BIOPHYSICAL ANALYSIS OF PROTEINS THAT CONTRIBUTE TO LIVER DAMAGE AND LIVER REGENERATION

by

Quentin Jerod Tekel Florence

(Under the Direction of John Patrick Rose)

ABSTRACT

The liver is a vital organ and is unique in that it is capable of regeneration. The work presented herein is focused on the biophysical and structural characterization of the hepatitis B virus surface protein HBsAg (the current hepatitis b vaccine) and the protein Augmenter of Liver Regeneration (a growth factor isolated from weanling or regenerating liver).

Hepatitis b is a blood-born-associated virus that has infected over 2 billion people worldwide and can lead to liver cancer. Thus, one of the overall goals of this study is to provide biophysical and structural information for the hepatitis b vaccine antigen HBsAg. The lack of an HBsAg structure has prevented us from understanding the effects of mutations of the viral polymerase, which can lead to vaccine escape and has impeded the development of bivalent vaccines. The goal of this work is to provide biochemical and structural information about HBsAg that can be used in the design of new and better vaccines.

Liver damage and regeneration have been the focus of research studies for decades. From these studies a growth factor called Augmenter of Liver Regeneration (ALR) was identified that augmented the regeneration process. ALR was then shown to be homologous the yeast *Saccharomyces cerevisiae* protein Erv1 (Essential for Growth

and Respiration protein 1) which has been demonstrated to be involved in the mitochondrial intermembrane space import (IMS) pathway. The IMS pathway has been well characterized in yeast and shows that Erv1 interacts with mitochondrial intermembrane assembly protein 40 (Mia40) and cytochrome c forming a Mia40 – ERV1 – cytochrome c redox chain. The goal of these studies is focused on characterizing the mammalian equivalent of the Mia40 – ALR – cytochrome c redox chain using *Rattus norvegicus* as a model. These studies will include the recombinant cloning, expression, purification of rat Mia40 and ALR and the charaterization of interactions between the Mia40, ERV1 and cytochrome c proteins in solution.

Together these studies should provide a foundation for the development of new tools that can be used to prevent liver damage and/or promote liver repair.

INDEX WORDS: Augmenter of Liver Regeneration, Hepatitis B Surface Antigen, Liver Disease, Liver Regeneration, Mitochondrial Intermembrane Space Import and Assembly Pathway, Crystal Structure, Small Angle X-ray Scattering, Surface Plasmon Resonance

BIOPHYSICAL ANALYSIS OF PROTEINS THAT CONTRIBUTE TO LIVER DAMAGE AND LIVER REGENERATION

By

Quentin Jerod Tekel Florence

B.S., The University of Georgia, 2003

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

BIOPHYSICAL ANALYSIS OF PROTEINS THAT CONTRIBUTE TO LIVER DAMAGE AND LIVER REGENERATION

by

Quentin Jerod Tekel Florence

Major Professor: John Patrick Rose

Committee: Bi-Cheng Wang
William Lanzilotta

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2010

DEDICATION

For Dr. Natalie Florence

He who finds a wife finds what is good (Proverbs 18:22). Natalie there are not words that I can use that would justify how much you mean to me and how thankful I am for all the love you show me. The work described in this dissertation could not have been accomplished without you. Thank you for listening when I complained. Thank for you for raising my spirits when I was down. Thank you for loving me unconditionally. Most of all, thank you for allowing me to be your husband. Natalie you are truly the greatest blessing God has placed in my life. I know that wherever we go from here will truly be blessed. **I Love You**

For Victor F. Florence, Carnetta D. Florence, Rodriquez A. Florence, and Victoria J. Florence

For God said, 'Honor your father and mother' (Matthew 15:4) Mom and Dad, I remember when I was 12 and I told you I would become a doctor. You said go for it and you have supported me ever since. Mom and Dad you truly have shown me how to be a man and how to never stop pursuing my dreams. Thank you for all that you have done for me and the things that are yet to come. **I Love You** and I will continue to honor all of your teachings.

As iron sharpens iron, so one person sharpens another (Proverbs 27:17) I must say that I am truly blessed to have a brother and sister that love and care about me. Rod and Victoria, you all do not know how proud I am to have you as my brother and sister.

You all have accomplished so much in your young lives that it motivates me to keep pushing. Thank you for keeping me humble and giving me lots of laughter. I know we will keep sharpening each other throughout the rest of our lives. I pray that all of your dreams come true and **I Love You** both.

ACKNOWLEDGMENTS

The completion of a doctorate degree is the culmination of a period of your life that you will never forget. During my time in graduate school, I have been fortunate enough to work with some of the greatest scientists and professionals in the fields of biochemistry and crystallography. This journey would not have been possible without the aid of Dr. David Puett, who was the biochemistry department head while I was an undergraduate. I would like to thank Dr. Puett for writing the most significant letter of recommendation I have ever received. Your letter of recommendation convinced the graduate committee to give me an opportunity to pursue a graduate degree and for that I will be eternally grateful. In addition to the help I received from Dr. Puett, the letters and guidance I received from Dr. Bi-Cheng Wang and Dr. Peter Horanyi were invaluable in my pursuit of a graduate degree. Dr. Wang provided me with an opportunity to pursue my unique ideas and grow as scientist, which has allowed me to come to this point. Peter, there are not words to describe how much your help meant to me. You believed I could be a good scientist before I even thought about being a scientist. Dr. John Rose has been a great mentor and support system during my time in graduate school. Thank you Dr. Rose because I know I could not have accomplished this without you. To all other members of my lab and the biochemistry department, I wan to say thank you for all the laughs and support. I look forward to seeing everyone grow into wonderful professionals. To all other friends and relatives, thank you for all of your support and encouragement. There is no greater feeling than the one you get when you realize that you have people that care about you. Thanks for everything.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER 1

INTRODUCTION

Purpose of the Study

Liver damage and regeneration have been the focus of numerous research studies, that have lead to a wealth of knowledge about the diseases that cause liver damage and the key components in the liver regeneration process that aids the liver in recovery. However, there remain several aspects of liver disease and liver regeneration that needs further investigation. Among these are better understandings of the molecules that stimulate the liver regeneration and their role in the regeneration process and the development of new and better preventative methods, such as vaccines. With this in mind, the research described herein focuses on (1) the biophysical characterization of the hepatoprotein augmenter of liver regeneration (ALR) (*1, 2*) and its newly discovered role in the mitochondrial intermambrane space import and assembly (IMS) pathway (*3*) and (2) providing structural information for the hepatitis b surface antigen (HBsAg) protein needed for the design of new and better prophylactics.

ALR is a protein that has been shown to increase the rate of liver regeneration (*4*) ALR is considered an augmenter because although it cannot initiate or terminate the liver regeneration process it increases the rate of liver regeneration. Previous studies have shown that ALR is located in multiple organs in the body, can be expressed as both a 125 (short form) and 198 (long form) residue protein and functions as a sulfhydryl oxidase (*5, 6*).

Recent studies have indicated that the long form of ALR likely functions in reduction oxidation reactions that occur in the IMS pathway in mammals (*3*). The

IMS pathway from *S. cerevisiae* was originally identified using immunoprecipitation assays in which yeast Mia40 (mitochondrial intermembrane space assembly protein 40) and Erv1 (essential for respiration and viability protein) bound to Mia40 antibodies (*7*). In addition, cytochrome c was identified as another vital component of the IMS pathway in yeast (*3*).

It has been shown that ALR is the mammalian homolog of Erv1 (*8*) and sequence analyses have shown that Mia40 and cytochrome c homologs exist in all higher eukaryotic organisms. Thus, a major goal of the research presented is to (1) establish that the IMS pathway in rat and (2) to characterize the interaction between rat Mia40, ALR and cytochrome c using common biochemical/biophysical techniques.

Specifically, the central hypothesis of this research is that a pathway similar to the IMS pathway identified in yeast is also present in R. norvegicus. The results from this study will aide in understanding the significance of the IMS pathway in mammals and the mechanism as to how ALR augments the liver regeneration process. In addition, the IMS pathway has been shown to play a significant role in the maturation of several key translocases of the inner membrane (TIM). The lack of TIM function has been linked to the human deafness dystonia syndrome and cell death (*9, 10*). Thus, a better understanding the IMS pathway is a necessary step in developing methods for the prevention of this syndrome. Finally, since regeneration is a phenomenon that is only observed in the liver, the more insight that can be gained into how it works and how different factors affect it could lead to the novel therapeutic techniques.

The hepatitis b virus is a blood-born-associated virus that has infected over 2 billion people worldwide (*11*). The chronic form of hepatitis b can lead to liver

cancer (*11*). One of the overall goals of this study is to provide structural information for the HBsAg protein, the antigenic component of the currently licensed vaccine for hepatitis b virus. Despite its worldwide usage and its application as a carrier for epitope presentation, the tertiary structure of this important protein is still unknown. In addition, the lack of a HBsAg structure has (1) prevented us from understanding the effects of mutations on polymerase drug-resistance and on the antigenic properties of the HBsAg major neutralizing antigenic epitope and (2) impeded the development of bivalent vaccines where HBsAg is used as a carrier for foreign antigenic epitopes. The results of the proposed studies will provide information for aiding in the understanding of how various mutations and fusion of foreign epitopes may affect the tertiary structure of the HBsAg neutralizing epitope. In addition, the studies should provide sufficient understanding of molecular recognition between the immunogenic site ('a'-determinant region) of HBsAg and antibodies. Such knowledge is necessary for the development of improved HBV vaccines, bivalent vaccines and understanding the role of drug-resistant mutations in HBV immune escape.

The research study described within this document provides fundamental insight into the cellular function of ALR and the structure of HBsAg. This research provides initial insight into the binding kinetics and protein-protein complex structures of components the IMS pathway along with solution structural information on HBsAg. These studies revealed information that will aide in the study of liver damage and liver regeneration.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

The liver is the largest internal vital organ in a majority of mammalian organisms. Structurally, the liver consists mainly of two major lobes, left and right, separated by the falciform ligament (Figure 2.1). The two lobes are primarily composed of hepatocyte cells that perform several functions including: bile production, nutrient storage, nutrient interconversion, detoxification, phagocytosis, and protein synthesis (*12*). Bile production is a key function of the liver. Bile plays a major role in digestion by diluting and neutralizing stomach acids so that pancreatic enzymes can function properly. Bile also emulsifies fats so that they can be properly broken down (*12*). Liver hepatocytes can also remove glucose, fats, vitamins, copper, and iron from the blood and store them on a short-term basis. The interconversion of nutrients is another vital function of the liver because ingested material is not always in proportions needed by the tissues or the material may not be in a form that cells can use. An example of interconversion would be combination of ingested fats with phosphorous and choline to produce phospholipids that can be used in the cell membrane (*12*). Liver hepatocytes also aid the kidneys in detoxification by altering the structure of many substances like ammonia, which allows them to be eliminated from the body more easily. Kupffer cells, which lie along the endothelium of the liver, phagocytize worn out and dying red and white blood cells from the blood stream which circulates through the liver (*12*). In addition to these numerous functions, the liver also is the location for the synthesis of several key blood proteins such as albumins, fibrinogen, globulins, heparin, and clotting factors (*12*). The unique properties and diversity of functions that the liver performs make it an organ that has

been the subject of numerous research studies.

Similar to other organs, the majority of the research conducted on the liver has focused on investigating liver disease and liver damage. The term hepatitis refers an inflammation of the hepatocyte cells caused by a number of diseases and infections. Among the causative agents for hepatitis are the hepatotropic viruses, five unrelated viruses (hepatitis A, B, C, D and E) that lead to liver inflammation. Among these viruses, and a focus of my research, is a member of the hepadnavirus family, Hepatitis B virus (HBV). Hepatitis B is a blood-born-associated virus that has infected about 2 billion people worldwide (*11, 13*). Transmission of Hepatitis B is caused by exposure to infectious blood or body fluids. The most common causes of transmission are the reuse of contaminated needles, unprotected sexual activity and vertical transmission, (transmission from mother to child during childbirth) (*13-15*). Despite being discovered more than 30 years ago, hepatitis B remains a challenging global health problem.

During replication, the 42-nanometer Hepatitis B virion (Figure 2.2) binds to an unidentified receptor protein and enters the cell via endocytosis. Following endocytosis, fusion of the cellular and viral membranes is mediated by viral envelope proteins and this fusion results in the release of the viral capsid into the cytoplasm. The viral capsid, which contains the viral DNA and a nuclear localization signal, is transported toward the nucleus *via* microtubules (*16*). At the nuclear membrane, the viral capsid interacts with the adapter proteins importin alpha and beta and is translocated into a nuclear pore. Inside the nuclear pore, a nuclear basket is formed which facilitates the breakdown of the viral capsid and the release of partially double stranded DNA into the nucleus (*16*) where the partially double stranded viral DNA is converted to covalently closed circular DNA that

encodes for the four viral mRNA molecules (*13, 17, 18*). Host cells then synthesize the viral mRNAs, which are translated into the proteins that form new virions. The newly formed Hepatitis B virions are then released from the cell and the process is repeated. After the host cell has been infected, liver cell damage is caused by the host immune response system. The primary response elements utilized are cytotoxic T lymphocytes, which destroy infected host cells and produce antivirus cytokines that can breakdown virions that are in viable hepatocytes (*19*). Acute viral infection combined with the initial host immune response lead to a number of symptoms that include vomiting, liver inflammation, and eventually jaundice.

Approximately 90% of the people that develop acute Hepatitis B infections are able to fight off the virus without any treatment (*11*). However, an increase in the number of chronic Hepatitis B infections has been observed by the World Health Organization (*15, 20*). Because of the difficulty of treating chronic Hepatitis B infection, the primary treatment for the disease has been prevention using vaccines (*21*). The vaccines that have been developed rely on the use of the viral envelope protein called the hepatitis B surface antigen (HBsAg). HBsAg, produced in yeast cells, is isolated as a 226-residue soluble lipoprotein particle with a spherical radius of 20 nanometers, as shown in Figure 2.3. Sequence analysis (www.predictprotein.org) suggests that the protein has four trans-membrane helices and large inner and outer membrane loops (see Figure 2.4). The region spanning residues 100-165 on the outer membrane loop comprises the HBsAg antigenic region. HBsAg self-assembles into non-infectious particles (Figure 2.4) and once these particles (*22*) are introduced to the bloodstream, anti-HBsAg antibodies are produced as part of the immune response. Gilbert *et al.* revealed through cryo-electron

microscopy that these noninfectious HBsAg particles form octahedral structured particles containing 48 copies of HBsAg whose surface epitopes resemble that of an intact virus (*23*). The anti-HBsAg antibodies recognize the surface epitopes, which helps promote resistance to the disease. However, researchers have discovered HBsAg escape mutations, which cause vaccination resistance (*11, 20, 21, 24*). In addition, these mutations often occur in or near the open reading frame for the viral polymerase (the target for HBV polymerase inhibitor therapies), which causes these HBV variants to be drug resistant. The escape mutations likely cause a conformational change in the HBsAg protein in the region encoding for the HBV major conformation-dependent neutralizing epitope (*25*) resulting in vaccine resistance (see Figure 2.4). In spite of the fact that a low-resolution cryo-electron microscopy structure is available, understanding the effect of mutations on HBsAg conformation is limited by the resolution of the technique.

Thus, gaining additional structural knowledge of HBsAg is a necessary step towards understanding the effect of escape mutations on HBsAg conformation and drug resistance. In addition, the HBsAg structure would also provide a foundation for the development of new or better combination vaccines. Combination vaccines are a single-shot vaccine, which can provide immunity against multiple infections. Combination vaccines are preferred because they are more efficient and cost effective than traditional vaccines. In previous studies, a modified HBsAg carrying the immunodominant hypervariable region 1 (HVR1) of the hepatitis C virus envelope protein E2 within the exposed ‘a’-determinant region was shown to develop an immune response against the hepatitis C virus in mice (*25*). These results provided evidence that HBsAg has the potential to be engineered so that it can be used as a combination vaccine.

Although HBsAg medicines and other vaccines have proven to be an effective treatment for the prevention of liver diseases, a high percentage of patients recover from acute liver injuries without treatment due to the fact that the liver possesses the ability to regenerate itself. The mammalian liver is unique among organs in that it can completely restore its weight and architecture if it is injured or reduced to no more than 70% of its mass. Although the first description of liver regeneration can be attributed to the story of Prometheus from Greek mythology, the first observation of liver regeneration following partial organ removal (hepatectomy) was reported in 1931 (*26*). Since that discovery, more than a century ago, investigators have been trying to understand the unique and complex processes of liver regeneration and to identify the factors that are responsible for the liver's regenerative property. Liver regeneration is a complex process that consists of many stages (*27*). When a portion of the liver is damaged, hepatocyte metabolic activities increase and DNA replication starts, signaling the beginning of the priming stage of the regeneration process. The priming stage prepares the hepatocyte cells for regeneration by causing the cells to become competent and respond to growth factors (*27*). Research has shown that 185 genes are either induced or repressed during the priming process. Among the genes induced are those associated with the production of transcription factors, proteins that help to regulate cell proliferation and proteins that regulate the cell cycle (*28*). Two important cytokines, tumor necrosis factor alpha (TNFα) and interleukin 6 (IL-6), have been shown to be essential for the priming process. Produced by Kupffer cells, IL-6 binds to the gp130 receptor on the surface of hepatocyte cells. This binding activates Janus kinases and results in the activation of signal transducer and activator of transcription 3 (STAT3), which mediates the response of several genes during the

priming process (*29-31*). TNFα induces stress-activated protein kinases and augments the production of c-myc, which regulates cell proliferation, and IL-6. These functions allow TNFα to maintain specific gene expression during liver regeneration (*32-34*). Following the priming process, the hepatocyte cells can now respond to growth factors that induce them to begin the cell proliferation process. The main growth factors that are capable of directly inducing cell proliferation are transforming growth factor alpha (TGFα), epidermal growth factor (EGF), and hepatocyte growth factor (HGF). The growth factors accomplish this activation by binding to receptors with tyrosine kinase activity, which activate protein kinase cascades (*27*). High expression levels of TGFα have been shown to directly correlate with DNA synthesis in hepatocytes (*35*). The release of TGFα by hepatocytes is a signal for the beginning of the mitotic process in regeneration (*36, 37*). EGF and TGFα both interact with the EGF receptor, which allows them to stimulate both protein and DNA synthesis (*38*). HGF is considered to be the most important growth factor because it stimulates the synthesis of TGFα by binding to the protein kinase receptor c-Met. The binding of HGF to cMet leads to the activation of signaling pathways, which results in the proliferation of hepatocytes (*39, 40*). Even though the previously described major factors are now fairly well established, a large number of potentially important proteins and peptides remain unproven as *bona fide* growth factors for *in vivo* liver regeneration. This second tier of potential growth factors includes the protein named augmenter of liver regeneration (ALR) (*1*), another research focus of this dissertation. ALR is also known as hepatic stimulatory substance, (HSS) (*41*) and as hepatoprotein (HPO) (*42*).

In 1931, Higgins and Anderson first demonstrated liver regeneration after partial

hepatectomy (*26*). Studies of McJunkin and Breuhaus followed shortly thereafter which showed that the modest mitotic response to a limited (30% - 45%) hepatectomy was significantly enhanced by intraperitoneal post operative injection of homogenate made from weanling or regenerating rat livers (*43*). The homogenate injection was later found to be most effective when injected into weanling rats, which had naturally hyperplasic livers. The injections were also found to have no effect on the normal adult liver (*44, 45*). After a lengthy lapse in interest, LaBrecque and Pesch (*46*) described what they termed the "hepatic stimulatory substance" (HSS) found in the cytosol of weanling or regenerating rat livers. HSS did not affect the adult liver but it significantly augmented liver regeneration after partial hepatectomy. These findings have been extended to dog (*47*). In addition, HSS was shown to augment the heightened proliferation that is caused by portacaval shunt (Eck's fistula) (see Figure 2.5), and to prevent acute hepatocyte atrophy usually observed 4 days after surgery (*48*). This heat-stable cytosolic factor was renamed “augmenter of liver regeneration” (ALR) by Francavilla *et al.,* (1987) (*1*) who highly purified (830,000 fold) the suspect peptide from regenerating rat liver homogenate. Once identified, ALR was cloned and expressed in COS cells (*49*) and later identified, cloned and expressed from the mouse and human genomes (*50*). The chromosomal location (human 16p13.3-13.1) (*8*), genome structure (*8, 50*), tissue distribution of mRNA (*8, 49, 50*) and the 3-dimensional structure (*51*) of ALR are now known (Figure 2.6).

ALR is not restricted to the liver but is also found in high concentration (as indicated by Northern blots) in kidney and testes (*49*). This finding is interesting since 16p13 has been shown to contain genes responsible for some kidney diseases (*52*) and ALR mRNA

levels are high in kidney. ALR has been reported to reduce the activity of liver-resident natural killer (NK) cells (*4, 53*), which have been postulated to be involved in termination and initiation of liver regeneration (*54*). This effect on NK cells, which is also observed with HGF and insulin-like growth factor (IGF-II), occurs only *in vivo* and not *in vitro,* suggesting that this factor is only one component of a complex physiological system. In addition to its role in liver regeneration, there is now evidence that ALR may function *in vivo* as a more general trophic factor. Adams *et al.* (1998) (*55*) have published data demonstrating that ALR significantly enhances the success rate of fetal rodent pancreas transplantation. In this study it was found that ALR alone was approximately 100-times more potent on a per weight basis than IGF-I in the fetal rat pancreas model and that ALR plus IGF-I were not synergistic.

ALR has been shown to be involved in the regulation of mitochondrial gene expression through the stimulation of the production of the mitochondrial transcription factor mtTFA (*56*), involved in the export of iron/sulfur (Fe/S) clusters from the mitochondrial matrix (*57*), and plays a role in the early stages of spermatogenesis (*58*). In addition, results from a recent study investigating ALR influence on the proliferation of hepatocytes and hepatic tumor cells suggests that ALR could also play an important role in the development of herpatocellular carcinoma (*59*). The above data suggest that ALR's overall role *in vivo* is much broader than just as an augmenter of liver regeneration and, because of potential therapeutic applications related to these studies, a closer look at the molecular mechanism of ALR action was warranted.

ALR homologues (ERV1 and ERV2) have also been identified in yeast (*Saccharomyces cerevisiae*). ERV1 has recently been identified as a key component in a

pathway that is responsible for protein transport and folding in the mitochondrial intermembrane space. Many proteins found in the mitochondria are synthesized in the cytosol and then imported into the mitochondria where they are folded. These proteins are transported to the intermembrane space, the inner membrane, or the mitochondrial matrix (*60*). In *S. cerevisiae*, among these peptides are a group of small cysteine rich proteins, which are targeted to the intermembrane space. Many of these proteins are transported from the cytosol to the intermembrane space via the IMS Pathway (*61-63*). Once the peptides have been synthesized in the cytosol, they are targeted to a translocase of the outer membrane (TOM) channel, which translocates them through the outer membrane of the mitochondria. Once in the mitochondrial intermembrane space, the protein Mia40 binds to them and promotes the formation of intramolecular disulfide bonds leading to correct protein folding and disulfide pairing (*61, 64*). The process leaves Mia40 in a reduced state and the protein must be reoxidized in order carry out its function again on incoming peptides. It is thought that this is accomplished through an interaction with Erv1, which is a sulfhydryl oxidase (catalyzes the formation of disulfide bonds) that leaves ERV1 in a reduced state. It has been proposed that Erv1 is reoxidized via interaction with cytochrome C (*61, 64*). This pathway, shown in Figure 2.7, is essential in yeast because removal of any component leads to an accumulation of unfolded peptides within the intermembrane space (*63, 65*) leading to growth arrest and death.

With the identification of the IMS pathway in *S. cerevisiae*, a search for protein homologs located in other eukaryotic organisms was carried out using bioinformatic analyses. Sequence analyses showed that Erv1, Mia40, and cytochrome c all have homologs. Since these proteins are conserved in higher eukaryotic organisms it suggests

that IMS pathway could be conserved among eukaryotes. Although work has been done to characterize the eukaryotic homologues of Erv1, Mia40, and cytochrome c individually, which suggests that they could function in a eukaryotic IMS pathway (*5, 7, 61, 66, 67*), to date there has not been a study focused on identifying components and confirming the existence of the IMS pathway in, higher eukaryotic organisms.

ALR/Erv1

As stated above, ALR has been identified as the mammalian homologue of *S. cerevisiae* Erv1 (*68*). Interestingly, rat ALR contains multiple splicing sites indicating that it could be found in two forms. ALR exists in the cell as a long form version (lrALR) containing 195 residues and a smaller 126 residue (srALR) version. Erv1 shares a 34% and 40% sequence identity (Figure 2.8) with srALR and lrALR, respectively. (check this to see if it is correct) Both ERV1 and ALR have a similar predicted secondary structure (all-helix): a conserved CxxC motif and a non-covalently bound flavin adenine dinucleotide (FAD) molecule (*69*). From the 3-dimensional structure determined in our laboratory, some insight into ALR's possible function was gained (*51*). ALR contains a unique alpha helical FAD binding motif which explains its yellow color (*51*). In addition, spectroscopic experiments monitoring ALR action on disulfide bond formation in lysozyme showed that ALR, similar to Erv1, was a sulfhydryl oxidase (*5, 51*).

In order to test the possibility of functional homology between the mammalian ALR and yeast Erv1 proteins, a chimeric gene for expression of human ALR in *S. cerevisiae* was constructed (*8*). In this experiment conducted by Lisowsky *et al.*, the yeast mitochondrial targeting sequence was fused to residues 26-125 of the human ALR sequence (*8*). This gene was then introduced into dying yeast cells lacking Erv1 and the

cell division cycle was restored. Restoration of the yeast life cycle by the human ALR chimera indicated that ALR could be described as the functional equivalent of Erv1 (*8*). This Lisowsky study combined with the bioinformatic data suggests that ALR may be performing a role similar to Erv1 in mammals.

Mia40

Originally identified in 2004, Mia40 was determined to be essential for the import and proper folding of the small TIM (translocase of the intermembrane space) proteins of the mitochondrial intermembrane space (*60*). Sequence analysis shows that Mia40 is conserved among eukaryotic organisms (Figure 2.9), however characterization studies have mainly been conducted on the *S. cerevisiae* version of the protein. The majority of the research on Mia40 has focused on how Mia40 functions. In the initial experiments, *S. cerevisiae,* which contained a mutant form of Mia40, were not able to properly import TIM proteins into the intermembrane space (*60*). Following this study, immunoprecipitation and dithiothreitol sensitivity assays were conducted and they determined that Mia40 interacted with Erv1 via disulfide bonding (*7*). Later studies showed that Mia40 was reduced during the folding and disulfide pairing process and that Erv1 would then oxidize Mia40 resetting the system (*7, 61*). In addition to the functional studies, inductively coupled plasma atomic emission spectroscopy (ICP-AES) and trypsin digest assays indicated that recombinantly expressed Mia40 bound and was stabilized by zinc and copper ions (*66*). These findings lead to the belief that Mia40 may also play a role in inserting metal ions into TIM proteins. In addition to the studies done on Mia40 from yeast, a characterization study of the human form of the protein was recently conducted in which they concluded that, similar to the yeast version of the

protein, human Mia40 is also synthesized in the cytosol and the Cx_9C motif within it is essential for its transport into the mitochondria (*67*). The yeast and human studies suggest that Mia40 is likely functioning in a similar manner in both organisms. However, further studies are needed to confirm that Mia40 from these organisms is performing a similar function.

Cytochrome c

Cytochrome c has been characterized as an electron transport protein which functions in the mitochondrial electron transport chain (*70*). Cytochrome c is highly conserved (~90% sequence identity) among eukaryotic organisms and it is located in the intermembrane space. Based on the previous findings that cytochrome c oxidizes some sulfhydryl oxidases, a double mixing stopped flow experiment was conducted in which equine heart cytochrome c was shown to oxidize short-form human ALR (*71*). These findings along with the mitochondrial location of cytochrome c suggested that cytochrome c might be the final electron acceptor for the IMS pathway. In addition, through spectroscopic studies, it was established that Erv1 readily reduced equine heart cytochrome c after being reduced by DTT (*64*). The above studies, although interesting, show that further research on the role of these proteins in the IMS pathway is warranted.

Recently numerous laboratories have focused on gaining insight into the mammalian IMS pathway. Klissenbauer *et al*. provided evidence that the long form of ALR is the mitochondrial functional form of the protein (*58*). With this in mind, research has focused on providing evidence demonstrating electron transfer between mia40, long form ALR and cytochrome c. In 2009, Daithankar *et al.* published work showing transfer of electrons from human Mia40 to long form ALR to cytochrome c *in vitro*. The work

demonstrated that the long form of ALR was capable of accepting electrons from Mia40 while the short form, which is lacking an extra CxxC motif, did display this ability. The catalytic data indicated that the electrons were transferred from Mia40 to long form ALR at a rate of 13 min^{-1} with a catalytic efficiency of 11025 $M^{-1}s^{-1}$ (*3*). Additionally, it was shown that the transfers of electrons from long form ALR to cytochrome c is less efficient than the transfer from short form ALR (*3*). The study further speculated that the reduced electron transfer efficiency for long form ALR was due to steric effects produced by the 70 residue N-terminal extension found in long form ALR that inhibits cytochrome c from getting close to the flavin molecule. Although this work provided important evidence about the electron transfer pathway for the Mia40 – ALR – cytochrome system, more studies are needed to fully characterize the molecular interactions that occur between the proteins.

With this in mind, the focus of the research outlined herein will be aimed at (1) gaining a better understanding of the molecular interactions of key IMS pathway proteins Mia40, ALR and cytochrome c using *Rattus norvegicus* as a model organism and (2) providing structural information about HBsAg needed for the development of next generation vaccines against Hepatitis B and other liver impacting viruses

Figure 2.1: Illustration of the liver depicting the two major lobes, left and right. The lobes are separated by the coronary ligament and falciform ligament. Also the image shows that the gallbladder is located under the right lobe.

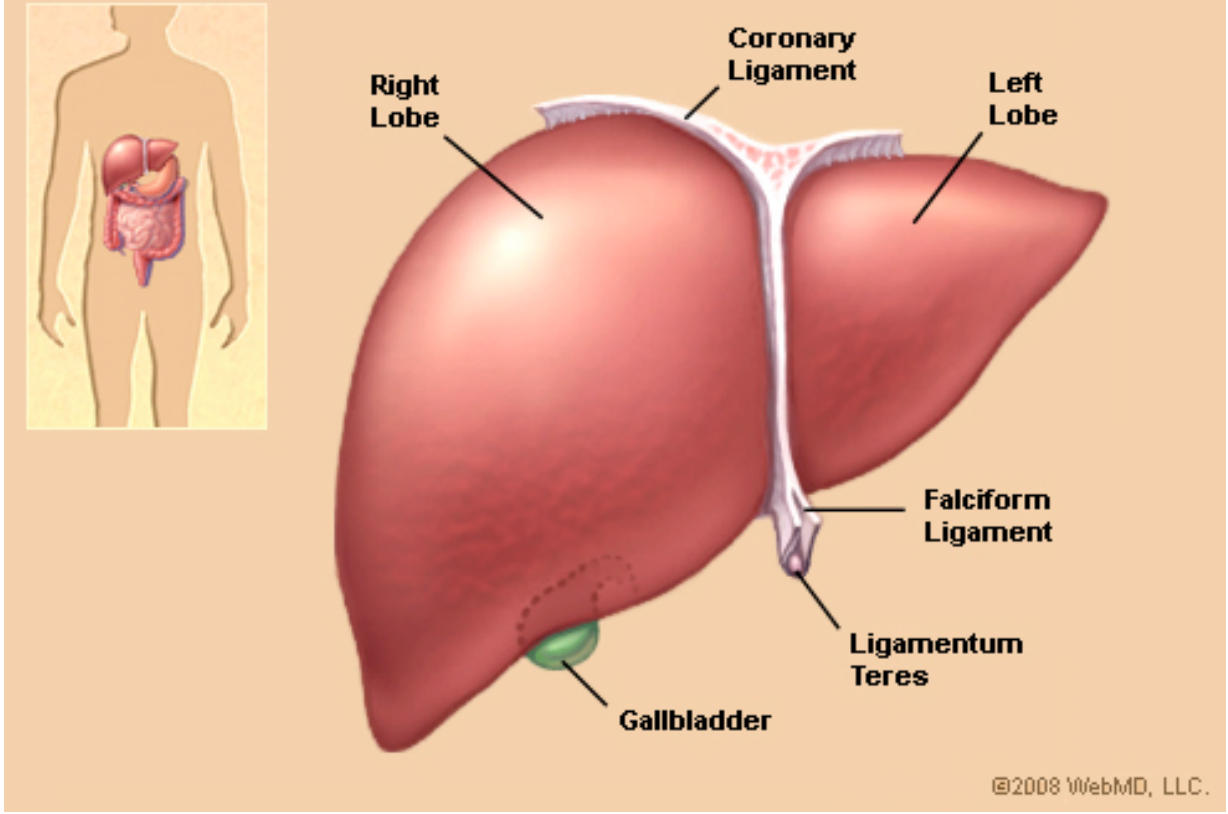


Figure 2.2: A cartoon representation of the Hepatitis B virion (8).

