DOUBLING ALBUMIN: *IN VIVO* CONSEQUENCES OF REITERATING ALBUMIN IN A SINGLE POLYPEPTIDE CHAIN

DOUBLING ALBUMIN: *IN VIVO* CONSEQUENCES OF REITERATING ALBUMIN IN A SINGLE POLYPEPTIDE CHAIN

By

TERESA ROSE. McCURDY, B.Sc.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the degree

Master of Science

McMaster University

Master of Science (2002) McMaster University
(Biology) Hamilton, Ontario

TITLE: Doubling Albumin: *In Vivo* Consequences of Reiterating Albumin in a Single Polypeptide Chain

AUTHOR: Teresa Rose McCurdy, B.Sc. (University of Waterloo)

SUPERVISOR: Dr. W.P. Sheffield

NUMBER OF PAGES: xv, 137

ABSTRACT

Objective. Albumin is an abundant and slowly-cleared plasma protein. Our laboratory previously incorporated albumin into recombinant fusion proteins to extend the plasma half-life of small proteins of potential therapeutic utility. We sought to determine if reiteration of albumin in a single polypeptide chain would further extend its half-life.

Design. Hexahistidine-tagged rabbit serum albumin (RSA) with Cys 34 altered to Ala to prevent disulphide-bonded dimerization was produced in yeast (H_6-RSA-C34A), and compared to a reiterated forma of H_6-RSA-C34A containing two domains of amino acids 1-584 separated by a hexaglycine spacer. Clearance of these purified iodinated proteins was compared in rabbits.

Materials and Methods. Site-directed mutagenesis employing PCR was used to alter the encoding plasmid. Proteins secreted from transformed *Pichia pastoris* yeast cell lines were purified using nickel-chelated affinity chromatography and radioiodinated by the Iodogen method. Labeled proteins were injected intravenously into rabbits and the residual acid-precipitable protein concentration in serial plasma samples was determined over time.

Results. *P. pastoris* cells transformed with the expression plasmids secreted 140 kDa H_6-diRSA and 70 kDa H_6-RSA-C34A proteins, which were purified to apparent homogeneity. Mean terminal catabolic half-lives (±SD) were 4.9 (±0.7) and 3.0 (±0.3)

days for H_6-RSA-C34A and H_6-diRSA, respectively. The n values were 9 for H_6-diRSA and 12 for H_6-RSA-C34A.

Conclusions. H_6-diRSA was cleared from the circulation more rapidly than H_6-RSA-C34A. We hypothesized that increased catabolism of the reiterated molecule could be due to an increased rate of cellular uptake and endocytosis of H_6-diRSA due to increased avidity for cellular binding sites. Non-reiterated albumin therefore appears to be optimal as a carrier protein for small recombinant blood products.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor, Bill Sheffield. I appreciate all the help you provided, the clear direction you gave, and the flexibility you allowed me that made it possible for me to complete this work.

I would also like to thank the members of my supervisory committee, Dr. Mark Hatton, and Dr. George Sorger, for their helpful suggestions and constructive criticism of the work contained in this thesis.

I would like to thank all my coworkers, especially Sharon Gataiance, for all her work with the rabbits, and Varsha Bhakta for her invaluable assistance with molecular techniques and troubleshooting. Thanks to Megan Begbie for helpful discussions about molecular biology in general and thesis writing in particular. The rest of the lab provided a welcome respite from the rigours of scientific work, in the form of coffee breaks including crossword puzzle solving and sparring. They were also most helpful in assuring me that I would in fact make it to this point and eventually graduate.

Finally I would like to thank my family. My mother, Marilyn Taracinski, was always sure I could meet any deadline, even when I wasn't. I appreciate all the support my husband, Chris, provided, and all the extra work he put in when I had deadlines. I hope that I have shown my boys, Andrew, Colin, Jacob, and Mitchell, that with perseverance and hard work you can achieve any goal.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

FIGURES:

TABLES:

ABBREVIATIONS

ATP	Adenosine triphosphate
BMM	Buffered Minimal Methanol medium
BSA	Bovine Serum Albumin
DNA	Deoxyribonucleic Acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ER	Endoplasmic Reticulum
HSA	Human Serum Albumin
IgG	Immunoglobulin G
IgM	Immunoglobulin M
kDa	Kilodaltons
LB	Luria Broth
MWCO	Molecular Weight Cutoff
MYA	Million Years Ago
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
ROS	Reactive Oxygen Species
RSA	Rabbit Serum Albumin

rRSA	recombinant Rabbit Serum Albumin
pdRSA	plasma-derived Rabbit Serum Albumin
SDS	Sodium Dodecyl Sulphate
TCA	Trichloroacetic acid
TFPI	Tissue Factor Pathway Inhibitor
tPA	Tissue Plasminogen Activator
Tris base	Tris(hydroxymethyl)aminomethane
TBST	Tris Buffered Saline with 0.05% Tween-20
YPD	Yeast Peptone Dextrose medium
YPDS	Yeast Peptone Dextrose with Sorbitol medium

INTRODUCTION

1.1 Albumin

1.1.1 General Features

Albumin is the most abundant plasma protein, and comprises 60% (w/v) of the protein in serum. The concentration of albumin in serum is 42 ± 3.5 g/L (Ohtani *et al.*, 1998). Albumin is a very soluble and stable protein which is relatively resistant to heat and acidic denaturation. It is highly negatively charged, carrying a net charge of -15 at pH 7.0, with an isoelectric point of 4.8 (Peters, 1985). Human serum albumin (HSA) is a 585 amino acid, single chain, non-glycosylated plasma protein (MW≈67 kDa) (Dugaiczyk *et al.*, 1982). Albumin is one of the few secreted plasma proteins that does not undergo glycosylation as a posttranslational modification. Translation of the mRNA produces a molecule with an N-terminal signal peptide and a "pro" sequence, which is cleaved before secretion (Dugaiczyk *et al.*, 1982). Albumin is synthesized by hepatocytes of the liver and is secreted into the circulatory system (Dugaiczyk *et al.*, 1982).

1.1.2 Evolution

The albumin gene (abbreviated ALB) is one member of a multigene family that has arisen through duplications of a single ancestral sequence. The other members of this

family of genes are α-fetoprotein (AFP), vitamin D binding protein (DBP, also known as group-specific component or GC), and α-albumin (ALF). It is estimated that gene duplication from an ancestral gene gave rise to the DBP gene some 580 Myr ago (MYA). About 295 MYA, a mutant DBP gene arose via gene duplication, yielding the albumin gene and a gene ancestral to both α-fetoprotein and α-albumin. Another duplication, about 250 MYA, separated these two youngest members of the albumin gene family.

All these proteins share a triple domain structure, consisting mostly of alpha helices with disulphide bonds forming many loops. The three domains of each protein share homology, as do the individual proteins. The loop structure found in these proteins is similar to that found in globins, such as myoglobin and the α-hemoglobin chain. Brown (1976) hypothesized that albumin evolved by repeated duplication of a fragment of the globin gene, due to the similarity of the structure of the antiparallel helices forming the loops. He found that the sequence similarities between the loops of the globins and the albumins were at least as similar as myoglobin and hemoglobin to each other.

Alexander *et al.* (1984) postulated that the ancestral gene of this family arose through amplification of a 27 base pair repeat. Their work with mouse albumin and α-fetoprotein resulted in the elucidation of a consensus sequence of 27 nucleotides thought to have duplicated repeatedly to form the progenitor domain. This progenitor domain consisted of 5 exons and was the basis of the triple domain structure of these related proteins.

These genes are all expressed in the liver, but are developmentally regulated. ALB, AFP, and DBP are all expressed in the fetal liver. α-fetoprotein is no longer

expressed after birth, while the expression of albumin increases and is maintained at high levels in adulthood. α-albumin is only expressed after birth and continues to be expressed in adulthood. These genes are located together on the q arm of chromosome 4 in humans (Nishio *et al.*, 1996). The protein products of these genes are secreted into the bloodstream and are components of blood plasma.

1.1.3 Physiological and Pharmacological Role

Maintenance of colloid osmotic pressure and ligand transport are the primary functions of albumin (Peters, 1985). Albumin is, in large part, responsible for preventing fluid leakage out of the vessels of the circulatory system by providing protein content to maintain the proper osmotic balance. Reductions in albumin levels can be pathological if not compensated.

Hypoalbuminemia, or lack of albumin, can have one of several causes. Many disease states are accompanied by depressed albumin levels, such as glomerulonephritis, or cirrhosis of the liver, due to increased excretion in the former, or reduction of synthesis in the latter. Another cause of low albumin content in plasma is a genetic condition called "analbuminemia", characterized by an almost total lack of albumin. Surprisingly there are few clinical symptoms associated with this condition, other than general edema (Minghetti *et al.*, 1985). Investigations have shown that this condition is compensated for by an increase in the quantity of other plasma proteins as well as a sharp decrease in the catabolic rate for any existing albumin, including that therapeutically infused (Freeman, 1969). Albumin infusion is somewhat controversial and efficacy of treatment is

dependent on medical condition. It has been shown to worsen some conditions in septic patients associated with respiratory distress by increasing pulmonary fluid levels (Byrne *et al.*, 1992). Other studies indicate its use in fluid replacement and inflammatory response reduction in hemorrhagic shock (Horstick *et al.*, 2002).

Many metabolites and drugs are transported to tissues bound to albumin. Fatty acids, which would normally be insoluble, are transported in the blood by albumin (He and Carter, 1992). Other, naturally occurring ligands of albumin are bilirubin, amino acids, metals, and hormones. The binding of ligands to albumin shows some variation between species, dependent upon amino acid composition. The histidine in position three binds copper and nickel in many mammalian species including human and rabbit, but is absent in such species as dog, pig, and chicken (Peters, 1985). The nickel binding properties of histidine 3 have been exploited in this project as a method of purification of the recombinant proteins and enhanced by the incorporation of additional histidine molecules.

The ligand binding properties of albumin have been extensively studied. In addition to the endogenous ligands mentioned above, albumin binds a wide variety of pharmacological agents. Sudlow and colleagues (1975, 1976) were the first to investigate and map the drug binding sites of albumin that bear her name. Sudlow's drug binding sites I and II are anionic drug binding sites. Some of the first drugs found to bind in these areas were warfarin in site I, and ibuprofen in site II. Other analgesics and anti-coagulants have been found to bind in these and other sites, as well as other compounds such as anaesthetics and tranquilizers. Much of the present research has focused on

binding sites of fatty acids, with which drugs often compete. The two primary binding sites have been found to be hydrophobic pockets located in subdomains IIA and IIIA, respectively. Recent studies have identified as many as 11 distinct binding locations for fatty acids distributed throughout the albumin molecule (Bhattacharya *et al.*, 2000). Researchers continue to elucidate the nature of the binding sites in albumin since this binding has a profound effect on the pharmacokinetics of the ligand, especially since often more than 90% of the drug binds to albumin (Curry *et al.*, 1998).

1.1.4 Structure

The structure of HSA has been determined by crystallography to consist of three homologous domains folded into a heart shape. Each domain contains 10 α helices and is divided into two subdomains. Each A and B subdomain encompasses 6 and 4 of the helices, respectively. The series of nine loop-link-loop structures shown in Figure 1 is approximately 67% α-helix, with the remaining amino acids acids forming the turns and extended polypeptide (He and Carter, 1992). This three dimensional structure of albumin is maintained by the 17 intramolecular disulphide bonds formed by 34 of the 35 cysteines in the molecule. The 35^{th} cysteine in albumin is a free thiol and can be the site of intermolecular disulphide bonding, as well as binding of such metal ligands as mercury and gold. This pattern of disulfide bonding is strictly conserved among mammals. High sequence identity is evident between many of these species as well. Equine serum albumin was crystallized and its structure compared to HSA and other mammalian albumins by Ho *et al.* (1993). They found 76% identity between equine and human

serum albumins, although the former is two amino acids smaller than the latter. The crystallographic studies showed high similarity, both for the structures and the binding properties of the molecules.

The free thiol in albumin, residue 34, can react to form a homodimer, which has been detected in commercial albumin preparations (Nakano *et al.*, 1982) and urine samples from patients with nephrotic diseases (Beilby *et al.*, 1985; Doman *et al.*, 1980). The proportion of albumin that naturally occurs as a dimer *in vivo* in blood is unknown and difficult to measure due to the possibility that *ex vivo* samples, exposed to air, may not reflect the intravascular situation. It can, however, be detected in plasma by SDS-PAGE under non-reducing conditions. One of the key steps in engineering the constructs for the proteins in this project was elimination of this free thiol, in order to prevent *in vivo* dimerization, which could potentially skew the clearance results.

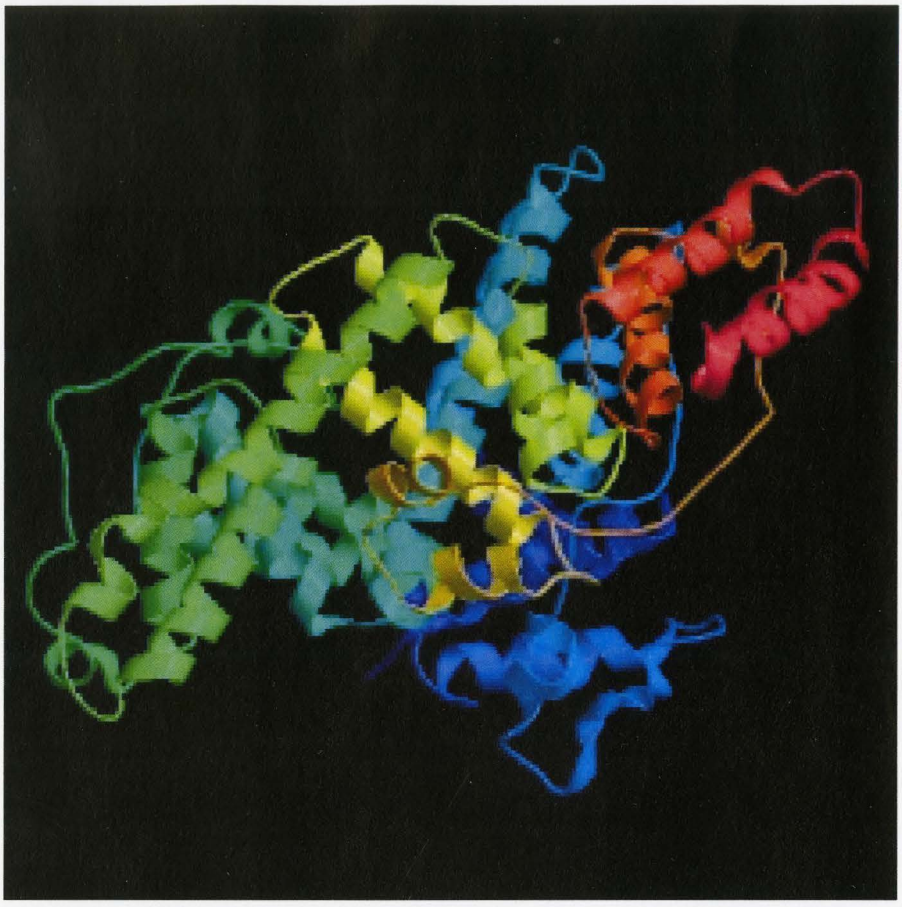
1.1.5 Rabbit Serum Albumin

Rabbit serum albumin (RSA) is a homologue of HSA. The two proteins share 73% sequence identity at the amino acid level. Mature RSA is smaller by one amino acid, the total of 584 is the same as that of other mammalian albumins such as rat, but one more than bovine and equine. The pre-pro sequence of rabbit albumin is identical to that of its human counterpart, and the 17 disulphide bonds, the single free cysteine, and the lone tryptophan are also conserved (Syed *et al.*, 1997).

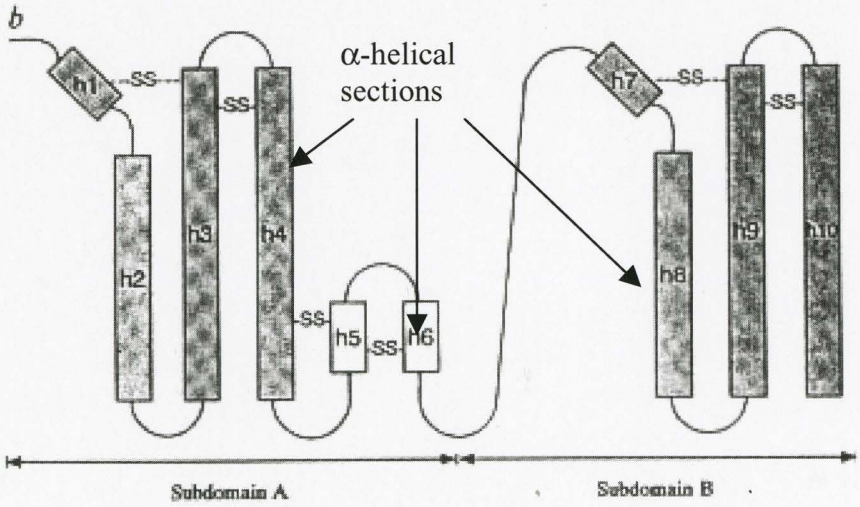
Figure 1: Structure of Albumin

Panel A shows a ribbon diagram of the molecular structure of Human Serum Albumin (HSA) from the Swissprot database. Panel B shows the loop-link-loop structure of the two subdomains contained within each albumin domain (modified from He and Carter, 1992).

A



B


b
α-helical
sections
h1
SS
SS
h2
h3
h4
SS
h5
SS
h6
h7
SS
SS
h8
h9
h10
Subdomain A
Subdomain B

1.2 Plasma Proteins

1.2.1 The Repertoire of Plasma Proteins

Many of the plasma proteins, other than albumin, tend to be either procoagulant or anti-coagulant. Most of these are enzymes and circulate as zymogens that must be cleaved to perform their enzymatic function. These proteins have been extensively studied in a number of different species in an effort to elucidate the process of haemostasis. They include such procoagulant factors as fibrinogen, prothrombin, factors VIII, IX and X, and tissue factor. Anti-coagulant factors include antithrombin, heparin cofactor II, and tissue factor pathway inhibitor (TFPI). Plasma also contains many immune system components, such as immunoglobulins and complement factors, that are also proteins. The internet plasma protein database (http://www-lecb.ncifcrf.gov/plasmaDB/) lists a total of 85 different plasma proteins identified by 2-D gel electrophoresis. Schwick and Heide (1977) quote the number of plasma proteins in excess of 200. As evidenced by those estimates, plasma consists of a complex mixture of proteins with a wide variety of functions.

1.2.2 Synthesis of Plasma Proteins

Most of the plasma proteins, including albumin, the coagulation and fibrinolytic factors, are synthesized in hepatocytes of the liver. The synthesis is directed to the surface of the endoplasmic reticulum (ER) by the signal peptide that these secreted proteins possess. This signal sequence is usually cleaved during protein synthesis by signal peptidases associated with the luminal membrane of the ER (Kleinsmith and Kish,

1995). Any N-linked glycosylation takes place here, during the formation of the polypeptide chain. The protein then moves into the Golgi body where any N-linked glycosylation may be further modified or O-linked glycosylation may take place. Membrane vesicles that bud off the Golgi complex transport the completed proteins to the surface of the cell where they are expelled by exocytosis. The pro sequence of some proteins, such as albumin, is cleaved in these vesicles, before secretion, during transport to the cell membrane (Kleinsmith and Kish, 1995). Many plasma proteins with enzymatic activity do not leave the cell fully mature, but are secreted as zymogens so that their function may be activated when needed by proteolytic cleavage. Other plasma proteins such as the immunoglobulins, are produced in plasma cells and lymphocytes (Franklin, 1970). In contrast to the constitutive expression of coagulation factors, immunoglobulin synthesis is governed by antigenic stimulation, with levels quite low in animals raised in a germ-free environment (Fahey and Sell, 1965).

1.2.3 General Catabolism of Plasma Proteins

Catabolism refers only to the breakdown of proteins into their component amino acids, and usually occurs within the lysosomes of cells. In general, there is evidence that size, charge and glycosylation affect the rate at which proteins are degraded. Dice and colleagues (1978) focused their attention on proteins that are intracellular residents. They suggested that under normal conditions larger cytoplasmic proteins were degraded more rapidly than smaller cytoplasmic ones. Catabolism was more rapid for glycoproteins than nonglycoproteins, and this was also true for acidic proteins versus either neutral or basic.

Of more relevance to this thesis, Dice and colleagues also reported a correlation between isoelectric point and rate of degradation for serum proteins, although it must be remembered that they worked with total serum rather than individual polypeptides. Other general factors affecting catabolism, that will be discussed further below, with reference to specific proteins, are metabolic rate and plasma concentration.

1.2.3.1 Immunoglobulins

Rate of catabolism in immunoglobulins, as well as other serum proteins is influenced by rate of metabolism. Smaller animals, such as mice, have a higher metabolic rate and their fractional catabolic rate per day for immunoglobulins is therefore higher. Large animals, such as cows and horses, have a lower catabolic rate due to their slower metabolism. Waldmann *et al.* (1970) have shown an inverse relationship between catabolism and body weight for IgG. There is a fair amount of variation among the different classes of immunoglobulins, and since IgG is the most plentiful, it will be mentioned with respect to research in this area. Another relationship noted by investigators of IgG is that between plasma concentration and catabolism. Freeman (1965) found that patients with low serum γ-globulin concentrations had decreased catabolic rates of this protein. Conversely, Lippincott *et al.* (1960) found that high serum concentrations of IgG resulted in progressively reduced half-lives (by an increase in catabolism).

1.2.3.2 Albumin

Factors affecting catabolism of albumin are similar to those mentioned for immunoglobulins. Plasma concentration of albumin affects the amount of this protein degraded each day. Kirsch *et al.* (1968) found that rats deprived of protein in their diet also experienced a drop in catabolic rate due to suppressed synthesis that lowered plasma concentration. This drop was gradually reversed when protein was added back into the diet. In rare cases of analbuminemia (genetic lack of albumin) an extremely low albumin catabolic rate has been noted, although the rate was found to gradually increase with the albumin pool size as infusions were given (Freeman, 1969). Dammacco *et al.* (1980) verified the lower catabolic rate by conducting albumin turnover studies in an analbuminemic patient. They found that ^{125}I-labeled albumin survived four times longer in the analbuminemic 30 year old female subject than it did in normal controls using experimental techniques similar to those contained in this thesis. Conversely, albumin infusion in normal humans or laboratory animals has been shown to result in an increase in catabolism. Hoffenberg (1970) points to a study of rabbits infused with surplus albumin over a 14-day period. Catabolism increased over the infusion period and continued after the infusions were stopped, until the pool size returned to normal. In this way the body can compensate for an increase in the pool of albumin or a decrease due to loss in a disease state. Immediate adjustments can also be made to plasma albumin concentration by shifting from one metabolic pool to another. A large amount of the total body albumin exists outside of the vascular space. Approximately 60% of the albumin in the human body is contained in the extravascular space (Dammacco *et al.*, 1980). This

pool can be shifted, on a temporary basis, to the intravascular pool, to aid in adjustment of plasma level.

1.2.3.3 Antithrombin

Antithrombin is a plasma protein whose function is inhibition of a number of coagulation factors. Its *in vivo* behavior has been the subject of many experiments, including a classic study by Carlson *et al.* (1984) who demonstrated its adherence to a three compartment model of clearance in the rabbit. These investigators showed that the behavior of antithrombin was consistent with its association with three physiologic pools, the plasma, the non-circulating vessel wall, and the extravascular space. Analysis of clearance data predicted catabolism from both the intravascular and extravascular compartments rather than degradation from the plasma alone.

1.3 Clearance of Proteins from the Circulation and the Body

1.3.1 Definition of Clearance

Clearance refers to the rate of loss of a specific protein from the plasma. A fraction of the concentration of a protein will extravasate from the intravascular space. In addition, a fraction will be degraded within the intravascular space. However, a fraction will return to the intravascular compartment from the extravascular space through the lymphatic system. Clearance can be analyzed using concepts in which proteins are regarded as residing in different compartments or pools within the body. This residence

is in a state of flux in which proteins can travel from one compartment to another reversibly and eventually are broken down irreversibly.

1.3.2 Factors Affecting Clearance

Clearance of proteins from plasma is a complex process, governed by a myriad of different factors. Protein size is one of the factors affecting residence time of proteins. Proteins under 40 kDa in size are usually short-lived in plasma due to filtration by the kidneys and excretion through the urine, unless they can bind tightly to larger proteins. Conventional wisdom dictates that the glomeruli of the kidney discriminates based on charge of the molecule as well as size, preferring to allow neutral and positively charged molecules to pass through into the urine (Brenner *et al.*, 1978). This theory has recently been disputed in a review by Russo *et al.* (2002). Much of the information obtained about charge selectivity was obtained from studies of charged polysaccharides during the 1970s, specifically, negatively charged dextran sulfate (Chang *et al.*, 1975, Bennett *et al.*, 1976). The dextran sulfate molecule was assumed to be excreted intact, but more recent findings point to the presence of desulphated dextran in the urine (Comper *et al.*, 1994). An alternate receptor-mediated route of clearance is suggested by Vyas *et al.* (1996). The apparent charge selectivity seen with dextran sulfate may in fact be attributable to uptake of this substance by glomerular endothelial cell receptors associated with heparan sulfate found on cell surfaces, and their subsequent desulphation by lysosomes. Once desulphated, the dextran is exocytosed and filtered by the kidney. This evidence supports the theory that the glomerular capillary wall discriminates only on the basis of

hydrodynamic size (Russo *et al.*, 2002). Megalin and cubilin appear to be the endocytic receptors responsible for uptake of various proteins in the proximal tubules of the kidney, including the relatively small fraction of albumin filtered by the kidneys. Only 25 mg of intact albumin are lost in the urine per day, and while the extent of albumin fragments in the urine is somewhat controversial, it likely does not exceed 20 grams per day; this compares to an estimated 125 grams of albumin in the circulation, and 250 grams in the tissues (Gburek *et al.* 2002; Russo *et al.*, 2002).

Other factors affecting clearance include the presence of receptors associated with the protein of interest, as well as susceptibility to proteases in the plasma. Figure 2 summarizes the alternate pathways by which a protein injected into the bloodstream may be cleared from the body. The presence of a receptor may hasten the clearance of a protein from the plasma compartment. Degradation of proteins by plasma proteases to fragments under the glomerular filtration threshold size results in fairly rapid excretion through the urine (Brenner *et al.*, 1978). Uptake by cells may result in catabolism with the amino acids generated by this process re-used in synthesis of other proteins.

1.3.3 Use of Radiolabeled Proteins to Study Clearance in Mammals

Use of radioactive isotopes such as ^{15}N to study plasma protein metabolism was used in early studies by researchers such as Schoenheimer and Rittenberg in the 1940s (Regoeczi, 1984). Such radioactive elements had serious drawbacks relating to their use to label proteins. Since nitrogen is reutilized in the body, it was difficult to follow the catabolism of only the originally labeled protein. In the 1950s researchers began to try to

combine proteins with radioactive forms of iodine. Iodination of proteins has several advantages which have made this technique the most popular for use in tracer experiments. Radiolabeling proteins with iodine occurs readily, it is stable, and there are several isotopes available for experiments which require double labeling (Regoeczi, 1984).

Radioactive iodine is also a useful label because it has a high specific activity and a small amount of labeled protein can be used to follow metabolic processes. The idea of tracer experiments is to introduce a small amount of labeled protein into the intravascular compartment and follow its disappearance from this pool. The intravascular compartment is regarded as the central compartment which is connected with the other compartments reversibly, although if excretion is thought of as a compartment, it is not a reversible one (Matthews, 1957). The small amount of injected material should not interfere with normal metabolic processes since it is only a fraction of the natural substance, but should reflect its activity. An injected, radiolabeled plasma protein should (and is assumed) to behave in the same manner as the native protein in respect to its compartmental distribution, intercompartmental rates of flux and rates of catabolism.

1.3.4 Compartmental Analysis of Clearance and Methodological Issues

The most commonly used clearance models for plasma proteins are the two- and three-compartment models. The choice of model is dictated by the shape of the curve derived from the disappearance of radioactivity over time from plasma of a labeled, injected protein. A protein that clears very quickly will likely only require two

exponential terms to describe its clearance, while a longer lived protein, such as albumin, will best be described by three. The equation describing such a curve is given by:

$$y(t)=C_1e^{-(a1)t} + C_2e^{-(a2)t} + C_3e^{-(a3)t},$$

where the exponents a_n are rate constants, with the first being the fastest and the last the slowest, and C_n are coefficients. Each term including C_n signifies the contribution to the curve of each compartment and the fitted curve is normalized so that $C_1 + C_2 + C_3$ equal unity (Carson and Jones, 1979).

