ABSTRACT

GARCIA, ENRIQUE. A Study of Amino Acid Linkers and their Effects on Activity and Structural Changes of $OleT_{JE}$-Superoxide Reductase Protein Fusions. (Under the direction of Dr. Amy Grunden and Dr. Robert Rose).

Interest in developing effective biofuel production methods has increased as evidence of climate change impacts that are in part resulting from the use of fossil fuels have gained global attention. One biofuel production process that has shown promise is the biological conversion of fatty acids to terminal alkenes that can be used as a drop-in replacement for transportation fuel. $OleT_{JE}$, a cytochrome P450 fatty acid peroxygenase produced by the bacterium *Jeotgallicoccus* sp., is an enzyme that can catalyze the conversion of fatty acid chains into terminal alkenes, a biofuel precursor, using hydrogen peroxide as a reaction mediator. While low concentrations of H_2O_2 are necessary for $OleT_{JE}$ activity, exogenous addition of H_2O_2 can result in irreversible heme degradation and enzyme inactivation. To help improve $OleT_{JE}$ efficiency and stability in a recombinant expression system, we sought to characterize an enzyme fusion between $OleT_{JE}$ and an enzyme that generates the reaction mediator H_2O_2. Superoxide reductase (SOR) is an antioxidant enzyme produced by the hyperthermophilic archaeon *Pyrococcus furiosus* that has been shown to mitigate oxidative stress in recombinant plants and bacteria. SOR catalyzes conversion of the superoxide anion into hydrogen peroxide and could provide the necessary hydrogen peroxide to support $OleT_{JE}$ activity. Preliminary studies indicated that when the fusion was expressed with $OleT_{JE}$ on the N-terminal end and SOR on the C-terminal end (OleT-SOR) that $OleT_{JE}$ had comparable activity to the unfused enzyme; however, the SOR-$OleT_{JE}$ fusion variant has very low activity, likely because of improper protein folding or steric hindrance. Amino acids linkers have been shown to improve fusion protein folding and activity in some cases. Therefore, we also sought to evaluate whether the addition of linker regions between the

$OleT_{JE}$ and SOR enzymes could improve $OleT_{JE}$ activity, and we used analytical ultracentrifugation and Small Angle X-ray Scattering (SAXS) techniques to examine OleT-SOR structural interactions.

Characterization of $OleT_{JE}$-SOR Fusion Enzymes for Use in Biofuel Production

by
Enrique Garcia

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Biochemistry

Raleigh, North Carolina
2021

APPROVED BY:

Dr. Amy Grunden Committee Chair	Dr. Robert Rose Co-Chair of Advisory Committee
Dr. Michael Goshe	

DEDICATION

I dedicate this work to my family members and friends who have helped me on this journey. Jose and Oralia who are my hardworking parents and taught me to persevere, my best friend Crystal for always being supportive and my siblings for pushing me to succeed. Lastly, I dedicate this to the Grunden Lab members, specifically Hannah Wapshott-Stehli and Dr. Amy Grunden for all their help.

BIOGRAPHY

My name is Enrique Garcia, and I am a first-generation college student. My life began in the rough neighborhoods of the eastern part of Los Angeles, California after my parents came to this country seeking work and a better life. That job hunt led them to a small town in North Carolina called Salisbury. Growing up in Salisbury was a great experience, it was a quiet place, and I had the chance to make friends. When I was only five, I met the girl that would have the most influence in my life and to this day she is my best friend. We went through high school and undergrad together at Catawba College. After obtaining my BS in biology I worked briefly at the LabCorp Regional Lab based in Burlington. Seeking a new challenge of my own, I applied to NC State for a biochemistry program called Interdisciplinary Biochemistry Master's Program (IBMP). Acceptance into this program meant that I had to make major move of my own to a big city. Raleigh has since been great to me, I have gotten the chance to meet great people from all over the world as well as be part of an amazing lab group. I owe a lot to the Grunden Lab because it has been a very diverse group of people that understand the complexities of being a graduate student.

ACKNOWLEDGMENTS

I would like to acknowledge Dr. Amy Grunden for her exceptional leadership and PhD candidate Hannah Wapshott-Stehli for all of her help and guidance throughout this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: A Study of Amino Acid Linkers and their Effects on Activity and Structural Changes of $OleT_{JE}$-Superoxide Reductase Protein Fusions.

1.1 Introduction

Finding a cost-effective way to produce biofuels has been an ongoing challenge that researchers have attempted to address in recent years. Several different types of biofuel production platforms have been developed and evaluated over the years (Nozzi et al., 2013). One such type of biofuel is known as first-generation biofuels which are characterized by their ability to be blended with petroleum-based fuels and that can combust in an existing internal combustion engine and redistributed through existing fuel infrastructure (Naik et al., 2009). The first-generation biofuel ethanol is produced from the fermentation of starch-based biomass such as corn, potatoes, sugarcane, and sugar beets. First generation biofuels are usually unprocessed and are typically used for heating, cooking or electricity production, while second generation biofuels are produced by processing and fermenting lignocellulosic biomass such as grasses, straws, and wood to produce alcohol-based fuels such as ethanol and butanol that can be used in light vehicles and various industrial processes (Nigam et al., 2009). Third generation biofuels are synthesized from lipid-producing biomass such as microalgae or macroalgae and oilseed crops. Third generation biofuels such as refined algal oils and biodiesel, and aviation fuels, unlike first- and second-generation fuels, are considered to be drop-in fuels that can be used to fuel an array of vehicles from gasoline-powered light-weight automobiles to diesel powered heavy-duty trucks and even supersonic jets (Nanda et al., 2018). **Figure 1-1** shows the difference in biomass needed for the three different generations of biofuel.

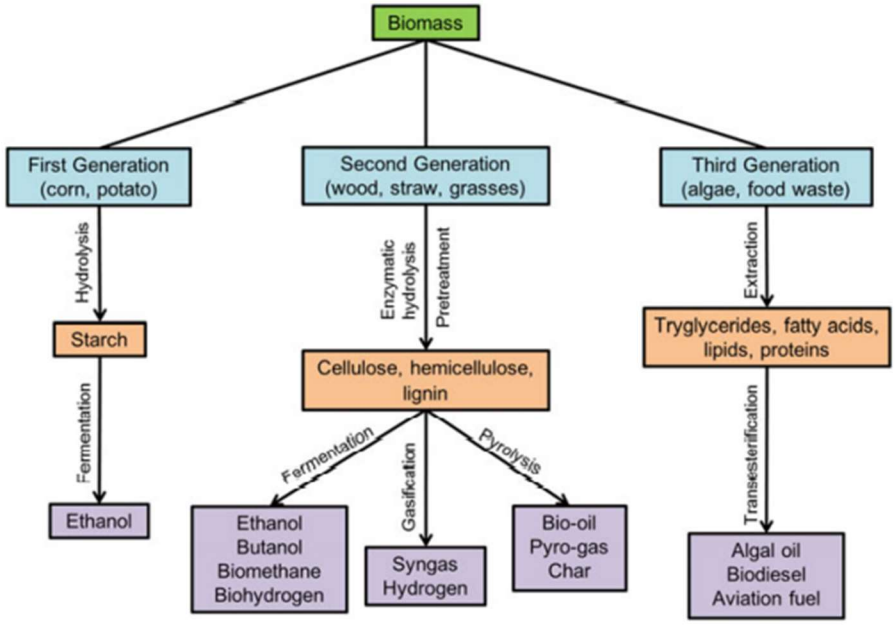


Figure 1-1. General overview of the basic steps in the production of first, second and third generation biofuels (from Nanda et al., 2018).

Since renewable energy is an alternative solution for growing energy demands, we must find a way to develop scalable, economically feasible renewable biofuel production platforms. One example of this is using algae as a biofuel production platform to produce bioethanol, biodiesel, and aviation fuels (Mannan et al., 2017). Another popular third generation biofuel feedstock are cyanobacteria which like algae can grow quickly, and fix carbon dioxide gas, but are typically much more genetically tractable than algae. Furthermore, cyanobacteria do not require fermentable sugars and do not require arable land to grow (Nozzi et.al 2013).

Both algae and cyanobacteria can serve as good sources of fatty acids for thermochemical conversion to alkanes and alkenes, which then undergo chemical cracking to produce different grades of gasolines (Nanda et al., 2018). Recent findings have revealed that enzymes such as $OleT_{JE}$, from bacterium *Jeotgalicoccus* subfamily CYP152 peroxygenase, can generate terminal alkenes from free fatty acids, which can replace the traditional expensive, metal-catalyst

requiring thermochemical conversion process (Rude et al., 2011). Alkenes are a particularly desirable biofuel precursor for drop-in transportation fuels because they can be used to produce gasolines as well as aviation fuels, and they can be generated from biomass using enzyme technologies. The most common biofuel conversion technology is thermochemical conversion. The advantage of thermochemical conversion is its high reaction efficiencies and compatibility with existing infrastructure facilities. The primary drawback to this type of technology is that it is relatively expensive, requiring costly metal catalysts (e.g. gold and platinum) that become fouled with use and must be replaced (Verma et al., 2011).

In our lab, we have shown that $OleT_{JE}$ can generate terminal alkenes from fatty acid substrates in the presence of hydrogen peroxide. Creating an enzyme system where there is a steady, low level of hydrogen peroxide provision to the peroxygenase is ideal for a biofuel production system that could remain active for extended periods of time. The way we chose to approach this aim was to fuse $OleT_{JE}$ with an enzyme that produces hydrogen peroxide from oxidative stress such as superoxide reductase (SOR) (Grunden et al., 2005). Linking $OleT_{JE}$ with SOR using different amino acid linkers that each have distinct structural characteristics should influence the interaction between the two proteins. (Chen et al., 2013). Ideally, the addition of the amino acid linkers would increase substrate conversion and produce terminal olefins that can be used for biofuel production. Changes in the structural interaction between $OleT_{JE}$ and SOR in the fusions can be evaluated using methods such as sedimentation velocity and small angle x-ray scattering (SAXS) analysis of the purified proteins.

1.2 Enzymes that can drive decarboxylation of fatty acids

Enzymes that can decarboxylate free fatty acids are important to investigate because they could potentially give us the solution towards making biofuels more cost effective and limiting our dependence on fossil fuels. UndA is a non-heme iron enzyme produced by *Pseudomonas fluorescens* that can generate 1-alkenes from fatty acids chains (C10-C14) using the reaction shown in **Figure 1-2** (Zhang et al., 2019).

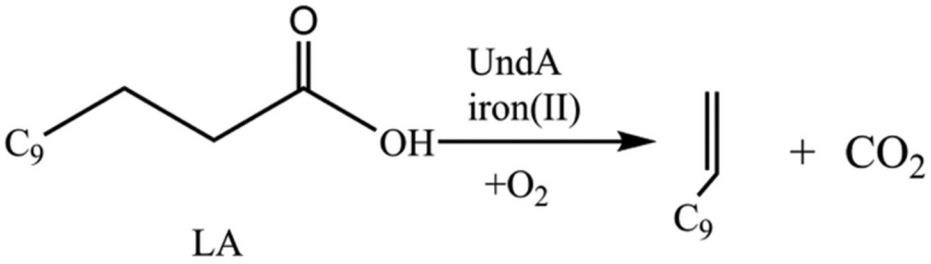


Figure 1-2. Decarboxylation reaction of lauric acid (LA) catalyzed by UndA (from Zhang et al., 2019).

Because of the electron imbalance during the oxidative decarboxylation of the substrate and the reduction of O_2, only single turnover reactions have been observed for UndA *in vitro* assays. Furthermore, the catalytic efficiency of UndA is quite low and only trace amounts of 1-alkenes are produced (Zhang et al., 2019). Another method of fatty acid decarboxylation is light-driven decarboxylation. This is an alternative approach for the synthesis of biodiesel that uses a photodecarboxylase from the photosynthetic bacterium *Chlorella variabilis NC64A (CvFAP)* that irreversibly makes alkanes from fatty acids and triglycerides as seen in **Figure 1-3** (Zhang et al., 2019).

Figure 1-3. *Cv*FAP-catalyzed photobiocatalytic decarboxylation reaction (from Zhang et al. 2019)

*Cv*FAP is a flavoenzyme that requires blue light (450 nm) to activate the catalytic FAD in the active site (Zhang et al., 2018). Admittedly newer studies have revealed that CvFAP has modest turnover numbers compared to other decarboxylation enzymes. However, results suggest that CvFAP exhibits a clear preference for long-chain fatty acids, thereby limiting its broad applicability (Zhang et al. 2019).

Cytochrome P450 enzymes are peroxygenases, meaning that they use hydrogen peroxide as a cofactor. CYP152 peroxygenases such as CYP-Aa162 and CYP-Sm46Δ29 also have been demonstrated to decarboxylate and hydroxylate free fatty acids. Gas Chromatography-Mass Spectrometry (GC-MS) analysis of catalytic activities revealed that CYP-Sm46Δ29 and $OleT_{JE}$ had similar conversion percentages of myristic acid (C14) at 73.2% and 74.2%, respectively but differed greatly in terms of product distribution. CYP-Sm46Δ29 only had 78.4% of the product generated as 1-tridecene, while $OleT_{JE}$ had 91.8% 1-tridecene production. CYP-Sm46Δ29 showed considerable β-hydroxylation at 33.2 % of the total product generated, while CYP-Aa162 had 78.9% of its total products as the α-hydroxylated product and only 16% as 1-tridecene (Xu et al., 2017). Even though there are many enzymes capable of decarboxylating fatty acids, $OleT_{JE}$ still seems to be the most promising fatty acid decarboxylation enzyme since

it forms alpha-alkenes that are useful for biofuel production as well as for making lubricants, polymers, and detergents. (Liu et al., 2014)

1.2.1 OleT$_{JE}$ catalyzed fatty acid decarboxylation.

OleT$_{JE}$ is a P450 enzyme that belongs to the CYP152 family of peroxygenases. The ability for OleT$_{JE}$ to decarboxylate biological fatty acids has brought attention to it because it can form terminal olefins (1-alkenes). In fact, terminal alkenes are the major product of OleT$_{JE}$ reacting with fatty acids, which sets it apart from other known fatty acid peroxygenases (Rude et al., 2011). OleT$_{JE}$ uses mainly the peroxide shunt pathway to decarboxylate free fatty acids of carbon change lengths ranging from C12-C20 (Matthews et al., 2017 & Zhang et al., 2019). The reaction uses H_2O_2 as a reaction mediator, which when bound to the iron center of OleT$_{JE}$ gives it enough redox potential to proceed with the decarboxylation reaction of a free fatty acid. **Figure 1-4** below shows an example of the interaction between OleT$_{JE}$ and H_2O_2.

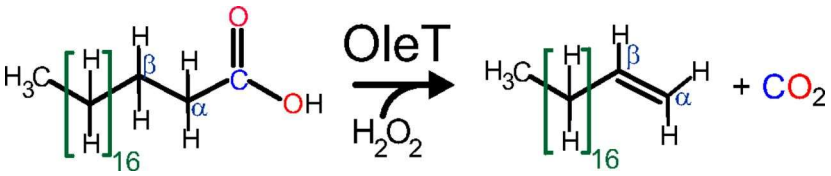


Figure 1-4. The decarboxylation reaction catalyzed by OleT$_{JE}$. Using a hydrogen peroxide co-substrate, OleT$_{JE}$ metabolizes a C$_n$ chain-length fatty acid to produce a C$_{n-1}$ alkene and carbon dioxide coproduct. (From Grant et al. 2016).

Terminal alkenes are important in today's world because they are considered as a source of biofuels and can have many uses depending on the length of the chain. Examples of this include anything from jet fuel to lubricants (Belcher et al., 2013). In an ideal scenario, enzymes such as OleT$_{JE}$ will be used to provide a "drop-in" solution to produce biofuels and thus have less of a reliance on the current non-renewable sources.

A previous study by Matthews et al. has evaluated a fusion of $OleT_{JE}$ with alditol oxidase (AldO). The fusions included different amino acid linkers such as $OleT_{JE}$-HRV3C-AldO, $OleT_{JE}$-alphahelix-AldO, and a cleaved fusion which compared $OleT_{JE}$ to AldO in a 1:1 protein ratio. AldO oxidizes polyols, which provides a source of H_2O_2 for $OleT_{JE}$. Findings suggest that there are high levels of substrate conversion in all three enzyme systems when in the presence of 0.1% and 1% glycerol (substrate conversion ranges from 81%, the lowest, to 98%, the highest). The major product in most cases is 1-tridecene when compared to the other two products which are the α-hydroxylated product (A-OH) and the β-hydroxylated product (B-OH) (Matthews et al., 2017). Finding other functional systems in which the ability of $OleT_{JE}$ to decarboxylate the fatty acid substrate into terminal alkene product has greatly increased could lead to a great source of a "drop-in" fuel production.