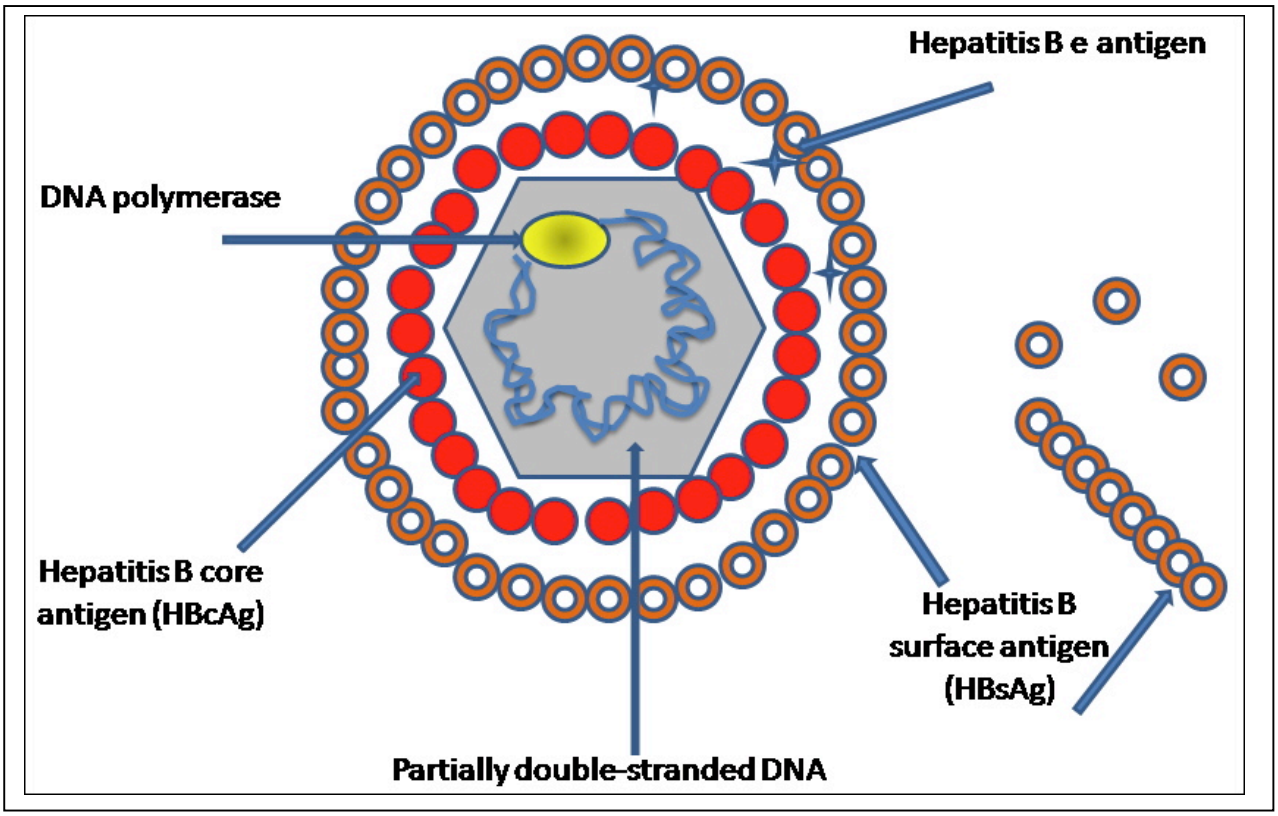


Figure 2.3: Electron micrographs of HBV particles observed in the blood stream. Left – the intact 42-nanometer double-shelled virus particle. Center – filaments of the extra envelop protein HBsAg. Right – 20 nanometer spherical HBsAg particle (11).

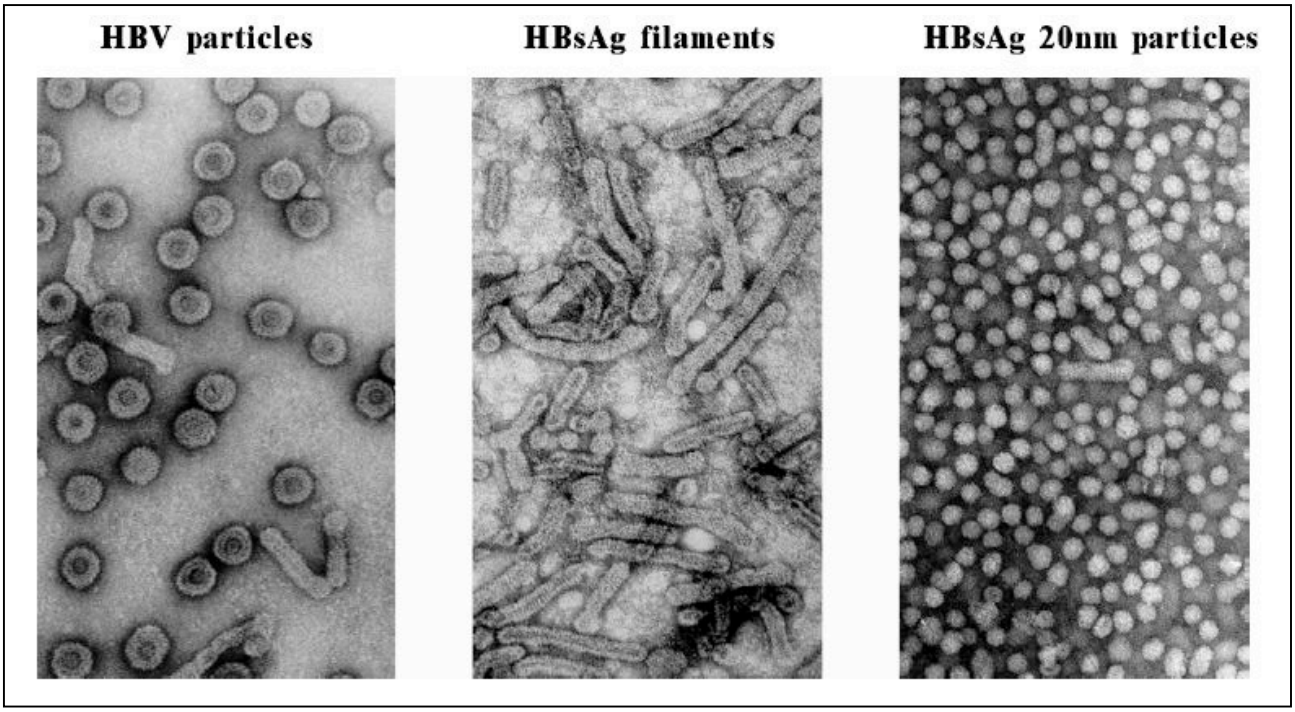


Figure 2.4: Top – amino acid sequence and structure prediction of HBsAg showing the proposed transmembrane helices and loop regions of the protein. Bottom left – sequence of the Rituximab escape mutant, mutations shown as red letters (after Schaefer, 2001). Bottom right – A cartoon representation of the HBsAg structure based on sequence and secondary structure (after Gerlich, 2003)

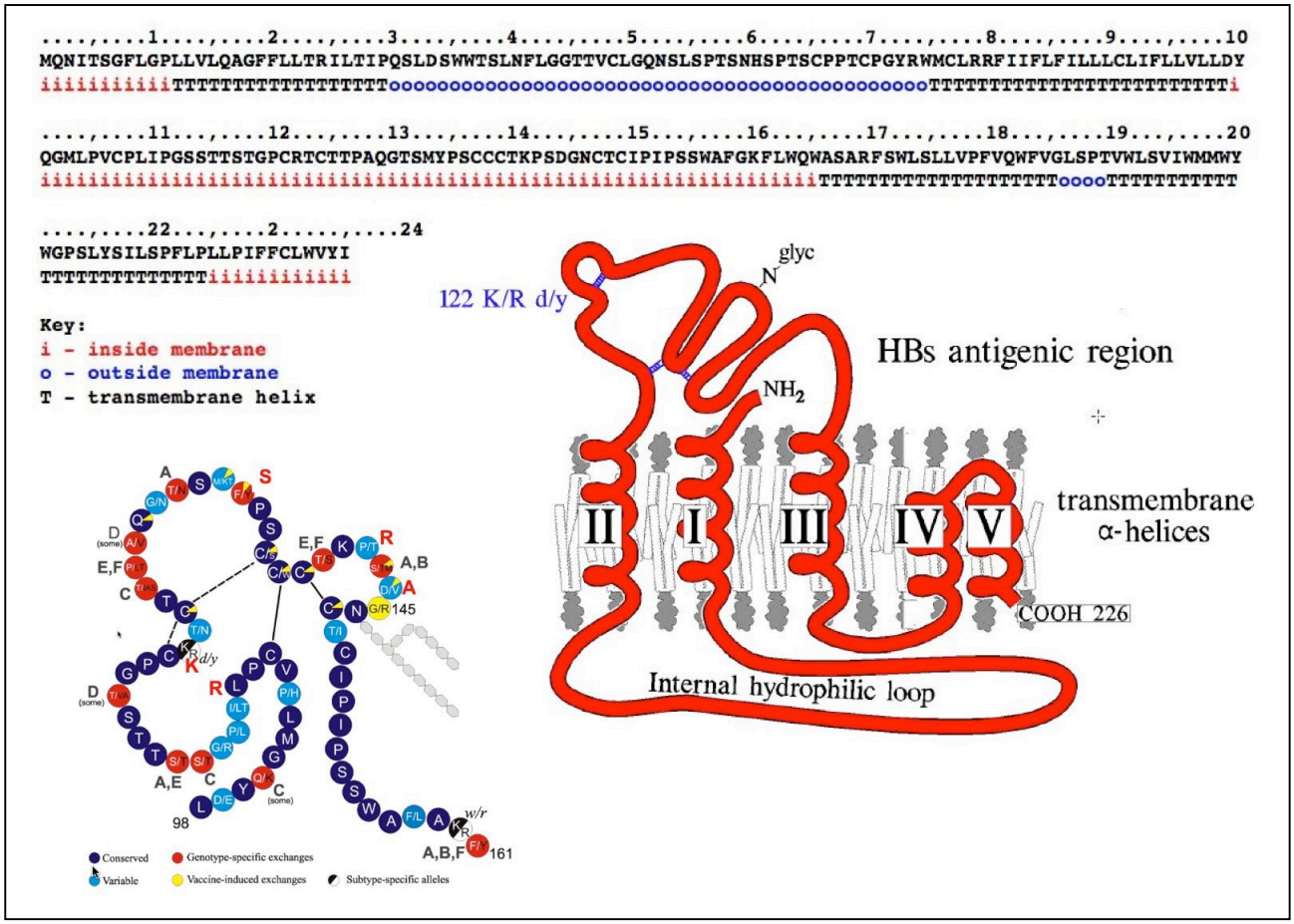


Figure 2.5: Portacavul shunt assay: A portacaval shunt is made when a connection is made between the portal vein (PV) and the inferior vena cava (IVC). The shunt cuts off blood and nutrient supply to the right lobe but not the left lobe. This allows nutrients, protein, and other substances to be added to the right lobe independent of the left lobe. In the assay, the left lobe is used as a control while the right lobe is observed to monitor how it responds to outside factors.

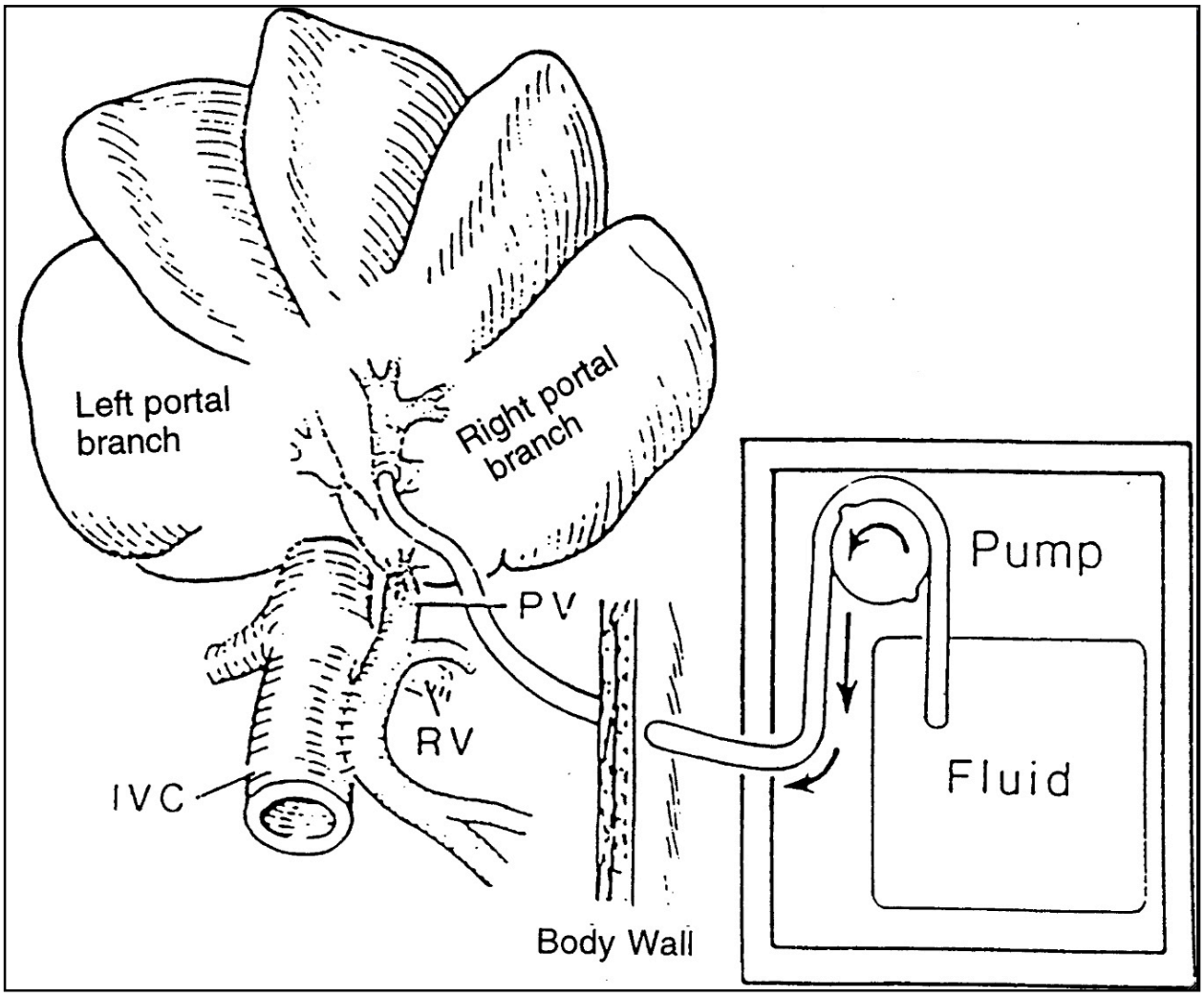


Figure 2.6: Structure of short rat ALR showing the FAD binding site at the mouth of the cone.

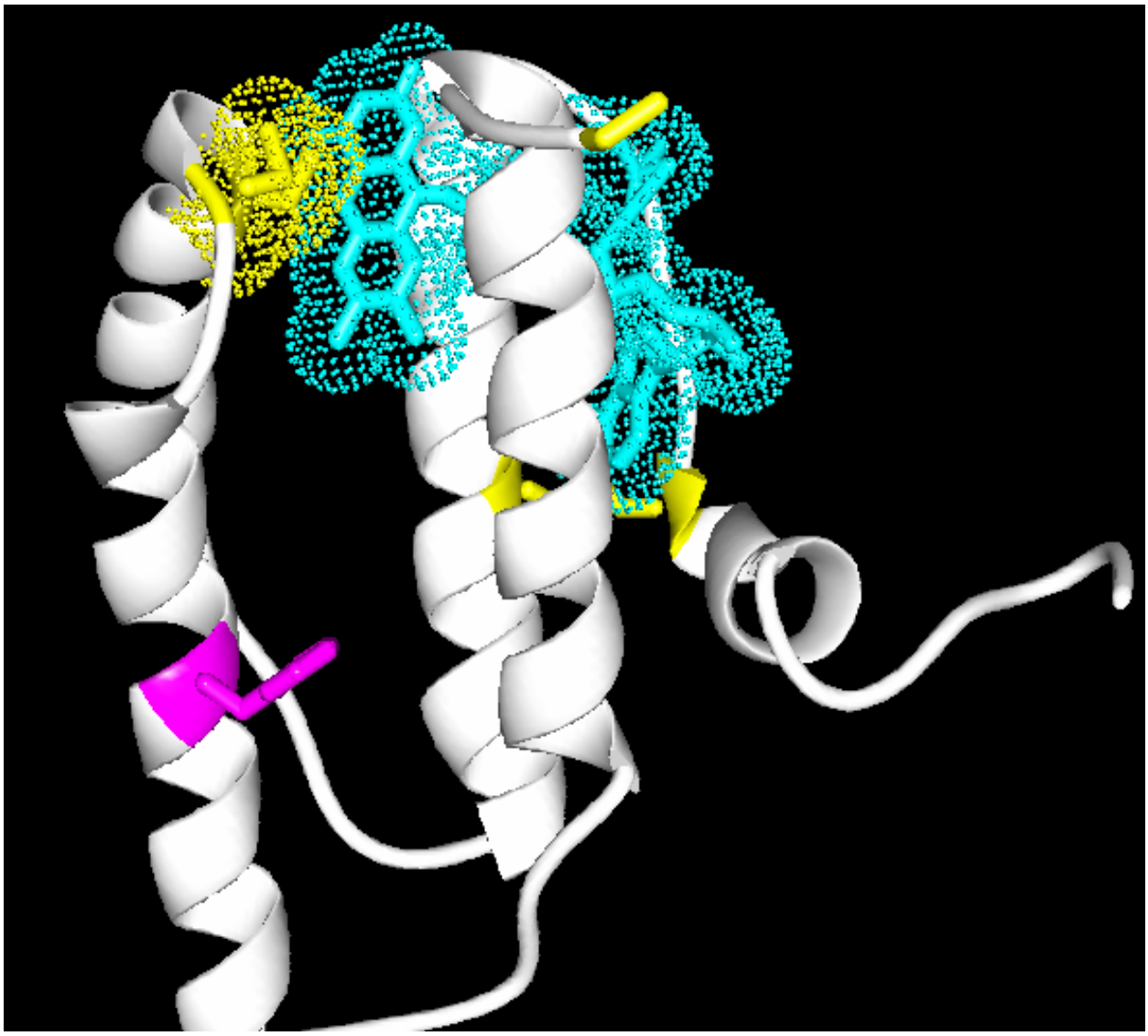


Figure 2.7: Depicts how imported cysteine rich proteins are folded by the IMS pathway in *Saccharomyces cerevisiae* (*63*).

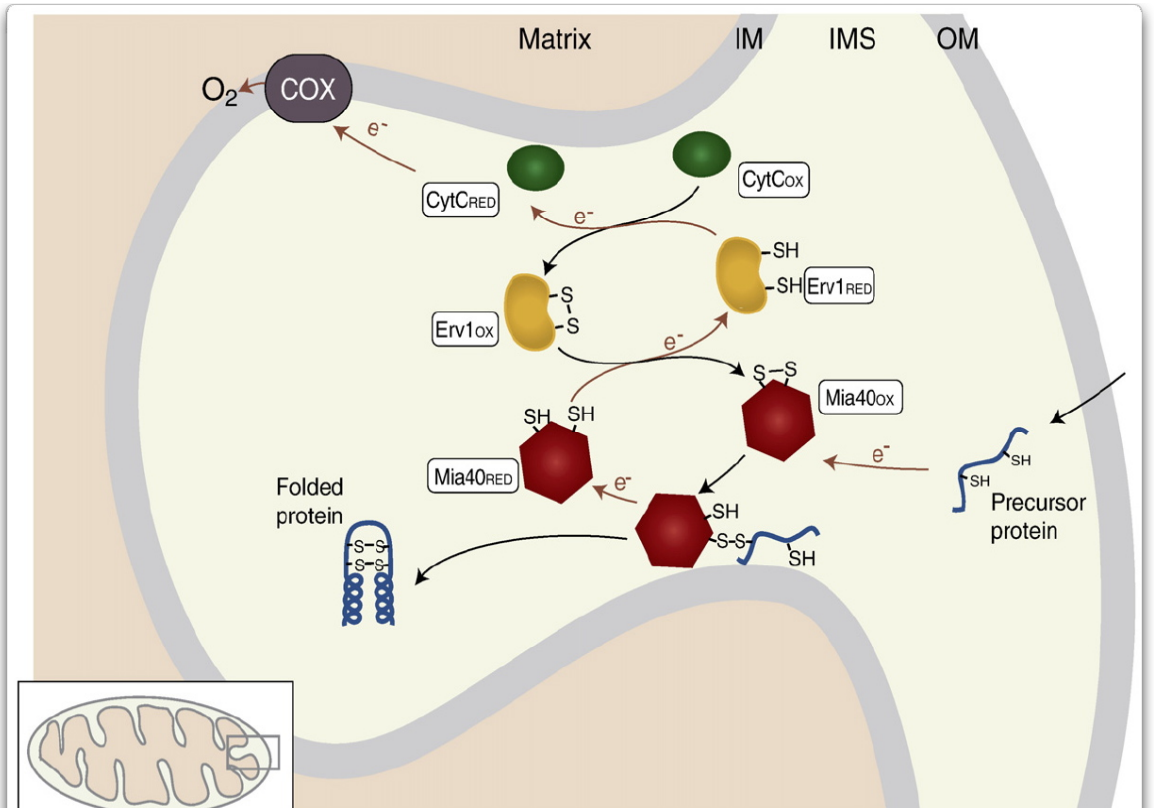


Figure 2.8: Amino acid sequence alignment of Erv1 from *S. cerevisiae* and both short and long forms of ALR from *R. norvegicus*. Identical cysteine residues highlighted yellow and the remaining identical amino acids highlighted red..

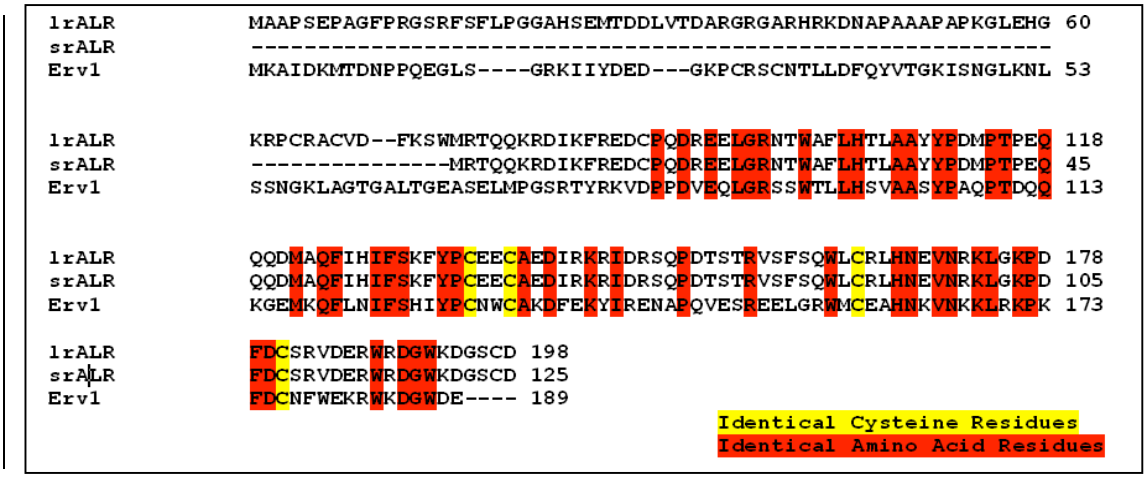
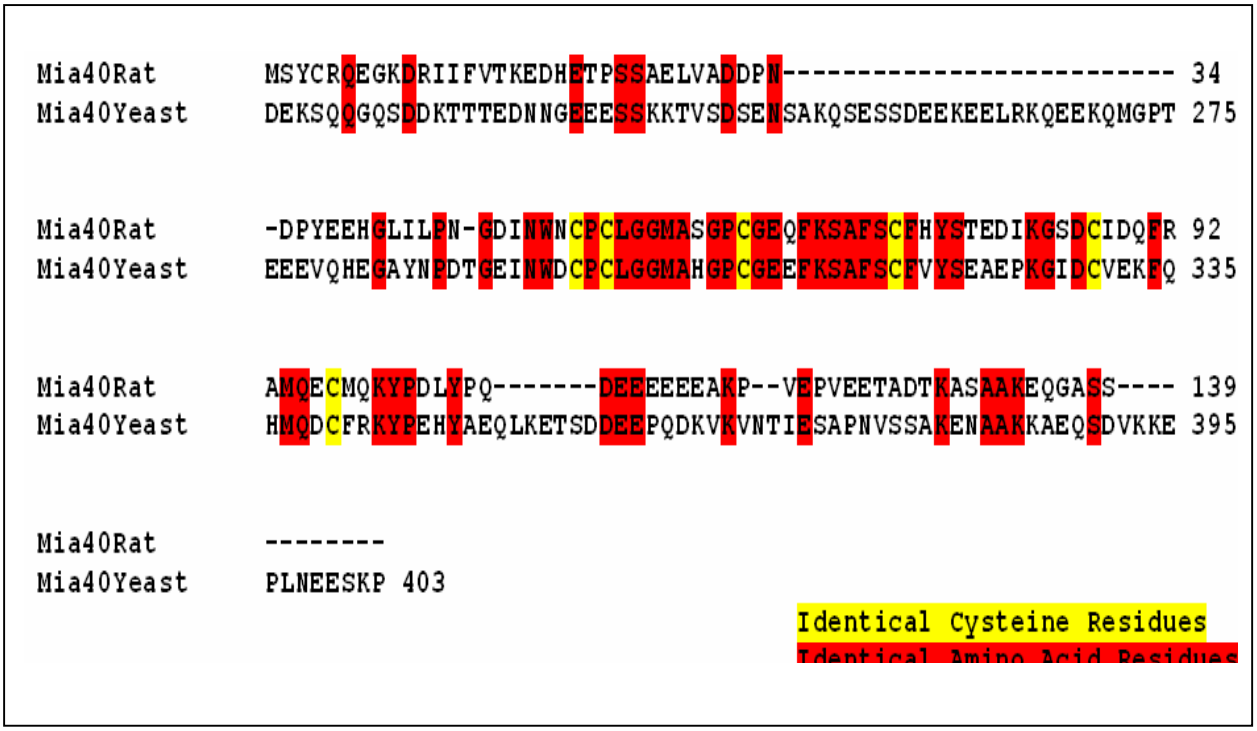


Figure 2.9: Sequence alignment of the C-terminal region (216-403) of Mia40 from *S. cerevisiae* and the Mia40 sequence from *R. norvegicus* with identical amino acids highlighted.



CHAPTER 3

EXPERIMENTAL METHODS

The following section describes the experiments that were conducted as part of the research described herein. In addition, experimental data is provided for the experiments which where not included in the accompanying manuscripts.

3.1 Cloning of *R. norvegicus* and *C. elegans* short and long form ALR

For the studies described, the cDNA needed for the expression of short and long form *R. norvegicus* ALR was a gift from Drs. Antionio Francavilla and Thomas Starzl, University of Pittsburgh Medical School. The cDNA encoding *C. elegans* long form ALR was a gift from Dr. Ming Luo, University of Alabama at Birmingham. The cDNA was used as a template for PCR (polymerase chain reaction). The reaction was carried out using a set of primers designed to incorporate the necessary restriction sites as well as an N-terminal hexahistidine purification tag followed by a tobacco etch virus (TEV) protease cleavage site to aid in purification (*72*). The PCR fragment was then ligated into a pTrcHis expression vector (Invitrogen) and the resulting vectors were named pHHlrALR (rat) and pHHcALR (*C. elegans*). The complete DNA sequence for all ALR constructs were verified by DNA sequencing.

3.2 Cloning of *R. norvegicus* Mia40

Template cDNA for rat Mia40 was purchased from ATCC (Manassas, VA) and used for PCR. Since Mia40 was to be cloned using Gateway® cloning (*73*), the primers used were engineered to contain attb1 and attb2 recombination sites required by the Gateway recombination reaction. The first step in Gateway® cloning, is the BP reaction in which BP clonases or flipases insert the PCR product into a DONR vector that allows

for rapid transfer into other various expression vectors. In this experiment, the Mia40 PCR product was inserted into pDONR 221, which is a holding vector with kanamycin resistance. The second step in Gateway® cloning is the LR reaction in which LR clonases or flipases insert the gene of interest into a destination vector for expression. Here, Mia40 was inserted into several destination vectors, which contained a variety of different purification/solubility tags such as a hexahistidine, maltose binding protein (MBP) or glutathione S-transferase (GST). All LR products were sequenced to ensure that the gene had been cloned correctly. In addition, Mia40 constructs were engineered to contain a TEV protease cleavage site between the purification tags and the Mia40 protein so that after purification the tags can be easily removed.

3.3 Expression and Purification of *R. norvegicus* and *C. elegans* short and long form ALR

To express the ALR proteins a 1-L culture of Circlegrow® media (MP Biomedicals) containing 100 mg of ampicillin was inoculated with *E. coli* JM109 that had been transformed with the selected plasmid. The culture was grown at 37°C for 18 hours. Cells were then harvested by centrifugation. The cell pellets were resuspended in 30 ml of a solution containing 20 mM Hepes pH 7.4, 150 mM NaCl, and 0.1mM PMSF (phenylmethylsulphonyl fluoride), sonicated, and then centrifuged to separate the soluble protein from the cell debris. The histadine tagged ALR proteins were isolated from the soluble protein fraction using metal affinity chromatography by running the supernatant over a column containing a nickel-NTA resin (GE Healthcare). After loading, the Ni-NTA column was washed (3 column volumes) using 20mM Hepes buffer, pH 7.4

containing 150 mM NaCl and 0.1 mM PMSF. The ALR protein was then eluted using the above buffer containing 300 mM imidazole. Following Ni-NTA chromatography, the protein was further purified using size exclusion chromatography. The purified protein fraction from the Ni-NTA column was loaded onto a Superdex200 column (GE Healthcare) and eluted using 20mM Hepes buffer, pH 7.4 containing 150 mM NaCl. The fractions containing the ALR protein (OD 270), were collected and concentrated by centrifugation using a Centracon concentrator (Millipore). In cases where removal of the histidine purification tag was required TEV protease cleavage was carried out. For the TEV protease cleavage reaction the Superdex200 purified ALR protein was combined with 0.5 mg of TEV protease (Sigma-Aldrich), 1mM dithiothreitol (DTT) and allowed to incubate at 30°C over night. Ni-NTA column chromatographyas described above was then used to separate the purified proteins away from the remaining histidine tagged TEV protease. Quality control employing sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and UV/vis spectroscopic monitoring was applied during various stages of the expression and purification process as shown in Figure 3.3.1. The above procedure generally yielded 5 mg of pure ALR protein

3.4 Expression and Purification of Mia40

To determine the optimal expression strategy, small-scale expression tests were carried out using the three purification/solubility tags described above (see Section 3.3). For these experiments, 1 mL cultures of Circlegrow® media (MP Biomedicals) containing 0.1 mg of ampicillin were inoculated with *E. coli* BL21DE3 cells which had been transformed with the His_6, MBP or GST containing plasmids. The cell cultures

where then allowed to grown for 1-2 hours at 37°C followed by induction with 1μL of 1M isopropyl β -D-1-thiogalactopyranoside (IPTG). After induction, the cells were allowed to grow for an additional 2-3 hours. The cells pellets were isolated by centrifugation and analyzed for protein expression using SDS PAGE. From the SDS PAGE analysis, it was determined that the only the pDEST 566 vector containing both an N-terminal hexahistidine-MBP tag successfully expressed soluble protein. Following small-scale expression, large-scale expression using 1L cultures were carried out to produce milligram quantities of the Mia40 protein. After large-scale expression the Mia40 protein was purified using the same buffers and procedures described previously in Section 3.3 for the ALR proteins (see Figure 3.1). Initial attempts to remove the MBP tag by TEV protease cleavage using the protocol described above (Section 3.3) failed. TEV protease cleavage was then attempted using different concentrations (.5 mM to 10mM) of DTT, and different TEV protease sources. In addition the cleavage reaction was carried out at different temperatures (4°C - 25°C). However, these attempts also failed to produce the cleaved Mia40 protein suggesting that that MBP-Mia40 construct might be folded in a way that restricts access to the TEV protease cleavage site. The above procedure generally yielded 7 mg of pure MBP tagged Mia40 protein.

3.5 Spectroscopic Studies of Components of the IMS pathway

In order to provide insight into Erv1/ALR mitigated electron transfer spectroscopic assays were developed to monitor thio group oxidation using reduced lysozyme (*5*). By monitoring absorbance peaks at 412 nm, these assays showed that the reduced thiol groups were oxidized and formed disulfide bonds in the presence of the ALR homologue Erv1 (*5*). Based on these assays and the spectroscopic studies described

by Wu, *et al* (2003) for short-form rat ALR (srALR), spectroscopic assays were developed to monitor the reduction of ALR by dithiothreitol (DTT). In addition, we used a similar spectroscopic approach to characterize the electron transfer from Mia40 to ALR to cytochrome c to provide evidence supporting ALR's role as part of the IMS pathway. The parameters and results for these experiments are described below.

UV-visible spectra were obtained using a Cary 1G spectrophotometer. The initial spectroscopic analysis of the ALR FAD reduction by dithiothreitol (DTT) was monitored with an Olis RSM 1000 spectrophotometer. The UV/visible spectrum (Figure 3.5.1) of purified srALR (120μM) features an absorbance peak at 450 nm, a characteristic feature of flavin molecules such as FAD. Once 1mM DTT was added to the ALR solution, reduction of srALR was observed as indicated by the loss of the oxidized flavin peak at 450 nm. The 450 nm peak did not return for at least 10 minutes (Figure 3.5.2). The experiment was then repeated with additional amounts of DTT added protein solution and the same results were obtained. With the reduction of srALR by DTT established, the assay could then be used to provide initial insight into ALR's possible role in the IMS pathway. In the IMS pathway from *Saccharomyces cerevisiae*, the electron transport pathway involves the oxidation of Mia40 by Erv1leaving Erv1 in a reduced state. Erv1 is then oxidized by cytochrome c leaving cytochrome c in a reduced state. Cytochrome c is then oxidized by cytochrome c oxidase resetting the redox chain.

