

**UNDERSTANDING THE FLAVOENZYME HUMAN AUGMENTER OF  
LIVER REGENERATION: BIOCHEMICAL AND STRUCTURAL  
PERSPECTIVES**

by

Stephanie Aron Ramadan

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry and Biochemistry

Fall 2013

© 2013 Stephanie Aron Ramadan  
All Rights Reserved

**UNDERSTANDING THE FLAVOENZYME HUMAN AUGMENTER OF  
LIVER REGENERATION: BIOCHEMICAL AND STRUCTURAL  
PERSPECTIVES**

by

Stephanie Aron Ramadan

Approved:

\_\_\_\_\_  
Murray Johnston, Ph.D.  
Chair of the Department of Chemistry and Biochemistry

Approved:

\_\_\_\_\_  
George H. Watson, Ph.D.  
Dean of the College of Arts and Sciences

Approved:

\_\_\_\_\_  
James G. Richards, Ph.D.  
Vice Provost for Graduate and Professional Education

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Colin Thorpe, Ph.D.  
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Brian Bahnson, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Thomas Hanson, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Sharon Rozovsky, Ph.D.  
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed:

---

Donald Watson, Ph.D.  
Member of dissertation committee

## ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisor Dr. Colin Thorpe for his advice and mentorship throughout my doctoral study.

I also had the privilege of working with Dr. Sharon Rozovsky for several projects. Thank you Sharon for your guidance, enthusiasm and mentorship on our collaborative projects.

I am very thankful for the hard work, dedication and insight of Dr. Brian Bahnson and Dr. Ming Dong for their collaborative work in solving a variety of crystal structures.

I would like to thank the following individuals: Dr. Vidyadhar Daithankar for his mentorship with the study of the R194H mutation in ALR, Renee Rubenstein and Wayne Wilkie for their help with the development of a method for the substitution of sulfur for selenium and Shawn Gannon for his preliminary work in the determination of the redox potential of ALR.

I would additionally like to thank the members of my Committee for their time and valuable input.

I would like to acknowledge the help and guidance of past and current members of the Thorpe and Rozovsky laboratories for their input into my research, especially the help of Benjamin Israel.

I would like to thank my husband, Dr. Danny Ramadan, for being an excellent mentor and teacher when I first joined the Thorpe Laboratory.

I dedicate this work to my loving husband and to my family.

## TABLE OF CONTENTS

LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
ABSTRACT .....	xv
Chapter	
1 OXIDATIVE PROTEIN FOLDING.....	1
1.1 Introduction .....	1
1.2 Disulfide Bond Formation in Gram-Negative Bacteria .....	3
1.3 Disulfide Bond Formation in Gram-Positive Bacteria .....	7
1.4 Disulfide Bond Formation in Eukaryotes.....	7
1.4.1 The ERO1/PDI Protein Folding Pathway in the ER .....	7
1.4.2 The QSOX Protein Folding Pathway in the ER .....	10
1.4.3 QSOX's Role in Cancer .....	15
1.4.4 Protein Folding in the Mitochondrial Intermembrane Space .....	15
REFERENCES .....	18
2 DISCOVERY OF AUGMENTER OF LIVER REGENERATION AND A NEW FAMILY OF FLAVIN-LINKED SULFHYDRYL OXIDASES .....	23
2.1 Introduction .....	23
2.2 Discovery of Augmenter of Liver Regeneration (ALR) .....	23
2.3 Assay of Hepatic Stimulator Substance (HSS) Activity .....	26
2.4 Sequence Determination of Augmenter of Liver Regeneration .....	27
2.5 Yeast ERV1 is a Human Homolog.....	30
2.6 Function of ERV1 and ERV2.....	32
2.7 ERV/ALR Protein Family: A Structural Comparison.....	33
2.7.1 Yeast ERV1 and ERV2 .....	36
2.7.2 Human Augmenter of Liver Regeneration .....	37
2.7.3 Plant ERV1 .....	37
2.7.4 Viral Members of the ERV Family .....	37
2.8 Summary .....	40

REFERENCES .....	41
3 AUGMENTER OF LIVER REGENERATION: IN VITRO AND IN VIVO STUDIES .....	45
3.1 Introduction .....	45
3.2 <i>In vitro</i> Characterization of ALR .....	46
3.2.1 Non-Physiological Substrates of ALR .....	49
3.2.2 Superoxide Generation by ALR .....	50
3.3 Tissue Specificity of ALR .....	50
3.4 Short-Form ALR .....	51
3.4.1 ALR is Protective Against Hydrogen Peroxide Induced Apoptosis .....	51
3.4.2 ALR is Protective Against Radiation-Induced Apoptosis .....	52
3.4.3 siRNA for ALR .....	52
3.4.4 Receptor for ALR .....	53
3.4.5 ALR Binds to JAB1 and Leads to the Activation of AP-1 .....	55
3.4.6 Treatment of Cells with Recombinant Human ALR .....	55
3.4.7 ALR's Role in Polyamine Levels .....	57
3.5 Long-Form ALR is Involved in Mitochondrial Fission and Dynamics ..	57
3.5.1 ALR is Present in Cancerous Tissues .....	60
3.5.2 Role in Organogenesis .....	60
3.5.3 ALR Interacts with Migration Inhibitory Factor (MIF) .....	60
3.5.4 NMR Studies of Long-Form ALR and Mia40 .....	61
3.6 Study of R194H Mutation in Human ALR .....	62
3.6.1 Materials and Methods .....	62
3.6.2 Primers .....	63
3.6.3 Mutation, Expression, and Purification of ALR and Arg194 Mutants .....	63
3.6.4 Crystallization and Data Collection .....	63
3.6.5 NMR Data .....	66
3.6.6 Crystal Structure, Solution and Refinement .....	68
3.6.7 In Silico Modeling .....	68
3.7 Location of R194 in Human ALR .....	69
3.8 Crystal Structure of Human Short-Form ALR .....	69

3.9	Comparison Between Wild Type and R194H Mutants by TROSY-HSQC NMR.....	78
3.10	Conclusions .....	80
REFERENCES .....		81
4	HUMAN AUGMENTER OF LIVER REGENERATION; PROBING THE CATALYTIC MECHANISM OF A FLAVIN-DEPENDENT SULFHYDRYL OXIDASE.....	88
4.1	Introduction .....	88
4.2	Materials and Methods .....	88
4.2.1	General Methods .....	89
4.2.2	Mutagenesis.....	89
4.2.3	Expression and Purification of sfALR .....	90
4.2.4	Preparation of the Apoprotein of sfALR and Reconstitution with 5-deaza-FAD .....	91
4.2.5	Redox Potential Experiments .....	91
4.2.6	Stopped-Flow Spectrophotometry.....	92
4.3	Redox Potential of the Proximal Disulfide.....	93
4.4	Overall Model for Catalysis in sfALR .....	100
4.5	Reductive Half-Reaction of sfALR Using DTT.....	102
4.6	Oxidative Half-Reaction of 2-Electron Reduced sfALR with Oxygen .....	106
4.7	Summary of Reductive and Oxidative Reaction .....	109
4.8	The Stabilization of Mixed Disulfide Intermediates in sfALR .....	109
REFERENCES .....		119
5	<sup>77</sup> SE ENRICHMENT OF PROTEINS EXPANDS THE BIOLOGICAL NMR TOOLBOX.....	124
5.1	Introduction .....	124
5.2	Materials and Methods .....	124
5.2.1	Expression of ALR and Trx .....	125
5.2.2	Purification of ALR and Trx .....	127
5.2.3	Mass Spectroscopy .....	128
5.2.4	UV-Vis Spectroscopy .....	128
5.2.5	Thermal and Chemically Induced Unfolding Transitions .....	129
5.2.6	Crystallization and Data Collection.....	129
5.2.7	Crystal Structure Solution and Refinement.....	130

5.2.8	NMR Spectroscopy .....	131
5.2.9	Accession Numbers .....	132
5.3	Method Development and Application .....	132
5.4	Method to Incorporate Selenium into Proteins.....	135
5.5	Thermal and Chemical Stability .....	144
5.6	Crystal Structure.....	146
5.7	NMR Spectroscopy .....	151
5.8	Conclusions .....	155
REFERENCES .....		157
6	SPECIFIC INCORPORATION OF SELENIUM INTO THE PROXIMAL DISULFIDE OF AUGMENTER OF LIVER REGENERATION .....	161
6.1	Introduction .....	161
6.2	Materials and Methods .....	163
6.2.1	Chemicals and Reagents.....	163
6.2.2	Design of Selenium Containing Constructs .....	163
6.2.3	Subcloning and Site-Directed Mutagenesis Primers .....	163
6.2.4	Expression of SecALR1 and SecALR2.....	164
6.2.5	Purification of SecALR1 and SecALR2.....	164
6.2.6	Expression and Purification of SecALR2 U145C .....	165
6.2.7	UV-Vis Spectroscopy .....	165
6.2.8	Turnover with DTT .....	165
6.3	Results and Discussion.....	166
6.3.1	Insertion of a SECIS Element into sfALR: Design of SecALR1 and SecALR2 Constructs .....	166
6.3.2	Subcloning.....	171
6.3.3	Expression and Purification of SecALR.....	172
6.3.4	Turnover of SecALR2 and SecALR2 U145C .....	177
6.3.5	SecALR's Reactivity with $\beta$ ME.....	177
6.4	Conclusions .....	182
REFERENCES .....		183
7	THE STUDY OF C142 MUTANTS OF AUGMENTER OF LIVER REGENERATION .....	188
7.1	Introduction .....	188

7.2	Material and Methods.....	189
7.2.1	Chemicals And Reagents.....	189
7.2.2	Mutagenesis.....	189
7.2.3	Expression and Purification of C142 Mutants.....	190
7.2.4	Crystallization and Data Collection.....	190
7.2.5	Structure Determination and Refinement.....	191
7.2.6	Structural Analyses.....	191
7.3	Results and Discussion.....	191
	REFERENCES.....	204
	Appendix	
	PERMISSION LETTERS.....	206

## LIST OF TABLES

Table 3.1 Acquisition parameters for the 2D $^1\text{H}$ - $^{15}\text{N}$ TROSY-HSQC NMR experiments .....	67
Table 3.2 Data collection and refinement statistics of human short-form ALR .....	71
Table 4.1 Redox potential and activity of the proximal disulfide in wild type and mutant sfALR proteins .....	99
Table 5.1 Sulfur-to-selenium substitution ratio and protein yield for growth media with varying ratios of sulfate and selenite .....	137
Table 5.2 X-ray data collection and refinement statistics .....	147
Table 5.3 Diselenide bond lengths and angles .....	150
Table 6.1 Yield and turnover number of wild type sfALR and SecALR2 all-cysteine variant .....	174
Table 7.1 Data collection and refinement statistics of human short-form ALR C142S .....	194
Table 7.2 Data collection and refinement statistics of human short-form ALR C142A .....	201

## LIST OF FIGURES

Figure 1.1 Selected methods that generate disulfide bonds .....	2
Figure 1.2 Disulfide bond formation and isomerization in the <i>E. coli</i> periplasm .....	5
Figure 1.3 Disulfide isomerization in the <i>E. coli</i> periplasm.....	6
Figure 1.4 The ERO1/PDI protein-folding pathway .....	9
Figure 1.5 The QSOX protein-folding pathway.....	11
Figure 1.6 Domain structure of QSOX.....	14
Figure 1.7 Flow of reducing equivalents in mammalian QSOX .....	14
Figure 1.8 Mia40 protein folding pathway.....	17
Figure 2.1 Stimulating the incorporation of tritiated thymidine into rat liver.....	25
Figure 2.2 Multiple sequence alignment of select ALR family members.....	29
Figure 2.3 Domain structures of human ALR and yeast enzymes .....	31
Figure 2.4 Comparison of non-viral ERV/ALR proteins .....	34
Figure 2.5 Comparison of viral ERV/ALR proteins with ERV2 .....	35
Figure 3.1 Disulfide connectivity for short- and long-form ALR.....	48
Figure 3.2 ALR's predicted role in liver regeneration.....	54
Figure 3.3 Overexpression of ALR results in reduced mitochondrial fission.....	59
Figure 3.4 Crystals of short-form ALR .....	65
Figure 3.5 Overall chain fold of dimeric human short-form ALR.....	72
Figure 3.6 Stereoview of the FAD prosthetic group in human short-form ALR .....	73
Figure 3.7 Overlay of rat and human short-form ALR monomers.....	74

Figure 3.8	A stereoview surrounding R194 in human short-form ALR and a minimized model of the R194H mutant.....	77
Figure 3.9	2D $^1\text{H}$ - $^{15}\text{N}$ TROSY-HSQC spectra of wild type and R194H mutant of short-form ALR .....	79
Figure 4.1	Determination of the redox potential of the proximal disulfide in sfALR substituted with 5-deaza-FAD .....	96
Figure 4.2	Schematic representation of intermediates observed during catalysis of the aerobic oxidation of DTT by sfALR and the involvement of mixed disulfide intermediates in the reductive half-reaction.....	101
Figure 4.3	Reduction of sfALR by DTT under anaerobic conditions.....	104
Figure 4.4	Rapid formation of charge-transfer intermediate when sfALR is mixed with DTT under anaerobic conditions .....	105
Figure 4.5	Oxidative half-reaction of sfALR .....	108
Figure 4.6	$\beta\text{ME}$ as a substrate of sfALR in the oxygen electrode .....	111
Figure 4.7	$\beta\text{ME}$ forms mixed disulfide intermediates with sfALR.....	112
Figure 4.8	The reductive half-reaction between sfALR and $\beta\text{ME}$ .....	113
Figure 4.9	Comparison of the nucleophilicity of DTT and $\beta\text{ME}$ towards DTNB ...	115
Figure 4.10	Spectral changes on the addition of monothiols to ALR at pH 7.5 and 9.0.....	117
Figure 5.1	Schematic representation of the two proteins used in this study .....	138
Figure 5.2	Mass spectroscopy of <i>E. coli</i> Trx as a function of selenite concentration in the growth media .....	139
Figure 5.3	Mass spectroscopy of ALR as a function of selenite concentration in the growth media.....	141
Figure 5.4	UV-Vis spectra of ALR with varying selenium composition.....	142
Figure 5.5	Reduction of selenium-substituted ALR by DTT .....	143
Figure 5.6	Thermal stabilities of native and selenium-substituted ALR.....	145

Figure 5.7 Electron density difference maps of the disulfide/diselenide and flavin in the active site of ALR refined with different ratios of selenium and sulfur occupancy.....	148
Figure 5.8 Structural characterization of selenium-rich ALR.....	149
Figure 5.9 <sup>77</sup> Se NMR spectroscopic characterization of flavin reduction in <sup>77</sup> Se-labeled ALR.....	154
Figure 6.1 Diagram of SECIS elements used in this study .....	169
Figure 6.2 Sequence comparison of constructs used in this study .....	170
Figure 6.3 UV-VIS spectrum of SecALR2 .....	176
Figure 6.4 Depiction of the charge-transfer complex between the electron-rich selenium at position 145 and the electron deficient FAD .....	177
Figure 6.5 Two possibilities of mixed disulfide formation between position 142 and DTT that results in a charge-transfer species .....	180
Figure 6.6 Formation of a charge-transfer intermediate with βME .....	181
Figure 6.7 Multiple sequence alignment of SecALR1, SecALR2 and wild type sfALR .....	186
Figure 6.8 Nucleotide multiple sequence alignment of wild type sfALR, SecALR1 and SecALR2.....	187
Figure 7.1 UV-Vis spectrum of sfALR C142S .....	193
Figure 7.2 Hydrogen bond network for two crystal structures of sfALR C142S.....	196
Figure 7.3 UV-Vis spectrum of sfALR C142A.....	198
Figure 7.4 Hydrogen bond network of sfALR C142A.....	202
Figure 7.5 Short-form ALR C142A and C142S grown in minimal media .....	203

## ABSTRACT

Augmenter of liver regeneration (ALR) is a multifaceted protein with biological roles including, but not limited to, disulfide bond formation, mitochondrial fission and fusion, spermatogenesis and activation of the MAP kinase pathway. This dissertation explores kinetic, thermodynamic and structural aspects of this protein. The first three Chapters provide background information on disulfide bond formation, the ERV/ALR family of proteins and biological roles of ALR, respectively. In Chapter 4, we determine the rate-limiting step of this flavin-dependent enzyme during the oxidation of the model substrate dithiothreitol by molecular oxygen. We also determined the redox potential of the redox-active disulfide proximal to the FAD cofactor and gained new insight into the formation and stabilization of the charge-transfer intermediate. Chapters 5 and 6 investigate the structural and functional consequence of replacing sulfur with selenium in both a random (Chapter 5) and specific (Chapter 6) manner. The final Chapter briefly describes the structure of two active site mutants of ALR, C142S and C142A. The C142S mutation crystallized with the charge-transfer interaction intact while the C142A construct was susceptible to oxidation that resulted in a cysteine sulfinic acid at position 145.

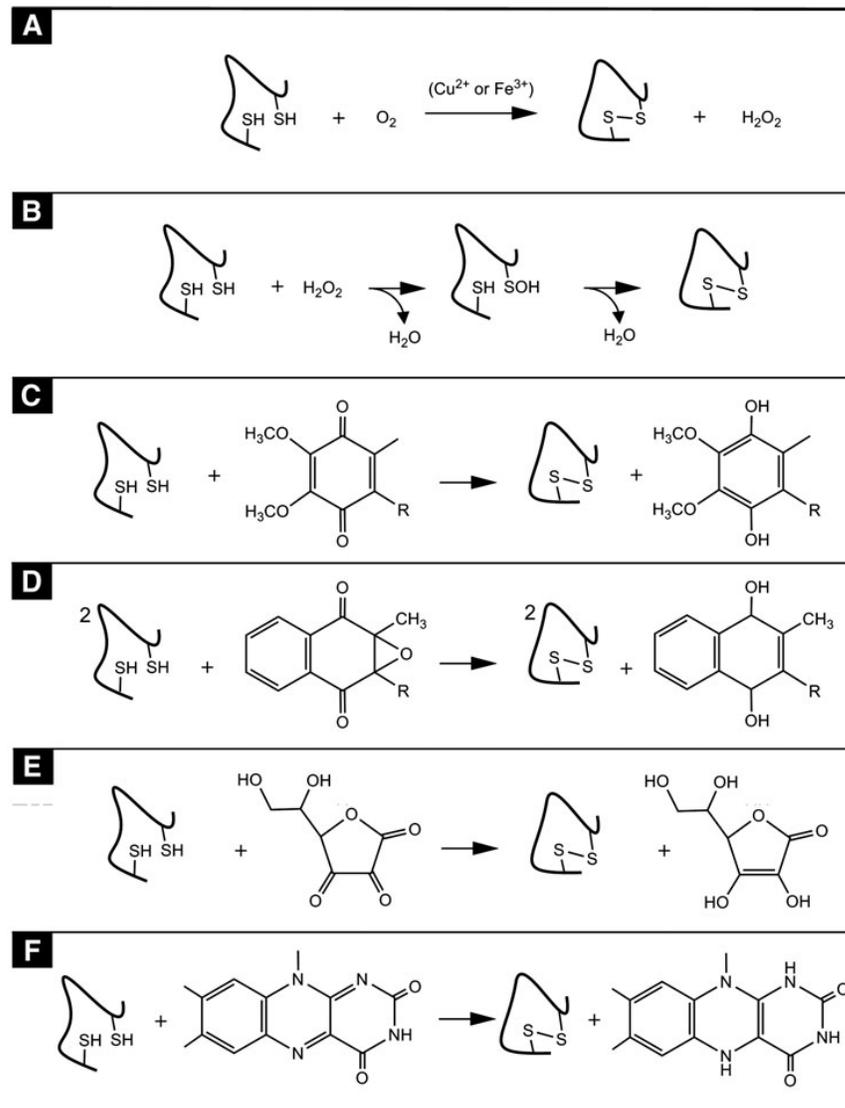
## Chapter 1

### OXIDATIVE PROTEIN FOLDING

#### 1.1 Introduction

Disulfide bonds are an important covalent post-translational modification, occurring between sulfur atoms of two adjacent cysteine residues. This process of inserting disulfide bonds is termed oxidative protein folding and is essential to insure the stability, structure and function of many proteins that are either resident in oxidizing regions of the cell or are secreted. If the incorrect disulfide pairs are retained the result is a misfolded, or aggregated, protein that will be targeted for degradation.

Some of the many routes to disulfide bonds are shown in Figure 1.1. These include A) aerobic oxidation in the presence of trace metals, B) via sulfenic acid intermediate formed on hydrogen peroxide exposure, C) using quinones in the active site of the bacterial oxidoreductase DsbB, D) in the presence of vitamin K epoxide reductase (VKOR) which utilize epoxide as the final electron acceptor, E) using dehydroascorbate, which can directly insert disulfide bonds or F) utilize a flavin adenine dinucleotide (FAD) cofactor to transfer electrons to or from a proximal redox active disulfide (Kodali and Thorpe, 2010a) (see Figure 1.1).



**Figure 1.1 Selected methods that generate disulfide bonds.** A) trace metals, B) hydrogen peroxide, C) quinones, D) vitamin K epoxide reductase (VKOR), E) dehydroascorbate and F) flavin adenine dinucleotide (FAD). Figure taken from (Kodali and Thorpe, 2010a) with permission from the publisher.

Our laboratory is specifically interested in FAD-dependent disulfide bond formation. Several proteins, which are FAD dependent sulfhydryl oxidases, will be discussed later in this Chapter, and the protein augmentor of liver regeneration, which is the focus of this dissertation, will be introduced more fully in Chapters 2 and 3.

## **1.2 Disulfide Bond Formation in Gram-Negative Bacteria**

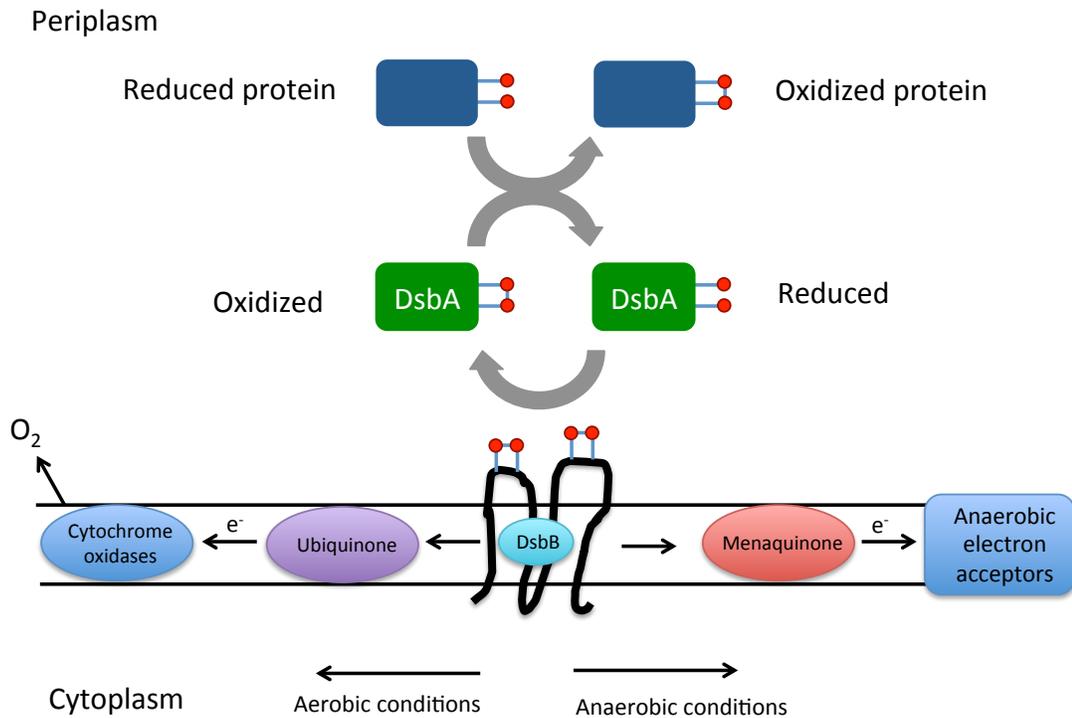
The formation of disulfide bonds in the bacterial periplasm has been well studied over the last decade. The Dsb (disulfide bond formation) family of proteins are key players in disulfide bond formation in *Escherichia coli*. There are two pairs of proteins which play separate, but critical, roles in oxidative folding. The first pair is DsbA and DsbB which work together to insert disulfide bonds into reduced client proteins (Figure 1.2) (Bardwell et al., 1991; Collet and Bardwell, 2002; Zapun et al., 1993). The second pair, DsbC and DsbD, isomerize the mispaired disulfides within the periplasm (Figures 1.3).

DsbA is a 21 kDa soluble protein with a thioredoxin-like fold, evident in the crystal structure (Martin et al., 1993). DsbA acts in the bacterial periplasm to insert disulfide bonds into reduced proteins. The active site of this protein contains a CPHC (Cys30 and Cys33) redox motif, and is one of the most oxidizing disulfide bond-containing proteins with a redox potential of -120 mV (Zapun et al., 1993). One of the reasons this protein is believed to be so oxidizing is that Cys30, which is N-terminally located and solvent exposed, has an extremely low pKa of ~3 (a typical cysteine thiol pKa is ~ 8.6). Hence Cys30 will be almost exclusively a thiolate anion at physiological pH values (Grauschopf et al., 1995). This thiolate can be stabilized by several interactions such as hydrogen bonds, but the electrostatic interactions between Cys30 and His32 are believed to be the most crucial (Guddat et al., 1997; Martin et al.,

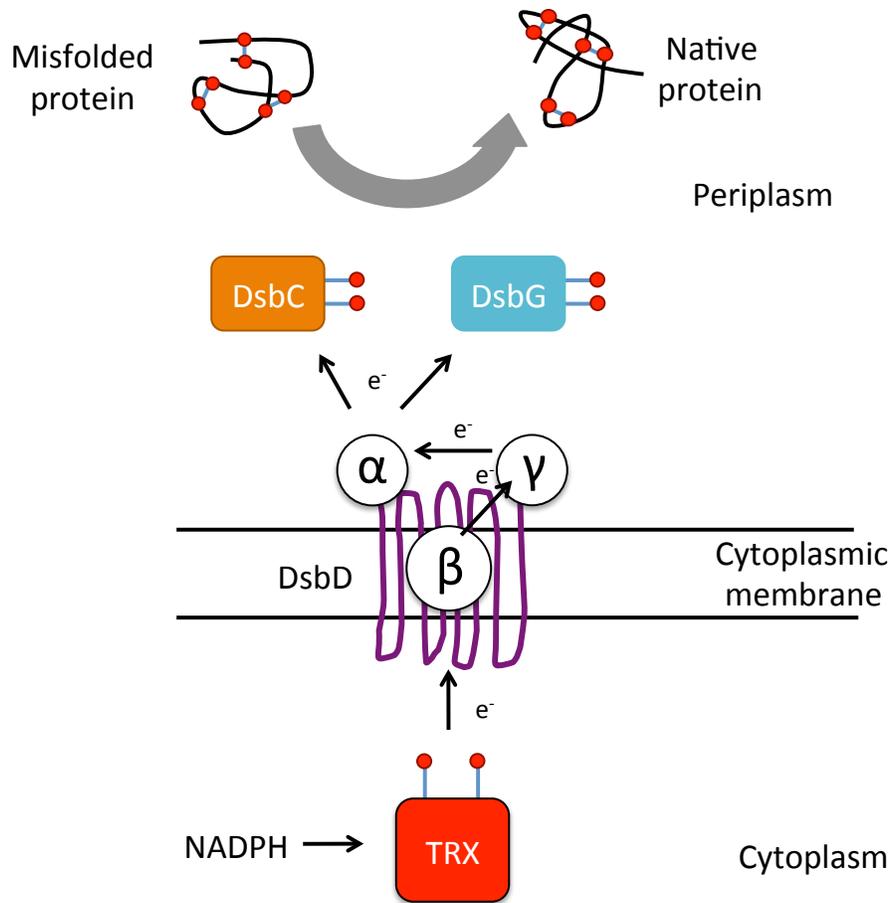
1993). As would be expected for the thermodynamic linkage between disulfide redox potentials and protein stability (Madden et al., 1984), the oxidized form of DsbA was found to be less stable than the reduced form of the protein (Wunderlich and Glockshuber, 1993).

It is the role of the 21 kDa plasma-membrane protein DsbB to regenerate DsbA with the ultimate reduction of quinone cofactors (Figure 1.2). DsbB has two periplasmic loops tethered between four transmembrane segments (Jander et al., 1994). Under anaerobic conditions, DsbB passes its electrons onto menaquinone and finally to electron acceptors such as fumarate reductase or nitrate reductase (Collet and Bardwell, 2002). Alternatively, under aerobic conditions, DsbB reduces ubiquinone with the eventual transfer to cytochrome *c* and utilization of molecular oxygen.

Since the insertion of disulfide bonds into client proteins by DsbA is error prone. A third catalyst, DsbC, acts as a disulfide isomerase to reshuffle mispaired disulfides. DsbC is a homodimer of 23 kDa subunits (Zapun et al., 1995). Each subunit contains a redox active CxxC motif (C98-C101) and one structural disulfide (Zapun et al., 1995). A crystal structure of DsbC has been solved to 1.9 Å resolution and shows that the protein appears V-shaped (McCarthy et al., 2000). It also should be mentioned that DsbG represents another protein disulfide isomerase in the bacterial periplasm. DsbG is 28% identical and 56% similar on the sequence level to DsbC (Andersen et al., 1997; Bessette et al., 1999). DsbC and DsbG are both kept in their active (reduced) forms by a 59 kDa inner-membrane protein called DsbD. DsbD, in turn, is reduced by cytosolic thioredoxin driven by NADPH and thioredoxin reductase.



**Figure 1.2 Disulfide bond formation and isomerization in the *E. coli* periplasm.** In the bacterial periplasm, DsbA oxidizes reduced protein substrates. DsbB returns DsbA to its active (oxidized) form after which electrons can proceed to ubiquinone (UQ) under aerobic conditions or menaquinone (MQ) under anaerobic conditions and finally onto molecular oxygen via the electron transport chain. The arrows indicate the flow of electrons. Figure adapted from (Collet and Bardwell, 2002).



**Figure 1.3 Disulfide isomerization in the *E. coli* periplasm.** Mismatched disulfide bonds are isomerized by DsbC and DsbG. The inner-membrane protein DsbD keeps DsbC and DsbG in their active (reduced) forms by receiving reducing equivalents from cytoplasmic reduced thioredoxin (TRX). The arrows depict the flow of electrons. Figure adapted from (Collet and Bardwell, 2002).

### **1.3 Disulfide Bond Formation in Gram-Positive Bacteria**

Unlike gram-negative bacteria, gram-positive bacteria lack a periplasmic space. Protein folding and disulfide bond formation take place in the area at the cell wall/membrane interface. Most studies of protein folding in gram-positive bacteria have utilized *Bacillus subtilis*, as a model system (Kouwen et al., 2007). Here, BsbA, BsbB, BsbC and BsbD are extracytoplasmic proteins believed to be responsible for the formation of disulfide bonds (Kouwen et al., 2007). BsbD is a homologue of DsbA and BsbB is a homologue of DsbC.

### **1.4 Disulfide Bond Formation in Eukaryotes**

Eukaryotic cells contain two relatively oxidizing compartments in which structural disulfide bonds can be introduced under normal conditions. Normally the cytosol is maintained with a highly reducing redox poise and disulfide bonds can only be introduced in localized areas devoted to oxidative folding of certain viral coat proteins (see Chapter 2). The de novo generation of disulfide bonds occurs in the endoplasmic reticulum (ER) and Golgi, as part of the cellular secretory apparatus, together with the intermembrane space of the mitochondrion. We first address the endoplasmic reticulum – the site of oxidative protein folding of perhaps 25% of all cellular proteins from higher eukaryotes. The mitochondrial system will be described later.

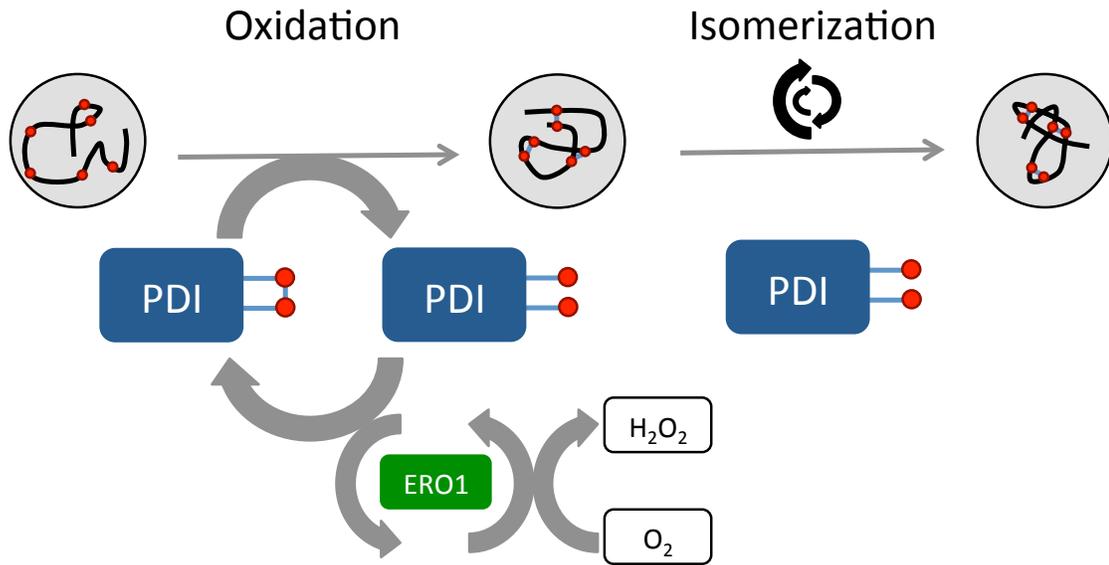
#### **1.4.1 The ERO1/PDI Protein Folding Pathway in the ER**

ERO1 (endoplasmic reticulum oxidoreductin 1) is an essential yeast glycoprotein discovered simultaneously by two research groups. Weissman and colleagues searched for genes encoded on a high-copy plasmid that would confer resistance to concentrations of DTT that would otherwise inhibit growth of *S.*

*cerevisiae* (Pollard et al., 1998). The second group screened 1200 temperature sensitive mutants in *S. cerevisiae* to find mutations that impact the export of ER secretory proteins (Frand and Kaiser, 1998). ERO1 is upregulated by the unfolded protein response pathway, which is typical of genes involved with protein folding in the endoplasmic reticulum (Frand and Kaiser, 1998; Pollard et al., 1998).

ERO1 is a 66 kDa protein containing 14 cysteine residues; 7 of these cysteines are highly conserved and most of which are disulfide bonded (Frand and Kaiser, 1998). The C-terminal domain of ERO1 is approximately 65% identical in sequence from yeast to humans (Frand and Kaiser, 1998). ERO1 and PDI play essential and distinct roles in disulfide bond formation in the ER (Frand and Kaiser, 1998; Pollard et al., 1998). It has previously been shown that PDI catalyzes the disulfide bond formation both *in vitro* and *in vivo* (LaMantia and Lennarz, 1993). After the discovery of the ER resident protein ERO1, it was shown that ERO1 and PDI interact (Frand and Kaiser, 1999, 2000). The ERO1/PDI system parallels the DsbB/DsbA pathway in the bacterial periplasm. Both PDI and DsbA act as the initial oxidant in their respective pathways, and ERO1 and DsbB restore PDI and DsbA to their active oxidized form, respectively.

In this pathway (Figure 1.4), PDI acts as both the oxidant by directly inserting the disulfide bonds into the client protein and the isomerase by reshuffling the disulfides to find the final, native conformation. The electrons from reduced PDI travel through two CxxC motifs in ERO1 to a FAD cofactor and finally to molecular oxygen as the final electron acceptor. ERO1 regenerates PDI to its active oxidized form with the net production of hydrogen peroxide.

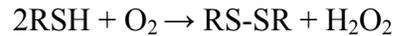


**Figure 1.4 The ERO1/PDI protein-folding pathway.** Here, reduced unfolded client proteins are oxidized by protein disulfide isomerase (PDI). Oxidized PDI is regenerated by the ER protein ERO1. Reduced PDI additionally acts as an isomerase to reshuffle disulfides to their native pairings. Figure adapted from (Kodali and Thorpe, 2010a).

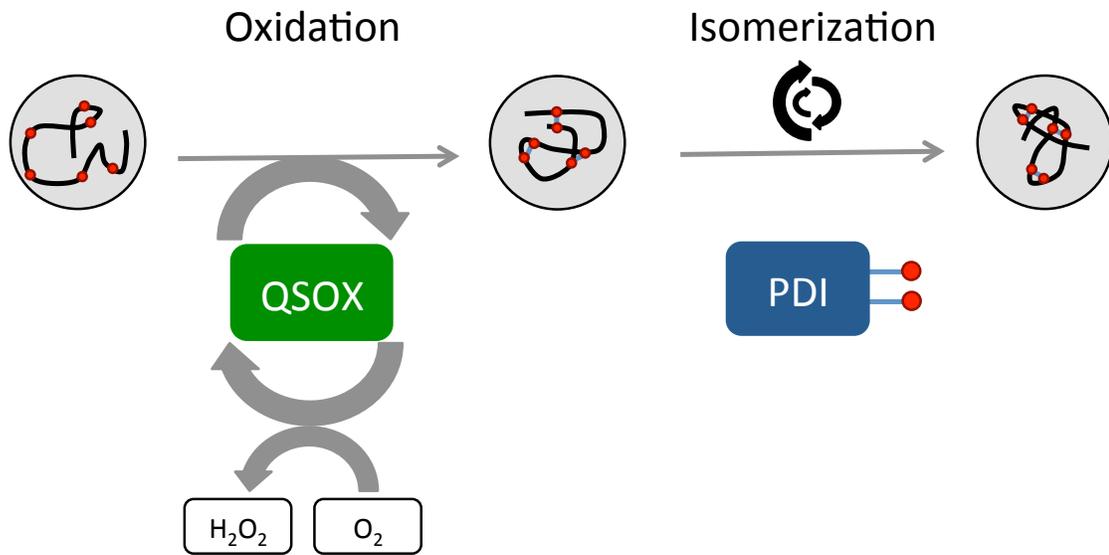
### 1.4.2 The QSOX Protein Folding Pathway in the ER

The isolation of previously unsuspected sulfhydryl oxidase activity in chicken egg white (Hooper et al., 1996) led to the discovery of the ERV/ALR family of flavoproteins detailed in Chapters 2 and 3 of this dissertation. Peptide sequencing of the egg white oxidase in 1998 showed that the closest matches were with a series of human growth factors including a protein named Quiescin Q6 (Hooper et al., 1999). This work led to the realization that Quiescin-sulfhydryl oxidase (QSOX) enzymes are widely distributed in non-fungal eukaryotes (Heckler et al., 2008a).

QSOX enzymes catalyze the following reaction:



*In vitro* the best substrates are unfolded reduced proteins and conformationally mobile peptides (Coddling et al., 2012). Unlike the ERO1 family of sulfhydryl oxidases, QSOX does not oxidize reduced PDI directly, leading to a proposed model in which the generation and isomerization of disulfides are functionally segregated (Figure 1.5). QSOX accomplishes the rapid direct oxidation of unfolded protein thiols because it is a fusion of thioredoxin and ERV/ALR domains (Figure 1.6).



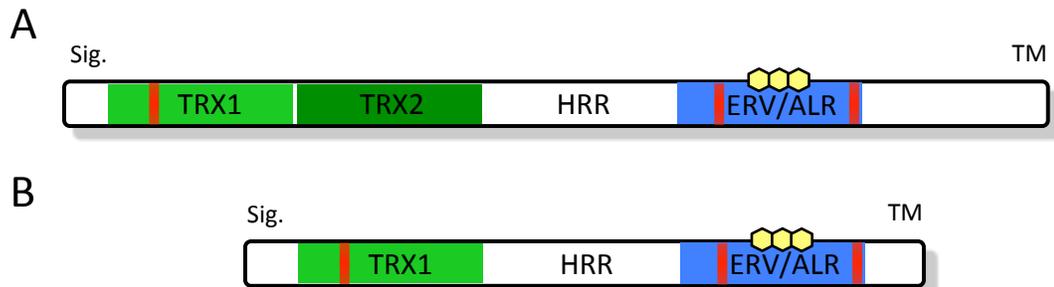
**Figure 1.5 The QSOX protein-folding pathway.** Reduced unfolded client proteins can be directly oxidized by QSOX. However, no known oxidant can correctly insert multiple disulfide bonds on the first try so a protein disulfide isomerase (PDI) is required to rearrange the oxidized protein into its native confirmation. Figure adapted from (Kodali and Thorpe, 2010a).

Humans have two isoforms of QSOX, QSOX1 and QSOX2. QSOXs have a domain structure consisting of one or two TRX domains (only one is redox-active), a helix rich region (five helix bundle similar to ERV/ALR but lacking a CxxC motif and FAD cofactor), an ERV/ALR domain and a transmembrane region (TM) (Figure 1.6). Mammalian QSOX has two TRX domains while algae, plants and protists contain only one TRX domain. The ERV/ALR domain of QSOX is also found as a stand-alone protein and this domain is homologous to the human ALR protein that will be the focus of this dissertation (see Chapter 2 for more details on the ALR/ERV family of proteins).

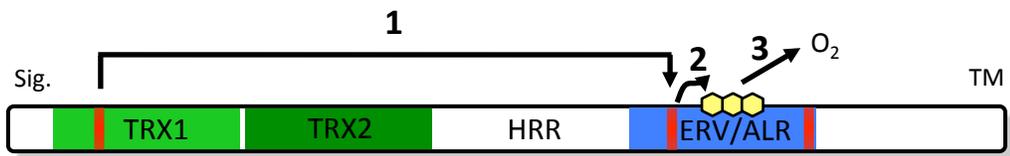
Studies of the human QSOX1 led to the proposed mechanism of the enzyme (Heckler et al., 2008b) followed by the characterization of the trypanosome QSOX mechanism (Kodali and Thorpe, 2010b) and an ongoing study of thermodynamic properties of TbQSOX by Benjamin Israel (submitted work). Figure 1.7 illustrates the flow of reducing equivalents in QSOX. Reducing equivalents are passed from the reduced unfolded substrate to the first TRX domain and are shuttled to the proximal CxxC motif in the ERV/ALR domain. Electrons then move to the FAD cofactor and on to molecular oxygen as the final electron acceptor.

QSOX is localized to the endoplasmic reticulum (Thorpe et al., 2002; Tury et al., 2004) and the Golgi (Chakravarthi et al., 2007; Mairet-Coello et al., 2004; Thorpe et al., 2002; Tury et al., 2004) and has also been found extracellularly (Amiot et al., 2004; Coppock et al., 2000; Mairet-Coello et al., 2005). The ER and Golgi are oxidizing compartments, in comparison to the cytosol, and assists in the formation of disulfide bonds.

Recently, crystal structures of QSOX from mouse (Alon et al., 2010) (PDB 3LLK) and trypanosome (PDB 3QCP and 3QD9 for the open and closed conformations, respectively) have been solved giving new insight into structural aspects of this protein (Alon et al., 2012).



**Figure 1.6 Domain structure of QSOX.** A) Domain structure of mammalian QSOX that harbors two TRX domains (only TRX1 is redox active), a helix rich region (HRR), a redox active ERV/ALR domain and finally a trans membrane region (TM) B) the QSOX domain structure for algae, plants and protists. Figure adapted from (Kodali and Thorpe, 2010a).



**Figure 1.7 Flow of reducing equivalents in mammalian QSOX.** 1) Reducing equivalents from the substrate will be transferred to the N-terminal CxxC motif in the first TRX domain. 2) These electrons will then be passed to the proximal CxxC motif in the ERV/ALR domain adjacent to the isoalloxazine ring of the FAD cofactor and 3) finally to molecular oxygen.

### **1.4.3 QSOX's Role in Cancer**

In addition to being a disulfide bond catalyst in the ER QSOX has also been implicated in various cancers. Two independent studies claim QSOX is down regulated by an increase in estrogen in breast cancer and endometrial cells (Inoue et al., 2002; Musard et al., 2001). Alternatively, QSOX has been shown to be up-regulated in pancreatic and prostate cancers (Antwi et al., 2009; Song et al., 2009) as well as play a role in pancreatic tumor cell invasion (Katchman et al., 2011). Additionally, the overexpression of QSOX has been shown to be protective against oxidative stress in a human breast cancer cell line (Morel et al., 2007). A recent paper by Lake and coworkers demonstrates QSOX also has the potential to be used as a biomarker by searching for a small peptide of QSOX in the plasma of patients (Antwi et al., 2009). Studies support QSOX is involved cancer but more studies need to be conducted to elucidate QSOX's exact role(s).

### **1.4.4 Protein Folding in the Mitochondrial Intermembrane Space**

Except for a few mitochondrially encoded proteins, almost all mitochondrial proteins are synthesized in the cytosol with signal sequences that direct proteins to the matrix, the inner and outer mitochondrial membranes or the intermembrane space (IMS) (Deponate and Hell, 2009).

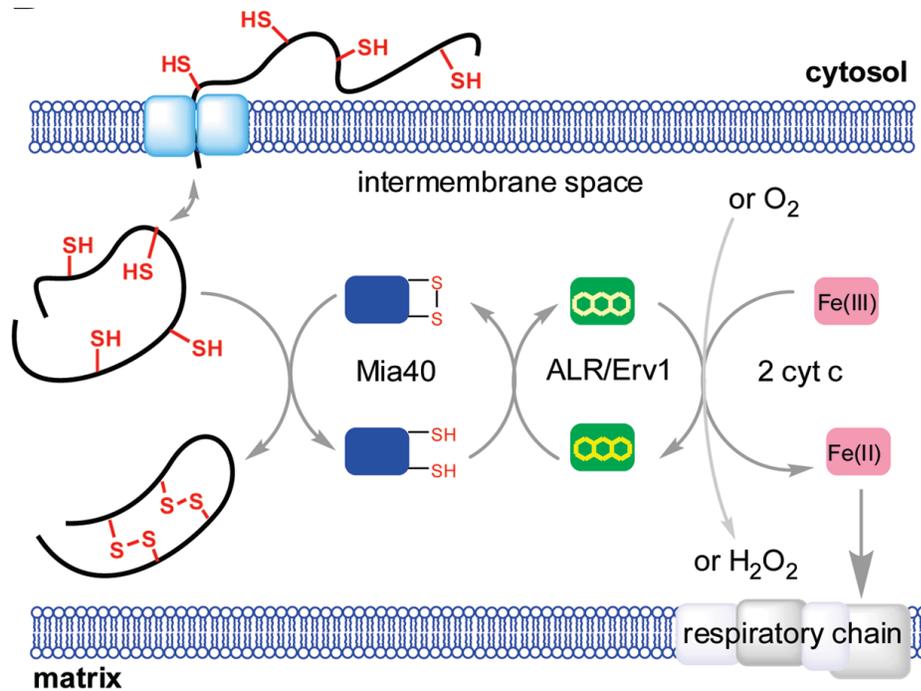
A number of proteins in the IMS contain either twin C<sub>X</sub>3C or twin C<sub>X</sub>9C motifs. Transporter of intermembrane (TIM) proteins twin C<sub>X</sub>3C motifs while the twin C<sub>X</sub>9C motifs contain proteins such as the copper chaperone Cox17 and Mia40 (discussed below). The TIM proteins act as chaperones in the transport of hydrophobic membrane proteins (Deponate and Hell, 2009). The twin C<sub>X</sub>3C and twin C<sub>X</sub>9C motif class of proteins both share a common fold involving two antiparallel  $\alpha$ -

helices which are stapled together using disulfides forming a hairpin-like structure (as demonstrated by the recently solved NMR structure of human Mia40 (Banci et al., 2009)). These disulfides are essential for the proper folding and structural stability of TIM proteins.

Following synthesis of the reduced protein in the cytosol the unfolded chains are imported by the translocase of the outer membrane (TOM) complex (Figure 1.8). Mia40 and ERV1 are two proteins involved in the folding and disulfide bond formation of proteins imported into the IMS via the TOM complex.

Once inside the IMS, these proteins encounter Mia40 and ERV1/ALR of the oxidative folding apparatus (Figure 1.8). The immediate oxidant is Mia40. Mia40 itself is a twin Cx<sub>9</sub>C protein and contains an additional CxC motif which shuttles reducing equivalents from the client proteins to the flavin-linked ERV1/ALR proteins and then to either molecular oxygen or cytochrome *c*. Mia40 is an essential protein in the IMS and has homologs in fungi, ameba, plants and animals. Mia40 is anchored to the inner membrane in yeast (Naoé et al., 2004) and is soluble in eukaryotes (Hofmann et al., 2005).

The next Chapter of this dissertation describes how ERV1 and ALR were discovered to be flavoproteins and presents background on structural characteristics of the ERV/ALR family.



**Figure 1.8 Mia40 protein folding pathway.** Reduced unfolded proteins enter through the TOM complex and are oxidized by Mia40 with the transfer of reducing equivalents to ERV1. Electrons are either passed on to molecular oxygen, to form H<sub>2</sub>O<sub>2</sub>, or to cytochrome *c* where they continue to the respiratory chain. Figure taken from (Daithankar et al., 2010) with permission from the publisher.

## REFERENCES

- Alon, A., Heckler, E.J., Thorpe, C., and Fass, D. (2010). QSOX contains a pseudo-dimer of functional and degenerate sulphhydryl oxidase domains. *FEBS Lett.* *584*, 1521–1525.
- Alon, A., Grossman, I., Gat, Y., Kodali, V.K., DiMaio, F., Mehlman, T., Haran, G., Baker, D., Thorpe, C., and Fass, D. (2012). The dynamic disulphide relay of quiescin sulphhydryl oxidase. *Nature* *488*, 414–418.
- Amiot, C., Musard, J.F., Hadjiyiassemis, M., Jouvenot, M., Fellmann, D., Risold, P.Y., and Adami, P. (2004). Expression of the secreted FAD-dependent sulphhydryl oxidase (QSOX) in the guinea pig central nervous system. *Brain Res. Mol. Brain Res.* *125*, 13–21.
- Andersen, C.L., Matthey-Dupraz, A., Missiakas, D., and Raina, S. (1997). A new *Escherichia coli* gene, *dsbG*, encodes a periplasmic protein involved in disulphide bond formation, required for recycling DsbA/DsbB and DsbC redox proteins. *Mol. Microbiol.* *26*, 121–132.
- Antwi, K., Hostetter, G., Demeure, M.J., Katchman, B.A., Decker, G.A., Ruiz, Y., Sielaff, T.D., Koep, L.J., and Lake, D.F. (2009). Analysis of the plasma peptidome from pancreas cancer patients connects a peptide in plasma to overexpression of the parent protein in tumors. *J. Proteome Res.* *8*, 4722–4731.
- Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Martinelli, M., Sideris, D.P., Katrakili, N., and Tokatlidis, K. (2009). MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nat. Struct. Mol. Biol.* *16*, 198–206.
- Bardwell, J.C.A., McGovern, K., and Beckwith, J. (1991). Identification of a protein required for disulfide bond formation in vivo. *Cell* *67*, 581–589.
- Bessette, P.H., Cotto, J.J., Gilbert, H.F., and Georgiou, G. (1999). In vivo and in vitro function of the *Escherichia coli* periplasmic cysteine oxidoreductase DsbG. *J. Biol. Chem.* *274*, 7784–7792.