1.3 Superoxide Reductase benefits in reducing oxidative stress

Superoxide Reductase (SOR) is a powerful antioxidant enzyme produced in hyperthermophilic, obligate anaerobic archaea such as *Pyrococcus furiosus* and has been found to reduce superoxide to hydrogen peroxide using the electrons from reduced rubredoxin, without the production of oxygen as happens with classic superoxide dismutase enzymes (**Figure 1-5;** Grunden et al., 2005).

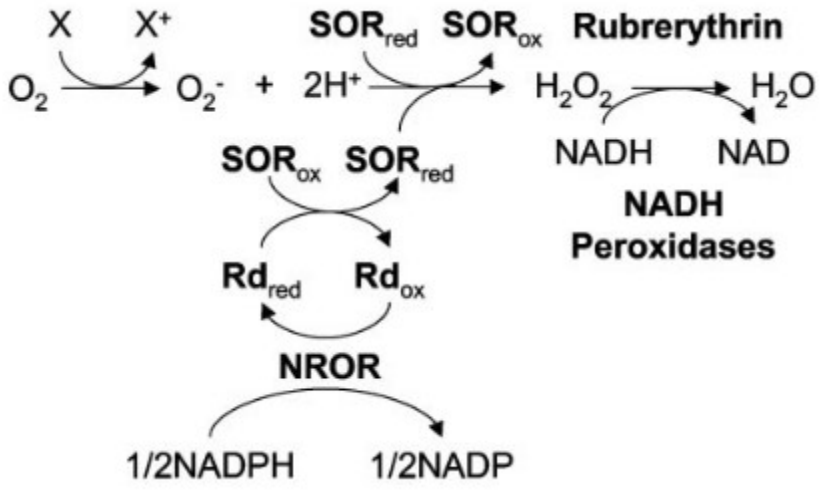


Figure 1-5: The NAD(P)H: rubredoxin oxidoreductase/ rubredoxin/ Superoxide reductase (NROR/Rd/SOR) pathway for superoxide oxidation by NADPH in *P. furiosus.* In a reducing environment, molecular oxygen will eventually be reduced to superoxide, often by reduced transition metals, such as iron (Symbolized by X and X^+). Superoxide is reduced by the reduced form of SOR. SOR uses electrons donated from rubrerythrin reduced in an NAD(P)H dependent reduction by NROR. The hydrogen peroxide that is produced is then reduced by the nonheme iron protein rubrerythrin, which has been shown to function as a peroxidase in *P. furiosus* (From Grunden et al., 2005).

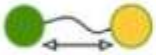
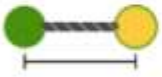
Characterization of this intriguing anaerobic antioxidant system prompted researchers to investigate the antioxidant properties and enhanced stress resilience afforded through the transgenic expression of *P. furiosus* SOR in bacterial and plant systems. The *P. furiosus* SOR gene has since been introduced into plants such as the Little Dogwood (*Cornus canadensis*) (Geng et al., 2016) and Arabidopsis (Im et al., 2009). Findings suggest that SOR aids in heat tolerance and drought tolerance. This is likely due to the ability of SOR to reduce reactive oxygen species (ROS) that are formed in response to abiotic stresses.

The ability of SOR to produce hydrogen peroxide from superoxide combined with $OleT_{JE}$'s ability to produce terminal alkenes from fatty acids provided the impetus to investigate whether a fusion combining these two enzymes could improve $OleT_{JE}$ activity and stability.

1.4 Amino Acid Linkers can increase bioactivity of fused enzymes.

The idea of fusing two or more proteins together is to incorporate distinct characteristics of each of the enzymes into the new fusion system. Generating functional enzyme fusions is not trivial, as it is difficult to know *a priori* whether fusion partners will fold properly, but it has been widely understood that introducing amino acid linkers adds to the stability and may increase interaction amongst the fused proteins (Chen et al., 2013). To this end, there are several options when selecting the type of amino acid linker to use. In brief, there are three main classifications of linkers that are typically considered for use in enzyme fusions: flexible, rigid, and cleavable linkers (**Table 1-1**).

Table 1-1: Summary of linker types (from Chen et al., 2013)

Linker		Advantages	Characteristics	Examples
Flexible		Allow for interaction between domains, or	Rich in small or hydrophilic amino acids	$(GGGGS)_n$, $(G)_n$
		Increase spatial separation between domains		
Rigid		Maintain distance between domains	Helical structure or rich in Pro	$(EAAAK)_n$, $(XP)_n$
Cleavable		Allow for *in vivo* separation of domains	Reductive or enzymatic cleavage	Disulfide, protease sensitive sequences

1.4.1 flexible linkers

Flexible linkers are typically composed of a chain of nonpolar amino acids such as Gly and have a polar amino acid group such as a Ser or Thr involved as well (Argos et al., 1990). An example of this would be the amino acid sequence $(GGGGS)_2$. As the name suggests, flexible amino acids permit a degree of movement between the two proteins. Flexibility only increases as the number of repeats increase.

Non-repeating sequences of amino acids also exist and provide the added function that they can be made to avoid large hydrophobic residues and maintain good solubility in aqueous solutions. The non-repeating sequence is less likely to be deleted during homologous recombination (Chen et al., 2013). An example of this is the amino acid sequence GSAGSAAGSGEF (Waldo et al., 1999). Flexible linkers can provide a fixed distance between the two proteins while also permitting a certain degree of movement and most likely allowing better protein folding which can be adjusted based on the length of the linker to reach an optimal biological activity between the two proteins.

1.4.2 Rigid Linkers

A potential pitfall of flexible linkers is that the lack of rigidity can prove to be detrimental and provide low expression or even less activity compared to the unfused proteins (Chen et al., 2013). Lack of rigidity may allow for too much movement. This is not the case with rigid linkers since they provide a fixed distance between the fused proteins. In a previous study by Amet et al. a flexible linker was used in a Tf-granulocyte colony stimulating factor fusion protein, but it proved ineffective. Applying a rigid linker instead provided the solution that allowed a fixed distance and each fusion protein to maintain its independent function (Amet et al., 2009). Rigid

linkers can be formed by amino acids that will adopt alpha helical structures such as $(EAAAK)_2$ (Bai et al., 2006). Another characteristic of rigid linkers is that they can be rich in proline residues such as $(PK)_n$. Proline in non-helical linkers can add increased stiffness and effective protein separation of protein domains (Chen et al., 2013). In either case, the linkers are providing a fixed distance between the two proteins. In a similar fashion, rigid linkers could aid in protein folding and increased interaction between the fused proteins by providing a fixed distance between them.

1.4.3 Cleavable linkers

This type of linker provides an important benefit of being able to separate the fused proteins when desired. This also provides researchers the opportunity to assess protein dynamics in a 1:1 ratio to observe how much activity each protein possesses individually as compared to their activity as part of a fusion. From an industrial application point of view this is important because it provides greater control over reaction dynamics. These types of linkers are special because they are cleaved only under specific conditions such as the presence of reducing agents or proteases (Chen et al., 2013). A characteristic of these linkers is that they allow for a precisely constructed homogenous product which can contain intramolecular disulfide bonds within the linker, for example, and capitalize on their reversible nature to cleave them when necessary. **Table 1-2** highlights some of the common outcomes that have occurred when adding amino acid linkers to different protein fusions.

Table 1-2: Amino acid linker additions and their effect on structure and biological activity.

Fusion Protein	Linker Type	Linker Sequence	Linker Function	Reference
Small-Chain variable fragment	Flexible	$(GGGGS)_3$	Increase Stability/Folding	Huston et al., 1988
Granulocyte colony-stimulating factor and transferrin	Flexible	$(GGGGS)_3$	Increase Stability/Folding Improve biological activity	Bai et al., 2006
Virus coat protein	Rigid	$(EAAAK)_3$	Increase Stability/Folding	Takamatsu et al., 1990
Beta-glucanase-xylanase	Rigid	$(EAAAK)_3$	Increase Stability/Folding	Lu et al., 2008
Human Serum Albumin and Interferon alpha2b	Flexible	GGGGS	Improve biological activity	Zhao et al., 2008
Human Serum Albumin and Interferon alpha2b	Rigid	PAPAP	Improve Biological Activity	Zhao et al., 2008
OleT and AldO	Cleavable	GSGLEVLFQGPGSGGGGS (HRV3C)	Increase substrate conversion	Matthews et al., 2017

A previous study showed that HRV3C linking $OleT_{JE}$ and AldO resulted in significant substrate conversion (Matthews et al., 2017), findings that prompted our interest in determining whether including the HRV3C linker in a $OleT_{JE}$-SOR fusion would also be beneficial. For our study we have focused on creating fusions using the following linkers as shown in **Table 1-3**. We wanted to include at least one of each type of linker so that we could evaluate the impacts of different linker types on $OleT_{JE}$-SOR fusion performance and structure.

Table 1-3: Amino acid linkers used in this study

Linker Name	Type	Linker Length (amino acids)
$(PK)_7$	Rigid	14
PAPAP	Rigid	5
$(EAAAK)_5$	Rigid	25
$(GGGGS)_2$	Flexible	10
GSAGSAAGSGEF **(Nonhomolog)**	Flexible	12
GSGLEVLFQGPGSGGGGS **(HRV3C)**	Cleavable	15
TGVS **(Bae1)**	Tether	4

1.5 Amino Acid Linkers in recombinant fusion proteins

Amino acids have been used as linkers in recombinant fusions for a while now to increase activity or stability between the two fused enzymes (Chen et al., 2013). The mechanism by which they interact is affected by the linker used. Some linkers can cause there to be cascade reactions, NAPDH recycling systems, electron transfer, and hydrogen peroxide supply for peroxidases as is the case for our fusion system (Aalbers et al., 2019). Creating multienzyme systems is a useful tool for biocatalytic applications.

For this study, $OleT_{JE}$ has been fused with SOR using several amino acid linkers that each have different structural properties. This will allow us to examine the impact of the different structural features of the fusion on $OleT_{JE}$ performance, which would have implication to the development of a robust fatty acid to alkene enzyme conversion system for biofuel production. The **Figure 1-6** outlines the overall interaction we want to have between $OleT_{JE}$ and SOR. Note that each $OleT_{JE}$-SOR fusion listed in Table 1-3 has a different amino acid linker and therefore may have difference performance characteristics.

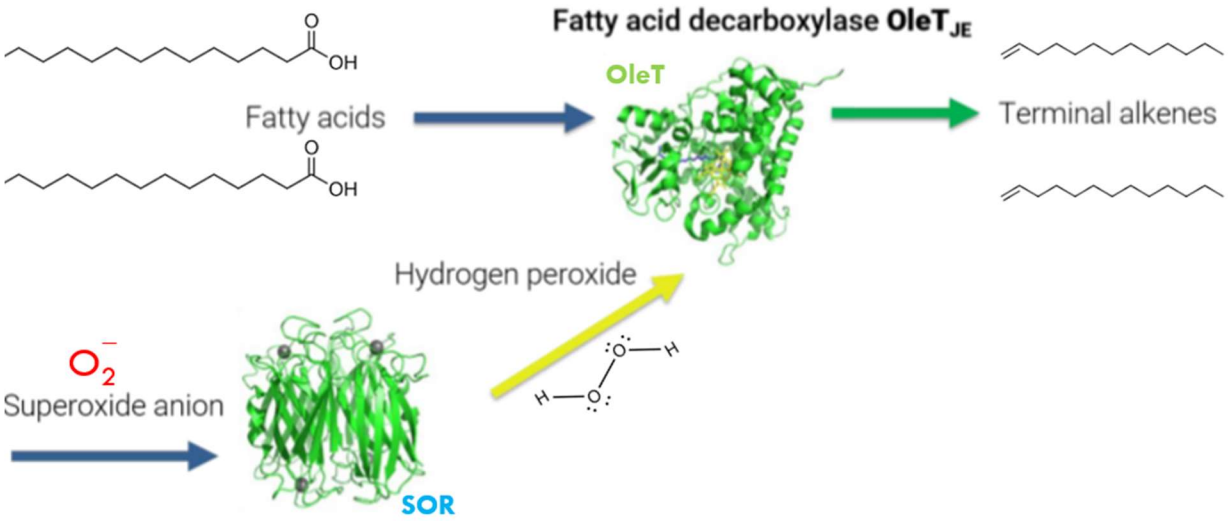


Figure 1-6: Proposed OleT$_{JE}$-SOR fusion reaction scheme.

1.6 Analytical assays to determine structural impact of the addition of amino acid linkers

There are various techniques to characterize protein structure and their interactions. Some examples are X-ray crystallography, analytical ultracentrifugation (AUC), and small angle x-ray scattering (SAXS). Each method provides unique and complementary information regarding protein structure and oligomerization that can be useful to study the effects of various linkers on our engineered protein fusions.

1.6.1. X-ray Crystallography.

X-ray Crystallography can be used to determine atomic and molecular structures of a crystal. Crystalline atoms cause the x-ray beams to diffract. The resulting diffraction data can be used to create a 3D image of the density of the electrons within the crystal (Ryu et al. 2017). An example of how x-ray crystallography has shown to be a beneficial technique in determining the structure of a fusion proteins is the crystallization of maltose-binding protein MCL1 fusion

which yielded a robust crystallography platform that generated the first apo MCL1 crystal structure and thus helped identify inhibitor binding areas that were never before accessed (Clifton et al. 2015). MCL1 is highly expressed in various tumors and can contribute to tumor development and resistance to chemotherapy (Zhang et al., 2002). Another example is GST-driven crystallization. A study done by Zhan et al., focused on using GST-fusion proteins as vehicles to determine the crystal structure of attached small peptides and biological regulatory domains. Since GST fusion proteins can be easily crystallized, they can be used as a technique for the structural determination of small regulatory domains (Zhan et al., 2001).

1.6.2 AUC: Sedimentation velocity and sedimentation equilibrium

Analytical ultracentrifugation (AUC) is an analytical technique that measures radial concentration gradients by the application of centrifugal force. Data produced using this technique can provide insight into the mass, size, and shape of a macromolecule in solution. (Cole et al., 2013). Sedimentation velocity (SV) analysis, which is a type of AUC, is often used to study reversible protein interactions such as characterization of self-association, heterogenous association, multiprotein complexes, binding constants, and determination of association constants (Brown et al., 2008). SV measures the rate at which molecules respond to centrifugal force. The samples are spun at very high speeds (>40,000 rpm) so that the sedimentation rate can be measured as it outcompetes diffusion. The rate of sedimentation provides information about mass and shape. (Cole et al., 2013). This method is used so frequently that AUC analysis is likely in reference to SV. Another AUC method that provides valuable information is sedimentation equilibrium (SE). In this technique, the protein is spun fast but not fast enough that it forms a pellet (~<20,000 rpm). Essentially what occurs is that over time, the forces of

sedimentation and diffusion “cancel” each other out and equilibrium is reached. This data can give insight into the native state of the protein (e.g., monomer or dimer) by allowing an accurate molecular weight to be determined. (Rhodes et al., 2009)

1.6.3 Small Angle X-ray Scattering (SAXS)

SAXS is a powerful method for the structural characterization of both ordered and disordered proteins in solution. SAXS provides low resolution information on the shape, conformation and assembly state of the proteins, nucleic acids and various macromolecular complexes (Kikhney et al., 2015). The sample is exposed to x-rays of a specific wavelength which scatter elastically from 0 to 5 degrees to produce a spatially averaged intensity distribution. Information about pore size, specific inner surface, surface to volume ratio, solution structure factor and lattice type can be observed depending on the type of sample analyzed. (Liu et al., 2013). SAXS is a similar technique to x-ray crystallography but does not require a crystal and can be done in solution. One complication, however, is that there is no orientational homogeneity, so scattering intensity depends solely on the scattering angle (Tamura et al., 2014). For this reason, SAXS is usually paired with other experiments such as AUC.