Since ALR is the mammalian Erv1 homolog, an experiment was designed to see if srALR oxidation rate in the above system is increased by the addition of cytochrome c. The experiment was carrier out in a similar manner as described above but 100μM cytochrome c was added to the ALR/DTT solution after ALR had been reduced.

Unfortunately, cytochrome c addition had no effect on the srALR spectrum. The only observable change in the spectra was the appearance of absorbance peaks at 415 nm, 520 nm, and 550 nm, a characteristic of reduced cytochrome c indicating that cytochrome c was immediately reduced by the DTT in the solution. A second experiment was then carried out in which 100μM cytochrome c and 120μM srALR were mixed together and 1 mM DTT was added to the solution (see Figure 3.5.3). In the first spectrum, which was recorded before the DTT addition, absorbance peaks were observed at 405 nm and 450 nm indicating that oxidized forms of srALR and cytochrome c were present in the solution. After the DTT addition, the 415 nm, 520 nm, and 550 nm reduced cytochrome c peaks that were observed as in the previous experiment. However, after 5 minutes the Figure 3.5.3 shows that peak at 450 nm, corresponding to oxidized flavin, begins to appear indicating that srALR was being re-oxidized. The 450 nm peak was observed to increase with time accompanied by peaks indicating the presence of reduced cytochrome c. The results of the experiment indicate that cytochrome c could play a role in the oxidation of ALR.

To further investigate the electron transfer between Mia40, ALR and cytochrome c, it was determined that the redox reactions need to be monitored in an anaerobic environment absent of oxygen. With this in mind, Mia40 and the short form of ALR where isolated using the previously described protocols (Section 3.3). Following purification, the proteins where degassed and purged with nitrogen gas before being placed in a anaerobic chamber (Coy Laboratory Products), which maintains an oxygen concentration of less then 2ppm. Since Mia40 must be in the reduced form for the electron transfer between Mia40 and ALR to occur, Mia40 was reduced by 1mM sodium

dithionite. The reducing agent was removed from the Mia40 solution by passing the solution over a 10mL Sephadex G25 column that had 5mL of Dowex resin (Sigma-Aldrich) layered on top of it. The sodium dithionite was removed from the solution by the Dowex resin leaving Mia40 in a reduced state. For the spectroscopic assays oxidized short rat ALR (100μM) was placed in an anaerobic cuvett and reduced Mia40 was added at increasing (10μM, 30μM, 60μM, and 90μM) concentrations. The results are presented in Figure 3.5.4 and shows that the addition of 90μM of reduced Mia40 fully reduced the FAD absorbance peak at 450nM implying electron transfer from Mia40 to srALR.

3.6 AlphaScreen Studies

Since the IMS pathway is based on oxidation/reduction reactions, which involve protein-protein interaction, experiments to study these possible interactions were carried out. The initial experiments aimed at studying ALR-cytochrome c interactions were conducted using the AlphaScreen technology (PerkinElmer) (*74, 75*). In an AlphaScreen experiment, one protein is bound to an acceptor bead and its putative binding partner is bound to a donor bead by reductive amidation or other techniques. The protein coated donor and acceptor beads are then mixed together in a variety of potential stabling solutions in a 384 well screening plate. An EnVision® (PerkinElmer) plate reader modified for AlphaScreen is used to analyze the solutions contained in the wells of the screening plate for the presence of an AlphaScreen signal at 520-620 nm. The plate reader excites each well in the screening plate using 680 nm photons that generates singlet oxygen on the donor bead. If the donor and acceptor beads are in close proximity

(less than 200 nm) the singlet oxygen can interact with the acceptor bead generating florescence in the 520-620 nm range, which is recorded by the reader.

AlphaScreen experiments were carried out to test for the interaction between horse heart cytochrome c and short-form rat ALR. For the analyses 20μM ALR and 20μM cytochrome c were covalently attached to the AlphaScreen acceptor and donor beads (four separate experiments) via reductive amination using the manufacturers protocols (*74, 75*). Reductive amination links the carbonyl groups on the peptide backbone of the protein to the surface of the respective bead. The protein-linked donor and acceptor beads were added to the wells of a 384-well screening plate containing, each containing one of the buffer systems to be analyzed. For each buffer condition analyzed, two experiments were carried out. The first experiment used ALR bound to the AlphaScreen donor bead and cytochrome c bound to the AlphaScreen acceptor bead. In the second experiment cytochrome c was bound to the AlphaScreen donor bead and ALR was bound to the acceptor bead. Each experiment was then carried out in triplicate.

The initial screening was used to identify buffer systems that promoted ALR-cytochrome c interaction. (Figure 3.6.1). For those buffer systems that showed that the two proteins were interacting (high AlphaScreen signal) further analyses were carried out. In these experiments unconjugated cytochrome c (or ALR) were was titrated into the solution to compete with the bound protein. The unconjugated protein should disrupt the interaction of the two beaded proteins and cause the AlphaScreen signal to decrease, as more protein is added, similar to a competitive binding experiment. Although in general theses analyses showed a decrease in AlphaScreen signal upon addition of unconjugated protein, only some of the competitive binding experiments showed the expected

continued decrease in AlphaScreen signal as more unconjugated protein was added to the system. An example of the contrasting results is presented in Figure 3.6.2. These results provide qualitative support for the interaction of cytochrome c with short form ALR.

3.7 Surface Plasmon Resonance

AlphaScreen provided initial evidence that a physical interaction occurs between ALR and cytochrome c. In order to quantitative data about the interaction between the two proteins, surface plasmon resonance (SPR) experiments where carried out utilizing a Biacore3000 (GE Healthcare). SPR occurs when light interacts with a metal surface, normally gold or silver, and causes the formation of surface plasmons (coherent electron oscillations) on the interface between the metal and another material (*76*). In the Biacore experiment, a ligand is immobilized to the metal material and solution of containing the analyte under study is passed over the ligand. If interaction between the ligand and analyte occurs, the mass of the ligand molecule interacting with the surface changes and this alters the distribution of the surface plasmons (*76*). The Biacore analyses utilizes this phenomena, in this case the change in the distribution of the surface plasmons due to a mass change of the ligand molecule interacting with the surface is measured by the change observed in the angle of light reflected from the surface. The change in angle of the reflected light is measured and recorded as resonance units by the Biacore3000 (*76*). By monitoring the real time changes of the resonance units, information can be gained about the association rate, dissociation rate, and equilibrium constants associated with the system under study.

Biacore experiments where designed to determine the binding kinetics of the

interaction between rat ALR (both short- and long-forms) and horse cytochrome c. In these experiments long-form ALR (120μM in 20mM HEPES buffer, containing 150mM NaCl, and 0.1% Triton-X 100,pH 7.4) was covalently linked to a Biacore Sensor chip CM4 (channel 2) using 70μL of a 1:1 mixture of N-hydroxysuc- cinimide (NHS) and N-ethyl-N'-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC) according to manufacturer protocols. The remaining activated dextran molecules on the CM4 chip where then blocked with 70μL of ethanolamine (*77*). The reaction resulted in a resonance unit (RU) change of 200 (100RU = 1ng/mm^2). For control purposes, channel 1 of the CM4 chip was activated with a 70μL 1:1 NHS-EDC mixture and blocked with 70μL of ethanolamine. The above immobilization process was then repeated using short-form ALR (158μM in 20mM HEPES, containing 150mM NaCl, and 0.1% Triton-X 100,pH 7.4). srALR was bound to channel 4 of the CM4 chip and channel 3 was used as the control for the short-form ALR interaction experiments. The srALR immobilization resulted in a 2000RU change.

The running buffer used for all experiments was 20mM HEPES buffer containing 150mM NaCl, and 0.1% Triton-X 100, pH 7.4. All experiments were run under oxidizing conditions at 25°C using the Biacore 3000 software. For the interaction analyses, cytochrome c (0.4μM, 0.8μM, 1μM, 2μM, 4μM, 5μM, 6μM, 8μM, 10μM and 12μM) dissolved in running buffer was injected into the appropriate channels of the CM4 chip using a flow rate of 15μL/min for 3 minutes to monitor association of the ALR ligand and cytochrome c analyte. After injection, running buffer alone was for 10 minutes to monitor the dissociation. After each experiment the chip was regenerated using a 30 second injection of 30μL of Glycine-HCl, pH of 3.96. All data analysis and binding

kinetic information was obtained using BIAevaluation software from GE Healthcare.

3.8 Small Angle X-Ray Scattering

Small-angle X-ray scattering (SAXS) is a simple experiment that can be used to characterize the solution properties of macromolecules, such as proteins, and their interactions (*78, 79*). Recently, SAXS has recently been used to characterize transient protein-protein complexes in solution (*80, 81*). Unlike an X-ray crystallography, a SAXS profile represents the simultaneous scattering measurement of the macromolecule in all orientations (*82*). The scattering data is transformed from reciprocal space to real space yielding the P(r)-distribution, which is a histogram of the interatomic vectors within the macromolecule. The P(r)-distribution (pair distance distribution) gives an estimate of the maximum dimension (d_{max}) of the particles in solution (*79*).

In addition, SAXS experiments allow for the calculation of the radius of gyration (R_g) of the molecules in solution. An increase in R_g, which provides an indication of the mass distribution of a macromolecule about its center of gravity, is an indication that the macromolecule is opening or unfolding (*82*). Additionally, SAXS data can be used to determine the volume of the scattering particle. Using these SAXS parameters, information can be gained about the dynamics of protein molecules in solution. Additionally, electron density envelopes are generated which give a visual model of the molecules in solution.

Since the SAXS experiments conducted for this research focused on monitoring the interaction of rat ALR with horse heart cytochrome c, several SAXS experiments were carried out using different protein combinations and buffer conditions as described

below. SAXS data were collected on beamline 12.3.1 (SIBYLS), Advanced light Source, Lawernce-Berkeley National Laboratory in collaboration with Dr. Jeffrey Habel (Lawrence-Berkely Laboratory). All experiments solutions were prepared under oxidizing conditions. For the SAXS analyses all proteins (short-form ALR, long-form ALR, cytochrome c, the short-form ALR-cytochrome c mixture, and long-form ALR-cytochrome c mixture) were diluted with 20mM HEPES buffer containing 150mM NaCl, pH 7.4 buffer to give a 20μL sample with a final protein concentration of 5mg mL^{-1}. For each protein sample analyzed, a buffer blank was also recorded (same exposure time) for buffer subtraction. For the initial experiments SAXS data (both protein solution and blank) were collected, in duplicate using a 0.5 sec exposure time

To analyze the ALR-cytochrome c interactions under reducing conditions solutions were prepared using 20mM HEPES buffer containing 150mM NaCl, 1mM tris(2-caboxyethyl) phosphine hydrochloride (TCEP), pH 7.4 and concentrated to give a protein concentration of 5mg mL^{-1}. As stated in Habel *et al*. (*83*), the buffer-subtracted data were analyzed using PRIMUS (*84*) and the GNOM (*85*). The P(r) output data was then used to calculate electron-density envelopes (see Figures 5.1 and 5.2). Each envelope is the product of 16 GASBOR (*86*) runs averaged with DAMMAVER (*87*).

3.9 Crystallization of long-form *C. elegans* ALR

Due to problems encountered in the crystallization of long-form ALR (both human and rat protein) alternate sources of the long-form protein were investigated. Secondary structure predictions using the long-form human ALR sequence showed that the N-terminal extension was largely unstructured. Sequences of long-form ALR like proteins

from other organisms were then investigated using the presence of predicted secondary structure in the N-terminal extension as a guide. The analysis showed that the *C. elegans* protein F56C11.3 (product of ORF F56C11.3) was a short 161 residue long-form ALR-like protein whose N-terminal extension contained a CxxC motif and some regions of predicted helical structure. Initial crystals of the long-form *C. elegans* ALR were obtained using the sitting drop vapor diffusion method. A 16 head Cartesian Honey Bee screening robot was used to set up 384 screening experiments using crystallization kits purchased from various sources ((Crystal Screens 1, 2, Cryo, Membfac, and Peg/Ion (Hampton Research), Wizard Screens 1 and 2 (Decode Genetics) and Classic screens 1 and 2 (Nextal Biotechnologies)) as previously reported (*88*). In the initial screening experiments, 200 nL sitting drops composed of equal volumes of protein and precipitant cocktail were set up using Greiner's square-well Crystal Quick plates (Cat.-No.: 609120). The plates were incubated at 18°C and imaged (day 1,3,7,15,30) using a CrystalFarm imaging system (Bruker AXS, Madison WI).

From the initial screen, crystals were observed in wells containing the Membfac #29 screen (0.1 M ammonium Sulfate, 0.1 M HEPES – Sodium, pH 7.5, 0.5 M di-sodium hydrogen Phosphate dihydrate/ 0.5 M di-Potassium hydrogen phosphate). The crystals, although large did not diffract, and could not be used for structure determination.

After obtaining the initial long-form *C. elegans* ALR crystals, more protein was purified to obtain better quality crystals. For these experiments, the 384 crystallization screens were set up using the Douglas Oryx8 robot, which prepares microbatch crystallization experiments (*89*). In microbatch crystallization, equal volumes (0.5 μL) of protein and precipitant cocktail are added to the bottom of a well of a NUNC MiniTray (

Cat. No.: 470378) plate (*89*) and then the aqueous droplet is covered by a mixture of 80% to 20% paraffin/silicon oil. The setups were then incubated at 18°C. As the plate incubates, the paraffin-silicone oil mixture will allow water to evaporate from the drop causing the concentration of the protein and precipitant to increase and hopefully as this happens crystals will form. From the microbatch set ups that were prepared, crystals (Figure 3.9.1 left) were obtained from Crystal Screen 1 #20 (100mM HEPES pH 7.5, 10% Isoproponal, 20% PEG 4000) but these crystals were small brittle needles which could not be mounted. Using this condition as a guide, crystal optimization experiments were set up. From these screens, crystals (Figure 3.9.1 right) with better morphology and size were obtained in multiple conditions but again, these crystals did not diffract using rotating anode X-rays.

3.10 Structure determination of hexahistidine tagged rat short-form ALR

The Rose lab has previously reported the crystallization of rat short-form ALR (*90*) containing a 14 residue (MGGSHHHHHHGMAS) N-terminal purification tag. These crystals were obtained by vapor diffusion using 2 μL hanging drops containing equal amounts of protein (20 mg mL^{-1}) and precipitant cocktail. The best crystals were obtained using a precipitant cocktail containing 20% PEG methyl ether (PEGMME) 2000, 0.1 M sodium acetate, pH 4.6 with 50 mM cadmium chloride added just prior to setup. A data set to 2.4Å consisting of 720 one-degree oscillation images recorded at cryogenic temperatures (~100K) on a MAR Research 300 mm image plate system was collected using rotating anode generated Cu Ka X-rays. The data were recorded in 30° wedges so the cadmium anomalous signal could be optimized. The data were indexed, integrated and scaled (in 90° batches) using HKL 1.9.1 (*91*). Data collection and

processing parameters are collected in Table 4.1.

The his-tagged ALR crystals were found to belong to space group $I4_1$ with a = 99.8Å and c = 113.4 with four molecules in the asymmetric unit, based on Matthews analysis (*92*). Analysis of the Bijvoet difference gave four Cd positions, however the structure could not be solved using cadmium signal present in the 2.4 Å single wavelength anomalous scattering (SAD) data set. Although a reasonable model could be obtained by Molecular Replacement (C.K. Wu personal communication) problems during refinement of the structure suggested that there were significant differences between the his-tagged structure and the ALR model (PDB entry 1OQC) used for the molecular replacement analysis. Thus, renewed attempts at solving the structure using the cadmium SAD data were initiated since and accurate crystal structure of his-tagged ALR would aid in the interpretation of the SAXS results.

3.11 Phasing Experiments

3.11.1 SAD Phasing

In order to complete the structure determination process, the phase problem must be addressed. For a review see Wang, 1985 (*93*). Briefly, in the X-ray diffraction experiment protein crystals are placed into an X-ray beam and the diffracted X-rays, which result when a lattice pass through the Ewald Sphere and Bragg's Law is satisfied, are collected with an X-ray area detector. From the recorded image the intensities of the reflections (defined by the Miller index hkl) and positions of the reflections in x, y and ϕ can be measured. The intensities are proportional to the square of the structure amplitudes (factors) $|F_{hkl}|$ for the given reflection.

$$I_{hkl} \propto |F_{hkl}|^2$$

The structure factor *F(hkl)* for a reflection *h, k, l* is a complex number and can be represented as follows:

$$\mathbf{F}_{hk\ell} = \sum_{j=1}^{N} f_j e^{2\pi i(hx_j + ky_j + \ell z_j)} = \sum_{j=1}^{N} f_j e^{i\phi_j}$$

where $2\pi(hx_j+ky_j+lz_j)$ is the phase (ϕ) of reflection hkl and *f (j)* is the scattering factor of atom *j* and depends on the element and the diffraction angle ($\sin\theta/\lambda$) of the corresponding reflection (*h,k,l)*. The above formula implies that if the atom types and positions (x,y,z) are known the structure factor can be calculated for each reflection (hkl) in units of electrons. Because the formula for the structure factor is a periodic function a Fourier transformation (FT) can be applied to it giving the electron density at any point (x,y,z) in the crystal lattice.

$$\rho_{(x,y,z)} = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} F_{(h,k,l)} \exp\left[-2\pi \cdot i(hx + ky + lz)\right]$$

In the Fourier transform, *reciprocal space* is being transformed into *real* or *direct space* (the electron density at a *real* point x,y,z in space). The value $\rho_{(xyz)}$ is the electron density at point x,y,z because *F(hkl)* is in units of electrons and the sum is divided by the cell volume *V*. In other words the diffraction pattern (an image of the reciprocal space) is transformed into the real space of electron density (an image of the protein). So in theory if the structure factors (obtained from the diffraction experiment) are known the structure (the density *of* the electrons in real space) corresponding to the diffraction pattern can be

determined provided we know the phases. In order to perform the Fourier transform, the complex structure factors $F_{(hkl)}$ are needed but only their magnitude $|F_{hkl}|$ can be measured since phase information is lost in the experiment. In terms of physics this means that the absolute value of the complex vector $F_{(hkl)}$ is known but not its phase, $\alpha_{(hkl)}$. This results in what is called the Phase Problem. So a better equation for the electron density is:

$$\rho_{(x,y,z)} = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l}\left|F_{(h,k,l)}\right|\exp\left[-2\pi \cdot i(hx + ky + lz - \alpha_{(h,k,l)})\right]$$

So every part of the above equation except $\alpha_{(h,k,l)}$ can be measured. To get an estimated of the phase however in order to calculate the phase other methods must be employed. Theses include multi-wavelength anomalous dispersion (MAD) (*94*) and single-wavelength anomalous scattering (SAD) (*93*). For the purpose of the research presented in this dissertation we will focus on SAD phasing using Wang's approach as an example.

In the SAD experiment, a single dataset collected at a specific wavelength will generate both a true phase and a false phase for each reflection (Figure 3.11.1). Wang has shown that the phase ambiguity associated with SAD data can be resolved using an iterative approach employing both real and reciprocal space filters and the fact that the electron density at any point in the lattice can not be less than zero. Wang's "Solvent flattening" approach resolves the SAD phase ambiguity problem by taking advantage of the fact that when heavy atoms are not in special positions or related by a center of symmetry, the Fourier map produced by the false phases will contain no structural information and will be in a form of general background noise. This background noise is generally lower than the correct protein density in the map. As seen in the flow chart

describing Wang's method (Figure 3.11.2) initial phases from the data are calculated and Fourier transformed to produce electron density. The key to solvent flattening is the process of accurately defining the boundary between a molecule and solvent in an electron density map. Programs that perform solvent flattening choose the boundary based on the assumption that except in certain cases the population and strengths of signals detected for the protein region are higher than that of the solvent. This boundary is a summation of density around a point. This summation is used as a mask and then the electron density outside of the mask, presumably only solvent, is flattened to 0 leaving only protein electron density. This improved electron density map is then inverse Fourier transformed back to calculated phases. Using a phase filter, original phases and the calculated phases are combined through phase averaging to generate improved phases. The improved phases are then cycled back with the structure factors from the data and the process cycles through again. This iterative cycling between real and reciprocal space occurs four times, and at the end of the fourth cycle, the electron density filter is saved and the process starts back up with the original phases and original amplitudes, as to not over bias the phases towards the inverse Fourier calculated phase value. This continues on for another four rounds and then that final density filter is used with the original phases and amplitudes. The density filter from the second cycling is used for a final four iterations to optimize the phase and the last step is three rounds of phase extension. The final phases, after phase extension, are used to calculate the final electron density map.

3.11.2 Cadmium-SAD phasing of the his-tagged short-form ALR structure

For the renewed attempt at Cd-SAD phasing the merged 720° SAD data set to 2.4Å

resolution described by Wu, *et al*., (2000) was used. The latest version (1.6.4) of the PHENIX software suite (*95*) was downloaded and installed on a Mac Pro computer (8 processors, 12 GB memory and 4 Tbytes disk storage) running OSX version 10.6.4. The PHENIX *Reflection file editor* was used to convert the SCALEPACK formatted (*96*) reflection file to MTZ format keeping the Bijvoet pairs separate and setting Rfree flags needed for refinement. The PHENIX *AutoSol* wizard was then used to setup and carry out the initial SAD phasing. A flow chart describing the inner workings of the AutoSol wizard is provided in Figure 3.11.3.

Phenix.Autosol takes in anomalous and/or isomorphous data and corrects it for directional dependence in diffraction quality (anisotropy) if it is necessary. The data is scaled and a substructure is located using phenix.hyss. The handedness of the substructure is scored and then phased using Phaser (*97*) or Solve (*98*). Statistical density modification and initial model building are completed using Resolve (*98*).

AutoSol was able to both phase the data and fit 441 of the 556 residues in the asymmetric unit. Briefly, SOLVE was able to identify 9 potential Cd sites (Table 3.11.2), four of which corresponded to the Cd sites reported by Wu, *et al*. (2000). The solution had an Overall Score of 44.98 +/-9.58 and a Figure of Merit of 0.43. RESOLVE and Buccaneer were used to fit the sequence into the resulting SOLVE phased electron density map. The initial model (441 residues, 13 segments and 85 solvent molecules) had an R value of 0.3208, an Rfree value of 0.3487 and a map to model correlation coefficient of 0.72. The model was then rebuilt using the PHENIX *AutoBuild*. After the rebuilding step the model (462 residues, 19 segments and 114 solvent molecules) had an R value of 0.2750, an Rfree value of 0.2950 and a map to model correlation coefficient of

0.74. At this point manual inspection using COOT (*99*) was carried out. Initial fit of the four FAD (**flavin** adenine dinucleotide) molecules in the asymmetric unit (one per monomer) was achieved using the 1OQC structure as a guide using CHIMERA (*100*) to superimpose the 1OQC structure onto the model produced by PHENIX *AutoBuild*. Correct placement of the FAD molecules was then confirmed using 2Fo-Fc electron density maps displayed using COOT. The *AutoBuild* structure also showed that there were two different Cd environments, three Cd-X_4 moieties and what appeared to be a Cd_2-X_6 moiety. Bijvoet difference Fourier analyses indicated that four of the atoms in the Cd_2-X_8 moiety were most probably chlorine since chloride was present in high concentration in the crystallization setup and chloride has a small but measurable (ΔF"= 0.84e^-) anomalous scattering signal using Cu Kα X-rays. In addition, the Cd-X bond lengths (~2.6Å) observed for these ligands agreed better with a Cd-Cl bond (bond length 2.6Å) than a Cd-O bond (bond length 2.3Å). Two of the remaining four cd ligands came from the side chains of Glu A 74 (XX) and Glu B 74 (XX). The remaining two ligands were modeled as water since they lacked anomalous scattering signal and the Cd-X bond distance of (2.3Å) fits a Cd-O bond.

3.11.3 Refinement of the his-tagged short-form ALR structure

The manually rebuilt ALR model including the four FAD molecules, three CdO_5 moieties and the $Cd_2Cl_4O_2$ cluster was then refined using the PHENIX *Refine* wizard. A flow chart describing the inner workings of the PHENIX *Refine* wizard is provided in Figure 3.11.4. Phenix.refine increases the quality of the structure by repeating a series of refinement tools. Among these tools used are bulk solvent, rigid body, atomic displacement

parameter, and occupancy refinement. These refinement tools utilize the refinement of individual atoms as well as sections of the protein to produce higher quality structures.

Figure 3.3.1: SDS PAGE analysis of purified srALR, lrALR, and Mia40 following the completion of the metal affinity and gel filtration chromatography..

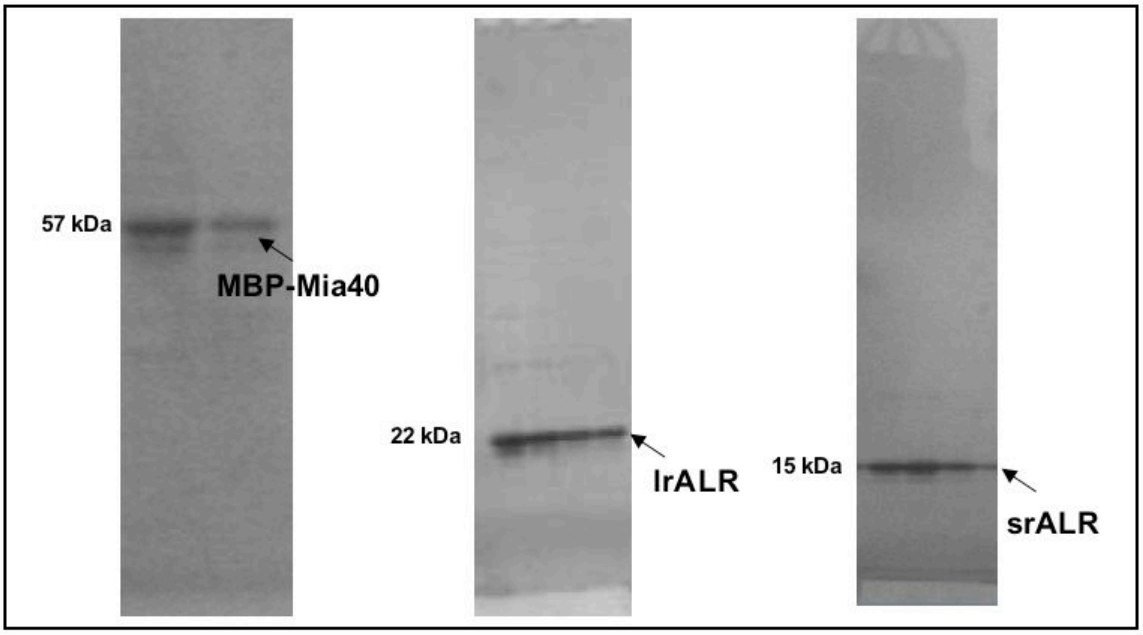


Figure 3.2.1: UV-VIS spectra of srALR showing the characteristic absorbance peak at 450nm. The peak corresponds to a flavin adenine dinucleotide that is coordinated by amino acids.

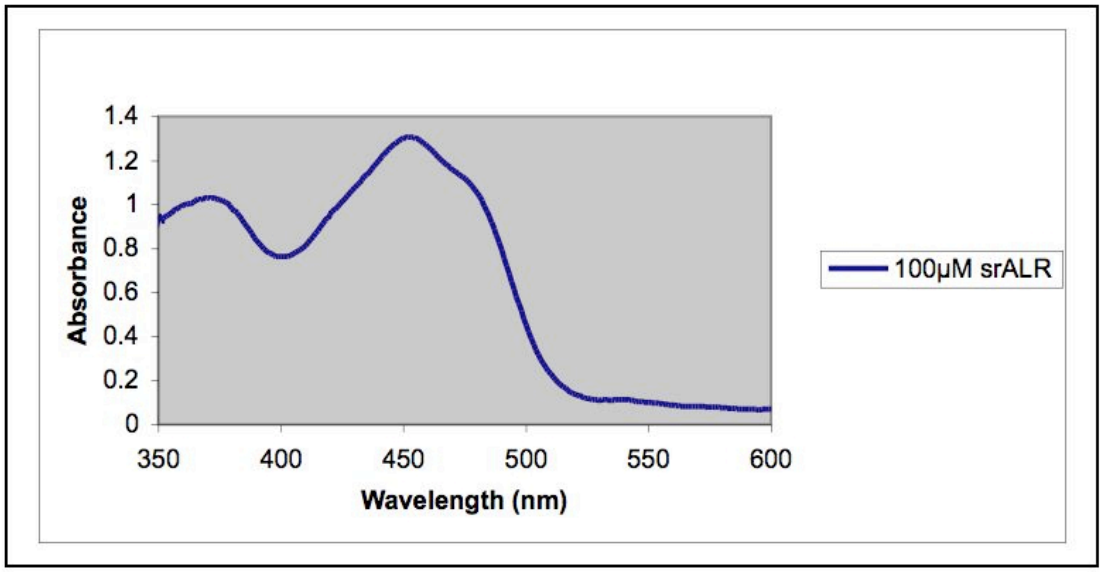


Figure 3.5.1: The spectra show the reduction of the FAD molecule in srALR by dithiothreitol. The absorption peak at 450nm is reduced and does not return for at least 10 minutes.

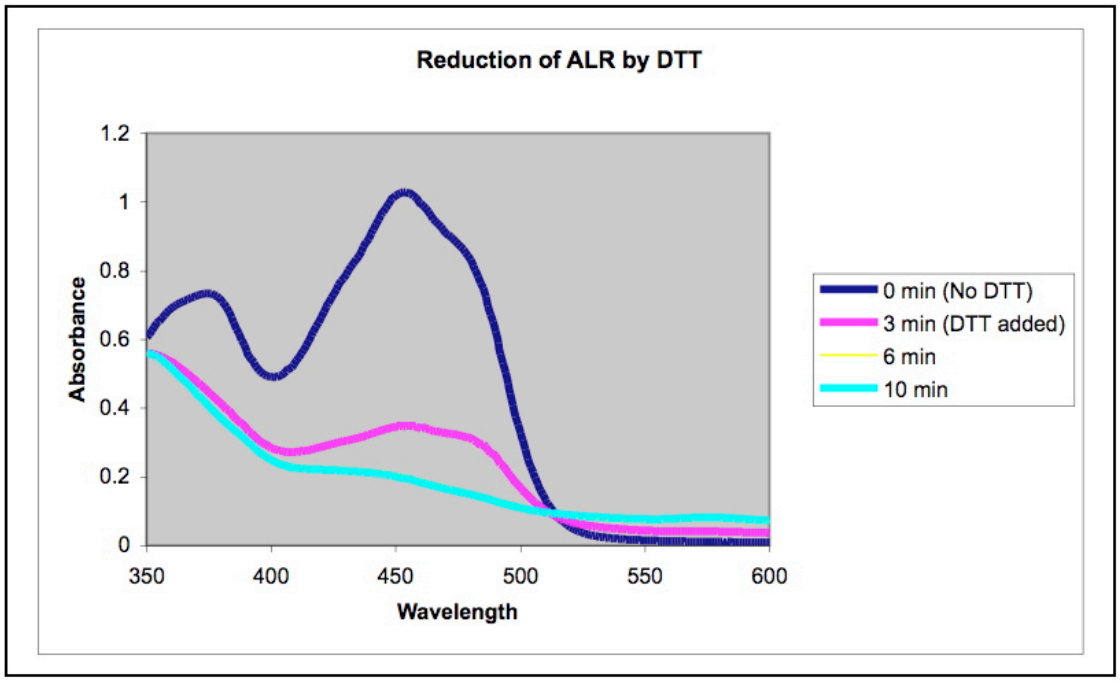


Figure 3.5.2: Absorption spectra for srALR and cytochrome c. When dithiothrietol is added to the sample, the 450nm peak for srALR is reduced and 415nm, 520nm, and 550nm peals appeared. These peaks correspond to cytochrome c being present in the reduced form. Unlike when srALR was reduced without cytochrome c present, the 450nm peak returns after 3 minutes and increases over time.

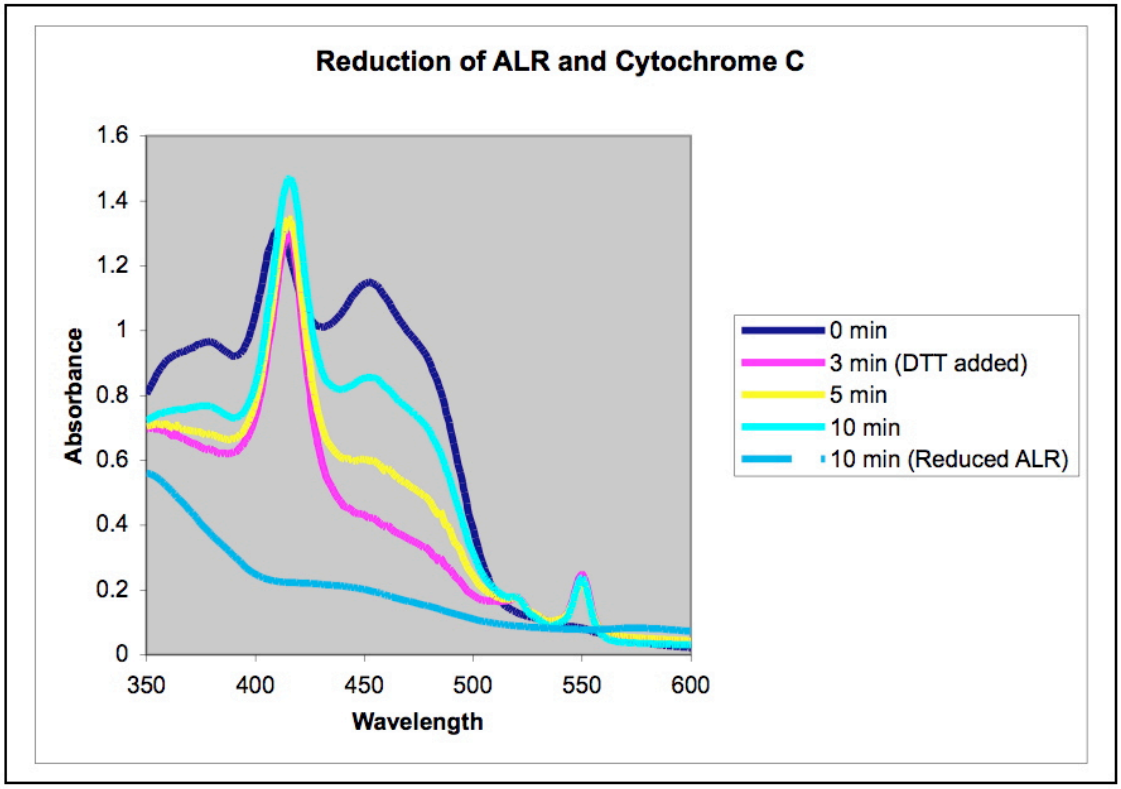


Figure 3.5.3: Absorption spectrum showing that srALR is reduced when reduced Mia40 is added too the solution.