In order to accurately assess the terminal catabolic half-life of the protein under study, blood samples must be taken and analyzed for a period long enough to determine the final exponential portion of the curve. This allows calculation of the final term because the contribution of the first two terms in the equation will be negligible at this point. Once this term is solved, it can be subtracted from the whole curve and the unknowns in the exponential term can be solved for by calculating the slope and y-intercept. By successively "peeling" these terms from the curve, all the terms in the equation can be solved (Carson and Jones, 1979).

Solution of the terms in the equation allows further calculation of values for pharmacokinetic parameters such as fractional catabolic rates, fractional distributions and half-lives for each compartment.

1.4. Artificial modifications of proteins that modulate clearance

1.4.1 Acceleration of clearance

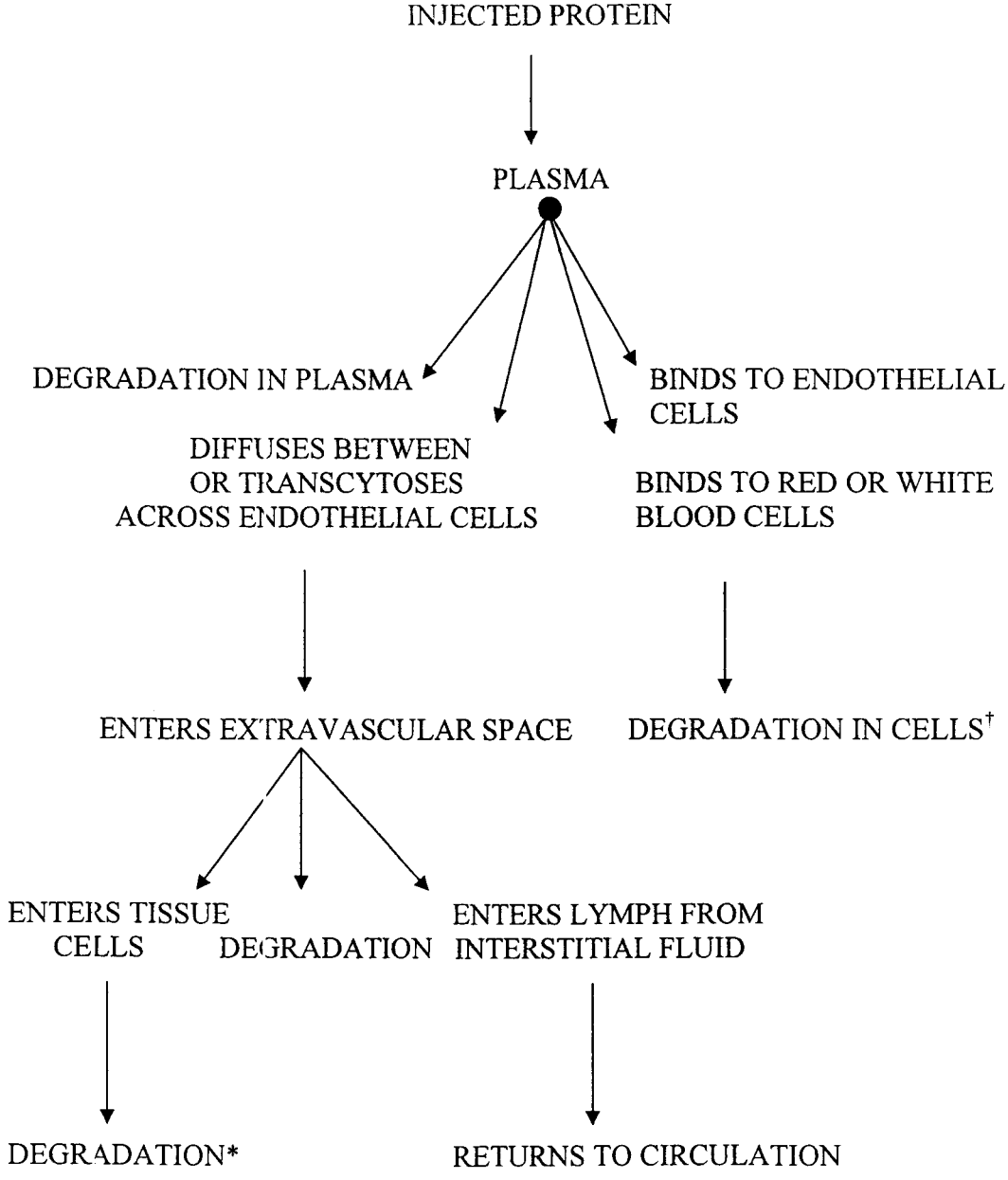
In most attempts at clearance modulation the goal is prolongation of clearance, but clearance acceleration has been achieved by chemical modification of some

FIGURE 2: Protein Clearance

This chart illustrates the variety of paths by which an injected protein may clear from the body.

PROTEIN CLEARANCE

POSSIBLE FATES OF AN INJECTED PROTEIN



*PARTIALLY OR FULLY DEGRADED PROTEIN LEAVES THE BLOOD IN THE KIDNEY AND ENTERS THE URINE

†ENDOCYTOSED PROTEIN CAN BE DEGRADED TO LIBERATE AMINO ACIDS FOR RE-USE

therapeutic proteins. This is often a secondary effect of the primary goal of targeting the therapeutic agent to a specific cell or organ. An example of this targeting is the experimental work by Yabe et al. (1999) in which catalase was succinylated, mannosylated and galactosylated in an effort to target this antioxidant to the site of reactive oxygen species (ROS) damage in ischemia/reperfusion injuries of the liver. The uptake of these modified catalase molecules by the mannose and scavenger receptors on the non-parenchymal cells was achieved, and the half-lives were significantly decreased in the process.

1.4.2 Slowing of clearance

Various methods of retarding clearance have been employed, including chemical crosslinking to larger molecules to increase the effective size of the molecule and manipulation of the gene to fuse with another moiety or remove a domain or residue(s) which is responsible for accelerated clearance (reviewed by Sheffield, 2001).

1.4.2.1. Chemical modifications

Chemical modifications have included PEGylation, addition of dextran molecules, and crosslinking to albumin. PEGylation involves addition of polyethylene glycol (PEG) chains to a protein. This approach was used by Hershfield *et al.* (1987) to prolong the half-life of adenosine deaminase to treat its genetic deficiency. This treatment has been used successfully on greater than 80 patients worldwide suffering from this rare disease.

Serum albumin is the ideal fusion partner for small therapeutic molecules since it has a long half-life, is widely distributed throughout the body, and is non-immunogenic. The use of yeast as an expression system has led to the production of large quantities of recombinant albumin Investigators have used a number of different yeasts to produce albumins, as well as other recombinant proteins, but *Pichia pastoris* appears to be the optimal system (Rosenfeld, 1999). *P. pastoris* is a methylotrophic yeast, able to use methanol as its sole carbon source. This fact, along with its ability to properly fold and secrete proteins, as well as carry out post-translational modifications has led to its widespread use. The most attractive feature of this expression system is the ability of the yeast to produce recombinant proteins on the scale of grams per litre by exploitation of the methanol inducible alcohol oxidase promoter (Rosenfeld, 1999). Ohtani *et al.* (1998) investigated HSA produced by the *P. pastoris* system and found no significant difference in physical or immunological properties from that of plasma derived HSA.

1.5.2 Rabbit serum albumin (RSA)

1.5.2.1 Altered clearance in diabetic rabbits

Uncontrolled diabetes results in a high concentration of glucose in the blood, in excess of 7.8 mM (Andreoli et al, 1993). Albumin is susceptible to glycation (covalent attachment of sugars, non-enzymatically) in the presence of this glucose. *In vivo* studies in rabbits have shown that the naturally glycated form of albumin is more rapidly cleared than the unglycated form in both normal and diabetic rabbits (Hatton *et al.*, 1993). The increased clearance is due to increased catabolism of glycated albumin. Conversely, unglycated albumin is catabolized more slowly in diabetic rabbits than in normal.

Glycation also causes increased association of the protein with the vascular wall and increased migration to the extravascular space. This suggests that alterations to the optimal conformation of plasma proteins such as albumin may result in their accelerated clearance.

1.5.2.2 Recombinant RSA (rRSA)

Like investigators working with human recombinant albumin, our laboratory found no difference in the properties of recombinant rabbit serum albumin (rRSA) from that of plasma-derived RSA (pdRSA). Superimposable clearance curves were obtained from *in vivo* experiments in rabbits (2.8-3.2 kg) with radiolabeled, COS cell-derived rRSA ($t_{1/2}$=4.32±0.25 days) and pdRSA ($t_{1/2}$=4.7±0.36 days) (Sheffield *et al.*, 2000). In a separate study comparing *P. pastoris* derived $RSAH_6$ (rRSA with a histidine tag at the C-terminal end) ($t_{1/2}$=4.2±0.3 days) with pdRSA ($t_{1/2}$=4.9±0.4 days) no significant difference in pharmacokinetic properties between these forms of RSA were found in rabbits (2.8-3.2 kg) (Sheffield *et al.*, 2001).

1.5.2.3 Effects of truncation or glycosylation

Experiments in our laboratory to investigate the effect on clearance of truncated or N-glycosylated albumin were conducted with recombinant radiolabeled rabbit proteins *in vivo* in rabbits (2.8-3.2 kg). Truncated versions of albumin were produced in COS cells using as the truncation point the division between interdomain boundaries. Three truncated proteins were produced and tested, one consisting of the first domain of

albumin (amino acids 1-185), one consisting of the first and second domains (1-377), and one using only the third domain (378-584). All of these mini-albumins exhibited accelerated clearance, from an average of 4 days for full-sized albumin to an average of 1.5 hours for the truncations. There was no significant difference in clearance time between the single domain protein and the protein containing two albumin domains. Organ distribution investigation suggested a renal route of clearance for all truncated forms (Sheffield *et al.*, 2000).

Mutations to albumin were introduced at two different sites to create the consensus sequence for N-linked glycosylation. Sites that required the least changes in the coding sequence were chosen. The site near the C-terminus mimicked the naturally occurring Albumin Casebrook D494N variant in humans and the N-terminal site mutated the valine at position 14 to threonine. The mean half-life of the D494N variant was reduced to 2.87 ± 0.23 days while V14T showed no significant difference in clearance. It was suggested that the former glycosylation rendered the albumin molecule more susceptible to proteolysis, while the latter site may not be accessible due to less surface exposure once protein folding has occurred (Sheffield *et al.*, 2000). Again, these results point to the possibility that any modifications in the structural features of albumin lead to increased clearance.

1.5.2.4. Genetic fusion to rRSA

1.5.2.4.1 Hirudin

Hirudin is a small protein derived from leech salivary glands. It is the most potent natural thrombin inhibitor known (Markwardt *et al.*, 1993). Hirudin is quickly cleared from the body via a renal route; its catabolic half-life is one hour or less. Our laboratory investigated a fusion protein based on the hypothesis that linking hirudin to albumin would prevent its rapid excretion by increasing its size above the glomerular filtration limit (and thereby retarding its clearance). Two geometries were initially tested for the hirudin-albumin fusion protein expression in COS-1 cells: hirudin linked through its C-terminus to the N-terminus of albumin (HLA, Hirudin-Linker-Albumin); and the reverse, hirudin linked through its N-terminus to the C-terminus of albumin (ALH, Albumin-Linker-Hirudin)..HLA was expressed at higher levels than ALH, exhibited thrombin inhibition similar to that of unfused hirudin, and acquired the delayed *in vivo* clearance profile of albumin (Syed *et al.*, 1997).

A yeast system was used in preference to the mammalian COS-1 expression system to obtain sufficient quantities of fusion protein to test anticoagulant function *in vivo*. HLA was shown to retain the expected anticoagulant activity of the hirudin moiety *in vitro*. The clearance behavior of yeast-derived HLA was shown to be indistinguishable from that of either rRSA or pdRSA (Sheffield *et al.*, 2001). Its clearance did not appear to be influenced by the presence of the unnatural hexaglycine linker whose purpose was to allow independent folding of the separate entities of the molecule.

1.5.2.4.2 Barbourin

To test if the clearance delay observed with the hirudin-albumin fusion protein was a general phenomenon, another small molecule with anti-thrombotic properties was fused to albumin. Barbourin, an anti- platelet disintegrin derived from a snake venom (*Sistrurus barbourii*) was attached to albumin in a similar configuration to that of HLA and produced in *P. pastoris*. *In vivo* studies of this novel protein provided evidence of a significantly retarded clearance rate when compared to recombinant barbourin expressed in the same yeast system. Both the recombinant barbourin and the fusion protein retained their full ability to inhibit platelet aggregation *in vitro* (Marques *et al.*, 2001). Again it appeared that the conformation of the albumin portion of the molecule remained sufficiently unchanged to confer delayed clearance on the product of the fusion of barbourin with albumin.

1.6 Purpose of investigation

Our laboratory has conducted several studies aimed at modifying the clearance of small proteins by their fusion to full length albumin. Fusion to albumin, at the DNA level, resulted in a decreased clearance time for each small protein tested. However, drastic alterations in the structure of albumin, such as truncation, resulted in accelerated clearance. Mixed clearance results were obtained by less severe manipulations such as glycosylation, depending on the site of alteration of the protein. Increasing the size of a protein above the glomerular filtration threshold has been shown to significantly retard clearance. Proteins above this threshold have variable catabolic half-lives: e.g. HSA with

a size of approximately 70 kDa at 19 days (Peters, 1985); IgG, approximately 150 kDa, at 28 days (Soloman *et al.*, 1963); and rabbit fibrinogen with an MW of 350 kDa at 2.2 days (in rabbits) (Hatton *et al.*, 1993). The purpose of this study, therefore, was to investigate the alteration in clearance of an albumin which is twice its normal size (i.e. from 67 kDa to approx. 135 kDa), with a view to further extending the half-life of fused peptides. In addition, from a physiological point of view, a secondary objective was to investigate the clearance of natural albumin by constructing a recombinant albumin incapable of disulphide-bonded dimerization. By investigating the *in vivo* clearance of this mutant protein, it was anticipated that insights as to the contribution of dimerization to the albumin clearance profile could be obtained.

MATERIALS AND METHODS

2.1 Sources of chemicals and reagents

Restriction and DNA modification enzymes were purchased from either MBI Fermentas Inc. (Hamilton, ON) or Fisher Scientific (Unionville, ON). The Institute for Molecular Biology and Biotechnology (MOBIX), McMaster University (Hamilton, ON) synthesized the oligonucleotides and provided DNA sequencing services. *Pichia pastoris* strain X-33, Zeocin™, ThermalAce™ high fidelity DNA polymerase, and competent *Escherichia coli* DH5α cells were all purchased from Invitrogen (Carlsbad, CA). The GeneEditor™ site-directed mutagenesis kit, plasmid pGEM5zf(+), and alkaline phosphatase conjugated anti-mouse IgG were bought from Promega Biotech (Madison, WI). Amersham Biosciences (Baie d'Urfé, PQ) was the source for T7 sequencing kits, T4 DNA ligase, mouse anti-hexahistidine antibody, and Sepharose 6B gel filtration chromatography resin. The Affi-Gel® Blue chromatography resin, gel filtration standards, and polyvinylidene difluoride (PVDF) membrane used in amino acid sequencing were purchased from Bio-Rad Laboratories (Hercules, CA). The Nickel-NTA affinity chromatography resin, DNA purification kits, and *Taq* DNA polymerase were obtained from Qiagen (Chatsworth, CA). The alkaline phosphatase-conjugated goat anti-chicken IgG was purchased from Kirkegaard and Perry Laboratories (Gaithersburg, MD). Pall Gelman (Ann Arbour, MI) was the supplier for the nitrocellulose membrane used for immunoblots. The EZ-Link™ PEO-Maleimide activated biotin was purchased

from Pierce Biotechnology (Rockford, IL). $Na^{125}I$ and $Na^{131}I$ used for protein labeling was purchased from Perkin Elmer Life Sciences (Boston, MA), while ^{35}S-ATP used for sequencing was purchased from Mandel Scientific (Toronto, ON). All chemicals not specifically listed above were of the highest grade commercially available.

2.2 Experimental Methods

2.2.1 Analysis and Manipulation of DNA

The majority of experimental tasks involving the manipulation of DNA involved standard protocols that have been previously described (Maniatis *et al.*, 1989). A more expanded description of these procedures, as well as a full description of any modifications or novel protocols, follows.

2.2.1.1 Growth of Bacteria and Purification of plasmid DNA

For standard molecular biological tasks such as the assembly of novel plasmid-borne DNA constructs, *Escherichia coli* (*E.coli*) DH5α cells were employed. These cells were purchased in competent form and transformed to ampicillin resistance following the manufacturer's instructions (Invitrogen), which involved three steps. These were incubation with plasmid DNA on ice, heat shock for 20 seconds at 37°, recovery in the presence of ampicillin-free Luria Broth (LB), and plating on LB/0.7% agar plates containing 0.1 mg/ml sodium ampicillin. A resulting, isolated colony, or, alternatively, a frozen, glycerinated culture derived from such a colony, was used to inoculate 2.5 ml of LB/0.1 mg/ml sodium ampicillin (LB/ampicillin) culture, which was grown overnight at

37°, rotating at 200 RPM in an environmental shaker. A rapid DNA mini-preparation procedure (Zhou *et al.*, 1990), a phenol-free alkaline lysis protocol, was used to isolate closed circular plasmid DNA for subsequent restriction analysis. Higher purity plasmid DNA, for example for automated DNA sequencing (see below), was obtained using a Qiaprep™ Spin Mini-Prep Kit as directed by the manufacturer (Qiagen). For applications in which larger quantities of plasmid DNA were required, for instance to serve as substrates for preparative restriction endonuclease digestions, a 200 ml LB/ampicillin culture was inoculated and grown overnight to stationary phase, and plasmid DNA purified using a Qiagen Plasmid Maxi Kit (Qiagen). All Qiagen DNA isolation products employ modified alkaline lysis protocols and a proprietary DNA-binding resin.

2.2.1.2 Enzymatic manipulation and electrophoresis of DNA

Purified plasmid DNA or PCR products (see below) were subjected to restriction endonuclease digestion according to the manufacturer's instructions (Fisher Promega), typically employing 2-10 units of enzyme per microgram of DNA, and under no circumstances using greater than 10% (vol/vol) enzyme. Prior to restriction endonuclease digestion, PCR products were ethanol precipitated using two volumes of absolute ethanol and 0.1 volumes of 7.0 M ammonium acetate in the presence of 10 μg/ml glycogen as a carrier. Restriction endonuclease digestion products were analyzed using agarose gel electrophoresis, typically employing 1% (weight/volume) agarose gels, using 89 mM Tris/89 mM acetic acid/2 mM EDTA, pH 8.0 (TAE buffer), in a submarine configuration.

For preparative isolation of restriction endonuclease digestion products, QIAquick™ gel purification kits (Qiagen) were used, following the manufacturer's instructions.

2.2.1.3 Amplification of DNA using the Polymerase Chain Reaction (PCR)

A reaction mixture consisting of 0.4 ng/μL template, 1.0 μM of each of two oligonucleotide primers, 1X *Taq* amplification buffer (providing 0.2 mM each deoxynucleotide triphosphate (dNTP), 1.5 mM $MgCl_2$, unspecified KCl and $(NH_4)_2SO_4$ and proprietary components) in a total volume of 50 μL was mixed and placed in a thermocycler. The mixture was heated to 95°C for 5 minutes before the addition of 5 units of *Taq* polymerase (Qiagen). This addition was followed by 40 cycles of: 1 minute of denaturation at 95°C, 1.5 minutes of annealing at 55°C, 2 minutes of extension at 72°C, and a final 10 minute extension at 72°C.

For some constructs, an alternative PCR procedure was employed, which differed from that described above only in the substitution of ThermalAce™ high fidelity DNA polymerase and amplification buffer (Invitrogen) for *Taq*. Thermocycling conditions were: initial denaturation, 3 minutes at 95°C, followed by 40 cycles of 30 seconds denaturation at the same temperature, 30 seconds of annealing at 55°C, 30 seconds of extension at 74°C, and the 10 minute final extension at the same temperature.

Additional details concerning which protocol was used for which construct, as well as the identity of the oligonucleotides employed in each case, are provided below.

2.2.1.4 Ligation of DNA

Ligation of DNA fragments was carried out either at 4^0C overnight or for 2 hours at 16°C. A 3:1 molar ratio of insert DNA to plasmid vector DNA was employed. Ligation reactions consisted of these linear DNA molecules, 0.1 - 0.3 Weiss units/ μL T4 DNA ligase, in 30 mM Tris-HCl (pH 7.8), 10 mM $MgCl_2$, 10 mM DTT, and 1 mM adenosine triphosphate (ATP) in a total of 20 μL reaction volume. Five μL of each 20 μL reaction was used to transform competent *E. coli* DH5α cells as described above.

2.2.1.5 DNA Sequencing

The Sanger method (1977) for DNA sequencing, which uses a dideoxy chain terminating procedure, was used to screen clones for introduced mutations, using a T7 Sequencing Kit (Amersham Biosciences) and α-[^{35}S]-dATP. Automated sequencing was also carried out on the albumin-encoding region of each cDNA construct by MOBIX. In both cases double-stranded plasmid DNA was sequenced, using chemical and thermal denaturation, respectively.

2.2.2 Construction of Expression Plasmids

Three novel expression plasmids were assembled in the course of this work. The term "expression plasmid" is used here to indicate a plasmid that will direct the heterologous expression of a specific protein, as distinct from an "expression vector" which requires the insertion of a specific cDNA to execute this function.

2.2.2.1 Generation of the pPZ-H_6RSA Expression Plasmid

The previously constructed pSG5-RSA expression vector (Syed *et al.*, 1997) was used as the template for the PCR-mediated addition of six His codons designed to lie in-frame to both an upstream prepro- sequence and a downstream RSA open reading frame. The *Taq*-based protocol described in Section *2.2.1.3* was employed, using forward (sense) primer 24294, 5'-TCT*CTCGAG*A AAAGACACCA TCACCATCAC CATAAAAGTG AGATTGCTCA C-3', and reverse primer 5775, 5'-GAGTCGACCA GACATGATA A-3'. The former added a *XhoI* restriction site (italicized and underlined) and the hexahistidine codons, while the latter flanked the cloning sites of pSG5. The resulting 2053 bp PCR product was inserted into the *EcoRV* site of pGEM5zf(+) (Fisher Scientific) after being made blunt-ended through repair of overhanging ends using the Klenow fragment of DNA polymerase I.

Following transformation of *E. coli DH5α* as described in Section 2.2.1.1, colonies that were candidates for those harbouring the desired novel plasmid were selected by the technique of blue-white screening (Maniatis *et al.*, 1989). Briefly, candidate colonies were white rather than blue on indicator plates (LB/agar/ampicillin plates surface treated with 100 mM isopropylthiogalactoside and 10 mM 5-bromo-4-chloro-3-indolyl-beta-D-galactoside). Double-digestion of mini-preparations of plasmid DNA with *SacI* and *EcoRI* identified a candidate clone that was designated pGEM5-H_6RSA, after confirmation of its DNA sequence from the beginning of the hexahistidine tag to the *HindIII* site. The corresponding 561 bp *XhoI* - *HindIII* restriction fragment of pGEM5-H_6RSA was combined with a 1322 bp *HindIII* - *EcoRI* fragment of pAlb192

(Syed *et al.*, 1997) and inserted between the *XhoI* and *EcoRI* sites of the *Pichia pastoris* yeast expression vector pPICZamp9ss (Sheffield *et al.*, 2001). The ligation reaction that was used was identical to that described in Section *2.2.1.4*, except that three DNA fragments were combined, using an insert:insert:vector ratio of 3:3:1. This approach minimized the amount of *Taq*-polymerase-amplified DNA that had to be validated by new sequencing, due to the error-prone nature of this polymerase. Following ligation and transformation, one resulting clone was characterized by restriction enzyme mapping and partial DNA sequencing, and designated pPZ-H_6RSA. This plasmid was used both to transform *Pichia pastoris* yeast to Zeocin resistance and to generate the pPZH_6RSA-C34A expression plasmid, as described, respectively, below.

2.2.2.2 Generation of the pPZ-H_6RSA (C34A) Expression plasmid

Plasmid pGEM5-H_6RSA was used as the template for PCR-mediated mutagenesis in which the reverse primer altered codon 34 from TGC (Cys) to GCC (Ala) without altering the nearby *SphI* restriction endonuclease site. The ThermalAce™ -based procedure described in Section 2.2.1.3 was employed, using oligonucleotide 24294 (described above) as the forward primer and oligonucleotide 24730 (5'-CTAACTTC*GC ATGC*TCTTCA TATGGGGCCT TCTGGAGATA CTGAG-3'), as the reverse, or antisense primer (with *SphI* site underlined) to yield a 151 bp PCR product. Following restriction enzyme digestion with *XhoI* and *SphI*, the resulting 138 bp fragment was combined with a 1780 bp partial *SphI* - *EcoRI* restriction fragment of pH6-RSA and inserted between the *XhoI* and *EcoRI* sites of pPICZamp9ss in a manner exactly

analogous to that employed in Section 2.2.2.1 for pPZ-H_6RSA. The resulting plasmid was validated by both manual and automated DNA sequencing, and designated pPZ-H_6RSA (C34A). It differed from pPZ-H_6RSA at only a single codon, codon 34.

2.2.2.3 Generation of the pPZ-H_6diRSA Expression Plasmid

The following steps were taken to produce a *Pichia pastoris* yeast expression plasmid directing the synthesis of two copies of the RSA cDNA, fused in-frame with no intervening termination codon. In this thesis, the protein product of this expression plasmid is referred to as "dialbumin", or "diRSA", abbreviated H_6diRSA. The first step in reiterating the H_6RSA-C34A cDNA involved introducing an *NcoI* site at the 3' end by PCR. The 24294 primer was again used as the forward (sense) primer, and oligonucleotide 5454, 5'-AG*CCATGG*CT TTACTTGATT CAAC-3' (*NcoI* site underlined), as the reverse (antisense) primer. The reaction conditions were the same as those in section *2.2.3*, ThermalAce™ protocol, except for a longer extension time of 2 minutes. In a separate reaction, pH_6RSA-C34A was again used as the template for a ThermalAce™-catalyzed PCR reaction intended to add an *NcoI* site to the 5' end, followed by a hexaglycine spacer. The reaction conditions were the same as those used to add the *NcoI* site to the 3' end, with the exception of the primers. Oligonucleotides 25085 (5'-GCAG*CCATGG* GAGGTGGCGG AGGTGGCGAA GCACATAAAA GTGAGATTGC-3', *NcoI* site underlined) and 11474 (5'-AACTGAGGAA CAGTCATGTC-3'; Sheffield *et al.*, 2001) were employed. *XhoI* and *NcoI* digestion of the former PCR product, and *NcoI* and *EcoRI* digestion of the latter PCR product yielded

1761 bp and 1885 bp restriction endonuclease fragments, respectively. The precipitated, digested, and gel-purified products from each PCR reaction were ligated together into the *XhoI, EcoRI* digested pPICZamp9ss vector, using the 3:3:1 molar ratio described above for three-part ligations. Following transformation and screening of resulting colonies by DNA mini-preparations using procedures described above, a clone was subjected to complete DNA sequencing and designated pPZ-H_6diRSA. The intended expression product of this plasmid was an 1175 amino acid protein comprised of five N-terminal His residues, codons 3-581 of RSA, a Met-$(Gly)_6$ hexapeptide, and codons 1-584 of RSA. Residues 1 and 2 were substituted by 5 His residues in order to make a hexahistidine sequence, with His3, the TLG of the first RSA moiety was deleted to avoid introduction of a helix-breaking Pro residue; the hexaglycine was introduced to separate the two RSA moieties; and the Met was introduced as a consequence of the use of the NcoI site in the DNA manipulations.