1.7 Summary and Perspective

In helping to solve the problem of non-renewable resources, we often find that the answers can come from nature. In our case, we turn to $OleT_{JE}$ which is an enzyme that has the capability to decarboxylate free fatty acids to generate terminal olefins and uses hydrogen peroxide as a reaction mediator (Rude et al., 2011). It is also known that SOR can reduce superoxide anions to generate hydrogen peroxide (Grunden et al., 2005). The focus of this

research is to fuse both proteins together with different amino acid linkers to create enzyme fusions in which the two proteins interact with each other in such a way that terminal alkenes can be effectively produced without building up too much hydrogen peroxide that could inactivate the catalytic heme group in $OleT_{JE}$.

The addition of amino acid linkers should give us an indication that we are reaching either one of those targets. Ideally, we should see high substrate conversion as well as a functional protein fusion that remains stable over time. Of course, this will require some trial and error in terms of finding the ideal linker type and length that meets the desired criteria. For this reason, it is important that we study the structure of the best fusions so we can further understand the structural changes that occur as a result of the linker. A combination of AUC and SAXS experiments will help us determine the mass and shape of the protein constructs which will yield valuable information regarding the changes between the different linker fusions. X-ray crystallography could also be a very useful method to study the structure of the fusions, however, growing the crystals has its own set of complications which is why we pursued the SAXS and AUC methods to generate our models since those can be done in solution.

References

Aalbers, F. S., & Fraaije, M. W. (2019). Enzyme Fusions in Biocatalysis: Coupling Reactions by Pairing Enzymes. *ChemBioChem*, *20*(1), 20–28. https://doi.org/10.1002/cbic.201800394Amet, N., Lee, H. F., & Shen, W. C. (2009). Insertion of the designed helical linker led to increased expression of Tf-based fusion proteins. *Pharmaceutical Research*, *26*(3), 523–528. https://doi.org/10.1007/s11095-008-9767-0

Amin-ul Mannan, M., Hazra, D., Karnwal, A., & Kannan, D. C. (2017). Algae as a platform for biofuel production - a sustainable perspective. *Algologia*, *27*(3), 337–356. https://doi.org/10.15407/alg27.03.337

Argos, P. (1990). An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. *Journal of Molecular Biology*, *211*(4), 943–958. https://doi.org/10.1016/0022-2836(90)90085-Z

Bai, Y., & Shen, W. C. (2006). Improving the oral efficacy of recombinant granulocyte colony-stimulating factor and transferrin fusion protein by spacer optimization. *Pharmaceutical Research*, *23*(9), 2116–2121. https://doi.org/10.1007/s11095-006-9059-5

Belcher, J., McLean, K. J., Matthews, S., Woodward, L. S., Fisher, K., Rigby, S. E. J., Nelson, D. R., Potts, D., Baynham, M. T., Parker, D. A., Leys, D., & Munro, A. W. (2014). Structure and biochemical properties of the alkene producing cytochrome p450 OleTJE (CYP152l1) from the jeotgalicoccus sp. 8456 bacterium. *Journal of Biological Chemistry*, *289*(10), 6535–6550. https://doi.org/10.1074/jbc.M113.527325

Brown, P. H., Balbo, A., & Schuck, P. (2008). Characterizing protein-protein interactions by sedimentation velocity analytical ultracentrifugation. *Current Protocols in Immunology*, *SUPPL. 81*, 1–36. https://doi.org/10.1002/0471142735.im1815s81

Chen, X., Zaro, J., & Shen, W.C (2013) Fusion Protein Linkers: Property, Design and Functionality. *Adv Drug Deliv* Rev. 2013 October 15; 65(10): 1357–1369. doi:10.1016/j.addr.2012.09.039.

Clifton, M. C., Dranow, D. M., Leed, A., Fulroth, B., Fairman, J. W., Abendroth, J., Atkins, K. A., Wallace, E., Fan, D., Xu, G., Ni, Z. J., Daniels, D., Van Drie, J., Wei, G., Burgin, A. B., Golub, T. R., Hubbard, B. K., & Serrano-Wu, M. H. (2015). A maltose-binding protein fusion construct yields a robust crystallography platform for MCL1. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0125010

Cole, J. L., & Sciences, M. (2013). *Analytical Ultracentrifugation Centrifugation | Analytical Ultracen- trifugation ☆ Characterizing biopharmaceuticals us- ing analytical ultracentrifugation*.Grant, J. L., Mitchell, M. E., & Makris, T. M. (2016). Catalytic strategy for carbon-carbon bond scission by the cytochrome p450 olet. *Proceedings of the*

National Academy of Sciences of the United States of America, *113*(36), 10049–10054. https://doi.org/10.1073/pnas.1606294113

Geng, X. M., Liu, X., Ji, M., Hoffmann, W. A., Grunden, A., & Xiang, Q. Y. J. (2016). Enhancing heat tolerance of the little dogwood cornuscanadensis L.f. with introduction of a superoxide reductase gene from the hyperthermophilic archaeon Pyrococcus furiosus. Frontiers in Plant Science, 7(JAN2016), 1–8. https://doi.org/10.3389/fpls.2016.00026

Grant, J. L., Mitchell, M. E., & Makris, T. M. (2016). Catalytic strategy for carbon-carbon bond scission by the cytochrome p450 olet. Proceedings of the National Academy of Sciences of the United States of America, 113(36), 10049–10054. https://doi.org/10.1073/pnas.1606294113

Grunden, A. M., Jenney, F. E., Ma, K., Ji, M., Weinberg, M. V., & Adams, M. W. W. (2005). In vitro reconstitution of an NADPH-dependent superoxide reduction pathway from *Pyrococcus furiosus*. *Applied and Environmental Microbiology*, *71*(3), 1522–1530. https://doi.org/10.1128/AEM.71.3.1522-1530.2005

Huijbers, M. M. E., Zhang, W., Tonin, F., & Hollmann, F. (2018). Light-Driven Enzymatic Decarboxylation of Fatty Acids. *Angewandte Chemie - International Edition*, *57*(41), 13648–13651. https://doi.org/10.1002/anie.201807119

Huston, J. S., Levinson, D., Mudgett-hunter, M., Tai, M., Novotný, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Huston, J. S., Levinson, D., Mudgett-huntert, M., Tai, M., Novotnv, J., Margolies, M. N., Ridge, R. J., Bruccolerit, R. E., Habert, E., Crea, R., & Oppermann, H. (2020). *Protein Engineering of Antibody Binding Sites : Recovery of Specific Activity in an Anti- Digoxin Single-Chain Fv Analogue Produced in Escherichia coli Roberto Crea and Hermann Oppermann Source : Proceedings of the National Academy of Sciences of the United States of America , Published by : National Academy of Sciences Stable URL : https://www.jstor.org/stable/32262*

Im, Y. J., Ji, M., Lee, A., Killens, R., Grunden, A. M., & Boss, W. F. (2009). Expression of Pyrococcus furiosus superoxide reductase in Arabidopsis enhances heat tolerance. Plant Physiology, 151(2), 893–904. https://doi.org/10.1104/pp.109.145409

Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Letters*, *589*(19), 2570–2577. https://doi.org/10.1016/j.febslet.2015.08.027

Liu, L., Boldon, L., Urquhart, M., & Wang, X. (2013). Small and Wide Angle X-ray Scattering studies of biological macromolecules in solution. *Journal of Visualized Experiments : JoVE*, *71*, 2–7. https://doi.org/10.3791/4160

Liu, Y., Wang, C., Yan, J., Zhang, W., Guan, W., Lu, X., & Li, S. (2014). Hydrogen peroxide-independent production of α-alkenes by OleT JE P450 fatty acid decarboxylase. *Biotechnology for Biofuels*, *7*(1), 1–12. https://doi.org/10.1186/1754-6834-7-28

Lu, P., & Feng, M. G. (2008). Bifunctional enhancement of a β-glucanase-xylanase fusion enzyme by optimization of peptide linkers. *Applied Microbiology and Biotechnology*, *79*(4), 579–587. https://doi.org/10.1007/s00253-008-1468-4Matthews, S., Tee, K. L., Rattray, N. J., McLean, K. J., Leys, D., Parker, D. A., Blankley, R. T., & Munro, A. W. (2017). Production of alkenes and novel secondary products by P450 OleTJE using novel H2O2-generating fusion protein systems. *FEBS Letters*, *591*(5), 737–750. https://doi.org/10.1002/1873-3468.12581

Matthews, S., Tee, K. L., Rattray, N. J., McLean, K. J., Leys, D., Parker, D. A., Blankley, R. T., & Munro, A. W. (2017). Production of alkenes and novel secondary products by P450 OleTJE using novel H2O2-generating fusion protein systems. FEBS Letters, 591(5), 737–750. https://doi.org/10.1002/1873-3468.12581

Naik, S. N., Goud, V. V., Rout, P. K., & Dalai, A. K. (2010). Production of first and second generation biofuels: A comprehensive review. *Renewable and Sustainable Energy Reviews*, *14*(2), 578–597. https://doi.org/10.1016/j.rser.2009.10.003

Nanda, S., Rana, R., Sarangi, P. K., Dalai, A. K., & Kozinski, J. A. (2018). A broad introduction to first-, second-, and third-generation biofuels. *Recent Advancements in Biofuels and Bioenergy Utilization*, 1–25. https://doi.org/10.1007/978-981-13-1307-3_1

Nigam, P. S., & Singh, A. (2011). Production of liquid biofuels from renewable resources. *Progress in Energy and Combustion Science*, *37*(1), 52–68. https://doi.org/10.1016/j.pecs.2010.01.003

Nozzi, N. E., Oliver, J. W. K., & Atsumi, S. (2013). Cyanobacteria as a Platform for Biofuel Production. *Frontiers in Bioengineering and Biotechnology*, *1*(September), 1–6. https://doi.org/10.3389/fbioe.2013.00007

Rhodes, D. G., Bossio, R. E., & Laue, T. M. (2009). Chapter 39 Determination of Size, Molecular Weight, and Presence of Subunits. In *Methods in Enzymology* (1st ed., Vol. 463, Issue C). Elsevier Inc. https://doi.org/10.1016/S0076-6879(09)63039-1

Rude, M. A., Baron, T. S., Brubaker, S., Alibhai, M., Del Cardayre, S. B., & Schirmer, A. (2011). Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from Jeotgalicoccus species. *Applied and Environmental Microbiology*, *77*(5), 1718–1727. https://doi.org/10.1128/AEM.02580-10

Ryu, W. (2017). *X-Ray Crystallography Understanding and utilizing the bio- molecule / nanosystems interface*.

Takamatsu, N., Watanabe, Y., Yanagi, H., Meshi, T., Shiba, T., & Okada, Y. (1990). Production of enkephalin in tobacco protoplasts using tobacco mosaic virus RNA vector. *FEBS Letters*, *269*(1), 73–76. https://doi.org/10.1016/0014-5793(90)81121-4

Tamura, N., & Gilbert, P. U. P. A. (2014). Small-Angle X-Ray Scattering Research Methods in Biomineraliza- tion Science.

Verma, M., Godbout, S., Brar, S. K., Solomatnikova, O., Lemay, S. P., & Larouche, J. P. (2012). Biofuels production from biomass by thermochemical conversion technologies. *International Journal of Chemical Engineering*, *2012*. https://doi.org/10.1155/2012/542426

Waldo, G. S., Standish, B. M., Berendzen, J., & Terwilliger, T. C. (1999). Rapid protein-folding assay using green fluorescent protein. *Nature Biotechnology*, *17*(7), 691–695. https://doi.org/10.1038/10904

Xu, H., Ning, L., Yang, W., Fang, B., Wang, C., Wang, Y., Xu, J., Collin, S., Laeuffer, F., Fourage, L., & Li, S. (2017). In vitro oxidative decarboxylation of free fatty acids to terminal alkenes by two new P450 peroxygenases. *Biotechnology for Biofuels*, *10*(1), 1–15. https://doi.org/10.1186/s13068-017-0894-x

Zhan, Y., Song, X., & Zhou, G. W. (2001). Structural analysis of regulatory protein domains using GST-fusion proteins. *Gene*, *281*(1–2), 1–9. https://doi.org/10.1016/S0378-1119(01)00797-1

Zhang, B., Gojo, I., & Fenton, R. G. (2002). Myeloid cell factor-1 is a critical survival factor for multiple myeloma. *Blood*, *99*(6), 1885–1893. https://doi.org/10.1182/blood.V99.6.1885

Zhang, S., & Liu, Y. (2019). Mechanism of fatty acid decarboxylation catalyzed by a non-heme iron oxidase (UndA): A QM/MM study. *Organic and Biomolecular Chemistry*, *17*(45), 9808–9818. https://doi.org/10.1039/c9ob02116g

Zhang, W., Ma, M., Huijbers, M. M. E., Filonenko, G. A., Pidko, E. A., Van Schie, M., De Boer, S., Burek, B. O., Bloh, J. Z., Van Berkel, W. J. H., Smith, W. A., & Hollmann, F. (2019). Hydrocarbon Synthesis via Photoenzymatic Decarboxylation of Carboxylic Acids. *Journal of the American Chemical Society*, *141*(7), 3116–3120. https://doi.org/10.1021/jacs.8b12282

Zhao, H. L., Yao, X. Q., Xue, C., Wang, Y., Xiong, X. H., & Liu, Z. M. (2008). Increasing the homogeneity, stability and activity of human serum albumin and interferon-α2b fusion protein by linker engineering. *Protein Expression and Purification*, *61*(1), 73–77. https://doi.org/10.1016/j.pep.2008.04.013

CHAPTER 2: Characterization of $OleT_{JE}$-SOR fusion enzymes for use in biofuel production

ABSTRACT

$OleT_{JE}$ is an enzyme that catalyzes the conversion of fatty acid chains into terminal alkenes, a biofuel precursor, using hydrogen peroxide as a reaction mediator. Superoxide reductase (SOR) is an antioxidant enzyme present in the hyperthermophilic archaeon, *Pyrococcus furiosus* and has been shown to mitigate oxidative stress in recombinant plants and bacteria. SOR catalyzes conversion of the superoxide anion into hydrogen peroxide and could provide the necessary hydrogen peroxide to efficiently support $OleT_{JE}$ activity. Otherwise, the hydrogen peroxide would have to be added exogenously, which is harmful to cells as well as the reactive heme center of $OleT_{JE}$. Past research has shown that fusions of these two enzymes with $OleT_{JE}$ on the N-terminal end and SOR on the C-terminal end are functional; however, the inverse fusion variant has very low activity, likely because of improper protein folding or steric hindrance. Therefore, we are evaluating whether the addition of a linker region between the $OleT_{JE}$ and SOR enzymes can improve bioactivity. Past research using amino acid linkers has shown that amino acid linkers could influence bioactivity and stabilization of the fusion constructs. We are using that knowledge to generate different fusions of $OleT_{JE}$ -SOR with different amino acid linkers. Gas Chromatography-Flame Ionization Detection (GC/FID) analysis of our protein fusions revealed that only fusion $OleT_{JE}$ -$(GGGGS)_2$-SOR was comparable to $OleT_{JE}$ in terms of product formation. *In vivo* data suggests that SOR is playing a protective role in protein function. Structural data via analytical ultracentrifugation (AUC) and small angle x-ray scattering (SAXS) was obtained to give us more insight into the interaction between the two proteins.

2.1 Introduction

Finding a cost-effective way to generate biofuels has been an area of research interest in the scientific community in recent years. Using enzymes that can decarboxylate free fatty acids into aliphatic hydrocarbons that can be refined into drop-in transportation fuels could be the solution towards making biofuels more cost-effective. *Jeotgalicoccus* sp. 8456 $OleT_{JE}$ (CYP152L1) is a fatty acid decarboxylase cytochrome P450 that uses hydrogen peroxide (H_2O_2) to catalyze production of terminal alkenes, which are industrially important chemicals with biofuel applications (Belcher et al., 2014, Rude et al., 2011, & Matthews et al., 2017). We want to utilize this process to generate the precursors for conversion to third generation biofuels.

Third generation biofuels are typically derived from algae, sewage sludge, or various food waste. These feedstocks are used generate fuels such as refined algal oils, biodiesel, and aviation fuels. Unlike first- and second-generation fuels, they can serve as drop-in fuel replacements using our current infrastructure, fueling an array of vehicles including gasoline-powered light-weight automobiles, diesel powered heavy-duty trucks, and even supersonic jets (Nanda et al., 2018 & Nigam et al., 2011). In order to keep up with energy demands, we must find a way to develop scalable and economically feasible renewable biofuel production platforms. Particular interest has arisen for the development of scalable algae-based biofuel production platforms to produce bioethanol, biodiesel, and aviation fuels because algae production would not compete with food production crops for arable land use, and algae grow rapidly under controlled growth conditions that maximize carbon dioxide to algal oil synthesis (Mannan et al., 2017 & Schenk et al., 2008).