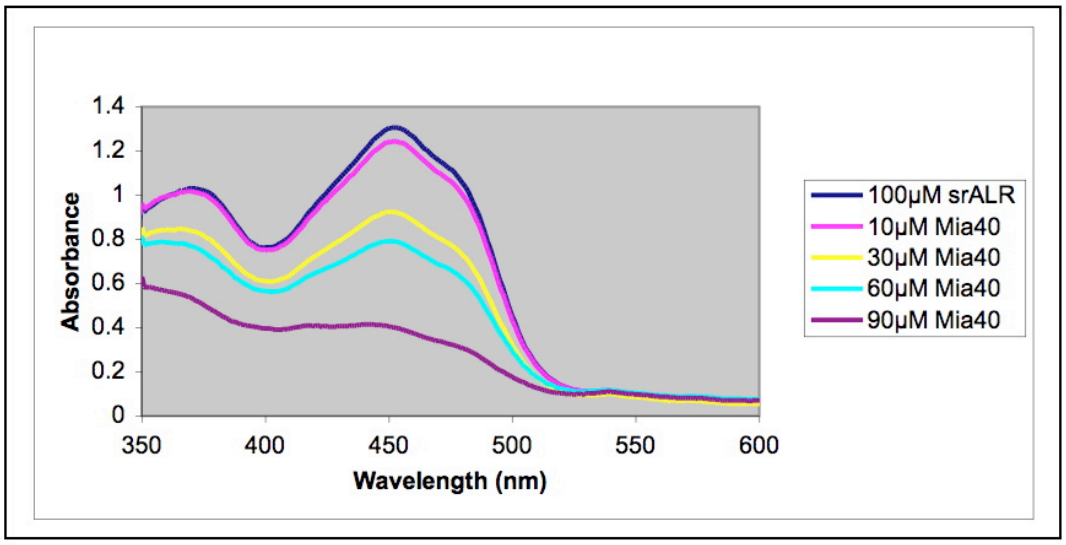


Figure 3.6.1: Alphascreen results for the interaction of srALR with cytochrome c. The fluorescence unit signals observed for srALR and cytochrome c are more than three times higher than the signals observed for control experiments using bovine serum albumin protein.

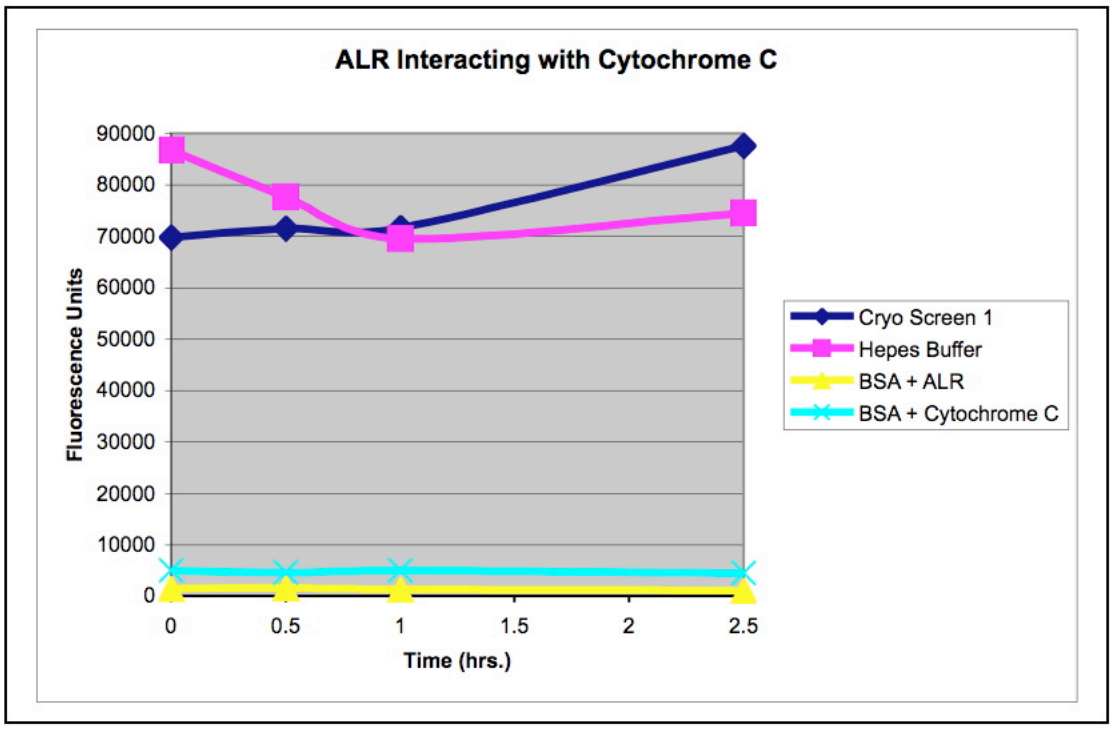


Figure 3.6.2: (Left) Alphascreen results showing the interaction of ALR and cytochrome c along with how the interaction is affected when unbeaded cytochromce c is added. In this condition, the signal is random. (Right) Results showing an alternative condition in which the signal decreases in a linear manner.

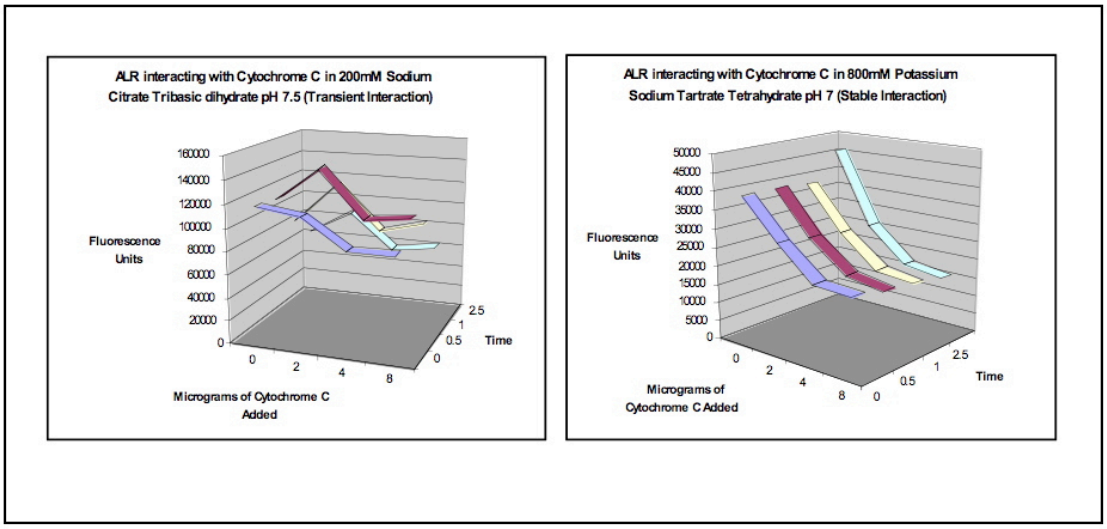


Figure 3.9.1: (Left) Initial crystals of long-form *C. elegans* ALR obtain from a microbatch experiment in 100mM HEPES pH 7.5,10% Isoproponal, 20% PEG 4000; (Right) Optimization of the initial condition in which crystals with better morphology are seen.(100mM HEPES pH 7.3,10% Iso-proponal, 15% PEG 4000)

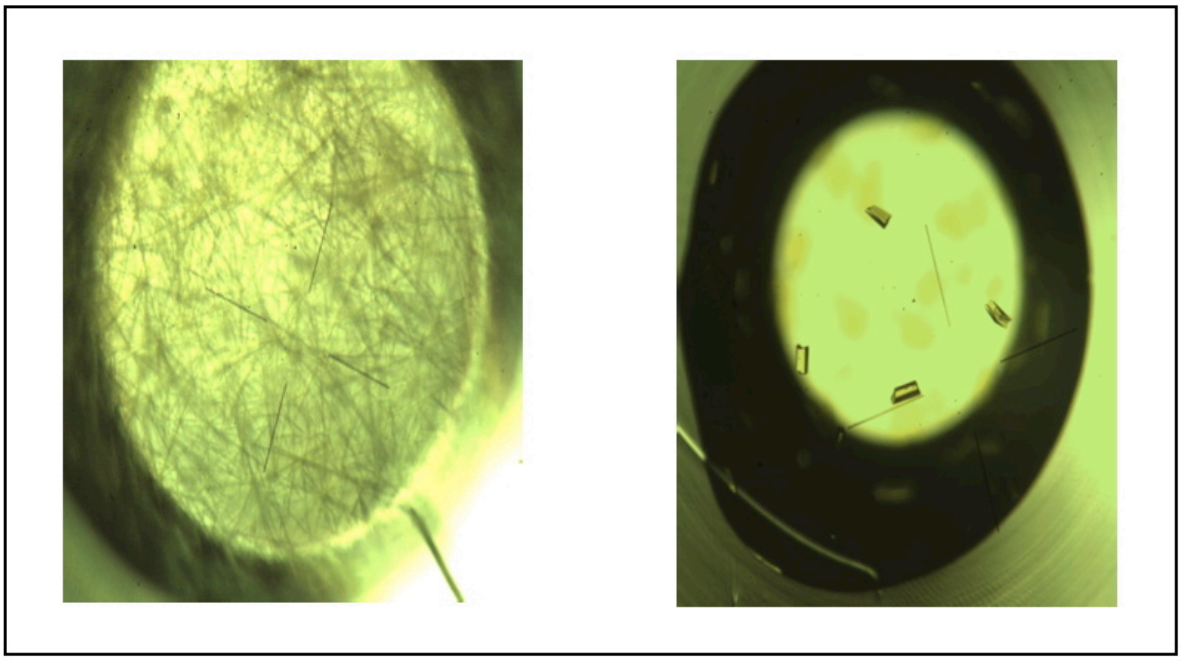


Figure 3.11.1: Adapted from Wang, 1985, this figure is a harker construction for phase calculation by single anomalous scattering (SAS) method.

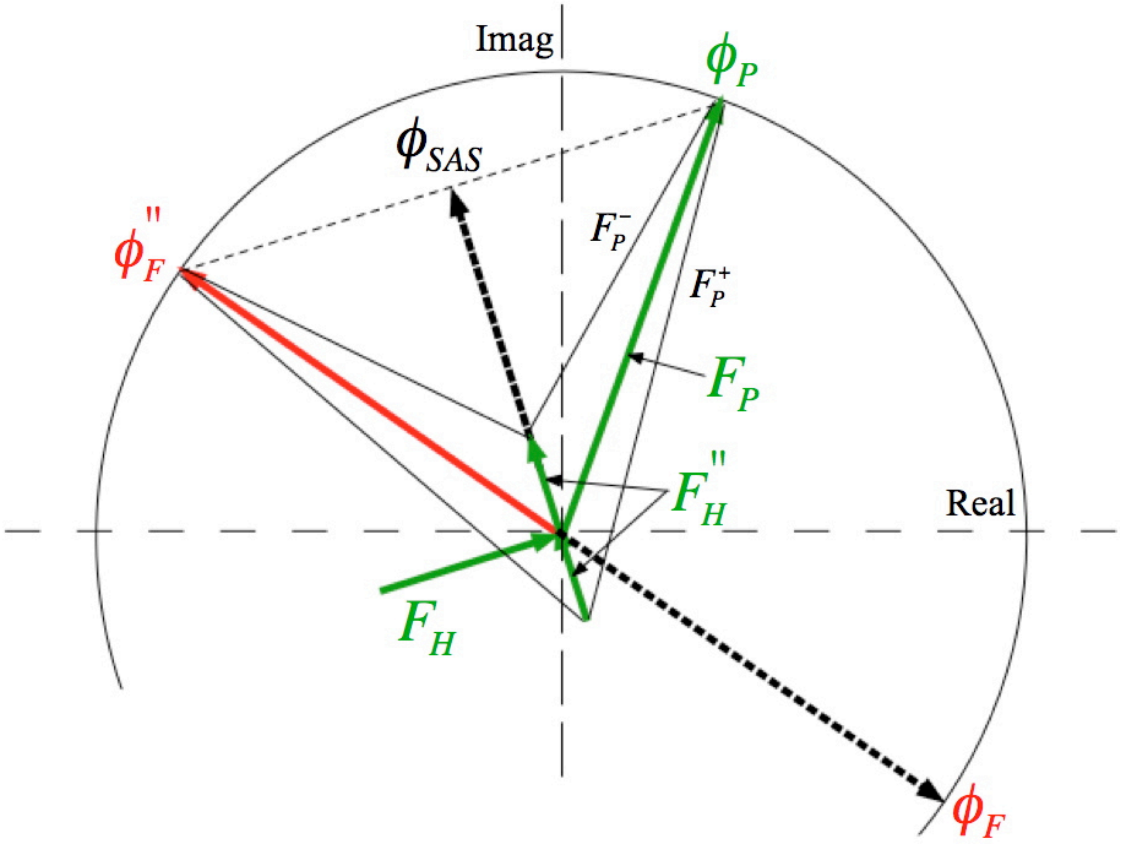


Figure 3.11.2 ISAS Flow Chart: Adapted from Wang's Methods in Enzymology manuscript. The flow chart is a diagram representing the various procedures that are utilized in the ISAS phasing method.

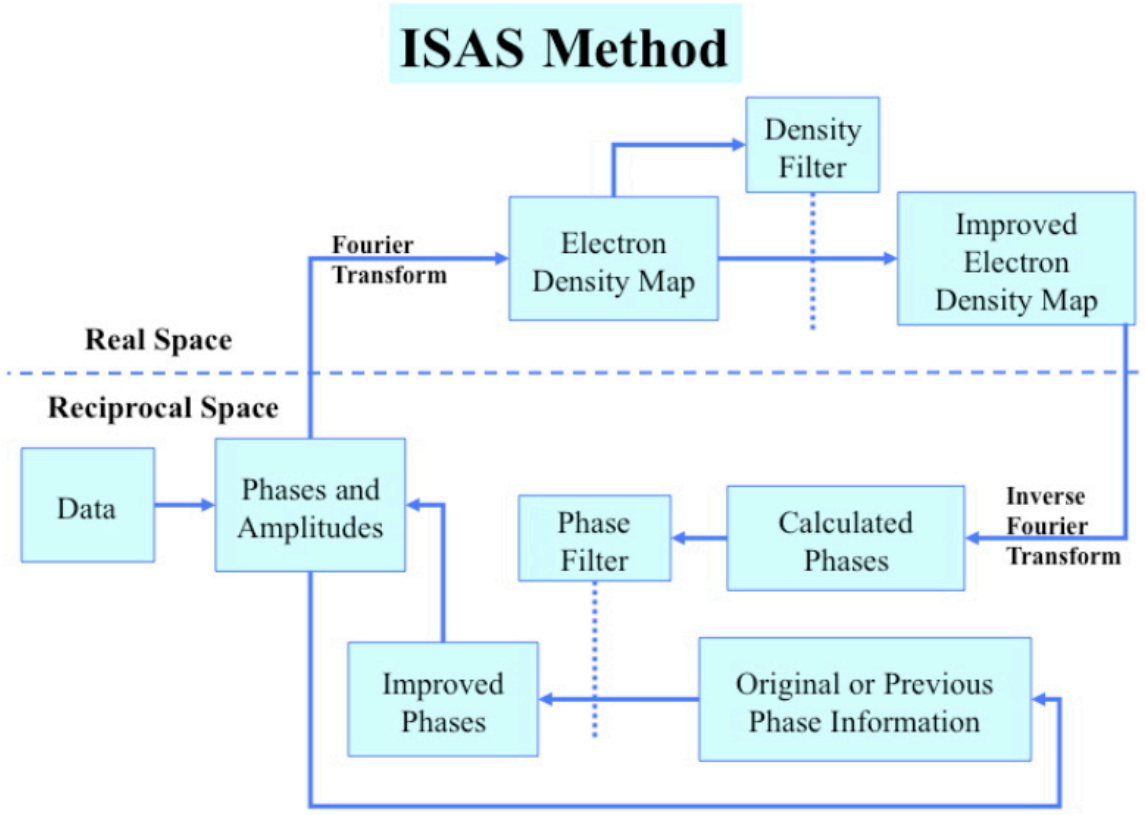


Figure 3.11.3: AutoSol Flow Chart showing the programs that are utilized to determine the initial model. After AutoSol scales the data, phenix.hyss is utilized to locate substructure. Heavy atoms are found and the substructure is phased using Phaser (*101, 102*) or Solve (*98*). Resolve produces the initial model that can then be refined.

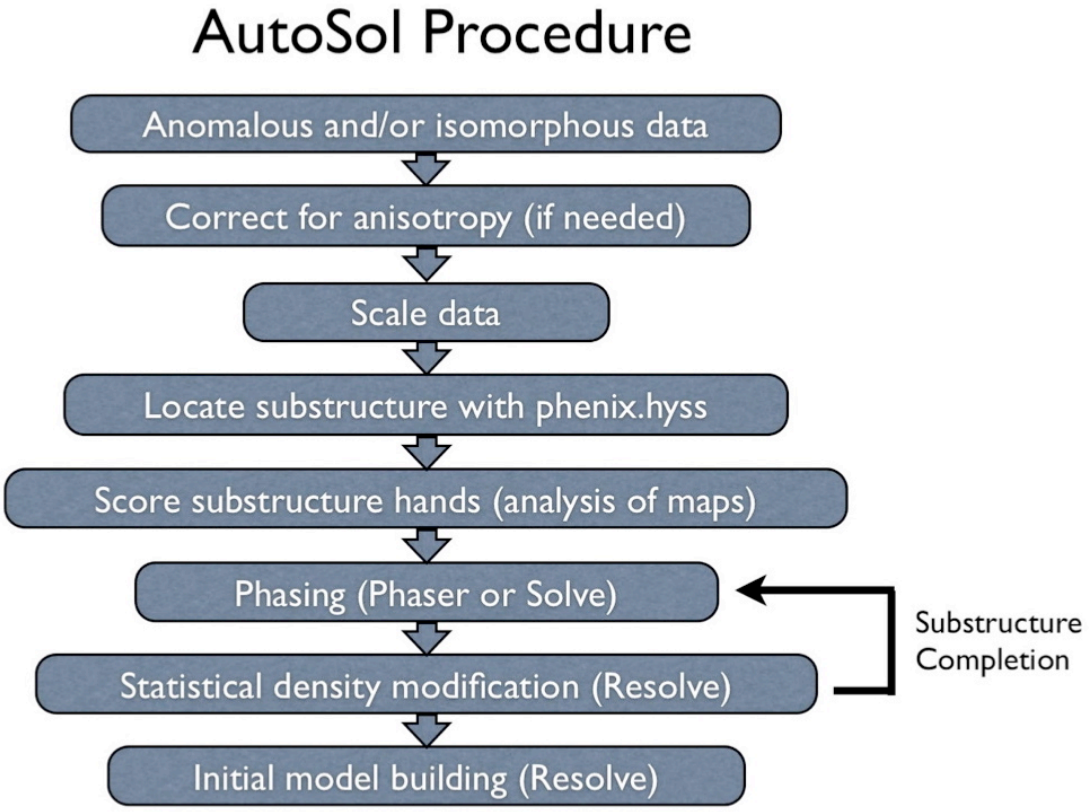
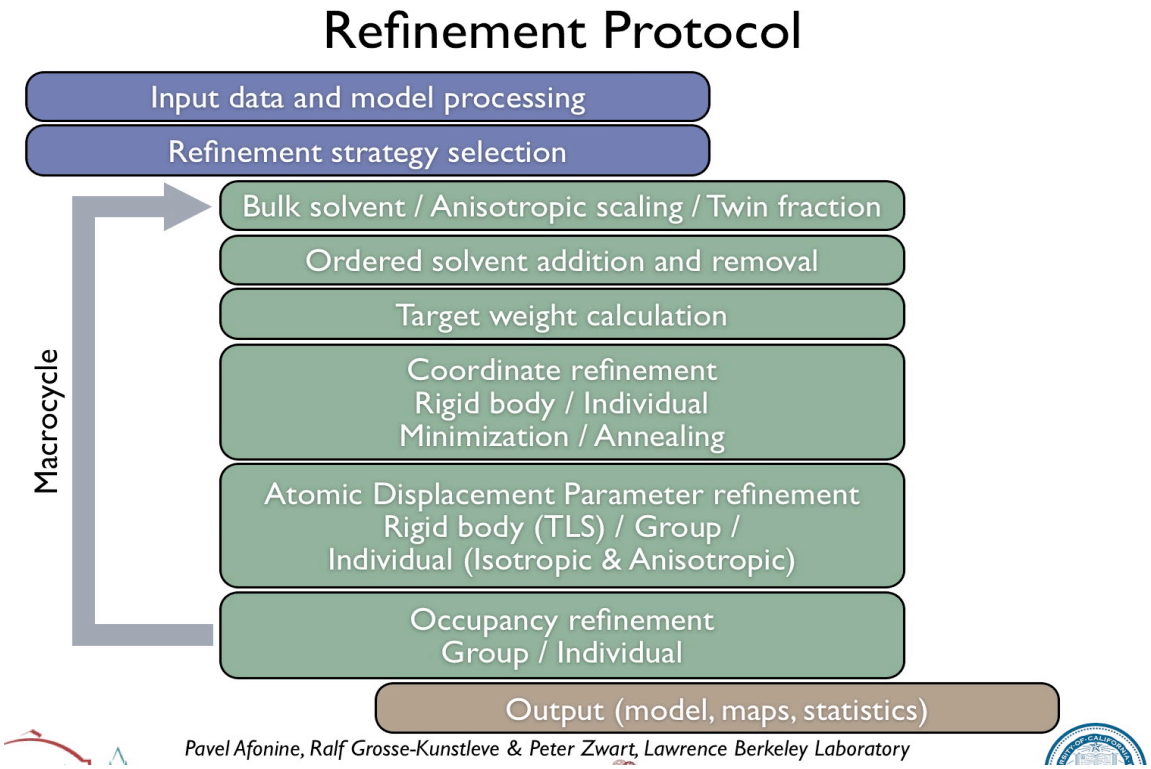


Figure 3.11.4: Phenix.Refine flow chart depicting the procedures that are carried out to refine the structure to a publishable quality. Phenix.refine allows users to choose multiple refinement strategies that refine individual atoms as well as fragments of the structure. The programs refine the structure based on calculated and observed data to produce the highest quality structure.



CHAPTER 4

SOLUTION STRUCTURE ANALYSIS OF THE HEPATITIS B SURFACE ANTIGEN PROTEIN

Hepatitis B virus (HBV) is a blood-born-associated virus that has infected over 2 billion people worldwide. It is also one of the major causes of liver cancer. The overall goal of this study is to derive the 3-dimensional crystal structure of the HBsAg protein, the antigenic component of the currently licensed vaccine for HBV. Despite its worldwide usage and its application as a carrier for epitope presentation, the tertiary structure of this important protein is still unknown. In addition, the lack of a HBsAg structure has (1) prevented us from understanding the effects of mutations on polymerase drug-resistance and on the antigenic properties of the HBsAg major neutralizing antigenic epitope and (2) impeded the development of bivalent vaccines where HBsAg is used as a carrier for foreign antigenic epitopes.

Due to the importance of the HBsAg structure the proposed research falls under a number of the Center for Disease Control's national health protection goals including (1) "Healthy People in Every Stage of Life" where new and better vaccines will impact disease in infants, children, adults and the aged, (2) "People Prepared for Emerging Health Threats" where again new and better vaccines will strengthen our resources to prevent, mitigate and control disease, and (3) "Healthy People in a Healthy World" where better vaccines developed based on the HBsAg structure would greatly impact the health of people in Asia and Africa.

In order to achieve the overall goal of the project, experiments were conducted to: (1) Produce diffraction quality crystals of HBsAg (either the native HBsAg protein or HBsAg – antibody complex) that can be used for X-ray diffraction studies. (2) Carryout single crystal X-ray diffraction experiments on the HBsAg crystals necessary to produce a crystal structure. (3) Refine the HBsAg crystal structure to the highest resolution possible

Crystallization of macromolecules requires homogenous samples of very high purity. However, SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) analysis of the initial HBsAg samples supplied by CDC showed that the protein samples were not pure enough for crystallization trials (data not shown). In addition, the amount of sample was small preventing further purification. The CDC lab then sent two new preparations: one sample purified by ultracentrifugation (HBsAg-UF) and another sample purified by hydrophobic interaction chromatography (HBsAg-HIC). The purity of theses samples was much improved (see Figure 4.1) and was deemed suitable for crystallization trials. The new samples were then concentrated to 6 mg/mL in 10 mM phosphate buffer, pH 8.5 using centrifuge concentrators.

Initial crystallization setups were prepared using the sitting drop vapor diffusion technique and a Cartesian Honeybee crystallization robot according to protocols (*88, 103*) developed by the Southeast Collaboratory for Structural Genomics (SECSG) (*104*). The SECSG screen is based on of eight popular commercial screens: Crystal Screen I & II, Peg/Ion, Cryo, and MembFac from Hampton Research, the Wizard I & II screens from Emerald Biosystems and Memsys from Molecular Dimensions Ltd. An advantage of this approach is that the protein requirements are low (0.384 mg) due to the 400-nanoliter

drop size afforded by the Honey Bee. The disadvantage of using the SECSG screen is that it was not designed for membrane proteins.

As an alternate approach the two HBsAg samples were sent to the High Throughput Screening Laboratory (HTSLAB) at the Hauptman Woodward Medical Research Institute where they were developing a membrane protein crystallization screening service. The HTSLAB uses the micro batch under oil technique (*89*) and automated solution handlers to quickly setup 1536 crystallization experiments in a single 1536 well plate. In the micro batch crystallization setup (see Figure 4.2), the protein and precipitant cocktail solutions (usually in a 1:1 ratio) are injected into one well of the 1536 well plate that has been partially filled with paraffin oil or a paraffin/silicone oil mixture. Once the all wells in the plate have been setup, the plate is then placed in an 18°C incubator. An image of each well in the plate is then recorded periodically over a period of six weeks and made available to the user via the Internet as illustrated in Figure 4.3.

From the HTSLAB screening (24,576 images) six conditions were observed to yield small crystals (see Figure 4.4 and Table 4.1). These crystals, although encouraging, were too small for diffraction analysis and these conditions must be further optimized. From Table 1, magnesium chloride hexahydrate is present in four out the five leads. Based on these results optimization experiments were set up utilizing the five leads that were obtained in from the HTSLAB screens. Optimization screens were setup utilizing the microbatch crystallization method (*105, 106*) and the conditions were optimized around pH, precipitant concentration, and protein concentration. Despite the initial success, these setups did not produce any diffraction quality protein crystals.

During the optimization process, additional crystallization screening was conducted using the HBsAg antibody in complex with HBsAg. Preparation of a protein-antibody complex is a common means of stabilizing a proteins conformation for crystallization trials. For example, the technique has recently been used in the structure determination of the Ebola virus glycoprotein (*107*). The complex was screened using the SECSG screen (*103*) conditions that were used on HBsAg protein. Similarly, these conditions did not lead to the production of diffraction quality crystals. In addition to the SECSG screens, microbatch and hanging drop crystallization experiments were setup-using conditions (MBC Class Suite (Qiagen), Membrane System (Molecular Dimensions), and Lipid Sponge Phase (Jena Bioscience)) that were engineered specifically for membrane proteins. These setups also did not produce any diffraction quality crystals. The failures that were experienced in crystallization trials for HBsAg redirected the immediate goals of the research project.

Although the overall goal of the project is to determine a high resolution X-ray crystal structure of HBsAg, the immediate focus of the project termed to attempting to gain some structural information about HBsAg. In order to accomplish this goal, small angle x-ray scattering experiments (SAXS) were conducted. Small-angle X-ray scattering (SAXS) is a simple experiment that can be used to characterize the solution properties of macromolecules, such as proteins, and their interactions (*78, 79*). Recently, SAXS has recently been used to characterize transient protein-protein complexes in solution (*80, 81*). Unlike in X-ray crystallography, a SAXS profile represents the simultaneous scattering measurement of the macromolecule in all orientations (*82*). The scattering data is transformed from reciprocal space to real space yielding the P(r)-distribution, which is

a histogram of the interatomic vectors within the macromolecule. The P(r)-distribution (pair distance distribution) gives an estimate of the maximum dimension (d_{max}) of the particles in solution (*79*).

In addition, SAXS experiments allow for the calculation of the radius of gyration (R_g) of the molecules in solution. An increase in $R_{g,}$ which provides an indication of the mass distribution of a macromolecule about its center of gravity, is an indication that the macromolecule is opening or unfolding (*82*). Additionally, SAXS data can be used to determine the volume of the scattering particle. Using these SAXS parameters, information can be gained about the dynamics of protein molecules in solution. Additionally, electron density envelopes are generated which give a visual model of the molecules in solution.

SAXS experiments for HBsAg were conducted in collaboration with researchers from Rigaku Americas located in The Woodlands Texas. SAXS data was generated using the Rigaku MicroMax-007 HFM source operating at 1.2 kW. The data was collected on the 200 mm multi-wire 2D detector at a distance of 1.5 m. All experiments solutions were prepared under oxidizing conditions. For the SAXS analyses all proteins (HBsAg and HBsAg antibody) were diluted with 1X Phosphate Buffered Saline (PBS) buffer containing 0.1% sodium azide, 40% glycerol, pH 7.4 buffer to give a 20μL sample with a final protein concentration of 2.875mg $mL^{-1}$and 5mg mL^{-1}. For each protein sample analyzed, a buffer blank was also recorded (same exposure time) for buffer subtraction. For the initial experiments SAXS data (both protein solution and blank) were collected, in duplicate using 30 minute and 60 minute exposure time. As stated in Habel *et al*. (*83*), the buffer-subtracted data were analyzed using PRIMUS (*108*) and the GNOM (*85*). The

P(r) output data was then used to calculate electron-density envelope (see Figure 4.5B). Each envelope is the product of 16 GASBOR (*109*) runs averaged with DAMMAVER (*110*).

As seen by the electron envelopes and the pair distribution plots of HBsAg (Figure 4.5A), the solution contained a 40 nm particle that is present at both concentrations. The HBsAg particle has a real space radius of gyration value of 140.78 ± 0.589. The large particle likely represents multiple protein molecules that are interacting with one another. The Kratky plot (Figure 4.6A), which can be used to look for aggregation and unfolding, reveals a prominent peak that indicates that majority of the protein is in a folded state while it is in solution alone. HBsAg in complex with the antibody Kratky analysis (Figure 4.6B) reveals a prominent peak with a high intensity 'tail' that indicates the complex is a mixture of folded and unfolded protein. The unfolded state of the protein and the binding of the antibody likely contribute to the real space radius gyration value being 146.44 ± 0.987. Electron density envelope structures of HBsAg in complex with the antibody have not been generated due the presence of a large amount of unfolded protein. These experiments reveal that the protein forms large particles in solution that would be difficult to crystallize.

HBsAg is a 24-kDa glycosylated membrane protein so crystallizing it would be challenging under any circumstances. SAXS experiments revealed that HBsAg forms large aggregates in solution that will further complicate the crystallization process. Crystallization is the controlled precipitation of protein molecules and the larger the molecule the more difficult the process becomes. The SAXS studies have provided solution structural information that can be used in the development of better

crystallization experiments and in comparison studies to cryogenic electron microscopy studies that have been done previously (*23, 111-113*). From these studies, the need for an alternative form of HBsAg arose. With this in mind, the future experiments will be focused on producing a version of HBsAg that may be more suitable for crystallization. In conjunction with the production of proteins, surface plasmon resonance studies will be utilized to characterize the kinetics of the HBsAg and antibody complex. These studies combined with the SAXS and crystallization experiments should provide a foundation for the development of better hepatitis b vaccines.

In conclusion, although diffraction quality crystals of HbsAg have yet to be produced knowledge has been gained about the solution and antibody binding properties of HBsAg, which will be invaluable for future research. A paper describing the SAXS studies on HBsAg and its antibody complex is in preparation.

Figure 4.1: SDS-Page analysis of HBsAg-HIC and HBsAg-UF samples.

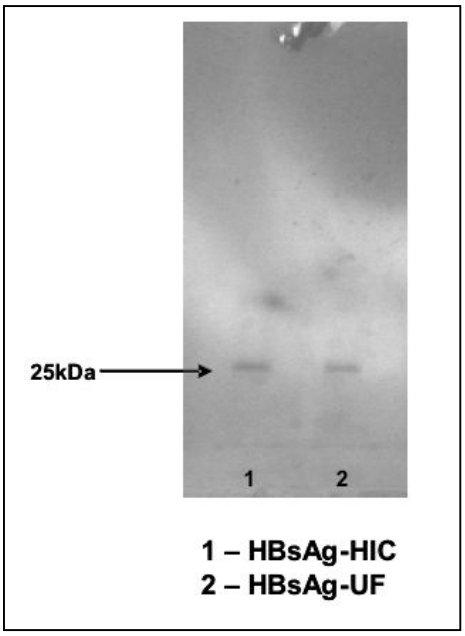


Figure 4.2. (Top) The steps involved the micro batch under oil crystallization setup and (bottom) the 1536 well HTSLAB crystallization plate.

