Metal Binding and Redox

Properties of Human Albumin

A Thesis

Submitted for the Degree of Doctor of Philosophy

in the

Faculty of Science and Engineering

by

Alan James Stewart BSc (Hons) GIBiol AMRSC

School of Chemistry

January 2003

Joseph Black Building

University of Edinburgh

Abstract

Albumin is the major protein in the blood and has a wide range of physiological functions including the involvement in the transport and metabolism of a variety of drugs, metal ions, hormones and fatty acids.

Cys^{34} of human albumin is known to be involved in a variety of physiologically important redox reactions and is a binding site for a range of therapeutics, particularly Au(I) antiarthritic drugs. The importance of His^{39} and Tyr^{84} in the control of the reactivity of Cys^{34} was investigated. The disulfide exchange reaction between Cys^{34} and DTNB was studied over a range of pH values (pH 6-10) for native rHA, octanoate-rHA and $His^{39}Leu$ and $Tyr^{84}Phe$ mutants. A large increase in reactivity toward DTNB was observed for the albumin mutant $Tyr^{84}Phe$, over this pH range. This shows that Tyr^{84} is important in the control of Cys^{34} reactivity. Wildtype albumin appears to become fully deprotonated around pH 8.4, whilst the $His^{39}Leu$ and $Tyr^{84}Phe$ mutants become fully deprotonated around pH 9.1. This suggests that the pK_a of the free thiol is higher in both mutants than in the wildtype protein and that both His^{39} and Tyr^{84} contribute to the stabilisation of the thiolate. Elevated fatty acid levels had no significant effect on Cys^{34} reactivity.

A novel switchable metal binding site at the interface of domains I and II has been identified. ^{111}Cd NMR studies show that Cd^{2+} ions bind at two sites on albumin, each giving rise to a distinct peak. Competition studies revealed that at one of these sites (site A) Zn^{2+}, Cu^{2+} and Ni^{2+} were able to displace $^{111}Cd^{2+}$. The chemical shift of this site

indicates the involvement of 2-3 histidinyl nitrogens for metal coordination. Analysis of the X-ray structure of albumin revealed that there is only one site on the molecule where 2 histidine residues lie close together. At this site, His^{67} and His^{247} are located within 5 Å of each other and additionally, Asn^{99} and Asp^{249} are nearby and are available as potential O-ligands for metal binding. It is shown that the His^{67}Ala mutation disrupts Cd^{2+} and Cu^{2+} binding by ^{111}Cd NMR and UV/vis spectrophotometric studies respectively. Weak Cl^- binding to the fifth coordination site was demonstrated for Cd^{2+} (log $K_d = -0.66$). Cd^{2+} binding was dramatically affected by high fatty-acid loading of albumin. Analysis of the X-ray structures suggests that fatty acid binding to site 2 triggers a spring-lock mechanism, which disengages the upper (His^{67}/Asn^{99}) and lower (His^{247}/Asp^{249}) halves of the metal site. These findings provide a possible mechanism whereby fatty acid (and perhaps other small molecule) binding to albumin could influence the transport and delivery of zinc in blood.

Albumin is a physiologically important antioxidant. Reactions of albumin with H_2O_2 and NO have been studied. Peptide mapping studies of albumin following reaction with H_2O_2, revealed that the residues Cys^{34}, Met^{123}, Met^{298} Met^{87} and Met^{329} are potential sites of peroxide-induced modification. ^1H and ^{15}N NMR studies are shown to be useful techniques in the study of the *S*-nitrosylation reactions between low-molecular weight thiols and acidified ^{15}N-labelled $NaNO_2$ but were unsuccessful in detecting *S*-nitrosylation of albumin. Mass spectrometry was found to be a suitable technique in the study of albumin-NO reactions. However, no transfer of NO from low-molecular weight *S*-nitrosothiols to native albumin occurs under the conditions used.

Albumin binding studies involving two cytotoxic ruthenium(II) complexes ($[(\eta^6$-p-cymene)RuCl(H$_2$NCH$_2$CH$_2$NH$_2$-N,N)]$^+$PF$_6^-$ and $[(\eta^6$-C$_6$H$_5$C$_6$H$_5$)RuCl(H$_2$NCH$_2$CH$_2$NH$_2$-N,N)]$^+$PF$_6^-$) and a novel platinum(II) complex ([Pt(NP$_3$)Cl]Cl) previously shown to have affinity toward low-molecular weight thiol compounds were carried out. Gel permeation chromatography showed that peaks corresponding to the free Ru(II) complexes, decreased upon the addition of albumin, indicative of binding. Following the addition of [Pt(NP$_3$)Cl]Cl, the absorbance of albumin at 398 nm increased, showing binding to have occurred.

Crystallisation trials with rHA, and of various albumin-Cu derivatives were carried out. Unfortunately, none of the crystals diffracted X-rays to greater than 5.45 Å.

"Research is to see what everybody else has seen, and to think what nobody else has thought."

Albert Szent-Györgi (1893-1986), U. S. biochemist

"The power of accurate observation is frequently called cynicism by those who don't have it."

George Bernard Shaw (1856-1950), Irish author and socialist

"Research is the process of going up alleys to see if they are blind."

Marston Bates (1906-1974), U. S. zoologist

"A lot of fellows nowadays have a B.A., M.D., or Ph.D. Unfortunately, they don't have a J.O.B."

"Fats" Domino, U. S. musician

For Sophie

Acknowledgements

School of Chemistry, University of Edinburgh

I would first of all like to thank Professor Peter Sadler for his excellent supervision and his ideas, advice and patience.

Dr. Claudia Blindauer for her help and expertise in all the NMR studies detailed in this thesis as well as the molecular modelling described in Chapter 5. The reactions detailed in Section 5.3.1 were carried out together with Frank Toner as part of his BSc Pharmacology Honours project.

The following people of, or formerly of, the Sadler group for their help and advice with various aspects of my work: Drs. Abraha Habtemarium, Socorro Murdoch, Stephen Paisey and John Parkinson. I would also like to thank all my friends in the Sadler group past and present (especially Kerry Bunyan), for the great atmosphere in the lab.

Delta Biotechnology Ltd., Nottingham

I would like to thank Dr. Steve Berezenko for his supervision, Dr. John Woodrow and Lee Blackwell for their efforts in enhancing my learning of protein purification techniques. Drs. Darrell Sleep and Les Evans for their help in synthesising the yeast mutants and Dr. Chris Finnis for synthesising the Cys^{34}Ala rHA mutant. Dave Tooth and Neil Dodsworth for their expertise in amino acid sequencing and mass spectrometry and Tony Greenfield his help in expressing the mutant proteins.

Table of Contents

Chapter 1

Introduction

1.1 Clinical Significance of Albumin

Human serum albumin (HSA) is the most abundant protein in the circulatory system (constituting around 60% of total plasma protein) and is one of the most extensively studied. Typically, albumin is present in the blood at a concentration of 42 g L^{-1} and contributes 80% to colloid osmotic blood pressure [1]. It is also thought to be primarily responsible for the maintenance of blood pH [2].

Albumin's most notable feature is its ability to bind reversibly to a vast array of ligands, such as fatty acids, hormones, metal ions and drugs. Albumin is the major transport protein in the blood and has been described as the most multi-functional transport protein known to date [1]. The drug-binding properties of albumin are very important to the pharmaceutical industry, as it is possible to alter the distribution, metabolism and efficacy of a drug in the body, solely by its affinity to albumin. This is in part due to the protein's high abundance in the blood, and also because it can transport and deliver a bound drug to its target site through the circulatory system. Many drugs are therefore designed not only to be potent in their desired function but also to reversibly bind to serum albumin. Some examples of therapeutic agents that show affinity towards albumin are, cisplatin [3], auranofin [4], warfarin [5], diazepam [6] and digitoxin [7].

A variety of albumin-associated human disorders exist. Hyperalbuminaemia is characterised by elevated levels of albumin in the blood, but is thought almost always to be due to dehydration rather than by an increased production of the protein. Hypoalbuminaemia also exists and can be the consequence of either a reduction in albumin synthesis (a symptom of some liver diseases), or of an increased loss of the protein, which may be seen in individuals suffering from nephritic syndrome and occasionally in patients with extensive burns. Hypoalbuminaemia may also be caused by malnutrition [8].

Another disorder, analbuminaemia is an extremely rare genetic trait in which the *alb* gene, which encodes albumin, is intact but modified in such a way as that its transcription product is non-functional [9]. Remarkably, around half of all patients diagnosed analbuminaemic, appear to be completely asymptomatic, but those who do suffer from chronic fatigue, lipodystrophy, a decrease in androgen synthesis and hyperlipidaemia [8]. The latter it is thought can be attributable (at least in part) to a compensatory increase in the levels of other plasma proteins [10].

Over 100 other genetic variants of albumin have been identified, usually through routine electrophoresis of serum or plasma samples. These variant albumins (or alloalbumins) usually carry point mutations in their coding or pre-coding sequence, altering their amino acid sequence and their secondary structure. There is evidence to suggest that these mutations do not occur completely randomly and that certain sites or hot spots are present within the genetic sequence of albumin, where mutations are found more frequently [11]. Examples of these hotspots are shown in Table 1.1.

Mutation	Codon Change	Variant Name
$Arg^{-2} \rightarrow His$	$CGT \rightarrow CAT$	Lille, Pollibauer, Somalia, Tokushima, Taipei, Fukuoka-2, Varese, Wu Yang, Mayo (EW220), Komagome-3 and Stirling.
$Arg^{-1} \rightarrow Gln$	$CGA \rightarrow CAA$	Christchurch, Gainesville, Y, Honolulu-2, Fukuoka-3, Mayo (JW180), Shizuoka and Kamloops.
$Glu^{570} \rightarrow Lys$	$GAG \rightarrow AAG$	Oliphant, Verona, Osaka-2, Phnom Penh, Nagano, Iowa City-3, Mayo (MT610, RW246 and SH420), Victoria (East India), Saitama-1, Ann Arbor, London (Ontario), Lübeck, Tokyo-1 and Shinanomachi-1.

Table 1.1 Mutations of human serum albumin and proalbumin. The table shows 35 variants that were given different names (usually the place of discovery), most probably because they were originally thought to be different. Detailed studies at the molecular level revealed that out of all these variants only 3 different alloalbumins exist. Source: ref [11].

Interestingly, a research group from the University of Hawaii identified a genetic variant of human albumin, which may have an application in the treatment of hyperthyroid diseases [12]. This variant appears to have an abnormally high affinity for thyroxin, the hormone that is over-expressed in hyperthyroidism. In preliminary tests this variant has been shown to be capable of quickly and effectively removing excess thyroxin from the body. Thyroxin is responsible for regulating heart rate, body temperature and caloric

consumption. Illnesses such as thyroid storm (where thyroxin production increases to dangerous levels during pregnancy) and Graves disease (where the body overproduces antibodies which stimulate the thyroid gland) can lead to heart disease and cause blood clots, often blocking major arteries. One of the main advantages of using this variant in the treatment of hyperthyroidism is that it cannot cross the placental barrier and so could be used to treat pregnant women without worry of affecting the foetus. It is also thought that because the protein has been found to exist naturally in humans, it may be a safer and more effective option than other treatments.

Albumin has also demonstrated functionality as an extracellular antioxidant [13] and has been found to have the ability of detoxifying several potentially hazardous chemicals in the blood (e.g. bilirubin [14] and cyanide [15]). There is, additionally, some evidence to suggest that in humans, serum albumin may inhibit apoptosis in endothelial cells [16]. The absence of albumin in rats makes them more susceptible to developing cancer [17]. Additionally, succinylated albumin has been reported as having anti-HIV activity [18]. Albumin has also been suggested as a possible source of amino acids for various tissues [19]. Other potential applications of albumin include the modification of the protein with ^{111}In for use as a selective contrast agent in the detection and treatment of tumours [20], whilst albumin chimeric molecules such as HSA-CD4 [21], HSA-Cu, Zn superoxide dismutase and rabbit serum albumin-hirudin [22] have been synthesised to increase the half-lives and distribution of these protein therapeutics, and reduce immunological responses towards them.

Clinically, human albumin is used in the treatment of patients with severe burns, shock or blood loss. Presently, it is produced commercially by extraction from human blood [23]

and because of this, periodic shortages of albumin have affected availability worldwide. Recently, Delta Biotechnology Ltd. has produced recombinant albumin (Recombumin®) in yeast that has successfully passed phase I clinical trials. An intramuscular study on 500 subjects was conducted under an FDA IND (Food and Drug Administration Investigation of a New Drug) in order to evaluate the tolerability and safety of Recombumin® compared with human serum albumin. Recombumin® was administered to 250 subjects in ascending doses of up to 65 mg. A significant safety margin was demonstrated in this study, considering that the expected highest single dose of albumin in currently approved pharmaceutical formulations is 15 mg. An intravenous study on 30 subjects was also completed to compare the tolerance, safety and pharmacodynamics of Recombumin® with human serum albumin. Recombumin® was administered to 15 subjects at increasing dose levels up to 50 g and was as well tolerated as the control drug human serum albumin [24]. This technology is designed to increase the availability of the protein and in certain cases, prevent the need for screening of native albumin for infections such as HIV, herpes and hepatitis.

A review of randomised controlled trials published by The Cochrane Injuries Group cast severe doubts regarding the appropriate use of albumin in the treatment of some medical conditions [25]. Their review concluded that the relative risk of death in albumin-treated patients suffering from hypovolaemia, burns or hypoproteinaemia was 6% higher than in those treated by other means. This figure constitutes 6 more deaths per 100 patients treated. The Cochrane study has had many critics but after the report's publication, albumin usage dropped dramatically in the UK. In February of 2000, the plasma products industry launched a £1.4M ($2.2M) international programme to promote albumin usage [26].

1.2 Albumin Structure

1.2.1 Primary, Secondary and Tertiary Structure

Human albumin is a 66.5 kDa single-chain protein of 585 amino acids. The chain contains 35 cysteine residues, which form 17 disulfide bridges, leaving a free thiol at Cys^{34}, and a single tryptophan (Trp^{214}). The protein also contains an unusually large number of charged amino acids and contains no sites for enzymatic glycosylation. In the human body approximately 10% of albumin is glycosylated non-enzymatically [27]. This, it is believed occurs primarily on Lys^{525} and is thought to alter its properties in such a way as to facilitate its ingestion into endothelial cells in preference to unmodified albumin [28].

The primary structure of the protein is characterised by a distinct loop-link-loop pattern, due to the nature of the protein's disulfide bonding (Figure 1.1). Sequence homology between human serum albumin and other mammalian albumins is relatively high (Table 1.2).

Several albumin-related proteins exist, such as α-fetoprotein, which is the major protein component in foetal plasma [29], and vitamin D binding protein [30]. Another protein member of particular interest, is Endo16, which is a large multi-domain Ca^{2+}-binding protein found in the sea urchin, both in the extracellular matrix and on the surface of endodermal cells during gastrulation [31]. This protein is believed to be the first albumin-related protein to be found in an invertebrate organism.

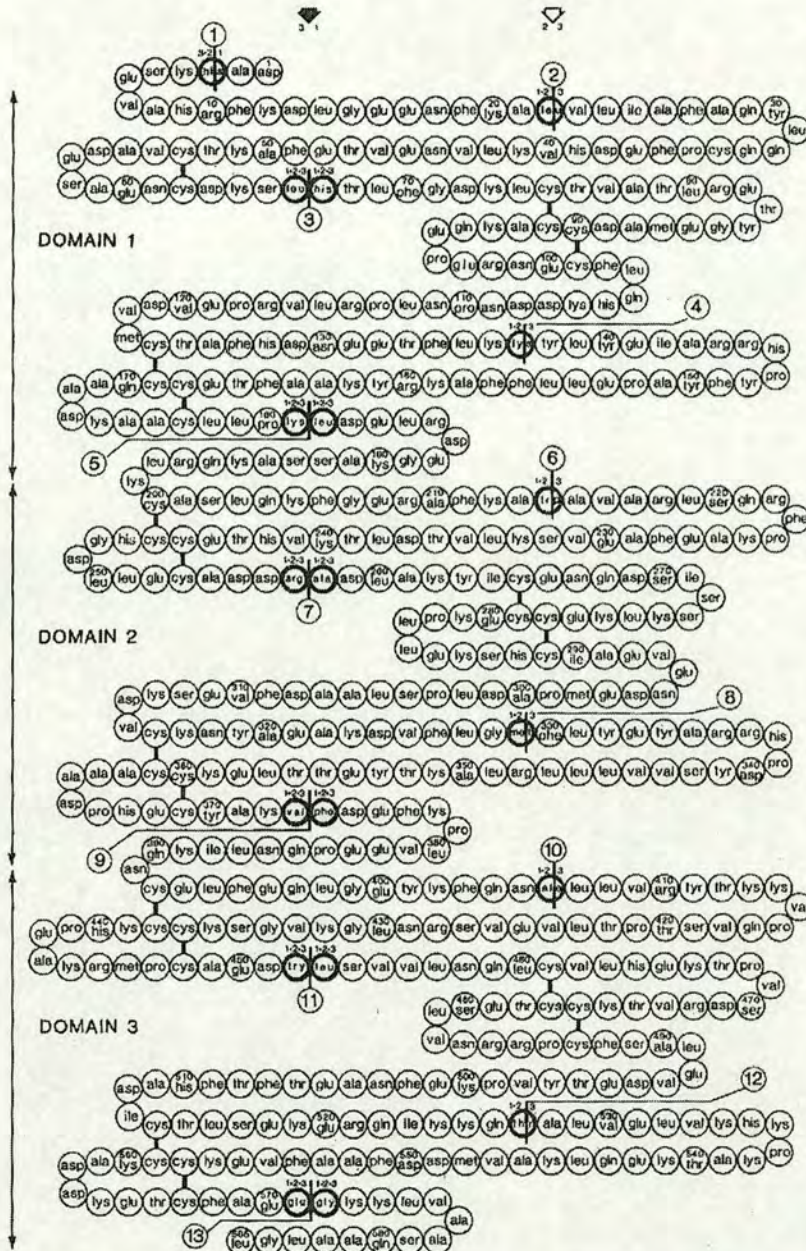**Figure 1.1 The loop-link-loop primary structure and amino acid sequence of human serum albumin.** The loops created by the disulfide bridges (indicated by the black lines between residues) are shown within albumin's three domains. Amino acids are identified by the standard 3-letter codes. The numbered vertical bars indicate the positions of exons in the *alb* gene. This illustration was taken from ref [9].

```
HUMAN     MKWVTFISLLLLFSSAYSRGVFRRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPF 60
MACAQUE   -------LLLYLFSSAYSRGVFRRDTHKSEVAHRFKDLGEEHFKGLVLVAFSQYLQQCPF 52
CANINE    MKWVTFISLLFLFSSAYSRGVVRREAYKSEIAHRYNDLGEEHFRGLVLVAFSQYLQQCPF 60
FELINE    MKWVTFISLLLLFSSAYSRGVTRREAHQSEIAHRFNDLGEEHFKGLVLVAFSQYLQQCPF 60
BOVINE    MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPF 60
SHEEP     MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFNDLGEENFQGLVLIAFSQYLQQCPF 60
PIG       --WVTFISLLFLFSSAYSRGVFRRDTYKSEIAHRFKDLGEQYFKGLVLIAFSQHLQQCPY 58
RABBIT    MKWVTFISLLFLFSSAYSRGVFRREAHKSEIAHRFNDVGEEHFIGLVLITFSQYLQKCPY 60
RAT       MKWVTFLLLLFISGSASRGVFRREAHKSEIAHRFKDLGEQHFKGLVLIAFSQYLQKCPY 60

HUMAN     EDHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEP 120
MACAQUE   EEHVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEP 112
CANINE    EDHVKLAKEVTEFAKACAAESGANCDKSLHTLFGDKLCTVASLRDKYGDMADCCEKQEP 120
FELINE    EDHVKLVNEVTEFAKGCVADQSAANCEKSLHELLGDKLCTVASLRDKYGEMADCCEKKEP 120
BOVINE    DEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEP 120
SHEEP     DEHVKLVKELTEFAKTCVADESHAGCDKSLHTLFGDELCKVATLRETYGDMADCCEKQEP 120
PIG       EEHVKLVREVTEFAKTCVADESAENCDKSIHTLFGDKLCAIPSLREHYGDLADCCEKEEP 118
RABBIT    EEHAKLVKEVTDLAKACVADESAANCDKSLHDIFGDKICALPSLRDTYGDVADCCEKKEP 120
RAT       EEHIKLVQEVTDEAKTCVADENAENCDKSIHTLFGDKLCAIPKLRDNYGELADCCAKQEP 120

HUMAN     ERNECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLF 180
MACAQUE   ERNECFLQHKDDNPNLPPLVRPEVDVMCTAFHDNEATFLKKYLYEVARRHPYFYAPELLF 172
CANINE    DRNECFLAHKDDNPGFPPLVAPEPDALCAAFQDNEQLFLGKYLYEIARRHPYFYAPELLY 180
FELINE    ERNECFLQHKDDNPGFGQLVTPEADAMCTAFHENEQRFLGKYLYEIARRHPYFYAPELLY 180
BOVINE    ERNECFLSHKDDSPDLPKLK-PDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLY 179
SHEEP     ERNECFLNHKDDSPDLPKLK-PEPDTLCAEFKADEKKFWGKYLYEVARRHPYFYAPELLY 179
PIG       ERNECFLQHKNDNPDIPKLK-PDPVALCADFQEDEQKFWGKYLYEIARRHPYFYAPELLY 177
RABBIT    ERNECFLHHKDDKPDLPPFARPEADVLCKAFHDDEKAFFGHYLYEVARRHPYFYAPELLY 180
RAT       ERNECFLQHKDDNPNLPPFEQRPEAEAMCTSFQENPTSFLGHYLHEVARRHPYFYAPELLY 180

HUMAN     FAKRYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAV 240
MACAQUE   FAARYKAAFAECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGDRAFKAWAV 232
CANINE    YAQQYKGVFAECCQAADKAACLGPKIEALREKVLLSSAKERFKCASLQKFGDRAFKAWSV 240
FELINE    YAEEYKGVFTECCEAADKAACLTPKVDALREKVLASSAKERLKCASLQKFGERAFKAWSV 240
BOVINE    YANKYNGVFQECCQAEDKGACLLPKIETMREKVLASSARQRLRCASIQKFGERALKAWSV 239
SHEEP     YANKYNGVFQECCQAEDKGACLLPKIDAMREKVLASSARQRLRCASIQKFGERALKAWSV 239
PIG       YAIIYKDVFSECCQAADKAACLLPKIEHLREKVLTSAAKQRLKCASIQKFGERAFKAWSL 237
RABBIT    YAQKYKAILTECCEAADKGACLTPKLDALEGKSLISAAQERLRCASIQKFGDRAYKAWAL 240
RAT       YAEKYNEVLTQCCTESDKAACLTPKLDAVKEKALVAAVRQRMKCSSMQRFGERAFKAWAV 240

HUMAN     ARLSQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLK 300
MACAQUE   ARLSQKFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYMCENQDSISSKLK 292
CANINE    ARLSQRFPKADFAEISKVVTDLTKVHKECCHGDLLECADDRADLAKYMCENQDSISTKLK 300
FELINE    ARLSQKFPKAEFAEISKLVTDLAKIHKECCHGDLLECADDRADLAKYICENQDSISTKLK 300
BOVINE    ARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDHQDALSSKLK 299
SHEEP     ARLSQKFPKADFTDVTKIVTDLTKVHKECCHGDLLECADDRADLAKYICDHQDALSSKLK 299
PIG       ARLSQRFPKADFTEISKIVTDLAKVHKECCHGDLLECADDRADLAKYMCEHQETISSHLK 297
RABBIT    VRLSQRFPKADFTDISKIVTDLTKVHKECCHGDLLECADDRADLAKYMCEHQETISSHLK 300
RAT       ARMSQRFPNAEFAEITKLATDVTKINKECCHGDLLECADDRAELAKYMCENQATISSKLQ 300

HUMAN     ECCEKPLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYAR 360
MACAQUE   ECCDKPLLEKSHCLAEVENDEMPADLPSLAADYVESKDVCKNYAEAKDVFLGMFLYEYAR 352
CANINE    ECCDKPVLEKSQCLAEVERDELPGDLPSLAADFVEDKEVCKNYQEAKDVFLGTFLYEYSR 360
FELINE    ECCGKPVLEKSHCISEVERDELPADLPPLAVDFVEDKEVCKNYQEAKDVFLGTFLYEYSR 360
BOVINE    ECCDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSR 359
SHEEP     ECCDKPVLEKSHCIAEVDKDAVPENLPPLTADFAEDKEVCKNYQEAKDVFLGSFLYEYSR 359
PIG       ECCDKPLLEKSHCIAEAKRDELPADLNPLEHDFVEDKEVCKNYKEAKDVFLGTFLYEYSR 357
RABBIT    ECCDKPTLEKAHCIYGLHNDEDTAGLPAVAEEEFVEDKDVCKNYEEAKDLFLGKFLYEYSR 360
RAT       ACCDKPVLQKSQCLAETEHDNIPADLPSIAADFVEDKEVCKNYAEAKDVFLGTFLYEYSR 360

HUMAN     RHPDYSVVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFE 420
MACAQUE   RHPDYSVMLLLRLAKAYEATLEKCCAAADPHECYAKVFDEFQPLVEEPQNLVKQNCELFE 412
CANINE    RHPEYSVSLLLRLAKEYEATLEKCCATDDPPTCYAKVLDEEKPLVDEPQNLVKTNCELFE 420
FELINE    RHPEYSVSLLLRLAKEYEATLEKCCATDDPPACYAHVFDEFKPLVEEPHNLVKTNCELFE 420
BOVINE    RHPEYAVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFE 419
SHEEP     RHPEYAVSVLLRLAKEYEATLEDCCAKEDPHACYATVFDKLKHLVDEPQNLIKKNCELFE 419
PIG       RHPDYSVSLLLRIAKIYEATLEDCCAKEDPPACYATVFDKFQPLVDEPKNLIKQNCELFE 417
RABBIT    RHPDYSVVLLLRLGKAYEATLKKCCATDDPHACYAKVIDEFQPLVDEPKNLVKQNCELYE 420
RAT       RHPDYSVSLLLRLAKKYEATLEKCCAEGDPPACYGTVLAEFQPLVEEPKNLVKTNCELYE 420
```

```
HUMAN     QLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVV  480
MACAQUE   QLGEYKFQNALLVRYTKKVPQVSTPTLVEVSRNLGKVGAKCCKLPEAKRMPCAEDYLSVV  472
CANINE    KLGEYGFQNALLVRYTKKAPQVSTPTLVEVSRKLGKVGTKCCKKPESERMSCADDFLSVV  480
FELINE    KLGEYGFQNALLVRYTKKVPQVSTPTLVEVSRSLGKVGSKCCTHPEAERLSCAEDYLSVV  480
BOVINE    KLGEYGFQNALIVRYTRKVPQVSTPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLI  479
SHEEP     KHGEYGFQNALIVRYTRKAPQVSTPTLVEISRSLGKVGTKCCAKPESERMPCTEDYLSLI  479
PIG       KLGEYGFQNALIVRYTKKVPQVSTPTLVEVARKLGLVGSRCCKRPEEERLSCAEDYLSLV  477
RABBIT    QLGDYNFQNALLVRYTKKVPQVSTPTLVEISRSLGKVGSKCCKHPEAERLPCVEDYLSVV  480
RAT       KLGEYGFQNAVLVRYTQKAPQVSTPTLVEAARNLGRVGTKCCTLPEAQRLPCVEDYLSAI  480

HUMAN     LNQLCVLHEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTL  540
MACAQUE   LNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALELDEAYVPKAFNAETFTFHADMCTL  532
CANINE    LNRLCVLHEKTPVSERVTKCCSESLVNRRPCFSGLEVDETYVPKEFNAETFTFHADLCTL  540
FELINE    LNRLCVLHEKTPVSERVTKCCTESLVNRRPCFSALQVDETYVPKEFSAETFTFHADLCTL  540
BOVINE    LNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALTPDETYVPKAFDEKLFTFHADICTL  539
SHEEP     LNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSDLTLDETYVPKPFDEKFFTFHADICTL  539
PIG       LNRLCVLHEKTPVSEKVTKCCTESLVNRRPCFSALTPDETYKPKEFVEGTFTFHADLCTL  537
RABBIT    LNRLCVLHEKTPVSEKVTKCCSESLSNRRPCFSALGPDETYVPKEFNAETFTFHADICTL  540
RAT       LNRLCVLHEKTPVSEKVTKCCSGSLVERRPCFSALTVDETYVPKEFKAETFTFHSDICTL  540

HUMAN     SEKERQIKKQTALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLV  600
MACAQUE   SEKEKQVKKQTALVELVKHKPKATKEQLKGVMDNFAAFVEKCCKADDKEACFAEEGPKFV  592
CANINE    PEAEKQVKKQTALVELLKHKPKATDEQLKTVMGDFGSFVDKCCAAEDKEACFAEEGPKLV  600
FELINE    PEAEKQIKKQSALVELLKHKPKATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLV  600
BOVINE    PDTEKQIKKQTALVELLKHKPKATEEQLKTVMENFVAFVDKCCAADDKEACFAVEGPKLV  599
SHEEP     PDTEKQIKKQTALVELLKHKPKATDEQLKTVMENFVAFVDKCCAADDKEGCFVLEGPKLV  599
PIG       PEDEKQIKKQTALVELLKHKPHATEEQLRTVLGNFAAFVQKCCAAPDHEACFAVEGPKFV  597
RABBIT    PETERKIKKQTALVELVKHKPHATNDQLKTVVGEFTALLDKCCSAEDKEACFAVEGPKLV  600
RAT       PDKEKQIKKQTALAELVKHKPKATEDQLKTVMGDFAQFVDKCCKAADKDNCFATEGPNLV  600

HUMAN     AASQAALGL  609
MACAQUE   AASQAALA-  600
CANINE    AAAQAALV-  608
FELINE    AAAQAALA-  608
BOVINE    VSTQTALA-  607
SHEEP     ASTQAALA-  607
PIG       IEIRGILA-  605
RABBIT    ESSKATLG-  608
RAT       ARSKEALA-  608
```

Table 1.2   Conservation of amino acid residues in mammalian pre-albumins.
Alignments were carried out using ClustalW (EMBL-EBI). Residues in red are part of the pre-albumin sequence and are cleaved prior to secretion. Colour code: yellow = fully conserved residues, green = residues with conserved properties and cyan = semi-conserved residues (in accordance to ClustalW rules). The numbering at the right-hand side includes the pre-albumin sequence and so does not represent the amino acid numbering of the native protein. Amino acid sequence accession numbers: Human, P02768; Macaque, A47391; Canine, S29749; Feline, S57632; Bovine, P02769; Sheep, P14639; Pig, P08835; Rabbit, P49065 and Rat, P02770.

The x-ray crystal structure of human albumin has been determined, in both free [32,33] and fatty acid-bound forms [34]. A summary of all albumin crystal structures solved to date is shown in Table 1.3. The secondary structure of human albumin is predominantly α-helical. The molecule itself is heart-shaped and consists of three repeating domains, constituted of two sub-domains. Figure 1.2 shows a ribbon representation of the albumin chain.

It is now widely accepted that the three albumin domains evolved from a single ancestral protein of about 190 amino acids, through a triplication event. Recently, the three albumin domains have been individually cloned, expressed in yeast and purified [35]. Each domain was then studied separately and the binding properties of several pharmaceutically important ligands, towards each were determined. The structural effects within each individual domain in acidic and basic environments [36] are discussed in Section 1.2.2.

Albumin is a flexible protein with the ability of changing shape rapidly in solution. This is thought to be due to the loop-link-loop nature of the disulfide arrangement within the molecule. Albumin has been described as having a "breathing" action [37] constantly opening and closing, revealing nooks to which ligands may bind.

PDB code	Resolution (Å)	Ligands bound	Reference
1AO6	2.50	None	[33]
1BJ5	2.50	5 myristates	[34]
1BKE	3.15	5 myristates, 2 triiodobenzoates	[34]
1BM0	2.50	None	[33]
1E78	2.60	None	[38]
1E7A	2.20	2 propofols	[38]
1E7B	2.38	3 halothanes	[38]
1E7C	2.40	5 myristates, 7 halothanes	[38]
1E7E	2.50	10 caprates	[39]
1E7F	2.43	8 laurates	[39]
1E7G	2.50	8 myristates	[39]
1E7H	2.43	7 palmitates	[39]
1E7I	2.70	7 stearates	[39]
1GNI	2.40	7 oleates	[40]
1GNJ	2.60	8 arachidonates	[40]
1H9Z	2.50	6 myristates, 1 R-(+) warfarin	[41]
1HA2	2.50	1 S-(-) warfarin	[41]
1UOR	2.80	None	[32]

Table 1.3   Summary of human albumin crystal structures available in the Brookhaven Protein Databank (October 2002).
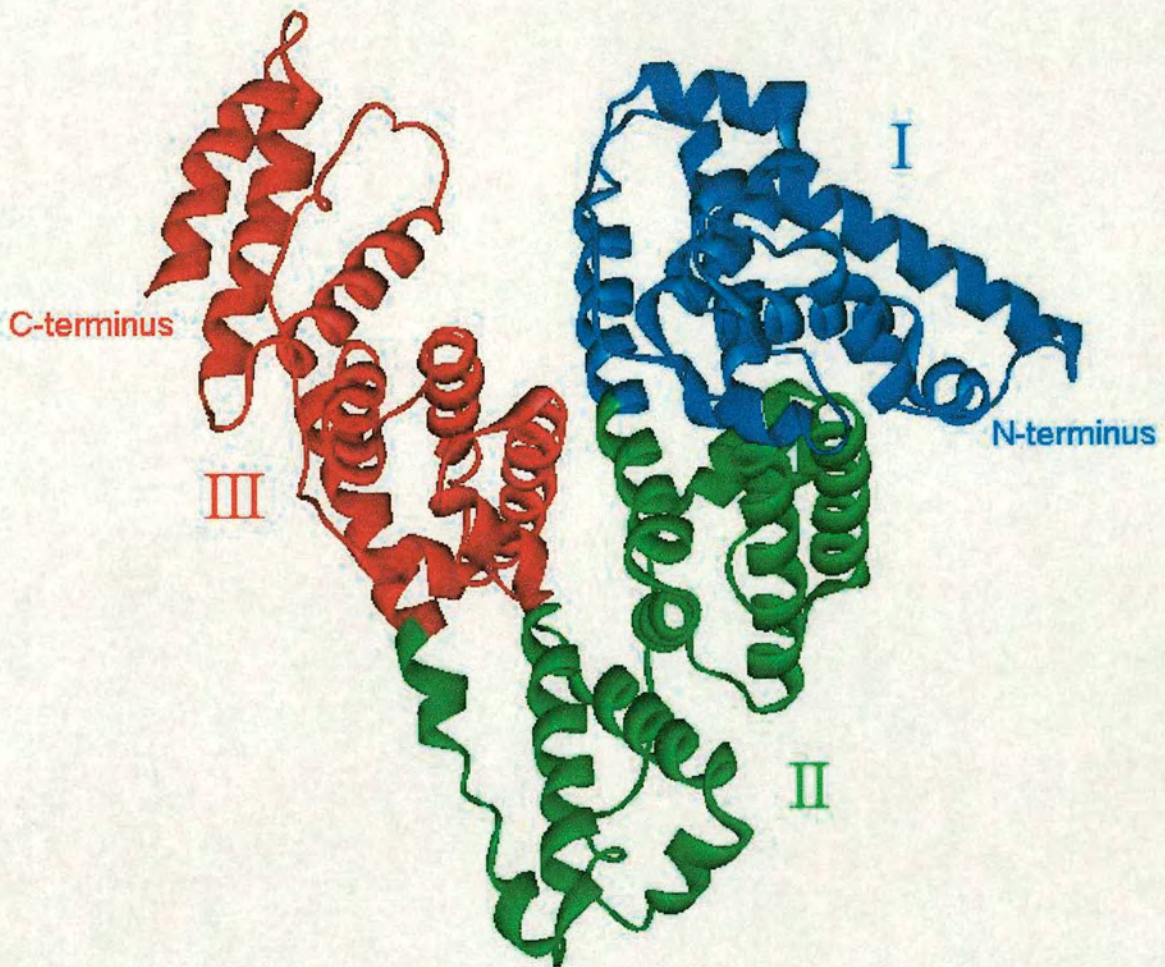
Figure 1.2 Tertiary structure of recombinant human albumin. Three-dimensional structure of human albumin coloured by domain (I-III). Note the predominantly α-helical structure. PDB code is 1AO6 [33].

1.2.2 pH-Dependent Structural Transitions

The 3-dimensional structure of albumin rapidly alters under non-physiological conditions, particularly with regard to changes in pH. Five pH-dependent isomeric forms have been recognised. Albumin's structure under physiological conditions is defined as being in the **N**, or normal form. At a pH of around 4, the albumin molecule denatures to what is known as the **F**, or fast form. Fluorescence studies of both Trp[214] [42] and a single bound 1-anilino-8-naphthalenesulfonate molecule [43] led to the conclusion that the **N** to **F** transition occurs in two steps. The normal protein first alters to an intermediate **F1** form, and then spontaneously expands to the **F** form. This change occurs at pH 4.3 and requires only about 100 ms to take place [44]. If the pH of an albumin solution is reduced further to below 3.5, it undergoes another change to become the **E** (expanded) form. At pH 2.5, the protein is believed to be as expanded as the molecular structure allows. The **E** form is characterised by an increase in intrinsic viscosity [45] and approximately three of the disulfide bonds can be reduced with 2-mercaptoethylamine [46].

Dockal and colleagues have isolated the three separate human albumin domains and investigated the effect of pH on each [36]. They found that the **N** to **F** transition occurs primarily in domain III with only minor changes in domain I. Domain II transforms into a molten globule-like state as the pH is reduced further, corresponding to the **E** form [36].

Denaturation of the **N** form of albumin also occurs when the pH is raised. Above pH 8 a "structural fluctuation" can be observed; the molecule loosens and loses rigidity. This structural isomer is termed the **B**, or basic form. Circular dichroism indicates that the **N** to **B** transition involves a relatively minor change in conformation, with only a 10% loss

in helix and an 8% increase in β-sheet [47]. Dockal found that during this transition, albumin domains I and II underwent minor structural isomerisation, whilst in domain III no structural changes were found [36].

Increasing the pH further (>10) results in the fifth structural isoform, termed the **A** or aged form. The name comes from a phenomenon first noticed by Sogami. He found that some defatted BSA that had been left in a salt-free solution at pH 9 for 3-4 days had an altered isoelectric point [48], thought to be the result of a sulfhydryl-catalyzed disulfide interchange reaction.

1.3 Ligand Binding to Serum Albumin

1.3.1 Investigations of Ligand Binding

It was once thought that albumin-ligand interactions are non-specific, with the protein behaving like a sponge and the ligands binding to the molecule's surface in a random manner. Binding studies are a prolific area of albumin research. Such studies have lead to the discovery of several distinct, high affinity ligand-binding sites.

Much of the early binding studies of albumin made use of equilibrium dialysis and spectroscopic methods. Other studies have assayed for ligand binding using radio- and fluorescent-labelling, through both direct and competitive binding studies. These types of analyses make it possible to measure binding constants, but do not give any information as to where on the albumin molecule the ligand is bound. The positions of binding sites in the albumin molecule have been examined by techniques such as site-directed mutagenesis and through peptide sequencing, following protein digestion in the

presence of ligand [49]. Crystallographic studies and modern spectroscopic methods have now revolutionised this area of albumin research. Crystal structures have now been reported for albumin bound to fatty acids [39,40], the aspirin analogue, 2,3,5-triiodobenzoate [34], warfarin [41], and the anaesthetics propofol and halothane (Brookhaven Protein Databank). NMR spectroscopy, HPLC and mass spectrometry have also proven to be invaluable methods for binding studies [50,51].

1.3.2 Fatty Acid Binding

The properties of fatty acid binding are probably the best understood of all known ligands. This is due in part to several crystal structures of human albumin determined by Curry and co-workers [34,39,40]. Fatty acids are an extremely important source of physiological energy in the body. They are released from adipocytes and lipoproteins and are transported to their target site, through the circulatory system bound to albumin. Most fatty acid in the blood is albumin bound; elevated localised free fatty acid levels can be associated with various diseases such as cancer and diabetes [52] and are a symptom of analbuminaemia [53].

In Curry *et al.*'s first fatty acid-albumin structure (Figure 1.3) [34], albumin is complexed with five myristate molecules, revealing five fatty acid binding sites arranged on the albumin molecule in an asymmetric manner. The first myristate (Myr 1) molecule binds to sub-domain IB via a salt bridge at Arg^{117}. Myr 2 binds at the interface between sub-domains IA, IB and IIA. Its methylene tail is predominantly held in place by sub-domain IA, while residues Arg257 and Ser287 on sub-domain IIA, and Tyr150 on IB, anchor its carboxylate end. Both Myr 3 and Myr 4 are located within sub-domain IIIA. The carboxylate group of Myr 3 is bound across the interface between sub-domains IIB and

IIIA to residues Ser[342], Arg[348] (from IIB) and Arg[485] (from IIIA). Myr 4 is also appears to be held in place at its carboxylate end by 3 hydrogen bond partners. These are Arg[410], Tyr[411] and Ser[485], which are all situated on sub-domain IIIA. The final myristate in the structure (Myr 5) is bound to sub-domain IIIB in a similar orientation as Myr 1 is to sub-domain IB, and is anchored at its carboxylate end by Lys[575].

Curry and colleagues also determined the crystal structures of albumin with long- and medium-chain fatty acids [39], as well as mono- and poly-unsaturated fatty acids [48]. These structures revealed a variety of other binding sites including sites I and II, which are discussed in section 1.3.3. Site II appears to be a high-affinity long-chain fatty acid binding site.

From comparisons of the crystal structures of fatty acid-bound albumin and fatty acid free albumin [33], it is clear that some domain rotations take place. Albumin, upon binding of the fatty acid molecules, appears to "open up". Domain II appears to facilitate this conformational change, acting as a spring to allow domains I and III to separate slightly (Figure 1.3). These structures show how albumin distrorts to accommodate ligands and provides the best indications to date as to how the fatty acids interact with the amino acid residues in the albumin chain. It is likely that because fatty acids invoke such a change in structure that binding of fatty acid to albumin may affect the binding of molecules at other sites. Structural information of this kind is invaluable in the design and development of novel therapeutics.
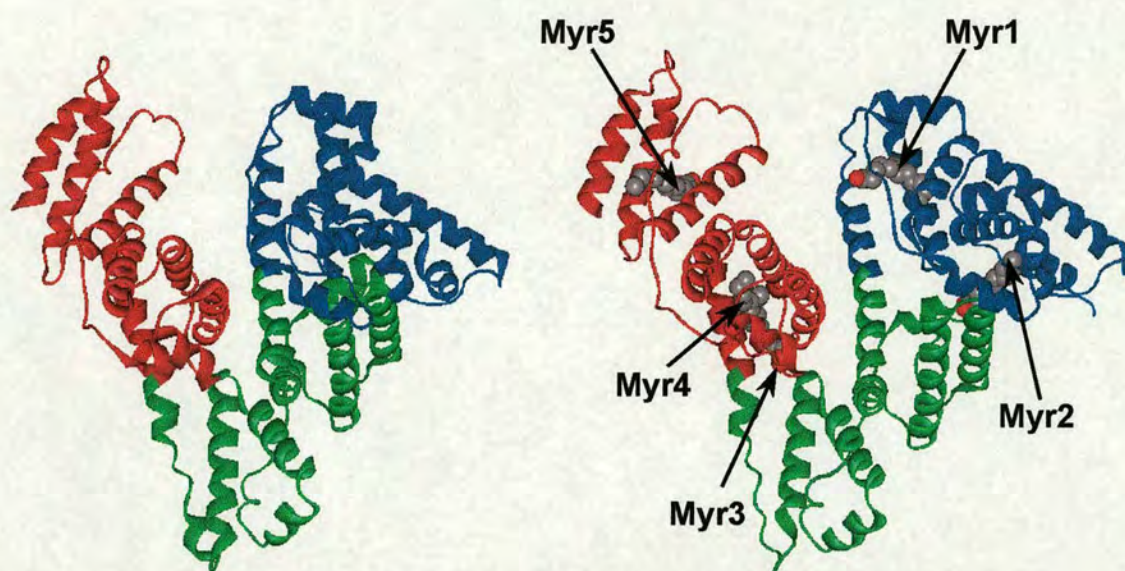
Figure 1.3 Comparison between unliganded (left) and myristic acid bound (right) human albumin. Domains I, II and III are coloured blue, green and red respectively; fatty acid myristate molecules are labelled 1-5. Both structures were drawn from the coordinates in the Brookhaven Protein Databank. The PDB code for unliganded albumin is 1AO6 [33] and for fatty acid bound albumin the code is 1BJ5 [34].

1.3.3 Ligand Binding Sites I and II

Two important ligand-binding sites referred to as sites I and II, were first identified by Sudlow and colleagues, through studies of competitive binding [54]. A large body of evidence indicates that these specialised regions are located within specialised cavities of subdomains IIA and IIIA respectively [9]. These sites are often thought of as being the main binding sites for small organic compounds. Each of these sites boasts the ability of binding a wide range of ligands. However site I appears to be more selective than site II. Examples of ligands, which bind to site I are warfarin, salicylate and the haemoglobin metabolite, bilirubin [55]. Others ligands that show specificity include haemin [56], dicarboxylic fatty acids [57], and a variety of dyes and indicator compounds [9]. It is difficult to define a typical site I ligand due to the diverse nature of the compounds that

show affinity, although most appear to be bulky heterocyclic anions, with a charge distribution centred within the molecule.

The 2.8 Å crystal structure of human albumin [32] revealed that the amino acids which line the specialised cavities of sites I and II have an asymmetric charge distribution. This causes them to have a hydrophobic surface on one side and a hydrophilic surface on the other. The cavities have been described as "sock-shaped" [1] with the "leg" region of the sock being the hydrophilic side and the more open "foot" region, the hydrophobic side. This explains why both sites are selective for the binding of anions. The crystal structures of human albumin with the R-(+) and S-(-) enantiomers of the anticoagulant warfarin bound have been elucidated [41]. Figure 1.4 shows the structure of warfarin.

The crystal structure confirms that the drug binds to site I, with both enantiomers adopting similar conformations, making many of the same specific contacts with the same amino acid side-chains. The packing of all six helices of subdomain IIA form the binding pocket. The pocket has two sub-chambers that accommodate different portions of the warfarin molecule. The benzyl group binds in a sub-chamber made by Phe^{211}, Trp^{214}, Leu^{219}, and Leu^{238}, with additional contacts from Arg^{218} and His^{242}. The coumarin moiety binds in the larger sub-chamber furthest from the entrance. Leu^{219}, Arg^{222}, Phe^{223}, Leu^{234}, Ile^{264}, Leu^{257} and Ile^{290} make up this compartment, with only the Ile^{260}, Ile^{264} and Ile^{290} residues making contact with the coumarin ring.
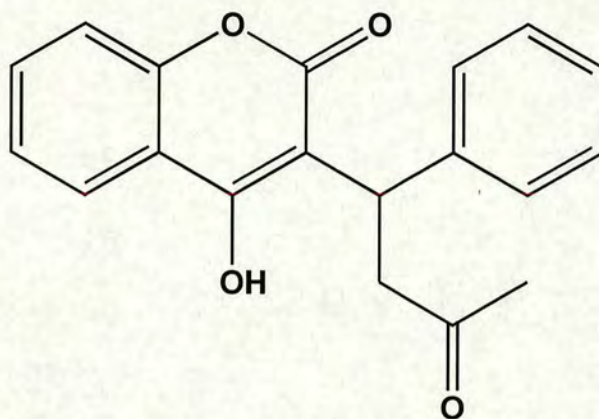
Figure 1.4   Structure of the anticoagulant, warfarin (3-(α-acetonylbenzyl)-4-hydroxycoumarin).

Site II is frequented by a much wider range drugs and small molecules than site I. In fact several molecules such as aspirin and triiodobenzoate bind to both sites [1]. In addition site II also binds fatty acids, tryptophan [58] and anions such as chloride, iodide, bromide and thiocyanate. Interestingly, site II as well as being a major ligand-binding site also shows esterase activity [59]. A variety of *p*-nitrophenyl esters has been cleaved at this site [60]. The esterase activity of serum albumin is greater in humans than in any other organism, suggesting that it has not evolved to carry out an important physiological function, but exists through a chance arrangement of the amino acids in this region of the human protein. The activity has been localised to Tyr[411], the same residue implicated in cyanide detoxification [61].

1.3.4 The Cys[34] Site

This site is arguably the most intriguing of all albumin's binding sites. Not only is it the binding site for a wide variety of biologically and clinically important small molecules, but it also provides antioxidant [13] and antiperoxidant [62] activities and is the site of

dimerisation. Cys^{34} is the only cysteine residue in the protein that is not involved in disulfide bridging. It is characterised by an extremely low pK_a of around 7 compared to a value of 8.5 for free cysteine [63]. Consequently, the thiol group of Cys^{34} is predominantly in the S^- form at physiological pH. Albumin is responsible for the largest fraction of free sulfhydryl in blood serum [1].

Much of the interest directed toward this site is due to potentially strong interactions with "soft" metal ions such as those in the anticancer drug, cisplatin, the antiarthritic drug, auranofin and mercurials. Other ligands, which show affinity toward this site, include heterocyclic, aromatic and small thiol-containing compounds (redox reactions to form disulfides). Recent crystallographic studies of albumin show that Cys^{34} is situated within a hydrophobic, sterically hindered pocket, approximately 9.5 Å deep, as shown in Figure 1.5. This agrees well with the previous observations, that Cys^{34} binds to heterocyclic compounds with greater affinity than to aromatic compounds [64], and that compounds with five-membered rings are more reactive than six-membered rings [65].