- Chakravarthi, S., Jessop, C.E., Willer, M., Stirling, C.J., and Bulleid, N.J. (2007). Intracellular catalysis of disulfide bond formation by the human sulfhydryl oxidase, QSOX1. *Biochem J* 404, 403–411.
- Codding, J.A., Israel, B.A., and Thorpe, C. (2012). Protein substrate discrimination in the quiescin sulfhydryl oxidase (QSOX) family. *Biochemistry* 51, 4226–4235.
- Collet, J.-F.F., and Bardwell, J.C.A. (2002). Oxidative protein folding in bacteria. *Mol Microbiol* 44, 1–8.
- Coppock, D., Kopman, C., Gudas, J., and Cina-Poppe, D.A. (2000). Regulation of the quiescence-induced genes: quiescin Q6, decorin, and ribosomal protein S29. *Biochem. Biophys. Res. Commun.* 269, 604–610.
- Daithankar, V., Schaefer, S.A., Dong, M., Bahnson, B.J., and Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augments liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* 49, 6737–6745.
- Deponte, M., and Hell, K. (2009). Disulfide bond formation in the intermembrane space of mitochondria. *J Biochem.*
- Frand, A.R., and Kaiser, C.A. (1998). The ERO1 gene of yeast is required for oxidation of protein dithiols in the endoplasmic reticulum. *Mol. Cell* 1, 161–170.
- Frand, A.R., and Kaiser, C.A. (1999). Ero1p oxidizes protein disulfide isomerase in a pathway for disulfide bond formation in the endoplasmic reticulum. *Mol. Cell* 4, 469–477.
- Frand, A.R., and Kaiser, C.A. (2000). Two pairs of conserved cysteines are required for the oxidative activity of Ero1p in protein disulfide bond formation in the endoplasmic reticulum. *Mol. Biol. Cell* 11, 2833–2843.
- Grauschopf, U., Winther, J.R., Korber, P., Zander, T., Dallinger, P., and Bardwell, J.C.A. (1995). Why is DsbA such an oxidizing disulfide catalyst? *Cell* 83, 947–955.
- Guddat, L.W., Bardwell, J.C.A., Glockshuber, R., Huber-Wunderlich, M., Zander, T., and Martin, J.L. (1997). Structural analysis of three His32 mutants of DsbA: support for an electrostatic role of His32 in DsbA stability. *Protein Sci.* 6, 1893–1900.
- Heckler, E.J., Rancy, P.C., Kodali, V.K., and Thorpe, C. (2008a). Generating disulfides with the Quiescin-sulfhydryl oxidases. *Biochim Biophys Acta* 1783, 567–577.

- Heckler, E.J., Alon, A., Fass, D., and Thorpe, C. (2008b). Human quiescin-sulfhydryl oxidase, QSOX1: probing internal redox steps by mutagenesis. *Biochemistry* *47*, 4955–4963.
- Hofmann, S., Rothbauer, U., Mühlenbein, N., Baiker, K., Hell, K., and Bauer, M.F. (2005). Functional and mutational characterization of human MIA40 acting during import into the mitochondrial intermembrane space. *J. Mol. Biol.* *353*, 517–528.
- Hooper, K.L., Joneja, B., White 3rd, H.B., and Thorpe, C. (1996). A sulfhydryl oxidase from chicken egg white. *J Biol Chem* *271*, 30510–30516.
- Hooper, K.L., Glynn, N.M., Burnside, J., Coppock, D.L., and Thorpe, C. (1999). Homology between egg white sulfhydryl oxidase and quiescin Q6 defines a new class of flavin-linked sulfhydryl oxidases. *J Biol Chem* *274*, 31759–31762.
- Inoue, A., Yoshida, N., Omoto, Y., Oguchi, S., Yamori, T., Kiyama, R., and Hayashi, S. (2002). Development of cDNA microarray for expression profiling of estrogen-responsive genes. *J. Mol. Endocrinol.* *29*, 175–192.
- Jander, G., Martin, N.L., and Beckwith, J. (1994). Two cysteines in each periplasmic domain of the membrane protein DsbB are required for its function in protein disulfide bond formation. *EMBO J.* *13*, 5121–5127.
- Katchman, B.A., Antwi, K., Hostetter, G., Demeure, M.J., Watanabe, A., Decker, G.A., Miller, L.J., Von Hoff, D.D., and Lake, D.F. (2011). Quiescin sulfhydryl oxidase 1 promotes invasion of pancreatic tumor cells mediated by matrix metalloproteinases. *Mol. Cancer Res.* *9*, 1621–1631.
- Kodali, V.K., and Thorpe, C. (2010a). Oxidative protein folding and the Quiescin-sulfhydryl oxidase family of flavoproteins. *Antioxid Redox Signal* *13*, 1217–1230.
- Kodali, V.K., and Thorpe, C. (2010b). Quiescin sulfhydryl oxidase from *Trypanosoma brucei*: catalytic activity and mechanism of a QSOX family member with a single thioredoxin domain. *Biochemistry* *49*, 2075–2085.
- Kouwen, T.R.H.M., van der Goot, A., Dorenbos, R., Winter, T., Antelmann, H., Plaisier, M.-C., Quax, W.J., van Dijl, J.M., and Dubois, J.-Y.F. (2007). Thiol-disulphide oxidoreductase modules in the low-GC Gram-positive bacteria. *Mol. Microbiol.* *64*, 984–999.
- LaMantia, M.L., and Lennarz, W.J. (1993). The essential function of yeast protein disulfide isomerase does not reside in its isomerase activity. *Cell* *74*, 899–908.

- Madden, M., Lau, S.M., and Thorpe, C. (1984). The influence of oxidation-reduction state on the kinetic stability of pig kidney general acyl-CoA dehydrogenase and other flavoproteins. *Biochem. J.* 224, 577–580.
- Mairet-Coello, G., Tury, A., Esnard-Feve, A., Fellmann, D., Risold, P.-Y., and Griffond, B. (2004). FAD-linked sulfhydryl oxidase QSOX: topographic, cellular, and subcellular immunolocalization in adult rat central nervous system. *J. Comp. Neurol.* 473, 334–363.
- Mairet-Coello, G., Tury, A., Fellmann, D., Risold, P.-Y., and Griffond, B. (2005). Ontogenesis of the sulfhydryl oxidase QSOX expression in rat brain. *J. Comp. Neurol.* 484, 403–417.
- Martin, J.L., Bardwell, J.C.A., and Kuriyan, J. (1993). Crystal structure of the DsbA protein required for disulphide bond formation in vivo. *Nature* 365, 464–468.
- McCarthy, A.A., Haebel, P.W., Törrönen, A., Rybin, V., Baker, E.N., and Metcalf, P. (2000). Crystal structure of the protein disulfide bond isomerase, DsbC, from *Escherichia coli*. *Nat. Struct. Biol.* 7, 196–199.
- Morel, C., Adami, P., Musard, J.-F., Duval, D., Radom, J., and Jouvenot, M. (2007). Involvement of sulfhydryl oxidase QSOX1 in the protection of cells against oxidative stress-induced apoptosis. *Exp. Cell Res.* 313, 3971–3982.
- Musard, J.F., Sallot, M., Dulieu, P., Fraïchard, A., Ordener, C., Remy-Martin, J.P., Jouvenot, M., and Adami, P. (2001). Identification and expression of a new sulfhydryl oxidase SOx-3 during the cell cycle and the estrus cycle in uterine cells. *Biochem. Biophys. Res. Commun.* 287, 83–91.
- Naoé, M., Ohwa, Y., Ishikawa, D., Ohshima, C., Nishikawa, S.-I., Yamamoto, H., and Endo, T. (2004). Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. *J. Biol. Chem.* 279, 47815–47821.
- Pollard, M.G., Travers, K.J., and Weissman, J.S. (1998). Ero1p: a novel and ubiquitous protein with an essential role in oxidative protein folding in the endoplasmic reticulum. *Mol. Cell* 1, 171–182.
- Song, H., Zhang, B., Watson, M.A., Humphrey, P.A., Lim, H., and Milbrandt, J. (2009). Loss of Nkx3.1 leads to the activation of discrete downstream target genes during prostate tumorigenesis. *Oncogene* 28, 3307–3319.

Thorpe, C., Hooper, K.L., Raje, S., Glynn, N.M., Burnside, J., Turi, G.K., and Coppock, D.L. (2002). Sulfhydryl oxidases: emerging catalysts of protein disulfide bond formation in eukaryotes. *Arch Biochem Biophys* 405, 1–12.

Tury, A., Mairet-Coello, G., Poncet, F., Jacquemard, C., Risold, P.Y., Fellmann, D., and Griffond, B. (2004). QSOX sulfhydryl oxidase in rat adenohypophysis: localization and regulation by estrogens. *J. Endocrinol.* 183, 353–363.

Wunderlich, M., and Glockshuber, R. (1993). Redox properties of protein disulfide isomerase (DsbA) from *Escherichia coli*. *Protein Sci* 2, 717–726.

Zapun, A., Bardwell, J.C.A., and Creighton, T.E. (1993). The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation in vivo. *Biochemistry* 32, 5083–5092.

Zapun, A., Missiakas, D., Raina, S., and Creighton, T.E. (1995). Structural and functional characterization of DsbC, a protein involved in disulfide bond formation in *Escherichia coli*. *Biochemistry* 34, 5075–5089.

## **Chapter 2**

### **DISCOVERY OF AUGMENTER OF LIVER REGENERATION AND A NEW FAMILY OF FLAVIN-LINKED SULFHYDRYL OXIDASES**

#### **2.1 Introduction**

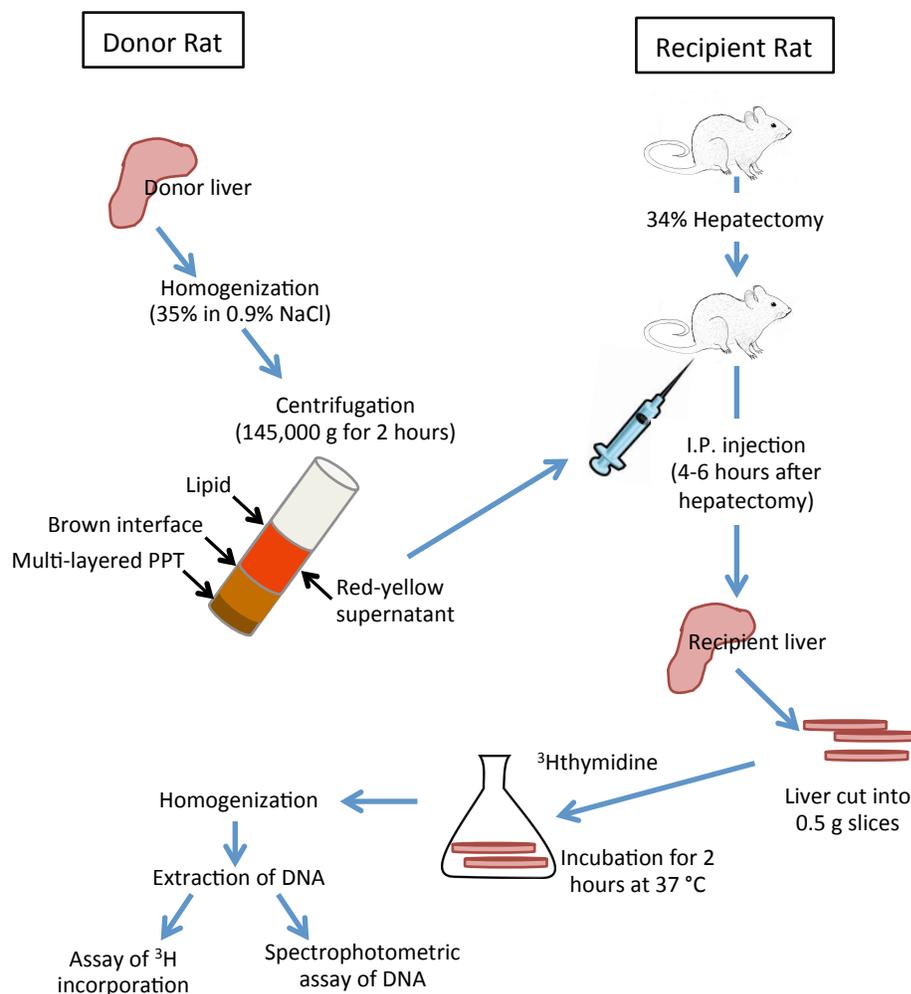
While the expanding physiological roles of augmenter of liver regeneration (ALR) are under active investigation, the long-form ALR is known to act as a sulfhydryl oxidase in the intermembrane space (IMS) of mitochondria and may function in the assembly of cytosolic Fe/S centers. The short-form of ALR is believed to play a role in the liver regeneration cascade. This Chapter will discuss the discovery of a factor which was found to stimulate liver regeneration and later shown to be the enzyme ALR. A discussion of structural characteristics of the ERV/ALR family of flavoproteins will also be addressed.

#### **2.2 Discovery of Augmenter of Liver Regeneration (ALR)**

The mammalian liver is one of the few organs with significant regenerative capacity. Following removal of the median and left lateral lobes (a partial hepatectomy amounting to about 30% of liver mass) the remaining lobes grow rapidly and the liver regains pre-surgery size in about one week in rats. One of the factors responsible for this remarkable regenerative ability was found by LaBrecque and Pesch to be released into the bloodstream from weanling and damaged rat liver, but not from adult healthy animals (LaBrecque and Pesch, 1975). They termed this factor

hepatic stimulator substance (HSS). A key aspect of their studies, and the subsequent biomedical work of others, was the assay used for HSS (Figure 2.1)

## Preparation and assay of hepatic regenerative stimulator substance



**Figure 2.1 Stimulating the incorporation of tritiated thymidine into rat liver.** The red-yellow supernatant extracted from donor rat livers was injected into recipient rats after partial hepatectomy. 4-6 hours after injection the recipient's liver was cut into thin slices and incubated with <sup>3</sup>H-thymidine, homogenized prior to analysis. Figure adapted from (LaBrecque and Pesch, 1975).

### **2.3 Assay of Hepatic Stimulator Substance (HSS) Activity**

Donor weanling Sprague Dawley rat livers were homogenized and a 35% w/v homogenate in 0.9% NaCl was centrifuged for 2 hours (140,000 g) (LaBrecque and Pesch, 1975). The red-yellow supernatant was decanted and 5 ml were injected into the body cavity of a recipient rat that received a 34% hepatectomy 4-6 hours prior to injection of the supernatant. 23 hours after partial hepatectomy, livers from recipient rats were removed. Slices (0.5 g) were then incubated with tritiated thymidine for cell proliferation assays (Figure 2.1).

Using this assay, LaBrecque and Pesch showed that intraperitoneal injection of an extract of weanling rat liver stimulated the regeneration of partially hepatectomized rat liver from 2.5-fold to 30-fold (LaBrecque and Pesch, 1975). Importantly, LaBrecque and Pesch were the first to report the presence of HSS to be age-dependent. Rats less than four weeks old produce detectable amounts of HSS, but by 6-8 weeks HSS was not detectable (LaBrecque and Pesch, 1975). While adult rats do not produce detectable amounts of HSS in their livers, 16 hours after 68% hepatectomy they were able to generate enough HSS to induce a 2.3-fold proliferative stimulation. After 26 hours stimulation was nearly 4-fold increased (LaBrecque and Pesch, 1975). Thereafter, the amount of HSS remained constant for 72 hours and then declined to undetectable levels by day 8 (LaBrecque and Pesch, 1975).

LaBrecque and Pesch showed that HSS was stable to heating at both 65 and 100 °C for 15 min, but lost its ability to stimulate liver regeneration following treatment with 10% perchloric acid. The authors did not attempt a purification of the factor and therefore could not conclude that HSS was a protein.

Subsequently, Starzl and colleagues isolated augmentor of liver regeneration (ALR) purifying the protein some 800,000-fold from the serum of partially

hepatectomized rats (Francavilla et al., 1987, 1991, 1993). The protein isolated was approximately 33 kDa which corresponds to the short-form ALR (Francavilla et al., 1994). To confirm the active agent in HSS was ALR, purified recombinant ALR (see below) was shown to prevent hepatocyte atrophy in dogs which is normally characteristic of Eck fistula (an Eck fistula is a tube that is inserted to connect the end of the hepatic portal vein and the side of the inferior vena cava) (Francavilla et al., 1994).

#### **2.4 Sequence Determination of Augmenter of Liver Regeneration**

The sequence of the rat ALR was determined by Hagiya et al. (Hagiya et al., 1994). A purified sample of rat ALR was run on an SDS-PAGE and two major bands were observed at approximately 48 kDa and 28 kDa under non-reducing conditions and 23 kDa and 14 kDa under reducing conditions suggesting that the rat ALR protein is a homodimer (Giorda et al., 1996; Hagiya et al., 1994). The 28 kDa band was digested with lysyl endopeptidase (*Achromobacter lyticus* protease I) and the peptides separated via HPLC using an octyldecyl silica column and sequenced. Primers for PCR amplification of the human ALR were chosen based on the partial amino acid sequences of rat ALR. The PCR amplified human DNA was then digested and sequenced. The corresponding translated protein sequence was 125 amino acids long and had a molecular weight of 15,081 Da. This sequence and mass corresponds to the short-form ALR.

Shortly after the sequence determination of rat ALR by Hagiya et al., the same group reported the sequence of the mouse and human ALR proteins. First the mouse genomic library was screened using cDNA from the rat. Positive plaques were purified, digested using restriction endonuclease and run on a gel. The DNA

fragments were subcloned and sequenced. Only a 48 kDa band was observed for the mouse ALR protein unlike the rat ALR which seems to have two forms of the protein or a cleavage site (Giorda et al., 1996). The mouse gene maps to chromosome 17 (Giorda et al., 1996).

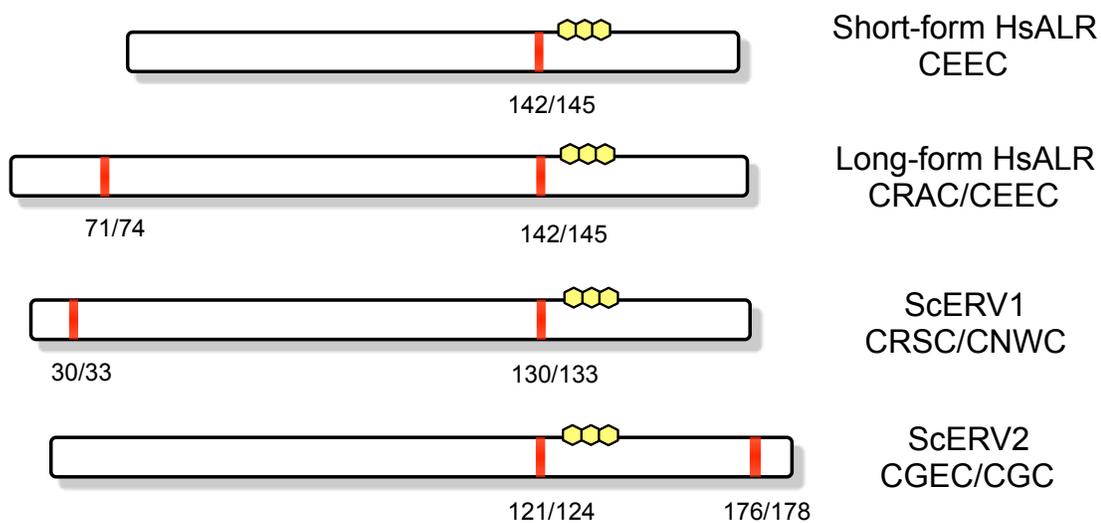
The human ALR sequence was determined using rat ALR cDNA as a probe in a manner similar to the mouse ALR sequence determination (Giorda et al., 1996). When a positive plaque was identified it was purified and the DNA was subcloned into an expression vector and sequenced (Giorda et al., 1996). The human enzyme was reported to be a covalently-linked homodimer which exists in two forms, a 15 kDa form localized to the nucleus and a 23 kDa form mainly in the cytoplasm (Li et al., 2002). The authors also suggested the possibility that ALR can form a heterodimer, of short and long-form subunits. With three mammalian sequences in hand, multiple sequence alignment showed high sequence conservation of the ALR proteins and an approximate 50% similarity to the yeast enzyme ERV1.



## 2.5 Yeast ERV1 is a Human Homolog

In 1992, Lisowsky and coworkers identified a gene in yeast that was essential for respiration and viability (ERV) (Lisowsky, 1992). Lisowsky et al. (Lisowsky et al., 1995) identified ALR as a structural and functional homologue of the yeast ERV1 having a sequence identity of 42%. The functional homology was determined by inserting an adapted human ALR sequence (amino acids 1-21 of scERV1 attached to 106-205 of HsALR) into a yeast vector to see if the human protein was able to rescue the yeast protein (Lisowsky et al., 1995). Their experiments showed that this replacement resulted in slower growth of the yeast cells, but did not severely affect viability.

The finding that the FAD-dependent sulfhydryl oxidase, QSOX, contained a C-terminal ERV domain (Hooper et al., 1999) led Lisowsky and coworkers to explore the outcome of the heterologous expression of their yeast growth factor in *Escherichia coli*. His-tagged ScERV1 was purified as a yellow, FAD-dependent enzyme with modest sulfhydryl oxidase activity towards both DTT and reduced lysozyme (Hofhaus et al., 2003; Lee et al., 2000). This finding not only established ScERV1 as the founding member of the small sulfhydryl oxidase family to be described more fully in this Chapter, but it also located the flavin binding domain in the larger QSOX enzymes. Lisowsky and colleagues then identified ALR as a mammalian ortholog of ScERV1 by expressing the protein in *E. coli* and finding flavin-dependent sulfhydryl oxidase activity. Finally, they identified ScERV2 as a paralog of yeast ERV1 and showed that it was also a FAD-linked sulfhydryl oxidase, but that it resided in the lumen of the endoplasmic reticulum (Gerber et al., 2001; Lisowsky et al., 1995).



**Figure 2.3 Domain structures of human ALR and yeast enzymes.** The red bars indicate the position of the redox-active disulfide motifs. The FAD cofactors are indicated in yellow. The CxxC sequences for each enzyme are shown to the right. The C-terminal redox active disulfide of ScERV2 is a CxC motif.

## 2.6 Function of ERV1 and ERV2

Not only does ERV1 play a role in folding of proteins in the IMS of mitochondria, as discussed in Chapter 1, but it was also identified as a requirement for the maturation of cytosolic Fe/S proteins (Lange et al., 2001). Lange et al. generated a temperature-sensitive mutant of ERV1 which causes a defect in maturation of cytosolic Fe/S proteins leading to the accumulation of “free” Fe in the mitochondria (Lange et al., 2001). A recent paper discovered an inhibitor of ERV1 which also reacts with ERV2 and HsALR. This compound 2,4-dichloro-6-((((phenylamino)phenyl)imino)methyl)phenol) or MitoBloCK-6 inhibits ScERV1, HsALR and ScERV2 with IC<sub>50</sub> values of 900 nM, 700 nM and 1.4 μM, respectively (Dabir et al., 2013).

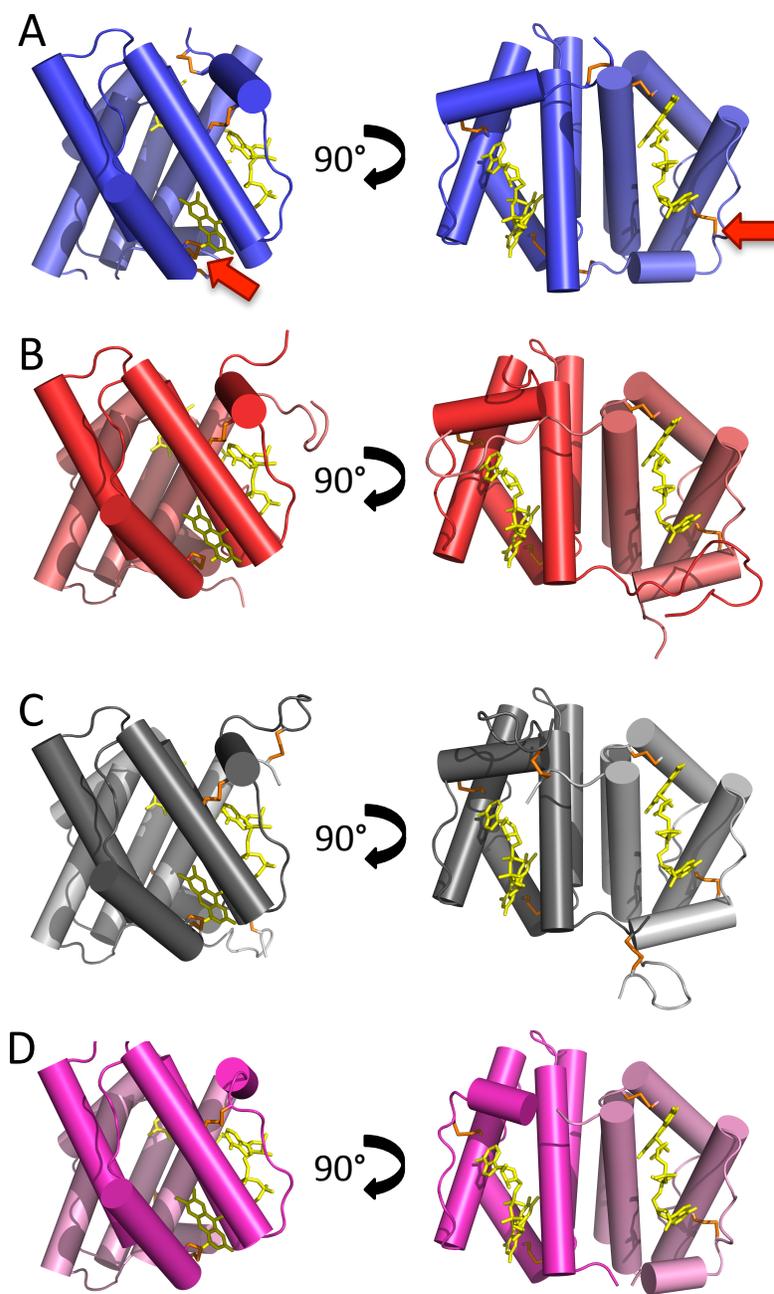
ERV2 was discovered when Lisowsky and colleagues took the sequence of ERV1 and searched the yeast genome to find similar sequences related to ERV1 (Stein and Lisowsky, 1998). ERV1 and ERV2 share only 20% sequence identity overall and 30% sequence identity in the conserved C-terminal domain (Stein and Lisowsky, 1998). ERV2 is not a functional homolog of ERV1 (Stein and Lisowsky, 1998) and is, in fact, not essential to the viability of yeast unlike ERV1 (Stein and Lisowsky, 1998). ERV1 shares approximately 42% sequence identity with human ALR (Lisowsky et al., 1995). While ERV2 is present throughout the fungal kingdom, it is not seen in metazoans, plants or protozoa (Vala et al., 2005). In contrast, QSOX is almost ubiquitous in higher eukaryotes but is not found in fungi. ERV2 was expressed recombinantly and shown to be a sulfhydryl oxidase (Gerber et al., 2001; Sevier et al., 2001). Both ERV2 and QSOX are localized to the ER (Sevier et al., 2001) and contribute to disulfide bond formation at that location. Currently, the physiological partners of ERV2 are unclear. Wang, Winther and Thorpe have characterized the

redox behavior of the enzyme (Wang et al., 2007). ERV2 can be inhibited by  $Zn^{2+}$  by forming a complex between the zinc and redox-active CGC motif near the N-terminus of ERV2 (Wang et al., 2007). It is also known that the length of the “tail” between the two redox-active disulfides is crucial for activity (Figure 2.3) (Vala et al., 2005).

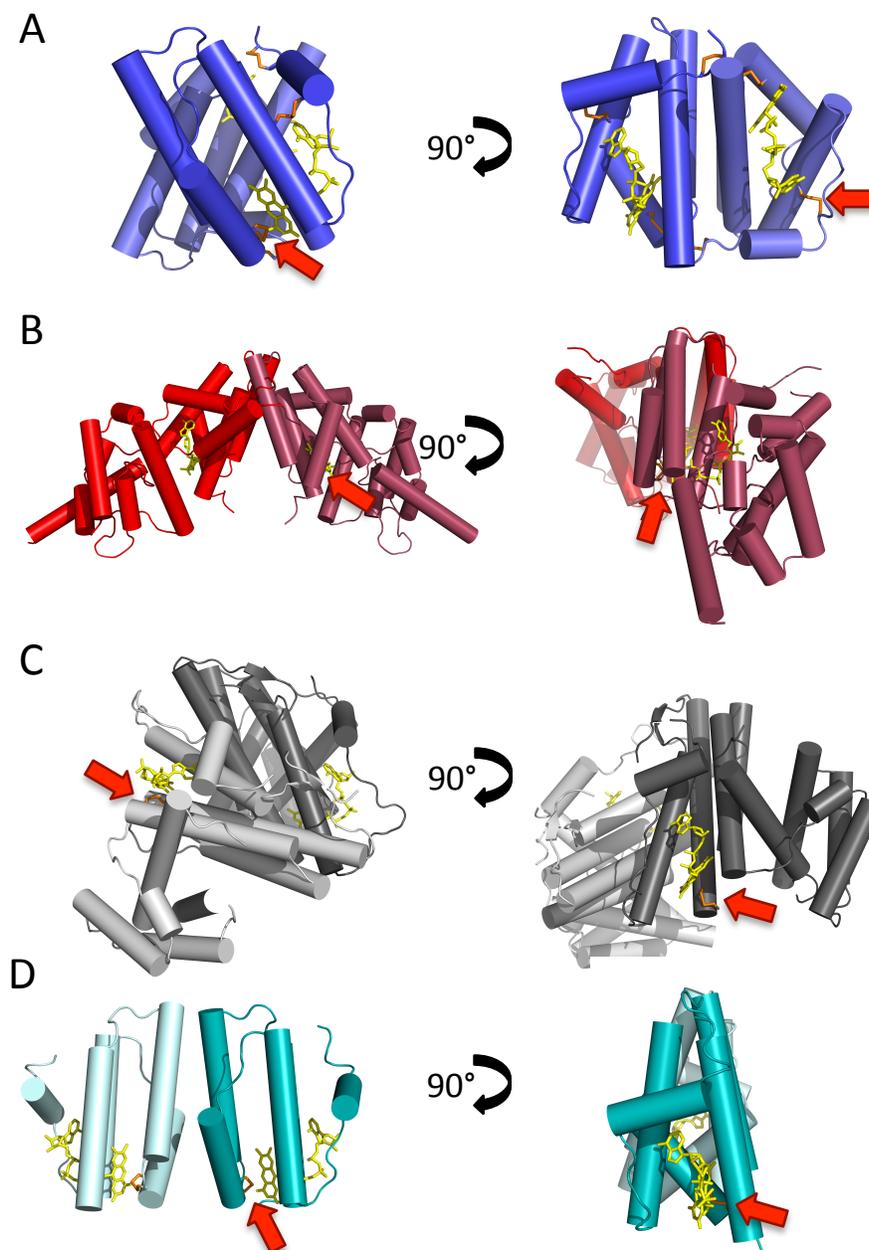
ERV2 exhibits activity with DTT, a common model substrate for these small sulfhydryl oxidases, but is less active with monothiols such as GSH, cysteine and 2-mercaptoethanol (Vala et al., 2005). PDI has been shown to be a modest substrate of ERV2, but DTT is a much better substrate (Vala et al., 2005).

## **2.7 ERV/ALR Protein Family: A Structural Comparison**

Since the discovery of ERV1 and ALR the family of proteins has expanded to include a wide variety of small sulfhydryl oxidases. This section will present an overview of structures from this family (Figures 2.4 and 2.5). All members of the ERV/ALR family of proteins are homodimers with each subunit comprising a tightly packed 4-helix bundle harboring a non-covalently linked FAD cofactor. Each monomer contains a redox-active CxxC motif proximal to the isoalloxazine ring of the FAD cofactor and generally an additional redox-active shuttle disulfide located on either the N- or C-terminus (Figure 2.3). Most of the structures that have been solved do not have electron density in the region of the shuttle disulfide consistent with a high level of conformational flexibility. This section will illustrate the many similar structural aspects of the ERV/ALR family while identifying the differences such as the location of the shuttle disulfide and variations in the dimer interface.



**Figure 2.4 Comparison of non-viral ERV/ALR proteins.** A) ScERV2 (PDB 1JR8, B) ScERV1 (PDB 3W4Y) C) short-form HsALR (PDB 3MBG) and D) AtERV1 (PDB 2HJ3). The FAD cofactor is shown in yellow. Disulfides are shown as orange sticks. The redox-active disulfide proximal to the isoalloxazine ring is depicted, for illustration, by a red arrow in the ERV2 structure.



**Figure 2.5 Comparison of viral ERV/ALR proteins with ERV2.** A) ScERV2 (PDB 1JR8) B) mimivirus R596 (PDB 3TD7) C) baculovirus Ac92 (PDB 3QZY) and D) African swine fever virus pB119L (PDB 3GWL). The FAD cofactor is shown in yellow. Disulfides are shown as orange sticks. The redox-active disulfide proximal to the isoalloxazine ring is depicted, for illustration, by a red arrow in the structures.

### 2.7.1 Yeast ERV1 and ERV2

ERV1 was the first enzyme of the family to be discovered and, as mentioned previously, this led to the discovery of an additional yeast paralog, ERV2. ScERV2 was the first reported ERV/ALR structure (Figures 2.4A and 2.5A) (Gross et al., 2002; Lisowsky, 1992). The structure shows a highly helical head-to-tail homodimer containing one FAD cofactor per monomer. This small helical FAD binding fold in ERV2 represented a novel flavin binding motif (Gross et al., 2002; Lisowsky, 1992) and served as the prototype for the structures to follow. The hydrophobic dimer interface represent  $\alpha 1$  and  $\alpha 2$  of both subunits and these two helices are almost antiparallel in orientation. The redox-active CxxC motifs are positioned adjacent to the isoalloxazine ring of the FAD at a turn/helix boundary facing away from the interface.

Recently the crystal structure of the ScERV1 was solved (Figure 2.4B) (Guo et al., 2012). ERV1 has sequence similarity to the long-form of HsALR as they both have N-terminal extensions harboring an additional redox-active CxxC motif. In order to crystallize this enzyme the authors made mutations to “trap” the flexible N-terminal extension to the proximal CxxC motif. ERV1 C30S/C133S was crystallized in the presence of an oxidizing redox buffer containing 5 mM GSSG:GSH (3:1) to capture a mixed disulfide comprising of the shuttle disulfide of one subunit communicating with the proximal disulfide of the other subunit. Like ERV2, the structure of ERV1 shows the same hydrophobic  $\alpha 1/\alpha 2$  dimer interface. This structure provides new insight into the possible conformation of the N-terminal extension of long-form HsALR.

### **2.7.2 Human Augmenter of Liver Regeneration**

The crystal structure of human ALR was solved in the Bahnsen laboratory in 2010 (Figure 2.4C) as part of an enzymological characterization of an autosomal recessive mutation in ALR (see Chapter 3 for details) (Daithankar et al., 2010). The human sequence is 86% identical to the rat enzyme so it was not surprising that the structures are essentially identical. ALR shares the same dimer interface ( $\alpha 1/\alpha 2$ ) as the yeast enzymes ERV1 and ERV2. To date, many attempts have been made in this laboratory to try crystallize the wild type long-form ALR with no success. Constructs were also designed which had cysteine mutations to try and “trap” the shuttle disulfide to the proximal disulfide but they have yet to yield crystals.

### **2.7.3 Plant ERV1**

*Arabidopsis thaliana* ERV1 (AtERV1) is a non-covalent homodimer that has three disulfides (Figure 2.4D) (Farver et al., 2009). Two of disulfides are redox-active and the third is structural. The structure of this enzyme has been solved (PDB 2HJ3) and shows that Cys119/Cys122 represents the proximal redox-active disulfide which communicates with the non-covalent FAD cofactor (Farver et al., 2009). The C-terminal extension harboring the second shuttle disulfide was too disordered to get any electron density information (Farver et al., 2009). These authors also noted that the structural disulfides in AtERV are accessible to reducing agents unlike the structural disulfide in the E10R poxvirus enzyme (see 2.7.4) which may need to be buried to prevent reduction from cytosolic GSH or TRX (Farver et al., 2009).

### **2.7.4 Viral Members of the ERV Family**

ERV/ALR family members are found in viruses as well and exhibit a wider range of structural diversity than their eukaryotic homologs. The viral sulfhydryl

oxidases are generally smaller than their eukaryote relatives and generally lack the non-conserved structural disulfides (Hakim and Fass, 2009). E10R is the product of an essential gene involved in the replication of vaccinia viruses (poxvirus) (Senkevich et al., 2000). E10R is a cytosolic membrane bound flavoprotein containing one disulfide and one free thiol (Senkevich et al., 2000). Studies suggest that E10R is able to oxidize the cytosolic poxvirus proteins L1R and F9L and complexes with A2.5L and possibly G4L glutaredoxin (Senkevich et al., 2000, 2002).

pB119L (Figure 2.5D) is a sulfhydryl oxidase from African Swine Fever Virus (ASFV), which is a large double stranded DNA virus similar to poxvirus and mimivirus (Hakim and Fass, 2009). These double stranded DNA viruses assemble in the cytosol of infected cells (Hakim and Fass, 2009). Even though these viruses are localized to the reducing environment of the cytosol they have intact disulfides (Hakim and Fass, 2009). One possible interacting partner of pB119L is the viral protein pA151R (Rodríguez et al., 2006). Fass and colleagues solved the crystal structure of pB119L $\Delta$ C (which lacks 16 amino acids on the C-terminus) (PDB 3GWL). pB119L and ERV2 homodimers share a very similar monomeric architecture but, notably, utilize different dimer interfaces. As mentioned earlier ERV1, ERV2 and ALR homodimers utilize  $\alpha$ 1/ $\alpha$ 2 helices to provide their intersubunit contacts. Surprisingly the dimer interface of pB119L comprises helices  $\alpha$ 2/ $\alpha$ 3, and the two FAD/proximal disulfide redox centers are in close proximity. There appears to be unusual plasticity in the dimer interfaces utilized by double-stranded DNA viruses. As noted below, other viral ALR proteins exploit an  $\alpha$ 1/ $\alpha$ 2 or an  $\alpha$ 3/ $\alpha$ 4 interface.

The next crystal structure determined for a viral ALR was that from the baculovirus *Autographa californica* (Figure 2.5C) (Hakim et al., 2011). Again this

protein is essential for virion assembly and virus propagation (Nie et al., 2011; Wu and Passarelli, 2010). Baculoviral ALR is a homodimer that utilizes an  $\alpha 3/\alpha 4$  subunit interface. More remarkably, in contrast to other ERV/ALR family members, the baculoviral homodimers are formed from subunits that each contain two 4-helical bundles (Figure 2.5C). The N-terminal bundle lacks a flavin and a proximal disulfide whereas the C-terminal helix bundle shows the FAD and redox active disulfide typical of ERV/ALR family members. It is striking that the interdomain surface between redox inactive and active helix bundles shows the same topology found with ERV1, ERV2 and ALR (involving  $\alpha 1/\alpha 2$  contact). Fass and colleagues have described the redox-active helix bundle as a pseudo-ERV domain (Alon et al., 2010) and have suggested it arose by gene duplication, fusion and loss of function. QSOX contains such a pairing of pseudo-ERV and ERV domains (HRR-ERV; Chapter 1) and it is thus intriguing that exactly the same arrangement appears in a viral sulfhydryl oxidase (Hakim and Fass, 2009).

The third distinct viral structure comes from mimivirus (Hakim et al., 2012). This R596 protein is a dimer of ~69 kDa consisting of a FAD containing sulfhydryl oxidase domain and a non-conserved ORFan domain per monomer (Hakim et al., 2012). R596 has a dimer interface of  $\alpha 1/\alpha 2$ , which is similar to ERV1, ERV2 and ALR (Hakim and Fass, 2009). In order to obtain soluble protein, 4 of the 8 cysteines (per monomer) were mutated to alanine to prevent protein aggregation during purification. Most viral ERV proteins lack the shuttle disulfide, however, that is not the case for R596. The mimiviral oxidase protein has a shuttle disulfide (Cx<sub>9</sub>C motif) from one monomer that transfers electrons to a nearby CxxC motif proximal to the FAD cofactor of the other subunit. The authors also tested a variety of substrates and

showed that DTT and the phosphine reductant THP are able to support turnover of the protein whereas reduced thioredoxin and reduced/denatured RNase A are not substrates.

## **2.8 Summary**

This Chapter has briefly described the discovery of ALR and the events that led to the realization that this growth factor was also a flavin linked sulfhydryl oxidase. The ERV/ALR domain comprises a four-helix bundle binding FAD with a redox-active disulfide proximal to the flavin. While all these small sulfhydryl oxidases are homodimers, viral members of this family are structurally more diverse. With this Chapter as an introduction to major family members, Chapter 3 will describe the biological aspects of mammalian ALR in more detail.

## REFERENCES

- Alon, A., Heckler, E.J., Thorpe, C., and Fass, D. (2010). QSOX contains a pseudo-dimer of functional and degenerate sulfhydryl oxidase domains. *FEBS Lett.* *584*, 1521–1525.
- Dabir, D. V, Hasson, S.A., Setoguchi, K., Johnson, M.E., Wongkongkathep, P., Douglas, C.J., Zimmerman, J., Damoiseaux, R., Teitell, M.A., and Koehler, C.M. (2013). A small molecule inhibitor of redox-regulated protein translocation into mitochondria. *Dev. Cell* *25*, 81–92.
- Daithankar, V., Schaefer, S.A., Dong, M., Bahnson, B.J., and Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augments liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* *49*, 6737–6745.
- Farver, O., Vitu, E., Wherland, S., Fass, D., and Pecht, I. (2009). Electron transfer reactivity of the *Arabidopsis thaliana* sulfhydryl oxidase AtErv1. *J Biol Chem* *284*, 2098–2105.
- Francavilla, A., Ove, P., Polimeno, L., Coetzee, M., Makowka, L., Rose, J., Van Thiel, D.H., and Starzl, T.E. (1987). Extraction and partial purification of a hepatic stimulatory substance in rats, mice, and dogs. *Cancer Res.* *47*, 5600–5605.
- Francavilla, A., Barone, M., Van Thiel, D.H., Mazzaferro, V., Prelich, J.G., and Starzl, T.E. (1991). Further steps of hepatic stimulatory substance purification. *Dig. Dis. Sci.* *36*, 674–680.
- Francavilla, A., Polimeno, L., Barone, M., Azzarone, A., and Starzl, T.E. (1993). Hepatic regeneration and growth factors. *J. Surg. Oncol. Suppl.* *3*, 1–7.
- Francavilla, A., Hagiya, M., Porter, K.A., Polimeno, L., Ihara, I., and Starzl, T.E. (1994). Augmenter of liver regeneration: its place in the universe of hepatic growth factors. *Hepatology* *20*, 747–757.
- Gerber, J., Muhlenhoff, U., Hofhaus, G., Lill, R., and Lisowsky, T. (2001). Yeast ERV2p is the first microsomal FAD-linked sulfhydryl oxidase of the Erv1p/Alrp protein family. *J Biol Chem* *276*, 23486–23491.

- Giorda, R., Hagiya, M., Seki, T., Shimonishi, M., Sakai, H., Michaelson, J., Francavilla, A., Starzl, T.E., and Trucco, M. (1996). Analysis of the structure and expression of the augmenter of liver regeneration (ALR) gene. *Mol Med* 2, 97–108.
- Gross, E., Sevier, C.S., Vala, A., Kaiser, C.A., and Fass, D. (2002). A new FAD-binding fold and intersubunit disulfide shuttle in the thiol oxidase Erv2p. *Nat Struct Biol* 9, 61–67.
- Guo, P.-C., Ma, J.-D., Jiang, Y.-L., Wang, S.-J., Bao, Z.-Z., Yu, X.-J., Chen, Y., and Zhou, C.-Z. (2012). Structure of yeast sulfhydryl oxidase Erv1 reveals electron transfer of the disulfide relay system in the mitochondrial intermembrane space. *J. Biol. Chem.*
- Hagiya, M., Francavilla, A., Polimeno, L., Ihara, I., Sakai, H., Seki, T., Shimonishi, M., Porter, K.A., and Starzl, T.E. (1994). Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: expression of biologically active recombinant ALR and demonstration of tissue distribution. *Proc Natl Acad Sci U S A* 91, 8142–8146.
- Hakim, M., and Fass, D. (2009). Dimer interface migration in a viral sulfhydryl oxidase. *J Mol Biol* 391, 758–768.
- Hakim, M., Mandelbaum, A., and Fass, D. (2011). Structure of a baculovirus sulfhydryl oxidase, a highly divergent member of the erv flavoenzyme family. *J. Virol.* 85, 9406–9413.
- Hakim, M., Ezerina, D., Alon, A., Vonshak, O., and Fass, D. (2012). Exploring ORFan domains in giant viruses: structure of mimivirus sulfhydryl oxidase R596. *PLoS One* 7, e50649.
- Hofhaus, G., Lee, J.E., Tews, I., Rosenberg, B., and Lisowsky, T. (2003). The N-terminal cysteine pair of yeast sulfhydryl oxidase Erv1p is essential for in vivo activity and interacts with the primary redox centre. *Eur J Biochem* 270, 1528–1535.
- Hooper, K.L., Glynn, N.M., Burnside, J., Coppock, D.L., and Thorpe, C. (1999). Homology between egg white sulfhydryl oxidase and quiescin Q6 defines a new class of flavin-linked sulfhydryl oxidases. *J Biol Chem* 274, 31759–31762.
- LaBrecque, D.R., and Pesch, L.A. (1975). Preparation and partial characterization of hepatic regenerative stimulator substance (SS) from rat liver. *J Physiol* 248, 273–284.

Lange, H., Lisowsky, T., Gerber, J., Muhlenhoff, U., Kispal, G., and Lill, R. (2001). An essential function of the mitochondrial sulfhydryl oxidase Erv1p/ALR in the maturation of cytosolic Fe/S proteins. *EMBO Rep* 2, 715–720.

Lee, J., Hofhaus, G., and Lisowsky, T. (2000). Erv1p from *Saccharomyces cerevisiae* is a FAD-linked sulfhydryl oxidase. *FEBS Lett.* 477, 62–66.

Li, Y., Wei, K., Lu, C., Li, M., Xing, G., Wei, H., Wang, Q., Chen, J., Wu, C., Chen, H., et al. (2002). Identification of hepatopoietin dimerization, its interacting regions and alternative splicing of its transcription. *Eur J Biochem* 269, 3888–3893.

Lisowsky, T. (1992). Dual function of a new nuclear gene for oxidative phosphorylation and vegetative growth in yeast. *Mol Gen Genet* 232, 58–64.

Lisowsky, T., Weinstat-Saslow, D.L., Barton, N., Reeders, S.T., and Schneider, M.C. (1995). A new human gene located in the PKD1 region of chromosome 16 is a functional homologue to ERV1 of yeast. *Genomics* 29, 690–697.

Nie, Y., Fang, M., and Theilmann, D.A. (2011). *Autographa californica* multiple nucleopolyhedrovirus core gene ac92 (p33) is required for efficient budded virus production. *Virology* 409, 38–45.

Rodríguez, I., Redrejo-Rodríguez, M., Rodríguez, J.M., Alejo, A., Salas, J., and Salas, M.L. (2006). African swine fever virus pB119L protein is a flavin adenine dinucleotide-linked sulfhydryl oxidase. *J. Virol.* 80, 3157–3166.

Senkevich, T.G., White, C.L., Koonin, E. V, and Moss, B. (2000). A viral member of the ERV1/ALR protein family participates in a cytoplasmic pathway of disulfide bond formation. *Proc Natl Acad Sci U S A* 97, 12068–12073.

Senkevich, T.G., White, C.L., Koonin, E. V, and Moss, B. (2002). Complete pathway for protein disulfide bond formation encoded by poxviruses. *Proc. Natl. Acad. Sci. U. S. A.* 99, 6667–6672.

Sevier, C.S., Cuozzo, J.W., Vala, A., Aslund, F., and Kaiser, C.A. (2001). A flavoprotein oxidase defines a new endoplasmic reticulum pathway for biosynthetic disulphide bond formation. *Nat. Cell Biol.* 3, 874–882.

Stein, G., and Lisowsky, T. (1998). Functional comparison of the yeast scERV1 and scERV2 genes. *Yeast* 14, 171–180.

Vala, A., Sevier, C.S., and Kaiser, C.A. (2005). Structural determinants of substrate access to the disulfide oxidase Erv2p. *J. Mol. Biol.* 354, 952–966.

Wang, W., Winther, J.R., and Thorpe, C. (2007). Ery2p: characterization of the redox behavior of a yeast sulfhydryl oxidase. *Biochemistry* 46, 3246–3254.

Wu, W., and Passarelli, A.L. (2010). *Autographa californica* multiple nucleopolyhedrovirus Ac92 (ORF92, P33) is required for budded virus production and multiply enveloped occlusion-derived virus formation. *J. Virol.* 84, 12351–12361.

