 41% beta sheet and 50% other structures, which are in agreement with the TvL X-ray crystal structure(Piontek et al. 2002) that has 11% helical, 37% beta sheet, and 52% other structures according to the secondary structure assignments from the program DSSP (Kabsch and Sander 1983).

When the DLac samples were dissolved in 8 M urea, a significant change in the CD specta was observed as there was an increase in unordered structure from 35% to 47% for the solubly expressed DLac and an increase from 39% to 52% for the refolded DLac. When the TvL was heat-treated at 90°C and cooled to 25°C, a CD spectrum (10% helical, 34% beta sheet) was observed that was very similar to what was seen with the DLac samples. After the heat treatment, TvL did not exhibit detectable enzymatic activity when tested with ABTS. In order to further compare TvL with the DLac samples, the TvL was chemically denatured and refolded using step-vise dialysis to mimic the DLac refolding protocol. Again, the CD spectrum of the denatured and refolded TvL was very similar to the refolded DLac protein as shown in Figs. 6C and 6D.

The thermodynamic stability of DLac was investigated at high ionic strength to prevent the aggregation of the protein. The refolded DLac, denatured and refolded TvL, and native TvL were denatured in urea and the unfolding was followed using CD spectroscopy. All three proteins had similar weakly sigmoidal unfolding curves (Fig. 7) suggesting similar stabilities. It is possible that the high (500 mM NaCl) ionic strength destabilizes the proteins. The free energy of folding for DLac and refolded TvL were -1.8 kcal/mol and -2.2 kcal/mol, respectively. The unfolding – folding transition of native TvL is not reversible under the conditions used, and thus the thermodynamic analysis cannot be applied.

DISCUSSION

We have shown how computational design could be used to engineer a large protein with little sequence similarity to any known sequence. Although current computational methods are still far from getting a folded and active protein of this size, our attempt highlights the important factors to consider in the design process. There are many examples of laccase and laccase-like proteins identified from prokaryotic organisms, and several of these have been recombinantly overexpressed in *E. coli*. Since the fungal laccase proteins generally have 3-D structures and active site architectures that are similar to the prokaryotic laccase enzymes, there is no obvious reason why the structure of a fungal laccase cannot similarly be expressed in a bacterial host. In fact, it has recently been reported that a fungal laccase (from *C. bulleri*) has been heterologously expressed in *E. coli* in a catalytically active form (Salony et al. 2008).

Computational protein design has been successfully used to identify primary amino acid sequences that will adopt a predetermined tertiary protein fold (Jaramillo et al. 2002; Ogata et al. 2003; Wernisch et al. 2000). We set out to use this approach to determine a primary sequence that would fold into the tertiary structure of the TvL glycoprotein. Ideally this novel designed protein would exhibit increased stability, and it would retain the high redox potential and activity of the parental protein which would be beneficial for a variety of applications.

187

Laccase enzymes follow a bi-bi ping-pong mechanism such that the substrate is oxidized by the T1 copper atom, the electron is conducted through the core of the protein to the T2/T3 site, and 4 electrons are used to reduce dioxygen to water. It is clear that the proper orientation and local environment around the copper atoms is essential for catalysis and therefore the computational design strategy left all amino acids within 12Å of the copper atoms unchanged. The final design allowed 339 of the 499 amino acids in the protein to be mutated making this one of the largest reported computational protein design efforts reported. The final designed DLac sequence had less than 50% sequence identity with that of parent TvL (Fig. 2), but it was correctly recognized by sequence alignment servers like BLAST to be a copper-oxidase.

One of the most interesting features of the new sequence is the substantial increase in tightly interacting amino acid pairs that were designed to form on the surface of the properly folded protein (Fig. 1C and D). This would result in a more extensive H-bond and salt bridge network on the surface of the properly folded protein, ideally leading to increased stability of the protein. This is a result of the computational methodology which tends to favor electrostatic interactions at the surface and not at the core (Jaramillo et al. 2002), but this increased polar network has not been found to be incompatible with folding (Suarez and Jaramillo 2009). The calculated theoretical pI of the DLac protein using ProtParam (Walker 2005) was 9.26 as compared to 5.87 for TvL (and an experimental value of 3.5 has been reported (Piontek et al. 2002)). The large change in the charges of the protein sequence may have contributed to the observed propensity for the polypeptide to aggregate, as it was found that in order for the expressed DLac protein to remain soluble, a high ionic strength buffer was required (500 mM NaCl). This is contrast to what was observed for the denatured and refolded native TvL protein, which did not require high salt to remain soluble.

Transformation of the designed gene in *E. coli* resulted in the high expression of protein that was preferentially located in IBs and several established methods were used to attempt to refold the insoluble protein. Protocols were developed to refold the DLac protein as well as to purify the small amount of solubly expressed DLac protein. Once purified, these samples behaved identically in every test performed, suggesting that they are in a similar conformation. The absence of the blue T1 copper combined with the lack of copper detected using GFAAS is consistent with lack of apparent kinetic activity of the enzymes under the experimental conditions tested.