So far, without exception, all mammalian and avian albumins sequenced to date contain a free cysteine residue. However, both salmon and cobra albumins differ at this and near-by residues [66,67]. In salmon albumin, Cys^{34} is replaced by serine, and in cobra albumin, it is replaced by alanine. Additionally in both these species, His^{39} (located close-by in the native protein, as can be seen in Figure 1.5) is substituted by leucine. Consequently albumin from these species should show a much-decreased affinity toward the aforementioned ligands.
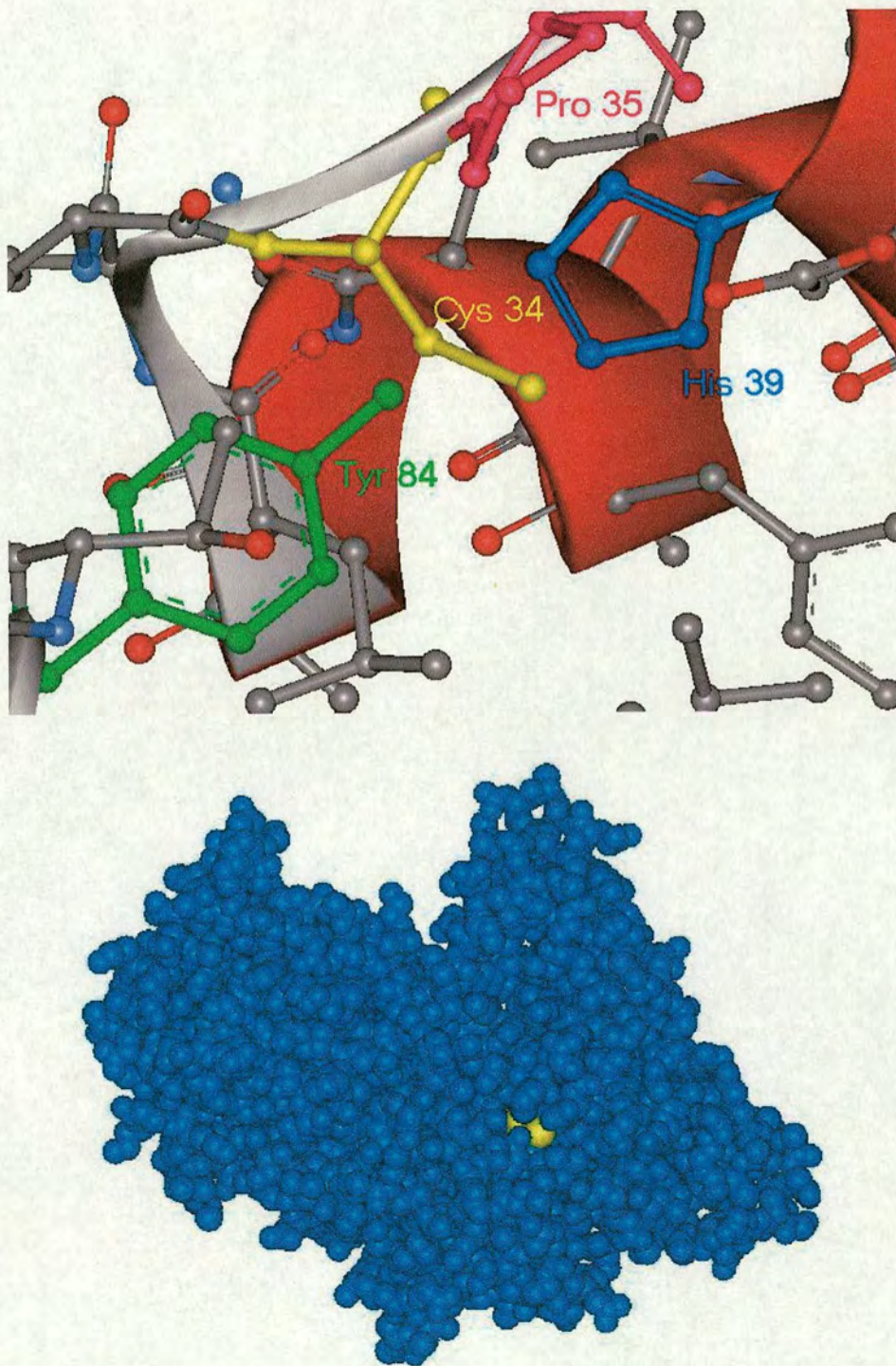
Figure 1.5 Environment of the Cys^{34} residue in human albumin. Above: Tyr^{84} and His^{39} residues are located near Cys^{34}. Below: The space-filling model of albumin (below) shows the Cys^{34} residue (yellow) to be buried within a crevice in the molecule.

Many preparations contain a significant proportion of albumin in dimeric form; it still remains unclear as to whether this form exists in blood plasma [9]. Dimerisation is thought to take place at Cys^{34}; blocking of the thiol group with agents such as iodoacetamide, cysteine or glutathione prevent albumin dimers from forming [8]. Alternatively, dimerisation is said to be induced by linking the single thiols of two albumin molecules with Hg^{2+} [68], but the dimerisation mechanism still remains poorly understood.

^{1}H NMR studies have revealed that upon ligand binding to the Cys^{34} thiol (or oxidation of the thiol to a sulfenic acid) a conformational change is invoked, where the residue moves from a buried to a solvent-exposed state [69,70]. This alteration in structure is coupled to a movement of His^3. The cysteine in the unbound reduced form may be stabilised in the buried state by the nearby His^{39}. A schematic representation of this conformational change upon binding of the antiarthritic drug, auranofin is shown in Figure 1.6.

Interestingly, a naturally occurring genetic variant of albumin has been identified which has no free thiol at Cys^{34} [71]. This variant discovered in Asola, Italy, contained a point mutation, which led to the introduction of a cysteine residue in place of Tyr^{140}. This new cysteine residue was found to form a new disulfide bridge with the thiol group of Cys^{34}, resulting in 18 disulfide bridges and consequently, no free thiol.
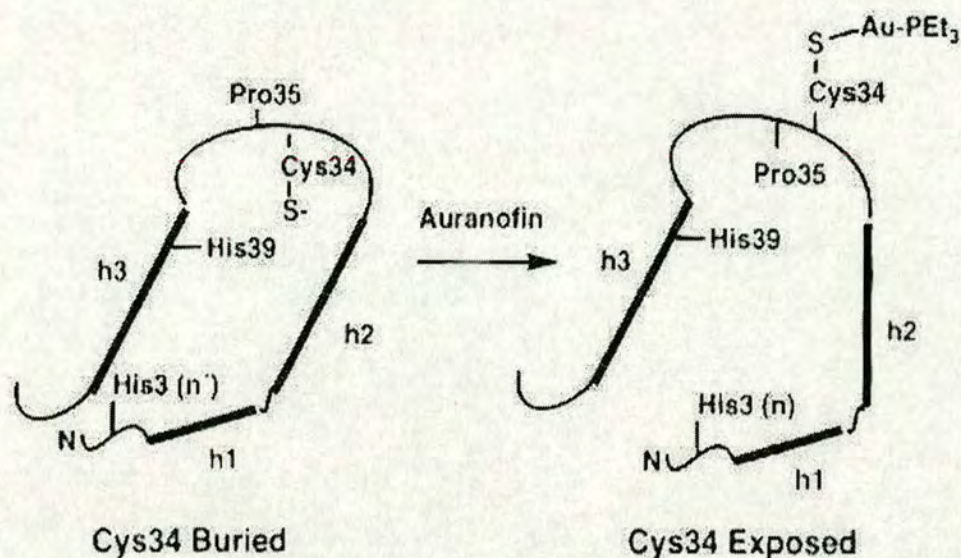
Figure 1.6 Change in the Cys[34] environment upon ligand binding. It is proposed that upon ligand binding or oxidation, the Cys[34] residue moves from a buried to a solvent-exposed state. In this example the ligand is the antiarthritic drug, auranofin. The conformational change appears to be coupled to a movement of the His[3] residue. Diagram taken from ref [72].

1.3.5 The N-terminal Binding Site

The N-terminus of human albumin is a strong binding site for Cu^{2+} and Ni^{2+} with association constants of 10^{16} M^{-1} [73] and 3.98×10^9 M^{-1} [74], respectively. Cu^{2+} was first known to bind at this site when it was discovered that the presence of $CuCl_2$ blocked the aspartyl α-amino group of albumin from reacting with the Sanger reagent, fluorodinitrobenzene (FDNB) [75]. Following this, peptide binding studies showed that Cu^{2+} binds to the BSA peptide fragment of residues 1-24 (and even to the BSA peptide 1-4, Asp-Thr-Ala-Lys) as strongly as to BSA itself [76].

37

The binding of Cu^{2+} to albumin has been investigated for several species and has been found to bind strongly only to those, which contain His[3] [9]. ^{1}H NMR studies have also implicated a lysine residue (presumably Lys[4]) as being a key residue in the arrangement of the Cu^{2+} and Ni^{2+} binding site [77]. Binding of Cu^{2+} and Ni^{2+} to this site is currently believed to involve a four-nitrogen coordination of the metal with the α-amino nitrogen of Asp[1], the two backbone amide nitrogens of Ala[2] and His[3] and an imidazole nitrogen of His[3]. The proposed structure of this site is shown in Figure 1.7.
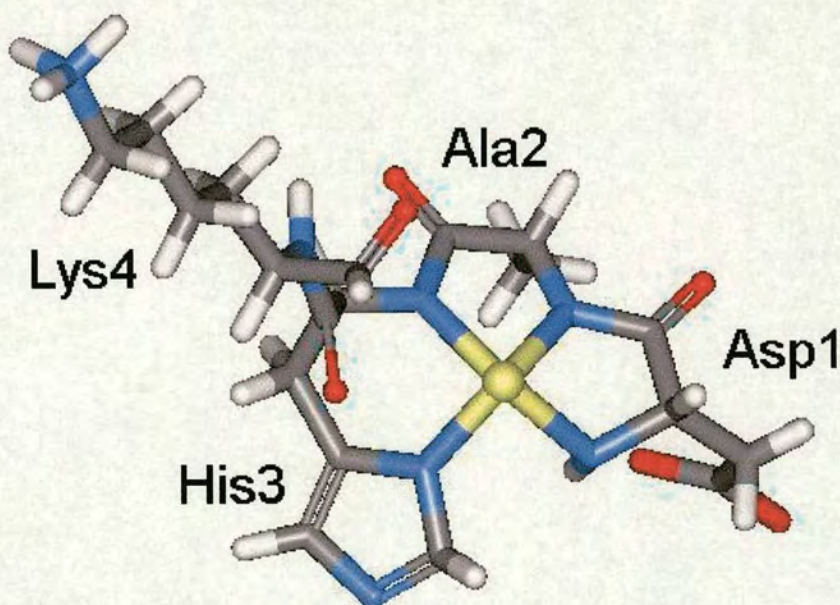
Figure 1.7 N-terminal binding site. A view of the proposed structure of the N-terminal Cu^{2+} and Ni^{2+} binding site of human albumin. The metal ion is coloured yellow and residues Asp[1], Ala[2], His[3] and Lys[4] are shown.

1.4 Aims of this Thesis

Most studies of the chemistry of human albumin have been carried out on albumin isolated from blood plasma. Such albumin is often highly heterogeneous, in particular a large proportion of Cys^{34} is often oxidised. The recent availability of recombinant albumin has provided an opportunity to study the chemistry of a more defined form of albumin, especially with regard the free thiol at Cys^{34}.

The aims of this work were as follows.

(1) To investigate the ligand binding properties of Cys^{34} with a range of different physiologically important molecules.

(2) To examine the roles neighbouring residues play in controlling the reactivity of Cys^{34}, using site-directed mutagenesis and recombinant methods.

(3) To investigate the metal binding properties of albumin and identify new metal-derived sites.

(4) To attempt to crystallise albumin in the hope of producing crystals of metal-derivatives that diffract x-rays at high-resolution for structural determination.

References: Chapter 1

1.   Carter, D. C. and Ho, J. X. (1994). Structure of serum albumin. *Adv. Prot. Chem.*, **45**: 153-203.

2.   Figge, J., Rossing, T. H. and Fencl, V. (1991). The role of serum proteins in acid-base equilibria. *J. Lab. Clin. Med.*, **117**: 453-467.

3.   Ivanov, A. I., Christodoulou, J., Parkinson, J. A., Barnham, K. J., Tucker, A., Woodrow, J. and Sadler, P. J. (1998). Cisplatin binding sites on human albumin. *J. Biol. Chem.*, **24**: 14721-14730.

4.   Shaw III, C. F. (1989). The protein chemistry of antiarthritic gold (I) thiolates and related complexes. *Comments Inorg. Chem.*, **8**: 233-267.

5.   Fehske, K. J., Schläfer, U., Wollert, U. and Müller, W. E. (1982). Characterisation of an important drug binding area on human serum albumin, including the high-affinity binding sites of warfarin and azapropazone. *Mol. Pharmacol.*, **21**: 387-393.

6.   Wilting, J., Hart, B. J. and De Gier, J. J. (1980). The role of albumin conformation in the binding of diazepam to human serum albumin. *Biochim. Biophys. Acta*, **626**: 291-298.

7.   Kragh-Hansen, U. (1981). Molecular aspects of ligand binding to serum albumin. *Pharmacol. Rev.*, **33**: 17-53.

8.   Peters, T., Jr. (1996). *All about albumin: Biochemistry, genetics, and medical applications.* Academic Press, New York.

9.   Peters, T., Jr. (1985). Serum albumin. *Adv. Prot. Chem.*, **37**: 161-245.

10. Russi, E. and Weigand, K. (1983). Analbuminemia. *Klin. Wochenschr.*, **61**: 541-545.

11.   Galliano, M., Kragh-Hansen, U., Tarnoky, A. L., Chapman, J.C., Campagnoli, M. and Minchiotti, L.(1999). Genetic variants showing apparent hot-spots in the human serum albumin gene. *Clin. Chim. Acta*, **289**: 45-55.

12.   Petersen, C. E., Ha, C.-E., Harohalli, K., Park, D. S., Feix, J. B., Isozaki, O. and Bhagavan, N. V. (1999).   Structural investigations of a new familial dysalbuminemic hypothyroxinemia genotype. *Clin. Chem.*, **45**: 1248-1254.

13.   Halliwell, B. (1988).   Albumin – an important extracellular antioxidant? *Biochem. Pharmacol.*, **37**: 569-571.

14.   Reed, R. G. (1977).   Kinetics of bilirubin binding to bovine serum albumin and the effects of palmitate. *J. Biol. Chem.*, **252**: 7483-7487.

15.   Jarabak, R. and Westley, J. (1986).   Serum albumin and cyanide detoxification. *J. Biol. Chem.*, **23**: 10793-10796.

16.   Zoellner, H., Höfler, M., Beckmann, R., Hufnag, P., Vanyek, E., Bielek, E., Wojta, J., Fabry, A., Lockie, S. and Binder, B. R. (1996).   Serum albumin is a specific inhibitor of apoptosis in human endothelial cells. *J. Cell Sci.*, **109**: 2571-2580.

17.   Kakizoe, T. and Sugimura, T. (1988).   Chemical carcinogenesis in analbuminemic rats. *Jpn. J. Cancer Res.*, **79**: 775-784.

18.   Kuipers, M. E., v. d. Berg, M., Swart, P. J., Laman, J. D., Meijer, D. K. F., Koppelman, M. H. G. M. and Huisman, H. (1999).   Mechanism of anti-HIV activity of succinylated human serum albumin. *Biochem. Pharmacol.*, **57**: 889-898.

19.   Peters, T., Jr. (1975). Serum albumin. *The Plasma Proteins 2nd Ed.*, Vol. 1, 133-181, Academic Press, New York.

20.   Wang, J., Chen, P., Su, Z.-F., Vallis, K., Sandhu, J., Cameron, R., Hendler, A. and Reilly, R. (2001). Amplified delivery of indium-111 to EGFR-positive human breast cancer cells. *Nuc. Med. Biol.*, **28:** 895-902.

21.   Yeh, P., Landais, D., Lemaître, M., Maury, M., Crenne, J.-Y., Becquart, J., Murry-Brelier, A., Boucher, F., Montay, G., Fleer, R., Hirel, P.-H., Mayaux, J.-F. and Klatzmann, D. (1992). Design of yeast-secreted albumin derivatives for human therapy: Biological and antiviral properties of a serum albumin-CD4 genetic conjugate. *Proc. Natl. Acad. Sci. USA*, **89:** 1904-1908.

22.   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:** 3243-3252.

23.   Tayot, J. L., Tardy, M., Gattel, P., Cueille, G. and Liautaud, J. (1987). Large scale use of spherosil ion exchangers in plasma fractionation. *Dev. Biol. Stand.*, **67:** 15-24.

24.   Aventis Behring (2000). Aventis Behring announces successful results for first recombinant albumin phase I clinical trial. *Press Release*, 4th October.

25.   Cochrane Injuries Group Albumin Reviewers (1998). Human albumin administration in critically ill patients: systematic review of randomised controlled trials. *BMJ*, **317:** 235-240.

26.   Yameh, G. (2000). Albumin industry launches global promotion. *BMJ*, **320:** 533.

27.   Garlick, R. L. and Mazer, J. S. (1983). The principal site of nonenzymatic glycosylation in human serum albumin *in* vivo. *J. Biol. Chem.*, **258:** 6124-6146.

28.   Williams, S. K., Devenny, J. J. and Bitensky, M. W. (1981). Micropinocytic ingestion of glycosylated albumin by isolated microvessels: possible role in

pathogenesis of diabetic microangiopathy. *Proc. Natl. Acad. Sci. USA*, **78:** 2393-2397.

29.   Gitlin, D. and Gitlin, J. D. (1975). Fetal and neonatal development of human plasma proteins. *The Plasma Proteins 2nd Ed.*, Vol. 2, 263-374, Academic Press, New York.

30.   Cooke, N. E. and David, E. V. (1985). Serum vitamin D-binding protein is a third member of the albumin and α-fetoprotein gene family. *J. Clin. Invest.*, **76:** 2420-2424.

31.   Soltysik-Española, M., Klinzing, D. C., Pfarr, K., Burke, R. D. and Ernst, S. G. (1994). Endo16, a large multidomain protein found on the surface and ECM of endodermal cells during sea urchin gastrulation, binds calcium. *Dev. Biol.*, **1165:** 73-85.

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

33.   Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999). Crystal structure of human serum albumin at 2.5 Å resolution. *Prot. Eng.*, **12:** 439-446.

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

35.   Dockal, M., Carter, D. C. and Rücker, F. (1999). The three recombinant domains of human serum albumin, structural characterization and ligand binding properties. *J. Biol. Chem.*, **274:** 29303-29310.

36.   Dockal, M., Carter, D. C. and Rücker, F. (2000). Conformational transitions of the three recombinant domains of human serum albumin depending on pH. *J. Biol. Chem.*, **275:** 3042-3050.

37.   Hvidt, A. and Wallevik, K. (1972). Conformational changes in human serum albumin as revealed by hydrogen-deuterium exchange studies. *J. Biol. Chem.*, **247:** 1530-1535.

38.   Bhattacharya, A. A., Curry, S. and Franks, N. P. (2000). Binding of the general anesthetics propofol and halothane to human serum albumin. High resolution crystal structures. *J. Biol.Chem.* **275:** 38731-38738.

39.   Bhattacharya, A. A., Grune, T. and 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.

40.   Petitpas, I., Grune, T., Bhattacharya, A. A. and Curry, S. (2001). Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *J. Mol. Biol.*, **314:** 955-960.

41.   Petitpas, I., Bhattacharya, A. A., Twine, S., East, M. and Curry, S. (2001). Crystal structure analysis of warfarin binding to human serum albumin. Anatomy of drug site I. *J. Biol. Chem.*, **276:** 22804-22809.

42.   Sogami, M., Era, S., Nagaoka, S. and Inouye, H. (1982). Circular dichroic and fluoropolarimetric studies on tryptophyl residues in acid-induced isomerization of bovine plasma albumin. *Int. J. Pept. Protein Res.*, **19:** 263-269.

43.   Era, S., Kuwata, K., Kida, K., Sogami, M. and Yoshida, A. (1985). Circular dichroic and fluorometric studies on the acid-induced isomerization of bovine plasma albumin-1-anilino-8-naphthalenesulfonate complex. *Int. J. Pept. Protein Res.*, **26:** 575-583.

44.   Taylor, R. P., Chau, V., Zenkowich, M. J. and Leake, L. H. (1978). Stopped-flow studies on the N-F transition in serum albumin. *Biophys. Chem.*, **7:** 293-299.

45.   Harrington, W. F., Johnson, P. and Ottewill, R. H. (1956). Bovine serum albumin and its behaviour in acid solution. *J. Am. Chem. Soc.*, **62:** 569-582.

46.   Alexander, P. and Hamilton, L. D. G. (1968). Changes in the reactivity of disulfide bonds in serum albumin on denaturation. *Arch. Biochem. Biophys.*, **88:** 128-135.

47.   Wilting, J., Weideman, M. M., Roomer, A. C. and Perrin, J. H. (1979). Conformational changes in human serum albumin around the neutral pH from circular dichroic measurements. *Biochim. Biophys. Acta*, **579:** 469-473.

48.   Sogami, M., Petersen, H. A. and Foster, J. F. (1969). The microheterogeneity of plasma albumins. V. Permutations in disulfide pairings as a probable source of microheterogeneity in bovine serum albumin. *Biochemistry*, **8:** 49-58.

49.   Sanger, F. (1963). Amino-acid sequences in the active centres of certain enzymes (Pedler Lecture). *Proc. Chem. Soc.*, **March:** 76-83.

50.   Sadler, P. J. and Viles, J. H. (1996). ^{1}H and ^{113}Cd NMR investigations of Cd^{2+} and Zn^{2+} binding sites on serum albumin: competition with Ca^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}. *Inorg. Chem.*, **35:** 4490-4496.

51.   Gu, C., Nikolic, D., Lai, J., Xu, X., van Breemen, R. B. (1999). Assays of ligand-human serum albumin binding using pulsed ultrafiltration and liquid chromatography-mass spectrometry. *Comb. Chem. High Throughput Screen.*, **2:** 353-359.

52.   Richieri, G. V., Anel, A. and Kleinfeld, A. M. (1993). Interactions of long-chain fatty acids and albumin: Determination of free fatty acid levels using the fluorescent probe ADIFAB. *Biochemistry*, **32:** 7574-7580.

53.   Bartter, F. C., Steinfeld, J. L., Waldmann, T. and Delea, C. S. (1961). Metabolism of infused serum albumin in the hypoproteinemia of gastrointestinal protein loss and in analbuminemia. *Trans. Assoc. Am. Physicians*, **74:** 180-194.

54.   Sudlow, G., Birkett, D. J. and Wade, D. N. (1975). The characterization of two specific drug binding sites on human serum albumin. *Mol. Pharmacol.*, **11:** 824-832.

55.   Kragh-Hansen, U. (1988). Evidence for a large and flexible region of human serum albumin possessing high affinity binding sites for salicylate, warfarin, and other ligands. *Mol. Pharmacol.*, **34:** 160-171.

56.   Adams, P. A. and Berman, M. C. (1980). Kinetics and mechanism of the interaction between human serum albumin and monomeric haemin. *Biochem. J.*, **191:** 95-102.

57.   Tonsgard, J. H. and Meredith, S. C. (1991). Characterization of the binding sites for dicarboxylic acids on bovine serum albumin. *Biochem. J.*, **276:** 569-575.

58.   Kragh-Hansen, U. (1991). Octanoate binding to the indole- and benzodiazepine-binding region of human serum albumin. *Biochem. J.*, **273:** 641-644.

59.   Koh, S. W. and Means, G. E. (1979). Characterization of a small apolar anion binding site of human serum albumin. *Arch. Biochem. Biophys.*, **192:** 459-464.

60.   Kurono, Y., Kushida, I., Tanaka, H. and Ikeda, K. (1992). Esterase-like activity of human serum albumin. VIII. Reaction with amino acid *p*-nitrophenyl esters. *Chem. Pharm. Bull.*, **40:** 2169-2172.

61.   Jarabak, R. and Westley, J. (1986). Localization of the sulfur-cyanolysis site of serum albumin to subdomain 3-AB. *J. Biochem. Toxicol.*, **6:** 65-70.

62.   Pirisino, R., Di Simplico, P., Ignesti, G., Bianchi, G. and Barbera, P. (1988). Sulfhydryl groups and peroxidase-like activity of albumin as a scavenger of organic peroxides. *Pharmacol. Res. Commun.*, **20:** 545-552.

63.   Pedersen, A. O. and Jacobsen, J. (1980). Reactivity of the thiol group in human and bovine albumin at pH 3-9, as measured by exchange with 2,2'-dithiodipyridine. *Eur. J. Biochem.*, **106:** 291-295.

64.   Mahieu, J.-P., Gosselet, N. M., Sebille, B., Garel, M. C. and Beuzard, Y. (1993). Reactivity of 42 disulfides with thiol group of human haemoglobin and human serum albumin. *Int. J. Biol. Macromol.*, **15:** 233-240.

65.   Gosselet, M., Mahieu, J.-P. and Sebille, B. (1990). Reactivity of aromatic and heterocyclic disulphides with thiol group of bovine serum albumin. *Int. J. Biol. Macromol.*, **10:** 241-247.

66.   Byrnes, L. and Gannon, F. (1990). Atlantic salmon (*Salmo salar*) serum albumin: cDNA sequence, evolution, and tissue expression. *DNA Cell Biol.*, **9:** 647-655.

67.   Havsteen, B., Wang, X. and Hansen, H. (1994). *N. naja* (Kauthia/Ate) CSA mRNA for serum albumin and translated products. Cobra serum albumin. GeneBank Database Accession No. 469860.

68.   Hughes, W. L. (1954). The proteins of blood plasma. *The Proteins,* Vol. IIB, 663-754, Academic Press, New York.

69.   Christodoulou, J., Sadler, P. J. and Tucker, A. (1994). A new structural transition of serum albumin dependent on the state of Cys34, detection by [1]NMR spectroscopy. *Eur. J. Biochem.*, **225:** 363-368.

70.   Christodoulou, J. (1997). *Ph.D. Thesis*, University of London.

71.   Minchiotti, L., Galliano, M., Kragh-Hansen, U., Watkins, S., Madison, J. and Putnam, F. W. (1995). A genetic variant of albumin (albumin Asola;

Tyr140→Cys) with no free –SH group but with an additional disulfide bridge. *Eur. J. Biochem.*, **228:** 155-159.

72.   Christodoulou, J., Sadler, P. J. and Tucker, A. (1995). ^{1}H NMR of human albumin in blood plasma: drug binding and redox reactions at Cys[34]. *FEBS Letters*, **376:** 1-5.

73.   Lau, S. J., Kruck, T. P. A. and Sarkar, B. (1974). A peptide molecule mimicking the copper(II) transport site of human serum albumin. A comparitive study between the synthetic site and albumin. *J. Biol. Chem.*, **249:** 5878.

74.   Glennon, J. D. and Sarker, B. (1982). Nickel(II) transport in human blood serum. Studies of nickel(II) binding to human albumin and to native-sequence peptide, and ternary-complex formation with L-histidine. *Biochem. J.*, **163:** 477.

75.   Peters, T., Jr. (1960). Interaction of one mole of copper with the alpha amino group of bovine serum albumin. *Biochim. Biophys. Acta*, **39:** 546-547.

76.   Bradshaw, R. A., Shearer, W. T. and Gurd, F. R. N. (1968). Sites of binding of copper(II) ion by peptide (1-24) of bovine serum albumin. *J. Biol. Chem.*, **243:** 3817-3825.

77.   Sadler, P. J., Tucker, A. and Viles, J. H. (1994). Involvement of a lysine residue in the N-terminal Ni^{2+} and Cu^{2+} binding site of serum albumins. Comparison with Co^{2+}, Cd^{2+} and Al^{3+}. *Eur. J. Biochem.*, **220:** 193-200.

Chapter 2

Experimental Techniques

2.1 Recombinant Albumin

Recombinant albumin was supplied by Delta Biotechnology Ltd as a 250 mg mL^{-1} solution containing 145 mM NaCl, 15 mg L^{-1} tween 80 and 40 mM octanoic acid. Prior to use, the solution was divided into 5 mL aliquots and stored at −20 °C. When required, an aliquot was thawed and dialysed using dialysis tubing with a 30 kDa molecular weight cut-off (BioDesign Inc.), in 5 L of 100 mM ammonium bicarbonate for 24 hours. The protein solution was then dialysed again in a fresh solution for a further 24 hours. This process is sufficient in the removal of salt from the preparation; it is thought that some of the octanoic acid (< 4 mol mol^{-1}) does remain bound to the protein. The solution was then frozen in liquid nitrogen and freeze-dried. The albumin powder was stored at 4 °C ready for use. Mutated albumins were also treated in this way following purification.

2.2 Albumin Concentration Determination

2.2.1 UV Absorption Spectrophotometry

Proteins are chains of 20 common amino acids that absorb light at a number of different wavelengths. The peptide bond itself absorbs in the 190 to 230 nm range. Protein quantification by this means is complicated because Asp, Glu, Asn, Gln, Arg and His side-chains also absorb in this region, as do many compounds present in buffers. Aromatic residues such as Phe, Tyr and Trp have side-chains that undergo π-π*

transitions, absorbing around 280 nm (see Table 2.1) [1]. Measurement of absorption around this wavelength is one of the most commonly used methods for protein concentration determination.

	Wavelength (nm)	Extinction (Absorbtion) Coefficient ($M^{-1} cm^{-1}$)
Phenylalanine	257.4	197
Tyrosine	274.6	1420
Tryptophan	279.8	5600

Table 2.1 Spectroscopic properties of the aromatic amino acids at neutral pH. Table drawn from [2].

In accordance with the Beer-Lambert Law, protein concentration is directly proportional to light absorbed at a fixed path length (the length of sample the light passes through). The law states that the absorbance, A, is equal to the product of the extinction coefficient, ε, the concentration of the protein, c and the path length, l, $A=\varepsilon cl$.

However, because this method uses the absorption of light from a small number of amino acid types only, it is very sensitive towards the amino acid composition of the protein. Surface amino acids absorb more than buried amino acids; therefore this technique is also sensitive to the tertiary structure of the molecule. These factors consequently mean that the extinction coefficient differs for all proteins and must be known (or calculated) prior to the assay.

Albumin concentrations were determined from the absorbance 279 nm using either a Shimadzu UV250 1PC or a Cary 300 scan spectrophotometer. The extinction coefficient was assumed to be 0.531 units mg^{-1} cm^{-1} [3], and for each assay a suitable blank (or control) solution was prepared. This consisted of another solution with the same composition without albumin present and was measured and subtracted from the albumin-containing sample, in order to eliminate any background absorbance.

2.2.2 Gel Permeation Chromatography

Gel permeation chromatography (GPC) is a well-defined method for separating proteins according to their molecular weight but it can also be used to determine protein concentrations indirectly.

The method uses a column containing finely divided porous particles made of polymeric resin with a continual flow of solvent passing through. The analyte is injected at the top of the column and washed through by the solvent. Protein molecules that are smaller than the pore sizes in the particles can enter the pores, and therefore have a longer path and longer transit time than the larger molecules that are unable to enter the pores. Thus, the larger molecules elute earlier, while the smaller molecules elute later [4]. After the protein leaves the column its presence is measured quantitatively by its absorbance (280 nm is the detectable wavelength for albumin). The resultant chromatogram (absorbance against time) contains peaks corresponding to the proteins injected. The peak height is directly proportional to its concentration in the sample injected. This linear relationship means that albumin solutions of known concentration can be run in order to relate peak height to concentration. This method has an added advantage of being able to detect both monomeric and dimeric albumin quantities as they are separated on the column.

For this method to be accurate it must be reproducible. Reproducibility of GPC can be affected by the flow rate of the solvent, temperature, solvent composition, column size and the particle size within the column. Therefore it is important that these variables remain constant for each set of samples.

Separations were performed using either a Shimadzu SCL-6B or a Hewlett-Packard 1100 series liquid chromatography system, both fitted with UV/vis spectrophotometric detectors. For each assay, 25 μL of sample was loaded onto a TSKgel 3000 SWXL column (TosoHaas), dimensions 7.8 mm x 300 mm, pore size 5 μm. The column temperature was kept at a constant 30 °C and the solvent flow rate was 1 mL min^{-1}. Assays were calibrated in triplicate against a 10 mg mL^{-1} albumin standard (Delta Biotechnology Ltd.). The running solvent was a phosphate buffer, pH 7.0 and consisted of 71 g L^{-1} Na_2SO_4 (0.5 M), 13.5 g L^{-1} $Na_2HPO_4.2H_2O$ (76 mM), 7.5 g L^{-1} $NaH_2PO_4.2H_2O$ (48 mM), 32.5 mg L^{-1} NaN_3 (500 μM).

2.2.3 Rocket Immunoelectrophoresis

This is a quantitative method, which has the advantage of being specific toward a single protein only as it relies on antibody-antigen recognition [5]. An antibody (Ab) is incorporated in a gel, and the antigen (Ag) mixture is added to the well and run into an electric field. This produces a rocket-shaped precipitation zone, the height of which is proportional to the concentration of Ag (or albumin) in the sample (see Figure 2.1). Albumin standards of known concentration can be run simultaneously in order to relate rocket height to concentration. This method was used exclusively for quantifying recombinant albumin produced in yeast cultures.
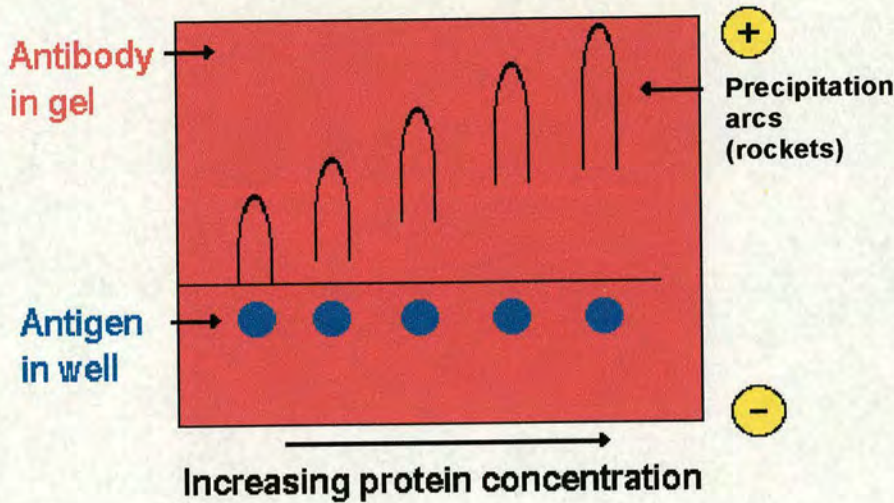
Figure 2.1   Diagramatic representation of rocket immunoelectrophoresis.   The electric field causes the protein (antigen) to migrate into the gel and form a rocket-shaped precipitate with the antibody.   The height of which is proportional to the initial antigen concentration in the well.

The gel consisted of 40 mL of 1% agarose, 0.01% Triton-X and was melted and cooled until clear. A vial of goat anti-albumin (Sigma, A-1151) was reconstituted in 5 mL of H_2O and 250 µL was added to the melted gel. The gel was poured onto Gelbond Film, 124 mm x 258 mm (Pharmacia) and allowed to set. Holes, 3 mm in diameter were then punched in the gel, 20 mm from the edge along its longest side. To each hole, 4 µL of sample or standard were added. Delta Biotechnology Ltd supplied the standards. The gel was placed in an electrophoresis tank containing rocket buffer. The rocket buffer consisted of 9.8 g L^{-1} trizma base (81 mM), 4.3 g L^{-1} tricine (24 mM), 1.06 g L^{-1} calcium lactate (3.4 mM) and 0.2 g L^{-1} sodium azide (3.1 mM). Wicks were used to allow a flow of charge from the buffer across the gel. The gel was run at 450 V for approximately 90 min. It was then blotted with absorbant paper to remove excess Ab and stained with coomassie brilliant blue gel stain. Coomassie brilliant blue, the most widely used stain for this purpose, is applied by soaking the gel in an acidic, alcohol solution containing the

53

dye. This fixes the protein in the gel by denaturing it and complexes the dye to the protein. Excess dye is removed by extensively washing the gel with an acidic solution [6]. Different gel stains have different sensitivities toward proteins. An example of this can be seen in Figure 3.7, where an SDS-PAGE gel (see Section 2.2.4) has been stained with coomassie brilliant blue and with silver stain. The silver stain allows much greater visual detection of low-level proteins than the coomassie stain. The gel was then destained with destain-solution and allowed to dry to the film. Albumin concentrations were then calculated by comparing the height of the rockets of samples with the standards.

2.2.4 Quantitative SDS-Gel Electrophoresis (SDS-PAGE)

Sodium dodecyl sulphate (SDS) is an anionic detergent, which denatures proteins by altering the balance of the weak non-bonding forces that maintain the native structure. SDS binds to proteins fairly specifically in a mass ratio of 1.4:1 and confers a negative charge to the polypeptide in proportion to its length [6]. This charge is so large that the intrinsic charge of the protein is masked. By loading wells in the gel with a sample containing protein and placing it in an electrical field it can be separated according to the charge it carries. Larger proteins carry more negative charge so move a shorter distance than smaller proteins. It is usually necessary to reduce disulphide bridges in proteins before they adopt the random-coil configuration necessary for separation by size. In denaturing SDS-PAGE separations therefore, migration is determined not by intrinsic electrical charge of the polypeptide, but by molecular weight. By running a sample against a solution containing proteins of known molecular weight, it is possible to determine the molecular weight of an unknown protein or indeed, identify a protein's presence within a mixture as a horizontal band on the gel. Proteins, as previously

mentioned can be visualised on a gel by staining. The intensity of the stain of a protein band is proportional to the concentration of protein present in the band.

Standards of recombinant albumin at 50, 100, 150 and 200 mg mL^{-1} (Delta Biotechnology Ltd.) and samples were diluted two-fold in 2x SDS buffer. The 2x SDS buffer consisted of 20 mM tris, pH 8.0, 2 mM EDTA, 5% (w/v) SDS and 0.01% (w/v) bromophenol blue. The samples were then boiled for 2 min and electrophoresed using the PhastSystem™ (Amersham). An aliquot (1 μL) was loaded onto a Phastgel® and the gel was run at 72 V for about 1 h. The gel was stained with coomassie brilliant blue gel stain solution and destained overnight. The gel was scanned using an Epson GT-9600 scanner and the band intensities measured and compared using GelWorks Advanced 4.01 (UltraViolet Products). This method (like with the rocket immunoelectrophoresis method) was only used for quantifying recombinant albumin produced in yeast cultures.

2.3 Albumin Thiol Determinations

The Cys-34 residue of human albumin makes up most of the free thiol of blood plasma and the importance of this group calls for special consideration when studying reactions at this site. A variety of ligands can bind to albumin at this site in the circulatory system. When isolated from plasma, approximately 28% of albumin molecules carry a half-cystine and 7% a half-glutathione as a mixed disulfide on this residue [7].

Recombinant albumin also contains heterogeneity at this site, although to a much lesser extent (Delta Biotechnology, unpublished data). This means that different preparations of human serum albumin and recombinant albumin contain different percentages of free

thiol. It is therefore important to know how many moles of free thiol are present per mole of albumin.

Thiol determinations were performed using a modified method based upon that of Ellman [8]. This is a spectrophotometric method, which uses 5,5'-dithiobis(2-nitrobenzoic acid), also known as DTNB or Ellman's reagent, to form a mixed disulfide with the free albumin thiol, Figure 2.2. The other product of this reaction is 5-thio(2-nitrobenzoic acid), TNB. This compound has a characteristic absorbance at 412 nm (ε = 13.9 mM^{-1} cm^{-1}). Measuring the TNB concentration produced during the reaction also gives the free thiol content of the sample because one molecule of TNB is formed for every molecule containing the mixed disulfide. This allows the number of moles of free thiol present per mole of albumin to be calculated.

For each thiol determination, a solution of DTNB in methanol was prepared (*ca.* 4000x more concentrated than the albumin to be analysed). The albumin (*ca.* 1 mg mL^{-1}) was prepared in 1 mL of 100 mM Tris buffer, pH 8.0. To each sample 10 µL of the DTNB solution (equal to 40 mol equivalents) was added and the reaction was allowed to proceed until the absorbance at 412 nm reached a maximum. A sample blank was also prepared with the same concentrations of Tris buffer, pH 8.0 and DTNB alone. The absorbance difference between the sample blank and the reaction was recorded and the thiol content was calculated.
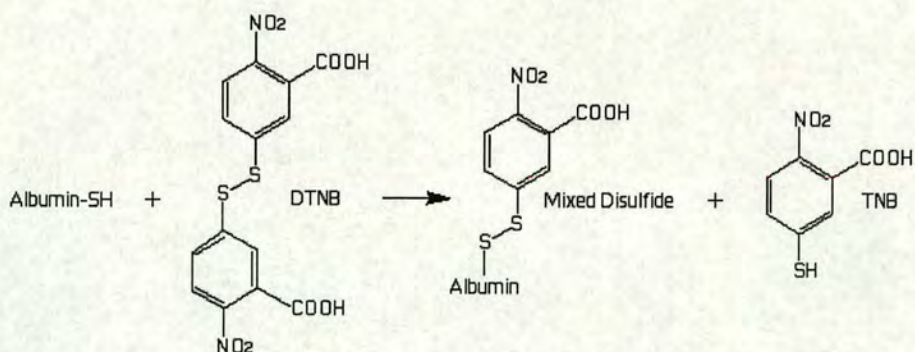
Figure 2.2 Reaction of albumin thiol with DTNB. DTNB reacts with the free thiol of albumin to form a mixed disulfide and 5-thio(2-nitrobenzoic acid) (TNB).

2.4 De-Fatting and Fatty Acid Loading

Albumin is the main transport protein for fatty acids in the circulatory system. Consequently, serum albumin isolated from plasma contains a significant amount of fatty acid bound. In addition to this, recombinant albumin contains a high proportion of octanoic acid, which is added to stabilise the molecule during purification. For some experiments in this thesis it was necessary to remove bound fatty acids from the albumin molecule and load with (sometimes several) equivalents of a particular fatty acid. This was particularly true during attempts to crystallise albumin. Fatty acids are known to invoke significant conformational changes in the molecule, see section 1.3.2. It is therefore important to be able to remove and add fatty acid to the albumin molecule.

2.4.1 Charcoal De-Fatting of Albumin

Albumin was modified using charcoal treatment, which was a modified version of that of Chen [9]. Approximately 5 g of albumin was dissolved in 100 mL of de-ionised H_2O and 3.75 g of activated charcoal (Sigma) washed with de-ionised H_2O was then added. The

preparation was mixed to give a homogenous suspension and the pH was adjusted to 2.75

by addition of HCl. The suspension was then incubated with stirring at 4 °C for 4 hours.

The mixture was then centrifuged at approximately 20,000 g for 30 min at 4 °C, the

supernatant was poured off and filtered through 0.2 μm filter. Following this, the pH was

raised to 5.5 with KOH to provide protection against acid ageing. Recovery of albumin

was typically between 75-80%. Salt was usually added to protect alterations in the

heterogeneity during storage. NaCl or KCl was added to give a final concentration of 30

mM.

2.4.2 Fatty Acid Loading of Albumin

Tetradecanoic (myristic, C14:0) acid (Sigma) was loaded on to albumin. Crystal

structure co-ordinates of albumin this these molecule bound are available in the

Brookhaven Protein Databank. The PDB code is 1BJ5 [10]. Fatty acid loading was

based upon the method used to obtain this structure.

The fatty acid was added to 20 mM phosphate, pH 7.5, to achieve a nominal

concentration of 2.5 mM. This suspension was heated to 55 °C to allow the dispersion of

the fatty acid. The solutions were cooled slightly before addition of the recombinant

albumin (Delta Biotechnology). Albumin-fatty acid complexes were prepared by mixing

50 mL of 2 mM albumin with 480 mL of the fatty acid suspension and allowed to bind at

room temperature for 1 hour. Excess fatty acid was removed by filtering through a 0.2

μm syringe filter. The solution was then concentrated to a final volume of 100 mL and

dialysed using a 10 kDa molecular weight cut-off LV centramate filter (Pall) coupled to a

peristaltic pump into a solution containing 0.1 mM of the respective fatty acid in 20 mM

phosphate, pH7.5. Solutions were concentrated to approximately 100 mg mL^{-1} using

Jumbosep centrifugal concentrators (Pall) with a molecular weight cut-off of 10,000 Da, filter sterilised and stored at –20 °C.

2.5 Mass Spectrometry

Aqueous solutions of test proteins were desalted/concentrated using either reversed phase (RP-) HPLC or reversed phase solid phase extraction (SPE) with recovered protein at concentrations of similar to 20 μM. The RP-HPLC method utilised a 5-70% gradient of acetonitrile in 0.1% TFA through a Brownlee Aquapore C8 column. Detection was by UV absorbance at 280nm. The SPE protocol employed mobile phase flow through cartridges (IST 25mg, 1000Å) under vacuum. Protein samples were loaded in aqueous solutions and eluted using 70% acetonitrile in 0.2% formic acid.

Solutions of proteins were introduced into a triple quadrupole mass spectrometer (Micromass Quattro), equipped with a conventional geometry AP-ESI source in positive ion mode, using flow injection analysis (FIA) and 20 scans were typically averaged. For protein analysis MS1 was calibrated against the protonated molecular ions of myoglobin (Sigma, horse heart) with resolution set to similar to 2000. Peptide samples were introduced using custom nanoelectrospray ion sources (both continuous flow/on-line and off-line nanovial configurations) in positive ion mode. MS and MSMS product ion spectra were acquired with significant scan averaging and the analysers were calibrated against protonated and sodiated ions from a mixture of polyethylene glycol (PEG).

2.6 Circular Dichroism

The phenomenon of circular dichroism is very sensitive to the secondary structure of polypeptides and proteins. Circular dichroism (CD) spectroscopy is a form of light absorption spectroscopy that measures the difference in absorbance of right- and left-circularly polarised light (rather than the commonly used absorbance of isotropic light) by a substance. It has been shown that CD spectra between 260 and approximately 180 nm can be analysed for the different secondary structural types: alpha helix, parallel and antiparallel beta sheet, turn, as well as others [11].

The simplest way of extracting secondary structure content from CD data is to assume that a spectrum is a linear combination of CD spectra of each contributing secondary structure type (e.g., "pure" alpha helix, "pure" beta strand etc.) weighted by its abundance in the protein's conformation [12]. A drawback of this approach is that there are no standard reference CD spectra for "pure" secondary structures. Synthetic homopolypeptides used to obtain reference spectra are in general, poor models for the secondary structures found in proteins. For example, the CD of an alpha helix has been shown to be length dependent and no homopolypeptide system has been found that is a good example of the beta sheet structure found in proteins. Several methods have been developed which analyse the experimental CD spectra using a database of reference protein CD spectra containing known amounts of secondary structure.

Albumin was diluted to approx. 1.5 mg mL^{-1} in 200 mM potassium phosphate, pH7.4. All spectra were recorded by Dr. Sharon Kelly (Scottish Circular Dichroism Facility, University of Glasgow). The instrument used was a JASCO J-600 spectropolarimeter. Secondary structure estimations were calculated using the SELCON procedure [13,14].

2.7 pH and Conductivity Measurements

2.7.1 pH Measurements

The pH measurements of buffers and solutions were performed using either a Corning pH meter 145 or a Corning pH meter 240 fitted with a glass calomel combination electrode (Aldrich) with a pH range of 0-14. This was with the exception that the pH measurements of NMR samples were performed using a Orion pH/ISE 710A meter fitted with a 4 mm diameter glass electrode (Aldrich). All electrodes were standardised with pH 4.0, 7.0 and 10.0 buffers (Aldrich). Values for ^2H_2O (D_2O) solutions are termed pH* and are uncorrected for the effect of deuterium on the glass electrode.

2.7.2 Conductivity Measurements

Conductivity measurements were performed at room temperature using a Radiometer CDM 83 conductivity meter fitted with a CDC 304T probe and was standardised with salt solutions of known conductivity.

References: Chapter 2

1.   Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. and Gray, T. (1995).   How to measure and predict the molar absorption coefficient of a protein. *Prot. Sci.*, **4:** 2411-2423.

2.   Bailey, J. E., Beaven, G. H., Chignell, D. A. and Gratzer, W. B. (1968).   An analysis of perturbations in the ultraviolet absorption spectra of proteins and model compounds. *Eur. J. Biochem.*, **7:** 8-14.

3.   Janatova, J., Fuller, J. K. and Hunter, M. J. (1968).   The heterogeneity of bovine albumin with respect to sulfhydryl and dimer content. *J. Biol. Chem.*, **243:** 3612-3622.

4.   Jenik, R. A. and Porter, J. W. (1981).   High-performance liquid chromatography of proteins by gel permeation chromatography. *Anal. Biochem.*, **111:** 184-188.

5.   Cann, J. R. (1975).   A phenomenological theory of rocket immunoelectrophoresis. *Biophys. Chem.*, **3:** 206-214.

6.   Voet, D. and Voet, J. G. (1995). *Biochemistry, 2nd Ed.* John Wiley & Sons, Inc.

7.   Andersson, L. O. (1966).   The heterogeneity of bovine serum albumin. *Biochim. Biophys. Acta.*, **117:** 115-133.

8.   Ellman, G. (1958).   A colorimetric method for determining low concentrations of mercaptans. *Arch. Biochem. Biophys.*, **74:** 443-450.

9.   Chen, R. F. (1967).   Removal of fatty acids from serum albumin by charcoal treatment. *J. Biol. Chem.*, **242:** 173-181.

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

11.   Johnson W. C., Jr. (1990).   Protein secondary structure and circular dichroism: a practical guide. *Proteins*, **7:** 205-214.

12.   Hennessey, J. P., Jr. and Johnson, W. C., Jr. (1981).   Information content in the circular dichroism of proteins. *Biochemistry*, **20:** 1085-1094.

13.   Sreerama, N. and Woody, R. W. (1993).   A self consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.*, **209:** 32-44.

14.   Sreerama, N. and Woody, R. W. (1994).   Poly(Pro)II type structure in globular proteins - Identification and CD analysis. *Biochemistry*, **33:** 10022-10025.

Chapter 3

Site-Directed Mutagenesis of rHA

3.1 Introduction

As discussed in Section 1.3.4, the Cys^{34} site is very important in the binding of a variety of metallodrugs to albumin. The binding properties at this site are largely attributed to the low pK_a of Cys^{34} (ca. 7) [1]. This abnormally low pK_a means that the free thiol group predominantly exists in the S^- form at physiological pH and hence is very reactive. Cys^{34} is deeply buried within the molecule but is thought to flip out upon ligand binding (Figure 1.6). The crystal structures of human albumin [2,3] reveal that residues His^{39} and Tyr^{84} are located very close to Cys^{34} in the non-liganded state and therefore may play a role in stabilising the S^- form of the thiol. The aim of this study was to alter the primary structure of human albumin through the technique of site-directed mutagenesis at these two candidate amino acids to ascertain their importance within this binding site and in albumin as a whole.

An important factor when deciding which amino acids were to be substituted within the Cys^{34} binding cleft is that a dramatic alteration could lead to a major disruption in protein conformation in this region. Therefore the change made had to be relatively subtle with regard to the substitution of amino acids. The amino acid chosen to replace His^{39} was leucine. Replacement of histidine with another positively charged amino acid such as lysine or arginine was ruled out as these residues have long hydrophobic chains, which could produce large conformational changes at the Cys^{34} site. Leucine was chosen because it is structurally similar to histidine; both amino

acids have a CH_2CH- branch but the leucine has two branched methyl groups attached rather than an imidazole ring. A choice of substitute for Tyr^{84} was more obvious. Phenylalanine differs from tyrosine only by the lack of a single oxygen atom, in the *para* position. A Cys^{34} to alanine mutant had previously been synthesised by Delta Biotechnology Ltd. and was also expressed and purified. All three substitutions are shown in Figure 3.1.