## **Chapter 3**

### **AUGMENTER OF LIVER REGENERATION: IN VITRO AND IN VIVO STUDIES**

#### **3.1 Introduction**

The previous Chapter introduced the discovery and isolation of ALR from rat liver and serum leading to the determination of ALR's amino acid sequence and the expression of recombinant ALR protein. Studies using recombinant ALR showed that ALR is a flavin-dependent sulfhydryl oxidase and is part of the conserved ERV/ALR family with a tightly packed 4-helix bundle structure. Human ALR exists in long- and short-forms. The long-form spliced variant has an N-terminal extension harboring a redox active CxxC motif similar to that of yeast ERV1. Both the long- and short-forms of the enzyme share a redox-active core containing a proximal CxxC motif adjacent to the isoalloxazine ring of the FAD cofactor (Figure 3.1).

Growth factor ERV1-like, or GFER, is the official gene name for augmenter of liver regeneration. This gene is located on chromosome 16 in the polycystic kidney disease (PKD1) region (Lisowsky et al., 1995). As mentioned in Chapter 1, the long-form ALR participates in disulfide bond formation in the mitochondrial intermembrane space (IMS) while the short-form ALR appears to act a growth factor initiating downstream events which lead to liver regeneration (Francavilla et al., 1994).

This Chapter will discuss some of the initial *in vitro* experiments performed in the Thorpe laboratory which provided a mechanistic basis for the work described in this dissertation then review key biological studies of the short- and long-forms of ALR. Finally, this Chapter concludes with a recent study on a human mutation in ALR that explored the biochemical consequences of the R194H mutation in human ALR.

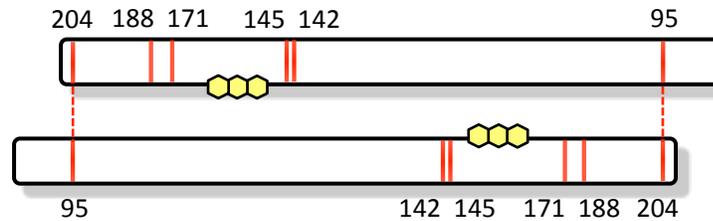
### **3.2 *In vitro* Characterization of ALR**

The long-form of ALR is 205 amino acids long (23 kDa monomer) with an additional 80 amino acid extension to the N-terminus relative to short-form ALR (125 amino acids, 15 kDa monomer) (Figure 3.1) (Li et al., 2002). Both the short- and long-form proteins are head-to-tail covalently-linked homodimers that rely on a non-covalently bound FAD cofactor for oxidase activity. This cofactor takes electrons from the proximal redox-active CxxC motif and transfers them to the final electron acceptor, molecular oxygen or cytochrome *c* (see Chapter 1 for more detail). The long-form ALR is localized predominantly to the IMS of mitochondria; the short-form is found intracellularly in the nucleus and cytosol and extracellularly as a circulating growth factor.

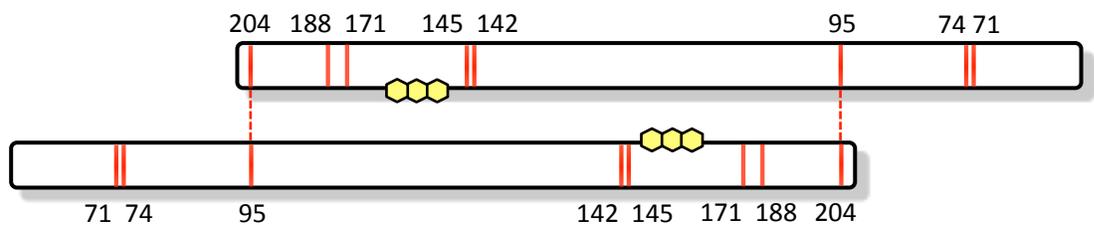
Both short- and long-form ALR share the same catalytic core, with a redox active CxxC motif proximal to the FAD cofactor where C142 is surface-exposed and communicates via a mixed disulfide bond with the substrate and C145 interacts with the FAD cofactor. When C142 is mutated to either an Ala or Ser the remaining cysteine, C145, forms a charge-transfer interaction with the flavin. However, the mutation of C145 to Ala or Ser does not result in a charge-transfer band because Ala and Ser are unable to communicate with the FAD. Any mutation to either C142 or

C145 results in an inactive enzyme. Additionally, the long-form ALR has an additional redox motif C71/C74 (Figure 3.1). Studies from this laboratory have demonstrated that the N-terminal redox motif allows the long-form to communicate with Mia40 (Daithankar et al., 2009). The disulfide connectivity shown in Figure 3.1 was not that originally proposed by Chen et al. for human ALR (Chen et al., 2003). Farrell and Thorpe showed that their connectivity was inconsistent with the crystal structure of the rat enzyme. The revised connectivity of Farrell and Thorpe was further supported by comparison of the sequences of other mammalian ALR enzymes and by analysis of the outcome of mutating cysteine residues (Farrell and Thorpe, 2005).

a. Human short-form ALR



b. Human long-form ALR



**Figure 3.1 Disulfide connectivity for short- and long-form ALR.** a. Short-form ALR begins with M81 and ends with D205. C95 from one subunit is disulfide bonded to C204 from the other subunit forming the covalent dimer. C142/C145 is the only redox active disulfide in the short-form ALR located proximal to the FAD cofactor. b. The long-form ALR has an 80 amino acid N-terminal extension relative to the short-form ALR beginning with M1 and ending with D205. This longer form of the enzyme has two redox active disulfides C71/C74 and C142/C145. The FAD cofactor is shown in yellow. Solid red bars represent cysteine residues. Dashed red lines represent intermolecular disulfides. Two additional cysteines (not shown) were mutated to alanines (C154A and C165A) in all constructs to prevent protein aggregation *in vitro* (Farrell and Thorpe, 2005).

A range of enzymological studies characterized the short-form ALR *in vitro*. Under anaerobic conditions using methyl viologen, as a redox mediator, it was shown to take 2.3 electrons to reduce the FAD cofactor in short-form ALR (Farrell and Thorpe, 2005). At the midpoint of FAD reduction, 25% of the short-form ALR accumulates as a blue-semiquinone; this is unexpected because flavin-linked oxidases are expected to form the red, anionic, flavosemiquinone, not the blue, neutral radical (Massey, 2000; Massey and Palmer, 1966). The redox potential of the FAD cofactor was determined to be  $-178 \pm 2$  mV by reducing ALR with dithionite in the presence of indicator dyes (Farrell and Thorpe, 2005). Since the flavin cofactor is essentially completely reduced on the addition of 2 reducing equivalence, the proximal disulfide must have a substantially more negative redox potential than the FAD (see Chapter 4). Farrell and Thorpe also showed that cytochrome *c* is a 100-fold better substrate for ALR than to oxygen (Farrell and Thorpe, 2005). Although the studies carried out by Farrell and Thorpe used the short-form ALR, a number of mechanistic conclusions are likely to be shared with the long-form of the enzyme since they share an identical catalytic core.

### 3.2.1 Non-Physiological Substrates of ALR

GSH, reduced RNase A and reduced lysozyme are very poor substrates of ALR, unlike QSOX, which prefers reduced unfolded proteins as substrates *in vitro* (Coddling et al., 2012; Daithankar et al., 2009; Lisowsky et al., 2001). Since the physiological substrate for short-form ALR is unclear, DTT is often used a model non-physiological substrate. Turnover with DTT yields  $K_M$  values near 1.6 mM and 3 mM and  $k_{cat}$  values of  $108 \text{ min}^{-1}$  and  $61 \text{ min}^{-1}$  for the short- and long-form ALR, respectively (Daithankar et al., 2010). Hence  $k_{cat}/K_M$  values show the long-form ALR

is 3x less catalytically proficient with DTT compared to the short-form (Daithankar et al., 2009).

### **3.2.2 Superoxide Generation by ALR**

The interest of this laboratory in the use and development of water-soluble phosphines as disulfide reductants (Cline et al., 2004) led to the discovery that several ALR and ERV family members release relatively high levels of the superoxide anion during aerobic turnover (Daithankar et al., 2012). This was unusual because flavin-dependent oxidases are expected to generate negligible levels of superoxide (Fridovich and Handler, 1961; Massey et al., 1969). Thus it was found that the model phosphine substrates THP and TCEP led to anomalously high turnover number when measured in the oxygen electrode, but that this anomaly disappeared with the inclusion of superoxide dismutase (Daithankar et al., 2012). Model studies with a xanthine/xanthine oxidase as a radical generation system confirmed that superoxide initiated a radical chain reaction that was driven by the oxidation of phosphines. In fact, ERV2 and short- and long-form ALR can release about 1/3 of reduced oxygen species as superoxide ion. In contrast, QSOX and glucose oxidase release negligible amounts of superoxide. While ALR and ERV release superoxide using DTT, this dithiol does not sustain the radical chain reaction and so the phenomenon previously escaped detection.

### **3.3 Tissue Specificity of ALR**

ALR is expressed at low levels in almost all tissues, for example: heart, spleen, lung, kidney, and skeletal muscle but is found in large amounts in the liver, in all regions of the brain (long-form only) and in testis (Gandhi, 2012; Hagiya et al., 1994;

Tury et al., 2005). ALR is secreted continuously by hepatocytes (Gandhi et al., 1999) and is expressed continuously in an inactive form and is activated upon secretion via an unknown mechanism during liver regeneration (LaBrecque and Pesch, 1975; Li et al., 2002). Additionally, ALR is overexpressed in cell proliferative diseases and conditions such as hepatocellular carcinoma (HCC), chronic hepatitis, intra hepatic cholangiocellular carcinoma or in regenerating liver after partial hepatectomy in rats (Polimeno et al., 2011; Yang et al., 1997; Yu et al., 2010).

### **3.4 Short-Form ALR**

This subsection will focus on the biological role of short-form ALR. ALR has been shown to be protective against multiple types of induced apoptosis. Li et al. showed that ALR activates the MAP kinase pathway by inducing phosphorylation of MAP kinase kinase and MAP kinase and by stimulating tyrosine phosphorylation of epidermal growth factor receptor (Li et al., 2000). The short-form ALR has a non-mitochondrial role in spermatogenesis (Klissenbauer et al., 2002).

Following 70% partial hepatectomy (PH) ALR levels in the serum increase as well as levels of other growth factors (Fausto et al., 1995; Francavilla et al., 1994; Michalopoulos, 1992). When the levels of ALR and other growth factors peak (approximately 12 hours post PH) a rapid and significant suppression of natural killer cells is observed (Francavilla et al., 1997). This suggests that ALR is protecting the liver against leukocytes during liver regeneration.

#### **3.4.1 ALR is Protective Against Hydrogen Peroxide Induced Apoptosis**

High levels of H<sub>2</sub>O<sub>2</sub> in eukaryotic cells causes lipid peroxidation and DNA damage eventually leading to apoptosis. ALR has been shown to be protective against

H<sub>2</sub>O<sub>2</sub>-induced apoptosis in human neuroblastoma and EEL-7402 hepatoma cells (Polimeno et al., 2009; Wu et al., 2007). Exogenous administration of recombinant ALR after PH down-regulates mRNA expression of Bax and Casp3, which are pro-apoptotic genes, and up-regulates Bcl-2 mRNA, which is an antiapoptotic gene that regulates mitochondrial apoptosis and prevents mitochondrial swelling. Concurrently, there is a decrease in malondialdehyde which measures lipid peroxidation, suggesting that ALR prevents protein oxidation (Polimeno et al., 2011). Additionally, reactive oxygen species (ROS), lipid peroxidation and protein carbonylation levels increase after PH and treatment with recombinant ALR reduces the levels of all of these (Polimeno et al., 2011).

#### **3.4.2 ALR is Protective Against Radiation-Induced Apoptosis**

ALR has been shown to be protective against radiation-induced apoptosis in SMMC-7721 hepatoma cell, and to preserve mitochondrial structure and reduce cytochrome *c* release (Cao et al., 2009). Normally these protective properties would be distinctly advantageous, but the overexpression of ALR in hepatocellular carcinoma (HCC) results in the cancer being less sensitive to radiation treatment (Cao et al., 2009). Silencing of ALR makes a radio-resistant HepG2 hepatoma cell strain sensitive to radiation making ALR a possible cancer target (Cao et al., 2009).

#### **3.4.3 siRNA for ALR**

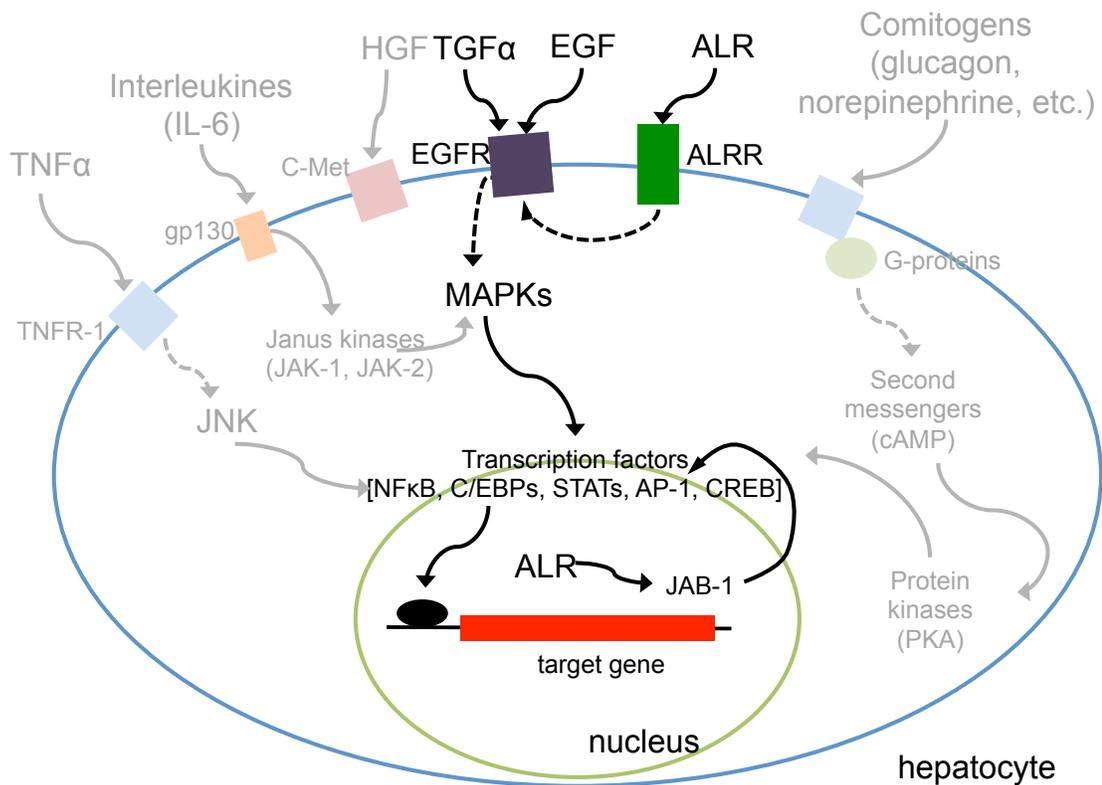
siRNA was used to significantly reduce the levels of ALR in glioma cells (Polimeno et al., 2012). There was a reduction in clusterin and Bcl-2 in cells treated with the siRNA/ALR and a significant increase in ROS and caspase-9, compared to the control cells (Polimeno et al., 2012). The oxidation from the ROS could be

counteracted by supplementing increasing doses of recombinant ALR (1-100 ng/mL) to the culture medium (Polimeno et al., 2012).

#### **3.4.4 Receptor for ALR**

Extracellular ALR activates MAPK/ERK (extracellular signal-regulated kinase) pathway by interacting with a receptor on the surface of hepatocytes which is reversible and specific for ALR (Figure 3.2). A maximum response was seen 5 and 10 min after ALR treatment, for ERK and MAPK pathways respectively before returning to basal levels after 30 min (Li et al., 2000; Wang et al., 1999). The  $K_d$  of the ALR receptor is reported to be 2 pM with 10,000 sites/cell for rat hepatocytes and 0.7 pM with 55,000 sites per cell for HepG2 cells (a human hepatocellular carcinoma cell line); the  $K_d$  values correlate nicely to half-maximum dose of ALR activity (Wang et al., 1999). Using a tyrosine kinase inhibitor the authors conclude that EGF and ALR use different mechanisms to promote the phosphorylation of EGFR (Li et al., 2000). While the ALR receptor can be found on normal (primary cultured) or abnormal hepatocytes (HepG2 for example) no receptor was found on other non-hepatocyte cell lines including COS-7 kidney cells and GLC-82 lung cells (Wang et al., 1999).

Wang et al. report that  $^{125}\text{I}$  short-form ALR can be specifically cross-linked to a receptor on the surface of rat and human hepatocytes using bio-sulfosuccinimidyl suberate (Wang et al., 1999). Following non-reducing SDS-PAGE they observed a 90 kDa band and inferred that the receptor is about 75 kDa by subtracting the molecular weight of the monomer (about 15 kDa). This appears to be in error, because ALR is a disulfide linked homodimer and hence an approximately 60 kDa receptor was covalently cross-linked to a 30 kDa ALR dimer in their experiments.



**Figure 3.2 ALR's predicted role in liver regeneration.** This figure illustrates the complex cascade of events that take place during liver regeneration. ALR first binds to its receptor (ALRR) and activates the epidermal growth factor receptor (EGFR) by tyrosine phosphorylation of EGFR. The MAPKs initiate transcription factors which can be controlled in part by ALR and its interaction with JAB1. Part of the image was greyed out for clarity. Figure adapted from (Pawlowski and Jura, 2006).

### **3.4.5 ALR Binds to JAB1 and Leads to the Activation of AP-1**

The N-terminal section of short-form ALR binds to JAB1 (Jun activation-domain binding protein), which is a subunit in the COP9 signalosome (Figure 3.2). JAB1 is a co-activator of transcription factor AP-1 (activator protein-1) and acts as a mediator between ALR and AP-1 (localized to the nucleus). A paper from 2001 shows evidence that the proximal redox active CxxC motif in short-form ALR (C142 and C145) is essential for JAB1 binding and activity (Chen et al., 2003; Lu et al., 2002). When ALR binds to JAB1 the phosphorylation of c-Jun is increased and this activates AP-1 (Lu et al., 2002). This pathway is independent of the MAPK/ERK pathway. Knockdown of ALR results in decreased levels of p27<sup>kip1</sup> (JAB1 inhibitor) and overexpression of ALR increased levels of p27<sup>kip1</sup> (Teng et al., 2011). ALR modulates hematopoietic stem cells by binding to JAB1 to counter JAB1 mediated nuclear export and destabilization of p27<sup>kip1</sup> (results in cell cycle inhibition) (Teng et al., 2011). The overexpression of ALR will decrease the places p27<sup>kip1</sup> can associate with JAB1 and the excessive amounts of p27<sup>kip1</sup> are exported from the nucleus to the cytoplasm for degradation (Teng et al., 2011). In multiple types of cancer, including chronic myeloid leukemia, a marker for poor prognosis is elevated JAB1 or decreased levels of cytoplasmic p27<sup>kip1</sup> (Tomoda et al., 2005). Additionally, a recent study shows ALR is a regulator of hematopoietic stem cell proliferation by binding to JAB1 and thus inhibiting the nuclear export of p27<sup>kip1</sup> (Sankar and Means, 2011).

### **3.4.6 Treatment of Cells with Recombinant Human ALR**

A significant increase in HepG2 and primary human hepatocyte proliferation was seen upon incubation with recombinant human ALR (rhALR) (Ilowski et al., 2010). The authors also observed a slow increase the phosphorylation of ERK1/2 and

Akt followed by decrease to basal levels in primary human hepatocytes upon exposure to rhALR over the course of an hour (Ilowski et al., 2010). The kinetics of phosphorylation of ERK1/2 and Akt with rhALR is different than with recombinant human EGF (Ilowski et al., 2010). Additionally, treatment of primary human hepatocytes with inhibitors for EGF receptor and MEK saw a 0.4 and 0.3 fold decrease in phosphorylated ERK1/2, respectively. Addition of rhALR did not increase the levels of p-ERK1/2 (Ilowski et al., 2010). ALR also seems to be liver specific unlike EGF and HGF (Ilowski et al., 2010).

A 4 hour pre-treatment of rhALR (750 ng/mL (~48 nM)) reduces the level of cytochrome *c* released into the cytosol when primary human hepatocytes were treated with 100 mM ethanol (Ilowski et al., 2011). rhALR was able to combat the apoptotic effects of several compounds such as TRAIL, anti-Apo, TGF- $\beta$  and actinomycin D in either HepG2 cells or primary human hepatocytes (Ilowski et al., 2011). Interestingly, the protective effect of rhALR against TRAIL is only in hepatocytes (Ilowski et al., 2011). The authors also tested bronchial (BC1), colonic (SW480), gastric (GC1) and pancreatic (L3.6PL) cell lines.

Activated peripheral blood lymphocytes (PBL) showed an increased viability and reduction in early apoptosis when incubated with rhALR (Wang et al., 2013). Additionally, rhALR significantly increased levels of Bcl-2 and mitochondrial cytochrome *c* and decreased the levels of BAX and cytosolic cytochrome *c* (Wang et al., 2013) in agreement with the findings of Polimeno et al. (Polimeno et al., 2009). These results indicate ALR's possible role in the down regulation of apoptotic cell death (Polimeno et al., 2009; Wang et al., 2013).

### **3.4.7 ALR's Role in Polyamine Levels**

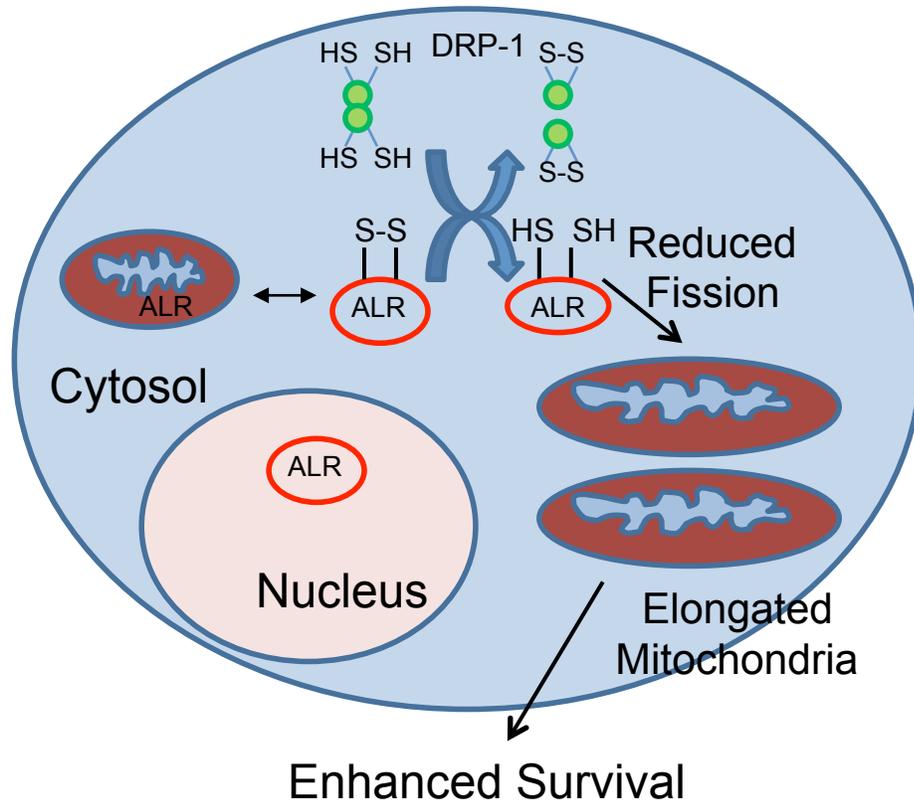
ALR may play a role in polyamine levels in human hepatocytes (Dayoub et al., 2006). Putrescine, spermidine and spermine are polyamines that are essential components in a variety of important biological pathways and are present in a variety of tissues and cell types. Dayoub et al. saw an increase in mRNA levels of ornithine decarboxylase (which produces putrescine and is the rate-limiting step in polyamine biosynthesis) and S-adenosylmethionine decarboxylase (which produces spermidine and spermine) (Dayoub et al., 2006). Putrescine levels increased 190% after cells were treated with rhALR for 9 hours (Dayoub et al., 2006). Polyamines can also be protective against liver damage caused by hepatotoxins (Tzirogiannis et al., 2004).

Exogenous treatment of hepatocytes with ALR also resulted in decreased levels of cytochrome P450 protein. This enzyme plays a key role in drug interactions and metabolism (Thasler et al., 2006). ALR is able to induce NF $\kappa$ -B activity and reduces constitutive androstane receptor expression in human hepatocytes (Thasler et al., 2005).

### **3.5 Long-Form ALR is Involved in Mitochondrial Fission and Dynamics**

A knockdown of ALR in embryonic stem cells resulted in abnormal mitochondria morphology and mitochondrial swelling or fragmentation, a change in the mitochondrial membrane potential and apoptosis (Todd et al., 2010a, 2010b). Sankar and coworkers found that levels of Drp1 (dynamin-like protein 1), which controls mitochondrial fission and dynamics, were elevated when ALR was down-regulated and levels of Drp1 were reduced when ALR was overexpressed (Figure 3.3) (Todd et al., 2010a, 2010b). Todd et al. hypothesize that ALR oxidizes the active, reduced form, of Drp1. With Drp1 inactivated, GTPase is, in turn, inactivated

resulting in enhanced survival of the cell (Todd et al., 2010a, 2010b). They further suggest that ALR may help maintain pluripotency or “stemness” of embryonic stem cells by preventing mitochondrial damage (Todd et al., 2010a, 2010b).



**Figure 3.3 Overexpression of ALR results in reduced mitochondrial fission.**

When ALR is overexpressed the levels of Drp1 are decreased resulting in reduced fission, elongation of the mitochondria and enhanced survival of the embryonic stem cells. Figure adapted from (Todd et al., 2010b).

### **3.5.1 ALR is Present in Cancerous Tissues**

Cytosolic and granular homogenates of normal and cirrhotic liver, as well as hepatocellular and cholangiocellular carcinoma tissues, show the expression of ALR (the 23 kDa isoform only) (Thasler et al., 2005). A variety of experiments using an antibody raised against the short-form ALR explored the localization of ALR in hepatocellular and cholangiocellular carcinomas (Thasler et al., 2005). Thasler et al. suggest that ALR may play a role in cirrhosis (Thasler et al., 2005).

### **3.5.2 Role in Organogenesis**

A recent paper studied the role ALR plays in vertebrate liver organogenesis. The authors utilized a morpholino antisense oligonucleotide to knockdown cytosolic and mitochondrial ALR (Li et al., 2012). Li et al. showed overexpression of wild type Zebrafish ALR and a C131S mutant (which corresponds to C145 in human ALR – the cysteine which directly communicates with the FAD cofactor) both promoted liver growth (Li et al., 2012). This observation led the authors to suggest that ALR utilizes an enzyme-dependent pathway (though AP-1) and an enzyme independent pathway (though MAPK) to promote liver growth (Li et al., 2012).

### **3.5.3 ALR Interacts with Migration Inhibitory Factor (MIF)**

Macrophage migration inhibitory factor (MIF) is a cytokine involved in immune response and inflammation. MIF (reductase activity) and ALR (oxidase activity) colocalize in the cytoplasm and have conserved CxxC motifs (Li et al., 2004). *In vitro* his pull-down and *in vivo* coimmunoprecipitation experiments suggest a direct interaction between MIF and ALR (Li et al., 2004). Both MIF and ALR can bind to JAB1 modulating the AP-1 pathway (Li et al., 2004). However, MIF and ALR have opposite actions as MIF reduces the growth promoting property of JAB1, and ALR

increases the effect of JAB1 on AP-1 (Li et al., 2004). In a yeast two-hybrid experiment it was shown that the binding ability of ALR to MIF was higher than that of ALR to JAB1 or MIF to JAB1 (Li et al., 2004). The overexpression of MIF increased the amount of ALR present in a dose dependent manner (Li et al., 2004).

#### **3.5.4 NMR Studies of Long-Form ALR and Mia40**

To date, no structure is available for the human long-form ALR. However, some useful studies have given more insight on the 80 amino acid N-terminal extension which is absent in the short-form of the enzyme (Figure 3.1). Surprisingly, the 80 amino acid peptide maintains substantially the same structure whether it is a peptide in solution or part of the long-form ALR as demonstrated by HSQC (Banci et al., 2013). Additionally, this same group showed only 16 of the 80 amino acids are required to interact with the 4-helix core (Banci et al., 2011). Even though it was previously demonstrated that the redox CxxC motif in the N-terminal region of long-form ALR was crucial for activity with Mia40 (Daithankar et al., 2009), Banci and coworkers set to demonstrate this using NMR techniques (Banci et al., 2011). Interestingly, these NMR experiments were done using yeast Mia40 and not human Mia40. It should be noted that yeast Mia40 and human Mia40 have widely differing redox potentials (-290 mV and -200 mV, respectively) and leading to the possibility that use of the non-cognate pair might mislead (Banci et al., 2009, 2011; Tienson et al., 2009).

### **3.6 Study of R194H Mutation in Human ALR**

It was recently found that three Moroccan children of related parents had a R194H mutation in both long and short-forms of ALR (Di Fonzo et al., 2009). The children presented with symptoms such as developmental delay, cataract, muscular hypotonia and hypotrophy. Comi and coworkers characterized this mutation in the homologous yeast protein ERV1, however, there are notable differences between the yeast and human proteins even though they share 42% sequence identity. For example, the yeast enzyme is not a covalent (disulfide linked) dimer unlike the human enzyme. Additionally, the distal (furthest from the FAD cofactor) CxxC motif in yeast ERV1 is located closer to the N-terminus as compared to the long-form of ALR. Using the human wild type enzyme and the enzyme with the R194H mutation, we conducted an enzymological characterization of the R194H mutant (Daithankar et al., 2010). Our work showed that the activity of the R194H enzyme is comparable to wild type, however, the mutant protein does not bind the FAD cofactor as tightly. The remaining sections of this Chapter describe the determination of the crystal structure of the human short-form ALR and its use in constructing a model of the R194H mutant. Finally, 2D NMR studies of both the wild type and mutant protein provide evidence for the surprising flexibility induced by the R194H mutation.

#### **3.6.1 Materials and Methods**

Reagents were obtained as described previously (Daithankar et al., 2009). Polyethylene glycol 8000 was from Thermo Fisher. Zinc acetate was from Allied Chemical and sodium cacodylate from SPI Chemical.

### **3.6.2 Primers**

Mutation of R194H for both lf- and short-form ALR used Quikchange site-directed mutagenesis (Stratagene). The primers (Integrated DNA Technologies) for mutating R194H were: forward: 5'- GCT CAA AAG TGG ATG AGC ACT GGC GCG ACG GCT GG -3' and reverse: 5'- CCA GCC GTC GCG CCA GTG CTC ATC CAC TTT TGA GC -3'

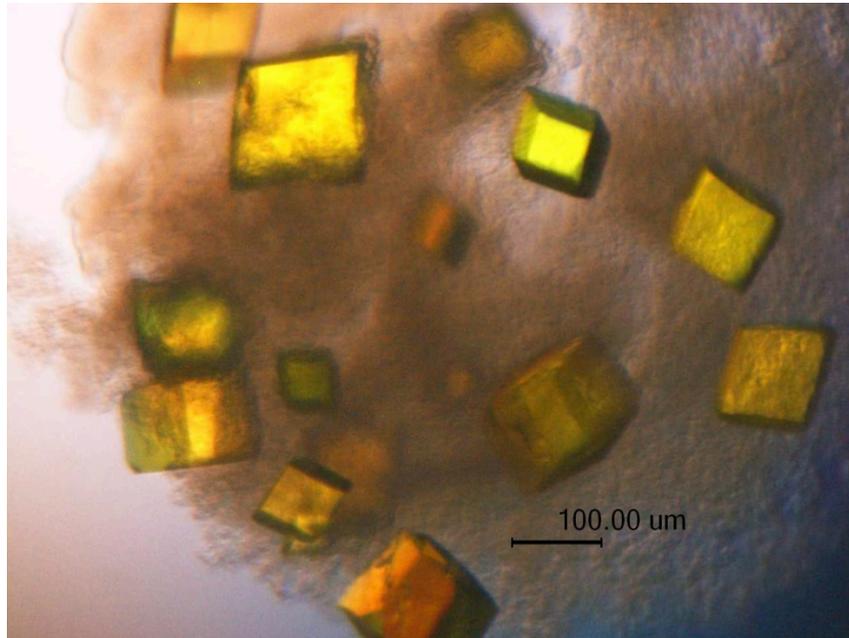
### **3.6.3 Mutation, Expression, and Purification of ALR and Arg194 Mutants**

The numbering of amino acid residues for short-form ALR followed that of long-form ALR. The full sequence of the histidine-tagged constructs of long-form and short-form ALR is shown in Figure 2.2. Mutagenesis was performed as described previously (Daithankar et al., 2009; Farrell and Thorpe, 2005) and confirmed by sequence analysis (Genewiz Inc.). As before, the two nonconserved and nonessential cysteines (C154 and C165) were mutated to alanine residues to avoid pronounced oxidative aggregation that was encountered in the early stages of work on human ALR (Daithankar et al., 2009). While two earlier papers (Daithankar et al., 2009; Farrell and Thorpe, 2005) designated these double mutants as ALR', we adopt the simpler abbreviation "ALR" here and refer to the former C154A/C165A protein as "wild type". Expression of <sup>15</sup>N- labeled human short-form ALR (using <sup>15</sup>NH<sub>4</sub>Cl, 99%; Cambridge Isotope Laboratories) followed the protocol of Marley et al. (Marley et al., 2001).

### **3.6.4 Crystallization and Data Collection**

Protein crystallization conditions were screened by hanging-drop vapor diffusion using Hampton Research crystal screening kits. Drops were generated at 25 °C by mixing 1 μL of the protein stock solution (8 mg/mL in 50 mM phosphate buffer,

pH 7.5) with 1  $\mu$ L of crystallization well solution. Promising conditions were refined to optimize crystal quality. The crystals used here (Figure 3.4) were grown using a solution of 18% (w/v) PEG 8000, 210 mM zinc acetate dihydrate, and 100 mM sodium cacodylate, pH 6.5, that had been filtered through a 0.2  $\mu$ m disposable filter unit (Nalgene). Crystals were dipped in a cryoprotectant mixture of the well solution containing 20% xylitol before flash-cooling in liquid nitrogen. Diffraction data were collected using in-house equipment (Rigaku RUH3R and R-AXIS IV). The crystal to detector distance was 100 mm, and the Cu radiation X-ray wavelength was 1.541 Å. Crystals were maintained at -180 °C, and data were collected for 15 min per 1° oscillation from a single crystal yielding a total of 180 diffraction images. The data were indexed, integrated, and scaled with the program HKL2000 (Otwinowski Z, 1997).



**Figure 3.4 Crystals of short-form ALR.** Crystals appeared within hours and were grown for 3 days at 25 °C in 18% w/v PEG 8000, 100 mM sodium cacodylate, pH 6.5, and 210 mM zinc acetate dihydrate. Figure used with permission from the publisher.

### 3.6.5 NMR Data

The  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC NMR spectra were recorded at 25 °C using a Bruker AV600 MHz spectrometer equipped with a cryoprobe operating at 600.13 and 60.81 MHz for  $^1\text{H}$  and  $^{15}\text{N}$ , respectively.  $^{15}\text{N}$ -Labeled wild type and R194H short-form ALR proteins were expressed and purified as described above. NMR samples were prepared in 5 mm tubes containing 130  $\mu\text{M}$  protein in 10 mM potassium phosphate buffer, pH 6.9, in 90%  $\text{H}_2\text{O}$ /10%  $\text{D}_2\text{O}$ . Acquisition parameters are listed in Table 3.1, and raw NMR data were processed using the NMRpipe program (Delaglio et al., 1995).

**Table 3.1 Acquisition parameters for the 2D  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC NMR experiments.**

Experiment	Points of acquired data		Spectral width (Hz)		Number of Scans
	$^1\text{H}$	$^{15}\text{N}$	$^1\text{H}$	$^{15}\text{N}$	
$^{15}\text{N}$ short-form ALR wt TROSY-HSQC	2048	400	9615.38	8389.26	128
$^{15}\text{N}$ short-form ALR R194H TROSY-HSQC	2048	400	9615.38	8389.26	128

### **3.6.6 Crystal Structure, Solution and Refinement.**

The human short-form ALR structure was solved by Dr. Ming Dong in the Bahnson laboratory using molecular replacement using the coordinates for the rat short-form ALR dimer (PDB code 1OQC). Molecular replacement was carried out using the program MOLREP from the CCP4 suite of programs (Bailey, 1994). Automated model building of human ALR employed ARP/wARP (Bailey, 1994), and 30 cycles of refinements utilized REFMAC5 and COOT (Emsley and Cowtan, 2004). Water molecules were placed during successive cycles of model building and refinement. The final human short-form ALR model (residues 81-205) had three subunits comprising one covalent dimer and half of another covalent dimer whose dimer interface resided on a crystallographic symmetry axis. The N-terminus of each subunit was disordered, with the final model including residues 94-205 from each subunit. A final  $2F_o - F_c$  electron density difference map confirmed the quality of the final model. The final  $R_{\text{working}}$  and  $R_{\text{free}}$  values were 0.189 and 0.235, respectively.

### **3.6.7 In Silico Modeling**

The R194H mutation of short-form ALR was modeled by Dr. Ming Dong of the Bahnson laboratory using the human short-form ALR structure using the program MODELER (Eswar et al., 2007; Pieper et al., 2009; Sali and Blundell, 1993), optimized with variable target function and molecular dynamics (Sali and Blundell, 1993), and further minimized using the CNS program (Brunger, 2007; Brünger et al., 1998). A total of 50 models were generated for the mutant R194H, and the structure with the lowest discrete optimized protein energy (DOPE) (Shen and Sali, 2006) was further minimized with CNS. DOPE values for the R194H model and the wild type structure were comparable. The minimized model has rmsd of 0.013 Å and 1.74° pre-

and post-CNS minimization, indicating good convergence. Solvent accessibilities for amino acid side chains and atoms were calculated using the NACCESS program (Hubbard and Thornton, 1993).

### **3.7 Location of R194 in Human ALR**

Our previous studies of human ALR (Daithankar et al., 2009; Farrell and Thorpe, 2005; Kay et al., 2006) have used a homology model for the short-form enzyme that was based on the rat short-form ALR dimer structure of Wu et al. (Wu et al., 2003). This model suggests that the R194 side chain of ALR would participate in several contacts, both with the ribose moiety of the FAD and with main chain atoms at the subunit interface (see later). This multiplicity of interactions suggested that an evaluation of the effect of the R194H mutation would be best undertaken with the cognate wild type protein: human ALR. While we have yet to be successful in crystallizing human long-form ALR, with its 80-residue N-terminal extension, crystals of the short-form enzyme were readily obtained (Figure 3.4) and used for the structural determination described below. Although it is possible that some effects of the R194H mutant are unique to the long-form of human ALR, none of the data presented here shows a selective impact between long and short versions of this flavoenzyme.

### **3.8 Crystal Structure of Human Short-Form ALR.**

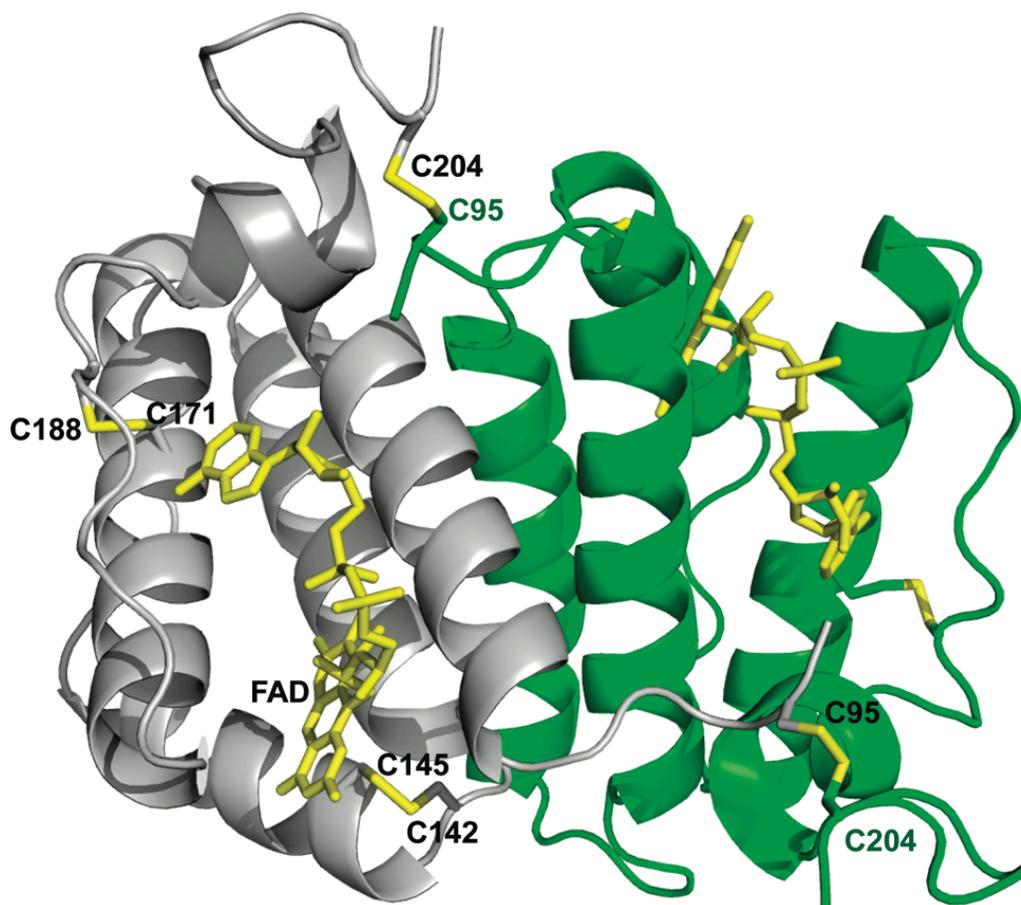
Human short-form ALR crystallized rapidly without removal of the N-terminal histidine affinity tag. The initial structure was solved using the rat short-form ALR (PDB code 1OQC) as a molecular replacement search model. The asymmetric unit consisted of three subunits: a disulfide-linked homodimer and half of an adjacent

dimer. In all subunits, the first 27 N-terminal residues of the short-form ALR construct were disordered. A summary of the data collection and refinement statistics for the crystal structure, encompassing residues D94-D205, is presented in Table 3.2. The overall fold of the covalent human short-form ALR dimer is shown in Figure 3.5. Figure 3.6, shows a difference electron density map surrounding the bound FAD prosthetic group ( $2F_o - F_c$ ) of this 1.85 Å structure. As would be expected from the level of protein sequence identity between rat and human short-form proteins (85%), both dimers overlay very well (with an RMSD of 0.531 Å; Figure 3.7).

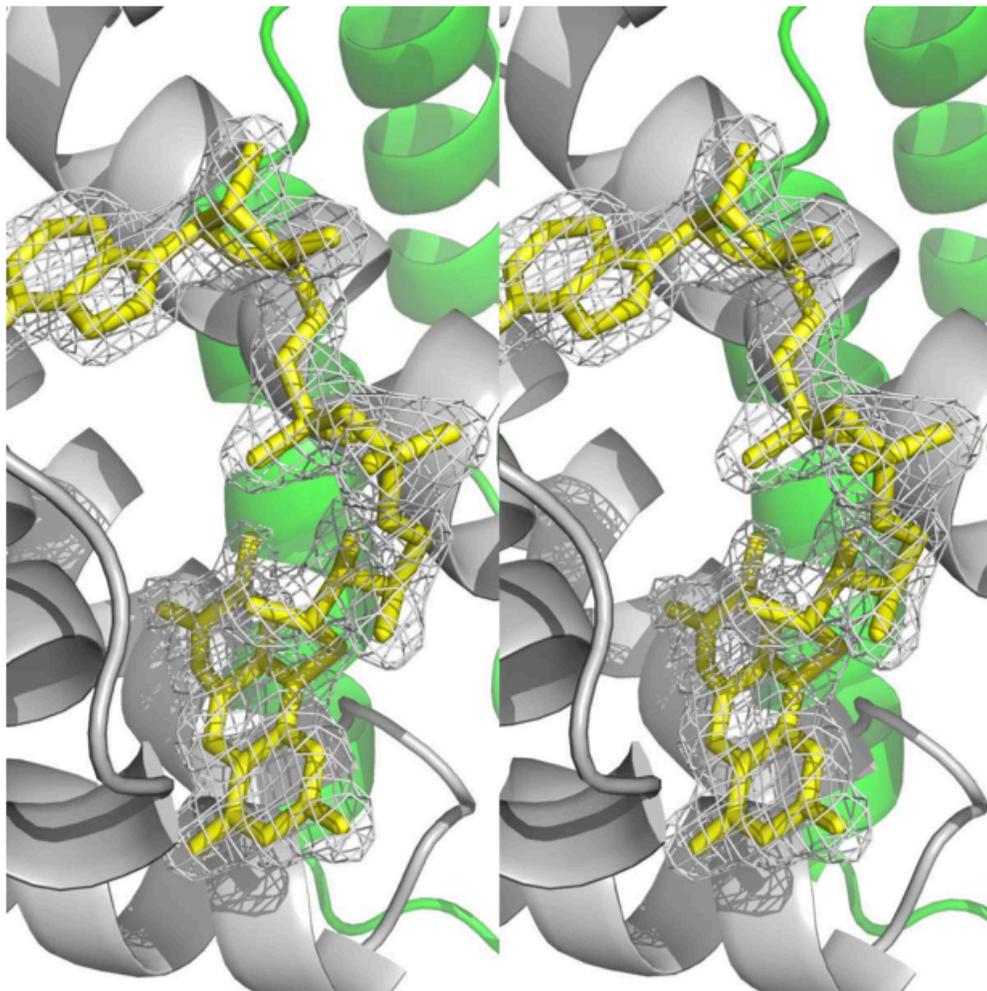
**Table 3.2 Data collection and refinement statistics of human short-form ALR<sup>a</sup>.**

Data Collection	
Space group	<i>C</i> 2
Unit cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	112.719, 65.145, 63.767
$\alpha$ , $\beta$ , $\gamma$ (deg)	90.0, 89.973, 90.0
Resolution (Å)	50.0-1.85 (1.92-1.85) <sup>b</sup>
Completeness (%)	99.9 (100.0)
Redundancy	3.6 (3.6)
<i>I</i> / $\sigma$ <i>I</i>	36.54 (3.13)
<i>R</i> <sub>merge</sub> linear <sup>c</sup>	0.044 (0.366)
Refinement	
Resolution (Å)	63.76-1.85
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> <sup>d</sup>	0.189/0.235
No. of atoms (non-hydrogen)	3429
Mean <i>B</i> value	31.9
RMSD bond lengths (Å)	0.017
RMSD bond angles (deg)	1.94

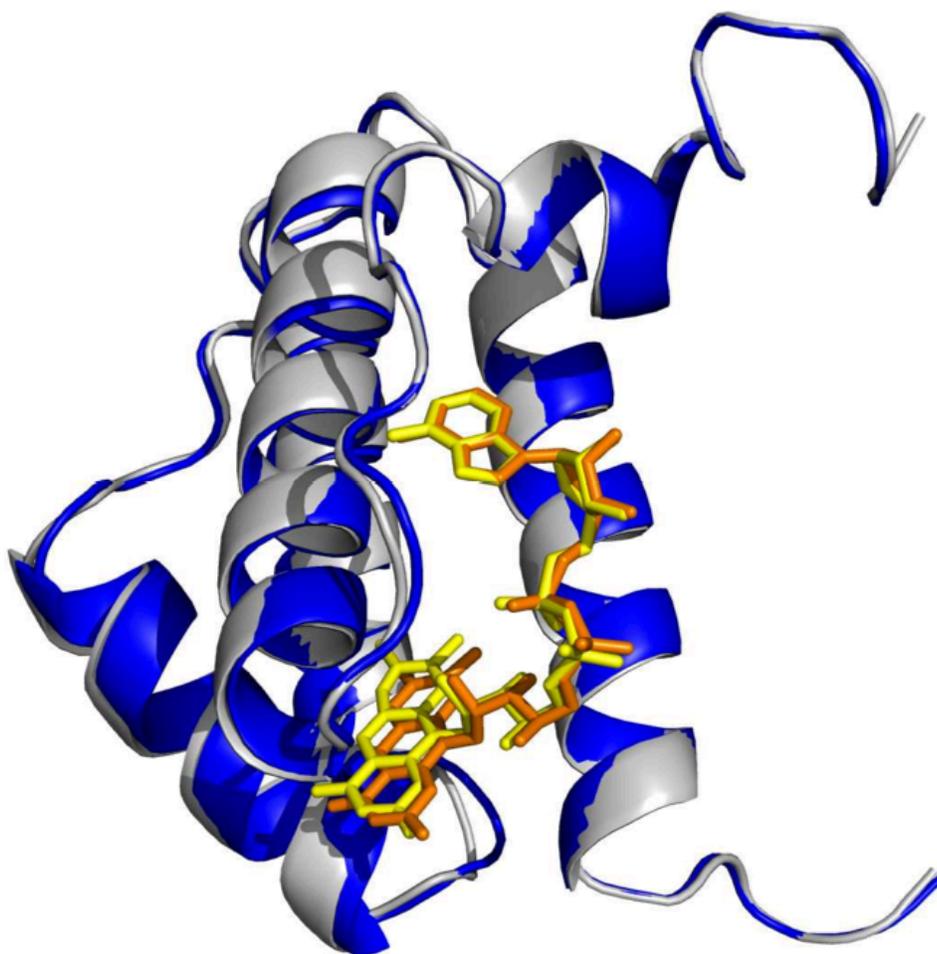
<sup>a</sup>Protein Data Bank accession code 3MBG. <sup>b</sup>Values in parentheses are for the highest resolution shell. <sup>c</sup> $R_{\text{merge}} = \sum |I_o - I_a| / \sum (I_a)$ , where *I*<sub>o</sub> is the observed intensity and *I*<sub>a</sub> is the average intensity, the sums being taken over all symmetry-related reflections. <sup>d</sup> $R_{\text{working}} = \sum |F_o - F_c| / \sum (F_a)$ , where *F*<sub>o</sub> is the observed amplitude and *F*<sub>c</sub> is the calculated amplitude. *R*<sub>free</sub> is the equivalent of *R*<sub>working</sub>, except it is calculated for a randomly chosen set of reflections that were omitted (5%) from the refinement process.