In order to further explore whether the DLac protein was properly folded, both SEC and CD measurements were performed. The SEC experiments demonstrated that both DLac samples as well as the denatured and refolded TvL elute at the same retention time, while the native (active) TvL elutes at a later time (Fig. 4). These results are consistent with the CD experiments where it was shown that both DLac samples have similar secondary structure contents. These spectra are very different from the native TvL spectrum, but are very similar to the spectrum of the chemically denatured and refolded TvL (which is no longer active) (Fig. 6). A similar CD spectrum has been also reported for another fungal laccase that was heat-treated (López-Cruz et al. 2006). Taken together these results suggest that the designed DLac polypeptide is able to adopt a similar soluble structure in both the cellular milieu, and following denaturation and refolding of protein accumulated in IBs. This appears to be a molten globule conformation that is very similar to the denatured and refolded TvL.

The weakly sigmoidal unfolding curves (Fig. 6) indicate that the resistance towards chemical denaturation is similar in all 3 protein samples. Moreover, it suggests that the proteins

studied have a low cooperativity in unfolding and thus a non two-state folding-unfolding process is expected. TvL is a comparatively large protein with three domains and unfolding may involve significant accumulation of intermediate states. Furthermore, the high ionic strength may destabilize the structure of native TvL. The folding – unfolding process for TvL is not reversible and thus thermodynamic analysis cannot be applied. However, the folding and unfolding of DLac and the *in vitro* folded TvL are reversible. A two-state model, that assumes no accumulation of an intermediate state, fit well to the unfolding curves. An unfolding free energy change of 1.8 kcal/mol for DLac and 2.2 kcal/mol for *in vitro* folded TvL was obtained. Thus, the thermodynamic stabilities of the two proteins are comparable and relatively low.

We speculate that this refolded structure could be a molten globular stable folding intermediate that cannot proceed to the native structure. This hypothesis is based on the observations that the protein can adopt a soluble 3-D structure; the structure is similar to the denatured and refolded TvL but not the native TvL, the protein does not bind copper and is inactive, and the protein can be reversibly unfolded with a low thermodynamic stability. One explanation for this could be limitations in copper incorporation into the DLac protein. In addition to their catalytic significance, bound coppers have been reported to have an important structural role in multicopper oxidases (Durão et al. 2006; Durão et al. 2008; Sedlak and Wittung-Stafshede 2007). It seems that native glycosylations are not sufficient to enable correct refolding as the native TvL sample (that is fully glycosylated) could not be refolded from the molten globular structure.

The successful chemical denaturation and refolding of some other laccases has previously been reported in the literature. The *B. subtilis* endospore coat protein, CotA, which is a prokaryotic laccase-like protein when expressed in *E. coli* was initially found to accumulate in IBs and initial refolding experiments were similarly unsuccessful (Martins et al. 2002), but conditions were obtained to enable the production of a soluble and functional CotA protein.(Sakasegawa et al. 2006) And, the fungal laccase from *C. bulleri* was successfully refolded and found to be active after subjecting it to several denaturing and metal chelating agents.(Salony et al. 2008) Unlike TvL, this laccase does not contain cysteine residues which may have simplified the refolding process for this protein.

Computational protein design provides an opportunity to test our understanding of protein folding rules and to validate these assumptions experimentally. Since proteins with structures very similar to TvL can fold into a functional form in prokaryotic hosts, there is no fundamental reason why a protein with an identical structure to TvL cannot similarly be expressed and folded in this environment. Therefore, future experiments will focus on determining if sequence features found in the prokaryotic laccases are missing in the DLac design, and this will help to better elucidate the folding pathway utilized by TvL and related multi-copper oxidases. In addition, future experiments will also explore the expression of the DLac protein in a eukarotic host, such as yeast, to see if this can enable functional folding and expression.

To the best of our knowledge, this is the largest protein that has ever been designed using these computational techniques. Laccases are multi-domain proteins generally comprised of 3 cupredoxin-like domains each with a greek key β-barrel topology (Giardina et al. 2010; Skalova et al. 2009). These domains likely evolved to be stable independent of the entire laccase protein. This suggests that future design efforts may also benefit from the independent modular design of the domains in conjunction with the design of the protein as a whole.

CONCLUSION

We have successfully expressed a computational designed laccase based on a fungal protein backbone in a bacterial host. The DLac protein is one of the largest computationally designed proteins expressed in an organism to date. The amount of soluble protein compared to the insoluble aggregates was small, but enough to conduct preliminary characterization experiments.

The soluble and refolded DLac proteins were apparently identical and neither of the proteins was found to be structurally comparable to the native TvL protein. Instead, the DLac proteins were apparently structurally similar to denatured and refolded TvL, which suggests that the native TvL requires a folding pathway that cannot be replicated following chemical denaturation. Future attempts to design this protein will likely require the incorporation of a folding pathway, which may be included by considering the stabilization of intermediate structures, although there are no templates for these presently available.

