Artificial Metabolons:

Design of Self-Assembled Bio-Complexes

Kristen E. Garcia

Submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2017

Abstract

Artificial Metabolons: Design of Self-Assembled Bio-Complexes

Kristen E. Garcia

Protein-protein interactions are vital to every living organism, and it is thought that most, if not all proteins interact in some way with other proteins for purposes including for cellular metabolism, signal transduction and DNA replication. These protein complexes can range in stability from permanent to transient, and they are driven by interactions at the protein-protein interfaces including hydrophobicity, hydrogen bonding, electrostatic interactions, van der Waals interactions and covalent disulfide bonding. Many complexes, such as transient complexes of sequential enzymes called metabolons, are poorly understood. In recent years, there have been many efforts to mimic nature and engineer new protein complexes with defined spatial arrangements with increased stability and more efficient transport of the enzymatic reaction intermediates. There is much to be understood in these complexes, including the role of substrate channeling. In this dissertation, we study a natural metabolon and engineer new protein complexes.

In our first study, we construct designed protein aggregates of the single enzyme small laccase (SLAC). SLAC is a multi-copper oxidase that can be easily genetically modified and is used as an oxygen-reduction catalyst on enzymatic bio-cathodes. A new dimeric interface is introduced, which, in combination with the threefold symmetry of the naturally trimeric SLAC, drives the self assembly of SLAC with two disulfide bonds in an oxidative environment. These enzymatically active aggregates form upon the addition of cupric ions to the purified protein, and electron microscopy shows the symmetry of the aggregates to be consistent with the design. We demonstrate improvements over the non-complexed enzyme including an increased resistance to

permanent thermal denaturation and a lower reaction overpotential and increased current density when employed on an oxygen-reduction bio-cathode with single-walled carbon nanotubes incorporated into the enzyme aggregates.

In our next line of work, we study a natural tricarboxylic acid (TCA) cycle metabolon, focusing on two enzymes: mitochondrial malate dehydrogenase (mMDH) and citrate synthase (CS). These enzymes have long been proposed to form a spatially organized complex that facilitates substrate channeling, a process in which a reaction intermediate is transferred directly from one enzyme active site to the next without first diffusing into the bulk through mechanisms such as electrostatic interactions. Structural evidence has been difficult to obtain due to the transient nature of many of these complexes. In Chapter 3, we examine the *in vitro* complex structure of the recombinant enzymes and find that it is similar to the recently proposed *in vivo* complex structure. Furthermore, there is evidence of a positively charged electrostatic channel connecting the enzyme active sites along which the oppositely charged reaction intermediate can travel by bounded diffusion. Site-directed mutagenesis along the channel on CS results in inhibited substrate channeling.

Finally, we develop a platform to study substrate channeling in engineered multi-enzyme complexes. Efforts to engineer multi-enzyme complexes in recent years have made use of protein and nucleic acid-based scaffolds. Many of these complexes exhibit increased coupled enzymatic activities, but there is a question of what effects are due to substrate channeling and how to apply these strategies to any enzyme pair. In this work, we attach CS and the non-channeling cytosolic malate dehydrogenase to DNA and engineered protein cage scaffolds. These assemblies retain their enzymatic activities, and these methods can be used to study substrate channeling in many enzyme pairs including the naturally channeling and inhibited channeling TCA cycle enzymes.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to take this opportunity to thank the many people whose help and support made this work possible. First, I want to thank my advisor, Scott Banta. Your support and advice over the last four and a half years has been invaluable, and I want to thank you for the opportunity to do this research in your lab.

Thank you to all my committee members, Drs. Alan West, Kyle Bishop, Oleg Gang and Henry Hess. I thank you for taking the time out of your busy schedules to serve on my committee, and I thank you in advance for your mercy.

Thank you to all of the Banta Lab members, past and present. You made this work possible. In particular, I want to acknowledge Beyza Bulutoglu, Harun Ozbakir and Tim Kernan for your constant support and friendship as well as Drs. Kevin Dooley and Tushar Patel for their guidance in my early days.

I would also like to thank the people that helped get me to graduate school: Drs. Plamen Atanassov, Maggie Werner-Washburne, Jared Roy, Bayo Falase and Rosalba Rincon.

Thank you to my parents, Lynda and Art, and my sisters, Weslie, Toma and Katya. You have always been in my corner cheering me on and supporting me in every way. You all are the best.

Finally, I want to thank my partner in life, Nick. You have been my rock, and I could not be here without you. Thank you.

Chapter 1

Introduction

1.1 Protein Complexes

Higher order protein complex formation is vital to all living organisms, and it has many functions in the cell including altering the specificity or kinetic properties of a protein, catalyzing metabolic reactions, signal transduction, chaperone-assisted protein folding and activation and suppression. Most proteins have more than one polypeptide chain, and many, if not all, proteins interact with other proteins through protein-protein interactions (PPIs) to form larger complexes. This higher order protein structure was first identified well before the primary, secondary and tertiary structures of proteins were discovered.[1,2] Interest in protein-protein complexes has risen significantly with the improvement of DNA sequencing technology, and PPIs have been reviewed extensively[3-7]

PPIs can be divided into several classifications: homo- and hetero-oligomeric, obligate and non-obligate, and transient and permanent complexes.[6,8,9] Homo-oligomers consist of multiple identical subunits, are symmetric in structure and are generally stable, while hetero-oligomers are comprised of non-identical subunits with varying levels of stability. Components of an obligate complex are not stable on their own *in vivo*, while components of non-obligate complexes can stably exist on their own. Classification of PPIs can further label them as either permanent or transient depending on the lifetime, or stability, of the complexes. Many non-obligate PPIs are classified as transient, and obligate PPIs are mostly permanent, with permanent and obligate classifications often used interchangeably throughout the literature.

PPIs are driven by the same interactions that are important to protein folding: hydrophobicity, hydrogen bonding, electrostatic interactions, van der Waals interactions and covalent disulfide bonding.[6] Protein-protein interfaces are geometrically and electrostatically

complementary across PPI classifications, and different types of interactions have varying levels of importance within each PPI classification. Protein-protein interfaces for permanent complexes tend to have larger interfacial areas with higher levels of hydrophobicity, and they may also include disulfide bonding. Transient complex interfaces tend to be less hydrophobic since the individual components are soluble and stable on their own, and salt-bridges and hydrogen bonds are more frequent in transient complexes than in permanent ones. Polar residues are common for transient complexes, and both permanent and transient complexes include charged residues in their protein-protein interfaces.

Equilibrium dissociation constants (K_D), defined as the ratio of the rate of complex dissociation and complex association (k_{off}/k_{on}), can be used to further differentiate between weak and strong transient complexes. Weak transient complexes are formed and broken continuously, and have K_D values in the millimolar to micromolar range, and strong transient complexes have nanomolar K_D values. Permanent complexes can have micromolar to femtomolar K_D ranges. These ranges in affinity and stability allow for PPIs to serve their wide range of functions throughout the cell.[9] Complexes are not present at all times, and PPIs, particularly in transient complexes, are largely dependent on their local environments.[3] This can make studying PPIs difficult, and weak transient complexes with millimolar K_D values are still poorly understood.

1.1.1 Engineering Protein Complexes

There have been significant efforts to engineer new protein-protein complexes with functions and and designs that are not found in nature. Many designs take advantage of the natural symmetry in

homo-oligomers.[10] The addition of small molecules and metal ions have made some of the designs dynamic, similar to what can be found in nature.[11,12] Design of these new protein complexes require introducing new protein-protein interfaces. Genetic fusions have been used to create assemblies with minimal design, however this strategy is limited when controlling for orientation.[13,14] The field has been greatly aided by computational developments. Components and designs can be chosen through docking, and residues at the chosen interfaces can be designed in order to create low-energy interfaces that drive assembly. These techniques have been used to create assemblies such as cages, dimers and three-dimensional crystals.[11,15-18]

1.1.2 Small Laccase

Laccases are multi-copper oxidases (MCOs) predominantly found in fungi and plants.[19] They have an affinity for have a wide range of phenolic compounds**.** Laccases have been used for oxygen-reduction cathodes due to their high redox potential, neutral pH range and the ability to participate in both mediated electron transfer, in which the electrons are shuttled between the electrode and enzyme active site by a chemical mediator, and direct electron transfer, in which the electrons are transferred directly to the enzyme active site.[19-23] Direct electron transfer heavily relies on the orientation of the enzyme. Fungal and plant laccases, however, are difficult to express in bacteria, making them less than ideal enzymes for protein engineering. The small bacterial laccase (SLAC) from *Streptomyces coelicolor*, named for its small size compared to fungal laccases due to its lack of the second of the three laccase domains, is easily expressed in *E. coli* and can be quickly engineered at a genetic level.[24]

SLAC is homo-oligomeric, with three identical subunits of 343 amino acids. The trimeric structure is integral to its activity as the active sites are located between monomers. Each trimer contains 12 copper ions, three trinuclear clusters (one type 2 and two type 3) and three of type 1, which give the enzyme its blue color. The coordination of the copper ions is heavily conserved across MCOs.[25] The substrate binding sites are located near the type 1 coppers where the substrate is oxidized. In total, four substrates are oxidized, and the four electrons are transferred to the trinuclear cluster by the triad His287-Cys288-His289, and an oxygen molecule is reduced to water at the trinuclear copper cluster.[26] SLAC has been genetically engineered in several ways including the site-directed mutagenesis for substrate speceficity,[27,28] the incorporation of a non-canonical amino acid for site-specific immobilization,[29] and the addition of DNA-binding and –helical domains for self-assembly through genetic fusions,[30,31] and it has been employed as an enzymatic catalyst for oxygen-reducing bio-cathodes with both mediated and direct electron transfer.[29,31,32]

1.2 Substrate Channeling in Multi-Enzyme Complexes

Enzyme complexes have been of interest for the past few decades, and these transient or permanent complexes of sequential enzymes have been styled "metabolons".[33] Multi-enzyme complexes have been discovered in primary and secondary pathways, and among the best-characterized metabolons are the pyruvate dehydrogenase, tryptophan synthase and glycine decarboxylase complexes, enzymes in the tricarboxylic acid (TCA) and Calvin cycles, and various components in fatty acid, nucleic acid and protein biosynthesis. A potential advantage of these defined spatial arrangements of sequential enzymes is substrate channeling, the process in which the reaction

intermediate is transferred directly from one enzyme active site to the next without diffusing in the bulk solution. Potential effects of substrate channeling can include the protection of unstable intermediates, the reduction of exposure of intermediates to competing pathways, the protection of the cell from toxic intermediates and the formation of local substrate concentrations that can serve to drive a reaction in a direction that may be otherwise thermodynamically unfavorable.[34,35]

Mechanisms for substrate channeling include intramolecular tunnels that sequester intermediates such as in tryptophan synthase,[36] swinging arms that transfer intermediates between active sites such as in the pyruvate dehydrogenase complex,[37] and electrostatic channeling, in which a charged intermediate can travel by bounded diffusion along a channel of oppositely charged residues such as in dihydrofolate reductase-thymidylate synthase (DHFR-TS).[38] Furthermore, while metabolon formation brings active sites into closer proximity, proximity alone is often insufficient for substrate channeling to occur without such mechanisms since diffusion is typically fast compared to reaction rates.[39-41] Leaky channeling, where some intermediate is not channeled, does occur. Confirming a substrate channeling hypothesis can be challenging due to the transient nature of many of these complexes, and, in the absence of structural evidence, multiple indirect kinetic methods to measure substrate channeling should be used.[35]

There have been several efforts to engineer new multi-enzyme complexes that mimic what is found in nature. Strategies to create these multi-enzyme complexes have ranged from simple fusion proteins to more complex protein-based scaffolds and nucleic acid scaffolds. Assemblies have been constructed *in vivo* and *in vitro*, and many have exhibited increases in coupled reaction activity upon assembly.[34,35,42] However, there remains some question of whether these complexes are truly channeling the reaction intermediates.

1.2.1 TCA Cycle Metabolon

The TCA cycle is the central metabolic pathway of aerobic organisms. Six out of eight mitochondrial enzymes in the cycle have been demonstrated to co-localize in the cell,[43] and it has been proposed that the enzymes form a spatially organized complex that enables substrate channeling and potentially drives the reaction in the forward direction. Mitochondrial malate dehydrogenase (mMDH) and citrate synthase (CS) in particular have been extensively studied for metabolon formation and substrate channeling[44-46] along with aconitase (Aco), which has been shown to associate with citrate synthase.[33]

mMDH is homo-oligomeric with two identical subunits of 314 amino acids.[47] It catalyzes the reversible oxidation of L-malate to oxaloacetate (OAA) using the NAD/NADH cofactor system and follows a sequential ordered mechanism with the cofactor binding to the enzyme first.[48] The reaction in the forward direction of malate to OAA has a highly unfavorable equilibrium constant ($K_{eq} = (2.86 \pm 0.12) \times 10^{-5}$),[49] which, along with the allosteric regulation of mMDH by citrate,[50] is further evidence that mMDH and CS participate in substrate channeling in order to drive the reaction in the forward direction.[51,52] An isoform of mMDH, cytosolic malate dehydrogenase (cMDH), which follows the same mechanism as mMDH, is often used for comparison when studying substrate channeling in TCA cycle enzymes.[53,54] The sequence homology of porcine mMDH and cMDH is only 19.6%, but the similarities increase when focusing on functionally important residues and quaternary structure.[55]

CS is the key enzyme of the TCA cycle catalyzing the Claisen condensation of acetyl-Coenzyme A with OAA to form citrate following an sequential ordered mechanism with OAA binding first.[56,57] CS is a homo-dimeric with 437 amino acids subunits. It has a high affinity for

OAA and undergoes a conformational change from an open form to a closed form upon binding to OAA.[58] The next enzyme in the cycle, Aco is monomeric with a [4Fe-4S] cluster to catalyze the two-step reversible dehydration and hydration of citrate to *cis*-aconitate to isocitrate,[59] and the overall reaction follows a three-step process (dehydrate, flip, rehydrate) where the flip is a multi-step 180° rotation of the intermediate *cis*-aconitate.[60] Each reaction follows Michaelis-Menten kinetics.

Indirect kinetic evidence and simulations have given support to the substrate channeling hypothesis by electrostatic channeling for mMDH and CS, but structural evidence of the TCA metabolon has been hindered due to the difficulties in isolating and purifying the transient protein complex. A new model has recently been proposed by Wu et al. which was determined by *in vivo* cross-linking and mass spectrometry.[61] This model is substantially different than what has been proposed previously.[33] With this new model, there is much that can be learned about the mechanism of substrate channeling in the complex.

1.3 Research Objectives

Enzyme complexes can have many benefits including increased stability and channeling of reaction intermediates. The design of novel enzyme complexes that mimic and expand what is seen in nature has been of considerable interest. However, much is still unknown and debated in the field. In this research, the main goals are to assemble and characterize single and multi-enzyme complexes. The specific goals of this thesis are outlined below along with their corresponding chapters within the thesis:

1. **Engineer and characterize a single-enzyme complex with long-range order.** In Chapter 2, we construct functional crystalline-like assemblies with the single enzyme SLAC by introducing new inter-protein disulfide bonds at the designed protein-protein interface to drive self assembly. This work investigates the kinetic activity of these complexes, and the electron transfer for these assemblies with incorporated carbon nanotubes was studied on oxygen-reducing bio-cathodes.

2. **Characterize the recombinant TCA cycle metabolon for substrate channeling.** In Chapter 3, we assemble a TCA cycle metabolon with recombinant enzymes and compare the metabolon structure to the recently discovered *in vivo* complex structure. Substrate channeling between mMDH and CS is probed by kinetic analysis in combination with site-directed mutagenesis.

3. **Construct a platform to study substrate channeling in engineered multi-enzyme complexes.** Chapter 4 discusses two strategies used to engineer multi-enzyme complexes on DNA and protein cage scaffolds. Complex assembly and enzymatic activities are evaluated.

Chapter 2

Designed Protein Aggregates Entrapping Carbon Nanotubes for Bioelectrochemical Oxygen Reduction

Note: A version of this chapter entitled "Designed Protein Aggregates Entrapping Carbon Nanotubes for Bioelectrochemical Oxygen Reduction" appeared in Biotechnology and Engineering volume 113, issue 11, pages 2321-2327.

Project Collaborators: Sofia Babanova, William Scheffler, Mansij Hans, David Baker, Plamen Atanassov, and Scott Banta. KEG expressed and purified all proteins, performed all colorimetric activity assays and analysis and prepared the manuscript for publication.

2.1 Abstract

The engineering of robust protein/nanomaterial interfaces is critical in the development of bioelectrocatalytic systems. We have used computational protein design to identify two amino acid mutations in the small laccase protein (SLAC) from *Streptomyces coelicolor* to introduce new inter-protein disulfide bonds. The new dimeric interface introduced by these disulfide bonds in combination with the natural trimeric structure drive the self-assembly of SLAC into functional aggregates. The mutations had a minimal effect on kinetic parameters, and the enzymatic assemblies exhibited an increased resistance to irreversible thermal denaturation. The SLAC assemblies were combined with single-walled carbon nanotubes (SWNTs), and explored for use in oxygen reduction electrodes. The incorporation of SWNTs into the SLAC aggregates enabled operation an elevated temperature and reduced the reaction overpotential. A current density of 1.1 mA/cm^2 at 0 V vs. Ag/AgCl was achieved in an air-breathing cathode system.

2.2 Introduction

A grand challenge in the development of efficient bioelectrocatalytic systems is the optimization of transport and reaction kinetics within robust electrode surface modifications. The incorporation of enzymes into these systems presents unique challenges including operational stability, poor compatibility with entrapping polymer matrices, and optimizing spatial arrangements with mediating or relay systems.[62,63] A variety of materials engineering approaches have been developed to better enable the incorporation of native enzymes into electrode systems,[64-66] and a few protein engineering approaches have been used to improve the protein/nanomaterial interface.[67-69]

Computational protein design has been successfully used to engineer almost every feature of proteins, including binding, stability and catalysis.[70-75] A recent trend is the use of protein design principles to create non-natural assemblies of protein structures to generate long-range order in self-assembling systems.[74,75] These unique protein structures show great promise for the development and improvement of bio/nano interfacial systems as these arrangements can lead to improved stability and optimized spatial orientations.

Multi-copper oxidase enzymes directly reduce oxygen and have been extensively explored for use in enzymatic bio-cathodes to be used in biological fuel cells (BFCs), bio-batteries and self-powered bio-sensors.[64,76,77] BFCs are devices that utilize biocatalysts for energy transformation.[78-81] Since the chemical energy stored in ubiquitous fuels available in the environment is transformed into electricity in BFCs, the practical application of these systems is often associated with energy harvesting. BFCs using enzymes as catalytic units are predominantly associated with the design of small devices, generating relatively high

power.[79] The critical drawback of enzymatic BFCs is their operational lifetime, which can be limited by the environmental stability of immobilized enzymes.

The fungal laccases have been well-studied for the development of bio-cathodes due to their high activity and redox potential, but the engineering of these proteins is more challenging as they generally do not express well in prokaryotic hosts.[82] Bacterial laccases, such as the small laccase (SLAC) from *Streptomyces coelicolor* can have lower redox potentials but are more active at neutral pH and are more amenable to protein engineering strategies.[83] Previously, the trimeric SLAC enzyme has been combined with single-walled carbon nanotubes (SWNTs) and also engineered to form self-assembling protein hydrogels with osmium redox centers, demonstrating its utility as a biocathode modification.[68,84,85] Here we use computational design to identify mutations for the introduction of new inter-protein disulfide bonds that would enable the SLAC enzyme to self-assemble into stable, functional crystalline-like assemblies.

2.3 Results and Discussion

2.3.1 SLAC Aggregate Design

The guiding principle in the assembly design process is to introduce a single new dimeric interface between SLAC protein molecules, which, in combination with the threefold symmetry of SLAC, will drive self-assembly of a material. Following the geometric constraints presented by Padilla et al.,[75] the SLAC enzyme can be assembled into a P 41 3 2 crystal architecture via a dimer interface between two trimers by constraining the angle

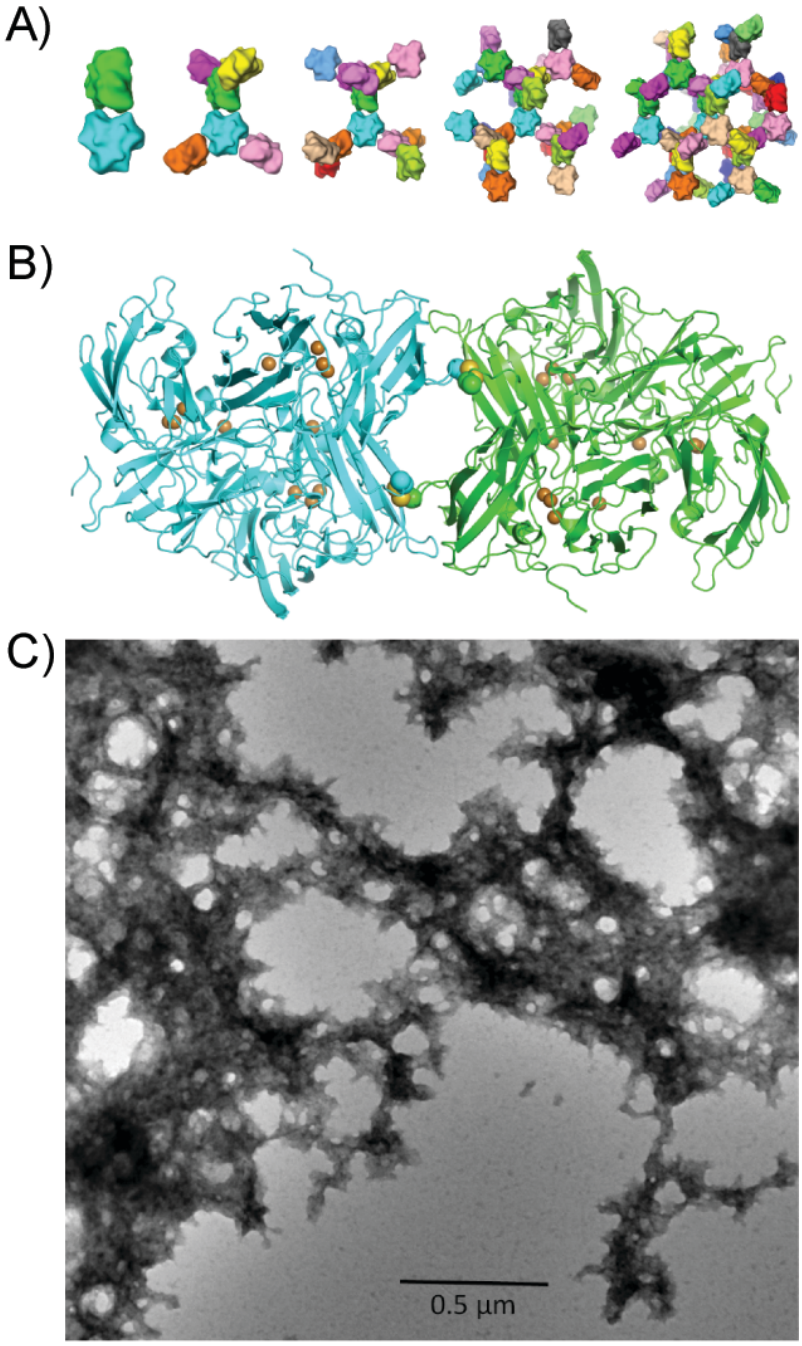

Figure 2.1: Design of the SLAC-DC-His assembly. (A) Illustration of one-interface design principle, wherein a single new interface gives rise to self-assembly. Shown is the stepwise application of the dimer interface to terminal SLAC-DC-His trimers, adding new trimers to the assembly until a fully connected crystalline architecture is attained. (B) Detail of the designed dimer interface between SLAC-DC-His trimers, showing the two disulfide linkages which form the interface as green, blue and yellow spheres. The active site copper atoms are also shown as gold spheres. (C)TEM image of SLAC-DC-His aggregates.

between the respective dimer and trimer symmetry axes to 35.3 degrees. The design principle is illustrated in Figure 2.1a. Beginning with two SLAC molecules in a two-fold symmetric arrangement, additional copies of SLAC bound with identical binding modes to the first pair, gradually build a coherent symmetric assembly.

Disulfide bonds were chosen to drive interface formation. This both simplifies the interface design problem and yields a very stable connection between subunits in an oxidizing environment. The symmetric angular constraint described above, along with the requirement that the trimers must be in contact with one another in order to form an interface, drastically limits the number of possible interfaces that can form the desired structure.[71] This limited set of compatible binding modes was searched exhaustively with Rosetta for shape complementarity and cross-interface positions capable of accommodating a disulfide bond.[70] One such disulfide-compatible geometry was found to be very similar to a crystal contact in the SLAC crystal structure (PDB id 3CG8), making it likely to be physically reasonable, and was selected as the design of choice. The chosen disulfide-mediated interface is shown in Figure 2.1b. The new mutations (G70C and A189C) were made to SLAC-His using site-directed mutagenesis and the resultant protein with a C-terminal polyhistidine tag was named SLAC-DC-His.

2.3.2 Characterization of SLAC-DC-His

Kinetic parameters for SLAC-DC-His and SLAC-His were determined in dilute solution with the substrate 2,6-dimethoxyphenol (DMP) (Table 2.1 and Figure S2.1). We have previously observed that the addition of an N-terminal polypeptide fusion can dramatically

decrease the turnover number (k_{cat}) compared to the wild-type while fusions to the C-terminus had a far less impact.[68,85] The double cysteine mutations in SLAC-DC-His did not significantly affect the turnover number compared to SLAC-His. However, the Michaelis constant K_M is larger for SLAC-DC-His, resulting in a decreased catalytic efficiency (k_{cat}/K_M).

Table 2.1: Kinetic parameters of SLAC-DC-His and SLAC-His for DMP oxidation[a]

Enzyme	K_M [mM]	k_{cat} [min^{-1}]	k_{cat}/K_M [$mM^{-1}min^{-1}$]
SLAC-DC-His	13 ± 2[b]	110 ± 10	8.2 ± 1.6[b]
SLAC-His	4.7 ± 0.4	120 ± 10	25 ± 2

[a] Measurements were performed in triplicate, and represented with their mean and standard deviation from the mean.

[b] $p < 0.01$, where statistical significance of SLAC-DC-His parameters compared to the control SLAC-His were calculated by unpaired two-sample t-test.

Purified SLAC-DC-His was studied for aggregate formation. Blue aggregates were seen to form in an oxidative environment when $CuSO_4$ was incubated with samples containing SLAC-DC-His (Figure S2.2a). TEM images of the aggregates showed assemblies with regions of 3-fold symmetry consistent with the computational design (Figure 2.1c). Protein aggregates were not seen to form when SLAC-His was exposed to the same conditions. SLAC-DC-His aggregates also formed in the presence of SWNTs after the addition of $CuSO_4$ (Figure S2.2b). Aggregates formed with and without SWNTs were found to be active with DMP (Figure S2.2d), and SWNTs were not found to significantly affect enzyme activity for DMP (Table S2.1). SLAC-DC-His aggregates were dissolved upon addition of the reducing agent dithiothreitol (DTT) (Figure S2.2c). Aggregates

exhibited a large variation of sizes (Figure S2.5). The aggregates did not appear to dissolve after storing them in buffer over a period of months.

2.3.3 RDE Experiments

In order to explore electrochemical function of the SLAC aggregates, rotating disk electrode (RDE) measurements were used to study the rate of oxygen reduction, which eliminates diffusional limitations and allows the study of oxygen reduction reaction activity in kinetically limited conditions. Three main electrochemical parameters were used for comparison: generated current density, onset and half-wave potentials.[86] The onset potential provides information for the thermodynamic aspect of the reaction, and a higher onset potential is an indication for decreased overpotential of the reaction or decreased activation energy. The generated current is descriptive for the kinetics of the oxygen reduction reaction, and higher current densities are indicative of improved reaction kinetics. The half-wave potential gives information for both the thermodynamics and kinetics of the process. It is the potential of a voltammetric curve at the point where the difference between the faradaic current and the non-faradaic current is equal to one-half of the limiting current.

To provide increased surface area and a better contact between the electrode and the enzyme aggregates, SWNTs were explored. This technique was applied to study the electrochemical activity of SLAC-DC-His aggregates when combined with SWNTs in three configurations: i) SLAC-DC-His aggregates were adsorbed on a surface modified with SWNTs; ii) pre-formed SLAC-DC-His aggregates were combined with SWNTs; iii) SWNTs were incorporated with SLAC-DC-His during aggregation. For each configuration,

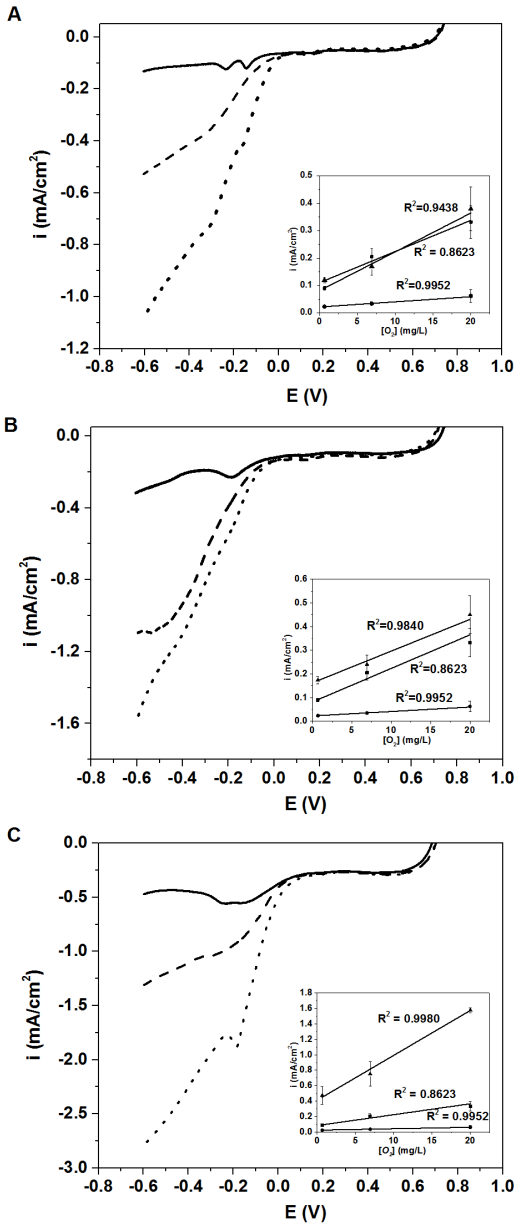

Figure 2.2: Oxygen-varied RDE measurements by electrode design. (A) SLAC-DC-His aggregates adsorbed on SWNTs, B) pre-formed SLAC-DC-His aggregates mixed with SWNT inks, and C) SWNTs incorporated into SLAC-DC-His aggregates. Experiments were performed in an oxygen depleted electrolyte (solid lines), oxygen saturated electrolyte (dash dotted lines) and an electrolyte exposed to air (dashed lines) (10 mV/s, 1600 rpm). Insets (reproduced for clarity in Figure S2.4) show the dependence of the current density at -0.15 V vs. Ag/AgCl on the oxygen concentration in the electrolyte with SLAC-DC-His aggregates (triangles), SLAC-His (squares) and SWNTs only (diamonds). Three independent identical electrodes for each cathode type were prepared and tested for reproducibility. The results are represented with the mean from the three measurements and the standard deviation from the mean.

current production from the oxygen reduction reaction was observed, and the current was linearly dependent on the oxygen concentration (Figure 2.2). Redox peaks observed with oxygen-depleted electrolyte were also observed without the enzyme catalyst and are likely due to impurities in the SWNTs (Figure S2.5). The enzymatic aggregates were stably absorbed on the rotating disk electrodes, and we did not visually observe loss of the deposited ink even at rotation rates of 1600 rpm.

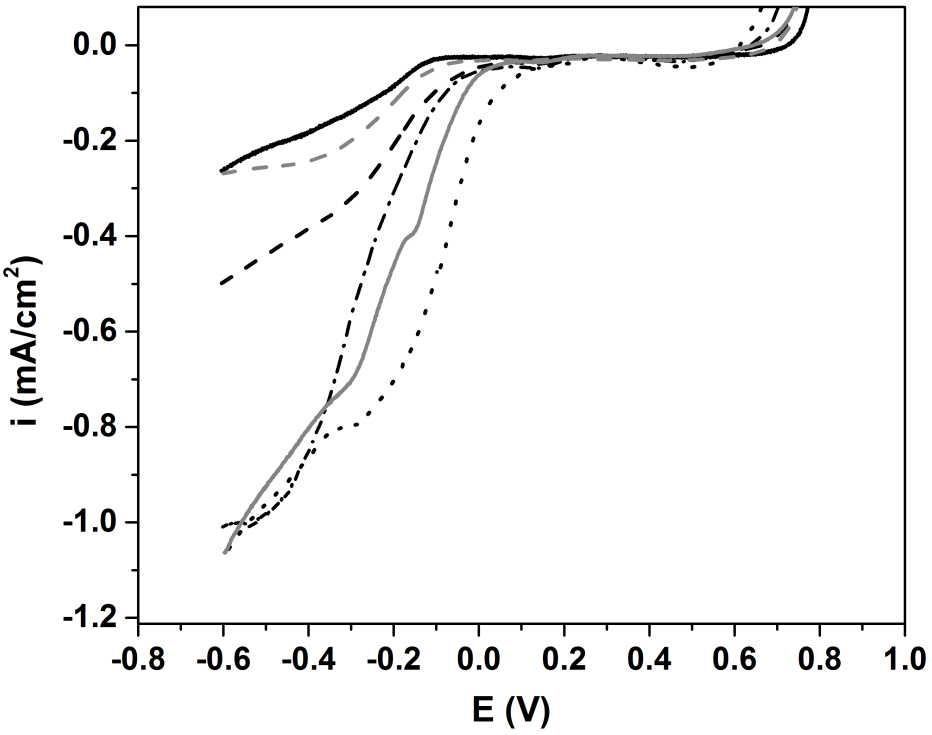

Figure 2.3: Comparison of RDE measurements. SWNTs only (solid black line), SLAC-His mixed with SWNTs (gray line), denatured SLAC-His mixed with SWNTs (gray dashed line), SLAC-DC-His aggregates adsorbed on SWNTs (dashed line) pre-formed SLAC-DC-His aggregates mixed with SWNT ink (dash dotted line) and SLAC-DC-His with incorporated SWNTs before aggregation (dotted line) in electrolyte exposed to ambient air (10 mV/s and 1600 rpm).

Comparing the electrochemical response of the SLAC-DC-His/SWNT composites, a sequential increase in the oxygen reduction reaction current densities can be seen,

demonstrating the importance of enzyme-nanomaterial interfacial interactions (Figure 2.3). The increase in current density observed when SLAC-DC-His aggregates are combined with SWNTs before being deposited onto the electrode is likely due to increased contacts between the enzymes and the SWNTs. Contact between the SWNTs and enzyme active sites is further increased in the case where SWNTs were introduced into a solution of purified SLAC-DC-His before aggregation was initiated. This enabled self-assembly to entrap SWNTs within the aggregates. This approach dramatically increased current densities during oxygen reduction. The generated current at -0.15 V and 19.8 ± 4.2 mg/mL oxygen content was 3.5 times higher than the preformed SLAC-DC-His aggregates mixed with SWNTs and 4.2 times higher than the physically adsorbed aggregates.

Table 2.2: Residual specific activities of SLAC-His and SLAC-DC-His for DMP oxidation[a]

Enzyme	25°C	70°C	98°C
SLAC-DC-His	1.0 ± 0.1	0.50 ± 0.08[b]	0.43 ± 0.08[b]
SLAC-His	1.0 ± 0.1	0.31 ± 0.05[b]	0.011 ± 0.004[b]

[a]Measurements were performed at 25°C after 30-minute incubation at elevated temperatures. Measurements were performed in triplicate, normalized by the activity at 25°C and represented with their mean and standard deviation from the mean.

[b]$p < 0.01$, where statistical significance compared to measurements at 25°C calculated by unpaired two-sample t-test.

The nonspecific incorporation of the SWNTs into SLAC-DC-His assemblies not only leads to enhanced current generation but also a significant positive shift in the onset

potential, from 0.075 V vs. Ag/AgCl for SLAC-His to 0.220 V vs. Ag/AgCl for SLAC-DC-His, and half-wave potential of approximately 100 mV, from - 0.168 V vs. Ag/AgCl for SLAC-His to -0.070 V vs. Ag/AgCl for SLAC-DC-His (Figure 2.3). These indicate a decrease of the reaction overpotential and improvement in kinetics are most likely due to the decreased tunneling distance and the orientation of the enzyme active site.[87] Based on the Butler-Volmer equation, a decrease in reaction overpotential by 100 mV leads to a 90 times increase in reaction kinetics and generated current.

2.3.4 Thermal Stability

SLAC-His and SLAC-DC-His aggregates were incubated at elevated temperatures and assayed in solution for DMP oxidation activity at room temperature in order to compare the resistance to irreversible thermal denaturation (Table 2.2). After incubation at 70°C, the residual activities of both samples decreased, with SLAC-His exhibiting a larger decrease in activity. This difference was more pronounced for the samples incubated at 98°C. The activity of the SLAC-His samples was reduced by 99% while the SLAC-DC-His samples retained 43% of the original activity.

Electrochemical activities were also measured at elevated temperatures for RDE electrodes (Table 2.3, Figure S2.6). At 50°C, there is no statistically significant difference in the retained activities of the SLAC-His/SWNTs and SLAC-DC-His/SWNTs composites. At 70°C, the SLAC-DC-His/SWNTs composite demonstrated higher electrochemical activity than the SLAC-His/SWNTs composite, retaining 47% activity. These results indicate that SLAC-DC-His aggregates are more resistant to thermal denaturation than

SLAC-His. Although it is unlikely that these electrodes would be operated at these higher temperatures, the increased resistance to thermal denaturation is likely due to the prevention of irreversible unfolding or the facilitated refolding enabled by the engineered disulfide bonds, and this would likely lead to increased operational lifetimes at ambient temperatures.

Table 3. Retained electrochemical activities[a]

Enzyme	25°C	50°C	70°C
SLAC-DC-His	1.0 ± 0.1	0.77 ± 0.10[b]	0.47 ± 0.03[a]
SLAC-His	1.0 ± 0.1	0.81 ± 0.02[b]	0.05 ± 0.04[a]

[a] Where SWNTs were incorporated in SLAC-DC-His aggregates prior to aggregation. Measurements were performed in triplicate, normalized by the activity at 25°C and represented with their mean and standard deviation from the mean.

[b] $p < 0.01$, where statistical significance compared to measurements at 25°C calculated by unpaired two-sample t-test. [b] $p < 0.05$, where statistical significance compared to measurements at 25°C calculated by unpaired two-sample t-test.