**Figure 3.5 Overall chain fold of dimeric human short-form ALR.** The two C95-C204 disulfides that join the gray and green subunits of the short-form ALR homodimer are shown in yellow, together with the redox-active proximal disulfide (C142-C145) and the structural disulfide (C171-C188) in the gray subunit. The non-covalent FAD is depicted in yellow. Figure used with permission from the publisher.



**Figure 3.6 Stereoview of the FAD prosthetic group in human short-form ALR.** A difference electron density map (coefficients  $2F_o - F_c$ ) is shown contoured at 1.5 sigma surrounding the flavin (yellow) in human short-form ALR. The two subunits are depicted by gray and green ribbons. Figure used with permission from the publisher.

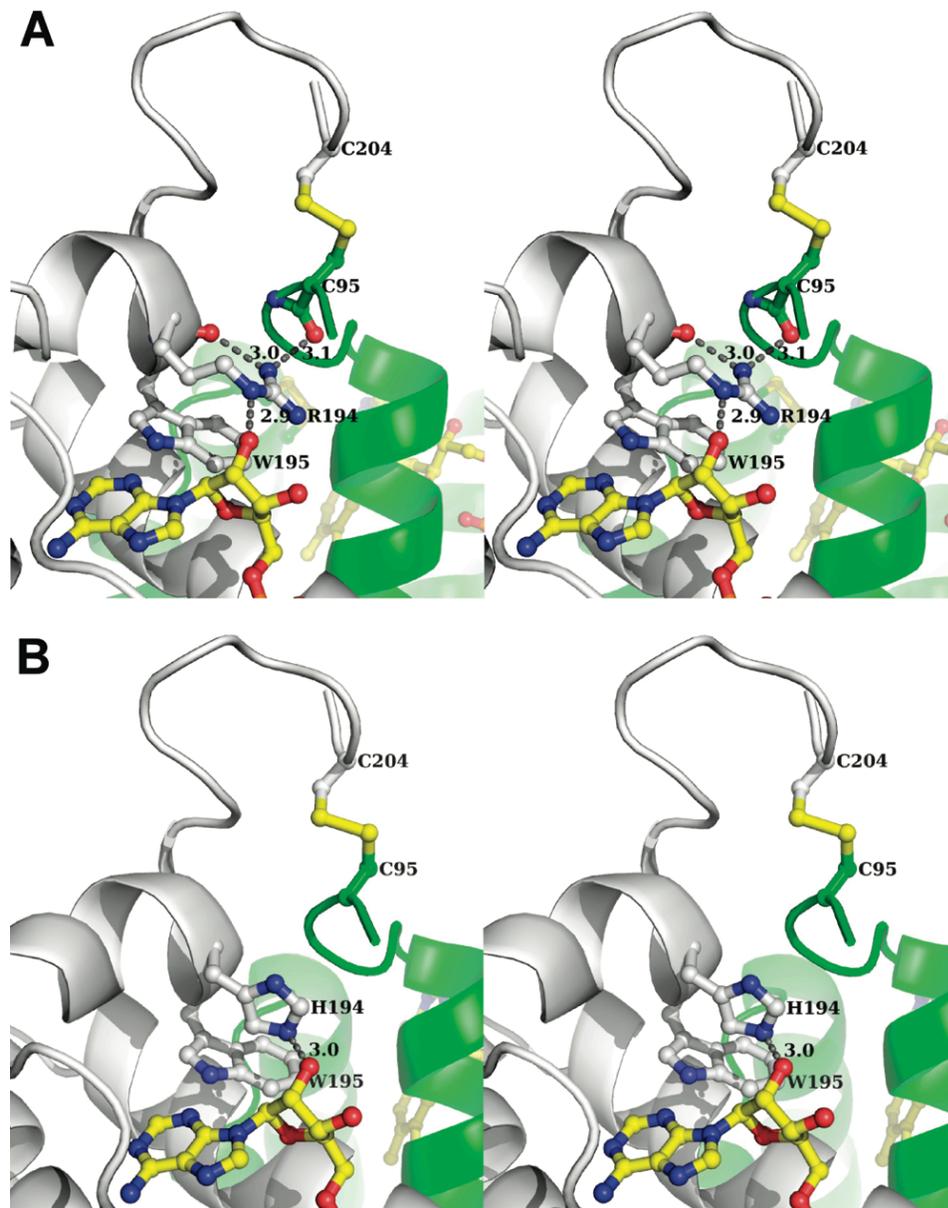


**Figure 3.7 Overlay of rat and human short-form ALR monomers.** The structure of a monomer of human short-form ALR from this work (gray ribbon, FAD in yellow; PDB code 3MBG) is superimposed on the corresponding rat short-form ALR structure (blue ribbon, FAD in orange; PDB code 1OQC). The two covalent dimers overlay closely with a RMSD of 0.531 Å. Figure used with permission from the publisher.

Figure 3.8A shows a stereoview of the human short-form ALR structure centered around R194. The residue is nearly coplanar with W195 suggestive of a cation- $\pi$  interaction. One of the terminal guanidino amino groups of R194 forms H-bonds with two main chain carbonyl oxygen atoms: one from R194 itself and the other contributed by C95 of the other subunit. This particular cysteine residue participates in the two interchain disulfides (C95-C204' and its counterpart, C204-C95') that maintain ALR as a covalent head-to-tail dimer (Figure 3.5). A third H-bond links the  $\delta$ -nitrogen of R194 to the 2'-OH of the ribose moiety of the bound FAD. This chain of interactions might reasonably be expected to link FAD binding with the conformational stabilization of the dimer interface in the vicinity of the interchain disulfide bridges.

Despite the expectation that the R194H mutation would prove structurally conservative, numerous crystallization screening attempts with the mutant protein proved unsuccessful. Indeed, the mutation has an unexpectedly large effect on protein stability and flavin binding (data not shown, (Daithankar et al., 2010)), and this probably contributes to our inability to crystallize the protein over a wide range of conditions. Hence, for comparison with the native protein, we prepared an energy-minimized model of R194H (Figure 3.8B). This model introduced very small changes to the backbone (with a  $C_{\alpha}$  RMSD of 0.244 Å) but generated significant perturbations in the region of residue 194. Here the imidazole ring of H194 is oriented at approximately right angles to the indole ring of W195. The model shows that H194 is now 3 Å from the 2'-OH of the FAD ribose in an orientation unfavorable for strong H-bond formation. While the importance of this potential H-bond in the R194H protein cannot be evaluated based on a modeled structure alone, the other two H-bonding

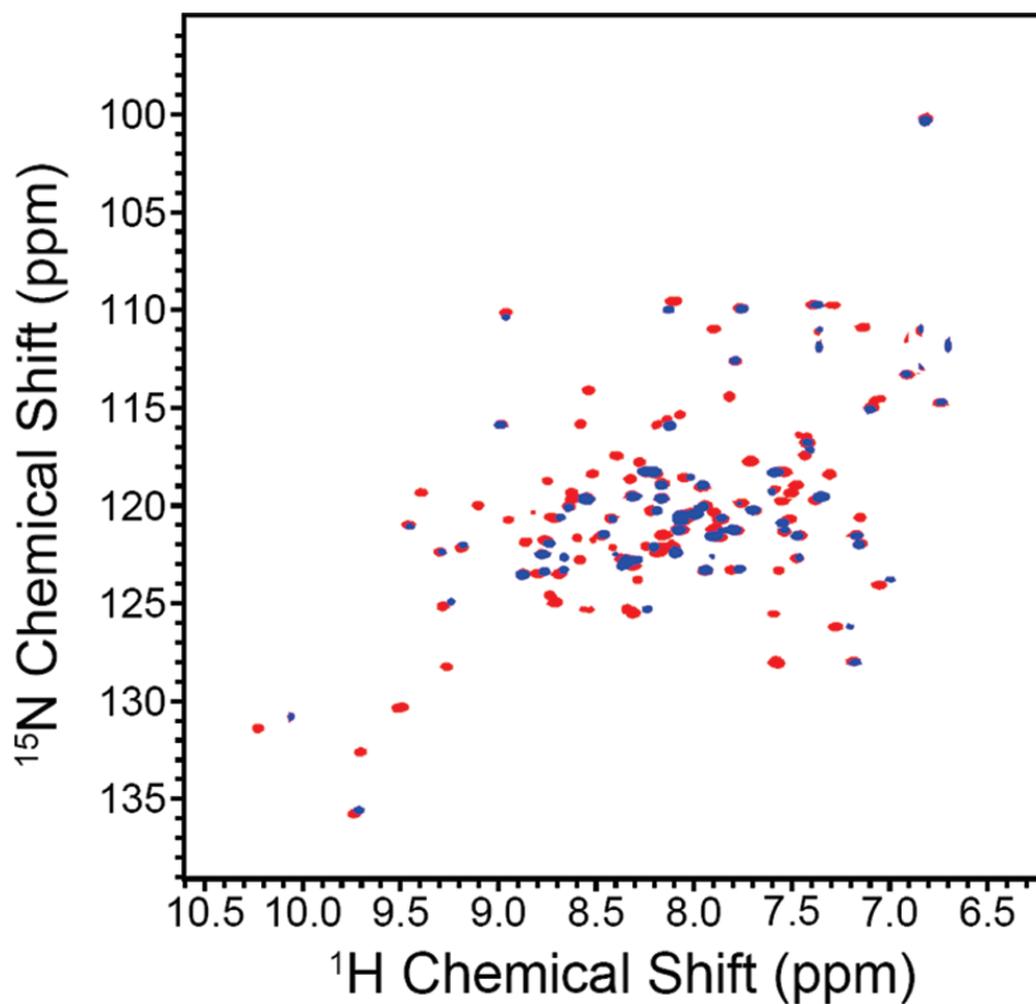
interactions identified for the wild type protein appear to be absent in the mutant (Figure 3.8B).



**Figure 3.8** A stereoview surrounding R194 in human short-form ALR and a minimized model of the R194H mutant. Panel A: R194 forms H-bonds with the 2'-OH of the ribose moiety of FAD, with its own main chain peptide carbonyl, and with the main chain carbonyl oxygen contributed by C95 of the other (green) subunit. Panel B represents a minimized model of the R194H short-form ALR mutant. Figure used with permission from the publisher.

### **3.9 Comparison Between Wild Type and R194H Mutants by TROSY-HSQC NMR.**

Figure 3.9 presents an overlay of 2D NMR spectra of  $^{15}\text{N}$ -labeled wild type (red) and R194H short-form ALR (blue) proteins acquired under identical conditions. The dispersion of chemical shifts shows that both proteins are well ordered, although the mutant shows considerably fewer resonances over the range depicted in Figure 3.9 (approximately 40% less). These absences are consistent with enhanced selective flexibility of the R194H mutant compared to the native protein.



**Figure 3.9** 2D  $^1\text{H}$ - $^{15}\text{N}$  TROSY-HSQC spectra of wild type and R194H mutant of short-form ALR. Wild type (red) and R194H (blue) spectra were acquired in 10 mM phosphate buffer, pH 6.9 at 25 °C. Figure used with permission from the publisher.

### 3.10 Conclusions

Recently, Di Fonzo et al. (Di Fonzo et al., 2009) described an R194H mutation of human ALR that led to cataract, progressive muscle hypotonia, and hearing loss in three children. The current work presents a structural characterization of the human wild type and the R194H ALR. R194 is located at the subunit interface of short-form ALR, close to the intersubunit disulfide bridges. The R194 guanidino moiety participates in three H-bonds: two main-chain carbonyl oxygen atoms (from R194 itself and from C95 of the intersubunit disulfide of the other protomer) and with the 2'-OH of the FAD ribose. The R194H mutation has minimal effect on the enzyme activity using model and physiological substrates of short and long ALR forms (Daithankar et al., 2010). However, the mutation adversely affects the stability of both ALR forms: e.g., by decreasing the melting temperature by about 10 °C, by increasing the rate of dissociation of FAD from the holoenzyme by about 45-fold, and by strongly enhancing the susceptibility of short-form ALR to partial proteolysis and to reduction of its intersubunit disulfide bridges by glutathione (data not shown; (Daithankar et al., 2010)). Finally, a comparison of the TROSY-HSQC 2D NMR spectra of wild type short-form ALR and its R194H mutant reveals a significant increase in conformational flexibility in the mutant protein. In sum, these *in vitro* data document the major impact of the seemingly conservative R194H mutation on the stability of dimeric ALR and provide an enzymological and structural rationalization of the *in vivo* observations of Di Fonzo et al. (Di Fonzo et al., 2009).

## REFERENCES

- Bailey, S. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* *50*, 760–763.
- Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Martinelli, M., Sideris, D.P., Katrakili, N., and Tokatlidis, K. (2009). MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. *Nat. Struct. Mol. Biol.* *16*, 198–206.
- Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Kallergi, E., Lionaki, E., Pozidis, C., and Tokatlidis, K. (2011). Molecular recognition and substrate mimicry drive the electron-transfer process between MIA40 and ALR. *Proc Natl Acad Sci U S A* *108*, 4811–4816.
- Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gajda, K., Felli, I.C., Gallo, A., Pavelkova, A., Kallergi, E., Andreadaki, M., et al. (2013). An intrinsically disordered domain has a dual function coupled to compartment-dependent redox control. *J. Mol. Biol.* *425*, 594–608.
- Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* *2*, 2728–2733.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* *54*, 905–921.
- Cao, Y., Fu, Y.L., Yu, M., Yue, P.B., Ge, C.H., Xu, W.X., Zhan, Y.Q., Li, C.Y., Li, W., Wang, X.H., et al. (2009). Human augments liver regeneration (hALR) is important for hepatoma cell viability and resistance to radiation-induced oxidative stress. *Free Radic Biol Med.*
- Chen, X., Li, Y., Wei, K., Li, L., Liu, W., Zhu, Y., Qiu, Z., and He, F. (2003). The potentiation role of hepatopoietin on activator protein-1 is dependent on its sulfhydryl oxidase activity. *J Biol Chem* *278*, 49022–49030.

- Cline, D.J., Redding, S.E., Brohawn, S.G., Psathas, J.N., Schneider, J.P., and Thorpe, C. (2004). New water-soluble phosphines as reductants of peptide and protein disulfide bonds: reactivity and membrane permeability. *Biochemistry* *43*, 15195–15203.
- Codding, J.A., Israel, B.A., and Thorpe, C. (2012). Protein substrate discrimination in the quiescin sulfhydryl oxidase (QSOX) family. *Biochemistry* *51*, 4226–4235.
- Daithankar, V., Farrell, S., and Thorpe, C. (2009). Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space. *Biochemistry* *48*, 4828–4837.
- Daithankar, V., Schaefer, S.A., Dong, M., Bahnson, B.J., and Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augmenter of liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* *49*, 6737–6745.
- Daithankar, V.N., Wang, W., Trujillo, J.R., and Thorpe, C. (2012). Flavin-linked Erv-family sulfhydryl oxidases release superoxide anion during catalytic turnover. *Biochemistry* *51*, 265–272.
- Dayoub, R., Thasler, W.E., Bosserhoff, A.K., Singer, T., Jauch, K.W., Schlitt, H.J., and Weiss, T.S. (2006). Regulation of polyamine synthesis in human hepatocytes by hepatotrophic factor augmenter of liver regeneration. *Biochem Biophys Res Commun* *345*, 181–187.
- Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* *6*, 277–293.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* *60*, 2126–2132.
- Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M.S., Eramian, D., Shen, M.-Y., Pieper, U., and Sali, A. (2007). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci. Chapter 2*, Unit 2.9.
- Farrell, S.R., and Thorpe, C. (2005). Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry* *44*, 1532–1541.
- Fausto, N., Laird, A.D., and Webber, E.M. (1995). Liver regeneration. 2. Role of growth factors and cytokines in hepatic regeneration. *FASEB J.* *9*, 1527–1536.

Di Fonzo, A., Ronchi, D., Lodi, T., Fassone, E., Tigano, M., Lamperti, C., Corti, S., Bordoni, A., Fortunato, F., Nizzardo, M., et al. (2009). The mitochondrial disulfide relay system protein GFER is mutated in autosomal-recessive myopathy with cataract and combined respiratory-chain deficiency. *Am J Hum Genet* 84, 594–604.

Francavilla, A., Hagiya, M., Porter, K.A., Polimeno, L., Ihara, I., and Starzl, T.E. (1994). Augmenter of liver regeneration: its place in the universe of hepatic growth factors. *Hepatology* 20, 747–757.

Francavilla, A., Vujanovic, N.L., Polimeno, L., Azzarone, A., Iacobellis, A., Deleo, A., Hagiya, M., Whiteside, T.L., and Starzl, T.E. (1997). The in vivo effect of hepatotrophic factors augmenter of liver regeneration, hepatocyte growth factor, and insulin-like growth factor-II on liver natural killer cell functions. *Hepatology* 25, 411–415.

Fridovich, I., and Handler, P. (1961). Detection of free radicals generated during enzymic oxidations by the initiation of sulfite oxidation. *J. Biol. Chem.* 236, 1836–1840.

Gandhi, C.R. (2012). Augmenter of liver regeneration. *Fibrogenesis Tissue Repair* 5, 10.

Gandhi, C.R., Kuddus, R., Subbotin, V.M., Prelich, J., Murase, N., Rao, A.S., Nalesnik, M.A., Watkins, S.C., DeLeo, A., Trucco, M., et al. (1999). A fresh look at augmenter of liver regeneration in rats. *Hepatology* 29, 1435–1445.

Hagiya, M., Francavilla, A., Polimeno, L., Ihara, I., Sakai, H., Seki, T., Shimonishi, M., Porter, K.A., and Starzl, T.E. (1994). Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: expression of biologically active recombinant ALR and demonstration of tissue distribution. *Proc Natl Acad Sci U S A* 91, 8142–8146.

Hubbard, S., and Thornton, J. (1993). NACCESS, computer program, Department of Biochemistry and Molecular Biology, University College, London.

Ilowski, M., Putz, C., Weiss, T.S., Brand, S., Jauch, K.W., Hengstler, J.G., and Thasler, W.E. (2010). Augmenter of liver regeneration causes different kinetics of ERK1/2 and Akt/PKB phosphorylation than EGF and induces hepatocyte proliferation in an EGF receptor independent and liver specific manner. *Biochem Biophys Res Commun* 394, 915–920.

- Ilowski, M., Kleespies, A., de Toni, E.N., Donabauer, B., Jauch, K.W., Hengstler, J.G., and Thasler, W.E. (2011). Augmenter of liver regeneration (ALR) protects human hepatocytes against apoptosis. *Biochem Biophys Res Commun* 404, 148–152.
- Kay, C.W.M., Elsässer, C., Bittl, R., Farrell, S.R., and Thorpe, C. (2006). Determination of the distance between the two neutral flavin radicals in augmenter of liver regeneration by pulsed ELDOR. *J. Am. Chem. Soc.* 128, 76–77.
- Klissenbauer, M., Winters, S., Heinlein, U.A., and Lisowsky, T. (2002). Accumulation of the mitochondrial form of the sulphydryl oxidase Erv1p/Alrp during the early stages of spermatogenesis. *J Exp Biol* 205, 1979–1986.
- LaBrecque, D.R., and Pesch, L.A. (1975). Preparation and partial characterization of hepatic regenerative stimulator substance (SS) from rat liver. *J Physiol* 248, 273–284.
- Li, Y., Li, M., Xing, G., Hu, Z., Wang, Q., Dong, C., Wei, H., Fan, G., Chen, J., Yang, X., et al. (2000). Stimulation of the mitogen-activated protein kinase cascade and tyrosine phosphorylation of the epidermal growth factor receptor by hepatopoietin. *J Biol Chem* 275, 37443–37447.
- Li, Y., Wei, K., Lu, C., Li, M., Xing, G., Wei, H., Wang, Q., Chen, J., Wu, C., Chen, H., et al. (2002). Identification of hepatopoietin dimerization, its interacting regions and alternative splicing of its transcription. *Eur J Biochem* 269, 3888–3893.
- Li, Y., Lu, C., Xing, G., Zhu, Y., and He, F. (2004). Macrophage migration inhibitory factor directly interacts with hepatopoietin and regulates the proliferation of hepatoma cell. *Exp Cell Res* 300, 379–387.
- Li, Y., Farooq, M., Sheng, D., Chandramouli, C., Lan, T., Mahajan, N.K., Kini, R.M., Hong, Y., Lisowsky, T., and Ge, R. (2012). Augmenter of liver regeneration (alr) promotes liver outgrowth during zebrafish hepatogenesis. *PLoS One* 7, e30835.
- Lisowsky, T., Weinstat-Saslow, D.L., Barton, N., Reeders, S.T., and Schneider, M.C. (1995). A new human gene located in the PKD1 region of chromosome 16 is a functional homologue to ERV1 of yeast. *Genomics* 29, 690–697.
- Lisowsky, T., Lee, J.E., Polimeno, L., Francavilla, A., and Hofhaus, G. (2001). Mammalian augmenter of liver regeneration protein is a sulphydryl oxidase. *Dig Liver Dis* 33, 173–180.
- Lu, C., Li, Y., Zhao, Y., Xing, G., Tang, F., Wang, Q., Sun, Y., Wei, H., Yang, X., Wu, C., et al. (2002). Intracrine hepatopoietin potentiates AP-1 activity through JAB1 independent of MAPK pathway. *FASEB J* 16, 90–92.

- Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* 20, 71–75.
- Massey, V. (2000). The chemical and biological versatility of riboflavin. *Biochem Soc Trans* 28, 283–296.
- Massey, V., and Palmer, G. (1966). On the existence of spectrally distinct classes of flavoprotein semiquinones. A new method for the quantitative production of flavoprotein semiquinones. *Biochemistry* 5, 3181–3189.
- Massey, V., Strickland, S., Mayhew, S.G., Howell, L.G., Engel, P.C., Matthews, R.G., Schuman, M., and Sullivan, P.A. (1969). The production of superoxide anion radicals in the reaction of reduced flavins and flavoproteins with molecular oxygen. *Biochem. Biophys. Res. Commun.* 36, 891–897.
- Michalopoulos, G. (1992). Liver regeneration and growth factors: old puzzles and new perspectives. *Lab. Invest.* 67, 413–415.
- Otwinowski Z, M.W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 276, 307–326.
- Pawlowski, R., and Jura, J. (2006). ALR and liver regeneration. *Mol Cell Biochem* 288, 159–169.
- Pieper, U., Eswar, N., Webb, B.M., Eramian, D., Kelly, L., Barkan, D.T., Carter, H., Mankoo, P., Karchin, R., Marti-Renom, M.A., et al. (2009). MODBASE, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* 37, D347–54.
- Polimeno, L., Pesetti, B., Lisowsky, T., Iannone, F., Resta, L., Giorgio, F., Mallamaci, R., Buttiglione, M., Santovito, D., Vitiello, F., et al. (2009). Protective effect of augments of liver regeneration on hydrogen peroxide-induced apoptosis in SH-SY5Y human neuroblastoma cells. *Free Radic Res* 43, 865–875.
- Polimeno, L., Pesetti, B., Annoscia, E., Giorgio, F., Francavilla, R., Lisowsky, T., Gentile, A., Rossi, R., Bucci, A., and Francavilla, A. (2011). Alrp, a survival factor that controls the apoptotic process of regenerating liver after partial hepatectomy in rats. *Free Radic. Res.* 45, 534–549.
- Polimeno, L., Pesetti, B., De Santis, F., Resta, L., Rossi, R., De Palma, A., Girardi, B., Amoroso, A., and Francavilla, A. (2012). Decreased expression of the augments of liver regeneration results in increased apoptosis and oxidative damage in human-derived glioma cells. *Cell Death Dis.* 3, e289.

- Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234, 779–815.
- Sankar, U., and Means, A.R. (2011). Gfer is a critical regulator of HSC proliferation. *Cell Cycle* 10, 2263–2268.
- Shen, M.-Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507–2524.
- Teng, E.C., Todd, L.R., Ribar, T.J., Lento, W., Dimascio, L., Means, A.R., and Sankar, U. (2011). Gfer inhibits Jab1-mediated degradation of p27kip1 to restrict proliferation of hematopoietic stem cells. *Mol. Biol. Cell* 22, 1312–1320.
- Thasler, W.E., Schlott, T., Thelen, P., Hellerbrand, C., Bataille, F., Lichtenauer, M., Schlitt, H.J., Jauch, K.W., and Weiss, T.S. (2005). Expression of augmenter of liver regeneration (ALR) in human liver cirrhosis and carcinoma. *Histopathology* 47, 57–66.
- Thasler, W.E., Dayoub, R., Muhlbauer, M., Hellerbrand, C., Singer, T., Grabe, A., Jauch, K.W., Schlitt, H.J., and Weiss, T.S. (2006). Repression of cytochrome P450 activity in human hepatocytes in vitro by a novel hepatotrophic factor, augmenter of liver regeneration. *J Pharmacol Exp Ther* 316, 822–829.
- Tienson, H.L., Dabir, D. V, Neal, S.E., Loo, R., Hasson, S.A., Boontheung, P., Kim, S.K., Loo, J.A., and Koehler, C.M. (2009). Reconstitution of the mia40-erv1 oxidative folding pathway for the small tim proteins. *Mol Biol Cell* 20, 3481–3490.
- Todd, L.R., Damin, M.N., Gomathinayagam, R., Horn, S.R., Means, A.R., and Sankar, U. (2010a). Growth Factor erv1-like Modulates Drp1 to Preserve Mitochondrial Dynamics and Function in Mouse Embryonic Stem Cells. *Mol Biol Cell*.
- Todd, L.R., Gomathinayagam, R., and Sankar, U. (2010b). A novel Gfer-Drp1 link in preserving mitochondrial dynamics and function in pluripotent stem cells. *Autophagy* 6, 821–822.
- Tomoda, K., Kato, J., Tatsumi, E., Takahashi, T., Matsuo, Y., and Yoneda-Kato, N. (2005). The Jab1/COP9 signalosome subcomplex is a downstream mediator of Bcr-Abl kinase activity and facilitates cell-cycle progression. *Blood* 105, 775–783.
- Tury, A., Mairet-Coello, G., Lisowsky, T., Griffond, B., and Fellmann, D. (2005). Expression of the sulfhydryl oxidase ALR (Augmenter of Liver Regeneration) in adult rat brain. *Brain Res* 1048, 87–97.

- Tzirogiannis, K.N., Panoutsopoulos, G.I., Demonakou, M.D., Papadimas, G.K., Kondyli, V.G., Kourentzi, K.T., Hereti, R.I., and Mykoniatis, M.G. (2004). The hepatoprotective effect of putrescine against cadmium-induced acute liver injury. *Arch. Toxicol.* 78, 321–329.
- Wang, G., Yang, X., Zhang, Y., Wang, Q., Chen, H., Wei, H., Xing, G., Xie, L., Hu, Z., Zhang, C., et al. (1999). Identification and characterization of receptor for mammalian hepatopoietin that is homologous to yeast ERV1. *J Biol Chem* 274, 11469–11472.
- Wang, N., Sun, H., Shen, Y., Li, X., Pan, T., Liu, G., and Liu, Q. (2013). Augmenter of liver regeneration inhibits apoptosis of activated human peripheral blood lymphocytes in vitro. *Immunopharmacol. Immunotoxicol.* 35, 257–263.
- Wu, C.K., Dailey, T.A., Dailey, H.A., Wang, B.C., and Rose, J.P. (2003). The crystal structure of augmenter of liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase. *Protein Sci* 12, 1109–1118.
- Wu, Y., Chen, L., Yu, H., Liu, H., and An, W. (2007). Transfection of hepatic stimulator substance gene desensitizes hepatoma cells to H<sub>2</sub>O<sub>2</sub>-induced cell apoptosis via preservation of mitochondria. *Arch. Biochem. Biophys.* 464, 48–56.
- Yang, X., Xie, L., Qiu, Z., Wu, Z., and He, F. (1997). Human augmenter of liver regeneration: Molecular cloning, biological activity and roles in liver regeneration. *Sci China C Life Sci* 40, 642–647.
- Yu, H.Y., Xiang, D.R., Huang, H.J., Li, J., and Sheng, J.F. (2010). Expression level of augmenter of liver regeneration in patients with hepatic failure and hepatocellular carcinoma. *Hepatobiliary Pancreat Dis Int* 9, 492–498.

## Chapter 4

# HUMAN AUGMENTER OF LIVER REGENERATION; PROBING THE CATALYTIC MECHANISM OF A FLAVIN-DEPENDENT SULFHYDRYL OXIDASE

### 4.1 Introduction

Chapters 1 and 2 introduced disulfide bond formation and the structure and function of ALR/ERV family members, while Chapter 3 focused on the function human ALR. This Chapter probes the catalytic mechanism of the short, cytokine, form of augmenter of liver regeneration (sfALR) using a model thiol substrate of the enzyme. The redox potential of the proximal CxxC motif was determined and is ~57 mV more reducing than the FAD cofactor. Additionally, we rationalized why glutathione is an ineffective monothiol substrate for ALR on the basis of the stringent steric requirements for thiol-disulfide exchange reactions.

### 4.2 Materials and Methods

Ampicillin, IPTG, DTT and PMSF were obtained from Gold Biotechnology Inc. Oxidized DTT, GnHCl and leupeptin were obtained from Sigma. Lysozyme and monobasic potassium phosphate were from Amresco. EDTA, LB broth and NaCl were from Fisher, and imidazole from Alfa Aesar.

#### 4.2.1 General Methods

The enzymatic turnover of sfALR was determined in duplicate or triplicate using a Clark-type oxygen electrode as previously described (Farrell and Thorpe, 2005). UV-VIS spectra were obtained as described previously (Daithankar et al., 2009, 2010; Farrell and Thorpe, 2005). Where necessary, turbidity correction software supplied with the HP8453 diode-array spectrophotometer was used to correct for slight light scattering of the spectra of wild type and mutant sfALR.

#### 4.2.2 Mutagenesis

Primers were purchased from Integrated DNA Technologies and mutagenesis was conducted as described previously (Daithankar et al., 2009; Farrell and Thorpe, 2005). DNA sequencing was performed by Genewiz Inc. The primers used in this work were: 5'- CTA AGT TTT ACC CCT GTG GGC CGT GTG CTG AAG ACC TAA G -3' and 5'- CCT AGG TCT TCA GCA CAC GGC CCA CAG GGG TAA AAC TTA G -3' for E143G/E144P;

5'- CTA AGT TTT ACC CCT GTC ATC GTT GTG CTG AAG ACC TAA G -3' and 5'- CCT AGG TCT TCA GCA CAG TGC GGA CAG GGG TAA AAC TTA G -3' for E143P/E144H;

5'- CTA AGT TTT ACC CCT GTA AGA AGT GTG CTG AAG ACC TAA G -3' and 5'- CCT AGG TCT TCA GCA CAC TTC TTA CAG GGG TAA AAC TTA G -3' for E143K/E144K.

5'- CTA AGT TTT ACC CCT GTA ACT GGT GTG CTG AAG ACC TAA G -3' and 5'- CCT AGG TCT TCA GCA CAC CAG TTA CAG GGG TAA AAC TTA G -3' for E143N/E144W;

5'- CTA AGT TTT ACC CCT GTG AGT GTG CTG AAG ACC TAA G -3'  
and 5'- CCT AGG TCT TCA GCA CAC TCA CAG GGG TAA AAC TTA G -3' for  
E143 deletion;

5'- CTA AGT TTT ACC CCT GTG AGG GCG AGT GTG CTG AAG ACC  
TAA G -3' and 5'- CCT AGG TCT TCA GCA CAC TCG CCC TCA CAG GGG  
TAA AAC TTA G -3' for E143 G E144 insertion; and

5'- CTA AGT TTT ACC CCT GTG AGT GGG AGT GTG CTG AAG ACC  
TAA G -3' and 5'- CCT AGG TCT TCA GCA CAC TCC CAC TCA CAG GGG  
TAA AAC TTA G -3' for E143 W E144 insertion.

#### **4.2.3 Expression and Purification of sfALR**

The pTrcHisA (Life Technologies, NY) expression vector (6xHis N-terminal affinity tag) encoded human short-form ALR (amino acids 81-205), in which two non-conserved and non-essential cysteine residues at positions 154 and 165 were mutated to alanine residues to minimize protein aggregation (Daithankar et al., 2009; Farrell and Thorpe, 2005). Expression was performed as previously described (Daithankar et al., 2010) utilizing 100 µg/mL ampicillin in the culture media. Cells were harvested at 5000 g for 8 min at 4 °C and immediately resuspended in 50 mM potassium phosphate, pH 7.5, containing 500 mM NaCl (binding buffer), flash frozen with liquid nitrogen and stored at -80 °C until purification.

Cells were thawed on ice and brought to 1 mM PMSF, 1 µM leupeptin and 0.1 mg/mL lysozyme immediately prior to two passes through a French pressure cell at 10000 psi. The lysed cells were sonicated briefly to shear DNA and the suspension centrifuged at 15000 g for 1 h. The lysate was applied to a 5 mL HisTrap FF column (GE Healthcare), pre-equilibrated with binding buffer, and the column washed with 15

column volumes of this buffer. ALR was eluted using binding buffer supplemented with 0.02, 0.05, 0.2, 0.5 and 1 M imidazole. Yellow fractions were combined and analyzed using 12% Tris-Glycine SDS PAGE under reducing and non-reducing conditions. ALR fractions were combined and dialyzed into 50 mM potassium phosphate, pH 7.5, containing 1 mM EDTA. Unless otherwise stated this buffer was used for subsequent characterization of sfALR.

#### **4.2.4 Preparation of the Apoprotein of sfALR and Reconstitution with 5-deaza-FAD**

Apoprotein was prepared as described by Daithankar et al. (Daithankar et al., 2010), exchanged with 50 mM potassium phosphate buffer, pH 7.0, containing 1 mM EDTA, and reconstituted with 1.2 equivalents of 5-deaza-FAD. After 15 min incubation on ice, excess flavin was removed with 4 washes of buffer (pH 7.0) in an Amicon Ultra 0.5 mL centrifugal filter device (Millipore). The extinction coefficient of 5-deaza-FAD substituted sfALR is  $11.5 \text{ mM}^{-1} \text{ cm}^{-1}$  at 408 nm and was determined as described previously (Daithankar et al., 2009).

#### **4.2.5 Redox Potential Experiments**

Solutions of 30  $\mu\text{M}$  5-deaza-FAD sfALR (0.13 mL in 50 mM potassium phosphate, 1 mM EDTA, pH 7.0, containing 20 mM oxidized DTT and 5 mM glucose) in a self-masking semimicro cuvette were incubated with 20 nM glucose oxidase and 5 nM catalase for 10 min to minimize subsequent aerobic oxidation of reduced DTT by residual traces of native ALR. Increments of 5 stock solutions of DTT (0.1, 1, 10, 20 and 100 mM) were added to yield final cuvette concentrations of up to 5.1 mM DTT. Spectral changes were complete 1 min after addition of DTT at 25 °C. Data were corrected for light scattering, reagent depletion and dilution and

plotted as a function of the ratio of reduced to oxidized DTT. The same procedure was followed for the mutant enzymes with the exception that 100 mM oxidized DTT was used for the CPHC mutant. Redox curves were fit using the Hill Fit equation in GraphPad Prism version 6.0c yielding  $K_{ox}$  and converted to redox potentials using the Nernst Equation using a value of -327 mV for the  $E^{\circ}$  of DTT (Lees and Whitesides, 1993).

#### **4.2.6 Stopped-Flow Spectrophotometry**

Stopped-flow experiments were carried out at 25 °C in a Hi-Tech Scientific SF-61 SX2 double mixing stopped-flow system using KinetAsyst software from TgK Scientific. Prior to anaerobic experiments, the flow cell and driving syringes were soaked overnight in anaerobic buffer containing 5 mM glucose and 50 nM glucose oxidase. For studies of the reductive half-reaction with DTT, sfALR (60  $\mu$ M in 50 mM phosphate buffer, pH 7.5 containing 1 mM EDTA, 5 mM glucose and 1 nM catalase) was deoxygenated in a tonometer prior to the anaerobic addition of 5 nM glucose oxidase. After further nitrogen/vacuum cycles, the tonometer was transferred to the stopped-flow instrument and mixed with various anaerobic solutions of DTT that had been deoxygenated in the same manner. DTT solutions were standardized using DTNB immediately prior to deoxygenation. The reduction of sfALR in the stopped-flow was followed in both diode-array and monochromator modes.

For the oxidative half-reaction, the 2-electron reduced protein (60  $\mu$ M ALR in 50 mM potassium phosphate, pH 7.5, containing 1 mM EDTA, 5 mM glucose and 1 nM glucose oxidase) was prepared anaerobically in a tonometer equipped with an ancillary quartz cuvette (Williams et al., 1979). A standardized solution of sodium dithionite was added from a gas-tight syringe and the progress of the titration was

followed spectrophotometrically to ensure generation of the 2-electron reduced enzyme without accumulation of excess reductant (Farrell and Thorpe, 2005). Syringes containing various oxygen concentrations in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA were prepared as in DuPlessis et al. (DuPlessis et al., 1998).

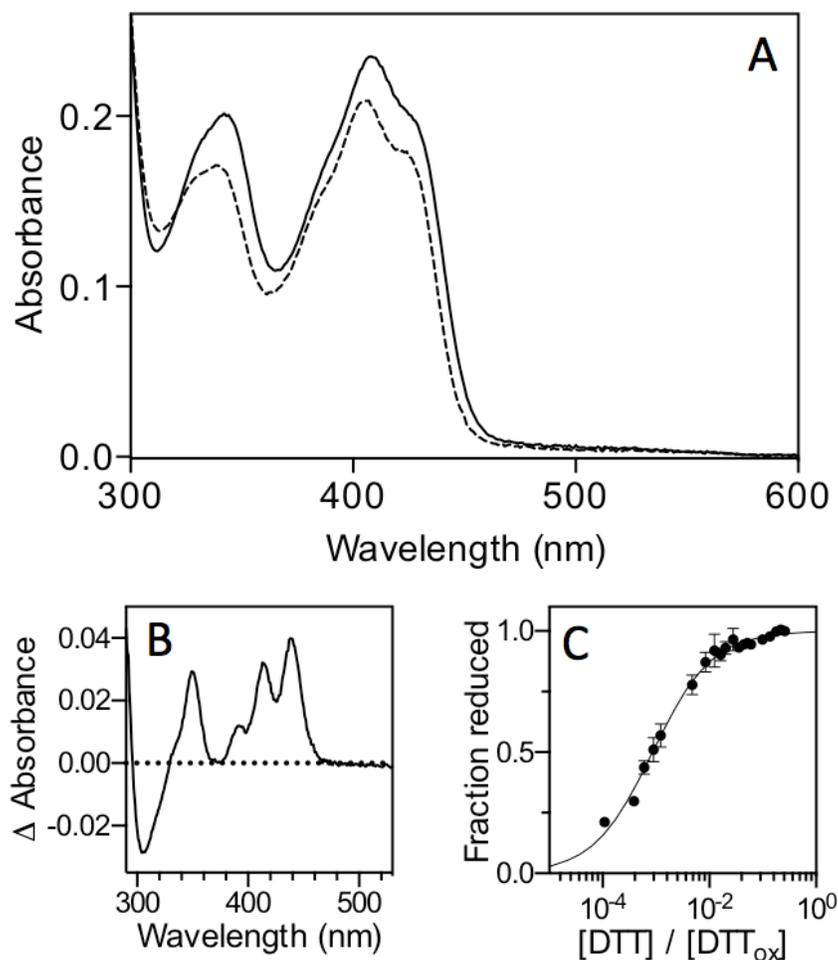
Rapid reaction data sets for sfALR were fit to the kinetic model presented in this Chapter using KinTek Explorer (Johnson et al., 2009a, 2009b). Observed rate constants are the average of at least 3 determinations. Analyses of the model were also run using Copasi Biochemical Network Simulator (Hoops et al., 2006).

### **4.3 Redox Potential of the Proximal Disulfide**

Anaerobic titrations of oxidized sfALR by sodium dithionite or with DTT both show full reduction of the bound FAD cofactor upon the addition of 2 electrons (Farrell and Thorpe, 2005). Similarly, back titration of reduced sfALR (prepared by photoreducing an anaerobic solution of the protein until the FAD chromophore was bleached) gave a stoichiometry of 2 molecules of the 1-electron oxidant ferricyanide for complete return of the flavin absorbance (Farrell and Thorpe, 2005). Collectively, these data show that the redox potential of the FAD is considerably more oxidizing than that of the proximal disulfide with which it communicates. Farrell and Thorpe measured the redox potential of the bound FAD moiety in sfALR, using equilibration with dyes, and found it to be  $-178 \pm 2$  mV (Farrell and Thorpe, 2005). In the context of understanding the redox communication between proximal disulfides and the flavin prosthetic groups of ERV/ALR family members, we then wanted to determine the redox potential of the redox active disulfide in sfALR. Since we wished to measure this potential in the context of oxidized flavin, we explored the preparation of sfALR

substituted with the much more reducing 5-deaza-FAD analog (Walsh et al., 1978). sfALR apoprotein was prepared by washing sfALR bound to Ni-IDA beads with 6 M GuHCl until the FAD content of eluates was negligible ((Daithankar et al., 2010); see Materials and Methods). The apoprotein was incubated with a slight molar excess of 5-deaza-FAD before excess free flavin was removed (see Materials and Methods). The spectrum of 5-deaza-FAD sfALR is shown in Figure 4.1. The oxidized flavin envelope shows a typical spectrum for a 5-deaza-flavin substituted protein with an absorbance maximum at 408 nm. While reduction with 5 mM DTT leads to a small decrease in the extinction coefficient of the oxidized 5-deaza-FAD, scission of the proximal disulfide leads to a 3 nm blue shift of the leading edge of the oxidized 5-deaza-flavin absorbance envelope (and to the difference spectrum Figure 4.1B). Such blue shifts upon removal of proximal disulfides, by mutagenesis or chemical modification, have been noted in a number of flavin-dependent disulfide oxidoreductases (for example, in lipoamide dehydrogenase (Thorpe and Williams, 1976a), QSOX (Brohawn et al., 2003; Heckler et al., 2008a), Erv2p (Wang et al., 2007) and ALR itself (Daithankar et al., 2009)). Panel C of Figure 4.1 follows this spectral change (at 439 nm) as a function of the composition of a redox buffer formed by mixing DTT with its oxidized counterpart (DTT<sub>ox</sub>; see Materials and Methods). The midpoint potential is  $-235 \pm 1.6$  mV at pH 7; a value some 57 mV more reducing than that of the normal flavin cofactor in sfALR (Farrell and Thorpe, 2005). Hence the equilibrium constant between 2-electron reduced forms of the enzyme in which either the FAD or the proximal disulfide undergoes 2-electron reduction is approximately 90-fold in favor of flavin reduction. This wide separation in redox potential provides confirmation of the results of reductive and oxidative titrations

described earlier in this section. Our attempts to repeat the redox potential measurements using GSH/GSSG buffers were thwarted by the extremely slow reduction of even the normal FAD-substituted ALR by GSH (see later).



**Figure 4.1 Determination of the redox potential of the proximal disulfide in sfALR substituted with 5-deaza-FAD.** Panel A shows the spectrum of 20  $\mu$ M 5-deaza-FAD enzyme (see Materials and Methods) in 50 mM phosphate buffer, pH 7.0, containing 1 mM EDTA recorded before and immediately after the addition of 5 mM DTT (solid and dashed lines respectively). Panel B shows the difference spectrum dominated at 439 nm by the blue shift in the leading edge of the oxidized 5-deazaflavin absorbance. Panel C follows the extent of these changes at 439 nm as a function of the redox poise of mixtures of reduced and oxidized DTT (see Materials and Methods).

While not the major focus of this Chapter, we also explored the consequences of making several active site mutations involving the intervening E143 and E144 residues in the CxxC motif. Surprisingly, substitution of the intervening EE dipeptide by the charge-reversed KK dipeptide showed minimal effect on the redox potential and the approximate 3-fold increase in catalytic efficiency is due to a change in the  $K_m$  for DTT (Table 4.1). Two other mutations were chosen because they represent strongly reducing and oxidizing sequences in another family of redox protein containing CxxC motifs. Thus in thioredoxin family members, CGPC is prototypical of the strongly reducing bacterial thioredoxins ( $E^{\circ'} = \sim -270$  mV; (Krause et al., 1991; Moore et al., 1964; Mössner et al., 1998) and, by contrast, the CPHC sequence is found in the highly oxidizing DsbA protein ( $E^{\circ'} = \sim -122$  mV; (Huber-Wunderlich and Glockshuber, 1998; Zapun et al., 1993)) located in the periplasm of gram negative bacteria. Mutation of the GP dipeptide in *E. coli* thioredoxin to PH leads to a sizable 66 mV more positive redox potential (Jonda et al., 1999). Reciprocally, replacement of the PH sequence in the oxidizing DsbA protein by the thioredoxin GP dipeptide generates a double mutant protein that is some 92 mV more reducing than the wild type protein (Aslund et al., 1997; Inaba and Ito, 2002). A comparable insertion of these two dipeptides in the context of the ALR fold produced more limited modulation in redox potential and little change in  $k_{cat}/K_m$  (Table 4.1).

Additionally, we wanted to explore the impact of an insertion or deletion of the intervening amino acids would have on the catalytic efficiency of ALR. Mutations were made which included the deletion of E143, which results in a CxC motif, and the insertion of a Gly or Trp between residues 143 and 144, resulting in a CxGxC or CxWxC motif. We assessed the activity of these constructs and to our surprise the

CEC mutant was just as active as wild type protein (Table 4.1). However, the insertion of a small or large amino acid (Gly or Trp, respectively) decreases  $k_{cat}/K_m$  by 120- and 470-fold, respectively.

**Table 4.1 Redox potential and activity of the proximal disulfide in wild type and mutant sfALR proteins<sup>a,b</sup>**

CxxC motif	Redox potential (mV)	$k_{cat}$ (min <sup>-1</sup> )	$K_m$ (mM)	$k_{cat}/K_m$ (M <sup>-1</sup> sec <sup>-1</sup> )
CEEC	-235 ± 1.6	107.8 ± 6.91	1.635 ± 0.38	1099
CGPC	-256 ± 1.2	85.94 ± 3.59	1.601 ± 0.23	895
CPHC	-211 ± 2.9	85.97 ± 3.69	1.152 ± 0.17	1243
CKKC	-239 ± 1.4	92.7 ± 2.19	0.4107 ± 0.04	3762
CNWC	n.d. <sup>c</sup>	103.7 ± 2.25	0.5194 ± 0.05	3328
CEC	n.d. <sup>c</sup>	28.48 ± 1.38	0.6059 ± 0.12	783
CEGEC	n.d. <sup>c</sup>	12.71 ± 0.64	23.5 ± 4.04	9
CEWEC	n.d. <sup>c</sup>	6.135 ± 0.53	43.37 ± 12.11	2

<sup>a</sup>Redox measurements were determined in 50 mM phosphate buffer, pH 7.0, containing 1 mM EDTA as described in Materials and Methods.

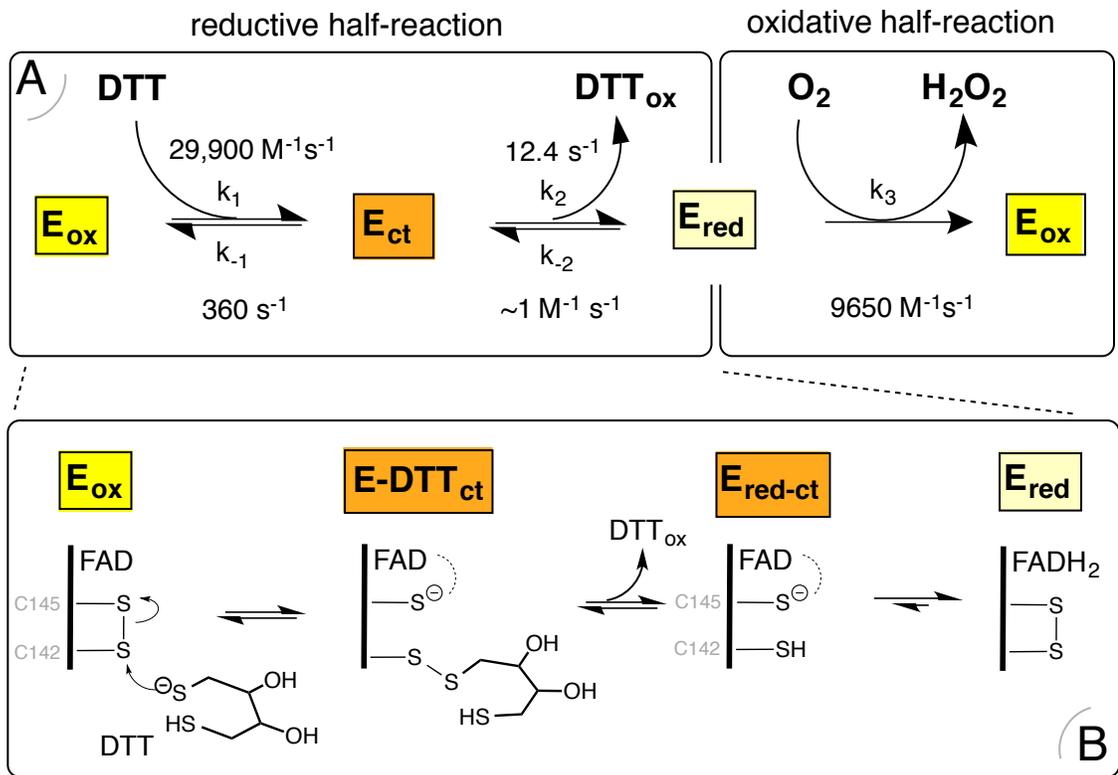
<sup>b</sup>turnover was determined in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA using DTT as a substrate.

<sup>c</sup>no data

Having confirmed the sizable difference in redox potentials between the two redox centers in wild type sfALR, we wished to explore the pre-steady state behavior of the enzyme in reductive and oxidative half-reactions and to reconcile these data with the steady-state kinetic parameters for the aerobic oxidation of DTT. This work provides new insights into the reactivity of the proximal CxxC motif in the context of the ALR fold.