Copper ions were not detected in the DLac protein samples and the lack of copper is consistent with the observed lack of kinetic activity. These results suggest that in the absence of the fungal cellular context, both DLac and TvL fold into the same soluble but inactive structure. These structures have a similar stability according to the unfolding curves, thus suggesting that the proteins are in a molten globule-like state.

ELECTRONIC SUPPLIMENTARY MATERIAL

The supplementary material available in the electronic edition contains the DNA sequence of the DLac gene, buffers and substrates used in the kinetics measurements, and a summary of the mass spectrometry results for the DLac proteins (Glykys et al Supp Mat.doc) A PDB file of the modeled DLac protein is also included in Supplementary Material.

ACKNOWLEDGEMENTS

REFERENCES

Alexandre G, Zhulin IB. (2000). Laccases are widespread in bacteria. Trends Biotechnol 18:41-2.

Barton SC, Gallaway J, Atanassov P. (2004). Enzymatic biofuel cells for implantable and microscale devices. Chem Rev 104:4867-86.

Brissos V, Pereira L, Munteanu FD, Cavaco-Paulo A, Martins LO. (2009). Expression system of CotA-laccase for directed evolution and high-throughput screenings for the oxidation of high-redox potential dyes. Biotechnol J 4:558-563.

Bulter T, Alcalde M, Sieber V, Meinhold P, Schlachtbauer C, Arnold FH. (2003). Functional expression of a fungal laccase in Saccharomyces cerevisiae by directed evolution. Appl Environ Microbiol 69:987-95.

Claus H. (2003). Laccases and their occurrence in prokaryotes. Arch Microbiol 179:145-50.

Dunbrack RL, Karplus M. (1993). Backbone-Dependent Rotamer Library for Proteins - Application to Side-Chain Prediction. Journal of Molecular Biology 230:543-574.

Durão P, Bento I, Fernandes A, Melo E, Lindley PF, Martins L. (2006). Perturbations of the T1 copper site in the CotA laccase from *Bacillus subtilis*: structural, biochemical, enzymatic and stability studies. J Biol Inorg Chem 11:514-26.

Durão P, Chen Z, Fernandes A, Hildebrandt P, Murgida D, Todorovic S, Pereira M, Melo E, Martins L. (2008). Copper incorporation into recombinant CotA laccase from *Bacillus subtilis*: characterization of fully copper loaded enzymes. J Biol Inorg Chem 13:183-193.

Endo K, Hayashi Y, Hibi T, Hosono K, Beppu T, Ueda K. (2003). Enzymological characterization of EpoA, a laccase-like phenol oxidase produced by Streptomyces griseus. J Biochem (Tokyo) 133:671-7.

Festa G, Autore F, Fraternali F, Giardina P, Sannia G. (2008). Development of new laccases by directed evolution: functional and computational analyses. Proteins 72:25-34.

Gallaway J, Wheeldon I, Rincon R, Atanassov P, Banta S, Barton SC. (2008). Oxygen-reducing enzyme cathodes produced from SLAC, a small laccase from Streptomyces coelicolor. Biosens Bioelectron 23:1229-35.

Gelo-Pujic M, Kim HH, Butlin NG, Palmore GT. (1999). Electrochemical studies of a truncated laccase produced in Pichia pastoris. Appl Environ Microbiol 65:5515-21.

Giardina P, Faraco V, Pezzella C, Piscitelli A, Vanhulle S, Sannia G. (2010). Laccases: a never-ending story. Cell Mol Life Sci 67:369-85.

Glykys DJ, Banta S. (2009). Metabolic control analysis of an enzymatic biofuel cell. Biotechnol Bioeng 102:1624-35.

Greenfield NJ. (2006). Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. Nat Protoc 1:2733-41.

Gupta N, Farinas ET. (2009). Narrowing laccase substrate specificity using active site saturation mutagenesis. Comb Chem High Throughput Screen 12:269-74.

Hudak NS, Barton SC. (2005). Mediated biocatalytic cathode for direct methanol membrane-electrode assemblies. Journal of the Electrochemical Society 152:A876-A881.

Jaramillo A, Wernisch L, Hery S, Wodak SJ. (2002). Folding free energy function selects native-like protein sequences in the core but not on the surface. Proceedings of the National Academy of Sciences of the United States of America 99:13554-13559.

Kabsch W, Sander C. (1983). Dictionary of Protein Secondary Structure - Pattern-Recognition of Hydrogen-Bonded and Geometrial Features. Biopolymers 22:2577-2637.

Kunamneni A, Camarero S, Garcia-Burgos C, Plou FJ, Ballesteros A, Alcalde M. (2008). Engineering and Applications of fungal laccases for organic synthesis. Microb Cell Fact 7:32.

Li X, Wei Z, Zhang M, Peng X, Yu G, Teng M, Gong W. (2007). Crystal structures of E. coli laccase CueO at different copper concentrations. Biochem Biophys Res Commun 354:21-6.