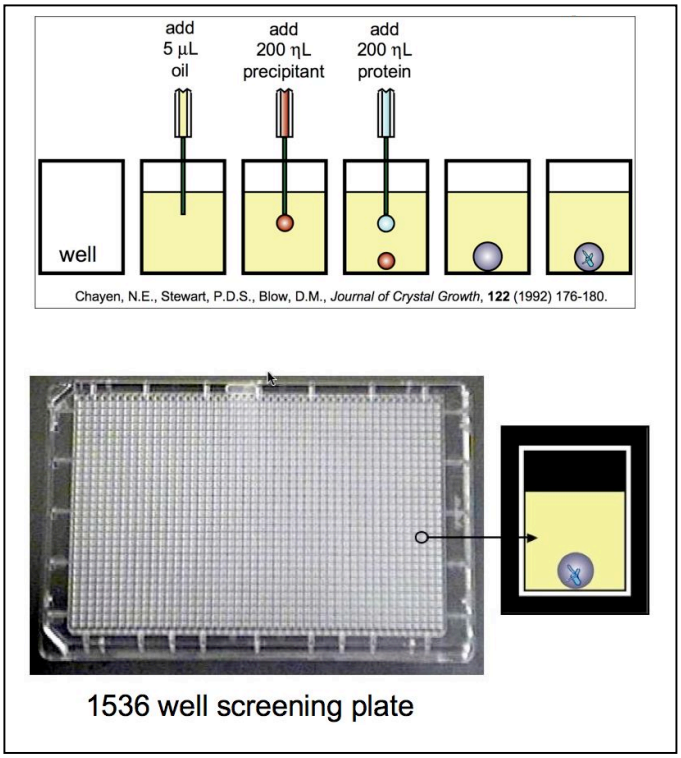


Figure 4.3: Snapshots of an HBsAg-HIC crystallization setup (screen condition #429) recorded at the HTSLAB.

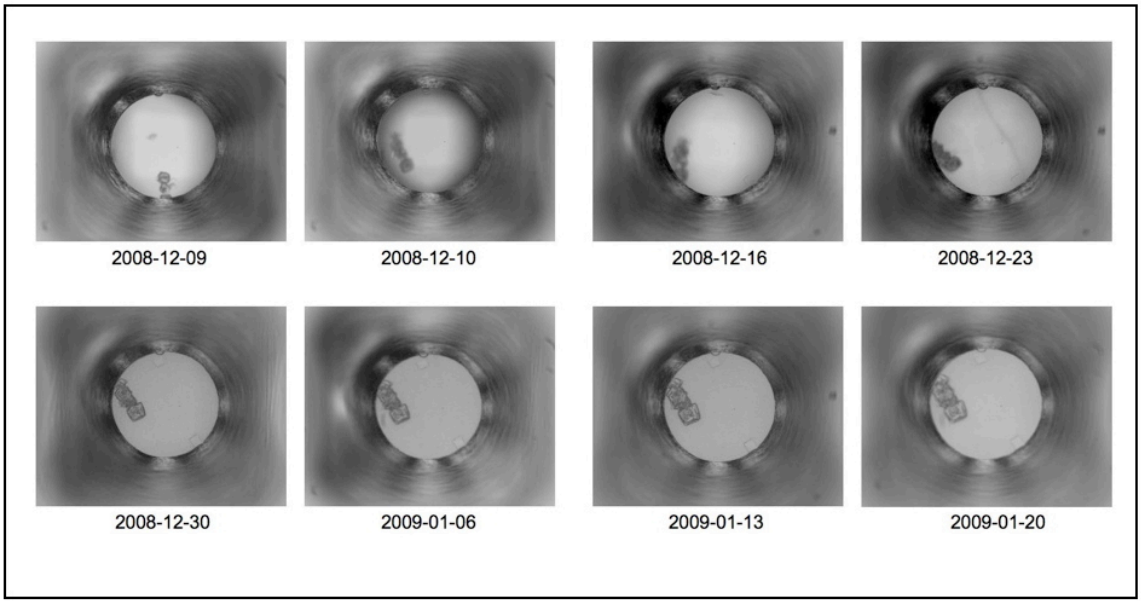


Figure 4.4: Images of the HWI crystallization setups, which produced crystals. Condition #429 produced crystals for both HBsAg-UF and HBsAg-HIC.

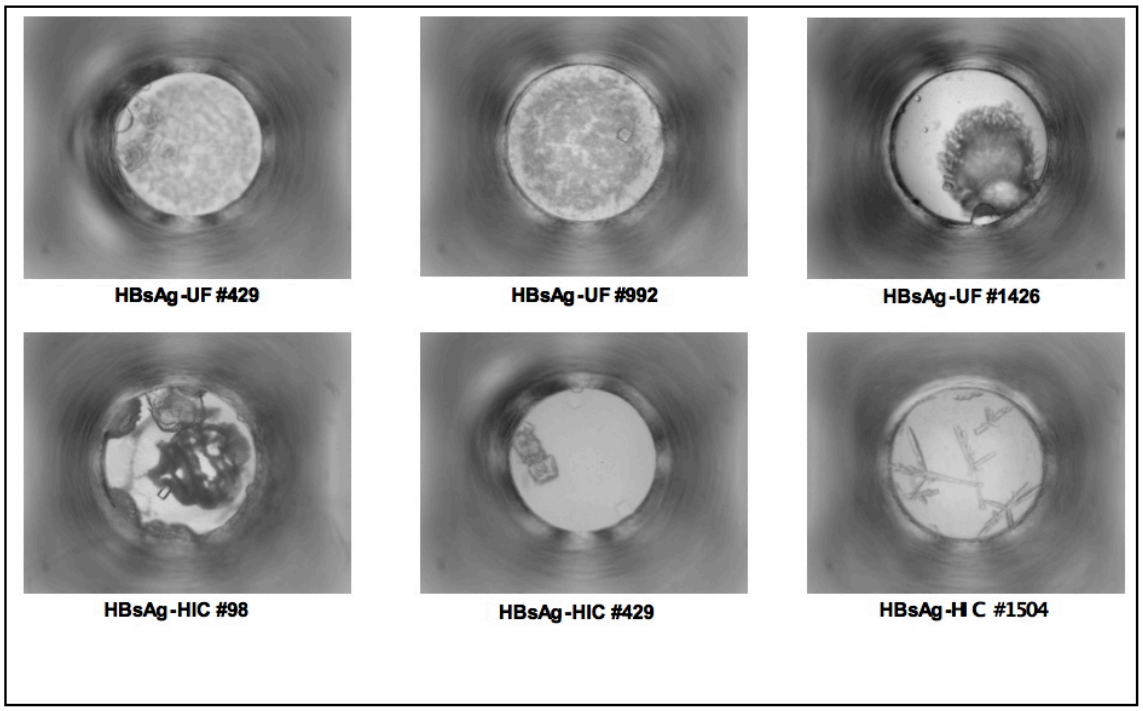
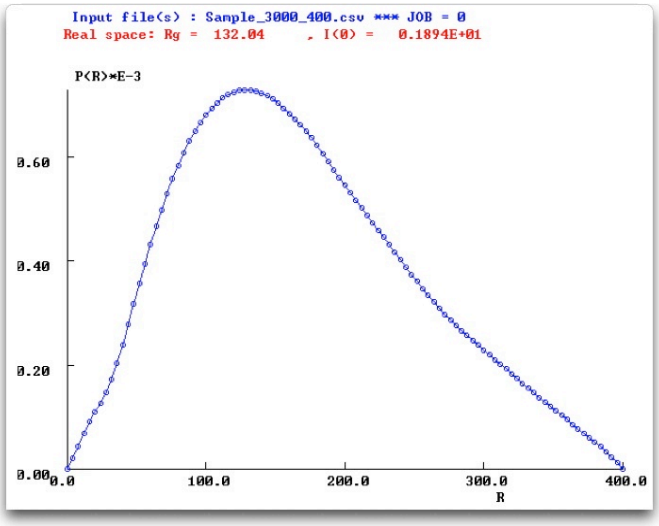


Figure 4.5: A) Pair distribution plot for HBsAg sample showing one large peak that falls back to the x-axis around 40nm. **B)** Electron density envelope model that shows the 40nm particle that is present in solution.

A)



B)

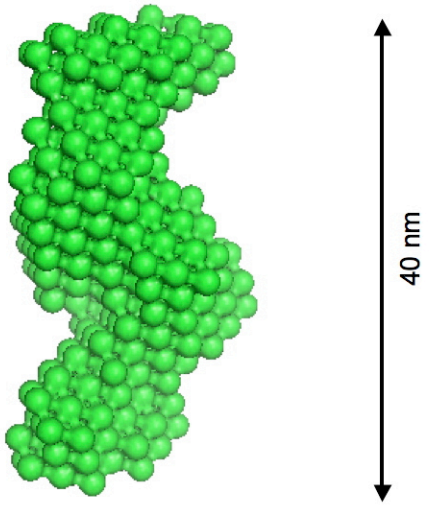
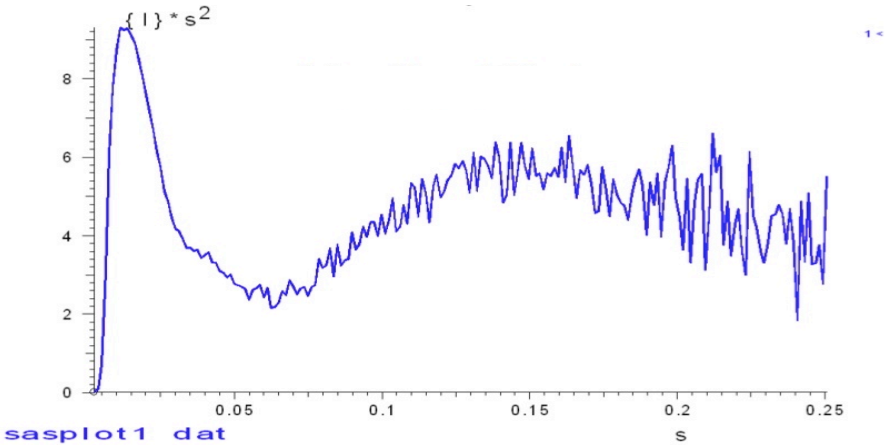


Figure 4.6: A) Kratky plot analysis for HBsAg showing one single prominent peak. This indicated that the predominant molecule in solution is folded. **B)** HBsAg-Antibody complex kratky analysis that shows a peak followed by a high intensity tail that indicates the presence of folded and unfolded protein.

A)



B)

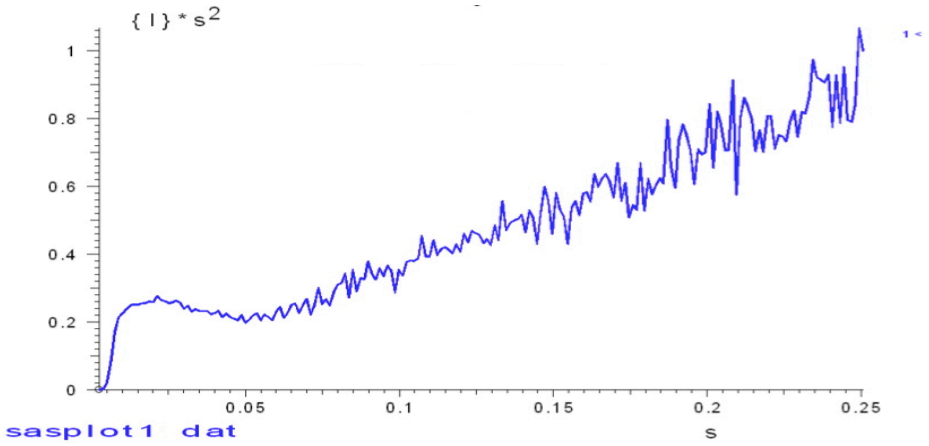


Table 4.1: Conditions that gave crystals from the HTSLAB screen

Screen	Conditions
98	0.2 M magnesium chloride hexahydrate, 0.1 M HEPES in 30% (w/v) PEG 8000, pH 7.0
429	0.4 M magnesium chloride hexahydrate, 0.1 M sodium citrate in 30% (v/v) PEG 400, pH 5.6
992	0.1 M magnesium chloride hexahydrate, 0.1 M HEPES in 8% (w/v) PEG 6000, pH 7.0
1426	0.4 M magnesium chloride hexahydrate, 0.1 M HEPES in 35% (w/v) PEG 3350, pH 7.0
1504	0.2 M lithium sulfate, 0.1 M sodium citrate in 28% (v/v) PEG 400, pH 3.5.
Screen	Conditions
98	0.2 M magnesium chloride hexahydrate, 0.1 M HEPES in 30% (w/v) PEG 8000, pH 7.0
429	0.4 M magnesium chloride hexahydrate, 0.1 M sodium citrate in 30% (v/v) PEG 400, pH 5.6
992	0.1 M magnesium chloride hexahydrate, 0.1 M HEPES in 8% (w/v) PEG 6000, pH 7.0
1426	0.4 M magnesium chloride hexahydrate, 0.1 M HEPES in 35% (w/v) PEG 3350, pH 7.0
1504	0.2 M lithium sulfate, 0.1 M sodium citrate in 28% (v/v) PEG 400, pH 3.5.

CHAPTER 5

THE STRUCTURE OF HEXAHISTIDINE-TAGGED FORM OF AUGMENTER OF LIVER REGENERATION REVEALS A NOVEL [$Cd_2Cl_4O_6$] CLUSTER THAT AIDS IN CRYSTAL PACKING

[1]Florence, Q., Wu, C.K., Habel, J., Wang, B.C. and Rose, J.

To be submitted to *Acta Crystallographica Section D*

Abstract

The crystal structure of the protein augmenter of liver regeneration containing a 14-residue hexahistidine purification tag (hsALR) to aid in purification has been determined to 2.4Å resolution and compared to solution small angle X-ray scattering data. The his-tagged protein, although easier to purify, resisted crystallization compared to the native protein reported previously (Wu *et al.*, 2000). Crystals were finally produced from solutions containing 50 mM $CdCl_2$ and the structure was determined using single wavelength anomalous scattering data collected on a rotating anode home source. The 2.4Å structure has an R_{work} of 0.1859, an R_{free} of 22.70 and good stereochemistry with RMSDs from ideality for bond lengths and bond angles of 0.007Å and 1.147°, respectively. The crystal structure revealed a novel $Cd_2Cl_4O_6$ cluster bridging two hsALR molecules that is believed to promote crystallization. The SAXS low-resolution electron density envelope confirms that hsALR is a dimer in solution.

Introduction

Augmenter of liver regeneration (ALR), as the name implies, was first identified in the cytosol from regenerating rat liver and researchers for over 20 years have been trying to elucidate the mechanism of ALR action at both the cellular and molecular levels. During this time ALR has been shown, in the presence of other factors, to augment the rate of regeneration as indicated by an increased rate of DNA synthesis and cell proliferation (Francavilla *et al*., 1987), to act as an immunosuppressant in reducing the activity of liver-resident natural killer cells (Francavilla *et al*., 1997; Tanigawa *et al*., 2000), to enhance the success rate of fetal pancreas transplantation in rodents (Adams *et al*., 1998), and potentially play an important role in the development and occurrence of herpatocellular carcinoma (Sun *et al*., 2005).

On the cellular level, ALR has been shown to be involved in regulation of mitochondrial gene expression (Polimeno *et al*., 2000), and other cellular processes including the export of iron/sulfur (Fe-S) clusters from the mitochondrial matrix (Lange *et al*., 2001). ALR has also recently been shown been shown to function in the mitochondrial intermembrane space (IMS) pathway important for the import of small cysteine rich proteins, such as the TIM-9 protein, into the IMS (Lisowsky *et al*., 2001; Chacinska *et al*., 2004; Allen *et al*., 2005; Lisowsky, 1996) via possible interactions with the proteins Mia40 (Tokatlidis, 2005) and cytochrome c (Farrell & Thorpe, 2005). This and other data suggest that ALR's overall role *in vivo* is much broader than just as an augmenter of liver regeneration.

ALR-like proteins can be found in organisms ranging from yeast to man. Interestingly, ALR-like proteins have also been observed in viruses (Senkevich *et al*.,

2000). In mammals, ALR exists as two alternatively spliced forms. The full length, 198 residue, 23 kDa form (lALR) that functions in IMS transport and a shorter 125 residue 15 kDa N-terminal truncated form (sALR) that is found in the cytosol during liver regeneration (Klissenbauer *et al*., 2002). The crystal structures of the sALR from rat (*R. norvegicus*) (Rose *et al*., 1999; Wu *et al*., 2003), Erv2, from yeast (*S. cerevisiae*) (Gross *et al*., 2002), an ALR homologue and human ALR (Daithankar *et al*., 2010) revealed a novel all helix FAD (flavin adenine dinucleotide) binding motif which differed considerably from the beta coordinated mode of FAD binding observed in other FAD containing structures.

We have previously reported the expression and purification of sALR containing an N-terminal MGGSHHHHHHGMAS purification tag that, unlike the native protein, required the presence of 50 mM $CdCl_2$ for crystallization (Wu *et al*., 2000). The need for cadmium ions is interesting since divalent cations have been shown to play crucial roles in facilitating protein crystallization (McPherson, 1982, 1990). Studies have shown that cadmium out performs other divalent cations in its comprehensive effect on the protein crystallization (Trakhanov & Quiocho, 1995) with cadmium ions frequently found promoting intermolecular interactions that facilitate molecular packing (Trakhanov *et al.*, 1998; Zanotti *et al.*, 1998) or stabilizing flexible regions of the protein (Qin *et al.*, 2006). Here we report the 2.4Å crystal structure and small angle X-ray scattering analysis of rat sALR protein described by Wu, *et al.* (2000). The crystal structure, determined by Cd-SAD (single wavelength anomalous scattering) revealed a novel $Cd_2Cl_4O_6$ cluster bridging two ALR molecules at a dimer-dimer interface.

Material and Methods

Crystallization and Data Collection: We have previously reported the crystallization and preliminary X-ray characterization of rat short-form ALR containing a hexahistidine N-terminal purification tag (Wu *et al.*, 2000) (hsALR). The crystals were obtained by vapor diffusion using 2 μL hanging drops containing equal amounts of protein (20 mg mL^{-1}) and a precipitant cocktail containing 20% PEG methyl ether (PEGMME) 2000, 0.1 M sodium acetate, pH 4.6 with 50 mM cadmium chloride added just prior to setup.

A data set to 2.4Å consisting of 720 one-degree oscillation images was recorded at cryogenic temperatures (~100K) on a MAR Research 30 cm image plate system using rotating anode generated Cu Ka X-rays. The data were collected using an inverse beam strategy (Naismith *et al.*, 1994) in 30° wedges in an attempt to optimize the cadmium anomalous signal. The data were indexed, integrated and scaled (in 90° batches) using HKL 1.9.1 (Otwinowski & Minor, 1997) to monitor the expected increase in the cadmium anomalous scattering signal, as more observations were added to the analysis, and to help detect/monitor crystal decay. Details of the data collection and processing are collected in Table 1.

Structure Determination and Refinement: Structure determination and refinement was carried out using PHENIX 1.6.4 (Keese *et al.*, 2010). *Phenix AutoSol* was used to determine the anomalous scattering substructure, protein phases and initial fit of sequence into the electron density. *Phenix AutoBuild* was used to improve and extend the initial *AutoSol* model. Manual rebuilding was then carried out using COOT (Emsley & Cowtan, 2004). The resulting model was then refined using *Phenix Refine.* The refined model and structure factors have been deposited in the Protein Data Bank (Berman *et al.*,

2002) as entry XXX. Details of the structure refinement are collected in Table 1.

Small Angle X-Ray Scattering analysis: SAXS data were collected on beamline 12.3.1 (SIBYLS), Advanced light Source, Lawrence-Berkeley National Laboratory. For the SAXS analyses the hsALR protein was diluted with 20mM HEPES buffer containing 150mM NaCl, pH 7.4 to give a 20μL sample with a final protein concentration of 5mg mL^{-1}. Data were collected in duplicate using a 0.5 sec exposure time. Solution blanks containing only buffer, needed for buffer subtraction, were also collected using the same 0.5 sec exposure time. The buffer-subtracted data were analyzed using PRIMUS (Konarev, 2003) and GNOM (Svergun, 1992). Pair-Distribution P(r) data were then used to calculate electron-density envelope. For the envelope calculations 16 GASBOR (Svergun *et al.*, 2001) runs were averaged using DAMMAVER (Volkov, 2003).

Results and Discussion

As previously reported crystals of hsALR belong to group $I4_1$ with cell constants of a = 99.8Å and c = 113.4Å, and contain four molecules per asymmetric unit (Wu *et al.*, 2000). Although a reasonable structure solution could be obtained by molecular replacement, problems during refinement of the structure suggested that there were significant differences between the hsALR structure and the native ALR model (PDB entry 1OQC) used for the molecular replacement analysis. Since an accurate model of hsALR structure was required for ongoing SAXS studies of the ALR-cytochrome c interactions in solution we explored solving the structure using the 720° cadmium SAD data set collected in 1998. Previous attempts at Cd-SAD phasing using this data were not met with success. Although we were able to determine the positions of four anomalous scatterers attempts at producing an interpretable electron density maps failed (Wu *et al.*,

2000).

Phasing and Refinement of the hsALR Structure: For the renewed attempt at Cd-SAD phasing the *Phenix Reflection file editor* was used to convert the SCALEPACK formatted (Otwinowski & Minor, 1997) reflection file to MTZ format (Bailey, 1994) keeping the Bijvoet pairs separate and setting the Rfree flags needed for refinement. The *Phenix AutoSol* wizard was then used to setup and carry out the initial SAD phasing. *AutoSol* was able to identify nine potential cadmium sites, four of which corresponded to the cadmium sites reported previously (Wu *et al.*, 2000). The solution had an Overall Score of 44.98 +/-9.58 and a Figure of Merit of 0.43. Resolve and Buccaneer were used to fit the sequence into the SOLVE phased electron density map (Zwart *et al.*, 2008). The initial model (4 chains, 441 residues, 13 segments and 85 solvent molecules) had an R_{Cryst} of 0.3208, an R_{free} of 0.3487 (Brunger, 1992) (using a 9.31% program generated test set) and a map to model correlation coefficient of 0.72. The model was then rebuilt using *Phenix AutoBuild*. After rebuilding the model (462 residues, 19 segments and 114 solvent molecules) had an R_{Cryst} of 0.2750, an R_{free} value of 0.2950 and a map to model correlation coefficient of 0.74. At this point manual inspection using COOT was carried out. Initial fit of the four FAD (flavin adenine dinucleotide) molecules in the asymmetric unit (one per monomer) was achieved by superimposing the native ALR structure (PDB entry 1OQC) onto the hsALR *AutoBuild* model using CHIMERA (Pettersen *et al.*, 2004). Correct placement of the FAD molecules was then confirmed using 2Fo-Fc electron density maps displayed using COOT. The *AutoBuild* structure showed that there were two different cadmium environments, three $Cd\text{-}X_5$ moieties and what appeared to be a $Cd_2\text{-}X_{10}$ moiety at the hsALR dimer-dimer interface. Bijvoet

difference Fourier analyses based on phases from the partial hsALR model indicated that four of the atoms in the Cd_2-X_{10} moiety were most probably chlorine since there was a anomalous scattering signal associated with these atoms and chloride was present in high concentration in the crystallization setup. Chloride has a small but measurable (Δf"= $0.84e^-$) anomalous scattering signal using Cu Kα X-rays. Four of the remaining Cd_2-X_{10} ligands came from the side chains of Glu A 74 (OE1 and OE2) and Glu B 74 (OE1 and OE2). The remaining two Cd_2-X_{10} ligands were modeled as water since they lacked anomalous scattering signal.

The manually rebuilt ALR model, including the four FAD molecules, three CdO_5 moieties and the $Cd_2Cl_4O_6$ cluster and 253 solvent molecules, was then refined using *Phenix Refine*. Experimental phases were used during all stages of refinement. Refinement of the occupancies for the cadmium, chlorine ions and oxygen atoms associated with the CdO_5 and $Cd_2Cl_4O_6$ moieties helped confirmed their identity. After several rounds of refinement (positional and atomic displacement parameters) a TLS model (three segments per chain) was calculated using the TLSMD server (Painter & Merritt, 2006) and included in the refinement. The final refined hsALR model contained residues 16-119 of monomer A, residues 16-122 of monomer B, residues 16-118 of monomer C, residues 16-62, 76-115, four FAD molecules, two CdO_5 cluster, a CdO_6 cluster, a $Cd_2Cl_4O_6$ cluster and 255 solvent molecules modeled as water. The 2.4Å refined model has an Rwork of 0.1882, an RFree of 0.2270 and good stereochemistry with RMSDs from ideality of 0.007Å for bond lengths, 1.147° for bond angles.

The hsALR structure: Similar to the rat sALR structure reported previously (Wu *et al.*, 2003), hsALR is a cone shaped five-helix antiparallel helical bundle with dimensions of

~32 x 26 x 20Å and a surface area of 5890 Å^2. The FAD cofactor sits at the mouth of the cone with N5 of the flavin mononucleotide moiety (FMN) located in close proximity to the Cys 62 – Cys 65 disulfide bond forming the enzyme's active site see Figure 1a,b.

The observed polypeptide chain begins at residue 16 located at the mouth of the cone and begins a short loop ending at α1 (residues 18–36). A β-turn (residues 37–41) follows and forms the tip of the cone. Helix α2 (residues 42–60) running antiparallel to α1, returns the chain back to the mouth of the cone. The region spanning residues 60-68 shows the most variability among the four molecules that comprise the crystallographic asymmetric unit as illustrated in Figure 1c. For chains A and C, a long loop connecting α2 to α3 (residues 68–76) is observed. Chain B exhibits a much shorter loop (residues 60–63) and helix α3 lengthens to span residues 63–76 that closely resembles the loop observed in the sALR structure (PDB entry 1OQC). Helix α3 is not observed in chain D since the electron density for residues 63–76 was not observed. The chain continues after α3 and loops back to the bottom of the cone where it enters helix α4 (residues 82–102). Helix α4 then brings the chain back to the mouth of the cone where it loops back and enters helix α5 (residues 107–116). After helix α5 the chain forms a loop that extends away from the cones surface.

Overall the four-hsALR monomers in the crystallographic asymmetric unit have a similar structure giving an RMSD of superposition (Meng *et al*., 2006) ranging from 0.38Å (superimposing chains A and B) to 0.62Å (superimposing chains A and B) with most of the variation observed in the loop connecting helices α2 and α3. A similar comparison to the Erv2 (PDB entry 1JR8) and human ALR (PDB entry 3MBG) structures again shows a very similar fold giving an RMSD of superposition ranging from of 0.274Å (superimposing chains hsALR-B and sALR-C) to 0.989Å (superimposing

chains hsALR-A and Erv2-A). Again, most of the variation observed is due to the unwinding of the loop connecting helices α2 and α3 in chains A and C of the hsALR structure.

The spatial orientation of the FAD molecule places its flavin mononucleotide (FMN) moiety at the mouth of the cone packed against helices α1, α2, and α3. Interactions between the FMN moiety and hsALR are for the most part hydrophobic. The FAD adenosine monophosphate (AMP) tail loops between helices α1 and α4, and packs against helices α4 and α5. Hydrogen bonding between the hsALR and the AMP tail anchors the FAD to the protein. The FAD conformation is very similar for the four FAD molecules observed in the hsALR structure. The four FAD molecules can be superimposed to give an average all atom RMSD 0.17Å (Pettersen *et al.*, 2004).

As noted above, the hsALR structure shows the most variability in the loop connecting helices α2 and α3 (residues 60 – 68), which defines the ALR active site. In addition, compared to the sALR structure there are fewer FAD - hsALR hydrogen bonds. In the sALR structure eight FAD - protein hydrogen bonds are observed. In the hsALR structure, between three to six FAD - protein hydrogen bonds are observed, depending on which of the hsALR monomers is used. FAD hydrogen bonding to residues in helix α4 (Cys 91, Asn 94, Asn 98) observed in sALR is not observed in the hsALR structure. However, even though the FAD binding environment is considerably different between the hsALR and sALR structures the FAD conformation observed in the two structures is remarkably similar giving an average RMSD of superposition (all atoms) of 0.21Å (Pettersen *et al.*, 2004). The variability of the FAD binding pocket and the structural variability observed for the 60 - 65 loop observed in the hsALR structure may be an

indicator of the dynamics of the ALR catalytic site needed to accommodate potential oxidoreduction partners.

The hsALR dimer: Similar to sALR, the his tagged ALR protein forms a head-to-tail disulfide linked dimer where Cys 15 and Cys 124 of one monomer forms a disulfide bond with Cys 124' and Cys 15', respectively with its dimer partner. The intermolecular disulfide bonds however are not critical for dimer formation as evidenced by the structure of Erv2 (Gross *et al.*, 2002), which lacks the N- and C-terminal cysteines required for intermolecular disulfide bond formation. In the Erv2 structure helix packing is the driving force for dimer formation. In addition, C15S and C124S mutagenesis studies show that intermolecular disulfide bond formation is not required for ALR dimer formation (Wu *et al.*, 2003) and the loss of the intermolecular disulfide bond does not significantly impact ALR oxidase activity (T. Dailey personal communication).

There are two ALR dimers in the crystallographic asymmetric unit, the A:B dimer composed of chains A and B and the C:D dimer composed of chains C and D. The A:B dimer interface (1210 $Å^2$) involves the interaction of 30 residues and is stabilized by hydrogen bonding (OD1 Asn 25A – ND2 Asn 25B and OE1 Gln 51A - NE2 Gln 51B) and two salt bridges (OE2 Glu 44A – NZ Lys 58B and OD2 Asp 48A – NZ Lys 58B) (Krissinel & Henrick, 2007). The C:D dimer interface (1188 $Å^2$) involves the interaction of 28 residues and is stabilized by hydrogen bonding (ND2 Asn 25C - OD1 Asn 25D, OD2 Asp 48C - NZ LYS 58D, OE1 Gln 51C - NE2 Gln 51D, NZ LYS 58C - OD2 Asp 48D) and two salt bridges (OD2 Asp 48C – NZ Lys 58D and NZ- Lys 58C – OD2 Asp 48D).

The ALR tetramer: Similar to the sALR structure the hsALR crystallographic

asymmetric unit contains a tetramer. The tetramer is made up of the A:B and C:D hsALR dimers bridged by the novel $Cd_2Cl_4O_6$ cluster shown in Figure 2a. The tetramer exhibits local-two fold symmetry centered at the midpoint of the $Cd_2Cl_4O_6$ cluster. Each cadmium ion of the cluster is coordinated by three chloride ions and three oxygen atoms as shown in Figure 2b. In the cluster, CD 1E is bound to chlorides CL 1E, CL 2E and CL3E, the side chain oxygen's (OD1 and OD2) of ASP 74C and a distal water molecule. Similarly, CD 2E is bound to chlorides CL 1E, CL 2E and CL4E, the side chain oxygen's (OD1 and OD2) of ASP 74A and a distal water molecule. The average bond lengths observed for atoms in the cluster are 2.61Å for the Cd – Cl bonds, 2.46Å for the Cd – O bonds and 3.82 for the Cd – Cd bond, which is in agreement with published data (Hines *et al.*, 2006). The distance between the CG atoms of the two bridging Asp 74 groups is 6.75Å. As a result, the cluster provides for long-range interaction between the two-hsALR dimers via coordination with Asp 74. A comparison of the hsALR and sALR structures shows that residues His 54 and Lys 58 have moved to accommodate the bulky cluster and allow coordination with Asp 74. The interaction of Asp 74 with the cluster may also stabilize the open conformation observed for the 60 - 68 loop in monomers A and C.

In addition to the $Cd_2Cl_4O_6$ cluster the dimer-dimer interface (554Å^2) involves the interaction of 17 residues and is stabilized by extensive hydrogen bonding (OD1 Asp 48A - NH1 Arg 75C, OD1 Asp 48A – NH2 Arg 75C, OE1 Gln 51A – NH1 Arg 75C, OE1 Gln 77A – NE2 His 54C, O Pro 78A – NE2 Gln 77C, NE2 His 54A – OE1 Gln 77C, NH1 Arg 75A – OD1 Asp 48C, NH1 Arg 75A – OE1 Gln 51C and NE2 Gln 77A – O Pro 78C) and four salt bridges (OD1 Asp 48A – NH1 Arg 75C, OD1 Asp 48A – NH2

Arg 75C, NH1 Arg 75A – OD1 Asp 48C and NH2 Arg 75A – OD1 Asp 48C).

The hsALR tetramer differs significantly from the tetramer observed in the sALR structure. In the sALR structure, the two ALR dimers associate in a manner where the N-terminal regions of the dimer face each other as shown in Figure 2c. One can envision that the addition of the 14-residue histidine tag to the N-terminus of the protein could potentially disrupt the dimer-dimer interface and crystal packing. This would explain why the hsALR protein failed to crystallize using the sALR conditions. In the hsALR structure the two ALR dimers, bridged by the $Cd_2Cl_4O_6$ cluster, associate in a way that places the N- and C-terminal residues facing away from the dimer-dimer interface. Thus, the addition of an N-terminal 14-residue histidine tag would be easily accommodated in the lattice. In addition to the $Cd_2Cl_4O_6$ cluster, three other Cd moieties are observed in the hsALR structure. A CdO_5 cluster associated with Glu 113C, a CdO_6 cluster associated with residues Glu 44C and Gln 47C, and a CdO_5 cluster associated with residues Asp 107B and Ser 109B. However, these clusters are not involved in crystal packing interactions.

hsALR SAXS analysis: From the SAXS analyses, the Kratky plot shown in Figure 3a has one prominent peak indicating that the protein is folded. The pair-distribution (P(r)) distribution function (Figure 3b) indicated that the real space radius of gyration for hsALR was 23.26 nm, which corresponds to a particle diameter of approximately 90Å. Figure 3c shows the fit (Pettersen *et al.*, 2004) of the hsALR dimer into the electron-density envelope. The electron density envelope was the average of 16 individual pair-distribution function calculations. The extra density observed in the envelope can be attributed to the 27 N-terminal amino acid residues that are missing from each of the

hsALR monomers in the crystal structure.