2.3.5 Gas-Diffusion Cathodes

Gas diffusion cathodes are advantageous due to the higher oxygen concentration available for reduction at the cathode compared to in solution. Therefore, the SLAC-DC-His assembly with incorporated SWNTs was integrated into the design of an air-breathing cathode and compared to a SLAC-His air- breathing cathode via steady state polarization curves (Figure 2.4). The current density at 0.0 V vs. Ag/AgCl recorded with the gas-diffusion SLAC-DC-His cathode was 1.1 mA/cm^2, which was significantly higher than the SLAC-His cathode. The SLAC-DC-His gas-diffusion cathode outperformed the SLAC-His cathode design throughout the potential region tested.

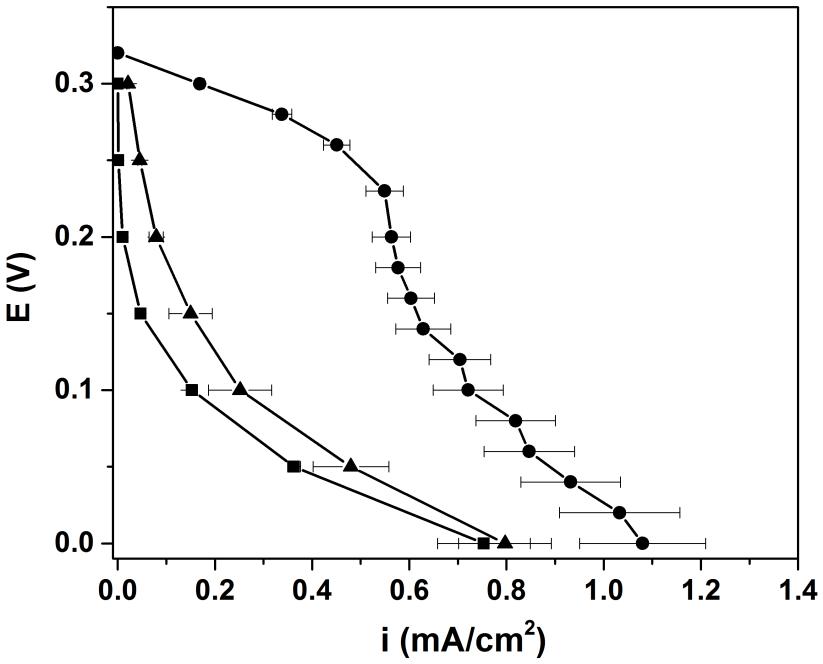

Figure 2.4: Potentiostatic polarization curves. SLAC-DC-His aggregates (circles), SLAC-His (triangles) and neat SWNTs (squares) when incorporated into the design of a gas-diffusion cathode with 9.2 μg of enzyme deposited on each cathode. Three independent identical electrodes for each cathode type were prepared and tested for reproducibility. The results in the figure are represented with the mean and the standard deviation from the mean.

This result can be compared to previous reports in the literature. A laccase-based gas-diffusion cathode with an identical design to the one used in this study generated 0.70 mA/cm^2 using air as the source of oxygen.[88] The interactions between the enzyme and the carbon black material relied on physical adsorption of the enzyme and direct electron transfer from the electrode to SLAC. To improve the direct electron transfer rate in the current design, we incorporated SWNTs along with the protein engineering approach and thus achieved 1.5 times higher current densities in comparison to the same miniature GDE exploring laccase as catalyst. Bilirubin oxidase (BOx) gas-diffusion cathode where BOx was physically adsorbed on the electrode surface demonstrated 0.35-0.50 mA/cm^2.[89] When an orienting agent (syringaldazine) and a tethering agent

(1-pyrenebutanoic acid, succinimidyl ester) were used for more efficient enzyme orientation, improved electron transfer and stability, the current densities of a BOx-GDE reached 0.70 mA/cm^2.[90] Using more complex tethering agent 4,4′-[(8,16-dihydro-8,16-dioxodibenzo[a,j]perylene-2,10-diyl)dioxy] dibutyric acid di(N-succinimidyl ester (DDPSE) and no orienting agent in the construction of laccase cathode demonstrated ~0.44 mA/cm^2.[64]

2.4 Conclusions

In conclusion, we have used computational protein design to create a SLAC double mutant that can self-assemble into active enzymatic crystalline-like assemblies. The enzyme aggregates were easily immobilized by physical adsorption, enabling high enzyme loadings. Additionally, the nonspecific incorporation of SWNTs into the aggregates led to improved kinetics from increased enzyme-SWNT contacts. The self-assembly increased resistance to thermal denaturation. A current density of 1.1 mA/cm^2 at 0 V vs. Ag/AgCl was achieved in an air-breathing cathode system. This designed self-assembly approach could be employed to create new biomaterials from other enzymes for use in biofuel cells as well as many other biocatalysis applications.

2.5 Materials and Methods

2.5.1 Materials

Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). All cloning enzymes and *E. coli* BL21 (DE3) competent cells were from New England Biolabs (Ipswich, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) was from Gold Biotechnology (St. Louis, MO). HALT protease inhibitor, Pierce Coomassie (Bradford) Protein Assay Kit, precast NuPAGE SDS-PAGE gels, NuPAGE SDS MES running buffer and Novex Sharp Pre-stained Protein Standard were from Thermo Fisher Scientific (Waltham, MA). HisTrap FF columns and the ÄKTA FPLC system were purchased from GE Healthcare (Piscataway, NJ). Uranyl acetate was purchased from Polysciences Inc. (Warrington, PA). Carbon black Vulcan X72 was purchased from Cabot Corporation (Boston, MA), Carbon coated 400 mush Cu/Rh grids were purchased from Ted Pella Inc (Redding, CA). All other chemicals were purchased from Sigma-Aldrich (St. Louis, MO) at the highest purity unless otherwise specified.

2.5.2 Construction, Expression and Purification of Enzymes

The pSLAC-His vector encodes for the SLAC enzyme from *Streptomyces coelicolor* with a C-terminal His_6-tag.[68] The Gly70 and Ala189 amino acids were mutated to cysteine by site-directed mutagenesis generating pSLAC-DC-His (Table S2.2). The vector pSLAC-DC-His was transformed into *E. coli* BL21(DE3) cells for SLAC-DC-His protein expression.

Cells were grown in 2xYT media at 37°C. At OD_{600} ~1.5, protein expression was induced by the addition of 0.4 mM IPTG, and the cells were grown an additional 20 hours at 25°C. Cells

were collected by centrifugation and stored at -20°C. Frozen cell pellets corresponding to 1 L of culture were thawed on ice for 30 minutes and the suspended in 50 mL of Buffer A (20 mM sodium phosphate buffer, 40 mM imidazole, 50 mM NaCl at 7.3 mM) with HALT EDTA-free protease inhibitor. Cells were sonicated for 6 minutes using a microtip probe, and lysates were centrifuged for 30 minutes at 15,000 x *g*. Lysates were applied to a HisTrap FF column equilibrated with Buffer A, and bound proteins were eluted with a linear gradient of 0 to 100% Buffer B (20 mM sodium phosphate buffer, 500 mM Imidazole, 500 mM NaCl at 7.3 mM) using an ÄKTA FPLC system. Fractions containing SLAC-His or SLAC-DC-His shown by SDS PAGE with bands at 38 kDa (Figure S2.7) were pooled, dialyzed four times against 10 mM sodium phosphate buffer at pH 7.3 and concentrated by ultrafiltration.

Purified SLAC-DC-His and SLAC-His solutions were incubated with $CuSO_4$ corresponding to five times the concentration of the enzyme at 4°C for several hours. SLAC-DC-His aggregates were collected by centrifugation (Figure S2.2a). SLAC-His was dialyzed against 50 mM ammonium bicarbonate buffer and lyophilized for long-term storage. Protein concentrations were determined by Bradford assay with bovine serum albumin standards for both SLAC-His and SLAC-DC-His and by absorbance at 590 nm with the extinction coefficient 4400 M^{-1} cm^{-1} for SLAC-His after the addition of $CuSO_4$.[83]

2.5.3 Kinetic Characterization by Colorimetric Assay

The activity of SLAC-DC-His aggregates for 2,6-Dimethylphenol (DMP) was verified visually after the addition of 10 mM DMP, and SLAC-DC-His aggregates formed in the presence of single-walled carbon nanotubes were also active (Figure S2.2d). 20 μL samples containing 1.8 mg of

SLAC-His and pre-formed SLAC-DC-His aggregates were incubated with an equal volume of SWNT ink for 1 hour at room temperature. Samples were also prepared in which SLAC-DC-His aggregates were formed after incubating SLAC-DC-His with the SWNT ink. Samples were then diluted 100x and assayed in 10 mM DMP. Measurements were performed in triplicate on a Spectramax M2 plate reader (Molecular Devices, Sunnyvale, CA), and reactions were monitored at 468 nm for the formation of the dimeric product 3,3',5,5'-tetramethoxydiphenylquinone (Figure S2.8) with the extinction coefficient 14,800 M^{-1} cm^{-1}. Activities were normalized by their activity without SWNTs (Table S2.1).

Samples containing 200 nM of SLAC-DC-His aggregates or SLAC-His in 10 mM sodium phosphate buffer, pH 7.3, were incubated at 25°C, 70°C, and 98°C for 30 minutes Samples were then cooled on ice for 10 minutes. Three samples at each temperature were assayed for residual activity in 10 mM DMP as described above. Average residual activities were normalized with respect to the residual activities at 25°C.

Dilute protein solutions were prepared to determine kinetic activities by adding $CuSO_4$ to 200 nM SLAC-DC-His. Protein solutions were incubated for several hours at 4°C with $CuSO_4$ corresponding to 5 times the concentration of enzyme. SLAC-DC-His activity was determined using concentrations ranging from 0 – 100 mM DMP in air-saturated 10 mM sodium phosphate buffer pH 7.3 with a final enzyme concentration of 20 nM. Measurements were performed in triplicate, and reactions were monitored at 468 nm. All data were fit to the Michaelis-Menten equation using SigmaPlot nonlinear regression software (Figure S2.1).

2.5.4 Transmission Electron Microscopy (TEM)

SLAC-DC-His aggregates with a final concentration of 0.7 mg/mL following a 10x dilution were negatively stained with 1% uranyl acetate in water on a carbon coated 400 mesh Cu/Rh grids that had been glow discharged for 1 minute before use. Samples were imaged on a Philips CM12 Tungsten Emission TEM (FEI, Eindhoven, Netherlands) at 120 kV with a Gatan 4 k x 2.67 k digital camera (Gatan Inc., Pleasanton, CA).

2.5.5 Aggregate Size Analysis

Samples of SLAC-DC-His with a concentration of 7 mg/mL were incubated for 20 minutes on ice, room temperature, and 50°C. $CuSO_4$ was then added and samples were incubated for 1 hour under the same conditions. Samples were diluted 100x then imaged using an Olympus CKX41 microscope with a Canon EOS 60D camera. At least 6 images per sample were analyzed with ImageJ to determine the area of 800 aggregates (Figure S2.3)

2.5.6 Rotating Disk Electrode (RDE) Measurements

Three different configurations for SLAC-DC-His immobilization and incorporation into the SWNTs matrix were utilized: i) SLAC-DC-His aggregates were adsorbed on a surface modified with SWNTs; ii) pre-formed SLAC-DC-His aggregates were combined with SWNTs; iii) SWNTs were incorporated with SLAC-DC-His during aggregation. For the first configuration, 30 μL of the SWNT ink was deposited on the surface of RDE and dried under nitrogen. Then 20 μL (0.9 mg) of SLAC-DC-His aggregates in 0.1 M phosphate buffer (pH 7.5) was dropped on the SWNTs

surface and dried at ambient conditions. For the second configuration, 20 μL of the SWNT ink was deposited on the surface of RDE and dried under nitrogen. 1.8 mg of SLAC-DC-His after their aggregation were dissolved in 20 μL of 0.1 M phosphate buffer (pH 7.5), mixed with 20 μL of the SWNT ink and left at ambient temperature for 1 hour. Then 20 μL of this enzyme-nanotubes assembly was dropped on the surface of the already placed SWNTs and dried under ambient conditions. For the final configuration in which the SWNTs were incorporated with SLAC-DC-His prior to their aggregation, 20 μL of the SWNT ink was deposited on the surface of the RDE and dried under nitrogen. 1.8 mg of SLAC-DC-His before aggregate formation was initiated in 20 μL of 0.1 M phosphate buffer (pH 7.5) was mixed with 20 μL of the SWNT ink and left at ambient temperature for 1 hour. $CuSO_4$ was added to the mixture and incubated so SLAC-DC-His aggregates can form in the presence of the SWNTs. Then 20 μL of this enzyme-nanotubes assembly was dropped on the surface of the already placed SWNTs and dried under ambient conditions.

An enzyme-free control electrode was prepared by depositing 20 μL of the SWNT ink on the surface of the RDE and then dried under nitrogen. A second control electrode with the control enzyme SLAC-His was also prepared. 20 μL of the SWNT ink was deposited on the surface of RDE and dried under nitrogen. 1.8 mg of SLAC-His were dissolved in 20 μL of 0.1 M phosphate buffer (pH 7.5), mixed with 20 μL of the SWNT ink and left at ambient temperature for 1 hour. Then 20 μL of this enzyme-nanotubes assembly was dropped on the surface of the already placed SWNTs and dried under ambient conditions. A third control electrode with denatured SLAC-His was also prepared. SLAC-His was incubated at 98°C for 1 hour in the presence of 8 M urea, dialyzed three times against 0.1 M phosphate buffer (pH 7.5) and checked for loss of activity for DMP. 20 μL containing 1.8 mg of denatured SLAC-His was mixed with 20 μL of the SWNT ink

and left at ambient temperature for 1 hour. Then 20 μL of this enzyme-nanotubes assembly was dropped on the surface of the already placed SWNTs and dried under ambient conditions. Each of the studied electrodes had 0.9 mg of enzyme per electrode.

All electrochemical measurements were performed in a three-electrode configuration with the cathode connected as working electrode, saturated Ag/AgCl and Pt-wire as reference and counter electrodes, respectively. 0.1 M phosphate buffer (pH 7.5) was used as electrolyte. For the RDE measurements, Pine Research Instrumentation rotator (Grove City, PA) coupled with a VersaSTAT 3 potentiostat (Princeton Applied Research, Oak Ridge, TN). Linear sweep voltammetry from 0.8 to -0.60 V vs. Ag/AgCl with 10 mV/s and different rotating rates (Figure S2.9) was carried out in oxygen depleted, oxygen saturated and an electrolyte with dissolved oxygen. Reverse scans were initially performed, but it was found that they did not provide any additional information and were excluded from subsequent measurements for simplicity. RDE measurements of SWNT-modified electrodes without enzyme were performed as a control (Figure S2.5). The oxygen concentration during the RDE measurements was monitored using DO probe. Three independent identical electrodes for each cathode type were prepared and tested for reproducibility.

2.5.7 Temperature-Controlled RDE Measurements

The modified RDE electrodes were introduced into electrochemical cell, and the temperature of the electrolyte was kept for 30 minutes at 25, 50 and 70°C, respectively, using a water jacket. Linear sweep voltammetry from 0.8 to -0.60 V vs. Ag/AgCl with 10 mV/s and 1600 rpm was carried out in oxygen saturated electrolyte (Figure S2.8). Current densities at -50 mV with the

SWNT control current density subtracted were used as comparison. Three independent identical cathodes were tested at each temperature for reproducibility.

2.5.8 Gas-Diffusion Cathodes

SLAC-DC-His was incorporated in the design of gas-diffusion cathodes. The cathodes were composed of plastic case, carbon black Vulcan X72, modified with 35% of polytetrafluoroethylene referred in this study as XC35 for the development of the gas-diffusion layer (GDL) and 100 mg of XC35 covered by a mixture of 50 mg XC35, 50 mg SWNTs and SLAC-DC-His aggregates formed in presence of SWNTs as it was described before, or SLAC-His explored as a catalytic layer (CL) (Figure S2.10). The CL was pressed at 500 psi for 5 min and circular discs with 0.15 cm diameter were cut for the cathode assembly. The final loading of SLAC-DC-His or SLAC-His per electrode was 9.2 μg of enzyme. Ni wire going from the catalytic layer through the GDL was placed for electrical contact. This cathode design allows manufacturing of multiple electrodes from a single XC35-carbon black-SLAC pressed tablet and thus increases the reproducibility of the preparation procedure. Pure oxygen was blown from the outer side of the GDL while potentiostatic polarization curves were obtained using a Gamry 600 potentiostat (Warminster, PA). Three independent identical electrodes for each cathode type were prepared and tested for reproducibility.

2.6 Supplemental Information

3.6.1 Supporting Figures

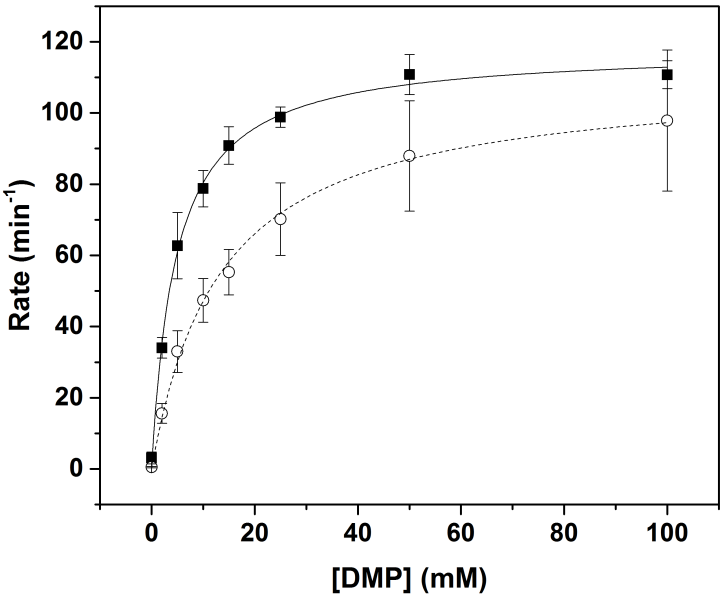

Figure S2.1: Dilute solution activity of SLAC-His (black circles) and SLAC-DC-His (white circles) with DMP in air-saturated solutions fit to the Michaelis-Menten equation. Measurements were performed in triplicate, and results are represented with the mean and the standard deviation from the mean.

\

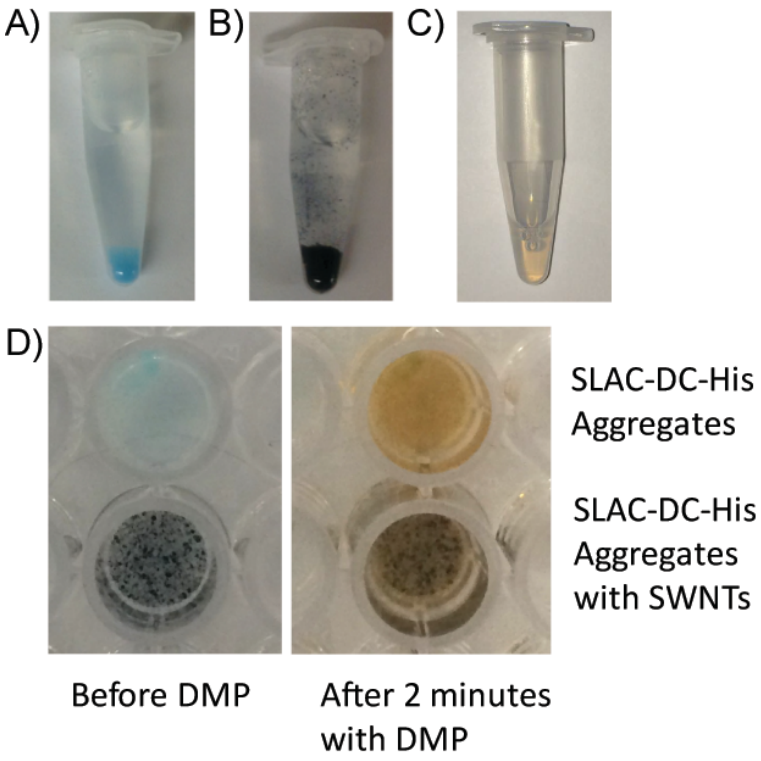

Figure S2.2: A) SLAC-DC-His aggregates, B) aggregates formed in the presence of 1% w/v SWNTs and C) aggregates disrupted by the reducing agent DTT at 100x molar excess. D) Active SLAC-DC-His aggregates with (bottom) and without (top) SWNTs before (left) and after (right) the addition of 10 mM DMP. Brown product formation is visible after DMP is added to the wells.

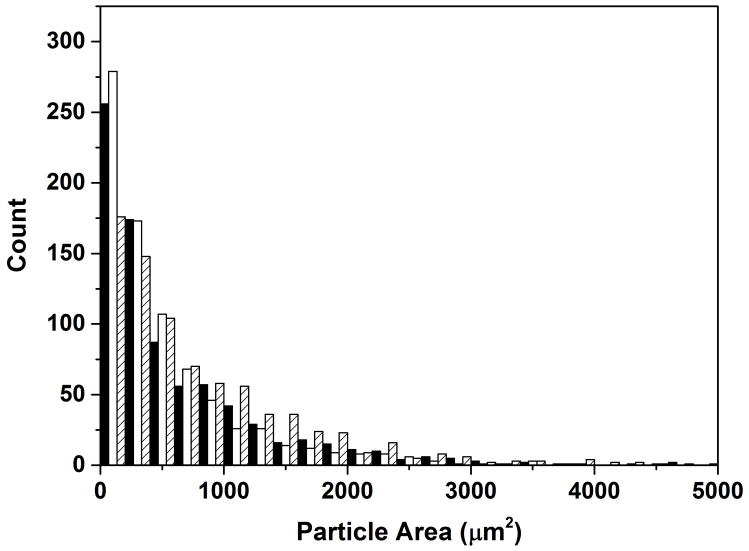

Figure S2.3: Histogram of SLAC-DC-His aggregate areas determined by image analysis with ImageJ for aggregates formed on ice (black), room temperature (white) and a water bath at 50°C (hatch pattern). One sample was prepared for each temperature and at least 6 images per sample were analyzed with ImageJ to determine the area of 800 aggregates.

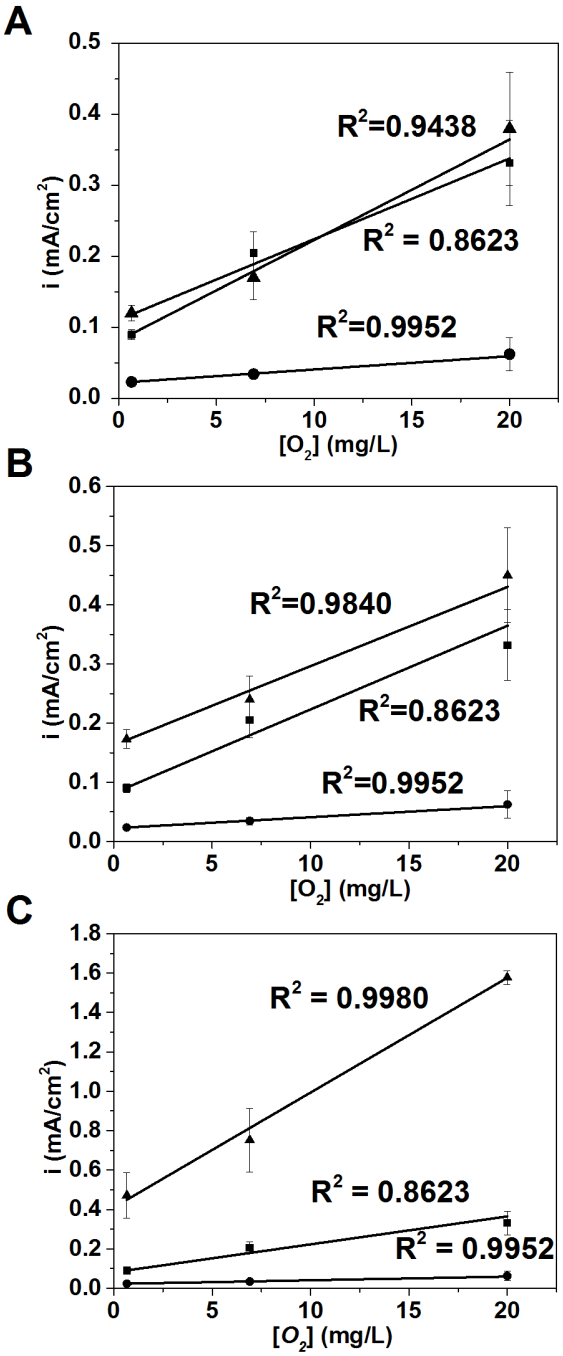

Figure S2.4: Insets reproduced from Figure 2.2 show the dependence of the current density from RDE measurements at -0.15 V vs. Ag/AgCl on the oxygen concentration in the electrolyte for SLAC-His (squares), SWNTs only (diamonds) and three configurations of SLAC-DC-His aggregates (triangles): A) SLAC-DC-His aggregates adsorbed on SWNTs B) pre-formed SLAC-DC-His aggregates mixed with SWNT ink, and C) SWNTs mixed with SLAC-DC-His before aggregation. The results are represented with the mean from the three independent measurements and the standard deviation from the mean.

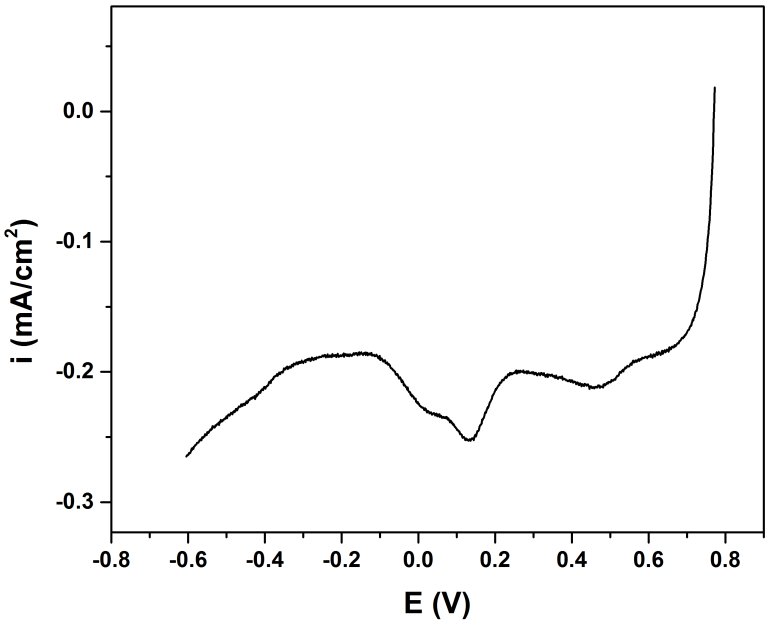

Figure S2.5: RDE measurements of SWNT ink in an oxygen-depleted electrolyte (10 mV/s, 1600 rpm).

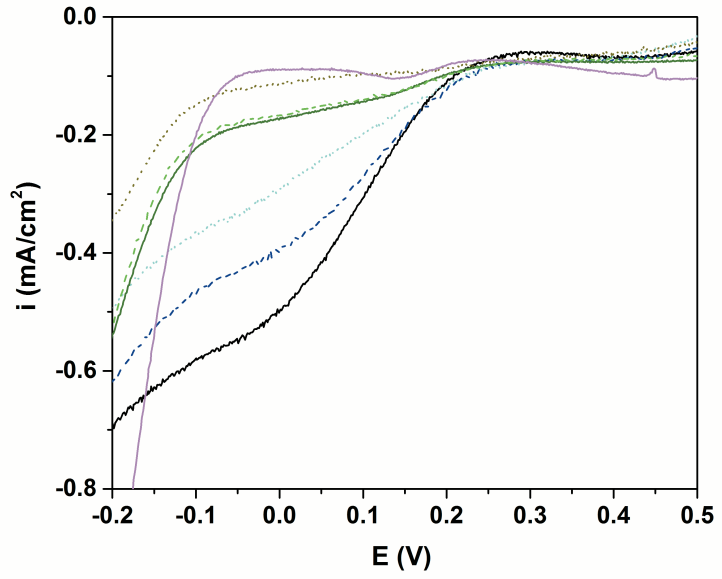

Figure S2.6: Temperature-controlled RDE measurements (10 mV/s and 1600 rpm) of: SLAC-DC-His/SWNTs where SWNTs were incorporated onto the SLAC-DC-His aggregates at 20°C and measured at 25°C (solid black), 50°C (dash dark blue) and 70°C (dotted light blue), SLAC-His/SWNTs composite at 25°C (solid dark green), 50°C (dash light green) and 70°C (dotted dark yellow) along with a control of SWNTs at 25°C (solid pink).

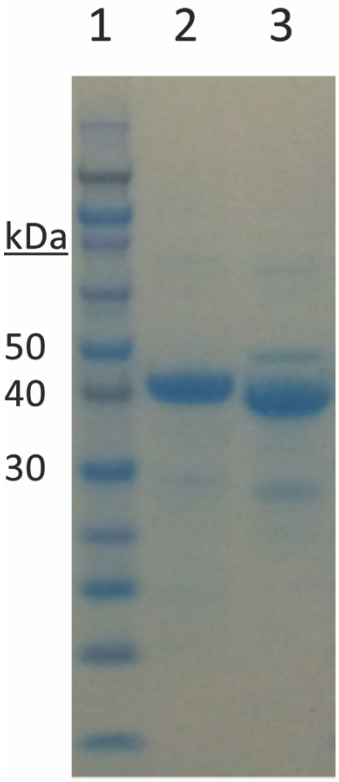

Figure. S2.7: 4-12% Bis-Tris SDS-PAGE of SLAC-His (Lane 2) and SLAC-DC-His (Lane 3) after purification. Proteins migrated to their expected molecular weights compared to Novex Sharp Protein Standard (Lane 1).

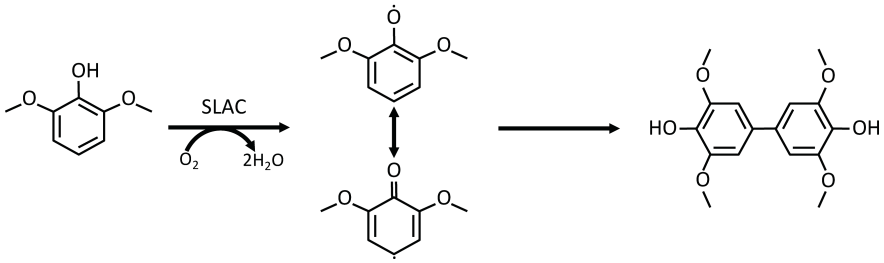

Figure S2.8: Scheme for the SLAC-mediated oxidation of DMP to the dimeric colored product 3,3',5,5'-tetramethoxydiphenylquinone.

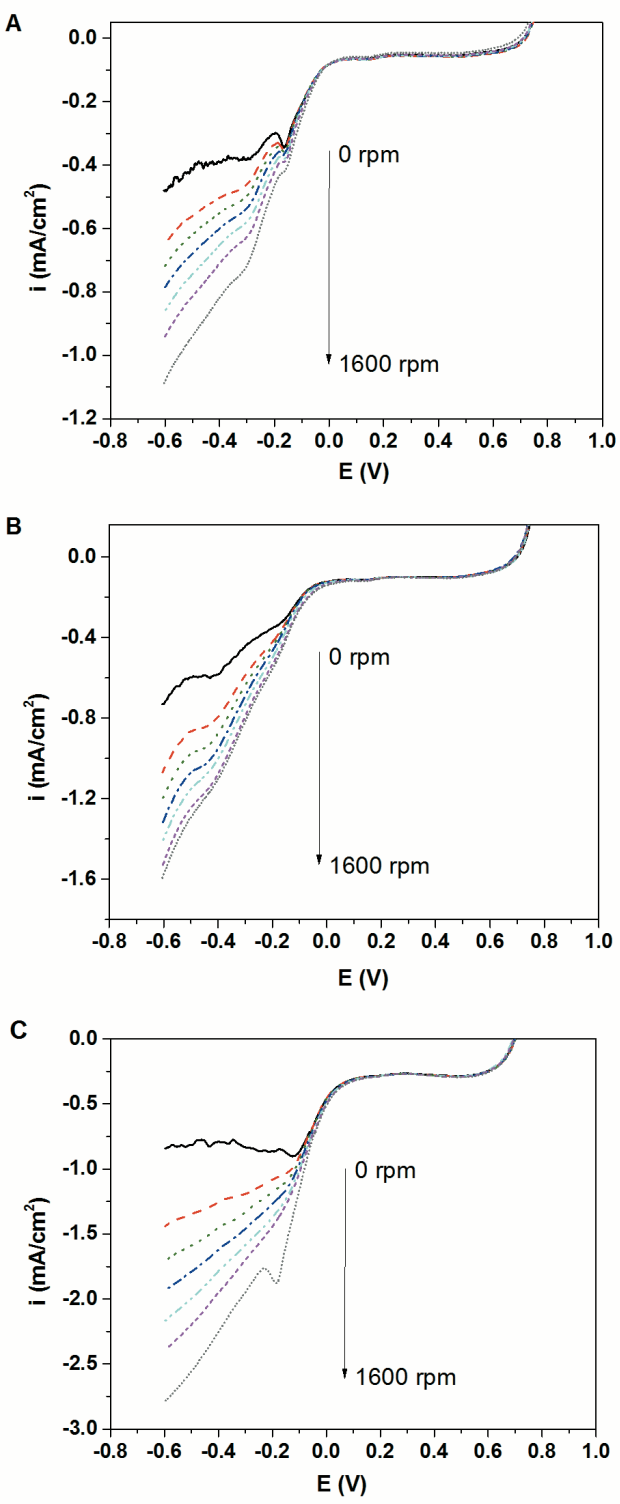

Figure S2.9: RDE measurements of A) adsorbed SLAC-DC-His aggregates on SWNTs, B) preformed SLAC-DC-His aggregates mixed with SWNT ink and C) SLAC-DC-His with incorporated SWNTs before aggregation, in oxygen-saturated electrolyte (10 mV/s, different rotating rates).

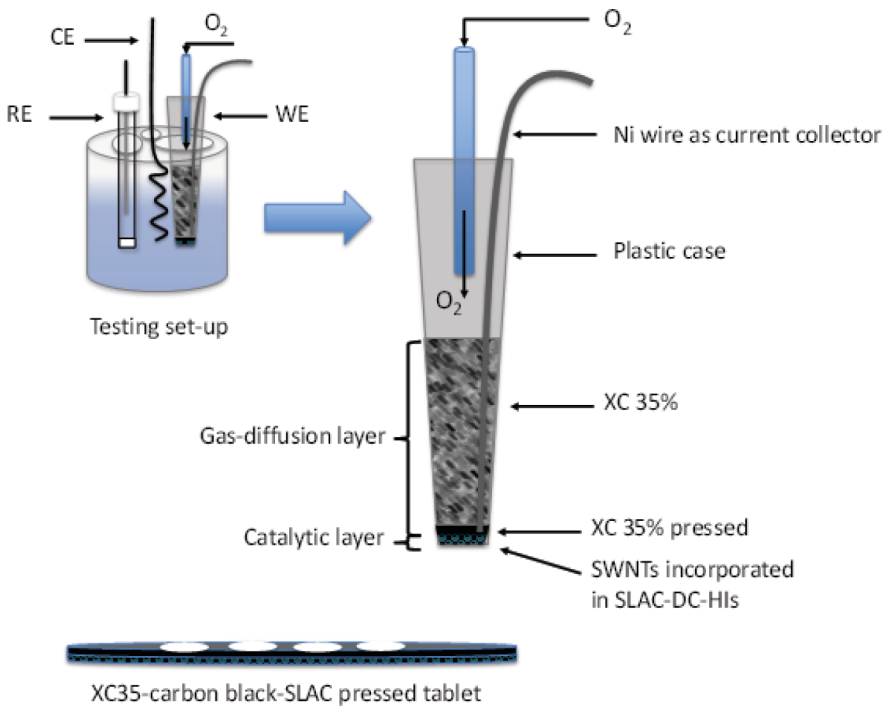

Figure S2.10: Schematic diagram of the gas diffusion electrode configuration and operation.

2.6.2 Supporting Tables

Table S2.2: Activity of SLAC-His and SLAC-DC-His for DMP in the presence of SWNTs and normalized to their activity without SWNTs.[a]

	No SWNTs	Mixed with SWNT ink	Mixed with SWNT ink before aggregation
SLAC-DC-His	1.0 ± 0.2	0.8 ± 0.1	0.8 ± 0.2
SLAC-His	1.00 ± 0.01	0.9 ± 0.1	N/A

[a] Measurements were performed in triplicate and represented with their mean and standard deviation from the mean.

Table S2.1: Oligos used for site-directed mutagenesis. Mutations are in italics.

Mutation	Oligo
G70Cs	5' - CGG CTT CGA GAA G*T*G CAA GGC GTC GG - 3'
G70Cas	5' – ACC GAC GCC TTG C*A*C TTC TCG AAG CC – 3'
A189Cs	5' - CGT GCT GCC GGA C*TG* CAC GCA CAC G– 3'
A189as	5' - GAT CGT GTG CGT G*CA* GTC CGG CAG C – 3'

Chapter 3

Direct Evidence for Metabolon Formation and Substrate Channeling in Recombinant TCA Cycle Enzymes

Note: A version of this chapter entitled "Direct evidence for metabolon formation and substrate channeling in recombinant TCA cycle enzymes" appeared in ACS Chemical Biology volume 11, issue 10, pages 2847-2853.

Project Collaborators: Beyza Bulutoglu, Fei Wu, Shelley D. Minteer and Scott Banta. KEG expressed and purified proteins, performed and analyzed all single-enzyme colorimetric activity assays, and assisted in analyzing transient time data and writing the manuscript.

3.1 Abstract

Supramolecular assembly of enzymes into metabolon structures is thought to enable efficient transport of reactants between active sites via substrate channeling. Recombinant versions of porcine citrate synthase (CS), mitochondrial malate dehydrogenase (mMDH) and aconitase (Aco) were found to adopt a homogeneous native-like metabolon structure *in vitro*. Site-directed mutagenesis performed on highly conserved arginine residues located in the positively-charged channel connecting mMDH and CS active sites led to the identification of CS(R65A) which retained high catalytic efficiency. Substrate channeling between the CS mutant and mMDH is severely impaired and the overall channeling probability decreased from 0.99 to 0.023. This work provides direct mechanistic evidence for the channeling of reaction intermediates and disruption of this interaction would have important implications on the control of flux in central carbon metabolism.