López-Cruz JI, Viniegra-Gonzalez G, Hernández-Arana A. (2006). Thermostability of native and pegylated *Myceliophthora thermophila* laccase in aqueous and mixed solvents. Bioconjug Chem 17:1093-8.

Machczynski MC, Vijgenboom E, Samyn B, Canters GW. (2004). Characterization of SLAC: a small laccase from Streptomyces coelicolor with unprecedented activity. Protein Sci 13:2388-97.

Madzak C, Mimmi MC, Caminade E, Brault A, Baumberger S, Briozzo P, Mougin C, Jolivalt C. (2006). Shifting the optimal pH of activity for a laccase from the fungus Trametes versicolor by structure-based mutagenesis. Protein Eng Des Sel 19:77-84.

Martins LO, Soares CM, Pereira MM, Teixeira M, Costa T, Jones GH, Henriques AO. (2002). Molecular and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the Bacillus subtilis endospore coat. J Biol Chem 277:18849-59.

Mayer AM, Staples RC. (2002). Laccase: new functions for an old enzyme. Phytochemistry 60:551-65.

Nakamura K, Go N. (2005). Function and molecular evolution of multicopper blue proteins. Cell Mol Life Sci 62:2050-66.

Ogata K, Jaramillo A, Cohen W, Briand JP, Connan F, Choppin J, Muller S, Wodak SJ. (2003). Automatic sequence design of major histocompatibility complex class I binding peptides impairing CD8(+) T cell recognition. Journal of Biological Chemistry 278:1281-1290.

Ooi T, Oobatake M, Nemethy G, Scheraga HA. (1987). Accesible Surface-Areas as a Measure of the Thermodynamic Parameters of Hydration of Peptides. Proceedings of the National Academy of Sciences of the United States of America 84:3086-3090.

Petrek M, Otyepka M, Banas P, Kosinova P, Koca J, Damborsky J. (2006). CAVER: A New Tool to Explore Routes from Protein Clefts, Pockets and Cavities. Bmc Bioinformatics 7:316.

Piontek K, Antorini M, Choinowski T. (2002). Crystal structure of a laccase from the fungus Trametes versicolor at 1.90-A resolution containing a full complement of coppers. J Biol Chem 277:37663-9.

Riva S. (2006). Laccases: blue enzymes for green chemistry. Trends Biotechnol 24:219-26.

Rodgers CJ, Blanford CF, Giddens SR, Skamnioti P, Armstrong FA, Gurr SJ. (2010). Designer laccases: a vogue for high-potential fungal enzymes? Trends Biotechnol 28:63-72.

Sakasegawa S, Ishikawa H, Imamura S, Sakuraba H, Goda S, Ohshima T. (2006). Bilirubin oxidase activity of Bacillus subtilis CotA. Applied and Environmental Microbiology 72:972-975.

Sakurai T, Kataoka K. (2007). Basic and applied features of multicopper oxidases, CueO, bilirubin oxidase, and laccase. Chem Rec 7:220-9.

Salony, Garg N, Baranwal R, Chhabra M, Mishra S, Chaudhuri TK, Bisaria VS. (2008). Laccase of *Cyathus bulleri*: structural, catalytic characterization and expression in *Escherichia coli*. Biochim Biophys Acta 1784:259-68.

Sedlak E, Wittung-Stafshede P. (2007). Discrete roles of copper ions in chemical unfolding of human ceruloplasmin. Biochemistry 46:9638-44.

Shleev S, Tkac J, Christenson A, Ruzgas T, Yaropolov AI, Whittaker JW, Gorton L. (2005). Direct electron transfer between copper-containing proteins and electrodes. Biosens Bioelectron 20:2517-54.

Skalova T, Dohnalek J, Ostergaard LH, Osteryaard PR, Kolenko P, Duskova J, Stepankova A, Hasek J. (2009). The Structure of the Small Laccase from Streptomyces coelicolor Reveals a Link between Laccases and Nitrite Reductases. Journal of Molecular Biology 385:1165-1178.

Solomon EI, Sundaram UM, Machonkin TE. (1996). Multicopper Oxidases and Oxygenases. Chem Rev 96:2563-2606.

Suarez M, Jaramillo A. (2009). Challenges in the computational design of proteins. J R Soc Interface 6 Suppl 4:S477-91.

Tortosa P, Jaramillo A. 2006. Active sites by Computational Protein Design. Proceedings of the II BIFI 2006 International Conference. p 96-101.

Villalobos A, Ness JE, Gustafsson C, Minshull J, Govindarajan S. (2006). Gene Designer: a synthetic biology tool for constructing artificial DNA segments. BMC Bioinformatics 7:285.

Walker JM. 2005. The proteomics protocols handbook. Totowa, N.J.: Humana Press. xviii, 988 p.

Wernisch L, Hery S, Wodak SJ. (2000). Automatic protein design with all atom force-fields by exact and heuristic optimization. Journal of Molecular Biology 301:713-736.