In conclusion, the structure of hsALR has revealed a novel $Cd_2Cl_4O_6$ cluster at the dimer-dimer interface, which enables long-range protein-protein interaction and significantly alters the dimer-dimer interaction observed in the sALR structure. The new mode of dimer-dimer association induced by $Cd_2Cl_4O_6$ cluster the can now accommodate the extra residues associated with the N-terminal hexahistidine purification tag and explains why hsALR failed to crystallize using the sALR crystallization protocol. The structure also provided a view of the dynamic nature of the $\alpha 2 - \alpha 3$ loop and the catalytic site that may provide insight into substrate binding. The results of the SAXS studies of hsALR in solution show the expected hsALR dimer. The low-resolution electron density envelope calculated from the SAXS pair-distribution function is able to accommodate the hsALR dimer including the extra 14 residue N-terminal histidine tag. The SAXS studies also showed that there is little structure to the first 27 residues of hsALR which agrees with the results of the crystal structure determination.

Table 5.1. Data collection and refinement details

Crystal	
Space group	$I4_1$
a (Å)	99.8
c (Å)	113.4
Data collection	
Source	Rigaku RU200
Detector	MAR Research 30 cm Image Plate
Wavelength (Å)	1.5418
Distance (mm)	175
2θ	0.0
Phi step (°):	1.0
Total Rotation (°)	720
Data processing:	HKL 1.9.1
Resolution (Å)	50 - 2.3
Completeness (%)	100 (100)
R_{sym}*	0.089 (0.425)
*outer shell (2.38 - 2.30 Å) values in parentheses	
Refinement	
Program	Phenix 1.6.4
Resolution (Å)	19.64 - 2.40
Completeness* (%)	99.28
R_{cryst}*	0.1859 (0.2043)
R_{free}*	0.2270 (0.2879)
*outer shell (2.49 - 2.40 Å) values in parentheses	
R.M.S. Deviations from ideality	
Bond lengths (Å)	0.007
Bond angles (°)	1.147
*Ramachandran analysis***	
Most favored (%)	97.90 (100.0)**
Disallowed (%)	0.0
**PROCHECK percent in all allowed regions shown in parentheses	
Final model	
Protein atoms	3334
Heterogen Atoms	232
Solvent atoms	255
PDB ID	XXX

Figure legends

Figure 5.1A: The hsALR monomer showing the bound FAD molecule and the Cys 62 – Cys 65 disulfide, components of the putative catalytic site. Image generated with CHIMERA (Pettersen *et al.*, 2004). The molecule is colored blue to red to denote sequence position.

Figure 5.1B: A map observed secondary structure for chain A on to the hsALR sequence generated by PDBSum (Laskowski *et al.*, 2005). Helices are denoted as magenta coils. Beta and gamma turns are highlighted by β and γ, respectively. Disulfide bond connections are highlighted in yellow. Residue contacts are denoted by red dots. Cadmium contacts are denoted by blue dots.

Figure 5.1C: A view (cross eye stereo) of the overlap of the four hsALR monomers in the crystallographic asymmetric unit generated by CHIMERA. The chains are colored as follows Chain A – white, chain B – magenta, chain C – cyan and chain D yellow. The variability in structure associated with the a2 – a3 loop (residues 60 – 68) can be seen at the bottom right of the structure.

Figure 5.2A: A view of the hsALR tetramer showing local 2-fold symmetry and the bridging $Cd_2Cl_4O_6$ cluster displayed with CHIMERA. The A:B and C:D dimers are colored as follows: Chain A – yellow, chain B – cyan, chain C – green and chain D – pink. Atoms making up the $Cd_2Cl_4O_6$ cluster are colored: Cd – yellow, Cl – green, O – red, C white.

Figure 5.2B: A detailed view of the $Cd_2Cl_4O_6$ cluster. Here each cadmium ion is hexadentate coordinating three chloride ions, one water and two oxygen atoms from Asp 74.

Figure 5.2C: A view of the sALR tetramer (PDB entry 1OQC). The two sALR dimers are colored yellow and cyan and tan and blue, respectively. Note that in the tetramer the two sALR dimers are positioned such that the N-termini of two of the sALR molecules are pointing into the dimer-dimer interface of the tetramer.

Figure 5.3A: A Kratky Plot revealing a single prominent peak that indicates that hsALR is folded in solution.

Figure 5.3B: Results from the pair-distribution calculations a P(r) plot, which shows that the real space radius of gyration for hsALR in solution is 23.26 nm. In addition, the curve also reveals that the diameter of the molecule is approximately 90Å.

Figure 5.3C: A fit (Pettersen *et al.*, 2004) of the hsALR dimer into the electron-density envelope calculated from an average of 16 GASBOR runs with P1 symmetry.

Fig. 5.1A

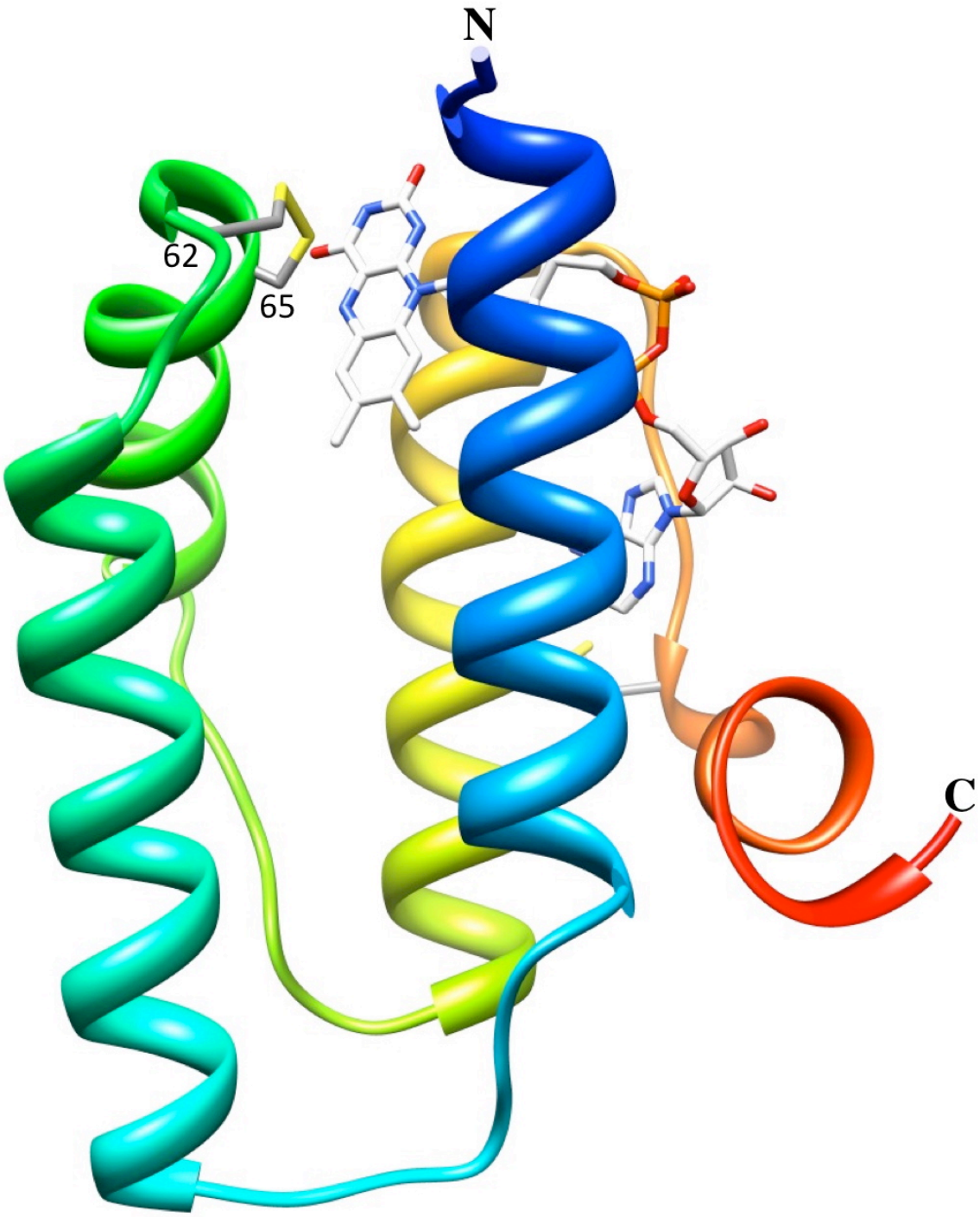


Fig. 5.1B

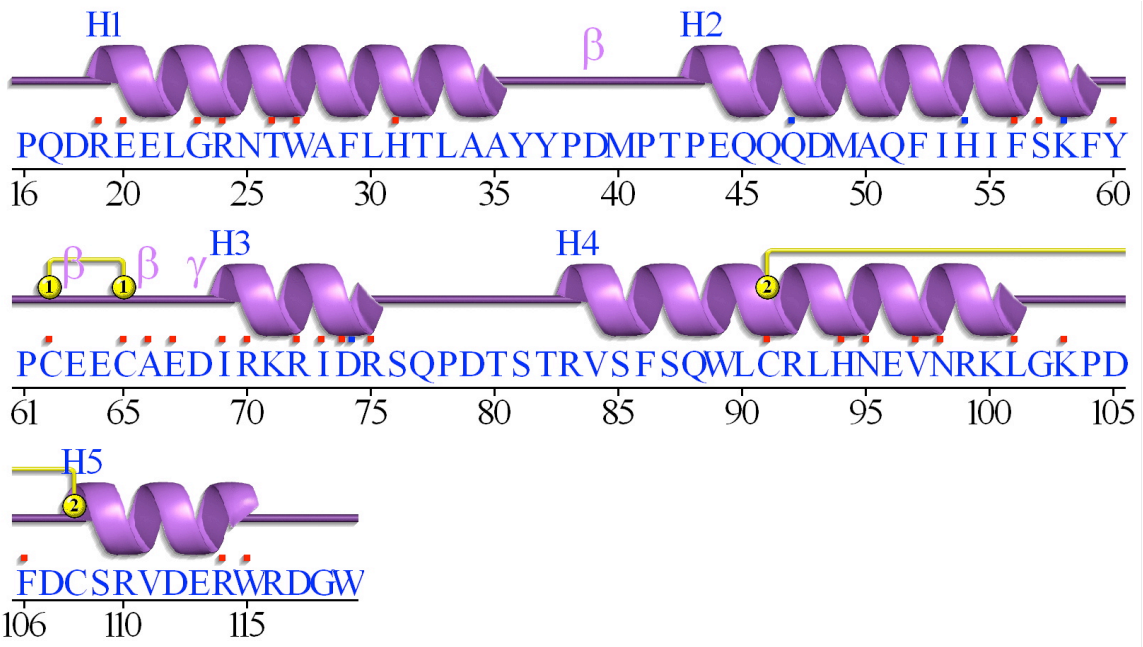

H1
β
H2
PQDREELGRNTWAFLHTLAAYYPDMPTPEQQQDMAQFIHIFSKFY
16 20 25 30 35 40 45 50 55 60
β β γ
H3
H4
PCEECAEDIRKRIDRSQPDTSTRVSFSQWLCRLHNEVNRKLGKPD
61 65 70 75 80 85 90 95 100 105
H5
FDCSRVDERWRDGW
106 110 115

Fig. 5.1C

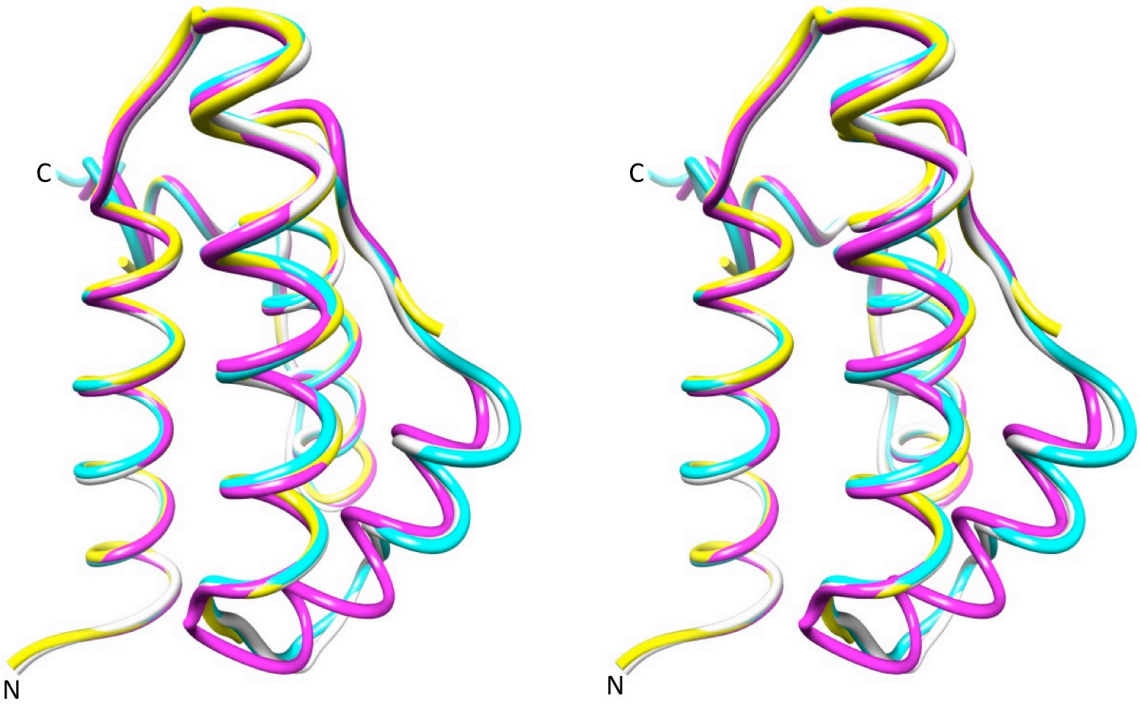


Fig. 5.2A

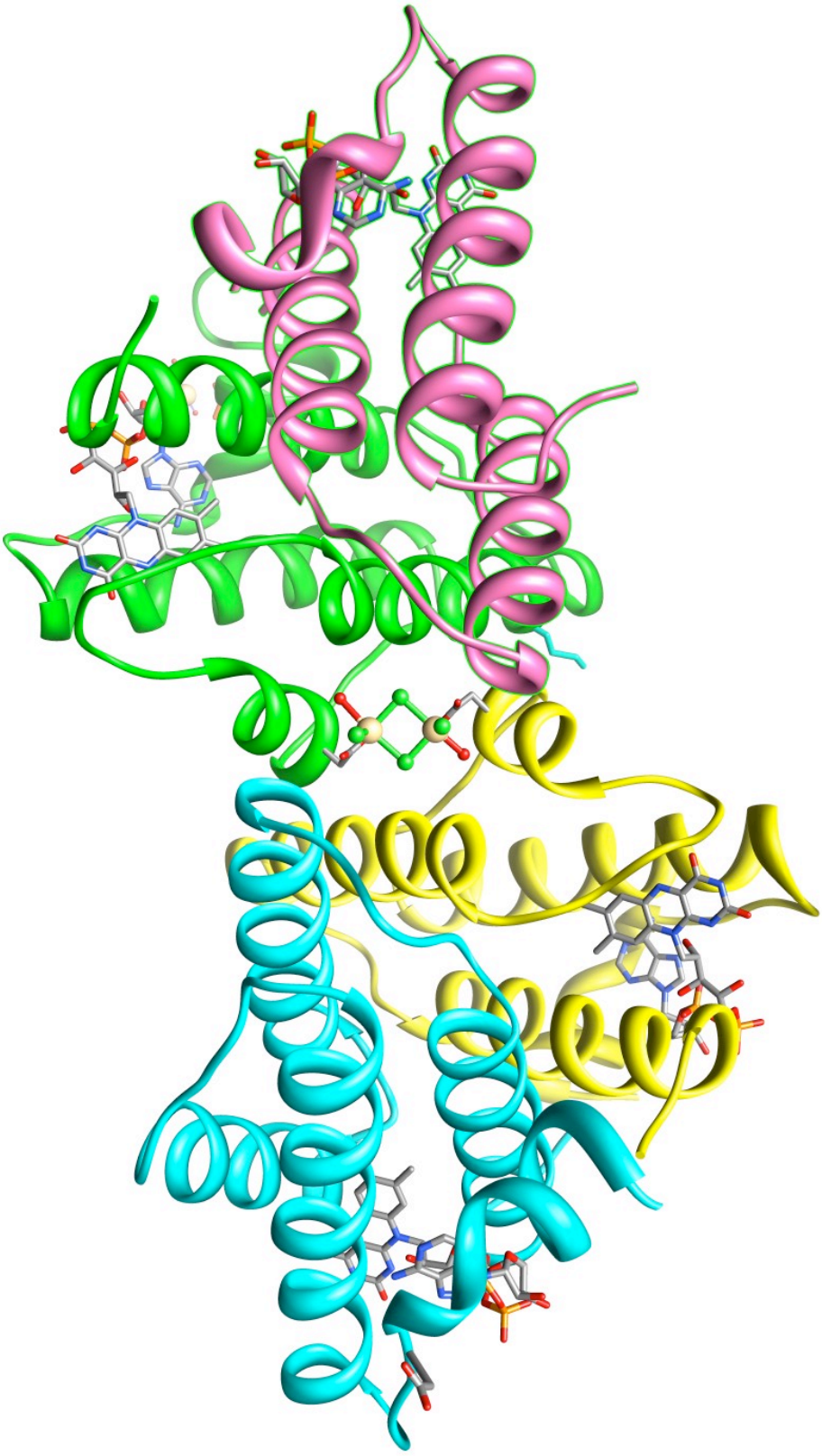


Fig. 5.2B

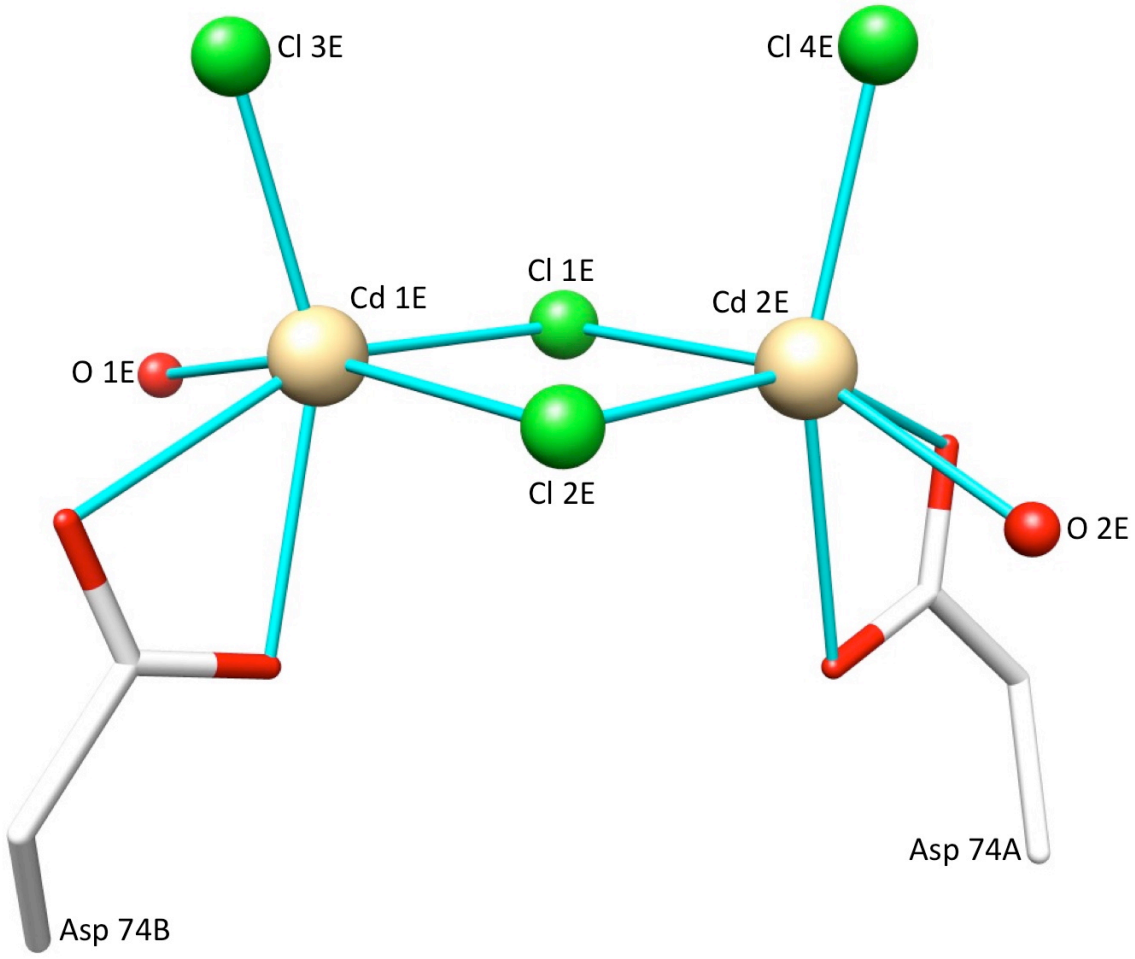


Fig. 5.2C

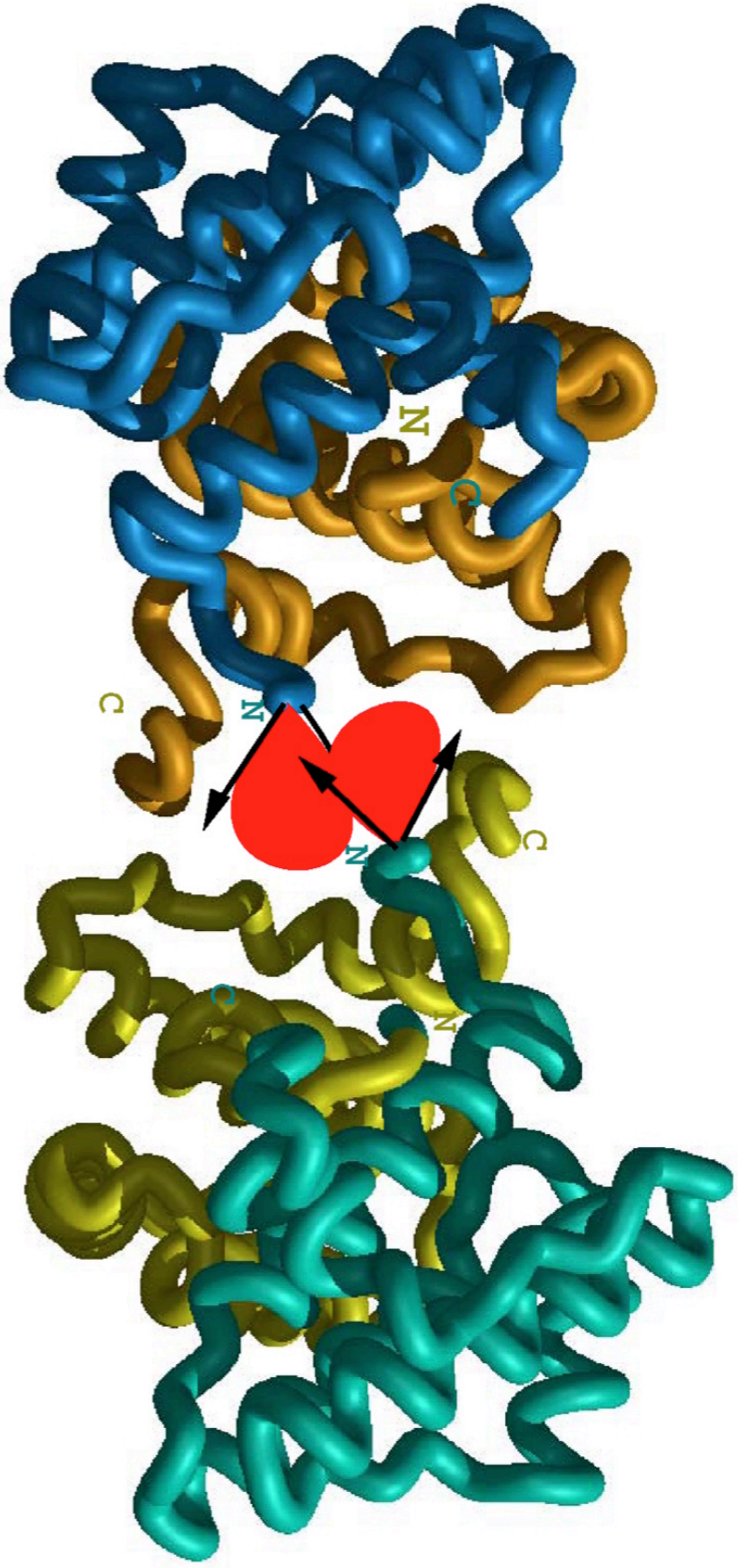


Fig. 5.3A

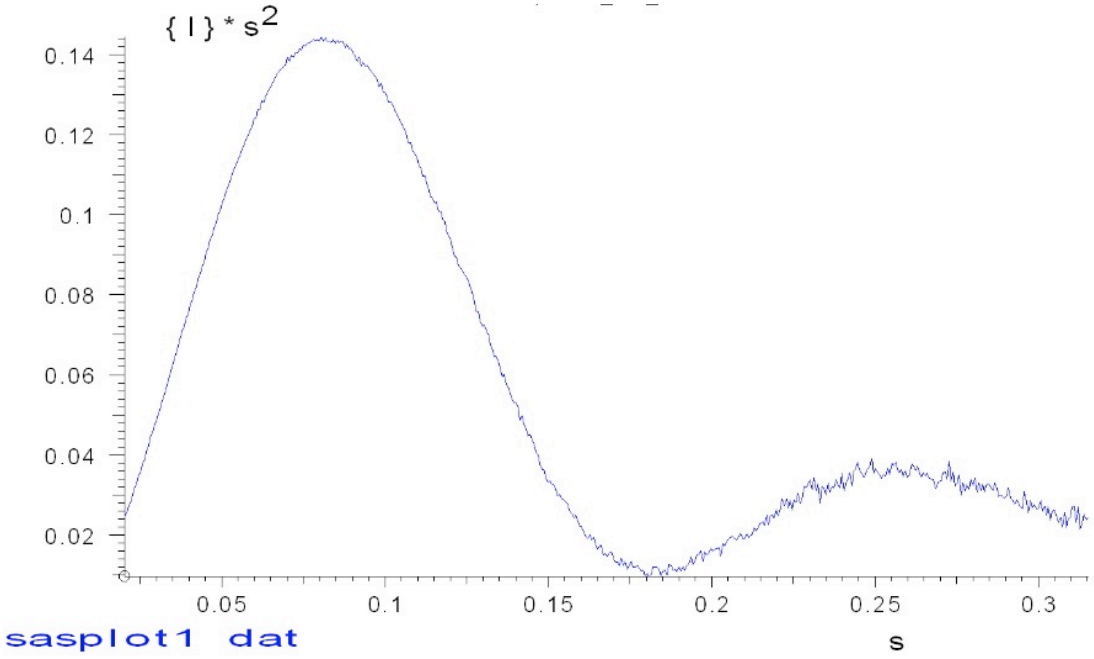

{I}*s2
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0.05
0.1
0.15
0.2
0.25
0.3
s
sasplot1 dat

Fig. 5.3B

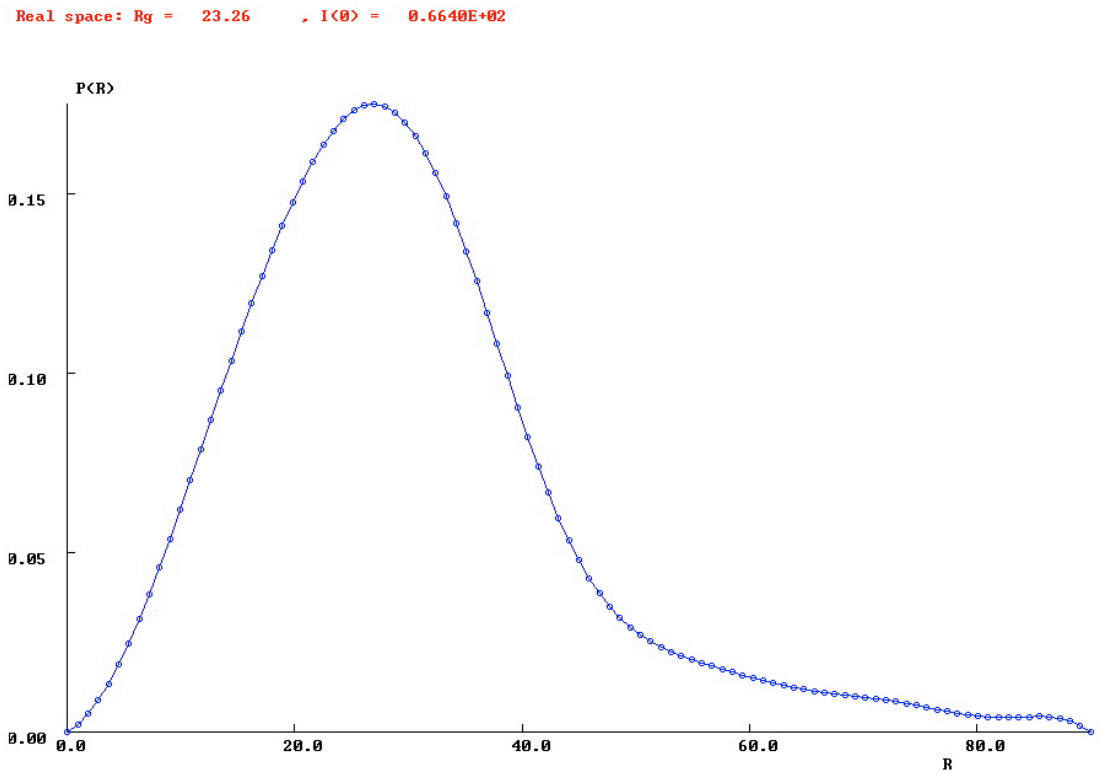
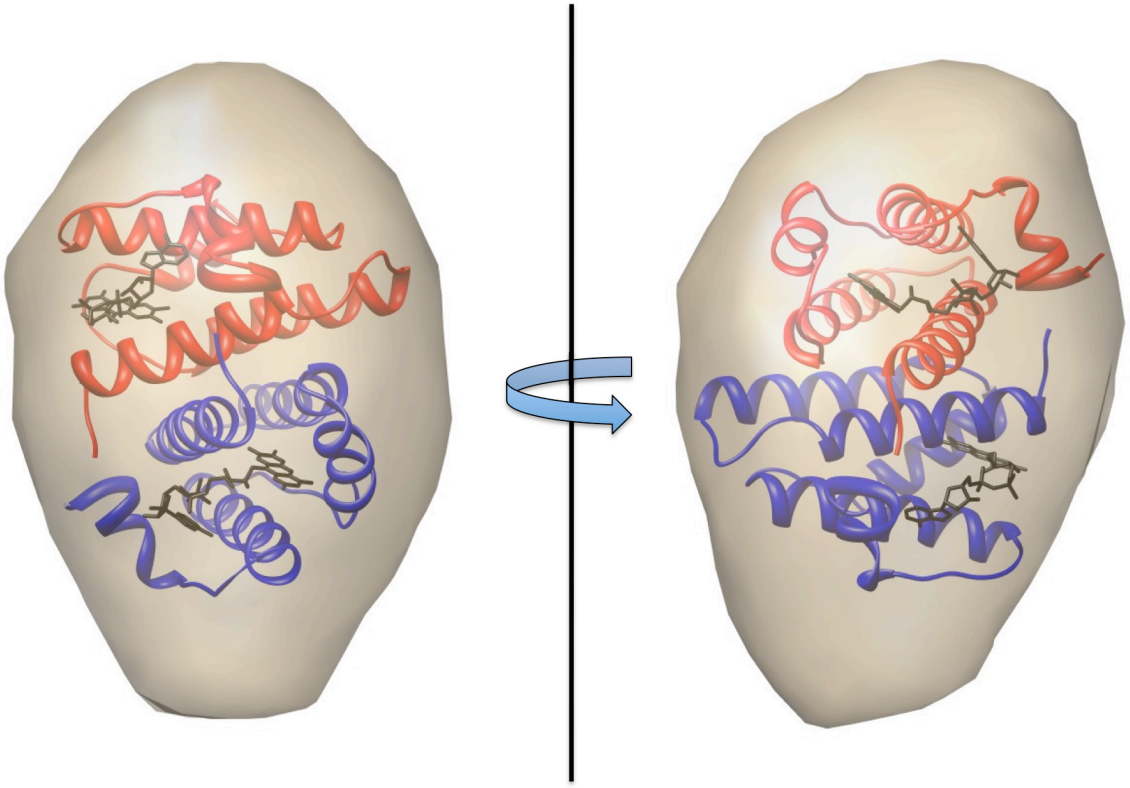

Real space: Rg = 23.26 , I(0) = 0.6640E+02
P(R)
0.15
0.10
0.05
0.00
0.0
20.0
40.0
60.0
80.0
R

Fig. 5.3C



Literature Cited

Adams, G. A., Maestri, M., Squiers, E. C., Alfrey, E. J., Starzl, T. E. & Dafoe, D. C. (1998). *Transplantation* **65**, 32-36.