Established chemical methods for the decarboxylation of fatty acids require rather harsh reaction conditions and rare metal catalysts (Huijbers et al., 2018 & Zhang et al., 2019). Since

OleT$_{JE}$ uses H_2O_2 as a reaction mediator (Rude et al., 2011 & Liu et al., 2014), there was interest in determining whether fusing OleT$_{JE}$ with an enzyme that can produce H_2O_2 could benefit OleT$_{JE}$ performance. Ideally the fusion would maintain the enzymes in close proximity to each other such that OleT$_{JE}$ could readily access H_2O_2 produced by SOR in such a way that terminal alkenes would be produced consistently but not allow too much H_2O_2 buildup which can cause oxidative damage to the heme group in OleT$_{JE}$. Using an enzyme-based fatty acid decarboxylase instead of an expensive metal catalyst to deoxygenate fatty acids for fuel precursor production could provide a more cost-effective system. A recent study by Matthews et al. evaluated the performance of fusions of OleT$_{JE}$-AldO in which alditol oxidase (AldO) oxidizes polyols to produce H_2O_2 that also had different amino acid linkers. They observed that the primary product is the desired terminal alkene with high percentages of substrate conversion to a product (In the 1% glycerol substrate conversion OleT$_{JE}$-HRV3C-AldO converted 96% of product while OleT$_{JE}$+AldO converted 88% and OleT$_{JE}$- α-helix-AldO converted 98%. In 0.1% glycerol OleT$_{JE}$-HRV3C-AldO converted 81%, OleT$_{JE}$+AldO converted 97% and OleT$_{JE}$- α-helix-AldO converted 93%); however, the fusions did not significantly out-perform OleT$_{JE}$ alone. (Matthews et al., 2017). For our study, we wanted to determine whether the fusion of OleT$_{JE}$ with several different linkers and the hydrogen peroxide-producing enzyme superoxide reductase (SOR) could improve OleT$_{JE}$ function compared to OleT$_{JE}$ alone.

Superoxide Reductase (SOR) is a powerful antioxidant enzyme produced in hyperthermophilic, obligate anaerobic archaea such as *Pyrococcus furiosus* and has been found to reduce superoxide to hydrogen peroxide using the electrons from reduced rubredoxin, without the production of oxygen as happens with classic superoxide dismutase enzymes (Grunden et al., 2005). The *P. furiosus* SOR gene has since been introduced into plants such as the Little

Dogwood (*Cornus canadensis*) (Geng et al., 2016) and Arabidopsis (Im et al., 2009). Findings suggest that SOR aids in heat tolerance and drought tolerance. This is likely due to the ability of SOR to reduce superoxide that is formed in response to abiotic stresses.

Understanding whether an enzyme fusion will work *a priori* is a difficult task because there are factors that will influence whether there will be proper folding and stability such as the linkers being flexible, rigid, or cleavable. A review by Chen et al. highlighted several amino acid linkers and whether they increased stability, folding or biological activity (Chen et al., 2013). **Table 2-1** shows the fusions were generated as part of this study using Gibson Assembly and traditional restriction digest cloning methods.

Table 2-1. $OleT_{JE}$-SOR amino acid linker fusions generated for this study.

Linker Name	Type	Fusion name
$(PK)_7$	Rigid	$OleT_{JE}$-$(PK)_7$-SOR
PAPAP	Rigid	$OleT_{JE}$-PAPAP-SOR
$(EAAAK)_2$	Rigid	$OleT_{JE}$-EAAAK-SOR
$(GGGGS)_2$	Flexible	$OleT_{JE}$-$(GGGGS)_2$-SOR
GSAGSAAGSGEF (nonhomolog)	Flexible	$OleT_{JE}$-NonHomolog-SOR
GSGLEVLFQGPGSGGGGS (HRV3C)	Cleavable	$OleT_{JE}$-HRV3C-SOR
TGVS (Bae1)	Tether	$OleT_{JE}$-Bae1-SOR

Reactions with the purified $OleT_{JE}$-SOR fusions were analyzed using GC/FID to determine the overall product conversion in comparison to $OleT_{JE}$ as well as to structurally analyze the enzyme fusions using small angle x-ray scattering (SAXS) and analytical ultracentrifugation (AUC). It was hoped that these methods of analysis could provide insight with regards to why some fusions were more comparable to $OleT_{JE}$ than others. Since these different fusions each contain amino acid linkers with different structural properties, it will allow us to examine the difference each type of structural feature has on the performance of $OleT_{JE}$,

which would have implications to the development of a robust fatty acid to alkene enzyme conversion system for biofuel production.

2.2 Methods

Lauric acid, myristic acid, palmitic acid, stearic acid, hydrogen peroxide solution, all gas chromatography standards, N, O-bistrifluoroacetamide + 1% trimethylchlorosilane, and 3,3',5,5'-tetramethylbenzidine (TMB) were purchased from MilliporeSigma USA (St. Louis, MO). Methods used in this research study were performed in conjunction with unpublished work from Ph.D. student Hannah Wapshott-Stehli as part of her 2021dissertation titled "Examination of $OleT_{JE}$-SOR fusions for fatty acid decarboxylation".

2.2.1 *Linker design and cloning and expression*

The different amino acid linkers used in this study were chosen based on information from Chen et al. (Chen et al., 2013). We already had an existing fusion of $OleT_{JE}$ -Bae1-SOR in pET28a available from previous research in our lab. We decided on two different cloning approaches, one was Gibson Assembly, and the other was traditional restriction digest using BamHI and BGlII. Forward and reverse primers for each linker were designed for Gibson assembly as well as a site directed mutagenesis primers to modify the original BaeI cut site to BamHI for traditional restriction digest. The primers used for generating the $OleT_{JE}$-SOR fusion constructs are listed in **Table 2-2**.

Table 2-2: List of amino acid linker primers used, and the cloning method used to generate the fusion.

Primer Name	Sequence 5'-3'	Cloning Method
PAPAP F	CCA GAT CTC CGG CAC CGG CACCGG GAT CCG G	Restriction digest
PAPAP R	CCG GAT CCC GGT GCC GGT GCCGGA GAT CTG G	Restriction digest
GGGGS2 F	CCA GAT CTG GTG GTG GTG GTAGCG GTG GTG GTG GTA GCG GATCCG G	Restriction digest
GGGGS2 R	CCG GAT CCG CTA CCA CCA CCACCG CTA CCA CCA CCA CCA GAT CTGG	Restriction digest
PK7 F	CCG GAT CTC CGA AAC CGA AACCGA AAC CGA AAC CGA AAC CGAAAC CGA AAG GAT CCG G	Restriction digest
PK7 R	CCG GAT CCT TTC GGT TTC GGT TTCGGT TTC GGT TTC GGT TTC GGT TTCGGA GAT CCG G	Restriction digest
EAAAK2 F	ACG TTC GTG AGG TTG TAG ACA GAA CAG AAG CGG CGG CGA AAG AAG	Gibson assembly
EAAAK2 R	GCC GCT ACG AAT GGT TTC ACT AAT CAT TTT CGC CGC CGC TTC TTT C	Gibson assembly
NONHOMOFLEX F	CCG GAT CTG GAT CCG CTG GCTCCG CTG CTG GTT CTG GCG AATTCG GAT CCG G	Restriction digest
NONHOMOFLEX R	CCG GAT CCG AAT TCG CCA GAACCA GCA GCG GAG CCA GCG GATCCA GAT CCG G	Restriction digest
HRV3C F	CCA GAT CTG GTA GCG GTC TGGAAG TAC TGT TCC AGG GTC CGG GTAGCG GTG GTG GTG GTA GCG GATCCG G	Restriction digest
HRV3C R	CCG GAT CCG CTA CCA CCA CCACCG CTA CCC GGA CCC TGG AACAGT ACT TCC AGA CCG CTA CCA GATCTG G	Restriction digest

For the Gibson assembly approach, the $OleT_{JE}$-BaeI-SOR vector was linearized with Bae1 according to the New England Biolabs protocol for this enzyme. The linearized $OleT_{JE}$-SOR product was electrophoresed through a 0.8% agarose gel. The vector DNA band was excised from the gel and was extracted and purified according to the Qiagen gel extraction protocol. The purified linearized vector DNA was quantified using a nanodrop instrument.

The linkers were generated using two single stranded DNA fragments that were allowed to anneal by being treated at 95˚C for 5 min in the thermocycler and then allowed to cool at room temperature on the bench overnight. The annealing buffer was 10 mM Tris-HCl, 50 mM NaCl, 1 mM EDTA, pH 8. The linker DNA sequences are shown in **Table 2-2**. A ratio of 1:3 vector-to-insert was used for Gibson Assembly by New England Biolabs. Once the Gibson Assembly reaction was finished, the constructs were transformed into competent XL-1 blue cells following the manufacturer's protocol. Colony PCR was used to screen for the addition of the linker. Colony PCR amplified a 400 bp region that included the linker area. If a linker were successfully added, a band shift could be detected in this region as illustrated in **Figure 2-1** below.

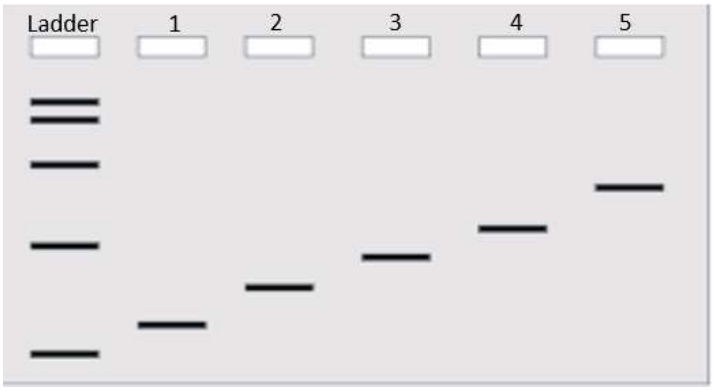


Figure 2-1: Example of how screening for linker length in the $OleT_{JE}$-SOR constructs was evaluated. The leftmost lane contains the DNA fragment sizing ladder. The second lane shows the fusion construct without a linker, and lanes 3-6 would be constructs containing linkers of progressively longer lengths. The observed band shift will correlate to size of the linker inserted.

$OleT_{JE}$-BaeI-SOR vectors were then modified to generate a BamHI cut site in place of BaeI using a Quickchange II site-directed mutagenesis kit (Agilent, Santa Clara, CA, USA). Nucleotides CGTA were deleted first using primers O_S CGTA DelF: 5'-GTAGACAGAACAACCGGTCTATGATTAGTGAAA C-3' and O_S CGTA DelR: 5'-

GTTTCACTAATCATAGACCGGTTGTTCTGTCTAC-3'. Nucleotides ATCC were then added using primers O_S ATCC AddF: 5'-GACAGAACAACCGGATCCTCTATGATTAGTGAAAC-3' and O_S ATCC AddR: 5'-CACTAATCATAGAGGATCCGGTTGTTCTGTC-3'. The new mutated fusion construct was digested with BamH1 to confirm the addition of the BamH1 site. Digestion by BamH1 was done according to the New England Biolabs protocol for the enzyme. Amino acid linker inserts were ligated into the linearized plasmid and transformed into *E. coli* XL-1 blue where colony PCR would be used on random transformants. A band shift such as shown **Figure 2-1** was used to identify constructs of interest and potential positive clones were sent off for sequencing confirmation.

2.2.2 Transformation into expression strain Arctic Express RIL (DE3) for screening fusion activity using crude extract assays.

Competent Arctic Express RIL (DE3) cell aliquots of 100 µl each were obtained from a previous stock and allowed to thaw on ice. The transformations were performed as per the Arctic Express RIL (DE3) protocol (Agilent, Santa Clara, CA). Overnight cultures were grown at 30°C in LB with 50 µg/ml kanamycin and 20 mg/ml gentamycin used to inoculate cultures the next day for expression. The inoculated cultures (50 ml) were then grown at 30°C in the same media until an OD of 0.6-0.8 was achieved. The temperature was then lowered to 16°C and 500 µM ALA and 500µM of FeS were added to aid in heme synthesis. Cells were then induced with 0.1 mM IPTG and grown at 16 ºC for 24 h. Cells were then centrifuged at 4144 x g at 4 ºC in a Fisher centrifuge. Cell pellets were re-suspended in 50 mM $NaPO_4$, 250 mM KNO_3, 1 mM

Sigma protease inhibitor cocktail for His-tagged proteins (Sigma, Milwaukee, WI), pH 7 before being frozen at -80°C.

2.2.3 Transformation into BL21 (DE3 for) large scale expression and purification.

Competent BL21 (DE3) *E. coli* cells were transformed according to the BL21 (DE3) protocol from New England Biolabs. Large scale expression required a 60 ml LB volume inoculated with a transformant of the desired fusion, grown overnight and with the addition of 50 µg/ml of kanamycin and 20 µg/ml gentamycin at 37 °C. The next day, 50 ml of the overnight culture was added to a 3L Fernbach flask containing 1 L LB supplemented with 1 ml kanamycin (50 µg/ml) and 1 mL gentamycin (20 µg/ml) to maintain the expression plasmids. The 1 L cultures were incubated at 37 °C until an OD of 0.06-0.08 was reached at which point 1 ml of 0.5 mM FeS, 1ml of 0.5 mM ALA and 200 µl of 0.5 M IPTG were added to help with cofactor synthesis and induce protein overexpression, respectively. The cultures were then incubated at 20 °C for 20 h. The cultures were then spun down at 4144 x g the supernatant discarded, and the pellets were resuspended in 50 µM $NaPO_4$, 250 mM KNO_3 pH 7 buffer and 1 mM Sigma protease inhibitor cocktail for His-tagged proteins (Sigma, Milwaukee, WI) and stored for future use in a -80 C freezer.

2.2.4 GC/FID crude assay sample preparation.

Cells were lysed using a French Pressure Cell set at 500 PSI by passing the cell mixture through three times. The broken cell suspensions were spun down at 28,000 x g at 4 °C, and the supernatants were collected for used in crude assays. One millimolar of Myristic Acid (diluted from 40mM stock) and 2 mM (diluted from 40mM stock) was then added to a 500 µl volume of

cell-free extract (CFE). The reactions were rocked slowly for 1 h before 37% HCl (v/v) was added to stop the reaction. Samples were then extracted for GC/FID analysis by adding an equal volume of ethyl acetate and vortexed for 30 s. Samples were spun down for 20 min at 13,300 x g at room temp. Top layers were then dried with magnesium sulfate, vortexed 5 s, and centrifuged for another 20 min at 13300 x g. Samples were then derivatized 1:1 v/v with N,O-bis (trimethylsilyl)trifluoroacetamide (BSTFA) at 70 °C for 20 min.

Derivatized fractions were loaded onto a Shimadzu GC/FID 2014. Split injection was used at a ratio of 1:10 to automatically load 1 µl of sample onto a Zebron ZB-5 column (30 m by 0.25 mm; 0.25-µm film) (Phenomenex). The inlet was set to 300˚C and the detector to 320˚C. The oven temperature and programmed gradient was set as in Rude *et al*. Peaks were identified using authentic standards from MilliporeSigma.