Wheeldon IR, Gallaway JW, Barton SC, Banta S. (2008). Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks. Proc Natl Acad Sci U S A 105:15275-80.

Xu F. (1997). Effects of redox potential and hydroxide inhibition on the pH activity profile of fungal laccases. J Biol Chem 272:924-8.

Xu F, Berka RM, Wahleithner JA, Nelson BA, Shuster JR, Brown SH, Palmer AE, Solomon EI. (1998). Site-directed mutations in fungal laccase: effect on redox potential, activity and pH profile. Biochem J 334 (Pt 1):63-70.

194

Xu F, Palmer AE, Yaver DS, Berka RM, Gambetta GA, Brown SH, Solomon EI. (1999). Targeted mutations in a Trametes villosa laccase. Axial perturbations of the T1 copper. J Biol Chem 274:12372-5.

Xu F, Shin W, Brown SH, Wahleithner JA, Sundaram UM, Solomon EI. (1996). A study of a series of recombinant fungal laccases and bilirubin oxidase that exhibit significant differences in redox potential, substrate specificity, and stability. Biochim Biophys Acta 1292:303-11.
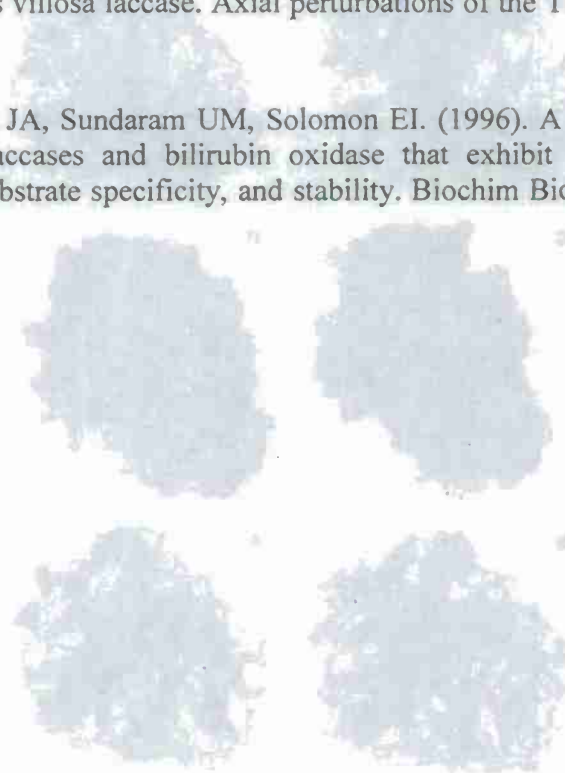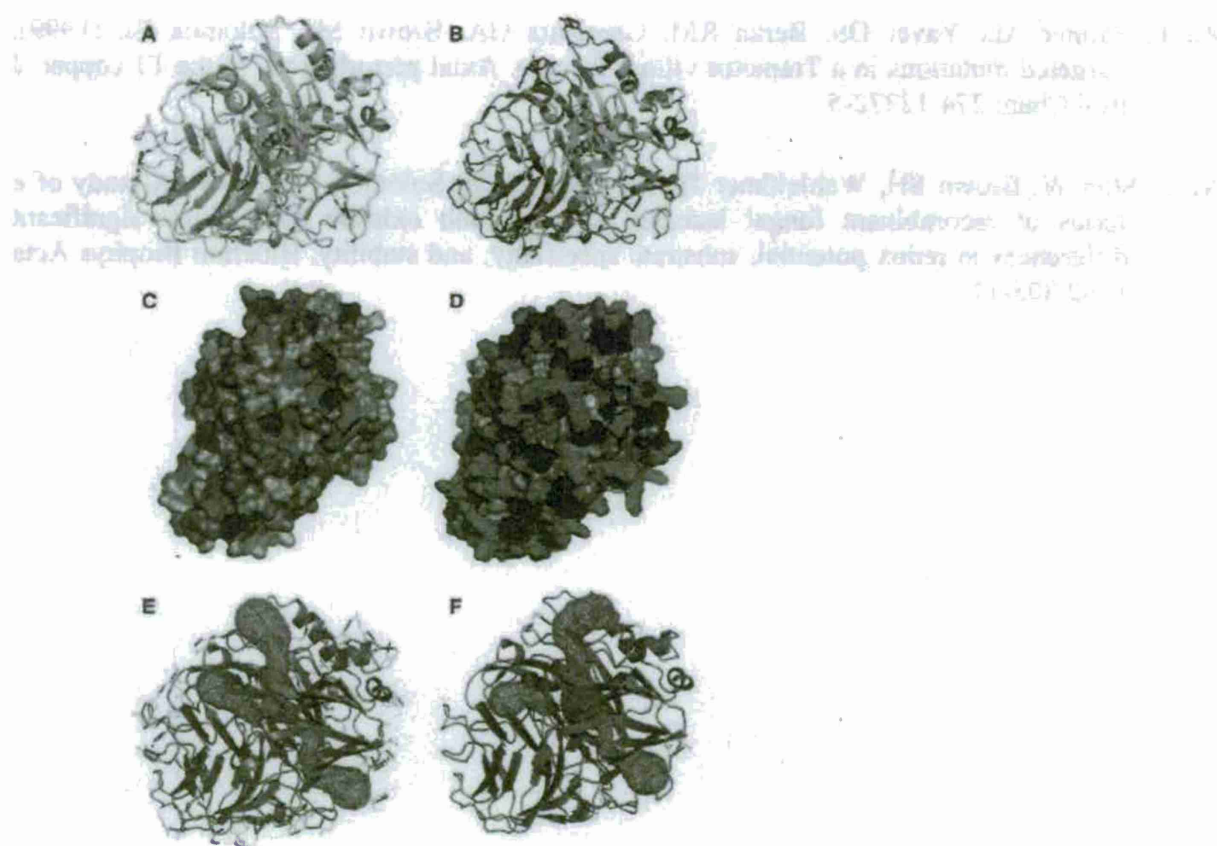
Figure 1. Structures of the TvL and DLac proteins. **A:** 3-D rendering of the native TvL structure (PDB access code 1GYC). Grey areas represent residues that were available for mutation in the DLac design while the green areas represent amino acids that were left unmutated. Blue lines indicate the bonds formed between tightly interacting amino acids (shown as sticks). **B:** 3-D rendering of the computational design of the DLac protein (The PDB code for DLac can be found in the Supplementary Information). Green areas represent amino acids that were left the same as compared to TvL. Grey areas indicate amino acids where mutation was allowed, but the wild type amino acid was selected. Purple areas indicate amino acids that are different from the TvL structure. Blue lines show tightly interacting amino acids. **C:** Space filling image of the wild type TvL structure. Orange amino acids are polar and grey amino acids are non-polar. Red amino acids are positively charged while blue are negatively charged. **D:** Space filling image of the designed structure of the DLac protein, with amino acids color coded the same as in panel C. **E:** Surface and cavity