Additionally, molecular modelling based on the crystal structure of albumin [3] (PDB 1AO6) suggested that a multi-metal binding site may involve a region at the interface of domains I and II, in which His67, Asn99, His247 and Asp249 are clustered (as discussed in Chapter 5). In order to determine whether this cluster is a metal ion binding site the His^{67} residue was mutated to an alanine, removing an imidazole ring and potential nitrogen ligand for metal coordination. This substitution is shown in Figure 3.2.
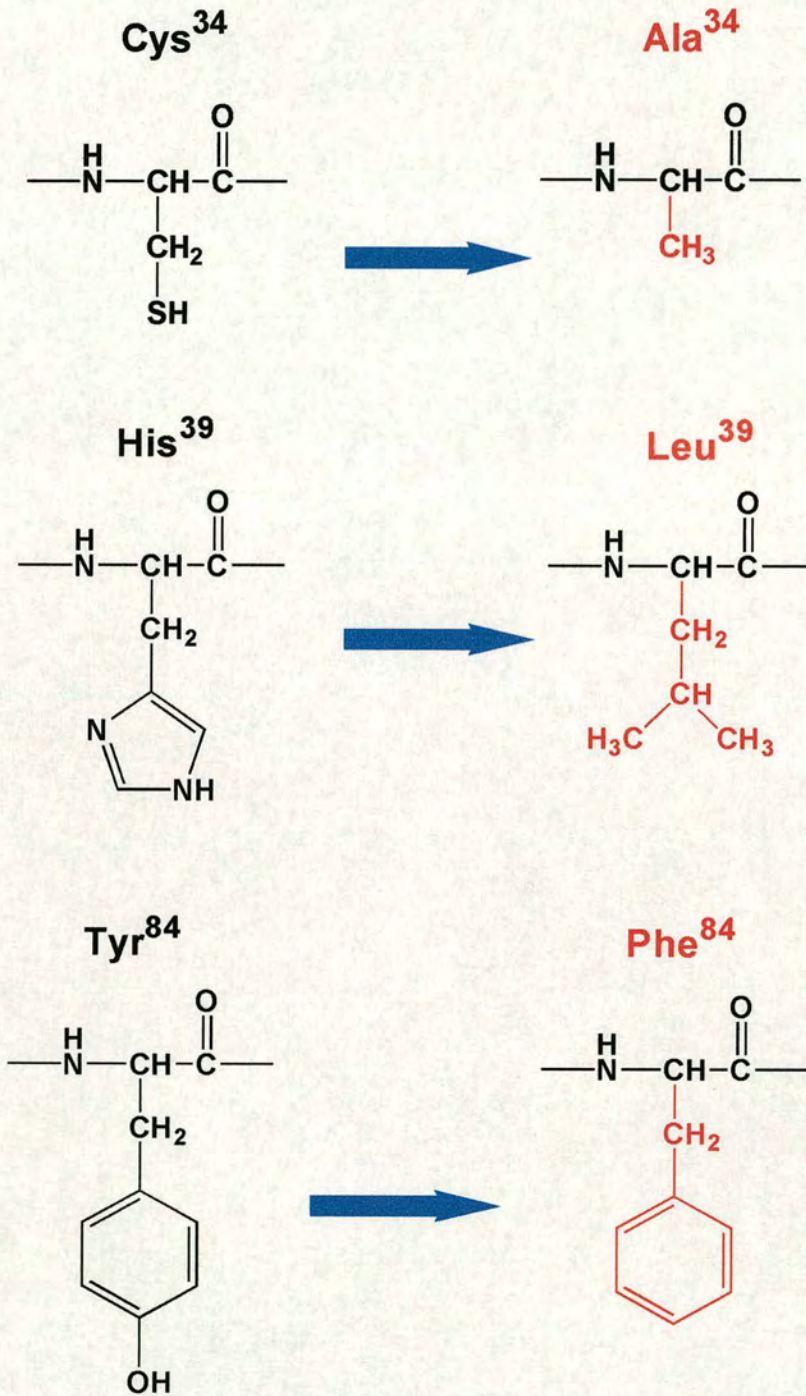
Figure 3.1 Amino acid substitutions around the Cys³⁴ site of human albumin.

The side-chains of the residues in each substitution are shown in red.

His[67]   Ala[67]
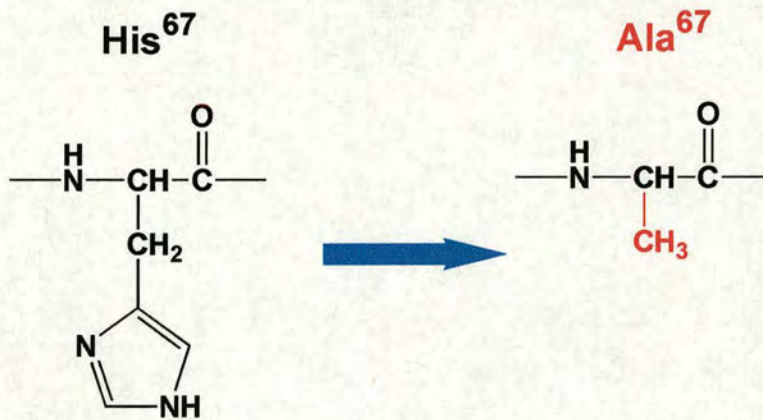
Figure 3.2 Amino acid substitution at His[67] of human albumin. The side-chain of the residue substituted is shown in red.

3.2 Synthesis of Mutated Human Albumin *E.coli* Expression Vectors

3.2.1 His[39]Leu and Tyr[84]Phe Mutant Albumins

The DNA sequence of the *alb* gene [4] was studied and it was found that CAT and TAT codons encode the His[39] and Tyr[84] residues respectively. Only a single base pair change in each codon was therefore required to encode Leu and Phe residues (CAT→CTT and TAT→TTT respectively).

The highly efficient method of site-directed mutagenesis described by Thomas Kunkel [5] was chosen, in order to minimise the need for DNA sequencing when screening for the mutated plasmid. This method is based upon the use of pQC262e, a single-stranded circular DNA template containing the albumin coding sequence and

an *E. coli* origin of replication (Figure 3.3). In order for the mutagenic reaction to occur, the template must have a proportion of uracil bases in place of thymine bases within its DNA sequence.

The uracils are incorporated by transforming the vector into a mutant strain of *E. coli*. This strain has 2 genes knocked out, *dut* and *ung*, which encode dUTPase and uracil *N*-glycosylase respectively. *E. coli* strains which lack the enzyme dUTPase contain elevated concentrations of dUTP [6,7], which effectively competes with dTTP for incorporation into DNA. *E. coli* strains, which also lack uracil *N*-glycosylase cannot remove incorporated uracil from DNA [8]. It should also be noted that for *in vitro* reactions uracil-containing templates are indistinguishable from normal templates since dUTP has the same coding potential as dTTP [9]. Uracil is non-mutagenic.

Following transformation of pQC262e and growth of the mutant strain, the uracil-containing vector can be isolated and used as a template for the production of a complementary strand that contains the desired sequence alteration but contains only TMP and no dUTP bases.

Mutagenic primers were designed for each desired mutation. Each mutagenic primer is a 35 bp DNA fragment, which is complementary to the *alb* gene at 34 of the 35 bp. The remaining base lies central within the primer and is complementary to the desired mutation. The mutagenic primers for the His[39]Leu and Tyr[84]Phe mutations were 5'-CATTCACTAATTTTACAAGATCTTCAAATGGACAC-3' and 5'-GCAGTCAGC CATTTCACCAAAGGTTTCACGAAGAG-3', respectively.
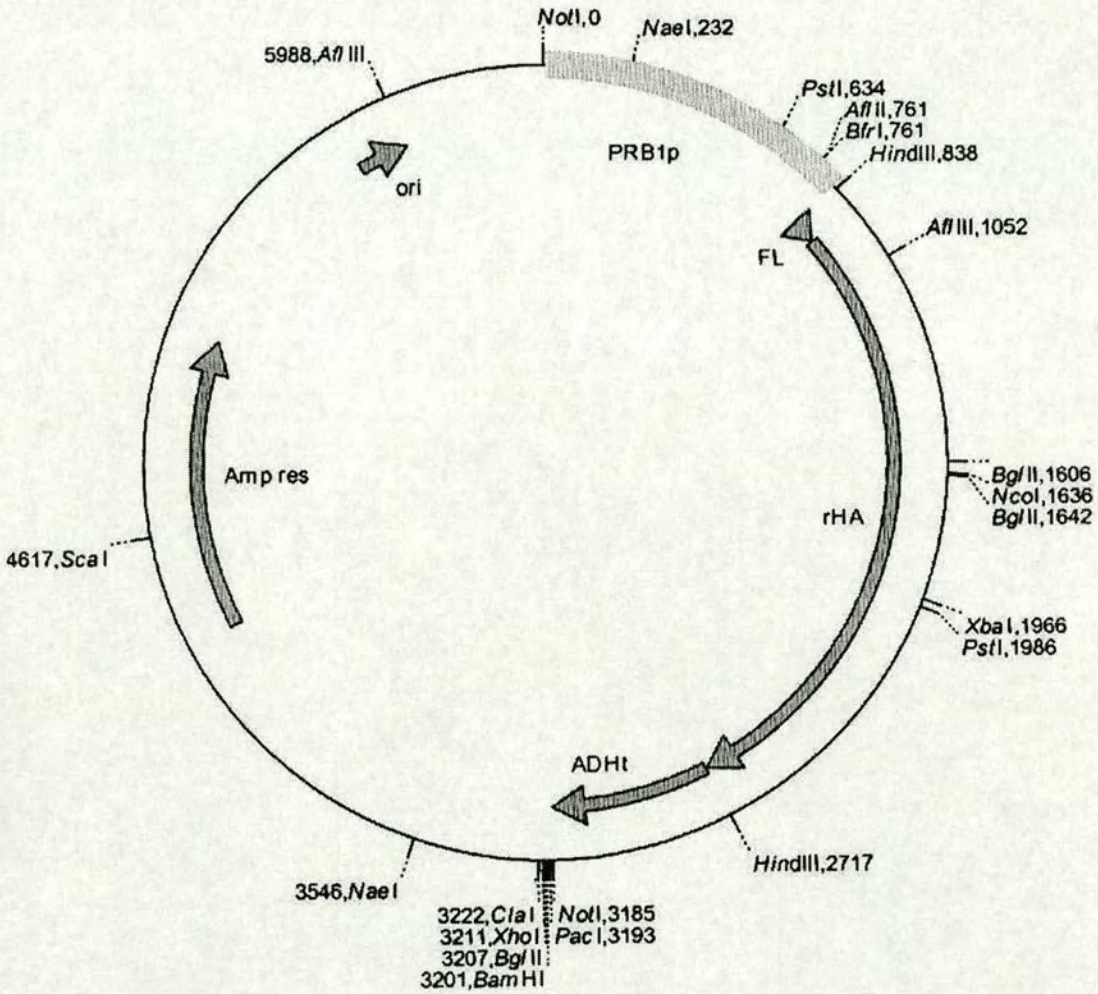
Figure 3.3   Map of pQC262e, the DNA template used for the mutagenesis reactions. The 6,397 base pair template contains the coding sequence for human albumin (rHA) with a promoter (PRB1p), terminator (ADHt) and fusion leader sequence (FL) conferring protein secretion in yeast. It also includes an *E. coli* origin for replication (ori) and confers ampicillin resistance (Amp res). Various restriction endonuclease sites are also shown.

3.2.1.1 Phosphorylation of Mutagenic Primers and Template Binding

The mutagenic primers (supplied by Delta Biotechnology Ltd.) were phosphorylated to allow later the 5' addition of dNTPs. For each mutation the following reaction mixture was prepared: 1 µL respective 20 µM primer solution, 2 µL T4 polynucleotide kinase x10 buffer solution, 1 µL 20 mM ATP, 1 µL T4 polynucleotide kinase, 15 µL H$_2$O. The reaction mixtures were then incubated at 37 °C for 60 min. Following the phosphorylation reaction 3 µL of 0.1 M ETDA (pH 7.5) was added, they were then heated at 65 °C for 10 min. The pQC262e template, 0.6 µL of 33 µM (Delta Biotechnology Ltd.) and 1.2 µL of 20x sodium citrate/sodium chloride (SSC) were then added to the mixtures. The mixtures were then placed in a beaker of water at 65 °C and allowed to cool at room temperature. The annealed primer/template mixtures were stored overnight at 20 °C.

3.2.1.2 Mutagenesis

The annealed primer/templates consist of a single-stranded loop of uracil-containing DNA with a single non-uracil-containing 35 bp complementary fragment bound. In order to make the loop double-stranded the following reaction mixture was prepared: 24.8 µL template/primer mixture, 10 µL 0.2 M Hepes, pH 7.8, 10 µL 20 mM dithiothreitol, 10 µL 0.1 M MgCl$_2$, 5 µL dNTPs (10 mM of each dATP, dCTP, dGTP and dTTP), 5 µL 20 mM ATP, 2 µL T4 DNA polymerase (2 units µL^{-1}), 1 µL T4 DNA ligase (5 units µL^{-1}), 35 µL H$_2$O. Reactions were placed on ice for 5 min then allowed to stand at room temperature for 5 min. This was followed by incubation at 37 °C for 2 h before the mixtures were again placed on ice.

Each double-stranded DNA was then transformed into DH5α (wild-type) *E. coli* cells. A 50 μL aliquot of each reaction mixture was added to 250 μL of competent DH5α cells and put in ice for 45 min. The cells were then heat shocked at 42 °C for 90 seconds to allow cellular uptake of the plasmid vector then placed on ice for 2 min. LB (700 μL) was then added to each mixture. The mixtures were then spread on LB agar plates containing ampicillin (100 μg mL^{-1}). The plates were then incubated overnight at 37 °C. These resulted in 12 possible His^{39} to Leu mutant colonies and 28 possible Tyr^{84}Phe mutant colonies.

Due to the presence of the ampicillin in the plates only DH5α cells, which have successfully incorporated the vector, can form colonies on the plate. In addition, DH5α cells contain both the *dut* and *ung* genes and therefore can make both dUTPase and more importantly uracil *N*-glycosylase. The glycosylase renders certain sites on the uracil-containing strands of the plasmids to become apyrimidinic (AP) sites [10]. These AP sites are lethal lesions, which block DNA synthesis [8] and are sites for incision by AP endonucleases which produce strand breaks [10]. These render most of the template strand to be biologically inactive. Thus the majority of the cells' progeny arise from the complementary strands, which contain the desired mutations. This results in a high efficiency of mutant production, typically >50% [5]. Screening for mutations was carried out by DNA sequence analysis.

3.2.1.3 Mutant Screening

Of the resultant colonies from the mutagenesis reaction, 4 were picked for each mutation. These were then grown overnight in 25 mL of LB containing 100 μg mL^{-1} ampicillin at 37 °C. The cells were harvested by brief centrifugation and the plasmids

were isolated using a QIAGEN QIAfilter plasmid kit.   The plasmids were then resuspended in 200 μL of H_2O and stored at -20 ºC.   A 20 μL aliquot of each plasmid DNA solution was diluted into 980 μL and the concentrations were then measured by absorbance at 260 nm.

In order to screen for successful mutations the coding region of the *alb* gene was cut out using HindIII restriction endonuclease from each of the resulting plasmid DNAs. HindIII restriction sites flank either end of the coding sequence (Figure 3.2). Digestion was also carried out with BglII restriction endonuclease as coincidently it was found that the histidine to leucine mutation created an extra BglII restriction site, 5'-AGATCT-3'.

For each digestion the following reaction mixture was prepared: 1 μg DNA, 1 μL 10 units μL^{-1} enzyme (HindII or BglII), 2 μL of x10 restriction enzyme buffer, H_2O was then added to bring the final volume to 20 μL. Reactions were incubated at 37 ºC for 2 hrs. 5 μL of gel loading buffer (containing 10 mg/mL RNase A) was added to each sample. 20 μL of each sample was loaded on a 1% agarose TBE gel containing ethidium bromide. DNA ladder (5 μL of 0.1 mg mL^{-1})was also loaded. The gel was run at 100 V for approximately 1 h.

The HindIII enzyme cut the DNA in 2 places yielding 2 fragments (a 4,383 bp fragment and an 1,820 bp fragment). The BglII enzyme cut the DNA in 3 places yielding 3 fragments (4,796, 1,371 and 36 bp fragments), except where the histidine to leucine mutation had occurred. In this case an additional band of 597 bp was seen with the 4,796 bp reduced to 4,199 bp.

Following the HindIII digest 2 visible bands were seen on the gel corresponding to the expected 4,383 and 1,820 bp fragments. The BglII digest yielded 2 bright bands corresponding to the 2 larger expected fragments. One colony had an additional smaller band, which was consistent with the 597 bp fragment expected for plasmids containing the histidine to leucine mutation. Note that on the BglIII gel the 36 bp fragment cannot be seen due to its size. This is because ethidium bromide, which is used to visualise the bands on the gel is a DNA intercalating agent and therefore the UV intensity of a band is dependent upon the size of the fragment, with larger bands always appearing brighter.

Polymerase chain reactions were performed in order to prepare the plasmid DNA for sequencing. The polymerase chain reaction (PCR) is a technique used to produce a large number of copies of a certain DNA sequence. The reaction uses a temperature-resistant polymerase from the bacterium *Thermus aquaticus*, *Taq* polymerase, which catalyses the growth of DNA primers [11]. Double stranded DNA is heat-denatured to generate single stranded templates. Lowering of the temperature allows complementary primers on opposite strands to extend in different directions, toward each other, in the presence of the components necessary for polymerisation. Repeated cycles of denaturation and synthesis lead to an exponential growth in the number of segments replicated. Figure 3.4 shows a representation of the PCR reaction.
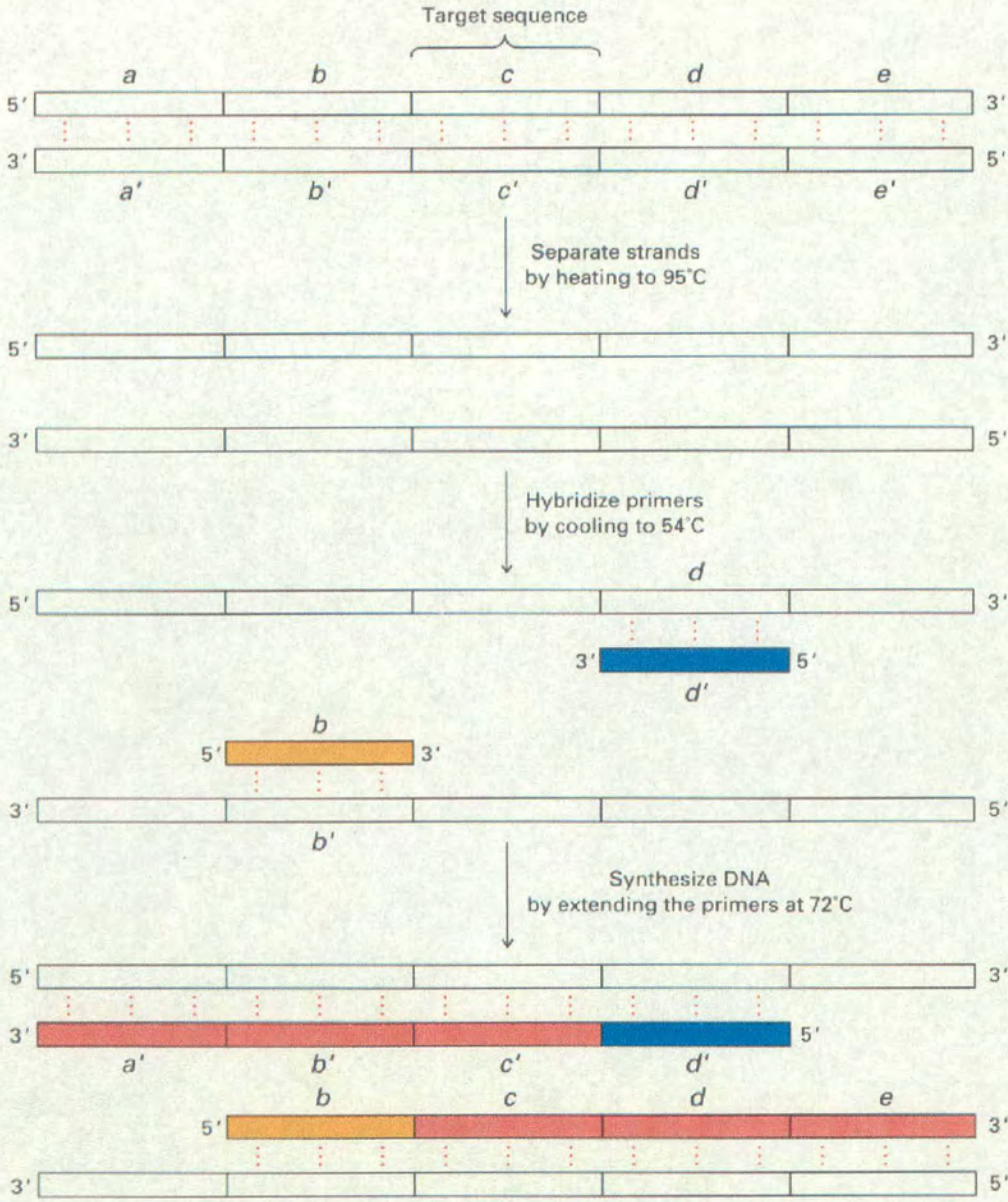
Figure 3.4 The polymerase chain reaction. The original DNA sequence is denoted by *abcde* and the complementary sequence by *a'b'c'd'e'*. Primer *b* is shown in yellow and primer *d'* in blue. New DNA is shown in red. Diagram taken from [12].

However in this case the PCR reaction performed was slightly different in that as well as containing the normal dNTPs which bind to the primers elongating the DNA, ddNTPs (dideoxy nucleotide triphosphates) were also present at a concentration 100 times lower than the dNTPs. Complementary ddNTPs stick to the primer in exactly the same way as dNTPs but stop the chain from getting any longer due to the absence of the 3' hydroxyl group, necessary to bond with the next dNTP. The consequence is that instead of generating millions of copies of the full coding sequence many different forms of "stunted" sequences are obtained. The elongation reaction proceeds so far before a ddNTP gets incorporated and the chain stops. By the laws of probability DNA chains are obtained which are a single base longer than the primer, 2 bases longer, 3 bases, 4, 5, 6 and so on for many hundreds of bases. Each type of ddNTP (i.e. ddATP, ddCTP, ddGTP and ddTTP) also has bound a different coloured fluorophore. The DNA sequencer separates all these oligonucleotides by their size by capillary electrophoresis. The capillary has a "window" at one end for fluorophore detection, smaller chains travel faster along the capillary than long chains. This means that the first chains to reach the window have a single additional base with a fluorophore attached, the next have an additional 2 bases (the latter with a flurophore), the next have an additional 3 bases (again, the latter with a flurophore) and so on. As each set of chains pass the capillary window the fluorophore, which is present in greatest abundance is detected. This allows the DNA sequencer to put together the full DNA sequence as each fluorophore represents a particular base (as shown in Figure 3.5).
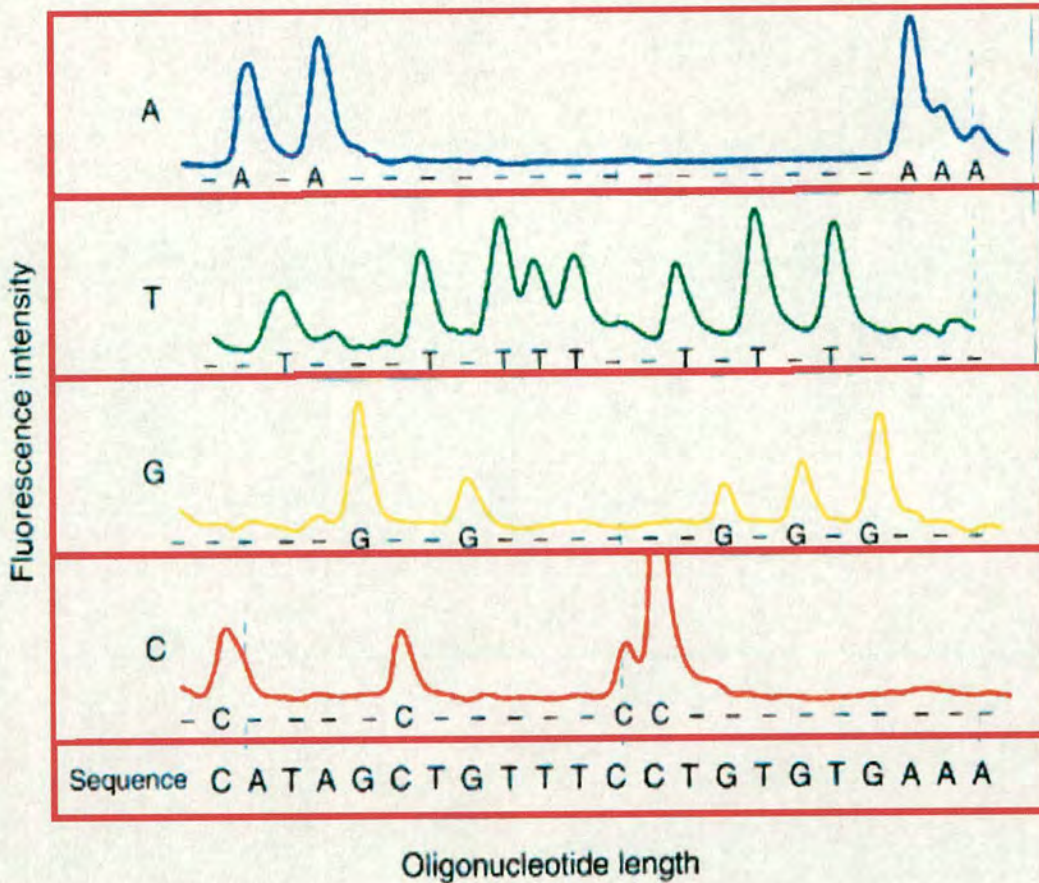
Figure 3.5 DNA fluorescent sequencing. Four different chromophores are used on the ddNTPs in the dideoxy termination reactions to put together a sequence. Diagram modified from [11].

Plasmid DNA from 4 potential tyrosine to phenylalanine mutant colonies was selected as was plasmid DNA from the colony yielding the extra BglII band as this inferred it to be a successful histidine to leucine mutant. PCR reactions with the selected plasmid DNAs were set up as follows: 8 µL ABI Prism™ big dye terminator cycle sequencing ready reaction mixture (with Ampli Taq polymerase for fluorescent sequencing), 25 µg mL^{-1} plasmid DNA, 40 nM primers (either TW2 and HSA7 (Delta Biotechnology Ltd.), which together cover the human albumin coding sequence), H$_2$O was added to bring the final volume 20 µL. The PCR reaction consisted of 25 cycles

of 96 °C for 10 s, 50 °C for 5 s and 60 °C for 4 min. The samples were then incubated at 4 °C.

Following the PCR reaction 2 μL of 3 M sodium acetate, pH 4.6 and 50 μL of 95% ethanol were added to the mixtures. The samples were then incubated at room temperature for 15 min and centrifuged at 13,000 rpm for 20 min. The supernatants were removed and the pellets were washed with 250 μL of 70% ethanol (by vortexing and centrifugation at 13,000 rpm for 5 min). The pellets were then dried in a 95 °C heating block and resuspended in 12 μL of template suppression agent. They were vortexed and spun down before sequencing on the Perkin-Elmer ABI Prism 310 Genetic Analyser. Plasmid DNA isolated from two colonies were found to have incorporated the tyrosine to phenylalanine mutation, the other two had not. The DNA isolated from the colony with the extra BglII band was confirmed to have incorporated the histidine to leucine mutation.

3.2.1.4 Insertion of Mutation into Fully Sequenced Plasmid

Before the respective mutant expression cassettes could be inserted into a yeast expression vector, it was important that no other mutations had taken place. The risk of which was minimised through the use of a plasmid, which had previously had all 6,397 bp sequenced (pDB2243). Both mutations are located in a region of the plasmid flanked at one side by a single NcoI restriction site and at the other a single BfrI restriction site (see Figure 3.2). In each of the mutant plasmids this region was sequenced, cleaved and ligated into the pDB2243.

The region of plasmid DNA from the 3 mutant colonies (1 His^{34}Leu mutant and 2

Tyr^{84}Phe mutants) to be cut and ligated, was sequenced as before, with the exception

that 3 additional primers (HSA8, HSA11 and HSA19 (Delta Biotechnology Ltd.))

were needed to flank the region which lies between the single NcoI and BfrI

restriction endonuclease sites on the plasmid. No additional mutations arose within

two of the colonies. However a plasmid from one of the tyrosine to phenylalanine

mutants had developed an A to T base change in the coding region of the *alb* gene.

Sequence analysis revealed however that this additional mutation is invisible at the

amino acid level. The respective codon would still code for the amino acid alanine at

Ala^{215}.

Restriction digests with the mutant plasmids and pDB2243 were set up as follows: 2

μL x10 restriction enzyme buffer, 4 μL 10 units μL^{-1} BfrI, 8 μL 10 units μL^{-1} NcoI, 1

μg respective plasmid DNA (0.1 μg of pDB2243), H$_2$O was added to bring the final

volume to 100 μL. Reaction mixtures were incubated at 37 °C for 2 hours. 40 μL (2

x 20 μL) of each were loaded on a 1% agarose TAE gel containing ethidium bromide

and run as before.

The larger bands from the pDB2243 digest and the smaller bands from the others

were cut from the gel with a sterile scalpel blade and the DNA was then purified using

the Geneclean II kit. To each volume of DNA/gel 3 volumes of sodium iodide

solution was added. The solutions were then incubated at 50 °C for 5 min to allow the

DNA/gel to dissolve. 5 μL of EZ-Glassmilk was added to each solution, before a

further incubation at room temperature for 5 min. The solutions were then

centrifuged at 13,000 rpm for 5 min; the supernatants were removed and washed 3

times with 200 μL of wash solution. The DNA was eluted from the pellet with 20 μL

of elution solution and the supernatants were collected. A 2 μL aliquot of each

supernatant was then run on a 1% agarose TBE gel containing ethidium bromide.

Ligation reactions were set up with each of the smaller mutated plasmid NcoI/BfrI

fragments with the large pDB2243 NcoI/BfrI fragment as follows: 2 μL 10x ligation

buffer, 2 μL 20 mM ATP, 2 μL pDB2243 (containing 0.5 μg DNA), 2 μL of

respective mutated fragment containing 0.2 μg DNA, 1 μL of DNA ligase and 8 μL

H_2O.

3.2.2 His^{67}Ala Mutant Albumin

The DNA sequence of the *alb* gene [4] was studied and it was found that a CAT

codon encodes the His^{67} residue. Therefore a double base pair change in the codon

was required to encode an Ala residue (CAT→GCT).

Unlike the Cys^{34} site mutations, for the His^{67} to alanine mutation, the QuikChange®

Site-Directed Mutagenesis Kit (Stratagene) was used. The kit incorporates a different

method than that of Kunkel, which allows the mutagenesis reaction to take place in a

double-stranded DNA template. Also this method is perfomed using PfuTurbo®

DNA polymerase, a high fidelity polymerase (6-fold higher fidelity than *Taq*

polymerase) that replicates both plasmid strands without displacing the mutant

oligonucleotides primers.

The basic procedure utilises a supercoiled double stranded DNA (dsDNA) vector

(based on pQC262e, figure 3.3) and two synthetic oligonucleotides primers containing

the desired mutation. The oligonucleotides primers, each complementary to opposite strands of the vector, are extended via the polymerase chain reaction (figure 3.4) by the PfuTurbo® polymerase. The two primers used for the His[67]Ala mutation were 5'-GCTGAAATTGTGACAAATCACTTGCTACCCTTTTTGGAGACAAATTATGC-3' and 5'-GCATAATTTGTCTCCAAAAAGGGTAGCAAGTGATTTGTCACAATT TTCAGC-3'. Incorporation of the oligonucleotides primers generates a mutated plasmid containing staggered nicks. Following the PCR, the product is treated with DpnI. The DpnI endonuclease (target sequence 5'-GAmTC-3') is specific for methylated and hemimethylated DNA and is used to digest the parental DNA template and to select for mutation-containing synthesised DNA [13]. DNA isolated from almost all *E. coli* strains is dam methylated (the methylase specified by the *dam* gene methylates the N^6-position of the adenine in the sequence 5'-GATC-3') [14] and is therefore susceptible to DpnI digestion. The nicked vector containing the desired mutations is then transformed into competent *E. coli* cells.

3.2.2.1 Mutagenesis

The following mutagenesis reaction mixture was prepared for PCR: 2µL of double stranded template (pQC262e), 5 µL of 10x reaction buffer, 1.25 µL of each primer (at 100 ng µL^{-1}), 1µL dNTPs (10 mM of each dATP, dCTP, dGTP and dTTP), 39.5 µL H$_2$O and 1 µL of PfuTurbo® DNA polymerase. The PCR reaction consisted of 1 cycle at 95 °C for 30 s, followed by 17 cycles of 95 °C for 30 s, 55 °C for 1 min and 68 °C for 14 min.

Following PCR, 1 µL of DpnI (10 units µL) was added to the reaction to remove the methylated, unmutated, parental strand. The reaction was then mixed and allowed to

digest at 37 °C for 1 hour. Following digestion, 1 μL of the reaction was then added to 50 μL of competent DH5α *E. coli* cells at 0 °C and incubated on ice for 30 min. The reaction was heat pulsed at 42 °C for 45 s and placed back on ice for 2 min to allow transformation to occur. LB broth (0.5 mL) was then added to the reaction and it was incubated at 37 °C for 20 min. The mixtures were then spread on LB agar plates containing ampicillin (50 μg mL^{-1}). The plates were then incubated overnight at 37 °C. Due to the presence of the ampicillin in the plates, only cells that have successfully incorporated the vector can form colonies on the plate. These resulted in 24 possible His^{67} to Ala mutant colonies.

3.2.2.2 Mutant Screening

Of the resultant colonies from the mutagenesis reaction, 3 were picked and were screened for successful mutations by fluorescent DNA sequencing. DNA was extracted and sequenced as described in Section 3.2.1.3. All were found to contain the desired mutation.

3.2.2.3 Insertion of Mutation into Fully Sequenced Plasmid

This was carried out using DNA from a colony of cells confirmed to carry the His^{67} to Ala mutation. The region of DNA carrying the mutation (between the NcoI and BfrI restriction endonuclease sites) was sequenced, confirmed to only carry the desired mutation, cleaved and inserted into pDB2243 as described in Section 3.2.1.4.

3.3 Synthesis of Yeast Vector and Transformation into *S. cerevisiae* cells

3.3.1 Introduction

The *S. cerevesiae* cells used were of the DXY1 strain, which contains disruptions in 2 of its genes, *Leu2* and *Yap3*. *Leu2 is* a gene encoding an essential enzyme in the leucine synthesis pathway. Yeast cells without this gene are unable to make leucine and hence cannot grow in the absence of leucine. The *Leu2* gene is present however on the yeast vector and so this disruption is used as a means of selection (only cells with this plasmid can grow). *Yap3* encodes a yeast aspartyl protease, which proteolytically clips the albumin at a specific site causing fragmentation [15].

3.3.2 Insertion of Expression Cassette into Yeast Vector

The yeast expression vector is a pUC9 vector with a single NotI site for insertion of the expression cassette (Figure 3.6). The pUC9 plasmid is a versatile vector, which can replicate in both *E. coli* and *S. cerevisiae* cells. Plasmids from all three mutant *E. coli* colonies were digested to allow insertion of the expression cassette into a yeast vector for expression. The expression cassette is the part of the *E. coli* plasmid, which carries the mutated albumin coding sequence and sequences required for transcription. The expression cassette is flanked at either side by a NotI restriction endonuclease site (Figure 3.3). Digestion with NotI alone would yield two fragments of 3,185 and 3,212 bp. These fragments are so similar in size that they would not be able to be separated on an agarose gel. Therefore the plasmids were also digested with ScaI, an

endonuclease that cuts the 3,212 bp fragment into 1,405 and 1,780 bp and allows easy

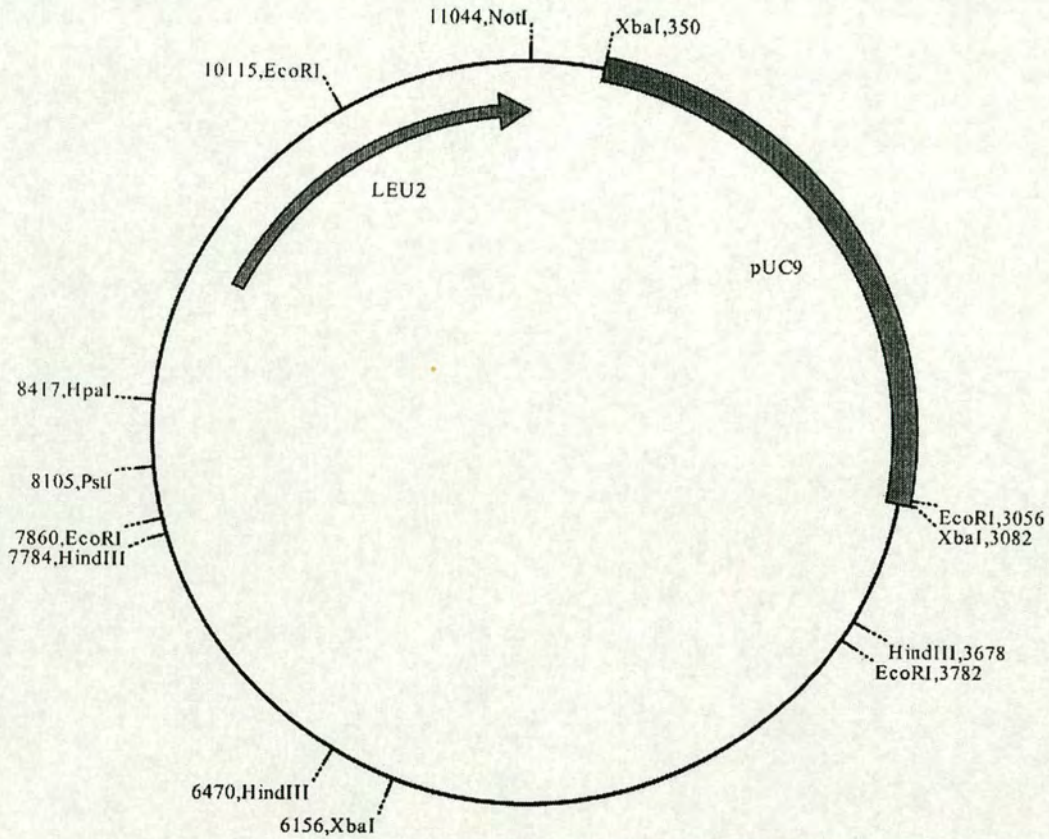recovery of the 3,185 bp expression cassette from the gel.

Figure 3.6 The Yeast expression vector, pSAC35. The vector contains the *Leu2*

gene as means of phenotypic selection and the pUC9 sequence. There is a NotI

endonuclease site for insertion of the mutated expression cassette. Various other

restriction endonuclease sites are also shown.

83

Restriction digests were set up as follows: 10 µL of 10x digestion buffer, 3 µL of respective plasmid DNA (1.5 µg mL^{-1}), 5 µL NotI (10 units µL^{-1}), 5 µL ScaI (10 units µL^{-1}) and 77 µL of H$_2$O. The digests were incubated at 37 °C for 1 hour and were ran on a 1% (w/v) agarose TAE gel containing ethidium bromide as previously described. The 3,185 bp fragment was cut from the gel and purified as described in Section 3.2.1.4.

The expression cassettes were then ligated into the pSAC35 yeast vector, which had previously been linearised with the NotI endonuclease. The reaction was set up as follows: 1 µL of 10x ligation buffer, 1 µL of pSAC35 (1.5 µg mL^{-1}), 5 µL of respective expression cassette (purified from gel), 1 µL of T4 DNA ligase and 2 µL H$_2$O. A control was also set up with 5 µL of H$_2$O instead of the expression cassette. Reactions were incubated at room temperature for 2 hours. To the ligation mixtiures, 100 µL of competent DH5α cells were added and the suspension was than placed on ice for 30 min. The cells were then heat-shocked at 42 °C for 40 s and placed back on ice for 2 min to allow transformation to occur. To each transformant, 200 µL of LB broth was added. They were then incubated at 37 °C for 15 min. The transformants were then plated onto LB agar plates containing 100 µg mL^{-1} ampicillin and incubated at 37 °C overnight.

Large numbers of colonies grew following ligation and transformation for each respective mutant. No colonies grew on the control plate, suggesting that most if not all of the colonies on the ligation/transformation plates had successfully incorporated the expression cassette. Twelve colonies from each respective mutation were then grown overnight in 25 mL of LB broth containing 100 µg mL^{-1} ampicillin at 37 °C.

The cells were harvested by brief centrifugation and the plasmids were isolated using a QIAGEN QIAfilter plasmid kit.

Restriction digests with NotI and EcoRI enzymes (set up as previous digests) were ran on a 1% (w/v) agarose TBE gel containing ethidium bromide and confirmed the presence of and fragments sizes expected of the final yeast vector.

3.3.3 Transformation of Yeast Vector into *S. cerevisiae* DXY1 Cells

S. cerevisiae DXY1 cells were grown in YEPD to late log phase/early stationary phase. The YEPD media consisted of 10 g L^{-1} yeast extract (Difco), 20 g L^{-1} Bactopeptone (Difco) and 2 % glucose. 20 mL of cells were harvested and washed three times with H_2O at 4 °C then resuspended in 2 mL of H_2O and cooled to 0 °C. Electroporation cuvettes (2 mm gap, BioRad) were cooled to 0 °C. In each of them 20 µL of respective vector, 100 µL of cells and 50 µL of 40% (w/v) polyethylene glycol 3000 were added.

Each was then pulsed (resistance = 1 kΩ, capacitance = 25 µFD and voltage = 800 V) using a Gene Pulsar Electroporator (BioRad). To each cuvette, 0.5 mL of 20% (w/v) sucrose was added and mixed. Each were then plated on BMMS agar plates. The BMMS plates consisted of 1.7 g L^{-1} yeast nitrogen base (w/o amino acids and ammonium sulfate), 5 g L^{-1} ammonium sulfate, 36 mL L^{-1} of 1 M citric acid, 126 mL L^{-1} of 1 M di-sodium hydrogen orthophosphate, 15 g L^{-1} bacteriological agar and 2% (w/v) sucrose. Plates were than incubated at 30 °C. After 4 days, approximately 20 to 100 colonies had grown on each plate.

3.3.4 Mutant Protein Expression

Protein expression of the yeast strains was checked using the techniques of rocket immunoelectrophoresis and quantitative SDS-PAGE. Yeast cells were inoculated into both YEPD and BMMS (made as BMMS agar but without the bacteriological agar) liquid media and the flasks were incubated at 30 °C with shaking for 4 days. Cultures were then spun at 3,000 rpm for 15 min and the supernatant was collected and assayed. Rocket immunoelectrophoresis was performed as described in section 2.2.3 and quantitative SDS-PAGE as described in Section 2.2.4. Expression levels were typically between 150 to 200 mg mutant albumin per mL of culture when grown in YEPD and 200 to 300 mg mutant albumin per mL of culture when grown in BMMS. All assayed colonies expressed the desired protein.

3.3.5 Trehalose Stocks for Fermentation

Yeast colonies were grown in BMMS liquid media to mid log phase to create trehalose stocks for fermentation. Following growth, 100 mL of each culture was spun at 3,000 rpm for 15 min, 75 mL of supernatant was removed and the pellet was resuspended. To the suspension, 25 mL of 40% D(+) trehalose was added. The suspension was then mixed, dispensed into 1 mL aliquots and stored at −80 °C ready for use.

3.4 Fermentation

To maximise yeast growth and protein production, it is first important to understand yeast growth. Four phases of growth are normally observed. First is the lag phase, the phase in which the cells adapt to their new environment. This would include induction of several metabolic pathways, before growth can start. This is followed by the exponential growth phase. Once the yeast has adapted to the new conditions the number of cells (or biomass) will increase exponentially until the carbon and/or nitrogen source has been consumed. After which, no growth occurs - this is the stationary phase. In *Saccharomyces cerevisiae* cultures following consumption of the sugar a phase, which looks like a stationary phase, is apparent. However, after a few hours the biomass increases exponentially again (although at a lower rate). This is due to the fact that during the first exponential phase, ethanol and acetate are produced and during the few hours of "stationary phase", the yeast adapts its metabolism by inducing enzymes that are needed for growth on these carbon sources. This type of growth is called diauxic growth. The final phase of yeast growth is the death phase where the cells lyse.

The production of the cloned protein is dependent on the biomass present within the fermenter. Yet if the cells have a vast supply of the carbon source, mass production of toxic fermentation products such as ethanol and acetate occurs. If the levels of these products become too high it is be sufficient to limit growth or even kill the cells. For this reason cells are grown under carbon limitation. This is achieved by having a batch medium and a feed medium. The batch medium is the initial source of nutrients for the cells after inoculation into the fermenter. When all of the sugars and ethanol in the batch medium have been used up, any acetate present is utilised by the cells. A

feed medium is then added at such a rate that the carbon source concentration is kept limiting. This therefore limits the rate of growth, preventing ethanol and acetate building up.

Yeast growth is also very sensitive to pH. As the yeast metabolise the carbon source they produce CO_2 as a by-product. CO_2 in solution forms a carbonic acid and hence the pH in the fermenter drops. The optimum pH for growth is 5.5; a solution of 17.5% ammonia is added when necessary to raise the pH. Conversely, during phases of acetate utilisation the pH increases. This is brought back down to optimum by the addition of 2 M H_2SO_4.

The batch medium was prepared in a 10 L fermenter as follows. An autoclaved 50% (w/v) carbon source, sucrose solution (200 mL) was added to 2560 mL distilled water and filtered using a 0.22 μm, SuporCap™ sterile capsule into the fermenter. A vitamin stock solution (50 mL, which contained 6 g L^{-1} calcium pantothenate, 10 g L^{-1} nicotinic acid, 6 g L^{-1} m-inositol, 1.5 g L^{-1} thiamine-HCl and 30 mg D-biotin) and 600 mL of a salt mixture (containing 28 g L^{-1} KOH, 40 g L^{-1} H_3PO_4, 7.17 g L^{-1} $MgSO_4{\cdot}7H_2O$, and 0.75 g L^{-1} $CaCl_2{\cdot}2H_2O$) were mixed with 40 mL of distilled water. The mixture was filtered using a 0.22 μm, Durapore filtration unit, then added to the fermenter. The filter was then flushed with 125 mL of distilled water. A trace element solution (45 mL, which contained 63.07 g L^{-1} H_2SO_4, 3 g L^{-1} $ZnSO_4{\cdot}7H_2O$, 10 g L^{-1} $FeSO_4{\cdot}7H_2O$, 2.42 g L^{-1} $MnSO_4{\cdot}H_2O$, 79 mg L^{-1} $CuSO_4{\cdot}5H_2O$, 0.5 g L^{-1} $Na_2MoO_4{\cdot}2H_2O$ and 0.56 g L^{-1} $CoCl_2{\cdot}6H_2O$) was filtered into the fermenter using the same device as used for the vitamin stock-salt solution mixture. The filter was flushed with 125 mL of distilled water.

Trehalose stocks of *S. cerevisiae* DXY1 cells, containing the respective mutated plasmids were inoculated into 100 mL of BMMS and grown to approximately 1 g cells per litre of culture. Approximately 60 mg of cells were then inoculated into the batch medium in the 10 L fermenter.

The feed medium was made as follows. Vitamin stock (120 mL), 250 mL of salt solution and 30 mL of distilled water were mixed and filtered using a 0.22 μm, Durapore filtration unit. The filter was then flushed with 100 mL of distilled water. Trace element solution (105 mL) was added to 295 mL of distilled water and filtered as was the vitamin stock-salt solution. The filter was again flushed with 100 mL of distilled water. The combined salts, vitamins and trace elements were added to 4 L of 62.5% (w/v) sucrose and made up to 5 L with sterile water. The feed was added following batch growth at a rate that increased to carbon limitation. The fermentation was completed when the entire feed medium had been added.

3.5 Purification

The *S. cerevisiae* DXY1 cell cultures, following fermentation were centrifuged at 3,000 rpm for 30 min. The supernatants were then removed and filtered. The recombinant protein was concentrated from the supernatant, using cation-exchange chromatography. An SP-sepharose Fast Flow cation-exchange column (column volume = 225 mL) was equilibrated with 4 column volumes of a 30 mM sodium acetate buffer, pH 5.5. The filtered supernatants were split into two batches of approximately 3 L. Both batches were purified for $His^{39}Phe$ and $Tyr^{84}Phe$ mutants but only one batch of $Cys^{34}Ala$ and $His^{67}Ala$ due to time restrictions. Sodium

octanoate was added (7.5 mL of a 2 M solution) to each batch and the pH was adjusted to 4.5 with acetic acid and the conductivity was adjusted to 5.5 mS cm^{-1} with deionised water before loading onto the column. After loading, the column was then washed with 8 column volumes of 50 mM acetate, 8 mM NaOH, pH 4.0 and 4 column volumes of a 27 mM sodium acetate buffer containing 2 M NaCl, pH 4.0. A third wash was carried out with 10 column volumes of the equilibration buffer. Finally the column was eluted with 2 column volumes of 85 mM sodium acetate containing 5 mM octanoic acid, pH 5.5.

The recombinant albumin content of the filtered supernatant and the eluent were determined by high-performance liquid chromatography (as decribed in Section 2.1.2). The average albumin concentrations of the supernatants and eluents were 2.5 g L^{-1} and 5.5 g L^{-1}, respectively. The mean purification yield was approximately 76%.

SP-sepharose Fast Flow eluents were then further purified by anion-exchange chromatography on a DEAE fast flow column (column volume = 167 mL). The column was equilibrated with 15 column volumes of 30 mM acetate, 27 mM NaOH, pH 5.5. The conductivity of SP-sepharose eluents was adjusted to 3.0 mS cm^{-1} with deionised water before loading onto the column. After loading the column was then washed with 5 column volumes of 15.7 mM $K_2B_4O_7 \cdot 4H_2O$, pH 9.2. The column was eluted with 0.75 column volumes of 85 mM acetate, 110 mM $K_2B_4O_7 \cdot 4H_2O$, pH 9.4. HPLC revealed the average albumin concentrations of the eluents to be 11.4 g L^{-1}. The mean purification yield was calculated to be approximately 107%. However, the HPLC method for albumin quantification measures monomer concentration only, and

during this purification step some dimeric albumin is thought converts back to monomer.

DEAE eluents were then purified further by affinity chromatography on a Delta Blue Agarose column (column volume = 423 mL). Delta Blue matrix was supplied by ACL. The column was equilibrated with 2 column volumes of a 250 mM ammonium acetate buffer, pH 8.9 before loading the DEAE eluent. After loading, the column was then washed with 5 column volumes of the equilibration buffer. The column was eluted with 2 column volumes of 50 mM phosphate buffer containing 2 M NaCl, pH 6.9. HPLC revealed the average albumin concentrations of the eluents to be 6.0 g L^{-1}. The mean purification yield was approximately 76%.