#### 4.4 Overall Model for Catalysis in sfALR

Figure 4.2 presents a scheme that incorporates all of the species that we have observed during the pre-steady state kinetic experiments to be described below. The covalent interaction of reduced DTT with oxidized sfALR yields, in principle, two charge-transfer species that are likely to have very similar spectra; both forms are combined within the orange box ( $E_{ct}$ ) in panel A of Figure 4.2. The first of these is the mixed disulfide shown more explicitly in panel B ( $E\text{-DTT}_{ct}$ ), involving attack of a DTT thiolate on the C142 interchange surface accessible sulfur of the proximal disulfide. After loss of  $\text{DTT}_{ox}$  the second charge-transfer species ( $E_{red-ct}$ ) rapidly yields  $E_{red}$  which then enters the oxidative half-reaction (Panel A) leading to the net reduction of oxygen and to the generation of hydrogen peroxide.

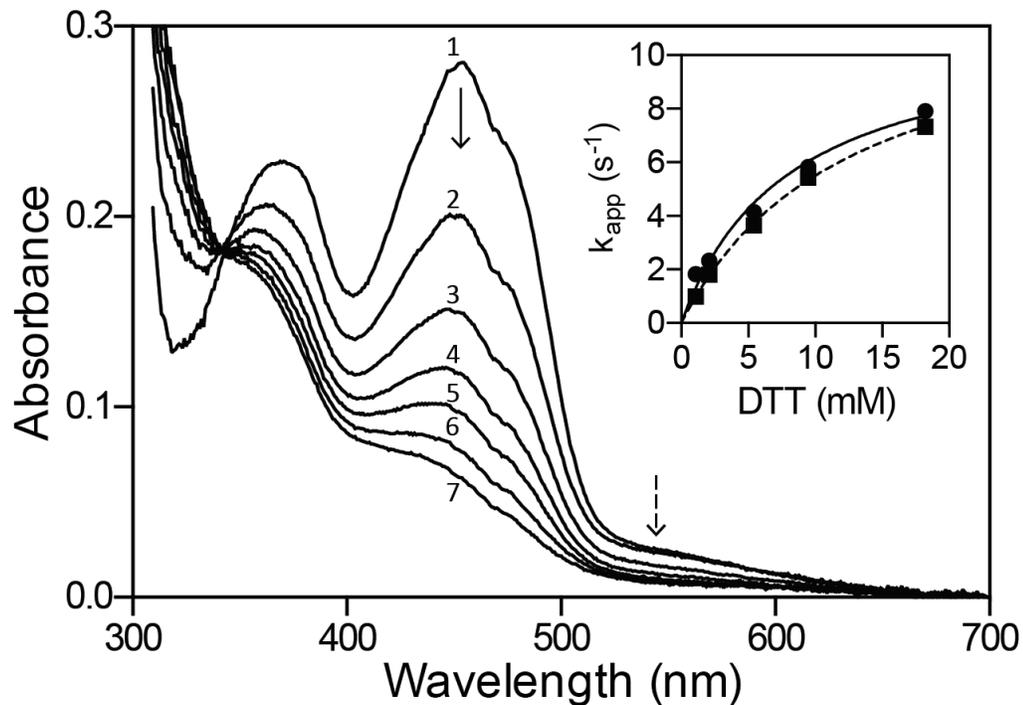


**Figure 4.2 Schematic representation of intermediates observed during catalysis of the aerobic oxidation of DTT by sfALR and the involvement of mixed disulfide intermediates in the reductive half-reaction.** Panel A shows the rate of interconversion of observable intermediates deduced from stopped flow measurements. Panel B shows the involvement of two charge-transfer intermediates during the reductive half-reaction.

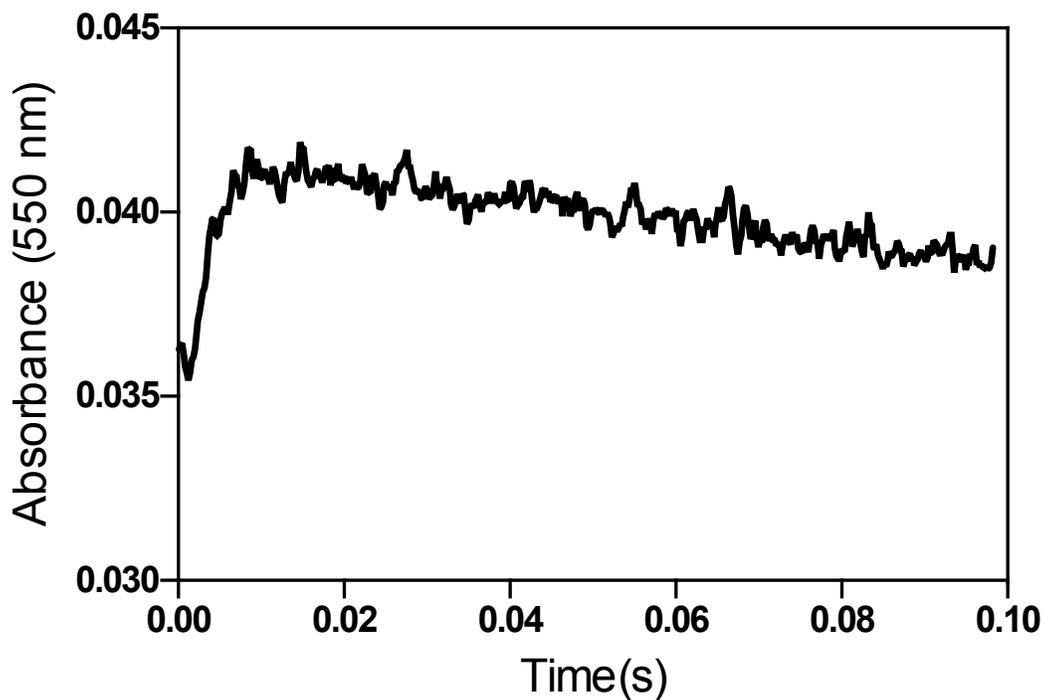
#### 4.5 Reductive Half-Reaction of sfALR Using DTT

We first investigated the interaction between sfALR and DTT under anaerobic conditions (see Materials and Methods). The progression of spectra upon mixing sfALR with 5.4 mM DTT is shown in Figure 4.3. The low intensity wedge-shaped charge-transfer absorbance band extending to about 650 nm is formed maximally at about 10 ms in the stopped flow under these conditions (Figure 4.4). Thereafter the feature decays with accumulation of the reduced enzyme ( $E_{\text{red}}$ ; see later). The  $E_{\text{red-ct}}$  form would be expected to show a long-wavelength charge-transfer species involving interaction between C145 (the charge-transfer thiol) and the oxidized flavin moiety (Figure 4.2). Precedent for this charge-transfer band in ALR comes from the static spectrum of C142A and C142S mutants ((Banci et al., 2012; Daithankar et al., 2009) and SAR unpublished observations) in which C145 thiolate would be within 3.1 – 3.2 Å of the  $C_{4a}$  locus of the isoalloxazine ring ((Banci et al., 2012) PDB 3U2L; Ramadan, Dong, Bahnson and Thorpe, PDB 3TK0 unpublished). However the  $E_{\text{red-ct}}$  species shown in Figure 4.2B is not observed upon 2-electron titration of wild type sfALR (Farrell and Thorpe, 2005) because the reduced flavin species  $E_{\text{red}}$  is overwhelmingly favored (as shown by the redox measurements described above). The other possible contributor to the charge-transfer absorbance feature seen in Figure 4.3 is the mixed disulfide species with DTT (Figure 4.2B;  $E\text{-DTT}_{\text{ct}}$ ). Such species have precedents in the mixed disulfides formed between two CxxC motifs in both Erv1p (Coppock and Thorpe, 2006; Hofhaus et al., 2003; Wang et al., 2007) and QSOX (Alon et al., 2012; Heckler et al., 2008a; Kodali and Thorpe, 2010). In addition, Williams and coworkers have shown that oxidized yeast glutathione reductase treated with GSH exists largely in the mixed disulfide form (Arscott et al., 2000). Later in this Chapter we will describe the surprising propensity of sfALR to form mixed disulfide species, but here

we are concerned with identifying the main features of catalytic turnover; hence we aggregate all the observable charge-transfer species into the  $E_{ct}$  form shown in Figure 4.2.



**Figure 4.3 Reduction of sfALR by DTT under anaerobic conditions.** The main panel shows diode-array traces recorded 0.005, 0.145, 0.295, 0.445, 0.625, 0.845, and 1.405 s after mixing oxidized sfALR with DTT to give final concentrations of 30  $\mu$ M and 5.4 mM respectively in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA (spectra 1-7 respectively; see Materials and Methods). The inset shows limiting apparent pseudo first-order rate constants of  $12.4 \pm 0.2 \text{ s}^{-1}$  and  $11.2 \pm 0.8 \text{ s}^{-1}$  for the reduction of FAD at 454 nm (solid line) and the disappearance of the charge-transfer intermediate at 550 nm (dashed line), respectively. Data are the average of triplicate runs.



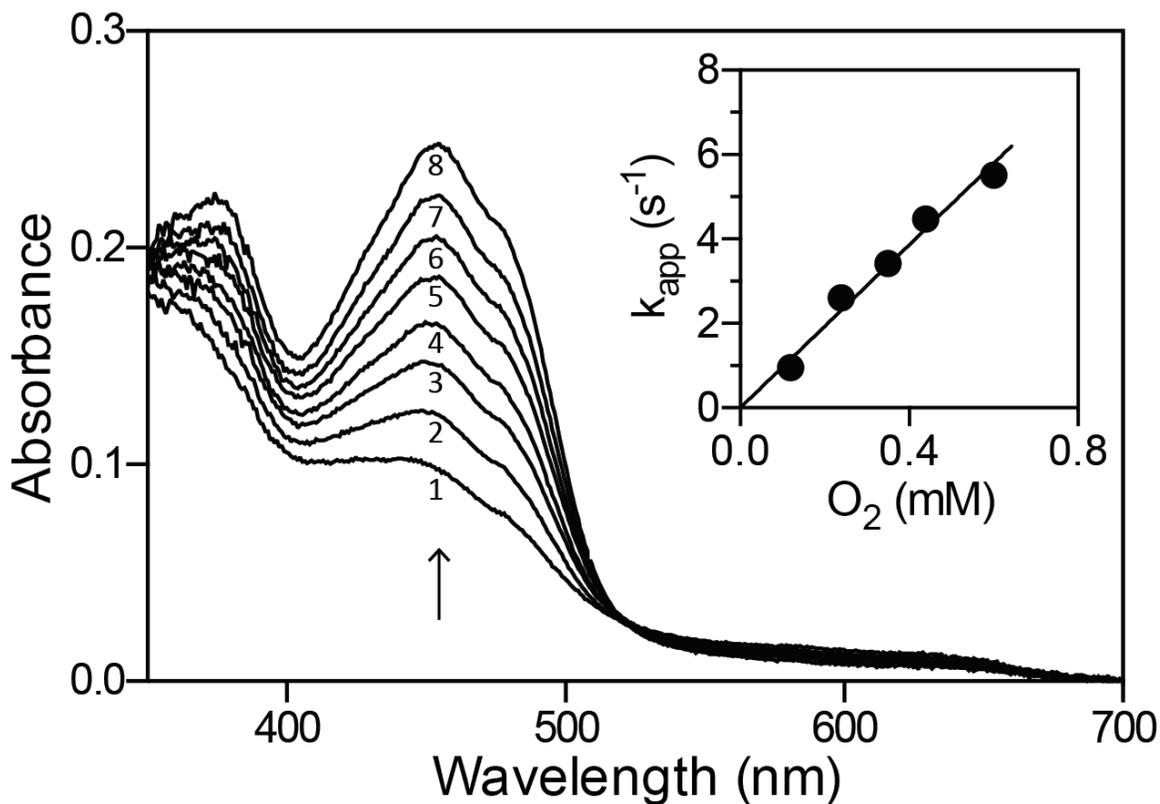
**Figure 4.4 Rapid formation of charge-transfer intermediate when sfALR is mixed with DTT under anaerobic conditions.** The trace records the absorbance at 550 nm when 2.1 mM DTT was mixed with 30  $\mu$ M sfALR in 50 mM phosphate buffer pH 7.5 containing 1 mM EDTA (see Materials and Methods). The observed pseudo first-order rate constants for the decline in absorbance at this wavelength are shown in the inset to Figure 4.3.

The DTT dependence of the apparent rate constants obtained from absorbance changes at 454 and 550 nm are comparable (Figure 4.3 with limiting  $k_{app}$  values of  $12.4 \pm 0.2 \text{ s}^{-1}$  and  $11.2 \pm 0.8 \text{ s}^{-1}$  respectively). Experiments conducted up to 18 mM DTT showed no detectable accumulation of the C<sub>4a</sub>-thiol-flavin adducts believed to be transitory intermediates in the reduction of flavins by thiols (Miller et al., 1990; O'Donnell and Williams, 1984; Thorpe and Williams, 1976b). The bimolecular reverse reaction  $k_{-2}$  (Figure 4.2) is negligible under the conditions of our experiments where, at most, stoichiometric levels of DTT<sub>ox</sub> are formed in the presence of a large concentration of competing DTT. This reverse reaction,  $k_{-2}$ , is set arbitrarily to  $1 \text{ M}^{-1}\text{s}^{-1}$  in Figure 4.2; simulations reveal that it could be increased up to  $10 \text{ M}^{-1}\text{s}^{-1}$  without significantly impacting the overall fits (see Materials and Methods).

#### **4.6 Oxidative Half-Reaction of 2-Electron Reduced sfALR with Oxygen**

ALR has been shown to be an atypical flavoprotein oxidase (Daithankar et al., 2009, 2012; Farrell and Thorpe, 2005). Firstly, ALR shows a high  $K_m$  for oxygen ( $\sim 240 \mu\text{M}$ ) and a distinct preference for cytochrome *c* as an electron acceptor of the enzyme (Daithankar et al., 2009; Farrell and Thorpe, 2005). While lfALR could utilize this cytochrome as a means to transfer reducing equivalents to the respiratory chain, sfALR functions extramitochondrially (see above) where cytochrome *c* is not normally resident. Hence oxygen is a logical physiological electron acceptor for sfALR. A second unusual feature of sfALR is that it releases more superoxide ion than would be expected for a simple flavoprotein oxidase (Daithankar et al., 2012). For these reasons we investigated the reaction of the reduced sfALR with molecular oxygen.

The 2-electron reduced enzyme  $E_{\text{red}}$  (Figure 4.2) was generated by dithionite titration in an anaerobic tonometer and subsequently mixed with various concentrations of dissolved oxygen (see Materials and Methods). The course of the reaction, followed in the diode-array stopped-flow spectrophotometer, showed a monotonic reappearance of the spectrum of oxidized ALR over a range of wavelengths (Figure 4.5). Prior to mixing with oxygen very small levels of the blue semiquinone had accumulated during the anaerobic dithionite titration (Farrell and Thorpe, 2005) and proved impractical to eliminate without over-reducing the enzyme. This species disappears during the subsequent reoxidation of the dihydroflavin form of sfALR monitored at 454 nm. The reappearance of  $E_{\text{ox}}$  yields a bimolecular rate constant of  $9650 \pm 150 \text{ M}^{-1}\text{s}^{-1}$  with an essentially linear dependence on oxygen concentration over the range studied (Figure 4.5, inset). This irreversible reaction is significantly slower than the  $10^5 - 10^6 \text{ M}^{-1}\text{s}^{-1}$  typically observed for flavoprotein oxidases (Chaiyen et al., 2012; Massey, 2002; Mattevi, 2006); additionally the rate constant is 50- to 90-fold slower than reoxidation of the ERV-domain in avian (Hooper and Thorpe, 1999) and *Trypanosoma brucei* QSOX (Kodali and Thorpe, 2010).



**Figure 4.5 Oxidative half-reaction of sfALR.** The 2-electron reduced enzyme (see Materials and Methods) was mixed with aerobic buffer to give final concentration of 30  $\mu$ M enzyme and 120  $\mu$ M oxygen. Spectra were recorded in the diode-array stopped-flow instrument 0.005, 0.055, 0.105, 0.155, 0.215, 0.285, 0.365 and 0.505 s after mixing (spectra 1-8 respectively). The inset plots the dependence on the apparent first-order rate constants for the return of the oxidized flavin at 454 nm as a function of dissolved oxygen concentration. Data points represent the average of 4 experiments at each oxygen concentration.

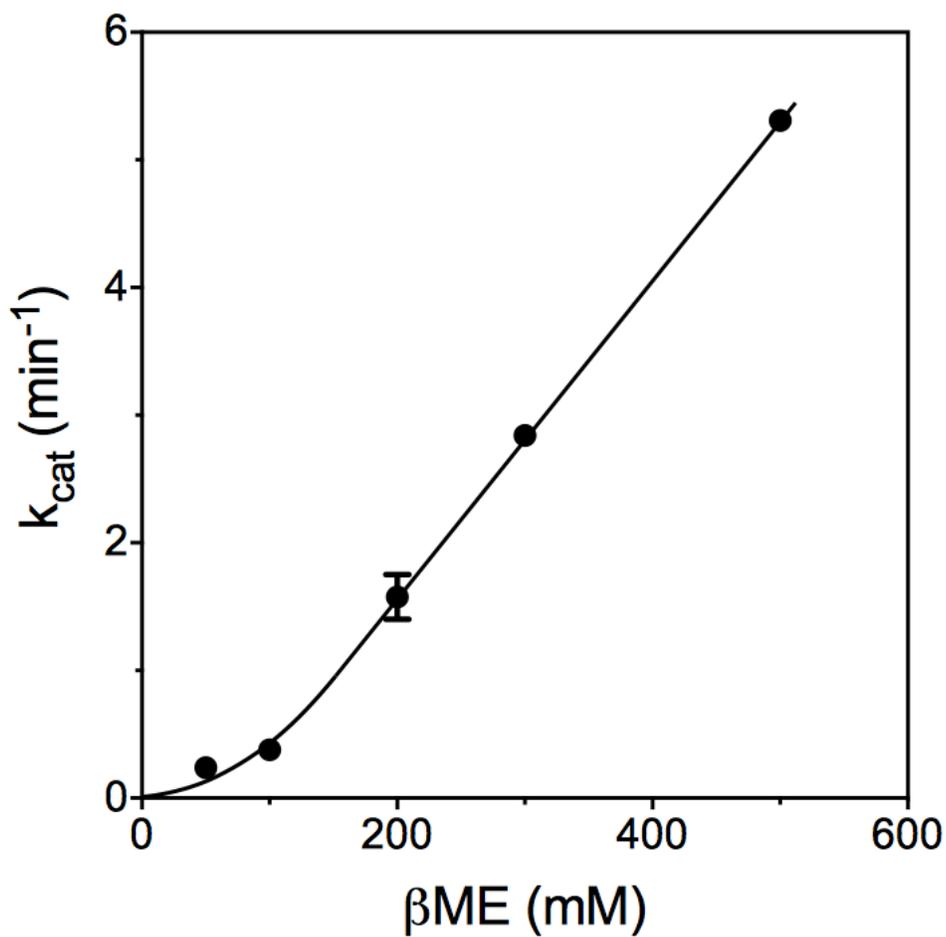
#### **4.7 Summary of Reductive and Oxidative Reaction**

The rate constants shown for the reductive half-reaction in Figure 4.2 predict steady-state kinetic parameters (see Materials and Methods) which are in good agreement with those determined experimentally. In terms of the  $k_{\text{cat}}$  and  $K_m$  values determined at 240  $\mu\text{M}$  oxygen, the observed and predicted values are as follows: 108  $\text{min}^{-1}$  and 99  $\text{min}^{-1}$  and 1.6 mM and 1.6 mM respectively. Farrell and Thorpe determined a  $K_m$  of 240  $\mu\text{M}$  for oxygen using 2 mM DTT (Farrell and Thorpe, 2005); compared to a predicted value of 212  $\mu\text{M}$ . While the very rapid formation of the small levels of the charge-transfer species (Figure 4.2 and Figure 4.3) made  $k_1$  impractical to determine experimentally, the values  $k_1$  and  $k_{-1}$  in Figure 4.2A provide good fits to the data (see Materials and Methods).

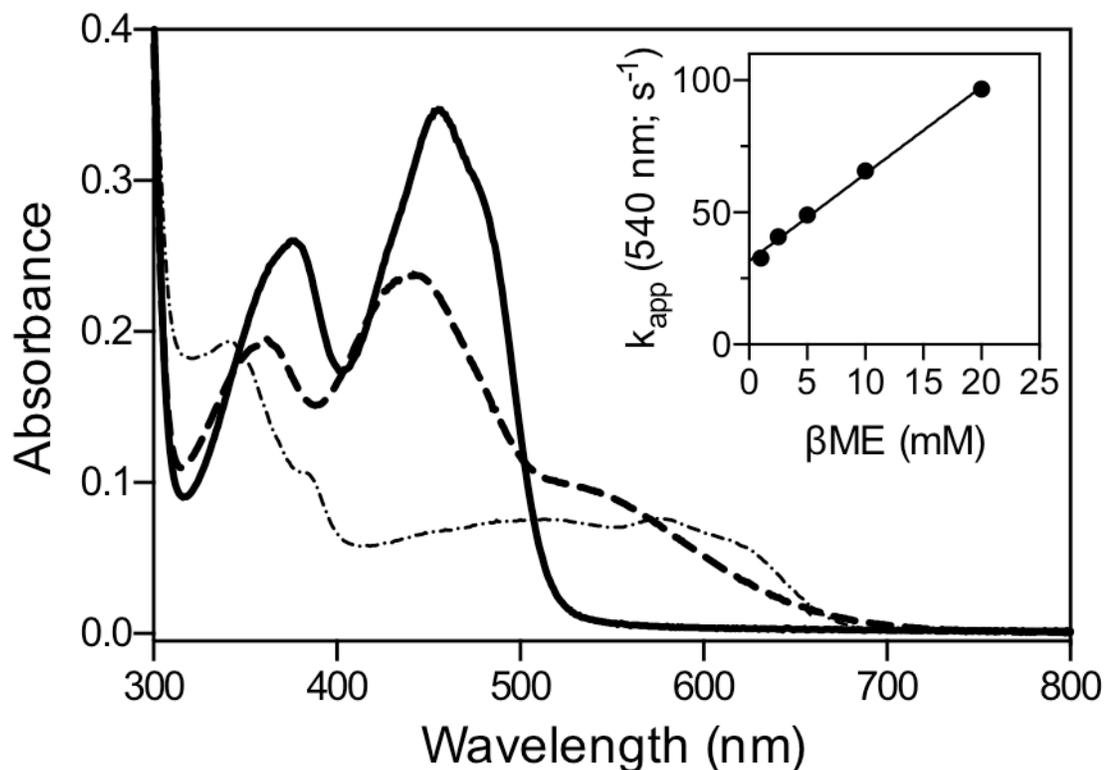
#### **4.8 The Stabilization of Mixed Disulfide Intermediates in sfALR**

We next consider whether a mixed disulfide intermediate with DTT could make a significant contribution to the aggregate  $E_{\text{ct}}$  species shown in Figure 4.2A. To explore the propensity of DTT to form mixed disulfide intermediates in sfALR we used the monothiol  $\beta\text{ME}$  which is structurally analogous to one half of DTT. We demonstrated that  $\beta\text{ME}$  is a very poor substrate showing a turnover number of  $\sim 0.3$  molecules of oxygen reduced/min at 100 mM thiol (Figure 4.6). Nevertheless  $\beta\text{ME}$  forms facile and surprisingly intense charge-transfer intermediates when mixed with oxidized sfALR (Figure 4.7). Upon prolonged incubation there is a very slow subsequent appearance of the blue flavin semiquinone which always appears when oxygen is depleted from solution of ALR reduced with thiol substrates (Daithankar et al., 2009; Farrell and Thorpe, 2005; Kay et al., 2006; Wu et al., 2003). The pseudo first-order appearance of the mixed disulfide charge-transfer intermediate was

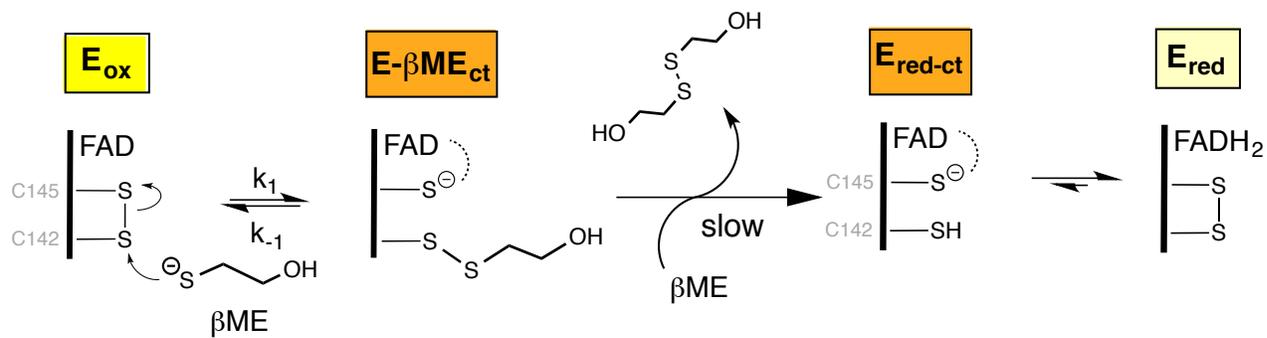
followed at 540 nm in the stopped flow spectrophotometer (inset Figure 4.7) as a function of  $\beta$ ME concentration giving rate constants of  $3295 \pm 59 \text{ M}^{-1}\text{s}^{-1}$  for  $k_1$  and  $31.6 \pm 0.6 \text{ s}^{-1}$  for  $k_{-1}$ , yielding an overall  $K_D$  of  $9.6 \pm 0.25 \text{ mM}$ ; see Materials and Methods). These data explain why plots of turnover versus  $\beta$ ME concentration exhibit upward curvature followed by a linear increase (Figure 4.6). Turnover of  $\beta$ ME cannot proceed unless the mixed disulfide intermediate shown in Figure 4.8 is captured by a second  $\beta$ ME molecule in a kinetically sluggish bimolecular encounter.



**Figure 4.6**  $\beta$ ME as a substrate of sfALR in the oxygen electrode. Turnover numbers were calculated using 10  $\mu\text{M}$  sfALR with increasing concentrations of  $\beta$ ME in 2.5 mL of 50 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA at 25  $^{\circ}\text{C}$ . Data points represent the average of two experiments.

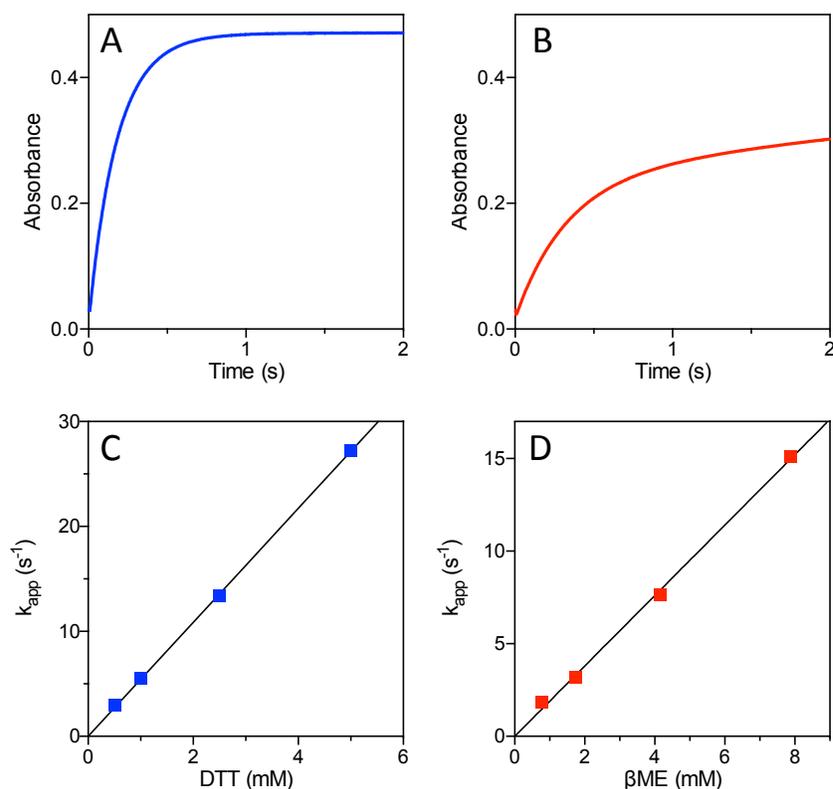


**Figure 4.7  $\beta$ ME forms mixed disulfide intermediates with sfALR.** The spectrum of sfALR (30  $\mu$ M) in air-saturated 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA was recorded before and 5 sec after the addition of 200 mM  $\beta$ ME in a diode-array spectrophotometer (solid and dashed lines respectively). After 7 min the progressive depletion of oxygen from solution leads to the appearance of the blue neutral semiquinone (see the text). The inset plots the pseudo first-order rate constants obtained from the increase in absorbance at 540 nm following mixing sfALR in the stopped flow spectrophotometer with increasing concentration of  $\beta$ ME (1 – 20 mM final). The slope of  $3295 \pm 59 \text{ M}^{-1}\text{s}^{-1}$  reflects the bimolecular association rate constant  $k_1$  and the intercept of  $31.6 \pm 0.6 \text{ s}^{-1}$  corresponds to  $k_{-1}$ . The corresponding equilibrium constant  $K_D = 9.6 \pm 0.25 \text{ mM}$ .



**Figure 4.8** The reductive half-reaction between sfALR and  $\beta$ ME. Rapid nucleophilic attack of the thiolate of  $\beta$ ME reversibly yields the mixed disulfide charge-transfer species  $E-\beta$ ME<sub>ct</sub> which is weakly reactive towards a second molecule of  $\beta$ ME.

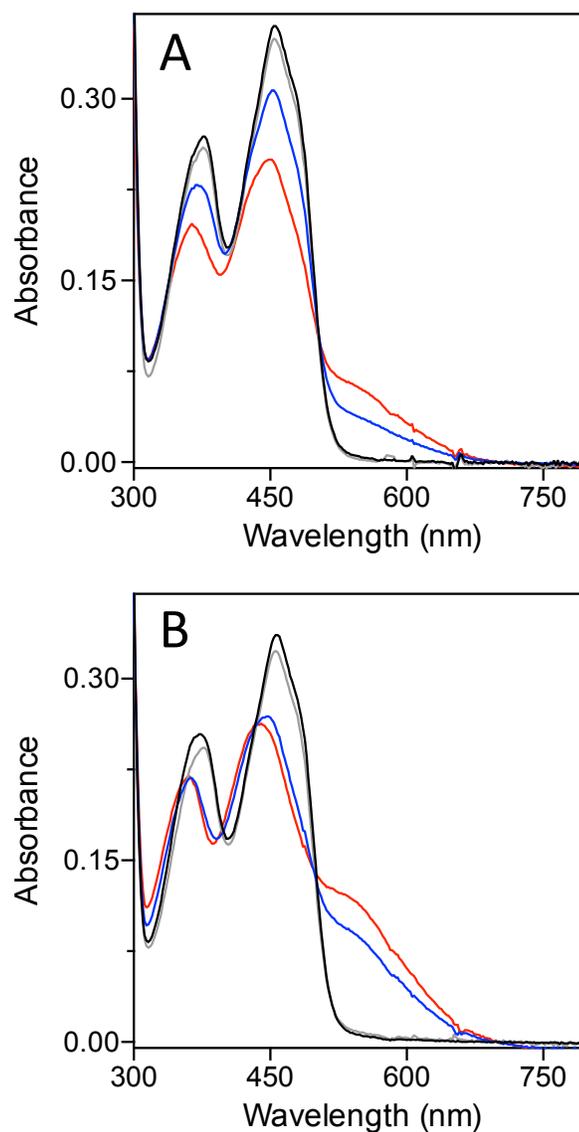
Since  $\beta$ ME is a monothiol structural mimic of half of DTT (Figure 4.2), there seems no reason to expect that these two reagents would not be comparable nucleophiles in the formation of the mixed disulfide intermediates shown in Figure 4.2B (Houk et al., 1987). For example, their thiol pK values are comparable (9.5 for  $\beta$ ME and 9.3 for the first pK of DTT (Houk et al., 1987)). As a further validation of the similarity between the reactivity of these two reagents we explored their reactivity towards the chromophoric disulfide DTNB in the stopped-flow instrument by following the release of TNB under pseudo first-order conditions (Figure 4.9). Here the encounter between thiols and DTNB is second-order overall with the expected zero intercept for a reaction that is essentially irreversible given the strongly oxidizing redox potential of the DTNB disulfide. As expected, these bimolecular rate constants are similar, when compared on a per thiol basis at pH 7.5, 25 °C ( $1.90 \pm 0.01 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$  for  $\beta$ ME; and  $2.72 \pm 0.007 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$  for DTT). In sum, these data suggest that a significant fraction of the charge-transfer band seen in Figure 4.3 represents a mixed disulfide with one thiol of DTT and that the relatively slow reduction of the flavin of  $12.4 \text{ s}^{-1}$  may reflect the resolution of the mixed disulfide by the second thiol moiety of DTT. Typically the subsequent transfer of reducing equivalents from the reduced proximal disulfide in charge transfer with an oxidized flavin is rapid ( $800 \text{ s}^{-1}$  for lipoamide dehydrogenase (Benen et al., 1992; Massey et al., 1960; Williams, 1992) and  $110 \text{ s}^{-1}$  for glutathione reductase (Rietveld et al., 1994)).



**Figure 4.9 Comparison of the nucleophilicity of DTT and  $\beta$ ME towards DTNB.**

Panel A shows a representative trace of the increase in absorbance at 412 nm upon mixing DTNB and DTT to final concentrations of 17  $\mu$ M and 1 mM respectively at 25  $^{\circ}$ C in 50 mM phosphate buffer, pH 7.5, containing 1 mM EDTA. The reaction is monophasic, and the dependence of the pseudo first-order rate constants (the average of 5 determinations) is plotted as a function of DTT concentrations in panel C. A reaction trace at 412 nm using 1.74 mM  $\beta$ ME is plotted in panel B under otherwise identical conditions. The first phase of the absorbance increase corresponds to attack of the monothiol on DTNB with the formation of the  $\beta$ ME-TNB mixed disulfide. A subsequent, much slower phase, represents the release of the second TNB from this mixed disulfide in the presence of excess thiol. The data were analyzed using two exponentials (see Materials and Methods) and 4 determinations of the pseudo first-order rate constants for the first phase were averaged and plotted as a function of  $\beta$ ME concentration.

These data with  $\beta$ ME suggest that ALR has an unexpected propensity to form mixed disulfides with a monothiol with which it would not have been expected to exhibit a particular affinity. Accordingly we surveyed several other monothiols (at 20 mM) for their abilities to generate these charge-transfer complexes at both pH 7.5 and pH 9 (Figure 4.10). While  $\beta$ ME, N-acetylcysteamine and cysteamine (not shown) are rather comparable in their behavior, glutathione is conspicuously unable to form significant levels of mixed disulfide. This behavior is also seen at pH 9, where the other monothiols generate strong charge-transfer bands. All these monothiols, including GSH, have rather comparable pK values (from 8.2 – 9.5) and therefore would be expected to have correspondingly comparable intrinsic nucleophilicities (Houk and Whitesides, 1987; Lees and Whitesides, 1993; Singh and Whitesides, 1993). However GSH is notably bulkier than the other thiols, suggesting that steric factors are likely to explain why this abundant cellular monothiol is an undetectable substrate of sfALR (Daithankar et al., 2009). Mixed disulfide bond formation requires in-line approach of the attacking nucleophilic thiolate along the disulfide axis (Bach et al., 2008; Fernandes and Ramos, 2004; Heckler et al., 2008b; Rosenfield et al., 1977). While this first step is already disfavored for GSH, the completion of catalysis would require a second glutathione to resolve this sterically congested mixed disulfide. Thus although reduction of the proximal disulfide ( $E^{\circ\prime}$  - 235 mV) by GSH ( $E^{\circ\prime}$  - 240 mV (Williams, 1992)) is certainly feasible from a thermodynamic perspective, it is strongly disfavored kinetically. Hence intracellular sfALR can avoid the potentially damaging generation of hydrogen peroxide driven by millimolar levels of GSH (Daithankar et al., 2012; Farrell and Thorpe, 2005).



**Figure 4.10 Spectral changes on the addition of monothiols to ALR at pH 7.5 and 9.0.** The spectrum of an aerobic solution of 30  $\mu\text{M}$  ALR (black line) was recorded 10 s before the addition of 20 mM concentrations of glutathione (grey line),  $\beta\text{ME}$  (red line) and N-acetylcysteamine (blue line). Panel A was conducted in phosphate buffer, pH 7.5, and panel B was recorded using Tris chloride buffer, pH 9.0.

Finally, we consider why monothiols can form significant complexes with sfALR when their bimolecular formation could be reversed intramolecularly by the locally enormous concentration of the neighboring C145 thiol shown in Figure 4.8. At least part of the answer may be the difference in pKs between the attacking monothiol and that of the C145 thiol that would be released upon mixed disulfide bond formation; as in all disulfide exchange reactions, the pK of attacking and leaving thiolates will profoundly modulate the position of the equilibrium (Houk and Whitesides, 1987; Lees and Whitesides, 1993; Singh and Whitesides, 1993; Winther and Thorpe, 2013). Thus the pK of  $\beta$ ME is 9.5 whereas the charge-transfer thiol has been reported to have a pK of 5.95 (Banci et al., 2012) biasing the mixed disulfide equilibrium in favor of the charge-transfer species in Figure 4.8. The surprising stability of these initial mixed disulfide species then would allow sufficient time for attainment of a conformation in which in-line approach of the resolving substrate thiolate leads to the transient generation of  $E_{\text{red-ct}}$  and to the subsequent rapid formation of the free reduced enzyme  $E_{\text{red}}$  (Figure 4.2).

## REFERENCES

- Alon, A., Grossman, I., Gat, Y., Kodali, V.K., DiMaio, F., Mehlman, T., Haran, G., Baker, D., Thorpe, C., and Fass, D. (2012). The dynamic disulphide relay of quiescin sulphhydryl oxidase. *Nature* *488*, 414–418.
- Arcscott, L.D., Veine, D.M., and Williams, C.H. (2000). Mixed disulfide with glutathione as an intermediate in the reaction catalyzed by glutathione reductase from yeast and as a major form of the enzyme in the cell. *Biochemistry* *39*, 4711–4721.
- Aslund, F., Berndt, K.D., and Holmgren, A. (1997). Redox potentials of glutaredoxins and other thiol-disulfide oxidoreductases of the thioredoxin superfamily determined by direct protein-protein redox equilibria. *J Biol Chem* *272*, 30780–30786.
- Bach, R.D., Dmitrenko, O., and Thorpe, C. (2008). Mechanism of thiolate-disulfide interchange reactions in biochemistry. *J. Org. Chem.* *73*, 12–21.
- Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., and Tokatlidis, K. (2012). An electron-transfer path through an extended disulfide relay system: the case of the redox protein ALR. *J. Am. Chem. Soc.* *134*, 1442–1445.
- Benen, J., van Berkel, W., Dieteren, N., Arcscott, D., Williams, C., Veeger, C., and de Kok, A. (1992). Lipoamide dehydrogenase from *Azotobacter vinelandii*: site-directed mutagenesis of the His450-Glu455 diad. Kinetics of wild-type and mutated enzymes. *Eur. J. Biochem.* *207*, 487–497.
- Brohawn, S.G., Miksa, I.R., and Thorpe, C. (2003). Avian sulfhydryl oxidase is not a metalloenzyme: adventitious binding of divalent metal ions to the enzyme. *Biochemistry* *42*, 11074–11082.
- Chaiyen, P., Fraaije, M.W., and Mattevi, A. (2012). The enigmatic reaction of flavins with oxygen. *Trends Biochem. Sci.* *37*, 373–380.
- Coppock, D.L., and Thorpe, C. (2006). Multidomain flavin-dependent sulfhydryl oxidases. *Antioxid Redox Signal* *8*, 300–311.

- Daithankar, V., Farrell, S., and Thorpe, C. (2009). Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space. *Biochemistry* *48*, 4828–4837.
- Daithankar, V., Schaefer, S.A., Dong, M., Bahnson, B.J., and Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augmenter of liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* *49*, 6737–6745.
- Daithankar, V.N., Wang, W., Trujillo, J.R., and Thorpe, C. (2012). Flavin-linked Erv-family sulfhydryl oxidases release superoxide anion during catalytic turnover. *Biochemistry* *51*, 265–272.
- DuPlessis, E.R., Pellett, J., Stankovich, M.T., and Thorpe, C. (1998). Oxidase activity of the acyl-CoA dehydrogenases. *Biochemistry* *37*, 10469–10477.
- Farrell, S.R., and Thorpe, C. (2005). Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry* *44*, 1532–1541.
- Fernandes, P.A., and Ramos, M.J. (2004). Theoretical insights into the mechanism for thiol/disulfide exchange. *Chemistry* *10*, 257–266.
- Heckler, E.J., Alon, A., Fass, D., and Thorpe, C. (2008a). Human quiescin-sulfhydryl oxidase, QSOX1: probing internal redox steps by mutagenesis. *Biochemistry* *47*, 4955–4963.
- Heckler, E.J., Rancy, P.C., Kodali, V.K., and Thorpe, C. (2008b). Generating disulfides with the Quiescin-sulfhydryl oxidases. *Biochim Biophys Acta* *1783*, 567–577.
- Hofhaus, G., Lee, J.E., Tews, I., Rosenberg, B., and Lisowsky, T. (2003). The N-terminal cysteine pair of yeast sulfhydryl oxidase Erv1p is essential for in vivo activity and interacts with the primary redox centre. *Eur J Biochem* *270*, 1528–1535.
- Hooper, K.L., and Thorpe, C. (1999). Egg white sulfhydryl oxidase: kinetic mechanism of the catalysis of disulfide bond formation. *Biochemistry* *38*, 3211–3217.
- Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P., and Kummer, U. (2006). COPASI--a COMplex PATHway SIMulator. *Bioinformatics* *22*, 3067–3074.

- Houk, J., and Whitesides, G.M. (1987). Structure-reactivity relations for thiol-disulfide interchange. *J. Am. Chem. Soc.* *109*, 6825–6836.
- Houk, J., Singh, R., and Whitesides, G.M. (1987). Measurement of thiol-disulfide interchange reactions and thiol pKa values. *Methods Enzymol.* *143*, 129–140.
- Huber-Wunderlich, M., and Glockshuber, R. (1998). A single dipeptide sequence modulates the redox properties of a whole enzyme family. *Fold Des* *3*, 161–171.
- Inaba, K., and Ito, K. (2002). Paradoxical redox properties of DsbB and DsbA in the protein disulfide-introducing reaction cascade. *EMBO J.* *21*, 2646–2654.
- Johnson, K.A., Simpson, Z.B., and Blom, T. (2009a). FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal. Biochem.* *387*, 30–41.
- Johnson, K.A., Simpson, Z.B., and Blom, T. (2009b). Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* *387*, 20–29.
- Jonda, S., Huber-Wunderlich, M., Glockshuber, R., and Mössner, E. (1999). Complementation of DsbA deficiency with secreted thioredoxin variants reveals the crucial role of an efficient dithiol oxidant for catalyzed protein folding in the bacterial periplasm. *EMBO J.* *18*, 3271–3281.
- Kay, C.W.M., Elsässer, C., Bittl, R., Farrell, S.R., and Thorpe, C. (2006). Determination of the distance between the two neutral flavin radicals in augments of liver regeneration by pulsed ELDOR. *J. Am. Chem. Soc.* *128*, 76–77.
- Kodali, V.K., and Thorpe, C. (2010). Quiescin sulfhydryl oxidase from *Trypanosoma brucei*: catalytic activity and mechanism of a QSOX family member with a single thioredoxin domain. *Biochemistry* *49*, 2075–2085.
- Krause, G., Lundström, J., Barea, J.L., Pueyo de la Cuesta, C., and Holmgren, A. (1991). Mimicking the active site of protein disulfide-isomerase by substitution of proline 34 in *Escherichia coli* thioredoxin. *J. Biol. Chem.* *266*, 9494–9500.
- Lees, W.J., and Whitesides, G.M. (1993). Equilibrium constants for thiol-disulfide interchange reactions: a coherent, corrected set. *J. Org. Chem.* *58*, 642–647.
- Massey, V. (2002). No Title. In *Flavins and Flavoproteins*, International Congress Series, N.S. Chapman, S., Perham, R., and Scrutton, ed. (Walter De Gruyter), pp. 3–11.

Massey, V., Gibson, Q.H., and Veeger, C. (1960). Intermediates in the catalytic action of lipoyl dehydrogenase (diaphorase). *Biochem. J.* 77, 341–351.

Mattevi, A. (2006). To be or not to be an oxidase: challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* 31, 276–283.

Miller, S.M., Massey, V., Ballou, D., Williams, C.H., Distefano, M.D., Moore, M.J., and Walsh, C.T. (1990). Use of a site-directed triple mutant to trap intermediates: demonstration that the flavin C(4a)-thiol adduct and reduced flavin are kinetically competent intermediates in mercuric ion reductase. *Biochemistry* 29, 2831–2841.

Moore, E.C., Reichard, P., and Thelander, L. (1964). Enzymatic synthesis of deoxyribonucleotides. V. purification and properties of thioredoxin reductase from *Escherichia coli* B. *J. Biol. Chem.* 239, 3445–3452.

Mössner, E., Huber-Wunderlich, M., and Glockshuber, R. (1998). Characterization of *Escherichia coli* thioredoxin variants mimicking the active-sites of other thiol/disulfide oxidoreductases. *Protein Sci.* 7, 1233–1244.

O'Donnell, M.E., and Williams, C.H. (1984). Reconstitution of *Escherichia coli* thioredoxin reductase with 1-deazaFAD. Evidence for 1-deazaFAD C-4a adduct formation linked to the ionization of an active site base. *J. Biol. Chem.* 259, 2243–2251.

Rietveld, P., Arscott, L.D., Berry, A., Scrutton, N.S., Deonarain, M.P., Perham, R.N., and Williams, C.H. (1994). Reductive and oxidative half-reactions of glutathione reductase from *Escherichia coli*. *Biochemistry* 33, 13888–13895.

Rosenfield, R.E., Parthasarathy, R., and Dunitz, J.D. (1977). Directional preferences of nonbonded atomic contacts with divalent sulfur. 1. Electrophiles and nucleophiles. *J. Am. Chem. Soc.* 99, 4860–4862.

Singh, R., and Whitesides, G.M. (1993). Thiol-disulfide Interchange. In *The Chemistry of Sulphur-Containing Functional Groups*, Patai, S., and Rappoport, Z., ed. (John Wiley), pp. 633–658.

Thorpe, C., and Williams, C.H. (1976a). Differential reactivity of the two active site cysteine residues generated on reduction of pig heart lipoamide dehydrogenase. *J. Biol. Chem.* 251, 3553–3557.

Thorpe, C., and Williams, C.H. (1976b). Spectral evidence for a flavin adduct in a monoalkylated derivative of pig heart lipoamide dehydrogenase. *J. Biol. Chem.* 251, 7726–7728.

Walsh, C., Fisher, J., Spencer, R., Graham, D.W., Ashton, W.T., Brown, J.E., Brown, R.D., and Rogers, E.F. (1978). Chemical and enzymatic properties of riboflavin analogues. *Biochemistry* 17, 1942–1951.

Wang, W., Winther, J.R., and Thorpe, C. (2007). Erv2p: characterization of the redox behavior of a yeast sulfhydryl oxidase. *Biochemistry* 46, 3246–3254.

Williams, C.H. (1992). Lipoamide dehydrogenase, glutathione reductase, thioredoxin reductase, and mercuric ion reductase-A family of flavoenzyme transhydrogenases. In *Chemistry and Biochemistry of Flavoenzymes*, F. Müller, ed. (CRC Press), pp. 121–211.

Williams, C.H., Arscott, L.D., Matthews, R.G., Thorpe, C., and Wilkinson, K.D. (1979). Methodology employed for anaerobic spectrophotometric titrations and for computer-assisted data analysis. *Methods Enzymol.* 62, 185–198.

Winther, J.R., and Thorpe, C. (2013). Quantification of thiols and disulfides. *Biochim. Biophys. Acta*.

Wu, C.K., Dailey, T.A., Dailey, H.A., Wang, B.C., and Rose, J.P. (2003). The crystal structure of augments liver regeneration: A mammalian FAD-dependent sulfhydryl oxidase. *Protein Sci* 12, 1109–1118.

Zapun, A., Bardwell, J.C.A., and Creighton, T.E. (1993). The reactive and destabilizing disulfide bond of DsbA, a protein required for protein disulfide bond formation in vivo. *Biochemistry* 32, 5083–5092.

## Chapter 5

# <sup>77</sup>SE ENRICHMENT OF PROTEINS EXPANDS THE BIOLOGICAL NMR TOOLBOX

### 5.1 Introduction

Sulfur, a key contributor to biological reactivity, is not amenable to investigations by biological NMR spectroscopy. To utilize selenium as a surrogate, the Rozovsky group has developed a generally applicable <sup>77</sup>Se isotopic enrichment method for heterologous proteins expressed in *Escherichia coli*. Here, we demonstrate <sup>77</sup>Se NMR spectroscopy of multiple selenocysteine and selenomethionine residues in the sulfhydryl oxidase augmentor of liver regeneration (ALR). The resonances of the active-site residues were assigned by comparing the NMR spectra of ALR bound to oxidized and reduced flavin adenine dinucleotide (FAD). An additional resonance appears only in the presence of the reducing agent and disappears readily upon exposure to air and subsequent reoxidation of the flavin. Hence, <sup>77</sup>Se NMR spectroscopy can be used to report the local electronic environment of reactive and structural sulfur sites, as well as changes taking place in those locations during catalysis.

### 5.2 Materials and Methods

Elemental <sup>77</sup>selenium (99.20%) was purchased from Isotopix USA (San Francisco, CA). Enzymes used for molecular biology were acquired from New

England Biolabs (Ipswich, MA). The pMHTDelta238 plasmid expressing tobacco etch virus (TEV) protease fused to the cytoplasmic maltose binding protein (Blommel and Fox, 2007) was purchased from the Protein Structure Initiative Material Repository. Chromatography media was supplied by GE Healthcare Biosciences Corporation (Pittsburgh, PA). Crystallization reagents were from Hampton Research (Aliso Viejo, CA). All other chemicals and reagents were supplied by SigmaAldrich (St. Louis, MO), Acros Organics (Geel, Belgium) and GoldBio (St. Louis, MO). All reagents and solvents were at least analytical grade and were used as supplied.