2.2.3 Transformation of Pichia pastoris with Expression Plasmids

Pichia pastoris yeast, strain X-33, was made competent using the supplier's protocol for electroporation (Invitrogen). Yeast cells were grown overnight in 1% yeast extract, 2% peptone, 2% dextrose (YPD) medium at 30°C, with shaking. Cells were harvested by centrifugation for 5 minutes at 1500 X g, decanted, and resuspended in 100 ml ice-cold sterile water. A second centrifugation and resuspension in cold water was followed by centrifugation and resuspension in 20 ml ice-cold 1.0 M sorbitol. A final centrifugation step was followed by resuspension in 1.0 ml of ice-cold 1.0 M sorbitol,

and storage on ice until use in electroporation. The pH_6RSA-C34A vector was linearized using *SacI*, precipitated, and resuspended in sterile distilled water. Aliquots of 80 µl of yeast cells in 1.0 M sorbitol were combined with 5 µg linearized expression plasmid DNA in 5 µl of sterile water in electroporation cuvettes. One cuvette contained yeast only, with no DNA, as a control. A Gene Pulser II (Bio-Rad) electroporator was employed, using pulse settings suggested by the supplier of the *Pichia pastoris* cell line (Invitrogen). The suggested settings were: 1.5 kV charging voltage; 25 µF capacitance; and resistance of 200. After electrical current was applied to the cuvettes, 1.0 mL of cold 1 M sorbitol was added to each and they were placed in the 30°C incubator for an hour with no shaking. 50 and 100 µL aliquots, respectively, of this yeast transformation reaction were spread on YPDS (YPD / 0.7% agar plates containing 1.0 M sorbitol) agar plates containing 0.1 mg/ml Zeocin. The plates were stored at 30°C and incubated for three to four days until colonies appeared. This protocol was repeated a number of times for the pH_6RSA vector and the diRSA vector in order to obtain sufficient colonies to test for expression of the protein of interest.

2.2.4 Protein Expression and Production

2.2.4.1 Trial-scale expression screening of Zeocin-resistant Pichia pastoris colonies

Zeocin-resistant colonies of transformed yeast were used to inoculate cultures that were grown overnight in 2.5 ml of YPD to mid-log phase, then centrifuged and resuspended in Buffered Minimal Methanol (BMM) medium (100mM potassium phosphate pH 6.0, 0.34% yeast-nitrogen base media with 1.0% ammonium sulphate, 4 x

10^{-5}% biotin, and 0.5% methanol) to an OD_{600nm} of 1.0. The cells were incubated at 30°C with shaking at 200 RPM, and methanol was added to 0.5% every twenty-four hours for three days. Samples (0.1 ml) were removed each day, centrifuged to pellet cells, and the supernatant was analyzed for recombinant protein expression by SDS-PAGE and immunoblotting. Colonies showing the highest levels of heterologous protein expression were chosen for production of protein preparations used for clearance studies.

2.2.4.2 Larger-scale expression of "non-reiterated" recombinant RSA

The H_6RSA and H_6RSA(C34A) proteins were produced using a procedure analogous to that described in the previous section, but using larger scale shaker-culture expression in two 2L Erlenmeyer flasks, each containing 500 ml of media, for each different protein. The conditioned media was removed from the cells by centrifugation at 2600 X g, and concentrated to 35 ml (for H_6RSA(C34A)), and 15 ml (for H_6RSA) with a 50 kDa MWCO filter under N_2 pressure in an Amicon stirred ultrafiltration device (Millipore, Bedford, MA). The H_6diRSA protein was produced at much lower levels than either H_6RSA or H_6RSA(C34A), necessitating a much larger scale-up of production of H_6diRSA.

2.2.4.3 Fermentor Production of H_6diRSA

Production of H_6diRSA was performed on a large scale using a 14 L glass reaction vessel in a New Brunswick Scientific Microferm Fermentor system (New Brunswick Scientific, Edison, NJ).

A starter culture of 500 ml YPD/0.1 mg/ml Zeocin was inoculated with cells transformed with pH_6diRSA and grown over the weekend to an OD_{600} of 9.85, centrifuged, and resuspended in 200 ml of sterile distilled water. This suspension was used to inoculate the autoclaved media in the fermentor, after the preparations described below were completed.

The basal salts media for fermentation (8L) was autoclaved inside the glass reaction vessel, and consisted of 2.3% Ortho-phosphoric acid (v/v), 4% glycerol (w/v), 5 mM calcium sulphate, 104 mM potassium sulphate, 60 mM magnesium sulphate (heptahydrate), and 74 mM potassium hydroxide. The pH and dissolved oxygen probes were anchored into the lid of the fermentation vessel before autoclaving to ensure they would introduce no contamination. The vessel was inserted into the fermentor, the temperature probe was inserted into its sheath, and the cooling water tubes and oxygen tank (containing 95% O_2 and 5% CO_2) were attached and the valves opened. The PTM_1 trace salts (24 mM cupric sulphate-$5H_20$, 0.5 mM sodium iodide, 18 mM manganese sulfate-H_20, 0.8 mM sodium molybdate dehydrate, 0.3 mM boric acid, 2 mM cobalt chloride, 147 mM zinc chloride, 234 mM ferrous sulphate heptahydrate, 0.8 mM biotin, and 0.5% sulphuric acid) and 0.375% (v/v) antifoam 289 (Sigma) were added once the media cooled to the set temperature of 30^0 C, and the pH was adjusted to 5 using ammonium hydroxide. Once the oxygen was turned on and the flow adjusted, the DO value was set at 100% saturation. The inoculum described in the previous paragraph was added at this point.

The dissolved oxygen value was monitored to establish the time at which the glycerol in the medium was exhausted; initially it was 100%, but fell as metabolic activity increased, then returned to full saturation when the carbon source was exhausted. Additional glycerol was fed in for five hours until the yeast reached optimal cell mass (approximately 120 g/L) and exhausted this additional fuel. At this point the carbon source was changed to methanol and the yeast were induced for 67 hours, before harvesting the conditioned media.

To adjust the conditioned media to conditions best suited to preserving and purifying its secreted proteins, the pH in the fermentation vessel was adjusted to 7.0, PMSF (1mM) and benzamidine (5mM) were added to inhibit proteases, and sodium azide (0.02%) was added as an antimicrobial agent. The contents of the fermentor were aliquoted into 1L bottles and centrifuged at 3000 rpm in a Sorvall RC-3 centrifuge. The conditioned media removed from the cells was concentrated to 600 mLs using a 30 kDa Molecular Weight Cut Off (MWCO) Tangential Flow Filtration cartridge (TFF-Millipore) attached to a high capacity peristaltic pump, prior to all subsequent purification steps.

2.2.5 Protein Purification

2.2.5.1 Purification of H_6RSA and H_6RSA(C34A)

Concentrated conditioned media containing either H_6RSA or H_6RSA(C34A) were separately applied to a 4 ml nickel-nitrilotriacetic acid (Ni-NTA) column. The column was washed with 50 mM sodium phosphate pH 7.4/300 mM NaCl/0.02% azide and

eluted with the same buffer supplemented with a linear gradient of imidazole. A 0-30 mM imidazole gradient was used for the H_6RSA(C34A) protein and 0-50 mM for its wild-type counterpart. Fractions of approximately 3 ml were collected and aliquots of these were analyzed by SDS-PAGE. Fractions of acceptable purity were pooled and concentrated in 10 kDa MWCO centrifugal filter devices (Millipore). The filters were also used to reduce the imidazole concentration and change the buffer of the samples to 20 mM sodium phosphate pH 7.4 for H_6RSA(C34A) and Tris-buffered saline (TBS) for the H_6RSA WT protein.

2.2.5.2 Purification of H_6diRSA

The first step in the purification process for dialbumin was the same as that for the other recombinant, *Pichia*-derived RSA proteins, with the exception of the column size and volume of elution fractions collected. A larger 125 mL Ni-NTA column was used to purify H_6diRSA and 10 mL fractions were collected from elution with the 0-50 mM imidazole gradient. The fractions containing the greatest proportion of dialbumin were pooled, dialyzed overnight against TBS, and concentrated using 10kDa MWCO spin concentrators. Concentrated dialbumin in TBS was next applied to a 5 mL Affigel Blue® (Cibacron Blue dye coupled to agarose beads) column and washed with 20 mM sodium phosphate, pH 7.1. Bound protein was eluted with a linear gradient of 0.4-1.4 M NaCl in 20 mM sodium phosphate, pH 7.1. Fractions containing H_6diRSA (assessed by SDS-PAGE) were pooled, concentrated and dialyzed into 20 mM sodium phosphate (pH 7.1) in 100 kDa MWCO centrifugal filtration devices. The final step in purification of

H_6diRSA was gel filtration in a 170 cm long X 1.5 cm diameter column containing Sepharose 6B chromatography resin (Amersham Biosciences). The sample was applied and eluted overnight with 1.0 M NaCl. Fractions of approximately 4.5 mLs were collected after 70 mLs of void volume was eluted, and aliquots were analyzed by SDS-PAGE (8% gels, under reducing conditions). Appropriate samples were pooled, concentrated and dialyzed against 20 mM sodium phosphate. All recombinant proteins were further characterized by immunoblot analysis and their concentration determined using a Bradford assay, described below.

2.2.6 Determination of Total Protein Concentration

Total protein concentration was determined by the method of Bradford (Bradford, 1976) using Bio-Rad protein assay reagents. A 2.0 mg/mL protein standard (bovine serum albumin) was serially diluted from 1 mg/mL to 0.05 mg/mL, in duplicate, in a microtitre plate. 5 µl of each protein sample was dispensed, in duplicate, into the wells following the standards. Coomassie Brilliant Blue G-250 reagent (200 µl) was aliquoted by multi-pipetter into the wells and the plate was analyzed in a BIO-TEK plate reader (Winooski, VT) at 630 nm. Protein quantification was calculated by the BIO-TEK program from the standard curve generated from the BSA samples.

2.2.7 Immunoblotting

Recombinant H_6RSA, H_6RSA (C34A), and H_6diRSA proteins were analyzed by immunoblot analysis as described by Towbin *et al.* (1979). Blots were probed with

affinity-purified chicken anti-RSA or mouse anti-hexahistidine as the primary antibody. Secondary antibodies were goat anti-chicken IgG (Kirkegaard and Perry) and anti-mouse IgG (Promega) respectively. Both secondary antibodies were conjugated to alkaline phosphatase for detection.

2.2.8 Molecular Size Evaluation of H_6diRSA by Gel Filtration

H_6diRSA (partially purified by Ni-NTA chromatography) was subjected to gel filtration with and without molecular weight standards in order to find the effective molecular volume occupied by this molecule. The molecular weight standards consisted of a mixture of thyroglobulin (670,000 Da), bovine gamma globulin (158,000 Da), ovalbumin (44,000 Da), myoglobin (17,000 Da), and vitamin B_{12} (1,350 Da). The standards or H_6diRSA sample were applied to the column individually, or in combination, and 3.25 mL fractions were collected overnight after 60 ml of void volume was eluted. The fractions containing myoglobin and vitamin B_{12} were observed visually since these standards were coloured. The other fractions were detected by spectrophotometry (at 280 nm) and by SDS-PAGE analysis. The absorbance at 280 nm was plotted against fraction number for molecular weight standards, molecular weight standards plus H_6diRSA, and H_6diRSA alone in order to observe peaks where the proteins eluted. Elution volume was calculated from these graphs and plotted against molecular weight of the standards on a log scale for standards and standards plus H_6diRSA.

2.2.9 Clearance Studies

2.2.9.1 Radiolabeling of proteins

H_6RSA, H_6RSA(C34A), and H_6diRSA proteins were labeled with either $Na^{125}I$-iodide or $Na^{131}I$-iodide by the Iodogen method (Fraker and Speck, 1978). Between 0.5 and 1.0 nmole of purified protein was labeled in each 0.4 ml reaction. Following exhaustive dialysis, the two labeled proteins were combined in order to deliver a dose of 2 X 10^7 dpm/kg respectively in a total volume of 1.0 ml phosphate-buffered saline. The specific activity of labeling ranged from 4.5 X 10^7 cpm/mg to 3.0 X 10^9 cpm/mg. Pairs of proteins were administered through a marginal ear vein to a total of 15 New Zealand White rabbits, male and female, ranging in size from 2.9 to 4.0 Kg. At time points from 6 minutes to 312 hours after injection, 0.9 ml of blood was drawn into a syringe containing 0.1 ml sodium citrate. Blood cells were removed by centrifugation at 14,000 x g, and 400 µl of plasma was precipitated with an equal volume of 20% trichloroacetic acid (TCA). Plasma samples were incubated on ice for 20 minutes, centrifuged, and the pellet (after removal of supernatant) placed in a gamma counter for determination of remaining radioactivity. The radioactive counts were mathematically adjusted to exclude the contribution of the ^{131}I isotope in the channel of ^{125}I, using singly labeled reference samples, and the adjusted value of the six minute post-injection blood sample was arbitrarily set as '100% value in plasma'. Semi-logarithmic graphs were plotted of the average percentage (of the first sample) of radioactivity remaining versus time after injection. All *in vivo* experiments were carried out under the terms of an Animal

Utilization Protocol approved by the Animal Research Ethics Board, Faculty of Health Sciences, McMaster University.

2.2.9.2 Organ distribution experiments

A total of 18 rabbits were injected with ^{125}I-labeled H_6RSA, H_6RSA(C34A), or H_6diRSA recombinant protein. Ten of the animals were euthanized after 24 hours and the remainder at 48 hours after injection. The liver, heart, spleen, kidney and lungs were carefully removed and placed in ice-cold saline. Each organ was weighed, then either divided equally into 1 cm X 1 cm X 1 cm cubes in the case of the small organs, or five such subsamples were removed for gamma counting. In either case, the excised pieces were each, weighed and placed in a vial for measurement of radioactivity in the gamma counter. The radioactivity in each organ was estimated by multiplying its weight by the counts per minute per gram determined from each piece and averaging the values. These values were added to obtain total radioactivity in 5 organs and the percentage contribution of each organ to this total was expressed graphically for the three proteins.

2.2.9.3 Pharmacokinetic Analysis

The clearance curve for each rabbit was plotted individually as described above using an iterative regression analysis referred to as "curve-peeling" (Carson *et al.*, 1979). The data from curve peeling was used to analyze the half-lives, compartmental distributions, and fractional catabolic rates of each protein. The necessary mathematical

manipulations and linear regressions were performed using a Microsoft Excel Version 5.0 environment (Microsoft, Redmond, WA).

Briefly, the logarithm of the residual radioactivity in the plasma was plotted against the time after injection in decimal days. The terminal exponential phase of the clearance was curve-fit using linear regression to obtain fits with an $R^2 > 0.9$. Subtraction of the equation describing this line from the remaining data yielded a biphasic curve. This property of the plasma curve indicated that a three compartment model best described the data, while the latter indicated the need for a three component model (Carson and Jones, 1979). After repetition of the curve-peeling process a single linear component remained on the semi-log plot, confirming the appropriateness of the three compartment approach. The plasma curve could be written as a 3-term exponential equation:

$$A_p(t) = C_1e^{-a1t} + C_2e^{a2t} + C_3e^{a3t}$$

In this model the compartments consist of the intravascular space, the vessel wall, and the extravascular space. The exponent a_n is found by measuring the slope of the regression line and is the rate constant. C_n signifies the compartmental distribution where the three terms add to unity, and can be found by extrapolating the regression line out to the y-axis. The amount of radioactivity remaining in the plasma at any time t is represented by the term $A_p(t)$, analogous to y(t), and can be used to calculate the half life by setting this value to 0.5 and solving for t (Carlson *et al.*, 1985), in other words:

$$T_{1/2}=\ln 2/a_1$$

Fractional catabolic rates (j_T, fraction of total body albumin or derivative catabolized per day, j_3, fraction of total body albumin or derivative catabolized per day in

the intravascular space, $j_{3,5}$, fraction of total body albumin or derivative catabolized per day in the intravascular space and non-circulating vessel wall), terminal catabolic half-lives, and distributions in blood, non-circulating vessel wall, and extravascular space, were calculated from C_n and a_n values derived from the curve peeling as described (Carlson *et al.*, 1984 and 1985). For comparison of pharmacokinetic parameters, one-way ANOVA with Tukey-Kramer's post test was performed using GraphPad InStat software version 3.01 for Windows, GraphPad Software Inc., San Diego, CA.

2.2.10 Biochemical analysis of purified proteins

2.2.10.1 Amino acid sequencing

Protein microsequencing was performed following SDS-PAGE and electrical blotting onto sequencing grade polyvinylydene difluoride membranes, essentially as described above for immunoblotting. Bands of interest were excised from the blot and subjected to the Edman degradation, performed with a Porton Gas-phase Microsequencer, model 2090, with on-line analysis, by the Protein Sequence Analysis Facility of the Hospital for Sick Children Biotechnology Service Centre (Toronto, ON).

2.2.10.2 MALDI-TOF Analysis

Matrix Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS) was performed by the McMaster University Department of Chemistry Regional Mass Spectrometry Facility (Hamilton, ON), using purified RSA-

related recombinant proteins at high concentration (10 mg/ml) dialyzed into sterile double-distilled water.

2.2.10.3 Thiol-reactive chemical modification

EZ-link™ PEO-maleimide activated biotin (Pierce, Rockford, IL), was used to detect free sulfhydryl groups in the non-reiterated recombinant albumins. Maleimide reacts primarily with free thiols under conditions of slightly acidic to neutral pH (6.5-7.5), but is reactive towards primary amines at higher pHs. The functional group is linked to biotin via a polyethylene oxide (PEO) spacer which can be detected on a blot with alkaline phosphatase conjugated to streptavidin. Approximately 60 ng of each of H_6RSA(C34A) and H_6RSA were reacted with the EZ-link product at pH 6.8 according to the protocol supplied by the manufacturer. After dialysis to remove excess reagent, aliquots of the reaction mixtures were electrophoresed under non-reducing conditions. The reacted proteins were then transferred to a nitrocellulose membrane and probed with AP conjugated streptavidin.

2.2.11 SDS-PAGE

Protein samples were analyzed by SDS-PAGE (Laemmli, 1970), typically using the Mini-Protean II™ system (Bio-Rad) and 10% SDS-polyacrylamide gels of 0.75 mm thickness. Prior to electrophoresis, samples were diluted by the addition of one-third volume of SDS-PAGE sample buffer (250 mM Tris-Cl pH 6.8/40% glycerol/ 8% SDS/0.5 % bromophenol blue/200 mM dithiothreitol). In some experiments, non-

reducing conditions were employed, using the same sample buffer lacking DTT. Samples were boiled for two minutes prior to loading onto the gel. Gels were stained with 35% methanol/10% glacial acetic acid/0.05% Coomassie Brilliant Blue R-250 and destained in the same solution lacking the dye. For autoradiography of iodinated samples, the same procedure was followed except that after destaining the gels were dried onto 3 MM Whatman paper using a Bio-Rad gel drying apparatus, and exposed to Kodak X-Omat XAR-1 film in an autoradiography cassette at -70°C.

RESULTS

3.1 DNA Manipulation and Transformation of Pichia pastoris

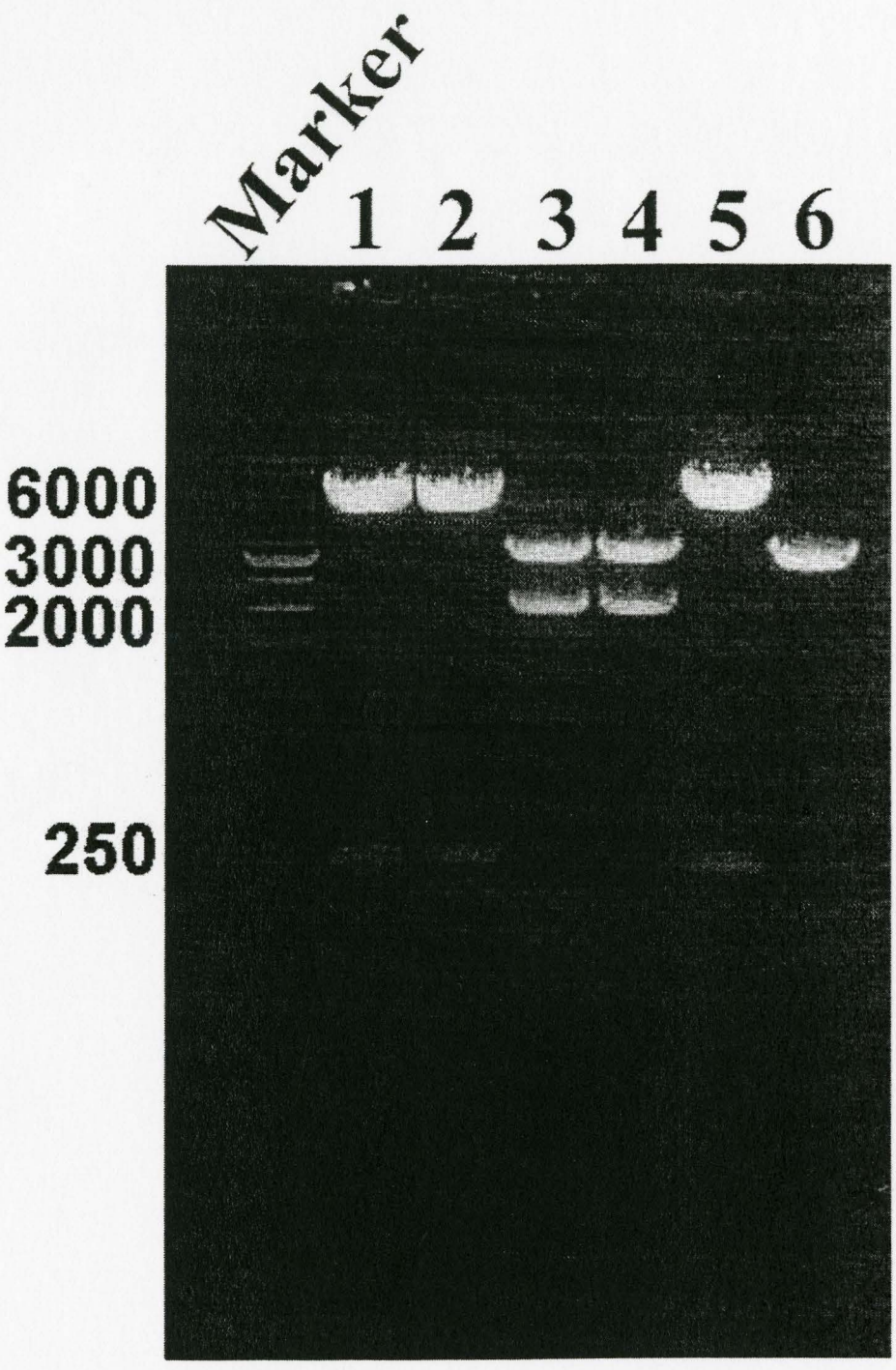
3.1.1 pPZ-H_6RSA Vector

For the purpose of generating a terminally His-tagged reiterated albumin, it was initially necessary to create an N-terminally modified RSA, as previous RSA expression vectors used in our laboratory were C-terminally modified (Syed *et al.* 1997). PCR-based mutagenesis was used to add codons specifying six His residues, in frame to both the pre-pro-α factor leader sequence and the subsequent RSA codons. The resultant PCR product was rendered blunt-ended and cloned into plasmid vector pGEM5zf(+) non-directionally. Five white transformed bacterial colonies and one blue colony were used to inoculate cultures and plasmid DNA was analyzed by restriction analysis, as shown in Figure 3. *SacI/EcoRI* double digestion of these DNA samples yielded two patterns: one comprised of a 4847 bp major fragment and a 206 minor fragment (lanes 1, 2 and 5); and the other comprised of a 3131 and a 1992 bp fragment (lanes 3 and 4). The unaltered plasmid derived from the blue colony (lane 6) exhibited a single 3003 bp band. As this pattern corresponded exactly to that expected for the two possible correct insertions of the amplified DNA into the pGEM vector, the plasmid corresponding to the sample in lane 1 was designated pGEM5-H_6RSA and used as the genetic starting material for all subsequent manipulations.

In order to minimize the amount of the novel H_6RSA coding region produced by PCR, and potentially containing misincorporations, a 561 bp *XhoI-HindIII* restriction

FIGURE 3: Agarose gel electrophoresis of combined *SacI/EcoRI* digestion of plasmid preparations of pGEM5-H_6RSA candidate colonies

One-fifth of each of six plasmid DNA mini-preparations was restricted with *SacI* and *EcoRI* and subjected to agarose gel electrophoresis in the presence of ethidium bromide on a 1% agarose gel. Samples shown in lanes 1-5 derive from colonies that were white on indicator plates; that shown in lane 6 derived from a blue colony. DNA molecular weight markers (GeneRuler 1Kb Marker, Fermentas) were applied in the "Marker" lane. The size of selected marker bands, in base pairs, is shown at left.


Marker
1
2
3
4
5
6
6000
3000
2000
250

digestion fragment was combined with a 1322 bp *HindIII-EcoRI* fragment of pALB192 (Syed *et al.*, 1997) and the major *XhoI-EcoRI* 4584 bp of pPICZamp9ss (Sheffield *et al.*, 2001) in a three part ligation reaction. This manipulation, shown schematically in Figure 4, limited the amount of DNA derived from PCR amplification of the first fragment. The resulting plasmid was characterized by restriction analysis analogous to that shown in Figure 3, and DNA sequencing in the amplified area, where no misincorporations were found. This plasmid was therefore designated pPZ-H_6RSA. The complete amino acid sequence of the RSA-related open reading frame of pPZ-H_6RSA, the mature H_6RSA amino acid sequence expected after cleavage and secretion, and the amino acid composition and mass of the expected H_6RSA protein, are shown in Figure 5. The genetic map of the final 6437 bp expression vector is shown in Figure 6. The results of automated sequencing of all three RSA-related constructs described here showed complete agreement with the published sequence and planned alterations, with two exceptions: Thr297-Pro298 was found to be Asp297-Thr298; and Ser1080-Asn1081 were found to be Val1080-Asp1081. Re-examination of records of the manual sequencing performed when RSA was cloned in our laboratory suggested that errors were made in manually transcribing the sequence into electronic files for submission to Genbank, and an erratum report to Genbank will be made in the near future.

3.1.2 Generation of C34A cDNA

Failure of the site directed mutagenesis kit (GeneEditor, Promega) to produce cDNA with the required change after three attempts prompted use of PCR to incorporate

Figure 4: Assembly of pPZ-H_6RSA vector

Schematic diagram of the steps used to create the yeast vector to express H_6RSA. The 561 bp piece from the *XhoI* to *HindIII* site of the amplified pSG5-RSA was ligated together with the *HindIII* to *EcoRI* site of the non-amplified pAlb192 vector into pPICZamp9ss.

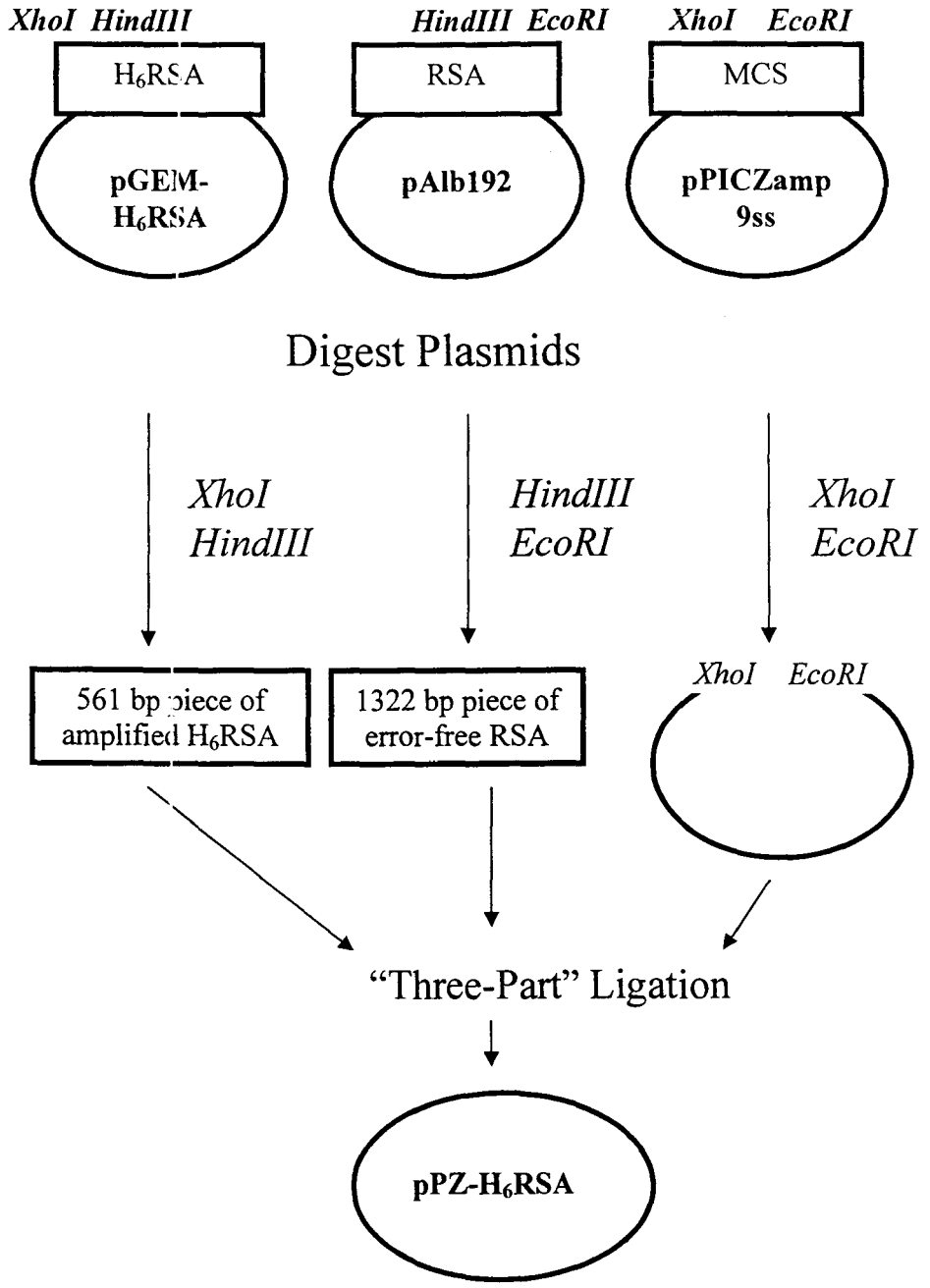

XhoI HindIII
H6RSA
pGEM-H6RSA
HindIII EcoRI
RSA
pAlb192
XhoI EcoRI
MCS
pPICZamp 9ss
Digest Plasmids
XhoI
HindIII
HindIII
EcoRI
XhoI
EcoRI
561 bp piece of amplified H6RSA
1322 bp piece of error-free RSA
XhoI EcoRI
"Three-Part" Ligation
pPZ-H6RSA

FIGURE 5: Primary sequence and amino acid composition of H_6RSA

The amino acid sequence of the modified RSA protein, deduced from sequencing reactions of its cDNA as described in Materials and Methods, in single letter code, is shown in panel A, with the numbers at right representing the position within the sequence. The histidine tag (HHHHHH) is shown at the beginning (N-terminus) of the mature (cleaved) molecule. The amino acid composition is given in panel B, by number of amino acids in the protein followed by percentage in the total protein, for each amino acid (given in single letter code). This was calculated using GenePRO software.