Delta Blue eluents were then concentrated using a 10 kD MWCO Pall Filtron LU Centramate filter connected to a peristaltic pump. Overall it was determined that 6.38 g of Cys^{34}Ala, 6.41 g of His^{39}Leu, 7.61 g of Tyr^{84}Phe and 4.25 g of His^{67}Ala albumins were recovered. A sample of concentrated solution from each of the purified products was diluted to 5 mg mL^{-1} and 10 μL of each were applied to an SDS-PAGE gel. Gels were made and run using the standard method of Laemmli [16]. The gels were stained with both coomassie blue stain and silver stain. A gel showing the His^{39}Leu and Tyr^{84}Phe purification products can be seen in Figure 3.7. The gels revealed no other proteins to be present at the 1% level (therefore proteins are approximately 99% pure).
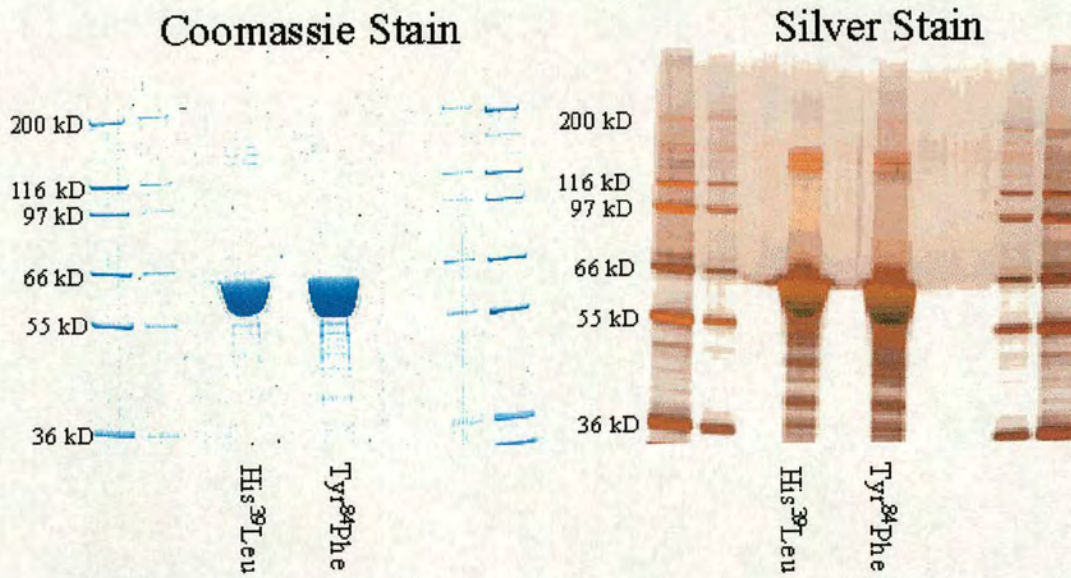
Figure 3.7 SDS-PAGE of purified products. The coomassie stained gel shows no other significant proteins to be present at the 1% level. The silver stained gel shows the presence of trace impurities within the purified products. The bands around 135 kD in the silver stained gel could be albumin dimer.

3.6 Characterisation of Mutant Albumins

3.6.1 Mass Spectrometry of Purified Proteins

In order to check that the proteins produced were of the expected molecular weight, ESI-mass spectrometry was carried out. Samples were diluted to 1 mg mL^{-1} and were prepared and analysed as described in Section 2.5. The resultant spectra are shown in Figure 3.8. The resultant masses were very close to the theoretical values and confirmed the mutant proteins to be the expected size in each case.
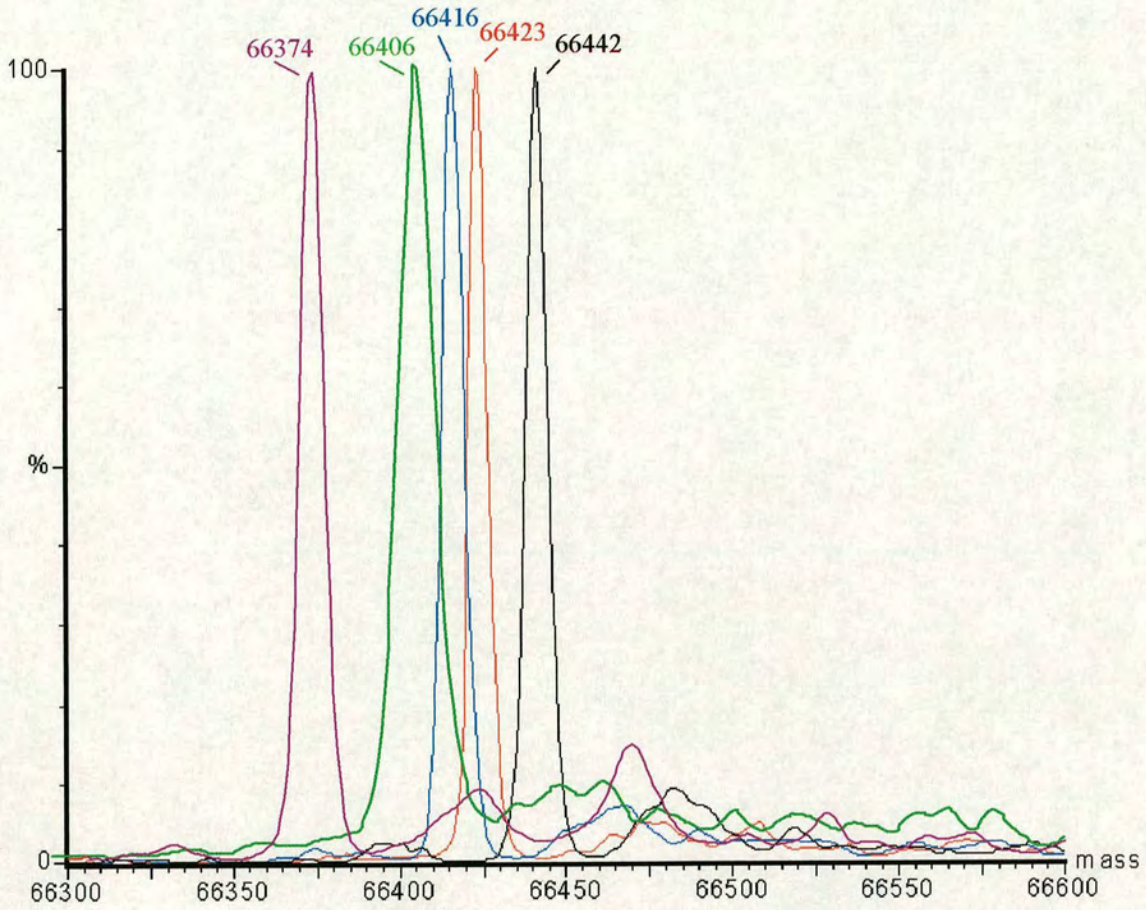
Figure 3.8 ESI-mass spectra of purified products. The spectra show measured masses of native and mutant proteins. Theoretical masses for native (black), Cys^{34}Ala (green), His^{39}Leu (blue), His^{67}Ala (purple) and Tyr^{84}Phe (red) albumins are 66438, 66406, 66418, 66372 and 66422, respectively.

3.6.2 Peptide Mapping of Mutant Albumins

In order to ensure the mutations had occurred at the correct site, peptide mapping of His^{39}Leu and Tyr^{84}Phe albumins was attempted.

3.6.2.1 Tryptic Digestion

Mutant albumins (0.5 mg of each) and reference standard in 100 μL were added to 100 μL of 6 M guanidine-HCl in 2 mM EDTA, 0.5 M Tris-HCl, pH 8.0. The samples were reduced by the addition of 5 μL of 100 mM dithiothreitol then were mixed and incubated for 1 hour at 37 °C. The samples were diluted by the addition of 200 μL of water. The digestion commenced with the addition of 10 μL of 1 mg mL^{-1} trypsin. The samples were incubated overnight at 37 °C in a shaking incubator. The following morning and then in the afternoon another 10 μL of 1 mg mL^{-1} trypsin was added and incubation of the samples was continued at 37 °C for 48 hours. They were centrifuged for 5 min, 400 μL of the supernatant was removed, transferred to a fresh tube, mixed with 400 μL of 0.1% TFA in water, mixed again and then centrifuged for a further 5 minutes.

3.6.2.2 HPLC Separation of Tryptic Fragments

The digests were chromatographed on a Waters HPLC system. The reversed-phase column was a Vydac C18 (dimensions 4.6 x 250 mm). Detection was performed at 214 nm. Two solvents were used solvent A (0.1% TFA) and solvent B (0.085% TFA in 70% acetonitrile). Initially sample was run with a linear gradient of 95% solvent A and 5% solvent B to 100% solvent B for 2 hours. Injection volume was 100 μL (equivalent to 65 μg of digest) and flow rate was 0.5 mL min^{-1}.

Some peak differences between the control and Tyr[84]Phe albumin were noted, notably the appearance of a novel peak at 53 min and the absence of a peak at 57 min in the Tyr[84]Phe rHA as shown in Figure 3.9. The novel peak was collected manually from the eluent of the HPLC detector for sequencing. The other mutant, His[39]Leu, was a little more elusive. About 6 peak changes were found but it is unknown which is caused by the residue change. The actual tryptic peptide containing the His[39]Leu mutation is 21 residues long, which means it elutes quite late in a crowded region of the chromatogram.
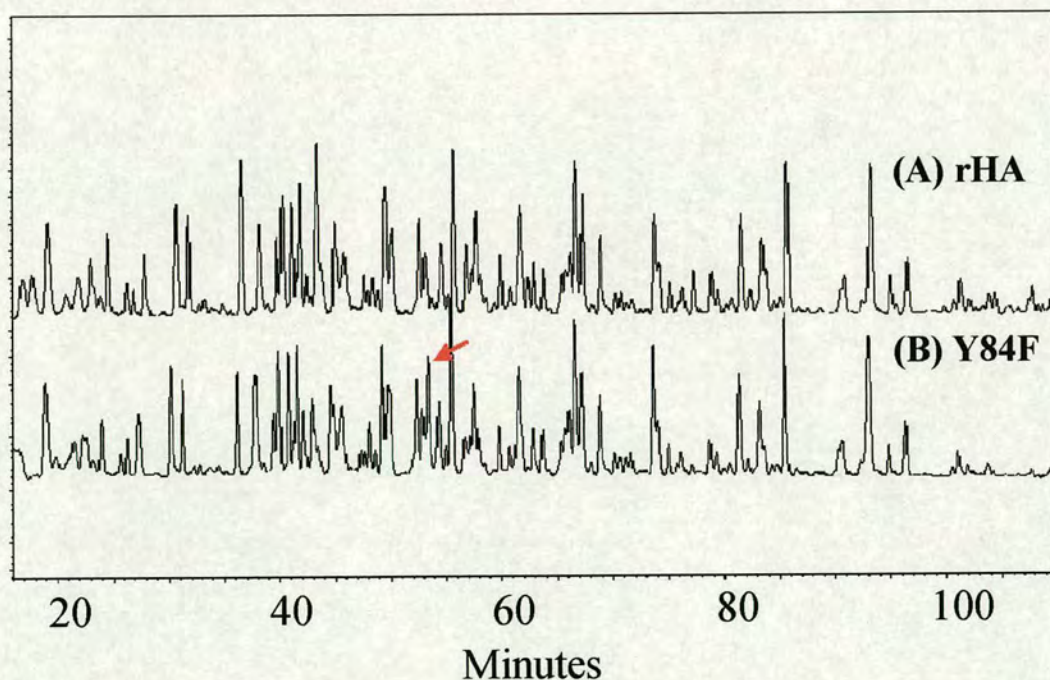
Figure 3.9 Reverse-phase HPLC trace of the (A) control rHA and (B) Y84F mutant digestion products. A new peak can be seen in the mutant digest, indicated by the arrow.

3.6.2.3 Amino-Terminal Sequence Analysis

Sequence analysis was performed on an Applied Biosystems 494 Procise sequencer.

A 60 μL aliquot of a collected peak was loaded onto a precycled filter and sequenced

using a pulsed-liquid cycle. The 53 min peptide, thought to correspond to the

Tyr[84]Phe mutation, was sequenced and the mutation was confirmed (Figure 3.10).

Two peaks suspected to be His[39]Leu mutant peptides were sequenced, but were found

to be of mixed sequences. This means that there was more than one peptide in the

peak, and a number of missed cleavage sites were also found, where the trypsin failed

to cleave the protein at a Lys or Arg residue.
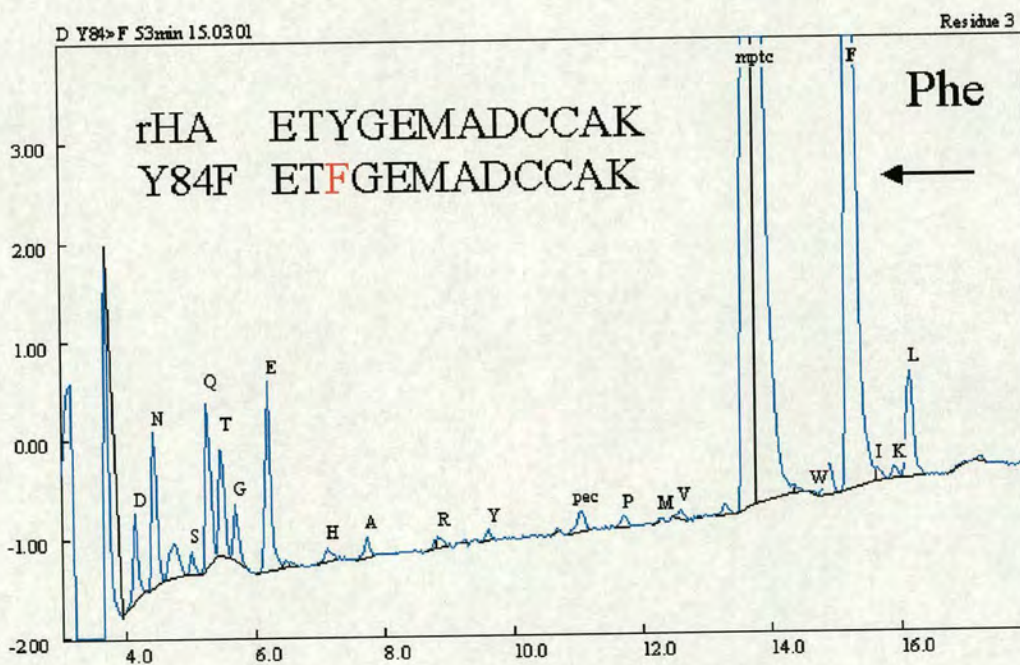
Figure 3.10 Sequence analysis of the novel peptide found in the Y84F mutant

digest. The elevated level of phenylalanine following cleavage of the third residue

confirms the mutation.

3.6.3 Circular Dichroism

Circular dichroism (CD) spectra were recorded of native, $Cys^{34}Ala$, $His^{39}Leu$, $His^{67}Ala$ and $Tyr^{84}Phe$ albumins (Figure 3.11) and secondary structure estimations were made (Table 3.1) as described in Section 2.6.

The secondary structure estimations calculated from the CD spectra reveal little difference in secondary structure between the native and mutated proteins.

Protein	α-Helix (%)	Anti-parallel β-Sheet (%)	Parallel β-Sheet (%)	Turns (%)	Other (%)	Total (%)
Native	44.3	8.3	7.7	14.2	25.4	99.8
	56.1	2.4	5.7	13.8	22.7	100.7
C34A	54.1	3.7	6.4	13.8	23.3	101.4
H39L	45.2	7.9	7.4	14.2	25.5	100.3
H67A	60.3	3.4	4.4	14.7	17.4	99.7
Y84F	45.8	7.9	7.7	14.1	25.8	101.3

Table 3.1 Secondary structure estimations of native and mutant albumins. The native protein structure content was calculated twice, once with $His^{39}Leu$ and $Tyr^{84}Phe$ mutants (top value) and with $Cys^{34}Ala$ and $His^{67}Ala$ mutants (bottom value).

Figure 3.11 Circular dichroism spectra of native and mutant albumins. In (a) circular dichroism spectra of wild type (pink), His[39]Leu (blue) and Tyr[84]Phe (red) albumin are shown. In (b) circular dichroism spectra of wild type (pink), Cys[34]Ala (blue) and His[67]Ala (red) albumin are shown.

References: Chapter 3

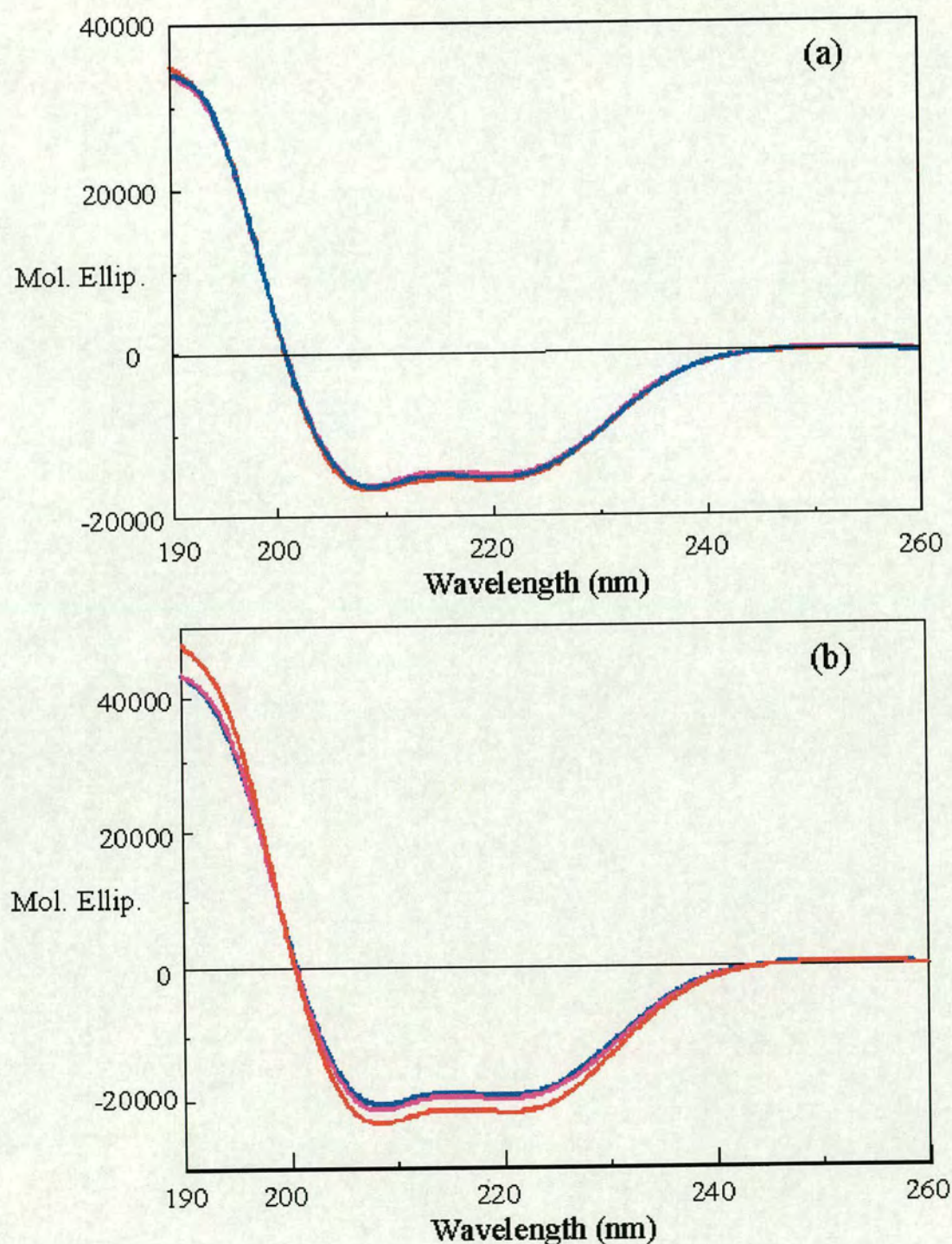1. Pedersen, A. O. and Jacobsen, J. (1980). Reactivity of the thiol group in human and bovine albumin at pH 3-9, as measured by exchange with 2,2'-dithiodipyridine. *Eur. J. Biochem.*, **106:** 291-295.

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

3. Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999). Crystal structure of human serum albumin at 2.5 Å resolution. *Prot. Eng.*, **12:** 439-446.

4. Minghetti, P. P., Ruffner, D. E., Kuang, W.-J., Dennison, O. E., Hawkins, J. W., Beattie, W. G. and Dugaiczyk, A. (1986). Molecular structure of the human albumin gene is revealed by nucleotide sequence within q11-22 of chromosome 4. *J. Biol. Chem.*, **261:** 6747-6757.

5. Kunkel, T. A. (1985). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA*, **82:** 488-492.

6. Hochhauser, S. J. and Weiss, B. (1978). *Escherichia coli* mutants deficient in deoxyuridine triphosphatase. *J. Bacteriol.* **134:** 157-166.

7. Konrad, E. B. and Lehman, I. R. (1975). Novel mutants of *Escherichia coli* that accumulate very small DNA replicative intermediates. *Proc. Natl. Acad. Sci. USA,* **72:** 2150-2154.

8. Sagher, D. and Strauss, B. (1983). Insertion of nucleotides opposite apurinic/apyrimidinic sites in deoxyribonucleic acid during *in vitro* synthesis: uniqueness of adenine nucleotides. *Biochemistry* **22:** 4518-4526.

9.   Shlomai, J. and Kornberg, A. (1978). Deoxyuridine triphosphatase of *Escherichia coli*. Purification, properties, and use as a reagent to reduce uracil incorporation into DNA. *J. Biol. Chem.* **253:** 3305-3312.

10.   Lindahl, T. (1982). DNA repair enzymes. *Annu. Rev. Biochem.* **51:** 61-87.

11.   Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin, R. C. and Gelbart, W. M. (1996). *An Introduction to Genetic Analysis, 6th Ed*. W. H. Freeman and Co.

12.   Stryer, L. (1995). *Biochemistry, 4th Ed*. W. H. Freeman and Co., New York.

13.   Nelson, M. and McClelland, M. (1992). Use of DNA methyltransferase/endonuclease enzyme combinations for megabase mapping of chromosomes. *Methods Enzymol.* **216:** 279-303.

14.   Hattman, S., Brooks, J. E. and Masurekar, M. (1978). Sequence specificity of the P1 modification methylase (M.*Eco* P1) and the DNA methylase (M.*Eco* dam) controlled by the *Escherichia coli dam* gene. *J. Mol. Biol.,***126:** 367-380.

15.   Kerry-Williams, S. M., Gilbert, S. C., Evans, L. R. and Ballance, D. J. (1997). Disruption of the *Saccharomyces cerivisiae YAP3* gene reduces the proteolytic degradation of secreted recombinant human albumin. *Yeast*, **14:** 161-169.

16.   Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227:** 680-685.

Chapter 4

Reactions at Cys^{34}

4.1 Introduction

As discussed in Section 1.3.4, Cys^{34} is the binding site for a wide variety of biologically and clinically important small molecules and is also responsible for the largest fraction of free sulfhydryl groups in blood serum [1]. Therefore, the kinetic and thermodynamic properties of the free thiol group are of great physiological significance.

Although a number of crystal structures of human albumin have been solved (as summarised in Table 1.3). These do not provide much information as to the dynamics of Cys^{34} and its surrounding residues in solution. Over the years, a wide range of techniques have been used to probe the microenvironment surrounding Cys^{34}, some examples are summarised in Table 4.1. Such studies have provided invaluable data with regard to the Cys^{34} site.

In solution, Cys^{34} has been shown to be accessible to Hg^{2+} [2] but to lie in a hydrophobic sterically hindered pocket, 9.5 Å deep [3-5]. It is proposed that upon ligand binding or oxidation, a conformational change is invoked with Cys^{34} moving from a buried to a solvent-exposed state. The conformational change appears to be coupled to a movement of the His^{3} residue, which could be observed using 2-dimensional ^{1}H NMR [6]. Interestingly, an *in vitro* reaction between albumin and

auranofin appeared to follow first order kinetics, with the reaction rate being independent of auranofin concentration [7]. This was proposed to be due to an altered albumin conformation where-by domain IA of the protein is open, increasing Cys[34] accessiblity and allowing gold binding. It was concluded that Cys[34] is on the surface of the protein following the binding of thiol-specific fluorescent probe, acrylodan [8] and fatty acid binding to the protein affects this site [8,9]. Perhaps the most interesting study is that of Kratz *et al.* [9]. They investigated binding of derivatives of the anti-cancer agent, doxorubicin hydrazone containing differing lengths of aliphatic spacers (2, 3, 5 and 7 carbons) as shown in Figure 4.1. Cys[34] reacted slower as the chain-length increased in the absence of fatty acid (2 > 3 > 5 > 7). In the presence of myristate the 5-carbon derivative reacted fastest (5 > 3 > 2 > 7). This shows that the binding of fatty acid at other sites on the molecule can alter the binding specificity of the Cys[34] site.
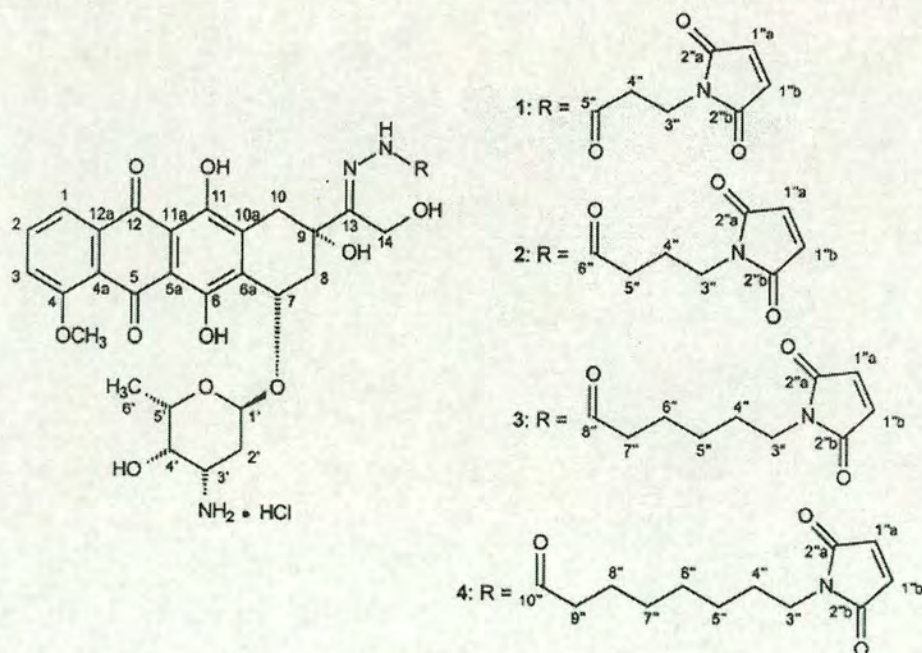

Figure 4.1 General structure of doxorubicin hydrazone derivatives. Structures 1-4 contain 2, 3, 5, and 7 carbon atoms respectively as aliphatic spacers. Modified from [9].

Technique / Method	Findings	Authors
Hg^{2+} binding	The free thiol of albumin is accessible to Hg^{2+} and can give dimeric albumins linked through the free thiol group.	Kay and Edsall, 1956 [2]
Absorbance and optical rotary dispersion (ORD) measurements	The free thiol group sits at a border between a polar helical segment and a hydrophobic area with two tyrosyl groups (one buried, one exposed) nearby.	Ohkubo, 1969 [3]
EPR spin-labelling	Using nitroxide compounds of differing chain-length, it was concluded that the free thiol lies in a crevice 9.5 Å deep.	Cornell *et al.*, 1981 [4]
Aromatic and heterocyclic disulfide binding	Compounds containing 5-membered heterocyclic rings react more readily than 6-membered aromatic rings, suggesting that Cys^{34} lies in a sterically restricted environment.	Gosselet *et al.*, 1990 [5]
^1H NMR	Binding at Cys^{34} envokes a change in His^3 proton resonances. From this it was concluded that upon binding, Cys^{34} moves from a buried to an exposed state.	Christodoulou *et al.*, 1995 [6]
Penefsky spin column chromatography and ^{31}P-NMR	The reaction between the anti-arthritic gold complex, auranofin is first order with respect to albumin. It was concluded that the rate-determining step involves a conformational change of the protein.	Roberts *et al.* 1996 [7]
Cys^{34}-specific fluorescent probe	Following the binding of acrylodan, Cys^{34} was found to be present on the surface of human albumin. Oleate binding affected the fluorescence of the acrylodan-albumin complex.	Narazaki *et al.*, 1997 [8]
Doxorubicin hydrazone derivatives as probes	Binding of doxorubicin hydrazone derivatives with differing lengths of aliphatic spacers were investigated (2, 3, 5 and 7 carbons). Cys^{34} reacted slower as the chain-length increased in the absence of fatty acid (2 > 3 > 5 > 7). In the presence of myristate the 5-carbon derivative reacted fastest (5 > 3 > 2 > 7), showing that fatty acid-binding affects the Cys^{34} site.	Kratz *et al.*, 2002 [9]

Table 4.1   Summary of some studies of the dynamic properties of the free thiol group of Cys^{34}.

4.2 Reactions with DTNB

4.2.1 Ionisation Behaviour of Cys[34] thiol

As previously mentioned, the kinetic and thermodynamic properties of the free thiol group are of great physiological significance. Thiol interchange can occur between the reduced Cys[34] thiol and an oxidised disulfide. The compounds 5,5'-dithio-bis(2-nitrobenzoic acid) (also known as Ellman's reagent [10] or DTNB) and 2,2'-dithiodipyridine (DTDP) are commonly used for protein thiol determinations and are useful for studying kinetics at the Cys[34] site, as the reactions can be quantified spectrophotometrically [11,12]. The structures of DTNB and DTDP are shown in Figure 4.2.
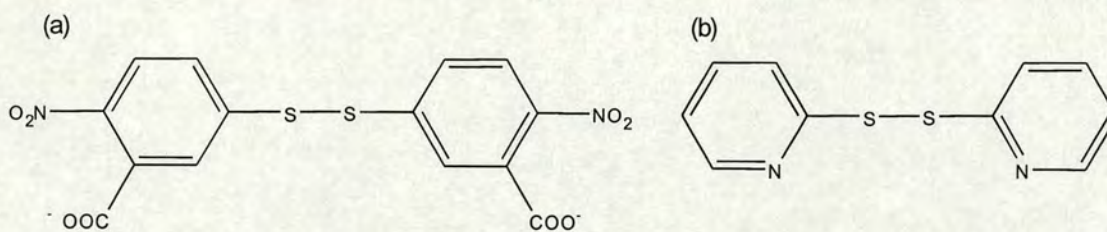
Figure 4.2 Structures of (a) DTNB and (b) DTDP.

During disulfide interchange (Equation 4.1), the formation of RS-SR' must be more stable (and therefore its formation more favourable) than RS-SR. This reaction is thermodynamically more favourable for disulfides with higher redox (reduction) potentials than albumin (Equation 4.2). The respective redox potentials at pH 7 ($E^{0'}$) of cystine/cysteine and GSSG/GSH are -245 and -262 mV, respectively, relative to the standard hydrogen electrode [13].

$$RS\text{-}SR + R\text{'}SH \longrightarrow RS\text{-}SR\text{'} + RSH \qquad\qquad [4.1]$$
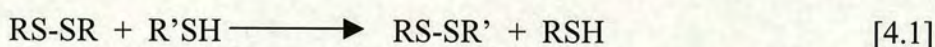
$$\Delta G = -nFE^0 \qquad\qquad [4.2]$$

where ΔG is the free energy change, n is the number of electrons, F is Faraday's constant (23.062 kcal) and E^0 is the redox potential.

A number of studies have used such disulfide interchange reactions to investigate the ionisation of the thiol of Cys[34] and have estimated pK_a values [12,14-16]. The results are summarised in Table 4.2.

Pedersen and Jacobsen investigated the kinetics for exchange between an aromatic disulfide (2,2'-dithiodipyridine, DTDP, Figure 4.2b) and the thiol group of human and bovine albumin over the pH range of 3-9 [12]. They found that the reactivity of Cys[34] in both albumins rapidly increases between pH 5 and 8 and suggested that Cys[34] of human albumin is characterised by an extremely low pK_a of approximately 7 compared to a pK_a value of 8-10, which is typical for a free cysteine side-chain [12]. Bovine and human albumins were found to have a similar reactivity at physiological pH, although the bovine protein reacted ~10 times faster than the human protein at pH >9.

A similar approach to Pedersen and Jacobsen [12] was followed by both Wilson *et al.* [15] and Svenson and Carlsson [16]. They investigated the kinetics for exchange between bovine albumin and the aromatic disulfides DTNB and DTDP, respectively, over the pH ranges 1 – 8 and 7 - 9. Both studies found that reactivity is dependent on

pH and both estimated pK_a values based upon their reactivity data of around 8.5. A value that is fairly typical for cysteinyl sulfur [16].

Albumin	Estimated thiol pK_a	Method used	Authors
Human and Bovine	~ 7 (both)	Reaction with 2,2'-dithiodipyridine over the pH range 3 – 9	Pedersen and Jacobsen, 1980 [12]
Bovine	< 5	Potentiometric difference titration with methyl methanethiosulfonate	Lewis *et al.*, 1980 [14]
Bovine	~ 8.5	Reaction with 5,5'-dithiobis(2-nitrobenzoic acid) over the pH range 7 – 9	Wilson *et al.*, 1980 [15]
Bovine	~ 8.5	Reaction with 2,2'-dithiodipyridine over the pH range 1 - 8	Svenson and Carlsson, 1975 [16]

Table 4.2 Reported pK_a values for albumin free thiol.

Lewis *et al.* speculated that the pK_a of bovine Cys[34] thiol is even lower, and is less than 5 [14]. They measured the ionisation behaviour of albumin Cys[34] (as well as glutathione, cysteine, 2-mercaptoethanol, 3-mercaptopropionic acid, 2-mercaptoethylamine, *cis*-2-mercaptocyclobutamine, 2-aminothiophenol and 5-mercapto-2-nitrobenzoic acid) by a potentiometric difference titration method they had previously used to measure the pK_a of the Cys[25] thiol of papain [17]. In order to determine the pK_a, they determined the protons that were released during the reaction of methyl methanethiosulfonate (MMTS) with the thiol group (Equation 4.3).

$$RSH + CH_3SO_2SCH_3 \longrightarrow RSSCH_3 + CH_3SO_2^- + H^+ \qquad [4.3]$$
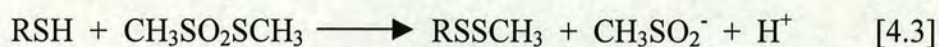
The pH dependence of the difference in proton content between an unmodifed albumin and the methylthio derivative was used to characterise the ionisation behaviour of the thiol. The protons released were measured by back-titrating with KOH using a pH electrode. They also found that indicator dyes could be successfully used for this purpose with the smaller thiol-containing compounds. The smaller thiol-containing compounds yielded sigmoidal curves with values relating exclusively to the ionisation state of the thiol group in each case. The mid-points of each agreed well with previously determined pK_a values in the literature. With bovine albumin their findings were not so straightforward. The best fit of their titration curves suggested that the ionisation of the Cys^{34} thiol also depends on three other groups, two with pK_a values of 7.9 and another with a pK_a >10 and that Cys^{34} itself has a pK_a <5. They could not however, eliminate the possibility that the changes in Cys^{34} ionisation they observed were due to pH-induced conformational changes, but do provide evidence that a decrease in rate of reaction may well reflect a decreased reactivity of the thiolate anion caused by the protonation of other groups. In light of this, it cannot therefore be assumed that kinetic data alone can be used to study the ionisation of thiol groups in proteins.

If the albumin thiol does have an abnoramlly low pK_a then it will be predominantly in the S^- form at physiological pH. The residues His^{39} and Tyr^{84} are located very close to Cys^{34} (Figure 1.5) in the X-ray crystal structure of albumin in the non-liganded state and therefore may play a role in stabilising the S^- form of the thiol. Interestingly,

studies have also shown that the binding of fatty acid molecules to albumin can increase the rate of Cys[34] oxidation [18,19].

The aims of this section are to determine whether His[39]Leu and Tyr[84]Phe mutations have an effect on Cys[34] reactivity. Previous studies have also shown fatty-acid binding to have an effect on Cys[34] reactivity [8,9], whilst several crystal structures show an increase in Cys[34] accessibility in the presence of bound fatty acid [20-22] (Figure 4.3). Recombinant human albumin in the presence of 10 mol equiv of octanoate was also examined to determine whether elevated fatty acid binding affects the kinetics of reaction of DTNB at this site. The rates of reaction of the dithiol, 5,5'-dithio-bis(2-nitrobenzoic acid) with rHA, high octanoate rHA (10 mol mol^{-1}), His[39]Leu and Tyr[84]Phe albumins were determined over a range of pH values between 6 and 10. The reaction is shown in Figure 4.4.
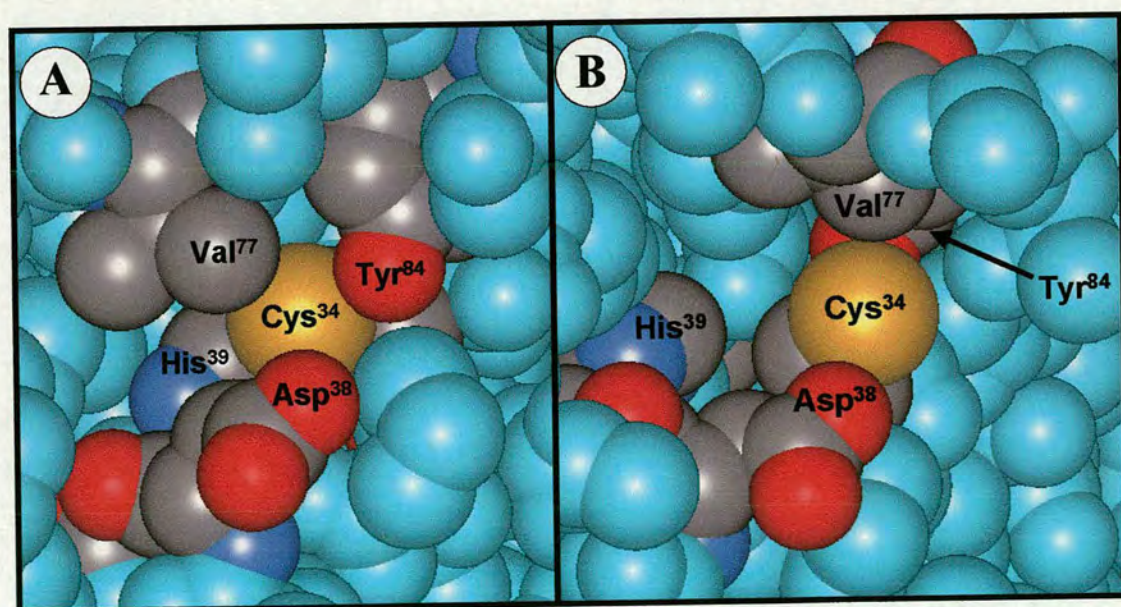
Figure 4.3 Space-filling representation of the environment around Cys[34] in (A) unliganded albumin and (B) albumin with 5 myristates bound. Drawn from PDB 1AO6 [23] and 1BJ5 [20], respectively.
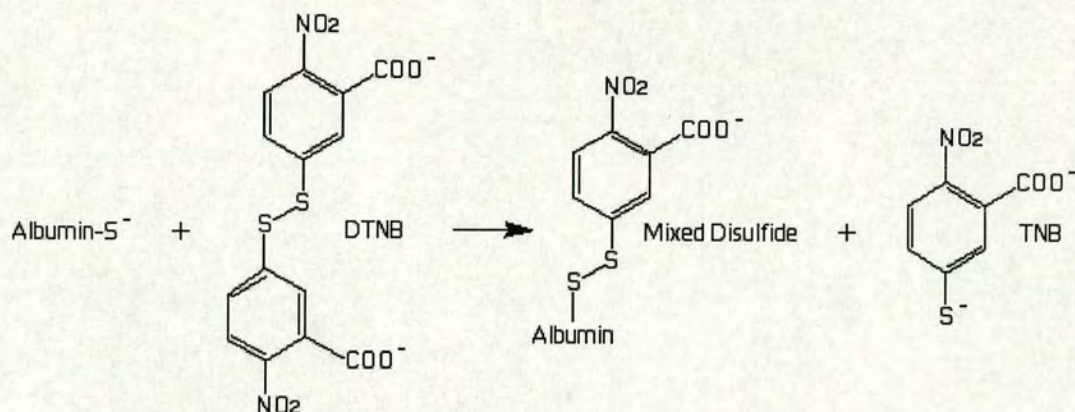
Figure 4.4 Reaction of albumin thiol with DTNB. DTNB reacts with the free thiol of albumin to form a mixed disulfide and 5-thio(2-nitrobenzoic acid) (TNB).

During a study of this kind it is important to be aware of the pK_a values of the (de)protonatable groups in the disulfide compounds. This is because the reaction rate is likely to be dependent on the overall charge on the molecule. Brocklehurst and Little have shown that the acid dissociation constants of DTNB are 1.57 and 2.15 [24] and the pK_a of the thiol of TNB is 4.50 [15], therefore the carboxylate groups will all be deprotonated and only the S⁻ form of the TNB product will be present over the pH range studied.

4.2.2 Experimental

Mass Spectrometry

rHA (15 µM) was incubated with 40 mol equiv of DTNB in 100 mM potassium phosphate, pH 8.0 for 1 hour. A mass spectrum of the product was recorded as detailed in Section 2.5.

Reactions of albumins with DTNB

Reactions mixtures were set up as follows: 1350 μL of respective buffer, 1350 μL of 50 mM KCl, 150 μL of 15 mM DTNB (dissolved in methanol). The following buffers were used: 0.2 M potassium phosphate (pH range 6.0-8.0), 0.2 M Tris-HCl (pH range 8.0-9.0) and 0.2 M CAPS, 3-(cyclohexylamino)-1-propanesulfonic acid (pH 9.0-10.0) range. The ionic strengths of the buffers were kept constant and were adjusted to that of phosphate (0.466 M) with KCl when necessary. The reaction mixture was equilibrated in the cuvette at 10 °C whilst mixing. Following equilibration, the reaction was started by addition of 150 μL of 150 μM albumin, at 10 °C. The reactions were performed at 10 °C in order to slow down the reaction rate as it was found that the rates at room temperature were extremely fast and difficult to measure accurately. All reactions were carried out in 1 cm pathlength cells with stirring and were followed by the change in absorbance at 412 nm on a Cary 300 Scan spectrophotometer fitted with a Peltier dual cell temperature controller. The thiol contents of the proteins were calculated as described in Section 2.3 and were 0.68, 0.80, 0.63 and 0.63 mol mol^{-1} for rHA, fatty rHA, His^{39}Leu and Tyr^{84}Phe albumins, respectively.

4.2.3 Results and Discussion

Mass spectometric analysis, following the reaction of rHA with DTNB, detected a peak of 66643 mass units (Figure 4.5). This is consistent with the formation of the albumin-TNB mixed disulfide, which has a theoretical mass of 66645 mass units.
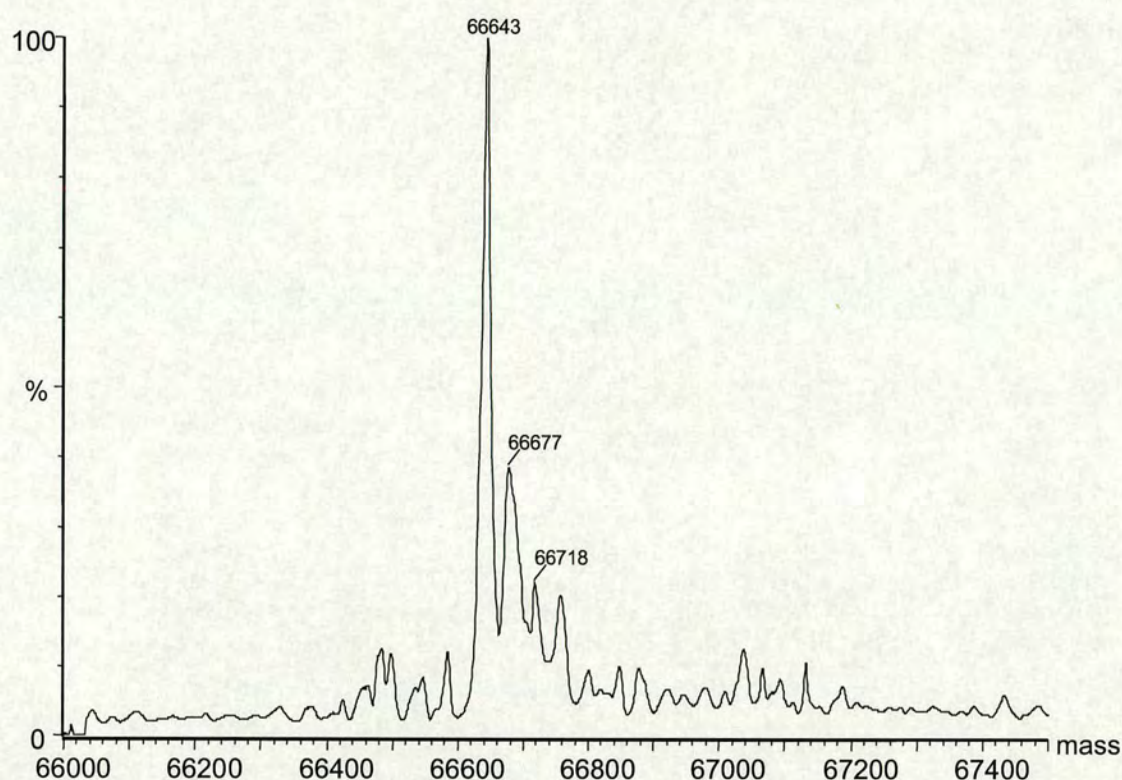
Figure 4.5   Mass spectrum of rHA following incubation with 40 mol equiv of

DTNB.   Reaction was carried out in 100 mM potassium phosphate, pH 8.0.   The

theoretical mass of the albumin-TNB mixed disulfide is 66645 mass units.

The reactions of DTNB with rHA, rHA in the presence of 10 mol mol^{-1} octanoate and

with His^{39}Leu and Tyr^{84}Phe albumin mutants were carried out at a range of pH values

(pH 6.0-10.2) at 10 °C.

The reaction between bovine albumin and DTNB has previously been shown to follow

second order kinetics [15].   The general second-order rate law for 1:1 stoichiometry is

the overall rate equation for a reaction that is first-order in I and first-order also in II.

$$\frac{-\mathrm{d}[I]}{\mathrm{dt}} = k_2[I][II] \qquad\qquad\qquad [4.4]$$

where [I] = albumin thiol concentration, [II] = DTNB concentration, k_2 = second order

rate constant and t = time.

Equation 4.4 approaches a limiting case if reactant II is in excess, then

$$\frac{-d[I]}{dt} = k_1[I] \qquad\qquad\qquad [4.5]$$

where k_1 = pseudo-first order rate constant (= k_2[II]). Integration yields:

$$k_1 = \frac{2.303}{t} \log \frac{[I]^0}{[I]} \qquad\qquad\qquad [4.6]$$

where $[I]^0$ = initial albumin thiol concentration.

A plot of $\log([I]^0/[I])$ against time gives a straight line with a slope equal to k_1. In all

cases the albumins were reacted with high excess of DTNB (100:1), and were treated

with the assumption of pseudo-first order kinetics. Appendix 1 shows the increase in

absorbance at 412 nm with time and a plot of $\log([I]^0/[I])$ vs time.

The pseudo-first order rate constants for the reaction between DTNB and the wildtype,

fatty-acid-bound and mutant albumins increased with pH and are tabulated in Table

4.3 and plotted in Figure 4.6.

pH	rate constant (min^{-1})			
	rHA	high octanoate rHA	His[39]Leu	Tyr[84]Phe
6.0	8.0×10^{-4}	1.0×10^{-3}	0	2.9×10^{-3}
6.2	9.0×10^{-4}	1.2×10^{-3}	0	4.0×10^{-3}
6.6	2.5×10^{-3}	3.0×10^{-3}	5.0×10^{-4}	9.4×10^{-3}
7.0	5.5×10^{-3}	6.0×10^{-3}	1.8×10^{-3}	2.2×10^{-2}
7.2	7.8×10^{-3}	9.7×10^{-3}	3.7×10^{-3}	2.8×10^{-2}
7.5	1.1×10^{-2}	1.5×10^{-2}	7.4×10^{-3}	5.0×10^{-2}
8.0	2.2×10^{-2}	2.7×10^{-2}	1.6×10^{-2}	1.5×10^{-1}
8.4	5.0×10^{-2}	5.8×10^{-2}	7.4×10^{-2}	1.0
8.8	5.2×10^{-2}	5.8×10^{-2}	1.4×10^{-1}	6.5
9.1	5.6×10^{-2}	5.9×10^{-2}	1.7×10^{-1}	8.2
9.5	6.1×10^{-2}	5.3×10^{-2}	1.9×10^{-1}	9.7
10.2	6.4×10^{-2}	9.2×10^{-2}	1.9×10^{-1}	10.8

Table 4.3   Effect of pH on the pseudo-first order rate constant for the reaction between DTNB and rHA, high octanoate rHA and mutant albumins.
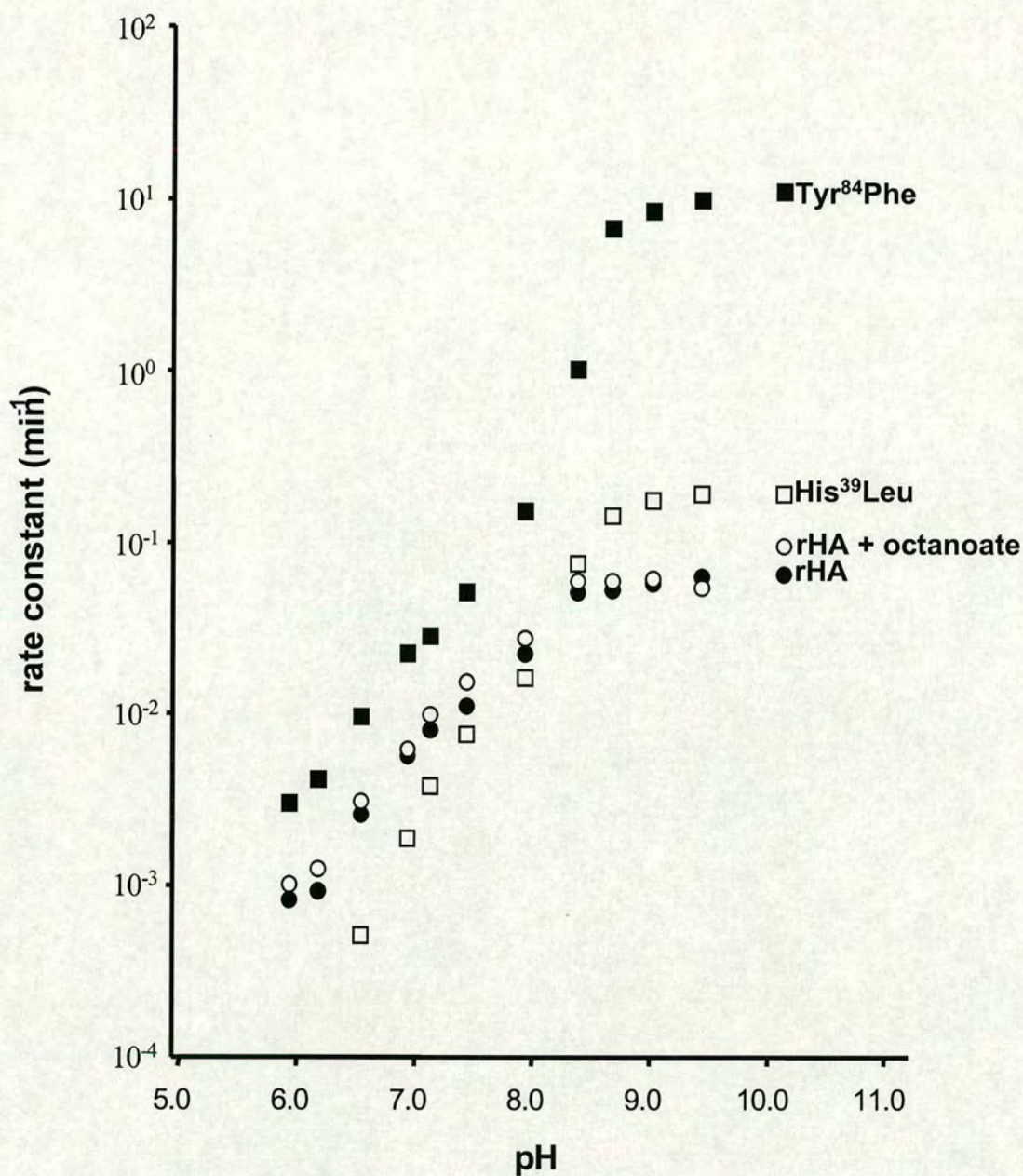
Figure 4.6 Effect of pH on the pseudo-first order rate constant for the reaction between DTNB and ● rHA (low octanoate), ○ high octanoate rHA, ■ Tyr[84]Phe and □ His[39]Leu. Non-logarithmic plots of the data are shown in Appendix 2.