analysis of the native TvL structure, blue lines indicate bonds formed between interacting amino acids on the surface of the protein, (shown as sticks). Channels leading to the T1, T2/T3 sites are shown in green. Tunnels have been computed using CAVER (Petrek et al. 2006) **F:** Surface and cavity analysis of the designed DLac protein, blue lines indicate bonds formed between interacting amino acids on the surface of the protein, (shown as sticks). Channels leading to the T1, T2/T3 sites are shown in green.

```
DLAC   MYIGPQABLEVKEKPLKPAGLELKLIVVKDRLPAPEILGKKGQEFRLRVKVQLRDKKLLKL
1GYC   AIGPAASLVVANAPVSPDGFLRDAIVVNGVFPSPLITGKKGDRFQLNVVDTLTNHTMLKS
       ***.*..*.*  *.*..*:..***;..;*;* *.****;.*;*.*  *.;;.;**

DLAC   TSIHWHGFKQKRTNWADGPAFVNQCPIVQGKEFEYKFRVPEQAGTFWYHSHLKTQYCVGL
1GYC   TSIHWHGFFQAGTNWADGPAFVNQCPIASGHSFLYDFHVPDQAGTFWYHSHLSTQYCDGL
       ********.* *.********.****.**.:*.*.;.*:.**.********.*****..**** **

DLAC   RGPFLVEDPRDPQRQKWKEWREKALLEIEDLLKQLAQEGPKFPRGLDALLLKGQGRAEEK
1GYC   RGPFVVYDPKDPHASRYDVDNESTVITLTEWYHTAARLGPRFPLGADATLINGLGRSAST
       ****;* **;**;.;;.. .*.;;; * *  ;   *;**;** * ** *;;* **; ..

DLAC   PEYRLAKVDVRQGIEKAIALVSISCDPNIKLKVDRLKLEVVEVDGIKVKPLEVKAVEIFA
1GYC   PTAALAVINVQHGKRYRFPLVSISCDPNYTFSIDGHNLTVIEVDGINSQPLLVDSIQIFA
       *  ** ;;*;;*.  ;  ;********  .;;.;  ;* *:*****;:** *.;;;***

DLAC   AQRIQAVQWARIEKDILWIKAEPKKGKPDWRRGIAAAYLILEGAPRREPKQREDREERPL
1GYC   AQRYSFVLNANQTVGNYWIRANPNPGTVGFAGGINSAILRYQGAPVAEPTTTQTTSVIPL
       ***..* *.   .::*:*: *..; ** ;* * ;*** *.. ;  ** **;* *. .

DLAC   EEDRLEPQQDKPVPGKPEPRGVRKWLRLDFEFEGDRFKIQRHEFYPPEEPVLKXIERGKK
1GYC   IETNLHPLARMPVPGSPTPGGVDKALNLAFNFNGTNFFINNASFTPPTVPVLLQILSGAQ
       **;.*..*  .*.*** *.*.*;;; ;.*;:;* .; . ** *** ;*  ; *

DLAC   EKKDLDPYGLVDPLPDRADIEIELPAKREAPGKPHPFHLHGHEFAVVKAERKRDLKQERP
1GYC   TAQDLLPAGSVYPLPAHSTIEITLPATALAPGAPHPFHLHGHAFAVVRSAGSTTYNYNDP
       ;:**.* * * *** ;; *** ***. *** ********** ****;;. ;;; : ;

DLAC   VLRDVVSTGEPRQRDQVKIKFRVDNPGPWFLHCHIDFHLDAGFAIVFAAAYEKVWEDVPK
1GYC   IFRDVVSTGTPAAGDNVTIRFQTDNPGPWFLHCHIDFHLEAGFAIVFAEDVADVKAANPV
       ;;******* *  *;** ;*;.*************.****.*******;  ;.*; *.

DLAC   PKEWDELCPDEQRKKSEDKLEHHHHHH
1GYC   PKAWSDLCPIYDGLSEANQ
       ** *.;*** ;  .. ;;
```