2.2.5 Protein Purification

Protein purification was performed as a 3-day process using the Bio-logic DuoFlow system from Bio-Rad USA. The first chromatography step made use of hydrophobic interaction, specifically a butyl sepharose column. Flow rate was 2 ml/min and samples were collected in 2 ml fractions. The gradient used was established by prior research done by Hannah Wapshott-Stehli dissertation work in our lab. The buffers used for this step were 50 mM $NaPO_4$, 250 mM KNO_3, 1 M NH_4SO_4 0.1% triton (Buffer A) as a binding buffer and a 50 mM $NaPO_4$, 250 mM KNO_3, 1% triton, pH 7, elution buffer (Buffer B). Fractions of the OleT$_{JE}$-SOR fusion proteins were collected and concentrated to 30 ml using a Vivaspin 20 ml 5000 MWCO centrifugal filter unit from Sartorius USA (Bohemia, NY) and dialyzed overnight in 4 L 5 mM $NaPO_4$, 500 mM KNO_3 pH 7 (Buffer C) using a 30 ml slide-a-lyzer G2 20,000 MWCO cassette from

ThermoFisher Scientific USA. The second column used was hydroxyapatite. The flow rate for this column was 1 ml/min and samples were collected in 1 ml fraction. The binding buffer was Buffer C and the elution buffer used was 500 mM $NaPO_4$, 250 mM KNO_3 pH 7 (Buffer D). $OleT_{JE}$-SOR fusion containing fractions were collected and concentrated down to 5 ml and dialyzed overnight in binding buffer for the size exclusion column, 100 mM KPO_4, 250 mM KNO_3 pH7 (Buffer E). As a final step, $OleT_{JE}$-SOR fusion samples were eluted through a size exclusion column (SEC). Flow rate was 1 ml/min and 1 ml fractions were eluted. Samples were collected and quantified using a pyridine hemachromagen assay (Barr et al., 2015). Lastly, samples were concentrated to a range of 50-100 µM using a 5 ml 5000 MWCO Amicon Ultra-4 Centrifugal filter unit from MilliporeSigma USA (St. Louis, MO) and stored in 15% glycerol in the -80 C for future use.

Throughout the purification process, fractions collected were quantified using a Bradford assay (BioRad USA), and polyacrylamide gel electrophoresis (PAGE) was performed for each fusion to verify band size.

2.2.6 GC/FID Analysis of fatty acid decarboxylation activity

Purified enzymes were used in fatty acid decarboxylation assays to compare the activity of $OleT_{JE}$ to that of the $OleT_{JE}$-SOR fusions. Reactions were prepared in 0.5 ml of 100 mM KPO_4, 300 mM KCl, pH 7, 1.5% Triton-X-100, 5% acetone (Buffer F) with 500 µM H_2O_2, 1 mM myristic acid, and 0.6µM of $OleT_{JE}$ or fusion enzyme. Reactions were incubated at 27˚C for 20 min while rotating at 150 rpm in a shaker platform before being quenched with 25 µl of 37% hydrochloric acid, spiked with 1 mM 1-heptadecene as an internal standard, and extracted using 400 µM ethyl acetate. Sample tubes were then spun at 13,300 x g for 20 min to separate the

layers. The top layer was extracted, dried with a pinch of anhydrous $MgSO_4$, and then derivatized 1:1 with BSTFA at 70°C for 20 min. Derivatized samples were collected and loaded onto the GC-FID instrument for reaction product quantification as described in section *2.2.4*.

2.2.7 $OleT_{JE}$ assays performed with methyl viologen to induce superoxide formation

Instead of 500 µM of H_2O_2, 2500 µg of *E. coli* NC906 cell free extract that had been re-suspended in 100 mM KPO_4, 300 mM KCl, pH 7, was added to the reaction along with 10 µM methyl viologen. Reactions were shaken at 150rpm at 27˚C as before, but were allowed to continue for 30 min instead of 20 min. This longer incubation time was meant to account for the slower generation of hydrogen peroxide from methyl viologen and SOR. Samples were extracted and run on the GC/FID instrument as described in section *2.2.4.*

2.2.8 E. coli NC906 growth rescue to confirm *in vivo* SOR activity

$OleT_{JE}$, SOR, and the fusion genes cloned into pBadHisA were transformed into the *sodAsodB* deletion strain *E. coli* NC906. Cultures were grown O/N in 5 ml of M63 minimal media containing glucose and casamino acids. Cells were spun down, washed twice with M63 minimal media containing glycerol instead of glucose and no amino acids, and then inoculated to an OD_{600} of ~0.06 in 30 ml of the same media. Cells were then grown at 37°C while shaking at 225 rpm. Cells were grown for five h before the initial OD_{600} reading was made, then an OD_{600} was taken every other hour until 15 h had passed.

2.2.9 Cytochrome C competition assay

SOR activity of the $OleT_{JE}$-SOR fusion proteins was evaluated using the cytochrome c competition assay in 0.05 M KPO_4, pH 7.8, with 0.01 M EDTA buffer. The standard reaction mixture contained 10 µM ferricytochrome c from horse heart, 50 µM xanthine, and enough xanthine oxidase to generate a slope of 0.025 absorbance units per min for the reduction of cytochrome c similar to a study done by Oneda et al. (Oneda et al., 2003). The reaction was monitored using a UV-2401PC UV-Vis Recording Spectrophotometer (Shimadzu, Kyoto, Kyoto Prefecture, JP) at 25°C. Once this slope was established, 959 µg of *E. coli* NC906 cellular extract (passed three times through a French Pressure cell and spun down) prepared from anaerobically grown cells was added to help re-reduce SOR. SOR or one of the fusions was then added until the slope of cytochrome c reduction decreased by 50% (to 0.0125). The specific activity is defined by that amount of SOR that was added to achieve this reduced slope. $OleT_{JE}$ was also assayed, and no inhibition of cytochrome c reduction was observed.

2.2.10 AUC sample preparation

Sedimentation velocity experiments were performed on $OleT_{JE}$ -BaeI-SOR, $OleT_{JE}$ – EAAAK-SOR, $OleT_{JE}$–$(GGGGS)_2$-SOR, $OleT_{JE}$-HRV3C-SOR, $OleT_{JE}$ and SOR. Samples were concentrated at 1 mg/ml, 129,000 x g and 150 scans per cell. All samples were in 100 mM KPO4, 250 mM KNO3, 15% glycerol pH7 buffer. AUC instrument used is a ProteomeLab XL-A by Beckman Coulter which uses ProteomeLab XL-A/XL-1 Graphical User Interface, version 6.2.

Sedimentation Equilibrium (SE) was run on fusion $OleT_{JE}$-$(GGGGS)_2$-SOR since we had the prior data indicating it was the most similar to $OleT_{JE}$ in terms of activity. This sample was

run at 3 different concentrations (1 mg/ml, 0.5 mg/ml, and 0.25 mg/ml) and at 3 different speeds 18,144 x g, 26,127 x g, and 32,256 x g.

These data were analyzed using a combination of sedfit, sednterp, UltraScan 4.0 and sedphat software. These tools were used to generate ellipses so that we could have a preliminary idea of the globular structure of each fusion.

2.2.11 SAXS analysis

$OleT_{JE}$ -BaeI-SOR, $OleT_{JE}$ –EAAAK-SOR, $OleT_{JE}$ –$(GGGGS)_2$-SOR, $OleT_{JE}$-HRV3C-SOR, $OleT_{JE}$ and SOR samples were sent to the Shared Materials Instrumentation Facility (SMIF) at Duke University, Durham, NC to acquire SAXS data using the Ganesha SAXS 2 instrument. Sample concentrations were at 1 mg/ml and in 100 µl aliquots. The resulting data were processed with Primus to measure radius of gyration (R_g) and for generating the Guinier plots and Kratky plots. The modelling software Modeller, Pymol and WinCoot were used to generate the putative protein structures.

2.3 Results

2.3.1 $OleT_{JE}$-SOR fusion Construction

An example of the screening process for the fusions prepared using the Gibson Assembly method is shown in **Figure 2-2**. Initial attempts at cloning did not yield very promising results, and it was not a very high throughput process because of low transformant count. Ultimately, the only clone generated using Gibson Assembly was $Olet_{JE}$-EAAAK-SOR, not including the original fusion $OleT_{JE}$-Bae1-SOR. After many troubleshooting attempts, an increase in

transformants was observed. Colony PCR allowed us to screen many transformants at a time, and once a band shift was detected in the gel, the clone would be isolated, the plasmid purified, and then restriction digested with BaeI to confirm the addition of the linker. If the linker was successfully added, it should not digest by BaeI any longer. Regardless, samples were sent for sequencing to confirm linker addition. One of the biggest benefits of using Gibson Assembly is that it can be a very effective cloning process if everything is calculated accurately such as the picamolar weights of vector and insert. It was observed that transformants that had passed the initial screen were usually positive clones.

Figure 2-2: Positive clones from Gibson Assembly screening on a DNA gel. This specific gel pertains to the fusion $OleT_{JE}$-EAAAK-SOR. **A)** Positive Clones are indicated by a band shift such as in transformant #4. (L=DNA ladder, C= control without linker addition). Unsuccessful transformants will not have a band shift. **B)** Positive $OleT_{JE}$ –EAAAK-SOR transformant #4 was digested with BaeI. Lack of digestion by BaeI means the removal of the BaeI cut site, likely indicating that it was replaced with the desired linker. Positive clones were confirmed by DNA sequencing. (D=digested sample, U= undigested control, Different intensities likely due to loading error).

Results from traditional restriction digest using BGlII and BamHI produced more transformants with an insert. However, a downside of this cloning method is that the inserts would often ligate in reverse conformation and/or adding too many iterations of the linkers giving a longer fusion than intended. Transformants that were "positive" for the addition of a single linker were isolated and sent for sequencing to confirm the addition of a single linker iteration and correct linker orientation (**Figure 2-3**). This process proved to be better and

produced the remaining fusions: Olet$_{JE}$-HRV3C-SOR, OleT$_{JE}$-(GGGGS)$_2$-SOR, OleT$_{JE}$-PAPAP-SOR, OleT$_{JE}$-(PK)$_7$-SOR, and OleT$_{JE}$-Nonhomolog-SOR. The plasmid constructs generated as part of this study, and the experiments each construct was used for are shown in **Table 2-3.**

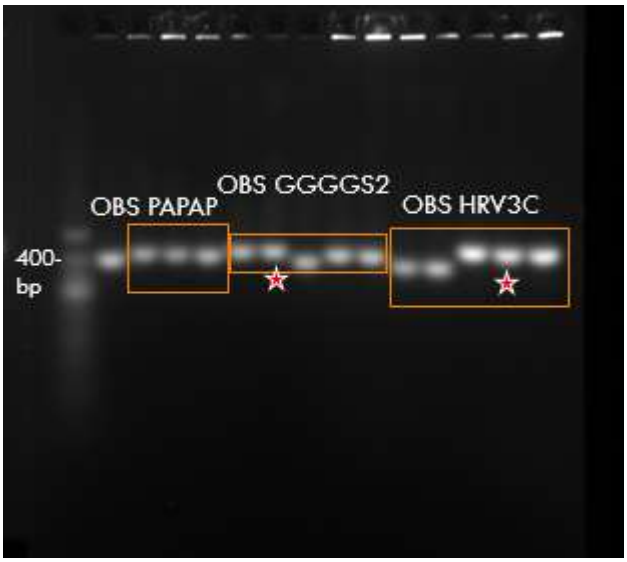


Figure 2-3: OleT$_{JE}$-SOR fusion constructs produced by restriction digest cloning. Positive clones are indicated by a band shift. Star indicates samples sent and confirmed by sequencing to be in the correct orientation.

Table 2-3. Plasmid constructs used in this study

ENZYME/LINKER NAME	CONSTRUCT	PLASMIDS	*E. COLI* EXPRESSION	EXPERIMENTS
OLET$_{JE}$	OleT$_{JE}$	pBadHisA	NC906	Growth rescue
		pET28a	Arctic Express (DE3)	SOR, OleT$_{JE}$ and methyl viologen assays
SOR	SOR	pBadHisA	NC906	Growth rescue
		pET28a	BL-21 (DE3)	SOR assays
BAE1	OleT$_{JE}$-BAE1-SOR	pBadHisA	NC906	Growth rescue
		pET28a	BL-21 (DE3)	SOR, OleT$_{JE}$, and methyl viologen assays
HRV3C	OleT$_{JE}$-HRV3C-SOR	pBadHisA	NC906	Growth rescue
		pET28a	BL-21 (DE3)	SOR, OleT$_{JE}$, and methyl viologen assays
GGGGS2	OleT$_{JE}$-(GGGGS)$_2$-SOR	pBadHisA	NC906	Growth rescue
		pET28a	BL-21 (DE3)	SOR, OleT$_{JE}$, and methyl viologen assays
EAAAK	OleT$_{JE}$-EAAAK-SOR	pBadHisA	NC906	Growth rescue
		pET28a	BL-21 (DE3)	SOR, OleT$_{JE}$, and methyl viologen assays

2.3.2 Activity screening of the OleT$_{JE}$-SOR fusion cell-free extracts using GC/FID analysis

Initially, several fusion constructs with different linkers were made and over-expressed in BL21(DE3) cells to produce cell-free extracts. **Figure 2-4** shows the amount of decarboxylated fatty acid product formed by the OleT$_{JE}$-SOR fusion enzymes in the cell-free extracts. 1-Tridecene product formation was determined for each OleT$_{JE}$-SOR fusion in order to confirm OleT$_{JE}$ activity and to inform the selection of the enzyme to over-express at large-scale for enzymatic and structural analysis. Initially, the four selected were OleT$_{JE}$-EAAAK-SOR, OleT$_{JE}$-(GGGG)$_2$-SOR, OleT$_{JE}$-NonHomolog-SORand OleT$_{JE}$-HRV3C-SOR along with OleT$_{JE}$ and SOR. However, at the large-scale expression stage, ONHS did not produce sufficient protein quantities for purification and characterization and was ultimately not included in further characterization.

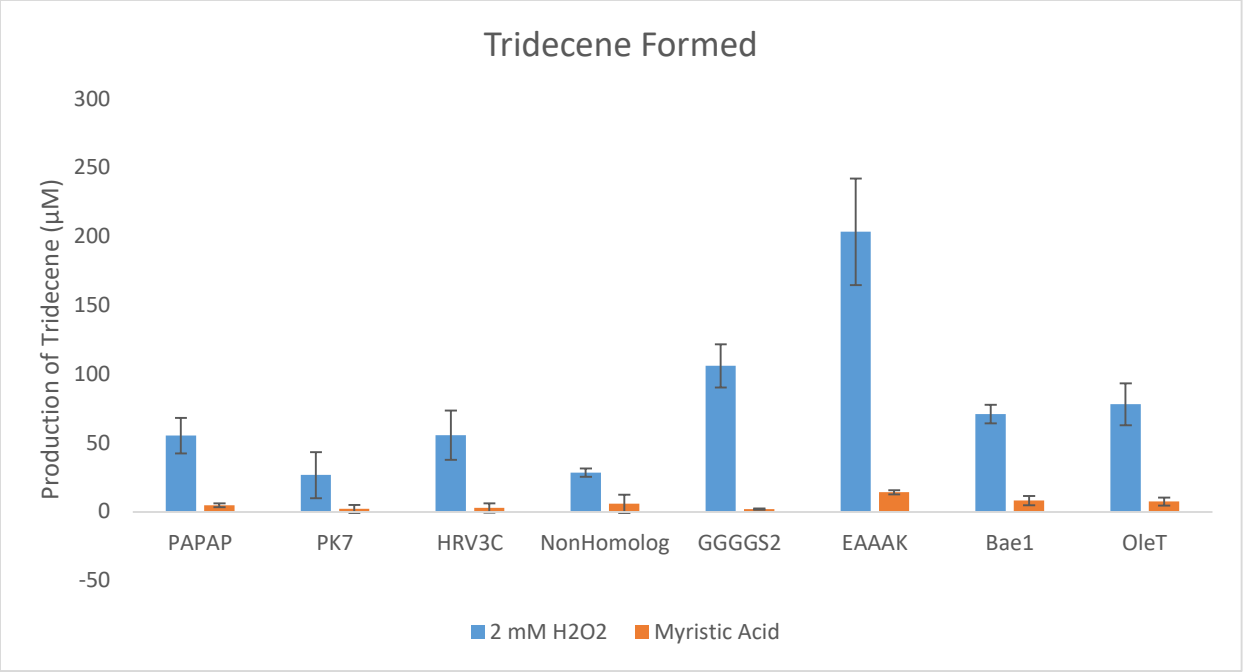


Figure 2-4: 1-Tridecene formation from the different fusion constructs, $OleT_{JE}$-BaeI-SOR and $OleT_{JE}$ controls. Letters between O and S indicate the amino acid linker used in the fusion. PAP=(PAPAP), PK7=$(PK)_7$, H=(HRV3C), NH=(GSAGSAAGSGEF), G=$(GGGGS)_2$, E=$(EAAAK)_2$. $OleT_{JE}$-EAAAK-SOR, $OleT_{JE}$-$(GGGG)_2$-SOR, and $OleT_{JE}$-HRV3C-SOR were selected as the best fusions. 40 µM of Myristic Acid used in the reaction. Error bars indicate SD (n=3).