The rate constants for the reaction of wild-type albumin with DTNB followed a similar trend to those observed by Pedersen and Jacobsen [12] in their kinetic study of the reaction of albumin with the uncharged DTDP molecule. The reactivity of Cys^{34} appears to rise steadily for all albumins between pH 6 and 8. The Tyr^{84}Phe mutation was found to have a marked effect on the reactivity of Cys^{34}. Over this pH range the rate of reaction is approximately 4 times greater for the Tyr^{84}Phe mutant than for the wild-type protein. At pH 6, the His^{39}Leu mutant reacts with DTNB much slower than the wild-type protein but has a similar rate constant at pH 8. Between pH 8 and 8.5 there is a larger increase in reactivity for each of the albumins. This is most pronounced for the Tyr^{84}Phe mutant. Cys^{34} reactivity in wild-type albumin appears to reach a maximum at around pH 8.4, whilst significant increases in Cys^{34} reactivity in the His^{39}Leu and Tyr^{84}Phe mutants are observed at pH values up to 9. At pH 10.2 the reaction is ~ 3 times faster with the His^{39}Leu mutant and ~ 170 times faster with the Tyr^{84}Phe mutant with the compared with that of the wild-type protein.

The binding of excess octanoate to albumin appeared to have very little effect on Cys^{34} reactivity. The rate constants for octanoate-bound albumin were generally slightly higher than that of rHA. This however, could be attributed to the higher thiol content of the octanoate-bound protein used in these experiments (0.80 mol mol^{-1} compared with 0.68 mol mol^{-1}). The rHA used in this study was dialysed prior to use and may still have contained as much as 4 mol mol^{-1} of octanoate, although it is not thought that this is a high enough level of fatty acid to affect the conformation of domain I dramatically. Fatty acid molecules bound to dialysed rHA were found not to perturb metal binding at a site on the interface of domains I and II, although 10 mol mol^{-1} octanoate did (see Section 5.2.7). Conformation changes affecting both the metal

binding site (Section 5.2.7) and the Cys^{34} site appear to be caused by fatty acid binding at so-called site 2 [20].

The rate of the reaction between albumin and DTNB is dependent upon a number of factors. First it depends upon the accessibility of the Cys^{34} sulfur to the DTNB. Secondly, the rate also depends upon the ionisation of the thiol. The thiolate anion is more reactive than the protonated thiol [12]. This is because the thiolate anion carries a negative charge and is therefore more likely to participate in nucleophilic attack. The rate also depends upon the environment of Cys^{34}. The charge distribution of nearby side-chains will influence the reaction rate; this in turn depends upon the 3-dimensional structure around the site.

The increase in activity of the albumins between pH 6 and 8 is probably due to the change in ionisation of the Cys^{34} sulfur and/or nearby residues. Albumin undergoes a structural transition at pH 8 from the N to B state (as discussed in Section 1.2.2). This transition is characterised by a 10% loss in the helical content of the protein [25] and may account for the steep rise in activity between pH 8 and 8.5. At high pH, the activity reaches a maximum for each of the albumins. Wildtype albumin reactivity appears to have reached its maximum around pH 8.4, while for both mutants the pH value of maximum activity is >9. This could mean that the pK$_a$ of the free thiol is higher in both mutants than for the wild-type protein and that both His^{39} and Tyr^{84} contribute in stabilising the thiolate anion. The structure of albumin is extremely flexible and it is possible that both residues could be involved in H-bonding to the Cys^{34} sulfur as shown in Figure 4.7. Likewise it is also possible that the effect on

reactivity is a consequence of the mutations changing the ionisable groups around Cys[34].

Analysis of the crystal structures of unliganded albumin (all structures are summarised in Table 1.2) shows that accessibility to the Cys[34] thiol is partially restricted by 3 residues, Asp[38], Val[77] and Tyr[84] (Figure 4.3). This at least partially explains why the DTNB reaction occurs so much faster with the Tyr[84]Phe mutant than with the wildtype protein. Mutation of tyrosine to a phenylalanine removes the oxygen atom partially covering the entrance to the Cys[34] sulfur, making the thiol more accessible to the DTNB.
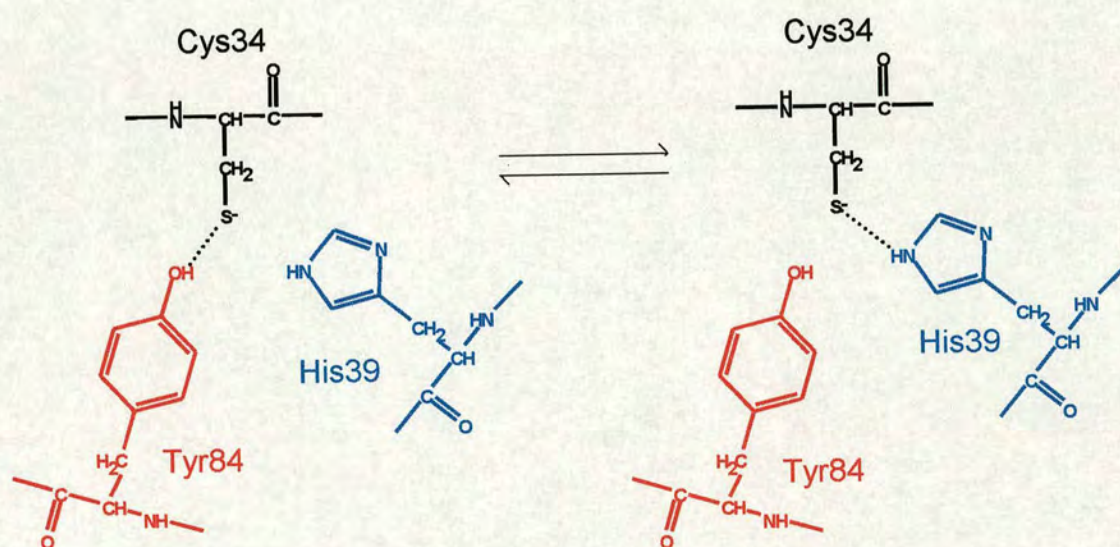

Figure 4.7 Possible H-bonding scheme for stabilisation of Cys[34] thiolate.

The lowered reactivity toward the His[39]Leu mutant at pH is < 8 may be due to a steric effect. The substitution of histidine by leucine could mean that Cys[34] becomes more buried. Also there may be a change in the ionisation of the thiol over this pH range.

The data presented here suggest that Tyr^{84} plays a major role in determining the reactivity of Cys^{34} towards the disulfide DTNB. At pH values between 6 and 8 the Tyr^{84}Phe mutant is *ca.* 4 times more reactive, and at pH 10.2 is 170 times more reactive. This can be explained by the existence of strong H-bonding between the Tyr^{84} OH group and the deprotonated Cys^{34} thiolate. The rate of reaction of both wildtype and mutant proteins studied here is highly sensitive to pH in the physiologically relevant range of pH 6 to 8. The pH control of Cys^{34} may therefore be important *in vivo*.

These findings show that His^{39} and Tyr^{84} play a role in controlling reactivity of DTNB at Cys^{34}. This has a number of implications, as it may be possible to synthesise mutant albumins, which have faster or slower reaction rates at this site. Such proteins could be useful for scavenging reactive oxygen species or could be administered to increase the efficacy of Cys^{34}-specific therapeutics.

4.3 Peroxidase Activity of Albumin

4.3.1 Production and Physiological Importance of Hydrogen Peroxide

Hydrogen peroxide (H_2O_2) is a product of many enzyme-catalysed reactions in the body. In particular, enzymes which break down certain amino acids and fatty acids such as D-amino acid oxidase [26] and acyl-CoA oxidase [27] make significant amounts of hydrogen peroxide. Hydrogen peroxide can form hydroxyl radicals that are damaging to normal tissue, and so these enzymes are kept inside specialised organelles in the cells called peroxisomes. The peroxisomes contain large amounts of

catalase to break down the hydrogen peroxide before it can escape [28]. Other enzymes that make significant amounts of hydrogen peroxide are plasma amine oxidase [29] and xanthine oxidase [30]. These reactions do not produce hydrogen peroxide directly, but rather superoxide ($O_2^{.-}$). In order to get rid of superoxide, the body also contains superoxide dismutase, an enzyme that converts $O_2^{.-}$ into H_2O_2 and H_2O [27].

Another interesting source of superoxide and peroxide (as well as other harmful oxygen species) in the body, is that produced by leukocytes (white blood cells) when they encounter harmful microorganisms. The leukocytes produce large amounts of superoxide, hydrogen peroxide, singlet oxygen ($^1O_2^{.-}$) and hypochlorite (OCl^-) in order to kill pathogenic bacteria [31].

Sulfur-containing amino acids in proteins have previously been shown to have hydrogen peroxide destroying activities [32]. The Cys^{34} residue of serum albumin has previously been implicated in the sequestration of hydrogen peroxide in blood [33]. Albumin has also been shown to catalyse the reaction between tert-butyl hydroperoxide and 2-nitro-5-mercaptobenzoic acid [34]. A study by Finch and colleagues identified 6 modified sites on the albumin molecule following treatment with hydrogen peroxide using peptide mapping [35]. They found that Cys^{34}, Met^{123}, Met^{298}, Met^{446} and Met^{548} residues were modified to sulfoxides.

In order to confirm these sites of modification, the experiment was repeated using a similar method. It was hoped however, that some extra sites would be identified by using a higher-resolving HPLC method (described in Section 3.5.2.2). The better the

separation of the tryptic fragments on the column, the easier it is to identify changes in peak pattern from the chromatogram and to isolate the fragments for amino acid sequencing and mass spectrometric analysis.

4.3.2 Experimental

Recombinant human albumin, rHA (2.5 mL of a 0.3 mM solution) was incubated at room temperature for 60 min, following addition of 50 μl of 10% (v/v) H_2O_2. A control rHA solution was also set up as above but with the addition of 50 μL of H_2O instead of H_2O_2. Prior to digestion, samples were desalted using solid phase extraction (SPE), employing mobile phase flow through cartridges (IST 25mg, 1000Å) under vacuum. Protein samples were loaded and eluted using 70% acetonitrile in 0.2% formic acid. Both samples were digested with trypsin as decribed in Section 3.5.2.1 and tryptic fragments were separated by HPLC as described in Section 3.5.2.2. Seven peaks from the H_2O_2-treated albumin digest were collected, which corresponded to new peaks not present in the control, as indicated in Figure 4.8. These fractions were identified by amino acid sequencing as described in Section 3.5.2.3 and analysed by mass spectrometry as described in Section 2.5.
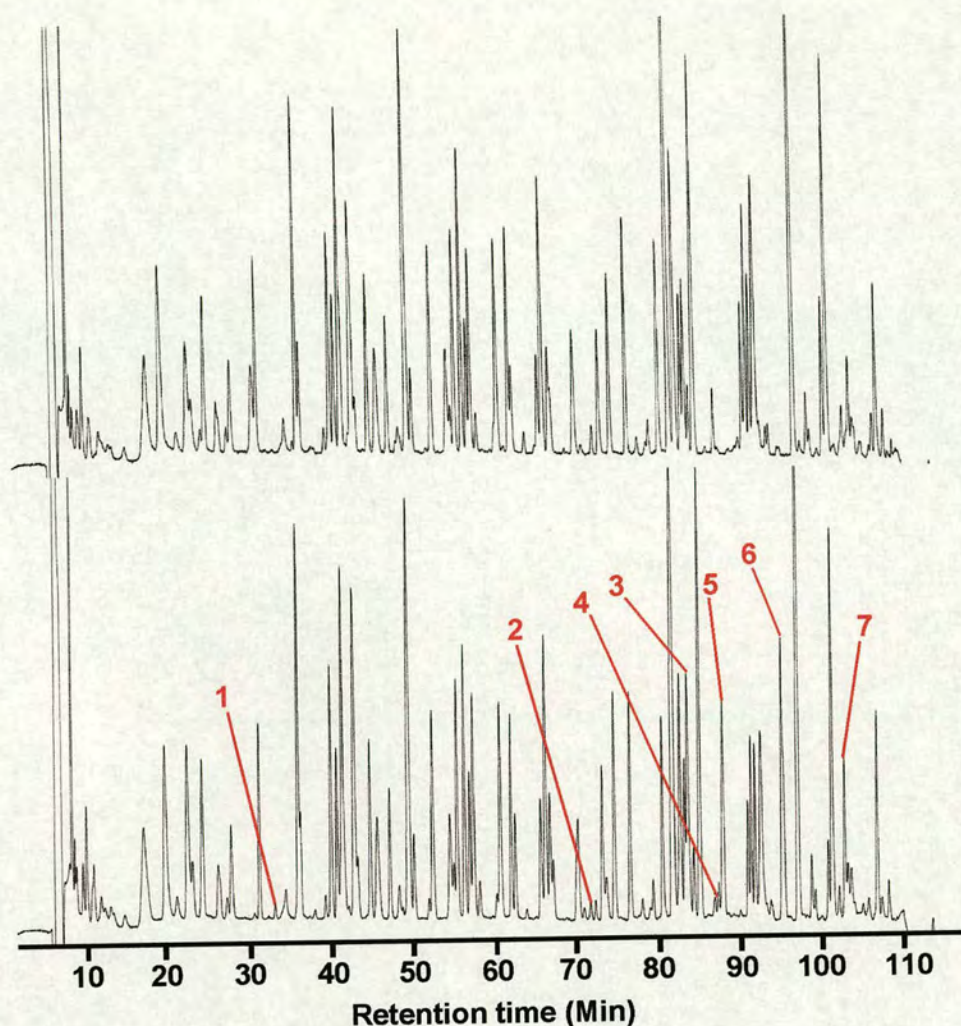
Figure 4.8 HPLC chromatograms of control (above) and H$_2$O$_2$-treated (below) tryptic peptides. The seven peaks collected from the H$_2$O$_2$-treated albumin digest are numbered (1-7).

4.3.3 Results and Discussion

Amino acid sequences from the seven peaks are shown in Table 4.4. The fraction containing peak 5 did not give good sequence data and so this peptide could not be identified. Peak 7 was found to correspond to the peptide containing Cys[34]. Of the other six peaks, four were found to correspond to methionine-containing peptides.

Peak 3 corresponds to a peptide containing neither a cysteine or methionine residue.

Mass spectrometry data for the six sequenced peaks are summarised in Table 4.4.

Peak	Sequence	Residue Numbers	Theoretical Mass (mass units)	Measured Mass (mass units)
1	ETYGEMADCCAK	82-93	1320.5	None detectable
2	LVRPEVDVMCTAFHDNEETFLK	115-136	2594.0	2791.5
3	HPDYSVVLLLR	338-348	1311.6	1466.4
4	SHCIAEVENDEMPADLPSLAADFVESK	287-313	2918.2	3487.8
6	DVFLGMFLYEYAR	324-336	1623.9	1637.8
7	ALVLIAFAQYLQQCPFEDHVK	21-41	2433.9	2516.7

Table 4.4 Summary of sequencing and mass spectrometry results. Amino acid sequences and the corresponding albumin residue numbers for the six sequenced peaks. Also theoretical and measured masses for the six peaks are compared.

With the exception of peak 1 for which no mass was detected, mass spectrometry reveals large differences between the theoretical and measured masses, suggesting that peroxide-induced modifications had taken place. As previously mentioned, hydrogen peroxide is a strong oxidising agent and has been shown to target the sulfur-containing amino acids of albumin [36]. Cysteine can be oxidised to sulfenic (RSOH), sulfinic (RSO_2H) or sulfonic (RSO_3H) acids. Whilst methionine, can be oxidised to a sulfoxide $(RSOCH_3)$ or sulfone (RSO_2CH_3). These oxidised derivatives are shown in Figure 4.9.
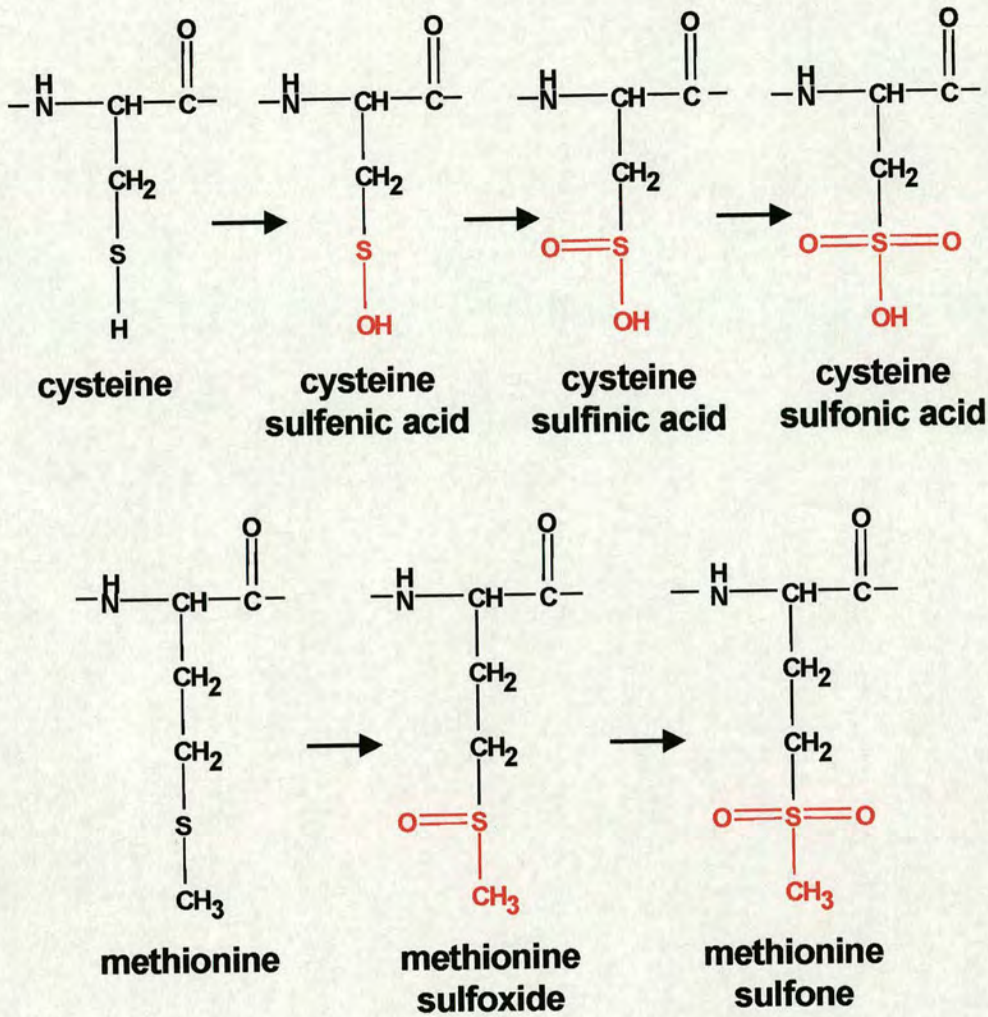
Figure 4.9 Oxidised derivatives of cysteine and methionine.

The mass increase of peak 6 is 13.9 mass units and is consistent, within error, of methionine oxidation to methionine sulfoxide. The mass increases of peaks 2, 3, 4 and 7 are too large to be attributed to a single modification. The increases of peaks 2 and 4 could partially be due to missed cleavage sites on the protein. Peroxide modifications on the protein could cause trypsin to fail in cleaving the peptide bond. Peaks 2 and 4 have mass differences of 197.5 and 569.6 mass units respectively. If tryptic cleavage failed to occur at Lys[136] then peak 2 would carry the extra lysine, Lys[137] which would correspond to an increase in 147.2 mass units. Likewise, if tryptic cleavage failed to occur at Lys[313] then the peptide corresponding to peak 4 would carry the extra amino

acids Asp[314], Val[315], Cys[316] and Lys[317] in its chain, corresponding to an increase in 521.6 mass units.

Further modifications may be attributed to radical-induced damage. Hydrogen peroxide can form hydroxyl radicals in the presence of trace amounts of metal ions such as copper [36] which the albumin solutions may have contained. It has been demonstrated that under certain circumstances an electron-transfer mechanism may exist, leading to the transfer of the radical species from the initial site of damage to amino acids located elsewhere in the protein and in particular to tyrosine, tryptophan and cysteine residues [37]. These result in the generation of phenoxyl, tryptophanyl and thiyl radicals, respectively.

Davies and colleagues investigated the effects of peroxide-derived radical-damage to bovine albumin using EPR spin trapping [38]. They found that when hydroxyl radicals react with bovine albumin, damage can be transferred from the thiol group of Cys[34] to carbon sites on the protein. Under the same conditions, blocking of the free thiol group markedly reduces the number of carbon-centred radicals. This process is represented in Figure 4.10. It has shown that free radical-induced damage of proteins can lead to a range of non-sulfur site protein modifications. For example, carbonyl derivatives of some amino acid residues are formed. Proline and arginine residues can be converted to glutamylsemialdehyde residues, lysine residues to 2-amino-adipylsemialdehyde residues, histidine residues can be converted to asparagine and/or aspartyl residues, whilst proline residues can be converted to glutamyl or pyroglutamyl residues [39]. Additionally, radicals are also known to induce crosslinking between tyrosine residues from different peptides [40]. Therefore the transfer of radical-

induced damage from Cys[34] to other sites on the molecule could explain the large mass increases observed and indeed why the non-sulfur containg peptide in peak 3 is modified.
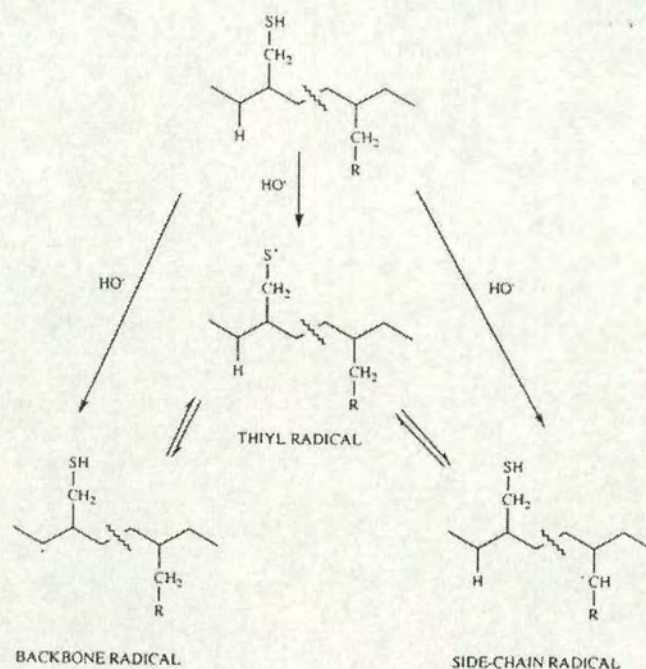

Figure 4.10   The thiol-dependent transfer of radical species to carbon sites. Radicals can be transferred from the thiol (or other initial site of damage) to both backbone and side-chain carbons. Modified from ref [38].

It should be noted that two of the four methionine residues (Met[123] and Met[298]) found within the sequenced peptides, were previously implicated as sites of modification by Finch *et al.* [35]. The other two (Met[87] and Met[329]) are potentially previously unidentified sites of modification. Analysis of the crystal structure of albumin [23] revealed that all four of the methionine residues highlighted in this study (as well as Cys[34]) would be easily accessible to oxidation. The locations of these residues on the albumin molecule are shown in Figure 4.11.
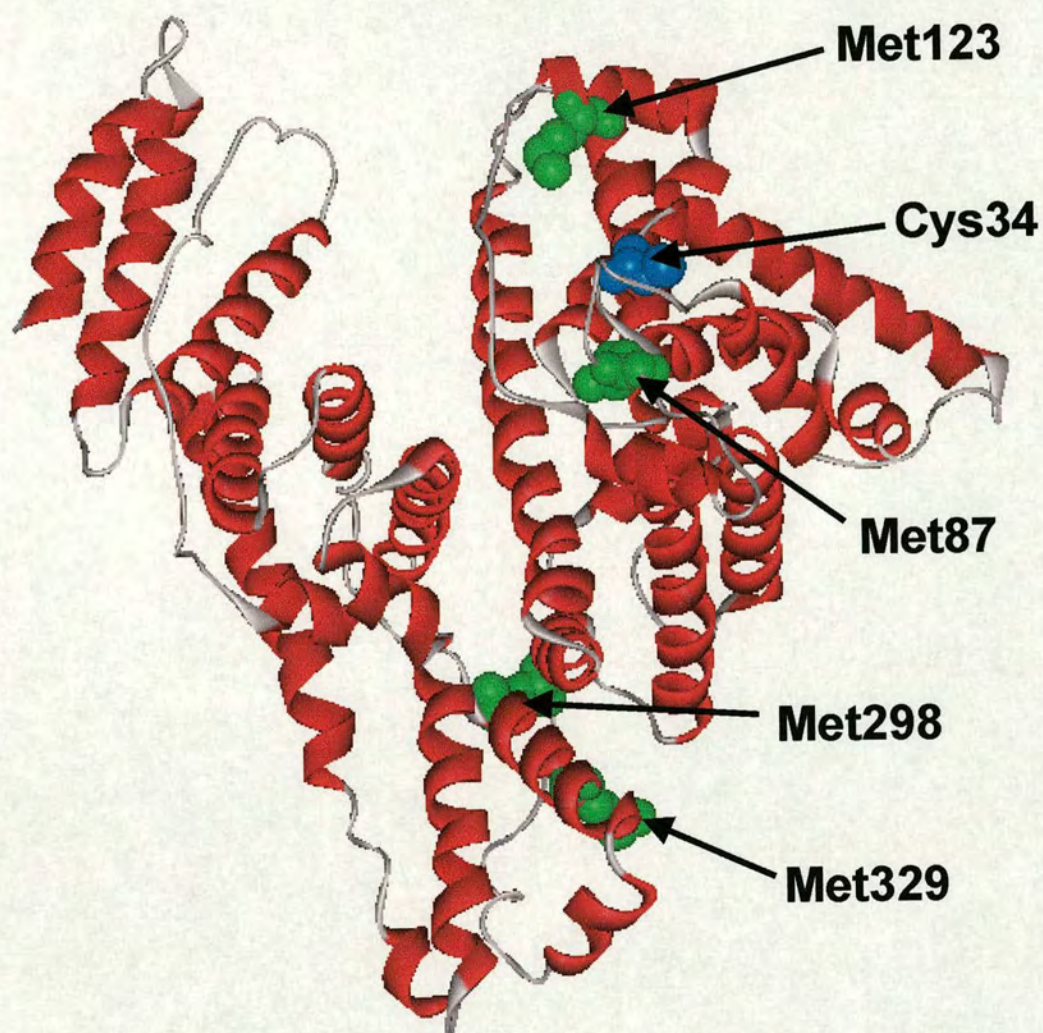
Figure 4.11 Sulfur-containing residues implicated as sites for hydroxyl radical-induced modifications. Drawn from PDB 1AO6 [23].

4.4 Reactions of Nitric Oxide with Low-Molecular Weight Thiols and Albumin

4.4.1 Generation of Nitric Oxide and its Physiological Importance

Nitric oxide (NO) is a key endogenous mediator in the cardiovascular system. Synthesised by nitric oxide synthase in the vascular endothelium from L-arginine [41], NO is associated with vasodilatory and antiplatelet bioactivities (Figure 4.12), as well as being a regulator of epithelial permeability. NO is also thought to be involved in smooth muscle cell relaxation and mitogenesis [42]. It is thought that these properties occur through the activation of guanylate cyclase, a haem-containing heterodimeric enzyme which catalyses the synthesis of cyclic GMP from GTP (Equation 4.7). NO has a high affinity for Fe^{2+} and activates the enzyme by binding to the iron centre in the haem moiety.

$$GTP \longrightarrow cyclic\ GMP + PPi \qquad [4.7]$$
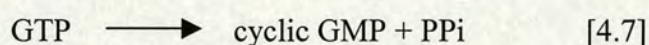
Inhibition of cGMP phosphodiesterase has been shown to down-regulate platelet adhesion [43]. There is also evidence to suggest that cyclic GMP is involved in regulating cervical epithelial permeability [44].

NO concentrations in the circulatory system are quite low due to scavenging by haemoglobin [45]. However, NO has also been shown to bind to the free thiol group in albumin *in vitro* [46] and is thought to exist primarily in the circulatory system as an

S-nitroso adduct of serum albumin [47]. *S*-nitrosoalbumin has been shown to have the vasodilatory [48] and antiplatelet [49] properties associated with NO. These results suggest that albumin may provide a means for delivery of nitric oxide to the haem centre of guanylate cyclase.
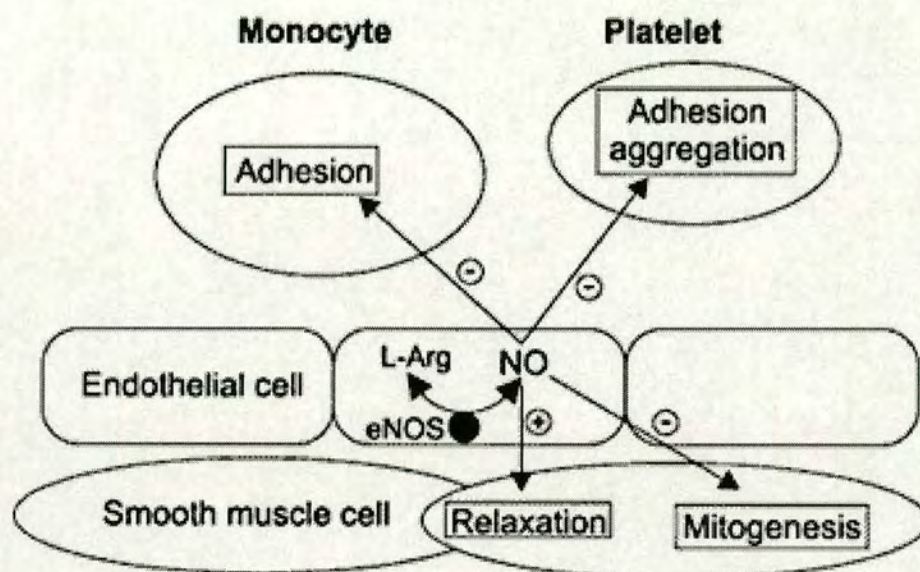
Figure 4.12   Synthesis of nitric oxide (NO) from L-arginine (L-Arg) by endothelial nitric oxide synthase (eNOS) and effects on vascular smooth muscle cells, platelets and monocytes. Taken from Ref: [42].

Additionally, smaller *S*-nitrosothiol compounds are easily prepared in solution which can spontaneously decompose producing NO [50]. Although these compounds are generally unstable in solution, the most stable appears to be *S*-nitroso-*N*-acetyl-penicillamine (SNAP) [51]. SNAP has been shown have similar properties to *S*-nitrosoalbumin being both a vasodilator [52] and able to prevent platelet aggregation [53].

The decomposition of *S*-nitrosothiols has been shown to be catalysed by trace amounts of Cu^{2+} (ca. < 10 μM) in solution [54]. *In vivo*, decomposition may also be accelerated by direct transfer of NO to reduced tissue thiols (transnitrosation) [55-58] to form unstable intermediates like *S*-nitrosocysteine or *S*-nitrosoglutathione, or via interaction with ascorbate [59,60] or even O_2^- [61]. The enzyme γ-glutamyl transpeptidase (γ-GT), has been shown to be involved in the decomposition of *S*-nitrosoglutathione (GSNO) [62]. Potential mechanisms of NO generation from this class of compound are shown in Figure 4.13.

$$\begin{array}{c}
Cu^{2+} \quad h\nu \quad \text{enzymes?} \\
\downarrow \quad \downarrow \quad \downarrow
\end{array}$$

2 RSNO $\longrightarrow$ **RS-SR + 2NO**

2 RSNO + 2CysSH $\longrightarrow$ **RS-SR + 2NO** $\xrightarrow{\text{fast}}$ **Cys-Cys + 2NO**

2 GSNO $\xrightarrow{\gamma\text{-GT}}$ **2 Glu + 2 Gly-Cys-NO** $\xrightarrow{\text{fast}}$ **2 NO + Gly-Cys-Cys-Gly**
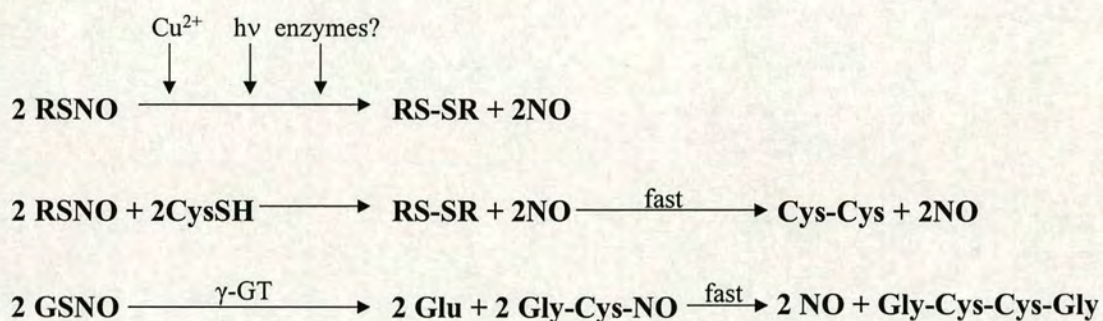
Figure 4.13 Potential mechanisms of NO generation from RSNO compounds. Taken from Ref: [42].

In order to determine the role thiols (particularly albumin) play in the delivery of NO to enzymes such as guanylate cyclase, thiol nitrosation of albumin and the low-molecular weight thiol compound, *N*-acetyl-L-cysteine were investigated. Studies were also carried out in order to elucidate whether transnitrosation can take place between low-molecular weight thiols and albumin.

4.4.2 *S*-Nitrosation Reactions of Albumin and low-molecular weight thiols

The technique of ^{1}H NMR spectroscopy was adopted to monitor the nitrosation of the low-molecular weight thiols, *N*-acetyl-L-cysteine (NAC) and glutathione (GSH) but not albumin due to the strong ^{1}H background signal that is present for large proteins. The nitrosation reaction utilised sodium nitrite as a source of NO (see Equation 4.8) at low pH (< pH 4) [63].

$$RSH + NaNO_2 + H^+ \longrightarrow RSNO + H_2O + Na^+ \qquad [4.8]$$
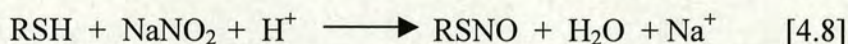
Also ^{15}N NMR spectroscopy, using ^{15}N-labelled sodium nitrite and mass spectrometric methods, were adopted to investigate nitrosation of low-molecular weight thiols and albumin and nitroso-interchange between the low molecular weight thiols *S*-nitrosoglutathione (GSNO) and *S*-nitroso-acetyl-penicillamine (SNAP) with albumin. Both GSNO and SNAP are known to have NO-donor properties [56].

4.4.2.1 Experimental

^{1}H NMR Studies

^{1}H-NMR studies on *N*-acetyl-L-cysteine (Sigma) and GSH (Sigma) were carried out using 10 mM solutions of the thiol (or cystine) with 100 mM $NaNO_2$ (BDH) in D_2O. The pH* was adjusted to 3 by drop-wise addition of DCl. pH* denotes a pH meter reading for a D_2O solution. ^{1}H NMR spectra of NAC and GSH were recorded before and after addition of $NaNO_2$. A ^{1}H NMR spectrum of oxidised cysteine (cystine) was

also recorded under the same conditions in order to ascertain whether NO catalysed the formation of the disulfide.

1D ^{1}H NMR spectra (500.13 MHz, Bruker DMX500) were acquired using a 5 mm TBI probe. Spectra were acquired over a sweep width of 20 ppm into 16 k complex data points, with a pulse width of 8.0 µs (90 °), 64 k transients, an acquisition time of 0.82 s and a recycle delay of 3 s. Prior to Fourier transformation, the FID was apodised by exponential multiplication (0.3 Hz line broadening).

^{15}N NMR Studies

^{1}H-NMR studies on NAC, GSH and rHA were carried out using 100 mM solutions of NAC or GSH (2 mM with rHA) with 10 mol equiv ^{15}N-labelled $NaNO_2$ (99% isotopic purity, Aldrich) in 10% D_2O, 90% H_2O. A spectrum was also recorded of 100 mM ^{15}N-labelled $NaNO_2$ under the same conditions. The pH* was adjusted to 3 by drop-wise addition of HCl. ^{15}N NMR spectra were recorded after addition of $NaNO_2$.

1D ^{15}N NMR spectra (50.71 MHz, Bruker DMX500) were acquired using a 5 mm TBI probe. Spectra were acquired over a sweep width of 200 ppm into 64 k complex data points, with a pulse width of 40.0 µs (90 °), 64 k transients, an acquisition time of 3.20 s and a recycle delay of 1 s. Prior to Fourier transformation, the FID was apodised by exponential multiplication (25 Hz line broadening).

Mass Spectrometry

Recombinant human albumin (150 µM) was nitrosylated by addition of 1.5 mM

$NaNO_2$ in H_2O. The pH was adjusted to 3 by drop-wise addition of HCl. The sample

was stored at –20 °C for 48 h. The sample was thawed and mass spectrum was

recorded as described in Section 2.5. Also 15 µM of wildtype, His39Leu and

Tyr84Phe rHAs were reacted with 1 mol equiv of either S-nitrosoglutathione (GSNO)

or S-nitroso-N-acetylpenicillamine (SNAP) (both purchased from Sigma) in 50 mM

phosphate buffer, pH 7.5. Mass spectra were recorded following addition of S-

nitrosothiol as described in Section 2.5.

4.4.2.2 Results and Discussion

^{1}H NMR spectra of NAC (Figure 4.14a) and GSH (Figure 4.14b) before and after

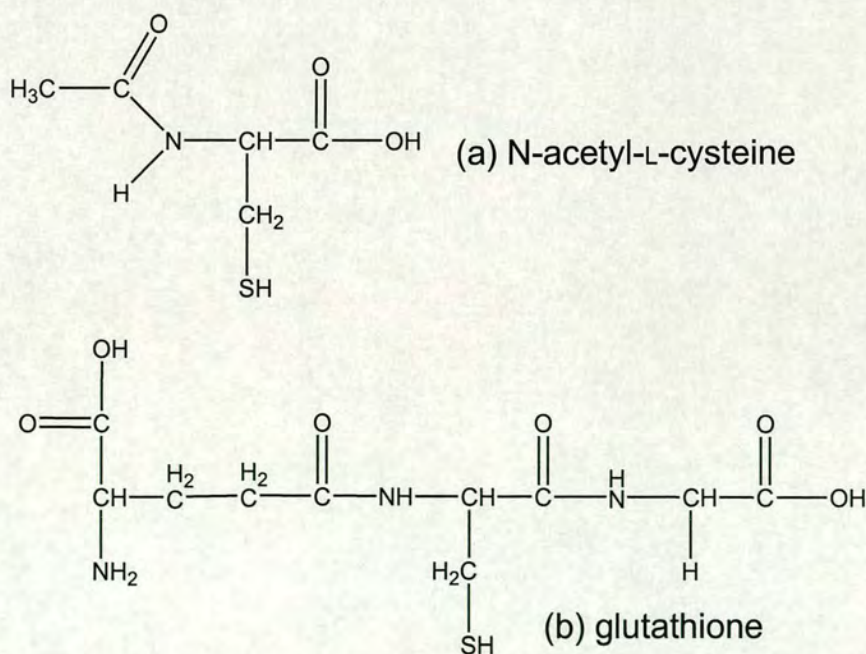addition of $NaNO_2$, and of cystine, are shown in Figure 4.15.

Figure 4.14 Structures of (a) N-acetyl-L-cysteine, NAC and (b) glutathione, GSH.

In the spectrum of NAC and GSH the βCH_2 protons give a multiplet signal at 2.90 ppm (shown in Figure 4.15 by the blue box). Following the addition of $NaNO_2$ to the NAC and GSH samples, two new broad signals appeared at 4.10 and 4.25 ppm (shown by the red box). These peaks are likely to be caused by a change in the environment of the βCH_2 protons, which would be consistent with nitrosation of the thiol group. Similarly, two doublets of doublets appeared at 3.05 and 3.40 ppm (shown by the green box). These signals are also present in the cystine spectrum suggesting that oxidation of the thiols of NAC and GSH to a disulfide occurs.
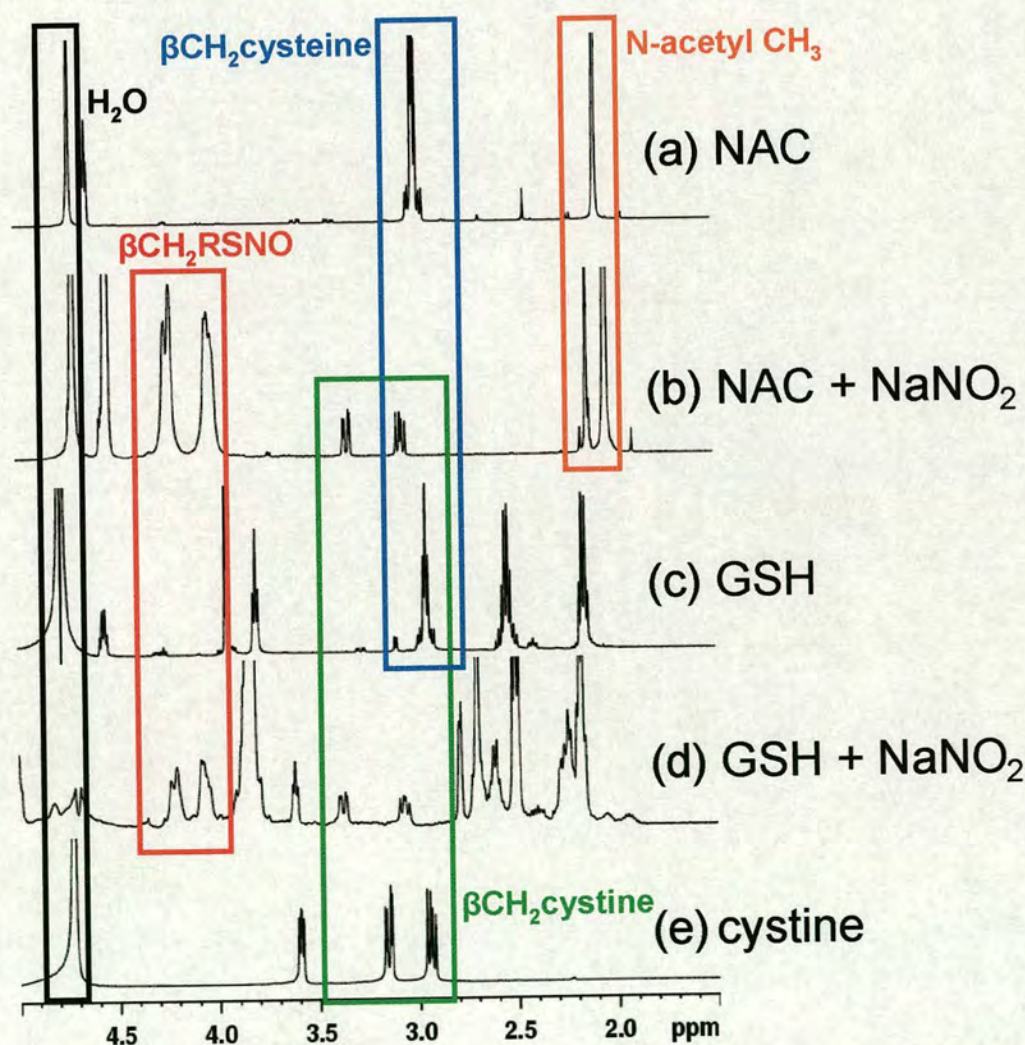

Figure 4.15 ^{1}H NMR spectra of (a) NAC, (b) NAC + $NaNO_2$, (c) GSH, (d) GSH + $NaNO_2$ and (e) cystine.

^{15}N NMR spectra of NAC, GSH and albumin before and after addition of ^{15}N-labelled-NaNO$_2$ and of 100 mM ^{15}N-labelled-NaNO$_2$ are shown in Figure 4.16. Both NAC and GSH in the presence of NaNO$_2$ gave rise to two peaks, a sharp peak around 580 ppm and a broader peak at around 745 ppm. These peaks correspond to free acidified NO$_2^-$ (see Figure 4.16c) [45,65] and the S-nitrosothiol compound respectively [66]. Unfortunately no peaks could be seen on the spectrum of albumin in the presence of NaNO$_2$. The detection of ^{15}N with albumin is most likely to be limited by the low concentration of the protein. This means that ^{15}N NMR is not suitable for investigating nitrosylation reactions with albumin under the conditions described here, in contrast to previous findings [46].
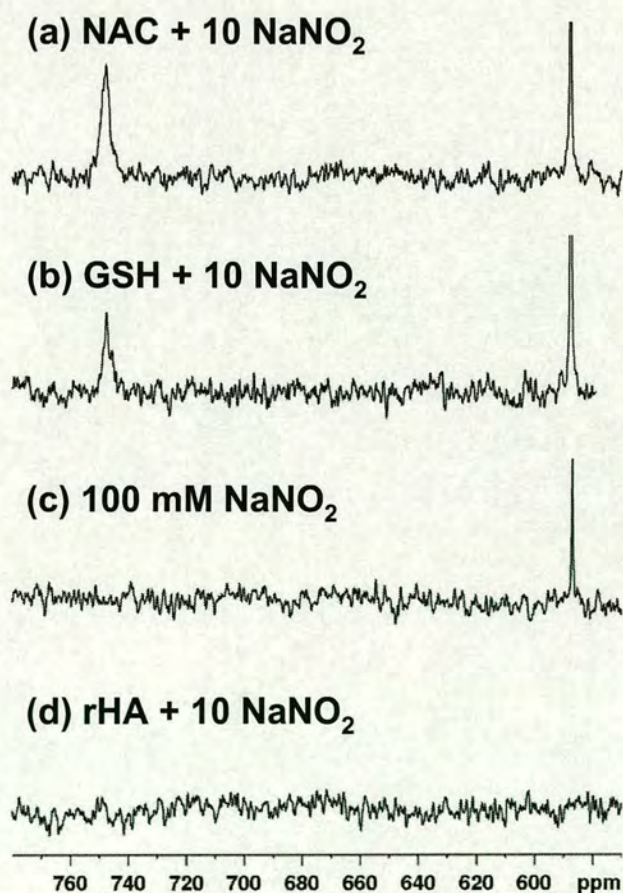

Figure 4.16 ^{15}N NMR spectra of (a) NAC + 10 NaNO$_2$, (b) GSH + 10 NaNO$_2$, (c) 100 mM NaNO$_2$ and (d) rHA + 10 NaNO$_2$.

Nitrosylation of albumin was however detectable by mass spectrometry. Figure 4.17 shows an ESI-mass spectrum of rHA following the addition of acidified NO_2^-. The spectrum shows that not all the albumin present has nitric oxide bound. This is most likely to be due to degeneration of the nitric oxide species, as the NO groups in *S*-nitrosothiol compounds are known to have short half-lives (S-nitroso BSA = ~24 h) [37].
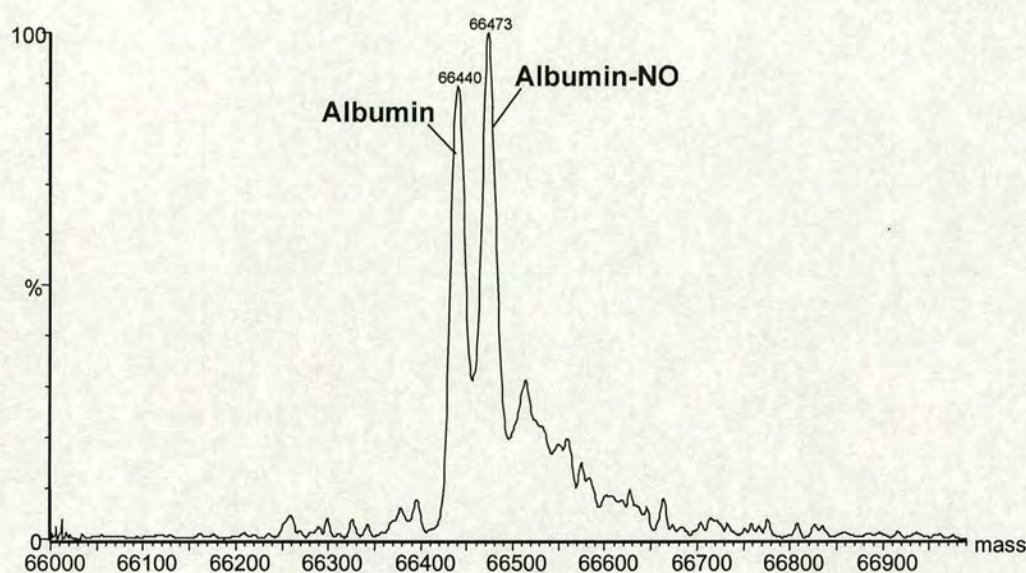
Figure 4.17 ESI-mass spectrum of rHA following the addition of acidified NO⁻₂. Theoretical masses of rHA and rHA-NO are 66438 and 66467 respectively.