A H6RSA

```
HHHHHHKSEI AHRFNDVGEE HFIGLVLITF SQYLQKCPYE EHAKLVKEVT DLAKACVADE    60
SAANCDKSLH DIFGDKICAL PSLRDTYGDV ADCCEKKEPE RNECFLHHKD DKPDLPPFAR   120
PEADVLCKAF HDDEKAFFGH YLYEVARRHP YFYAPELLYY AQKYKAILTE CCEAADKGAC   180
LTPKLDALEG KSLISAAQER LRCASIQKFG DRAYKAWALV RLSQRFPKAD FTDISKIVTD   240
LTKVHKECCH GDLLECADDR ADLAKYMCEH QETISSHLKE CCDKPILEKA HCIYGLHNDE   300
TPAGLPAVAE EFVEDKDVCK NYEEAKDLFL GKFLYEYSRR HPDYSVVLLL RLGKAYEATL   360
KKCCATDDPH ACYAKVLDEF QPLVDEPKNL VKQNCELYEQ LGDYNFQNAL LVRYTKKVPQ   420
VSTPTLVEIS RSLGKVGSKC CKHPEAERLP CVEDYLSVVL NRLCVLHEKT PVSEKVTKCC   480
SESLVDRRPC FSALGPDETY VPKEFNAETF TFHADICTLP ETERKIKKQT ALVELVKHKP   540
HATNDQLKTV VGEFTALLDK CCSAEDKEAC FAVEGPKLVE SSKATLG                 587
```

Composition from 1 to 587

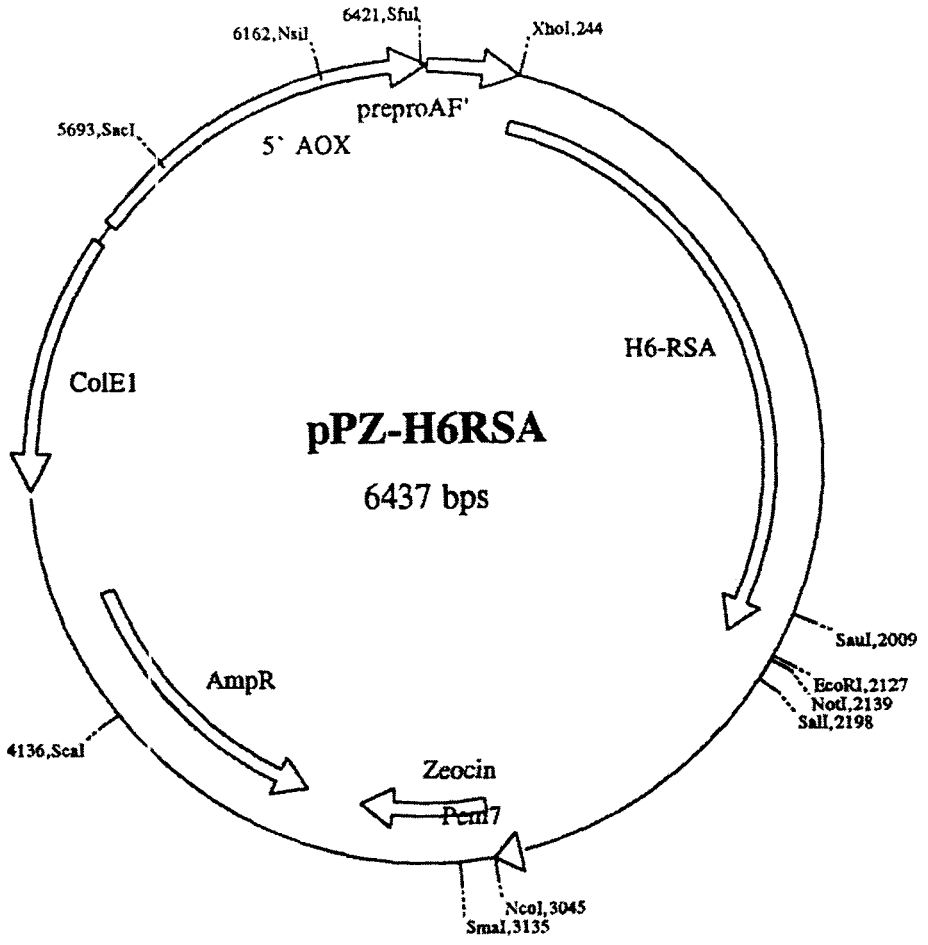
B H6RSA

A	53	9.0
C	35	6.0
D	43	7.3
E	55	9.4
F	24	4.1
G	20	3.4
H	28	4.8
I	16	2.7
K	57	9.7
L	62	10.6
M	1	0.2
N	12	2.0
P	29	4.9
Q	14	2.4
R	22	3.7
S	26	4.4
T	27	4.6
V	38	6.5
W	1	0.2
Y	24	4.1

587 residues, MW 66500

FIGURE 6: Schematic representation of pPICZamp9ss-H_6RSA (pPZ-H_6RSA)

The yeast expression vector pPICZamp9ss containing the α-factor signal sequence (prepro AF) and hexahistidine tagged RSA cDNA (H6-RSA) is shown. A number of genetic elements are depicted in the diagram that allow for protein expression, selection and growth in bacteria. These are: ColE1, *E. coli* origin of replication to allow maintenance of the vector in the bacteria; Amp R, Ampicillin resistance gene allowing selection of the plasmid in *E. coli*; Zeocin, Zeocin resistance gene that allows for selection of yeast containing the cDNA and Pem7, the constitutive promoter for this resistance gene; and 5'AOX, promoter for methanol-inducible protein expression.


6421,SfuI
6162,NsiI
XhoI,244
preproAF'
5' AOX
5693,SacI
H6-RSA
ColE1
pPZ-H6RSA
6437 bps
SauI,2009
EcoRI,2127
NotI,2139
SalI,2198
AmpR
4136,ScaI
Zeocin
Pem7
NcoI,3045
SmaI,3135

the codon change. The mutation in the PCR product was detected by analysis with the dideoxy chain termination method (Sanger *et al.*, 1977). This is shown in Figure 7, compared with the sequencing results for the same segment of pPZ-H_6RSA (designated WT). Again, a three-part ligation was used to reduce the chance of errors in the coding sequence, analogous to that used to construct the H_6RSA sequence. The timed, partial restriction enzyme digestion of the H_6RSA insert with *SphI* yielded fragments that were used both for the PCR procedure and as the non-amplified portion of the final plasmid. This procedure is depicted in Figure 8. Diagnostic restriction enzyme analysis was used to screen candidate colonies transformed with a portion of the ligation reaction. The amplified portion of the plasmid was sequenced and found to contain the expected sequence. This plasmid, designated pPZ-RSA-C34A and shown in Figure 9, was subsequently manipulated to produce the dialbumin encoding vector.

3.1.3 Dialbumin Construct

The H_6RSA-C34A construct was reiterated by manipulation of the 3' end of the first copy to add an *NcoI* site and the 5' end of the second copy to incorporate the hexaglycine spacer and an *NcoI* site. These additions were accomplished by PCR and restriction digests of the products were quantified and purified from an agarose gel, then used in the ligation shown diagrammatically in Fig. 10. Candidate colonies grown on LB Amp plates from the transformation of DH5α cells with this ligation mixture were screened for proper sized insert by restriction enzyme digestion and agarose gel electrophoresis as in experiments analogous to those shown in Figure 3. A schematic

FIGURE 7: Sequencing of wild-type and mutant RSA cDNAs

An autoradiogram of a 8% urea/polyacrylamide gel is shown, highlighting the mutation of codon 34 from TGC, which codes for cysteine in wild type RSA (panel B), to GCC, which codes for alanine in the mutant C34A (panel A). The columns in the autoradiograms correspond to A, C, G, and T, from left to right, in each panel.

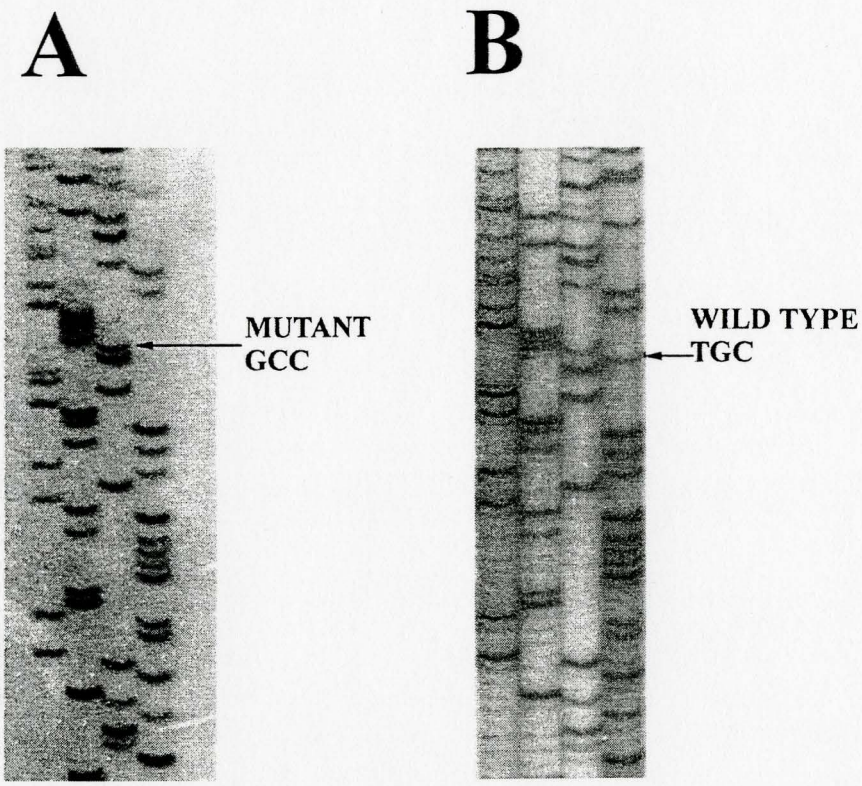

A
B
MUTANT
GCC
WILD TYPE
TGC

FIGURE 8: Cloning of construct for C34A into the yeast expression vector

A schematic diagram of the method used to clone hexahistidine tagged C34A cDNA into the pPICZamp9ss yeast vector is shown. The cysteine 34 to alanine mutation was introduced into H_6RSA cDNA by PCR using high fidelity *Pfu* polymerase. Only the first 150 bp piece was used in the ligation to reduce the chance of base misincorporation in the remainder of the cDNA. This piece was obtained by partial digestion with *XhoI* and *SphI*. The 1750 bp piece from the first *SphI* site to the end of the cDNA was liberated from the pSG5 vector containing the RSA cDNA by partial digestion with *SphI* and *EcoRI*. This RSA cDNA was previously sequenced and known to be error-free. The yeast expression vector pPICZamp9ss was digested with *XhoI* and *EcoRI*, and the two RSA cDNA pieces were ligated into the vector.

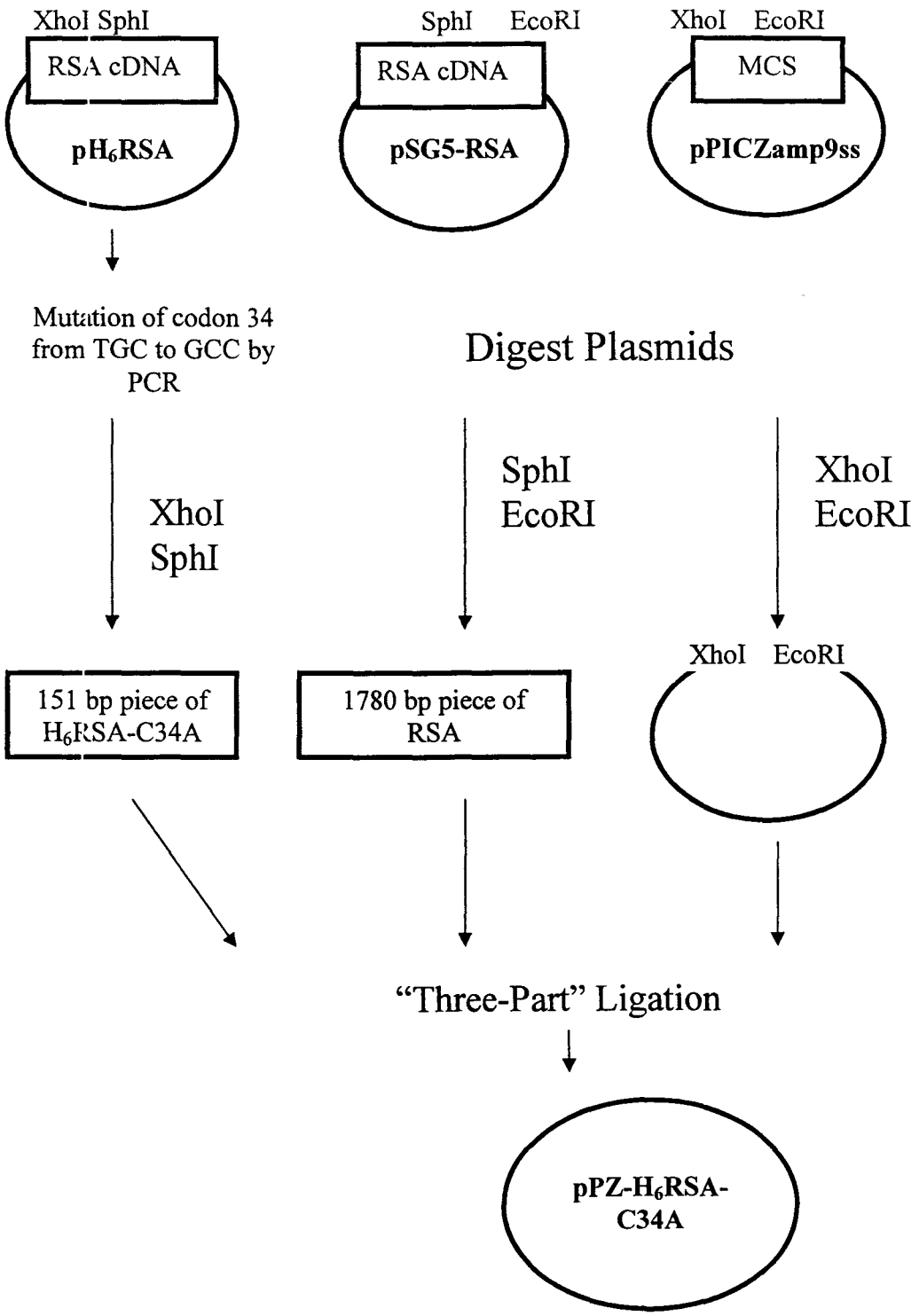

XhoI SphI
RSA cDNA
pH6RSA
SphI EcoRI
RSA cDNA
pSG5-RSA
XhoI EcoRI
MCS
pPICZamp9ss
Mutation of codon 34 from TGC to GCC by PCR
Digest Plasmids
XhoI
SphI
SphI
EcoRI
XhoI
EcoRI
151 bp piece of H6RSA-C34A
1780 bp piece of RSA
XhoI EcoRI
"Three-Part" Ligation
pPZ-H6RSA-C34A

Figure 9: Diagram representing pPZ-H_6RSAC34A expression vector

Genetic elements identified are the same as those described in Figure 6.

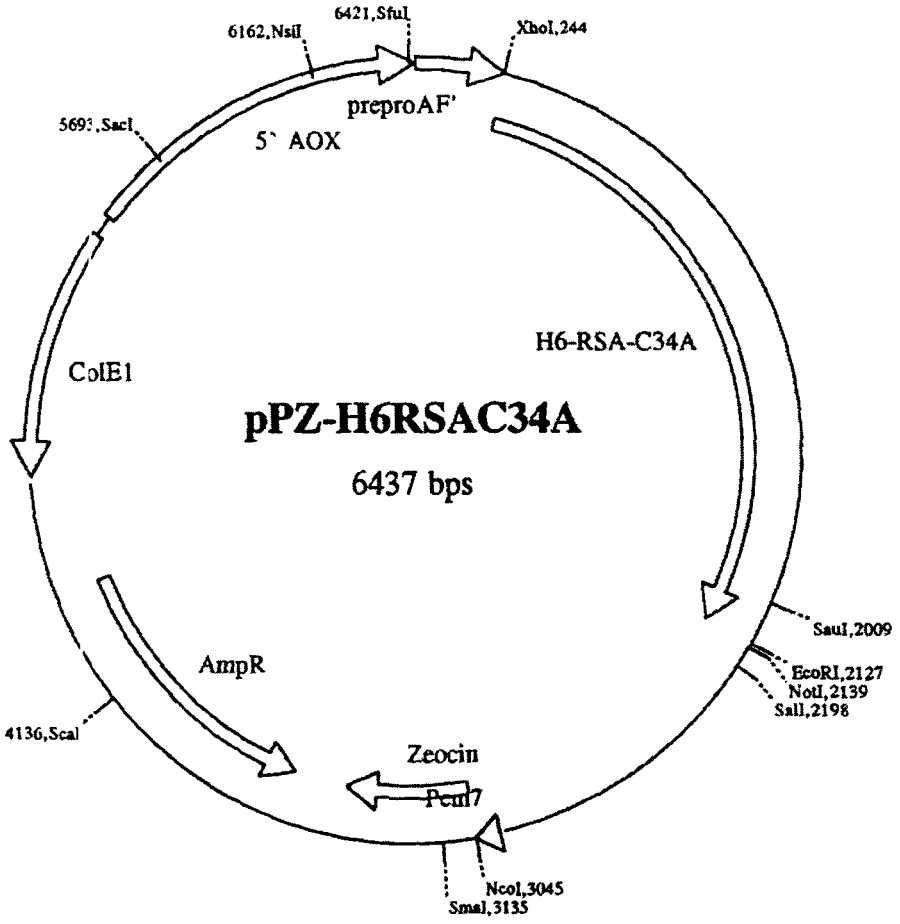

6421,SfuI
6162,NsiI
XhoI,244
preproAF'
5` AOX
5693,SacI
H6-RSA-C34A
ColE1
pPZ-H6RSAC34A
6437 bps
SauI,2009
EcoRI,2127
NotI,2139
SalI,2198
AmpR
4136,ScaI
Zeocin
Pcul7
NcoI,3045
SmaI,3135

FIGURE 10: Cloning of diRSA cDNA into yeast expression vector.

This schematic diagram shows two copies of H_6RSA cDNA with the cysteine 34 to alanine mutation, each altered by PCR, with an NcoI site added to the 3' end of the first copy, and an NcoI site and hexaglycine spacer added to the 5' end of the second copy. These pieces of cDNA were then ligated together into the pPICZamp9ss vector previously cut with XhoI and EcoRI.

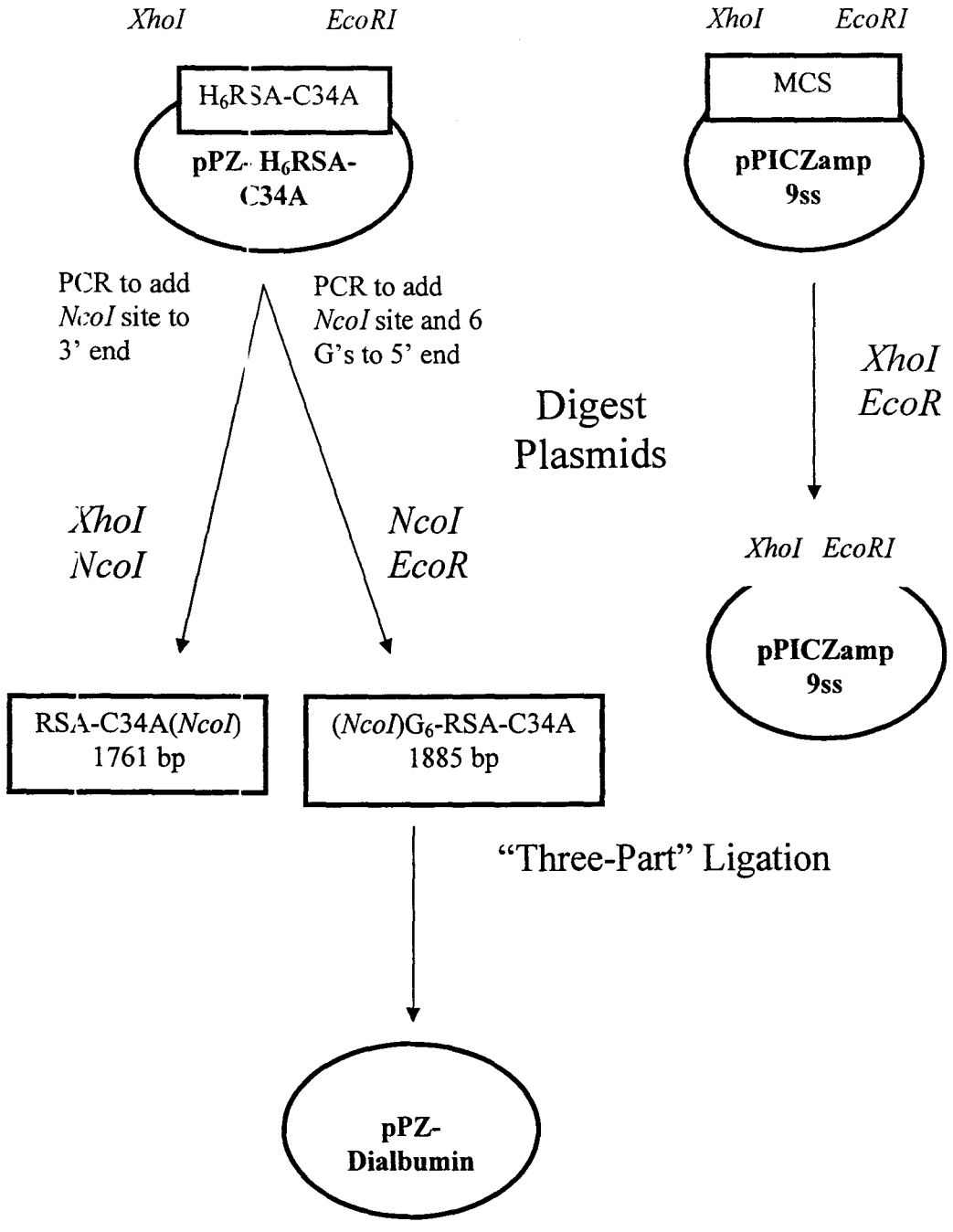

XhoI
EcoRI
H6RSA-C34A
pPZ- H6RSA-C34A
XhoI
EcoRI
MCS
pPICZamp 9ss
PCR to add NcoI site to 3' end
PCR to add NcoI site and 6 G's to 5' end
Digest Plasmids
XhoI EcoR
XhoI NcoI
NcoI EcoR
XhoI EcoRI
pPICZamp 9ss
RSA-C34A(NcoI) 1761 bp
(NcoI)G6-RSA-C34A 1885 bp
"Three-Part" Ligation
pPZ- Dialbumin

diagram of the pPZ-diRSA vector appears in Figure 11. The complete amino acid sequence of the dialbumin, its amino acid composition and mass, are shown in Figure 12.

3.1.4 Transformation of Pichia pastoris X-33 with Engineered Expression Vectors

All three constructed expression vectors were linearized with *SacI* and used to separately transform yeast cells by electroporation. These yeast expression plasmids were designed to permanently transform the yeast by integration into yeast chromosomes using homologous recombination. *SacI* linearizes the vector in the 5' AOX promoter region which is homologous to the region in this yeast regulated by methanol metabolism. This integration directs production of the encoded protein product when methanol is the only carbon source. A schematic diagram of this integration appears in Figure 25 in the Discussion. The direction of the recombinant proteins into the media was accomplished using the α-factor signal sequence from *Saccharomyces cerevisiae* which was engineered into the expression vector. This 80 amino acid sequence, which is cleaved after expression, allowed the *Pichia pastoris* yeast to secrete the novel proteins in order to simplify purification. Zeocin-resistant colonies from each transformation reaction were screened for expression of the correctly sized protein by SDS-PAGE and immunoblot analysis of conditioned media from a small scale methanol induction experiment (data not shown). Candidate colonies producing the greatest amount of recombinant protein (as shown by electrophoretic and immunological analysis) were chosen to be scaled up for production of protein to be purified and used in clearance

FIGURE 11: Schematic representation of pPICZamp9ss-H_6diRSA (pPZ- H_6diRSA)

See Figure 6 for identification of the genetic elements in the diagram.