3.2 Introduction

Enzymes frequently function in sequential, multi-step cascades and the co-localization of the enzymes in self-assembling clusters is often observed.[91-93] The term "metabolon" has been used to describe these non-covalent dynamic enzyme complexes.[94] These arrangements enable substrates to be channeled between active sites without escaping into the medium.[95] When the intermediate transport is not 100% efficient, leaky channeling can occur, but the intermediates are sequestered enough to prevent equilibrium with the surroundings.[95-97]

Substrate channeling within metabolons results in several metabolic advantages. High local substrate concentrations enable better fluxes through a pathway, despite unfavorable equilibrium constants. Intermediates can be protected from the bulk phase, hindering competition from alternative pathways and protecting the cell from toxicity. These effects on the mass transport allow the enzymes to operate at high efficiencies even when the average concentrations of intermediates in the bulk phase are low, resulting in the improvement of overall catalytic efficiency of the metabolic process.[96-100]

Metabolons exist in many pathways, including glycolysis, fatty acid oxidation, amino acid metabolism, lipid biosynthesis and the tricarboxylic acid (TCA) cycle.[95,101] Several enzymes of the TCA cycle participate in metabolon formation, including citrate synthase (CS), mitochondrial malate dehydrogenase (mMDH) and aconitase (Aco).[102-105] Since they play a central role in cellular energy generation, metabolons of the TCA cycle have been well-studied.[44,45,51,52,106-109] The CS/mMDH interactions are of particular interest since the free oxaloacetate (OAA) intermediate concentration in the cell is thought to be too low to sustain the experimentally determined cycle rate and the mMDH reaction has an unfavorable equilibrium constant in the

forward direction of the cycle.[51,52]

Most metabolon investigations employ indirect techniques to infer channeling.[98,104,108] Fundamental characterizations of substrate channeling have focused on enclosed channels such as the tunneling that occurs in tryptophan synthase.[36,110,111] However, the bounded diffusion mechanism within the metabolon is more relevant to most biological systems. And, it is becoming clear that these "leaky channeling" systems are inspiring new approaches in biocatalysis where coupled reaction/transport systems are being engineered with biomimetic substrate channeling pathways.[97] Recently, the first structural characterization of a native mMDH-CS-Aco TCA cycle complex was reported.[112] In this work, we created and characterized recombinant versions of the TCA cycle enzymes and this enables the use of site-directed mutagenesis to explore structural determinants of substrate channeling *in vitro*.

3.3 Results and Discussion

The TCA cycle enzymes are a canonical example of the importance of substrate channeling. Characterizing the leaky channeling within this system has been difficult due to limited experimental tools. The first structural evidence for natural metabolon formation and subsequent electrostatic substrate channeling within these enzymes was recently obtained by resolving the three-dimensional structure of the mMDH-CS-Aco complex by *in vivo* chemical cross-linking, mass spectrometry and protein docking.[112] Here, we characterize a synthetic metabolon formed *in vitro* by three recombinant versions of these enzymes of the cycle, which form a similar conformation as the natural complex *in vivo*. Substrate channeling is further investigated by site-directed mutagenesis and channeling can be significantly impaired by a single site-directed

mutation.

The mMDH and CS enzymes are dimers composed of identical subunits, weighing 34 kDa and 49 kDa respectively, whereas the next enzyme, Aco (85 kDa), is monomeric.[59,113,114] mMDH catalyzes the reversible NAD(H)-dependent conversion of L-malate and OAA. CS converts OAA and acetyl coenzyme A (acetyl-CoA) to citrate and Coenzyme A (CoA). Aco catalyzes the dehydration-rehydration of citrate to iso-citrate, with cis-aconitate being the intermediate. These enzymes have been individually characterized, with porcine heart being the most extensively studied variants, which share >95% sequence homology with bovine heart enzymes. Here, three different enzyme groups were investigated: native tissue enzymes isolated from the intact bovine mitochondria, commercially available porcine wild-type enzymes and recombinantly produced porcine enzymes. For the recombinant enzymes, codon optimized synthetic genes coding for porcine heart mMDH, CS and Aco were expressed in *E. coli*. Enzymes were purified to >90% (Figure S3.1) and the protein yields were 5 mg L^{-1}, 75 mg $L^{-1}$and 50 mg $L^{-1}$for mMDH, CS and Aco, respectively.

3.3.1 *In Vitro* Metabolon Conformations

Over the last decade, cross-linking/mass spectrometry analysis has been a common and standardized technique for studying protein complexes that can't be evaluated by X-ray or NMR analysis.[115,116] *In vitro* chemical cross-linking of protein-protein interactions between recombinant mMDH and CS was performed in the presence of Aco. SDS-PAGE analysis of commercial and recombinant enzyme mixtures demonstrated the formation of higher-ordered complexes after incubation with disuccinimidyl glutarate (DSG), indicated by a set of intense bands above 100 kDa

(Figure S3.2). In-gel tryptic digestion was conducted on the large complex bands, and the extracted peptide fragments were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS).

In the mitochondrial matrix, compartmentalized TCA cycle enzymes diffuse slowly, and their dynamic association is stabilized by the crowded environment. On the contrary, the apparent diffusion coefficient of enzymes in dilute solution is approximately two orders of magnitude higher, and the random molecular collision occurs much faster[101,117-120] than DSG cross-linking chemistry, so it is not possible to isolate and purify a stable *in vitro* complex without cross-linking and there are minimal experimental artificats from random molecular collisions without strong intermolecular interactions. However, the *in vitro* system only contains the three enzymes mMDH-CS-Aco, so the mass spectrometric data analysis and protein docking is easier, because the identity of all of the enzymes participating in the complex are known. Matching experimentally detected tryptic peptides to protein databases using the *Mascot* search engine identified the three enzymes in both cross-linked and non-cross-linked sample bands (Table S3.1). Cross-linked peptide candidates were determined by comparing cross-linked and non-cross-linked mass spectra, and matching additional masses after cross-linking to a manually-built theoretical mass database. Using the distance restraint (25 Å) on potentially DSG-linked residues (Figure S3.3), a hybrid protein docking method was utilized to elucidate the interactions between mMDH and CS *in vitro*, as previously described.[27]

For the complexes of commercial enzymes, a number of structures were found to meet the selection criteria and bear at least three identified cross- links (Figure S3.4). These structures exhibited distinct conformations, implying that mMDH and CS without any modifications interact in a random manner in dilute solution. Compared to the native tissue mitochondrial TCA cycle

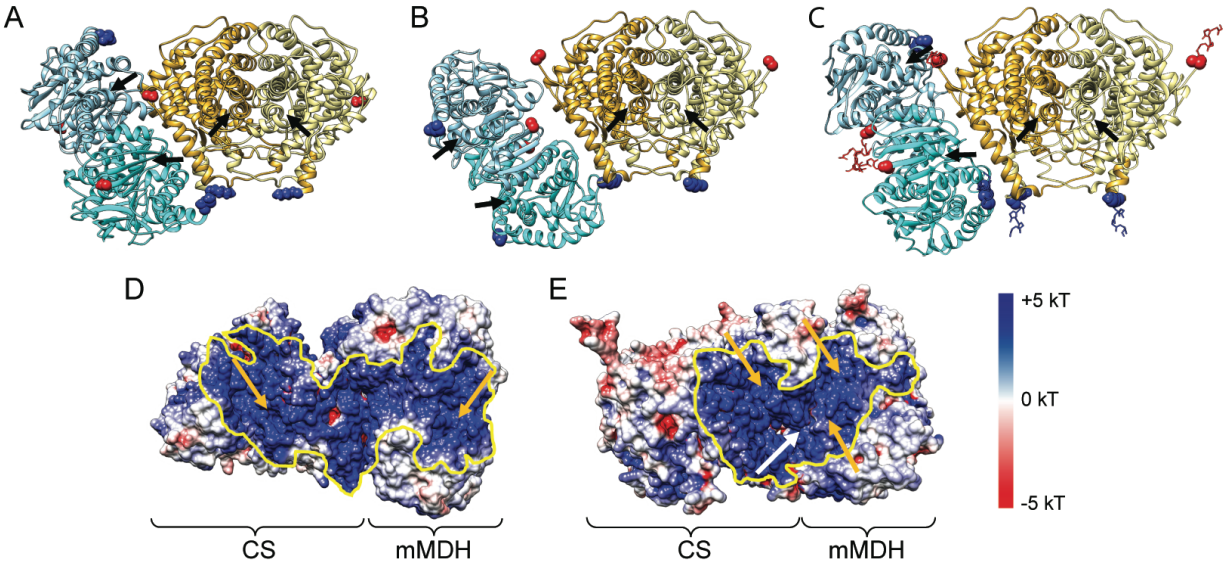

Figure 3.1: Structures of the mMDH-CS complex and the simulated electrostatic potential on the surface of the complex. (*A*) Native tissue mitochondrial metabolon. (*B*) *In vitro* complex formed by commercial enzymes. (*C*) *In vitro* complex formed by recombinant enzymes. N- and C-termini are represented by red and blue spheres, respectively. FLAG-tag and Polyhistidine-tag are represented by red and blue sticks, respectively. mMDH and CS active sites are denoted by black arrows. (*D*) *In vitro* complex formed by commercial enzymes. (*E*) *In vitro* complex formed by recombinant enzymes. Surface regions of positive potential and negative potential are colored in blue and red, respectively. The electrostatic channeling path for OAA is highlighted by the yellow edge. Orange arrows indicate the active sites. The white arrow indicates the location of Arg65 and Arg67. The surface ESP was calculated with water molecules at pH 7.4.

metabolon (Figure 3.1a), one model of the complex of commercial mMDH-CS showed up with the most structural similarity. In this structure (Figure 3.1b), three DSG cross-links (matching three MS peaks, Table S3.2) were obtained between mMDH Lys 191, Lys 277, Lys 283 and CS Lys 325. The α-helices of CS at Ala1-His 28 and Ser426-Lys437 are buried in the interface that covers the inter-subunit domain of mMDH. Compared to the native tissue mitochondrial metabolon *in vivo*, however, mMDH is flipped around the axis parallel to the binding interface by about 180 degrees. As a result, the two N-termini of mMDH are in close proximity with CS while the C-termini are pointing outward. This flipping may be less favorable for channeling of OAA, as the

mMDH active site clefts are open to the bulk phase and separated from CS active sites by a longer distance (73 Å) than that in the native tissue mitochondrial metabolon (35 Å).[112]

In contrast to the random association of commercial enzymes that yielded a number of complex conformations, the recombinant enzyme interactions were more restricted and resulted in a unique structure bearing seven DSG cross-links (matching four MS peaks, Table S3.3) between mMDH Lys81, Lys217, Lys304, Lys305 and CS Lys76, Lys325, Lys432. The α-helices of CS near its N- and C-termini again participate in the binding interface with the inter-subunit region of mMDH. Although mMDH was rotated around the axis perpendicular to the interface by approximately 30 degrees as compared to the metabolon *in vivo*, the final structure maintains most of the natural features (Figure 3.1c). No termini are buried in the interface, possibly due to spatial hindrance from additional amino acids not present in commercial or native mitochondrial enzymes appended to the termini (FLAG-tag (8 amino acids on N-terminus) and polyhistidine purification tag (6 amino acids on C-terminus)), but the relative locations of termini around the interface are not significantly altered. The two N-termini of the mMDH dimer point away from CS and the mMDH and CS active sites are brought within a closer proximity (40 Å) than what was observed in the complexes of commercial mMDH-CS (73 Å). With this shorter pathway, OAA transfer was expected to be faster than that in the wild-type complexes. In addition, the interfacial areas in the recombinant mMDH-CS complexes are about 12,100 $Å^2$, suggesting that they may be more thermodynamically stable than either the native tissue mitochondrial metabolons (10,000 $Å^2$)[112] or the complexes of commercial mMDH-CS (11,300 $Å^2$). Of course these measurements are based on docking of rigid protein models, and subtle conformational changes in the proteins upon self-assembly could lead to changes in the interaction areas.

Formation of commercial and recombinant mMDH-CS-Aco complexes *in vitro* was also

examined by docking Aco onto solved mMDH-CS complexes (Figure S3.5). Three DSG cross-links (matching three MS peaks) were identified between CS Lys16, Lys76, Lys80 and Aco Lys4, Lys117 in the tri-enzyme complex formed by commercial enzymes (Table S3.4). It was found that CS Lys76 and Lys80 were located in the mMDH-CS interface, implying that mMDH and Aco would not be present on the same subunit of CS. Therefore, each CS dimer can only bind one mMDH dimer and one Aco monomer (Figure S3.5a). In the recombinant mMDH-CS-Aco complexes, there were four DSG cross-links between CS Lys294, Lys300 and Aco Lys709, Lys712, exhibiting no conflict with the identified cross-links between mMDH and CS, although only one MS peak was matched (Table S3.5). This result is consistent with that obtained from *in vivo* cross-linking of CS and Aco. Residues within C-terminal region of the recombinant Aco were recognized at the CS-Aco interface. A groove formed between mMDH and CS appeared to accommodate the C-terminal of Aco (Figure S3.5b). Hence, the resulted tri-enzyme association was found to be more compact than that of the metabolon formed by commercial mMDH and CS. Similarly to the native tissue mitochondrial metabolon, a recombinant octamer comprised of one CS dimer, two mMDH dimers and two Aco monomers could possibly form *in vitr*o.[112]

A model of the three-enzyme system was built on the assumption that inclusion of Aco does not alter the complex formation between mMDH and CS. The interaction between Aco and CS (or mMDH) is weaker than that between mMDH and CS according to previous observations *in vivo*, even though there is a lower chance that CS (or mMDH) binds Aco prior to mMDH (or CS) in dilute solution. Therefore, the contribution of Aco to the structural assignment between mMDH and CS was of lesser interest, and the remainder of the experimental efforts were focused on the investigation of the channeling of OAA within the mMDH-CS complex which is consistent with most of the related research in the literature.[44,45]

3.3.2 Simulated Electrostatic Channeling

Elcock and McCammon previously demonstrated through Brownian dynamics simulations that electrostatic forces at the surface of a yeast mMDH-CS fusion protein greatly improved the OAA transfer efficiency.[44,54] In their fusion protein, a continuous surface of positive electrostatic potential bridged the active sites implying an important role of surface charge in the directed transport of OAA. In the natural TCA cycle metabolon, theoretical evidence for electrostatic channeling was also found between mMDH and CS active sites using simulation tools.[112] To further investigate electrostatic channeling in the mMDH-CS complex formed *in vitro*, the electrostatic surface potential (ESP) was examined using the Poisson-Boltzmann equation in the presence of water molecules at pH 7.4. As illustrated in Figure 3.1d, a long and broad band of positive potential covers the majority of the complex surface of the commercial enzymes on one side, connecting the active sites. In the recombinant mMDH-CS complex (Figure 3.1e), the positive patch connecting active sites was reduced due to their increased proximity and relative orientation. Taken together with previous results from simulation of surface ESP of free enzymes, the formation of substrate channeling in such dynamic complexes is a product of electrostatic protein-protein interactions and rearrangement of surface charges upon association. Charged surface residues, especially positively charged arginines and lysines, likely play an essential role in directed transport of negatively charged OAA.

To explore this, interfacial residues of the recombinant complex were identified by screening surface arginine and lysine residues within a distance of 20 Å from each other. As a result, CS Arg65 and Arg67 were estimated to be important for the formation of the positive channel. Site-directed mutagenesis was performed at these positions and six different CS mutants were explored, where Arg65 and Arg67 were replaced by either alanine or aspartic acid: R65A,

R67A, R65A/R67A, R65D, R67D and R65D/R67D. Prior mutational studies of CS have involved the active site residues,[121-124] aiming for the improvement of enzyme catalysis while the two residues mutated in this work (Arg65 and Arg67) are not located near the active site (Figure S3.6, Table S3.6).

Specific activities of the CS mutants were determined (Table S3.7). Arg67 was found to be crucial for the enzymatic activity of the recombinant CS. Any mutation of this side chain decreased or eliminated the enzymatic activity. In addition, CS(R67D) and CS(R65D/R67D) were found to be structurally affected by the mutations as shown in (Figure S3.7). CS(R67A), CS(R65A/R67A), CS(R65D) exhibited two or three orders of magnitude reductions of specific activities in comparison to the recombinant wild-type CS. CS(R67D) and CS(R65D/R67D) enzymatic activities could not be determined. CS(R65A) had similar enzymatic activity compared to recombinant wild-type CS. However, this mutation resulted in the dissociation of the mMDH-CS complex as implied by the disappearance of the mMDH-CS complex bands in native PAGE gels. (Figure S3.8). Residues 65 and 67 were compared among different species, and Arg67 is highly conserved whereas Arg65 is generally well conserved suggesting the importance of positive charge at these positions (Table S3.8).

3.3.3 Kinetic Analysis of Recombinant Enzymes

The steady state kinetics of the recombinant mMDH, recombinant CS and mutant CS enzymes were evaluated. Both mMDH and CS follow the ordered bi-bi kinetic mechanism, where mMDH binds to its cofactor and CS binds to OAA first.[113,114] The full steady state kinetic parameters were determined for recombinant mMDH, CS and CS(R65A) (Table 3.1). Recombinant CS had a

$k_{cat}/K_{M,acetyl\text{-}CoA}$ value of 4.3 $\mu M^{-1}s^{-1}$ and $k_{cat}/K_{M,OAA}$ value of and 11 $\mu M^{-1}s^{-1}$ whereas CS(R65A) had a $k_{cat}/K_{M,acetyl\text{-}CoA}$ value of 5.6 $\mu M^{-1}s^{-1}$ and a $k_{cat}/K_{M,OAA}$ value of 39 $\mu M^{-1}s^{-1}$. These values indicate that R65A mutation did not impair the kinetic behavior of the enzyme.

Table 3.1: Kinetic parameters of recombinant enzymes.[a]

Enzyme	k_{cat} (s^{-1})	$K_{i,A}$ (μM)[b]	$K_{M,A}$ (μM)[b]	$K_{M,B}$ (μM)[c]
mMDH, fwd	31 ± 2	420 ± 80	130 ± 30	830 ± 100
mMDH, rev	870 ± 140	15 ± 2	87 ± 22	33 ± 7.9
CS	88 ± 4	4.6 ± 0.8	7.9 ± 1.2	21 ± 3
CS(R65A)	44 ± 3	1.7 ± 0.1	1.1 ± 0.3	7.8 ± 2.3

[a]Data is given as mean values ± s.d. from at least three independent measurements. [b]Substrate A is NAD^+ for mMDH, fwd, NADH for mMDH, rev and OAA for both CS and CS(R65A). [c]Substrate B is L-malate for mMDH, fwd, OAA for mMDH, rev and Acetyl-CoA for both CS and CS(R65A).

3.3.4 Channeling in the Presence of a Competing Enzyme and Viscous Solution

A common method to probe substrate channeling is to introduce an enzyme competing for the same intermediate.[125] As illustrated in Figure 3.2a, the mMDH-CS complex catalyzes sequential conversion of L-malate to citrate via OAA as the intermediate, using NAD^+ and acetyl-CoA as cofactors. A competitive pathway was introduced with aspartate aminotransferase (AAT), which catalyzes the conversion of OAA and L-glutamate to aspartate and α-ketoglutarate. In the presence of AAT, the resultant rate of citrate generation measured with crude lysate containing the native tissue mitochondrial mMDH-CS complex showed little change (Figure 3.2b). Approximately 88% of the recombinant mMDH-CS activity was retained, whereas 72% of the commercial mMDH-CS

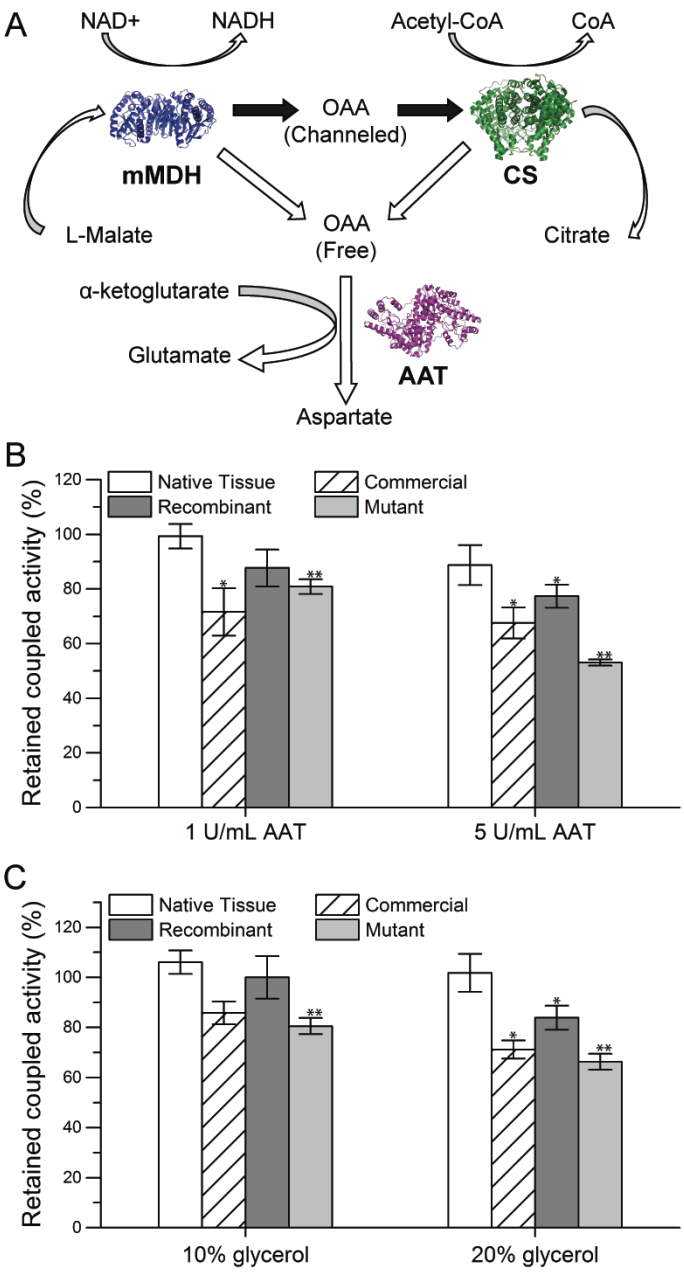

Figure 3.2: Demonstration of channeling of OAA within the mMDH-CS complex. (*A*) Schematic of the coupled mMDH-CS catalysis in the presence of competing enzyme AAT. White and black arrows represent diffusion and directed channeling, respectively. Crystal structure of AAT is obtained from Protein Data Bank (PDB ID: 1AAT). (*B*) Coupled activity retention calculated as the ratio of reaction rate before and after adding AAT measured with 100 μM mMDH-CS complex or crude mitochondrial lysate (about 5 μM protein in total). (*C*) Coupled activity retention in the presence of glycerol. Error bars represent standard deviation calculated from three independent experiments. Statistical significance (p-value) with respect to samples without AAT was calculated by a two-sample t-test with Welch correction for unequal variances: * = $p < 0.05$, ** = $p < 0.01$.

activity was retained in the presence of 1 U mL^{-1} of AAT. As the AAT concentration was increased to 5 U mL^{-1}, 89%, 77% and 68% of the coupled activity remained in the native tissue mitochondrial, recombinant, and commercial enzyme complexes, respectively (Figure 3.2b). The mutant recombinant complex (mMDH-CS(R65A)) only retained 53% of the coupled activity for 5 U mL^{-1} AAT. To minimize potential mMDH-AAT interactions, coupled catalysis was also explored with complexes immobilized in modified chitosan polymers. This ensured that AAT in bulk solution was physically separated from the complex and its interaction with the metabolon could be minimized. Similar results were obtained compared to the free complexes in solution (Figure S3.9).

Glycerol was added into the assay solution in order to explore a more viscous environment, and the efficiency of mass transport in different mMDH-CS complexes were compared. As Figure 3.2c demonstrates, coupled activity of the native tissue mitochondrial mMDH-CS complex in crude lysate was not affected by the increased viscosity. For the recombinant complex, catalysis was similar in 10% (v/v) glycerol, but decreased by 16% in 20% glycerol. This indicates that substrate channeling in the artificial complex functions in a "leaky" fashion. When compared to the complex of commercial enzymes and the mutant recombinant complex (mMDH-CS(R65A)), which respectively lost 15% and 20% of coupled activity in 10% glycerol and 30% and 35% of coupled activity in 20% glycerol, the recombinant complex (mMDH-CS) was less affected by increases in viscosity, indicating improved mass transport.

In a non-channeling system, OAA escaping into the bulk phase would be consumed by AAT, thus reducing the production of citrate. The coupled production of citrate by the recombinant complex was significantly less affected by the presence of a competing enzyme AAT, thus reducing the production of citrate. The coupled production of citrate by the recombinant complex

was significantly less affected by the presence of a competing enzyme (AAT) or a viscous reagent (glycerol). Taken together with the structural evidence, these results demonstrate that assembly of sequential enzymes is important for efficient substrate channeling. In this work, the ESP of commercial and recombinant mMDH-CS complexes were calculated, and it was shown that an electrostatic channeling path bridging active sites forms at the surface of commercial as well as recombinant enzymes. However, the recombinant complex with its active sites closely facing each other provides a channeling advantage over the complex of commercial enzymes whose active sites are further apart and facing oppositely. Moreover, the R65A mutation prevents association and the formation of the electrostatic channel in the metabolon. Mutated recombinant complexes retained only 50% of citrate production in the presence of 5 U mL^{-1} AAT demonstrating that intermediate transport in the mMDH-CS(R65A) is more prone to being interrupted by competing pathways.

3.3.5 Transient-Time Analysis

Transient time (τ) is used to describe the lifetime of intermediates in coupled catalysis. For enzymes that do not channel intermediates, a longer time is required to reach the steady state flux of the reaction intermediate.[126] Here, the ordered bi bi enzymatic reactions were treated as pseudo first-order by saturating the cofactors in the system. The overall initial reaction rates were limited by the OAA transport, which is dependent on the diffusion coefficient (D_i) of intermediates and the diffusing length (l) between active sites. Assuming that diffusion coefficients of OAA were not altered, the transient time is related to the diffusion distance length between the mMDH and CS active sites.

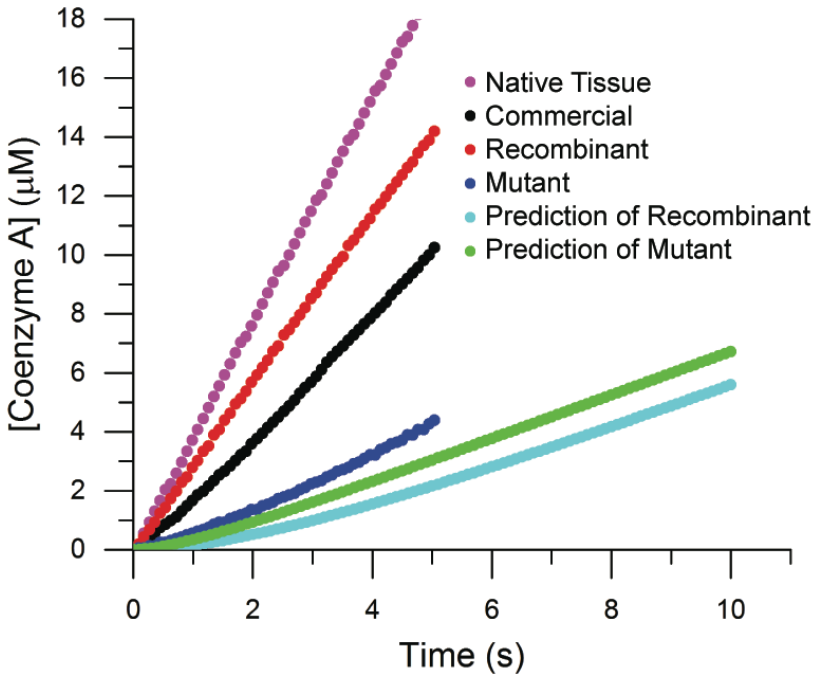

Figure 3.3: Characterization of OAA channeling via transient-time analysis. Citrate production was measured over time. Predictions of recombinant and mutant complexes were determined using rate equations for each enzyme. Transient time for each enzyme sample was determined by extrapolating the linear line fitted to the curve to the time axis as indicated by arrows (Table 3.2).

The transient times of OAA were measured in the native tissue mitochondrial metabolon and metabolons formed by commercial, recombinant and mutant mMDH and CS (Figure 3.3, Table 3.2). The simulated electrostatic channeling pathway (40 Å) for OAA was shorter in the mMDH-CS complex formed by recombinant enzymes, compared to commercial enzymes (73 Å). Consequently, the transient time of OAA in the recombinant complex was measured to be 30 ± 11 ms, which is comparable to the value of 40 ± 5 ms measured with crude lysate containing the native tissue mitochondrial TCA cycle metabolon. In comparison, the transient time of OAA in the complex of commercial enzymes was almost one order of magnitude higher, with a value of 290 ± 40 ms. In the case of the mutant mMDH-CS(R65A) complex, the resultant transient time

was increased to 880 ± 60 ms, approaching 1 s. According to these results, the channeling of OAA was fastest in the native tissue mitochondrial complex followed by the recombinant complex, complex of commercial enzymes and the mutant mMDH-CS(R65A) complex.

Agreeing with previous experiments, transient-time analysis also demonstrated that the mMDH-CS complex formed by recombinant enzymes in solution achieved similar channeling characteristics as the native tissue mitochondrial metabolon. The transient time of OAA in the recombinant mMDH-CS complex was reduced by 90% compared to the complex of commercial enzymes, indicating a higher channeling efficiency. The reduction of transient time in the recombinant complex, as compared to complex of commercial enzymes, was larger than the expected value of 70%, according to the equation for one-dimensional random walk, $l^2 = 2D_i$.[127] This is likely a result of the better orientation of the active sites in the recombinant complex. In the complex of commercial enzymes, the active site clefts in mMDH are facing oppositely to CS and are open to the bulk phase, resulting in an increased chance of OAA escape. Although substrate channeling in both systems was found to be "leaky", it is evident that recombinant enzymes exhibit higher catalysis coupling efficiency.

An analytical approach has been developed, which relates the substrate channeling phenomena to the Michaelis-Menten parameters of the enzymes.[54,126,128] A channeling probability parameter has been defined as $p_c\, p_r$, which can be obtained using the transient time, K_M and V_{max} of the second enzyme in the complex. The equation applicable to it mMDH-CS system is defined as the following:[54]

$$\tau = \frac{K_{M,app,OAA}\,(1-p_c p_r)}{V_{max,app}} \qquad (3.1)$$

By using the ordered bi bi rate equation and the kinetic parameters of the recombinant mMDH, CS and CS(R65A) enzymes, simulated rates of citrate formation by the unassembled enzymes were calculated. From these, the predictions of transient times of OAA in recombinant mMDH-CS and mMDH-CS(R65A) systems were estimated to be 2.5 s and 0.90 s, respectively (Table 3.2). These values represent the transient times with no interaction and no channeling, thus the probability parameter, $p_c\, p_r$, can be taken to be equal to 0. As shown by Eq. 1, $K_{M,app}/V_{max,app}$ is equal to the transient time in the case of no channeling. When this $K_{M,app}/V_{max,app}$ parameter is used together with the observed transient times of the complexes formed by the enzymes (0.03 s and 0.88 s for the recombinant and mutant complex respectively), the combined channeling parameters are calculated to be 0.99 and 0.023 for the recombinant mMDH-CS and mMDH-CS(R65A) respectively.

This analysis further confirms the channel formation is disrupted in the mutant mMDH-CS(R65A) complex. The predicted and measured transient times are very similar for mMDH-CS(R65A) (0.90 s vs 0.88 s), which indicates that the intermediate OAA is channeling poorly after the arginine mutation. On the contrary, the measured time lag for the recombinant complex was 0.03 s, indicating efficient channeling of OAA within the metabolon. The probability parameter, $p_c\, p_r$, should approach 1 as intermediates are efficiently channeled[54] and this value was found to be 0.99 and 0.023 for mMDH-CS and for mMDH-CS(R65A) respectively. A comparison of the $k_{cat}/K_{M,AcCoA}$ values for the recombinant wild type (4.3 $\mu M^{-1}s^{-1}$) and mutant CS (5.6 $\mu M^{-1}s^{-1}$) as well as the $k_{cat}/K_{M,OAA}$ values (11 $\mu M^{-1}s^{-1}$ for wild type and 39 $M^{-1}s^{-1}$ for the mutant CS) indicate that these results arise from changes in the transport efficiency of the complexes and are not due to major changes in kinetic activities.

Table 3.2: Measured and predicted transient times of OAA.[a]

Sample	τ (ms)
Native Tissue	40 ± 5
Commercial	290 ± 40
Recombinant	30 ± 11
Mutant	880 ± 60
Prediction of Recombinant	2500
Prediction of Mutant	900

[a]Transient times were determined from linear fits from Figure 3.3. Mean and standard deviation for experimental values were calculated from three independent experiments.

Metabolic Control Analysis (MCA) provides a framework to understand how metabolic fluxes are regulated by enzymatic activities.[129,130] The TCA cycle is a highly regulated network in central metabolism. To further investigate the potential implications of the R65A mutation, the elasticity coefficients of CS and CS(R65A) with respect to substrate OAA. These parameters indicate how much the reaction rates are affected by changes in OAA concentrations, and were estimated to be 0.79 for CS and 0.63 for CS(R65A) (Equation S3.2) based on the steady-state substrate concentrations obtained with our model. Thus, the sensitivity of CS to OAA concentrations was decreased by the mutation. Taking the connectivity theorem into account, it can be concluded that this mutation would lead to an increased flux control coefficient indicating a potentially increased role in regulating metabolic control. The potential impact on the flux control may explain why this mutation is rarely observed in nature.

3.4 Conclusions

Complex metabolic pathways, such as the TCA cycle, involve multiple enzymatic steps that require efficient mass transfer of intermediates between active sites. The metabolon formation within the TCA cycle and the interactions between malate dehydrogenase and citrate synthase in particular, have been a major research focus. In this work, we presented the first direct evidence for metabolon formation among recombinantly produced mMDH and CS. The structural and kinetic analyses demonstrated that the recombinant versions of these enzymes self-assemble *in vitro*, similar to their native counterparts *in vivo*. Important residues for the enzyme interactions were identified and site-directed mutational analysis was performed for the first time to investigate the substrate channeling among these enzymes. A single mutation in CS, R65A, along the positively charged patch connecting the active sites, disrupted the transport of the negatively-charged intermediate, decreasing the overall channeling probability from 0.99 to 0.023. These results demonstrate the importance of substrate channeling in this critical biological pathway.

3.5 Materials and Methods

3.5.1 Materials

Synthetic genes coding for the porcine heart enzymes were synthesized by Genscript (Piscataway, NJ). All genes have a Flag-tag at the N-terminus and a 6xHis-tag at the C-terminus, for identification and purification purposes, respectively. Restriction enzymes for DNA cloning were purchased from New England Biolabs (Ipswich, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) and ampicillin sodium salt, were purchased from Gold Biotechnology (St. Louis, MO). Amicon centrifugal filters were purchased from Millipore. Disuccinimidyl glutarate (DSG), sodium dodecyl sulfate polyacrylamide electrophoresis gels (SDS-PAGE) and running buffers were purchased from Invitrogen-Life Technologies (Carlsbad, CA). *E. coli* BL21 and BL21(DE3) cell lines were purchased from Bioline (Taunton, MA). Chaperon plasmid pGro7 was purchased from Clontech Laboratories–Takara (Mountain View, CA). Ala-chitosan was prepared as previously described.[131] Fresh bovine heart was purchased from a local slaughterhouse and used immediately. All other reagents and materials were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise stated.

3.5.2 Construction of Recombinant Enzymes

Cloning of the synthetic genes into expression plasmids

The genes coding for CS and Aco were cloned into pET-20b(+) backbone using the NdeI and HindIII restriction sites. Resulting plasmids were transformed into BL21(DE3) cells. mMDH was inserted into pMAL-c4e expression vector via the same restriction sites. The resulting construct

and the chaperon plasmid pGro7 were co-transformed into BL21 cells for expression.

Site-directed mutagenesis of recombinant CS

CS Arg65 and Arg67 were mutated to alanine and aspartic acid to create single and double mutants via site-directed mutagenesis: CS(R65A), CS(R67A), CS(R65A/R67A), CS(R65D), CS(R67D), CS(R65D/R67D). CS(R65A) was used as the template to mutate the Arg67 to alanine and CS(R65D) was used as the template to mutate Arg67 to aspartic acid. Corresponding primer sequences used during the PCR reaction are given in Table S3.9.

3.5.3 Expression and Purification of Enzymes

Expression and purification of the recombinant enzymes

All constructs were expressed in 1L of sterilized Terrific Broth, inoculated with 10 ml overnight culture. For Aco and CS, the media was supplemented with 100 μg/mL ampicillin. M was added to mMDH cultures in addition to the ampicillin. The cells were grown to an OD_{600} of 0.6 while shaking at 37°C, and protein expression was induced with 0.5 mM IPTG for mMDH and CS, and with 0.6 mM IPTG for Aco. Expression was carried out for 18–20 h at 25°C. Cells were harvested by centrifugation at 5000 × g for 10 min and resuspended in 50 mL HisTrap binding buffer (20 mM Tris, 150 mM NaCl and 20 mM imidazole, pH 7.4) per L of culture supplemented with HALT EDTA-free protease inhibitor. Soluble proteins were collected via centrifugation at 15000 × g for 30 min after the cells were lysed by sonication with an ultrasonication probe in an ice bath for 6 min (5 s on pulse and 2 s off pulse). Enzymes of interest were purified by immobilized metal affinity chromatography using a HisTrap columns (GE Healthcare Life Sciences, Pittsburgh, PA),

where bound enzymes were eluted with the elution buffer (20mM Tris, 150mM NaCl, 500mM imidazole (pH 7.4). mMDH was buffer exchanged into 20mM Tris-HCl (pH 8.7) and purified via anion exchange chromatography where the enzyme was eluted using a linear NaCl gradient from 0 to 1 M NaCl. All enzymes were further purified with size exclusion chromatography after buffer exchanging into 50mM Tris, 150mM NaCl (pH 7.4). Amicon filters (Milipore, Billerica, MA) with 30kDa (for mMDH and CS) and 50kDa (for Aco) molecular weight cutoff were used in order to concentrate the protein solutions as well as to exchange the buffer in between different purification steps.

Preparation of crude mitochondrial lysate.

Extraction of the bovine heart mitochondria was done according to the procedure described by Rogers *et al.* with some modifications.[132] Bovine heart cubes were blended with cold isolation buffer (70 mM sucrose, 210 mM mannitol, 5 mM HEPES, 1 mM EGTA and 0.5% (w/v) BSA, pH 7.2) in a Waring laboratory blender. Meat suspension was centrifuged at 500 × g for 10 min, and the supernatant was centrifuged at 26000 × g for 20 min. Pellet was homogenized in the isolation buffer and centrifuged twice again at 500 × g for 10 min. Supernatant was filtered through a double-layer cheesecloth and centrifuged at 1000 × g for 20 min. The mitochondria pellet was washed with lysis buffer (50 mM Tris, 150 mM NaCl, 2 mM EDTA and 1 mM PMSF, pH 7.4) at 26000 × g for 10 min. Pellet resuspended in the lysis buffer was sonicated with an ultrasonication probe in ice bath for 4 min (5 s on pulse and 15 s off pulse). The crude lysate was initially cleared at 5000 × g for 30 min. EDTA and PMSF were removed through the pre-packed Sephadex™ G-25M column (GE Healthcare Life Sciences). Protein concentration in the mitochondrial lysate was determined to be 1 mg/mL by BCA assay.