The transfer of nitric oxide from the low-molecular weight *S*-nitrosothiols, GSNO and SNAP to albumin and the His[39]Leu and Tyr[84]Phe mutants were also investigated using mass spectrometry (Figure 4.18). Nitrosylation of albumin thiol leads to a mass increase of 29 Da. Such an increase can only be seen in the Tyr[84]Phe mutant on addition of SNAP, the smaller of the two *S*-nitrosothiols (GSNO and SNAP have molecular masses of 336.3 and 220.2 Da respectively). As discussed in Section 4.1 Tyr[84] is extremely critical in controlling the reactivity of the Cys[34] thiol. Hence

DTNB reacts at the Cys[34] thiol of the Tyr[84]Phe mutant at a far greater rate than wildtype. These data do not demonstrate that the transfer of NO from *S*-nitrosoglutathione or *S*-nitroso-*N*-acetyl-penicillamine to native albumin occurs.

Figure 4.18   ESI-mass spectra of rHA and the His[39]Leu and Tyr[84]Phe mutant albumins following the addition of GSNO or SNAP. Nitrosylation of albumin thiol would lead to a mass increase of 29 amu.
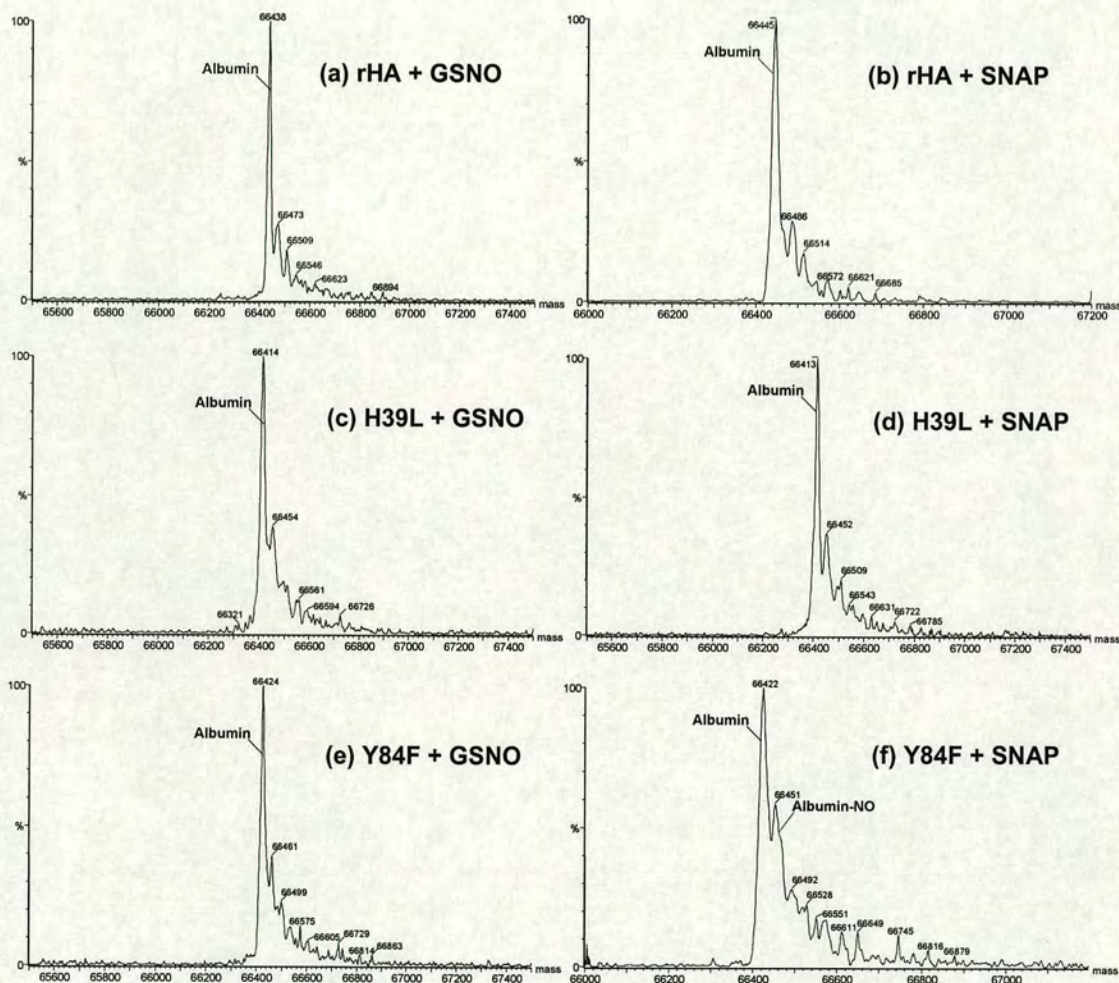
136

4.5 Conclusions

The disulfide exchange reaction between Cys^{34} and DTNB was studied over a range of pH values (pH 6.0-10.2) for native rHA, octanoate-rHA and His^{39}Leu and Tyr^{84}Phe mutants. A large increase in reactivity toward DTNB was observed for the albumin mutant Tyr^{84}Phe, over this pH range. Tyr^{84} is therefore important in the control of Cys^{34} reactivity. Wildtype albumin appears to become fully deprotonated around pH 8.4, whilst the His^{39}Leu and Tyr^{84}Phe mutants become fully deprotonated around pH 9.1. This suggests that the pK$_a$ of the free thiol may be higher in both mutants than in the wildtype protein and that both His^{39} and Tyr^{84} contribute to the stabilisation of the thiolate. Elevated fatty acid levels had no effect on significant effect on Cys^{34} reactivity.

Albumin is a physiologically important antioxidant. Reactions of albumin with H$_2$O$_2$ and NO have been studied. Peptide mapping studies of albumin following reaction with H$_2$O$_2$, revealed that the residues Cys^{34}, Met^{123}, Met^{298} Met^{87} and Met^{329} are potentially sites of peroxide-induced modification. ^{1}H and ^{15}N NMR studies are shown to be useful techniques in the study of the S-nitrosylation reactions between low-molecular weight thiols and acidified ^{15}N-labelled NaNO$_2$ but were unsuccessful in detecting S-nitrosylation of albumin. Mass spectrometry was found to be a suitable technique in the study of albumin-NO reactions. However, no transfer of NO from the low-molecular weight S-nitrosothiols, GSNO or SNAP to native albumin was observed for the conditions described.

References: Chapter 4

1.   Carter, D. C. and Ho, J. X. (1994).   Structure of serum albumin.   *Adv. Prot. Chem.*, **45**: 153-203.

2.   Kay, C. M. and Edsall, J. T. (1956).   Dimerization of mercaptalbumin in the presence of mercurials.   III. Bovine mercaptalbumin in water and in concentrated urea solutions. *Arch. Biochem. Biophys.*, **65**: 354-399.

3.   Okhubo, A. (1969).   On the conformation around the sulfhydryl group in human serum albumin. *J. Biochem.*, **65**: 879-888.

4.   Cornell, C. N., Chang, R. and Kaplan, L. J. (1981).   The environment of the sulfhydryl group in human plasma albumin as determined by spin labeling. *Arch. Biochem. Biophys.*, **209**: 1-6.

5.   Gosselet, M., Mahieu, J.-P. and Sebille, B. (1990).   Reactivity of aromatic and heterocyclic disulphides with thiol group of bovine serum albumin. *Int. J. Biol. Macromol.*, **10**: 241-247.

6.   Christodoulou, J., Sadler, P. J. and Tucker, A. (1995). ^{1}H NMR of human albumin in blood plasma: drug binding and redox reactions at Cys^{34}. *FEBS Letters*, **376**: 1-5.

7.   Roberts, J. R., Xiao, J., Schliesman, B., Parsons, D. J. and Shaw III, C. F. (1996).   Kinetics and mechanism of the reaction between serum albumin and auranofin (and its isopropyl analogue) *in vitro*. *Inorg. Chem.*, **35**: 424-433.

8.   Narazaki, R., Maruyama, T. and Otagiri, M. (1997).   Probing the cysteine 34 residue in human serum albumin using fluorescence techniques. *Biochim. Biophys. Acta*, **1338**: 275-281.

9.   Kratz, F., Warnecke, A., Scheuermann, K., Stockmar, C., Schwab, J., Lazar, P., Druckes, P., Esser, N., Drevs, J., Rognan, D., Bissantz, C., Hinderling, C.,

Folkers, G., Fichtner, I. and Unger. C. (2002). Probing the cysteine-34 position of endogenous serum albumin with thiol-binding doxorubicin derivatives. Improved efficacy of an acid-sensitive doxorubicin derivative with specific albumin-binding propetires compared to that of the parent compound. *J. Med. Chem.*, **45:** 5523-5533.

10.   Ellman, G. L. (1959). Tissue sulfhydryl groups. *Arch. Biochem. Biophys.*, **82:** 70-77.

11.   Ashworth, M. R. F. (1976). *The determination of sulphur-containing groups, Vol. 2. Analytical methods for thiol groups*. Academic Press, London.

12.   Pedersen, A. O. and Jacobsen, J. (1980). Reactivity of the thiol group in human and bovine albumin at pH 3-9, as measured by exchange with 2,2'-dithiodipyridine. *Eur. J. Biochem.*, **106:** 291-295.

13.   Millis, K. K., Weaver, K. H. and Rabenstein, D. L. (1993). Oxidation/reduction potential of glutathione. *J. Org. Chem.*, **58:** 4144-4146.

14.   Lewis, S. D., Misra, D. C. and Shafer, J. A. (1980). Determination of interactive thiol ionisations in bovine serum albumin, glutathione, and other thiols by potentiometric difference titration. *Biochemistry*, **19:** 6129-6137.

15.   Wilson, J. M., Wu, D., Motiu-DeGrood and Hupe, D. J. (1980). A spectrophotometric method for studying the rates of reaction of disulfides with protein thiol groups applied to bovine serum albumin. *J. Am. Chem. Soc.*, **102:** 359-365.

16.   Svenson, A. and Carlsson, J. (1975). The thiol group of bovine serum albumin. High reactivity at acidic pH as measured by the reaction with 2,2'-dipyridyl disulfide. *Biochim. Piophys. Acta*, **400:** 433-438.

17.   Lewis, D. C., Johnson, F. A. and Shafer, J. A. (1974).   Potentiometric determination of ionizations at the active site of papain. *Biochemistry*, **15:** 5009-5017.

18.   Noel, J. K. and Hunter, M. J. (1972).   Bovine mercaptalbumin and non-mercaptalbumin monomers.   Interconversions and structural differences. *J. Biol. Chem.*, **247:** 7391-7406.

19.   Takabayashi, K., Imada, T., Saito, Y. and Inada, Y. (1983).   Coupling between fatty acid binding and sulfhydryl oxidation in bovine serum albumin. *Eur. J. Biochem.*, **136:** 291-295.

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

21.   Bhattacharya, A. A., Grune, T. and 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.

22.   Petitpas, I., Grune, T., Bhattacharya, A. A. and Curry, S. (2001).   Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *J. Mol. Biol.*, **314:** 955-960.

23.   Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999).   Crystal structure of human serum albumin at 2.5 Å resolution. *Prot. Eng.*, **12:** 439-446.

24.   Brocklehurst, K. and Little, G. (1973).   Reactions of papain and of low-molecular-weight thiols with some aromatic disulfides.   2,2'-dipyridyl disulphide as a convenient active-site titrant for papain even in the presence of other thiols. *Biochem. J.*, **133:** 67-80.

25.   Wilting, J., Weideman, M. M., Roomer, A. C. and Perrin, J. H. (1979). Conformational changes in human serum albumin around the neutral pH from circular dichroic measurements. *Biochim. Biophys. Acta*, **579:** 469-473.

26.   Umhau, S., Pollegioni, L., Molla, G., Diederichs, K., Welte, W., Pilone, M. S. and Ghisla, S. (2000). The X-ray structure of D-amino acid oxidase at very high resolution identifies the chemical mechanism of flavin-dependent substrate dehydrogenation. *Proc. Natl. Acad. Sci. USA,* **97:** 12463-12468.

27.   Yeldandi, A. V., Rao, M. S. and Reddy, J. K. (2000). Hydrogen peroxide generation in peroxisome proliferator-induced oncogenesis. *Mutat. Res.*, **448:** 159-177.

28.   Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K. and Watson, J. D. (1994). *The molecular biology of the cell, 3rd Ed.* Garland, New York.

29.   Sharmin, S., Sakata, K., Kashiwagi, K., Ueda, S., Iwasaki, S., Shirahata, A. and Igarashi, K. (2001). Polyamine cytotoxicity in the presence of bovine serum amine oxidase. *Biochem. Biophys. Res. Commun.*, **282:** 228-235.

30.   Stryer, L. (1995). *Biochemistry, 4th Ed.* W. H. Freeman and Co., New York.

31.   Janeway, C. A. and Travers, P. (1997). *Immunobiology: The immune system in health and disease.* Current Biology Ltd./Garland, New York.

32.   Lim, Y. S., Cha, M. K., Kim, M. K., Uhm, T. B., Park, J. W., Kim, K. and Kim, I. H. (1993). Removals of hydrogen peroxide and hydroxyl radical by thiol specific antioxidant protein as a possible role *in vivo. Biochem. Biophys. Res. Commun.* **192:** 273-280.

33.   Pirisino, R., Di Simplicio, P., Ignesti, G., Bianchi, G. and Barbera, P. (1988). Sulfhydryl groups and peroxidase-like activity of albumin as scavenger of organic peroxides. *Pharmacol. Res. Commun.*, **20:** 545-552.

34.   Ignesti, G., Banchelli, M. G., Pirisino, R., Raimondi, L. and Buffoni, F. (1983). Catalytic properties of serum albumin. *Pharmacol. Res. Commun.*, **15:** 569-579.

35.   Finch, J. W., Crouch, R. K, Knapp, D. R. and Schey, K. L. (1993). Mass spectrometric identification of modifications to human serum albumin treated with hydrogen peroxide. *Arch. Biochem. Biophys.*, **305:** 595-599.

36.   Ozawa, T, Ueda, J. and Hanaki, A. (1993). Copper(II)-albumin complex can activate hydrogen peroxide in the presence of biological reductants: first EPR evidence for the formation of hydroxyl radical. *Biochem. Mol. Biol. Int.*, **29:** 247-253.

37.   Prutz, W. A. (1990). Free radical transfer involving sulphur peptide functions, in *Sulfur-Centred Reaction Intermediates in Chemistry and Biology.* Chatgilialoglu, C. and Asmus, K. D, Eds., Plenium Press, New York.

38.   Davies, M. J., Gilbert, B. C. and Haywood, R. M. (1993). Radical-induced damage to bovine serum albumin: Role of the cysteine residue. *Free Rad. Res. Comms.*, **18:** 353-367.

39.   Stadtman, E. R. (1990). Metal ion-catalyzed oxidation of proteins: biochemical mechanism and biological consequences. *Free Radic. Biol. Med.*, **9:** 315-325.

40.   Hanan, T. and Shaklai, N. (1995). The role of H_2O_2-generated myoglobin radical in crosslinking of myosin. *Free Radic. Res.*, **22:** 215-227.

41.   Palmer, R. M. J., Ashton, D. and Moncada, S. (1988). Vascular endothelial cells synthesise nitric oxide from L-argenine. *Nature*, **333:** 664-666.

42.   Megson, I. R. (2000). Nitric oxide donor drugs. *Drugs Fut.*, **25:** 701-715.

43.   Radomski, M. W., Palmer, R. M. J. and Moncada, S. (1987). The role of nitric oxide and cGMP in platelet adhesion to vascular endothelium. *Biochem. Biophys. Res. Comm.*, **148:** 1482-1489.

44.   Gorodeski, G. I. (2000). Role of nitric oxide and cyclic guanosine 3',5'-monophosphate in the estrogen regulation of cervical epithelial permeability. *Endocrinology*, **141:** 1658-1666.

45.   Butler, A. R., Megson, I. R. and Wright, P. G. (1998). Diffusion of nitric oxide and scavenging by blood in the vasculature. *Biochim. Biophys. Acta*, **1425:** 168-176.

46.   Stamler, J. S., Simon, D. I., Osborne, J. A., Mullins, M. E., Jaraki, O., Michel, T., Singel, D. J. and Loscalzo, J. (1992). *S*-Nitrosylation of proteins with nitric oxide synthesis and characterisation of biologically active compounds. *Proc. Natl. Acad. Sci. USA*, **89:** 444-448.

47.   Stamler, J. S., Jaraki, O, Osborne, J., Simon, D. I., Keany, J., Vita, J., Singel, D., Valeri, R. and Loscalzo, J. (1992). Nitric oxide circulates in mammalian plasma primarily as an *S*-nitroso adduct of serum albumin. *Proc. Natl. Acad. Sci. USA*, **89:** 7674-7677.

48.   Gaston, B., Drazen, J. M., Jansen, A., Sugarbaker, D. A., Loscalzo, J., Richards, W. and Stamler, J. S. (1994). Relaxation of human bronchiol smooth muscle by S-nitrosothiols *in vitro*. *J. Pharmacol. Exp. Ther.*, **268:** 978-984.

49.   Simon, D. I., Stamler, J. S., Jaraki, O., Keaney, J. F., Osborne, J. A., Francis, S. A., Singel, D. J. and Loscalzo, J. (1993). Antiplatelet properties of protein *S*-nitrosothiols derived from nitric oxide and endothelium-derived relaxing factor. *Arterio. Thromb.*, **13:** 791-799.

50.   Williams, D. L. H. (1985).   S-Nitrosation and the reactions of S-nitroso compounds. *Chem. Soc. Rev.*, **14:** 171-196.

51.   Field, L., Dilts, R. V., Ramanthan, R., Lenhert, P. G. and Carnahan, G. E. (1978). An unusually stable thionitrite from N-acetyl-D,L-penicillamine; X-ray crystal and molecular structure of 2-(acetylamino)-2-carboxy-1,1-dimethylethyl thionitrite. *Chem. Comm.*, **1978:** 249-250.

52.   Butler, A. R., Williams, D. L. H. (1993).   The physiological role of nitric oxide. *Chem. Soc. Rev.*, **22:** 233-241.

53.   Ignarro, L., Buga, G. N., Byrns, R. E., Wood, K. S. and Chaudhuri, G. (1988). Endothelium-derived relaxing factor and nitric oxide possess identical pharmacologic properties as relaxants of bovine arterial and venous smooth muscle. *J. Pharm. Exp. Ther.*, **246:** 218-226.

54.   McAninly, J., Williams, D. L. H., Askew, S. C., Butler, A. R. and Russell, C. (1993).   Metal ion catalysis in nitrosothiol (RSNO) decomposition. *Chem. Comm.*, **1993:** 1758-1759.

55.   Meyer, D. J., Kramer, H., Ozer, N., Coles, B., Ketterer, B. (1994).   Kinetics and equilibria of *S*-nitrosothiol-thiol exchange between glutathione, cysteine, penicillamines and serum albumin. *FEBS Lett.*, **345:** 177-178.

56.   Askew, S. C., Butler, A. R., Flitney, F. W., Kemp, G. D., Megson, I. L. (1995). Chemical mechanism underlying the vasodilator and platelet anti-aggregating properties of *S*-nitroso-*N*-acetyl-D,L-penicillamine and *S*-nitrosoglutathione. *Bioorg. Med. Chem.*, **3:** 1-9.

57.   Singh, R. J., Hogg, N., Joseph, J., Kalyanaraman, B. (1996). Mechanism of nitric oxide release from *S*-nitrosothiols. *J. Biol. Chem.*, **271:** 18596-18603.

58.   Liu, Z., Rudd, M. A., Freedman, J. E., Loscalzo, J. (1998). *S*-Transnitrosation
      reactions are involved in the metabolic fate and biological actions of nitric
      oxide. *J. Pharmacol. Exp. Ther.*, **284:** 526-534.

59.   Radomski, M. W., Rees, D. D., Dutra, A., Moncada, S (1992). S-Nitroso-
      glutathione inhibits platelet aggregation in vitro and in vivo. *Br. J.
      Pharmacol.*, **107:** 745-749.

60.   Holmes, A.J., Williams, D. L. H. (1998).   Reaction of *S*-nitrosothiols with
      ascorbate: Clear evidence of two reactions. *Chem. Comm.*, **1998:** 1711-1712.

61.   Jourd'heuil, D., Mai, C.T., Laroux, F.S., Wink, D.A., Grisham, M.B. (1998).
      The reaction of *S*-nitrosoglutathione with superoxide. *Biochem Biophys Res
      Commun.*, **246:** 525-530.

62.   Hogg, N., Singh, R. J., Konorev, E., Joseph, J. and Kalyanaraman, B. (1997).
      S-nitrosoglutathione as a substrate for γ-glutamyl transpeptidase. *Biochem. J.*,
      323: 477-481.

63.   Park, J. W. (1993).   *S*-Nitrosylation of sulfhydryl groups in albumin by
      nitrosating agents. *Arch. Pharm. Res.*, **16:** 1-5.

64.   Jansen, A., Drazen, J, Osborne, J. A., Brown, R., Loscalzo, J. and Stamler, J. S.
      (1992).   The relaxant properties in guinea pig airways of S-nitrosothiols. *J.
      Pharmacol. Exp. Ther.*, **261:** 154-160.

65.   Simon, D. I., Mullins, M. E., Jia, M., Gaston, B., Singel, D. J. and Stamler, J.
      S. (1996).   Polynitrosylated proteins: Characterization, bioactivity, and
      functional consequences. *Proc. Natl. Acad. Sci. USA*, **93:** 4736-4741.

66.   Wang, K., Hou, Y., Zhang, W., Ksebati, M. B., Xian, M., Cheng, J.-P. and
      Wang, P. G. (1999).   ^{15}N NMR and electronic properties of *S*-nitrosothiols.
      Bioorg. Med. Chem. Lett., **9:** 2897-2902.

Chapter 5

Interactions of Metals with Albumin

5.1 Introduction

About 25 chemical elements are thought to be essential for mammalian life, 15 of these elements are metals. These metals perform a wide range of important physiological functions within the body including the maintenance of ionic gradients, the carrying of oxygen in the blood as well as playing key roles at the active sites of many enzymes. These are summarised in Table 5.1. In addition to these essential metal elements, a variety of metallodrugs, metal-containing therapeutics and diagnostic agents are commonly used in modern medicine (Table 5.2). Albumin is known to transport metal ions and metallodrugs in the circulatory system [1] and the binding of these to the protein has been well studied.

5.1.1 Metal Ion Binding

Perhaps the most notable example of metal ion binding to albumin is that of Cu^{2+} and Ni^{2+} at the N-terminus. These metals bind more specifically and more tightly to albumin than other metals [1]. Cu^{2+} and Ni^{2+} N-terminal binding is discussed in Section 1.3.5.

Calcium and magnesium are also known to bind to albumin. The first study of Ca^{2+} ion binding to albumin was carried out in 1934 [2]. Albumin binds 1-2 Ca^{2+} ions per molecule and although the interaction is relatively weak ($K_a = 1.5$ x 10^3 M^{-1} [3]) it is highly significant from a physiological and clinical standpoint.

Non-Metals	Metals
H, C, N, O, F, Si, P, S, Cl, Se and I	Na, Mg, K, Ca, V, (Cr,) Mn, Fe, Co, Ni, Cu, Zn, Mo and Sn

Table 5.1   The elements thought to be essential for mammals. Source: ref [4]. Chromium is shown in brackets as there are conflicting views on whether it is an essential element [5-7].

Therapeutic/Diagnostic Agent	Metals Used
radiopharmaceuticals	^{99m}Tc and ^{186}Re
diagnostic agents	MRI (e.g. Gd, Mn) X-ray (e.g. Ba)
anti-infective agents	Ag, Sb and Fe
superoxide dismutase mimics	Mn and Fe
cardiovascular therapeutics	Fe and Ru
insulin mimetics	V and Cr
photodynamic therapeutics	Sn and Lu
anti-cancer agents	Pt and Ru
anti-arthritic drugs	Au
anti-ulcer drugs	Bi

Table 5.2 Some metals and their use in medicine. Source: ref [8].

It is thought that 45% of the 2.4 mM of circulating calcium is free, another 45% is albumin bound, whilst 10% is complexed to small molecules such as citrate and phosphate [1]. Magnesium is bound to albumin with less affinity than calcium ($K_a = 1 \times 10^2$ M^{-1} [9]). Also the concentration of magnesium is about half that of calcium, and again about 45% is bound by proteins. The binding of magnesium to albumin has received very little attention but is thought to parallel calcium in terms of binding site [1].

Another essential metal in the body is zinc present in plasma at a concentration of approximately 19 μM. Zinc is present in over 300 enzymes, including alcohol dehydrogenase and a number of alkaline phosphatases. Zinc binds to albumin with an association constant of 3.4×10^7 M^{-1} [3]. Around 65% of zinc present in plasma is loosely bound to albumin [10] and albumin has been shown to modulate the uptake of zinc into endothelial cells [11].

It has been shown that 1.08 mol Zn^{2+} and 2.08 mol Cd^{2+} bind per mol albumin [9]. Although, Zn^{2+} and Cd^{2+} sites are uncharacterised, ^{113}Cd NMR studies have revealed that zinc competes with cadmium at one site. The ^{113}Cd chemical shift suggests that this site contains 2-3 histidine residues. The other cadmium site is thought to be composed mainly of oxygen atoms [12]. Viles and Sadler have shown that both these cadmium sites are conserved in several mammalian albumins with similar ^{113}Cd chemical shifts for human, bovine, murine, canine, sheep and equine albumins [13]. Although cobalt concentrations in blood plasma are less than 0.01 μM, it is an important component of vitamin B_{12} and is often used as a probe for Zn^{2+}. Several studies have investigated cobalt-albumin binding. Co^{2+} has been found to bind to albumin ($K_a = 6.5 \times 10^3$ M^{-1}

148

[14]) and has 3 principal binding sites [15]. ^{1}H NMR perturbations in the presence of Co^{2+} are consistent with binding at the N-terminal copper-nickel site [16].

Al^{3+} and Mn^{2+} ions also bind to albumin but have a much stronger affinity toward serum transferrin [1,17]. Normally Al^{3+} concentration in plasma is less than 0.4 μM but can increase to as much as 5 μM in patients on renal dialysis. In such extreme cases as much as 34% of Al^{3+} has been shown to be albumin bound [18]. The K_a for Mn^{2+} binding to albumin is 2.4 x 10^4 M^{-1} [19]; only about 5% of added ^{54}Mn has been found to be associated with albumin in plasma. Additionally Hg^{2+} has been shown to bind at Cys^{34} and reportedly is able to link single albumin thiols forming dimers [20].

5.1.2 Metallodrug Binding

A variety of metallodrugs rely on albumin to carry them through the circulatory system to their target site(s). Perhaps the longest-established of these compounds are the Au(I) thiolates used in the treatment of rheumatoid arthritis [21]. Early in the last century gold complexes were used in the treatment of tuberculosis. In 1929, a French physician noted that these compounds had a greater efficacy in the treatment of arthritis [22]. These early therapeutics were soon abandoned due to their serious toxicity and were replaced by cortisone as the preferred treatment. Cortisone itself had serious side effects and with the advent of available atomic absorption spectroscopy to monitor serum gold levels, gold therapies began resurgence in the 1970s [21].

Perhaps the most notable of these therapies are the anti-arthritic drugs, auranofin, myochrysine and solganol. The structures of these 3 drugs are shown in Figure 5.1, the latter two form gold(I)-sulfur oligomers. Auranofin by contrast is a well-defined

monomeric complex with triethylphosphine (Et$_3$P) and 2,3,4,6-tetra-O-acetyl-β-1-D-thioglucose (Atg) ligands. The mode of action of these drugs still remains unknown, mainly because ligand exchange reactions at Au(I) occur rapidly [23]. The Cys[34] residue of albumin has been shown to be the preferred binding site for auranofin, forming a triethyl phosphine Au(I) adduct at the cysteinyl sulfur (see Reaction 5.1) [23].

$$AlbSH + Et_3PAuSAtg \rightarrow AlbSAuPEt_3 + HSAtg \qquad (5.1)$$

The Au(I)-sulfur oligomeric complexes, myochrysine and solganol, are also thought to bind at Cys[34]. None of these three drugs undergo ligand exchange reactions with methionines or disulfide bridges in proteins in the body, although auranofin has been shown to react with albumin disulfide bridges *in vitro* if the drug is in excess [23]. Competition studies have eliminated the possibility of the drug binding at sites I and II [21].

Albumin also plays an important role in the metabolism of the platinum anti-cancer drug cisplatin (*cis*-[PtCl$_2$(NH$_3$)$_2$]). Cisplatin is used extensively for the treatment of ovarian and testicular cancers as well as a variety of solid tumours [24]. The cytotoxic effect of this drug is due to attack on DNA bases following hydrolysis of the complex, inducing apoptosis within cancer cells [25]. Albumin's importance in the efficacy of these drugs has been highlighted by a number of studies. Hypoalbuminaemic patients have shown poor response to cisplatin treatment [26]. Whilst the administration of preformed albumin-cisplatin complexes has shown cytotoxicity, increasing patient survival time [27].
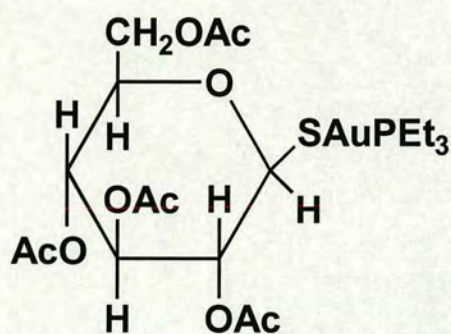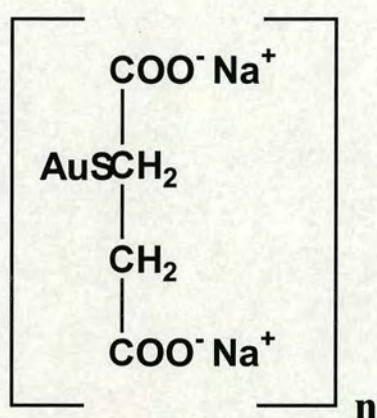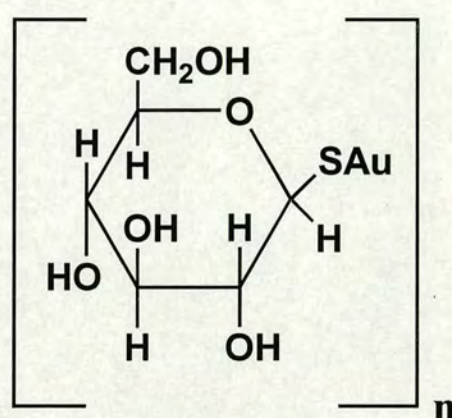
CH$_2$OAc

O

H

H

SAuPEt$_3$

OAc   H

H

AcO

H   OAc

auranofin

COO$^-$ Na$^+$

AuSCH$_2$

CH$_2$

COO$^-$ Na$^+$

n

myochrysine

CH$_2$OH

O

H

H

SAu

OH   H

H

HO

H   OH

n

solganol

Figure 5.1   Three clinically used anti-arthritic gold(I) complexes. Note that myochrysine and solganol form oligomeric chains through their bridging thiolate sulfurs.

^{15}N NMR studies of the reaction between ^{15}N-labelled cisplatin and Cys, Met and His modified albumins have revealed that the major cisplatin binding sites on the protein involve methionine residues. Binding was also detected at Cys[34], and at a site involving a nitrogen ligand, in the form of an S, N macrochelate. In the later stages of reaction the release of NH$_3$ was found to occur due to the strong *trans* influence of the Met sulfur, weakening the Pt-NH$_3$ bonds. Protein cross-linking through platinum was also observed [24]. These reactions are represented in Figure 5.2.
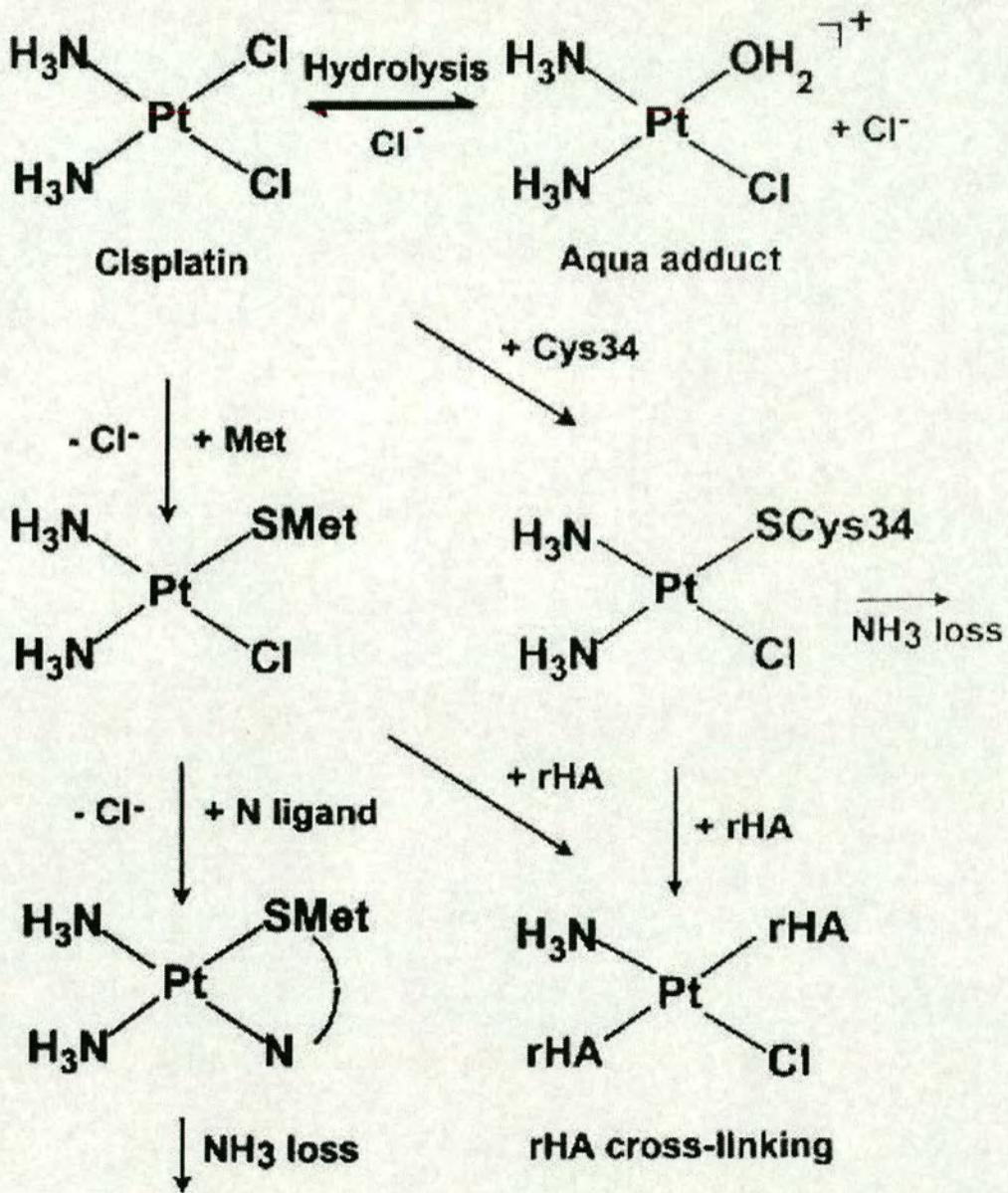
Figure 5.2 Reactions of cisplatin with albumin. Modified from ref [24].

5.2 Identification of a Binding Site on Human Albumin for Zinc and Other Metal Ions

5.2.1 Background

Zinc is among the most important of the trace elements in human nutrition. The remarkable ability of this metal to participate in flexible and easily exchangeable ligand binding with organic molecules underlies the extraordinary extent to which zinc has been incorporated into a wide range of biological systems. Zinc is essential for gene expression and nucleic acid metabolism [28], which accounts in part for its importance for cellular growth and differentiation, in which it may actually have a regulatory role. Its ligand binding properties are utilised effectively at the catalytic sites of a wide range of enzymes [29,30]. Zinc also has many structural roles in biological membranes, cell receptors, enzymes and other proteins. One much-quoted example is the "zinc finger" that is present in certain transcription proteins [28].

Albumin transports zinc and copper in the blood. Albumin has been shown to modulate zinc uptake by endothelial cells [11] and receptor-mediated vesicular co-transport across the endothelium has been demonstrated with albumin-zinc complexes *in vitro* [31]. The binding sites for Zn^{2+} on albumin have not been specifically located, even though albumin is believed to be the main zinc transport protein in the circulation.

5.2.2 Zinc Binding

An indication of where Zn^{2+} binds on albumin has come from [113]Cd-NMR studies [12,13,32,33]. Several mammalian albumins have two strong binding sites for Cd^{2+} with chemical shifts characteristic of N/O coordination (Figure 5.3). For human albumin, [113]Cd shifts of 24 and 114 ppm (relative to $Cd(ClO_4)$) [13] are indicative of sites containing a single imidazole nitrogen and 2-3 imidazole nitrogens, respectively. Zn^{2+}, Cu^{2+} and Ni^{2+} ions can displace Cd^{2+} from the latter of these sites on human albumin [13].

Examination of the X-ray crystal structure of fatty-acid free human albumin (PDB 1AO6) revealed that only one potential site on the albumin molecule has 2 histidine side-chains within 5 Å from each other, His67 and His247. The identification of other residues around this site revealed that Asn99 and Asp249 were also close enough to provide oxygen ligands for metal binding. Asn99 could also potentially provide a nitrogen ligand from the amide group of its side chain. (Figure 5.4). These 4 residues are highly conserved in all mammalian albumins sequenced to date (Table 1.2).

The aim of this study was to use site-directed mutagenesis and [111]Cd NMR spectroscopy to determine whether the cluster of amino acid side-chains, His^{67}, Asn^{99}, His^{247} and Asp^{249} (Figure 5.4) is involved in Zn^{2+} binding.
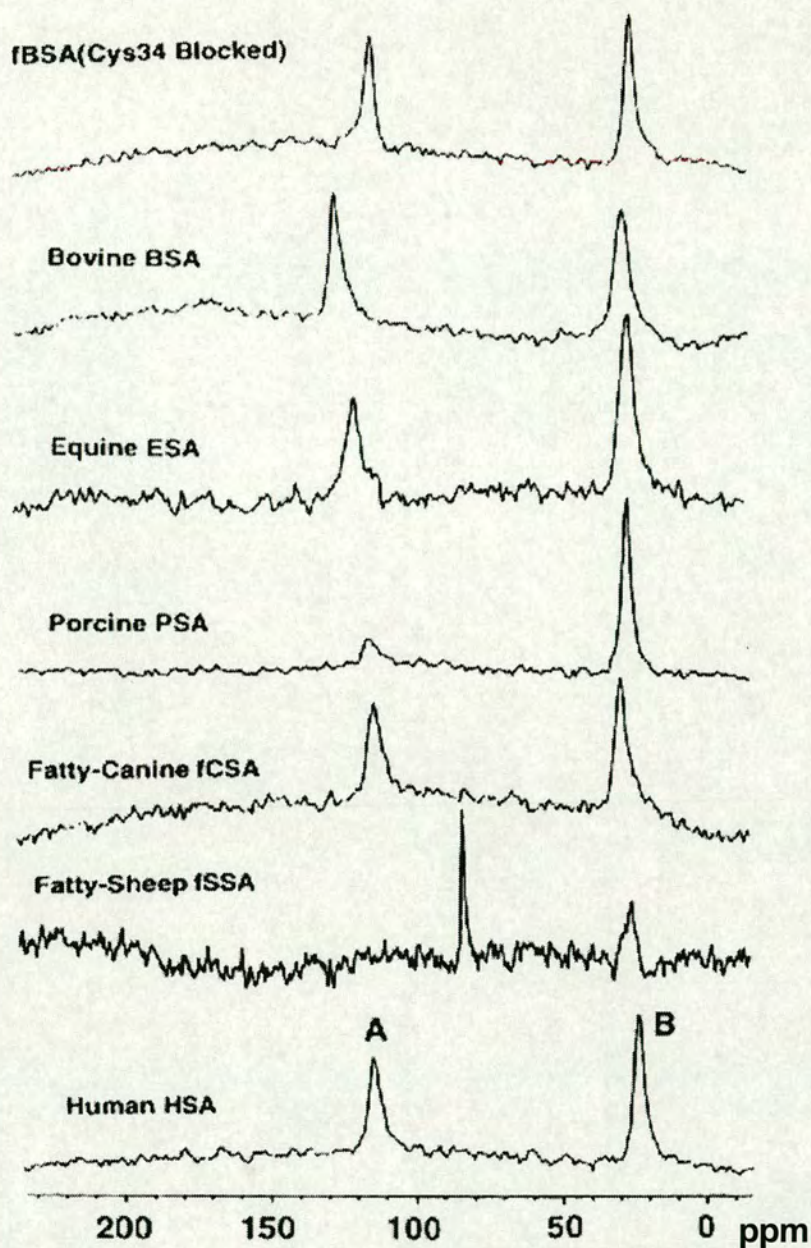
Figure 5.3 ^{113}**Cd NMR spectra of mammalian albumin (2 mM) in the presence of 2**

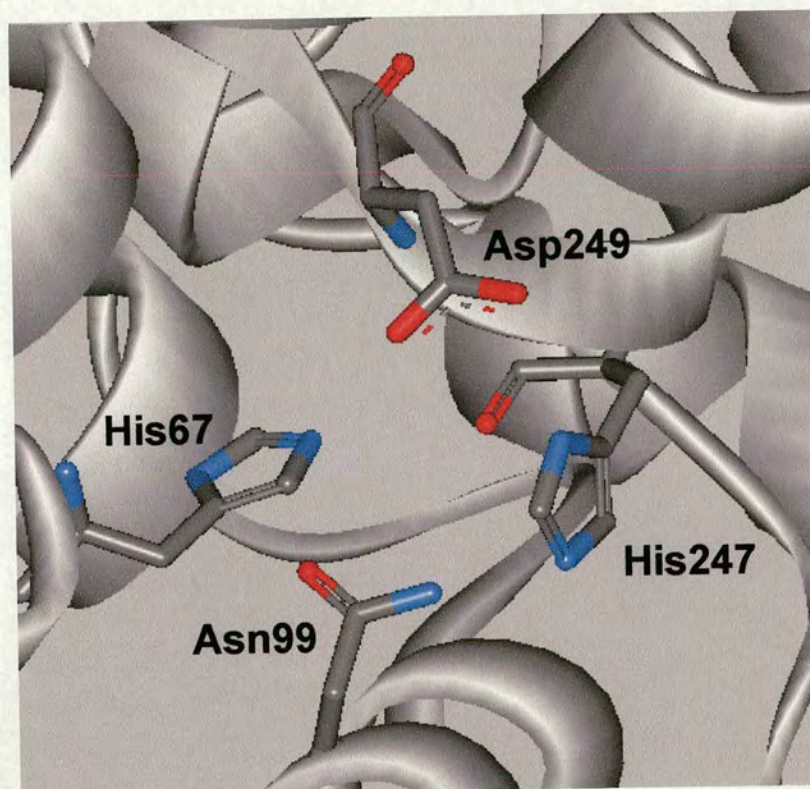**mol equiv of** ^{113}**CdClO$_4$.** Taken from ref [13].

Figure 5.4 The structure of human albumin as reported in PDB 1AO6, in the region of the proposed metal binding site. CPK colours are used for the amino acid side-chains.

5.2.2.1 Experimental

Synthesis of Mutant Albumins

Mutant albumins were synthesized and characterized as described in Chapter 3.

[111]Cd NMR Studies

[111]Cd-NMR studies were carried out using 1.5 mM recombinant human albumin (rHA) or respective mutant protein (Cys[34]Ala or His[67]Ala) at the same concentration, in 50 mM Tris, pH 7.1, 100 mM NaCl, 10% D_2O with 2 mol equiv of [111]CdCl$_2$ at 21 °C. Various equivalents of ZnCl$_2$ (Acros) were added for metal titration experiments, the pH was

checked and adjusted (if required) after each addition. Dissolution of ^{111}CdO (95.11% isotopic purity, Oak Ridge National Laboratory, Tennessee, USA), generated ^{111}CdCl$_2$ and ^{111}Cd(ClO$_4$)$_2$ in the appropriate amount of 1 M HCl or HClO$_4$, respectively.

1D ^{111}Cd-{^{1}H} NMR spectra (106.04 MHz, Bruker DMX500) were acquired using a 10 mm BBO (direct observe) probe head with the help of Dr. Claudia Blindauer. Inverse-gated proton decoupling was achieved by composite pulse decoupling using GARP [34]. Typically, spectra were acquired over a sweep width of 30 kHz (280 ppm) into 4 k complex data points, with a ^{111}Cd pulse width of 17.5 μs (90º), 36 k transients, an acquisition time of 0.10 s, and a recycle delay of 0.30 s. Prior to Fourier transformation, data were zero-filled to 16 k data points and apodised by exponential multiplication (120 Hz line broadening).

5.2.2.2 Results and Discussion

The ^{111}Cd-NMR studies on 1.5 mM recombinant human albumin (rHA), in 50 mM Tris, pH 7.1 with 2 mol equiv of ^{111}CdCl$_2$ confirmed binding at 2 sites (A and B) with peaks at 131 and 27 ppm (relative to Cd(ClO$_4$)), respectively. Under the same conditions the His^{67}Ala mutant gave rise to a single peak at 29 ppm (Figure 5.5). The Cys^{34}Ala, His^{39}Leu and Tyr^{84}Phe mutants gave rise to two peaks identical to that of the wild type (Figure 5.6). Addition of 0.5 and 1 mol equiv of ZnCl$_2$ to rHA in the presence of 2 mol equiv of ^{111}Cd^{2+} resulted in a decrease in intensity of the peak at 131 ppm (Figure 5.5). These results confirm that zinc can displace cadmium from site A. They also show that Cys^{34} is not involved in the binding of Cd^{2+}. Cadmium does not bind to the His^{67}Ala mutant at siteA, which implicates His^{67} to be a key residue in the binding of cadmium

and zinc at this site and provides evidence for the aforementioned cluster of amino acids
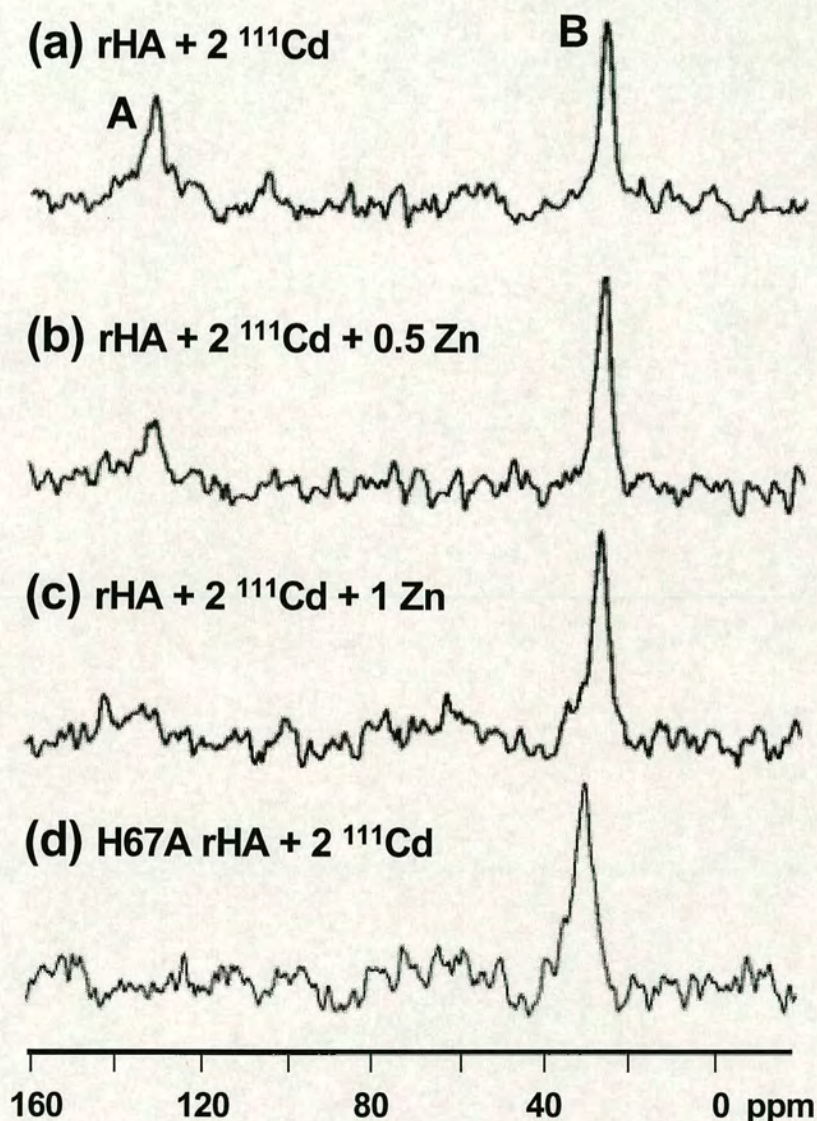
(shown in Figure 5.4) to be the site of binding.

Figure 5.5 ^{111}**Cd NMR spectra of native rHA with various mol equiv of ZnCl$_2$ and**

mutant rHA. All are in the presence of 2 mol equiv of ^{111}CdCl$_2$, in 50 mM Tris, 100

mM NaCl, pH 7.1. (a) Native rHA, (b) rHA plus 0.5 mol equiv ZnCl$_2$, (c) rHA plus 1

mol equiv ZnCl$_2$, and (d) mutant His^{67}Ala rHA.
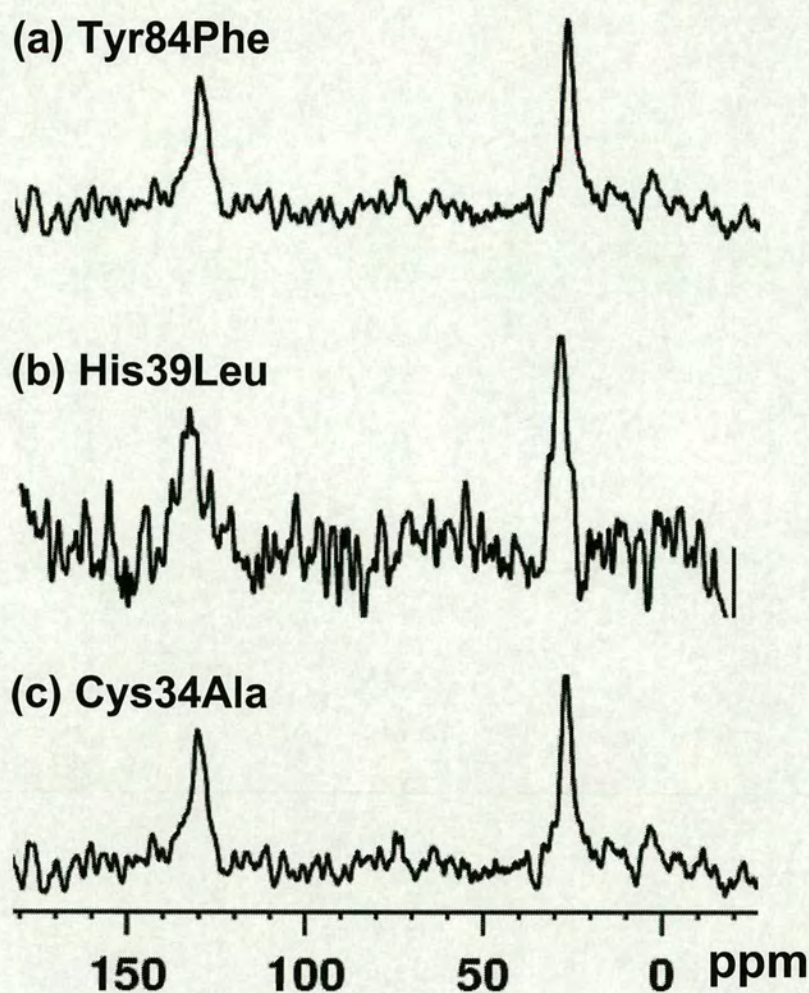
(a) Tyr84Phe

(b) His39Leu

(c) Cys34Ala

150   100   50   0 ppm

Figure 5.6 ^{111}**Cd NMR spectra of mutant rHAs.** All are in the presence of 2 mol equiv of ^{111}CdCl$_2$, in 50 mM Tris, 100 mM NaCl, pH 7.1. (a) Tyr^{84}Phe rHA, (b) His^{39}Leu rHA, (c) Cys^{34}Ala rHA. All are similar to the ^{111}Cd spectrum of native rHA shown in Figure 5.5(a).