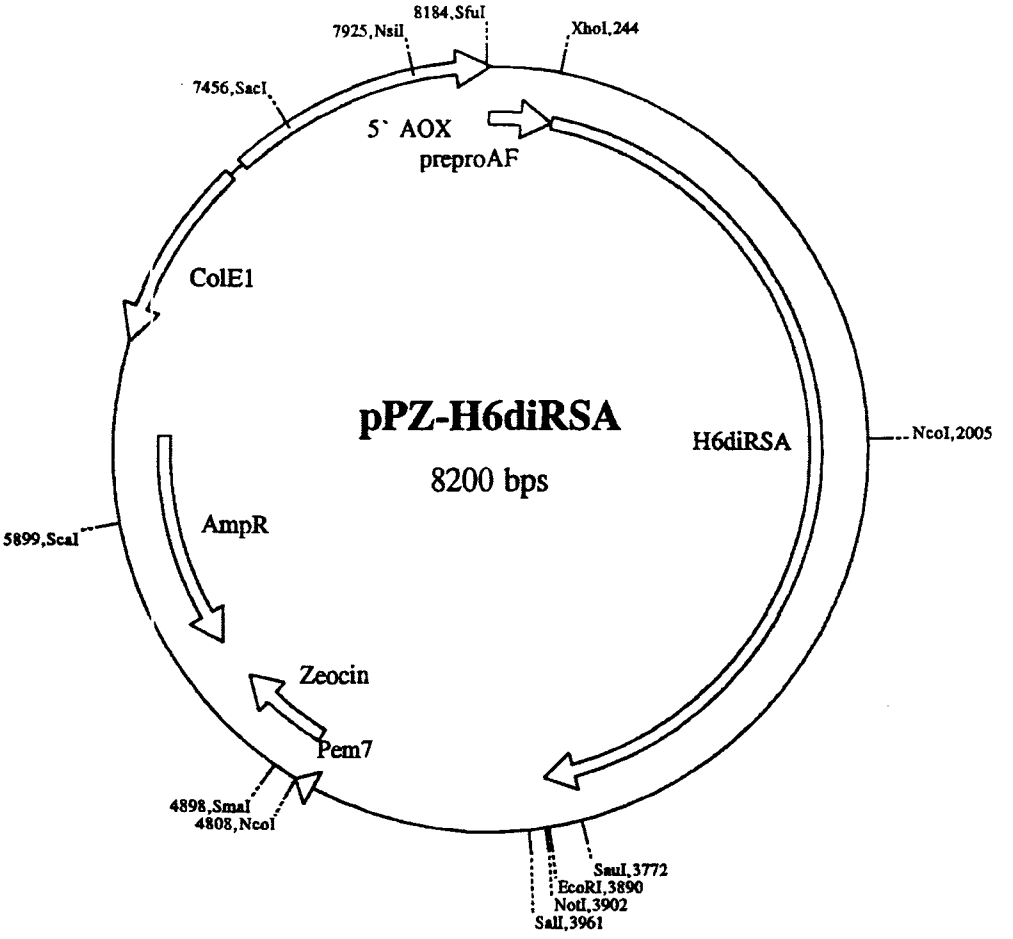

8184,SfuI
7925,NsiI
XhoI,244
7456,SacI
5` AOX
preproAF
ColE1
pPZ-H6diRSA
8200 bps
H6diRSA
NcoI,2005
AmpR
5899,ScaI
Zeocin
Pem7
4898,SmaI
4808,NcoI
SauI,3772
EcoRI,3890
NotI,3902
SalI,3961

FIGURE 12: Primary sequence and amino acid composition of Dialbumin

The amino acid sequence of the modified, reiterated RSA protein, as deduced from sequencing reactions of its cDNA, as described in Materials and Methods, in single letter code, is shown in panel A. The numbers at right represent the position within the sequence. The histidine tag (HHHHHH) is shown at the beginning (N-terminus) of the mature (cleaved) molecule, and the glycine spacer (GGGGGG) is shown in between the two modified copies (mutated cysteines at position 34 in each copy). The amino acid composition is given in panel B, by number of amino acids in the protein followed by percentage in the total protein, for each amino acid (given in single letter code). Again, this was calculated using GenePRO software.

A

Dialbumin

HHHHHHKSEI AHRFNDVGEE HFIGLVLITF SQYLQKAPYE EHAKLVKEVT DLAKACVADE 60
SAANCDKSLH DIFGDKICAL PSLRDTYGDV ADCCEKKEPE RNECFLHHKD DKPDLPPFAR 120
PEADVLCKAF EDDEKAFFGH YLYEVARRHP YFYAPELLYY AQKYKAILTE CCEAADKGAC 180
LTPKLDALEG KSLISAAQER LRCASIQKFG DRAYKAWALV RLSQRFPKAD FTDISKIVTD 240
LTKVHKECCH GDLLECADDR ADLAKYMCEH QETISSHLKE CCDKPILEKA HCIYGLHNDE 300
TPAGLPAVAE EFVEDKDVCK NYEEAKDLFL GKFLYEYSRR HPDYSVVLLL RLGKAYEATL 360
KKCCATDDPH ACYAKVLDEF QPLVDEPKNL VKQNCELYEQ LGDYNFQNAL LVRYTKKVPQ 420
VSTPTLVEIS RSLGKVGSKC CKHPEAERLP CVEDYLSVVL NRLCVLHEKT PVSEKVTKCC 480
SESLVDRRPC FSALGPDETY VPKEFNAETF TFHADICTLP ETERKIKKQT ALVELVKHKP 540
HATNDQLKTV VGEFTALLDK CCSAEDKEAC FAVEGPKLVE SSKAMGGGGG GEAHKSEIAH 600
RFNDVGEEHF IGLVLITFSQ YLQKAPYEEH AKLVKEVTDL AKACVADESA ANCDKSLHDI 660
FGDKICALPS LRDTYGDVAD CCEKKEPERN ECFLHHKDDK PDLPPFARPE ADVLCKAFHD 720
DEKAFFGHYL YEVARRHPYF YAPELLYYAQ KYKAILTECC EAADKGACLT PKLDALEGKS 780
LISAAQERLR CASIQKFGDR AYKAWALVRL SQRFPKADFT DISKIVTDLT KVHKECCHGD 840
LLECADDRAD LAKYMCEHQE TISSHLKECC DKPILEKAHC IYGLHNDETP AGLPAVAEEF 900
VEDKDVCKNY EEAKDLFLGK FLYEYSRRHP DYSVVLLLRL GKAYEATLKK CCATDDPHAC 960
YAKVLDEFQP LVDEPKNLVK QNCELYEQLG DYNFQNALLV RYTKKVPQVS TPTLVEISRS 1020
LGKVGSKCCK HPEAERLPCV EDYLSVVLNR LCVLHEKTPV SEKVTKCCSE SLVDRRPCFS 1080
ALGPDETYVP KEFNAETFTF HADICTLPET ERKIKKQTAL VELVKHKPHA TNDQLKTVVG 1140
EFTALLDKCC SAEDKEACFA VEGPKLVESS KATLG 1175

B

Composition from 1 to 1175

Dialbumin

A	109	9.3
C	68	5.8
D	86	7.3
E	111	9.4
F	48	4.1
G	45	3.8
H	51	4.3
I	32	2.7
K	114	9.7
L	123	10.5
M	3	0.3
N	24	2.0
P	58	4.9
Q	28	2.4
R	44	3.7
S	52	4.4
T	53	4.5
V	76	6.5
W	2	0.2
Y	48	4.1

1175 residues, MW 132634

experiments. A linear representation of RSA and the recombinant proteins encoded by the DNA inserted into the expression vectors is given in Figure 13.

3.2 Protein Expression

3.2.1 Shaker Culture Expression

Zeocin-resistant, transformed yeast cells were used to start cultures for protein expression. After centrifugation and resuspension in BMM, the cells were induced with additional methanol over a 72 hour period. Analysis of the conditioned media showed a band of approximately 70 kDa from the yeast transformed with pPZ-H_6RSA and pPZ-H_6RSA-C34A, and a band of approximately 140 kDa from the yeast transformed with pPZ-diRSA that did not appear in the lanes with only yeast culture (before induction). The time course of these fermentations by SDS-PAGE analysis of conditioned media is shown in Figure 14. These proteins reacted with both anti-histidine antibodies and anti-RSA antibodies by immunological analysis (data not shown). The supernatant from these fermentations was collected after centrifugation and concentrated with a 50-kDa filter under nitrogen pressure.

3.2.2 Fermentor Production

It was apparent from electrophoretic analysis that yeast transformed with the pPZ-diRSA expression vector produced far less protein, upon induction, than those transformed with either the pPZ-H_6RSA or pPZ-H_6RSA-C34A (Fig. 14A: compare "WT-72 hrs lane" to Fig. 14B, "C34A-72 hrs" and Fig. 14B, "Dialb-72 hrs"). For this reason

FIGURE 13: Rabbit serum albumin and *P. pastoris*-derived recombinant proteins

Schematic linear representations of rabbit serum albumin and the recombinant proteins expressed by *P. pastoris* in this project are shown. $\mathbf{H_6}$ represents the hexahistidine tag added to the N-termini of the recombinant proteins. $\mathbf{G_6}$ represents the hexaglycine spacer inserted between the two copies of RSA in the dialbumin molecule. The column at the right shows the size of the recombinant proteins in amino acids (aa).

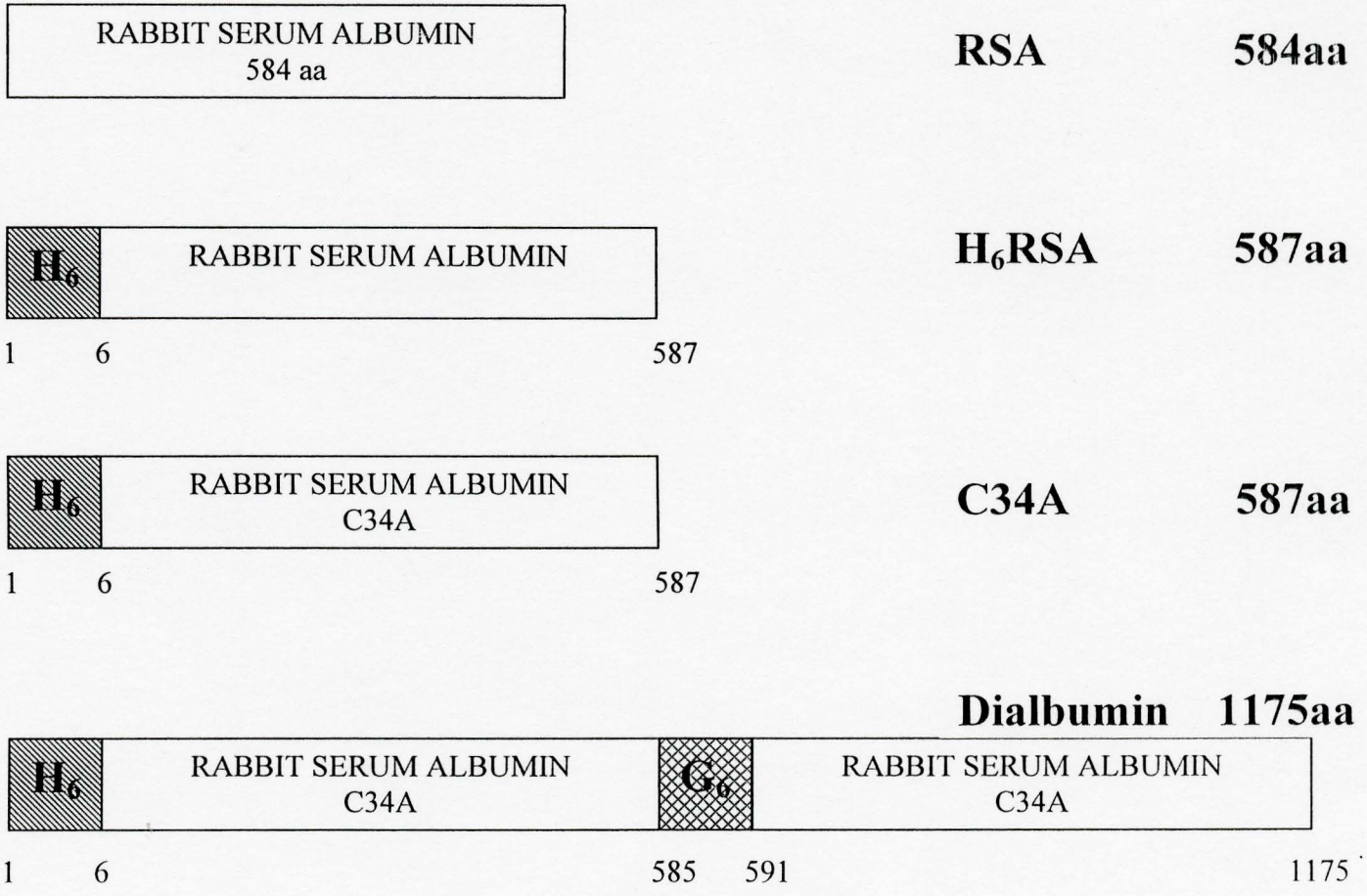
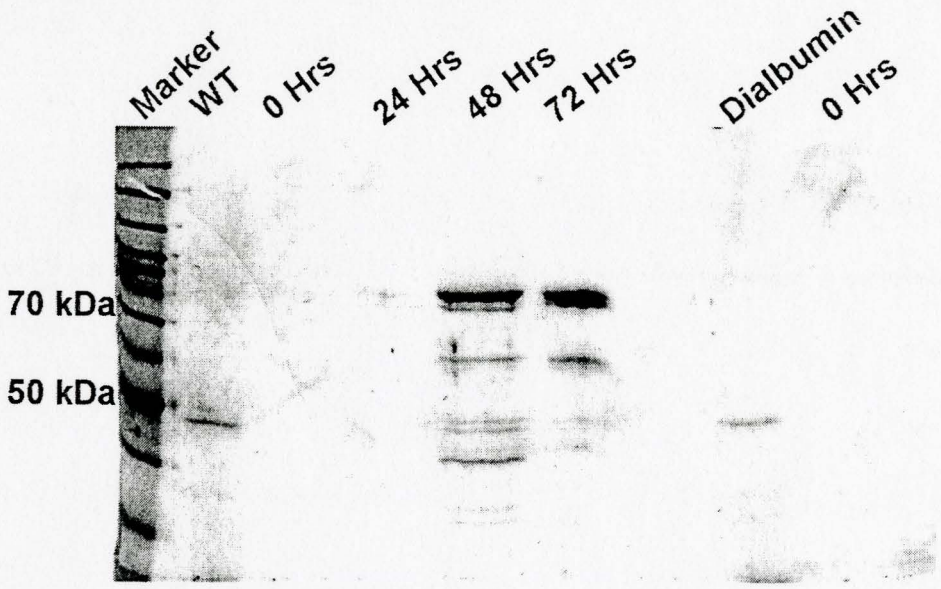

RABBIT SERUM ALBUMIN
584 aa
RSA
584aa
H6
RABBIT SERUM ALBUMIN
H6RSA
587aa
1
6
587
H6
RABBIT SERUM ALBUMIN
C34A
C34A
587aa
1
6
587
Dialbumin
1175aa
H6
RABBIT SERUM ALBUMIN
C34A
G6
RABBIT SERUM ALBUMIN
C34A
1
6
585
591
1175

FIGURE 14: Time course of protein expression by methanol induction of transformed *Pichia pastoris*

Pichia pastoris was transformed with either the pPZ-H_6RSA, pPZ-H_6RSAC34A, or pPZ-H_6diRSA vector and grown overnight at 30^0C, with shaking, in YPD media containing Zeocin. The cell suspension was centrifuged and resuspended in BMM media to an OD_{600} of 1.0. The BMM suspended yeast was allowed to grow for 72 hours at 30^0C with 0.05% methanol added every 24 hours to induce expression of the protein encoded by the vector. An aliquot of each original yeast culture in YPD was collected, as well as each conditioned media at 0, 24, 48, and 72 hours after addition of methanol. These samples were electrophoresed on an 8% SDS-PAGE gel under reducing conditions and stained with Coomassie Blue. Invitrogen Benchmark™ protein ladder was used in the marker lane, with relevant band sizes noted at left in panel A. Wild type H_6RSA in YPD, at 0, 24, 48, and 72 hours, as well as dialbumin in YPD and at 0 hours after induction are shown at right in panel A. Panel B consists of H_6RSAC34A at 0, 24, 48, and 72 hours after the start of induction, as well as dialbumin at 24, 48, and 72 hours, and finally, C34A in YPD.

A


Marker
WT
0 Hrs
24 Hrs
48 Hrs
72 Hrs
Dialbumin
0 Hrs
70 kDa
50 kDa

B

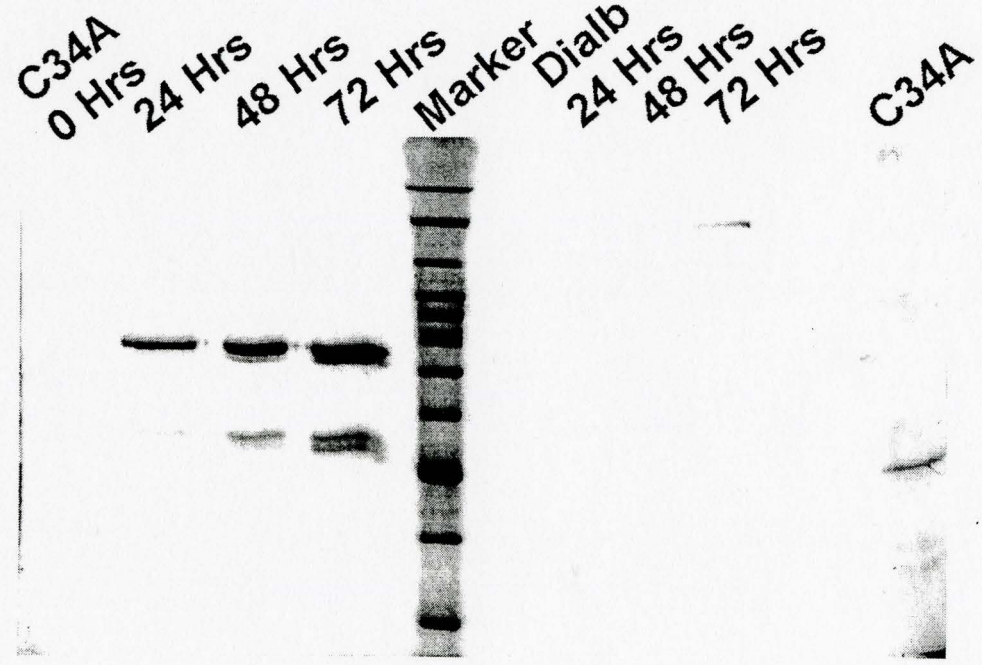
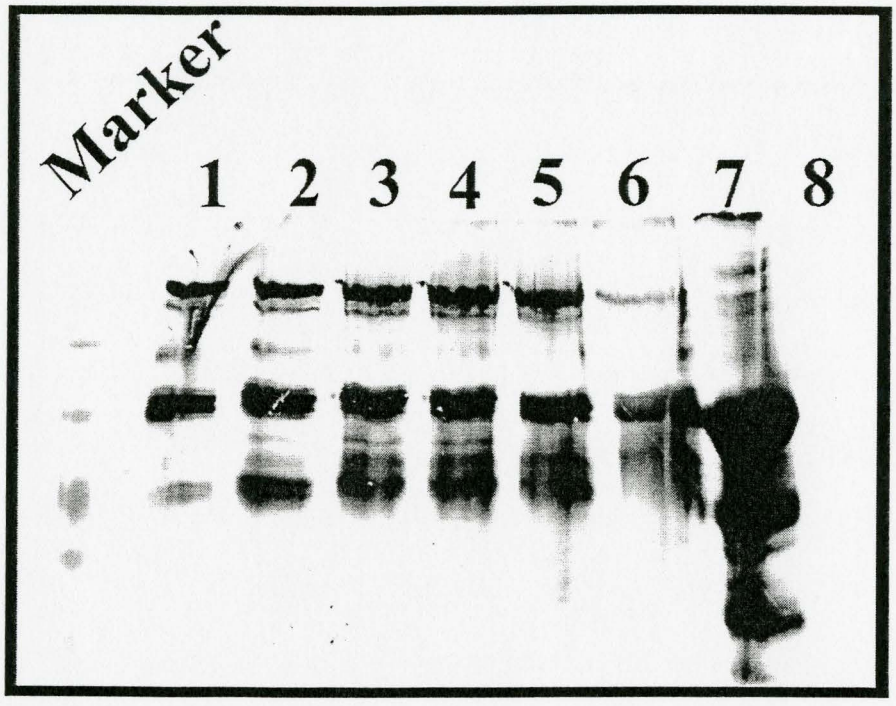

C34A
0 Hrs
24 Hrs
48 Hrs
72 Hrs
Marker
Dialb
24 Hrs
48 Hrs
72 Hrs
C34A

Figure 15 : Immunoblot of conditioned media samples with pPZ-diRSA

Samples of conditioned media from a diRSA fermentation were electrophoresed on a reducing SDS-PAGE gel and transferred to a nitrocellulose membrane. The blot was then probed with anti-RSA antibodies. Lane 1: Conditioned media from a fermentation sample taken 12.5 hours after the start of induction. Lane 2: Sample taken 17.5 hours after start of induction. Lanes 3 and 4: Samples taken 37.5 and 43 hours after start of induction, respectively. Lane 5: Sample taken 60 hours after start of induction. Lanes 6 and 7: Positive controls from previous fermentations under the same conditions (diRSA and H_6RSA(C34A), respectively); lane 8: Negative control (Heparin cofactor II). Benchmark™ Prestained Protein Ladder (Invitrogen) was used, the visible marker bands are, from top to bottom, 113.7 kDa, 80.9 kDa, 63.8 kDa, 49.5 kDa, and 37.4 kDa. The highest weight molecular marker co-migrates with the H_6diRSA bands, is barely visible, and is given as 182.9 kDa.


Marker
1 2 3 4 5 6 7 8

the 14L New Brunswick Scientific Microferm Fermentor was employed to produce the larger recombinant protein. 8L of fermentation basal salts medium was used to grow the yeast in the inoculum. Once the glycerol in the original medium was exhausted (as shown by a drop in the dissolved oxygen), a glycerol feed was established to increase the yeast mass. After five hours of glycerol feed, the yeast was allowed to exhaust the remaining glycerol, then the food source was switched to methanol. This activated the methanol metabolism pathway and resulted in the appearance of an approximately 140 kDa band on an SDS-PAGE gel. Immunological testing showed a reaction with anti-hexahistidine antibodies (data not shown) and also with anti-RSA (shown in Figure 15). Conditioned media was removed at the time points, after the start of induction, shown in Figure 15, and analyzed for reactivity with anti-RSA antibodies. This figure illustrates the extent of proteolysis of diRSA that occurred during the fermentation process.

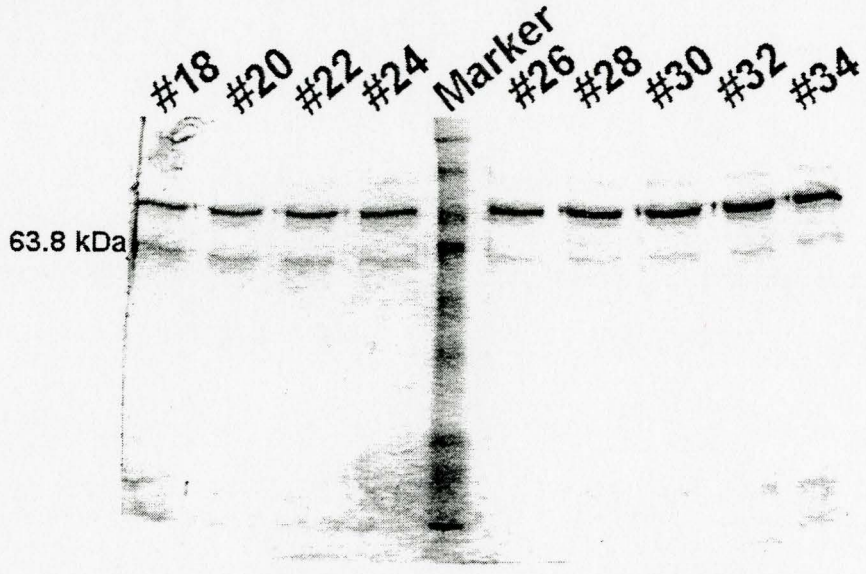
3.2.3 Purification of Novel Proteins

Nickel chelate affinity chromatography was used to purify the *Pichia pastoris* derived concentrated proteins. A series of eluted fractions electrophoresed by SDS-PAGE and stained with Coomassie Blue are shown to demonstrate the relative purity of H_6RSA and H_6RSA(C34A) proteins in Figure 16. As described below, this step resulted in two proteins of sufficient purity for radiolabeling. The dialbumin protein was also chromatographed using Ni^{2+}-NTA affinity resin but this resulted in a much more complex elution pattern. A band with greater mobility than expected for diRSA was observed, as shown in Panel A in Fig. 17. The elution profile for the pooled, concentrated and dialyzed

FIGURE 16: Purification of H_6RSA and H_6RSA-C34A by Ni^{2+}-NTA Affinity Chromatography

Aliquots of elution fractions from Ni^{2+} columns were electrophoresed by reducing SDS-PAGE (10% gels). Panel A shows fractions of WT (H_6RSA) protein eluted from a 4 mL Ni^{2+}-NTA column, while Panel B shows H_6RSA(C34A) fractions eluted from the same column (after regeneration). Invitrogen Benchmark™ prestained protein ladder was used as the marker in panel A with the position of the 63.8 kDa band noted at left. The position of the 50 kDa band of the Benchmark™ protein ladder appears at left for panel B.

A


#18
#20
#22
#24
Marker
#26
#28
#30
#32
#34
63.8 kDa

B

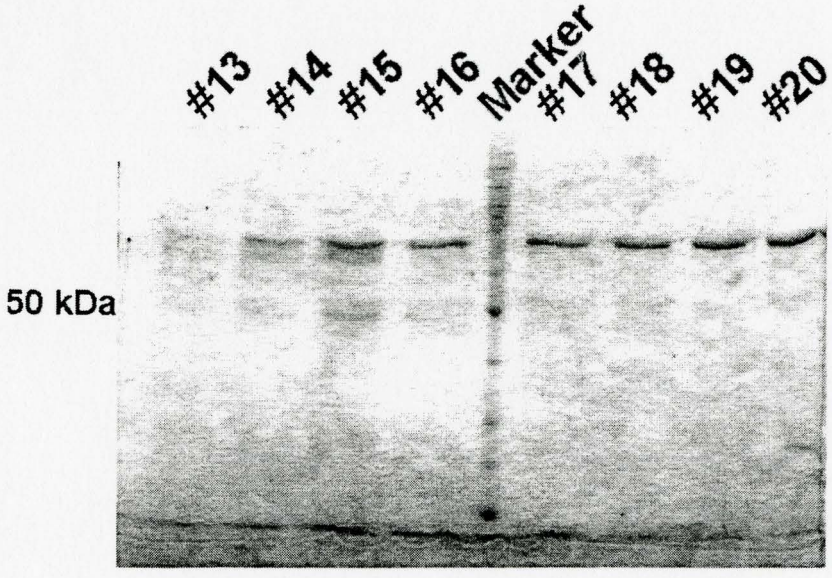
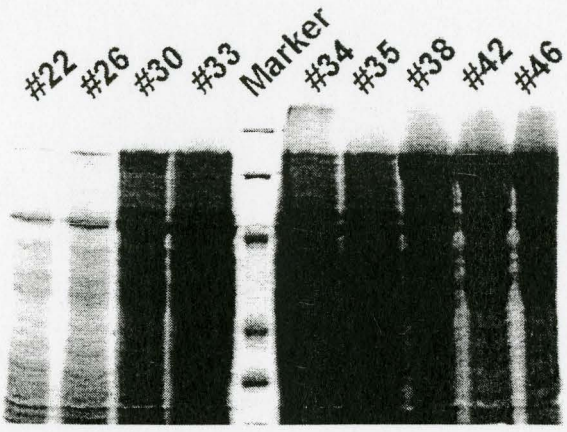

#13
#14
#15
#16
Marker
#17
#18
#19
#20
50 kDa

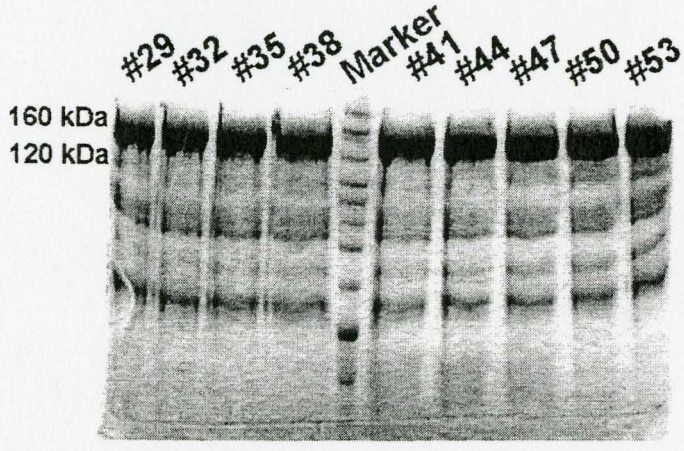
FIGURE 17: SDS PAGE for eluted fractions containing dialbumin protein

P. pastoris derived dialbumin was applied to a Ni^{2+}-NTA column, fractions collected, and an aliquot of protein-containing fractions electrophoresed by reduced SDS-PAGE (Panel A). Fractions 38-46 were pooled, concentrated and dialyzed against 20 mM Na phosphate for application to an Affigel Blue column. Aliquots of protein-containing fractions were again electrophoresed under the same conditions and this Coomassie Blue stained gel appears in Panel B. Fractions 29-50 were pooled, concentrated and dialyzed against 50 mM Na Phosphate (150 mM NaCl) for application to the sepharose 6B (Pharmacia) size exclusion column. Electrophoresed aliquots of the protein-containing fractions from this column are shown in panel C. MBI Fermentas protein molecular weight marker was used in panel A. Invitrogen Benchmark™ protein ladders were used as markers in panels B and C.