### **5.2.1 Expression of ALR and Trx**

The pTrcHisA (Life Technologies, Grand Island, NY) expression vector containing the gene encoding the short-form of human ALR (residues 81–205) with mutations C154A and C165A was as previously described (Farrell and Thorpe, 2005). For this study, a TEV cleavage site was introduced between the N-terminal hexahistidine tag and the first methionine. Following cleavage with TEV protease, a serine is present before residues 81–205. *E. coli* thioredoxin (Trx) fused to a C-terminal hexahistidine tag was expressed in pET32a (Addgene plasmid 11516). The two proteins were expressed using the BL21(DE3) cell line. Cells were grown in a defined media adapted from Studier (Studier, 2005) C-750501 minimal media designed for <sup>13</sup>C labeling and abbreviated here as MC-750501. The modified medium contained 50 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM KH<sub>2</sub>PO<sub>4</sub>, 10 mM NaCl, 50 mM NH<sub>4</sub>Cl, 2 mM MgCl<sub>2</sub>, 0.2x metals, 1x vitamins, 0.4% glucose and 200 μM CaCl<sub>2</sub>, supplemented with antibiotics. 1x trace metal solution contained 50 μM FeCl<sub>3</sub>, 20 μM CaCl<sub>2</sub>, 10 μM MnCl<sub>2</sub>, 10 μM ZnCl<sub>2</sub>, 2 μM CoCl<sub>2</sub>, 2 μM CuCl<sub>2</sub>, 2 μM NiCl<sub>2</sub>, 2 μM Na<sub>2</sub>MoO<sub>4</sub> and 2 μM H<sub>3</sub>BO<sub>3</sub>. The recipe for 1000x vitamin solution was as detailed by Studier

(Studier, 2005). A 20 mL starter culture of MC-750501, supplemented with 5 mM  $\text{Na}_2\text{SO}_4$  and antibiotics selection was grown for about 9 hours. A 1 mL starter culture was used to inoculate 1 L growth media supplemented with 50  $\mu\text{M}$   $\text{Na}_2\text{SO}_4$  in 2.8 L baffled flasks. For constructs that required the use of kanamycin sulfate, we reduced the concentration of the kanamycin to 50  $\mu\text{M}$  and used the antibiotics as the sole source of sulfur in the main incubation. Cells were grown at 37 °C, with good aeration and an antibiotics selection until the cell density no longer increased. Typically, this value is reached after 14 hours with an optical density close to 0.8 at 600 nm. Note that the doubling time of the cells in the MC-7504501 growth media is about 60 minutes. At this point, protein expression was induced with 0.5 mM IPTG. At induction, the cells were supplied with 50  $\mu\text{M}$  mixture of  $\text{Na}_2\text{SO}_4/\text{Na}_2\text{SeO}_3$  at the desired ratio. Following 4 hours of growth, a second 50  $\mu\text{M}$  aliquot was added. Cells were harvested 9–12 hours past induction by centrifuging them at 5000 g for 10 minutes at 4 °C. In general, the optimal expression time is approximately double that of the expression time in LB, due to the longer doubling times in the relatively spartan growth media. The cell paste was resuspended in 50 mM potassium phosphate buffer (pH 7.5) containing 500 mM NaCl [immobilized metal affinity chromatography (IMAC) buffer], flash frozen using liquid nitrogen and stored at –80 °C. For NMR experiments, the appropriate amount of elemental  $^{77}\text{Se}$  was oxidized to  $^{77}\text{selenite}$  in a minimal volume of nitric acid (Gladyshev, 1994). The expected selenite was confirmed by solution  $^{77}\text{Se}$  NMR. The volume was kept to a minimum in order to reduce the amount of acid added to the growth media.

### 5.2.2 Purification of ALR and Trx

The cell paste was thawed and lysed in IMAC buffer supplemented with 1 mM PMSF and 0.5 mM benzimidazole using a high pressure homogenizer (EmulsiFlex-C5; Avestin, Ottawa, Canada). Cell debris was removed by centrifugation at 15,000 g for 1 hour at 4 °C. The clear supernatant was loaded onto a 5 mL HisTrap FF column, and the column was washed with IMAC buffer, supplemented with 20 mM imidazole. ALR was eluted using 5 mL fractions of IMAC buffer supplemented with 50 mM, 200 mM, 500 mM and 1 M imidazole. Yellow fractions containing ALR were combined and the buffer was exchanged into 20 mM Tris (pH 7.5), 1 mM EDTA by dialysis and the hexahistidine tag cleaved with TEV protease, at 4 °C for 18 hours (Blommel and Fox, 2007; Kapust et al., 2001). The resulting protein was then dialyzed against 20 mM Tris and 1 mM EDTA (pH 7.5) at 4 °C. The protein was applied to a 5 mL HiTrap SP HP column to remove the TEV and then loaded on a Source 15Q column and eluted with a salt gradient between 0 and 100 mM NaCl over 20 column volumes. Purification of *E. coli* thioredoxin was by identical IMAC chromatography using IMAC buffer with 400 mM imidazole. The buffer was then exchanged to 20 mM Tris (pH 8.0), 1 mM EDTA and the protein loaded on a 5 mL HiTrapQ HP and eluted with a salt gradient between 0 and 300 mM NaCl over 20 column volumes. For both ALR and Trx, the fractions containing proteins were combined. The protein was dialyzed against 20 mM Tris (pH 7.5) and 1 mM EDTA, then concentrated to 20 mg/mL and flash frozen using liquid N<sub>2</sub> and stored at -80 °C until further use. Protein purity, as determined by 15% SDS-PAGE Tris-glycine gels, was higher than 99%. Protein concentration was determined using the molar extinction coefficients for FAD bound ALR of 11.6 mM<sup>-1</sup> cm<sup>-1</sup> at 456 nm (Farrell and Thorpe, 2005). The ratio of sulfur and selenium was determined by liquid chromatography– electrospray ionization–time

of flight mass spectrometry and inductively coupled plasma spectroscopy. While we have optimized the protocol for a final concentration of 100  $\mu\text{M}$  selenite, we have found for  $^{77}\text{Se}$  incorporation that a lower concentration of selenium (70  $\mu\text{M}$ ) yields identical selenium incorporation ratio but a higher yield (sometimes by a factor of 2). For simultaneous incorporation of  $^{13}\text{C}$ , regarded to be the most costly isotopes discussed here, we have tested the Marley method in which an initial culture is grown to exponential phase in LB, washed with saline and resuspended in growth media containing isotopically labeled metabolic precursors (Marley et al., 2001). We found that this method yields identical results as long as the cells are allowed to equilibrate after suspension and cell density monitored to ensure that the internal sulfur has been depleted. If  $^{13}\text{C}$  and  $^{15}\text{N}$  is not considered, then it is also possible to enrich the growth media with additional amino acids (Studier, 2005).

### **5.2.3 Mass Spectroscopy**

Mass spectra were obtained using a QTOF Ultima (Waters, Milford, MA) operating under positive electrospray ionization mode connected to an LC-20AD (Shimadzu, Kyoto, Japan). ALR samples were separated from small molecules by reverse-phase chromatography on a C4 column (WatersXBridgeBEH300) using an acetonitrile gradient from 30% to 71.4% with 0.1% trifluoroacetic acid as the mobile phase in 25 minute. Data were acquired from  $m/z$  350 to  $m/z$  25,000 at a rate of 1 second/scan.

### **5.2.4 UV-Vis Spectroscopy**

UV-Vis spectra were recorded using a HP8453 diode array spectrophotometer (Hewlett-Packard, Palo Alto, CA). DTT and ALR were prepared in 20 mM Tris and 1

mM EDTA (pH 7.5) and mixed to give final concentrations of 20  $\mu$ M ALR and 5 mM DTT. DTT solutions were standardized with 5,5'-dithiobis-(2-nitrobenzoate).

### **5.2.5 Thermal and Chemically Induced Unfolding Transitions**

Thermal unfolding transitions of 10  $\mu$ M ALR in 10 mM phosphate (pH 7.5) were followed by circular dichroism using a J-810 circular dichroism spectropolarimeter (Jasco, Essex, UK). Spectra were collected between 190 and 260 nm, in 2  $^{\circ}$ C increments from 40 to 96  $^{\circ}$ C using a 10 mm cell. Data were analyzed using the mean residue ellipticities ( $\text{deg cm}^2/\text{dmol}$ ) at 222 nm as a function of temperature using a two-state model. Data were an average of two independent measurements. Flavin dissociation was monitored by recording the decrease in absorbance at 496 nm in the presence of 5 M guanidine hydrochloride. Spectra were recorded at room temperature using 10  $\mu$ M ALR in 20 mM Tris buffer, 1 mM EDTA and 5 M guanidine hydrochloride (pH 7.5).

### **5.2.6 Crystallization and Data Collection**

Protein crystals were obtained by hanging-drop vapor diffusion at 22  $^{\circ}$ C. The selenium-rich ALR reproducibly crystallized in a broad range of protein and precipitant concentrations. The best crystals formed in a reservoir solution of 20 mM MES buffer pH 6.5 and 16% polyethylene glycol 8000. Crystallization drops were prepared by mixing 1  $\mu$ L of selenium-rich ALR (10 mg/mL protein in 20 mM MES pH 6.5) with 1  $\mu$ L of the reservoir solution and placing over 250  $\mu$ L of the reservoir solution. Crystals appeared within hours and reached their maximum size in 2 weeks. Crystals were flash-cooled in liquid nitrogen following incubation of the crystal in a cryo-solution (25% w/v xylitol) made from the reservoir well solution. Diffraction

data were collected by Dr. Ming Dong using an in-house Rigaku RUH3R rotating anode generator with a RAXIS IV image plate area detector. X-ray diffraction data were collected at  $-180\text{ }^{\circ}\text{C}$  with a total of 180 one-degree oscillations (15 minute) from one crystal. The diffraction data were indexed and scaled with the program HKL2000 (Otwinowski Z, 1997).

### 5.2.7 Crystal Structure Solution and Refinement

The crystal of the selenium-rich ALR was indexed in the space group  $C222_1$  with cell dimensions a, b and c of 50.9, 76.9 and 63.5 Å, respectively. There was one protein subunit in the asymmetric unit and the solvent content was 34.7%. The crystal structure was solved by Dr. Ming Dong utilizing molecular replacement using the program MOLREP of CCP4 (Bailey, 1994) using the structure of human ALR (PDB code 3MBG) as a search model (Daithankar et al., 2010). During model building, refinement was carried out using the program REFMAC5 of CCP4. Model modifications were performed using the graphics program Coot (Emsley and Cowtan, 2004). In contrast to the search model ALR in which the electron density was clear only starting with residue Asp94, the entire N-terminus was built in the selenium-rich model, starting with Ser80. (Residue S80 is an extra residue at the N-terminus that remains after protease cleavage of the hexahistidine tag.) To facilitate the refinement of 90% Se and 10% S partial occupancy for cysteine and methionine residues, we used the program PHENIX (Adams et al., 2010) for the final rounds of refinement. Water molecules were placed during successive cycles of model building and refinement. The final model contains 126 amino acid residues, 1 FAD molecule and 112 water molecules. A final  $2F_o - F_c$  difference electron density map (Figure 5.8) for the model

confirmed the validity of the final model. The final  $R_{\text{working}}$  and  $R_{\text{free}}$  values were 0.200 and 0.221, respectively.

### 5.2.8 NMR Spectroscopy

Solution-state NMR spectra were acquired on a 14.1 T Bruker AV600 spectrometer equipped with a 5 mm broadband outer-coil liquid-state probe (Bruker) operating at a frequency of 114.493 MHz. Proton-decoupled  $^{77}\text{Se}$  spectra of selenium were recorded using spectral width of 96.1 KHz, an acquisition time of 0.085 second, with 250,000 scans, pulse delay of 0.5 second, a  $90^\circ$  pulse width of 15 microseconds and WALTZ decoupling with a 2.78 kHz field. While the spin-lattice relaxation time,  $T_1$ , was not measured for ALR, values for selenocysteine in a variety of other selenoproteins range between 0.9 and 1.5 seconds at a field of 14.1 T. NMR spectra were processed using backward linear prediction of the first four points of the free induction decay and 100 Hz exponential apodization.  $^{77}\text{Se}$  chemical shifts are reported with respect to diphenyl diselenide used as an external secondary chemical shift reference standard, set at 463.0 ppm (dimethyl selenide as primary reference at 0 ppm) (Duddeck, 1995). Sample temperature was 25 °C, controlled to within 0.1 °C. The NMR sample contained 2 mM ALR in 50 mM phosphate, 1 mM EDTA and 10%  $\text{D}_2\text{O}$  (pH 7.5) in a volume of 350  $\mu\text{L}$  in a Shigemi NMR tube (Shigemi Inc., Allison Park, PA). ALR was enriched to 50%  $^{77}\text{Se}$ . The sample was reduced by adding sodium dithionite to the NMR tube to a final concentration of 6 mM in a gas-tight Shigemi NMR tube (Wilmad LabGlass, New Jersey).

### 5.2.9 Accession Numbers

The atomic coordinates for the structure of selenium-substituted ALR have been deposited in the PDB (PDB code 3U5S).

### 5.3 Method Development and Application

While detection of carbon, nitrogen and proton resonances is routine in biological NMR, the only NMR-active isotope of sulfur,  $^{33}\text{S}$ , is a low-sensitivity quadrupolar nucleus unsuited for detection in biological systems. The NMR-active isotope,  $^{77}\text{Se}$ , may be used as a surrogate in order to gain insight into the multifaceted roles of sulfur in biology. Selenium is located below sulfur in the Periodic Table and shares many physicochemical properties, including electronegativity, van der Waals radius and redox states (Wessjohann et al., 2007). Substitution of cysteine's sulfur by selenium, generating selenocysteine, occurs in nature (Lobanov et al., 2009), does not generally compromise function, and can even generate enzymes with new properties (Johansson et al., 2005).

The full potential of selenium NMR in biology has yet to be realized. While selenium is easily detected with conventional hardware and experiments, its large chemical shielding response brings about efficient relaxation routes resulting in short transverse relaxation rates that broaden  $^{77}\text{Se}$  lines in solution (Gettins and Wardlaw, 1991). This problem is not pronounced for L-selenomethionine whose chemical shift tensor span is  $\Omega=580$  ppm (Potrzebowski et al., 1999) and can be detected with  $^{77}\text{Se}$  at natural abundance (7.5%) (Zhang and Vogel, 1994). However, isotopic enrichment is necessary for selenocysteine whose resonance widths are reported to be broader in biological systems (see below) (Gettins and Wardlaw, 1991). The span of selenium chemical shift tensors in diselenide bonds of organic molecules range between 500

and 900 ppm (Demko and Wasylishen, 2009) and the span of the dimeric form of selenocysteine, L-selenocystine, is  $\Omega=601.5$  ppm (Rozovsky laboratory, unpublished measurements). However, the span of the chemical shift tensor of selenylsulfide bonds is expected to be larger (Gettins and Wardlaw, 1991). While the chemical shift tensor of selenocysteine has not been reported, measurements in the Rozovsky group and others suggest it is indeed a more difficult target to detect (the resonance line widths at half-height are reported to be 150-500 Hz at a magnetic field of 14.1 T).

Earlier NMR studies were hindered by the availability of selenium-rich proteins (specifically selenocysteine enriched proteins) and by low sensitivity in the absence of isotopic enrichment (Gettins and Crews, 1991; House et al., 1992; Luthra et al., 1982). Selenoproteins were prepared by chemical reactions with small seleno-compounds (Luthra et al., 1982), or by using an animal diet enriched with  $^{77}\text{Se}$  (Gettins and Crews, 1991). These earlier studies described direct detection of specific biological species such as selenenic ( $\text{ESeO}_2\text{H}$ ), selenenylsulfide ( $\text{ESeSR}$ ) and the selenolate form ( $\text{ESe}^-$ ) in selenosubtilisin (House et al., 1992), denatured ribonuclease A, lysozyme (Luthra et al., 1982) and glutathione peroxidase (Gettins and Crews, 1991). Relaxation properties were described for  $^{77}\text{Se}$  introduced by reacting free thiol groups of proteins with the selenium analog of Ellman's reagent [5,5'-dithiobis-(2-nitrobenzoate)] (Gettins and Wardlaw, 1991). More recently,  $^{77}\text{Se}$ -selenocysteine-containing peptides were prepared by solid-phase peptide synthesis (Mobli et al., 2009, 2011) allowing assignment of diselenide connectivity in a 37-residue spider toxin (Mobli et al., 2009) and determination of  $\text{pK}_a$  of individual selenocysteine residues in a bioactive peptide hormone and neurotransmitter (Mobli et al., 2011).

These studies reaffirmed the sensitivity of  $^{77}\text{Se}$  NMR to its environment and established its ability to provide structural information for small peptides.

As the paucity of previous biological selenium NMR data testifies, current methods for preparation of  $^{77}\text{Se}$  selenocysteine-rich proteins are not broadly applicable. Solid-state synthesis allows labeling of residues at specific sites but has a molecular weight limit, and along with chemical ligation requires specialized equipment and expertise. Incorporation of selenocysteine in proteins by *E. coli*'s genetic incorporation machinery offers specificity but usually necessitates introducing mutations in the target protein (Johansson et al., 2004). Depending on the specific system it may also suffer from low yield (Bar-Noy and Moskovitz, 2002; Su et al., 2005). Consequently, it has yet to be used for the large-scale production necessary for NMR experiments. The enrichment of proteins with seleno-amino acids by supplementing *E. coli*'s growth media with the respective amino acids (Salgado et al., 2011; Strub et al., 2003) is an efficient and high yield method. However, for isotopic labeling with  $^{77}\text{Se}$ , the  $^{77}\text{Se}$ -enriched amino acids are not commercially available and the cost of their custom synthesis presents an obstacle for their routine utilization. Further, in-house synthesis of  $^{77}\text{Se}$  selenocysteine (Moroder, 2005; Stocking et al., 1997; Wessjohann and Schneider, 2008) is not generally accessible to the non-specialist given selenium's toxicity. Hence, none of the current methods to isotopically enrich proteins with  $^{77}\text{Se}$  offer an inexpensive, nonspecialized procedure generally applicable to proteins irrespective of their amino acid sequence and molecular weight. To gain additional insight into the capabilities of  $^{77}\text{Se}$  NMR and electron paramagnetic resonance in larger biological systems we have developed such a procedure for the enrichment of proteins with  $^{77}\text{Se}$  by heterologous expression in *E. coli*. Here, we

describe a broadly applicable, cost effective, selenium enrichment method that circumvents the need for costly (and hazardous) chemical synthesis. Furthermore, the selenium enrichment method described here is not only compatible with  $^{77}\text{Se}$  isotopic enrichment but also with incorporation of  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^2\text{H}$ . It also allows the ratio of sulfur-to-selenium substitution to be controlled. Therefore, this new method can tailor proteins with different properties by fine-tuning the percent of selenium incorporation. Thus, for disulfide-containing proteins, partial substitution of sulfur with selenium generates proteins rich in selenylsulfides (S-Se) bonds, while diselenides (Se-Se) predominate at higher substitution ratios. Since the redox potential of sulfur and selenium differs (Metanis et al., 2006) this could potentially lead to subtle differences in their reactivity (see below).

#### **5.4 Method to Incorporate Selenium into Proteins**

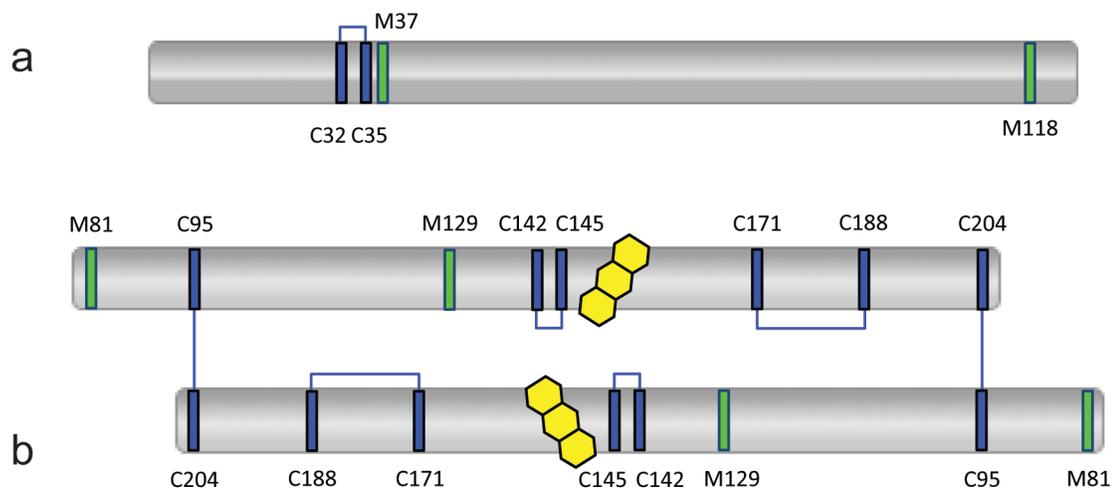
Because of the toxicity of selenium, previous investigations of Se uptake and its random incorporation in *E. coli* proteins focused only on the effect of limited amounts of selenium (Cowie and Cohen, 1957; Müller et al., 1997; Tuve and Williams, 1961). We have tested several selenium concentrations and means of supplementation to optimize a defined growth medium compatible with both  $^{77}\text{Se}$  labeling of sulfur sites in methionine and cysteine and the incorporation of  $^{13}\text{C}$ ,  $^{15}\text{N}$ , and  $^2\text{H}$ . A conspicuous benefit of our method is that it utilizes  $^{77}\text{selenite}$ , which is readily obtained from the cost-effective elemental  $^{77}\text{Se}$ . Hence the cost per liter of growth media is comparable to that of  $^{15}\text{N}$ . *E. coli* cultures are grown in the presence of limited amounts of sulfur, so that the sulfur pool is exhausted prior to induction. Once growth is stalled, protein expression is induced, and the requisite mixture of sulfate/selenite is added to a final concentration of 100  $\mu\text{M}$  in two steps, spaced 4

hours apart, to achieve the desired sulfur-to-selenium substitution ratio (see details in Materials and Methods). Table 5.1 shows that this method affords workable expression levels and reproducible ratios of substitution for two structurally diverse proteins. *E. coli* thioredoxin contains two cysteine residues, forming a redox-active disulfide bond, and two methionine residues (Figures 5.1 and 5.2). As described elsewhere in this dissertation, the structurally more complex protein ALR is a FAD-linked sulfhydryl oxidase that is involved in signal transduction and oxidative protein folding in the mitochondrial intermembrane space (Daithankar et al., 2009; Farrell and Thorpe, 2005). The short-form of this protein contains two catalytic disulfides, four structural disulfides, and four methionine residues per homodimer of 32 kDa (Figure 5.1). Interestingly, we have not detected a noticeable bias in selenium incorporation at specific sites even though the incorporation of selenium at different positions may influence the folding path in *E. coli* (Lees, 2012; Metanis and Hilvert, 2012).

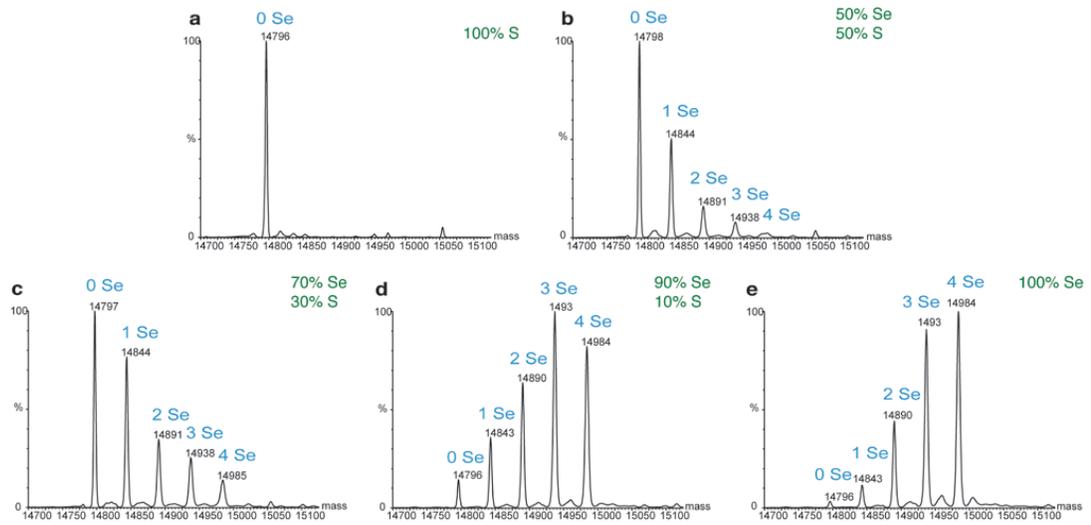
**Table 5.1 Sulfur-to-selenium substitution ratio and protein yield for growth media with varying ratios of sulfate and selenite.**

Protein	Sulfate in growth media ( $\mu\text{M}$ )	Selenite in growth media ( $\mu\text{M}$ )	Yield (mg/L)	Selenium incorporation (% Se)
Trx	5000	0	67 $\pm$ 3	0 <sup>a</sup>
	70	30	47 $\pm$ 7	12 $\pm$ 3
	50	50	46 $\pm$ 8	22 $\pm$ 5
	30	70	35 $\pm$ 4	25 $\pm$ 1
	10	90	28 $\pm$ 2	68 $\pm$ 1
	0	100	10 $\pm$ 4	78 $\pm$ 2
ALR	100	0	23 $\pm$ 3	0 <sup>a</sup>
	20	80	12.5 $\pm$ 1	56 $\pm$ 6
	0	100	4 $\pm$ 0.5	87 $\pm$ 7

Error was estimated based on the range of two independent measurements of yield and at least three independent measurements of selenium incorporation from different batches. <sup>a</sup> Within the measurement accuracy.

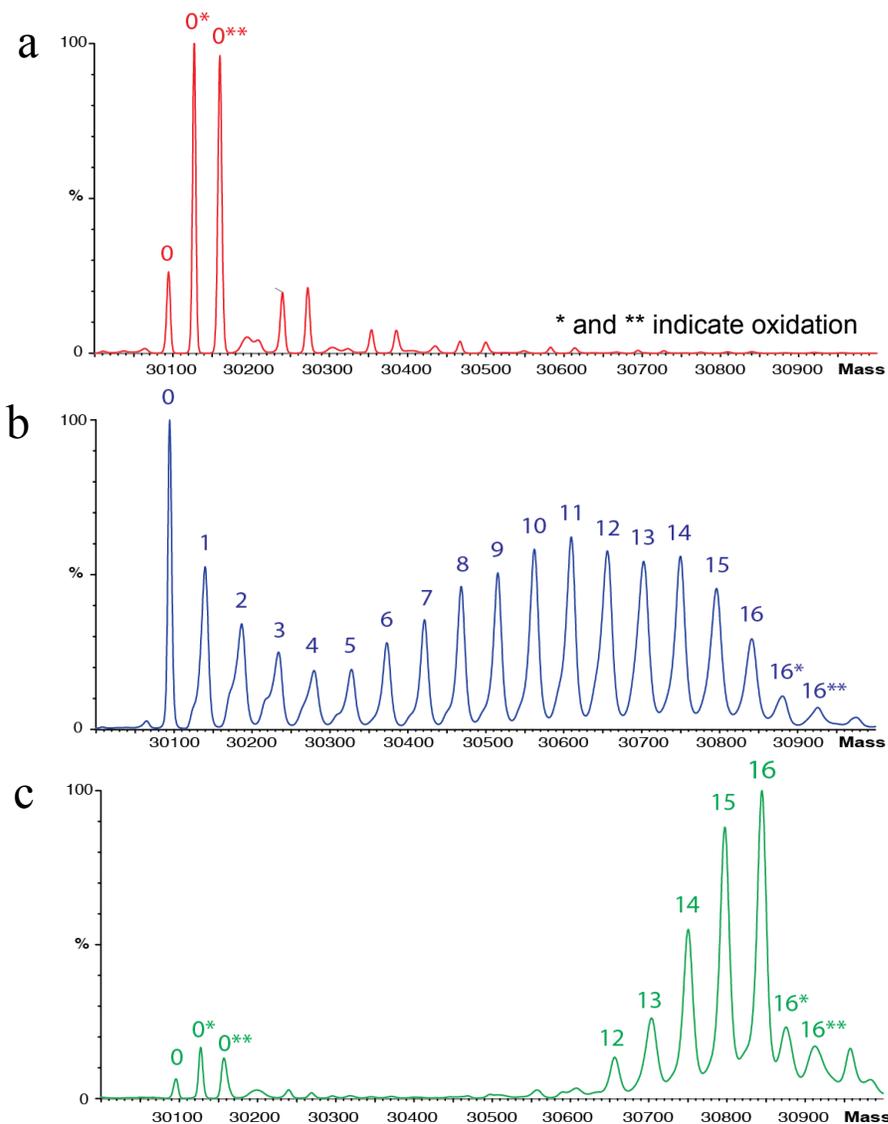


**Figure 5.1 Schematic representation of the two proteins used in this study.** a. In addition to two cysteines forming the catalytic disulfide (C32-C35) there are two methionine residues in *E. coli* Trx. b. Depiction of the short-form of human ALR used in these studies. Each subunit of the covalent homodimer has two methionines and six cysteines (green and blue bars respectively). The catalytic cysteines, C142-C145, interact with the flavin cofactor, shown in yellow. Figure used with permission from the publisher.

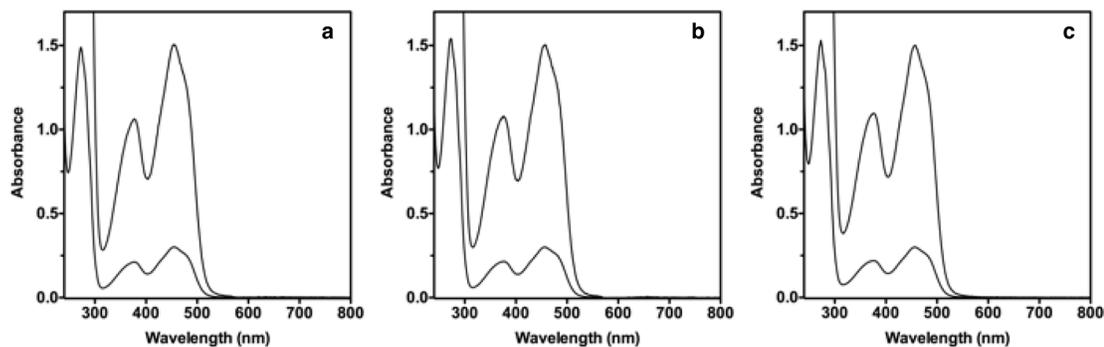


**Figure 5.2** Mass spectrometry of *E. coli* Trx as a function of selenite concentration in the growth media. a. Mass spectrometry of Trx grown solely on sulfate, b. Trx grown on 50  $\mu$ M selenite and 50  $\mu$ M sulfate, c. Trx grown on 70  $\mu$ M selenite and 30  $\mu$ M sulfate, d. Trx grown on 90  $\mu$ M selenite and 10  $\mu$ M sulfate, and e. Trx grown on 100  $\mu$ M selenite. Figure used with permission from the publisher.

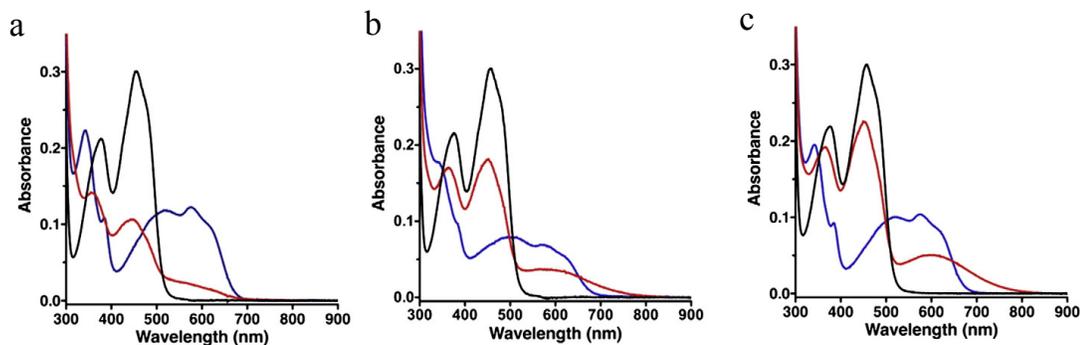
We use the relatively cysteine-rich ALR to evaluate the effect of selenium substitution on the enzymatic activity, stability, and three-dimensional structure of this cytokine-like protein. Incorporation of up to ~90% selenium into ALR (Figure 5.3), still showed UV-Vis spectra of the purified oxidized enzyme comparable to that of the native protein (Figure 5.4). Selenium-substituted ALR was also enzymatically active. Thus the flavin cofactor is readily reducible with the formation of the expected blue neutral flavosemiquinone upon aerobic incubation with the model substrate DTT (Figure 5.5) (Farrell and Thorpe, 2005). Further, replacement of the active site disulfide by a diselenide forms a new charge-transfer intermediate with the flavin prosthetic group (visualized 5 seconds after mixing), with an absorbance band extending  $> 750$  nm (Figure 5.5c).



**Figure 5.3 Mass spectroscopy of ALR as a function of selenite concentration in the growth media.** a. ALR grown solely on sulfate. The star denotes oxidation. Unlabeled peaks at increasing molecular weight are trifluoroacetic acid adducts. b. ALR grown in a 1:4 sulfate/selenite molar ratio mixture yielding a selenium incorporation of ~60%. c. ALR grown solely on selenite, where ~90% of the sulfur atoms are replaced by selenium. Figure adapted with permission from the publisher.



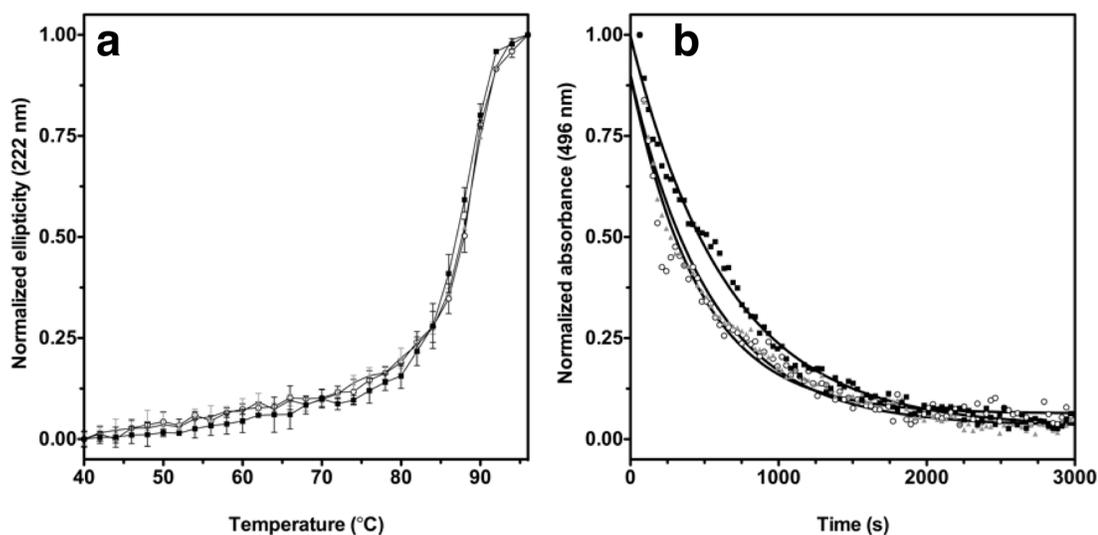
**Figure 5.4 UV-Vis spectra of ALR with varying selenium composition.** a. UV-Vis spectrum of the native ALR, b. UV-Vis spectrum of ALR grown on a 1:4 sulfate/selenite molar ratio mixture yielding a selenium incorporation of ~60%, c. UV-Vis spectrum of ALR grown solely on selenite where ~90% of the sulfur atoms are replaced by selenium. A 5-fold magnification of the visible region is also presented to show details of flavin absorbance spectra. Figure used with permission from the publisher.



**Figure 5.5 Reduction of selenium-substituted ALR by DTT.** The black line represents the enzyme before addition of DTT, red line 5 seconds and blue line 5 minutes after adding DTT. a. The spectra at 5 minutes show native and selenium-substituted enzymes form the blue flavosemiquinone. This species is developing at 5 seconds in panel a for native ALR, whereas panels b and c reveal a featureless long-wavelength charge-transfer band with the oxidized flavin. b. ALR with ~60% selenium incorporation (~48% of the selenium atoms are in selenylsulfide bonds) has a charge-transfer band, easily discerned at wavelengths > 550 nm. c. Selenium-rich ALR with ~90% selenium incorporation (~81% of the selenium atoms are in diselenide bonds) displays the highest levels of charge-transfer complex. Figure used with permission from the publisher.

## **5.5 Thermal and Chemical Stability**

In addition to the role of selenium in modulating the behavior of the redox-active cysteines, ALR contains both intra- and inter-chain structural disulfides (Figure 5.1b). It was therefore of interest to determine the impact of selenium substitution on both the global thermal stability of ALR and on the rate of flavin release rate from the enzyme (Figure 5.6). Both approaches show that the stability of selenium-based ALR is only slightly lower than that of the native ALR (Daithankar et al., 2010).



**Figure 5.6 Thermal stabilities of native and selenium-substituted ALR.** a. Thermal denaturation of native and selenium-substituted ALR proteins followed by CD. Spectra were recorded in 10 mM phosphate buffer, pH 7.5, from 40 to 96 °C for ALR with 0, 60 and 90% selenium substitution (■, ▲ and ○ respectively). The data represent the average of two determinations (error bars show standard error mean). b. Stability of native and selenium-substituted ALR monitored by FAD release. The proteins were diluted into 5 M guanidine HCl at pH 7.5 and the release of FAD was monitored by decreased absorbance at 496 nm for ALR with 0, 60 and 90% selenium substitution (■, ▲ and ○ respectively). The average of two datasets was expressed as a percentage of the total absorbance change and the solid lines are fit to a single exponential decay. Figure used with permission from the publisher.

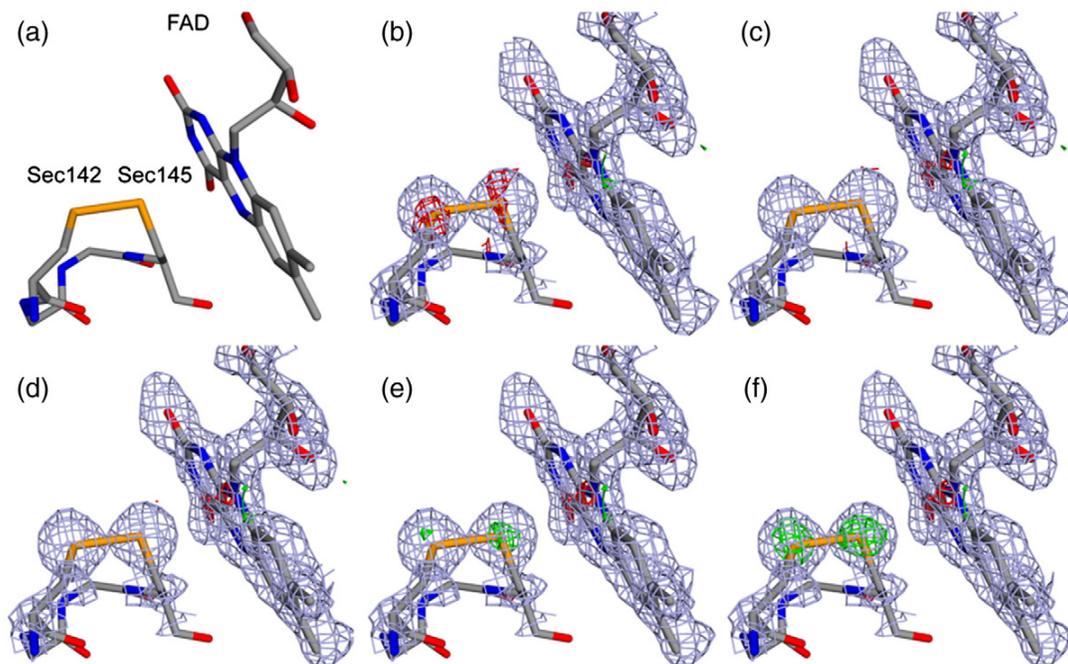
## 5.6 Crystal Structure

A more detailed picture of the effect of substitution emerges from the structure of the selenium-rich ALR solved by X-ray crystallography (resolution of 1.5 Å,  $R_{\text{free}}$  of 0.221, PDB code 3U5S, Table 5.2). The high-resolution data provided the first opportunity to refine a protein X-ray structure with occupancy of sulfur and selenium using the substitution ratio determined by mass spectrometry and inductively coupled plasma spectroscopy (Figures 5.7). Alignment of the native ALR structure against selenium-rich ALR shows an RMSD of 0.305 Å for all atoms (Figure 5.8). The only significant structural changes reflect the longer Se-Se and Se-C bonds and the corresponding adjustments in dihedral angles (Table 5.3). For selenium-rich ALR, these torsion angles are 2-4° larger than those of the native sulfur-containing protein.

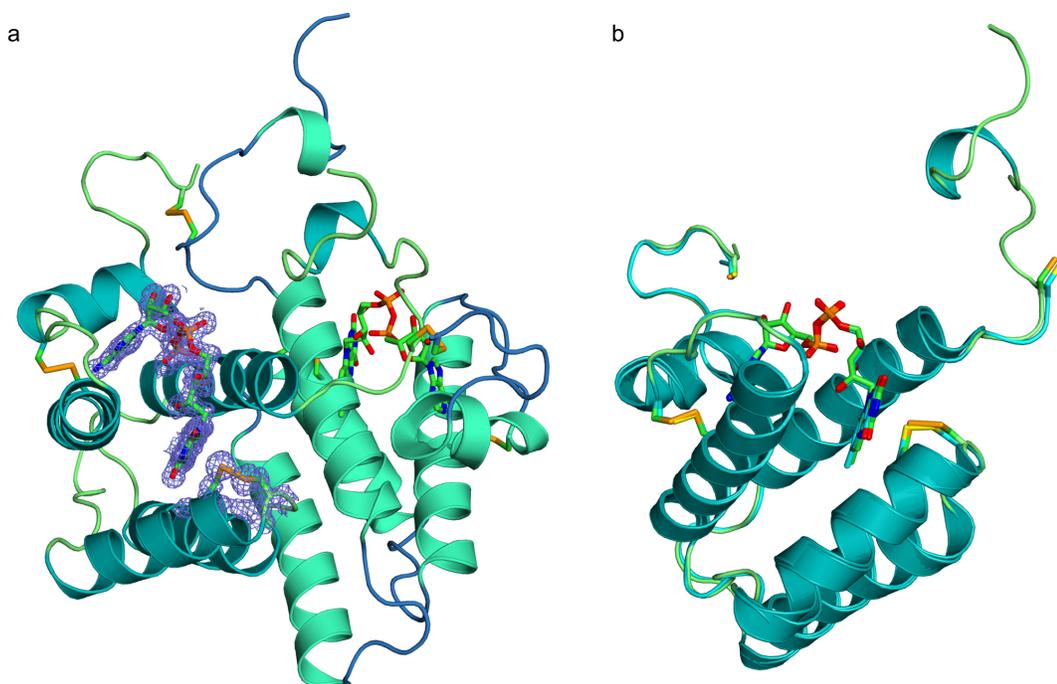
**Table 5.2 X-ray data collection and refinement statistics.**

<b>Crystal parameters</b>	
Space group	$C222_1$
Unit cell dimensions <i>a, b, c</i> (Å)	50.90, 76.88, 63.46
<b>Data collection</b>	
Resolution (Å)	50.0-1.50 (1.55-1.50) <sup>a</sup>
Completeness (%)	80.8 (36.4)
Redundancy	6.2 (2.0)
$I / \sigma I$	39.0 (2.4)
$R_{\text{merge}}$ linear	0.050 (0.254)
<b>Refinement statistics</b>	
Refinement program	Phenix
Resolution (Å)	21.5-1.50
$R_{\text{working}} / R_{\text{free}}$	0.200/0.221
Number of atoms (non-hydrogen)	1245
Mean B value	19.4
RMSD bond lengths (Å)	0.015
RMSD bond angles (°)	2.15
Ramachandran plot, most-favored	94.6%
Ramachandran plot, additionally and generously allowed region	5.4%

<sup>a</sup> The values in parentheses are for the highest resolution shell of 1.55 – 1.50 Å.



**Figure 5.7** Electron density difference maps of the disulfide/diselenide and flavin in the active site of ALR refined with different ratios of selenium and sulfur occupancy. a. A view of the active site residues near the flavin in CPK coloring. Panels b-f: the difference electron density maps computed following refinement with 60-100% selenium occupancy are shown for the redox diselenide (Sec142-Sec145) and flavin group of selenium-rich ALR. The  $2F_o-F_c$  map is drawn in grey contoured at  $1.2 \sigma$ . The  $F_o-F_c$  maps were contoured at  $3.0 \sigma$ , with positive difference density shown in green and negative difference density shown in red. b. 100% Se, 0% S. c. 90% Se, 10% S. d. 80% Se, 20% S. e. 70% Se, 30% S. f. 60% Se, 40% S. Figure used with permission from the publisher.



**Figure 5.8 Structural characterization of selenium-rich ALR.** a. The electron density difference map ( $2F_o-F_c$ ) is displayed around FAD and the adjacent selenium atoms in selenium-rich ALR dimer (PDB code 3U5S). The predominant diselenide bonds are shown as orange sticks. b. Overlay of the selenium-rich (diselenide bonds are shown as orange sticks and FAD in green) and native ALR (PDB code 3MBG, disulfide bonds are shown as yellow sticks and FAD in cyan). Figure used with permission from the publisher.

**Table 5.3 Diselenide bond lengths and angles. <sup>a</sup>**

Diselenide bond	Bond lengths selenium-rich ALR (Å)	Bond lengths native ALR (Å)	Dihedral angles selenium-rich ALR (°)	Dihedral angles native ALR (°)
Cys95-Cys204 (intermolecular)	2.31	2.01	-84.48	-81.73
Cys171-Cys188 (structural disulfide)	2.30	2.05	-86.88	-88.95
Cys142-Cys145 (active site)	2.34	2.03	65.65	61.97

<sup>a</sup> Bond lengths here represent an average of ~ 81% Se-Se, 9% Se-S, 9% S-Se, 1% S-S.

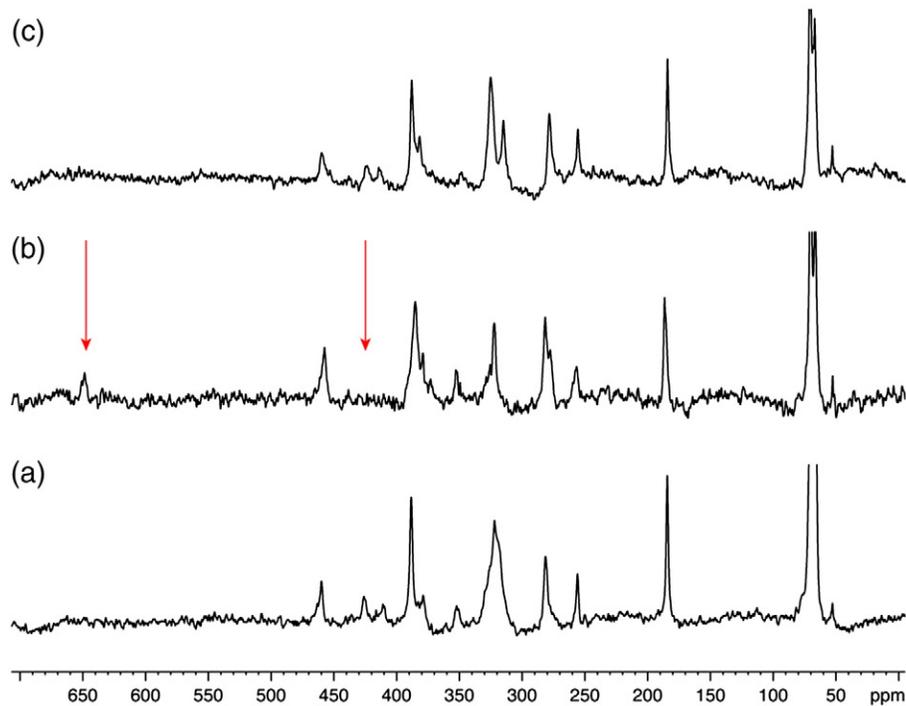
## 5.7 NMR Spectroscopy

An important facet of this selenium enrichment method is the ability to incorporate  $^{77}\text{Se}$  into all the sulfur-containing residues, hence allowing detection by NMR spectroscopy.  $^{77}\text{Se}$ , a spin  $I=1/2$  nucleus, has a pronounced chemical shielding response (Duddeck, 2004). The latter renders it highly sensitive to changes in the electronic environment, such as modification in bonding and conformation. Consequently,  $^{77}\text{Se}$  NMR measurements, in combination with theoretical calculations of the magnetic shielding tensor, will provide new approaches to probe the role of the local environment in shaping the reactivity of cysteines. Further,  $^{77}\text{Se}$  NMR could be used to investigate oxidative damage to selenocysteine and selenomethionine residues; address the binding of metals, ligands and proteins and detect changes in the local environment during catalysis, folding and conformational changes.