5.2.3 Copper and Nickel Binding

Cu^{2+} and Ni^{2+} are known to bind to a second site on the albumin molecule as well as the N-terminus [35]. ^{111}Cd NMR spectroscopy and UV/vis spectrophotometry were used to determine whether these metals are also able to bind at the zinc and cadmium binding site described above.

5.2.3.1 Experimental

[111]Cd NMR Studies

These studies were carried out as described in Section 5.2.2.1 with the exception that aliquots of 0.1 M $CuCl_2$ (Fisons) or $NiCl_2$ (Acros) solution were added to the samples in the metal titration experiments.

UV-Vis Spectrophotometry

Aliquots of $CuCl_2$ solution (in 0.2 mol eq steps) were added to 2 mM solutions of rHA and the H67A mutant in 200 mM potassium phosphate, pH 7.4. Spectra were recorded using a Shimadzu UV250 1PC spectrophotometer between 400 to 800 nm using 1 cm pathlength cells at 25 °C.

5.2.3.2 Results and Discussion

Copper Binding

The addition of 1 mol equiv of $CuCl_2$ did not affect Cd^{2+} binding. This is most likely due to the high affinity of Cu^{2+} for the N-terminus, with Cd^{2+} displacement occurring only after saturation of binding at this site. Addition of 2 and 3 mol equiv of $CuCl_2$ to rHA in the presence of 2 mol equiv of [111]$CdCl_2$ appeared to affect Cd^{2+} binding at site A and led to the formation of a new [111]Cd peak at 37 ppm (Figure 5.7).
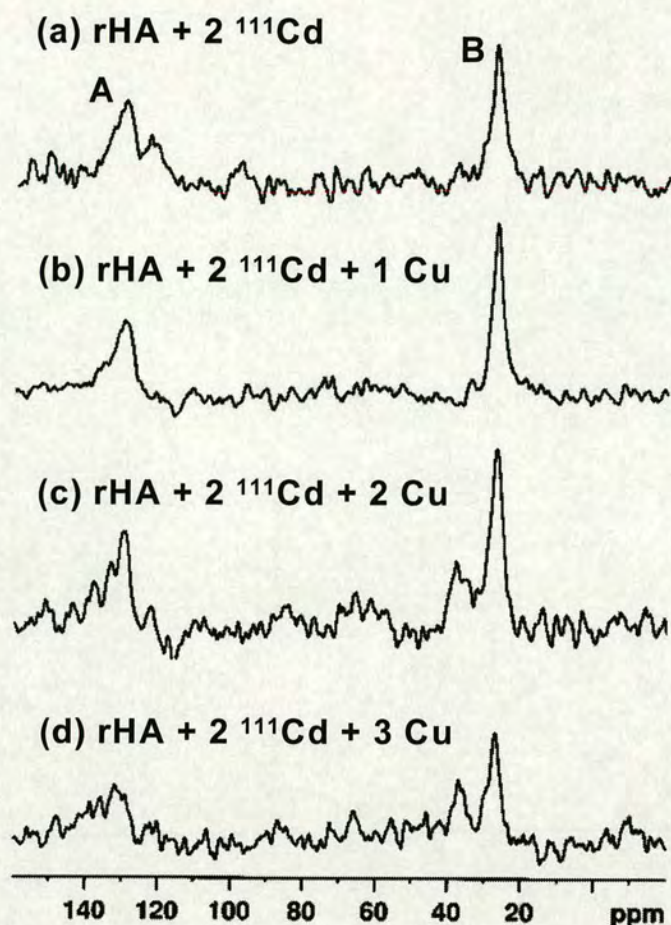
(a) rHA + 2 ^{111}Cd

A

B

(b) rHA + 2 ^{111}Cd + 1 Cu

(c) rHA + 2 ^{111}Cd + 2 Cu

(d) rHA + 2 ^{111}Cd + 3 Cu

140   120   100   80   60   40   20   ppm

Figure 5.7 ^{111}**Cd NMR spectra of rHA in the presence of 2 mol equiv of** ^{111}**CdCl$_2$, in 50 mM Tris, 100 mM NaCl, pH 7.1 with increasing concentrations of Cu^{2+}.** (a) Native rHA, (b) rHA plus 1 mol equiv CuCl$_2$, (c) rHA plus 2 mol equiv CuCl$_2$, and (d) rHA plus 3 mol equiv CuCl$_2$.

The number of nitrogen ligands coordinating to Cu^{2+} in peptides [36] and proteins [37] is known to affect the wavelength of the d-d absorption bands of these complexes. An absorption band at 525 nm appeared after the first addition of CuCl$_2$ (Figure 5.8), indicative of N-terminal loading of the proteins by Cu^{2+}, characteristic of 4 N coordination to Cu^{2+} [36]. However a marked difference in absorption was observed after the further addition of 1 mol equiv CuCl$_2$ to each of the proteins. The native protein developed a second absorption band at 625 nm (Figure 5.8A) and the mutant a much

161

broader band at 750 nm (Figure 5.8B). These bands suggest coordination of Cu^{2+} to 2 N

and 1 N respectively [37]. Absorption maxima were identified by normalising the 0.2

mol equiv spectrum with the 2 mol equiv spectrum for that of the first species then by

subtraction of the normalised spectrum from the 2 mol equiv spectrum for that of the

second species. These results show that site A has a greater affinity for Zn^{2+} than Cd^{2+},

and that Cu^{2+} binds competitively at this site. The results and also suggest the

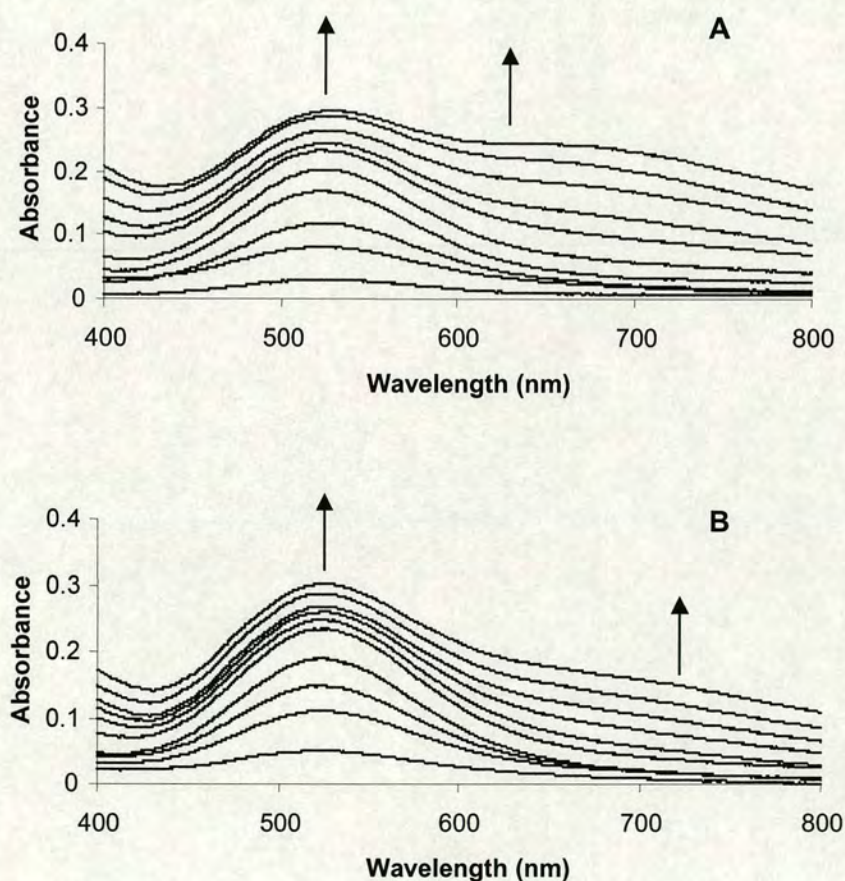involvement of His^{67} in metal coordination.

Figure 5.8 UV-vis absorption spectra of native rHA (A) and H67A rHA (B) with 0.2

to 2 mol equiv of CuCl₂, in 0.2 mol equiv steps (bottom to top) in 200 mM potassium

phosphate, pH 7.4. The arrows highlight the wavelengths of the absorption maxima.

Nickel Binding

Addition of 2 and 3 mol equiv of $NiCl_2$ to rHA in the presence of 2 mol equiv of $^{111}CdCl_2$ had a similar effect to Cu^{2+} on Cd^{2+} binding at site A (Figure 5.9). However, addition of Ni^{2+} did not lead to the formation of a new ^{111}Cd peak. As with the copper salt, the addition of 1 mol equiv of $NiCl_2$ did not affect Cd^{2+} binding. This (as with copper) can be attributed to the high affinity of Ni^{2+} for the N-terminus, with Cd^{2+} displacement occurring only after saturation of binding at this site. It is therefore apparent that Cd^{2+} binding site A is a secondary binding site for Cu^{2+} and Ni^{2+} as well as being a site for Zn^{2+}. Whereas the addition of Cu^{2+} led to the formation of a new peak, the addition of Ni^{2+} did not.
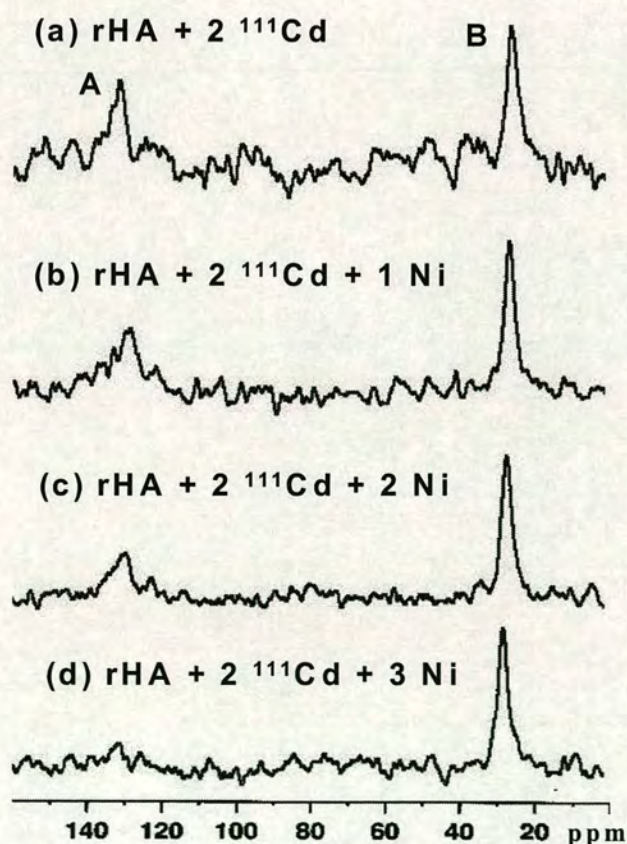

Figure 5.9 ^{111}**Cd NMR spectra of rHA in the presence of 2 mol equiv of** ^{111}**CdCl₂, pH 7.1 with increasing concentrations of Ni²⁺.** (a) Native rHA, (b) rHA plus 1 mol equiv NiCl₂, (c) rHA plus 2 mol equiv NiCl₂, and (d) rHAplus 3 mol equiv NiCl₂.

5.2.4 Chloride Coordination to Albumin-Bound Cadmium

It was noticed that the [111]Cd chemical shift of peak A in these experiments differed from the shift in the [113]Cd albumin studies of Sadler and Viles [13]. Peak A was found to have δ = 131 whilst they found δ = 114. Since the magnetic properties of [111]Cd and [113]Cd isotopes are so similar [38], no differences in chemical shifts would be expected. Closer investigation of the Sadler and Viles studies revealed that their studies were carried out without added NaCl in their buffered albumin solution. In order to ascertain whether the [111]Cd chemical shift of peak A was dependent on chloride concentration (and whether chloride itself could co-ordinate to the metal) spectra were recorded at a range of NaCl concentrations.

5.2.4.1 Experimental

[111]Cd NMR studies were carried out as described in Section 5.2.2.1 with the exception that 1.5 mM rHA was dissolved in 100 mM potassium phosphate, pH 7.1, 10% D_2O with 2 mol equiv of [111]$CdCl_2$ at a range of NaCl concentrations. The chemical shift of peak A in each case was recorded.

5.2.4.2 Results and Discussion

Titrations of Cd_2-rHA with chloride ions suggested that a fifth ligand is involved in site A. Chloride induced large low-field shifts of peak A, but not peak B (Figure 5.10) consistent with weak Cl⁻ binding (log K = 0.23±0.05) at this site. Since the affinity of Cl⁻ towards Zn^{2+} in low-molecular weight complexes [39] is about an order of magnitude smaller than for Cd^{2+}, it is likely that physiologically, water would dominate as the fifth ligand for Zn^{2+} bound to albumin (Cl⁻ concentration in blood *ca.* 104 mM). It should also

be noted that the values for the chloride-induced shifts are consistent with the findings of
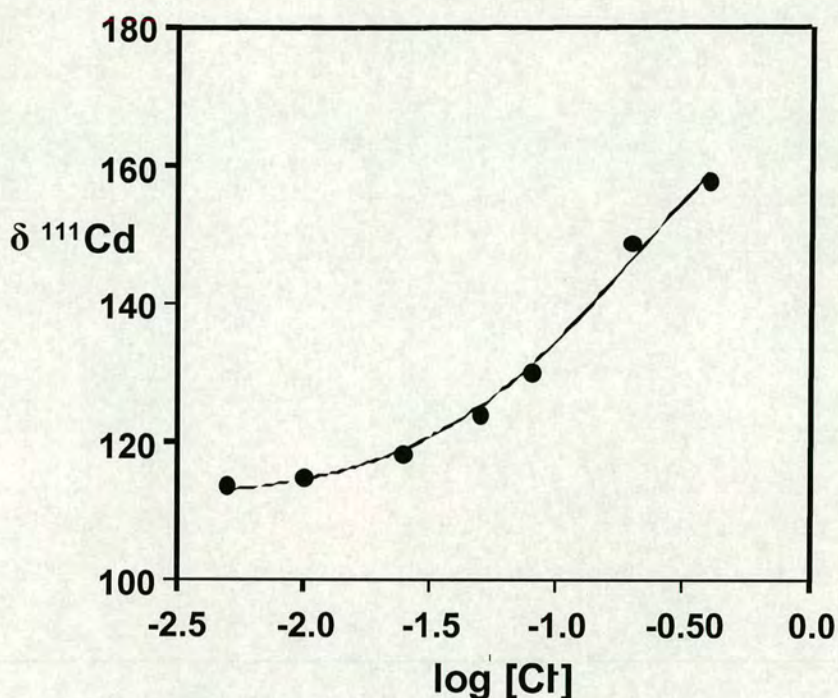
Sadler and Viles [13].

Figure 5.10 Effect of chloride concentration on the [111]Cd NMR chemical shift of peak A of Cd$_2$-rHA. The binding is weak (fast exchange on the NMR time scale) and analysis of the curve as a 1:1 complex (by Dr. Claudia Blindauer) gave $K_a = 0.66 \pm 0.27$. Peak B moves to low field by only 4.7 ppm over the same range of chloride concentrations.

5.2.5 Molecular Modeling

Using the results from the [111]Cd NMR studies, a model of the proposed site with a zinc atom bound was created with the help of Dr Claudia Blindauer. An initial model of Zn-containing albumin was built based on the published crystal structure [40] (pdb accession code 1AO6) using WebLab Viewer v4.0 (Accelrys). The zinc site was modelled as 5-coordinate with the 4 amino acid cluster and with water as the fifth ligand.

The model was imported into Sybyl v6.8 (TRIPOS Inc.) for energy minimisation to optimise geometry, using the TRIPOS force field, after some specific parameters for zinc had been defined. Bond lengths for Zn^{2+} bound to histidine (2.00 Å) and aspartate (2.00 Å), and water (2.06 Å) were taken from ref [41], and a Zn-O bond length for the Asn-Zn^{2+} interaction (2.15 Å) was estimated based on the crystal structures of calcineurin, 5'-endonucleotidase and kidney bean purple acid phosphatase, which were obtained from the Brookhaven Protein Databank (pdb accession codes 4KPB, 1AUI and 1TCO). Force constants were taken from the TRIPOS force field. Bond angles around zinc were not constrained at all, because for Zn^{2+} with a coordination number of 5, no regular or uniform angles were to be expected.

In a first step, the geometry around the zinc was optimised by 200 steps of energy minimisation of the zinc atom, the four protein ligand residues, and the water molecule only. A further 10 steps of energy minimisation were then employed on the entire protein to remove bad geometries and Van der Waals contacts which had been introduced through the atom movements in the first step. The r.m.s.d. values between the original protein structure and the modified model is 0.13 Å for all atoms, and 1.21 Å for the ligands residue atoms only. The final model is shown in Figure 5.11.

The building of a model of a tetrahedral Zn^{2+} site containing only N ligands from the side-chains of His[67] and His[247], and O ligands from Asn[99] and Asp[249] was also attempted by force-field based energy minimization based on the published crystal structure. Despite applying angle constraints, this yielded a geometry that resembled a distorted trigonal bipyramid. This suggests that the 5-coordinate model would be the most stable.

These findings provide a rational for the design of therapeutic recombinant albumins for controlling the levels of available zinc in blood and its delivery to target tissues. Mutagenesis of appropriate residues around this site could be used to produce albumins with decreased or increased Zn affinity. This could have particular implications for the control of zinc deficiencies, the activity of infectious agents [42,43] and for the design of zinc-activated pharmaceuticals [44,45].
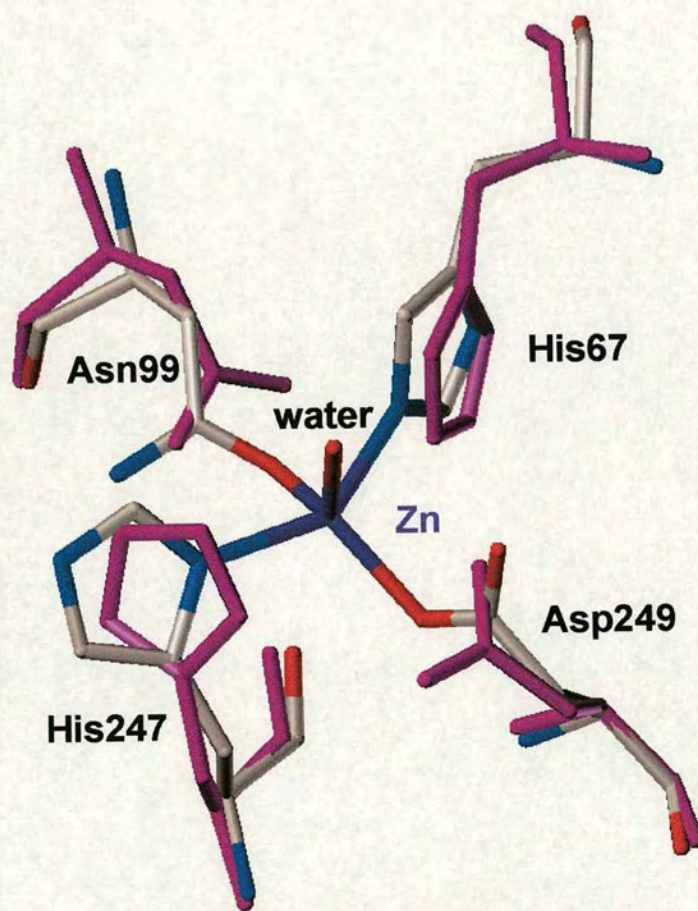

Figure 5.11 Energy-minimized model of the proposed Zn^{2+} binding site on human albumin (CPK colors). Zn^{2+} binding requires only small movements of the amino acid side-chains, as can be seen from the comparison with the superpositioned X-ray structure of apo-human albumin (PDB 1AO6, magenta).

5.2.6 Cadmium Binding to Albumins from Other Species

Although Cd^{2+} binds at two sites (A and B) on a range of mammalian albumins (Figure 5.3), there are some exceptions. Porcine albumin appears to bind Cd^{2+} significantly only at site B, while fatty acid-bound sheep albumin has 2 binding sites but the chemical shift of site A shifts is approximately 30 ppm upfield [13]. These observations suggest that porcine albumin lacks site A and that the cadmium bound at this site to fatty acid-sheep albumin is in a different environment from the other mammalian albumins. All 4 of the residues implicated at site A are completely conserved in both pig and sheep albumin. Sequence comparisons between different species revealed the 4 amino acid cluster to be poorly conserved in non-mammalian albumins. The aims of this study were to confirm that porcine albumin lacks site A, to determine whether the fatty acid bound to sheep albumin in Sadler and Viles' experiments [13] had an effect on the Cd^{2+} binding and to determine whether Cd^{2+} can bind to non-mammalian albumins.

5.2.6.1 Experimental

[111]Cd NMR studies were carried out as described in Section 5.2.2.1 with the exception that 1.5 mM of the respective albumin was dissolved in 100 mM potassium phosphate, pH 7.1, 10% D_2O, 50 mM NaCl with 2 mol equiv of [111]$CdCl_2$.

Fatty acid free porcine and sheep albumins were purchased from Sigma (A-1173 and A-6289, respectively). Of the non-mammalian albumins in the protein sequence database only chicken albumin (NBS Biologicals) was easy to obtain commercially. The chicken albumin purchased was Cohn fraction V [46], purified from chicken blood and is therefore likely to have fatty acid molecules bound. The chicken albumin was dialysed

as described in Section 2.1 and lyophilised prior to use. Porcine and sheep albumins were used as bought.

5.2.6.2 Results and Discussion

[111]Cd NMR spectra of chicken albumin and fatty acid free porcine and sheep albumins in the presence of 2 mol equiv of [111]CdCl$_2$ are shown in Figure 5.12. Neither free [111]Cd^{2+} or the two [111]Cd peaks associated with human albumin were observed for chicken albumin, suggesting that [111]Cd^{2+} ions bind to chicken albumin but are in intermediate exchange (on the NMR timescale) between two or more sites.

(a) Chicken HSA

(b) Porcine HSA

(c) Sheep HSA

160   140   120   100   80   60   40   20   0   ppm

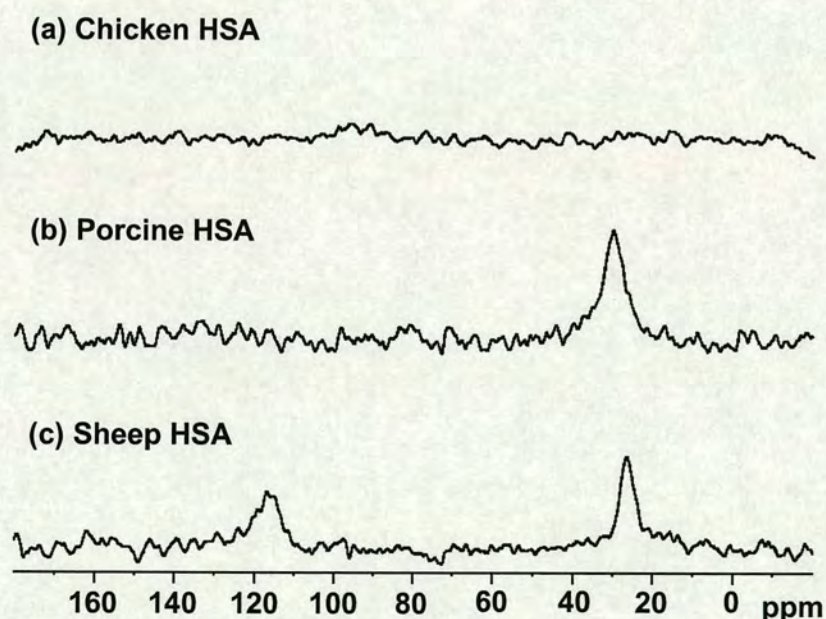**Figure 5.12** [111]**Cd NMR spectra of (a) chicken, (b) porcine and (c) sheep albumins in the presence of 2 mol equiv of** [111]**CdCl$_2$, pH 7.1.**

The [111]Cd NMR spectra of fatty-acid-free porcine albumin again revealed only Cd^{2+} binding at site B, despite having all the residues associated with metal binding at site A conserved. Closer analysis of the residues around the proposed site A in human albumin

reveals that Tyr[30] (located only 3-4 Å away) is a histidine in the pig sequence. Tyr[30] is within H-bonding distance of Asn[99] and is therefore likely to play a role in holding the Asn[99] side-chain in place for metal binding. The lack of this H-bond in the pig albumin may disorientate the binding site. Alternatively, another histidine around this site could provide an alternative ligand for metal binding. This could mean that $^{111}Cd^{2+}$ still binds to the proposed site in porcine albumin but is in intermediate exchange with other sites and cannot be seen by NMR.

The ^{111}Cd NMR spectrum of fatty acid free sheep albumin was identical to human albumin. The upfield shift of peak A observed by Sadler and Viles [13] was most likely an effect caused by fatty acid binding. The effects of fatty acid binding are discussed in the next section.

5.2.7 Fatty Acid-Induced Switch of Site A

It has been noted in published studies that the fatty acid content of serum albumin influences the appearance of ^{113}Cd peak A [13]. The aim of this section was to compare ^{111}Cd NMR spectra of Cd_2-rHA with high (ca. 8 mol mol^{-1}) and low (< 4 mol mol^{-1}) octanoate content.

5.2.7.1 Experimental

^{111}Cd NMR studies were carried out as described in Section 5.2.2.1 with the exception that each rHA (both supplied by Delta Biotechnology Ltd) was dissolved to a concentration of 1.5 mM in 100 mM potassium phosphate, pH 7.1, 10% D_2O with 2 mol equiv of $^{111}CdCl_2$.

5.2.4.2 Results and Discussion

It was observed that peak B was not affected by the variation in fatty acid loading, but peak A was not present in the high octanoate samples (Figure 5.13). The 12 X-ray crystal structures of fatty-acid loaded albumin were examined and it was found that the preorganisation of site A is disrupted in all of them. This arises from the binding of a fatty acid anion in so-called site 2 [47]. Fatty-acid-binding at this site also affects the Cys^{34} site (Section 4.2.2).
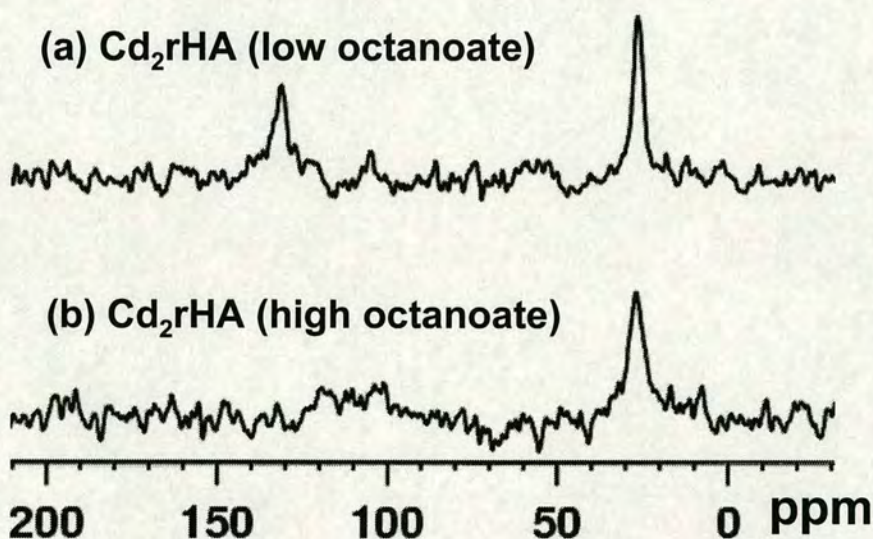
(a) Cd$_2$rHA (low octanoate)

(b) Cd$_2$rHA (high octanoate)

200   150   100   50   0 ppm

Figure 5.13 ^{111}**Cd NMR spectra of rHA in the presence of 2 mol equiv of** ^{111}**CdCl$_2$, pH 7.1.** (a) rHA with a low octanoate content, (b) rHA with a high octanoate content.

The carboxylate end of fatty acid in site 2 interacts with Arg^{257} and Ser^{287} (both subdomain IIA), Tyr^{150} (subdomain IB) and the methyl end with subdomain IA. To accommodate a fatty acid anion, the long helix connecting domains I and II bends and the two half sites in unliganded rHA move by more than 10 Å to form a continuous cavity [47]. This results in a movement of residues His^{247} and Asp^{249} (Figure 5.14) by 4-7 Å away from the other two residues in the proposed Zn^{2+} site, His^{67} and Asn^{99}.

171

Asp^{249} also appears to change its side-chain conformation in order to maintain the H bond to Nε2 of His^{67}, and forms an additional H bond to Asn^{99} (Figure 5.15). His^{247}, which is H-bonded to Asn^{99} in the unliganded structure, forms an H-bond with Glu^{100} in the fatty-acid bound structures.
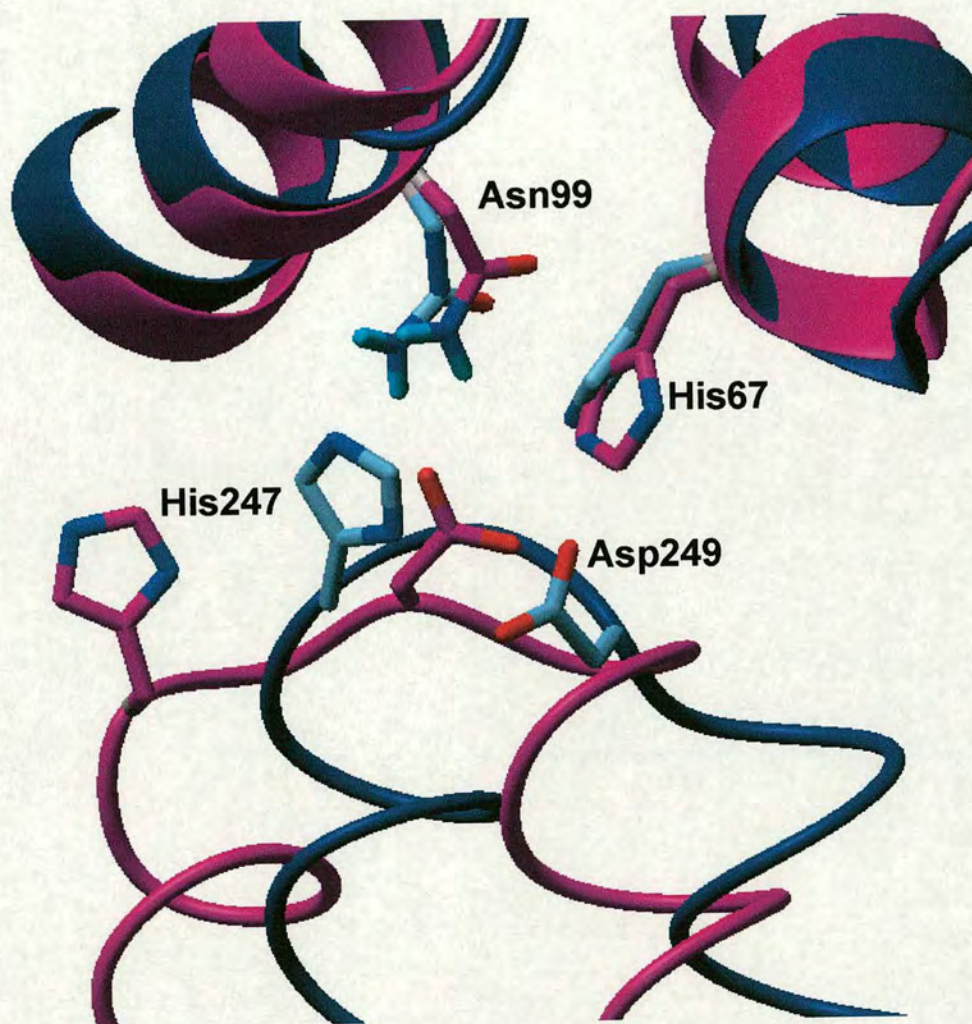

Figure 5.14 Overlay of the structures of human albumin without (blue, PDB accession code 1AO6) and with bound myristate (magenta, PDB accession code 1BJ5). The 4-7 Å movement of His^{247} and Asp^{249}, which are situated in the loop connecting the two helices, away from the metal site. The overlay was generated (by Dr. Claudia Blindauer) in Swiss PDB Viewer (v. 3.6) by aligning the backbone atoms of His^{67} and Asn^{99}.
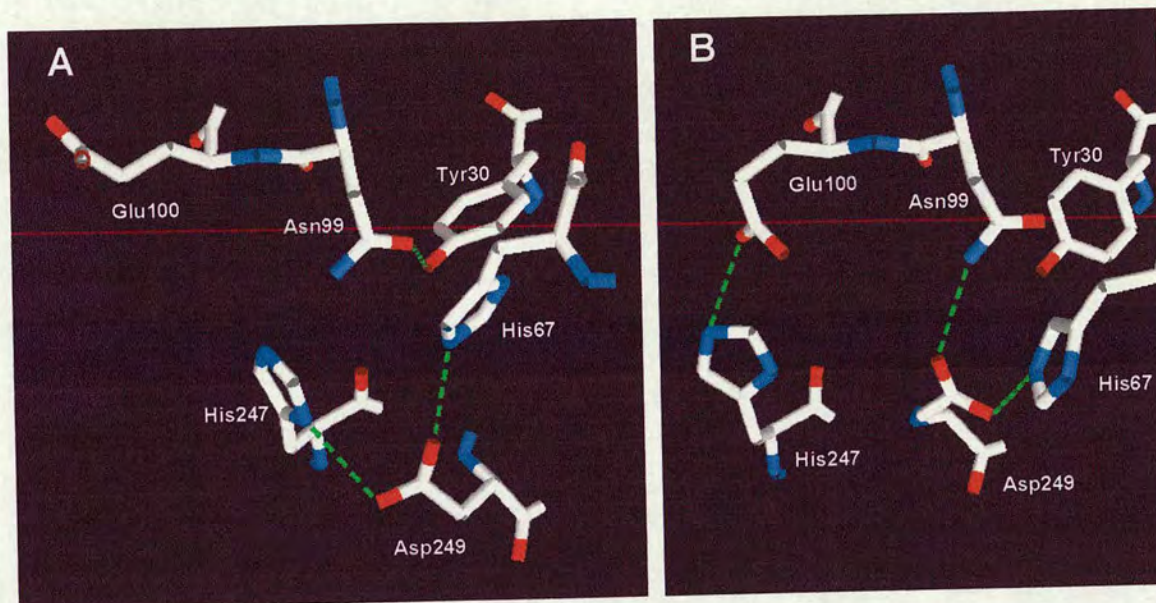
Figure 5.15 H bonding network around the proposed Zn^{2+} binding site (A) without and (B) with fatty acid bound . The orientation of Asn^{99} is the same in both figures.

The network of disulfide bonds (Figure 5.16) partially rigidifies the frameworks on which the proposed Zn ligands are based and also provides a mechanism for inter-helix communication of binding events in this region. The result is a spring-lock mechanism, which engages or disengages the His^{67}/Asn^{99} and His^{247}/Asp^{249} upper and lower jaws of the metal binding site. Under normal physiological conditions, human serum albumin carries 1-2 molecules of long-chain fatty acid (C_{16}-C_{20}) [1], but up to 8 binding sites have been identified [48]. Two of the three high-affinity sites are thought to be located in domain III, and one in domain I. Fatty acid site 2 appears to be a low-affinity site that only becomes populated when fatty acid anions are abundant.
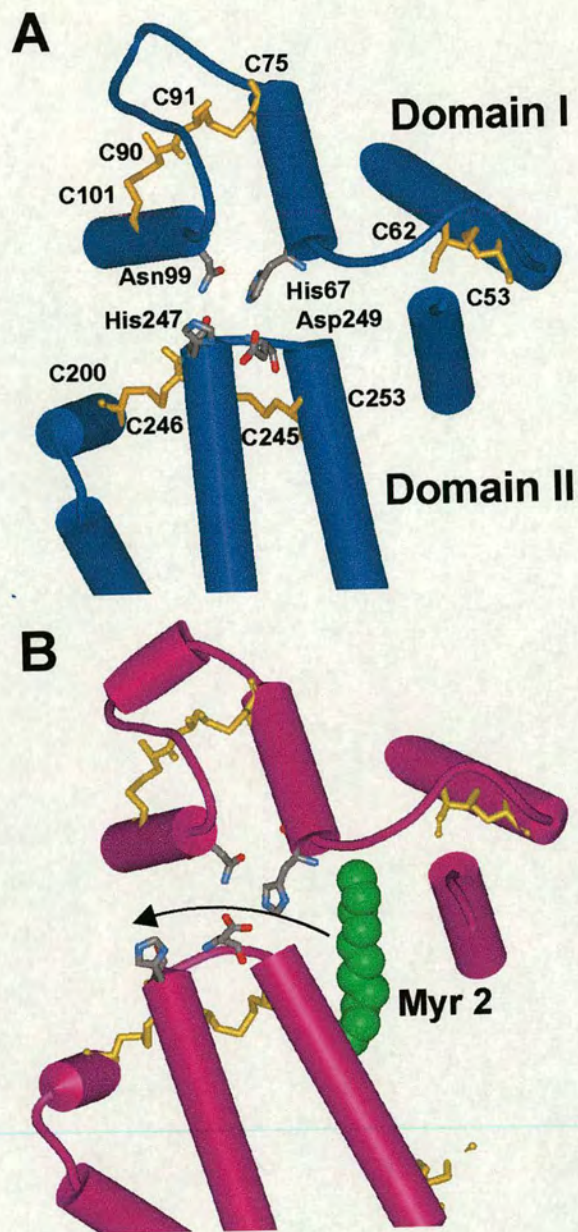
Figure 5.16 Effect of fatty acid binding to site 2 on the proposed metal binding site in human albumin without (blue, PDB accession code 1AO6) and with bound myristate (magenta, PDB accession code 1BJ5). (A) Domain and secondary structure plus cystine-bridge network around the metal site. The cystine residues (yellow) and the zinc ligands (CPK) are shown as stick models. (B) The tilting of the two-helix-bundle in domain II induced by fatty acid binding in myristate site 2 (1BJ5); myristate (green) is shown as a space-filling model.

5.3 Metal Complex Binding to Albumin

5.3.1 Binding of Two Cytotoxic Ruthenium Complexes to Albumin

5.3.1.1 Introduction to RM116 and RM175

Investigations into ruthenium complexes as potential anticancer drugs are of much current interest. This is largely due to the drug NAMI-A ($Na[transRuCl_4Me_2SO(Im)]$), which is currently undergoing clinical trials. NAMI-A, may become the first anti-metastatic drug on the market [49,50].

Two ruthenium(II) complexes (RM116, $[(\eta^6\text{-p-cymene})RuCl(H_2NCH_2CH_2NH_2\text{-}N,N)]^+PF_6^-$ and RM175, $[(\eta^6\text{-}C_6H_5C_6H_5)RuCl(H_2NCH_2CH_2NH_2\text{-}N,N)]^+PF_6^-$), have demonstrated cytotoxicity, even toward cisplatin resistant tumour cells [51]. The mechanism of action of RM116 and RM175 is unknown but may involve an interaction with DNA, like cisplatin. The structures of RM116 and RM175 are shown in Figure 5.17.

Albumin is recognised as being a key protein in the delivery of drugs through the circulatory system (see Section 1.1). A wide range of metallodrugs and compounds are known to bind to the Cys^{34} residue of albumin. The aim of this study was to determine whether either of the two complexes bind to albumin and hence determine whether or not albumin may be involved in their delivery to tumour cells.
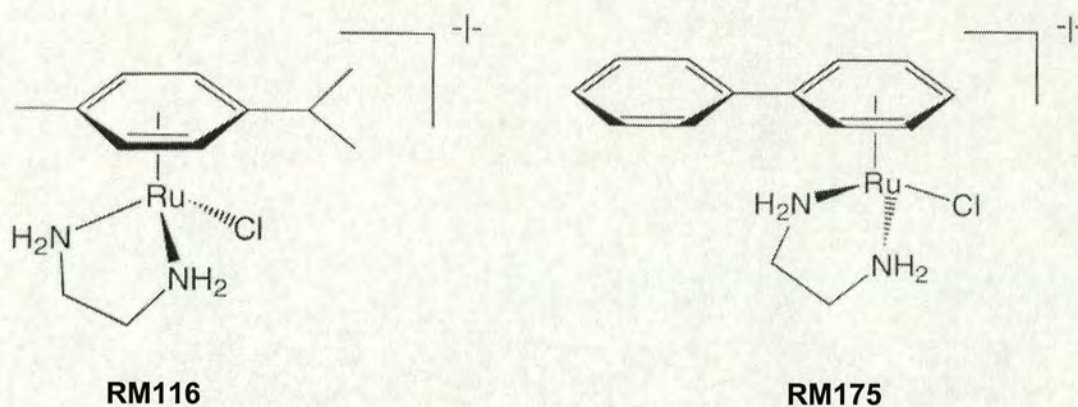
RM116   RM175

Figure 5.17 Structures of the cations in the Ru^{2+} arene complexes RM116 and RM117, which contain p-cymene and biphenyl ligands respectively. PF$_6$ acts as the counter ion to both complexes.

5.3.1.2 Experimental

The binding of RM116 and RM175 (synthesised by Haimei Chen) to rHA was investigated by HPLC. The thiol content of the recombinant human albumin (Delta Biotechnology Ltd.) used was determined via a standard DTNB assay (see Section 2.3) as 0.7 mol mol^{-1}.

The HPLC was carried out using a Shodex KW-803 porous silica 8 mm x 300 mm gel permeation column on a Hewlett-Packard 1100 series HPLC. Detection was measured at 320 nm as both complexes absorb at this wavelength. Note that the protein is not detected at this wavelength. The mobile phase was the same as that used for the albumin quantification assay, Section 2.2.2. By this method, the protein separates from the complex due to its large size, and binding can be seen as a decrease in the free complex.

Solutions containing various concentrations of albumin in the SH-form (from 13.1 µM to 105.0 µM) and RM116 (52.5 µM) in water were incubated for 24 h at 37˚C. Aliquots (100 µl) of each reaction were loaded onto the column. The above reactions were repeated with various concentrations of free thiol human albumin (from 12.6 µM to 100.8 µM) and RM175 (50.4 µM).

5.3.1.3 Results and Discussion

The resultant HPLC chromatograms show the decrease of free RM116 and RM175, respectively in the presence of increasing concentrations of albumin (Figure 5.18). RM116 is present as a single peak with a retention time of *ca.* 13.7 min in A. RM175 appears to be represented by two peaks with retention times of 13.5 and 15.0 minutes respectively in B. The first of these two peaks is most likely to be a phosphate complex (Dr Fuyi Wang, personal communication). From these results it is clear that both RM116 and RM175 bind to human albumin.

Further work carried out in our laboratory by a BSc project student, Frank Toner using ICP-AES (inductively coupled plasma atomic emission spectrometry) to measure albumin-bound ruthenium gave K_a values of 2.7×10^9 M^{-1} and 6.2×10^9 M^{-1} for RM116 and RM175 respectively. He also determined the respective B_{max} values of 4.98 nmol mg^{-1} and 3.35 nmol mg^{-1} (Toner, unpublished data). These K_a and B_{max} values are only valid if the ruthenium compounds bind albumin at a single site. It is possible that either of the compounds may bind at more than one site with different affinities. Toner also investigated competitive binding of both ruthenium compounds with DTNB, see Section 4.2. He found that the DTNB greatly inhibited binding of RM116 to albumin but had no effect upon the binding of RM175 to albumin (Toner, unpublished data). This result

suggests that Cys[34] is the main binding site for RM116. It is possible that RM175 binds to either or both of binding sites I and II. These sites are often thought of as being the main binding sites for small organic compounds and are described in more detail in Section 1.3.3. RM175 is likely to be too large to be able to fit into the hydrophobic crevice associated with the Cys[34] site because of the extra phenyl group, not found in RM116.
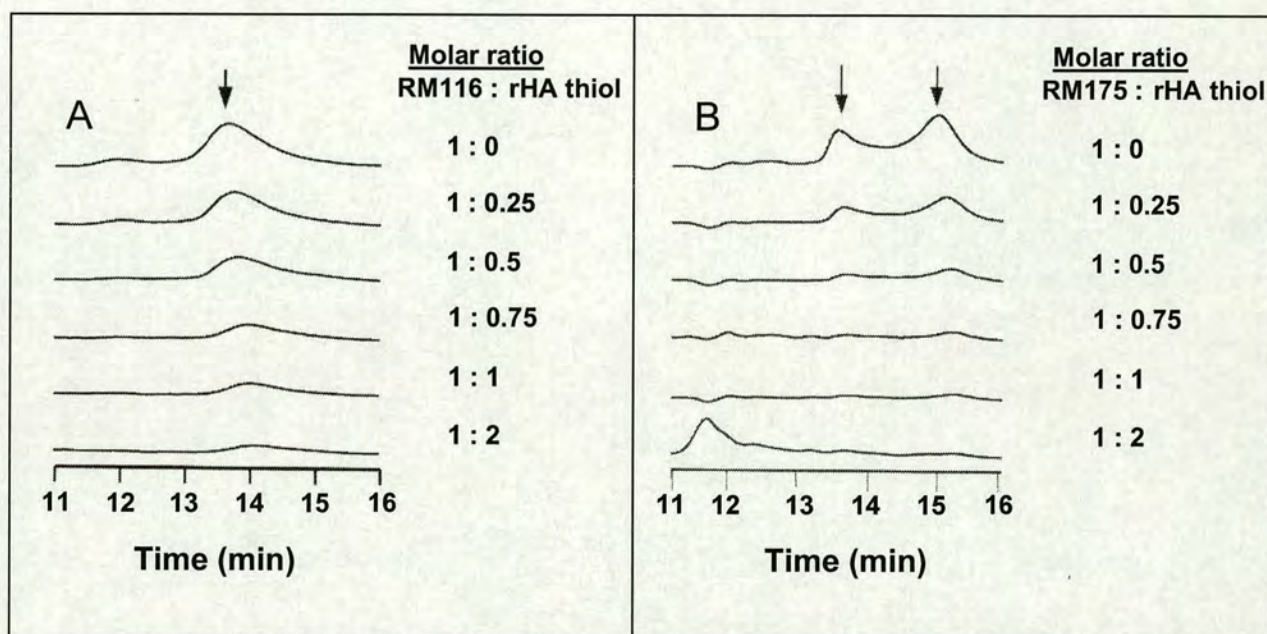
Figure 5.18 HPLC analysis of ruthenium complex binding. A and B show the decrease in free RM116 and RM175 respectively upon addition of of albumin.

5.3.2 Binding of an aminophosphine platinum(II) complex to albumin

5.3.2.1 Introduction to [Pt(NP$_3$)Cl]Cl

Aminophosphine Pt(II) complexes can bind to the DNA bases guanine and thymine *via* ring-opening reactions and are cytotoxic to cancer cells [52,53]. The target for cytotoxic platinum complexes is usually thought to be DNA, but they can also react with many

other biomolecules such as the sulfur-containing amino acids methionine and cysteine. The cysteine-containing tripeptide glutathione (γ-L-Glu-L-Cys-Gly, GSH) is known to interact with cisplatin and its analogues and competes with nucleobases for platinum binding [54]. Additionally, sulfur-donor ligands can be used as rescue agents to minimise the toxic side effects of these therapeutics [55].

The square-planar platinum(II) complex, [Pt(NP$_3$)Cl]Cl, where NP$_3$ = N(CH$_2$CH$_2$PPh$_2$)$_3$, has previously been shown to react with thiol-containing biomolecules such as N-acetyl-L-cysteine (NAC) and GSH [56] and to generate new strong absorption bands which are potential probes for such interactions. The aim of this study was to determine whether this platinum complex shows affinity toward the free thiol of albumin.

5.3.2.1 Experimental

The complex [Pt(NP$_3$)Cl]Cl has previously been shown to have an extinction coefficient of 20700 M^{-1} cm^{-1} at 256 nm. The complex upon binding to cysteine has extinction coefficients of 1775.0, 3567.5 and 10020 M^{-1} cm^{-1} at 500, 398 and 325 nm, respectively [56].

A small amount of [Pt(NP$_3$)Cl]Cl (synthesised by Inés García-Seijo) was dissolved in methanol, the concentration of the solution was determined spectrophotometrically to be 1.38 mM. This solution was diluted to 230 μM with methanol and 300 μL was reacted with 300 μL of 15 μM rHA in aqueous 100 mM Tris-HCl, pH 7.4 at 25 °C. The albumin thiol content was determined to be 0.65 mol mol^{-1} as described in Section 2.3. The formation of an albumin-complex product was observed at 398 nm against a blank

containing 230 µM $[Pt(NP_3)Cl]Cl$ in 50:50 methanol:100 mM Tris-HCl, pH 7.4 using a Shimadzu UV250 1PC spectrophotometer.

5.3.2.2 Results and Discussion

The absorbance at 398 nm was found to increase over time indicating binding of $[Pt(NP_3)Cl]Cl$ to albumin, Figure 5.19. If binding only at Cys^{34} is assumed then under the conditions described, with $[Pt(NP_3)Cl]Cl$ in high excess, the reaction should follow pseudo-first order kinetics and a rate constant could be calculated. This assumption would mean that only one molecule of complex could bind per free thiol group. Using the extinction coefficient for cysteine-complex binding (at 398 nm), it seems unlikely that the above assumption is correct, as the change in absorbance is too large. After 70 minutes, 81.6 nmol of Cys^{34}-complex would have to be present. This is not possible, as only approximately 1.5 nmol of free-thiol containing albumin is initially present.