3.5.4 Structural Analysis

In vitro chemical cross-linking of mMDH, CS and Aco

Commercially available enzymes purchased from Sigma-Adrich and recombinant enzymes were cleaned up by a pre-packed Sephadex^{TM} G-25M column into 10 mM phosphate buffer (pH 7.4) to remove ammonium sulfate and other salts containing primary amine. Afterwards, mMDH, CS and Aco were mixed equally to a total protein concentration of 20 μM in 10 mM phosphate buffer (pH 7.4). DSG dissolved in 50 μL of DMF was added to the enzyme mixture to a final concentration of 1 mM. The approximate DSG/protein molar ratio was 50:1 to ensure an efficient capture of weak protein-protein interactions in dilute solution without dramatic loss of enzyme activity. As the non-cross-linked control, 50 μL of DMF containing no DSG was used. Cross-linking was carried out at room temperature for 30 min under gentle shaking and quenched by adding 2 M Tris buffer (pH 8.3) to a final concentration of 20 mM.

Separation and in-gel digestion of enzyme complexes

Enzyme mixtures were washed with 50 mM Tris buffer (pH 7.4) in filter-incorporated Amicon tubes with a mass cutoff at 10 kDa at 5000 × g for 15 min to remove phosphates and extra DSG. Afterwards, cross-linked and non-cross-linked samples were directly separated by reducing SDS PAGE, which was performed on a 4–20% gradient gel according to the protocol provided by the manufacturer. Gel bands of interest were excised and de-stained twice in 1 mL of 50% methanol with 50 mM ammonium bicarbonate at room temperature, under gentle vortexing for 1 h. The gel slices were rehydrated in 1 mL of 50 mM ammonium bicarbonate at room temperature for 30 min, and the gel bands/spots of interest were cut into several pieces. These gel pieces were rehydrated

in 1 mL of 100% acetonitrile at room temperature under gentle shaking for 30 min. Acetonitrile was carefully removed from the gel pieces with a pipette tip prior to trypsin digestion. The gel pieces were incubated with 10–20 μL of sequence-grade modified trypsin (20 ng/μL, Promega) in 50 mM ammonium bicarbonate overnight at 37 °C. Digestion was quenched by adding 20 μL of 1% formic acid. Then, the solution was allowed to stand, and peptides that dissolved in the 1% formic solution were extracted and collected. Further extraction of peptides from the gel material was performed twice by adding 50% acetonitrile with 1% formic acid and sonicating at 37 °C for 20 min. All these solutions were collected and combined. A final complete dehydration of the gel pieces was accomplished by adding 20 μL of 100% acetonitrile followed by incubation at 37 °C for 20 min. The combined supernatant solutions of extracted peptides were dried in a vacuum centrifuge (Speed-Vac). The peptides were reconstituted in 100 μL of 5% acetonitrile with 0.1% formic acid for mass spectrometric analysis.

Mass spectrometric instrumentation

Peptides were analyzed using a nano-liquid chromatography-tandem mass spectrometry (LC-MS/MS) system comprised of a nano-LC pump (Eksigent, Dublin, CA) and a LTQ-FT mass spectrometer Thermo Fisher Scientific (Waltham, MA). The LTQ-FT is a hybrid mass spectrometer with a linear ion trap used typically for MS/MS fragmentation (i.e. peptide sequence) and a Fourier transform ion-cyclotron resonance (FT-ICR) mass spectrometer used for primary accurate mass measurement of peptide ions. The LTQ-FT is equipped with a nanospray ion source (Thermo Fisher Scientific). Approximately 5 to 20 fM of tryptic-digested or phosphopeptide-enriched samples were dissolved in 5% acetonitrile with 0.1% formic acid and injected onto a homemade C18 nanobore LC column for nano-LC-MS/MS. A linear gradient LC profile was used

to separate and elute peptides, consisting of 5 to 70% solvent B in 78 min with a flow rate of 350 nL/min (solvent A: 5% acetonitrile with 0.1% formic acid; solvent B: 80% acetonitrile with 0.1% formic acid). The LTQ-FT mass spectrometer was operated in the data-dependent acquisition mode controlled by *Xcalibur 1.4* software, in which the "top 10" most intense peaks observed in an FT primary scan (i.e. MS survey spectrum) were determined by the computer on-the-fly and each peak was subsequently trapped for MS/MS analysis and sequenced through peptide fragmentation by collision-induced dissociation. Spectra in the FT-ICR were acquired from m/z 400 to 1700 at 50000 resolving power with about 3 ppm mass accuracy. The LTQ linear ion trap was operated with the following parameters: precursor activation time was 30 ms and activation Q was 0.25; collision energy was set at 35%; dynamic exclusion width was set at low mass of 0.1 Da with one repeat count and duration of 10 s.

Mascot database searches

LTQ-FT MS raw data files were processed to peak lists with *BioworksBrowser 3.2* software (Thermo Fisher Scientific). Processing parameters used to generate peak lists were as followed: precursor mass was between 401–5500 Da; grouping was enabled to allow five intermediate MS/MS scans; precursor mass tolerance was set at 5 ppm; minimum ion count in MS/MS was set to 15, and minimum group count was set to 1. Resulting DTA files from each data acquisition were merged and searched against the NCBI or custom databases for identified proteins, using *MASCOT* search engine (Matrix Science Ltd; version 2.2.1; in-house licensed). Searches were done with tryptic specificity, allowing two missed cleavages or "non-specific cleavage" and a mass error tolerance of 5 ppm in MS spectra (i.e. FT-ICR data) and 0.5 Da for MS/MS ions (i.e. LTQ Linear ion trap). Identified peptides were generally accepted only when the *MASCOT* ion score value

exceeded 20.

Identification of cross-linked peptides

Mass spectrometric raw files were analyzed via *Thermo Xcalibur* software and peptide peaks of interest were picked manually. A theoretical mass database of potential inter-protein cross-links was built up using a spreadsheet by combining two peptides, which were identified in individual (non-cross-linked) enzymes but missed in the cross-linked enzyme complex by *MASCOT* database search. Additional peptide peaks only found in cross-linked spectra were screened against the mass database. Cross-link candidates were selected by the following rules: trypsin did not cut at the C-terminus of modified lysines or lysines with proline on the C-terminus; up to two missed cleavages were allowed, but non-specific cuttings were not considered; peptide length was 5–30 amino acids; each cross-linked peptide had at least one lysine for cross-linking as well as a lysine or arginine at C-terminal; peaks showed up in at least duplicate experiments; mass error = 5 ppm. Flexible modifications that might be obtained by oxidation or during SDS-PAGE running were applied to specific residues for identification and the respective mass variations were previously summarized.[112] Identified cross-links were examined by *Mascot* automated target-decoy search against NCBI database to estimate false-discovery rate (FDR) and no protein hits were reported above identity threshold ($p = 0.05$)

Hybrid protein docking

Global docking and local docking were carried out to solve the structure of the mMDH-CS-Aco complex. In global docking, an automated protein docking web server, *Cluspro*

(http://cluspro.bu.edu/), was utilized.[133-135] Cross-linked lysines identified by manual search were set as attracting residues. All proteins were treated as rigid bodies with their "open" conformations obtained from crystal structures, giving top 100 ~ 120 structures of highest score based on surface shape complementarity and free energies of desolvation and electrostatic interactions. The crystal structure of Aco, mMDH, and CS were obtained from the Protein Data Bank (PDB ID: 1MLD, 1CTS, and 7ACN). The 24 AA sequences at N-terminus are signal peptides, which are cut off from mature enzymes, so they were excluded from simulation. However, for the recombinant proteins, the FLAG-tag and His-tag was included in model as well as in the final structures. Prior to local docking, all model candidates were screened by *Xwalk* software suite to filter out false positives by distance constraints.[136] Maximum Euclidean distance limit was set to 25 Å, resulting from a combination of DSG spacer arm length (7.7 Å), lysine side chain length (6 Å × 2) and backbone flexibility. In addition to Euclidean distance limit, solvent accessible surface (SAS) distance was set to 30 Å to mimic molecular flexibility of DSG. Solvent radius was 1.4 Å by default and set to 2 Å for SAS distance calculation. Rotamers were removed and only the distance of Cβ-Cβ between two lysines was calculated. A pair of lysines on two proteins in global candidates were considered as a potential cross-link, if their *Xwalk*-calculated separation is no more than the limits.[137] After distance filtering, global candidates bearing at least two potential cross-links were subject to local docking by another protein docking web server, *Rosetta* (^{90}).[138-140] Derived from each starting global structure, 10 local candidates of lowest interface energy were screened again by *Xwalk*. Final complex structures were chosen based on two criteria: local candidates of lowest interface energy were clustered around a single position on the energy landscape and the structure had the highest number of potential cross-linkers in agreement with experimental results. Interfacial residues in final structures were determined when the measured Euclidean distance was less than

20 Å.

Simulation of electrostatic surface potential

Prior to simulation, docked structures were modified by *PDB2PQR* web server (http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/) to add missing hydrogens and/or heavy atoms and to estimate their titration states.[141,142] Protein complexes were protonated with favorable hydrogen bonds. Charges and radius were assigned from Amber force field. *PROPKA* was used to predict pK_a shifts in complexes at pH 7.4.[143] Calculation of surface ESP by Poisson-Boltzmann equation was done by *APBS* web server (http://www.poissonboltzmann.org/docs/apbs-installation/)[144,145] with the following parameter settings: water molecules were not removed; no additional ions were added at zero ionic strength; biomolecular dielectric constant was set at 2; and solvent dielectric constant was set at 78.54.

3.5.5 Kinetic Analysis

Kinetic analysis of the recombinant enzymes

mMDH and CS activity measurements were carried out as in Shatalin *et al*. with some modifications.[45] mMDH was measured for activity with different substrate concentrations for the forward and reverse reactions, in 100 mM potassium phosphate buffer (pH 7.4) in a 96-well plate. L-malate, NAD^+, oxaloacetate (OAA) and NADH concentrations were varied from zero to 3 mM, 4 mM, 0.1 mM and 0.1 mM, respectively. NADH concentration was measured spectrophotomerically at 340 nm after the addition of 1 nM and 0.1 nM mMDH for the forward and reverse reactions, respectively. CS activity was determined in 100 mM potassium phosphate buffer (pH 7.4) as well, via monitoring the coenzyme A (CoA) production at 412 nm in the

presence of 1 mM DTNB (5,5'-dithiobis(2-nitrobenzoate)) with 1 nM of enzyme in a 96-well plate. OAA and acetyl coenzyme A (acetyl-CoA) concentrations were varied from zero to 0.5 mM and 0.2 mM, respectively. NADH production/consumption was calculated using the extinction coefficients 6220 $M^{-1}cm^{-1}$. Production of citrate can be spectrophotometrically monitored through a subsequent reaction of CoA and DTNB, which yields a di-anion (TNB^{2-}) absorbing at 412 nm. Citrate production rate was determined from the maximum linear slope of the curve of absorbance over time. Extinction coefficient of TNB^{2-} at 412 nm was 14,150 $M^{-1}cm^{-1}$, and the light path length was 0.56 cm. All enzyme concentrations were determined by Bradford Assay (Thermo Scientific) following the protocol provided by the manufacturer and a SpectraMax M2 (Molecular Devices, Sunnyvale, CA) was used for absorbance readings. Obtained data was fitted into ordered bi-bi equation (Equation 3.2) in order to calculate the kinetic parameters of the enzymes:

$$v = \frac{V_{max}[A][B]}{K_{iA}K_B + K_B[A] + K_A[B] + [A][B]} \tag{3.2}$$

Coupled activity assays of the mMDH-CS complex in solution with AAT or glycerol

Equal amounts of mMDH and CS were mixed in 10 mM PBS (pH 7.4) to a final total protein concentration of 20 μM, and incubated under gentle shaking at room temperature for 30 min. The coupled activity of the mMDH-CS complex (100 nM) or the crude lysate (0.5 mg/mL) was assayed in a 96-well plate with 1 mM L-malate, 2 mM NAD^+, 0.1 mM acetyl coenzyme A, 0.2 mM DTNB and10 mM glutamate in 200 μL of 100 mM potassium phosphate buffer (pH 7.4) in the presence of 1 or 5 U mL^{-1} AAT. Control experiments were done without adding AAT. Glycerol was added to the enzyme and substrate solutions to 10% and 20% (v/v) prior to mixing. Then the coupled

activity of the mMDH-CS complex (100 nM) or the crude lysate (0.5 mg mL^{-1}) was assayed in a 96-well plate with 1 mM L-malate, 2 mM NAD^+, 0.1 mM acetyl-CoA and 0.2 mM DTNB in 200 μL of 100 mM potassium phosphate buffer (pH 7.4). Control experiments were done without adding glycerol. The absorbance increase at 412 nm was monitored by Synergy™ HTX multi-mode microplate reader (BioTek) over 1 min at 1 s intervals. One unit (U) of enzyme activity was defined as 1 μmole of product formed in one minute.

Coupled activity assay of immobilized mMDH-CS complex with AAT

The mMDH-CS complex solutions (4 μM) and ala-chitosan solution (10 mg mL^{-1}) were mixed at a volume ratio of 2:1, and incubated on vortex at room temperature for 15 min. Cross-linked samples were prepared by incubating mixtures of mMDH and CS at 20 μM with 0.2 mM DSG under gentle shaking at room temperature for 30 min, followed by quenching with 2 M Tris (pH 8.3). The cross-linking ratio of DSG:protein was lowered to 10:1 to minimize potential deactivation of enzymes by excessive cross-linkers. 25 μL of the enzyme/polymer suspension was pipetted to the bottom of a polystyrene cuvette (1 cm for light path length) and dried in a vacuum at room temperature for 2 h. Coupled activity of immobilized enzyme complex was assayed in 1 mL of 100 mM potassium phosphate buffer (pH 7.4) containing 1 mM L-malate, 2 mM NAD^+, 0.1 mM acetyl-CoA, 0.2 mM DTNB, 10 mM glutamate and AAT at 0 or 1 U mL^{-1}. The absorbance change at 412 nm was monitored by a UV-Vis spectrophotometer (Evolution 260 Bio, Thermo Scientific).

Transient time measurement by fast kinetic study

Fast kinetic experiment was carried out in a 96-well plate measured by the plate reader equipped with a dual injection module. Before assays, 10 μL of mMDH (20 μM) and 10 μL of CS (20 μM) were mixed in 10 mM PBS (pH 7.4) and incubated under gentle shaking at room temperature for 30 min, followed by dilution to 2 mL in 100 mM potassium phosphate buffer (pH 7.4). The crude lysate was directly used without further dilution. Substrate solution was prepared in 2 mL of potassium phosphate buffer containing 2 mM L-malate, 4 mM NAD^+, 0.2 mM acetyl-CoA and 0.4 mM DTNB. To setup the assay condition, enzyme and substrate solutions were respectively injected by two separate syringes at a flow rate of 250 μL s^{-1}. Total assay volume was 200 μL s^{-1} per well. Absorbance at 412 nm was read every 90 ms over 1 min. Transient time of OAA was determined by extrapolating the linear line fitted to the absorbance curve within the first recorded 5 s.

Predicted transient time calculations

In addition to the experimental fast kinetic study of mMDH-CS complex, transient time analysis was performed using the experimentally obtained kinetic parameters of individual recombinant mMDH, recombinant CS and mutant CS(R65A). In Matlab, a function “xprime” is defined as the change in substrate / product concentrations with respect to time and ode45 function is used to solve the differential equations describing these substrate (OAA and acetyl-CoA) consumptions / product (citrate and CoA) formations. Obtained data is fitted using Excel and transient time of OAA was determined by extrapolating the linear line fitted to the time ver sus product concentration plot. The Matlab code is given below:

```
function xprime = concentrations(t,x);

xprime=[(v0-((V1*x(1)*x(2))/((KiA*KB)+(KA*x(2))+(KB*x(1))+(x(1)*x(2)))));(-
1)*(V1*x(1)*x(2))/((KiA*KB)+(KA*x(2))+(KB*x(1))+(x(1)*x(2)));(V1*x(1)*x(2))/((KiA*KB)+(KA*x(
2))+(KB*x(1))+(x(1)*x(2)));(V1*x(1)*x(2))/((KiA*KB)+(KA*x(2))+(KB*x(1))+(x(1)*x(2)))];
```

where x(1) is OAA and x(2) is acetyl-CoA.

Elasticity coefficient calculations

In order to calculate the elasticity coefficient of CS with respect to OAA, following equation is used (18):

$$\varepsilon_{OAA}^{CS} = \frac{\partial v_{CS}}{\partial [OAA]} \frac{[OAA]}{v_{CS}} \qquad (2)$$

where v is the reaction rate of CS and $[OAA]$ is the concentration of substrate OAA. After taking the derivative of the rate equation with respect to OAA, kinetic parameters and steady state OAA concentration belonging to CS and CS(R65A) are fitted to the Eq. 2, in order to obtain the steady state elasticity coefficients.[129]

3.6 Supplementary Information

3.6.1 Supporting Figures

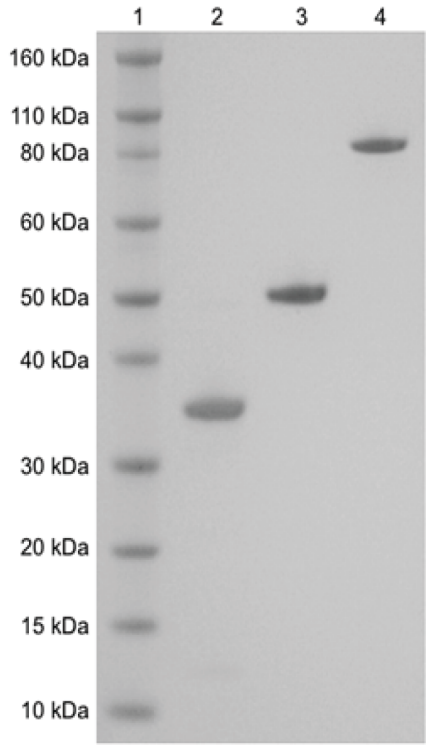

Figure S3.1: SDS-PAGE analysis of recombinant enzymes. (1) Protein ladder. (2) mMDH - 34 kDa. (3) CS - 49 kDa. (4) Aco - 85 kDa.

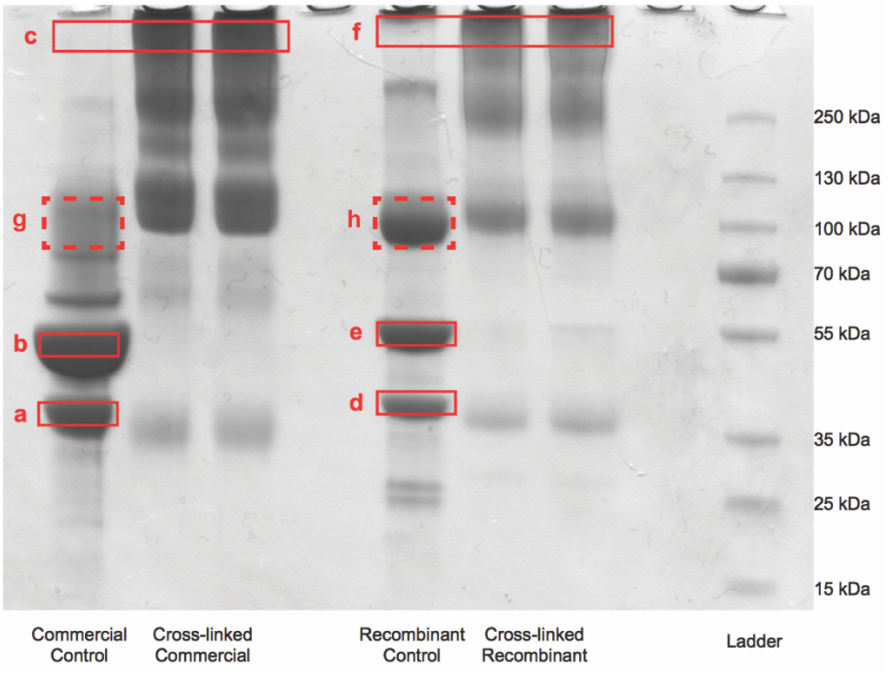

Figure S3.2: SDS-PAGE analysis of DSG-cross-linked and non-cross-linked (control) mMDH-CS complexes *in vitro* for commercial and recombinant enzymes. (a) Commercial mMDH subunit. (b) Commercial CS subunit. (c) Commercial mMDH-CS complex. (d) Recombinant mMDH subunit. (e) Recombinant CS subunit. (f) Recombinant mMDH-CS complex. (g) Commercial Aco. (h) Recombinant Aco. Gel bands of interest for subsequent digestion and analysis are indicated by solid red squares.

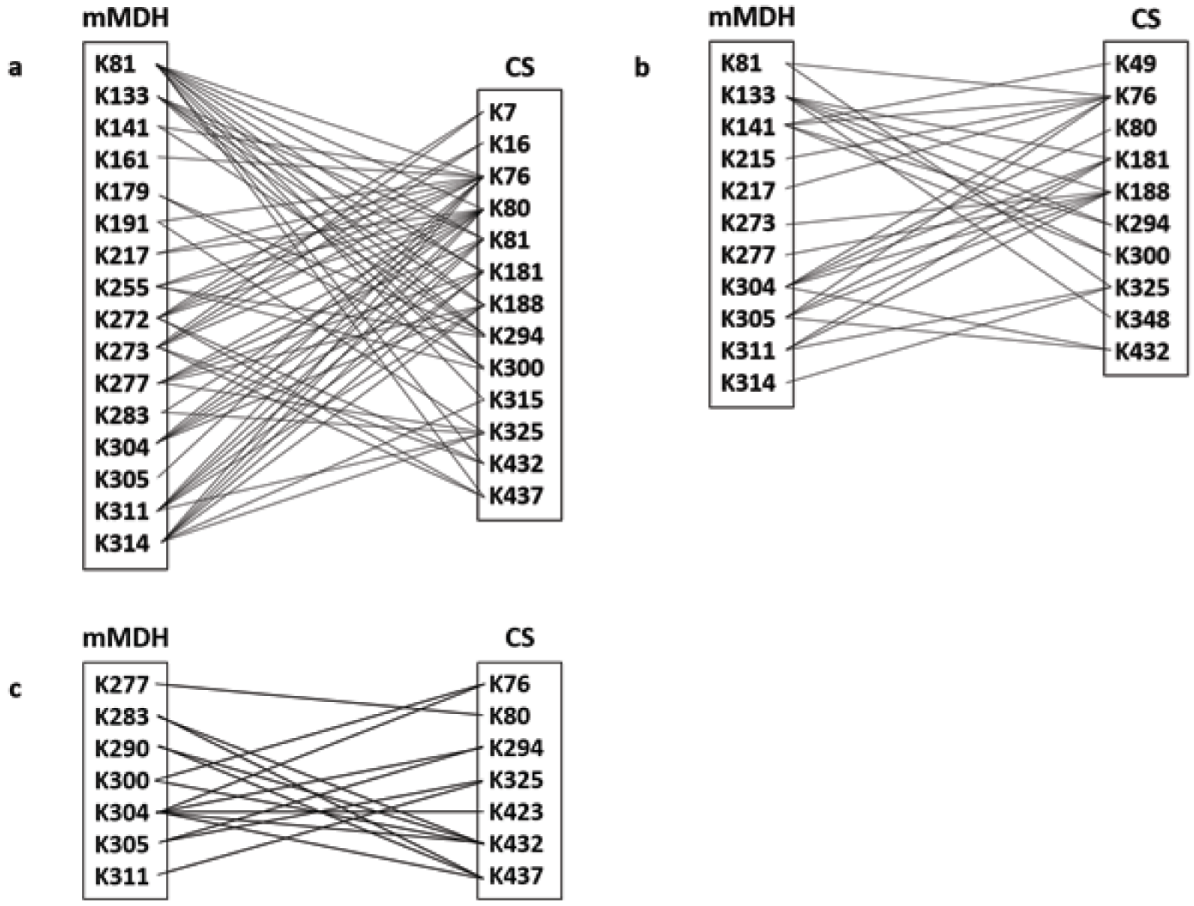

Figure S3.3: Summary of all possibly cross-linked lysine residues (represented by K) (**a**) in the *in vitro* mMDH-CS complex formed by commercially available enzymes**.** (**b**) in the *in vitro* mMDH-CS complex formed by recombinant enzymes. (**c**) in the *in vivo* mMDH-CS complex formed by native tissue mitochondrial enzymes.

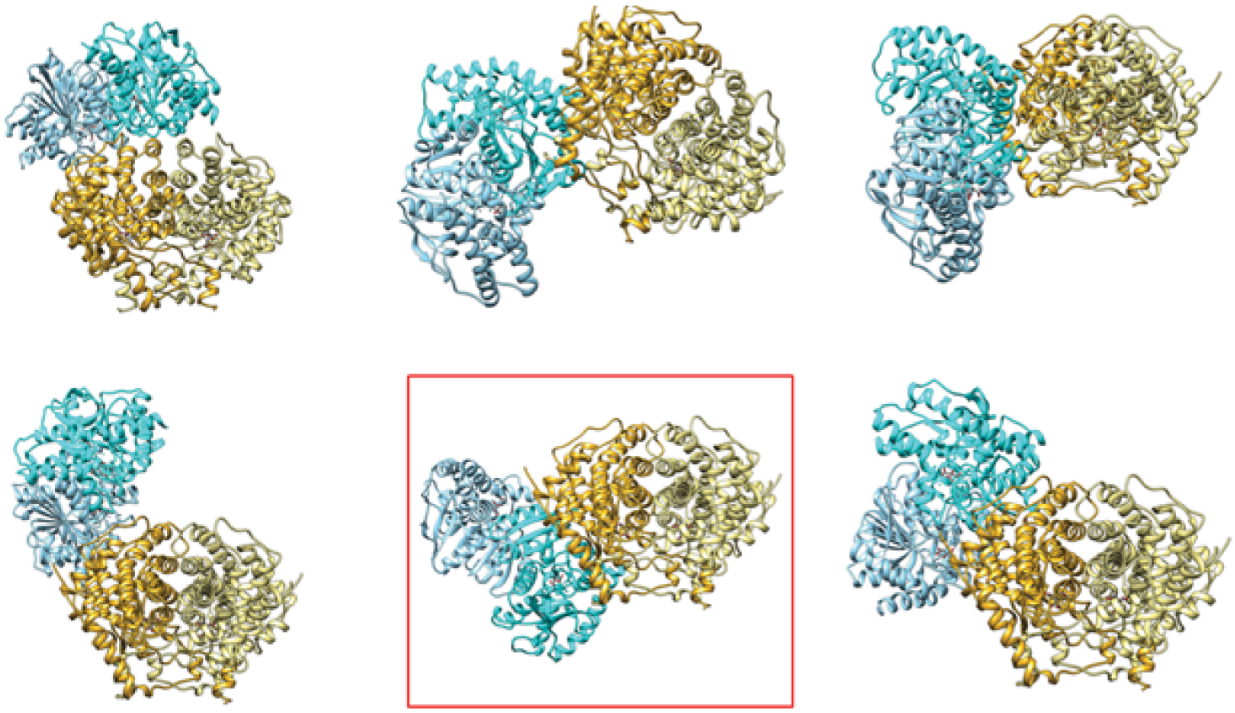

Figure S3.4: mMDH-CS metabolon structures. All possible complex structures formed by commercial mMDH (colored in sky blue and cyan) and CS (colored in gold and yellow) in solution. The crystal structure of mMDH and CS were obtained from the Protein Data Bank (PDB ID: 1MLD and 1CTS).The structure with the most similarity to the natural metabolon is highlighted in the red square.

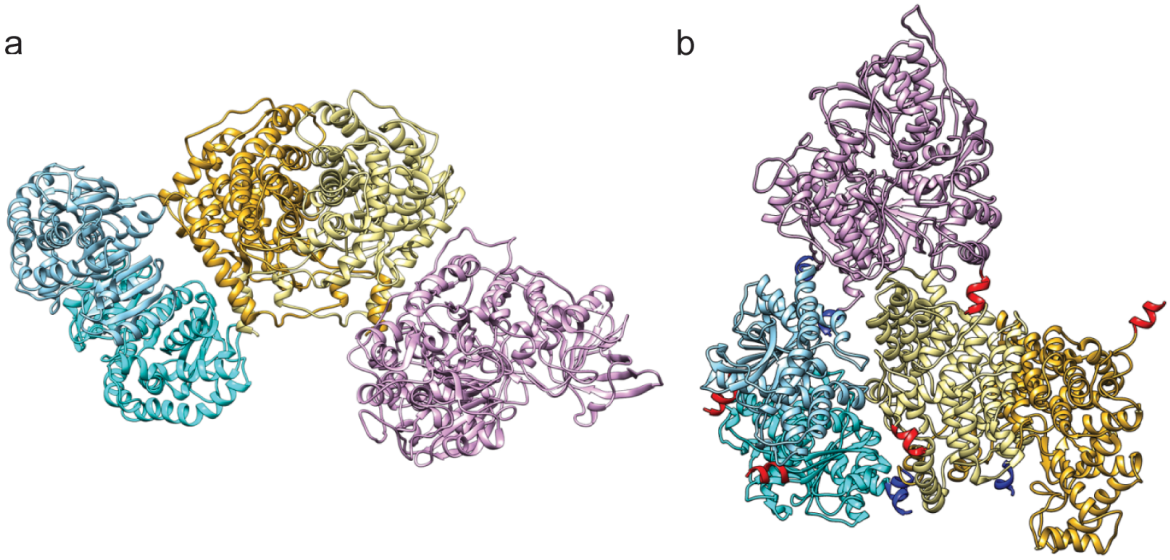

Figure S3.5: mMDH-CS-Aco complex structures**.** (**a**) Structures of the complex of commercial enzymes and (**b**) recombinant mMDH-CS-Aco complexes formed *in vitro*. The crystal structure of Aco, mMDH, and CS were obtained from the Protein Data Bank (PDB ID: 1MLD, 1CTS, and 7ACN). Aco is colored in purple. FLAG-tagged N-terminus and His-tagged C-terminus are highlighted in red and blue, respectively.

Figure S3.6: Location of mutations in CS**.** Structure of a porcine CS subunit (PDB ID: 1CTS). The residues shown in purple represent the mutation sites found in literature (Table S3.6).[121,124] The residues shown in yellow represent the Arg65 and Arg67 positions

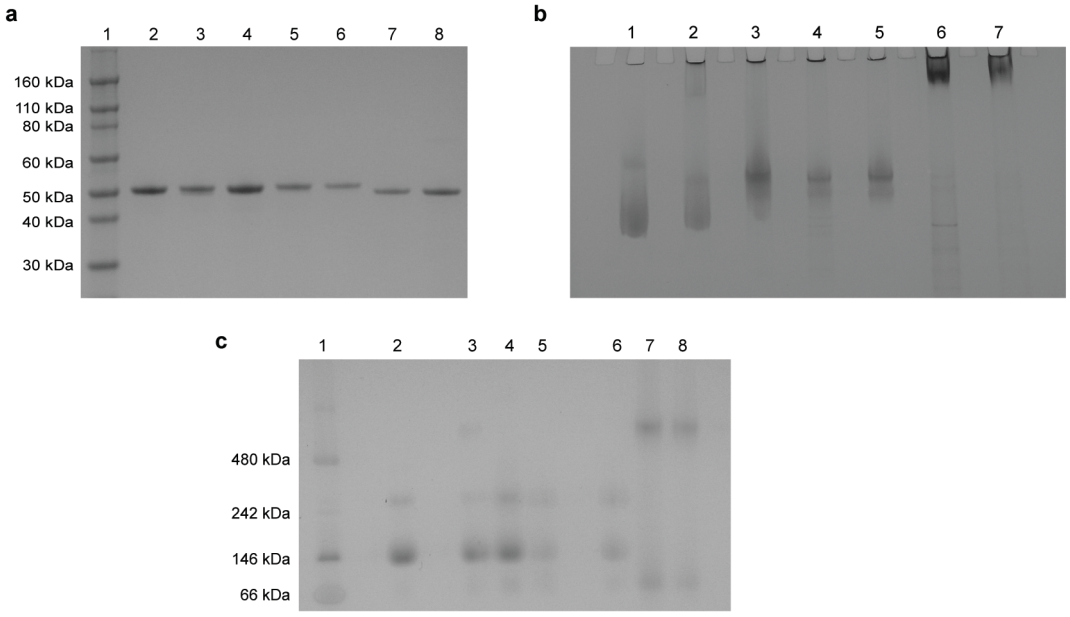

Figure S3.7: Protein gel pictures of CS mutants. (a) SDS-PAGE analysis: (1) Protein ladder. (2) recombinant CS. (3) CS(R65A). (4) CS(R67A). (5) CS(R65A-R67A). (6) CS(R65D). (7) CS(R67D). (8) CS(R65D-R67D). All mutants have a molecular weight of 49 kDa. (b) Native-PAGE analysis: (1) recombinant CS. (2) CS(R65A). (3) CS(R67A). (4) CS(R65A-R67A). (5) CS(R65D). (6) CS(R67D). (7) CS(R65D-R67D). (c) Blue Native-PAGE analysis: (1) Protein ladder. (2) recombinant CS. (3) CS(R65A). (4) CS(R67A). (5) CS(R65A-R67A). (6) CS(R65D). (7) CS(R67D). (8) CS(R65D-R67D).

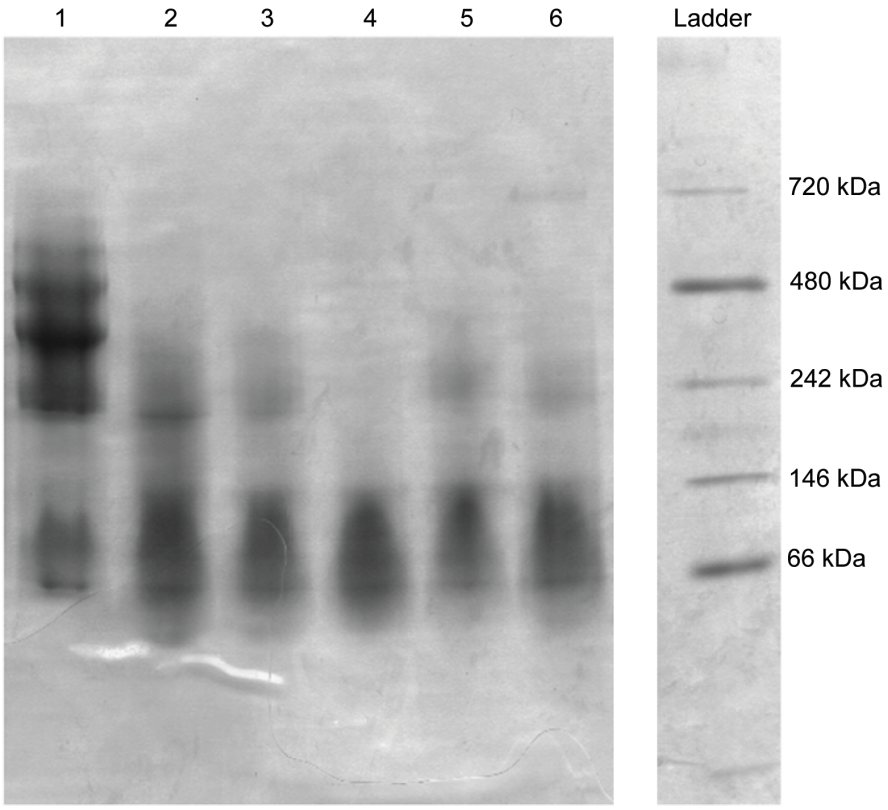

Figure S3.8: Blue Native-PAGE analysis. (1) the recombinant mMDH-CS complex, (2) the recombinant mMDH-CS(R65D) complex, (3) the recombinant mMDH-CS(R67A) complex, (4) the recombinant mMDH-CS(R67D) complex, (5) the recombinant mMDH-CS(R65A) complex, (6) the recombinant mMDH-CS(R65A/R67A) complex and (7) protein ladders. All the complexes were at 20 μM.

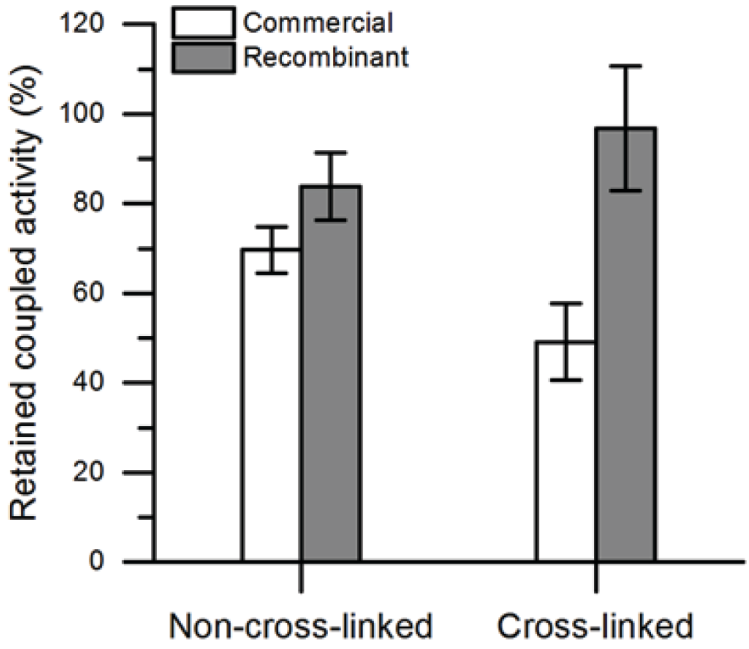

Figure S3.9: Coupled activity retention of the commercial and recombinant mMDH-CS complexes immobilized in Ala-chitosan in the presence of 1 U/mL AAT. Error bars represent standard deviation calculated from three independent experiments. Statistical significance (p-value) was calculated by a two-sample t-test with Welch correction for unequal variances: * = $p < 0.05$, ** = $p < 0.01$.

3.6.2 Supporting Tables

Table S3.1: Identified Krebs cycle enzymes in the cross-linked and non-cross-linked (control) protein bands based on Mascot search results (mass tolerance: 5 ppm). Same sequences matched to database with variable modifications were counted as one unique peptide.