Figure 2. Comparison of DLac and TvL (1GYC) primary sequences. There is less than 50% sequence homology between the two proteins. In order to express and purify the DLac sequence, an N-term methionine was added to the sequence as well as a C-term hexahistidine tag (LEHHHHHH).

Figure 3. SDS-PAGE results of the DLac expression and purification. Lane A: MW Ladder, Lane B: Whole cell extract of the uninduced cell culture, Lane C: Whole cell extract of the induced cell culture, Lane D: Soluble protein fraction from induced cells, Lane E: Insoluble protein fraction from induced cells solubilized in 8 M urea, Lane F: Refolded DLac protein from an IMAC column, Lane G: Soluble DLac protein after SEC, Lane H: Refolded DLac protein after SEC, Lane I: Native TvL after SEC column, Lane J: Denatured and refolded TvL protein after SEC column, Lane K: MW Ladder.
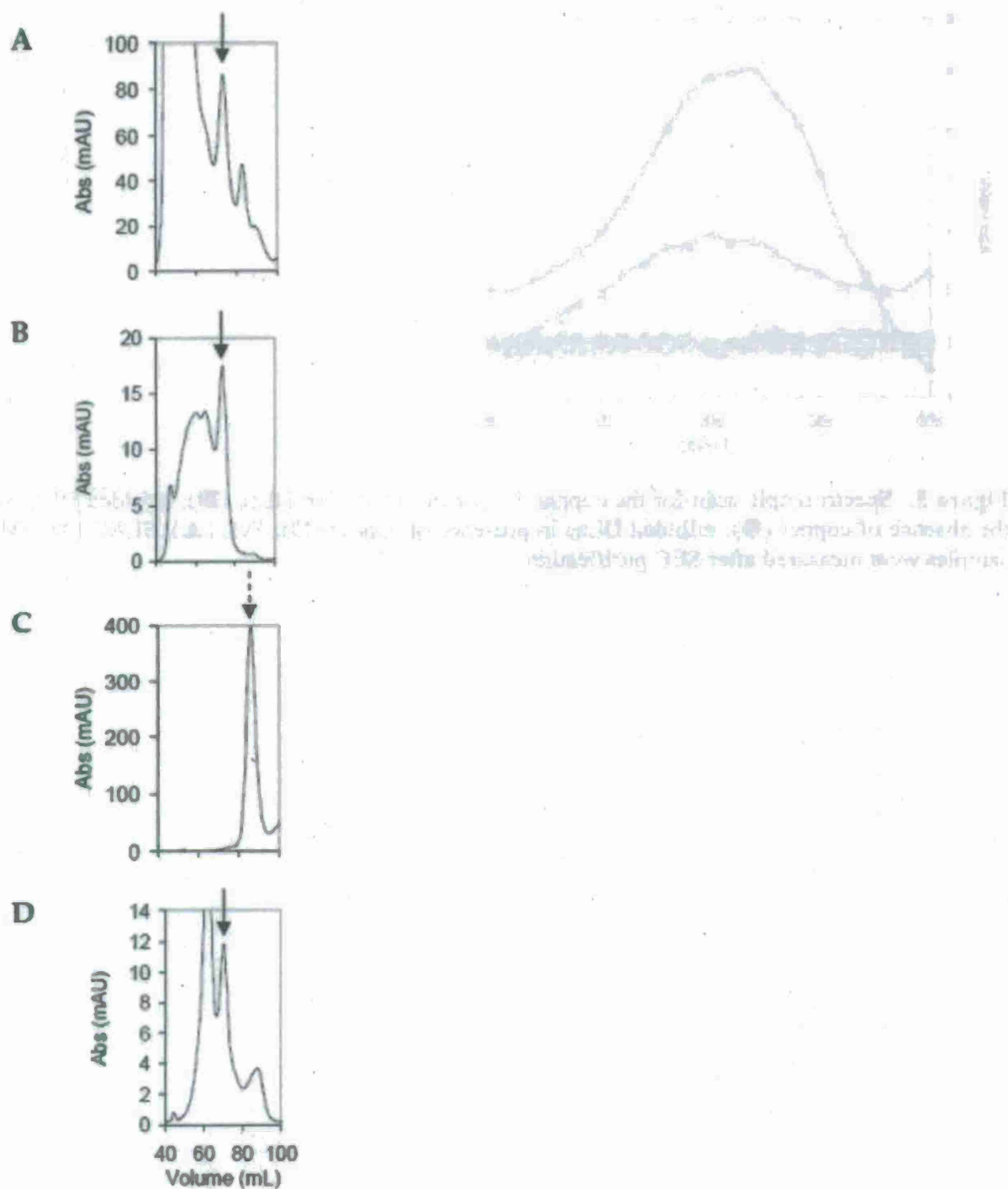
Figure 4. SEC chromatograms for DLac and TvL. Panel A: Soluble DLac protein, Panel B: Refolded DLac protein, Panel C: Native TvL protein, Panel D: Denatured and refolded TvL (with the refolding occurring in the SEC column). For the DLac and refolded TvL proteins, the peak that eluted at ~73 mL was used (black arrows). For the native TvL, the active protein eluted later (dashed arrow).
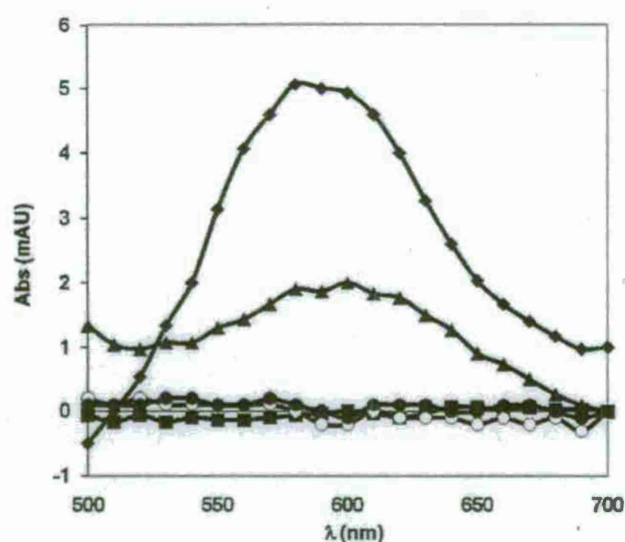
Figure 5. Spectroscopic scan for the copper T1 site of the soluble DLac (■), refolded DLac in the absence of copper (●), refolded DLac in presence of copper (O), TvL (▲), SLAC (◆). All samples were measured after SEC purification.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
31-01-2012	Final	From 01/05/2006 - To 30/10/2011