2.3.3 $OleT_{JE}$-SOR fusion purification and activity analysis

The $OleT_{JE}$ –EAAAK-SOR, $OleT_{JE}$ –$(GGGGS)_2$-SOR, and $OleT_{JE}$-HRV3C-SOR fusion proteins were chosen for "large-scale" expression (3 L) and purification. **Figure 2.5** shows the PAGE analysis of representative purifications for each of the $OleT_{JE}$-SOR fusion enzymes.

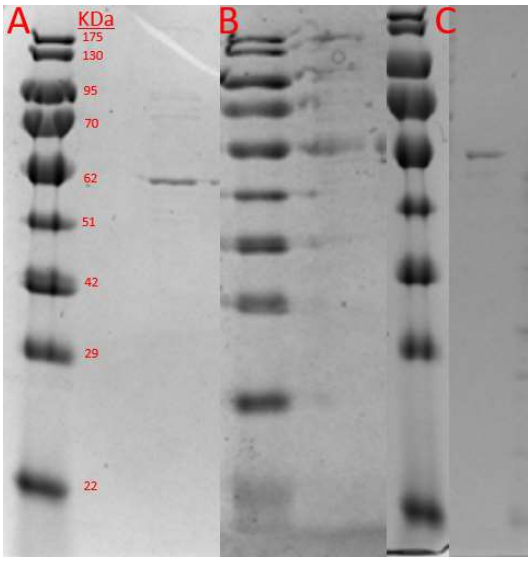


Figure 2-5: PAGE analysis of protein fusions. (A) $OleT_{JE}$-EAAAK-SOR, (B) $OleT_{JE}$-$(GGGGS)_2$-SOR and (C) $OleT_{JE}$-HRV3C-SOR post-size exclusion chromatography (SEC) purification using large scale expression (3 L). The expected band size is 65.2 KDa for $OleT_{JE}$ – EAAAK-SOR, 63.1 KDa for $OleT_{JE}$-$(GGGGS)_2$-SOR and 64.7 KDa for $OleT_{JE}$-HRV3C-SOR. Visualized using coomassie blue stain.

GC/FID analysis of the purified enzymes revealed that the fusion $OleT_{JE}$-$(GGGGS)_2$-SOR is the only one with activities comparable to $OleT_{JE}$. As indicated in **Figure 2-6**, $OleT_{JE}$ produced the most 1-tridecene per mole of enzyme when in reaction with hydrogen peroxide when considering the decarboxylase activity expressed as the total turnover number (TTN), which is calculated as moles of product over moles of enzyme and provides an indication of the number of times an enzyme is able to turnover a substrate before it loses activity. There was no statistically significant change between the TTN of $OleT_{JE}$-$(GGGGS)_2$-SOR and $OleT_{JE}$, however, the rest of the fusions demonstrated lower TTNs. **Table 2-4** shows the alcohol product

percentage of the total turnover number for each fusion. No significant difference could be observed between any of the fusions when compared to $OleT_{JE}$.

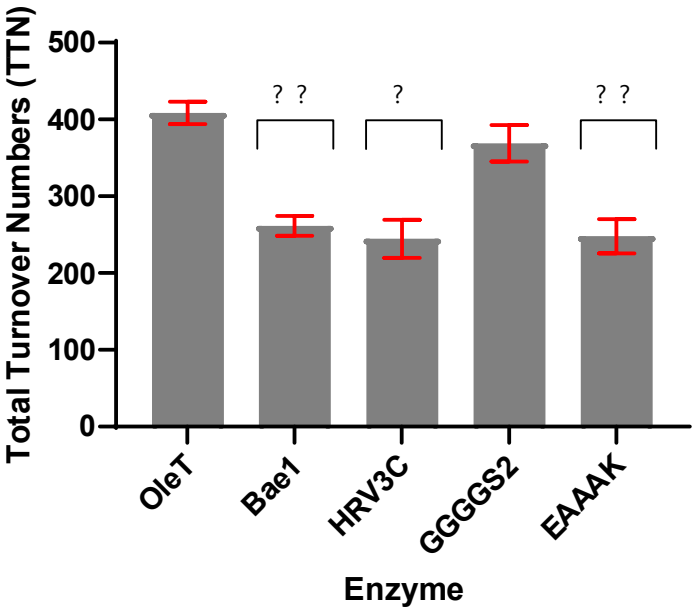


Figure 2-6: Specific activity of $OleT_{JE}$ and fusions with myristic acid. $OleT_{JE}$ and fusions were reacted with myristic acid (40 µM) in the presence of 500 µM hydrogen peroxide. A comparison of the total turnover numbers (TTN) of each enzyme is presented (n=3 replicates for each sample). * indicates $p < 0.05$, and ** indicates $p < 0.01$ when calculated TTNs were analyzed through one-way Brown-Forsythe and Welch ANOVA tests with Dunnets T3 multiple comparisons test.

Table 2-4: Percentage of TTN that was alcohol product over alkene product generated by each fusion. No statistical significance was found among percent alcohol produced by the $OleT_{JE}$-SOR fusion enzymes compared to $OleT_{JE}$.

Enzyme/Fusion	% Alcohol Product
$OleT_{JE}$	34.6 ± 0.63%
$OleT_{JE}$-Bae1-SOR	32.8 ± 0.08%
$OleT_{JE}$-HRV3C-SOR	35.1 ± 3.16%
$OleT_{JE}(GGGGS)_2$-SOR	34.2 ± 0.13%
$OleT_{JE}$-EAAAK-SOR	33.2 ± 0.43%

Next, the $OleT_{JE}$-SOR fusion enzymes activities were compared to $OleT_{JE}$ with the addition of varying molar ratios of SOR in cell free extract (CFE) and with methyl viologen (MV). In the presence of CFE, MV produces superoxide. Essentially, most fusions underperformed when compared to $OleT_{JE}$ (**Figure 2-7**). The only fusion with a comparable trend is once again $OleT_{JE}$-$(GGGGS)_2$-SOR. Similar activity levels (measured as TTN) are observed for $OleT_{JE}$-$(GGGGS)_2$-SOR compared to $OleT_{JE}$, while all the other fusions have lower TTNs. It is worth noting that the $OleT_{JE}$ –EAAAK-SOR fusion is the only one that showed improvement when in the presence of MV and higher SOR molar ratios.

The direct comparison between the inclusion and exclusion of MV is shown in **Figure 2-8**. Again, it is observed that for all samples, having MV did not aid in increasing TTNs except for the $OleT_{JE}$ –EAAAK-SOR fusion. For $OleT_{JE}$ reactions supplemented with 1:5 and 1:10 SOR, there is even a significant drop-in activity in the presence of MV. In addition, the activity of $OleT_{JE,}$ and the fusions responds to different amounts of SOR with and without MV, **(Figure 2-9)** indicating that increasing amounts of SOR has an overall increase in 1-Tridecene formation without MV and the opposite effect when MV is present. The main instance where this not true is

with the fusion OleTJE-EAAAK-SOR and that may be due to the fact that this reaction was optimized using that fusion.

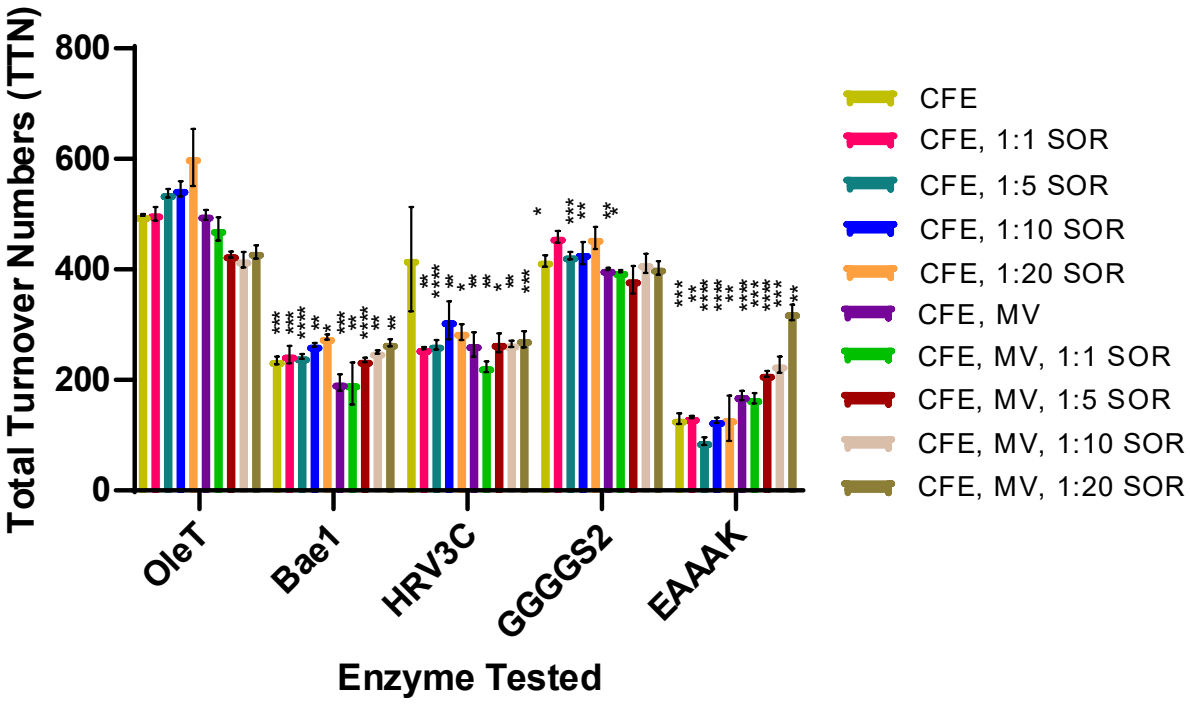


Figure 2-7: $OleT_{JE}$ activity in the presence of NC906 CFE and methyl viologen. Anaerobic NC906 cell-free extract (CFE) and methyl viologen (MV) were used in the standard $OleT_{JE}$ reaction in place of hydrogen peroxide. Activity without MV and in the presence of greater molar ratios of SOR is also shown. Statistical analysis was performed via one-way Brown-Forsythe and Welch ANOVA tests, and Dunnett's T3 multiple comparisons test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and ****$p < 0.0001$ when compared with $OleT_{JE}$ under the same conditions.

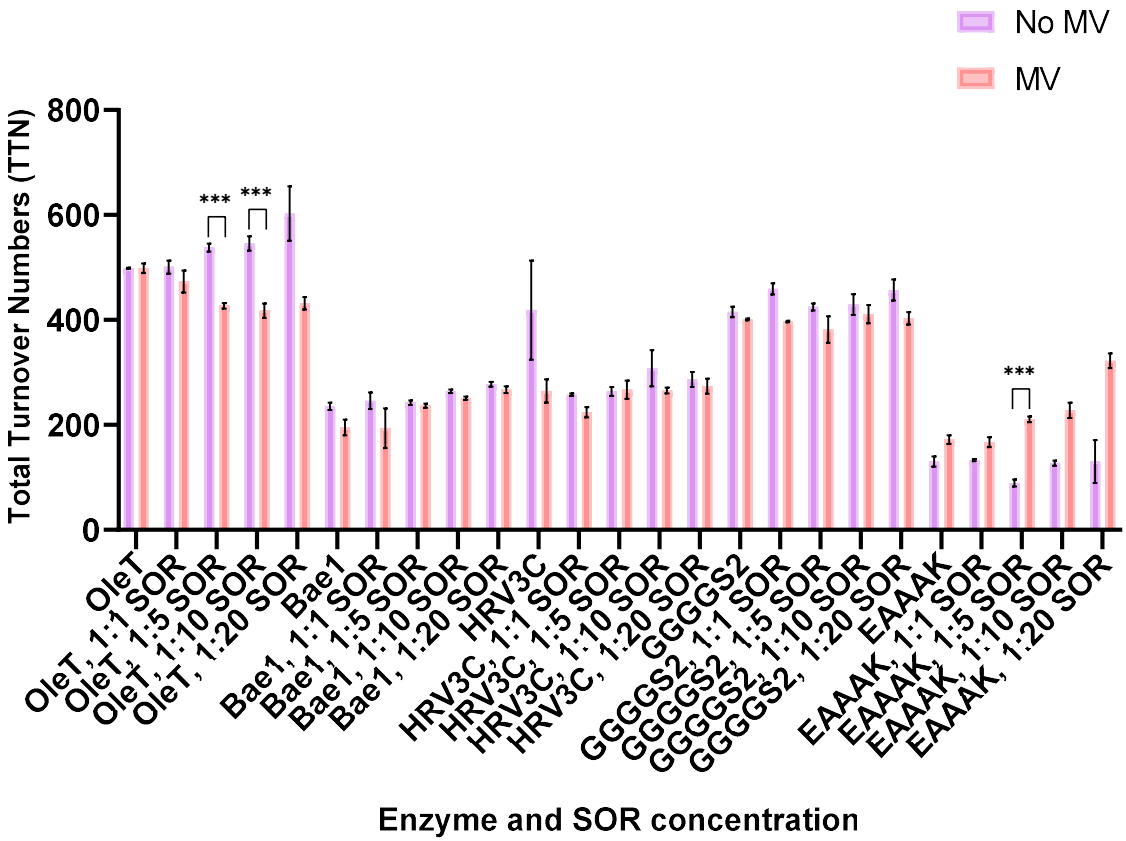


Figure 2-8: Comparison of each decarboxylase reaction with and without methyl viologen. Enzymes in the same reaction condition are plotted with their TTNs both in the presence of and without methyl viologen. *** indicates a $p < 0.05$ when comparing no MV vs MV. Using the Welch multiple t-tests performed with Holm-Sídák correction.

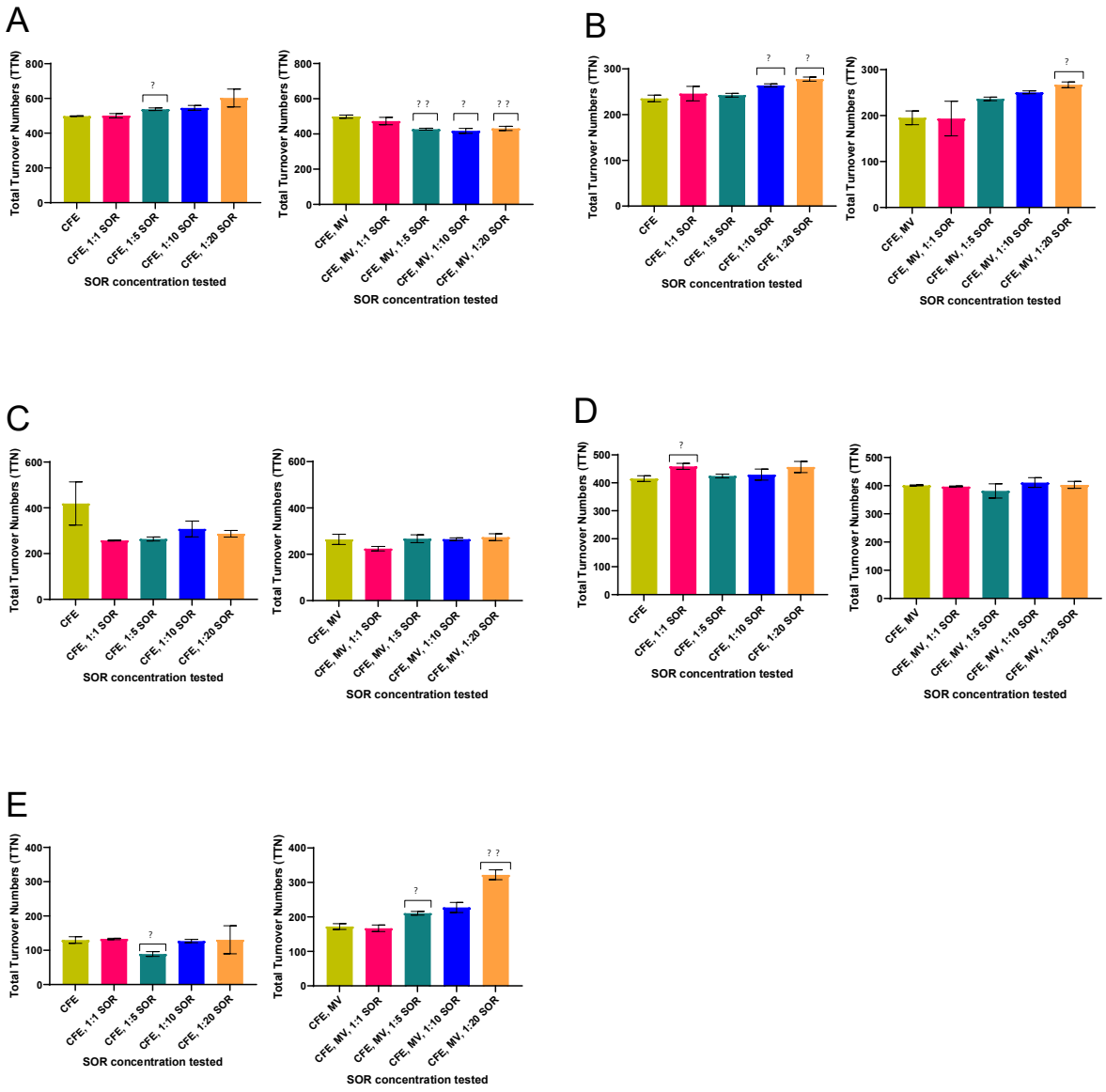


Figure 2-9: Analysis of TTN values with molar excess amounts of SOR. TTN of enzymes with or without the presence of methyl viologen in different molar excess amounts of SOR were compared to $OleT_{JE}$ or fusion with no SOR excess. $OleT_{JE}$ is panel A, Bae1 is panel B, HRV3C is panel C, GGGGS2 is panel D, and EAAAK is panel E. * indicates a $p < 0.05$ and ** indicates a $p < 0.01$. TTNs were compared using a one-way Brown-Forsythe and Welch ANOVA with Dunnett's T3 multiple comparisons test.