Band name	Enzyme name	Score	Queries matched	Unique peptides	Sequence coverage
Non-cross-linked commercial	mMDH	1598	162	32	67%
	CS	893	63	17	25%
	Aco	805	32	20	30%
Cross-linked commercial	mMDH	782	172	12	44%
	CS	557	424	10	22%
	Aco	97	3	2	3%
Non-cross-linked recombinant	mMDH	82	42	14	48%
	CS	831	96	16	30%
	Aco	2672	197	56	52%
Cross-linked recombinant	mMDH	507	16	9	35%
	CS	449	35	10	18%
	Aco	1223	69	26	32%

Table S3.2: List of identified DSG cross-linked peptides from complexes of commercially purchased mMDH and CS

mMDH peptide	**CS peptide**	**m/z**	**M (expt.)**	**M (calc.)**	**Mass error (ppm)**
GIEKNLGIGK(HDSG)ISPFEEK or GIEK(HDSG)NLGIGKISPFEEK	KTDPRYTC(PAM)QR	1136.2471	3405.7230	3405.7239	-0.26
TIIPLISQCTPKVDFPQDQLSTLTGR	KTDPRYTCQR	1059.2951	4233.1512	4233.1561	-1.16

Table S3.3: List of identified DSG-cross-linked peptides in the recombinant mMDH-CS complex

mMDH peptide	**CS peptide**	**m/z**	**M(expt.)**	**M(calc.)**	**Mass error (ppm)**
GC(PAM)DVVVIPAGVPRK PGM(OX)TR	KTDPR	917.4877	2749.4412	2749.4366	1.67
ASIKK(HDSG)GEEFVK or ASIK(HDSG)KGEEFVK	SM(OX)STD GLIKLVDS K	985.5126	2953.5159	2953.5208	-2.78
GC(PAM)DVVVIPAGVPRK PGMTR	FRGYSIPEC (PAM)QKM LPK	997.2728	3985.0620	3985.0559	1.08
IQEAGTEVVKAK(HDSG)A GAGSATLSMAYAGAR or IQEAGTEVVK(HDSG)AKA GAGSATLSMAYAGAR	GYSIPEC(C AD)QKM(O X)LPK	874.8433	4369.1800	4369.1740	1.73
IQEAGTEVVK(HDSG)AKA GAGSATLSM(OX)AYAGAR or IQEAGTEVVKAK(HDSG)A GAGSATLSM(OX)AYAGAR	GYSIPEC(C AD)QKMLP K	874.8433	4369.1800	4369.1740	1.73

Table S3.4: List of identified DSG cross-links of recombinant Aco and CS peptides

Aco peptide	CS peptide	m/z	M (expt.)	M (calc.)	Mass error (ppm)
RAKDINQEVYNFLATAGAK	GYSIPECQKM(OX)LPK	929.2218	3712.8580	3712.8592	-0.32
AKVAMSHFEPHEYIRYDLLEK	DILADLIPKEQAR	1039.0386	4152.1252	4152.1351	-2.38
AKDINQEVYNFLATAGAK	GYSIPEC(CAD)QKM(OX)LPKAK	954.2368	3812.9180	3812.9101	2.07

Table S3.5: List of identified DSG cross-links of recombinant Aco and CS peptides

Aco peptide	CS peptide	m/z	M (expt.)	M (calc.)	Mass error (ppm)
DFAPGKPCIIK	EVGKDVSDEKLR				
or	or	967.1876	2898.5409	2898.5520	-3.83
DFAPGKPCIIK	EVGKDVSDEKLR				

Table S3.6: CS mutations found in literature

Residue	Mutations	Location
H235[a]	Q	
N242[a]	E	
H274[b]	G, R	OAA / acetyl-CoA binding site
G275[a]	A, V	
H320[a]	R, N, Q, G	OAA binding site
D327[a]	N	
D375[b]	N, Q, E, G	acetyl-CoA binding site
R401[a]	G, H, K	OAA binding site

[a]Evans, et al.[121] [b]Alter, et al.[124]

Table S3.7: Kinetic properties of recombinant CS mutants[a]

Enzyme	Enzyme Concentration (nM)	Specific Activity (s^{-1})	Relative Specific Activity
Recombinant CS	3	28.5	1.00
CS(R65A)	3	24.5	0.86
CS(R67A)	300	0.59	0.02
CS(R65A/R67A)	3000	0.12	0.004
CS(R65D)	300	0.24	0.008
CS(R67D)		N/A	
CS(R65D/R67D)		N/A	

[a]The specific activities were measured in 50 mM Tris-HCl (pH 7.7) with 0.2 mM acetyl-CoA, 0.5 mM OAA and 1 mM DTNB.

Table S3.8: Conservation of Arg65 and Arg67 in CS[a]

Organism	Sequence ID	Match (%)	Residue at Position 65 / 67
Capitella teleta	ELU15766.1	76	E / R
Metaseiulus occidentalis	XP_003739390.1	73	A / R
Ixodes scapularis	XP_002411280.1	73	A / R
Sarcoptes scabiei	KPM02928.1	73	S / R
Limulus polyphemus	XP_013775370.1	72	E / R
Nematostella vectensis	XP_001641037.1	72	E / R
Stegodyphus mimosarum	KFM64927.1	68	E / R
Rhizophagus irregularis	EXX70579.1	65	E / R

[a]The data includes the mismatches between CS Arg65 and Arg67 among the first 500 hits obtained via Basic Local Alignment Search Tool.[146] Sequence identity varied between 64 – 100% for the search**.**

Table S3.9: Primers used to perform site-directed mutagenesis on CS

Mutation	Primer	Sequence
CS(R65A)	Forward	5’ CGGACGAAGGCATTGCTTTTCGCGGTTATTC 3’
	Reverse	5’ GAATAACCGCGAAAAGCAATGCCTTCGTCCG 3’
CS(R67A)	Forward	5’ GAAGGCATTCGTTTTGCCGGTTATTCGATCCC 3’
	Reverse	5’ GGGATCGAATAACCGGCAAAACGAATGCCTTC 3’
CS(R65D)	Forward	5’ CCGGACGAAGGCATTGATTTTCGCGGTTATTCG 3’
	Reverse	5’ CGAATAACCGCGAAAATCAATGCCTTCGTCCGG 3’
CS(R67D)	Forward	5’ CGAAGGCATTCGTTTTGACGGTTATTCGATCCC 3’
	Reverse	5’ GGGATCGAATAACCGTCAAAACGAATGCCTTCG 3’
CS(R65A/R67A)	Forward	5’ GAAGGCATTGCTTTTGCCGGTTATTCGATCCC 3’
	Reverse	5’ GGGATCGAATAACCGGCAAAAGCAATGCCT TC 3’
CS(R65D/67D)	Forward	5’ CGAAGGCATTGATTTTGACGGTTATTCGATCCCG 3’
	Reverse	5’ CGGGATCGAATAACCGTCAAAATCAATGCCTTCG 3’

Chapter 4

Development of a Platform to Study Substrate Channeling in Two-Enzyme Complexes on DNA and Protein Cage Scaffolds

Note: A version of this chapter is being prepared for publication.

Project Collaborators: Beyza Bulutoglu and Scott Banta. KEG performed the genetic cloning for all cMDH variants and Spy-CS-His, designed and performed all experiments, analyzed the data and prepared the manuscript.

4.1 Abstract

There has been recent interest in engineering multi-enzyme complexes with spatially defined enzyme assembly in order to promote substrate channeling and enhanced cascade kinetics. However, much is still unknown about substrate channeling in these engineered systems. Two enzymes that have not been found to naturally channel, cytosolic malate dehydrogenase (cMDH) and citrate synthase (CS) were assembled on DNA and protein cage scaffolds. Engineered two-component icosahedral cages were covalently labeled with cMDH and CS before and after cage assembly with the SpyTag/SpyCatcher system. Zinc-finger DNA-binding proteins (ZFPs) and chemical cross-linking were investigated as methods of enzyme attachment to the DNA scaffolds. The cMDH/ZFP fusion protein was determined to have a dissociation constant of 1.12 ± 0.49 μM for its target DNA, and it was assembled on two DNA scaffolds cross-linked to CS. All multi-enzyme assemblies on the DNA and protein cage scaffolds retained their individual and coupled enzymatic activities. These assemblies can be utilized to study substrate channeling in engineered multi-enzyme complexes.

4.2 Introduction

Engineering complexes of sequential enzymes with defined spatial arrangements has been an area of focus in recent years in order to mimic multi-enzyme complexes called metabolons that naturally form in the cell.[34,35,147] Many natural metabolons have been studied such as tryptophan synthase,[36,148] thymidylate synthase-dihydrofolate reductase (TS-DHFR)[38] and the pyruvate dehydrogenase complex[37] in addition to many more in primary and secondary pathways.[149-152] The formation of these metabolons can have many benefits in the cell including substrate channeling.[3] In substrate channeling, the reaction intermediate of two enzymes is transported directly from one enzyme active site to the next enzyme active site without diffusing in the bulk solution. Substrate channeling can serve to protect unstable intermediates, reduce exposure of the intermediate to competitive enzymes and result in the formation of local substrate concentrations that can drive the reaction in a direction that may otherwise be thermodynamically unfavorable.

Enzyme orientation and active site proximity influence substrate channeling, but proximity alone is often not sufficient for substrate channeling to occur.[40] Diffusion is often fast compared to reaction rates,[39] and Brownian dynamics simulations have predicted that at active distances greater than 1 nm other mechanisms must play a role.[41] There are several mechanisms for substrate channeling including channeling through intramolecular tunnels[36,111,153] and electrostatic channeling,[54,154] in which a network of charged surface residues connects the enzyme active sites along which the oppositely charged intermediate is transported via bounded diffusion. This can result in leaky channeling, in which some of the intermediate is not channeled but is released into the bulk. Confirming a substrate channeling hypothesis can be challenging due to the transient nature of many of these complexes, and, in the absence of structural evidence, multiple indirect kinetic methods to measure substrate channeling should be used.[35]

A natural metabolon that channels by an electrostatic channeling mechanism is comprised of several enzymes from the TCA cycle including mitochondrial malate dehydrogenase (mMDH)[55] and citrate synthase (CS).[25] mMDH catalyzes the reversible conversion of L-malate and oxaloacetate (OAA) utilizing the NAD/NADH cofactor system, and CS catalyzes the condensation of OAA and acetyl-Coenzyme A (acetyl-CoA) to citrate and Coenzyme A (CoA) .[33,61,103,104] Several efforts have been made to study the TCA cycle metabolon structure and its mechanism of channeling, and the non-channeling isoform of mMDH, cytosolic malate dehydrogenase (cMDH), is often used for comparison.[53,54,108] We recently found direct evidence for substrate channeling in a metabolon comprised of recombinant mMDH and CS, and determined that channeling of the intermediate oxaloacetate (OAA) was inhibited with a single R65A mutation in CS along the proposed channel connecting the two enzyme active sites.[61,154]

Most efforts to engineer new multi-enzyme complexes have made use of protein and nucleic acid scaffolds to create *in vivo* an *in vitro* assemblies of sequential enzymes.[155-161] Factors such as defined spatial arrangements, enzyme distance and stoichiometry have been studied. Protein scaffold assemblies have mostly focused on cohesin-dockerin protein-protein interaction domains,[162-164] and mixed substrate channeling results were found for these complexes depending on the enzyme cascades assembled.[165] Nucleic acid scaffolds have been popular due to the ability to control the spatial arrangements at the nanometer scale by changing the nucleic acid sequence. Glucose oxidase (GOx) and horseradish peroxidase (HRP) have been frequently studied as a two-enzyme cascade on DNA scaffolds, including studying the cascade activity as a function of interenzyme distance[166] and the addition of a non-catalytic protein "bridge" between GOx and HRP to facilitate substrate channeling of the intermediate.[39] Many of these arrangements on protein and nucleic acid scaffolds have found improvements in cascade activities, but there

remains a discussion of whether these improvements are a due to channeling of the intermediates or changes in individual enzyme kinetics and what role the scaffold itself may have.[42,147,167,168]

In this work, we focus on a set of four enzymes that perform the same coupled reaction in pairs with different natural channeling capabilities in order to study substrate channeling in engineered multi-enzyme complexes (Figure 4.1a): naturally channeling mMDH and CS, non-channeling cMDH and CS, and mMDH and CS(R65A) in which channeling was inhibited by the R65A mutation in CS. While complexes that include each pairing will be studied, cMDH and CS were used to develop assembly methods due to their lack of natural substrate channeling and high expression rates. Methods for enzyme assembly include DNA and protein cage scaffolds, both of which are discussed in the following sections. By using this set of enzymes and scaffolds, we may be able to better-understand the benefits of constructing these defined spatial arrangements of sequential enzymes and what factors influence substrate channeling.

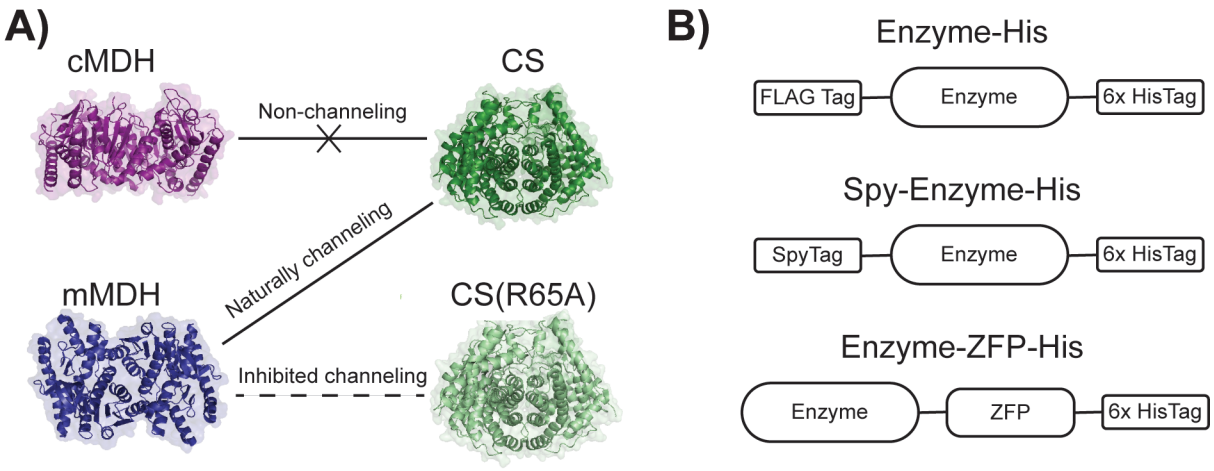

Figure 4.1: Enzymes to study substrate channeling. (A) Structures of enzymes depicting whether substrate channeling occurs between pairs. Crystal structures of cMDH, mMDH and CS were obtained from the Protein Data Bank (PDB ID: 5MDH, 1MLD and 1CTS). (B) Cartoons of enzyme fusion proteins used in the study. Cartoons are not representative of relative sizes.

4.3 Protein Cage Scaffold for Multi-Enzyme Complex Assembly

A computationally designed two-component icosahedral cage was chosen as a scaffold for multi-enzyme complex assembly (Figure 4.2a).[16] This I53-50 cage is named for its icosahedral architecture and naturally pentameric and trimeric components. Each cage is comprised of 20 trimers (A) and 12 pentamers (B) which have been mutated along their threefold and fivefold icosahedral axes, respectively, in order to drive self-assembly. Characterization of cages assembled from purified proteins by Bale et al. showed that they assembled precisely into the designed cage structure.[16] Strategies for enzyme attachment to the protein cage scaffold include genetic fusions to the cage proteins and post-translational labeling of the cage proteins. Post-translational labeling with the SpyTag/SpyCatcher system was chosen so that cage proteins could easily be labeled by a variety of enzymes before or after cage assembly.[169,170] When the SpyTag peptide is mixed with the SpyCatcher protein, an irreversible isopeptide bond rapidly forms between the reactive aspartic acid residue in the SpyTag peptide and the reactive lysine residue in the SpyCatcher protein (Figure 4.3a).

The SpyTag peptide was genetically fused to the N-termini of the codon optimized cMDH-His and CS-His, resulting in the enzymes Spy-cMDH-His and Spy-CS-His (Figure 4.1b), which resulted in similar yields as cMDH-His and CS-His when expressed in *E. coli*. The trimeric cage component (A) and the pentameric cage component (B) were expressed with N-terminal genetic fusions of the SpyCatcher protein, referred to as Catch-A and Catch-B in the following report. When assembled, cages comprised of Catch-A and Catch-B (A/B) are 4.2 MDa in size, and they can be labeled with up to 120 units of enzyme per cage. In this way, we can create ordered, highly dense enzyme complexes on the protein cage scaffold.

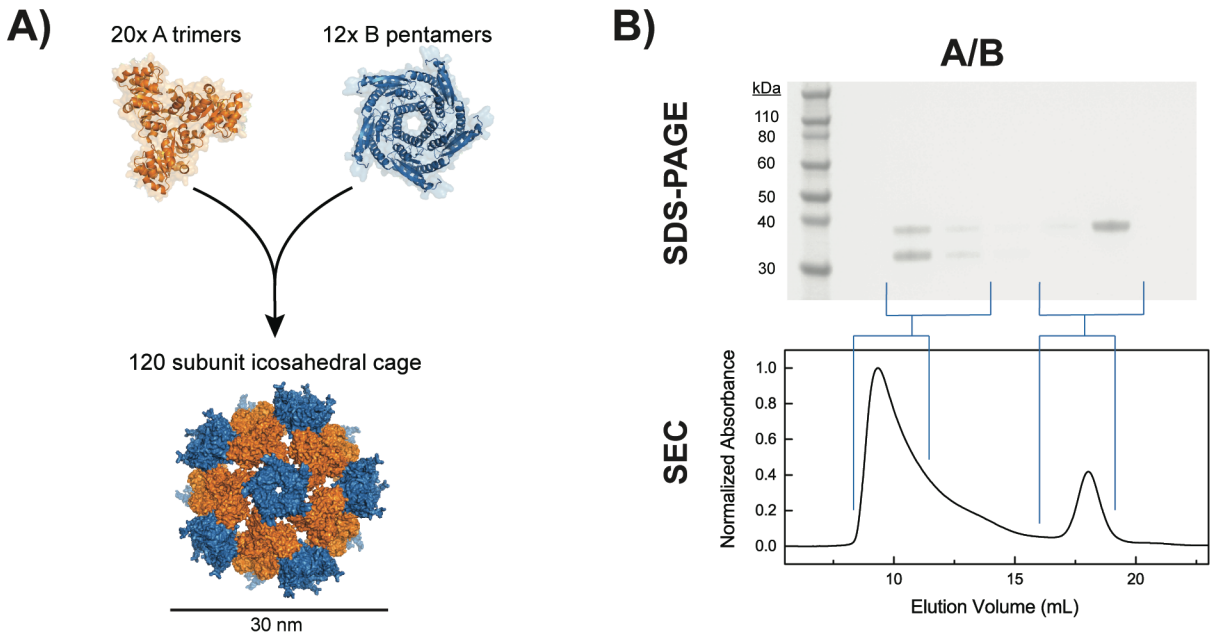

Figure 4.2: Cage assembly. (A) Cage component and assembled cage structures.[16] A and B components self-assemble into the icosahedral cage when mixed. Scale bar only applies to assembled cage. (B) SEC chromatogram (bottom) and SDS-PAGE (top) of assembled unlabeled cage purification. Chromatograms have been normalized by the maximum absorbance. Corresponding fractions between the SEC chromatograms and SDS-PAGE are indicated.

4.3.1 Assembly of Enzyme-Labeled Cage Complexes

Unlabeled cage scaffold assembly

Unlabeled icosahedral cages, referred to as A/B, were assembled by mixing Catch-A and Catch-B at equimolar concentrations and purified by size exclusion chromatography (SEC) (Figure 4.2b). SDS-PAGE of the purification fractions displayed two bands corresponding to the molecular weights of Catch-A (37 kDa) and Catch-B (32 kDa), which self-assemble through non-covalent interactions. These results are similar to what was previously found for the purification of the assembled cage without SpyCatcher fusions, in which cage formation was confirmed by electron microscopy (EM) and small-angle x-ray scattering (SAXS) in addition to SEC chromatograms and

SDS-PAGE.[16] Excess Catch-A was observed in fractions with larger elution volumes during the purification of A/B complexes. No excess Catch-B was observed in later elution volumes even though a similar amount of excess Catch-B as excess Catch-A would be expected. When purified under the same conditions without Catch-B (Figure S4.1), unlabeled Catch-A purified at larger elution volumes than the A/B complexes but smaller than the elution volume of excess Catch-A during A/B complex purification. For unlabeled Catch-B without Catch-A (Figure S4.1) however, two major SEC peaks were observed, the first of which had a similar elution volume as the A/B cage complexes. These results suggest that Catch-B may be forming complexes larger than the expected natural pentamer, and, when combined with Catch-A, these Catch-B complexes cannot form the designed icosahedral cages and are instead purified along with with the A/B cage complexes.

Labeling of individual cage components

The individual cage components Catch-A and Catch-B were labeled with Spy-cMDH-His and Spy-CS-His via the adaptor protein SpyCatcher that is genetically fused to the N-terminus of each cage protein. All four combinations were explored: Catch-A and Spy-cMDH-His formed A-Spy-cMDH, Catch-A and Spy-CS-His formed A-Spy-CS, Catch-B and Spy-cMDH-His formed B-Spy-cMDH and Catch-B and Spy-CS-His formed B-Spy-CS. Cage proteins and enzymes were incubated at equimolar concentrations and conjugation products were purified by SEC (Figure 4.3b). SDS-PAGE analysis of the combined peak fractions exhibited the expected increases in molecular weights for each of the purified conjugation products (Figure 4.3c). Some impurities from unconjugated enzyme and cage components were present in every purified conjugation, and they may be due to incomplete conjugation of all subunits for the naturally multimeric enzymes

and cage components.

Assembly of pre-labeled cage components

Complexes were also assembled with Catch-A and Catch-B that had been pre-labeled with Spy-cMDH-His and Spy-CS-His. Combinations of labeled and unlabeled cage components were prepared and purified by SEC, including complexes pre-labeled with both enzymes (Figure 4.4) and complexes pre-labeled with a single enzyme (Figure S4.2). These include A-Spy-cMDH/B and A/B-Spy-cMDH for complexes pre-labeled with Spy-cMDH-His only, A-Spy-CS/B and A/B-Spy-CS for complexes pre-labeled with Spy-CS-His only and A-Spy-cMDH/B-Spy-CS, A-Spy-CS/B-Spy-cMDH and A-Spy-cMDH/A-Spy-CS/B-Spy-cMDH/B-Spy-CS for complexes pre-labeled with both Spy-cMDH-His and Spy-CS-His. In each case, there are major bands corresponding to the expected molecular weights of each labeled or unlabeled component at similar elution volumes, suggesting that multi-component complexes are being formed. Similar to unlabeled Catch-B, B-Spy-cMDH, A-Spy-CS and B-Spy-CS, which were not combined with a second cage component, were observed to purify at elution volumes similar to the full complexes (Figure S4.3). Complexes may be forming with individual labeled cage components as both enzymes are naturally dimeric and the cage components are also naturally multimeric. For these reasons, it is unclear what percentages of the labeled components formed multi-component complexes or if they assembled with the designed icosahedral cage architecture.

Labeling of pre-assembled cages

Preformed A/B cages were labeled with Spy-cMDH-His and Spy-CS-His post-assembly. Purified

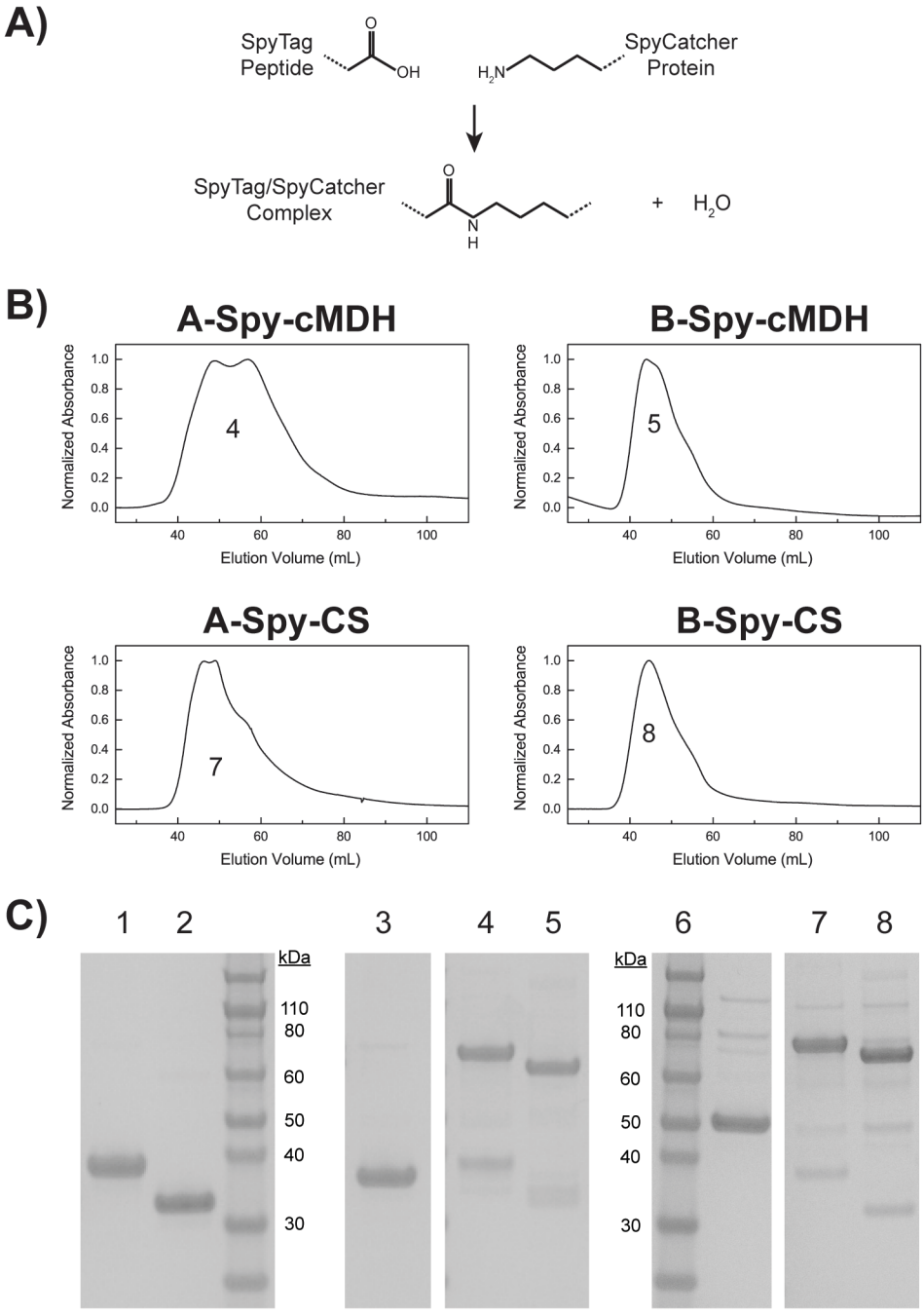

Figure 4.3: Labeling of individual cage components with the SpyTag/SpyCatcher system. (A) Conjugation reaction were the reactive Asp in the SpyTag peptide reacts with the reactive Lys in the SpyCatcher protein to form an isopeptide bond. (B) SEC chromatograms of conjugated Spy-cMDH-His and Spy-CS-His to Catch-A and Catch-B. Chromatograms were normalized by the maximum absorbance. Peak fractions were collected and analyzed on a (C) SDS-PAGE, where peaks are labeled to correspond with gel lane labels. SDS-PAGE lanes are (1) Catch-A, (2) Catch-B, (3) Spy-cMDH-His, (4) A-Spy-cMDH, (5) B-Spy-cMDH, (6) Spy-CS-His, (7) A-Spy-CS and (8) B-Spy-CS.

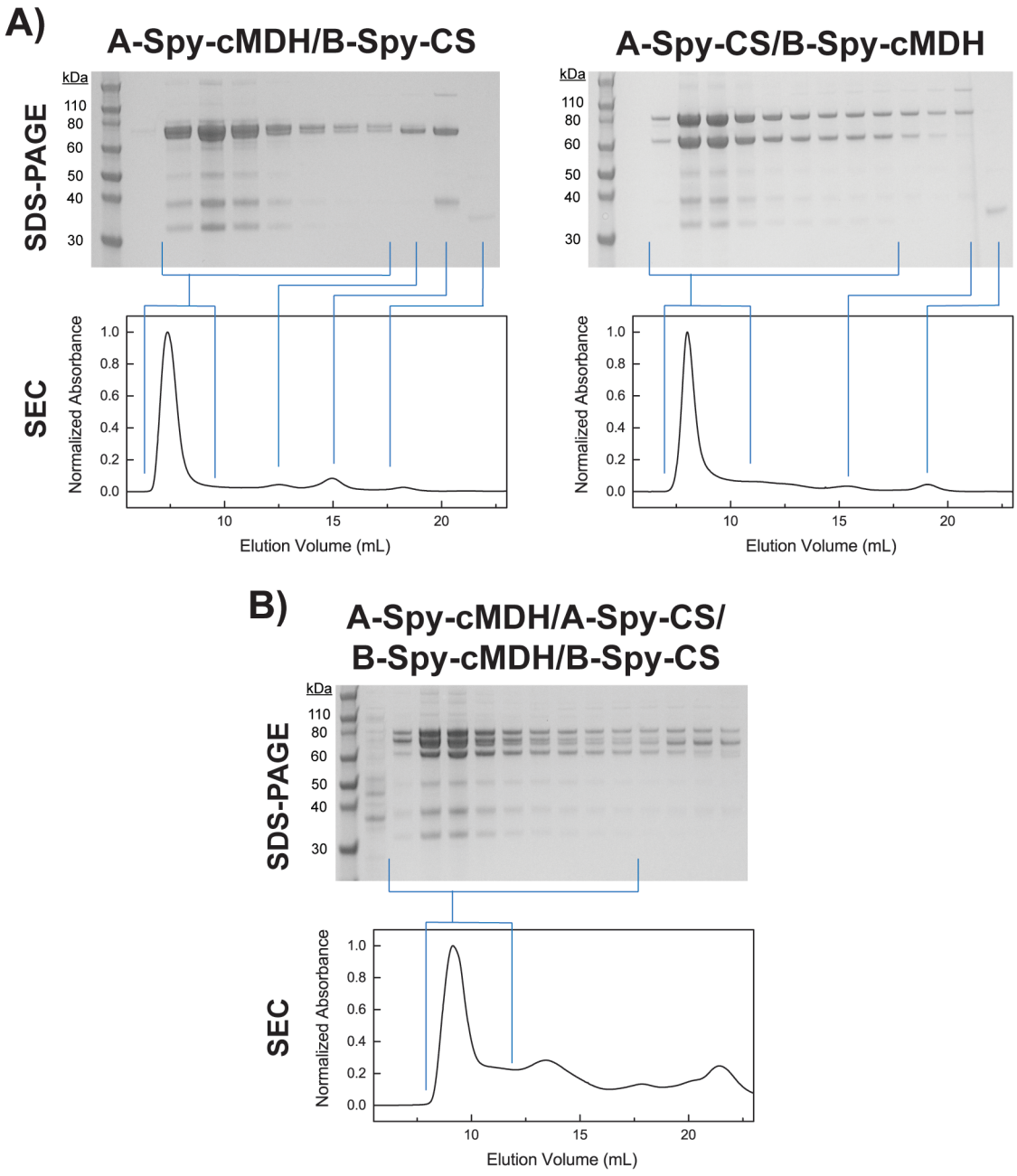

Figure 4.4: Pre-labeled dual-enzyme complexes. Purification of cages assembled with pre-labeled cage components with Spy-cMDH-His and Spy-CS-His. (A) Each component was labeled with one enzyme each. (B) Both components were labeled with both enzymes. Labeled cage components were incubated for self-assembly before purification. Corresponding fractions between the SEC chromatograms (bottom) and SDS-PAGE (top) are indicated, and chromatograms have been normalized by the maximum absorbance.

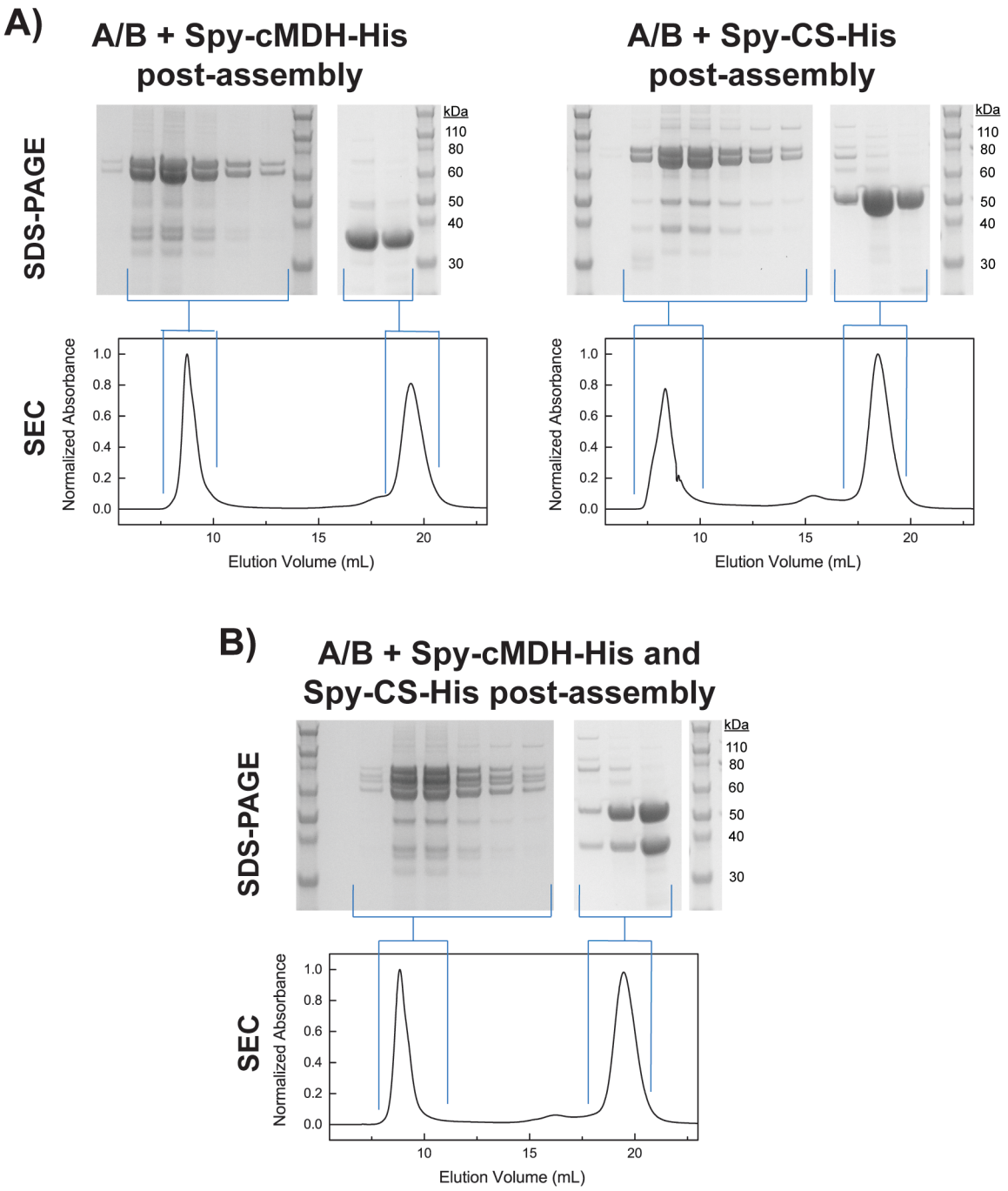

Figure 4.5: Cage labeling post-assembly. Purification of cages labeled after the unlabeled cages were first purified. Purified complexes were incubated with Spy-cMDH-His and Spy-CS-His, and labeled cages were then purified. SEC chromatogram (bottom) and SDS-PAGE (top) for (A) cages labeled with a single enzyme (B) cages labeled with a mixture of both enzymes. Chromatograms have been normalized by the maximum absorbance. Corresponding fractions between the SEC chromatograms and SDS-PAGE are indicated.

A/B complexes were incubated with Spy-cMDH-His, Spy-CS-His and an equimolar mixture of both Spy-cMDH-His and Spy-CS-His, and conjugation products were purified by SEC (Figure 4.5). For A/B complexes labeled with Spy-cMDH-His post-assembly, two major bands were observed matching the expected molecular weights of A-Spy-cMDH and B-Spy-cMDH, the conjugation products of Spy-cMDH-His to Catch-A and Catch-B, respectively. Similar results were observed for A/B complexes labeled with Spy-CS-His post-assembly, in which two major bands corresponding to A-Spy-CS and B-Spy-CS, the conjugation products of Spy-cMDH-His to Catch-A and Catch-B, respectively, are seen. When the A/B complexes were labeled with an equal mixture of Spy-cMDH-His and Spy-CS-His, four major bands were seen corresponding to the four conjugation products. For each case, the observed bands in each purification appear to be of similar sizes, indicating that each expected conjugation reaction is occurring at similar conversion rates. Purification of the labeled complexes showed large separations between the labeled complexes and the excess enzyme in all cases. For each case, some impurities corresponding to non-conjugated enzymes and cage proteins were observed. Impurities may be due to the multimeric structure of each of the components, where not every subunit is conjugated.