To demonstrate our ability to detect and resolve multiple selenium sites in proteins at positions normally occupied by sulfur, Figure 5.9 presents solution-state NMR detection of selenium in ALR at a magnetic field of 14.1 T ( $^1\text{H}$  frequency of 600 MHz,  $^{77}\text{Se}$  frequency of 114.493 MHz). Here, ALR was enriched to ~50% selenium to generate a combination of diselenide and selenylsulfide bonds. Based on the count of sulfur-containing amino acids in ALR – six cysteines and three methionines – we expect resonances of selenium in six selenylsulfide, six diselenide and three seleno-ether bonds. As shown, the resonances of most species are readily observed by direct detection of selenium. The half-height line width of selenium resonances ranges from 180 to 320 Hz, due to chemical shift anisotropy driven relaxation. The differences in line width suggest that the  $T_2$  relaxation time may be influenced by local protein dynamics. The isotropic chemical shifts of selenium in selenomethionine are between 50 and 70 ppm, a common range for selenomethionine in proteins (Zhang and Vogel,

1994). The isotropic chemical shifts of selenium in selenocysteine range from 185 to 460 ppm, where 10 resonances are resolved. The sample did not have resonances corresponding to oxidized selenium residues (Se-OH, Se-O<sub>2</sub>H and SeO<sub>3</sub>H at 1100-1300 ppm (Gettins and Crews, 1991)) or reduced selenocysteine (approximately -220 ppm (House et al., 1992)). This wide range for selenium resonances in diselenide and selenylsulfide bonds in the same protein is at first glance surprising, considering the fact that ALR's three selenylsulfide/diselenide bonds are all at water-exposed positions. However, studies in the Rozovsky laboratory indicate that even for a single selenocysteine within a given redox motif (such as Cys-Gly-Ala-Sec), it is possible to detect resonances that differ in chemical shifts by about 135 ppm and by a 7-fold difference in line width. Similar differences in chemical shift values were recently reported for Selenocysteine positioned in different locations in a peptide (Mobli et al., 2011). The isotropic chemical shifts are likely to depend on the dihedral angle at which the selenium is located (Potrzebowski et al., 1995; Tanioku et al., 2009), nearby charges and so on. Assignment of specific resonances is however not straightforward since information on how these factors affect selenocysteine resonances in proteins is sparse. While the chemical shifts of ALR have been reported and are available via the Biological Magnetic Resonance Data Bank (Banci et al., 2011), correlating Se resonances with other nuclei via two-dimensional experiments is complicated by the fast T<sub>2</sub> relaxation time. (The line width of individual resonances reported here range from 50 to 300 Hz.) In parallel, the assignment and interpretation of ALR selenium NMR spectra is continuing via genetic incorporation of selenocysteine in unique positions using *E. coli*'s innate insertion machinery (Arnér et al., 1999) as well as site-specific mutagenesis. It would be interesting to employ chemical shielding

calculations to further study the local environment of individual selenocysteine residues using the high-resolution structural information obtained by the X-ray refinement described above (see Figure 5.7 and Table 5.2).



**Figure 5.9**  $^{77}\text{Se}$  NMR spectroscopic characterization of flavin reduction in  $^{77}\text{Se}$ -labeled ALR. Proton-decoupled  $^{77}\text{Se}$  spectra of ALR with 50% selenium enrichment acquired at 14.1 T. a. Spectra of  $^{77}\text{Se}$ -labeled ALR with oxidized FAD (selenomethionine resonances are truncated for clarity). b. Spectra of  $^{77}\text{Se}$ -labeled ALR with reduced FAD under anaerobic conditions. Upon addition of a reducing agent, the resonance peaks at 412 and 426 ppm disappears, while a new resonance peak at 651 ppm appears (highlighted with an arrow). c. Same sample shown in panel **b** following exposure to air and removal of the reducing agent. Figure used with permission from the publisher.

Further assignment of the resonances and an additional assessment for  $^{77}\text{Se}$  NMR sensitivity is possible by comparing the spectra of ALR when bound to oxidized and reduced flavin (Figure 5.9). When the FAD is reduced with sodium dithionite (Daithankar et al., 2009; Farrell and Thorpe, 2005), two resonances at 412 and 426 ppm disappear. These two resonances were assigned to be selenylsulfide bonds by comparing the NMR spectra of ALR with different selenium incorporation levels (data not shown). The resonances reappear upon reoxidation of FAD and removal of the reducing agent (Figure 5.9c). Hence, we conclude that the resonances are likely to be the selenylsulfide bonds in the active site residues proximal to the flavin prosthetic group (Figure 5.8a). In addition to the changes in the 412 to 426 ppm region, a new resonance is detected at 651 ppm. The modification is reversible, as discerned from the disappearance of the 651 ppm resonance upon removal of the reducing agent and reoxidation of FAD (Figure 5.9c).

## 5.8 Conclusions

In sum, the reported substitution method provides a robust, flexible and general platform that can be used to enrich heterologous proteins in *E. coli* with selenium, without the need to modify expression systems, synthesize selenium compounds or calibrate sulfur/selenium consumption. We demonstrate the ability to acquire spectra of multiple selenium atoms in a mid-sized protein in a time-effective fashion. It is clear that selenium NMR can be used as a reporter for multiple sulfur/selenium locations without the need to introduce selenium in a site-specific fashion. This allows us to directly probe the environment, conformation and interactions of the site of interest and the capacity to directly detect modifications in selenocysteine residues and changes in the local environment taking place during catalysis. Hence,  $^{77}\text{Se}$

provides a powerful spectroscopic probe for the multiple roles of cysteine and methionine in protein structure, stability and function.

## REFERENCES

- Adams, P.D., Afonine, P. V, Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* *66*, 213–221.
- Arnér, E.S., Sarioglu, H., Lottspeich, F., Holmgren, A., and Böck, A. (1999). High-level expression in *Escherichia coli* of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the *selA*, *selB* and *selC* genes. *J. Mol. Biol.* *292*, 1003–1016.
- Bailey, S. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* *50*, 760–763.
- Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Kallergi, E., Lionaki, E., Pozidis, C., and Tokatlidis, K. (2011). Molecular recognition and substrate mimicry drive the electron-transfer process between MIA40 and ALR. *Proc Natl Acad Sci U S A* *108*, 4811–4816.
- Bar-Noy, S., and Moskovitz, J. (2002). Mouse methionine sulfoxide reductase B: effect of selenocysteine incorporation on its activity and expression of the seleno-containing enzyme in bacterial and mammalian cells. *Biochem. Biophys. Res. Commun.* *297*, 956–961.
- Blommel, P.G., and Fox, B.G. (2007). A combined approach to improving large-scale production of tobacco etch virus protease. *Protein Expr. Purif.* *55*, 53–68.
- Cowie, D.B., and Cohen, G.N. (1957). Biosynthesis by *Escherichia coli* of active altered proteins containing selenium instead of sulfur. *Biochim. Biophys. Acta* *26*, 252–261.
- Daithankar, V., Farrell, S., and Thorpe, C. (2009). Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space. *Biochemistry* *48*, 4828–4837.
- Daithankar, V., Schaefer, S.A., Dong, M., Bahnson, B.J., and Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augmenter of liver regeneration and

- characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* *49*, 6737–6745.
- Demko, B.A., and Wasylishen, R.E. (2009). Solid-state selenium-77 NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* *54*, 208–238.
- Duddeck, H. (1995). Se-77 nuclear magnetic resonance spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.* 1–323.
- Duddeck, H. (2004). Se-77 NMR spectroscopy and its applications in chemistry. *Reports NMR Spectrosc.* *52*, 105–166.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* *60*, 2126–2132.
- Farrell, S.R., and Thorpe, C. (2005). Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry* *44*, 1532–1541.
- Gettins, P., and Crews, B.C. (1991). <sup>77</sup>Se NMR characterization of <sup>77</sup>Se-labeled ovine erythrocyte glutathione peroxidase. *J. Biol. Chem.* *266*, 4804–4809.
- Gettins, P., and Wardlaw, S.A. (1991). NMR relaxation properties of <sup>77</sup>Se-labeled proteins. *J. Biol. Chem.* *266*, 3422–3426.
- Gladyshev, V.N. (1994). Nicotinic Acid Hydroxylase from *Clostridium barkeri*: Electron Paramagnetic Resonance Studies show that Selenium is Coordinated with Molybdenum in the Catalytically Active Selenium-Dependent Enzyme. *Proc. Natl. Acad. Sci.* *91*, 232–236.
- House, K.L., Dunlap, R.B., Odom, J.D., Wu, Z.P., and Hilvert, D. (1992). Structural characterization of selenosubtilisin by selenium-77 NMR spectroscopy. *J. Am. Chem. Soc.* *114*, 8573–8579.
- Johansson, L., Chen, C., Thorell, J.-O., Fredriksson, A., Stone-Elander, S., Gafvelin, G., and Arnér, E.S.J. (2004). Exploiting the 21st amino acid-purifying and labeling proteins by selenolate targeting. *Nat. Methods* *1*, 61–66.
- Johansson, L., Gafvelin, G., and Arnér, E.S.J. (2005). Selenocysteine in proteins-properties and biotechnological use. *Biochim. Biophys. Acta* *1726*, 1–13.
- Kapust, R.B., Tozser, J., Fox, J.D., Anderson, D.E., Cherry, S., Copeland, T.D., and Waugh, D.S. (2001). Tobacco etch virus protease: mechanism of autolysis and rational

design of stable mutants with wild-type catalytic proficiency. *Protein Eng. Des. Sel.* *14*, 993–1000.

Lees, W.J. (2012). Going non-native to improve oxidative protein folding. *Chembiochem* *13*, 1725–1727.

Lobanov, A. V, Hatfield, D.L., and Gladyshev, V.N. (2009). Eukaryotic selenoproteins and selenoproteomes. *Biochim. Biophys. Acta* *1790*, 1424–1428.

Luthra, N.P., Costello, R.C., Odom, J.D., and Dunlap, R.B. (1982). Demonstration of the feasibility of observing nuclear magnetic resonance signals of <sup>77</sup>Se covalently attached to proteins. *J. Biol. Chem.* *257*, 1142–1144.

Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* *20*, 71–75.

Metanis, N., and Hilvert, D. (2012). Strategic use of non-native diselenide bridges to steer oxidative protein folding. *Angew. Chem. Int. Ed. Engl.* *51*, 5585–5588.

Metanis, N., Keinan, E., and Dawson, P.E. (2006). Synthetic seleno-glutaredoxin 3 analogues are highly reducing oxidoreductases with enhanced catalytic efficiency. *J. Am. Chem. Soc.* *128*, 16684–16691.

Mobli, M., de Araújo, A.D., Lambert, L.K., Pierens, G.K., Windley, M.J., Nicholson, G.M., Alewood, P.F., and King, G.F. (2009). Direct visualization of disulfide bonds through diselenide proxies using <sup>77</sup>Se NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* *48*, 9312–9314.

Mobli, M., Morgenstern, D., King, G.F., Alewood, P.F., and Muttenthaler, M. (2011). Site-specific pK(a) determination of selenocysteine residues in selenovasopressin by using <sup>77</sup>Se NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* *50*, 11952–11955.

Moroder, L. (2005). Isosteric replacement of sulfur with other chalcogens in peptides and proteins. *J. Pept. Sci.* *11*, 187–214.

Müller, S., Heider, J., and Böck, A. (1997). The path of unspecific incorporation of selenium in *Escherichia coli*. *Arch. Microbiol.* *168*, 421–427.

Otwinowski Z, M.W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* *276*, 307–326.

Potrzebowski, M.J., Michalska, M., Blaszczyk, J., Wieczorek, M.W., Ciesielski, W., Kazmierski, S., and Pluskowski, J. (1995). *Molecular Modeling, X-ray Diffraction,*

- and  $^{13}\text{C}$ ,  $^{77}\text{Se}$  CP/MAS NMR Studies of Bis(2,3,4,6-tetra-O-acetyl- $\beta$ -D-glucopyranosyl) Diselenide and Disulfide. *J. Org. Chem.* *60*, 3139–3148.
- Potrzebowski, M.J., Katarzynski, R., and Ciesielski, W. (1999). Selenium-77 and carbon-13 high-resolution solid-state NMR studies of selenomethionine. In *Magnetic Resonance in Chemistry*, pp. 173–181.
- Salgado, P.S., Taylor, J.D., Cota, E., and Matthews, S.J. (2011). Extending the usability of the phasing power of diselenide bonds: SeCys SAD phasing of CsgC using a non-auxotrophic strain. *Acta Crystallogr. D. Biol. Crystallogr.* *67*, 8–13.
- Stocking, E.M., Schwarz, J.N., Senn, H., Salzmann, M., and Silks, L.A. (1997). Synthesis of L-selenocystine, L-[ $^{77}\text{Se}$ ]selenocystine and L-tellurocystine. *J. Chem. Soc. Perkin Trans. 1* 2443–2448.
- Strub, M.P., Hoh, F., Sanchez, J.F., Strub, J.M., Böck, A., Aumelas, A., and Dumas, C. (2003). Selenomethionine and selenocysteine double labeling strategy for crystallographic phasing. *Structure* *11*, 1359–1367.
- Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* *41*, 207–234.
- Su, D., Li, Y., and Gladyshev, V.N. (2005). Selenocysteine insertion directed by the 3'-UTR SECIS element in *Escherichia coli*. *Nucleic Acids Res.* *33*, 2486–2492.
- Tanioku, A., Hayashi, S., and Nakanishi, W. (2009). Analysis of one-bond Se-Se nuclear couplings in diselenides and 1,2-diselenoles on the basis of molecular orbital theory: torsional angular dependence, electron density influence, and origin in  $j(\text{Se}, \text{Se})$ . *Bioinorg. Chem. Appl.* 381925.
- Tuve, T., and Williams, H.H. (1961). Metabolism of selenium by *Escherichia coli*: biosynthesis of selenomethionine. *J. Biol. Chem.* *236*, 597–601.
- Wessjohann, L.A., and Schneider, A. (2008). Synthesis of selenocysteine and its derivatives with an emphasis on selenenylsulfide (-Se-S-) formation. *Chem. Biodivers.* *5*, 375–388.
- Wessjohann, L.A., Schneider, A., Abbas, M., and Brandt, W. (2007). Selenium in chemistry and biochemistry in comparison to sulfur. *Biol. Chem.* *388*, 997–1006.
- Zhang, M., and Vogel, H.J. (1994). Two-dimensional NMR studies of selenomethionyl calmodulin. *J. Mol. Biol.* *239*, 545–554.

## Chapter 6

### SPECIFIC INCORPORATION OF SELENIUM INTO THE PROXIMAL DISULFIDE OF AUGMENTER OF LIVER REGENERATION

#### 6.1 Introduction

As described earlier in this dissertation, human augmenter of liver regeneration (ALR) is a redox active flavoprotein that exists in two spliced forms. The long-form ALR (lfALR) is involved in the disulfide bond formation pathway in the intermembrane space of mitochondrion while the short-form (sfALR) is predicted to play a role in the complex liver regeneration cascade (Chapters 2 and 3). Both forms share a core flavin-binding domain, which contains a redox active CxxC motif proximal to the isoalloxazine ring. The mechanism of ALR involves transfer of reducing equivalents from the thiol substrate to the proximal CxxC motif to the FAD cofactor and finally to oxygen or cytochrome *c* as the final electron acceptor. Details of the structure, function and kinetic mechanism of ALR can be found in Chapters 2, 3 and 4.

Rozovksy and colleagues demonstrated that sfALR can be used as a stable platform for the substitution of sulfur for selenium (Schaefer et al., 2013). Chapter 5 shows the characterization of the selenium-containing sfALR used in that study and demonstrates that the structure of sfALR was essentially unaltered by the substitution of 90% of the sulfur atoms by selenium. Intriguingly, a new long-wavelength intermediate species was observed when DTT was added to the selenium-containing protein. Such feature is expected to arise from the interaction of a selenolate (E-Se<sup>-</sup>)

and the FAD. However, because the extensive substitution of sulfur for selenium occurs in all sulfur-containing amino acids, the identity of the species that contributed to the charge-transfer intermediate was unclear. This Chapter reports investigations of this feature by incorporation of selenium into sfALR in a site-specific manner. We were particularly interested in characterizing an ALR mutant in which the charge-transfer cysteine (C145) was mutated to a selenocysteine residue. We wished to examine whether the spectral behavior of this CxxU motif coincided with that of the predominantly UxxU form derived from the indiscriminant substitution of selenium into sfALR.

Here we describe the substitution of the sulfur closest to the FAD (C145) by selenium. This method required several mutations to the primary amino acid sequence of ALR to accommodate a selenocysteine insertion sequence (SECIS), which we characterize in the context of both the sulfur and selenium-containing sfALR. A SECIS element is a stem-loop feature in the mRNA of protein that allows the tRNA for selenocysteine to tightly bind the ribosome with the help of adaptor proteins (Hüttenhofer et al., 1996; Liu et al., 1998; Zinoni et al., 1990). This selenium containing ALR protein was expressed in low yield from *E. coli* but proved stable and enzymatically active. Finally spectrophotometric characterization of this mono-substituted sfALR derivative shows that the charge-transfer intermediate observed previously represents a charge-transfer interaction between selenocysteine 145 and the protein's FAD.

## **6.2 Materials and Methods**

### **6.2.1 Chemicals and Reagents**

All chemicals and reagents were supplied by Sigma-Aldrich (St. Louis, MO), Acros Organics (Geel, Belgium), Fisher Scientific (Pittsburg, PA), New England Biolabs (Ipswich, MA) and GoldBio (St. Louis, MO). All reagents and solvents were at least analytical grade and were used as supplied.

### **6.2.2 Design of Selenium Containing Constructs**

A sfALR gene was codon optimized for expression in *E. coli* and synthesized by DNA2.0 in the pJexpress414 vector. The gene contained the following mutations: C145U, R150V, K151R, C154D and C165S relative to the wild type sfALR. This construct was named SecALR1. A second construct, SecALR2, contained the following mutations: C145U, R150V, K151A, C154H and C165S (relative to wild type ALR). SecALR2 was generated using site-directed mutagenesis using SecALR1 as the template. The amino acids and nucleotide sequences of all constructs are provided in Supplemental Figures 6.7 and 6.8.

### **6.2.3 Subcloning and Site-Directed Mutagenesis Primers**

The following primers (IDT or Sigma) were used for subcloning the SecALR1 gene into the pTrcHisA and pET-28a vectors: (NheI site) 5'- AAA TTT GCT AGC ATG CGC ACC CAA CAA -3' and (HindIII site) 5'- ACC GAA AAG CTT AAT CGC AGG AAC CG -3'. The following mutagenesis primers were used for converting SecALR1 to SecALR2: 5'- GAC CTG GTT GCA CGC ACC GGA ACC AC -3' and 5'- GTG GTT CCG GTC CGT GCA ACC AGG TC -3'. To generate the SecALR2 U145C the following mutagenesis primers were used: 5'- CCG TGT GAG

GAG TGC GCT GAA GAC CTG G -3' and 5'- CCA GGT CTT CAG CGC ACT CCT CAC ACG G -3'. Subcloning and mutagenesis were as previously reported (Daithankar et al., 2009; Farrell and Thorpe, 2005) and confirmed using Genewiz Inc.

#### **6.2.4 Expression of SecALR1 and SecALR2**

Cells were grown in Terrific Broth that was inoculated with 10 mL overnight cultures, with 100 µg/mL ampicillin (to maintain the pJexpress414 vector) and 34 µg/mL chloramphenicol (to maintain the pSUABC vector; see below). Cells were grown at 37 °C to an OD<sub>600</sub> = ~2.2 and the temperature was reduced to 18 °C approximately 30 min prior to induction. Protein expression was induced with 1 mM IPTG and the media was supplemented with 5 µM Na<sub>2</sub>SeO<sub>3</sub> and 100 µg/mL L-cysteine. The cells were grown for 24 hours, harvested at 5000 g for 8 min and re-suspended in 50 mM potassium phosphate (pH 7.5) supplemented with 500 mM NaCl (binding buffer) before freezing.

#### **6.2.5 Purification of SecALR1 and SecALR2**

Purification of the selenium containing constructs was as previously reported (Schaefer et al., 2013) with the exception that approximately 100 µL of Ni-IDA (Invitrogen) resin was used per 3 L of culture medium to minimize non-specific binding (instead of using a 5 mL HisTrap FF column from GE Healthcare). The resin was washed with 10 mL of binding buffer followed by five 1 mL volumes of binding buffer containing 20 mM imidazole, four 1 mL volumes containing 200 mM imidazole, two 1 mL volumes containing 500 mM imidazole and two 1 mL volumes containing 1 M imidazole. Yellow fractions were combined and dialyzed against 50 mM potassium phosphate (pH 7.5), 1 mM EDTA. After dialysis, the sample was

centrifuged briefly to remove slight turbidity and analyzed for purity on reducing and non-reducing 16% tris-tricine SDS-PAGE.

#### **6.2.6 Expression and Purification of SecALR2 U145C**

Expression and purification of SecALR2 U145C in pJexpress414 vector was identical to expression and purification of wild type sfALR as previously reported (Daithankar et al., 2010).

#### **6.2.7 UV-Vis Spectroscopy**

UV-Vis spectra were recorded using an HP8453 diode-array spectrophotometer (Hewlett-Packard, Palo Alto, CA) as described previously (Daithankar et al., 2009, 2010; Farrell and Thorpe, 2005). Where necessary, turbidity correction software supplied with the diode-array spectrophotometer was used to correct for slight light scattering. SecALR2 and SecALR2 U145C were stored in 50 mM potassium phosphate (pH 7.5) containing 1 mM EDTA and mixed to give final concentrations of ~30  $\mu$ M ALR and 20 mM  $\beta$ -mercaptoethanol ( $\beta$ ME). Experiments performed at pH 9.0 used 20 mM Tris (pH 9.0) supplemented with 1 mM EDTA.

#### **6.2.8 Turnover with DTT**

The enzymatic activity was measured using a Clark-type oxygen electrode as described previously (Daithankar et al., 2010; Farrell and Thorpe, 2005).

## 6.3 Results and Discussion

### 6.3.1 Insertion of a SECIS Element into sfALR: Design of SecALR1 and SecALR2 Constructs

Selenocysteine is the 21<sup>st</sup> amino acid encoded by the UGA codon. Genetic incorporation methods utilize a SECIS element that allows for site-specific incorporation of selenocysteine (for reviews please see (Donovan and Copeland, 2010; Turanov et al., 2011)). In eukaryotes, the SECIS element is in the 3' untranslated region, while in bacteria the SECIS element is in the mRNA coding region, downstream of the UGA codon. Though this study utilizes the human ALR protein, we design a SECIS, similar to that of a bacterial selenoprotein, to incorporate selenocysteine into the active site of ALR. (For reviews of eukaryotic selenocysteine incorporation please see (Lobanov et al., 2009; Xu et al., 2007).)

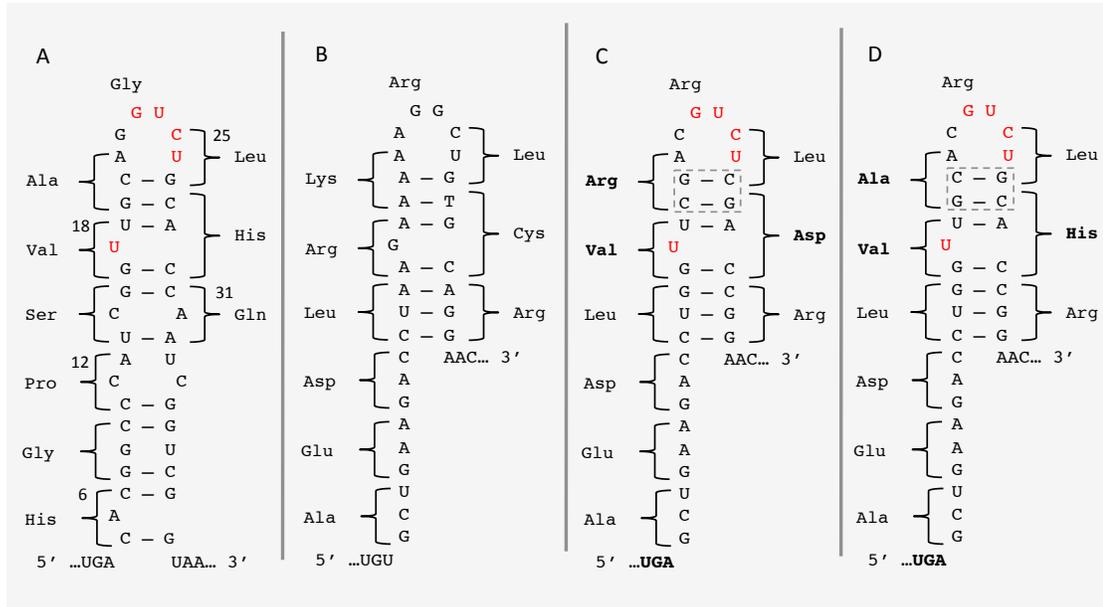
The incorporation of selenocysteine into a protein's amino acid sequence requires a specific tRNA for selenocysteine, selenocysteyl-tRNA<sup>Sec</sup>. In bacteria, the tRNA<sup>Sec</sup> (SelC) is charged with a serine by serine-tRNA synthetase (Leinfelder et al., 1988). This is followed by the conversion of seryl-tRNA<sup>Sec</sup> into selenocysteyl-tRNA<sup>Sec</sup> by selenocysteine synthase (SelA) (Forchhammer et al., 1991). The source of selenium is selenophosphate (Forchhammer et al., 1991; Veres et al., 1992), which is synthesized by selenophosphate synthetase (SelD) utilizing ATP and selenide (Glass et al., 1993). SelB, *E. coli*'s selenium specific elongation factor, binds specifically to selenocysteyl-tRNA<sup>Sec</sup> (Forchhammer et al., 1989) and the SelB binds to the loop of the SECIS element (Kromayer et al., 1996).

Using the SECIS element from *E. coli*'s formate dehydrogenase H (*fdhF* gene) as a model (Figure 6.1a), Liu and colleagues described which components of the SECIS were essential for selenocysteine incorporation (Liu et al., 1998). Their work

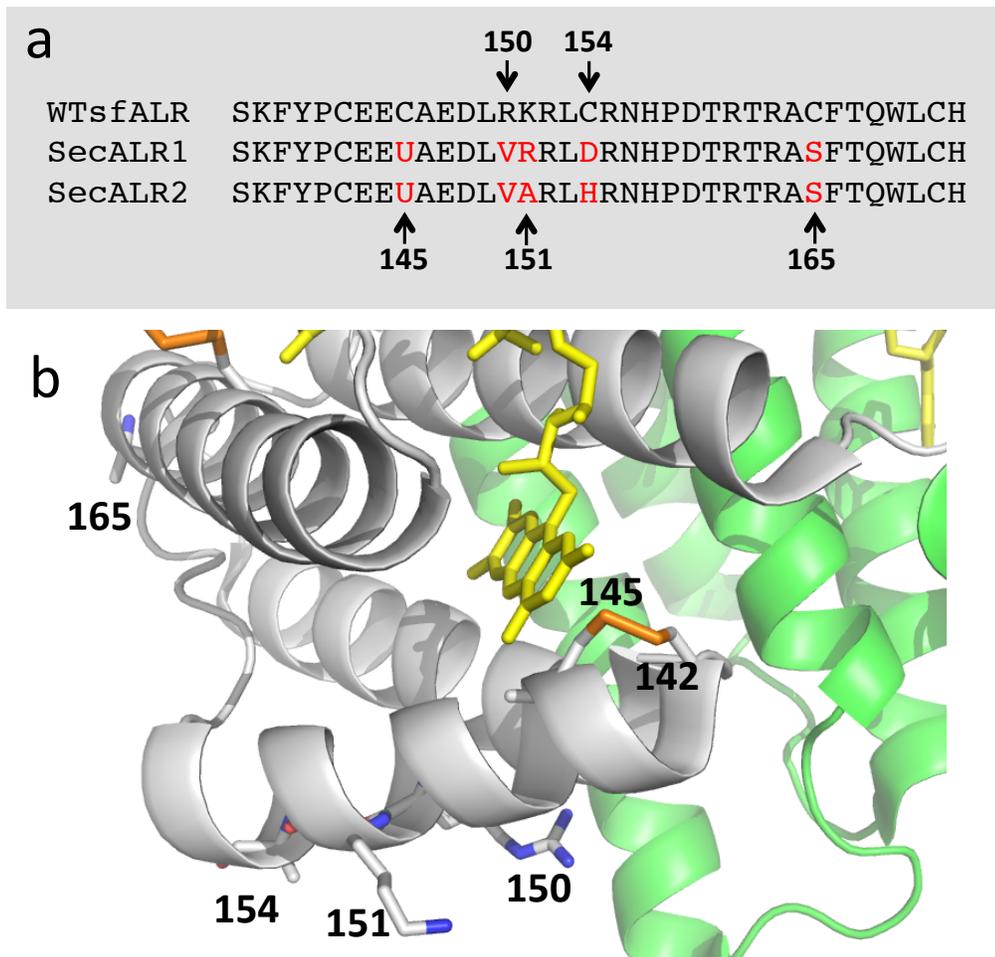
showed the 11 nucleotides after the UGA (where U is nucleotide number 1) do not have to be base paired. This is in contrast to the SECIS element present in wild type formate dehydrogenase H (Liu et al., 1998). A minimal “mini upper-stem loop”, consisting of 5 base pairs that harbor a U nucleotide bulge at position 17 followed by a 6 nucleotide loop region, is required for selenocysteine incorporation (Klug et al., 1997; Liu et al., 1998). The pairing of U18/A29 was demonstrated to be important, but it was also shown that the inversion mutation of A18/U29 still allows the incorporation of selenium (Liu et al., 1998). Heider and colleagues reported the inversion of the nucleotide pair C20/G27 to G20/C27 decreased protein expression by 30%, but the inversion of G19/C28 to C19/G28 had no effect on protein expression (Heider et al., 1992). Another study showed the inversion of the nucleotides at positions 20/27 in formate dehydrogenase H abolishes selenium incorporation (Liu et al., 1998). In formate dehydrogenase H, the guanine nucleotide at position 22 was demonstrated to be essential for efficient read through of the UGA codon (Heider et al., 1992). However, a recent study by Aldag and coworkers showed that a cytosine at the same location was tolerated by cytochrome P450 (Aldag et al., 2009). Additionally, the base pairings of non-specific nucleotides at positions 15/31, 16/30, 18/29 and 19/28 are required (Liu et al., 1998). Also, nucleotides G23, U24 and U17/U18 directly interact with elongation factor, SelB, and must be conserved (Baron et al., 1993; Heider et al., 1992; Hüttenhofer et al., 1996; Klug et al., 1997).

Taking the information described above into consideration, we created two selenium constructs, a construct with a substantially modified SECIS sequence and fewer mutations called SecALR1 and a construct with more conserved SECIS sequence and subsequently more mutations called SecALR2. Figure 6.1 contrasts the

original formate dehydrogenase H SECIS sequence with that introduced into the two ALR constructs (the nucleotides shown in red on Figure 6.1A must be completely conserved in our designs). The SecALR1 construct (Figure 6.1C) resulted in the following mutations: C145U, R150V, K151R, C154D and C165S in the context of wild type sfALR. There were 13 nucleotide mutations that resulted in four silent and five missense mutations. Figure 6.2 shows the location of the various mutations relative to the active site. R150 and K151 (from wild type ALR) are both located at solvent-exposed positions however the mutation of K151R is more conservative than R150V since both Lys and Arg are positively charged. In order to retain a charged amino acid at position 151 in ALR (resulting in K151R) nucleotides 19 and 20 (G and C, respectively) from the formate dehydrogenase stem loop are inverted and to maintain the proper Watson-Crick base pairings, nucleotides 27 and 28 also needed to be inverted (resulting in C154D). As described above, this could have severe consequences on the read through of the UGA codon (Heider et al., 1992; Liu et al., 1998). The R150V results in a Val at the same position in the conserved loop as in formate dehydrogenase H (compare nucleotides 16-18 on Figure 6.1A and C) and this allows for the proper U nucleotide bulge at position 17. The surface-accessible amino acid side chains at positions 154 and 165 were replaced by D and S, respectively. The mutations of C154D and C165S were expected to be tolerated because they are present at these positions in the highly homologous rat protein. The sfALR protein used for *in vitro* studies typically contains C154A and C165A mutations to prevent protein aggregation during purification as previously described (Farrell and Thorpe, 2005).



**Figure 6.1 Diagram of SECIS elements used in this study.** A) The SECIS element from *E. coli* formate dehydrogenase H. Nucleic acids numbered starting with the U in the UGA codon as number 1, B) wild type sfALR primary sequence oriented in a stem-loop for comparison, C) SecALR1 nucleotide sequence arranged in a stem-loop diagram, and D) SecALR2 nucleotide sequence arranged in a stem-loop feature. Locations highlighted in red are necessary for binding of *E. coli*'s elongation factor, SelB and can not be alternated. The dashed boxes indicate the differences between SecALR1(C) and SecALR2 (D). Bolded amino acids indicate a change from the wild type sequence. The UGA codon (selenocysteine) is bolded.



**Figure 6.2 Sequence comparison of constructs used in this study.** a) The sequences of wild type sfALR, SecALR1 and SecALR2. Mutations in the various constructs relative to the wild type sfALR are indicated in red. b) Location of residues mutated in this study. The redox active site disulfide (C142/C145) is in communication with the non-covalently bound FAD cofactor (yellow).

Using SecALR1 as a template, we designed a more conservative second construct using site-directed mutagenesis (Figure 6.1D). This construct, termed SecALR2, contained the following mutations: C145U, R150V, K151A, C154H and C165S (relative to wild type ALR) (Figure 6.2). SecALR2 contained 12 nucleotide mutations relative to wild type ALR with the nucleotides shown in red being conserved from the formate dehydrogenase H stem-loop shown in Figure 6.1A. Four of the nucleotide mutations resulted in three silent mutations and eight of the nucleotide mutations resulted in four missense mutations. Like the SecALR1 construct, SecALR2 has C145U, R150V and C165S mutations, but the mutations at positions 151 and 154 are different between the two constructs. In this second construct, Val, Ala, Leu and His at positions 150, 151, 153 and 154, respectively, are a result of the overall conservation of nucleotides 15-31 from formate dehydrogenase H (with the exception of guanine to cytosine mutation at nucleotide 22 to conserve the charged amino acid R152) making this second design more conservative when compared to SecALR1. A direct comparison of SecALR1 to SecALR2 will show the only difference is the inversion of nucleotide pairs 19/27 and 20/28 (see the dashed boxes in Figure 6.1C and D).

### **6.3.2 Subcloning**

The pJexpress414 SecALR1 construct contained a C-terminal 6xHis tag. We subcloned SecALR1 into two additional expression vectors with different promoters and affinity tag locations to test expression levels and ease of purification. We chose pTrcHisA with a trp-lac promoter (Invitrogen) because this vector has an N-terminal 6xHis affinity tag. This vector is used for wild type sfALR expression and the yield is typically around 30 mg/L. The pET-28a vector with a T7 promoter (Novagen) was

also selected because it also contains an N-terminal 6xHis affinity tag that typically increase protein yield (Waugh, 2005). After the sequence of SecALR1, in both pTrcHisA and pET-28a vectors, was confirmed, mutagenesis primers were utilized to convert SecALR1 into SecALR2 (see Materials and Methods).

In order to increase selenium incorporation efficiency and overall protein yield during expression, a second vector, pSUABC, was also transfected into the BL21(DE3) cell line. The vector pSUABC allows for the co-expression of SelA (selenocysteine synthase), SelB (elongation factor SelB) and SelC (tRNA<sup>Sec</sup>) (Arnér et al., 1999). The use of this additional vector can increase translation efficiency by up to 5 fold (Rengby et al., 2004).

### **6.3.3 Expression and Purification of SecALR**

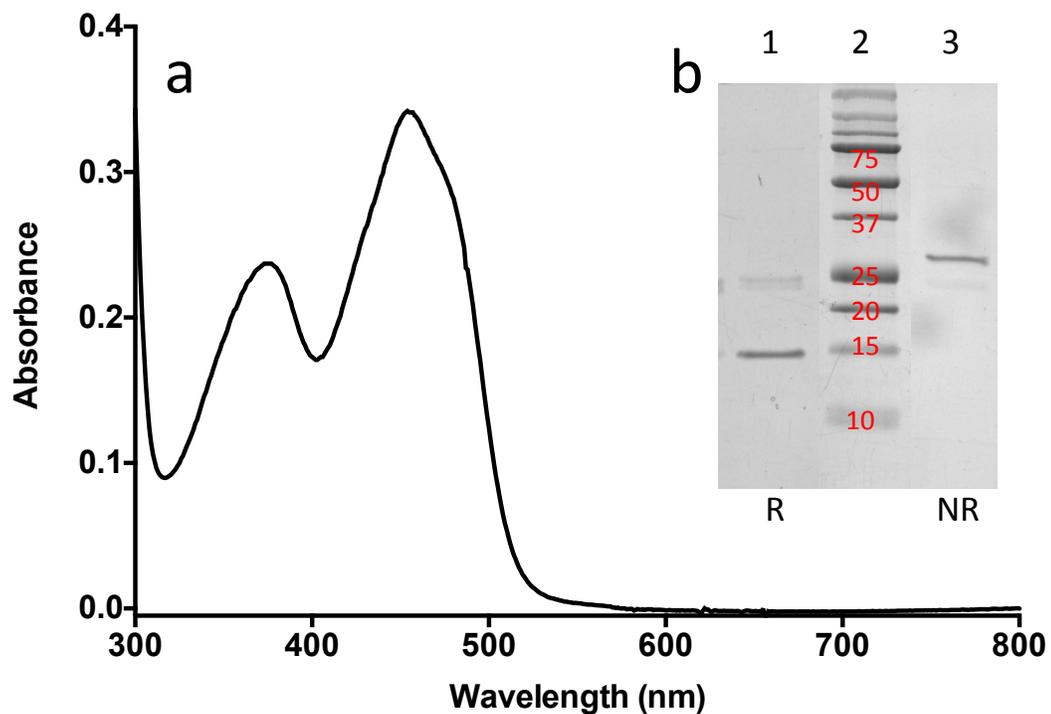
Expression and purification trials were performed for SecALR1 in several vectors (pJexpress414, pTrcHisA and pET-28a) using the cell line BL21(DE3). The expression of the protein (in any vector) could not be detected with SDS-PAGE (not shown). We first attempted a small-scale expression and purification using SecALR1 in the pJexpress414 vector and we detected full length SecALR1 using a Western blot (6x-His Epitope Tag Antibody, Pierce). However, when a large-scale expression and purification was attempted there was a considerable level of insoluble full-length protein and so initial efforts were made to reduce, denature and purify this material by Ni-IDA resins. However, refolding of the protein after exposure to denaturants proved to be problematic and negligible amounts of the folded protein (supplemented with FAD) were recovered. In summary, the SecALR1 construct proved to have low yield and the rest of the study employed the second design, SecALR2.

SecALR2 (in the pJexpress414 vector with the C-terminal 6xHis tag) could be purified under oxidizing and non-denaturing conditions with very little aggregated protein. Table 6.1 details the yield of sfALR, SecALR2 and the control SecALR2 U145C. This all-cysteine control allowed us to determine the impact of the three additional mutations that were the consequence of introducing the SECIS element (Figure 6.1). This construct was expressed using normal conditions in LB medium that contains no supplemented selenium (see Materials and Methods). Importantly, this triple mutant was expressed at ~8 mg/L (Table 6.1) and showed a spectrum characteristic of wild type sfALR. Hence, the additional mutations to the sfALR framework do not result major differences from the wild type protein. Next, we consider the consequence of replacing the charge-transfer cysteine (C145) with selenocysteine.

**Table 6.1 Yield and turnover number of wild type sfALR and SecALR2 all-cysteine variant.** Turnover number averages of two experiments. The yield for SecALR2 U145C was obtained from a single large-scale growth experiment.

Construct	Vector	Yield per Liter (mg)	Turnover with 5 mM DTT ( $\text{min}^{-1}$ )
sfALR	pTrcHisA	18.4	49
SecALR2	pJexpress414	0.7	8
SecALR2 U145C	pJexpress414	8.4	30

Figure 6.3 shows the UV-Vis spectrum of SecALR2. The inset shows SecALR2 run on SDS-PAGE under reducing and non-reducing conditions. We observed bands near the expected molecular weight of the monomer and dimer (15802 Da and 31604 Da, respectively) for SecALR2. Hence, we have demonstrated that it is possible to express and purify SecALR2 and that the resulting protein is relatively pure and well-folded. We next characterized the activity of this selenium-containing protein.



**Figure 6.3 UV-VIS spectrum of SecALR2.** a) The UV-Visible spectrum of the oxidized SecALR2 exhibited a typical flavin envelope for a well-behaved flavoprotein. b) SDS-PAGE analysis of SecALR2. Lane 1: SecALR2 – reducing conditions, Lane 2: protein ladder (molecular weights indicated in kDa), Lane 3: SecALR2 – non-reducing conditions. The expected molecular weight of the full length SecALR2 in pJexpress414 vector is 15802 Da and 31604 Da for the monomer (reducing conditions) and homodimer (non-reducing conditions), respectively. The truncated protein molecular weight is 7728 Da and the heterodimer (one full length monomer and one truncated monomer) would have a molecular weight of 23530 Da (non-reducing conditions). Most of the truncated protein was found in the insoluble fraction (not shown).

### 6.3.4 Turnover of SecALR2 and SecALR2 U145C

Table 6.1 summarizes the turnover information of wild type ALR, SecALR2 and SecALR2 U145C using 5 mM DTT. The wild type protein has a turnover number of  $49 \text{ min}^{-1}$ . Under these conditions, the SecALR2 construct and the control SecALR2 U145C have turnover numbers of  $8 \text{ min}^{-1}$  and  $30 \text{ min}^{-1}$ , respectively. Thus, the mutations required to introduce the SECIS element cause an approximate 40% decline in turnover number with DTT. When benchmarked against the control SecALR2 U145C, the conversion of C145U led to an approximate 4-fold reduction in enzymatic activity.

### 6.3.5 SecALR's Reactivity with $\beta$ ME

During characterization of the 90% selenium-enriched sfALR that contained mostly diselenide bonds in the active site (Chapter 5; Figure 5.5c), we observed a new intermediate when the protein was mixed with 5 mM DTT. At that time, we interpreted that feature as likely representing a charge-transfer species like that shown in Figure 6.4.



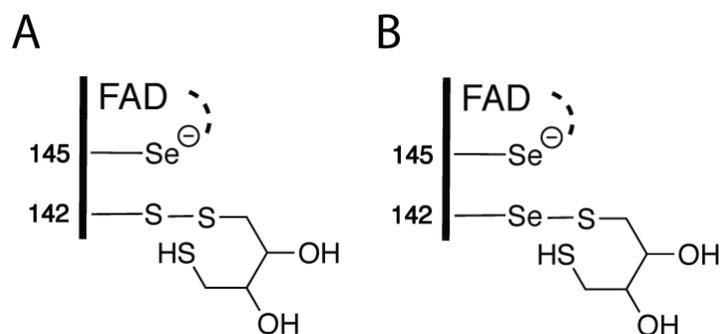
**Figure 6.4 Depiction of the charge-transfer complex between the electron-rich selenium at position 145 and the electron deficient FAD.**

This however would be rather surprising because it would be expected to rapidly reoxidize via transfer of reducing equivalents first to the flavin cofactor and then to molecular oxygen. Furthermore, it is unclear to what extent the new charge-transfer species was affected by the neighboring U142 residue. The work described in Chapter 4 allowed us to answer this issue unequivocally using  $\beta$ ME. There we showed that  $\beta$ ME is a poor substrate of sfALR but was able to rapidly generate mixed disulfide intermediates that formed prominent charge-transfer bands with the oxidized flavoprotein. This demonstrated that the charge-transfer species is almost exclusively due to the formation of a mixed disulfide, like that shown in Figure 6.5.

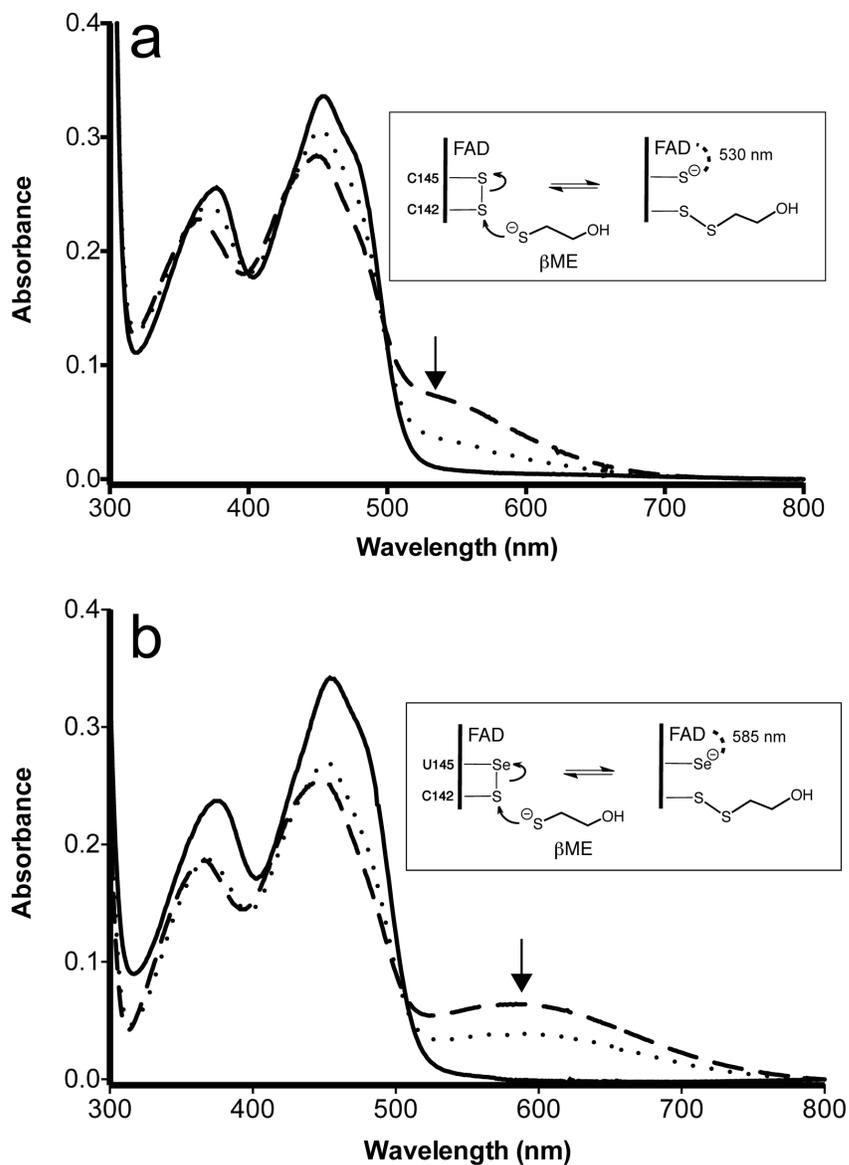
Even though DTT is a better model substrate for measuring ALR's activity, using DTT to study charge-transfer formation with CxxC (wild type ALR) or CxxU (SecALR2) motif is difficult due to the rapid resolution of the mixed disulfide (Figure 6.5A). As shown in Chapter 4, this intermediate species can only be detected using the stopped-flow spectrophotometer on the millisecond timescale. Attempts to observe the charge-transfer with SecALR2 using 5 mM DTT were unsuccessful, presumably due to the short lifetime of the selenolate-FAD interaction. However, as demonstrated in Chapter 4, the monothiol  $\beta$ ME can induce a long-lived charge-transfer interaction. Therefore we also examined the charge-transfer complex of a CxxU containing sfALR with  $\beta$ ME.

The experiment with  $\beta$ ME was carried out for the control protein construct (SecALR2 U145C) in Figure 6.6. The wedge-shaped absorbance feature extending from 700 nm to the leading edge of the oxidized flavin envelope is qualitatively similar to the wild type sfALR at pH 9.0 but is approximately half the intensity. The intensity loss presumably reflects the impact of the 3 mutations required to generate

the SECIS element. When examining the crystal structure of human sfALR (PDB 3MBG) we can see Arg 150 is making stabilizing interactions with Asp 147 and Glu 175 and Lys 151 also interacts with Glu 175. Mutation of these two positive residues at positions 150 and 151 could cause electrostatic repulsion that could somehow be affecting the protein's activity. At pH 9.0, where much more  $\beta$ ME is in the thiolate form ( $pK = 9.6$ ) than at pH 7.5 ( $\sim 25\%$  compared to  $\sim 0.8\%$ ), the long-wavelength band is more pronounced. While the long-wavelength band overlays with the flavin absorbance, it is clear that its maximum will be at about 530 nm. In contrast, the charge-transfer feature from SecALR2 shows a distinct peak at 585 nm, easily resolved from the flavin envelope. Thus, it is clear that the long-wavelength feature seen in the UV-Vis spectrum of the randomly substituted Se-rich ALR is dominated by a selenide to oxidized flavin complex (Chapter 5; Figure 5.5c), with a maximal charge-transfer absorbance essentially identical to that observed for the SecALR2 protein.



**Figure 6.5 Two possibilities of mixed disulfide formation between position 142 and DTT that results in a charge-transfer species.** Panel A depicts a mixed disulfide between C142 and DTT. The Se-FAD charge-transfer is short lived due to the rapid resolution of the mixed disulfide by the second thiol of DTT. This would be representative of SecALR2 with DTT. Panel B illustrates a mixed disulfide between U142 and DTT in the protein that also contains selenocysteine at position U145. The charge-transfer intermediate for this protein has a longer lifetime due to the slower resolution of the Se-DTT mixed disulfide. This is what is observed with the selenium-rich ALR from Chapter 5.



**Figure 6.6 Formation of a charge-transfer intermediate with  $\beta$ ME.** a) The formation of a charge-transfer between a thiolate and the proximal FAD cofactor with SecALR2 C145U construct. The solid line shows the spectrum of the oxidized protein at pH 7.5. The dotted line is 10 sec after the addition of 20 mM  $\beta$ ME at pH 7.5. The dashed line was recorded after comparable addition at pH 9.0. The inset depicts mixed disulfide bond formation as described in Chapter 4. b) The charge-transfer formation in the SecALR2 construct. The solid line shows the spectrum of oxidized protein at pH 7.5. The dotted and dashed lines represent the 10 sec after the addition of 20 mM  $\beta$ ME at pH 7.5 and 9.0, respectively.

These data, taken with the results in Chapter 4, strongly suggest that the reason why the “new” intermediate in Se-rich ALR from Chapter 5 was readily observed with DTT is because the adduct species forms rapidly in the presence of 5 mM DTT but decays much more slowly than in the situation where there is a conventional mixed disulfide between C142 and DTT (Figure 6.5).

#### **6.4 Conclusions**

We have demonstrated the successful design, expression and purification of a single selenocysteine containing ALR. To our knowledge this is the first demonstration of a charge-transfer complex between a selenide and oxidized flavin. It should be noted that in studies of sarcosine oxidase, Jorns and colleagues investigated the spectral effects of replacing the oxygen atom of the ether functional group in methyl sarcosine by sulfur, selenium and tellurium (Wagner et al., 2000). They found an approximate 60 nm red shift of the charge-transfer complex on substitution of sulfur by selenium. In our case the selenium at position 145 is positioned 3.2 Å directly above the C4a position of the isoalloxazine ring. The ~55 nm red-shift in the charge-transfer absorbance is certainly consistent with the expectation that the increased electron density and polarizability of Se over S would make the selenide a better charge-transfer donor and hence yield complexes which are red shifted compared to their sulfur counterparts (Ghisla and Massey, 1989).