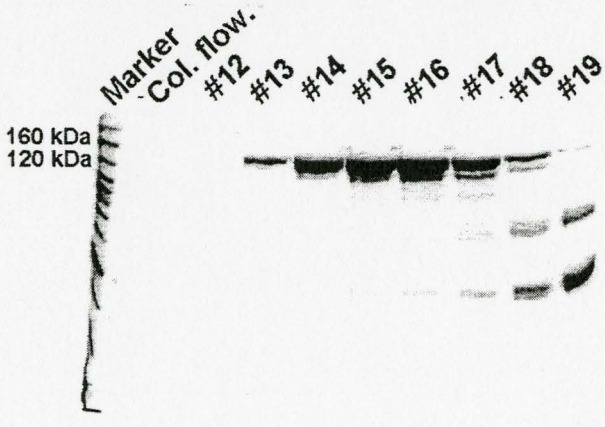
A


#22
#26
#30
#33
Marker
#34
#35
#38
#42
#46

B


#29
#32
#35
#38
Marker
#41
#44
#47
#50
#53
160 kDa
120 kDa

C


Marker
Col. flow.
#12
#13
#14
#15
#16
#17
#18
#19
160 kDa
120 kDa

dialbumin-containing fractions from the Ni-NTA column chromatographed on Affigel Blue resin can be seen in Panel B of Figure 17. While the approximately 140 kDa full-length diRSA protein was enriched, the degree of purity was still less than that of the non-reiterated RSAs. A subsequent gel filtration step with sepharose 6B (Pharmacia) was therefore used with 1M NaCl as the eluant. The elution profile for this chromatographic step can be seen in Panel C of Figure 17. All of the purified proteins were then pooled, concentrated and dialyzed against 20mM Na phosphate to remove the eluants.

3.2.4 Characterization of Recombinant Proteins

Aliquots of the concentrated, purified proteins were analyzed by electophoresis on an 8% SDS-polyacrylamide gel under reducing and non-reducing conditions. The novel proteins were compared with rabbit plasma that was reduced and unreduced. The proteins were also analyzed by immunological means using both an anti-histidine antibody and an anti-RSA antibody. Histidine tagged α_1-proteinase inhibitor (α_1-PI, also known as α_1-antitrypsin) was used as a negative control for the anti-RSA immunoblot and a positive control for the anti-hexahistidine immunoblot, while the rabbit plasma served as a positive control for the anti-RSA immunoblot and a negative control for the anti-hexahistidine immunoblot. The Coomassie Blue stained SDS-PAGE gel is shown in Figure 18, while the immunoblots of the purified proteins are shown in Figure 19 (anti-His_6 in panel A and anti-RSA in Panel B). Proteins of the expected size for H_6RSA, H_6RSA-C34A, and dialbumin were expressed by *Pichia pastoris* as shown by SDS-

FIGURE 18: Composite gel of H_6RSA(C34A), H_6RSA, dialbumin, and plasma, reduced and unreduced.

Purified recombinant H_6RSA(C34A), Wild type (H_6RSA), and dialbumin, all grown in *Pichia pastoris* yeast were analyzed by 8% SDS-PAGE electrophoresis, and compared with rabbit plasma. Each protein and the plasma were electrophoresed under reducing and non-reducing (unred.) conditions. Approximately 1μg of each protein was loaded onto the gel (excluding plasma). The gel was stained with Coomassie Blue after electrophoresis. Benchmark™ Protein Ladder (Invitrogen) was used, with the size of selected marker bands, in kDa, shown at left

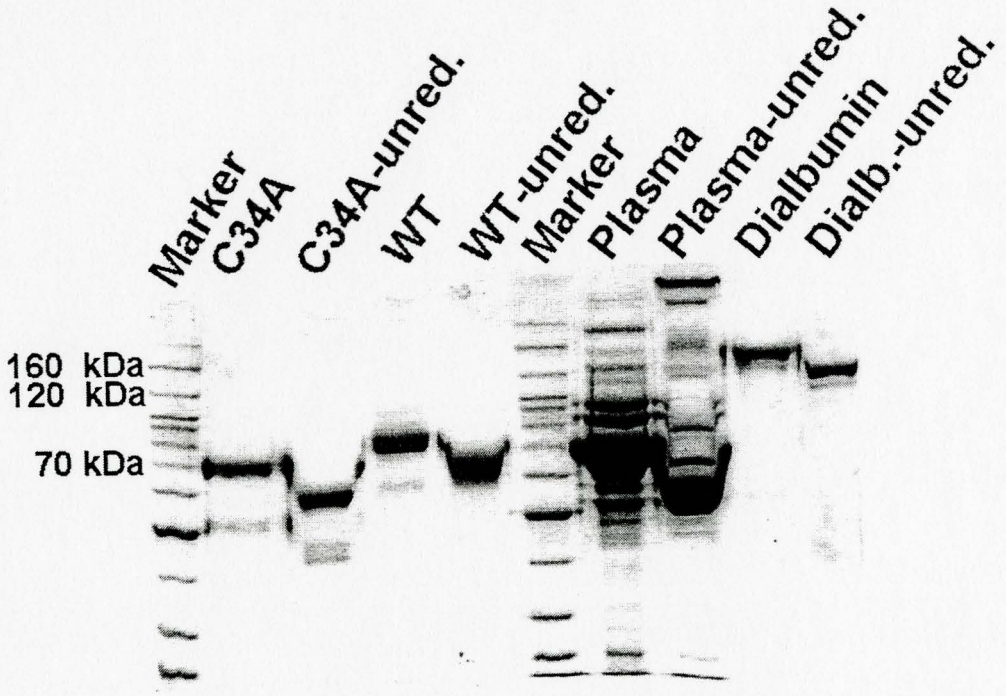
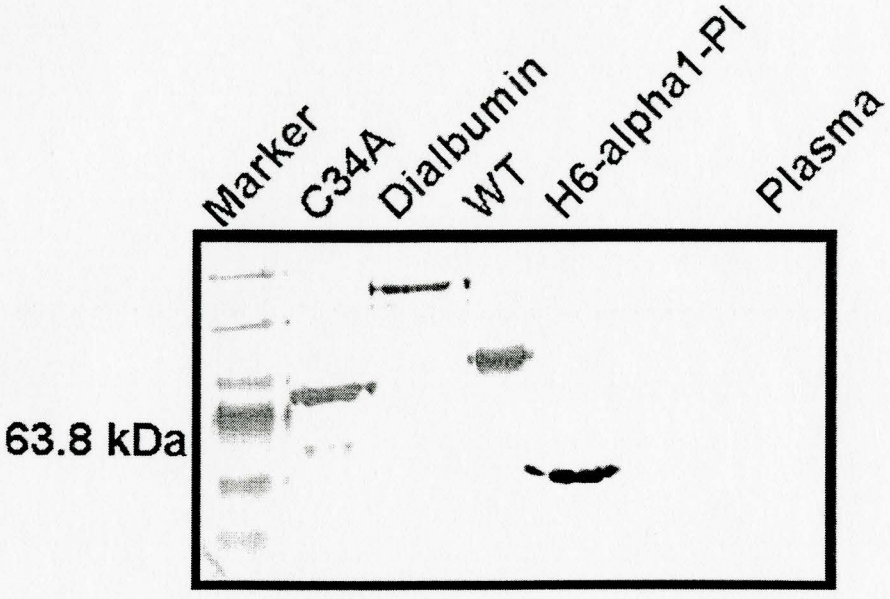

Marker
C34A
C34A-unred.
WT
WT-unred.
Marker
Plasma
Plasma-unred.
Dialbumin
Dialb.-unred.
160 kDa
120 kDa
70 kDa

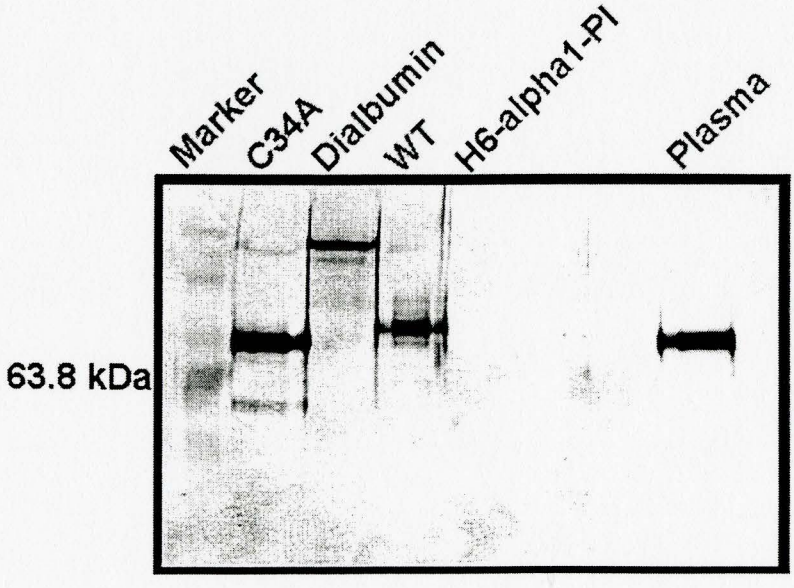
FIGURE 19: Immunoblots of recombinant proteins

Three recombinant proteins were analyzed for reaction with anti-RSA and anti-hexahistidine antibodies. of purified wild type RSA (WT), C34A-RSA (C34A), and dialbumin were compared with the same amount of hexahistidine tagged α_1-proteinase inhibitor (H6-alpha1-PI) and the equivalent of 1.0 μL of rabbit plasma, after electrophoresis on a reducing 8% SDS-PAGE gel and transfer to a nitrocellulose membrane. The immunoblot in panel A was decorated with anti-hexahistidine antibodies, while the blot in panel B was decorated with anti-RSA. Benchmark™ Prestained Protein Ladder (Invitrogen) was used, with the size of the orientation band, in kDa, shown at left.

A


Marker
C34A
Dialbumin
WT
H6-alpha1-PI
Plasma
63.8 kDa

B


Marker
C34A
Dialbumin
WT
H6-alpha1-PI
Plasma
63.8 kDa

PAGE analysis. Immunoblots showed them to be reactive with anti-RSA and anti-hexahistidine antibodies. Based on this evidence it was concluded that the purified proteins were actually the WT (H_6RSA), C34A (H_6RSA(C34A)), and dialbumin (H_6diRSA) encoded by the engineered constructs. Due to the sensitivity of the immunoblot assay, minor residual contaminating bands that are also immunoreactive can be seen on the blots.

The sizes of the H_6RSA(C34A) and dialbumin proteins were further confirmed by MALDI-TOF mass spectrometry. The mass of the C34A protein was determined to be 65,900 Da, and the dialbumin, 131,500 Da, as expected. The calculated sizes of these molecules are 66,461 Da and 132,634 Da, respectively. Although both proteins were prepared for amino acid sequencing as described in Materials and Methods, the sequencing facility was unable to obtain a sequence following automated Edman degradation.

Both the WT(H_6RSA) and H_6RSA(C34A) proteins were reacted with biotinylated PEO-maleimide and examined for evidence of subsequent biotinylation. The expected results were that the WT(H_6RSA) protein would be biotinylated through its free thiol and that the H_6RSA(C34A) mutant would be unreactive. Biotinylation was detected using band detection via decoration with streptavidin-alkaline phosphatase coloured reaction products. In each of three independent experiments, more intense staining of the WT(H_6RSA) protein than the mutant was observed (data not shown).

3.2.5 Quantification of Recombinant Proteins

The purified, pooled, concentrated, and dialyzed elution fractions from the chromatographic columns were quantified using a Bradford assay in a microtitre plate. The proteins were serially diluted to fall within the range of the standards and the program calculated protein amount by comparison with a standard curve and adjustment for the dilution factor. The overall yield for each of the three RSA-related proteins, H_6RSA, H_6RSA(C34A), and diRSA, respectively, was: 4.9, 6.8, and 0.5 mg of purified protein/L of culture.

3.2.6 Gel Filtration of Dialbumin

Albumin is a globular protein. Since dialbumin is a reiterated albumin, a gel filtration experiment was performed to investigate whether or not this molecule acts as a globular protein as well. The standards (thyroglobulin, 670,000 Da; gamma globulin, 158,000 Da; ovalbumin, 44,000 Da; myoglobin, 17,000 Da; and vitamin B_{12}, 1,350 Da) were applied to the 170 cm length, 1.5 cm diameter column of Sepharose 6B (Pharmacia), followed by an aliquot of partially purified dialbumin. The absorbance of the samples was read at 280 nm and the elution volume was calculated for the peak absorbance of each standard. These values were plotted versus molecular weight on a log scale. This graph is shown in Fig. 20. Each protein-containing sample was also analyzed on an 8% SDS-PAGE gel to investigate where the dialbumin eluted. The dialbumin was found to co-elute with the bovine gamma globulin of molecular weight 158,000 Da. It was

FIGURE 20: Graph of elution volume versus molecular weight for standards and dialbumin by gel filtration

Molecular weight standards were combined with partially purified dialbumin protein and applied to the gel filtration column. The spectrophotometric reading (at 280 nm) was measured for each elution fraction. The elution volume was calculated for each of the molecular weight protein standards and this value was plotted against molecular weight on a log scale. The point of elution for pure dialbumin (as seen on an SDS-PAGE gel) is shown by the arrow.

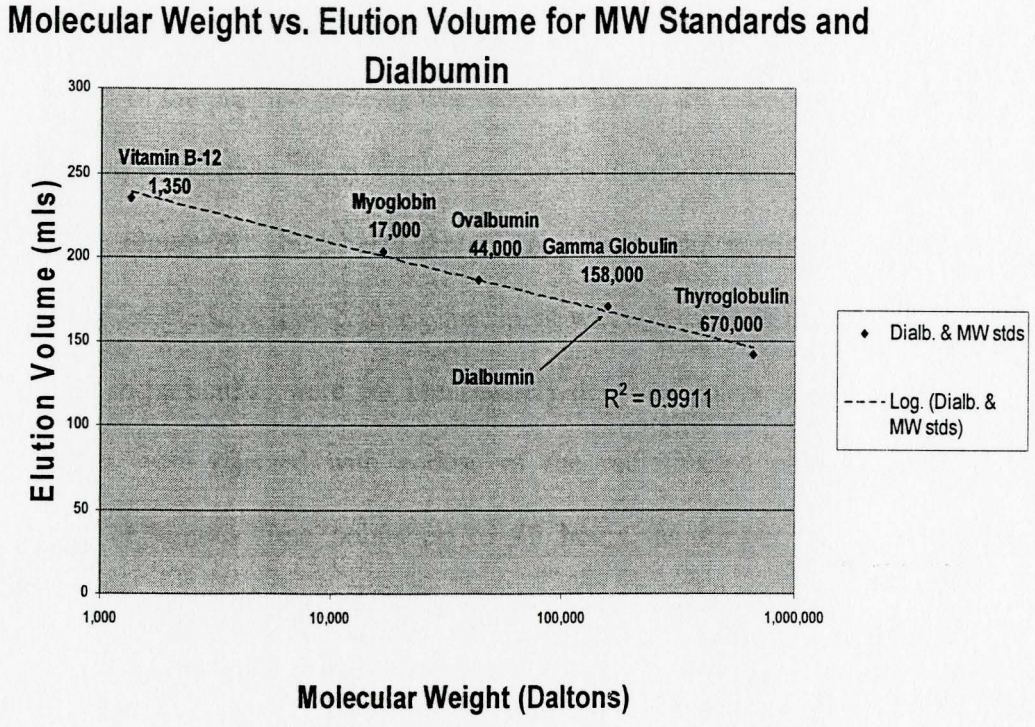

Molecular Weight vs. Elution Volume for MW Standards and Dialbumin
Elution Volume (mls)
300
250
200
150
100
50
0
1,000
10,000
100,000
1,000,000
Molecular Weight (Daltons)
Vitamin B-12
1,350
Myoglobin
17,000
Ovalbumin
44,000
Gamma Globulin
158,000
Thyroglobulin
670,000
Dialbumin
R2 = 0.9911
Dialb. & MW stds
Log. (Dialb. & MW stds)

FIGURE 21: Autoradiogram of purified proteins

The recombinant proteins expressed in *P. pastoris* were purified and labeled with either ^{125}I (C34A and WT) or ^{131}I (Dialbumin). After electrophoresis on a reducing 8% (Dialbumin) or 10% (WT and C34A) SDS-PAGE gel, the gel was dried and exposed to film to create the autoradiogram shown at right.

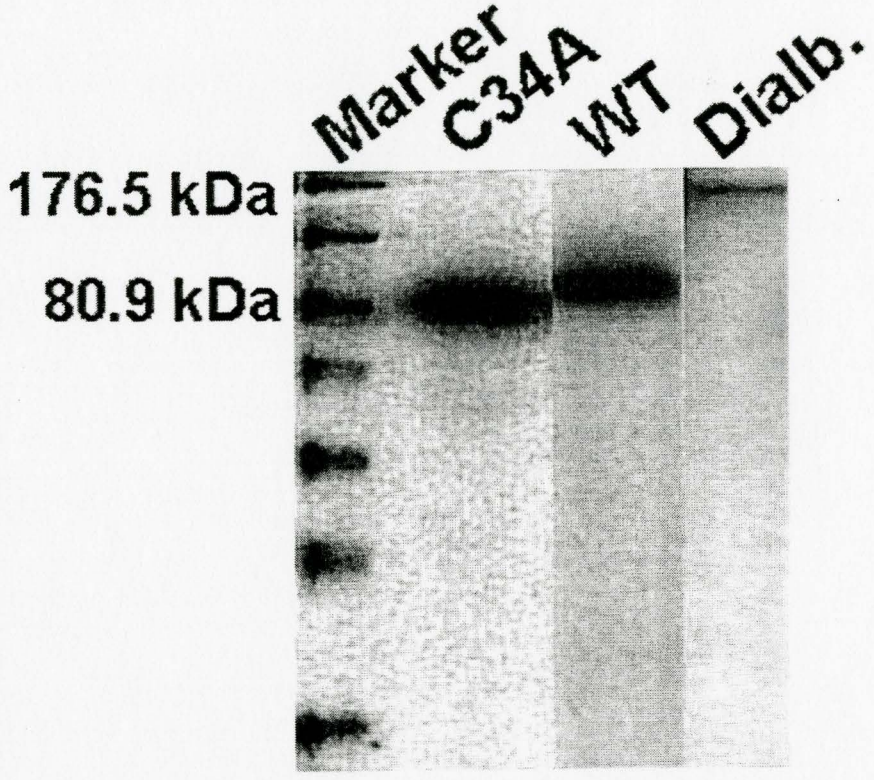
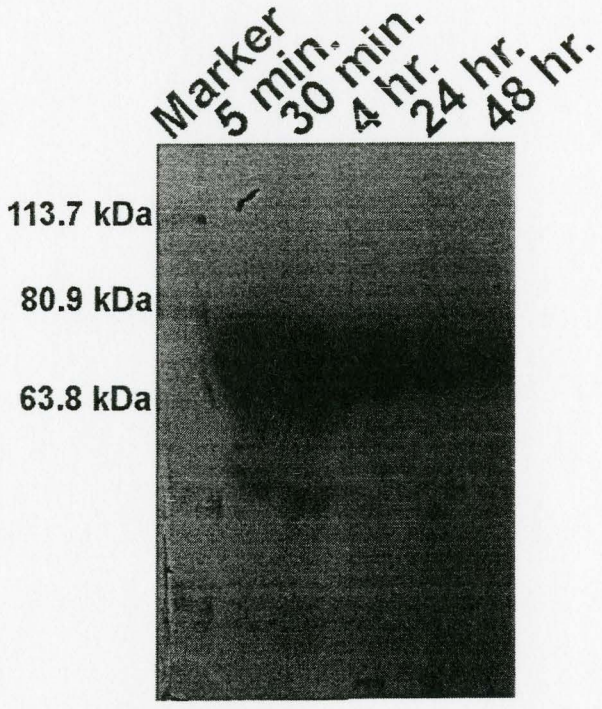

Marker
C34A
WT
Dialb.
176.5 kDa
80.9 kDa

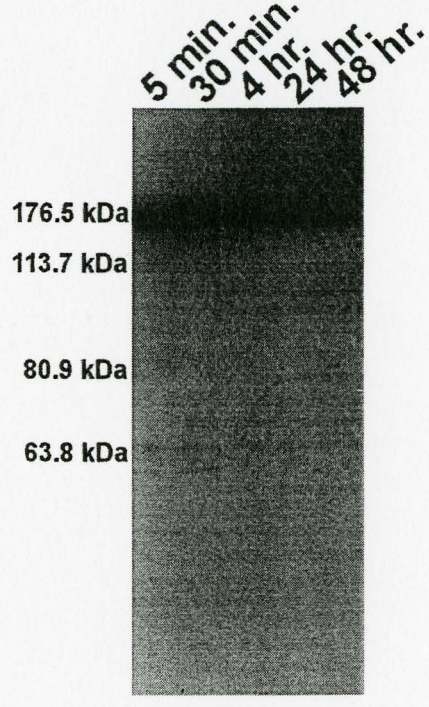
FIGURE 22: Autoradiogram of rabbit plasma to 48 hours after injection with radiolabeled recombinant protein

Rabbits were injected with ^{125}I labeled C34A, WT, or dialbumin. Blood samples were taken at the various time points shown and the cells were removed by centrifugation. 10μL of the remaining plasma was electrophoresed on a reducing 10% SDS-PAGE (C34A protein Panel A, and WT protein, Panel C) gel under reducing conditions. The dialbumin protein in Panel B was electrophoresed on an 8% SDS-PAGE gel. The gels were dried and used to expose the films.

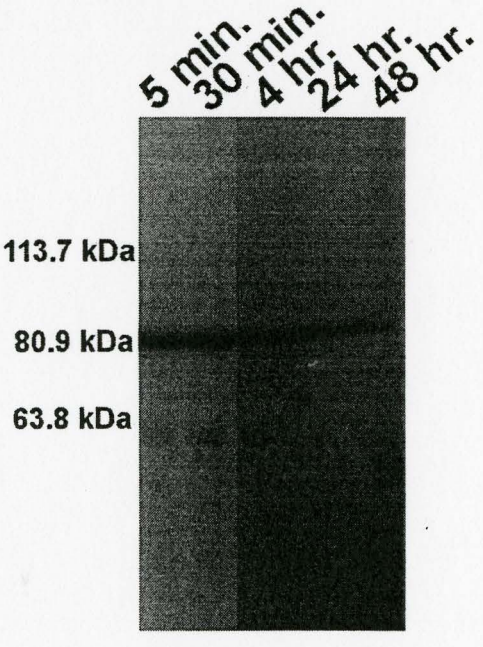
A


Marker
5 min.
30 min.
4 hr.
24 hr.
48 hr.
113.7 kDa
80.9 kDa
63.8 kDa

B


5 min.
30 min.
4 hr.
24 hr.
48 hr.
176.5 kDa
113.7 kDa
80.9 kDa
63.8 kDa

C


5 min.
30 min.
4 hr.
24 hr.
48 hr.
113.7 kDa
80.9 kDa
63.8 kDa

3.3.2 In Vivo Data Collection and Analysis

Studies were conducted comparing the clearances of two proteins per experiment. One of the proteins in each experiment was labeled with ^{125}I-sodium iodide, while the other was labeled with ^{131}I-sodium iodide. A total of fifteen rabbits were injected with radiolabeled proteins; 6 were injected with WT, 9 with dialbumin and 12 with C34A.

The New Zealand white rabbits (7 females and 8 males) of weight ranging from 2.9-4.0 kg (mean wt. 3.5 kg) were injected with the proteins. Trichloroacetic acid-precipitated plasma samples from each time point were quantified using the gamma counter to determine remaining radioactivity. The values for the counts had to first be "deconvoluted", that is, adjusted for spillover by the more energetic ^{131}I ion. Briefly, the "spillover" percentage was calculated, added to the ^{131}I column and subtracted from the ^{125}I column. The count for the first sample taken after injection, at 6 minutes, was set as 100% radioactivity (after "deconvolution"). A graph of the disappearance of radioactivity over time for data compiled from all three proteins is shown for the first four hours in Fig. 24, Panel A. Panel B shows the graph of the compiled clearance data for all time points. The graph in Panel A (first four hours) shows a similar trend for all three proteins, with no significant difference between any of the proteins. The superimposable curves for H_6RSA(C34A) and wild type (H_6RSA) on the graph in Panel B (all time points to 312 hours) are as expected. However, the dialbumin curve diverged from the other two after approximately 24 hours and exhibited a more rapid reduction. In other words, the protein cleared significantly more rapidly.

3.3.3 Pharmacokinetic Analysis

Clearance curves, similar to that in Fig.24, Panel B, were constructed for each rabbit, for each injected protein. These curves were subjected to iterative regression analysis, to obtain the α, β, and γ half-lives, compartmental distributions, and fractional catabolic rates for each protein. The averaged values (± SD), are presented completely in Table 1. A terminal catabolic half-life of 3.0 ± 0.3 days was determined for dialbumin, as compared to 4.9 ± 0.7 days for C34A, a 1.6-fold difference. Statistical determinations using one-way analysis of variance (ANOVA) found this to be an extremely significant difference ($p<0.0001$). Significant differences were also found between C34A and dialbumin for the j3 ($p<0.001$), j3.5 ($p<0.001$), and jT ($p<0.001$) values as well as the value for aP ($p<0.05$). None of the values in the chart were significantly different between WT (H_6RSA) and C34A (H_6RSA(C34A)).

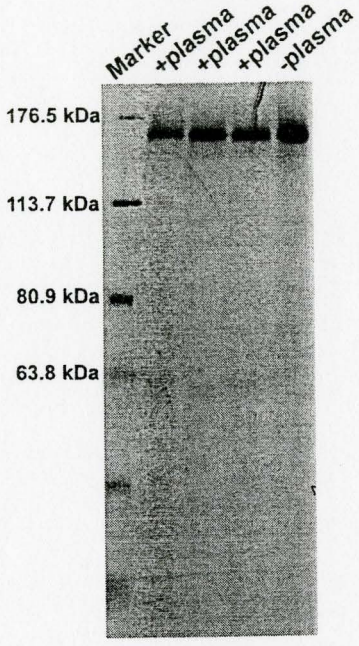
3.3.4 Organ Distribution Experiments

^{125}I-labeled proteins were injected into rabbits, and they were sacrificed after either 24, or 48 hours. The New Zealand white rabbits (10 females and 8 males) used in these experiments ranged in weight from 2.3-3.7 kg (mean weight 3.1 kg). Quantity of radiation present in five different organs was determined. These organs were the liver, lung, heart, spleen, and kidney. Figure 25 graphically depicts the findings. Panel A depicts the results after 24 hrs (n=4 for dialbumin, n=3 for WT and C34A), while the results after 48 hours (n=4 for dialbumin, n=3 for C34A, and n=1 for WT) are shown in Panel B. The radioactivity in each organ was calculated as a percentage of the total

FIGURE 23: Autoradiogram comparing electrophoretic profiles of radiolabeled diRSA with or without normal rabbit plasma

Panel A shows an autoradiogram of an SDS gel containing 10,000 cpm per lane of ^{125}I-labeled diRSA, either alone (-plasma) or combined with 20 μL of normal rabbit plasma (+plasma) prior to electrophoresis. The autoradiogram was exposed for one hour Panel B is the same, but exposed for 5 days in order that any minor bands present in the purified protein could be visualized.