4.3.2 Characterization of Cage Assemblies

All individual components and multi-component complexes were analyzed on a non-denaturing agarose gel to shed light on whether multi-component complexes were formed or if the individual components are only being co-purified (Figure 4.6a and Figure S4.4). In the case of the unlabeled components, there was an observable difference between the purified unlabeled A/B complexes and the unlabeled individual components, Catch-A and Catch-B. The lane for A/B complexes

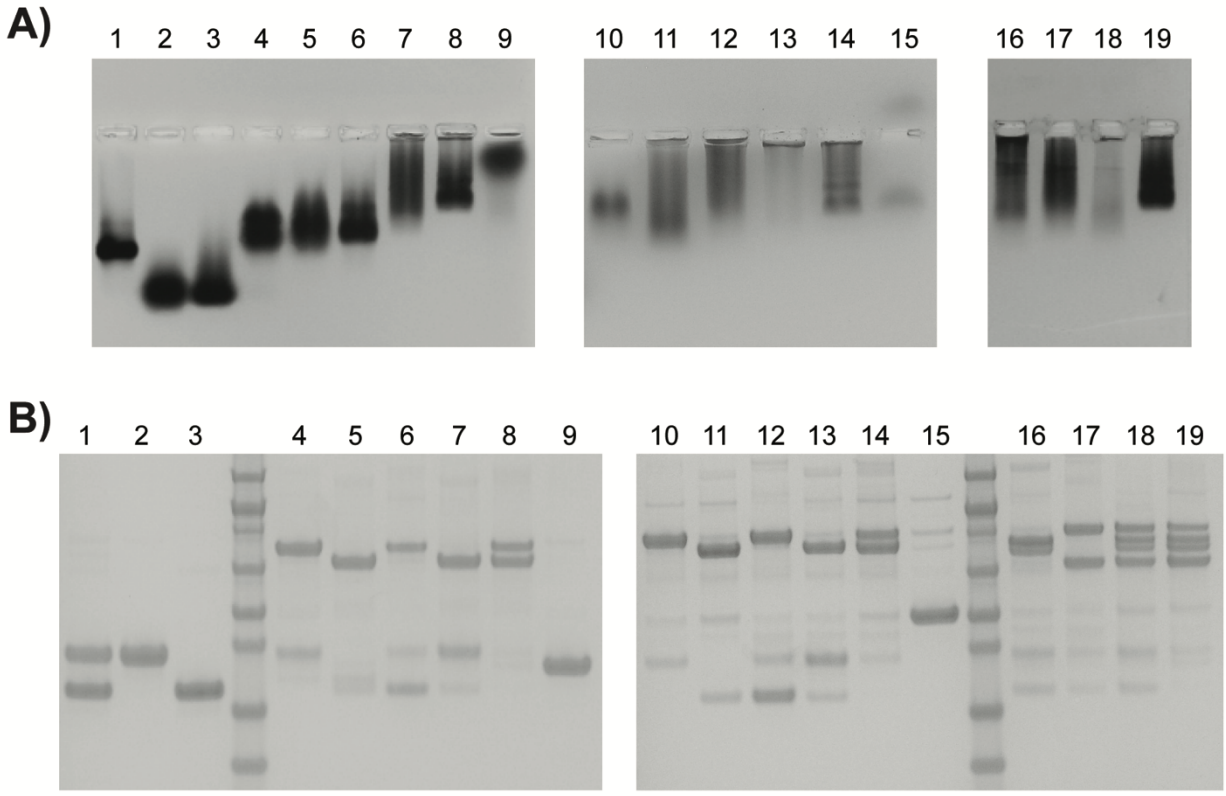

Figure 4.6: Analysis of purified cage complexes. (A) native agarose gels and (B) SDS-PAGE. Lane numbers correspond to the same samples: (1) A/B, (2) Catch-A, (3) Catch-B, (4) A-Spy-cMDH, (5) B-Spy-cMDH, (6) A-Spy-cMDH/B, (7) A/B-Spy-cMDH, (8) A/B + Spy-cMDH-His post-assembly, (9) Spy-cMDH-His, (10) A-Spy-CS, (11) B-Spy-CS, (12) A-Spy-CS/B, (13) A/B-Spy-CS, (14) A/B + Spy-CS-His post-assembly (15) Spy-CS-His, (16) A-Spy-cMDH/B-Spy-CS, (17) A-Spy-CS/B-Spy-cMDH, (18) A/B + Spy-cMDH-His and Spy-CS-His post-assembly. Full native agarose gels are presented in Figure S4.4.

shows one major band that is discernibly higher than those for Catch-A and Catch-B, further indicating that the A/B cage complex was formed. Results for labeled complexes were less clear. Lanes for the complexes A/B-Spy-cMDH, A/B labeled with Spy-cMDH-His post-assembly, A-Spy-CS/B, A/B-Spy-CS, A/B labeled with Spy-CS-His post-assembly, A-Spy-cMDH/B-Spy-CS, A-Spy-CS/B-Spy-cMDH, A-Spy-cMDH/A-Spy-CS/B-Spy-cMDH/B-Spy-CS and A/B labeled

Table 4.1: Specific activities of all cage assemblies and individual labeled components.[a]

Sample Name	MDH (U/mg)	CS (U/mg)	Coupled Enzyme (U/mg)
A/B	1.0 ± 1.1	0.2 ± 0.1	0.0 ± 0.1
A-Spy-cMDH	59.6 ± 1.4	–	–
B-Spy-cMDH	66.2 ± 3.2	–	–
A-Spy-cMDH/B	52.4 ± 3.3	–	–
A/B-Spy-cMDH	62.1 ± 1.8	–	–
A/B + Spy-cMDH-His post-assembly	71.1 ± 2.4	–	–
Spy-cMDH-His	69.3 ± 3.4	–	–
A-Spy-CS	–	38.7 ± 0.2	–
B-Spy-CS	–	47.4 ± 1.2	–
A-Spy-CS/B	–	49.0 ± 0.5	–
A/B-Spy-CS	–	52.5 ± 1.1	–
A/B + Spy-CS-His post-assembly	–	69.9 ± 1.3	–
Spy-CS-His	–	62.1 ± 1.4	–
A-Spy-cMDH/B-Spy-CS	47.7 ± 1.0	54.9 ± 1.5	5.0 ± 0.1
A-Spy-CS/B-Spy-cMDH	58.0 ± 2.8	37.5 ± 1.0	5.5 ± 0.1
A-Spy-cMDH/A-Spy-CS/ B-Spy-cMDH/B-Spy-CS	45.6 ± 1.4	33.5 ± 1.2	4.5 ± 0.1
A/B + Spy-cMDH-His and Spy-CS-His post-assembly	46.0 ± 3.1	28.4 ± 0.6	6.1 ± 0.2

[a] Measurements were performed in triplicate and represented with their mean and standard deviation from the mean.

with Spy-cMDH-His and Spy-CS-His post-assembly all smeared over ranges extending from the loading wells. The individual component B-Spy-CS also smeared from the loading well, making it difficult to differentiate between B-Spy-CS and any multicomponent complex it was included

in. However, the B-Spy-CS band extended further down the lane than those for any of the multi-component complexes it is a part of, indicating that some multi-component complexes may be formed. Differences between individual components and multi-component complexes were more easily discernable in samples that didn't include B-Spy-CS, such as A-Spy-cMDH/B, A/B-Spy-cMDH, A/B labeled with Spy-cMDH-His post-assembly and A-Spy-CS/B-Spy-cMDH. Further analysis of multi-component complexes such as electron microscopy will be required to confirm complex formation and structure.

Samples were assayed for individual enzymatic activities of MDH and CS (Table 4.1). As expected, Catch-A, Catch-B and A/B complexes did not exhibit enzymatic activity for either assay. All conjugated enzymes retained their enzymatic activities, but at this time it is not known what effect conjugation and multi-component assembly have on their enzymatic activities. The coupled enzymatic activities were also measured for multi-enzyme complexes. As expected, all of the multi-enzyme complexes retained exhibited coupled enzymatic activity. Further kinetic analysis is underway, and multiple methods to probe substrate channeling will be performed for each complex.

4.4 DNA Scaffold for Multi-Enzyme Complex Assembly

4.4.1 DNA Attachment with Zinc-Finger Proteins

Non-covalent binding interactions can be employed for the site-specific attachment of enzymes to a DNA scaffold. Zinc-finger proteins (ZFPs) are a class of DNA binding proteins that have been studied extensively, and ZFPs have been engineered to selectively bind to a wide variety of DNA

sequences through techniques such as the construction of chimeras of known ZFPs, phage display, and the rational design of novel ZFPs.[171-174] The three fingered ZPF from the mouse transcription factor Zif268 (3ZF)[175,176] and the designed ZFP AZP4 both have been reported to have nanomolar equilibrium dissociation constants for their ten base pair DNA target sequences (Table S4.1).[171] Both ZFPs have previously been utilized as DNA binding adaptors for the site-specific positioning of fluorophores to DNA scaffolds.[177] We previously genetically fused 3ZF to small laccase (SLAC) from *Streptomyces coelicolor*, creating the bifunctional enzyme SLAC-3ZF-His, for use as an oxygen-reduction catalyst on an enzymatic biocathode.[30] The enzyme/ZFP fusion protein retained both its enzymatic and binding activities. Here, we further explore ZFPs as a way to site-specifically attach multiple enzymes to the DNA scaffold to study substrate channeling.

Table 4.2: Specific activities of enzyme/ZFP fusion proteins[a]

Enzyme	Normalized specific activity
cMDH-3ZF-His	0.79 ± 0.01
CS-AZP4-His	0.81 ± 0.03

[a] Values were normalized by the specific activity of cMDH-His for cMDH-3ZF-His and CS-His for CS-AZP4-His. Measurements were performed in triplicate and represented with their mean and standard deviation from the mean.

The two ZFPs, 3ZF and AZP4, were genetically fused to the C-termini of cMDH and CS, with cMDH fused to 3ZF and CS fused to AZP4 (Figure 4.1b). Each ZFP was followed by a C-terminal polyhistidine tag, resulting in the enzyme/ZFP fusion proteins cMDH-3ZF-His and CS-AZP4-His (Figure S4.5). While the yields for CS-AZP4-His were not low, yields for CS-His (75 mg L^{-1}), [154] the recombinant enzyme without the ZFP domain, were nearly two-fold higher than

those for CS-AZP4-His while yields for cMDH-3ZF-His was similar to cMDH-His (70 mg L^{-1}). The catalytic activities were measured for cMDH-3ZF-His and CS-AZP4-His, and both zinc-finger enzymes exhibited 20% reduced specific activity when compared to cMDH-His and CS-His (Table 4.2). Full kinetic analysis for the ordered bi-bi mechanisms of the enzymes is underway.

Binding of cMDH-3ZF-His and CS-AZP4-His to their DNA target sequences was first confirmed with gel shift assays (Figure 4.7). With a 0.75 μM total DNA concentration, 3ZF-DNA/cMDH-3ZF-His complexes appeared at 0.15 μM, the lowest enzyme concentration tested, and no free DNA was detected at the highest concentration tested, 18.75 μM. cMDH-3ZF-His was also found to bind to the AZP4-DNA and Control-DNA targets with much lower affinities than for the 3ZF-DNA target. CS-AZP4-His also bound to its target DNA sequence AZP4-DNA with lower affinity than what was observed for cMDH-3ZF-His to 3ZF-DNA. The DNA/CS-AZP4-His complex was first detectable at a higher enzyme concentration, 0.75 μM. CS-AZP4-His was not seen to bind to the 3ZF-DNA or Control-DNA targets in the measured concentration range. Neither cMDH-His or CS-His were shown to bind to any DNA target. Two groups of DNA/enzyme complexes can be discerned, particularly for cMDH-3ZF-His. The lower band, likely corresponding to one DNA probe bound to one unit of the enzyme dimer, appears at lower enzyme concentrations and decreases in size at higher enzyme concentrations. The higher band, likely corresponding to two DNA probes with each bound to the two enzyme subunits, emerges at higher enzyme concentrations. The dimeric structure of cMDH, mMDH and CS adds complexity to the system.[55,57] This is especially important when designing DNA scaffolds for enzyme assembly. When using multiple binding sites on the DNA scaffold, larger complexes could form through the binding of multiple strands of DNA on each dimeric enzyme/ZFP fusion protein. This is not specific to the ZFP immobilization strategy, and the formation of these larger complexes must be

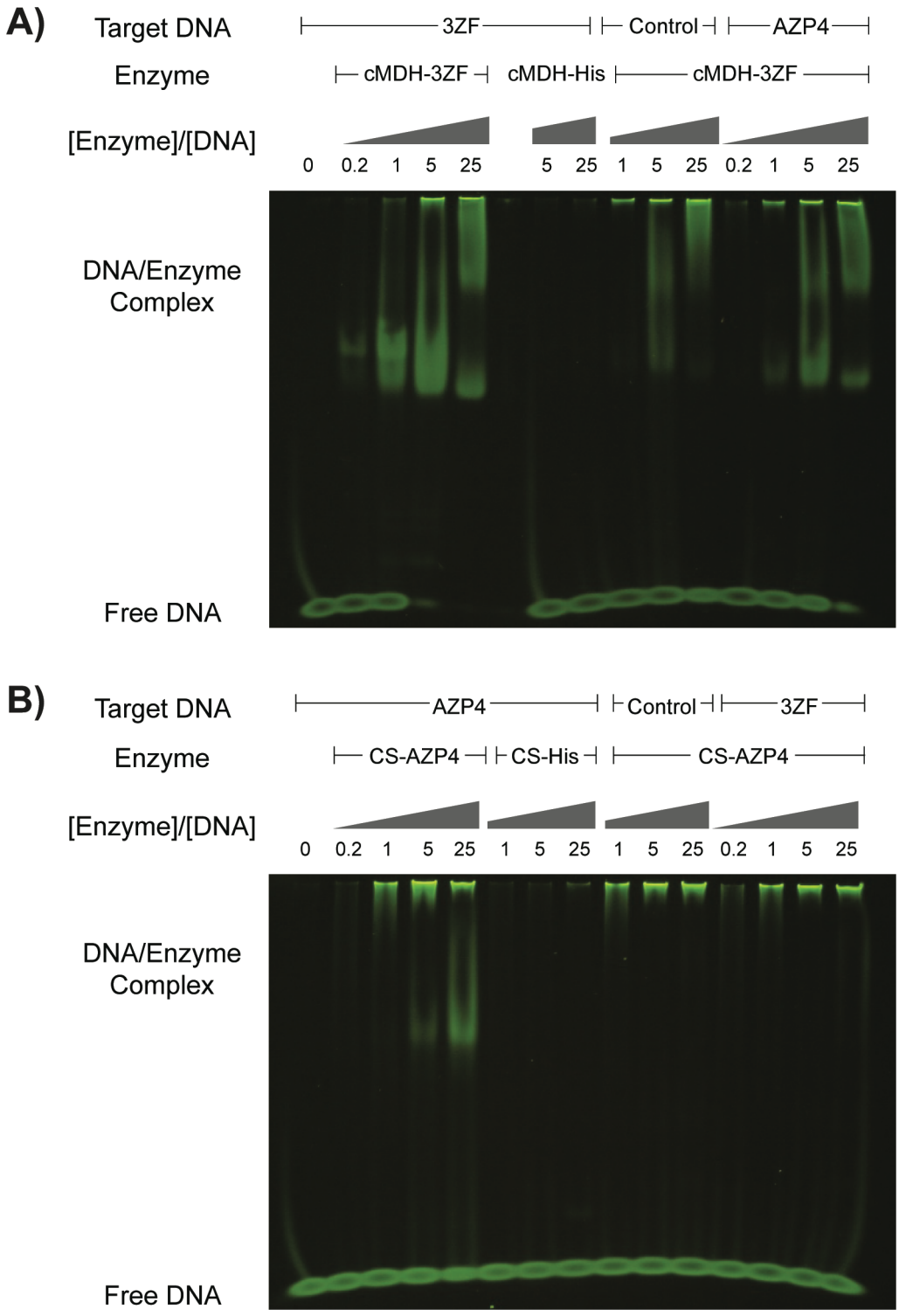

Figure 4.7: Gel shift assays. (A) cMDH-3ZF-His and (B) CS-AZP4-His. 6% DNA retardation gels were run for 85 minutes at 100 V with 0.5x TBE and were stained with SYBR Green.

considered when studying substrate channeling in multi-enzyme complexes assembled on a DNA scaffold.

DNA-binding activities of cMDH-3ZF-His and CS-AZP4-His were further assessed using a magnetic bead capture assay paired with an enzymatic activity assay with mixed results (Figure 4.8).[30] More cMDH-3ZF-His bound to the Bt-3ZF-DNA target sequence than the Bt-AZP4-DNA or Bt-Control-DNA sequences over the concentration range assayed. The dissociation constant, K_D, for the Bt-3ZF-DNA/cMDH-3ZF-His binding isotherm was determined to be 1.12 ± 0.49 μM from a fit of the data. These results are different than reported K_D values for 3ZF (Zif268) ranging from 0.01 to 6.5 nM[171,172] or for a biotin-modified 3ZF protein, which was determined to have a dissociation constant of 63 ± 18 nM.[177] Previous results for SLAC-3ZF-His are similar to what was found for cMDH-3ZF-His.[30] These results indicate that the binding affinity of the ZFP is reduced when expressed as a fusion protein. The total number of available binding sites, 1400 ± 400 pmol/mg beads, was calculated from the fit of the data and specific activity of free cMDH-3ZF-His, which was assumed to remain constant after binding to the DNA. This calculated number of binding sites is higher than the binding capacity of the beads listed by the manufacturer, where greater than 500 pmol/mg of biotin-labeled oligonucleotides and 1000 pmol/mg of free biotin may bind. It is possible that there are fewer binding sites than what was calculated due to an increase in enzymatic activity of bound cMDH-3ZF-His, but this is not yet know.

CS-AZP4-His was not shown to specifically bind to the Bt-AZP4-DNA target under the same conditions as cMDH-3ZF-His over the concentrations tested during the magnetic bead capture assay. (Figure 4.8b). Other binding conditions were explored to improve the binding activity and specificity. The addition of bovine serum albumin (BSA) to reduce nonspecific binding resulted in a two-fold increase in cMDH-3ZF-His to the Bt-3ZF-DNA target (Table S4.2),

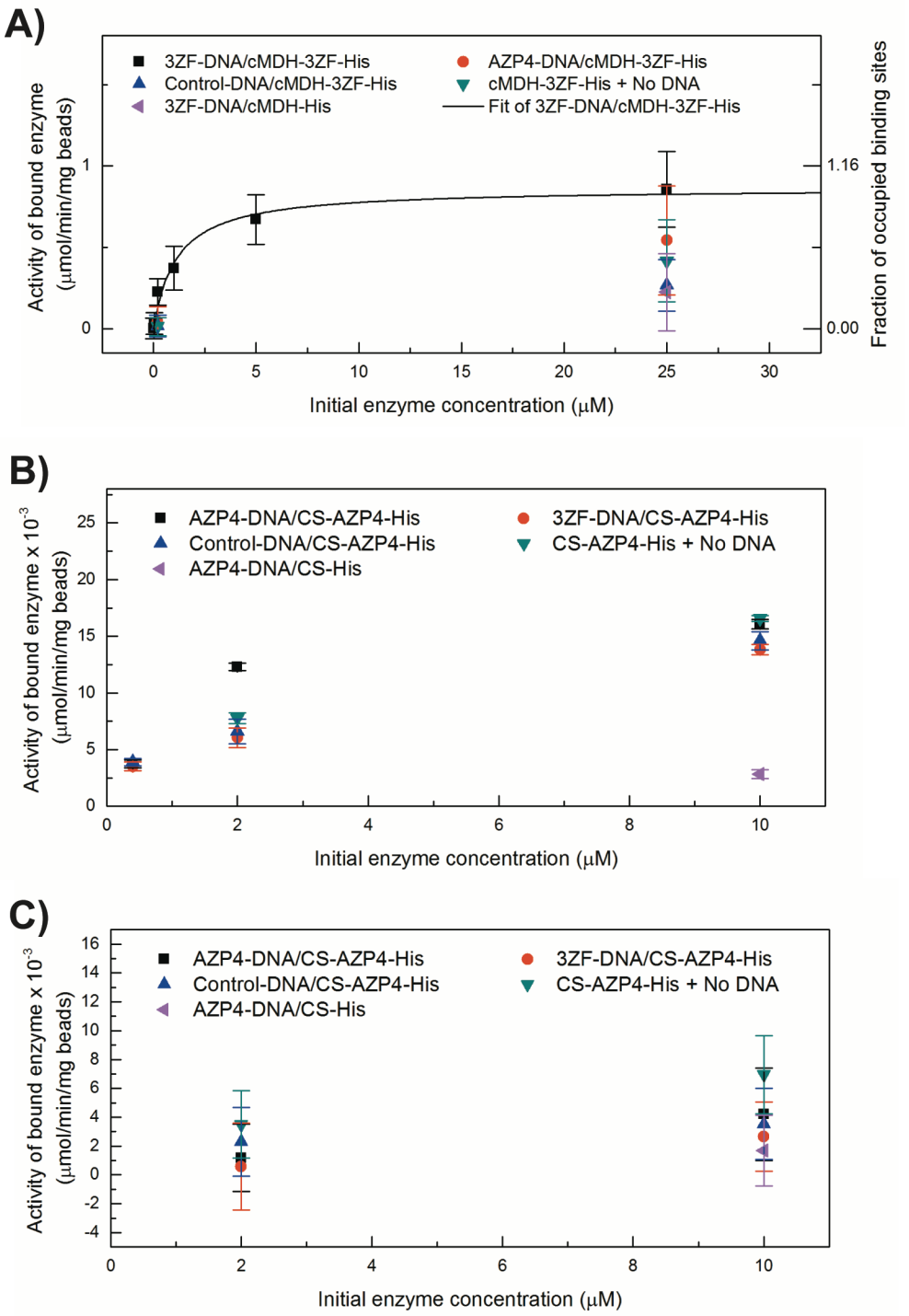

Figure 4.8: DNA binding assay with magnetic bead capture. (A) cMDH-3ZF-His in Bead Assay Buffer, (B) CS-AZP4-His in Bead Assay buffer and (C) CS-AZP4-His in Bead Assay Buffer with 0.5% BSA. The measured enzymatic activities of bound zinc-finger enzymes is on the left axis. For (A), the right axis, % enzyme bound to available binding sites was calculated from the fit of cMDH-3ZF/3ZF-DNA. Measurements were performed in triplicate and represented with their mean and standard deviation from the mean.

but the binding specificity did not improve as there were also large increases in the amount of enzyme bound to the Bt-Control-DNA and Bt-AZP4-DNA targets. The addition of sheared salmon sperm DNA did not improve the binding activity or specificity for cMDH-3ZF-His. The magnetic bead capture biding assay was repeated for CS-AZP4-His with the addition of 0.5% BSA (Figure 4.8c), and no improvement was seen due to the addition of BSA. For both assay conditions, the amount of bound CS-AZP4-His was two orders of magnitude lower than for cMDH-3ZF-His, and it is likely that the binding affinity is lower than what is appropriate for the assay. These results do not agree with previous studies in which the K_D for the AZP4 ZFP was determined to be 11 nM compared to 4 nM for 3ZF.[171,177] It is possible that alternative zinc finger proteins with higher binding affinity may have increased binding affinity than AZP4 when fused to CS. This and alternate immobilization strategies are being explored.

4.4.2 Enzyme/Oligonucleotide Cross-Linking

Covalent attachment of enzymes to DNA oligonucleotides is a common strategy used to assemble multi-enzyme complexes on a DNA scaffold.[39,159,160,178] Modification of lysine and cysteine residues are the most common strategies, and they can be conjugated via chemical crosslinking to oligomers modified with a variety of functional groups.[179] A major disadvantage in these strategies is the lack of control over the stoichiometry. Frequently, there are multiple lysine and cysteine residues that can be modified, which may be a disadvantage when site-specific attachment is required. Several strategies can be employed in order to acquire site-specific oligomer attachment each with their own advantages and disadvantages.[170,180,181]

There are eight cysteine residues and 25 lysine residues in mMDH, five cysteine residues

and 31 lysine residues in cMDH and four cysteine residues and 25 lysine residues in CS and CS(R65A), with a large percentage being surface-accessible for modification. An additional two lysine residues reside in the FLAG-tag on the N-termini of mMDH-His, cMDH-His, CS-His and CS(R65A)-His. Previous work assembling multi-enzyme complexes with a DNA scaffold attached the DNA nonspecifically to surface lysine residues, and these resulted in multiple oligonucleotides attached to the enzymes.[159] This nonspecific attachment results in a lack control over enzyme orientation and number of scaffold attachments, and strategies targeting one residue are preferred.

In order to establish the number of oligonucleotides that may be attached to unmodified cMDH-His and CS-His via thiol modification, the enzymes were incubated with methoxypolyethlene glycol maleimide (PEG-maleimide). For both cMDH-His and CS-His, shifts in the molecular weights of the 39 kDa cMDH-His and 52 kDa CS-His were observed (Figure S4.6) corresponding to one attachment to CS-His. The observed increase in molecular weight was slightly higher than 5 kDa, the molecular weight of the PEG-maleimide, but shifts of up to 15 kDa have been seen when only one cysteine was available for modification. Under these conditions, it is not clear which cysteine residue is being modified or if multiple residues are being modified on separate enzyme units. Cys 110 of cMDH and Cys 184 for CS are potential candidates as they are surface accessible and do not appear to form disulfide bonds within the enzyme dimers (Figure S4.7).

In order to demonstrate CS-His/oligonucleotide attachment for DNA scaffold-directed enzyme assembly, an amino-modified oligonucleotide previously annealed to form the DNA assembly scaffolds S1-DNA and S2-DNA was attached to CS-His through thiol modification with the bifunctional cross-linker, N-ε-maleimidocaproyl-oxysulfosuccinimide ester (Sulfo-EMCS) (Figure 4.9a,b), forming S1-DNA/CS and S2-DNA/CS. The cross-linking efficiencies to CS-His

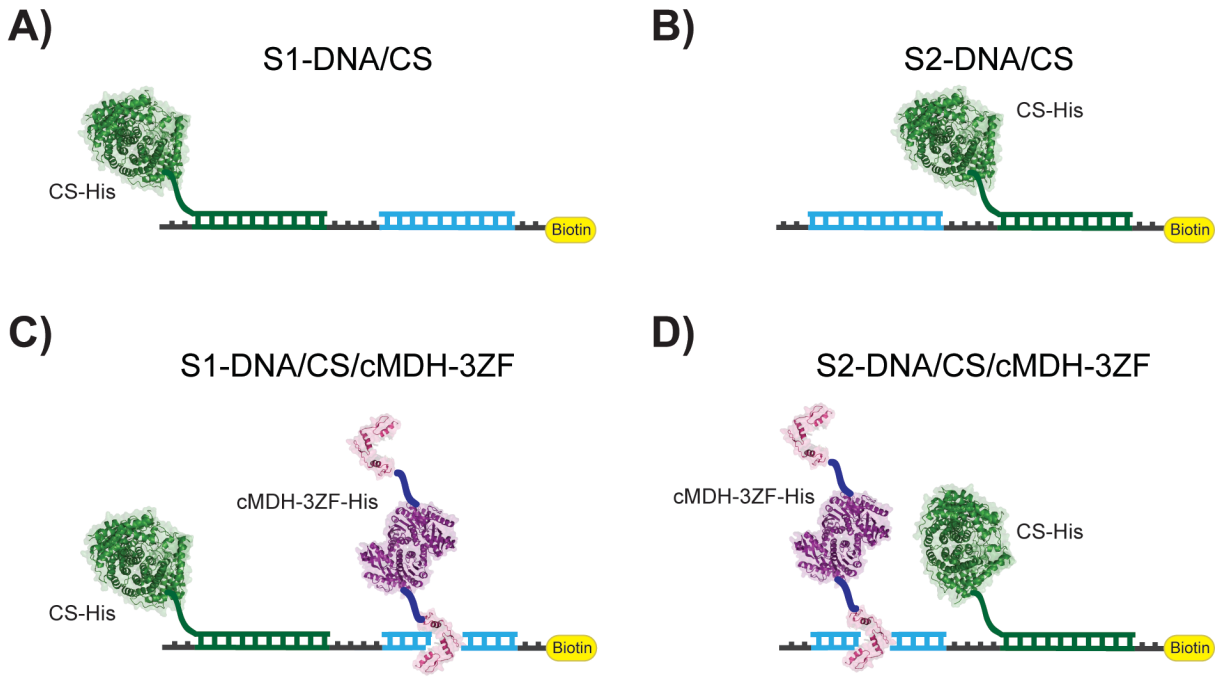

Figure 4.9: Cartoon of designed enzyme/DNA assemblies. (A) S1-DNA/CS-His, (B) S2-DNA/CS-His, (C) S1-DNA/CS-His/cMDH-3ZF-His and (D) S2-DNA/CS-His/cMDH-3ZF-His. Crystal structures of cMDH and CS were obtained from the Protein Data Bank (PDB ID: 5MDH and 1CTS).

Table 4.3: Specific activities for assemblies on S1-DNA and S2-DNA.[a]

Specific activity	S1-DNA/CS	S1-DNA/CS/ cMDH-3ZF	S2-DNA/CS	S2-DNA/CS/ cMDH-3ZF
MDH (U/mg)	–	0.60 ± 0.12	–	0.67 ± 0.06
CS (U/mg)	1.33 ± 0.02	1.06 ± 0.01	1.39 ± 0.01	1.41 ± 0.09
Coupled Enzyme (U/g)	–	26 ± 5	–	32 ± 5

[a] Measurements were performed in triplicate and represented with their mean and standard deviation from the mean.

were assumed to be low as was the case for the PEG-maleimide attachment to CS-His, and non-modified CS-His was removed by washing after the biotin-modified scaffolds were immobilized through magnetic affinity bead capture. CS-His attachment in S1-DNA/CS and S2-DNA/CS was confirmed by measuring the enzymatic activities (Table 4.3). For S1-DNA/CS, 1.33 ± 0.02 U/mg beads of CS activity was measured, and 1.39 ± 0.01 U/mg beads of CS activity was measured for S2-DNA/CS. Assuming the specific activity did not change upon cross-linking to the DNA, these correspond to 117 ± 2 pmol/mg of CS-His immobilized for S1-DNA/CS and 122 ± 1 pmol/mg of CS-His immobilized for S2-DNA/CS. These values are lower than the listed binding capacity of greater than 500 pmol of biotin-modified oligonucleotides for the streptavidin coated magnetic beads. This disagreement may be due to the presence of DNA not cross-linked to CS-His or due to a reduction in CS activity upon DNA cross-linking.

4.4.3 Multi-Enzyme Complexes on DNA Scaffolds

In order to show assembly of cMDH and CS on a DNA scaffold, S1-DNA/CS and S2-DNA/CS were incubated with cMDH-3ZF-His to form S1-DNA/CS/cMDH-3ZF and S2-DNA/CS/cMDH-3ZF (Figure 4.9c,d). Both DNA scaffolds included the 3ZF target sequence on which cMDH-3ZF-His can site-specifically bind. S1-DNA and S2-DNA only differ by the location of the 3ZF target sequence compared to the oligonucleotide cross-linked to CS-His. After washing, 0.58 ± 0.12 U/mg beads of bound cMDH-3ZF-His remained for S1-DNA/CS/cMDH-3ZF, and 0.67 ± 0.06 remained bound for S2-DNA/CS/cMDH-3ZF (Table 4.3). Assuming the activity of bound cMDH-3ZF-His does not differ from free cMDH-3ZF-His, this corresponds to 700 ± 140 and 810 ± 70 pmol/mg beads for S1-DNA/CS/cMDH-3ZF and S2-DNA/CS/cMDH-3ZF, respectively. These

values are within the ranges calculated for the assay conditions and the calculated K_D of cMDH-3ZF-His binding to Bt-3ZF-DNA and are six-fold higher than the calculated amount of CS-His in S1-DNA/CS and S2-DNA/CS, assuming that CS enzymatic activity did not change upon cross-linking. The coupled enzyme activity was also evaluated for both assemblies, with 26 ± 5 U/g of beads measured for S1-DNA/CS/cMDH-3ZF and 32 ± 5 U/g of beads measured for S2-DNA/CS/cMDH-3ZF. A full detailed kinetic analysis of each enzyme in the assemblies and several indirect methods to characterize substrate channeling in S1-DNA/CS/cMDH-3ZF, S2-DNA/CS/cMDH-3ZF and other assemblies on DNA scaffolds is ongoing.

4.5 Future Directions

Multi-enzyme complexes were assembled on DNA and protein cage scaffolds, with several arrangements of cMDH and CS for each scaffold type. For assemblies on DNA scaffolds, there are many more arrangements that can be explored with these immobilization strategies. The DNA scaffold can be used as a molecular ruler in order to study the effects of distance, and additional proteins with a range of surface charges can be included in the assembly to interact with the charged intermediate OAA. Multiple methods of enzyme attachment to the DNA scaffold can be studied by using ZFPs and covalent oligonucleotide attachment in order to explore the role that they may play on channeling, such as interactions with the channeling intermediates. The SpyTag/SpyCatcher attachment system can also be applied to the DNA scaffold assemblies in order to simplify the system for a large number of enzyme pairs. The SpyCatcher protein could be site-specifically attached to the oligonucleotides and then labeled with SpyTag/enzyme fusion proteins.

An important consideration going forward is the quaternary structure of mMDH, cMDH, CS and CS(R65A). For both the DNA and protein cage scaffolds, the dimeric enzymes may form larger, more complex structures than what is intended. It is important that conditions are used to assemble the preferred complexes. Careful analysis of the protein cage complex will be performed to verify that cages are assembling in their designed icosahedral form. The comparison of multi-component cage complexes that have been labeled with both Spy-cMDH-His and Spy-CS-His post-assembly and complexes of pre-labeled cage components will be compared in order to optimize assembly strategies for the formation of the designed cage structure.

Now that multi-enzyme complexes with cMDH and CS have been assembled, the same complexes will be assembled with the other two enzymes, mMDH and CS(R65A). Combinations of the four enzymes will be compared for the different arrangements: the naturally channeling enzymes mMDH and CS, the naturally non-channeling enzymes cMDH and CS, and mMDH and CS(R65A) in which substrate channeling has been inhibited by the R65A mutation. With this set of enzymes, structural differences of the two MDH isoforms can also be explored, and the results from assembly on these scaffolds can be compared to the naturally channeling mMDH/CS complexes.

For each of the multi-enzyme complexes assembled, substrate channeling between the enzymes will be characterized. Several methods will be employed including the comparison of initial coupled enzymatic rates, transient time analysis and the addition of the competitive enzyme aspartate aminotransferase.[154] Several outcomes are possible including the support or the rejection of a substrate channeling hypothesis for each of the arrangements studied. For example, we may find increased coupled enzymatic activities compared to the freely diffusing enzymes, but find that these improvements are not due to channeling but to changes in the individual enzymatic activities.

A careful kinetic analysis of these assemblies may lead to a better understanding of leaky substrate channeling in engineered complexes to be better-understood.

4.6 Conclusions

We have developed a platform to study substrate channeling in engineered multi-enzyme complexes. The naturally non-channeling enzymes cMDH and CS were assembled on DNA and protein cage scaffolds, which were designed so that cMDH and CS can be exchanged for the other two pairs of enzymes in the set, mMDH and CS, which naturally channel, and mMDH and CS(R65A), in which channeling was found to be inhibited. Engineered two-component protein cage complexes were covalently labeled with cMDH and CS in different arrangements through the SpyTag/SpyCatcher labeling system, with proteins labeled before and after cage complex assembly. For DNA directed assembly, cMDH-3ZF-His was shown to bind to its target DNA with micromolar affinity, and it was attached two DNA scaffolds cross-linked to CS-His. While each complex retained their individual and coupled enzymatic activities, substrate channeling has yet not been demonstrated in any of the multi-enzyme complexes and is currently being studied further. By using this platform to study a set of enzymes that includes channeling and non-channeling enzymes that catalyze the same reactions and multiple methods to assemble those enzymes, we will be able to better-understand substrate channeling in engineered multi-enzyme complexes.

4.7 Materials and Methods

4.7.1 Materials

Synthetic genes for cMDH and AZP4 were synthesized by Genscript (Piscataway, NJ). Restriction enzymes for DNA cloning were purchased from New England Biolabs (Ipswich, MA). Isopropyl β-D-1-thiogalactopyranoside (IPTG) and ampicillin sodium salt, were purchased from Gold Biotechnology (St. Louis, MO). All purification columns were purchased from GE Healthcare Life Sciences (Pittsburgh, PA). Amicon centrifugal filters were purchased from Millipore. Disuccinimidyl glutarate (DSG), sodium dodecyl sulfate polyacrylamide electrophoresis gels (SDS-PAGE), TBE DNA gels and running buffers were purchased from Invitrogen-Life Technologies (Carlsbad, CA). *E. coli* BL21 and BL21(DE3) cell lines were purchased from Bioline (Taunton, MA). Chaperon plasmid pGro7 was purchased from Clontech Laboratories–Takara (Mountain View, CA). All other reagents and materials were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise stated.

4.7.2 Construction of Recombinant Enzymes

Constructs encoding mMDH-His and CS-His were assembled previously.[154] Constructs encoding the genes for Catch-A and Catch-B were a kind gift from David Baker (University of Washington). The gene encoding the codon optimized cMDH with a Flag-tag at the N-terminus was PCR amplified using primers P1 and P2 (Table S4.3). The reverse primer introduced a C-terminal 6xHis-tag. The resulting NdeI/HindIII-digested PCR product was ligated into a NdeI/HindIII-digested pET-20b(+) backbone. The resulting construct was transformed into BL21(DE3) cells for

cMDH-His expression (Figure 4.1b).

The gene encoding 3ZF was PCR amplified from SLAC-3ZF-His with primers P3 and P4 (Table S4.3).[30] The codon optimized gene for AZP4 was flanked by HindIII restriction sites. A construct encoding the gene for mMDH-3ZF-His was first constructed by amplifying mMDH with primers P5 and P6, which was digested with NdeI and HindIII and then ligated into pET-20b(+). The resulting construct was digested with HindIII, and the HindIII-digested 3ZF gene was inserted into this site. In order to construct cMDH-3ZF-His, cMDH was first PCR amplified with P7 and P8, and the NdeI/HindIII-digested PCR product was ligated into the pet20-b(+) backbone. The resulting construct was digested with HindIII and was ligated 3ZF gene cut from mMDH-3ZF-His by with HindIII. CS-AZP4-His was constructed by PCR amplifying CS with P9 and P10, which was restriction digested with NdeI and HindIII and ligated into the pET-20b(+) backbone. The resulting construct and the codon optimized gene for AZP4 were digested with HindIII and ligated. The resulting constructs were transformed into BL21(DE3) cells for cMDH-3ZF-His and CS-AZP4-His expression (Figure 4.1b).

To construct Spy-cMDH-His and Spy-CS-His, cMDH-His and CS-His were PCR amplified with primers P11 and P12 for cMDH-His, P13 and P14 for CS-His (Table S4.3). The forward primers, P11 and P13 included sequences encoding the SpyTag peptide (AHIVMVDAYKPTK) followed by a six amino acid linker preceded by a NdeI restriction site. Nde/HindIII-digested genes for Spy-cMDH-His and Spy-CS-His were ligated into the pET-20b(+) backbone. The resulting constructs were transformed into BL21(DE3) cells for Spy-cMDH-His and Spy-CS-His expression (Figure 4.1b).

4.7.3 Expression and Purification of Enzymes

For all cMDH and CS constructs, cells were grown in 1 L of sterilized Terrific Broth, inoculated with 10 mL of overnight culture. All cultures were supplemented with 100 µg/mL ampicillin, and cultures containing mMDH-His and mMDH-3ZF-His were also supplemented with 35 µg/mL chloramphenicol. Cells were grown to an OD_{600} of 0.6 at 37˚C at which time protein expression was induced with 0.5 mM IPTG and continued for 20 hours at 25˚C. Cells were collected by centrifugation and stored at -20˚C. Thawed cell pellets corresponding to 1 L of culture was resuspended in 50 mL of Enzyme HisTrap Binding buffer (20 mM Tris-HCl, 150 mM NaCl, 20 mM imidazole, pH 7.4) supplemented HALT EDTA-free protease inhibitor. Cells were lysed by sonication for six minutes (5 s on pulse and 2 s off pulse) with a microtip probe in an ice bath, and lysates were centrifuged for 30 minutes at 15,000 × g. Clarified lysates were applied to HisTrap FF columns, and proteins of interest were eluted with a gradient of the Enzyme HisTrap Elution Buffer (20 mM Tris-HCl, 150 mM NaCl, 500 mM imidazole, pH 7.4). Fractions containing the enzymes of interest as identified by enzymatic activity and SDS-PAGE were collected and buffer exchanged by ultrafiltration into the next required buffer using Amicon filters. All enzyme concentrations were measured by Bradford assay.

Plasmids encoding the genes for the trimeric (Catch-A) and pentameric (Catch-B) cage I53-50 cage components with the SpyCatcher protein on the N terminus and a 6xHis-tag on the C terminus of both proteins. Plasmids were transformed into BL21(DE3) cells for protein expression. Flasks with 1L LB supplemented with 50 µg/mL kanamycin were induced with 30 mL of overnight culture. Cells were grown to an OD_{600} of 0.6 at 37˚C. Protein expression was induced with 1 mM IPTG and continued for 3 hours at 37˚C for Catch-A and 18˚C for 5 hours for Catch-B. Cells were collected by centrifugation and pellets corresponding to 2 L of culture was resuspended in 50 mL

of Cage HisTrap Binding Buffer (50 mM Tris-HCl, 0.5 M NaCl, 20 mM imidazole, 0.75% CHAPS, 1 mM DTT, pH 8.0) supplemented with HALT EDTA-free protease inhibitor. Once resuspended, cells were stored at -20˚C until later use. Cells were lysed by sonication for 6 minutes (5 s on and 2 s off) in an ice bath with a microtip probe and centrifuged for 30 minutes at 15,000 × g. Clarified lysates were applied to HisTrap FF columns, and proteins of interest were eluted with the stepwise addition of Cage HisTrap Elution Buffer (50 mM Tris-HCl, 0.5 M NaCl, 0.5 M imidazole, 0.75% CHAPS, 1 mM DTT, pH 8.0). Fractions containing the desired proteins were combined, concentrated by ultrafiltration and proteins were further purified on the same day on a size exclusion column equilibrated with Cage Assay Buffer (50 mM Tris-HCl, 0.5 M NaCl, 1 mM DTT, pH 8.0). Fractions containing the proteins of interest identified by SDS-PAGE were collected, concentrated by ultrafiltration and stored for later use.