Allen, S., Balabanidou, V., Sideris, D. P., Lisowsky, T. & Tokatlidis, K. (2005). *J Mol Biol* **353**, 937-944.

Bailey, S. (1994). *Acta Crystallogr D Biol Crystallogr* **50**, 760-763.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D. & Zardecki, C. (2002). *Acta Crystallogr D Biol Crystallogr* **58**, 899-907.

Brunger, A. T. (1992). *Nature* **355**, 472-475.

Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L. K., Schulze-Specking, A., Truscott, K. N., Guiard, B., Meisinger, C. & Pfanner, N. (2004). *The EMBO Journal* **23**, 3735-3746.

Daithankar, V. N., Schaefer, S. A., Dong, M., Bahnson, B. J. & Thorpe, C. (2010). *Biochemistry* **49**, 6737-6745.

Emsley, P. & Cowtan, K. (2004). *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132.

Farrell, S. R. & Thorpe, C. (2005). *Biochemistry* **44**, 1532-1541.

Francavilla, A., Ove, P., Polimeno, L., Coetzee, M., Makowka, L., Rose, J., Van Thiel, D. H. & Starzl, T. E. (1987). *Cancer Res* **47**, 5600-5605.

Francavilla, A., Vujanovic, N. L., Polimeno, L., Azzarone, A., Iacobellis, A., Deleo, A., Hagiya, M., Whiteside, T. L. & Starzl, T. E. (1997). *Hepatology (Baltimore, Md* **25**, 411-415.

Gross, E., Sevier, C. S., Vala, A., Kaiser, C. A. & Fass, D. (2002). *Nat Struct Biol* **9**, 61-67.

Hines, C. C., Reichert, W. M., Griffin, S. T., Bond, A. H., Snowwhite, P. E. & R.D., R. (2006). *J.Mol.Struct.* **796**, 76.

Keese, A. M., Schut, G. J., Ouhammouch, M., Adams, M. W. & Thomm, M. (2010). *J Bacteriol* **192**, 1292-1298.

Klissenbauer, M., Winters, S., Heinlein, U. A. & Lisowsky, T. (2002). *J Exp Biol* **205**, 1979-1986.

Konarev, P. V., Volkov, A.V., Sokolova, M.H.J., Koch, and D.I. Svergun (2003). *Journal of Applied Crystallography* **35**, 1277-1282.

Krissinel, E. & Henrick, K. (2007). *J Mol Biol* **372**, 774-797.

Lange, H., Lisowsky, T., Gerber, J., Muhlenhoff, U., Kispal, G. & Lill, R. (2001). *EMBO Rep* **2**, 715-720.

Laskowski, R. A., Chistyakov, V. V. & Thornton, J. M. (2005). *Nucleic Acids Res Biol Crystallogr*, D **266**-268.

Lisowsky, T. (1996). *Yeast* **12**, 1501-1510.

Lisowsky, T., Lee, J. E., Polimeno, L., Francavilla, A. & Hofhaus, G. (2001). *Dig Liver Dis* **33**, 173-180.

McPherson, A. (1982). *Preparation and analysis of protein crystals*. John Wiley & Sons, Inc.

McPherson, A. (1990). *Eur J Biochem* **189**, 1-23.

Meng, E. C., Pettersen, E. F., Couch, G. S., Huang, C. C. & Ferrin, T. E. (2006). *BMC Bioinformatics* **7**, 339.

Naismith, J. H., Emmerich, C., Habash, J., Harrop, S. J., Helliwell, J. R., Hunter, W. N., Raftery, J., Kalb, A. J. & Yariv, J. (1994). *Acta Crystallogr D Biol Crystallogr* **50**, 847-858.

Otwinowski, Z. & Minor, W. (1997). *Methods in Enzymology* **A276**, 307-326.

Painter, J. & Merritt, E. A. (2006). *Acta Crystallogr D Biol Crystallogr* **62**, 439-450.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J Comput Chem* **25**, 1605-1612.

Polimeno, L., Capuano, F., Marangi, L. C., Margiotta, M., Lisowsky, T., Ierardi, E., Francavilla, R. & Francavilla, A. (2000). *Dig Liver Dis* **32**, 510-517.

Qin, L., Hiser, C., Mulichak, A., Garavito, R. M. & Ferguson-Miller, S. (2006). *Proc Natl Acad Sci U S A* **103**, 16117-16122.

Rose, J. P., Wu, C. K., Dailey, T. A., Dailey, H.A. & Wang, B. C. (1999). *General Assembly, XVIII Congress of the International Union of Crystallography*

Senkevich, T. G., White, C. L., Koonin, E. V. & Moss, B. (2000). *Proc Natl Acad Sci U S A* **97**, 12068-12073.

Sun, H., Yu, H. F., Wu, C. X., Guan, X. Q. & Liu, Q. (2005). *Zhonghua Gan Zang Bing Za Zhi* **13**, 205-208.

Svergun, D. I. (1992). *Journal of Applied Crystallography* **25**, 495-503.

Svergun, D. I., Petoukhov, M. V. & Koch, M. H. (2001). *Biophysical Journal* **80**, 2946-2953.

Tanigawa, K., Sakaida, I., Masuhara, M., Hagiya, M. & Okita, K. (2000). *J Gastroenterol* **35**, 112-119.

Tokatlidis, K. (2005). *Cell* **121**, 965-967.

Trakhanov, S., Kreimer, D. I., Parkin, S., Ames, G. F. & Rupp, B. (1998). *Protein Sci* **7**, 600-604.

Trakhanov, S. & Quiocho, F. A. (1995). *Protein Sci* **4**, 1914-1919.

Volkov, V. V. a. S., D.I. (2003). *Journal of Applied Crystallography* **36**, 860-864.

Wu, C.-K., Dailey, T. A., Dailey, H. A., Francavilla, A., Starzl, T. E., Wang, B. C. & Rose, J. P. (2000). *J. Prot. Pept. Letts.* **7**, 25-32.

Wu, C. K., Dailey, T. A., Dailey, H. A., Wang, B. C. & Rose, J. P. (2003). *Protein Sci* **12**, 1109-1118.

Zanotti, G., Panzalorto, M., Marcato, A., Malpeli, G., Folli, C. & Berni, R. (1998). *Acta Crystallogr D Biol Crystallogr* **54**, 1049-1052.

Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C. & Adams, P. D. (2008). *Methods Mol Biol* **426**, 419-435.

CHAPTER 6

BIOPHYSICAL ANALYSIS OF THE INTERACTION OF AUGMENTER OF LIVER REGENERATION WITH CYTOCHROME C

[1]Florence, Q., Xu, H., Habel, J., Tiusanu, S. and Rose, J.

To be submitted to *Biochemistry*

Abstract

Augmenter of liver regeneration (ALR) is a protein that exists in two alternatively spliced forms, a short cytosolic form (srALR) and a long mitochondrial form (lrALR), inside the cell (*1, 2*). Long form ALR and cytochrome c have recently been identified as key components of the mammalian mitochondrial intermembrane space import and assembly pathway, which imports cysteine rich proteins synthesized in the cytosol into the mitochondrial intermembrane space (*3*). In this pathway, ALR becomes reduced when it oxidizes the protein mitochondrial import and assembly protein 40. In order for the pathway to continue, ALR transfers electrons to cytochrome c so that ALR is returned to an oxidized state (*3-5*). The rate of electron transfer from ALR to cytochrome c has been well characterized but the interactions between the two molecules have not been studied. This work details studies that were conducted to gain kinetic and structural information about the ALR-cytochrome c studies. Small Angle X-ray Scattering (SAXS) experiments displayed solution envelope structures of srALR-cytochrome c and lrALR-cytochrome c complexes in solution. The envelope structure for lrALR-cytochrome c complex was the same in both oxidizing and reducing conditions while the srALR-cytochrome c complex showed a large shift between oxidizing and reducing conditions. These experiments provided evidence supporting that the lrALR-cytochrome c complex forms independently of electron transfer. Additionally, surface plasmon resonance studies displayed equilibrium constants for both complexes that indicated the formation of transient complexes. These results provide supporting evidence to the enzymatic studies that have been conducted on the complexes.

Introduction

Augmenter of liver regeneration (ALR) is a protein that has been shown to stimulate the liver regeneration process and function as a sulfhydryl oxidase (*6, 7*) Originally identified in regenerating rat livers, ALR belongs to the Erv1/ALR protein family that has sequence homology from prokaryotic to eukaryotic organisms(*1, 2, 7*). Mammalian ALR is found in two alternatively spliced forms. The 15 kDa short form ALR (srALR) is mainly located in the cellular cytoplasm while the 23 kDa long form (lrALR) is primarily located in the intermembrane space of the mitochondria (*1*). Studies from *Saccharomyces cerevisiae* have indicated that the ALR homologue Erv1 functions in the intermembrane space import and assembly pathway (*8*). In the pathway, mitochondrial intermembrane assembly protein 40 (Mia40) interacts via intermediate disulfide bond formation with small cysteine rich peptides that are being transported across a translocase of the outer membrane. Through this interaction, the peptides become properly folded and may coordinate a metal ion. After performing this function, Mia40 is in a reduced state and is oxidized by Erv1. The two electrons that are received by Erv1 are transferred to two cytochrome c molecules, which pass them to the electron transport chain (*9, 10*). This process cycles as new peptides are transported across the membrane.

Previous studies by Daithankar *et al.* have shown that mammalian Mia40 transfers electrons to the lrALR but not to srALR (*3*). The data indicates that this is due to the additional CxxC motif located in the N-terminal extension of the lrALR. In addition to studying the electron transfer between Mia40 and lrALR, the Daithankar study also showed that lrALR has a 12-fold lower catalytic efficiency for cytochrome c than srALR

(*3*). This may be due to the manner in which lrALR and cytochrome c interact. To date there has not been a study of the interaction ALR and cyochrome c. Numerous groups have conducted both *in vitro* and *in vivo* studies to investigate the interaction between Mia40 and Erv1/ALR. However, rather little is known about the nature of the interaction of ALR with cytochrome c. As stated previously the transfer of electrons between the two proteins is well established but the binding kinetics and manner in which they interact still needs to be investigated more (*4, 5, 11, 12*).

This work represents the initial binding kinetics study of the interaction between ALR and cytochrome c from *Rattus Norvegicus*. Surface plasmon resonance studies are used to compare the binding rates for lrALR and srALR with cytochrome c. These studies help establish how the additional N-terminal extension on lrALR affects the rates of interaction for both proteins. The study is further enhanced by Small angle X-ray scattering studies, which provide the initial envelope structures depicting the solution structures of the complexes. This works suggests that lrALR and srALR interact with cytochrome c independent of electron transfer. The work provides additional support for lrALR being an enzymatic substrate for cytochrome c.

Experimental Methods

Expression and Purification of ALR

For the various forms of ALR, cDNA was isolated for expression vector construction obtained from *R. norvegicus* and used as a template for PCR (polymerase

chain reaction). The reaction was carried out using a set of primers designed to incorporate restriction sites and six histidines with a tobacco etch virus (TEV) protease site was added at the amino-terminus. The PCR fragment was then ligated into a pTrcHis expression vector (Invitrogen) and the resulting vectors were named pHHhALR (human), pHHrALR (rat) and pHHcALR (C. elegans). The complete DNA sequence for all ALR constructs were verified.

To express the protein a 1-L culture of Circlegrow® (MP Biomedicals) media containing 100 mg of ampicillin was inoculated with E. coli JM109 that had been transformed with the selected plasmid. The culture was grown at 37°C for 18 hours and cells were harvested by centrifugation. The cell pellets were resuspended in 30 ml of a solution containing 20 mM Hepes pH 7.5, 200 mM NaCl and 0.1mM PMSF (phenylmethylsulphonyl fluoride), sonicated, and then centrifuged to separate the soluble protein from the cell debris. Initially, the proteins were isolated from the soluble protein fraction using metal affinity chromatography by running the fraction over a column containing a nickel-NTA resin. After the fraction was washed, the protein was eluted with a buffer containing 20mM Hepes pH 7.5, 200 mM NaCl and 0.1 mM PMSF and 300 mM imidazole. Following the nickel column, the protein was further purified using size exclusion chromatography. The eluted fraction was loaded onto a Superdex75 column containing 20mM Hepes pH 7.5, 200 mM NaCl buffer. The fractions which contained the protein were collected and concentrated. After the protein eluted, purity was assessed by SDS PAGE and UV-visible spectra. All cytochrome c used in the experiments was purchased from Sigma-Aldrich (C2506).

SAXS experiments

SAXS data were collected on ALS beamline 12.3.1. In SAXS experiments under oxidizing conditions, srALR, lrALR, cytochrome c, srALR-cytochrome c mixtures, and lrALR-cytochrome c mixtures were diluted with 20mM Hepes, 150mM NaCl pH 7.4 buffer to give 20μL samples at final concentration of 5mg mL^{-1}. For each protein sample and a buffer blank, SAXS data were collected using a 0.5 s, a 5 s and a second 0.5 s exposure. To test the interactions under reducing conditions, dilutions with 20mM Hepes, 150mM NaCl, 1mM tris(2-caboxyethyl) phosphine hydrochloride (TCEP) pH 7.4 buffer to give 5mg mL^{-1} were prepared and SAXS experiments were repeated under these conditions. As stated in Habel et al., buffer-subtracted files were analyzed using PRIMUS (*13*) and the GNOM (*14*) P(r) output file was used to calculate electron-density envelopes. Each envelope is the product of 16 GASBOR (*15*) runs averaged with DAMMAVER (*16*). All envelope structures were fit with the crystal structures of his-tagged srALR (hsALR) and equine heart cytochrome c (*17*) using the CHIMERA (*18*) software. In all envelope overlays, at least 90 percent of the crystal structures were successfully fit into the envelope.

Surface Plasmon Resonance

Biacore experiments where designed to determine the binding kinetics of the interaction between lrALR and cytochrome c along with the interaction between srALR and cytochrome c. lrALR at 120μM in 20mM Hepes, 150mM NaCl, and 0.1% Tritin-X 100 at pH 7.4 was covalently linked in channel 2 of the sensorchip surface with a 70μL 1:1 mixture of N-hydroxysuc- cinimide (NHS) and N-ethyl-N'-(3-dimethylaminopropyl)-

carbodiimide hydro- chloride (EDC). The remaining activated dextran molecules on he chip where then immobilized with 70μL of ethanolamine (*19*). The reaction resulted in a resonance unit (RU) change of 200 (100RU=1ng/mm^2). For control purposes, channel 1 was activated with a 70μL 1:1 NHS-EDC mixture and blocked with 70μL of ethanolamine. The immobilization steps where repeated with 158μM srALR in channel 4 and channel 3 was used as the srALR control channel. The srALR immobilization resulted in a 2000RU change. 20mM Hepes, 150mM NaCl, and 0.1% Tritin-X 100 at pH 7.4 was used as the running buffer for the experiments which where run under oxidizing conditions. Cytochrome c, which is the analyte in the experiment, was diluted to various concentrations (.4μM, .8μM, 1μM, 2μM, 4μM, 5μM, 6μM, 8μM, 10μM and 12μM) using the running buffer. The various concentrations of cytochrome c were injected over the surface at a flow rate of 15μL/min for 3 minutes at 25°C. After the injection period, running buffer was injected over the surface for 10 minutes to monitor the dissociation. The chip was regenerated for the next injection with a 30 second injection of 30μL of Glycine-HCl at a pH of 3.96. Kinetic information was obtained using BIAevaluation software from GE Healthcare.

Results and Discussion

Cytochrome c has been shown to be the terminal electron acceptor in the intermembrane space import and assembly pathway by various techniques including double mixing experiments and spectroscopic assays (*3, 5*). This work attempted to look at the pathway based on the interactions that occur during the electron transfer within the pathway. Initially, small angle x-ray scattering experiments were conducted to determine

if any structural information could be gained about the interaction of ALR with cytochrome c. In addition, to looking at the physiologically relevant interaction of lrALR with cytochrome c, experiments were conducted on the interaction of srALR with cytochrome c to determine if additional insight can be gained into why srALR does not function in the pathway.

Small Angle X-Ray Scattering

Envelope structures for the individual proteins, srALR-cytochrome c complex, and the lrALR-cytochrome c complex under reducing and oxidizing conditions where determined to 15Å. Envelope structures for the individual proteins revealed structures of globular molecules under both oxidizing and reducing conditions (figure 6.1A-F). lrALR is the only protein that has a narrow region within it's structure under both oxidizing and reducing conditions (figure 6.1C and 6.1D). The narrow region is normally an indication that the protein contains a random coiled or unfolded region. When comparing the structure with the srALR structure (figure 6.1A and 6.1B), the 80 residue N-terminal region of lrALR greatly alters the overall envelope structure but the hsALR fits into the globular portion envelope in a manner similar to the srALR structure. The flexibility of the N-terminal region makes it ideal for interaction. Interestingly, the envelope structure of srALR under reducing conditions (figure 6.1B) displayed the formation of a narrow tail on one end of the envelope. This indicates that under reducing conditions, srALR may undergo conformational changes that make it ideal for interacting with other proteins. This is supported by the envelope structures that were determined for the complexes.

As seen in figure 6.2A, the envelope structure under oxidizing conditions for lrALR-cytochrome c complex revealed a multi-lobed envelope that appeared to display the proteins interacting with one another. From these fit studies, the envelope structure provides evidence supporting the ALR dimmer interacting with one molecule of cytochrome c. To test if electron transfers would alter the envelope structure, the experiments were repeated with reducing agent, TCEP, added to the solution. There was not a significant shift observed in the envelope structure (Figure 6.2B). The lrALR and cytochrome c appeared to interact in the same manner as previously observed under aerobic conditions. This result indicated that the interaction of lrALR-cytochrome c is independent of electron transfer between the two molecules.

Previous studies have shown that electron transfer between srALR and cytochrome c is actually more efficient than electron transfer from lrALR to cytochrome c (*3*). SAXS experiments were conducted on srALR-cytochrome c to determine if any structural evidence could be observed to support this observation. The envelope structures determined for srALR-cytochrome c under oxidizing conditions produced a similar multi-lobed structure. However, a narrow region connected the two lobes and this indicates that the interaction is mediated by a flexible region (Figure 6.3A). However, when the complex was subjected to reducing conditions there was a shift in the envelope structure. The two lobes of the complex moved close together eliminating the narrow region that appeared to link them (Figure 6.3B). Unlike the lrALR-cytochrome c complex which does not shift, the shift that occurs with srALR-cytochrome c likely brings the FAD molecule from srALR and the heme group from cytochrome c within a close proximity of one another. This change in proximity could make the transfer of electrons

from srALR to cytochrome c more efficient. In addition the N-terminal region of the lrALR likely inhibits this shift from occurring and thus the electrons likely have to travel from the FAD to the CXXC motif and then to the heme group, which is what is observed in previous studies. These SAXS envelope structures demonstrate that lrALR-cytochrome c and srALR-cytochrome c complexes can be formed but the interaction between lrALR and cytochrome c does not appear to change under oxidizing or reducing conditions. The N-terminal region of lrALR seems to be key in controlling the interaction between ALR and cytochrome c.

Surface Plasmon Resonance

In order to gain more insight into the interaction of ALR and cytochrome c surface plasmon resonance studies were conducted using the Biacore3000 apparatus. In Biacore experiments, a ligand is immobilized to the metal material and an analyte in a buffer condition is passed over the ligand. If interaction between the ligand and analyte occurs, the mass of the molecules interacting with the surface changes and this alters the distribution of the surface plasmons (*20, 21*). Biacore utilizes the fact that if there is a change in the distribution of the surface plasmons due to a mass change of the molecule interacting with the surface then the angle of the reflecting light changes. The change in the reflecting light is measured and recorded as resonance units by the Biacore3000 (*20, 21*). Through the real time measurements of the changes in these resonance units, information can be gained about the association rate, dissociation rate, and equilibrium constants of an interaction.

In the experiment srALR and lrALR were immobilized to the sensor chip in different channels while cytochrome c was used as the analyte. Representative biacore

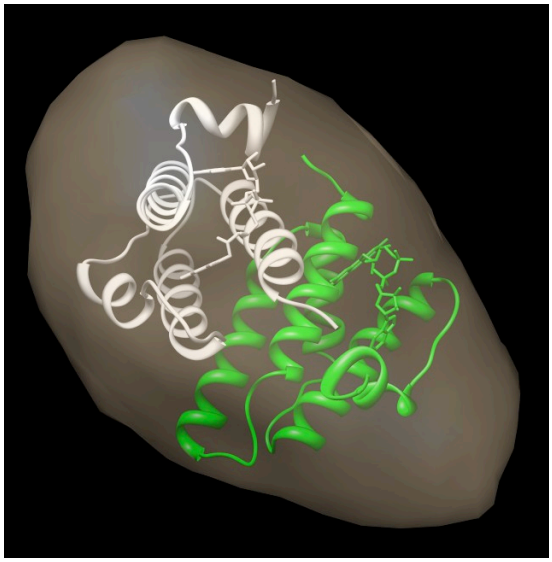
curves that were used to calculate the binding kinetics are shown in figure 6.4A and 6.4B. As seen in table 1, the association constant (ka), dissociation constant (kd), and equilibrium constant (KD) for the lrALR-cytochrome c interaction are 1.84 x 10^3 Ms^{-1}, 2.1 x 10^{-3} s^{-1}, and 1.45 μM. The equilibrium constant value provides an indication into the strength of the interactions between the two proteins. KD values below 10 nM signals that the proteins form stable complexes with one another. While values above this, normally indicate more transient interactions. The KD value obtain for the lrALR-cytochrome c complex indicates the two proteins form a transient complex. A transient complex would be necessary for the proteins to function properly in the IMS import and assembly pathway. Unlike the lrALR-cytochrome c complex, srALR appears to form a more stable complex with cytochrome c. The association constant (ka), dissociation constant (kd), and equilibrium constant (KD) for the srALR-cytochrome c complex are 4.4 x 10^3 Ms^{-1}, 1.3 x 10^{-3} s^{-1}, and 300 nM (Table 6.1). Although the interaction between srALR and cytochrome c is tighter than the lrALR-cytochrome c complex, a KD of 300 nM indicates that the interaction is still a weaker interaction. The difference between the equilibrium constants is likely due to the fact that the N-terminal region of lrALR inhibits cytochrome c from binding in the same manner that it binds to srALR. As stated previously, this may be due to the fact that the 80 amino acids that compose the N-terminal region have a calculated pI of 8.9 and cytochrome c has a pI of 10 (*17*). The pI values indicate that these two molecules may repel each other. In addition, the region that cytochrome c binds to on lrALR is likely different than the srALR binding region. More studies need to be done to determine the amino acids that promote the binding of the two molecules to better understand the interaction.

Conclusion

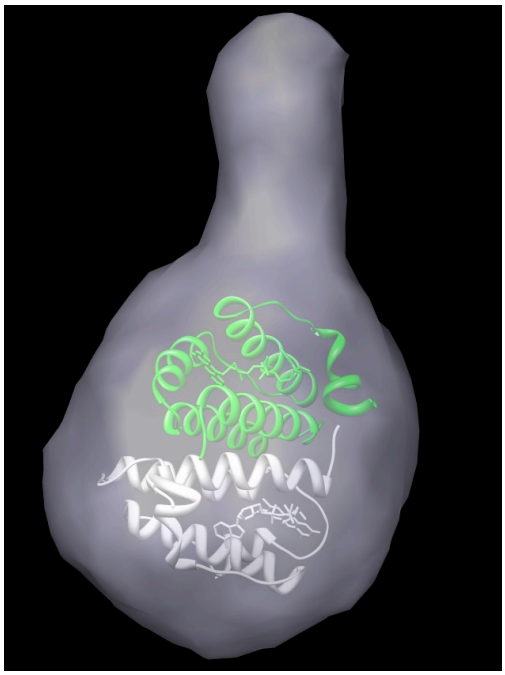
Redox reactions are prevalent throughout the cell and are essential to the function of many pathways (*22*). The electron transfer between the two molecules normally coincides with some physical interaction. These interactions occur quickly and are usually difficult to monitor. Fortunately in some cases, the protein molecules will interact with one another independent of the electron transfer. Studying these interactions can provide additional insight into the mechanism for these proteins. Here we report a study that has identified the ability of lrALR and srALR to interact with cytochrome c in an electron transfer independent manner. From our SAXS analysis, envelope structures were determined to give the initial structural information about the complex. While Biacore surface plasmon resonance analysis was utilized to conduct the initial binding kinetic study of the interactions. Our studies reveal that these proteins interact with one another in a transient manner that is ideal for redox reactions. Additionally, this work further supports previous studies that all suggest that lrALR is the version of augmenter of liver regeneration that functions in the mitochondrial intermembrane space import and assembly pathway. Further studies need to be conducted on the additional protein interactions that are in the pathway. Understanding these interactions could provide a foundation for studies that are attempting to gain additional insight into how ALR augments the liver regeneration process.

Figure 6.1: Envelope structures of the individual proteins with the crystal structures of hsALR and cytochrome c fit into the envelopes. A) Globular envelope structure of srALR under oxidizing conditions. The hsALR dimer fits into the envelope with any outliers. B) hsALR fit into srALR envelope structure determined under reducing conditions. A formation of a narrow tail can be observed in this structure. C) hsALR fit into lrALR envelope structure determined under oxidizing conditions. lrALR envelope structure reveals a large flexible region. D) hsALR fit into lrALR envelope structure determined under reducing conditions. The structure looks similar to the one observed under oxidizing conditions. E) Cytochrome c envelope structure under oxidizing conditions with cytochrome c crystal structure fit into the envelope. F) Cytochrome c envelope structure under reducing conditions. The envelope structure looks similar to that observed under oxidizing conditions.

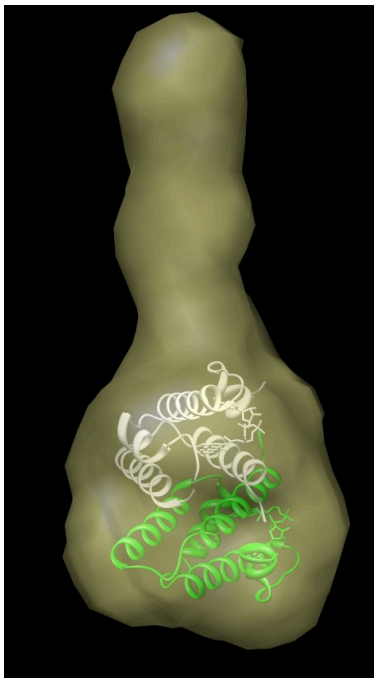
A)



B)

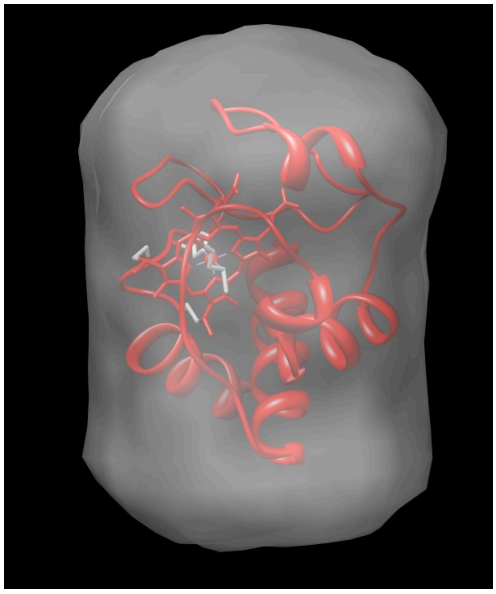


C)



D)

E)



F)

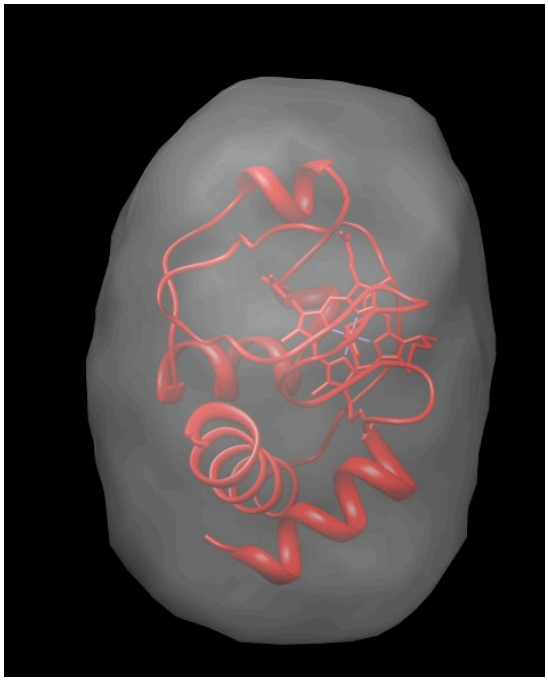
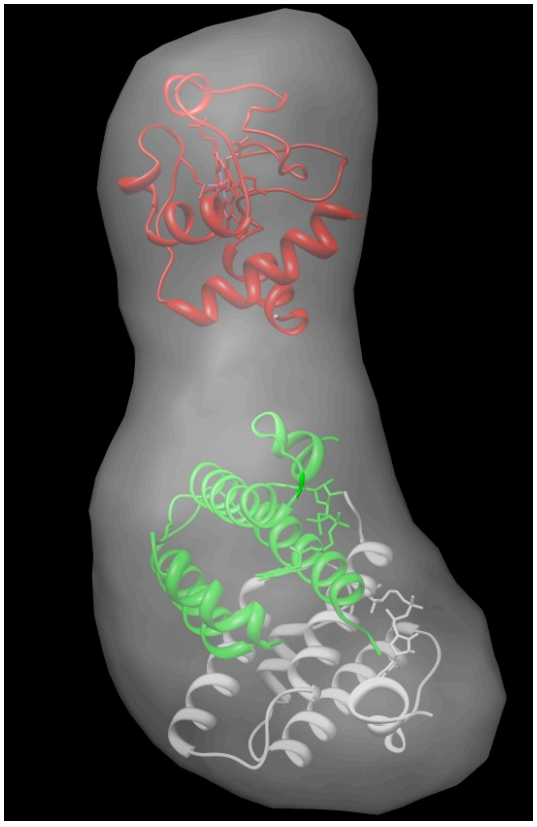


Figure 6.2: SAXS envelope structures of the lrALR-cytochrome c protein complex. A) Envelope structure of the lrALR-cytochrome c under oxidizing conditions. The image is an overlay of two structure determination runs. B) The envelope structure of lrALR-cytochrome c complex in reducing conditions looks similar to the structure under oxidizing conditions.

A)



B)

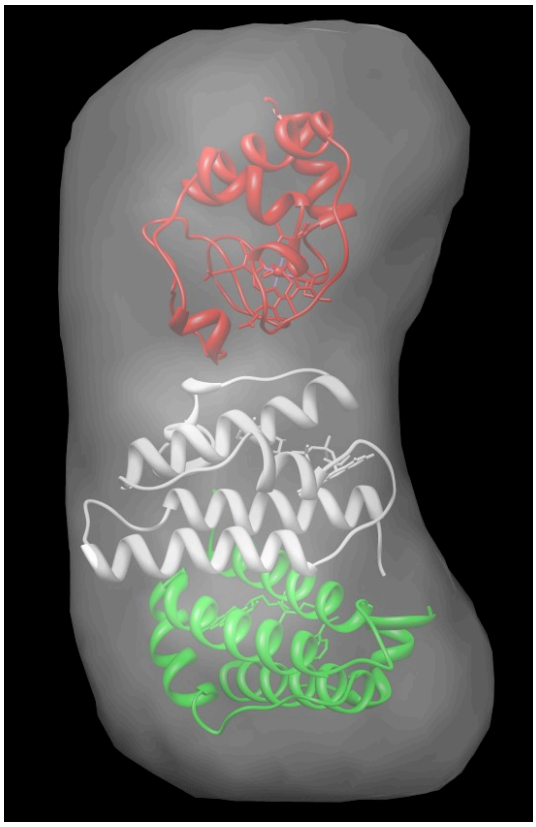
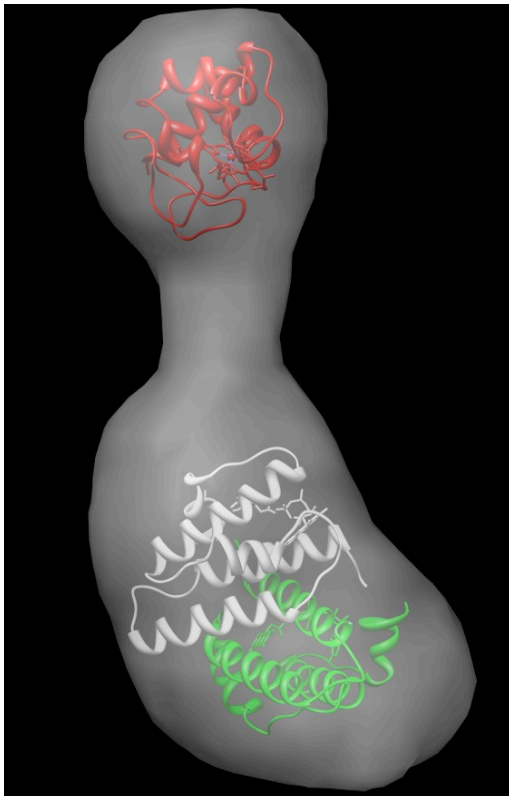


Figure 6.3: SAXS envelope structures of the srALR-cytochrome c protein complex. A) Envelope structure of the srALR-cytochrome c under oxidizing conditions. The image is an overlay of two structure determination runs. B) A large shift is observed in the envelope structure of srALR-cytochrome c under reducing conditions. It appears that the FAD moiety from srALR and the heme group from cytochrome c are brought closer together for electron transfer.