4. TITLE AND SUBTITLE
MURI: Fundamentals and Bioengineering of Enzymatic Fuel Cells

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-06-1-0264

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Plamen Atanassov, University of New Mexico
Scott Banta, Columbia University
Shelley Minteer, University of Utah
Scott Calabrese Barton, Michigan State University
Michael Cooney, University of Hawaii
Sanjeev Mukerjee, Northeastern University

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF NEW MEXICO
1700 LOMAS BLVD NE STE 2200
MSC 01 1247, 1 UNIVERSITY OF MEXIC
ALBUQUERQUE NM 83701-0001

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF, AFRL DUNS 143574726
AF OFFICE OF SCIENTIFIC RESEARCH
875 N. RANDOLPH ST. ROOM 3112
ARLINGTON VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0676

12. DISTRIBUTION/AVAILABILITY STATEMENT

A-Approve for public Release.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The MURI research has been comprised of four interconnected Thrust Areas, : Enzyme Engineering of Bio-electrocatalysts, Oxidation of Complex Biofuel, Electron Transfer Fundamentals and Interface Architecture and Transport Phenomena in Bio-electrodes. The program established in summary that enzymatic biofuel cells present a viable candidate for ultimate miniaturization of power sources because it is a class of energy conversion devices that employ molecular electrocatalysts – enzymes and hierarchical surface architectures developed in the course of the advancement of bio-nano interface technology. Processes at the enzymatic biofuel cells take place on several hierarchical scales: molecular-to-nanoscale (enzyme catalysis), nanoscale-to-macroscale for the transport and macroscale for device integration. The figure on the right illustrates this hierarchy of processes. Enzymatic biofuel cells present a viable candidate for ultimate miniaturization because it is a class of energy conversion devices that employ molecular electrocatalysts – enzymes and hierarchical surface architectures developed in the course of the advancement of bio-nano interface technology.

15. SUBJECT TERMS
Biological Fuel Cells, Enzymatic Fuel Cells, Bio-nano Technology, Bio-devices for Energy Harvestig

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
UU	UU	UU	UU		19b. TELEPHONE NUMBER (Include area code)