4.7.4 Characterization of Kinetic Activities

Individual specific activities

Specific activities of all cMDH variants were measured in Activity Assay Buffer (100 mM potassium phosphate buffer, pH 7.4) with 0.1 mM (OAA) and 0.1 mM NADH. NADH consumption was monitored spectrophotometrically at 340 nm. Enzyme concentrations were typically 0.1-10 nM. One unit of and cMDH activity was defined as the amount of enzyme necessary to consume 1 μmol of NADH per minute. Specific activities of all CS variants were measures in assay buffer with 0.1 mM acetyl-coenzyme A, 0.1 mM OAA and 0.2 mM DTNB (5,5'-dithiobis(2-nitrobenzoate)). Production of coenzyme A was monitored through the subsequent reaction of coenzyme A and DTNB, yielding TNB^{2-}, which absorbs at 412 nm. One

unit of CS activity was defined as the amount of enzyme necessary to produce 1 μmol of coenzyme A per minute. All spectrophotometric assays were performed on a SpectraMax M2 (Molecular Devices, Sunnyvale, CA). All measurements were performed with at least three replicates. Extinction coefficients of 14,150 M^{-1} cm^{-1} for DTNB and 6220 M^{-1} cm^{-1} for NADH were used for calculations.

Coupled enzymatic activity

Coupled enzymatic activity was measured in Activity Assay Buffer with 1 mM L-malate, 2 mM NAD^+, 0.1 mM acetyl-coenzyme A and 0.2 mM DTNB. The reaction was initiated by the addition of 10% of the reaction volume of bead assemblies or controls, and the reaction was monitored by measuring the subsequent reaction of a final product coenzyme A and DTNB at 412 nm. One unit (U) of enzymatic activity was defined as 1 μmol of coenzyme A produced per minute.

4.7.5 Cage Assembly

All cages were assembled in Cage Assay Buffer. The A and B components, unlabeled or labeled with the enzymes, were combined at equimolar concentrations. When unlabeled cages were assembled to be labeled post-assembly were combined at 200 μM each. For all other assemblies, components were combined at 20 μM each. Samples were incubated at 4˚C with rotation overnight and purified on a Superose 6 10/300 gel filtration column. Fractions were analyzed with SDS-PAGE, and those containing proteins in the desired assembly state were collected, concentrated by ultrafiltration and stored at 4˚C for later analysis.

4.7.6 SpyTag/SpyCatcher Conjugation

Labeling of individual cage components

For SpyTag/SpyCatcher conjugation, Spy-cMDH-His or Spy-CS-His was combined with Catch-A or Catch-B at 200 μM each in Cage Assay Buffer. Samples were incubated overnight at 4˚C with rotation. Samples were applied to a HiLoad 16/60 Superdex 200 size exclusion column equilibrated with Cage Assay Buffer for purification. Fractions were analyzed by SDS-PAGE, and those with the purified conjugation product were collected and concentrated by ultrafiltration.

Post-assembly labeling of cages

Purified assembled cages were incubated with Spy-cMDH-His, Spy-CS-His or an equimolar mixture of the two enzymes. Cage complexes with concentrations corresponding to 20 μM of both Catch-A and Catch-B were mixed with 200 μM of the total enzyme. Samples were incubated overnight at 4˚C with rotation and purified with a Superose 6 10/300 gel filtration column. Fractions were analyzed by SDS-PAGE, and fractions containing the desired proteins were collected, concentrated by ultrafiltration and stored at 4˚C until for analysis.

4.7.7 Characterization of Cage Assemblies

Native agarose gel for large complex separation

All individual proteins and cage assemblies were analyzed on an agarose gel under native conditions. The gels were prepared with 1% agarose with TB running buffer (90 mM Tris-Borate, pH 8.5) was allowed to set for several hours at 4˚C. Loading buffer (0.12 M Tris base, 20%

glycerol) was added to an equal volume of 1 μg/μL protein. For each sample, 10 μg of protein was loaded per well. Each gel included Catch-A, Catch-B and A/B complexes for reference. Gels were run at 45 V for 5 hours at 4˚C with TB running buffer. Gels were stained for 1 hour with 0.125% Coomassie R-250 in 40% methanol and 10% acetic acid. Gels were destained in several washes in 40% methanol and 10% acetic acid and imaged for analysis.

Kinetic characterization of cage assemblies

The individual enzyme activities were measured for cMDH and CS activity for all individual cage components and complexes as previously described. Coupled activities were measured for all multi-enzyme complexes. Unlabeled A/B was also assayed for coupled enzyme activity.

4.7.8 Gel Shift Assays

Protein binding to DNA was analyzed using a fluorescence-based gel shift assay. 3ZF-DNA and AZP4-DNA include target sequences for 3ZF and AZP4, respectively, and Control-DNA includes a control sequence. Target Target DNA was prepared by combining equimolar amounts of forward and reverse oligos (O1 and O3 for 3ZF-DNA, O4 and O6 for AZP4-DNA and O7 and O8 for Control DNA, Table S4.4) in Annealing Buffer (10 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA, pH 8.0). Oligos were incubated in a heat block at 95˚C for ten minutes and allowed to cool in the heat block for two hours at room temperature. DNA annealing was confirmed by running annealed DNA samples and individual oligos on a 20% polyacrylamide TBE gel. 3ZF-DNA, AZP4-DNA and Control-DNA at 0.75 μM final concentration were incubated with cMDH-3ZF-His and CS-AZP4-His at several concentrations ranging from 0 to 18.75 μM in Binding Assay Buffer (50 mM

Tris-HCl, 0.5 M NaCl, 0.05% IGEPAL, 1 mM DTT, 0.1 mM $ZnSO_4$, pH 8.0) for two hours at 4˚C. Target DNA was also incubated with cMDH-His and CS-His as a control. Samples were mixed with 5x Loading Buffer (Binding Assay Buffer with 50% glycerol), and loaded onto a 6% DNA retardation gel prepared with 0.5x TBE. Gels were run with 0.5x TBE running buffer at 4˚C for 85 minutes at 100 V. Gels were stained with SYBR Green (Thermo Fisher Scientific) nucleic acid stain according to the manufacturer's instructions and imaged on a UV light box.

4.7.9 DNA Binding with Magnetic Bead Capture

Protein binding to DNA targets were characterized with a bead capture assay using streptavidin coated magnetic beads (New England Biolabs). Bt-3ZF-DNA, Bt-AZP4-DNA Bt-Control-DNA includes a control sequence consist of the same sequences as 3ZF-DNA, AZP4-DNA and Control-DNA with the addition of biotin to the 5' end of the forward primer (O2 for Bt-3ZF-DNA, O5 for Bt-AZP4-DNA and O8 for Bt-Contol-DNA. Annealed DNA was prepared as described previously. 0.8 mg of magnetic beads were washed three times and incubated in 1 mL of Blocking Buffer (50 mM Tris-HCl, 150 mM NaCl, 0.5% BSA) for 1 hour at 4˚C with rotation. Beads were resuspended in 0.8 mL of Binding Assay Buffer with 0.75 μM of Bt-3ZF-DNA, Bt-AZP4-DNA or Bt-Control-DNA and incubated overnight at 4˚C with rotation. Beads were washed six times with Binding Assay Buffer and resuspended in Binding Assay Buffer. Beads were aliquoted, and enzyme was added to the samples at various concentrations in 200 μL total volume. For cMDH-3ZF-His, 0.01 mg of beads per 200 μM sample. When buffer conditions were compared in cMDH-3ZF-His, 0.1 mg of cMDH-3ZF-His was used in each 200 μL with 10 μM of enzyme. Buffer additions compared included 0.5% BSA, 100 μg/mL sheared salmon sperm DNA (Invitrogen) and both BSA and

sheared salmon sperm DNA. For CS-AZP4-His, 0.1 mg of beads was used per 200 µL sample. CS-AZP4-His was assayed in Binding Assay buffer with no additions and Binding Assay Buffer with 0.5% BSA.

4.7.10 Enzyme/Oligonucleotide Attachment

PEG- maleimide attachment

In order to evaluate the number of cysteine residues available for modification, cMDH-His and CS-His were first attached to PEG-maleimide. 10 µM cMDH-His was incubated with 0.1, 10 and 100 µM PEG-maleimide overnight at 4˚C. 100 µM CS-His was incubated with 1 mM PEG-maleimide and incubated overnight at 4˚C. Samples were analyzed by SDS-PAGE (Figure S4.6).

Oligonucleotide attachment to CS

Scaffold DNA was first assembled. Oligos were combined in Annealing Buffer with ratios of O10:O11:O12 of 2:2:1 for S1-DNA and O10:O11:O13 of 2:2:1 for S2-DNA (Table S4.4). Oligos O10 (3ZF binding site) and O11 (amimo-modified oligonucleotide for CS attachment) were supplied in excess, and O12 and O13 included biotin on their 5' ends. Samples were annealed at 70˚C for 10 minutes on a heat block and allowed to cool with the heat block to room temperature. Samples were incubated at 4˚C before use. Before enzyme attachment, 100 µL of S1-DNA and S2-DNA (75 µM final concentration of O11 and O12) were buffer exchanged six times into Activity Assay Buffer. Volumes were brought to 0.9 mL, and 100 µL of freshly made 10 mM Sulfo-EMCS was added to each sample. Samples were incubated for 2 hours at 4˚C with rotation, then buffer exchanged four times into Activity Assay Buffer to remove excess cross-linker.

Volumes were brought to 1 mL with a final concentration of 75 μM CS-His. Samples were incubated for 2 hours at 4˚C with rotation then buffer exchanged six times into Activity Assay Buffer to remove non-cross-linked DNA. Volumes were brought to 500 μL. DNA-attached CS-His samples, now referred to as S1-DNA/CS and S2-DNA/CS, were immobilized on streptavidin coated magnetic beads. 400 μL of S1-DNA/CS and S2-DNA/CS were mixed with 400 μL of beads in Binding Assay Buffer. Beads were previously washed three times and blocked for 1 hour at 4˚C with rotation in Blocking Buffer. 0.8 mg of beads total were added to each sample, and they were incubated overnight at 4˚C with rotation. Beads were washed six times with Binding Assay Buffer. For kinetic analysis, bead-immobilized S1-DNA/CS and S2-DNA/CS were resuspended in Activity Assay Buffer.

4.7.11 DNA/Multi-Enzyme Complex Assembly and Characterization

Complex Assembly

S1/CS and S2/CS were aliquoted into samples containing 0.1 mg of beads. Samples were incubated with 25 μM cMDH-3ZF-His in Binding Assay Buffer with 200 μL total volume. Samples were incubated for five hours at 4˚C with rotation. Beads were washed three times with Binding Assay Buffer and resuspended in Activity Assay Buffer.

Kinetic Characterization of Assemblies

The individual enzyme activities were measured for individual and coupled enzyme activities in S1-DNA/CS, S2-DNA/CS, S1-DNA/CS/cMDH-3ZF and S2-DNA/CS/cMDH-3ZF assemblies as previously described. Typically, 1 μg of beads were assayed.

4.8 Supplemental Information

4.8.1 Supporting Figures

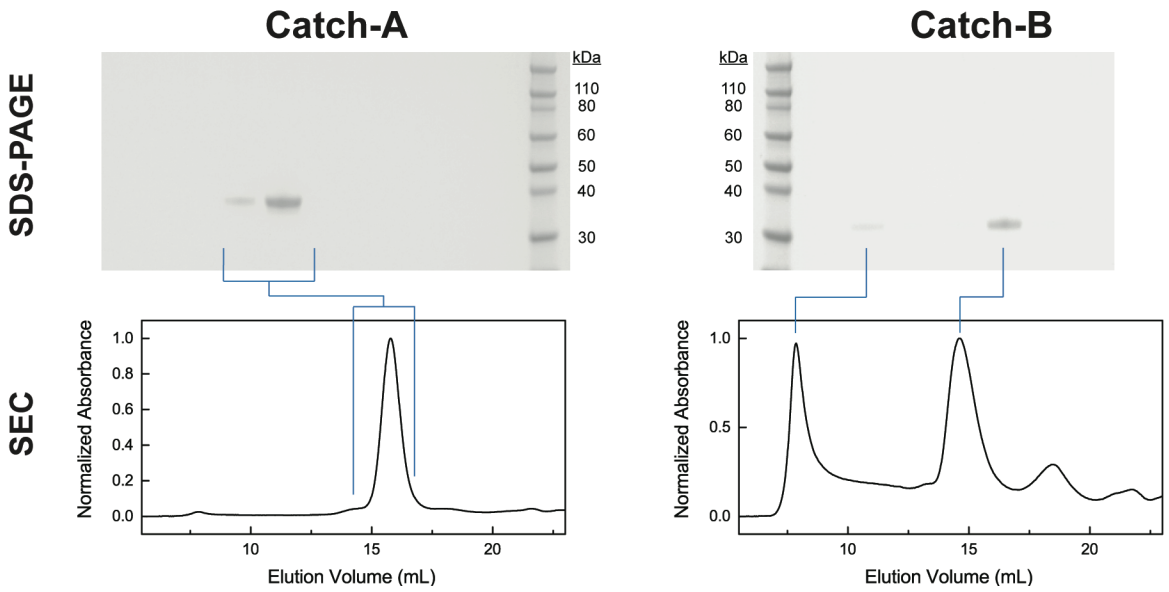

Figure S4.1: Unlabeled cage purification. SEC chromatogram (bottom) and SDS-PAGE (top) for individual components. Chromatograms have been normalized by the maximum absorbance. Corresponding fractions between the SEC chromatograms and SDS-PAGE are indicated.

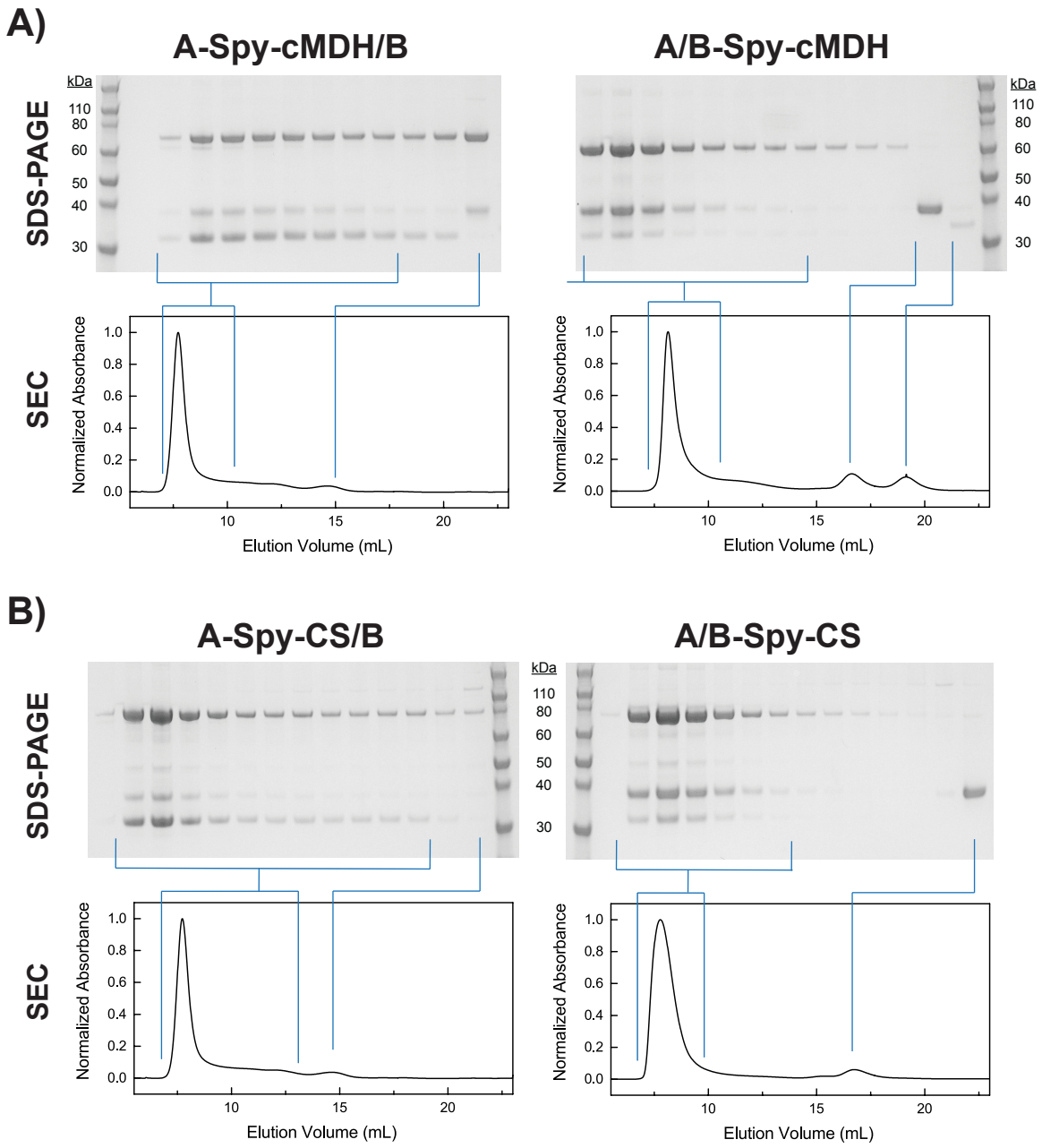

Figure S4.2: Single-enzyme complexes. Purification of cages assembled with one cage component labeled with (A) Spy-cMDH-His and (B) Spy-CS-His. Labeled and unlabeled cage components were incubated for self-assembly before purification. Corresponding fractions between the SEC chromatograms (bottom) and SDS-PAGE (top) are indicated, and chromatograms have been normalized by the maximum absorbance.

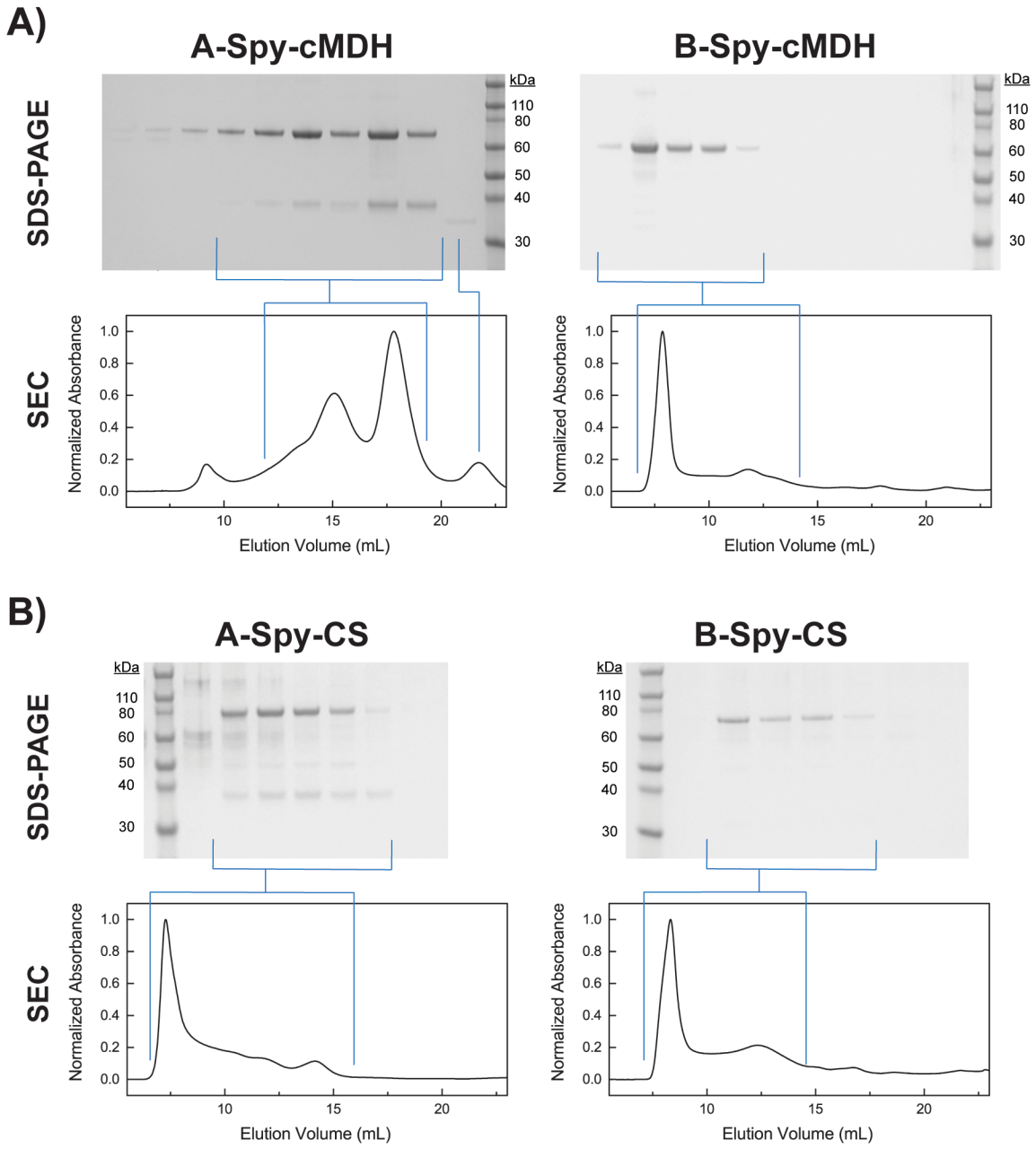

Figure S4.3: Individual labeled cage components purification. Cage components were labeled with (A) Spy-cMDH-His and (B) Spy-CS-His and purified without incubation with a second component. Corresponding fractions between the SEC chromatograms (bottom) and SDS-PAGE (top) are indicated, and chromatograms have been normalized by the maximum absorbance.

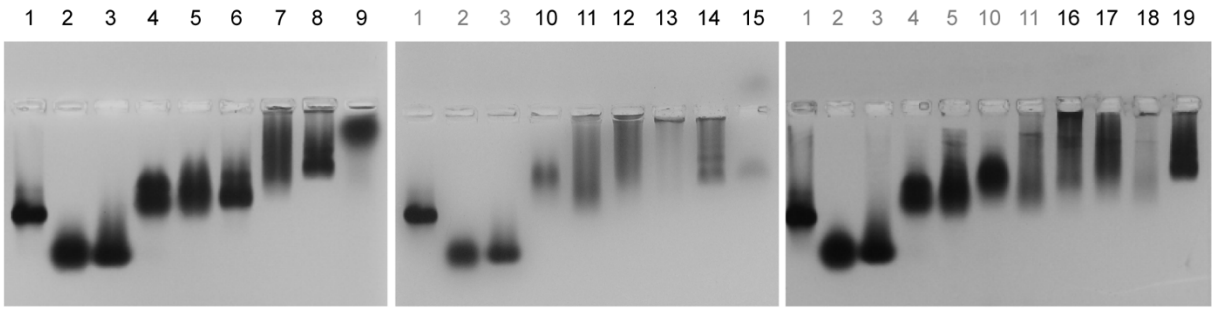

Figure S4.4: Full native agarose gels from Figure 4.10a. The first three lanes include the same samples for each gel: (1) A/B, (2) Catch-A, (3) Catch-B, (4) A-Spy-cMDH, (5) B-Spy-cMDH, (6) A-Spy-cMDH/B, (7) A/B-Spy-cMDH, (8) A/B + Spy-cMDH-His post-assembly, (9) Spy-cMDH-His, (10) A-Spy-CS, (11) B-Spy-CS, (12) A-Spy-CS/B, (13) A/B-Spy-CS, (14) A/B + Spy-CS-His post-assembly (15) Spy-CS-His, (16) A-Spy-cMDH/B-Spy-CS, (17) A-Spy-CS/B-Spy-cMDH, (18) A/B + Spy-cMDH-His and Spy-CS-His post-assembly. All repeated lanes are labeled in grey.

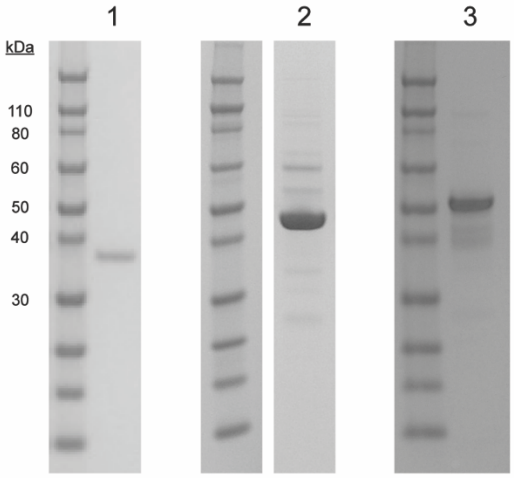

Figure S4.5: SDS-PAGE of (1) cMDH-His, (2) cMDH-3ZF-His and (3) CS-AZP4-His.

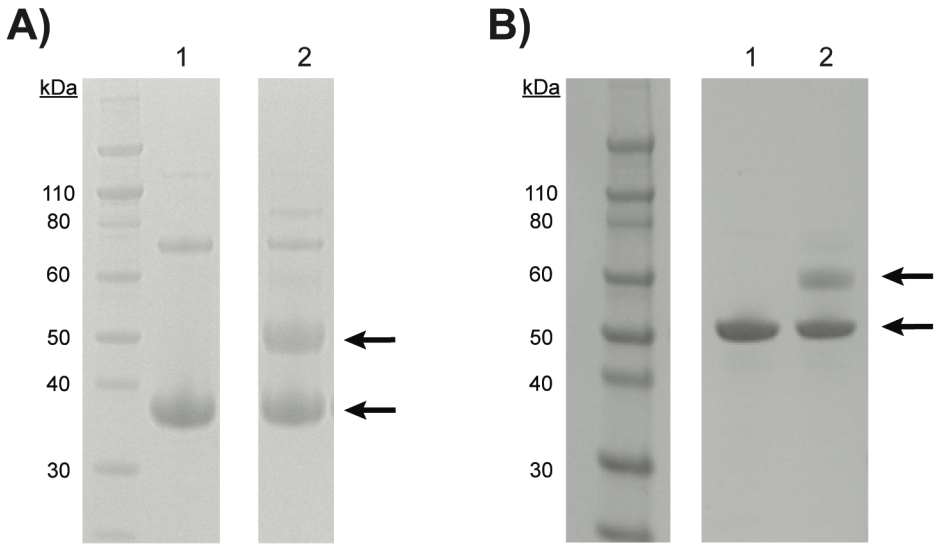

Figure S4.6 SDS-PAGE of PEG-maleimide attachment to (A) cMDH-His and (B) CS-His. For both (A) and (B): (1) enzyme only and (2) enzyme incubated with 10x molar concentration of PEG-maleimide. Lower arrows show enzyme monomers, and upper arrows show one PEG-maleimide attachment.

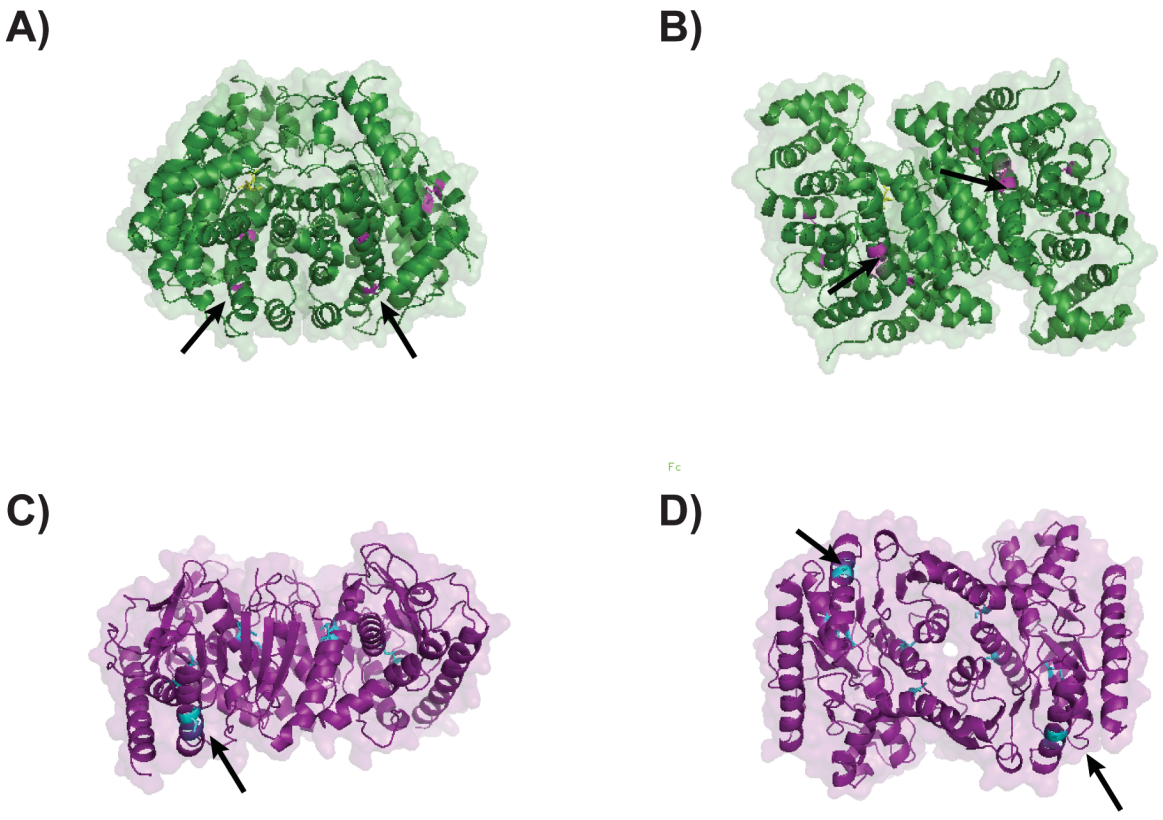

Figure S4.7 Cysteine residues in (A-B) cMDH (PDB ID: 5MDH) and (C-D) CS (PDB ID: 1CTS) dimers for two views. The five cysteine residues of cMDH are highlighted in cyan, and the four cysteine residues of CS are highlighted in magenta. Black arrows indicate Cys110 of cMDH and Cys184 of CS.

4.8.2 Supporting Tables

Table S4.1: Protein sequences for 3ZF and AZP4 and DNA target sequence.[a]

Zinc-finger protein	Zif1 -1 2 3 6	Zif2 -1 2 3 6	Zif3 -1 2 3 6	Target DNA sequence	K_D (nM)
ZIF268 (3ZF)	R D E R	R D H T	R D E R	5'-GCG TGG GCG T-3'	4
AZP4	Q N D R	R D S R	E D N T	5'-TAC GTG GCA T-3'	11

[a] All table contents are taken from Sera et al.[171]

Table S4.2: Bead assay buffer comparison for cMDH-3ZF-His.[a]

Target DNA	No additions	BSA	Salmon sperm DNA	BSA + Salmon sperm DNA
3ZF-DNA	1.0 ± 0.2	1.9 ± 0.2	1.0 ± 0.3	2.0 ± 0.3
AZP4-DNA	1.0 ± 1.0	1.3 ± 1.3	1.2 ± 0.8	2.3 ± 2.4
Control-DNA	1.0 ± 0.3	1.9 ± 1.4	1.5 ± 0.5	1.1 ± 1.0
None	1.0 ± 0.5	1.6 ± 1.3	0.6 ± 0.4	1.3 ± 1.0

[a] Binding assay with magnetic bead capture for cMDH-3ZF-His in Bead Assay Buffer with additions of 0.5% BSA and 100 µg/mL sheared salmon sperm DNA. Specific activities of the bound enzyme after washing was measures and normalized by the specific activities of samples with no buffer additions. Measurements were performed in triplicate and represented with their mean and standard deviation from the mean.

Table S4.3: DNA oligos used for PCR.[a]

Oligo name	Sequence
P1	5'- GAG CAC *CAT ATG* GAC TAC AAG GAC GAT GAT GAT AAG GG -3'
P2	5'- GAG CAC *AAG CTT* TTA GTG ATG GTG ATG ATG ATG CGC GCT GCT CAG GAA CTC GAA CGC -3'
P3	5'- ATT ATA *AAG CTT* GAT GAC GAT GAC AAA GGC GGT TCA GGC GGT GG -3'
P4	5'- TTA AAT *AAG CTT* TCA GTG GTG GTG GTG GTG GTG -3'
P7	5'- ATT ATA *CAT ATG* GAC TAT AAA GAC GAT GAC GAT AAA GGC ATG AAG GTC – 3'
P8	5'- TTA AAT *AAG CTT* GCC TTT CAT GTT CTT GAC AAA TTC TTC GCC C -3'
P9	5'- GAG CAC *CAT ATG* AGC GAA CCG ATC CGT GTG C -3'
P10	5'- GAG CAC *CAT ATG* AGC GAA CCG ATC CGT GTG C -3'
P11	5'- ATT ATA *CAT ATG* GAC TAT AAA GAC GAT GAC GAT AAA GGC ATG AGC AG -3'
P12	5'- TTA AAT *AAG CTT* TTT GCT ATC GAC CAG TTT AAT CAG ACC G TCC -3'
P13	5'- GAG CAC *CAT ATG* GGA GCC CAC ATC GTG ATG GTG GAC GCC TAC AAG CCG ACG AAG GGT AGT GGT GAA AGT GGT ATG AGC GAA CCG ATC CGT GTG C -3'
P14	5'- GAG CAC *AAG CTT* TTA GTG ATG GTG ATG ATG GTG ACC CGC -3'
P15	5'- GAG CAC *CAT ATG* GGA GCC CAC ATC GTG ATG GTG GAC GCC TAC AAG CCG ACG AAG GGT AGT GGT GAA AGT GGT ATG AGC AGC ACG AAC CTG AAA GAC ATC C -3'
P16	5'- GAG CAG *AAG CTT* TTA GTG ATG GTG ATG ATG ATG GCC TTT GC -3'

[a] Restriction sites are indicated by italics.

Table S4.4: Oligos for binding assays

Oligo name	Sequence
O1	5’- TATGGATCCTACCATGGAGCGTGGGCGTAAGCTTAT -3’
O2	Biotin- 5'- TAT GGA TCC TAC CAT GGA GCG TGG GCG TAA GCT TAT -3’
O3	5’- ATA AGC TTA CGC CCA CGC TCC ATG GTA GGA TCC ATA -3’
O4	5’- TAT GGA TCC TAC CAT GGA TAC GTG GCA TAA GCT TAT -3’
O5	Biotin- 5’- TAT GGA TCC TAC CAT GGA TAC GTG GCA TAA GCT TAT -3’
O6	5’- ATA AGC TTA TGC CAC GTA TCC ATG GTA GGA TCC ATA -3’
O7	5’- TAT GGA TCC TAC CAT GGA CCT ATG TGC TAA GCT TAT -3’
O8	Biotin- 5’- TAT GGA TCC TAC CAT GGA CCT ATG TGC TAA GCT TAT -3’
O9	5’- ATA AGC TTA GCA CAT AGG TCC ATG GTA GGA TCC ATA -3’
O10	5’- TTA CGC CCA CGC TCC -3’
O11	Amino- 5’- CCC CCC AAA AAT GGA TGG ATG GA -3’
O12	Biotin- 5’- TGC TTG GGA GCG TGG GCG TAA CAC TCA CTC ATC TCC ATC CAT CCA CTT GCT -3’
O13	Biotin- 5’- TGC TTG TCC ATC CAT CCA CAC TCA CTC ATC GGA GCG TGG GCG TAA CTT GCT -3’

Chapter 5

Summary and Future Directions

5.1 Summary

In this work, we described our efforts to engineer and characterize single and multi-enzyme complexes. Chapter 1 provided an overview of natural and engineered enzyme complexes and introduced the enzymes studied in the following chapters. We began by discussing the the importance of protein complexes for all living organisms and the types of protein-protein interactions (PPIs) that are observed. Many proteins are oligomeric, and most proteins interact in some way with other proteins. These PPIs can be divided into the classifications of permanent, weak transient and strong transient based on their stabilities and affinities. Weak transient complexes, such as multi-enzyme complexes referred to as metabolons, are difficult to study since their complex formation is highly dependent on their environment. The direct transfer of reaction intermediates in these complexes of sequential enzymes has been of great interest. This substrate channeling, in which the substrate is channeled by mechanisms such as intramolecular tunnels and electrostatic interactions without first diffusing in the bulk solution, can result in benefits such as the protection of intermediates and the creation of local substrate concentrations. Efforts have been made to engineer new protein complexes mimic nature. Much is still unknown about substrate channeling in natural and engineered systems.

We also provided an overview of the single and multi-enzyme systems discussed in the following chapters, small laccase (SLAC) and the TCA cycle enzymes mitochondrial malate dehydrogenase (mMDH), citrate synthase (CS) and aconitase (Aco). SLAC is a naturally trimeric multicopper oxidase enzyme that has been studied as an oxygen-reduction catalyst on enzymatic bio-cathodes with mediated and direct electron transfer Unlike many other laccases, SLAC is easily expressed in *E. coli*, and it has been genetically engineered previously by Wheeldon, Gallaway and Szilvay.[30-32] Six out the eight TCA cycle enzymes are thought to for a spatially

defined complex that facilitates the transport of reaction intermediates and drives the reaction in the forward direction. This TCA cycle metabolon has been proposed for decades, but confirmation with structural evidence has only been recently proposed by collaborators Wu and Minteer.[61]

In Chapter 2, we described our efforts engineering SLAC to self-assemble into functional, crystalline-like assemblies. A new dimeric interface was introduced between SLAC trimers in a symmetric design chosen due to its similarity to the crystal architecture of SLAC. Two cysteine residues (G70C and A189C) were introduced at the new interface to drive self-assembly with disulfide bonding. Upon the addition of cupric ions, the purified enzyme, SLAC-DC-His, formed blue, enzymatically active aggregates that were shown to disappear with the addition of a reducing agent. These aggregates showed increased resistance to permanent thermal denaturation, and electron microscopy of the aggregates showed architectures consistent with the computational design. Spectroscopic kinetic characterization in dilute solution showed a similar turnover number, k_{cat} as the non-mutated enzyme SLAC-His. However, the Michaelis constant, K_M increased, leading to a three-fold reduction in catalytic efficiency, k_{cat}/K_M. The aggregates were studied electrochemically combined with single-walled carbon nanotubes (SWNTs) on a rotating disk electrode with three arrangements: aggregates adsorbed on SWNT-modified electrodes, SWNTs mixed with pre-formed aggregates and aggregates formed in the presence of SWNTs. Increases in oxygen-reduction current were observed with increased incorporation of SWNTs, likely due to increased contact with the enzyme active sites, and a decrease in overpotential for aggregates formed in the presence of SWNTs indicated better enzyme orientation. The SLAC-DC-His aggregates were also found to outperform SLAC-His on an air-breathing gas-diffusion cathode for the entire potential range tested.