A


Marker
+plasma
+plasma
+plasma
-plasma
176.5 kDa
113.7 kDa
80.9 kDa
63.8 kDa

B

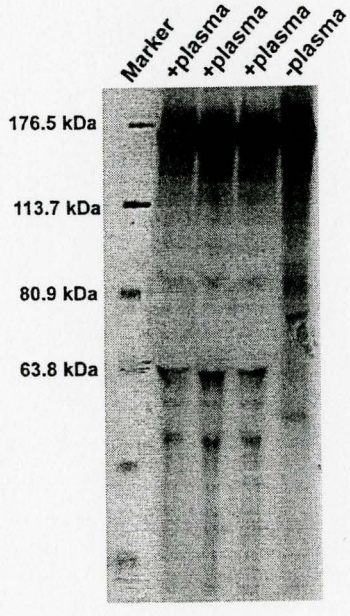
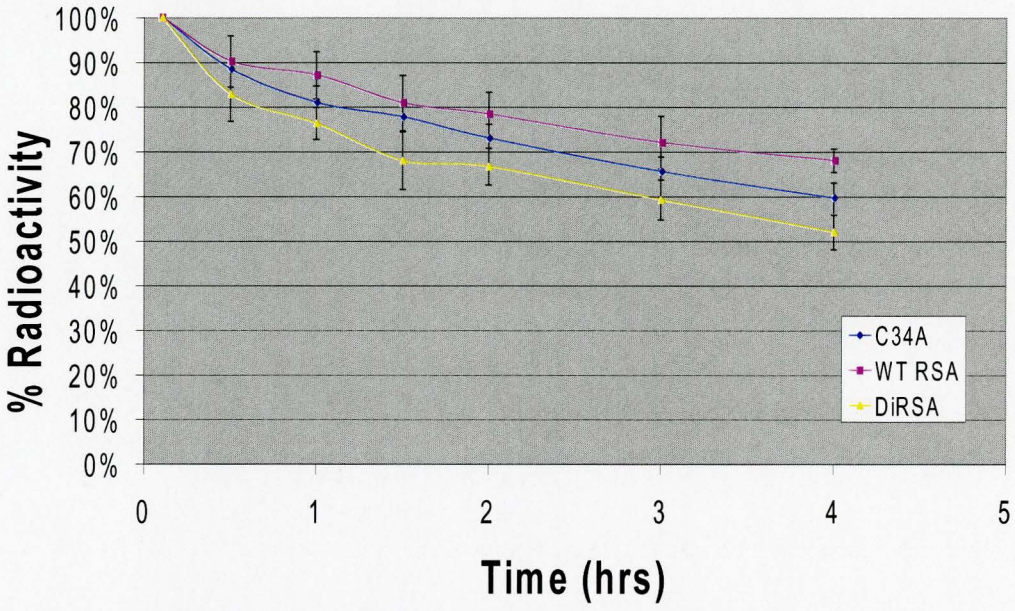

Marker
+plasma
+plasma
+plasma
-plasma
176.5 kDa
113.7 kDa
80.9 kDa
63.8 kDa

FIGURE 24: *In vivo* clearance of recombinant ^{125}I and ^{131}I-labeled proteins in rabbits

Rabbits were injected with 10 μg (approx. 30 million counts) radiolabeled H_6diRSA (dialbumin), H_6RSA(C34A), and/or H_6RSA (WT) in a typical experiment and blood samples were taken at various time points over a 13 day period. The acid-precipitable radioactivity was quantified using a gamma counter and plotted as the percentage remaining radioactivity (compared to the 6 minute value) versus time after injection (n=6 for WT; n=9 for dialbumin; n=12 for C34A, unpaired data, $\pm$ SD).

A

Clearance of Radioactively Labeled Proteins - First 4 hours



B

Clearance of Radioactively Labeled Proteins

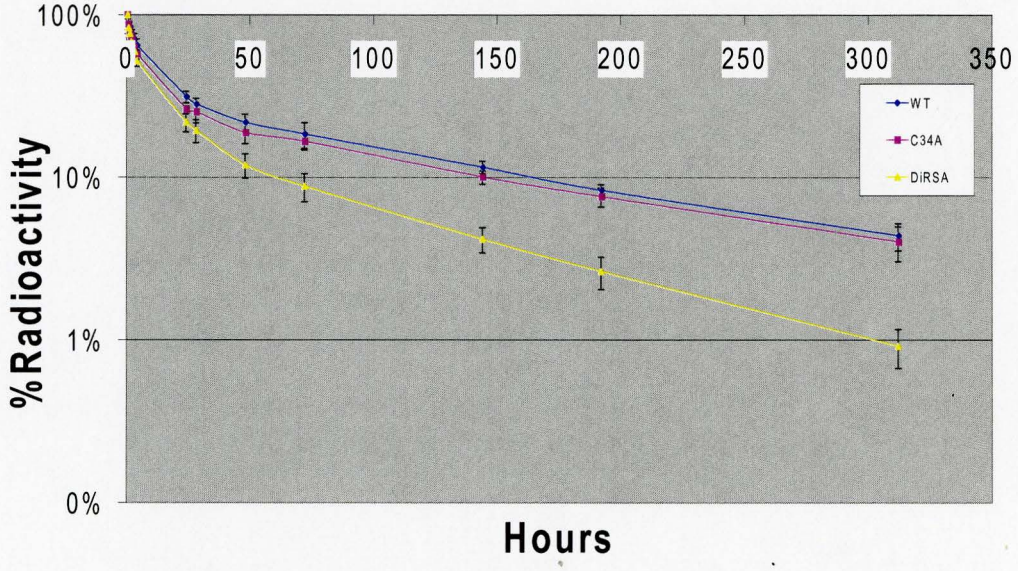


TABLE 1: Pharmacokinetic properties of radiolabeled recombinant proteins

Half-lives, compartmental distributions, and fractional catabolic rates of recombinant ^{125}I and ^{131}I-labeled dialbumin, C34A and WT RSA calculated from individual plasma curves. A three- compartment model was used. Symbols: α, circulating plasma half-life; β, half-life of protein interacting with the vessel wall; γ, terminal catabolic half-life; j_T, fraction of recombinant protein catabolized in the total body; $j_{3.5}$, fraction of protein catabolized in extravascular space and vessel wall; j_3, fraction of protein catabolized in plasma; aP, fraction of protein in plasma compartment; aW, fraction of protein interacting with vessel wall; aE, fraction of protein in extravascular space. (n=6 for WT, n=9 for dialbumin, n=12 for C34A, $\pm$ SD)

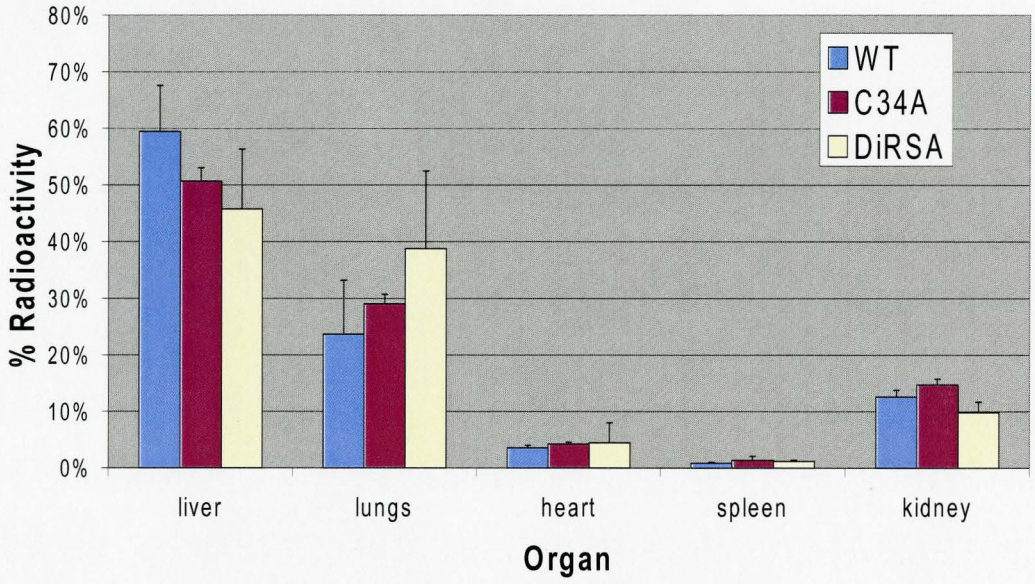
	C34A	DiRSA	WT
T1/2γ (d)	**4.9** ± 0.7	**3.0** ± 0.3*,**	**4.8** ± 0.5
T1/2β	0.26 ± 0.09	0.37 ± 0.14	0.35 ± 0.08
T1/2α	0.03 ± 0.01	0.02 ± 0.01	0.02 ± 0.01
aP	0.27 ± 0.02	0.23 ± 0.04*,**	0.31 ± 0.04
aW	0.10 ± 0.04	0.11 ± 0.04	0.08 ± 0.04
aE	0.63 ± 0.05	0.66 ± 0.06	0.61 ± 0.03
j3	0.54 ± 0.07	1.05 ± 0.21*,**	0.48± 0.05
j3.5	0.39 ± 0.05	0.71 ± 0.17*,**	0.37 ± 0.03
jT	0.14 ± 0.02	0.23 ± 0.02*,**	0.15 ± 0.02

$p < 0.05$; DiRSA vs. C34A, ,*; DiRSA vs WT, **.

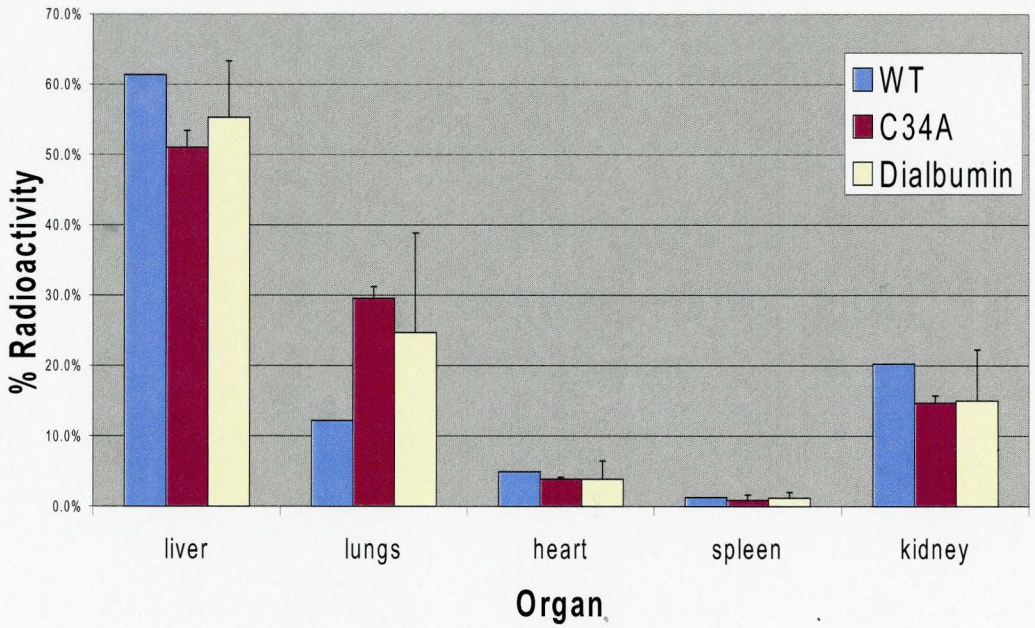
FIGURE 25: Organ distribution of recombinant radiolabeled proteins

The graphs illustrate the organ-bound radioactivity as a percentage of the total in five whole organs: liver, lung, heart, spleen, and kidney. A total of 10 rabbits were injected, then sacrificed after 24 hours (Panel A). The organs were excised, weighed, and pieces of each were placed in the gamma counter to determine residual radioactivity. The average count per minute per gram of each organ was calculated and multiplied by the weight of the organ. The total counts for the five organs were determined and counts of each organ expressed as a percentage of the total. Panel B shows the results of the 8 rabbits sacrificed after 48 hours.

A **Distribution of Radioactivity in Organs at 24 Hours**



B **Distribution of Radioactivity in Organs at 48 Hours**



amount of radioactivity determined for all five organs. As can be seen in Figure 25, the greatest accumulation of radioactivity was found in the liver for all three proteins after both 24 and 48 hours.

DISCUSSION

In this study the effects of two alterations to the structure of RSA on *in vivo* clearance in rabbits were assessed. The first alteration was the conversion of the free thiol, Cys 34, to Ala, while the second alteration was the expansion of albumin from a three-domain to a six-domain polypeptide, by reiteration of the RSA sequence in a single chain. While the physiological importance of albumin as an abundant plasma protein that maintains osmotic pressure and carries numerous other molecules has made it one of the most studied of the plasma proteins (Peters, 1985), relatively little is known concerning the mechanisms and determinants of its clearance and catabolism. Acquiring such information seemed of both biological interest, and of potential clinical relevance due to the use of albumin as a therapeutic product administered to individuals such as burn patients (Bellomo, 2002).

This investigation was a natural extension of previous work in our laboratory (Sheffield *et al.*, 2000; Sheffield *et al.*, 2001). It had been shown that recombinant C-terminally hexahistidine-tagged RSA exhibited indistinguishable clearance properties *in vivo* compared to its plasma-derived counterpart, whether expressed in African green monkey kidney (COS-1) cells or the yeast *Pichia pastoris*. Moreover, increasing the molecular volume of either hirudin (Syed *et al.*, 1997; Sheffield *et al.*, 2001) or barbourin (Marques *et al.*, 2001) by fusion to RSA increased the half-life of these molecules by 160- and 30-fold, respectively. Given the example of human immunoglobulin G, a plasma protein of approximately double the molecular size of HSA and with a 50%

longer half-life (Solomon, 1963), we hypothesized that by doubling the RSA polypeptide chain to form a larger protein would decrease its rate of clearance from the intravascular space. If so, then further extensions in the half-life of small therapeutically relevant proteins might be expected following fusion to the reiterated albumin.

It had previously been established that any truncations of RSA resulted in greatly accelerated clearance (Sheffield *et al.*, 2000), indicating a size effect in at least one direction. However, additions of N-linked glycans to RSA were either without effect, or resulted in a moderate decrease, of approximately 30%, in terminal catabolic half-life (Sheffield *et al.*, 2000). Insufficient amounts of the mutant albumins were available for a detailed structural determination of the N-glycans that were attached during albumin biosynthesis in the COS-1 cell system employed. While the additions of N-linked glycans might be expected to increase the molecular volume of RSA, they might also invoke the action of glycan-specific receptors. The reiteration strategy was designed to test the size hypothesis without introducing these additional considerations.

A factor, which could have potentially confounded our analysis, was the existence of the free thiol at residue Cys 34 in RSA. Because of this conserved feature, all mammalian albumins can dimerize via formation of a disulphide bond between two molecules. The extent of in vivo dimerization is thought to be in the range of 5-30% (Scorza *et al.*, 1998), although any manipulations to blood or plasma samples that are required to make this estimate (e.g. dilution and SDS treatment prior to electrophoresis, or protracted gel filtration experiments) could alter this value. We therefore mutated Cys 34 to Ala, a small hydrophobic residue frequently employed in alanine scanning

mutagenesis experiments (DeLano, 2002), and one incapable of disulphide bond formation.

4.1 Expression of H6-RSA-related recombinant proteins

4.1.1 Expression Strategy

Pichia pastoris was chosen as the expression system for the recombinant proteins under investigation for a number of reasons. Yeast cells are able to carry out most posttranslational modifications performed by higher eukaryotes and undertake protein folding while yielding higher quantities of protein than mammalian cells (Rosenfeld, 1999). Modifications carried out by these lower eukaryotes include appropriate disulphide bond formation, phosphorylation, myristoylation, and both O- and N-linked glycosylation, although in the latter instance yeast do not process high mannose sugars and tend to hyperglycosylate (Cregg, 1999). The ability to induce secretion of the protein into the media with methanol provides ease of purification, especially since there are few secreted yeast proteins. Another benefit of this system is the capacity to produce greater amounts of protein by scaling up through fermentation. The expression vectors provided by the supplier (Invitrogen) were designed to promote homologous recombination with the yeast genome. They lack a yeast origin of replication, and thus must integrate into the genome in order to be maintained by the yeast. By combining a dominant selectable marker, the Zeocin resistance gene, the system allows selection of yeast in which recombination has taken place. Because homologous recombination is favoured in yeast over non-homologous recombination (Cregg, 1999), especially when the recombination

is directed by the generation of free ends through linearization, the majority of Zeocin-resistant colonies are expected to yield the outcome shown in Figure 26. A double cross-over, for instance involving both the 5' AOX and the AOX transcription termination (AOTT) regions, would result in the loss of Zeocin resistance and would not be detected. A non-homologous insertion into the yeast genome would leave the engineered RSA gene either without a promoter, or under the control of a promoter fortuitously close to the site of insertion. In the latter case, expression would not be methanol-inducible. Therefore, the observation that H_6RSA, H_6RSA(C34A), and H_6diRSA were produced in a methanol-inducible fashion by Zeocin-resistant cells suggests that in each case homologous recombination took place. Direct demonstration, by Southern blotting, of this genetic event was not undertaken since the desired outcome, of protein production, was obtained.

The constructs were designed to create an RSA unit as close as possible to that found physiologically. The third residue of mature RSA, histidine, has a natural affinity for nickel. This was exploited by adding five additional histidine codons upstream of the naturally occurring one to aid purification by strengthening the binding of the tagged protein to the nickel column. The hexaglycine sequence was introduced between the reiterated albumins in order to allow proper folding of each separate entity, as was done in the previous hirudin-albumin and barbourin-albumin work. To avoid making the molecule too much larger than the naturally occurring dimer, three residues were removed from the C-terminal end of the first copy.

4.1.2 Levels of expression of non-reiterated and reiterated albumins in P. pastoris

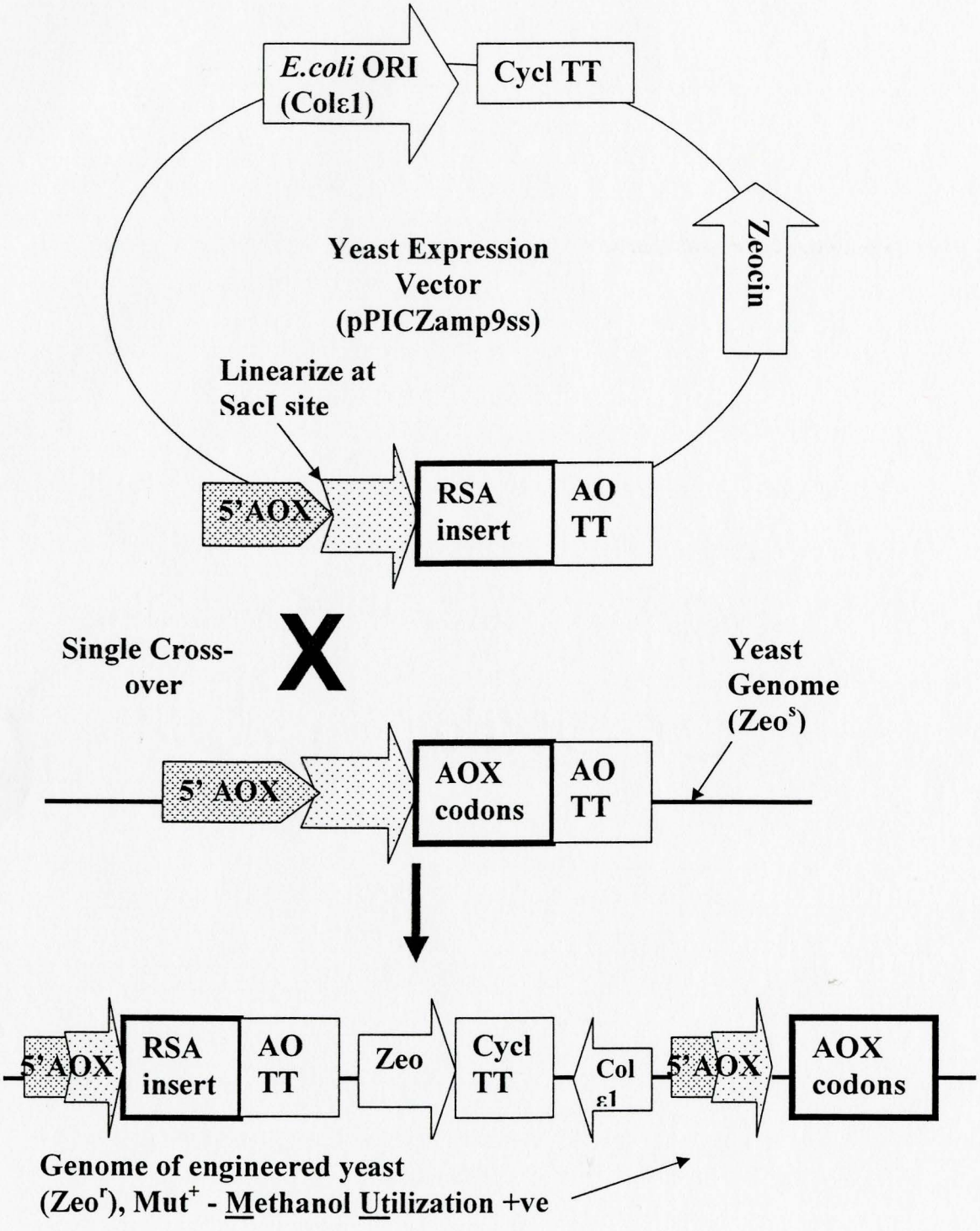
Recombinant albumin in both the H_6RSA and H_6RSA(C34A) forms was produced in high quantities, similar to levels produced previously for $HLAH_6$ and $BLAH_6$ (Sheffield *et al.*, 2001 and Marques *et al.*, 2001). DiRSA, however, was produced in much lower quantities, as can be seen in Figure 14, where equivalent quantities of conditioned media from time course fermentations were electrophoresed. Possible factors implicated in this low expression are reduced efficiency of transcription or translation, or difficulty encountered in secretion of this much larger protein. One factor that was shown to reduce the quantity of diRSA produced was proteolysis. Figure 14 (in the results section) shows the large amount of proteins of various sizes smaller than diRSA that reacted with anti-RSA antibodies by immunoblot in fractions of the conditioned media from yeast transformed with the pPZdiRSA vector. Although the sample in Lane 2 was from early on in the induction (approx. 12 hours after start of induction), the breakdown products are evident at a concentration at least equal to the concentration of diRSA. Protease inhibitors were added at the end of the fermentation to the conditioned media containing yeast, but proteolysis had already occurred to an appreciable extent as observed in Figure 14 as previously mentioned.

4.1.3 Physical properties of non-reiterated and reiterated albumins in P. pastoris

The H_6RSA(C34A) protein migrated at approximately 70 kDa on SDS-PAGE gels as expected. DiRSA, which consists of two copies of RSA(C34A), migrated at

FIGURE 26: Schematic representation of DNA insertion into the yeast genome by homologous recombination

Yeast expression vectors are first linearized using SacI, then transformed into the yeast by electroporation. Homologous sequences in the vector (located at the cut ends) promote recombination events that incorporate the gene of interest into the yeast genome in the proper area to provide protein expression when yeast metabolizes methanol.


E.coli ORI
(Colε1)
Cycl TT
Zeocin
Yeast Expression
Vector
(pPICZamp9ss)
Linearize at
SacI site
5'AOX
RSA
insert
AO
TT
Single Cross-
over
X
Yeast
Genome
(Zeo^s)
5' AOX
AOX
codons
AO
TT
5'AOX
RSA
insert
AO
TT
Zeo
Cycl
TT
Col
ε1
5'AOX
AOX
codons
Genome of engineered yeast
(Zeo^r), Mut^+ - Methanol Utilization +ve

approximately 140 kDa, also as expected. These results were also confirmed by Matrix-Assisted Laser Desorption/Ionization-Time of Flight Mass Spectrometry (MALDI-TOF MS), with the sizes given as 65,900 Da and 131,500 Da, respectively. This compares favourably with the theoretical masses of 66,461 and 132,634, with the percentage differences being 0.85% and 0.87%, respectively. These differences may result from error inherent in the MALDI-TOF technique, particularly in this high mass range. Alternatively, they could correspond to a loss, through proteolysis, of 5 (for H_6RSA(C34A)) or 10 (for H_6diRSA) amino acids from the C-terminus. N-terminal degradation can be ruled out from the Ni-NTA data, since both proteins behaved as if they retained an intact His tag. These purified proteins also reacted with anti-hexahistidine antibodies on immunoblots.

The H_6RSA protein unexpectedly migrated slightly more slowly on SDS-PAGE, appearing 3-5 kDa larger. The primary sequences of the H_6RSA and H_6RSA(C34A) only differ in the 34th codon, thus there may be a conformational or charge effect associated with this difference resulting in the varied migration. Other cases of dissimilar migrations for proteins with a single amino acid difference have been cited in the literature. Shirotani *et al.* (1999) documented a difference in migration between WT antithrombin and the Utah-mutant which differs by a single amino acid. Most of the known albumin mutants were first detected by a difference in electrophoretic mobility, many of them varying only by a single amino acid substitution. Gitlin *et al.* (1961) investigated cases of "double-albuminemia", a genetic condition in which two electrophoretically distinct albumins appeared in serum from the same individual. The

anomaly was found to be caused by a lysine residue that was present in the B form and not in the A form or normal albumin. Since the H_6RSA(C34A) molecule resembled the composition of H_6diRSA more closely than H_6RSA and its calculated and observed mass (by MALDI-TOF) were similar, and it showed an expected electrophoretic migration, the comparison of the behaviour of these two molecules *in vivo* can be made with some confidence (pending exploration of the H_6RSA's mobility).

An attempt was made to confirm that the alteration of residue 34 from Cys to Ala had actually left the C34A mutant protein devoid of reactive thiols. While the alteration at the DNA level was confirmed, and the disulphide bonding pattern of mammalian albumins is exquisitely conserved (Brown, 1976), this additional property would have provided independent confirmation of the nature of the mutant. Unfortunately, the PEO-maleimide agent is not completely thiol-specific, and likely has some reactivity against amines, according to its manufacturer (Pierce). The more intense reaction of this chemical with the wild-type RSA than its mutant counterpart is therefore suggestive, but not conclusive of the lack of a free thiol in the mutant.. Future further refinement of the experimental procedure may provide a higher degree of specificity for thiols versus amines and definitively show the presence of the WT thiol and its absence in C34A.

4.1.4 Clearance of the recombinant RSAs in vivo

Radiolabeling of the recombinant albumins resulted in essentially radiochemically pure preparations for *in vivo* studies as can be seen in Figure 21. This was an essential prerequisite for such studies, as it allowed us to assume that acid-precipitable counts

corresponded to the full-length protein of interest. Since albumin lacks widely-reported enzymatic activities (but see Hurst *et al.*, 1999 and Drmanovic *et al.*, 1999), it was not possible to assess whether iodination altered the structure of the molecule by determining if its activity was unaltered. However, the extended clearance profiles of all three RSA-related proteins examined in this study argue against substantial alterations, which would be expected to cause very early removal from the circulation by macrophages and macrophage-laden tissues, the so-called reticulo-endothelial system, which efficiently scavenge misfolded and/or foreign proteins.

While the initial purity and integrity of the iodinated RSAs and diRSA was important, it was equally important for accurate interpretation of the clearance results that it be determined whether there was significant proteolytic alteration or conversion of the proteins in the plasma, *in vivo*. Electrophoretic and radiographic analysis of plasma samples taken from rabbits injected with all three proteins showed that the proteins remained intact in plasma, and that any degradation happened elsewhere in the body. If degradation occurred in plasma, acid-precipitable intermediate products were not detected.

Complete purity can rarely be achieved in protein purification. Overexposures of autoradiograms of plasma samples from rabbits injected with diRSA showed some faint bands that migrated more rapidly than diRSA. However, appropriate exposures of iodinated diRSA electrophoresed alone, electrophoresed after incubation in normal rabbit plasma, or diRSA recovered from rabbit plasma after injection in vivo showed the same pattern and extent of trace breakdown products. This finding suggests that these proteins

were present in the initial injection material and did not form in vivo, further suggesting that the clearance analysis was appropriate and relates in large measure to the behaviour of intact diRSA.

Good agreement was shown between the results of this study and previous work on RSA clearance. The catabolic half-life of plasma-derived RSA ranged between 4.7 (± 0.4) days and 4.9 (± 0.4) days, and that of C-terminally histidine-tagged RSA ($RSAH_6$), produced in *P. pastoris*, 4.2 (± 0.3) days in a recent study in our laboratory (Sheffield *et al.*, 2000 and 2001). The $RSAH_6$ produced in the COS-1 expression system exhibited a half-life of 4.0 (±0.1) days (Sheffield *et al.*, 2001). The half-life values obtained in this project for wild type N-terminally His-tagged (H_6RSA) albumin and the H_6RSA(C34A) mutant were 4.8 (±0.5) days and 4.9 (±0.7) days, respectively. The non-significant variation between previous results and those presented here may be due to size of the experimental animals. Prior experiments employed rabbits between 2.8 and 3.2 Kg, while those used for this work were larger (average 3.5 Kg), with only two of fifteen within the smaller range. The effect of older, larger rabbits on plasma half-life of plasma-derived RSA has been seen previously in the extended time of 5.6 (±0.1) days observed by Hatton *et al.* (1993) using 1 year old rabbits of average weight 5.1 Kg.

These results suggest that the hexahistidine tag functioned as desired, as a minimal alteration to RSA structure that did not affect its clearance. This conclusion can be drawn since neither N-terminally His-tagged RSA nor C-terminally His-tagged RSA differed significantly in terminal catabolic half-life or in any other clearance parameter either from each other or from plasma-derived RSA. Extending the comparison to

H_6RSA(C34A), its clearance parameters did not differ from those of H_6RSA. This finding has important ramifications for the question of the relevance of the albumin dimer. The C34A alteration renders impossible albumin dimerization, unless previously described patterns of disulphide bond formation are incorrect. This is independently suggested by our findings with the greatly reduced reactivity of this mutant with a largely thiol-reactive reagent. Since the C34A mutant's clearance is not different from that of the wild-type tagged RSA, and as the latter can form dimers, this suggests that albumin dimerization does not affect its clearance. This lack of effect could arise either because the dimerization is fleeting, with dimers constantly forming and dissolving, or because the clearance of the dimer is identical to that of the monomer.