## REFERENCES

- Aldag, C., Gromov, I.A., Garcia-Rubio, I., von Koenig, K., Schlichting, I., Jaun, B., and Hilvert, D. (2009). Probing the role of the proximal heme ligand in cytochrome P450cam by recombinant incorporation of selenocysteine. *Proc Natl Acad Sci U S A* *106*, 5481–5486.
- Arnér, E.S., Sarioglu, H., Lottspeich, F., Holmgren, A., and Böck, A. (1999). High-level expression in *Escherichia coli* of selenocysteine-containing rat thioredoxin reductase utilizing gene fusions with engineered bacterial-type SECIS elements and co-expression with the *selA*, *selB* and *selC* genes. *J. Mol. Biol.* *292*, 1003–1016.
- Baron, C., Heider, J., and Böck, A. (1993). Interaction of translation factor SELB with the formate dehydrogenase H selenopolypeptide mRNA. *Proc. Natl. Acad. Sci. U. S. A.* *90*, 4181–4185.
- Daithankar, V., Farrell, S., and Thorpe, C. (2009). Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space. *Biochemistry* *48*, 4828–4837.
- Daithankar, V., Schaefer, S.A., Dong, M., Bahnson, B.J., and Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augmenter of liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry* *49*, 6737–6745.
- Donovan, J., and Copeland, P.R. (2010). Threading the needle: getting selenocysteine into proteins. *Antioxid Redox Signal* *12*, 881–892.
- Farrell, S.R., and Thorpe, C. (2005). Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry* *44*, 1532–1541.
- Forchhammer, K., Leinfelder, W., and Böck, A. (1989). Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. *Nature* *342*, 453–456.

- Forchhammer, K., Leinfelder, W., Boesmiller, K., Veprek, B., and Böck, A. (1991). Selenocysteine synthase from *Escherichia coli*. Nucleotide sequence of the gene (*selA*) and purification of the protein. *J. Biol. Chem.* *266*, 6318–6323.
- Ghisla, S., and Massey, V. (1989). Mechanisms of flavoprotein-catalyzed reactions. *Eur. J. Biochem.* *181*, 1–17.
- Glass, R.S., Singh, W.P., Jung, W., Veres, Z., Scholz, T.D., and Stadtman, T.C. (1993). Monoselenophosphate: synthesis, characterization, and identity with the prokaryotic biological selenium donor, compound SePX. *Biochemistry* *32*, 12555–12559.
- Heider, J., Baron, C., and Böck, A. (1992). Coding from a distance: dissection of the mRNA determinants required for the incorporation of selenocysteine into protein. *EMBO J.* *11*, 3759–3766.
- Hüttenhofer, A., Westhof, E., and Böck, A. (1996). Solution structure of mRNA hairpins promoting selenocysteine incorporation in *Escherichia coli* and their base-specific interaction with special elongation factor SelB. *RNA* *2*, 354–366.
- Klug, S.J., Hüttenhofer, A., Kromayer, M., and Famulok, M. (1997). In vitro and in vivo characterization of novel mRNA motifs that bind special elongation factor SelB. *Proc. Natl. Acad. Sci. U. S. A.* *94*, 6676–6681.
- Kromayer, M., Wilting, R., Tormay, P., and Böck, A. (1996). Domain structure of the prokaryotic selenocysteine-specific elongation factor SelB. *J. Mol. Biol.* *262*, 413–420.
- Leinfelder, W., Zehelein, E., Mandrand-Berthelot, M.A., and Böck, A. (1988). Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* *331*, 723–725.
- Liu, Z., Reches, M., Groisman, I., and Engelberg-Kulka, H. (1998). The nature of the minimal “selenocysteine insertion sequence” (SECIS) in *Escherichia coli*. *Nucleic Acids Res.* *26*, 896–902.
- Lobanov, A. V., Hatfield, D.L., and Gladyshev, V.N. (2009). Eukaryotic selenoproteins and selenoproteomes. *Biochim. Biophys. Acta* *1790*, 1424–1428.
- Rengby, O., Johansson, L., Carlson, L.A., Serini, E., Vlamis-Gardikas, A., Kårsnäs, P., and Arnér, E.S.J. (2004). Assessment of production conditions for efficient use of *Escherichia coli* in high-yield heterologous recombinant selenoprotein synthesis. *Appl. Environ. Microbiol.* *70*, 5159–5167.

Schaefer, S.A., Dong, M., Rubenstein, R.P., Wilkie, W.A., Bahnson, B.J., Thorpe, C., and Rozovsky, S. (2013). (77)Se enrichment of proteins expands the biological NMR toolbox. *J. Mol. Biol.* 425, 222–231.

Turanov, A.A., Xu, X.-M., Carlson, B.A., Yoo, M.-H., Gladyshev, V.N., and Hatfield, D.L. (2011). Biosynthesis of selenocysteine, the 21st amino acid in the genetic code, and a novel pathway for cysteine biosynthesis. *Adv. Nutr.* 2, 122–128.

Veres, Z., Tsai, L., Scholz, T.D., Politino, M., Balaban, R.S., and Stadtman, T.C. (1992). Synthesis of 5-methylaminomethyl-2-selenouridine in tRNAs: 31P NMR studies show the labile selenium donor synthesized by the selD gene product contains selenium bonded to phosphorus. *Proc. Natl. Acad. Sci. U. S. A.* 89, 2975–2979.

Wagner, M.A., Trickey, P., Chen, Z.W., Mathews, F.S., and Jorns, M.S. (2000). Monomeric sarcosine oxidase: 1. Flavin reactivity and active site binding determinants. *Biochemistry* 39, 8813–8824.

Waugh, D.S. (2005). Making the most of affinity tags. *Trends Biotechnol.* 23, 316–320.

Xu, X.-M., Carlson, B.A., Mix, H., Zhang, Y., Saira, K., Glass, R.S., Berry, M.J., Gladyshev, V.N., and Hatfield, D.L. (2007). Biosynthesis of selenocysteine on its tRNA in eukaryotes. *PLoS Biol.* 5, e4.

Zinoni, F., Heider, J., and Böck, A. (1990). Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. *Proc. Natl. Acad. Sci. U. S. A.* 87, 4660–4664.

## SUPPLEMENTAL FIGURES

```

wtsfALR      MGGSHHHHHHGMASMRTOQKRDTKFREDCPPDREELGRHSWAVLHTLAAYYDDLPTPEQQ
SecALR1     -----MRTQKRDTKFREDCPPDREELGRHSWAVLHTLAAYYDDLPTPEQQ
SecALR2     -----MRTQKRDTKFREDCPPDREELGRHSWAVLHTLAAYYDDLPTPEQQ

wtsfALR      QDMAQFIHLFSKFYPCEECAEDLRKRLCRNHDPDRTRACFTQWLCHLNNEVNRKLGKPDF
SecALR1     QDMAQFIHLFSKFYPCEEUAEDLVRRLDRNHDPDRTRASFTQWLCHLNNEVNRKLGKPDF
SecALR2     QDMAQFIHLFSKFYPCEEUAEDLVARLHRNHDPDRTRASFTQWLCHLNNEVNRKLGKPDF

wtsfALR      DCSKVDERWRDGGKDGSCD-----
SecALR1     DCSKVDERWRDGGKDGSCDHHHHHH
SecALR2     DCSKVDERWRDGGKDGSCDHHHHHH

```

**Figure 6.7 Multiple sequence alignment of SecALR1, SecALR2 and wild type sfALR.** SecALR1 and SecALR2 have a C-terminal histidine tags (pJexpress414 vector) and wild type sfALR with N-terminal histidine tag (pTrcHisA vector). The native CEEC motif was replaced by a CEEU motif. The amino acids shown in red are the result of mutations at the nucleotide level to allow the proper Watson-Crick base pairings for the SECIS element (see Figure 6.1). Sequences for SecALR1 and SecALR2 have been codon optimized for expression in *E. coli* by DNA 2.0 (see Materials and Methods).

```

wtsfALR      ATGGGGGGTTCTCATCATCATCATCATCATGGTATGGCTAGCATGCGGACGCAGCAGAAG
SecALR1      -----ATGCGCACCCAACAAAAA
SecALR2      -----ATGCGCACCCAACAAAAA
                                     ***** ** *.**.*.

wtsfALR      CGGGACACCAAGTTTAGGGAGGACTGCCCGCCGGATCGCGAGGAAGCTGGGCCGCCACAGC
SecALR1      CGTGACACCAAGTTCGGTGAAGATTGCCCGCCAGATCGTGAAGAGCTGGGTCCGCATAGC
SecALR2      CGTGACACCAAGTTCGGTGAAGATTGCCCGCCAGATCGTGAAGAGCTGGGTCCGCATAGC
** ***** . * **.* ** *****.****** **.*.****** ** **

wtsfALR      TGGGCTGTCCCTCCACACCTTGGCCGCTACTACCCGACCTGCCACCCAGAACAGCAG
SecALR1      TGGGCAGTTCTGCACACTTTGGCCGCGTACTACCCGGATCTGCCGACGCCGGAGCAACAG
SecALR2      TGGGCAGTTCTGCACACTTTGGCCGCGTACTACCCGGATCTGCCGACGCCGGAGCAACAG
*****:* ** ** ***** ***** ***** ** ***** ** **.*.*.*.

wtsfALR      CAAGACATGGCCAGTTCATACATTTATTTCTAAGTTTACCCCTGTGAGGAGTGTGCT
SecALR1      CAGGACATGGCGCAGTTCATCCACCTGTTTAGCAAATCTATCCGTGTGAGGAGTGAGCT
SecALR2      CAGGACATGGCGCAGTTCATCCACCTGTTTAGCAAATCTATCCGTGTGAGGAGTGAGCT
**.****** *****.* ** *.***: **.* ** ** *******.*

wtsfALR      GAAGACCTAAGAAAAAGGCTGTGCAGGAACCCAGACACCCGCACCCGGGCATGCTTC
SecALR1      GAAGACCTGGTTCGACGCTCTCGACCAGAACCCCTGATACGCGTACCCGTGCGAGCTTT
SecALR2      GAAGACCTGGTTCGACGCTCTCGACCAGAACCCCTGATACGCGTACCCGTGCGAGCTTT
*****.*. : *.** ** .*.*****.* ** ** ** ***** **.*

wtsfALR      ACACAGTGGCTGTGCCACCTGCACAATGAAGTGAACCGCAAGCTGGGCAAACCGTACTTC
SecALR1      ACCCAGTGGCTGTGTCATCTGCACAACGAGGTGAATCGCAAGCTGGGCAAACCGTACTTC
SecALR2      ACCCAGTGGCTGTGTCATCTGCACAACGAGGTGAATCGCAAGCTGGGCAAACCGTACTTC
**.****** ** ***** **.****** *******.*

wtsfALR      GACTGCTCAAAAGTGGATGAGCGCTGGCGCGACGGCTGGAAGGATGGCTCCTGTGACTAG
SecALR1      GATTGCAGCAAAGTCGATGAACGTTGGCGTGACGGCTGGAAGGACGGTTCCTGCGATCAC
SecALR2      GATTGCAGCAAAGTCGATGAACGTTGGCGTGACGGCTGGAAGGACGGTTCCTGCGATCAC
** ***: .***** *****.* ** ***** *******.*

wtsfALR      -----
SecALR1      CACCATCATCATCAC
SecALR2      CACCATCATCATCAC

```

**Figure 6.8 Nucleotide multiple sequence alignment of wild type sfALR, SecALR1 and SecALR2.** The wild type ALR sequence is not *E. coli* optimized, unlike SecALR1 and SecALR2 (see Materials and Methods). The red nucleotides represent mutations made to generate the SECIS stem-loop feature.

## Chapter 7

# THE STUDY OF C142 MUTANTS OF AUGMENTER OF LIVER REGENERATION

### 7.1 Introduction

In several places in this dissertation we have reported the presence of charge-transfer complexes between cysteine 145 and the oxidized flavin of sfALR. For example, Chapter 4 reports the rapid formation of a charge-transfer intermediate with dithiothreitol and a variety of monothiols, including  $\beta$ -mercaptoethanol ( $\beta$ ME). Additionally, Chapters 5 and 6 demonstrate the formation of this intermediate species when selenium is proximal to the FAD cofactor. As our work progressed we were interested in studying the properties of the C142S and C142A mutations to generate species in which the charge-transfer complex might be fully formed. We also wished to determine the pK of the charge transfer thiolate, because such species would be expected to have anomalously low pK values (down to a pK values of 3.7 (Arcscott, Thorpe, & Williams, 1981)). However, during the course of this work Tokatlidis and coworkers reported a pK for the C142S mutant of 5.95 and further presented a 1.8 Å structure of this orange form of the protein (Banci et al., 2012). We too had succeeded in obtaining a 1.6 Å structure of the same mutant and so we briefly present an analysis of this high-resolution structure in this final Chapter. During this work we were able to crystallize the C142A mutant protein, but we were surprised at the lability of the charge-transfer complex under the crystallization conditions. While the initial solution

was a distinct orange color, reflecting the long-wavelength component of the charge-transfer complex, the protein turned yellow after several hours in the crystallization medium. The resulting crystal structure showed that C145 had undergone oxidation to the sulfinic acid state resulting in a loss of the charge-transfer complex. This work rationalizes the puzzling observation that the C142A mutant is isolated as an orange protein when *E. coli* is grown in a rich medium but shows no charge-transfer absorbance when the recombinant protein is isolated from cells cultivated in minimal medium. The marked difference in stability towards oxidation between C142A and C142S thiolate complexes of sfALR is attributed, in part, to the formation of a hydrogen bond between the hydroxyl group of S142 and the thiolate of C145.

## **7.2 Material and Methods**

### **7.2.1 Chemicals And Reagents**

Chemicals and reagents used in this study were as previously reported (Daithankar, Schaefer, Dong, Bahnson, & Thorpe, 2010; Schaefer et al., 2013).

### **7.2.2 Mutagenesis**

Primers for the mutation of C142A: 5'- CTA AGT TTT ACC CCG CTG AGG AGT GTG CTG -3' and 5'- CAG CAC ACT CCT CAG CGG GGT AAA ACT TAG -3' and primers for C142S: 5'- CTA AGT TTT ACC CCA GCG AGG AGT GTG CTG -3' and 5'- CAG CAC ACT CCT CGC TGG GGT AAA ACT TAG -3'.

Mutagenesis was performed as previously described (Daithankar, Farrell, & Thorpe, 2009; Farrell & Thorpe, 2005) and mutations were confirmed using Genewiz Inc. All proteins contain the previously described C154A and C165A mutations (Daithankar et al., 2009; Farrell & Thorpe, 2005).

### **7.2.3 Expression and Purification of C142 Mutants**

Expression of wild type short-form ALR, C142A and C142S mutants were as described previously (Daithankar et al., 2010). Cell paste was lysed in the presence of 1 mM PMSF, 0.1 mg/mL lysozyme and 1  $\mu$ M leupeptin by two passes through a French press (10000 psi) followed by brief sonication. The lysed cells were then centrifuged for 1 hr at 15000 g at 4 °C and the supernatant was flowed over a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated with 50 mM potassium phosphate, 500 mM NaCl, pH 7.5. The column was washed with several column volumes of binding buffer followed by 5 mL elution fractions of binding buffer supplemented with 50 mM, 200 mM, 500 mM and 1 M imidazole. The fractions eluted with the 200 mM to 1 M imidazole were then pooled and dialyzed against 50 mM potassium phosphate, 1 mM EDTA, pH 7.5.

The protein used for crystallography underwent cleavage of the histidine tag with TEV protease as outlined previously (Schaefer et al., 2013). The minimal media used in this study was as previously reported (Schaefer et al., 2013).

### **7.2.4 Crystallization and Data Collection**

Crystals were obtained using hanging-drop vapor diffusion. 1  $\mu$ L of the protein (~10 mg/mL in 20 mM Mes buffer, pH 6.5) was mixed with 1  $\mu$ L of the well solution containing 20 mM Mes pH 6.5 and 18% (w/v) PEG 8000. Crystals were dipped in a 20% xylitol mixture that was used as a cryoprotectant before freezing in liquid nitrogen. Diffraction data were collected using in-house equipment (Rigaku RUH3R and R-AXIS IV). The distance from the crystal to the detector was 100 mm and the Cu radiation X-ray wavelength was 1.54 Å. Data were collected from a single crystal for 10/15 min per 1° oscillation at -180 °C yielding 180 diffraction images.

The program HKL2000 was used to index, integrate and scale the data (Otwinowski Z, 1997).

### **7.2.5 Structure Determination and Refinement**

The structures for the C142S and C142A were solved by Dr. Brian Bahnson and Dr. Ming Dong at the University of Delaware by molecular replacement using the wild type human short-form ALR (PDB 3MBG). Molecular replacement was carried out using the program MOLREP from the CCP4 suite of programs (Bailey, 1994). 30 cycles of refinements were completed with the programs REFMAC5 and COOT (Emsley & Cowtan, 2004). Water molecules were placed during successive cycles of model building and refinement. The final models of C142S and C142A (residues 81-205 in addition to S80 which is not part of the ALR protein sequence but remains after the TEV cleavage) had one subunit comprising half of one covalent dimer. A final  $2F_o - F_c$  electron density difference map confirmed the quality of the final model. The final  $R_{\text{working}}$  and  $R_{\text{free}}$  values were 0.182 and 0.226 for C142S and 0.198 and 0.266 for C142A, respectively.

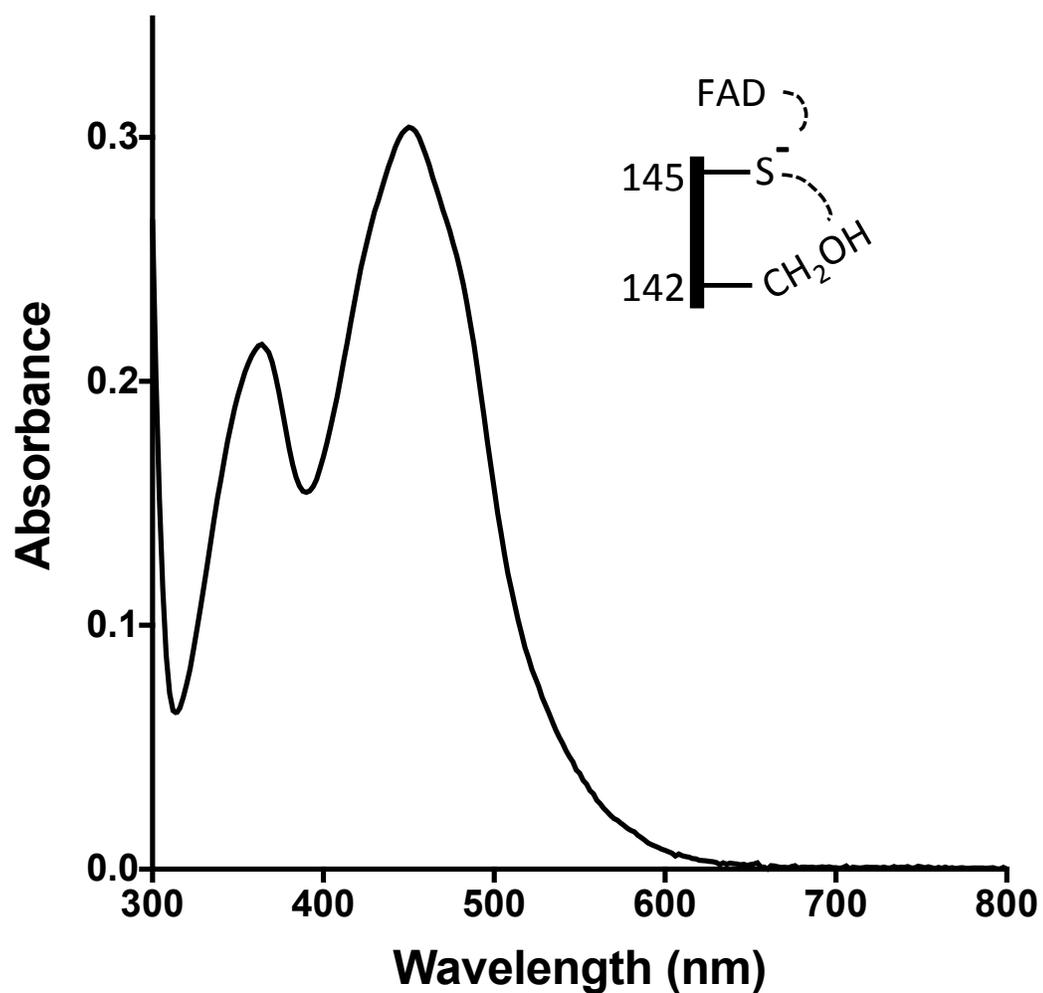
### **7.2.6 Structural Analyses**

Crystal structures of sfALR C142A and C142S have been deposited into the Protein Data Bank (PDB) with the codes 4LDK and 3TK0, respectively.

## **7.3 Results and Discussion**

The recombinant sfALR C142S was purified as an orange protein from LB medium. Figure 7.1 shows the UV-Vis spectrum of this serine mutant of ALR. The long-wavelength feature (>600 nm) is attributed to the formation of a charge-transfer complex that is formed when the electron-rich thiolate (C145) interacts with the

electron-deficient FAD (Figure 7.1 inset). This serine mutant was readily crystallized using hanging drop vapor diffusion and the 1.61 Å structure was solved in collaboration with Dr. Brian Bahnson and Dr. Ming Dong at the University of Delaware and the coordinates were submitted to the Protein Data Bank in August of 2011 (PDB 3TK0). Data refinement and statistics for this sfALR mutant are reported in Table 7.1. During the development of this project, Banci et al. published a communication that reported the pK of C145 to be 5.95 together with a 1.8 Å structure of sfALR C142S (PDB 3U2L) (Banci et al., 2012). The two structures are very similar with an RMSD of 0.128 Å. Figure 7.2 illustrates the hydrogen-bonding network surrounding thiolate 145 for both serine structures. The distance from the hydroxyl group of S142 to the thiolate group at position 145 is 3.1 Å (the structure reported by Banci et al. has two conformations of S142 with distances of 2.8 and 4.2 Å to the thiolate group; panel B) and a distance of 3.2 Å (3.1 Å for Banci et al.) from the thiolate to the C4a position on the isoalloxazine ring. Both structures shown in Figure 7.2 have Y138 making a stabilizing interaction with an active site water. However, the unpublished structure from the Thorpe group (Figure 7.2A) shows an additional interaction between this active site water and the thiolate from C145 as well as stabilizing interactions to E144 and R98.

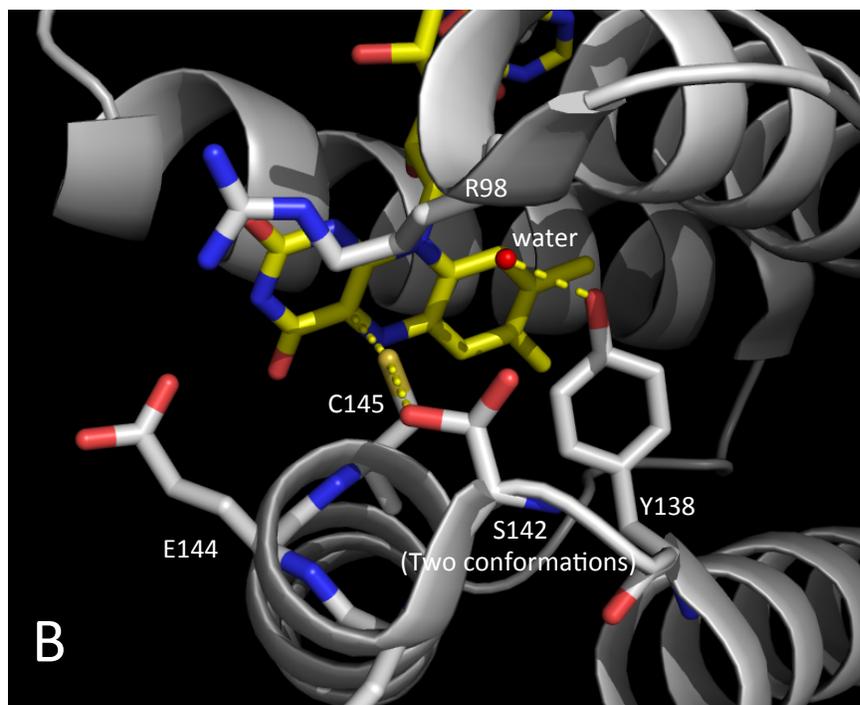
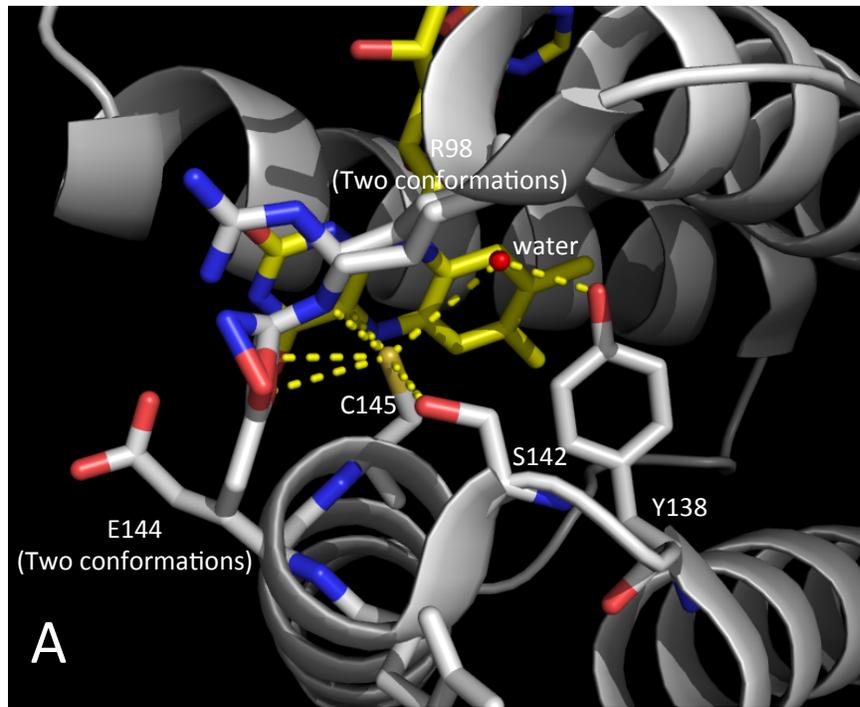


**Figure 7.1 UV-Vis spectrum of sfALR C142S.** A small charge-transfer band is seen at wavelengths  $>500$  nm. The spectrum was recorded in 50 mM potassium phosphate, pH 7.5, supplemented with 1 mM EDTA. The inset shows the thiolate at position 145 interacting with the proximal FAD cofactor as well as the hydroxyl group of S142.

**Table 7.1 Data collection and refinement statistics of human short-form ALR C142S<sup>a</sup>.**

Data Collection	
Space group	C2221
Unit cell dimensions	50.853, 76.968, 63.289
<i>a</i> , <i>b</i> , <i>c</i> (Å), $\alpha$ , $\beta$ , $\gamma$ (deg)	90.0, 89.973, 90.0
Resolution (Å)	50.0-1.6
Completeness (%)	97.3
Redundancy	4.7
I/ $\sigma$ I	48.32
R <sub>merge</sub> linear <sup>b</sup>	0.044
Refinement	
Resolution (Å)	31.645-1.611
R <sub>work</sub> /R <sub>free</sub> <sup>c</sup>	0.182/0.226
Mean <i>B</i> value	20.12
RMSD bond lengths (Å)	0.026
RMSD bond angles (deg)	2.337

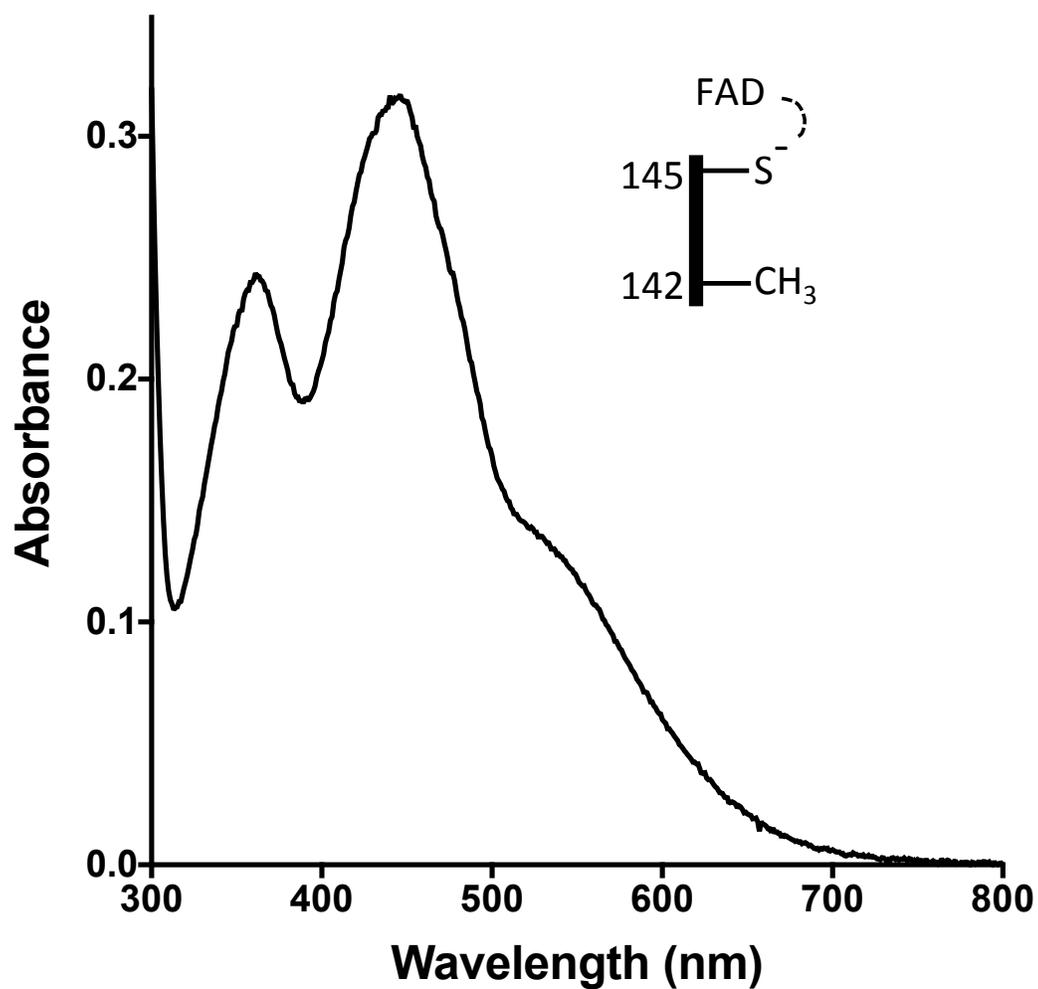
<sup>a</sup>Protein Data Bank accession code 3TK0 <sup>b</sup> $R_{\text{merge}} = \sum |I_o - I_a| / \sum (I_a)$ , where  $I_o$  is the observed intensity and  $I_a$  is the average intensity, the sums being taken over all symmetry-related reflections. <sup>c</sup> $R_{\text{working}} = \sum |F_o - F_c| / \sum (F_a)$ , where  $F_o$  is the observed amplitude and  $F_c$  is the calculated amplitude.  $R_{\text{free}}$  is the equivalent of  $R_{\text{working}}$ , except it is calculated for a randomly chosen set of reflections that were omitted (5%) from the refinement process.



**Figure 7.2 Hydrogen bond network for two crystal structures of sfALR C142S.**

Panel A shows the 1.61 Å crystal structure of sfALR C142S (PDB 3TK0) submitted by Bahnson and colleagues (unpublished). Both R98 and E144 have two conformations where one conformation hydrogen bonds to the thiolate of C145. Panel B illustrates the hydrogen bond network of the 1.8 Å crystal structure of sfALR C142S from Banci et al. Here S142 adopts two conformations where one conformation is able to make a stabilizing interaction with C145. Both structures have active site waters which are stabilized by Y138. However, in our structure (panel A) the water (red dot) also interacts with the C145 thiolate.

To continue our study of charge-transfer complexes in ALR we made a second mutation at position 142 to an alanine. This sfALR C142A construct exhibited a more substantial long-wavelength charge transfer feature shown in Figure 7.3. Alanine at position 142 results in a more intense charge-transfer complex when compared to serine at the same position (Figure 7.1) presumably because of the ability of the serine to H-bond to the thiolate (Figure 7.2) thereby decreasing its ability to act as a charge transfer donor. Comparable spectral effects have also been observed between alanine and serine mutations in human QSOX1 (Heckler, Alon, Fass, & Thorpe, 2008).



**Figure 7.3 UV-Vis spectrum of sfALR C142A.** A substantial charge-transfer complex is formed with maximal absorption around 520 nm. The spectrum of this mutant protein was measured in 50 mM potassium phosphate, 1 mM EDTA, pH 7.5. The inset is a schematic of the thiolate to FAD charge-transfer.

We were successful in obtaining protein crystals of the sfALR C142A that diffracted to 2.04 Å resolution using the same crystallographic conditions as those reported for the C142S protein (Table 7.2). Even though this protein was orange to begin with (due to the prominent charge-transfer complex; Figure 7.3), upon mixing with the reservoir solution the charge-transfer complex disappeared and yellow crystals appeared. Again, in collaboration with Dr. Brian Bahnson and Dr. Ming Dong, the crystal structure of this mutant enzyme was solved. However, upon close examination of the active site residues, extra electron density was observed surrounding the sulfur atom of C145. Figure 7.4 illustrates that C145 had undergone oxidation to a sulfinic acid which explains the loss of the charge-transfer complex. The two oxygen atoms covalently attached to the sulfur of C145 are able to simultaneously hydrogen bond to the water molecule in the active site and to R98. However, in this C142A structure, (unlike the C142S) residue 142 is unable to participate in the stabilization of C145.

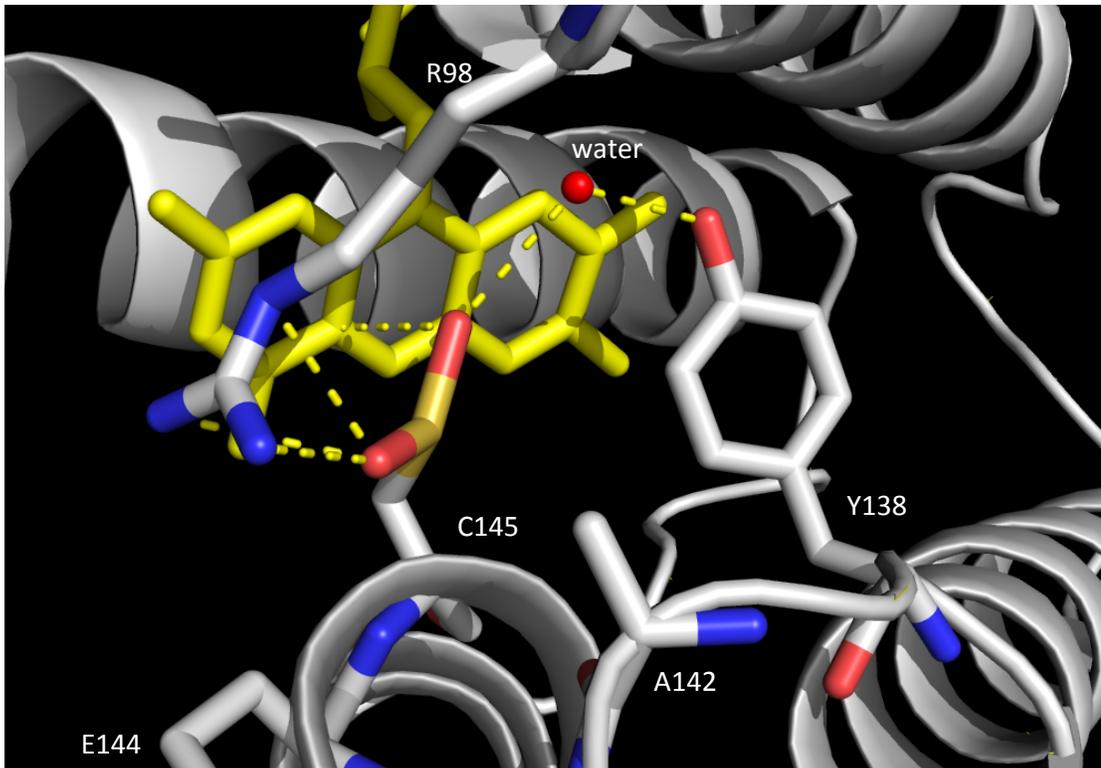
There are currently 103 cysteine sulfinic acid (abbreviated CSW) modifications in the PDB with 26 of those entries for human proteins. An example of the physiological importance of sulfinic acids is the role of DJ-1 as an oxidative stress response protein that plays a role in Parkinson's disease (Canet-Avilés et al., 2004). We do not think that the ALR's C145 sulfinic acid plays a physiological role, but we do predict that the polyethyleneglycol (PEG), used as a precipitant in our crystallography conditions, was a source of peroxides and caused the oxidation of C145 (Kumar & Kalonia, 2006; Ray & Puvathingal, 1985). In addition we expressed both the sfALR C142S and C142A mutants in a minimal medium and observed the C142S construct was isolated as an orange protein while the C142A construct was

yellow (Figure 7.5). These observations suggest that minimal media seemingly induces oxidative stress in *E. coli* which leads to the conversion of C145 to a sulfinate in the context of the C142A mutation. Strikingly a serine at position 145 seems to partially protect C145 from this oxidation by stabilizing the thiolate. In the future we hope to explore a range of additional C142 mutations for their stability during expression and purification as well as investigating selected mutants by UV-Vis spectroscopy and X-ray crystallography.

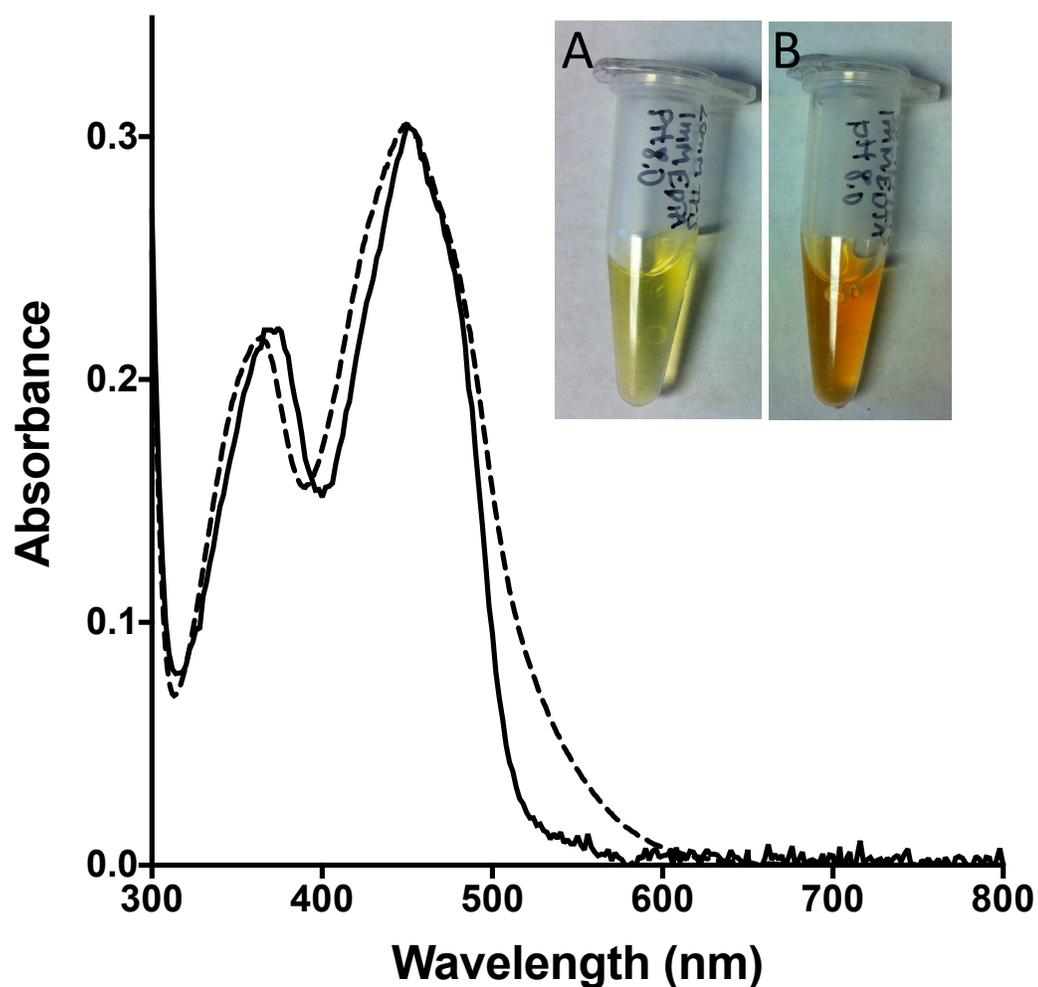
**Table 7.2 Data collection and refinement statistics of human short-form ALR C142A<sup>a</sup>.**

Data Collection	
Space group	C2221
Unit cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å), $\alpha$ , $\beta$ , $\gamma$ (deg)	50.897, 77.77, 62.772 90.0, 89.973, 90.0
Resolution (Å)	50.0-2.04
Completeness (%)	99.9
Redundancy	6.8
I/ $\sigma$ I	18.56
R <sub>merge</sub> linear <sup>b</sup>	0.106
Refinement	
Resolution (Å)	23.58-2.04
R <sub>work</sub> /R <sub>free</sub> <sup>c</sup>	0.198/0.266
Mean <i>B</i> value	28.592
RMSD bond lengths (Å)	0.02
RMSD bond angles (deg)	1.981

<sup>a</sup>Protein Data Bank accession code 4LDK <sup>b</sup>R<sub>merge</sub> =  $\sum |I_o - I_a| / \sum (I_a)$ , where *I<sub>o</sub>* is the observed intensity and *I<sub>a</sub>* is the average intensity, the sums being taken over all symmetry-related reflections. <sup>c</sup>R<sub>working</sub> =  $\sum |F_o - F_c| / \sum (F_a)$ , where *F<sub>o</sub>* is the observed amplitude and *F<sub>c</sub>* is the calculated amplitude. R<sub>free</sub> is the equivalent of R<sub>working</sub>, except it is calculated for a randomly chosen set of reflections that were omitted (5%) from the refinement process.



**Figure 7.4 Hydrogen bond network of sfALR C142A.** The 2.04 Å crystal structure of sfALR C142A (PDB 4LDK) submitted by Bahnson and colleagues (unpublished) shows the sulfinic form of cysteine 145. Consistent with the C142S structure (Figure 7.2A), R98 and an active site water (red dot) participate in hydrogen bond interactions with residue 145. However, in this structure there are two oxygen atoms covalently attached to the sulfur at this position.



**Figure 7.5 Short-form ALR C142A and C142S grown in minimal media.** UV-Vis spectra of sfALR C142A (black line) and C142S (dashed line). A charge-transfer absorbance can be seen in the C142S spectrum. The photographs in the inset represent the purified C142A and C142S mutants grown in minimal medium in panels A and B, respectively. The spectra were recorded in 20 mM Tris, 1 mM EDTA pH 8.0, and are normalized at 456 nm for ready comparison.

## REFERENCES

- Arscott, L. D., Thorpe, C., & Williams, C. H. (1981). Glutathione reductase from yeast. Differential reactivity of the nascent thiols in two-electron reduced enzyme and properties of a monoalkylated derivative. *Biochemistry*, *20*(6), 1513–20. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7013796>
- Bailey, S. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr*, *50*(Pt 5), 760–763. doi:10.1107/S0907444994003112 S0907444994003112 [pii]
- Banci, L., Bertini, I., Calderone, V., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., & Tokatlidis, K. (2012). An electron-transfer path through an extended disulfide relay system: the case of the redox protein ALR. *Journal of the American Chemical Society*, *134*(3), 1442–5. doi:10.1021/ja209881f
- Canet-Avilés, R. M., Wilson, M. A., Miller, D. W., Ahmad, R., McLendon, C., Bandyopadhyay, S., ... Cookson, M. R. (2004). The Parkinson's disease protein DJ-1 is neuroprotective due to cysteine-sulfinic acid-driven mitochondrial localization. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(24), 9103–8. doi:10.1073/pnas.0402959101
- Daithankar, V., Farrell, S., & Thorpe, C. (2009). Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space. *Biochemistry*, *48*(22), 4828–4837. doi:10.1021/bi900347v
- Daithankar, V., Schaefer, S. A., Dong, M., Bahnson, B. J., & Thorpe, C. (2010). Structure of the human sulfhydryl oxidase augmenter of liver regeneration and characterization of a human mutation causing an autosomal recessive myopathy. *Biochemistry*, *49*(31), 6737–6745. doi:10.1021/bi100912m
- Emsley, P., & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr*, *60*(Pt 12 Pt 1), 2126–2132. doi:S0907444904019158 [pii] 10.1107/S0907444904019158

- Farrell, S. R., & Thorpe, C. (2005). Augmenter of liver regeneration: a flavin-dependent sulfhydryl oxidase with cytochrome c reductase activity. *Biochemistry*, 44(5), 1532–1541. doi:10.1021/bi0479555
- Heckler, E. J., Alon, A., Fass, D., & Thorpe, C. (2008). Human quiescin-sulfhydryl oxidase, QSOX1: probing internal redox steps by mutagenesis. *Biochemistry*, 47(17), 4955–4963. doi:10.1021/bi702522q
- Kumar, V., & Kalonia, D. S. (2006). Removal of peroxides in polyethylene glycols by vacuum drying: implications in the stability of biotech and pharmaceutical formulations. *AAPS PharmSciTech*, 7(3), 62. doi:10.1208/pt070362
- Otwinowski Z, M. W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology*, 276, 307–326.
- Ray, W. J., & Puvathingal, J. M. (1985). A simple procedure for removing contaminating aldehydes and peroxides from aqueous solutions of polyethylene glycols and of nonionic detergents that are based on the polyoxyethylene linkage. *Analytical biochemistry*, 146(2), 307–12. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/4025798>
- Schaefer, S. A., Dong, M., Rubenstein, R. P., Wilkie, W. A., Bahnson, B. J., Thorpe, C., & Rozovsky, S. (2013). (77)Se enrichment of proteins expands the biological NMR toolbox. *Journal of molecular biology*, 425(2), 222–31. doi:10.1016/j.jmb.2012.11.011

## PERMISSION LETTERS

**Ballen, Karen** <[KBallen@liebertpub.com](mailto:KBallen@liebertpub.com)>   
To: Stephanie Schaefer <[saschaef@UDel.Edu](mailto:saschaef@UDel.Edu)>  
RE: use of figure in my thesis

September 24, 2013 3:09 PM  
[Hide Details](#)

1 Attachment, 452 bytes

Dear Stephanie:  
Copyright permission is granted for use of this figure in your thesis.  
Kind regards,  
Karen Ballen  
Manager, Copyright Permissions

---

**From:** Stephanie Schaefer [<mailto:saschaef@udel.edu>]  
**Sent:** Tuesday, September 24, 2013 9:59 AM  
**To:** Ballen, Karen  
**Subject:** use of figure in my thesis

Hello,  
I am looking to get permission to use one figure (from a recent paper published in my group) for my thesis.  
[Oxidative Protein Folding and the Quiescin-Sulfhydryl Oxidase Family of Flavoproteins](#)   
Vamsi K. Kodali, Colin Thorpe  
Antioxidants & Redox Signaling. October 15, 2010, 13(8): 1217–1230.

However when I fill out the information for reprints and permissions it says I have to pay over \$100. Is this correct? There is no option to request permission for a thesis so I selected book.

Thank you for your help,  
Stephanie Ramadan



**Title:** Structure of the Human Sulphydryl Oxidase Augmenter of Liver Regeneration and Characterization of a Human Mutation Causing an Autosomal Recessive Myopathy,  
**Author:** Vidyadhar N. Daithankar, Stephanie A. Schaefer, Ming Dong, Brian J. Bahnson, and Colin Thorpe  
**Publication:** Biochemistry  
**Publisher:** American Chemical Society  
**Date:** Aug 1, 2010

Copyright © 2010, American Chemical Society

User ID
<input type="text"/>
Password
<input type="text"/>
<input type="checkbox"/> Enable Auto Login
<input type="button" value="LOGIN"/>
<a href="#">Forgot Password/User ID?</a>

If you're a **copyright.com** user, you can login to RightsLink using your copyright.com credentials. Already a **RightsLink** user or want to [learn more?](#)

## PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

If credit is given to another source for the material you requested, permission must be obtained from that source.

[BACK](#)[CLOSE WINDOW](#)



**Title:** Se Enrichment of Proteins Expands the Biological NMR Toolbox

**Author:** Stephanie A. Schaefer, Ming Dong, Renee P. Rubenstein, Wayne A. Wilkie, Brian J. Bahnson, Colin Thorpe, Sharon Rozovsky

**Publication:** Journal of Molecular Biology

**Publisher:** Elsevier

**Date:** Jan 23, 2013

Logged in as:  
Stephanie Ramadan  
Account #:  
3000706785

[LOGOUT](#)

Copyright © 2013, Elsevier

### Order Completed

Thank you very much for your order.

This is a License Agreement between Stephanie Ramadan ("You") and Elsevier ("Elsevier") The license consists of your order details, the terms and conditions provided by Elsevier, and the [payment terms and conditions](#).

License number	Reference confirmation email for license number
License date	Oct 12, 2013
Licensed content publisher	Elsevier
Licensed content publication	Journal of Molecular Biology
Licensed content title	Se Enrichment of Proteins Expands the Biological NMR Toolbox
Licensed content author	Stephanie A. Schaefer, Ming Dong, Renee P. Rubenstein, Wayne A. Wilkie, Brian J. Bahnson, Colin Thorpe, Sharon Rozovsky
Licensed content date	23 January 2013
Licensed content volume number	425
Licensed content issue number	2
Number of pages	10
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	All
Actual number of figures/tables/illustrations	12
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Understanding the flavoenzyme human augmenter of liver regeneration: biochemical and structural perspectives
Expected completion date	Nov 2013
Elsevier VAT number	GB 494 6272 12
Billing Type	Invoice
Billing address	204 Drake Hall University of Delaware