2.3.4 Growth recovery assays to assess SOR function for the OleT$_{JE}$-SOR fusions

Growth recovery assays were performed for a wild type *E. coli* strain (NC905) and a mutant strain that has its manganese (*sodA*) and iron (*sodB*) superoxide dismutase genes deleted to determine whether SOR in the OleT$_{JE}$-SOR fusions can functionally complement the superoxide dismutase (SOD) defects in NC906. The *E. coli*/pBad NC905 control expresses SOD which protects iron-containing branched-chain amino acid synthesis enzymes, which are very sensitive to oxidative stress. Thus, it was expected the NC905 strain would grow well aerobically in glucose minimal media that was not supplemented with amino acids, whereas it was expected that the SOD mutant strain, NC906, would experience slower, less robust growth. It has previously been documented that *P. furiosus* SOR can complement for the SOD defect in *E. coli* NC906 (Ji, 2007). Therefore, the growth rescue assay was performed to evaluate whether the SOR expressed in the different OleT$_{JE}$-SOR fusions can function in place of SOD to protect the sensitive amino acid synthesis enzymes. As anticipated, fusions containing SOR were capable of growth recovery, and NC906 transformed with either the pBAD-OleT$_{JE}$ expression vector or the empty vector pBAD control exhibited low growth compared to the wildtype NC905 strain or NC906 transformed with the pBAD-SOR or OleT$_{JE}$-SOR plasmids (**Figure 2-10**).

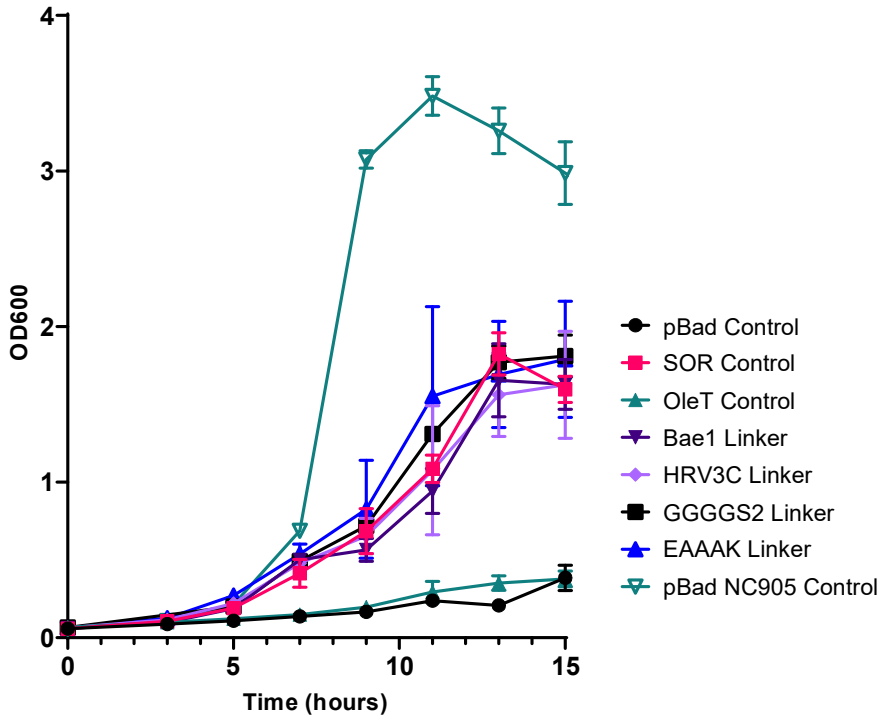


Figure 2-10 *E. coli* NC 906 growth rescue when SOR or the $OleT_{JE}$-SOR fusions are expressed. The optical density at 600 nm is recorded for each NC905 or NC906 transformant over the course of 15 h. Cells were grown in triplicate, and the error bars represent the standard deviation of triplicates.

2.3.5 *In vitro* SOR competition assay

The SOR competition assay with cytochrome c revealed interesting results that are not necessarily consistent with the *in vivo* growth rescue assay described in section 2.3.4. *In vitro* SOR activity was only observed for the $OleT_{JE}$-HRV3C-SOR fusion and not observed with the $OleT_{JE}$ –EAAAK-SOR and $OleT_{JE}$ -$(GGGGS)_2$-SOR (**Figure 2-11)**.

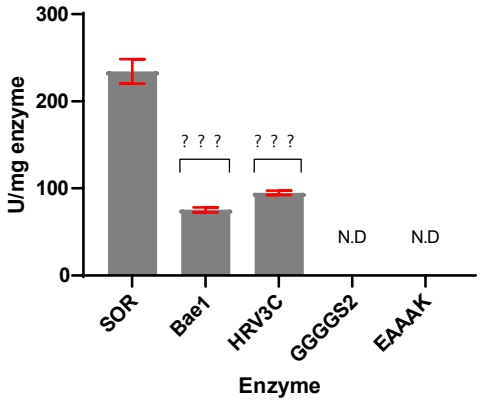


Figure 2-11: Specific activity (U/mg) of SOR and OleT$_{JE}$-SOR fusions. The specific activity of SOR and SOR in the fusions was compared using a competition assay with cytochrome c. A higher U/mg implies that a smaller concentration of enzyme was needed to decrease the rate of cytochrome c reduction by half. *** indicates a p value < 0.001 when fusions are compared with SOR using a one-way Brown-Forsythe and Welch ANOVA tests with multiple comparisons. Error bars are for SE, n=3).

2.3.6 Analytical Ultracentrifugation analysis of OleT$_{JE}$, SOR, and the OleT$_{JE}$-SOR fusions

AUC sedimentation analysis was done using different processing software such as Sedfit, which aids in modeling the different data that was collected. As the centrifuge spins, snapshots of each different sample at the 280 nm wavelength are recorded throughout the run. Each piece of data collected has information on the trajectory of the sedimentation and is fit by generating a graph that plots rate of sedimentation (Cole et al., 2008). The data is then fit to the molecular weight amongst other factors such as temperature to give an approximate frictional ratio. This analysis provides a general idea of a potential frictional ratio for each of the OleT$_{JE}$-SOR fusions. **Figure 2-12** showcases an example of the data that is generated using Sedfit, in this case for the

fusion $OleT_{JE}$ –EAAAK-SOR. **Table 2-5** is a compilation of the data resulting from the Sedfit analysis output. The enzyme $OleT_{JE}$, $OleT_{JE}$ –EAAAK-SOR, and $OleT_{JE}$-HRV3C-SOR fusions appear to exist as monomers in solution, while the $OleT_{JE}$-$(GGGGS)_2$-SOR and $OleT_{\mathbf{JE}}$ -Bae1-SOR fusions seem to be dimers. **Figure 2-13** provides a rough estimate of what the globular structure may look like given the different frictional ratios. A noteworthy finding is that $OleT_{JE}$-$(GGGGS)_2$-SOR has been behaving most like $OleT_{JE}$ in other assays in terms of activity so we are hoping AUC data would give us structural insight as to why. Sedimentation Velocity (SV) data suggests that $OleT_{JE}$-$(GGGGS)_2$-SOR is a mixture of a dimer and monomer, and this is further verified by sedimentation equilibrium (SE) analysis (**Figure 2-14)**. Sedimentation equilibrium (SE), which can give accurate information about molecular weight, and the collected data suggests that $OleT_{JE}$-$(GGGGS)_2$-SOR is a mixture of a monomer and dimer approximately in a ratio of 2:3.

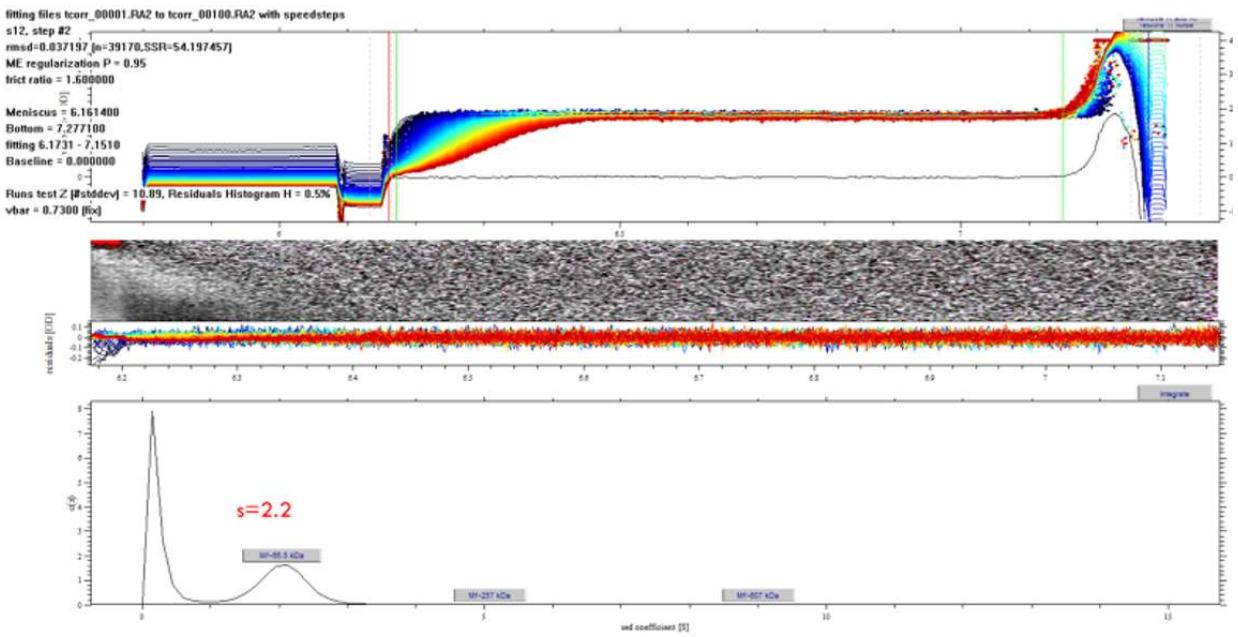
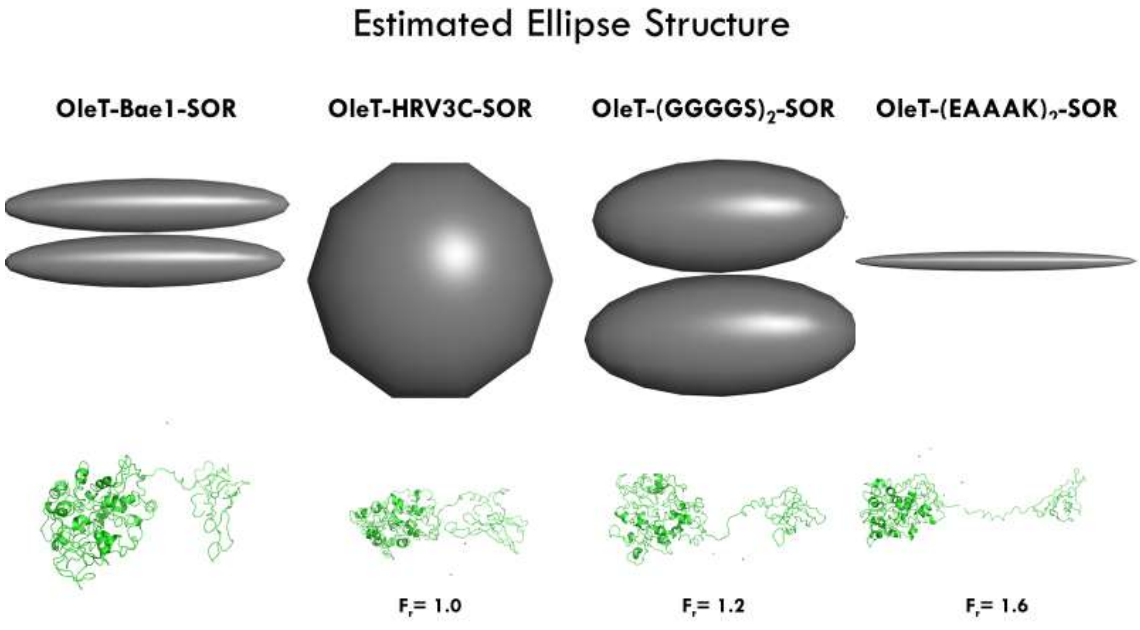


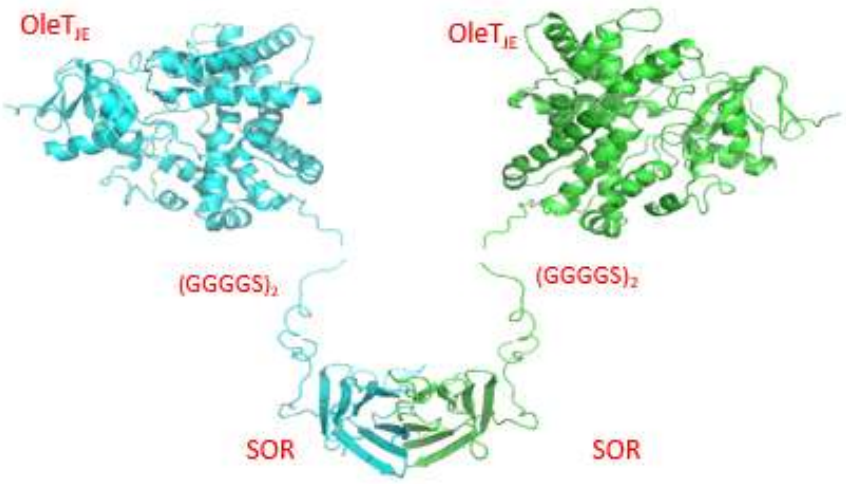
Figure 2-12: Sedimentation Velocity data using Sedfit (version 2.0) for fusion $OleT_{JE}$ – EAAAK-SOR to determine the approximate frictional ratio and thus give insight into the relative shape of the fusion structure. Sedimentation coefficient (s)= 2.2 in correspondence with M= 65 KDa suggests this sample is a monomer.

Table 2-5: Compilation of sedimentation velocity data for the $OleT_{JE}$-SOR fusions and $OleT_{JE}$.

Fusion Name	F_r	Sedimentation Coefficient	Calculated MW kDa	Shape (prolate oblate, sphere)	Oligomerization State
$OleT_{JE}$	1.000	3.7	48.4	Sphere	Monomer
$OleT_{JE}$-HRV3C-SOR	1.0	3.1	64.7	Sphere	Monomer
$OleT_{JE}$-$(GGGGS)_2$-SOR	1.2	3.9	63.2 126.4	Prolate	Monomer Dimer mix
$OleT_{JE}$-EAAAK-SOR	1.6	2.2	65.3	Prolate	Monomer
$OleT_{JE}$-Bae1-SOR	1.23	4.1	63.1 126.1	Prolate	Monomer Dimer mix



A)



B)

Figure 2-13: Estimated ellipsoid structure based on preliminary AUC data for the $OleT_{JE}$-SOR fusions. **Panel A** shows a compilation of the proposed globular and cartoon structures for the fusions. **Panel B** showcases a proposed ribbon structure for the dimer formed by the $OleT_{JE}$-$(GGGGS)_2$-SOR fusion. Figures generated using Pymol (version1.7.0.5).