A)



B)

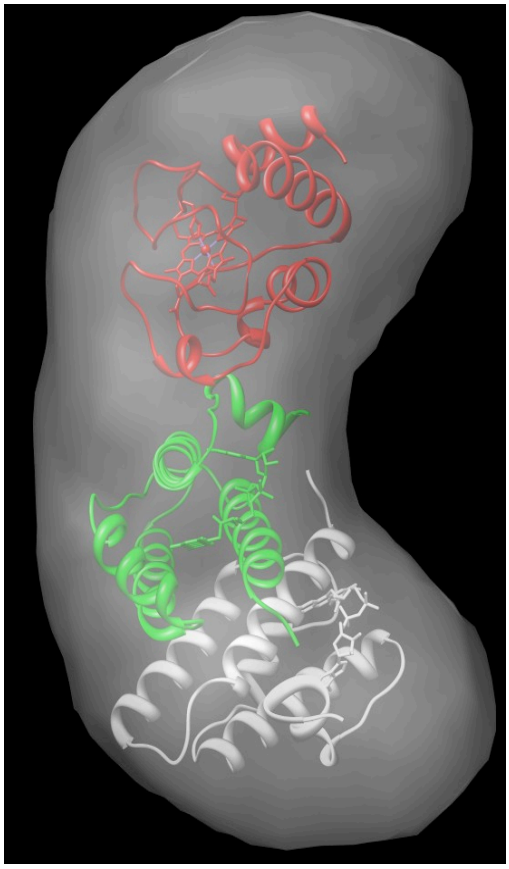
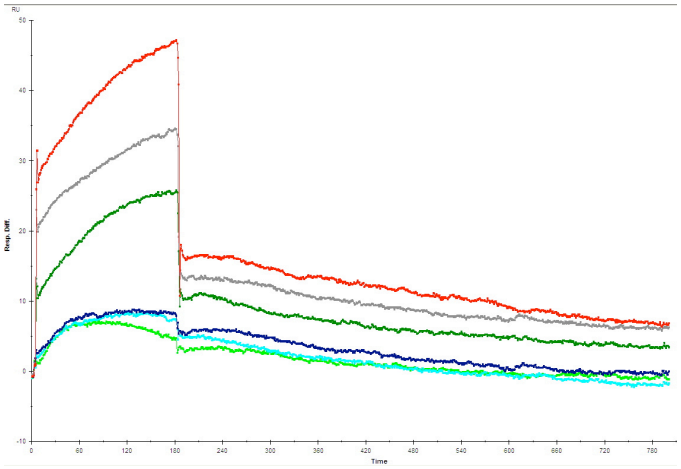


Figure 6.4: Suface plasmon resonance curves showing the change in the resonance signal as cytochrome c is injected over the chip. (A) The curves display the resonance signals for the interaction of cytochrome c with lrALR following injections of cytochrome c at concentrations of .4 µM, .8 µM, 1 µM, 2 µM, 4 µM, 8 µM, and10 µM. (B) The curves display the resonance signals for the interaction of cytochrome c with srALR following injections of cytochrome c at concentrations of .4 µM, .8 µM, 1 µM, 2 µM, 4 µM, 5µM, 8 µM, and10 µM.

A)



B)

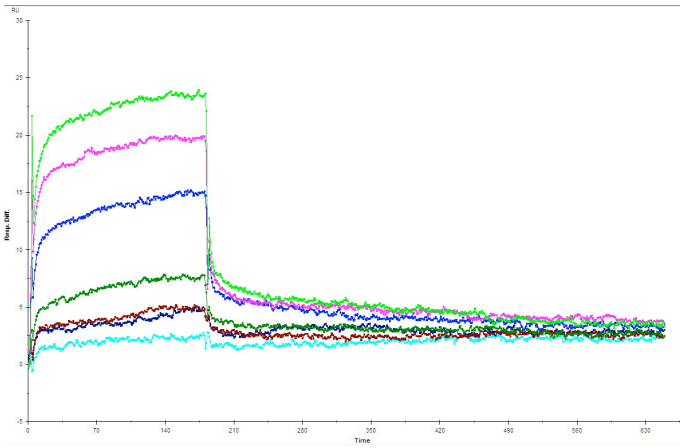


Table 6.1: Kinetic Calculations for the interactions of cytochrome c with lrALR and srALR.

	k_a (1/Ms)	k_d (1/s)	KA (1/M)	KD (M)
lrALR	1.84×10^{3}	2.1×10^{-3}	8.77×10^{5}	1.14×10^{-6}
srALR	4.4×10^{3}	1.3×10^{-3}	3.3×10^{6}	3.0×10^{-7}

Literature Cited

1. 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.
2. Hagiya, M., Francavilla, A., Polimeno, L., Ihara, I., Sakai, H., Seki, T., Shimonishi, M., Porter, K. A., and Starzl, T. E. (1995) Cloning and sequence analysis of the rat augmenter of liver regeneration (ALR) gene: expression of biologically active recombinant ALR and demonstration of tissue distribution, *Proceedings of the National Academy of Sciences of the United States of America 92*, 3076.
3. Daithankar, V. N., Farrell, S. R., and Thorpe, C. (2009) Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space, *Biochemistry 48*, 4828-4837.
4. Muller, J. M., Milenkovic, D., Guiard, B., Pfanner, N., and Chacinska, A. (2008) Precursor oxidation by mia40 and erv1 promotes vectorial transport of proteins into the mitochondrial intermembrane space, *Molecular biology of the cell 19*, 226-236.
5. Hell, K. (2008) The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria, *Biochimica et biophysica acta 1783*, 601-609.
6. Lisowsky, T., Lee, J. E., Polimeno, L., Francavilla, A., and Hofhaus, G. (2001) Mammalian augmenter of liver regeneration protein is a sulfhydryl oxidase, *Dig Liver Dis 33*, 173-180.

7. 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.

8. Allen, S., Balabanidou, V., Sideris, D. P., Lisowsky, T., and Tokatlidis, K. (2005) Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome c, *Journal of molecular biology 353*, 937-944.

9. Rissler, M., Wiedemann, N., Pfannschmidt, S., Gabriel, K., Guiard, B., Pfanner, N., and Chacinska, A. (2005) The essential mitochondrial protein Erv1 cooperates with Mia40 in biogenesis of intermembrane space proteins, *Journal of molecular biology 353*, 485-492.

10. Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J. M. (2005) A disulfide relay system in the intermembrane space of mitochondria that mediates protein import, *Cell 121*, 1059-1069.

11. Terziyska, N., Grumbt, B., Bien, M., Neupert, W., Herrmann, J. M., and Hell, K. (2007) The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway, *FEBS letters 581*, 1098-1102.

12. Gabriel, K., Milenkovic, D., Chacinska, A., Muller, J., Guiard, B., Pfanner, N., and Meisinger, C. (2007) Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway, *Journal of molecular biology 365*, 612-620.

13. Konarev, P. V., Volkov, A.V., Sokolova, M.H.J., Koch, and D.I. Svergun (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *Journal of Applied Crystallography 35*, 1277-1282.

14. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria *J. Appl. Cryst. 25*, 495-503.

15. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. J. (2001) Determination of Domain Structure of Proteins from X-Ray Solution Scattering, *80*, 2946-2953.

16. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering, *Journal of Applied Crystallography 36*, 860-864.

17. Bushnell, G. W., Louie, G. V., and Brayer, G. D. (1990) High-resolution three-dimensional structure of horse heart cytochrome c, *Journal of molecular biology 214*, 585-595.

18. Goddard, T. D., Huang, C. C., and Ferrin, T. E. (2007) Visualizing density maps with UCSF Chimera, *J Struct Biol 157*, 281-287.

19. Patel, R., Pollner, R., de Keczer, S., Pease, J., Pirio, M., DeChene, N., Dafforn, A., and Rose, S. (2000) Quantification of DNA using the luminescent oxygen channeling assay, *Clinical chemistry 46*, 1471-1477.

20. Karlsson, R., Michaelsson, A., and Mattsson, L. (1991) Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system, *Journal of immunological methods 145*, 229-240.

21. Murphy, M., Jason-Moller, L., and Bruno, J. (2006) Using Biacore to measure the binding kinetics of an antibody-antigen interaction, *Curr Protoc Protein Sci Chapter 19*, Unit 19 14.

22. Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L. K., Schulze-Specking, A., Truscott, K. N., Guiard, B., Meisinger, C., and Pfanner, N. (2004) Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins, *The EMBO journal 23*, 3735-3746.

CHAPTER 7

CONCLUSIONS

The study of liver damage and liver regeneration has evolved over several decades. Although knowledge has been gained into how these processes work, more research needs to be conducted to provide a foundation for developing therapeutics that would take advantage of the liver's regeneration process and prevent liver damage. The research conducted in this work directly contributes to this field of study and provides a framework for additional studies on liver damage and regeneration.

The research projects described within this dissertation provide a solid foundation for future studies on liver damage and regeneration. The biophysical studies that were conducted on the hepatitis b surface antigen protein (HBsAg) provide initial solution structural information that could be utilized in future experiments for the development of better preventative and therapeutic methods for hepatitis b. Since the SAXS studies revealed the presence of a large 40 nm particle in solution, these studies revealed more insight into why determining high-resolution structural information on the HBsAg protein has proven to de difficult. Although SAXS studies of HBsAg in complex with an antibody has not produced an envelope structure, the SAXS studies did reveal that the interaction of HbsAg with the antibody causes the protein to be present in both folded and unfolded states in solution. These studies have led to the experiments aimed at the preparation of alternative forms of HBsAg that will hopefully be more convenient for crystallization and still produce the same antigenic response. These HBsAg proteins will likely be in the form of mutants that are outside of the antigenic region but they will

prevent the protein from aggregating into 40nm particles in solution. In addition to producing mutants for structural studies, the production of mutants that are located in the antigenic region could add to studies on the protein-antibody interaction of HBsAg escape mutants. Surface plasmon resonance (SPR) studies could be used as a tool for addressing this need. SPR studies would be used to readily assess the strength of the protein-antibody interaction of any mutants that are produced. Although diffraction quality crystals of HbsAg have yet to be produced, knowledge was gained about the solution and antibody binding properties of HBsAg, which will be invaluable for future research and development of better vaccines to hepatitis b.

The goal of the second research project that was conducted focused on providing evidence that supported the existence of the mitochondrial intermembrane space import and assembly pathway in mammals. Specifically, the project focused on testing the hypothesis that the augmenter of liver regeneration (ALR) replaces the protein essential for respiration and viability 1 (Erv1) in the mammalian version of the pathway. In order to accomplish this task, several biochemical assays were carried out. Aerobic and anaerobic spectroscopic assays revealed that electron transfer between the mammalian forms of Mia40 (mitochondrial import and assembly protein 40), ALR, and cytochrome c does occur. Alphascreen and SPR experiments revealed that both forms of ALR interact with cytochrome c. SPR also revealed that although the complexes formed were transient, srALR has a stronger binding affinity for cytochrome c even though it does not function in the pathway. SAXS studies revealed solution structural information about srALR, lrALR, cytochrome c, srALR-cytochrome c complex, and lrALR-cytochrome c complex under both oxidizing and reducing conditions. In addition, the SAXS revealed that the

proteins interact with one another independent of electron transfer. Lastly, crystallographic studies were conducted on his-tagged short form ALR. The crystal structure revealed a structure that was similar to the previously reported native structure of ALR. However, the variability of the FAD binding pocket and the structural variability observed for the 60 - 65 loop observed in the his-tagged short form ALR structure may be an indicator of the dynamics of the ALR catalytic site needed to accommodate potential oxidoreduction partners. The crystal structure of the his-tagged short form ALR also supplemented the SAXS studies by providing a high-resolution structure that could be modeled into the SAXS envelope structures. The results from all the experiments that were conducted provide support the hypothesis that ALR functions in the mammalian mitochondrial intermembrane space assembly and import pathway. These experiments have laid the groundwork for future experiments that can be conducted to investigate the additional protein interactions and electron transfers that occur in the pathway.

Liver damage and the subsequent regeneration that occurs afterwards is a complex and well-studied process. Despite the numerous studies into the proteins, peptides, and other macromolecules involved, more information needs to be gained into liver damage and regeneration so that the impact of diseases like hepatitis b can be curbed in the future. This study has provided additional knowledge that can lead to development of therapeutic techniques that would aide in accomplishing this task.

BIBLIOGRAPHY

1. 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.
2. 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 (Baltimore, Md 20*, 747-757.
3. Daithankar, V. N., Farrell, S. R., and Thorpe, C. (2009) Augmenter of liver regeneration: substrate specificity of a flavin-dependent oxidoreductase from the mitochondrial intermembrane space, *Biochemistry 48*, 4828-4837.
4. 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 (Baltimore, Md 25*, 411-415.
5. Lisowsky, T., Lee, J. E., Polimeno, L., Francavilla, A., and Hofhaus, G. (2001) Mammalian augmenter of liver regeneration protein is a sulfhydryl oxidase, *Dig Liver Dis 33*, 173-180.
6. Hagiya, M., Francavilla, A., Polimeno, L., Ihara, I., Sakai, H., Seki, T., Shimonishi, M., Porter, K. A., and Starzl, T. E. (1995) 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 92*, 3076.

7. Allen, S., Balabanidou, V., Sideris, D. P., Lisowsky, T., and Tokatlidis, K. (2005) Erv1 mediates the Mia40-dependent protein import pathway and provides a functional link to the respiratory chain by shuttling electrons to cytochrome c, *J Mol Biol 353*, 937-944.

8. 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.

9. Jin, X., Bauer, D. E., Tuttleton, S. E., Lewin, S., Gettie, A., Blanchard, J., Irwin, C. E., Safrit, J. T., Mittler, J., Weinberger, L., Kostrikis, L. G., Zhang, L., Perelson, A. S., and Ho, D. D. (1999) Dramatic rise in plasma viremia after CD8(+) T cell depletion in simian immunodeficiency virus-infected macaques, *J Exp Med 189*, 991-998.

10. Tranebjaerg, L., Hamel, B. C., Gabreels, F. J., Renier, W. O., and Van Ghelue, M. (2000) A de novo missense mutation in a critical domain of the X-linked DDP gene causes the typical deafness-dystonia-optic atrophy syndrome, *Eur J Hum Genet 8*, 464-467.

11. Williams, R. (2006) Global challenges in liver disease, *Hepatology (Baltimore, Md 44*, 521-526.

12. Seeley, R., Stephens, T., Tate, P. (2000) Anatomy and Physiology, *Fifth*.

13. Chang, M. H. (2007) Hepatitis B virus infection, *Semin Fetal Neonatal Med 12*, 160-167.

14. Barker, L. F., Shulman, N. R., Murray, R., Hirschman, R. J., Ratner, F., Diefenbach, W. C., and Geller, H. M. (1996) Transmission of serum hepatitis. 1970, *Jama 276*, 841-844.

15. Alter, M. J. (2003) Epidemiology and prevention of hepatitis B, *Semin Liver Dis 23*, 39-46.

16. Glebe, D. (2007) Recent advances in hepatitis B virus research: a German point of view, *World J Gastroenterol 13*, 8-13.

17. Bruss, V. (2007) Hepatitis B virus morphogenesis, *World J Gastroenterol 13*, 65-73.

18. Beck, J., and Nassal, M. (2007) Hepatitis B virus replication, *World J Gastroenterol 13*, 48-64.

19. Iannacone, M., Sitia, G., Isogawa, M., Marchese, P., Castro, M. G., Lowenstein, P. R., Chisari, F. V., Ruggeri, Z. M., and Guidotti, L. G. (2005) Platelets mediate cytotoxic T lymphocyte-induced liver damage, *Nat Med 11*, 1167-1169.

20. Vierling, J. M. (2007) The immunology of hepatitis B, *Clin Liver Dis 11*, 727-759, vii-viii.

21. Zuckerman, J. N. (2006) Vaccination against hepatitis A and B: developments, deployment and delusions, *Curr Opin Infect Dis 19*, 456-459.

22. Doerr, H. W., and Gerlich, W. H. (2002) *Medizinische Virologie*, Georg Thieme Verlag KG, Stuttgart, Germany.

23. Gilbert, R. J., Beales, L., Blond, D., Simon, M. N., Lin, B. Y., Chisari, F. V., Stuart, D. I., and Rowlands, D. J. (2005) Hepatitis B small surface antigen particles are octahedral, *Proc Natl Acad Sci U S A 102*, 14783-14788.

24. Coleman, P. F. (2006) Detecting hepatitis B surface antigen mutants, *Emerg Infect Dis 12*, 198-203.

25. Netter, H. J., Woo, W. P., Tindle, R., Macfarlan, R. I., and Gowans, E. J. (2003) Immunogenicity of recombinant HBsAg/HCV particles in mice pre-immunised with hepatitis B virus-specific vaccine, *Vaccine 21*, 2692-2697.

26. Higgins, G. M., and Anderson, R. M. (1931) Restoration of the liver of white rat following partial surgical removal *Arch. Pathol. 12*, 186-202.

27. Pawlowski, R., and Jura, J. (2006) ALR and liver regeneration, *Molecular and cellular biochemistry 288*, 159-169.

28. Su, A. I., Guidotti, L. G., Pezacki, J. P., Chisari, F. V., and Schultz, P. G. (2002) Gene expression during the priming phase of liver regeneration after partial hepatectomy in mice, *Proc Natl Acad Sci U S A 99*, 11181-11186.

29. Stahl, N., Boulton, T. G., Farruggella, T., Ip, N. Y., Davis, S., Witthuhn, B. A., Quelle, F. W., Silvennoinen, O., Barbieri, G., Pellegrini, S., and et al. (1994) Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components, *Science 263*, 92-95.

30. Braun, P., Hu, Y., Shen, B., Halleck, A., Koundinya, M., Harlow, E., and LaBaer, J. (2002) Proteome-scale purification of human proteins from bacteria, *Proc Natl Acad Sci U S A 99*, 2654-2659.

31. Hemmann, U., Gerhartz, C., Heesel, B., Sasse, J., Kurapkat, G., Grotzinger, J., Wollmer, A., Zhong, Z., Darnell, J. E., Jr., Graeve, L., Heinrich, P. C., and Horn, F. (1996) Differential activation of acute phase response factor/Stat3 and Stat1 via the cytoplasmic domain of the interleukin 6 signal transducer gp130. II. Src

homology SH2 domains define the specificity of stat factor activation, *J Biol Chem 271*, 12999-13007.

32. Webber, E. M., Bruix, J., Pierce, R. H., and Fausto, N. (1998) Tumor necrosis factor primes hepatocytes for DNA replication in the rat, *Hepatology (Baltimore, Md 28*, 1226-1234.

33. Kirillova, I., Chaisson, M., and Fausto, N. (1999) Tumor necrosis factor induces DNA replication in hepatic cells through nuclear factor kappaB activation, *Cell Growth Differ 10*, 819-828.

34. Diehl, A. M., Yang, S. Q., Yin, M., Lin, H. Z., Nelson, S., and Bagby, G. (1995) Tumor necrosis factor-alpha modulates CCAAT/enhancer binding proteins-DNA binding activities and promotes hepatocyte-specific gene expression during liver regeneration, *Hepatology (Baltimore, Md 22*, 252-261.

35. Webber, E. M., Wu, J. C., Wang, L., Merlino, G., and Fausto, N. (1994) Overexpression of transforming growth factor-alpha causes liver enlargement and increased hepatocyte proliferation in transgenic mice, *Am J Pathol 145*, 398-408.

36. Gallucci, R. M., Simeonova, P. P., Toriumi, W., and Luster, M. I. (2000) TNF-alpha regulates transforming growth factor-alpha expression in regenerating murine liver and isolated hepatocytes, *J Immunol 164*, 872-878.

37. Argast, G. M., Campbell, J. S., Brooling, J. T., and Fausto, N. (2004) Epidermal growth factor receptor transactivation mediates tumor necrosis factor-induced hepatocyte replication, *J Biol Chem 279*, 34530-34536.

38. McTigue, M. A., Wickersham, J. A., Pinko, C., Showalter, R. E., Parast, C. V., Tempczyk-Russell, A., Gehring, M. R., Mroczkowski, B., Kan, C. C., Villafranca,

J. E., and Appelt, K. (1999) Crystal structure of the kinase domain of human vascular endothelial growth factor receptor 2: a key enzyme in angiogenesis, *Structure Fold Des 7*, 319-330.

39. Jo, M., Stolz, D. B., Esplen, J. E., Dorko, K., Michalopoulos, G. K., and Strom, S. C. (2000) Cross-talk between epidermal growth factor receptor and c-Met signal pathways in transformed cells, *J Biol Chem 275*, 8806-8811.

40. de Juan, C., Benito, M., Alvarez, A., and Fabregat, I. (1992) Differential proliferative response of cultured fetal and regenerating hepatocytes to growth factors and hormones, *Experimental cell research 202*, 495-500.

41. 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.

42. He, F., Wu, C., Tu, Q., and Xing, G. (1993) Human hepatic stimulator substance: a product of gene expression of human fetal liver tissue, *Hepatology (Baltimore, Md 17*, 225-229.

43. McJunkin, F. A., and Breuhaus, H. C. (1931) Homologous liver as a stimulus to hepatic regeneration, *Arch. Pathol. 12*, 900-908.

44. Teir, H., and Ravanti, K. (1953) Mitotic activity and growth factors in the liver of the whole rat., *Exp. Cell Res. 5*, 500-507.

45. Blomqvist, K. (1957) Growth stimulation in the liver and tumour development following intraperitoneal injections of liver homogenates in the rat, *Acta Pathol Microbiol Scand 121*, 1-65.

46. LeBrecque, D. R., and Pesch, L. A. (1975) preparation and partial characterization of hepatic regenerative stimulator substance (HSS) from rat liver., *J. Physiol. 248*, 273-284.

47. Terblanche, J., Porter, K. A., Starzl, T. E., Moore, J., Patzelt, L., and Hayashida, N. (1980) Stimulation of hepatic regeneration after partial hepatectomy by infusion of a cytosol extract from regenerating dog liver, *Surg Gynecol Obstet 151*, 538-544.

48. Starzl, T. E., Jones, A. F., Terblanche, J., Usui, S., Porter, K. A., and Mazzoni, G. (1979) Growth-stimulating factor in regenerating canine liver, *Lancet 1*, 127-130.

49. 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.

50. 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.

51. 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.

52. Consortium, T. E. P. K. D. (1994) The polycystic kidney disease 1 gene encodes a 14 kb transcript and lies within a duplicated region on chromosome 16. The European Polycystic Kidney Disease Consortium, *Cell 77*, 881-894.

53. Tanigawa, K., Sakaida, I., Masuhara, M., Hagiya, M., and Okita, K. (2000) Augmenter of liver regeneration (ALR) may promote liver regeneration by reducing natural killer (NK) cell activity in human liver diseases, *J Gastroenterol 35*, 112-119.

54. Itoh, H., Abo, T., Sugawara, S., Kanno, A., and Kumagai, K. (1988) Age-related variation in the proportion and activity of murine liver natural killer cells and their cytotoxicity against regenerating hepatocytes, *J Immunol 141*, 315-323.

55. Adams, G. A., Maestri, M., Squiers, E. C., Alfrey, E. J., Starzl, T. E., and Dafoe, D. C. (1998) Augmenter of liver regeneration enhances the success rate of fetal pancreas transplantation in rodents, *Transplantation 65*, 32-36.

56. Polimeno, L., Capuano, F., Marangi, L. C., Margiotta, M., Lisowsky, T., Ierardi, E., Francavilla, R., and Francavilla, A. (2000) The augmenter of liver regeneration induces mitochondrial gene expression in rat liver and enhances oxidative phosphorylation capacity of liver mitochondria, *Dig Liver Dis 32*, 510-517.

57. 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.

58. 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.

59. Sun, H., Yu, H. F., Wu, C. X., Guan, X. Q., and Liu, Q. (2005) [Expression of augmenter of liver regeneration in hepatic tumor cells and its clinical significance.], *Zhonghua Gan Zang Bing Za Zhi 13*, 205-208.

60. Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L. K., Schulze-Specking, A., Truscott, K. N., Guiard, B., Meisinger, C., and Pfanner, N. (2004) Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins, *The EMBO journal 23*, 3735-3746.

61. Tokatlidis, K. (2005) A disulfide relay system in mitochondria, *Cell 121*, 965-967.

62. Gabriel, K., Milenkovic, D., Chacinska, A., Muller, J., Guiard, B., Pfanner, N., and Meisinger, C. (2007) Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway, *J Mol Biol 365*, 612-620.

63. Hell, K. (2008) The Erv1-Mia40 disulfide relay system in the intermembrane space of mitochondria, *Biochim Biophys Acta 1783*, 601-609.

64. Bihlmaier, K., Mesecke, N., Terziyska, N., Bien, M., Hell, K., and Herrmann, J. M. (2007) The disulfide relay system of mitochondria is connected to the respiratory chain, *J Cell Biol 179*, 389-395.

65. Muller, J. M., Milenkovic, D., Guiard, B., Pfanner, N., and Chacinska, A. (2008) Precursor oxidation by mia40 and erv1 promotes vectorial transport of proteins

into the mitochondrial intermembrane space, *Molecular biology of the cell 19*, 226-236.

66. Terziyska, N., Lutz, T., Kozany, C., Mokranjac, D., Mesecke, N., Neupert, W., Herrmann, J. M., and Hell, K. (2005) Mia40, a novel factor for protein import into the intermembrane space of mitochondria is able to bind metal ions, *FEBS Lett 579*, 179-184.

67. Hofmann, S., Rothbauer, U., Muhlenbein, 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.

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

69. Lisowsky, T. (1994) ERV1 is involved in the cell-division cycle and the maintenance of mitochondrial genomes in Saccharomyces cerevisiae, *Curr Genet 26*, 15-20.

70. Bushnell, G. W., Louie, G. V., and Brayer, G. D. (1990) High-resolution three-dimensional structure of horse heart cytochrome c, *J Mol Biol 214*, 585-595.

71. 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.

72. 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 engineering 14*, 993-1000.

73. Hartley, J. L., Temple, G. F., and Brasch, M. A. (2000) DNA cloning using in vitro site-specific recombination, *Genome Res 10*, 1788-1795.

74. Zhang, Y. M., Wu, B., Zheng, J., and Rock, C. O. (2003) Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction, *J Biol Chem 278*, 52935-52943.

75. Patel, R., Pollner, R., de Keczer, S., Pease, J., Pirio, M., DeChene, N., Dafforn, A., and Rose, S. (2000) Quantification of DNA using the luminescent oxygen channeling assay, *Clinical chemistry 46*, 1471-1477.

76. Karlsson, R., Michaelsson, A., and Mattsson, L. (1991) Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system, *J Immunol Methods 145*, 229-240.

77. Karlsson, R., and Larsson, A. (2004) Affinity measurement using surface plasmon resonance, *Methods Mol Biol 248*, 389-415.

78. Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., 2nd, Tsutakawa, S. E., Jenney, F. E., Jr., Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S. J., Scott, J. W., Dillard, B. D., Adams, M. W., and Tainer, J. A. (2009) Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS), *Nat Methods 6*, 606-612.

79. Putnam, C. D., Hammel, M., Hura, G. L., and Tainer, J. A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining

accurate macromolecular structures, conformations and assemblies in solution, *Q Rev Biophys 40*, 191-285.

80. Shiozawa, K., Konarev, P. V., Neufeld, C., Wilmanns, M., and Svergun, D. I. (2009) Solution structure of human Pex5.Pex14.PTS1 protein complexes obtained by small angle X-ray scattering, *J Biol Chem 284*, 25334-25342.

81. Blobel, J., Bernado, P., Svergun, D. I., Tauler, R., and Pons, M. (2009) Low-resolution structures of transient protein-protein complexes using small-angle X-ray scattering, *J Am Chem Soc 131*, 4378-4386.

82. Rambo, R. P., and Tainer, J. A. Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering, *Current opinion in structural biology 20*, 128-137.

83. Habel, J. E., Bursey, E. H., Rho, B. S., Kim, C. Y., Segelke, B. W., Rupp, B., Park, M. S., Terwilliger, T. C., and Hung, L. W. (2010) Structure of Rv1848 (UreA), the Mycobacterium tuberculosis urease gamma subunit, *Acta Crystallogr Sect F Struct Biol Cryst Commun 66*, 781-786.

84. Konarev, P. V., Volkov, A.V., Sokolova, M.H.J., Koch, and D.I. Svergun (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 35*, 1277-1282.

85. Svergun, D. I. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria *J. Appl. Cryst. 25*, 495-503.

86. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering, *Biophysical journal 80*, 2946-2953.

87. Volkov, V. V., and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering, *J Appl Crystallogr 36*, 860-864.

88. Liu, Z. J., Tempel, W., Ng, J. D., Lin, D., Shah, A. K., Chen, L., Horanyi, P. S., Habel, J. E., Kataeva, I. A., Xu, H., Yang, H., Chang, J. C., Huang, L., Chang, S. H., Zhou, W., Lee, D., Praissman, J. L., Zhang, H., Newton, M. G., Rose, J. P., Richardson, J. S., Richardson, D. C., and Wang, B. C. (2005) The high-throughput protein-to-structure pipeline at SECSG, *Acta Crystallogr D Biol Crystallogr 61*, 679-684.

89. Chayen, N. E., Stewart, P. D. S., and Baldock, P. (1994) New Developments of the Impax Small-Volume Automated Crystallization System, *Acta Crystallogr D 50*, 456-458.

90. Wu, C.-K., Dailey, T. A., Dailey, H. A., Francavilla, A., Starzl, T. E., Wang, B. C., and Rose, J. P. (2000) Expression, Purification, Crystallization and Preliminary X-ray Analysis of the Augmenter of Liver Regeneration, *J. Prot. Pept. Letts. 7*, 25-32.

91. Szebenyi, D. M., Arvai, A., Ealick, S., Laiuppa, J. M., and Nielsen, C. (1997) A system for integrated collection and analysis of crystallographic diffraction data, *J Synchrotron Radiat 4*, 128-135.

92. Matthews, B. W. (1968) Solvent content of protein crystals, *J Mol Biol 33*, 491-497.

93. Wang, B. C. (1985) Resolution of phase ambiguity in macromolecular crystallography, *Methods Enzymol 115*, 90-112.

94. Hendrickson, W. A., Pahler, A., Smith, J. L., Satow, Y., Merritt, E. A., and Phizackerley, R. P. (1989) Crystal structure of core streptavidin determined from multiwavelength anomalous diffraction of synchrotron radiation, *Proc Natl Acad Sci U S A 86*, 2190-2194.

95. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.

96. Borek, D., Minor, W., and Otwinowski, Z. (2003) Measurement errors and their consequences in protein crystallography, *Acta Crystallogr D Biol Crystallogr 59*, 2031-2038.

97. Read, R. J. (2003) Strengthening molecular replacement with maximum likelihood in Beast, *Crystallography Reviews 9*, 33-41.

98. Terwilliger, T. C. (2003) SOLVE and RESOLVE: automated structure solution and density modification, *Methods Enzymol 374*, 22-37.

99. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.

100. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis, *J Comput Chem 25*, 1605-1612.

101. Storoni, L. C., McCoy, A. J., and Read, R. J. (2004) Likelihood-enhanced fast rotation functions, *Acta Crystallogr D Biol Crystallogr 60*, 432-438.

102. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C., and Adams, P. D. (2008) Automated structure solution with the PHENIX suite, *Methods Mol Biol 426*, 419-435.

103. Wang, B. C., Adams, M. W., Dailey, H., DeLucas, L., Luo, M., Rose, J., Bunzel, R., Dailey, T., Habel, J., Horanyi, P., Jenney, F. E., Jr., Kataeva, I., Lee, H. S., Li, S., Li, T., Lin, D., Liu, Z. J., Luan, C. H., Mayer, M., Nagy, L., Newton, M. G., Ng, J., Poole, F. L., 2nd, Shah, A., Shah, C., Sugar, F. J., and Xu, H. (2005) Protein production and crystallization at SECSG -- an overview, *J Struct Funct Genomics 6*, 233-243.

104. Adams, M. W., Dailey, H. A., DeLucas, L. J., Luo, M., Prestegard, J. H., Rose, J. P., and Wang, B. C. (2003) The Southeast Collaboratory for Structural Genomics: a high-throughput gene to structure factory, *Acc Chem Res 36*, 191-198.

105. Baldock, P., Mills, V., and Shaw Stewart, P. D. (1996) A comparison of microbatch and vapor diffusion for initial screening of crystallization conditions, *J. Crystal Growth 168*, 170-174.

106. Chayen, N. E., Stewart, P. D. S., Maeder, D. L., and Blow, D. M. (1990) An Automated-System for Microbatch Protein Crystallization and Screening, *J Appl Crystallogr 23*, 297-302.

107. Lee, J. E., Fusco, M. L., Hessell, A. J., Oswald, W. B., Burton, D. R., and Saphire, E. O. (2008) Structure of the Ebola virus glycoprotein bound to an antibody from a human survivor, *Nature 454*, 177-182.

108. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis, *J Appl Crystallogr 36*, 1277-1282.

109. Svergun, D. I., Petoukhov, M. V., and Koch, M. H. J. (2001) Determination of Domain Structure of Proteins from X-Ray Solution Scattering, *80*, 2946-2953.

110. Volkov, V. V. a. S., D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering *J Appl Crystallogr 36*, 860-864.

111. Dryden, K. A., Wieland, S. F., Whitten-Bauer, C., Gerin, J. L., Chisari, F. V., and Yeager, M. (2006) Native hepatitis B virions and capsids visualized by electron cryomicroscopy, *Mol Cell 22*, 843-850.

112. Harris, A., Belnap, D. M., Watts, N. R., Conway, J. F., Cheng, N., Stahl, S. J., Vethanayagam, J. G., Wingfield, P. T., and Steven, A. C. (2006) Epitope diversity of hepatitis B virus capsids: quasi-equivalent variations in spike epitopes and binding of different antibodies to the same epitope, *J Mol Biol 355*, 562-576.

113. Seitz, S., Urban, S., Antoni, C., and Bottcher, B. (2007) Cryo-electron microscopy of hepatitis B virions reveals variability in envelope capsid interactions, *The EMBO journal 26*, 4160-4167.