In Chapter 3, we characterized a recombinant TCA cycle metabolon, comparing it to the

natural metabolon isolated from intact mitochondria. The structure of in the recombinant metabolon was found to be similar to the structure of the *in vivo* metabolon, with a 30-degree rotation in mMDH around the axis perpendicular to the protein-protein interface. Electrostatic surface potentials showed a positive "channel" connecting mMDH and CS active sites along which the negatively charged reaction intermediate is proposed to channel by a bounded diffusion mechanism. Mutations from arginine to alanine and aspartic acid residues were made along this patch in CS. Most mutations were detrimental to the enzymatic activity, but one mutant, CS(R65A), was found to have similar kinetic parameters as recombinant CS for a sequential ordered mechanism. Further kinetic analyses for substrate channeling showed that the recombinant complex performed similarly to the natural metabolon. The mutant complex, however, exhibited increased transient times and decreased coupled enzymatic activities in the presence of a competing enzyme that could consume non-channeled intermediate in the bulk solution, indicating that channeling was inhibited by the R65A mutation in CS.

Chapter 4 discussed our efforts to develop a platform to study substrate channeling in engineered two-component multi-enzyme complexes. In order to develop the platform, CS and the naturally non-channeling isoform of mMDH, cytosolic were first studied for assembly, looking to extend it to study the naturally channeling pair, mMDH and CS, as well as mMDH and CS(R65A), in which channeling was found to be inhibited. In Chapter 4.3, we assembled multi-enzyme complexes on an engineered two-component icosahedral cage with a SpyTag/Spy-Catcher attachment strategy, in which the SpyTag peptide formed a stable isopeptide bond with the SpyCatcher protein upon simple mixing. The cage components are naturally trimeric and pentameric, and there are 120 enzyme-attachment sites on each assembled cage through genetic fusions to the SpyCatcher protein. Pre-assembled cage complexes and individual cage

components, which were later assembled into complexes, were labeled with the SpyTag/enzyme fusion proteins. Complex formation was observed, but it is not yet known if the components corresctly assembled into the designed cage due to the multimeric structure of each of the enzymes and cage components. Chapter 4.4 described two methods of enzyme attachment to the DNA scaffolds through zinc-finger DNA-binding proteins (ZFPs) and chemical crosslinking. To assemble the multi-enzyme complexes on DNA scaffolds, CS was attached to the DNA via cysteine residue cross-linking, and cMDH was attached with a genetic fusion to a zinc-finger DNA-binding protein that was determined to have a K_D of 1.12 ± 0.49 μM for its target DNA. All assemblies on the DNA and protein cage scaffolds retained their individual and coupled activities, but substrate channeling has yet to be observed in the complexes. Strategies to improve assembly on the DNA scaffolds and characterizing the complexes for substrate channeling were discussed.

5.2 Future Directions

While much has been accomplished in this work, there are areas that can be further explored and improved upon. In Chapter 2, we engineered crystalline-like assemblies with a single enzyme, SLAC, by introducing a new protein-protein interface. These SLAC-DC-His aggregates showed increased current densities compared to SLAC-His, and further characterization of the lifetimes of these cathodes would be beneficial. In addition, this work could potentially be expanded to assemblies of multiple enzymes for a compatible pair of sequential enzymes.

In Chapter 3, we characterized a recombinant TCA metabolon for complex formation and substrate channeling, and inhibited channeling with an R65A mutation in CS. This work was expanded in Chapter 4, in which we developed a platform to study substrate channeling in two-

enzyme assemblies on protein cage and DNA scaffolds. By using these scaffolds, we can study the coupled reactions of the three enzyme pairs: naturally channeling mMDH and CS, non-channeling cMDH and CS and channeling-inhibited mMDH and CS(R65A). In addition to the two DNA assemblies that were discussed, a variety of assemblies could be studied including those with different spacing between enzyme anchor points and additional charged proteins or peptides to interact with the reaction intermediate. A detailed kinetic analysis of each enzyme should be performed so that changes in the coupled enzyme kinetics can be accurately compared. In this way we may be able to learn more about the substrate channeling mechanism in these enzymes.

There are several ways that the assemblies in Chapter 4 can be improved and expanded. A better method of attachment to the DNA scaffold may be through the stable SpyTag/SpyCatcher system used to attach cMDH and CS to the protein cage scaffold. The SpyCatcher protein could be site-specifically attached to the oligonucleotides, and, after conjugation to the SpyTag/enzyme fusion proteins, the enzymes could be assembled on the DNA scaffolds. A library could be built with a variety of DNA assembly architectures and other two-component protein assemblies such as other designed protein cages and protein layers. Additionally, a second enzyme-attachment strategy such as the HaloTag can be used in combination with the SpyTag/SpyCatcher system. By using two stable attachment strategies, the two component scaffolds could be pre-assembled and then site-specifically labeled with enzyme pairs. In this way, we could easily use genetic fusions of enzymes to these labeling tags to study a large number of enzyme pairs on a variety of architectures.

6 References

1. Sanger, F. & Thompson, E. O. P. The amino-acid sequence in the glycyl chain of insulin. I. The identification of lower peptides from partial hydrolysates. *Biochem. J.* **53,** 353–366 (1953).

2. Perutz, M. F. Structure of hemoglobin. *Brookhaven Symp. Biol.* **13,** 165–183 (1960).

3. Peng, X., Wang, J., Peng, W., Wu, F.-X. & Pan, Y. Protein–protein interactions: detection, reliability assessment and applications. *Briefings Bioinform.*, 1-22 (2016).

4. Janin, J., Bahadur, R. P. & Chakrabarti, P. Protein–protein interaction and quaternary structure. *Quart. Rev. Biophys.* **41,** 133-180 (2008).

5. Stites, W. E. Protein−Protein Interactions: Interface Structure, Binding Thermodynamics, and Mutational Analysis. *Chem. Rev.* **97,** 1233–1250 (1997).

6. Acuner Ozbabacan, S. E., Engin, H. B., Gursoy, A. & Keskin, O. Transient protein-protein interactions. *Protein Eng., Des. Sel.* **24,** 635–648 (2011).

7. Valdar, W. S. J. & Thornton, J. M. Protein–protein interfaces: Analysis of amino acid conservation in homodimers. *Proteins: Struct., Funct., Genet.* **42,** 108–124 (2001).

8. Park, S., Reyes, J. A., Gilbert, D. R., Kim, J. & Kim, S. Prediction of protein-protein interaction types using association rule based classification. *BMC Bioinf.* **10,** 36–15 (2009).

9. Nooren, I. M. A. & Thornton, J. M. Diversity of protein–protein interactions. *EMBO J.* **22,** 3486–3492 (2003).

10. King, N. P. & Lai, Y.-T. Practical approaches to designing novel protein assemblies. *Curr. Opin. Struct. Biol.* **23,** 632–638 (2013).

11. Der, B. S. *et al.* Metal-mediated affinity and orientation specificity in a computationally designed protein homodimer. *J. Am. Chem. Soc.* **134,** 375–385 (2012).

12. Fegan, A., White, B., Carlson, J. C. T. & Wagner, C. R. Chemically controlled protein assembly: techniques and applications. *Chem. Rev.* **110,** 3315–3336 (2010).

13. Sinclair, J. C., Davies, K. M., Vénien-Bryan, C. & Noble, M. E. M. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nat. Nanotechnol.* **6,** 558–562 (2011).

14. Boyle, A. L. *et al.* Squaring the circle in peptide assembly: from fibers to discrete nanostructures by de novo design. *J. Am. Chem. Soc.* **134,** 15457–15467 (2012).

15. King, N. P. *et al.* Accurate design of co-assembling multi-component protein

nanomaterials. *Nature* **510,** 103–108 (2014).

16. Bale, J. B. *et al.* Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **353,** 389–394 (2016).

17. Stranges, P. B., Machius, M., Miley, M. J., Tripathy, A. & Kuhlman, B. Computational design of a symmetric homodimer using β-strand assembly. *Proc. Natl. Acad. Sci. U.S.A.* **108,** 20562–20567 (2011).

18. Lanci, C. J. *et al.* Computational design of a protein crystal. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 7304–7309 (2012).

19. Shleev, S. *et al.* Direct electron transfer between copper-containing proteins and electrodes. *Biosens. and Bioelectron.* **20,** 2517–2554 (2005).

20. Gallaway, J. W. & Calabrese Barton, S. A. Effect of redox polymer synthesis on the performance of a mediated laccase oxygen cathode. *J. of Electroanal. Chem.* **626,** 149–155 (2009).

21. Gupta, G. *et al.* Direct bio-electrocatalysis by multi-copper oxidases: Gas-diffusion laccase-catalyzed cathodes for biofuel cells. *Electrochim. Acta* **56,** 10767–10771 (2011).

22. Hudak, N. S., Gallaway, J. W. & Barton, S. C. Formation of mediated biocatalytic cathodes by electrodeposition of a redox polymer and laccase. *J. Electroanal. Chem.* **629,** 57–62 (2009).

23. Gallaway, J. W. & Calabrese Barton, S. A. Kinetics of redox polymer-mediated enzyme electrodes. *J. Am. Chem. Soc.* **130,** 8527–8536 (2008).

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

25. Jones, S. M. & Solomon, E. I. Electron transfer and reaction mechanism of laccases. *CMLS, Cell. Mol. Life Sci.* **72,** 869–883 (2015).

26. Skálová, T. *et al.* The Structure of the Small Laccase from Streptomyces coelicolor Reveals a Link between Laccases and Nitrite Reductases. *J. Mol. Biol.* **385,** 1165–1178 (2009).

27. Toscano, M. D., De Maria, L., Lobedanz, S. & Østergaard, L. H. Optimization of a Small Laccase by Active-Site Redesign. *Chembiochem* **14,** 1209–1211 (2013).

28. Sherif, M. *et al.* Biochemical studies of the multicopper oxidase (small laccase) from Streptomyces coelicolorusing bioactive phytochemicals and site-directed mutagenesis. *Microb. Biotechnol.* **6,** 588–597 (2013).

29. Guan, D., Kurra, Y., Liu, W. & Chen, Z. A click chemistry approach to site-specific

immobilization of a small laccase enables efficient direct electron transfer in a biocathode. *Chem. Commun.* **51,** 2522–2525 (2015).

30. Szilvay, G. R. *et al.* Engineering of a redox protein for DNA-directed assembly. *Chem. Commun.* **47,** 7464–3 (2011).

31. Wheeldon, I. R., Gallaway, J. W., Barton, S. C. & Banta, S. Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 15275–15280 (2008).

32. Gallaway, J. *et al.* Oxygen-reducing enzyme cathodes produced from SLAC, a small laccase from Streptomyces coelicolor. *Biosens. Bioelectron.* **23,** 1229–1235 (2008).

33. Velot, C., Mixon, M. B., Teige, M. & Srere, P. A. Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry* **36,** 14271–14276 (1997).

34. Zhang, Y. H. P. Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv.* **29,** 715–725 (2011).

35. Wheeldon, I. *et al.* Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8,** 299–309 (2016).

36. Hyde, C. C., Ahmed, S. A., Padlan, E. A., Miles, E. W. & Davies, D. R. Three-dimensional structure of the tryptophan synthase alpha 2 beta 2 multienzyme complex from Salmonella typhimurium. *J. Biol. Chem.* **263,** 17857–17871 (1988).

37. Perham, R. N. Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* **69,** 961–1004 (2000).

38. Elcock, A. H., Potter, M. J., Matthews, D. A., Knighton, D. R. & McCammon, J. A. Electrostatic channeling in the bifunctional enzyme dihydrofolate reductase-thymidylate synthase. *J. Mol. Biol.* **262,** 370–374 (1996).

39. Fu, J., Liu, M., Liu, Y., Woodbury, N. W. & Yan, H. Interenzyme Substrate Diffusion for an Enzyme Cascade Organized on Spatially Addressable DNA Nanostructures. *J. Am. Chem. Soc.* **134,** 5516–5519 (2012).

40. Bauler, P., Huber, G., Leyh, T. & McCammon, J. A. Channeling by Proximity: The Catalytic Advantages of Active Site Colocalization Using Brownian Dynamics. *J. Phys. Chem. Lett.* **1,** 1332–1335 (2010).

41. Eun, C., Kekenes-Huskey, P. M., Metzger, V. T. & McCammon, J. A. A model study of sequential enzyme reactions and electrostatic channeling. *J. Chem. Phys.* **140,** 105101–13 (2014).

42. Lin, J.-L. & Wheeldon, I. Kinetic Enhancements in DNA–Enzyme Nanostructures Mimic the Sabatier Principle. *ACS Catal.* **3,** 560–564 (2013).

43. Morgunov, I. & Srere, P. A. Interaction between citrate synthase and malate dehydrogenase - Substrate channeling of oxaloacetate. *J. of Biol. Chem.* **273,** 29540–29544 (1998).

44. Elcock, A. H. & McCammon, J. A. Evidence for Electrostatic Channeling in a Fusion Protein of Malate Dehydrogenase and Citrate Synthase. *Biochemistry* **35,** 12652–12658 (1996).

45. Shatalin, K., Lebreton, S., Rault-Leonardon, M., Velot, C. & Srere, P. A. Electrostatic channeling of oxaloacetate in a fusion protein of porcine citrate synthase and porcine mitochondrial malate dehydrogenase. *Biochemistry* **38,** 881–889 (1999).

46. Halper, L. A. & Srere, P. A. Interaction between citrate synthase and mitochondrial malate dehydrogenase in the presence of polyethylene glycol. *Arch. of Biochem. Biophys.* **184,** 529–534 (1977).

47. Gleason, W. B., Fu, Z., Birktoft, J. & Banaszak, L. Refined crystal structure of mitochondrial malate dehydrogenase from porcine heart and the consensus structure for dicarboxylic acid oxidoreductases. *Biochemistry* **33,** 2078–2088 (1994).

48. Silverstein, E. & Sulebele, G. Catalytic mechanism of pig heart mitochondrial malate dehydrogenase studied by kinetics at equilibrium. *Biochemistry* **8,** 2543–2550 (1969).

49. Guynn, R. W., Gelberg, H. J. & Veech, R. L. Equilibrium Constants of Malate Dehydrogenase, Citrate Synthase, Citrate Lyase, and Acetyl Coenzyme-a Hydrolysis Reactions Under Physiological Conditions. *J. Biol. Chem.* **248,** 6957–6965 (1973).

50. Gelpí, J. L., Dordal, A., Montserrat, J., Mazo, A. & Cortés, A. Kinetic-Studies of the Regulation of Mitochondrial Malate-Dehydrogenase by Citrate. *Biochem. J.* **283,** 289–297 (1992).

51. Beeckmans, S. & Kanarek, L. Demonstration of physical interactions between consecutive enzymes of the citric acid cycle and of the aspartate-malate shuttle. A study involving fumarase, malate dehydrogenase, citrate synthesis and aspartate aminotransferase. *Eur. J. Biochem.* **117,** 527–535 (1981).

52. Robinson, J. B., Inman, L., Sumegi, B. & Srere, P. A. Further Characterization of the Krebs Tricarboxylic-Acid Cycle Metabolon. *J. Biol. Chem.* **262,** 1786–1790 (1987).

53. Shatalin, K., Lebreton, S., Rault-Leonardon, M., Vélot, C. & Srere, P. A. Electrostatic Channeling of Oxaloacetate in a Fusion Protein of Porcine Citrate Synthase and Porcine Mitochondrial Malate Dehydrogenase. *Biochemistry* **38,** 881–889 (1999).

54. Adrian H Elcock, Gary A Huber, A. & McCammon, J. A. Electrostatic Channeling of Substrates between Enzyme Active Sites: Comparison of Simulation and Experiment. *Biochemistry* **36,** 16049–16058 (1997).

55. Minarik, P., Tomaskova, N., Kollarova, M. & Antalik, M. Malate dehydrogenases -

Structure and function. *Gen. Physiol. Biophys.* **21,** 257–265 (2002).

56. Johansson, C. J., Mahlen, A. & Petterson, G. Kinetic studies on citrate synthase from pig heart. *Biochim. Biophys. Acta* **309,** 466–472 (1973).

57. Wiegand, G. & Remington, S. J. Citrate Synthase - Structure, Control, and Mechanism. *Annu. Rev. Biophys. Biophys. Chem.* **15,** 97–117 (1986).

58. Remington, S., Wiegand, G. & Huber, R. Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 A resolution. *J. Mol. Biol.* **158,** 111–152 (1982).

59. Beinert, H., Kennedy, M. C. & Stout, C. D. Aconitase as Iron−Sulfur Protein, Enzyme, and Iron-Regulatory Protein. *Chem. Rev.* **96,** 2335–2374 (1996).

60. Lloyd, S. J., Lauble, H., Prasad, G. S. & Stout, C. D. The mechanism of aconitase: 1.8 Å resolution crystal structure of the S642A: citrate complex. *Protein Sci.* **8,** 2655–2662 (1999).

61. Wu, F. & Minteer, S. Krebs cycle metabolon: structural evidence of substrate channeling revealed by cross-linking and mass spectrometry. *Angew. Chem. Int. Ed.* **54,** 1851–1854 (2015).

62. Kim, J., Jia, H. & Wang, P. Challenges in biocatalysis for enzyme-based biofuel cells. *Biotechnol. Adv.* **24,** 296–308 (2006).

63. Minteer, S. D., Liaw, B. Y. & Cooney, M. J. Enzyme-based biofuel cells. *Current Opinion in Biotechnology* **18,** 228–234 (2007).

64. Lau, C. *et al.* Design of Carbon Nanotube-Based Gas-Diffusion Cathode for O2 Reduction by Multicopper Oxidases. *Adv. Energy Mater.* **2,** 162–168 (2011).

65. Addo, P. K., Arechederra, R. L. & Minteer, S. D. Evaluating Enzyme Cascades for Methanol/Air Biofuel Cells Based on NAD+-Dependent Enzymes. *Electroanalysis* **22,** 807–812 (2010).

66. Lim, J., Cirigliano, N., Wang, J. & Dunn, B. Direct electron transfer in nanostructured sol?gel electrodes containing bilirubin oxidase. *Phys. Chem. Chem. Phys.* **9,** 1809–6 (2007).

67. Holland, J. T., Lau, C., Brozik, S., Atanassov, P. & Banta, S. Engineering of Glucose Oxidase for Direct Electron Transfer via Site-Specific Gold Nanoparticle Conjugation. *J. Am. Chem. Soc.* **133,** 19262–19265 (2011).

68. Szilvay, G. R. *et al.* Engineering of a redox protein for DNA-directed assembly. *Chem. Commun.* **47,** 7464–3 (2011).

69. Wong, T. S. & Schwaneberg, U. Protein engineering in bioelectrocatalysis. *Curr. Opin.*

Biotechnol. **14,** 590–596 (2003).

70. King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510,** 103–108 (2014).

71. King, N. P. *et al.* Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **336,** 1171–1174 (2012).

72. Jiang, L. *et al.* De Novo Computational Design of Retro-Aldol Enzymes. *Science* **319,** 1387–1391 (2008).

73. Korkegian, A., Black, M. E., Baker, D. & Stoddard, B. L. Computational thermostabilization of an enzyme. *Science* **308,** 857–860 (2005).

74. Usui, K. *et al.* Nanoscale elongating control of the self-assembled protein filament with the cysteine-introduced building blocks. *Protein Sci.* **18,** 960–969 (2009).

75. Padilla, J. E., Colovos, C. & Yeates, T. O. Nanohedra: using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 2217–2221 (2001).

76. Babanova, S., Artyushkova, K., Ulyanova, Y., Singhal, S. & Atanassov, P. Design of experiments and principal component analysis as approaches for enhancing performance of gas-diffusional air-breathing bilirubin oxidase cathode. *J. Power Sources* **245,** 389–397 (2014).

77. Tsujimura, S., Kamitaka, Y. & Kano, K. Diffusion-Controlled Oxygen Reduction on Multi-Copper Oxidase-Adsorbed Carbon Aerogel Electrodes without Mediator. *Fuel Cells* **7,** 463–469 (2007).

78. Cooney, M. J., Svoboda, V., Lau, C., Martin, G. & Minteer, S. D. Enzyme catalysed biofuel cells. *Energy Environ. Sci.* **1,** 320–18 (2008).

79. Luckarift, H. R., Atanassov, P. & Johnson, G. R. *Enzymatic Fuel Cells: From Fundamentals to Applications*. *Enzymatic Fuel Cells: From Fundamentals to Applications.* (John Wiley & Sons, Inc., 2014).

80. Calabrese Barton, S., Gallaway, J. & Atanassov, P. Enzymatic Biofuel Cells for Implantable and Microscale Devices. *Chem. Rev.* **104,** 4867–4886 (2004).

81. Logan, B. E. *et al.* Microbial fuel cells: Methodology and technology. *Environ. Sci. Technol.* **40,** 5181–5192 (2006).

82. Glykys, D. J. *et al.* Pushing the limits of automatic computational protein design: design, expression, and characterization of a large synthetic protein based on a fungal laccase scaffold. *Syst. Synth. Biol.* **5,** 45–58 (2011).

83. Machczynski, M. C., Vijgenboom, E., Samyn, B. & Canters, G. W. Characterization of

SLAC: A small laccase from Streptomyces coelicolorwith unprecedented activity. *Protein Sci.* **13,** 2388–2397 (2004).

84. Gallaway, J. *et al.* Oxygen-reducing enzyme cathodes produced from SLAC, a small laccase from Streptomyces coelicolor. *Biosens. Bioelectron.* **23,** 1229–1235 (2008).

85. Wheeldon, I. R., Gallaway, J. W., Barton, S. C. & Banta, S. Bioelectrocatalytic hydrogels from electron-conducting metallopolypeptides coassembled with bifunctional enzymatic building blocks. *Proc. Natl. Acad. Sci. U.S.A.* **105,** 15275–15280 (2008).

86. Bard, A. & Faulkner, K. M. *Electrochemical methods - Fundamentals and Applications*. (John Wiley & Sons, Inc., 2001).

87. Ramírez, P. *et al.* Direct electron transfer from graphite and functionalized gold electrodes to T1 and T2/T3 copper centers of bilirubin oxidase. *Biochim. Biophys. Acta, Bioenerg.* **1777,** 1364–1369 (2008).

88. Gupta, G. *et al.* Direct bio-electrocatalysis by multi-copper oxidases: Gas-diffusion laccase-catalyzed cathodes for biofuel cells. *Electrochim. Acta* **56,** 10767–10771 (2011).

89. Gupta, G. *et al.* Direct electron transfer catalyzed by bilirubin oxidase for air breathing gas-diffusion electrodes. *Electrochem. Commun.* **13,** 247–249 (2011).

90. Ulyanova, Y. *et al.* Effect of enzymatic orientation through the use of syringaldazine molecules on multiple multi-copper oxidase enzymes. *Phys. Chem. Chem. Phys.* **16,** 13367–9 (2014).

91. Schoffelen, S. & van Hest, J. C. M. Multi-enzyme systems: Bringing enzymes together in vitro. *Soft Matter* **8,** 1736–1746 (2012).

92. Ovádi, J. & Sreret, P. A. Macromolecular compartmentation and channeling. *Int. Rev. Cytol.* (1999).

93. Srere, P. A. & Mosbach, K. Metabolic compartmentation: symbiotic, organellar, multienzymic, and microenvironmental. *Annu. Rev. of Microbiol.* **28,** 61–83 (1974).

94. Srere, P. A. The metabolon. *Trends Biochem. Sci.* **10,** 109–110 (1985).

95. Ovádi, J. Physiological Significance of Metabolic Channeling. *J. Theor. Biol.* **152,** 1–22 (1991).

96. Spivey, H. O. & Ovádi, J. Substrate channeling. *Methods* **19,** 306–321 (1999).

97. Wheeldon, I. *et al.* Substrate channelling as an approach to cascade reactions. *Nat. Chem.* **8,** 299–309 (2016).

98. Morgunov, I. & Srere, P. A. Interaction between citrate synthase and malate dehydrogenase. Substrate channeling of oxaloacetate. *J. Biol. Chem.* **273,** 29540–29544

(1998).

99. Huang, X., Holden, H. M. & Raushel, F. M. Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annu. Rev. Biochem.* **70,** 149–180 (2001).

100. Conrado, R. J., Varner, J. D. & DeLisa, M. P. Engineering the spatial organization of metabolic enzymes: mimicking nature's synergy. *Curr. Opin. Biotechnol.* **19,** 492–499 (2008).

101. Srere, P. A. Complexes of sequential metabolic enzymes. *Annu. Rev. Biochem.* **56,** 89–124 (1987).

102. Moore, G. E., Gadol, S. M., Robinson, J. B. & Srere, P. A. Binding of citrate synthase and malate dehydrogenase to mitochondrial inner membranes: tissue distribution and metabolite effects. *Biochem. Biophys. Res. Commun.* **121,** 612–618 (1984).

103. D'Souza, S. F. & Srere, P. A. Cross-linking of mitochondrial matrix proteins in situ. *Biochim. Biophys. Acta.* **724,** 40–51 (1983).

104. Robinson, J. B. & Srere, P. A. Organization of Krebs tricarboxylic acid cycle enzymes in mitochondria. *J. Biol. Chem.* **260,** 10800–10805 (1985).

105. Velot, C., Mixon, M. B., Teige, M. & Srere, P. A. Model of a quinary structure between Krebs TCA cycle enzymes: a model for the metabolon. *Biochemistry* **36,** 14271–14276 (1997).

106. Fernie, A. R., Carrari, F. & Sweetlove, L. J. Respiratory metabolism: glycolysis, the TCA cycle and mitochondrial electron transport. *Curr. Opin. Plant Biol.* **7,** 254–261 (2004).

107. Lyubarev, A. E. & Kurganov, B. I. Supramolecular organization of tricarboxylic acid cycle enzymes. *BioSystems* **22,** 91–102 (1989).

108. Lindbladh, C. *et al.* Preparation and Kinetic Characterization of a Fusion Protein of Yeast Mitochondrial Citrate Synthase and Malate-Dehydrogenase. *Biochemistry* **33,** 11692–11698 (1994).

109. Srere, P. A., Mattiasson, B. & Mosbach, K. An immobilized three-enzyme system: a model for microenvironmental compartmentation in mitochondria. *Proc. Natl. Acad. Sci.* **70,** 2534–2538 (1973).

110. Miles, E. W., Rhee, S. & Davies, D. R. The Molecular Basis of Substrate Channeling. *J. Biol. Chem.* **274,** 12193–12196 (1999).

111. Raushel, F. M., Thoden, J. B. & Holden, H. M. Enzymes with molecular tunnels. *Acc. Chem. Res.* **36,** 539–548 (2003).

112. Wu, F. & Minteer, S. Krebs cycle metabolon: structural evidence of substrate channeling

revealed by cross-linking and mass spectrometry. *Angew. Chem. Int. Ed.* **54,** 1851–1854 (2015).

113. Minarik, P., Tomaskova, N., Kollarova, M. & Antalik, M. Malate dehydrogenases-structure and function. *Gen. Physiol. Biophys.* **21,** 257–265 (2002).

114. Wiegand, G. & Remington, S. J. Citrate synthase: structure, control, and mechanism. *Annu. Rev. Biophys. Biophys. Chem.* **15,** 97–117 (1986).

115. Jin Lee, Y. Mass spectrometric analysis of cross-linking sites for the structure of proteins and protein complexes. *Mol. BioSyst.* **4,** 816–823 (2008).

116. Rappsilber, J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J. Struct. Biol.* **173,** 530–540 (2011).

117. Haggie, P. M. & Verkman, A. S. Diffusion of tricarboxylic acid cycle enzymes in the mitochondrial matrix in vivo. Evidence for restricted mobility of a multienzyme complex. *J. Biol. Chem.* **277,** 40782–40788 (2002).

118. Saks, V., Beraud, N. & Wallimann, T. Metabolic compartmentation - A system level property of muscle cells: Real problems of diffusion in living cells. *Int. J. Mol. Sci.* **9,** 751–767 (2008).

119. Senske, M. *et al.* Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc.* **136,** 9036–9041 (2014).

120. Zhou, H.-X., Rivas, G. & Minton, A. P. Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **37,** 375–397 (2008).

121. Evans, C. T., Kurz, L. C., Remington, S. J. & Srere, P. A. Active site mutants of pig citrate synthase: effects of mutations on the enzyme catalytic and structural properties. *Biochemistry* **35,** 10661–10672 (1996).

122. Evans, C. T., Owens, D. D., Slaughter, C. A. & Srere, P. A. Characterization of Mutant Tmk368k Pig Citrate Synthase Expressed in and Isolated From Escherichia-Coli. *Biochem. Biophys. Res. Commun.* **157,** 1231–1238 (1988).

123. Linda C Kurz *et al.* Effects of Changes in Three Catalytic Residues on the Relative Stabilities of Some of the Intermediates and Transition States in the Citrate Synthase Reaction. *Biochemistry* **37,** 9724–9737 (1998).

124. Alter, G. M. *et al.* Mutation of essential catalytic residues in pig citrate synthase. *Biochemistry* **29,** 7557–7563 (1990).

125. Lin, J.-L., Palomec, L. & Wheeldon, I. Design and Analysis of Enhanced Catalysis in Scaffolded Multienzyme Cascade Reactions. *ACS Catal.* **4,** 505–511 (2014).

126. Ovádi, J. *et al.* Transient-time analysis of substrate-channelling in interacting enzyme systems. *Biochem. J.* **257,** 187–190 (1989).

127. Bard, A. J. & Faulkner, L. R. *Electrochemical methods: fundamentals and applications*. (John Wiley & Sons, Inc, 2001).

128. Easterby, J. S. A generalized theory of the transition time for sequential enzyme reactions. *Biochem. J.* **199,** 155–161 (1981).

129. Fell, D. A. Metabolic control analysis: a survey of its theoretical and experimental development. *Biochem. J.* **286,** 313–330 (1992).

130. Kell, D. B. & Westerhoff, H. V. Metabolic control theory: its role in microbiology and biotechnology. *FEMS Microbiol. Lett.* **39,** 305–320 (1986).

131. Martin, G. L., Minteer, S. D. & Cooney, M. J. Spatial Distribution of Malate Dehydrogenase in Chitosan Scaffolds. *ACS Appl. Mater. Interfaces* **1,** 367–372 (2009).

132. Rogers, G. W. *et al.* High Throughput Microplate Respiratory Measurements Using Minimal Quantities Of Isolated Mitochondria. *PLoS ONE* **6,** e21746 (2011).

133. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: a fully automated algorithm for protein-protein docking. *Nucleic Acids Res.* **32,** W96–9 (2004).

134. Comeau, S. R., Gatchell, D. W., Vajda, S. & Camacho, C. J. ClusPro: an automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics* **20,** 45–50 (2004).

135. Kozakov, D. *et al.* How good is automated protein docking? *Proteins: Struct., Funct., Bioinf.* **81,** 2159–2166 (2013).

136. Kahraman, A., Malmström, L. & Aebersold, R. Xwalk: computing and visualizing distances in cross-linking experiments. *Bioinformatics* **27,** 2163–2164 (2011).

137. Herzog, F. *et al.* Structural Probing of a Protein Phosphatase 2A Network by Chemical Cross-Linking and Mass Spectrometry. *Science* **337,** 1348–1352 (2012).

138. Chaudhury, S. *et al.* Benchmarking and Analysis of Protein Docking Performance in Rosetta v3.2. *PLoS ONE* **6,** e22477 (2011).

139. Lyskov, S. *et al.* Serverification of Molecular Modeling Applications: The Rosetta Online Server That Includes Everyone (ROSIE). *PLoS ONE* **8,** e63906 (2013).

140. Lyskov, S. & Gray, J. J. The RosettaDock server for local protein-protein docking. *Nucleic Acids Res.* **36,** W233–8 (2008).

141. Dolinsky, T. J. *et al.* PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35,** W522–5

(2007).

142. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32,** W665–7 (2004).

143. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J. Chem. Theory Comput.* **7,** 525–537 (2011).

144. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci.* **98,** 10037–10041 (2001).

145. Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M. & Jensen, J. H. Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pKa Values. *J. Chem. Theory Comput.* **7,** 2284–2295 (2011).

146. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402 (1997).

147. Idan, O. & Hess, H. Engineering enzymatic cascades on nanoscale scaffolds. *Curr. Opin. Biotechnol.* **24,** 606–611 (2013).

148. Dunn, M. F. Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. *Arch. Biochem. Biophys.* **519,** 154–166 (2012).

149. Haggie, P. M. & Verkman, A. S. Diffusion of Tricarboxylic Acid Cycle Enzymes in the Mitochondrial Matrix in Vivo: Evidence for Restricted Mobility of Multienzyme Complex. *J. Biol. Chem.* **277,** 40782–40788 (2002).

150. Campanella, M. E., Chu, H. Y. & Low, P. S. Assembly and regulation of a glycolytic enzyme complex on the human erythrocyte membrane. *Proc. Natl. Acad. Sci.* **102,** 2402–2407 (2005).

151. Suss, K. H., Arkona, C., Manteuffel, R. & Adler, K. Calvin Cycle Multienzyme Complexes Are Bound to Chloroplast Thylakoid Membranes of Higher-Plants Insitu. *Proc. Natl. Acad. Sci.* **90,** 5514–5518 (1993).

152. Winkel-Shirley, B. Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. *Physiol. Plant.* **107,** 142–149 (1999).

153. Miles, E. W., Rhee, S. & Davies, D. R. The molecular basis of substrate channeling. *J. Biol. Chem.* **274,** 12193–12196 (1999).

154. Bulutoglu, B., Garcia, K. E., Wu, F., Minteer, S. D. & Banta, S. Direct Evidence for Metabolon Formation and Substrate Channeling in Recombinant TCA Cycle Enzymes. *ACS Chem. Biol.* **11,** 2847-2853 (2016).

155. Dueber, J. E. *et al.* Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* **27,** 753–759 (2009).

156. Conrado, R. J. *et al.* DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acids Research* **40,** 1879–1889 (2012).

157. Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333,** 470–474 (2011).

158. Fu, J. *et al.* Multi-enzyme complexes on DNA scaffolds capable of substrate channelling with an artificial swinging arm. *Nat. Nanotechnol.* **9,** 531–536 (2014).

159. Van Nguyen, K., Giroud, F. & Minteer, S. D. Improved Bioelectrocatalytic Oxidation of Sucrose in a Biofuel Cell with an Enzyme Cascade Assembled on a DNA Scaffold. *J. Electrochem. Soc.* **161,** H930–H933 (2014).

160. Xin, L., Zhou, C., Yang, Z. & Liu, D. Regulation of an enzyme cascade reaction by a DNA machine. *Small* **9,** 3088–3091 (2013).

161. Sun, Q., Madan, B., Tsai, S.-L., DeLisa, M. P. & Chen, W. Creation of artificial cellulosomes on DNA scaffolds by zinc finger protein-guided assembly for efficient cellulose hydrolysis. *Chem. Commun.* **50,** 1423–1425 (2014).

162. Liu, F., Banta, S. & Chen, W. Functional assembly of a multi-enzyme methanol oxidation cascade on a surface-displayed trifunctional scaffold for enhanced NADH production. *Chem. Commun.* **49,** 3766 (2013).

163. You, C., Myung, S. & Zhang, Y. H. P. Facilitated Substrate Channeling in a Self-Assembled Trifunctional Enzyme Complex. *Angew. Chem. Int. Ed.* **51,** 8787–8790 (2012).

164. You, C. & Zhang, Y. H. P. Self-Assembly of Synthetic Metabolons through Synthetic Protein Scaffolds: One-Step Purification, Co-immobilization, and Substrate Channeling. *ACS Synth. Biol.* **2,** 102–110 (2013).

165. You, C. *et al.* Enzymatic transformation of nonfood biomass to starch. *Proc. Natl. Acad. of Sci.* **110,** 7182–7187 (2013).

166. Wilner, O. I., Shimron, S., Weizmann, Y., Wang, Z.-G. & Willner, I. Self-Assembly of Enzymes on DNA Scaffolds: En Route to Biocatalytic Cascades and the Synthesis of Metallic Nanowires. *Nano Lett.* **9,** 2040–2043 (2009).

167. Idan, O. & Hess, H. Origins of Activity Enhancement in Enzyme Cascades on Scaffolds. *ACS Nano* **7,** 8658–8665 (2013).

168. Idan, O. & Hess, H. Diffusive transport phenomena in artificial enzyme cascades on scaffolds. *Nat. Nanotechnol.* **7,** 769–770 (2012).

169. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. of Sci.* **109,** E690–E697 (2012).

170. Li, L., Fierer, J. O., Rapoport, T. A. & Howarth, M. Structural Analysis and Optimization of the Covalent Association between SpyCatcher and a Peptide Tag. *J. Mol. Biol.* **426,** 309–317 (2014).

171. Sera, T. & Uranga, C. Rational Design of Artificial Zinc-Finger Proteins Using a Nondegenerate Recognition Code Table. *Biochemistry* **41,** 7074–7081 (2002).

172. Elrod-Erickson, M. & Pabo, C. O. Binding Studies with Mutants of Zif268: Contributions of Individual Side Chains to Binding Affinity and Specificity in the in the Zif268 Zinc Finger DNA Complex. *J. Biol. Chem.* **274,** 19281–19285 (1999).

173. Greisman, H. A. & Pabo, C. O. A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* **275,** 657–661 (1997).

174. Pomerantz, J. L., Wolfe, S. A. & Pabo, C. O. Structure-based design of a dimeric zinc finger protein. *Biochemistry* **37,** 965–970 (1998).

175. Elrod-Erickson, M., Benson, T. E. & Pabo, C. O. High-resolution structures of variant Zif268-DNA complexes: implications for understanding zinc finger DNA recognition. *Structure* **6,** 451–464 (1998).

176. ElrodErickson, M., Rould, M. A., Nekludova, L. & Pabo, C. O. Zif268 protein-DNA complex refined at 1.6 angstrom: A model system for understanding zinc finger-DNA interactions. *Structure* **4,** 1171–1180 (1996).

177. Nakata, E. *et al.* Zinc-Finger Proteins for Site-Specific Protein Positioning on DNA-Origami Structures. *Angew. Chem. Int. Ed.* **51,** 2421–2424 (2012).

178. Fu, J., Liu, M., Liu, Y. & Yan, H. Spatially-Interactive Biomolecular Networks Organized by Nucleic Acid Nanostructures. *Acc. Chem. Res.* **45,** 1215–1226 (2012).

179. Niemeyer, C. M. Semisynthetic DNA-protein conjugates for biosensing and nanofabrication. *Angew. Chem. Int. Ed.* **49,** 1200–1216 (2010).

180. Tanrikulu, I. C., Schmitt, E., Mechulam, Y., Goddard, W. A. & Tirrell, D. A. Discovery of Escherichia coli methionyl-tRNA synthetase mutants for efficient labeling of proteins with azidonorleucine in vivo. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 15285–15290 (2009).

181. Chalker, J. M., Bernardes, G. J. L., Lin, Y. A. & Davis, B. G. Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. - Asian J.* **4,** 630–640 (2009).