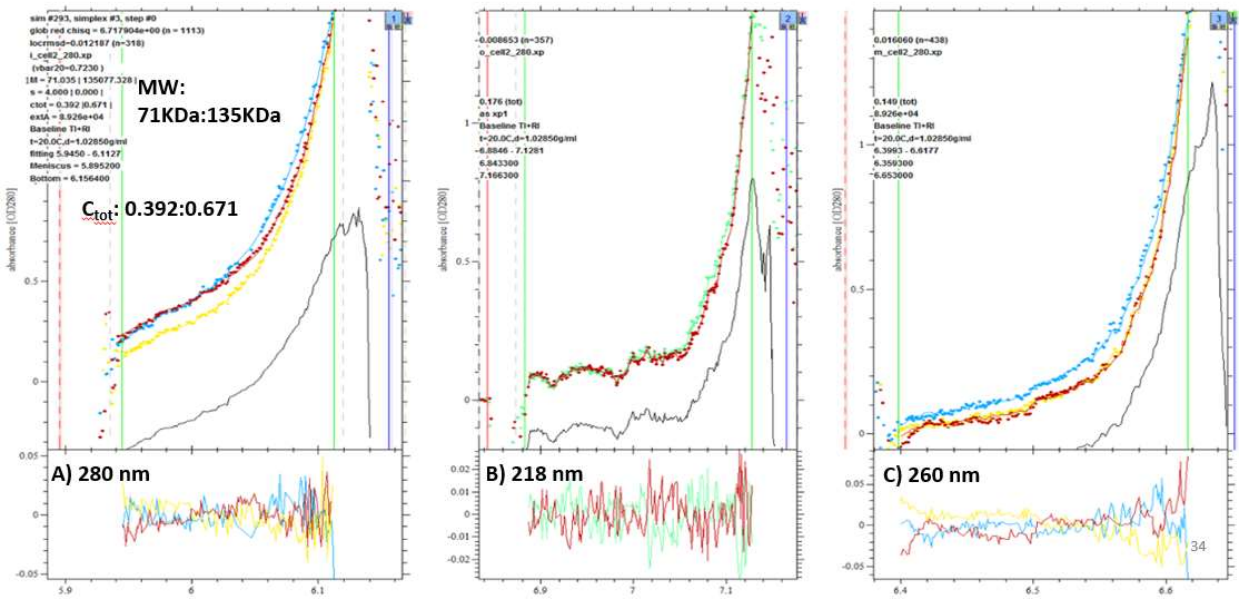


Figure 2-14: SE data for fusion $OleT_{JE}$-$(GGGGS)_2$-SOR suggesting it is a mixture of monomer and dimer. Data collected at three different concentrations of $OleT_{JE}$-$(GGGGS)_2$-SOR (1 mg/ml, 0.5 mg/ml, and 0.25 mg/ml) indicated by the different colored lines. Line color is irrelevant, the figure highlights that the sample reached equilibrium at 3 these three concentrations and at three different wavelengths (280, 218, 260 mn) as shown in panel A, B, and C, respectively. Approximate monomer to dimer ratio is simplified to 2:3.

2.3.7 Small Angle X-ray Scattering (SAXS) analysis of $OleT_{JE}$, SOR, and the $OleT_{JE}$-SOR fusions

SAXS is an experimental technique for structural characterization of macromolecules in solution which enables time-resolved analysis of conformational changes under physiological conditions. SAXS analysis provides low resolution scattering intensities which requires the integration of SAXS data into molecular dynamics simulations to generate computationally efficient native structure models (Mertens et al., 2010, & Weiel et al., 2019). In general, the overall parameters of proteins in solution such as R_g, volume, molecular mass, and folding state

can all be directly computed from SAXS data (Kikhney et al., 2015). Radius of gyration values (R_g) is a measure of the radius of the molecule and gives reference to a potential size. Models can then be generated and fit using the different R_g values that SAXS provides.

SAXS data for this study were collected at the Shared Materials Instrumentation Facility (SMIF) facility at Duke University by R&D Engineer, Justin Gladman and analyzed in collaboration with Dr. Robert Rose and Dr. Greg Buhrman from the Molecular and Structural Biochemistry Department at North Carolina State University (Raleigh, NC). The samples provided for SAXS analysis were all 100 µl at 1 mg/ml concentrations. Based on SAXS data analysis using the Primus program, R_g values of 302 Å for $OleT_{JE}$ –EAAAK-SOR, 84.5 Å for $OleT_{JE}$ –$(GGGGS)_2$-SOR, 88.3 Å for $OleT_{JE}$-HRV3C-SOR, and 91.45 Å $OleT_{JE}$ -BaeI-SOR were calculated. These R_g values were deemed to be too large to be valid. Having R_g values that are too large makes it difficult to accurately fit the data. For this reason, it was not possible to generate what would be considered valid models for any of the $OleT_{JE}$-SOR fusions. The high R_g values could be the result of sample quality or because of signal interference from Fe present in $OleT_{JE}$ and/or SOR protein preparations. Additional SAXS analysis would be needed for $OleT_{JE}$-SOR fusions model generation.

2.4 Discussion and conclusions

In the course of this study, four $OleT_{JE}$-SOR fusions were successfully cloned and over-expressed and purified ($OleT_{JE}$ -BaeI-SOR, $OleT_{JE}$ –EAAAK-SOR, $OleT_{JE}$ –$(GGGGS)_2$-SOR, $OleT_{JE}$-HRV3C-SOR). The initial activity screening and large-scale protein expression trials were used to select the $OleT_{JE}$-SOR fusions that were biochemically and structurally characterized.

The GC/FID analysis of the different purified fusions provided interesting results. At the outset, it was not clear whether any of the fusions would perform as well as or better than $OleT_{JE}$. Results from this study revealed that only the $OleT_{JE}$-$(GGGGS)_2$-SOR fusion performed similarly to $OleT_{JE}$, whereas the other $OleT_{JE}$-SOR fusions yielded lower total turnover numbers (**Figure 2-6**). This result prompted us to try different reactions such as the addition of cell free extract which provides reduction to support SOR function, which generates superoxide for SOR catalysis as well as different $OleT_{JE}$-SOR molar ratios. GC-FID results indicated that that the MV addition to the reaction mix does not increase activity compared to the activity detected with the addition of hydrogen peroxide (**Figures 2-7** and **2-8**) except in the case of the $OleT_{JE}$ – EAAAK-SOR fusion. This makes sense because $OleT_{JE}$ –EAAAK-SOR was the fusion that was used to establish what MV concentrations with which to supplement the reactions. In hindsight, we could have determined the best MV concentration for each individual fusion and then perform the GC/FID analysis; however, for industrial applications, the use of MV is not a viable option because it is a very toxic chemical.

Structure analysis of the $OleT_{JE}$-SOR fusions was also attempted to provide a deeper understanding of why $OleT_{JE}$-$(GGGGS)_2$-SOR appeared to function comparably to $OleT_{JE}$ and why the other fusions exhibited lower activity. Although, AUC-SV data was collected for all of the purified $OleT_{JE}$-SOR fusions, due to time constraints, SE data was only collected for $OleT_{JE}$-$(GGGGS)_2$-SOR since it showed the highest activity for any of the $OleT_{JE}$-SOR fusions. The resulting AUC data suggests that $OleT_{JE}$-$(GGGGS)_2$-SOR is a mostly a dimer ($OleT_{\mathbf{JE}}$ -Bae1-SOR is also mostly a dimer-monomer mix) which is interesting because it is the only dimer amongst the amino acid fusions (**Figure 2-13** and **Table 2-5**), whereas $OleT_{JE}$ –EAAAK-SOR, and $OleT_{JE}$-HRV3C-SOR appear to be monomers. These findings suggests that perhaps SOR

provides better functionality to the fusion when it can adopt a dimer confirmation. This can be further validated by some of the GC/FID data with increasing SOR molar ratios (**Figure 2-9**). Too much SOR seems to be detrimental to the fusion construct because it may generate too much hydrogen peroxide, but adding lower amounts of SOR such as a 1:1 ratio could prove to be a solid avenue for a future direction to try with the rest of the fusions that did not originally make it to the purification stage.

In conclusion, it does not seem that making a fusion between $OleT_{JE}$ and SOR is better than adding exogenous molar ratios of SOR to $OleT_{JE}$. However, the idea of generating a superior performing fusion system is still a desirable goal. Future studies could include exploring different linker lengths particularly with the $OleT_{JE}$-$(GGGGS)_2$-SOR fusion, as it performed best among the fusions analyzed as part of this study. This line of reason is further supported by plasma-based catalysis data generated by Hannah Wapshott-Stehli of the Grunden Lab for the $OleT_{JE}$-SOR fusions where $OleT_{JE}$-$(GGGGS)_2$-SOR had higher TTN values compared to $OleT_{JE}$ when using a helium atmosphere in the plasma jet to produce reactive oxygen species to mediate the $OleT_{JE}$ reaction.

References

Amin-ul Mannan, M., Hazra, D., Karnwal, A., & Kannan, D. C. (2017). Algae as a platform for biofuel production - a sustainable perspective. Algologia, 27(3), 337–356. https://doi.org/10.15407/alg27.03.337

Barr, I., Guo, F., (2015) Pyridine Hemochromagen Assay for Determining the Concentration of Heme in Purified Protein Solutions. *Bio Protoc*. 2015; 5(18):

Belcher, J., McLean, K. J., Matthews, S., Woodward, L. S., Fisher, K., Rigby, S. E. J., Nelson, D. R., Potts, D., Baynham, M. T., Parker, D. A., Leys, D., & Munro, A. W. (2014). Structure and biochemical properties of the alkene producing cytochrome p450 OleTJE (CYP152l1) from the *jeotgalicoccus* sp. 8456 bacterium. Journal of Biological Chemistry, 289(10), 6535–6550. https://doi.org/10.1074/jbc.M113.527325

Chen, X., Zaro, J., & Shen, W.C (2013) Fusion Protein Linkers: Property, Design and Functionality. Adv Drug Deliv Rev. 2013 October 15; 65(10): 1357–1369. doi:10 Matthews, S., Tee, K. L., Rattray, N. J., McLean, K. J., Leys, D., Parker, D. A., Blankley, R. T., & Munro, A. W. (2017). Production of alkenes and novel secondary products by P450 OleTJE using novel H2O2-generating fusion protein systems. FEBS Letters, 591(5), 737–750. https://doi.org/10.1002/1873-3468.12581.1016/j.addr.2012.09.039.

Cole, J. L., Lary, J. W., Moody, J., Laue, T. M., (2008) Analytical Ultracentrifugation: Sedimentation Velocity and Sedimentation Equilibrium. *Methods Cell Biol*. 2008 ; 84: 143–179. doi:10.1016/S0091-679X(07)84006-4.

Geng, X. M., Liu, X., Ji, M., Hoffmann, W. A., Grunden, A., & Xiang, Q. Y. J. (2016). Enhancing heat tolerance of the little dogwood *Cornus canadensis* L.f. with introduction of a superoxide reductase gene from the hyperthermophilic archaeon *Pyrococcus furiosus*. Frontiers in Plant Science, 7(JAN2016), 1–8. https://doi.org/10.3389/fpls.2016.00026

Grunden, A. M., Jenney, F. E., Ma, K., Ji, M., Weinberg, M. V., & Adams, M. W. W. (2005). In vitro reconstitution of an NADPH-dependent superoxide reduction pathway from *Pyrococcus furiosus*. Applied and Environmental Microbiology, 71(3), 1522–1530. https://doi.org/10.1128/AEM.71.3.1522-1530.2005

Huijbers, M. M. E., Zhang, W., Tonin, F., & Hollmann, F. (2018). Light-Driven Enzymatic Decarboxylation of Fatty Acids. Angewandte Chemie - International Edition, 57(41), 13648–13651. https://doi.org/10.1002/anie.201807119

Im, Y. J., Ji, M., Lee, A., Killens, R., Grunden, A. M., & Boss, W. F. (2009). Expression of *Pyrococcus furiosus* superoxide reductase in Arabidopsis enhances heat tolerance. Plant Physiology, 151(2), 893–904. https://doi.org/10.1104/pp.109.145409

Ji. M, (2007) *In Vitro* reconstitution and in vivo complementation studies of the *P. furiosus* superoxide detoxification system. Doctoral dissertation. North Carolina State University

Kikhney, A. G., & Svergun, D. I. (2015). A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS Letters, 589(19), 2570–2577. https://doi.org/10.1016/j.febslet.2015.08.027

Liu, Y., Wang, C., Yan, J., Zhang, W., Guan, W., Lu, X., & Li, S. (2014). Hydrogen peroxide-independent production of α-alkenes by OleT JE P450 fatty acid decarboxylase. Biotechnology for Biofuels, 7(1), 1–12. https://doi.org/10.1186/1754-6834-7-28

Matthews, S., Tee, K. L., Rattray, N. J., McLean, K. J., Leys, D., Parker, D. A., Blankley, R. T., & Munro, A. W. (2017). Production of alkenes and novel secondary products by P450 OleTJE using novel H2O2-generating fusion protein systems. FEBS Letters, 591(5), 737–750. https://doi.org/10.1002/1873-3468.12581

Mertens, H. D. T., & Svergun, D. I. (2010). Structural characterization of proteins and complexes using small-angle X-ray solution scattering. Journal of Structural Biology, 172(1), 128–141. https://doi.org/10.1016/j.jsb.2010.06.012

Nanda, S., Rana, R., Sarangi, P. K., Dalai, A. K., & Kozinski, J. A. (2018). A broad introduction to first-, second-, and third-generation biofuels. Recent Advancements in Biofuels and Bioenergy Utilization, 1–25. https://doi.org/10.1007/978-981-13-1307-3_1

Nigam, P. S., & Singh, A. (2011). Production of liquid biofuels from renewable resources. *Progress in Energy and Combustion Science*, *37*(1), 52–68. https://doi.org/10.1016/j.pecs.2010.01.003

Oneda, H., & Inouye, K. (2003). Effect of Nitration on the Activity of Bovine Erythrocyte Cu,Zn-Superoxide Dismutase (BESOD) and a Kinetic Analysis of Its Dimerization-Dissociation Reaction as Examined by Subunit Exchange between the Native and Nitrated BESODs. Journal of Biochemistry, 134(5), 683–690. https://doi.org/10.1093/jb/mvg193

Schenk, P. M., Thomas-Hall, S. R., Stephens, E., Marx, U. C., Mussgnug, J. H., Posten, C., Kruse, O., & Hankamer, B. (2008). Second Generation Biofuels: High-Efficiency Microalgae for Biodiesel Production. *BioEnergy Research*, *1*(1), 20–43. https://doi.org/10.1007/s12155-008-9008-8

Rude, M. A., Baron, T. S., Brubaker, S., Alibhai, M., Del Cardayre, S. B., & Schirmer, A. (2011). Terminal olefin (1-alkene) biosynthesis by a novel P450 fatty acid decarboxylase from Jeotgalicoccus species. Applied and Environmental Microbiology, 77(5), 1718–1727. https://doi.org/10.1128/AEM.02580-10

Wapshott-Stehli, H., (2020) Examination of OleT-SOR fusions for fatty acid decarboxylation. Unpublished doctoral dissertation. North Carolina State University.

Weiel, M., Reinartz, I., & Schug, A. (2019). Rapid interpretation of small-angle X-ray scattering data. PLoS Computational Biology, 15(3), 1–27. https://doi.org/10.1371/journal.pcbi.1006900

Zhang, W., Ma, M., Huijbers, M. M. E., Filonenko, G. A., Pidko, E. A., Van Schie, M., De Boer, S., Burek, B. O., Bloh, J. Z., Van Berkel, W. J. H., Smith, W. A., & Hollmann, F. (2019). Hydrocarbon Synthesis via Photoenzymatic Decarboxylation of Carboxylic Acids. *Journal of the American Chemical Society*, *141*(7), 3116–3120. https://doi.org/10.1021/jacs.8b12282