The diRSA protein had a shorter terminal catabolic half-life of 3.0 ($\pm$0.3) days. The clearance curve for diRSA diverged from the H_6RSA and H_6RSAC34A curves after 24 hours, but in the early part of the curve the clearance of the three proteins was indistinguishable. The distribution of diRSA was significantly lower in the plasma compartment than H_6RSA or H_6RSAC34A, but this difference was distributed between the vessel wall and extravascular space and these numbers were not significantly different from the other two proteins. The three fractional catabolic rates for diRSA were all significantly faster than the other two proteins. This faster breakdown of protein however, was not apparent in a significantly different pattern of distribution throughout the organs investigated.

Given that diRSA cleared more rapidly than either the C34A mutant or the wild-type hexahistidine-tagged recombinant RSA, either it does not accurately represent the

natural albumin dimer or the albumin dimer is not relevant to protein clearance. While we are without a crystal structure of the natural albumin dimer and of diRSA, it is difficult to assess the degree to which the two entities resemble each other, the latter explanation is also suggested by our results with the C34A mutant.

We cannot rule out the possibility that part of the explanation for the clearance profile of diRSA involves dosage effects. Varying the specific activity of labeling over a 30-fold range did not alter the clearance profile that was observed, an observation which suggests that iodination did not significantly alter the structure of the protein. With respect to the potential dosage effects, these are also unlikely. It was previously demonstrated that repetition of clearance experiments with a tracer dose of the hirudin-albumin fusion protein, using the same tracer level supplemented with 1 mg/kg unlabeled fusion protein showed no dose dependence (Sheffield *et al.*, 2001). This is likely to be the case with diRSA, and for that matter with all recombinant albumins, as they are metabolized in concert with a huge level of endogenous albumin.

4.1.5 Possible mechanisms of the accelerated clearance of diRSA

There is no convincing evidence of the existence of a receptor specific for albumin in the vasculature (but see Schnitzer *et al.*, 1988). The megalin and cubilin found in the kidney tubules do not affect the majority of albumin in the circulation, only that which is degraded and cleared by a renal route (Russo *et al.*, 2002). Albumin is known to interact with the endothelium of the vessel wall and pass through it to the extravascular space. The non-specific receptors (e.g. cell-surface fatty acids) that may

govern the uptake of albumin in the endothelium and subsequently in the tissues may have an affinity for certain sites on the albumin molecule. In the case of diRSA the number of sites with which to interact is doubled. Although on a unit basis this increase may not be relevant, there may be an effect of increased avidity. This concept is exemplified by antibody binding, in particular by pentameric immunoglobulin M (IgM). IgM - antigen interactions can be of relatively low affinity, but IgM can nevertheless trap antigen effectively because of its avidity. This term refers to the greater than additive effects of having physically linked low affinity binding sites in close proximity (Roitt and Delves, 2001).

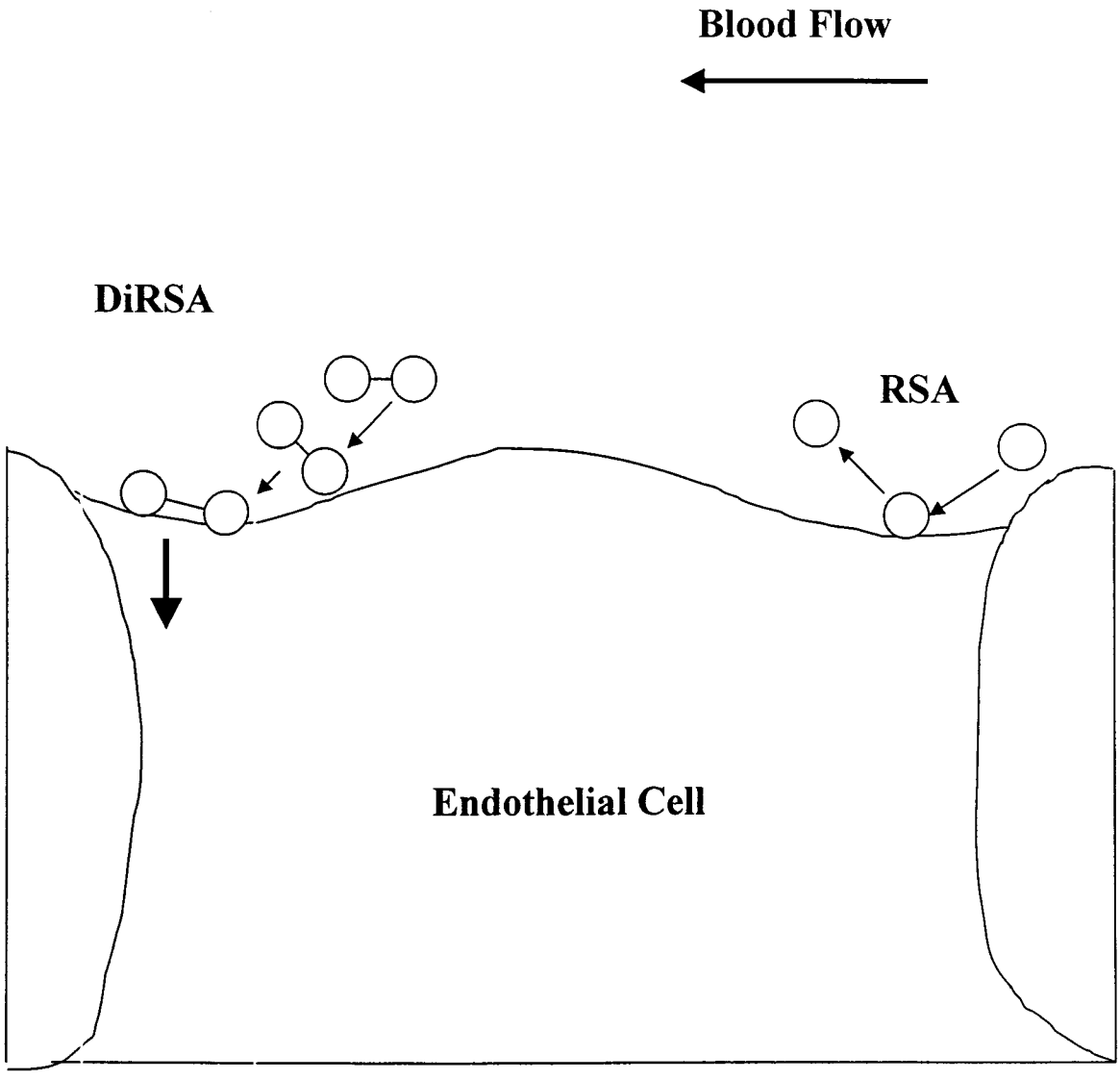
With reference to diRSA, if albumin binds to cell surfaces through low affinity interactions, then the initial binding of one portion of the diRSA will bring the second RSA into closer proximity to the cell surface. This could facilitate its binding, as shown schematically in Fig. 27. The same is possibly true of the reversal of binding; if one of the coupled RSAs is released, its physical association with a bound RSA moiety could facilitate its re-binding. These possible mechanisms, in total, could allow sufficiently increased residency time at the cell surface, and a resultant increase in internalization and degradation. This increase would not have to be enormous in nature to explain the approximately two-fold increase in terminal catabolic rate that was observed for diRSA.

4.1.6 *Possible uses of the reiterated albumin*

Although the diRSA cleared at a faster rate than the single albumins, the half-life of 3.0 days is still much longer than many of the plasma proteins. For example,

FIGURE 27: Pictorial representation of diRSA uptake.

The theorized mechanism of uptake of diRSA versus that of non-reiterated RSA is shown at right. Two hypothetical situations are depicted. For RSA, (at right), a low affinity binding event is insufficient to lead to endocytosis. In contrast, for diRSA (at left), the initial low affinity binding of one of the two linked RSAs promotes the binding of the second linked RSA. Possibly the increased avidity stabilizes the binding and permits endocytosis.


Blood Flow
DiRSA
RSA
Endothelial Cell

prothrombin, which is similar in size to albumin, at 68 kDa clears with a half-life of 1.7 days in the rabbit (Hatton *et al.*, 1995). Heparin cofactor II is also similar in size to albumin, at 70-72 kDa, but remains in the body of the rabbit for only about a third as long, with a half-life of 1.61 days (Hatton *et al.*, 1997). The larger size of the diRSA, coupled with only a 40% reduction of the time in the circulation suggests that it could be used as molecular bridge with active sites at each end of the molecule. This would be of use if two active groups needed separation to function effectively. Another way this molecule could be utilized is in cases of thrombocytopenia where there are fewer platelets than normal. Fibrinogen provides a model in this regard. Each fibrinogen molecule contains six polypeptides with an $\alpha_2\beta_2\gamma_2$ quaternary structure. Platelets aggregate when a fibrinogen or von Willebrand factor molecule engages a binding site on one platelet with one of its ends, and a second site on another platelet with another. Equipping diRSA with N- and C-terminal fibrinogen receptor binding sequences (e.g. RGDS peptides) could create an artificial fibrinogen-like molecule. This could serve as a "lead compound"-type protein for development as a way of promoting blood clotting in individuals prone to bleeding. An important consideration in this regard was that we observed no evidence of immunogenicity of diRSA; had there been an immune response in the rabbits that were treated, the terminal phase of clearance would have exhibited a discontinuity after 4-7 days.

4.1.7 Implications of this study for the significance of the albumin cys-cys dimer

Clearance of monomeric albumin occurs with a half-life between 4.6-5.5 days. The engineered dimer in this study cleared more rapidly, with a half-life of 3.0 days. This further implies that the naturally occurring dimer pool does not contribute significantly to clearance. Although we found no evidence suggesting that the albumin dimer contributed to the overall catabolism of albumin in a way different to that of monomeric albumin in normal rabbits, it could still contribute to the control of albumin concentration in situations in which albumin levels were increased in the blood. This increased concentration would likely promote dimerization. If the natural dimer acted like diRSA, it would be cleared more rapidly, dropping the albumin pool and bringing the overall concentration of albumin down into the normal range. This mechanism would provide a rapid means of adjusting albumin levels to changes in fluid balance.

4.2 *Future experiments*

The possible mechanisms suggested above to account for the increased clearance of diRSA should be tested experimentally. This could be done with relative ease by initiating studies of cultured cells of endothelial origin, as well as, in parallel studies, of the monocyte lineage which mature into macrophages. Radiolabeled H_6RSA(C34A) and radiolabeled diRSA could be compared in terms of the time course and extent of their binding to the cell surface. Initially this could be done simply by following the disappearance of label from the conditioned media, and subsequently controlled by including inhibitors of endocytosis, like chloroquine, in control plates. If the hypothesis

outlined above is correct, a more rapid uptake of diRSA than H_6RSAC34A should be observed.

With respect to possible technical refinements, strains of *P. pastoris* have been developed which are partially deficient in some proteases. These could be tested for potentially increased yields of the relatively unstable diRSA. In addition, some researchers have reported increased yields of proteins when comparing heterologous expression in *Pichia methanolicus* as opposed to *Pichia pastoris*. This approach could also be explored, although expression plasmids would have to be re-constructed using vectors optimized for use in this different yeast species.

One interpretation of the findings presented in this thesis is that albumin gene duplication stopped at the level of 3 homologous domains in mammals because, as the mammalian circulatory system was evolving, 3 homologous domains best fitted the protein to survive in the circulation for an optimum time to carry out its function and give sufficient return to the host for the energy expended during it synthesis. The finding that diRSA is more rapidly cleared than non-reiterated RSAs could mean that a further gene duplication of albumin was not an option for selection, as this would simply lead to a lower steady state level of albumin in the circulation and confer no advantage. This idea could be tested by an evolutionary biologist, keeping in mind the need to explain the further expansion of albumin in the lamprey (Gray and Doolittle, 1992), in future experiments or analyses.

Finally, the potential use of diRSA as a molecular bridge could be explored, as described above.

REFERENCES

Alexander, F., Young, P.R., Tilghman, S.M. (1984). Evolution of the Albumin: α-Fetoprotein Ancestral Gene from the Amplification of a 27 Nucleotide Sequence. *J Mol Biol*, 173:159-176.

Andreoli, T.E., Bennett, J.C., Carpenter, C.C.J., Plum, F., Smith, L.J.Jr. (1993). Cecil Essentials of Medicine. W.B. Saunders Company, Philadelphia, Pennsylvania.

Bellomo, R. (2002). Resuscitation: Colloids vs. Crystalloids. *Blood Purif*, 20:239-42.

Bennett, C.M., Glassock, R.J., Chang, R.L., Deen, W.M., Robertson, C.R, Brenner, B.M., Troy, J.L., Ueki I.F., Rasmussen, B (1976). Permselectivity of the glomerular capillary wall. Studies of experimental glomerulonephritis in the rat using dextran sulphate. *J Clin Invest*, 57:1287-94.

Bhattacharya, A., Grune, T., Curry, S. (2000). Crystallographic Analysis Reveals Common Modes of Binding of Medium and Long-Chain Fatty Acids to Human Serum Albumin. *J Mol Biol*, 303:721-732.

Bielby, J.P., Chin, C., Garcia-Webb, P., Bhagat, C.I. (1985). An Albumin Dimer in Urine. *Clinical Chemistry,* 31(3):478-479.

Brenner, B.M., Hostetter, T.H., Humes, H.D. (1978). Glomerular permselectivity: barrier function based on discrimination of molecular size and charge. *Am J Physiol,* 234:F455-F460.

Brown, J.R. (1976). Structural origins of mammalian albumin. *Federation Proc*, 35:2141-2144.

Byrne, K, Tatum, J.L., Henry, D.A., Hirsch, J.I., Crossland, M., Barnes, T., Thompson, J.A., Young, J., Sugerman, H.J. (1992). Increased morbidity with Increased pulmonary albumin flux in sepsis-related adult respiratory distress Syndrome. *Crit Care Med*: 20:28-34.

Carlson, T.H., Atencio, A.C., Simon, T.L. (1984). *In vivo* behavior of radioiodinated rabbit antithrombin III. Demonstration of a noncirculating vascular compartment. *J Clin Invest,* 74:191-9.

Carson, E.R., Jones, E.A. (1979). Use of kinetic analysis and mathematical modeling in the study of metabolic pathways *in vivo*. *N Eng J Med,* 300:1016-27.

Chang, R.L., Deen, W.M., Robertson, C.R., Brenner, B.M. (1975). Permselectivity of the glomerular capillary wall:III. Restricted transport of polyanions. *Kidney Int*, 8:212-18.

Comper, W.D., Tay, M., Wells, X., Dawes, J. (1994). Desulphation of dextran sulphate during kidney ultrafiltration. *Biochem J*, 297:31-4

Cregg, J.M. In:Gene Expression Systems: Using Nature for the Art of Expression. Fernandez, J.M., Hoeffler, J.P., eds. San Diego, California: Academic Press 1999.

Curry, S., Mandelkow, H., Brick, P., Franks, N. (1998). Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nature Structural Biology,* 5(9):827-835.

DeLano, W.L. (2002). Unraveling hot spots in binding interfaces:progress and Challenges. *Curr Opin Struct Biol*, 12:14-20.

Dice, J.F., Walker, C.D., Byrne, B., Cardiel, A. (1978). General characteristics of protein degradation in diabetes and starvation. *Proc Natl Acad Sci USA*, 75:2093-2097.

Doman, J., Fox, J., Hardwicke, J. (1980). Conditions for Albumin Dimerisation in Nephrotic Urines. Dependence on an Ultrafilterable Factor. *Clinica Chimica Acta,* 108:201-210.

Drmanovic, Z., Voyatzi, S., Kouretas, D., Sanpazidou, D., Papageorgiou, A., Antonoglou, O. (1999). Albumin possesses intrinsic enolase activity towards dihydrotestosterone which can differentiate benign from malignant breast tumors. *Anticancer Res*, 19:4113-24.

Dugaiczyk, A., Law, S.W., Dennison, O.E. (1982). Nucleotide sequence and the encoded amino acids of human serum albumin mRNA. *Proc Natl Acad Sci USA*, 79:71-75.

Fahey, J.L., Sell, S. (1965). The Immunoglobulins of Mice V. The Metabolic (Catabolic) Properties of Five Immunoglobulin Classes. *J Exp Med*, 122:41.

Fleck, A. (1985). Computer models for metabolic studies on plasma proteins. *Ann Clin Biochem,* 22:33-49.

Fraker, P.J., Speck, J C. (1978). Protein and cell membrane iodination with a sparingly soluble chloramide-1,3,4,6-tetrachloro-3,6-diphenylglycoluril. *Biochem Biophys Res Comm*, 80-849-52.

Franklin, E.C. In: Plasma Protein Metabolism: Regulation of Synthesis, Distribution, and Degradation. Rothschild, M.A., Waldmann, T., eds. New York: Academic Press 1970.

Freeman, T. (1969). Protein turnover studies in heavy proteinuria. *Br J Urol*, S:109-114.

Freeman, T.(1965). Gamma globulin metabolism in normal humans and in patients. *Ser Haematol*, 4:76.

Gburek, J., Verroust, P.J., Willnow, T.E., Fyfe, J.C., Nowacki, W., Jacobsen, C., Moestrup, S.K., Christensen, E.I. (2002). Megalin and Cubilin are Endocytic Receptors Involved in the Renal Clearance of Hemoglobin. *J Am Soc Nephrol*, 13;423-30.

Gibbs, P.E.M., Witke, W.F., Dugaiczyk, A. (1998). The Molecular Clock Runs at Different Rates Among Closely Related Members of a Gene Family. *J Mol Evol*, 46:552-561.

Gitlin, D., Schmid, K., Earle, D.P., Givelber, H. (1961). Observations on Double Albumin. II. A Peptide Difference Between Two Genetically Determined Human Serum Albumins. *J Clin Invest*, 40:820.

Gray, J.E., Doolittle, R.F. (1992). Characterization, primary structure, and evolution of lamprey albumin. *Protein Sci*, 1:289-302.

Hatton, M.W.C., Richardson, M., Winocour, P.D. (1993). On Glucose Transport and Non-Enzymic Glycation of Proteins *In Vivo*. *J Theor Biol*, 161:481-90.

Hatton, M.W., Southward, S., Ross-Ouellet, B., Richardson, M., Winocour, P. (1993). Catabolism of unglycated and naturally glycated forms of rabbit fibrinogen: The interaction with the healthy and deendothelialized aorta wall in normal and diabetic rabbits. *J Lab Clin Med*, 12:588-97.

Hatton, M.W.C., Southward, S.M.R., Serebrin, S.D., Kulczycky, M., Blajchman, M.A. (1995). Catabolism of rabbit prothrombin in rabbits: Uptake of prothrombin by aorta wall before and after a de-endothelializing injury in vivo. *J Lab Clin Med*, 126:521-9.

Hatton, M.W.C., Hoogendoorn, H., Southward, S.M.R., Ross, B., Blajchman, M.A. (1997). Comparative metabolism and distribution of rabbit heparin cofactor II and rabbit antithrombin in rabbits. *Am J Physiol*, 272:E824-E831.

He, X.M., Carter, D.C.(1992). Atomic structure and chemistry of human serum albumin. *Nature*, 358:209-215.

Hershfield, M.S., Buckley, R.H., Greenberg, M.L., Melton, A.L., Schiff, R., Hatem, C., Kurtzberg, J., Markert, M.L., Kobayashi, R.H., Kobayashi, A.L. (1987). Treatment of adenosine deaminase deficiency with polyethylene glycol-modified adenosine deaminase. *N Engl J Med*, 316:589-96.

Ho, J.X., Holowachuk, E.W., Norton, E.J., Twigg, P.D., Carter, D.C. (1993). X-ray and primary structure of horse serum albumin (*Equus caballus*) at 0.27-nm resolution. *Eur J Biochem*, 215:205-12.

Hoffenberg, R. In: Plasma Protein Metabolism: Regulation of Synthesis, Distribution, and Degradation. Rothschild, M.A., Waldmann, T., eds. New York:Academic Press 1970.

Horstick, G., Lauterbach, M., Kempf, T., Bhakdi, S., Heimann, A., Horstick, M., Meyer, J., Kempski, O. (2002). Early albumin infusion improves global and local hemodynamics and reduces inflammatory response in hemorrhagic shock. *Crit Care Med*, 30:851-855.

Hurst, R., Bao, Y., Ridley, S., Williamson, G. (1999). Phospholipid hydroperoxide cysteine peroxidase activity of human serum albumin. *Biochem J*, 338:723-8.

Keyt, B.A., Paoni, N.F., Refino, C.J., Berleau, L., Nguyen, H., Chow, A., Lai, J., Pena, L., Pater, C., Ogez, J., Etcheverry, T., Botstein, D., Bennett, W.F. (1994). A faster-acting and more potent form of tissue plasminogen activator. *Proc Natl Acad Sci USA*, 91:3670-74.

Kleinsmith, L.J., Kish, V.M. (1995). Principles of Cell and Molecular Biology. HarperCollins College Publishers, New York, New York.

Kirsch, R., Frith, L., Black, E., Hoffenberg, R. (1968). Regulation of Albumin Synthesis and Catabolism by Alteration of Dietary Protein. *Nature*, 217:578-9.

Lippincott, S.W., Korman, S., Fong, C., Stickley, E., Wolins, W., Hughes, W.L. (1960). Turnover of Labeled Normal Gamma Globulin in Multiple Myeloma. *J Clin Invest*, 39:565.

Markwardt,, F., Richter, M., Walsmann, P., Riesener, G., Paintz, M.(1990). Preparation of dextran-bound recombinant hirudin and its pharmacokinetic Behavior. *Biomed Biochim Acta*, 49:1103-8.

Marques, J., George, J. K., Smith, I.J., Bhakta, V., Sheffield, W.P. (2001). A Barbourin-albumin Fusion Protein that is Slowly Cleared *In Vivo* Retains the Ability to Inhibit Platelet Aggregation *In Vitro*. *Thromb Haemost*, 86:902-8.

Matthews, C.M.E.(1952). The theory of tracer experiments with ^{131}I-labeled plasma proteins. *Physiol Med*, 2:36-53.

Minghetti, P.P., Law, S.W., Dugaiczyk, A. (1985). The Rate of Molecular Evolution of α-Fetoprotein Approaches That of Pseudogenes. *Mol Biol Evol*, 2:347-58.

Nakano, N., Shimamori, Y., Yamaguchi, S. (1982). Binding Capacities of Human Serum Albumin Monomer and Dimer by Continuous Frontal Affinity Chromatography. *Journal of Chromatography*, 237:225-232.

Nishikawa, M., Yamashita, F., Takakura, Y., HashidaM., Sezaki, H. (1992). Demonstration of the receptor-mediated hepatic uptake ofdextran in mice. *J Pharm Pharmacol*, 44:396-401.

Nishio, H., Heiskanen, M., Palotie, A., Belanger, L., Dugaiczyk, A. (1996). Tandem Arrangement of the Human Serum Albumin Multigene Family in the Sub-Centromeric Region of 4q:Evolution and Chromosomal Direction of Transcription. *J Mol Biol*, 259(1):113-119.

Ohtani,W., Nawa, Y., Takeshima, K., Kamuro, H., Kobayashi, K., Ohmura, T. (1998). Physicochemical and Immunological Properties of Recombinant Human Serum Albumin from *Pichia pastoris*. *Analytical Biochemistry*, 256:56-62.

Peters, T. (1985). Serum Albumin. *Advances in Protein Chemistry*, 37:161-245.

Poznansky, M.J., Halford, J., Taylor, D. (1988). Growth hormone-albumin conjugates. Reduced renal toxicity and altered plasma clearance. *FEBS Lett*, 239:18-22.

Regoeczi, E. (1984). Iodine-Labeled Plasma Proteins. CRC Press, Inc., Boca Raton, Florida.

Roitt, I.M., Delves, P.J. (2001). Roitt's Essential Immunology Blackwell Science Ltd., Malden, Massachussettes.

Rosenfeld, S. (1999). Use of *Pichia pastoris* for Expression of Recombinant Proteins. *Methods in Enzymology*, 306:154-169.

Rothschild, M.A., Waldmann, T. (1970). Plasma Protein Metabolism: Regulation of Synthesis, Distribution, and Degradation. Academic Press, New York, New York.

Rothschild, M.A., Oratz, M, Evans, C., Schreiber, S.S. (1964). Alterations in Albumin Metabolism after Serum and Albumin Infusions. *J Clin Invest*, 43:1874-1879.

Russo, L., Bakris, G.L., Comper, W.D. (2002). Renal handling of albumin: A critical review of basic concepts and perspective. *Am J Kidney Dis*, 39:899-919. 39:

Sanger, F., Nicklen, S., Coulson, A.R. (1977). DNA sequencing with chain terminating Inhibitors. *Proc Natl Acad Sci USA*, 74:5463-7.

Schnitzer, J.E., Carley, W.W., Palade, G.E. (1988). Albumin interacts specifically with a 60-kDa microvascular endothelial glycoprotein. *Proc Natl Acad Sci USA*, 85:6773-77.

Schwick, H.G., Heide, K. (1977). Trends in human plasma protein research, *Trends Biochem Sci*, 2:125.

Scorza, G., Minetti, M. (1998). One-electron oxidation pathway of thiols by peroxynitrite in biological fluids: bicarbonate and ascorbate promote the formation of albumin disulphide dimers in human blood plasma. *Biochem J*, 329:405-413.

Sheffield, W.P., Smith, I.J., Syed, S., Bhakta, V. (2001). Prolonged *in vivo* anticoagulant activity of a hirudin-albumin fusion protein secreted from *Pichia pastoris*. *Blood Coagul Fibrinolysis*, 12:433-443.

Sheffield, W.P. (2001). Modification of Clearance of Therapeutic and Potentially Therapeutic Proteins. *Cardiovasc & Haemat Dis*, 1:1-22.

Sheffield, W.P., Marques, J.A., Bhakta, V., Smith, I.J. (2000). Modulation of Clearance of Recombinant Serum Albumin by Either Glycosylation or Truncation. *Thrombosis Research*, 99:613-621.

Shirotani, H., Tokunaga, F., Koide, T. (1999). Cellular and Functional Characterization of Three Recombinant Antithrombin Mutants That Caused Pleiotropic Effect-Type Deficiency. *J Biochem*, 125:253-62.

Sudlow, G., Birkett, D.J., Wade, D.N.(1975). The characterization of two specific drug binding sites on human serum albumin. *Mol Pharmacol*, 11:824-32

Sudlow, G., Birkett, D.J., Wade, D.N. (1976). Further characterization of specific drug binding sites on human serum albumin. *Mol Pharmacol*, 12:1052-61

Syed, S., Schuyler, P.D., Kulczycky, M., Sheffield, W.P. (1997). Potent antithrombin activity and delayed clearance from the circulation characterize recombinant hirudin genetically fused to albumin. *Blood*, 89(9):3243-52.

Thompson, A.R. (1982). Radioimmunoassay of factor IX in the hemophilias. *Methods Hematol*, 5:122-131.

Vyas, G.N., Fudenberg, H.H. (1969). Am(1), the first genetic marker of human immunoglobulin A. *Proc Natl Acad Sci USA*, 64:1211.

Waldman, T. In: Plasma Protein Metabolism: Regulation of Synthesis, Distribution, and Degradation. Rothschild, M.A., Waldmann, T., eds. New York: Academic Press 1970.

Widmann, F.K., Itatani, C.A. (1998). An Introduction to Clinical Immunology and Serology. F.A. Davis Company, Philadelphia, Pensylvania.

Yabe, Y., Nishikawa, M., Tamada, A., Takakura, Y., Hashida, M. (1999). Targeted Delivery and Improved Therapeutic Potential of Catalase by Chemical Modification: Combination with Superoxide Dismutase Derivatives. *The Journal of Phamacology and Experimental Therapeutics*, 289:1176-84.

Yeh, P., Landais, D., Lemaitre, M., Maury, I., Crenne, J., Becquart, J., Murry-Brelier, A., Boucher, F., Montay, G., Fleer, R., Hirel, P., Mayaux, J., Klatzmann, D. (1992). *Proc Natl Acad Sci USA*, 89:1904-1908

Zhou, C., Yang, Y., Jong, A.Y. (1990). Mini-prep in ten minutes. *Biotechniques*, 8:172-173.