The change in the UV spectrum indicates that the complex does bind to albumin. However, the value of the extinction coefficient suggests that binding at Cys^{34} is only a minor contributor to this change in absorbance (if at all) and that a different site on albumin is involved in binding of the complex.
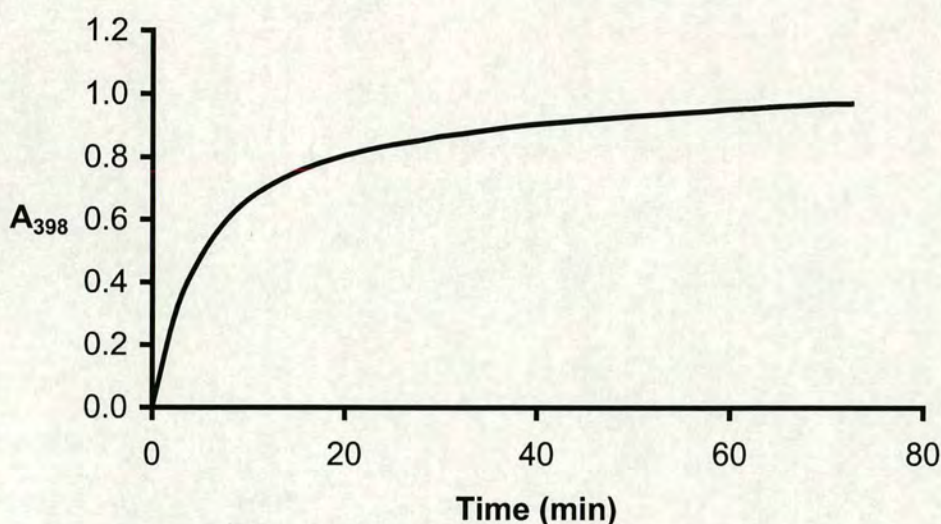
Figure 5.19 Reaction of rHA with 20 mol eqiv [Pt(NP₃)Cl]Cl per free thiol in 100 mM Tris-HCl, pH 7.4. As monitored by a change in absorbance at 398 nm.

5.4 Conclusions

Albumin is the major transport protein in blood for Zn^{2+}, a metal ion required for a variety of physiological processes and which various drugs and toxins recruit. It has been proposed that zinc binding to albumin involves at least 2 histidine ligands, but no structural model of the site had been reported. Here Zn^{2+} and other metal ion binding to recombinant human albumin have been studied through competition with $^{111}Cd^{2+}$ using ^{111}Cd NMR.

Analysis of the X-ray structures of unliganded human albumin in the PDB led to the identification of only one potential metal site containing at least two His residues. It lies at the interface of domains I and II, and two additional potential ligands (Asp and Asn) are situated nearby, poised to form a 5-coordinate site with a low molecular weight ligand such as water or Cl⁻ in the fifth coordination site. Weak Cl⁻ binding to Cd^{2+} in this site

was detected for rHA. The amino acids contributing to this site are highly conserved in mammalian albumins. The trigonal bipyramidal site is apparently "preformed" in unliganded albumin, although none of the reported X-ray structures appear to contain a metal ion bound here. Mutation of His67 to Ala disrupted the binding of both Cd^{2+} and Cu^{2+}, known competitors for the major zinc site.

Analysis of the available X-ray crystal structures of albumin and the [111]Cd NMR data presented here suggest that fatty acids disrupt metal binding to site A. This triggers a movement of two of the proposed ligands (His247 and Asp249) by about 5 Å away from His67 and Asn99, the two other proposed protein ligands in the Zn^{2+} site. It will now be of interest to explore the possible consequences of the interactive binding of Zn^{2+} and fatty acids to albumin. The proposed switching of the zinc site in human albumin by fatty acid binding is an intriguing example of an allosteric interaction between an organic nutrient and an essential metal ion. It might be possible to exploit this in drug design and therapy.

Albumin binding studies involving two cytotoxic ruthenium(II) complexes ($[(\eta^6$-p-cymene)RuCl$(H_2NCH_2CH_2NH_2$-N,N)$]^+PF_6^-$ and $[(\eta^6$-$C_6H_5C_6H_5)$RuCl$(H_2NCH_2CH_2NH_2$-N,N)$]^+PF_6^-$) and a novel platinum(II) complex ($[Pt(NP_3)Cl]Cl$) previously shown to have affinity toward low-molecular weight thiol compounds were carried out. Gel permeation chromatography showed that peaks corresponding to the free Ru(II) complexes, decreased upon the addition of albumin, indicative of binding. The absorbance of albumin at 398 nm following the addition of $[Pt(NP_3)Cl]Cl$ increased, showing binding to have occurred.

References: Chapter 5

1.   Peters, T., Jr. (1996). All about albumin: Biochemistry, genetics, and medical applications. Academic Press, New York.

2.   McLean, F. C. and Hastings, A. B. (1934). A biological method for the estimation of calcium ion concentration. *J. Biol. Chem.*, **107**: 337-350.

3.   Masuoka, J., Hegenaur, J., Van dyke, B. R. and Saltman, P. (1993). Intrinsic stoichiometric equilibrium constants for the binding of zinc(II) and copper(II) to the high affinity site of serum albumin. *J. Biol. Chem.*, **268**: 21533-21537.

4.   Berlow, P. P., Burton D. J. and Routh J. I. (1974). *Introduction to the Chemistry of Life.* Saunders College Publishing, New York.

5.   Stearns, D. M. (2000). Is chromium a trace essential metal? *Biofactors*, **11**: 149-162.

6.   Andersson, R. A (1997). Chromium as an essential nutrient for humans. *Regul. Toxicol. Pharmacol.*, **26**: S35-41.

7.   Morris, B., Gray, T. and MacNeil, S. (1995). Evidence for chromium acting as an essential trace element in insulin-dependent glucose uptake in cultured mouse myotubes. *J. Endocrinol.*, **144**: 135-141.

8.   Guo, Z. and Sadler, P. J. (1999). Metals in Medicine. *Angew. Chem. Int. Ed.*, **38**: 1512-1531.

9.   Pedersen, K. O (1972). Binding of calcium to serum albumin. III. Influence of ionic strength and ionic medium. *Scand. J. Clin. Lab. Invest.*, **29**: 427-432.

10.   Vallee, B. L. and Falchuk, K. H. (1993). The biochemical basis of zinc physiology. *Physiol. Rev.*, **73**: 79-118.

11.   Rowe, D. J. and Bobilya, D. J. (2000). Albumin facilitates zinc acquisition by endothelial cells. *Proc. Soc. Exp. Biol. Med.*, **224**: 178-186.

12.   Goumakos, W., Laussac, J. P., Sarkar, B. (1991). Binding of cadmium(II) and zinc(II) to human and dog serum albumins. An equilibrium dialysis and ^{113}Cd-NMR study. *Biochem. Cell Biol.*, **69**: 809-820.

13.   Sadler, P. J. and Viles, J. H. (1996). ^{1}H and ^{113}Cd NMR investigations of Cd^{2+} and Zn^{2+} binding sites on serum albumin: competition with Ca^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+}. *Inorg. Chem.*, **35**: 4490-4496.

14.   Nandedkar, A. K., Hong, M. S. and Friedberg, F. (1974). Co^{2+} binding by plasma albumin. *Biochem. Med.*, **9**: 177-183.

15.   Bian, H.-D., Liang, H., Tu, C., Zhang, H. and Shen, P. (2000). Binding equilibrium studies between Co^{2+} and HSA or BSA. *Chem. Res. Chinese Uni.*, **16**: 276-279.

16.   Sadler, P. J., Tucker, A. and Viles, J. H. (1994). Involvement of a lysine residue in the N-terminal Ni^{2+} and Cu^{2+} binding site of serum albumins. Comparison with Co^{2+}, Cd^{2+} and Al^{3+}. *Eur. J. Biochem.*, **220**: 193-200.

17.   Scheuhammer, A. M. and Cherian, M. G. (1985). Binding of manganese in human and rat plasma. *Biochim. Biophys. Acta*, **840**: 163-169.

18.   Fatemi, S. J., Kadhir, F. H. and Moore, G. R. (1991). Aluminium transport in blood serum. Binding of aluminium by human transferring in the presence of human albumin and citrate. *Biochem. J.*, **280**: 527-532.

19.   Nandedkar, A. K. N., Nurse, C. E. and Friedberg, F. (1973). Mn^{++} binding by plasma proteins. *Int. J. Pept. Protein Res.*, **5**: 279-281.

20.   Hughes, W. L. and Dintzis, H. M. (1954). Crystallization of the mercury dimers of human albumin and bovine mercatalbumin. *J. Biol. Chem.*, **239**: 845-849.

21.   Shaw III, C. F. (1989). The protein chemistry of antiarthritic gold (I) thiolates and related complexes. *Comments Inorg. Chem.*, **8**: 233-267.

22.   Forestier, J. (1929).   La chrysotherapie dans les rhumatismes chroniques. *Bull. Mem. Soc. Med. Hop. Paris*, **53**: 323-329.

23.   Roberts, J. R., Xiao, J., Schliesman, B., Parsons, D. J. and Shaw III, C. F. (1996). Kinetics and mechanism of reaction between serum albumin and auranofin (and its isopropyl analogue) *in vitro*. *Inorg. Chem.*, **35**: 424-433.

24.   Ivanov, A. I., Christodoulou, J., Parkinson, J. A., Barnham, K. J., Tucker, A., Woodrow, J. and Sadler, P. J. (1998). Cisplatin binding sites on human albumin. *J. Biol. Chem.*, **273**: 14721-14730.

25.   Gelasko, A. and Lippard, S. J. (1998).   NMR solution structure of a DNA dodecamer duplex containing a *cis*-diammineplatinum(II) d(GpG) intrastrand cross-link, the major adduct of the anticancer drug cisplatin. *Biochemistry*, **37**: 9230-9239.

26.   Espinosa, E., Feliu, J., Zamora, P., González Barón, M., Sánchez, J. J., Ordóñez, A. and Espinosa, J. (1995). Serum albumin and other prognostic factors related to response and survival in patients with advanced non-small cell lung cancer. *Lung Cancer*, **12**: 67-76.

27.   Holding, J. D., Lindup, W. E., van Laer, C., Vreeburg, G. C., Schilling, V., Wilson, J.A. and Stell, P. M. (1992). Phase I trial of a cisplatin-albumin complex for the treatment of cancer of the head and neck. *Br. J. Clin. Pharmacol.*, **33**: 75-81.

28.   Molkentin, J. D. (2000). The zinc finger-containing transcription factors GATA-4, -5, and -6. Ubiquitously expressed regulators of tissue-specific gene expression. *J. Biol. Chem.*, **275**: 38949-38952.

29.   Lipscomb, W. N. and Sträter, N. (1996). Recent advances in zinc enzymology. *Chem. Rev.*, **96**: 2375-2433.

30.   Karlin, S. and Zhu, Z.-Y. (1997). Classification of mononuclear zinc metal sites in protein structures. *Proc. Natl. Acad. Sci. USA*, **94**: 14231-14236.

31.   Tibaduiza, E. C. and Bobilya, D. J. (1996). Zinc transport across an endothelium includes vesicular cotransport with albumin. *J. Cell. Physiol.*, **167**: 539-547.

32.   Martins, E. O. and Drakenberg, T. (1982). Cadmium(II), zinc(II), and copper(II) ions binding to bovine serum albumin. A cadmium-113 NMR study. *Inorg. Chim. Acta*, **67**: 71-74.

33.   Viles, J. H., Tucker, A., Patel, S. U. and Sadler, P. J. (1996). ^{1}H and ^{113}Cd NMR studies of metal binding to isolated albumin and albumin in blood plasma. *Bull. Magn. Reson.*, **18**: 182-183.

34.   Shaka, A. J., Barker, P. B. and Freeman, R. (1985). Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.*, **64**: 547-552.

35.   Bal, W., Christodoulou, J., Sadler, P. J. and Tucker , A. (1998). Multi-metal binding site of serum albumin. *J. Inorg. Biochem.*, **70**: 33-39.

36.   Pettit, L. D., Pyburn, S., Bal, W., Kozlowski, H. and Bataille, M. (1990). A study of the comparative donor properties to copper(II) of the terminal amino and imidazole nitogen in peptides. *J. Chem. Soc. Dalton. Trans.*, 3565-3570.

37.   Zhang, Y. and Wilcox, D. E. (2002). Thermodynamic and spectroscopic study of Cu(II) and Ni(II) binding to bovine serum albumin. *J. Biol. Inorg. Chem.*, **7**: 327-337.

38.   Öz, G., Pountney, D. L. and Armitage, I. M. (1998). NMR spectroscopic studies of $I = \frac{1}{2}$ metal ions in biological systems. *Biochem. Cell Biol*, **76**: 223-234.

39.   Pettit, G. and Pettit, L. D. (1993). *IUPAC Stability Constants Database*, IUPAC and Academic Software, Otley, UK.

40.   Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999). Crystal structure of human serum albumin at 2.5 Å resolution. *Prot. Eng.*, **12:** 439-446.

41.   Harding, M. M. (2001). Geometry of metal-ligand interactions in proteins. *Acta Cryst.*, **D57:** 401-411.

42.   Klimpel, K. R., Arora, N. and Leppla, S. H. (1994). Anthrax toxin lethal factor contains a zinc metalloprotease consensus sequence which is required for lethal toxin activity. *Mol. Microbiol.*, **13:** 1093-1100.

43.   Håkansson, M., Antonsson, P., Björk, P. and Svensson, L. A. (2001). Cooperative zinc binding in a staphylococcal enterotoxin A mutant mimics the SEA-MHC class II interaction. *J. Biol. Inorg. Chem.*, **6:** 757-762.

44.   Katz, B. A. and Luong, C. (1999). Recruiting Zn^{2+} to mediate potent, specific inhibition of serine proteases. *J. Mol. Biol.*, **292:** 669-684.

45.   Janc, J. W., Clark, J. M., Warne, R. L., Elrod, K. C., Katz, B. A. and Moore, W. R. (2000). A novel approach to serine protease inhibition: kinetic characterization of inhibitors whose potencies and selectivities are dramatically enhanced by zinc(II). *Biochemistry*, **39:** 4792-4800.

46.   Cohn, E. J., Strong, L. E., Hughes, W. L., Jr., Mulford, D. J., Ashworth, J. N., Melin, M. and Taylor, H. L. (1946). Preparation and properties of serum and plasma proteins. IV. A system for the separation into fractions of the protein and lipoprotein components and biological tissues and fluids. *J. Am. Chem. Soc.*, **68:** 459-475.

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

48.   Petitpas, I., Grune, T., Bhattacharya, A. A. and Curry, S. (2001). Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *J. Mol. Biol.*, **314:** 955-960.

49.   Bergamo, A., Cocchietto, M., Capozzi, I., Mestroni, G., Alessio, E. and Sava, G. (1996). Treatment of residual metastases with Na[trans-RuCl$_4$(DMSO)Im] and ruthenium uptake by tumour cells. *Anti-Cancer Drugs*, **7:** 697-702.

50.   Messori, L., Orioli, P., Vullo, D., Alessio, E. and Lengo, E. (2000). A spectroscopic study of the reaction of NAMI, a novel ruthenium(III) anti-neoplastic complex, with bovine serum albumin. *Eur. J. Biochem.*, **267:** 1206-1213.

51.   Morris, R. E., Aird, R. E., Murdoch, P. S., Chen, H., Cummings, J., Hughes, N. D., Parsons, S., Parkin, A., Boyd, G., Jodrell, D. I. and Sadler, Peter J. (2001). Inhibition of cancer cell growth by ruthenium(II) arene complexes. *J. Med. Chem.*, **44:** 3616-3621.

52.   Margiotta, N., Habtemariam, A. and Sadler, P. J. (1997). Strong, rapid binding of a platinum complex to thymine and uracil under physiological conditions. *Angew. Chem. Int. Ed. Engl.*, **36:** 1185-1187.

53.   Habtemariam, A., Watchman, B., Potter, B. S., Palmer, R., Parsons, S., Parkin, A. and Sadler, P. J. (2001). Control of aminophosphine chelate ring-opening in Pt(II) and Pd(II) complexes: Potential dual-mode anticancer agents. *Dalton Trans.*, 1306-1318.

54.   Reedijk, J. (1996). Improved understanding in platinum antitumor chemistry. *Chem Comm.*, 801-806.

55.   Reedijk, J. (1999). Why does cisplatin reach guanine-N7 with competing S-donor ligands available in the cell? *Chem. Rev.*, **99:** 2499-2510.

56.   García-Seijo, M. I. (2000). *PhD Thesis*, University of Santiago de Compostela,

Spain.

Chapter 6

Crystallisation of rHA

6.1 Protein Crystallisation

6.1.1 Introduction

The main method of elucidating the 3-dimensional structure of a molecule as large as human albumin is by X-ray diffraction. As previously mentioned in Chapter 1, several X-ray crystal structures of human albumin have been elucidated, in both free [1,2] and fatty acid bound forms [3-5] and are summarised in Table 1.3. However, to date no structures of albumin have been solved of albumin with either a ligand bound at Cys^{34}, a metal bound, or of a mutant albumin. The main aim of this chapter was to develop a method of crystallisation so that such structures could be elucidated.

6.1.2 Crystallisation Theory

The main hurdle in solving the structure of a protein as large as albumin, is the production of a highly diffracting (<3.0 Å resolution) protein crystal. The general processes by which proteins crystallise are similar to that of salts and small organic compounds. To grow crystals, molecules must be brought into a thermodynamically unstable state known as supersaturation. This is usually achieved by the gradual removal of solvent and proceeds via three stages: nucleation, growth, and cessation of growth.

Nucleation is the process by which molecules or non-crystalline aggregates that are free in solution come together in such a way as to produce a thermodynamically stable aggregate with a repeating lattice. Crystallisation is known to lower the free energy of proteins by approximately 3-6 kcal mol^{-1} relative to the solution state [6]. The formation of crystalline aggregates from supersaturated solutions does not necessarily lead to the formation of crystals. The aggregate first needs to exceed a specific size (the critical size) defined by the competition of the ratio of the surface area of the aggregate to its volume [7,8]. Once this critical size is reached, the aggregate becomes a nucleus capable of further growth. If the nucleus decreases in size so that it is smaller than the critical size, dissolution occurs. The process of formation of crystalline aggregates and non-crystalline precipitates does not involve the competition between surface area and volume and so can occur much faster than crystallisation.

The degree to which nucleation occurs is determined by the degree of supersaturation of the solutes in the solution. The extent of supersaturation is in turn related to the overall solubility of the potentially crystallising molecule. The solubility of a protein can be described by a phase diagram (Figure 6.1). Higher solubility allows for a greater number of diffusional collisions and so higher degrees of supersaturation produce more stable aggregates (due to the higher probability of collision between diffusing molecules) and so increase the chances of the formation of stable nuclei. When there is a finite number of solute molecules, the production of a large number of small crystals usually results. At lower solute concentrations, the formation of individual stable nuclei increases in rarity, thus favoring the formation of single crystals [9].
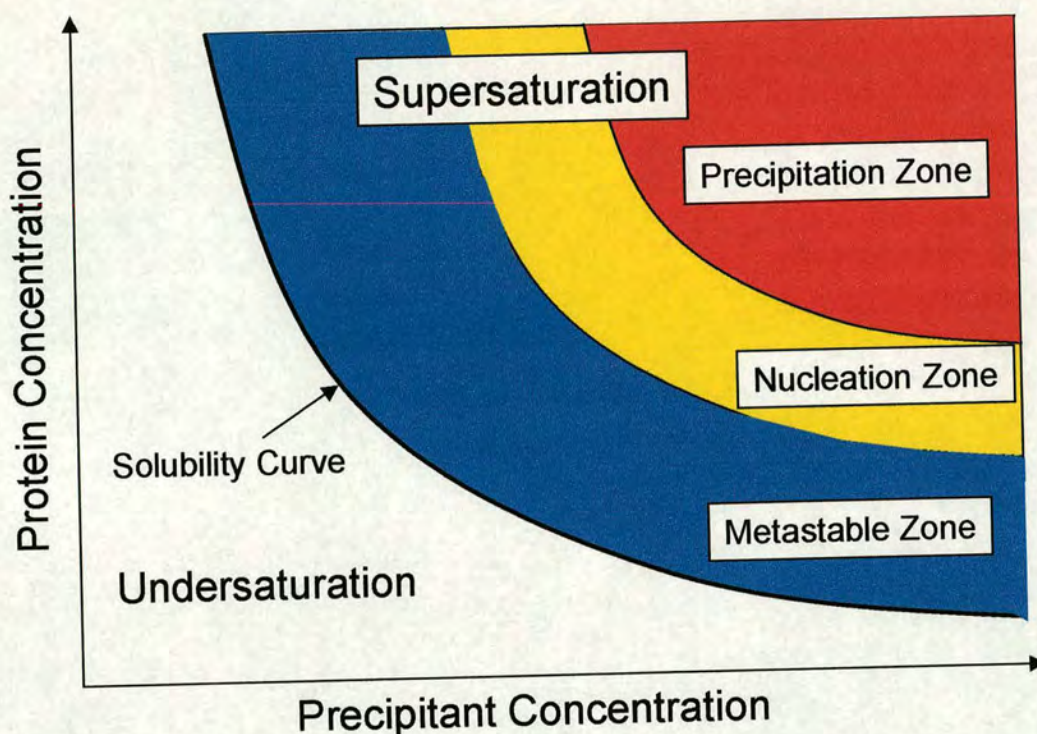
Figure 6.1 Typical phase diagram of a protein. The solubility curve divides undersaturated from supersaturated areas. On the curve, proteins are in equilibrium between solution and crystal, they will not crystallise in the area of undersaturation. The supersaturated region is divided into three zones. The metastable zone is where the solution may not nucleate for a long time but this zone will sustain growth. It is frequently necessary to add a seed crystal. The nucleation zone is where protein crystals nucleate and grow. The precipitation zone is where proteins do not nucleate or grow in this region but precipitate out of solution.

Crystal growth generally starts at solute concentrations sufficient for nucleation to occur, and can continue at concentrations lower than the nucleation threshold [9]. The purity of the protein is one of the most important factors in crystal growth. Proteins need not be absolutely pure to give crystals, although crystals from impure solutions frequently give poor diffraction data. To obtain good quality crystals, proteins need

not only to be reasonably free of contaminating proteins, but be "conformationally pure" [9]. Albumin is a particularly difficult protein to crystallise due to its heterogeneity and flexibility. Albumin preparations contain a significant amount of dimer. Also when isolated from plasma, approximately 28% of albumin molecules carry a half-cystine and 7% a half-glutathione as a mixed disulfide on Cys^{34} [10]. An important purity issue is that all traces of protease must be removed from the samples, as incubation times are frequently long (days to months). Lin and coworkers recommend the use of FPLC (or HPLC) for general purification of proteins and to assure homogeneity on both macroscopic and microscopic levels [11]. Lin notes several cases where the use of FPLC techniques has improved the reproducibility of crystallization as well as the maximum resolution to which protein crystals diffract. In order to combat these purity issues, only highly pure recombinant human albumin (supplied by Delta Biotechnology Ltd.) was used in the experiments described. Other factors, which can affect crystal growth, include pH and temperature [9].

Cessation of growth of crystals can occur for a multitude of reasons. The most obvious is the decrease in concentration of the crystallising solute to the point where the solid and solution phases reach exchange equilibrium. In this case, the addition of more solute can result in continued crystal growth. However, some crystals reach a certain size beyond which growth does not proceed irrespective of the solute concentration. This may be a result either of lattice strain effects or poisoning of the growth surface [9].

Lattice strain effects in tetragonal lysozyme crystals have been demonstrated by Feher and Kam [7]. Halved crystals of hen egg white lysozyme, when placed in fresh crystallisation solutions, grew to exactly the same size as the original crystal. This suggests that the strain in the lattice effectively prevents addition of molecules to the

surface once a certain critical volume is reached. Crystals affected by lattice strain are therefore size limited.

Poisoning of growing faces occurs when foreign or damaged molecules are incorporated into the growing crystal face resulting in successive defects that interrupt the crystal lattice. An example of this might be the incorporation of a proteolytically knicked protein onto the face of an otherwise perfect protein crystal. If the nicked molecule is unable to form the same lattice contacts with newly added molecules, as would the perfect protein, then its incorporation will cause defects in the growing lattice. Since the growth of crystal lattices typically selects for perfect molecules (rather than damaged or incorrect molecules) the concentration of these defective molecules relative to perfect molecules tends to increase as growth proceeds. Thus, as crystals grow larger the likelihood of incorporation of defective molecules into the lattice increases (also the increase in surface area contributes). Sato and coworkers have used laser scattering tomography to identify lattice defects in large crystals of orthogonal hen egg white lysozyme [12]. Their results showed the occurrence of defects not only at the surface, but also within the bulk of the crystal itself.

6.1.3 Methods of Protein Crystallisation

As previously mentioned, crystal growth is usually achieved through the gradual removal of solvent from a protein solution. A common method of doing this is by vapour diffusion. In vapour diffusion methods a droplet of protein solution containing precipitant is equilibrated with a reservoir of precipitant of higher concentration so that water preferentially evaporates from the drop. If conditions are right this will produce

a gradual increase in protein and precipitant concentrations so that a few crystals may form. The most common of these are the hanging and sitting drop methods. Representations of these are shown in Figure 6.2.
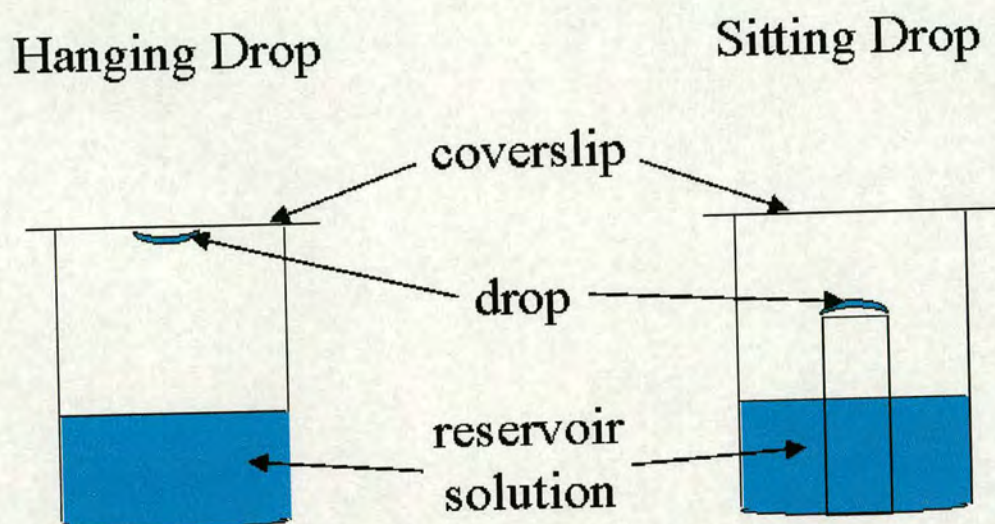

Figure 6.2 Diagrammatic representation of hanging and sitting drop vapour diffusion. Both techniques utilise a reservoir, containing a buffered precipitate solution. The drop also contains the same components as the reservoir (usually at half the concentration) as well as the protein to be crystallised.

6.2 Crystallisation Experiments

In the first crystallisation experiment, the crystallisation method of Sugio *et al.* was adopted [2]. This is a standard hanging drop procedure. Sugio and coworkers managed to crystallise albumin from a drop containing 150-255 mg mL^{-1} protein, 50 mM potassium phosphate, (pH 5.0-5.5), 30-38% polyethylene glycol 400 and 5 mM sodium azide over 2-3 weeks. This crystallisation attempt as well as all others was set up in 26-well Linbro plates (ICN). Reservoir solutions were set up as shown in Appendix 3. A 2 µL volume of 200 mg mL^{-1} rHA (all albumins in this Chapter were

supplied by Delta Biotechnolgy Ltd. and were de-fatted as described in Section 2.4.1

prior to use) was mixed with 2 µL of each reservoir solution. These mixtures formed

the drops, which were spotted onto the siliconised coverslips (Molecular Dimensions

Ltd.). Each coverslip was sealed to the top of the well with vacuum grease as shown

in Figure 6.2. The plate was stored at 18 °C. No protein crystals grew over a period of

months. No precipitation was seen.

The second crystallisation experiment followed the method of Carter *et al.* [13]. This

method again utilises the hanging drop procedure. Reservoir solutions were set up as

shown in Appendix 4. A 2 µL volume of 200 mg mL^{-1} rHA was mixed with 2 µL of

each reservoir solution to form the drops. The plate was stored at 18 °C. No protein

crystals grew over a period of months. Precipitation was seen in some of the wells.

Crystallisation experiment 3 was slightly more successful. The hanging drop method

was adopted and the plate was set up as shown in Appendix 5. A 5 µL volume of 150

mg mL^{-1} rHA was mixed with 5 µL of each reservoir solution to form the drops. The

plate was stored at 4 °C. After approximately 2 weeks a tetragonal crystal was found

in well B5. The X-ray diffraction pattern of the crystal was recorded by Prof.

Malcolm Walkinshaw (Institute of Cell and Molecular Biology, University of

Edinburgh). It was found to diffract X-rays only to 6 Å resolution.

The poor resolution of the X-ray data may have been due to a number of factors.

Because the recombinant albumin used is relatively homogeneous the most likely of

these are due to its dimer content and "conformational purity" (a consequence of the

albumin's highly flexible structure). Crystallisation was next attempted using the

method of Curry [3-5]. It was thought that the presence of fatty acid within the structure of albumin would help to rigidify the structure and lead to the formation of higher quality X-ray diffracting crystals.

In this experiment rHA, was loaded with myristic acid (C14:0) as described in Section 2.4. This time the sitting drop method was adopted and the plate was set up as shown in Appendix 6. A 5 μL volume of 110 mg mL^{-1} myristate-rHA was mixed with 5 μL of each reservoir solution to form the drops. The drops were placed in the sitting drop microbridges (Crystal Microsystems) and as with the hanging drops, coverslips were sealed on the top of each well with vacuum grease. The plate was stored at 4 °C. Only very small rod-shaped crystals grew after a couple of weeks, too small for diffraction studies. The size of the crystals may be limited by lattice strain or by surface poisoning.

Crystallisation of native rHA was attempted with a higher molecular weight polyethylene glycol, PEG 6000 by the hanging drop method. The plate was set up with polyethylene glycol 6000 varying from 1-10% in each well and 50 mM potassium phosphate, pH 7.5. A 5 μL volume of 50 mg mL^{-1} rHA was mixed with 5 μL of each reservoir solution to form the drops. The albumin in each of the drops precipitated out of solution almost immediately.

The next crystallisation experiment used a crystal screen (Structure Screen 1, Molecular Dimensions Ltd.). Such screens are often used to determine initial crystallisation conditions and to establish the solubility of a macromolecule in a varying range of pH and precipitants. They take the form of a number of solutions

selected from previously-published works in which crystallisation of a protein has been achieved.

The screen contained 50 different solutions. Of each solution 1 mL was used as the reservoir solution for both hanging and sitting drop experiments. A 5 μL volume of 100 mg mL^{-1} rHA was mixed with 5 μL of each reservoir solution to form the drops. The plates were stored at 4 °C. A summary of the solutions used and the results of the experiment are shown in Appendix 7. The sitting drop well set up with solution 34 had an aggregate present, which appeared to be crystalline.

Well 34 had present 30% PEG 400, 0.1 M Tris pH 8.5 and 0.2 M tri-sodum citrate dihydrate. A fresh sitting drop plate was set up with buffer and precipitant concentrations varied around those above as shown in Appendix 8. A 5 μL volume of 100 mg mL^{-1} rHA was mixed with 5 μL of each reservoir solution to form the drops. The plate was stored at 4 °C. No crystals grew.

Up to this point the only experiment from which a suitable crystal was obtained for diffraction studies was experiment 3. Further crystallisation experiments utilised the conditions set up in this experiment. Crystallisation experiments were set up in order to grow crystals of the His^{39}Leu and Tyr^{84}Phe rHA mutants. Prior to crystallisation the mutant proteins were defatted as described in Section 2.4. The plate was set up as shown in Appendix 9. A 5 μL volume of 150 mg mL^{-1} the respective mutant rHA was mixed with 5 μL of each reservoir solution to form the drops. The His^{39}Leu mutant protein was present in drops in colums 1-3 on the plate and the Tyr^{84}Phe in 4-6. The plate was stored at 4 °C. After approximately 2 weeks tetragonal crystals were found

in both His[39]Leu and Tyr[84]Phe mutant-containing wells. The crystals ranged in size

from *ca.* 0.3 mm across to 1.0 mm across and 0.1 mm deep.

After harvesting, the crystals were frozen in liquid nitrogen and X-ray diffraction was

measured by Steve Robinson in the Department of Structural Biochemistry, University

of Edinburgh. Data were collected for 40 min using a mar345 image plate at 400 mm

from the crystal on a Nonius 571 rotating anode X-ray generator. The X-ray source

was a copper anode, wavelength = 1.542 Å. It was found that only 3 of crystals (A, B

and C) produced X-ray diffraction patterns all of which were of the His[39]Leu mutant.

The space group of the unit cell of crystal A was primitive monoclinic with

dimensions of a=76.08 Å, b=185.88 Å and c=182.85 Å and angles of α=90.47°,

β=93.62° and γ=87.78°. The a, b and c values indicate the length of the unit cell in

each of the three dimensions (or axes). The metric tensor distortion index = 0.93%,

which agrees with 1AO6 [2] in the PDB fairly well (Table 6.1). The best resolution

was 5.55 Å, but was expected be complete only to 7 Å or so if collected. The
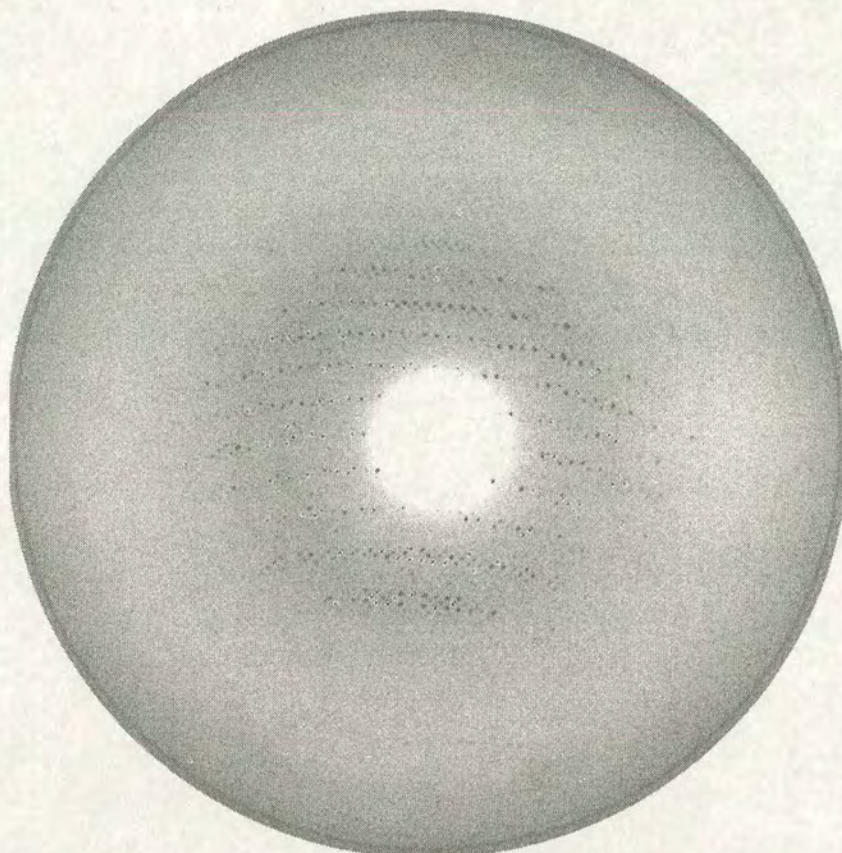
diffraction pattern is shown in Figure 6.3.

Figure 6.3 Diffraction pattern of crystal A. Note that a hole is cut in the film to reduce the overexposure caused by the very intense beam compared to the diffracted beams. Diffracted X-rays can be seen as black spots on the film.

The space group of the unit cell of crystal B was primitive monoclinic with dimensions of a=75.35 Å, b=182.74 Å and c=180.95 Å and angles of α=91.85°, β=94.21° and γ=92.51°. The metric tensor distortion index = 1.33% and the best resolution following diffraction is 5.45 Å, but was expected to be complete to less if collected. The diffraction pattern is shown in Figure 6.4.
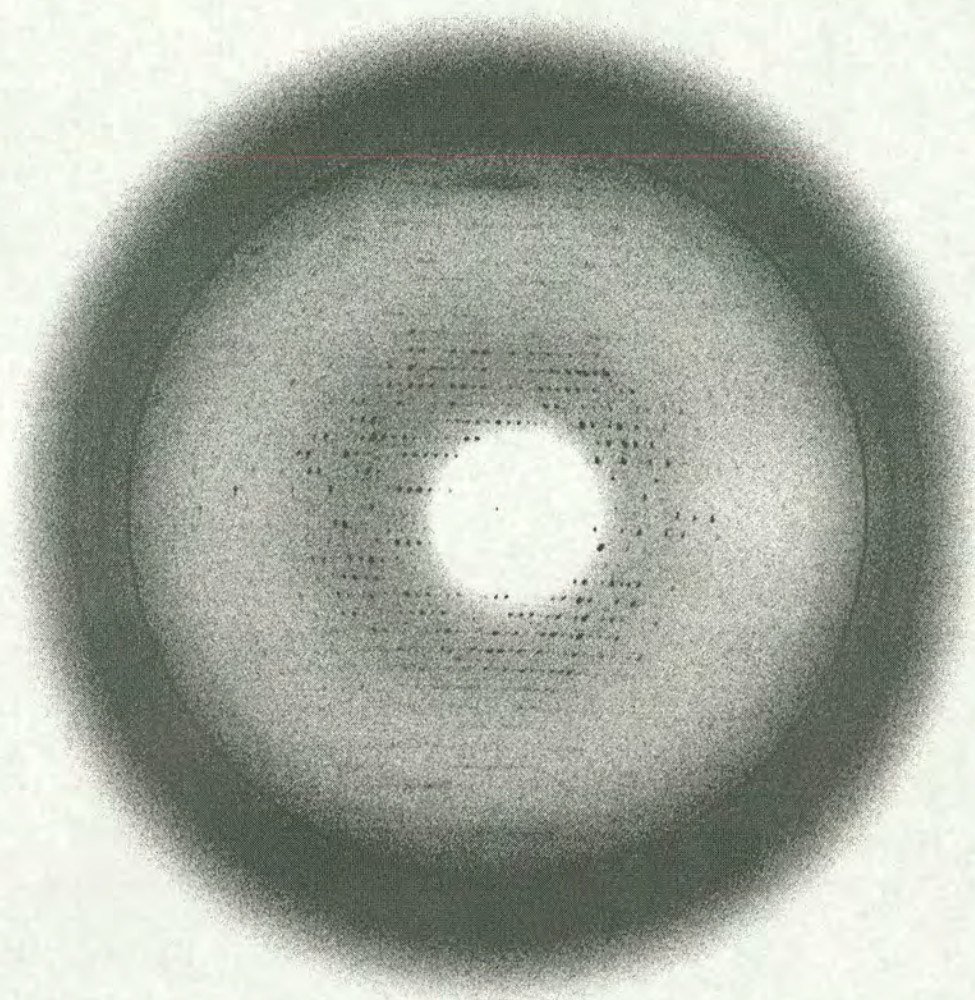
Figure 6.4 Diffraction pattern of crystal B.

The space group of the unit cell of crystal C was primitive orthorhombic with dimensions of a=158.66 Å, b=179.87 Å and c=184.60 Å and angles of α=90.17°, β=89.69° and γ=88.09°. The metric tensor distortion index = 0.80% and the best resolution following diffraction is 7.90 Å, but will likely be complete only to 10 Å or so if collected. The diffraction pattern is shown in Figure 6.5.
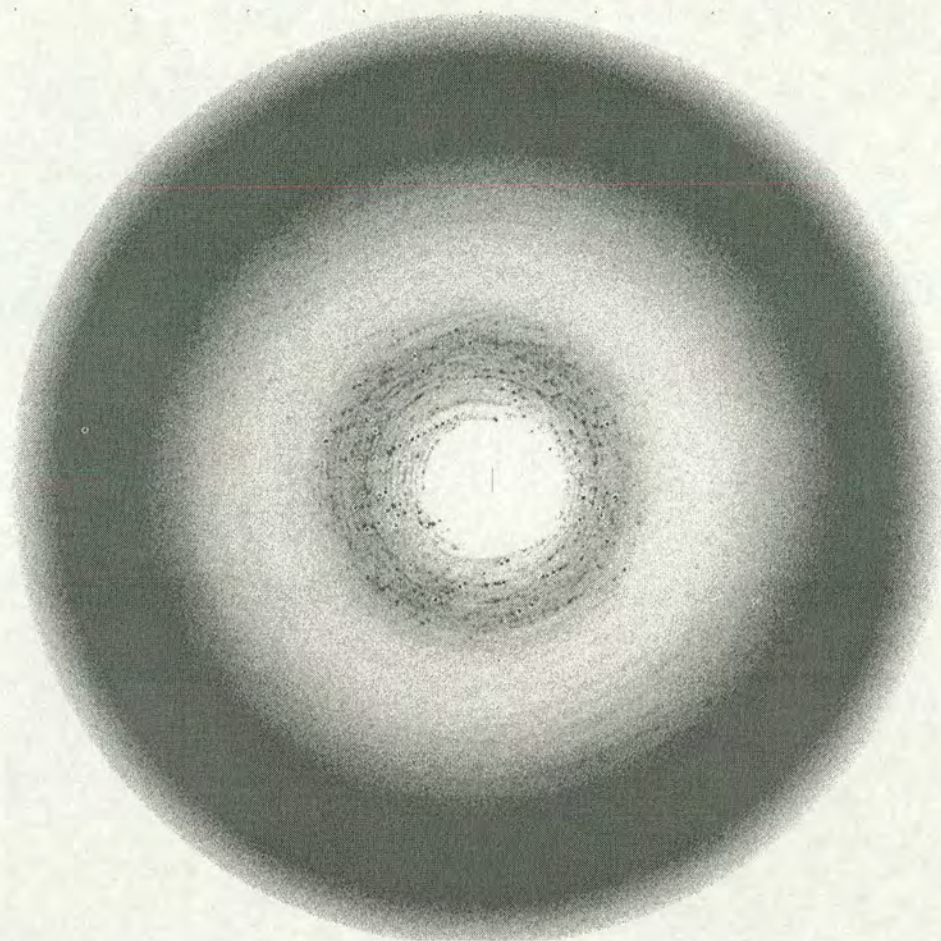
Figure 6.5 Diffraction pattern of crystal C.

The poor resolution of the crystals again is most likely due to the presence of dimer or

of poor "conformational purity" of the protein because the recombinant albumin used

is of the highest purity.

An attempt was also made to crystallise albumin with Cu^{2+} bound at the N-terminus of

albumin from a number of mammalian species. A plate for a hanging drop experiment

was set up with reservoir solutions containing 40% PEG 400 and 50 mM potassium

phosphate, pH 5.5. In these experiments 1 mol eqiv of $CuCl_2$ was added to a 150 mg

mL^{-1} solution of the following fatty-acid-free albumins: rHA, bovine (Sigma, A-0281),

murine (Sigma, A-1056), porcine (Sigma, A-1173) and canine albumins (Sigma, A-3184).

A 5 μL volume of 150 mg mL^{-1} of the respective albumin-Cu complex was mixed with 5 μL of each reservoir solution to form the drops. On the plate, the drops on the coverslips of wells in columns 1 and 2 contained rHA, column 3 contained bovine, 4 murine, 5 porcine and column 6 contained canine albumins. The plate was incubated at 4 °C. After 2 weeks, tiny yellowish rod-like crystals formed in all the drops except those containing canine albumin. The result is interesting, as it is well known that dog albumin has poor affinity toward copper as they lack the His[3] (Table 1.2) which provides 2 N ligands for Cu^{2+} binding at the N-terminal site. The crystals were re-incubated at 4 °C but unfortunately even after a period of months never became large enough for diffraction studies to be carried out.

References: Chapter 6

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

2. Sugio, S., Kashima, A., Mochizuki, S., Noda, M. and Kobayashi, K. (1999). Crystal structure of human serum albumin at 2.5 Å resolution. *Prot. Eng.*, **12:** 439-446.

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

4. Bhattacharya, A. A., Grune, T. and 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.

5. Petitpas, I., Grune, T., Bhattacharya, A. A. and Curry, S. (2001). Crystal structures of human serum albumin complexed with monounsaturated and polyunsaturated fatty acids. *J. Mol. Biol.*, **314:** 955-960.

6. Drenth, J. and Haas, C. (1992). Protein crystals and their stability. *J. Cryst. Growth.*, **122:** 107-109.

7. Feher, G. and Kam, Z. (1985). Nucleation and growth of protein crystals: general principles and assays. *Meth. Enzymol.*, **114:** 77-111.

8. Boistelle, R. and Astier, J. P. (1988). Crystallization mechanisms in solution. *J. Cryst. Growth.*, **90:** 14-30.

9. Berry, M. B. (1995). *Ph.D. Thesis*, Rice University, Texas, U.S.A.

10. Andersson, L. O. (1966). The heterogeneity of bovine serum albumin. *Biochim. Biophys. Acta.*, **117:** 115-133.

11.   Lin, S.-X., Sailofsky, B., Lapointe, J., Zhou, M. (1992).   Preparative fast purification procedure of various proteins for crystallization.   *J. Cryst. Growth.*, **122:** 242-245.

12.   Sato, K., Fukuba, Y., Mitsuda, T., Hirai, K., and Moriya, K. (1992). Observation of lattice defects in orthorhombic hen-egg white lysozyme crystals with laser scattering tomography.   *J. Cryst. Growth.*, **122:** 87-94.

13.   Carter, D. C., He, X. M., Munson, S. H., Twigg, P. D., Gernert, K. M., Broom, M. B. and Miller, T. Y. (1989).   Three-dimensional structure of human serum albumin.   *Science*, **244:** 1195-1198.

14.   Appleton, D. W. and Sarkar, B. (1971).   The Absence of Specific Copper (II)-binding Site in Dog Albumin.   *J. Biol Chem.*, **246:** 5040-5045.

Appendices

Appendix 1

The Increase in Absorbance at 412 nm with Time and a Plot of Log($[I]^0$/[I]) vs Time in the Determination of the Pseudo-First Order Rate Constant for the Reactions Described in Section 4.1

As discussed in Section 4.1.1, for the reactions under the conditions described, a plot of log ($[I]^0$/[I]) against time gives a straight line with a slope equal to the pseudo-first order rate constant (where [I] = albumin thiol concentration and $[I]^0$ = initial albumin thiol concentration).

For the reaction of rHA with DTNB at pH 8.0, $[I]^0$ = 4.1 x 10^{-6} M. This value is calculated from the final absorbance, once all the albumin molecules containing a free thiol group have been oxidised. Oxidation of Cys^{34} with DTNB results in the stoichiometric (1:1) production of TNB, as the DTNB is reduced (see Section 2.3). Therefore the production of TNB and indeed thiol oxidation can be quantified using the extinction coefficient of TNB, ε = 13.9 units mM^{-1} cm^{-1} at 412 nm. The absorption at 412 nm during this reaction is shown in Figure A1.

The [I] value at each recorded timepoint during the reaction was calculated in a similar way as $[I]^0$, in order to plot log ($[I]^0$/[I]) against time. As the reaction approached the end, this plot yielded a straight line as shown in Figure A2.
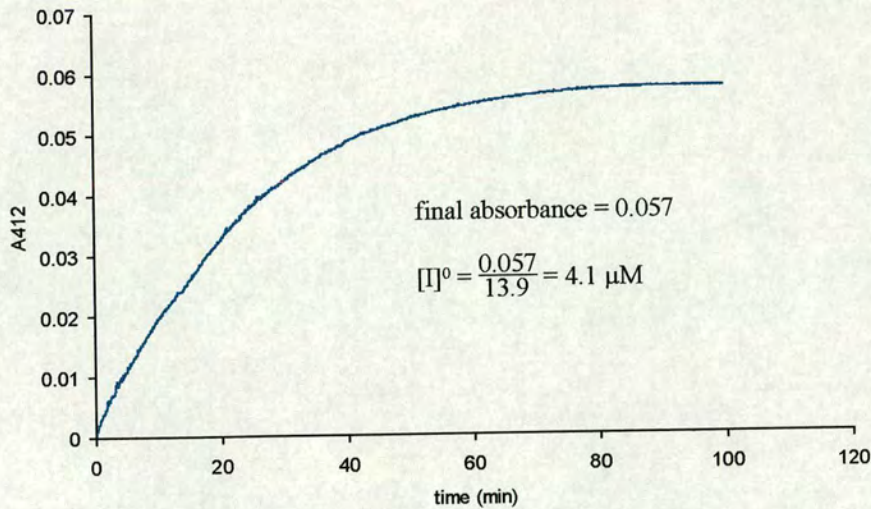
Figure A1 The absorption at 412 nm during the reaction between rHA and DTNB at pH 8.0.
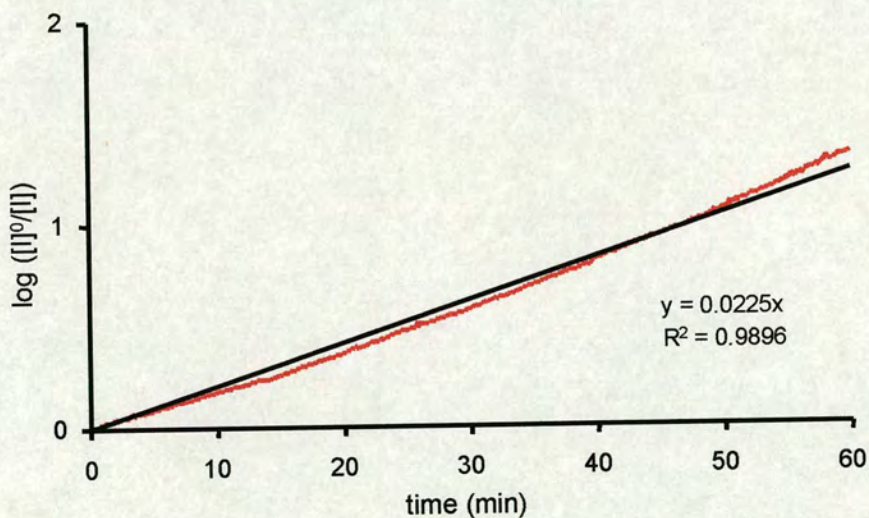
Figure A2 A plot of log ([I]^0/[I]) against time for the reaction between rHA and DTNB at pH 8.0. The black line shows a trendline for the data shown in red. The trendline is a fit of the data to a straight line, allowing a gradient (and therefore the rate constant) to be easily calculated. The resultant equation and R^2 value of this line are displayed.

The gradient of the trendline = 0.0225, therefore the pseudo-first order rate constant for this reaction = 2.25×10^{-2} min^{-1}.

Appendix 2

Non-Logarithmic Plots of Rate Constant vs pH for the reaction of rHA, high octanoate rHA and His[39]Leu and Tyr[84]Phe mutant albumins with 100 mol equiv of DTNB. As described in Section 4.2.2.

rHA

rHA (high octanoate)
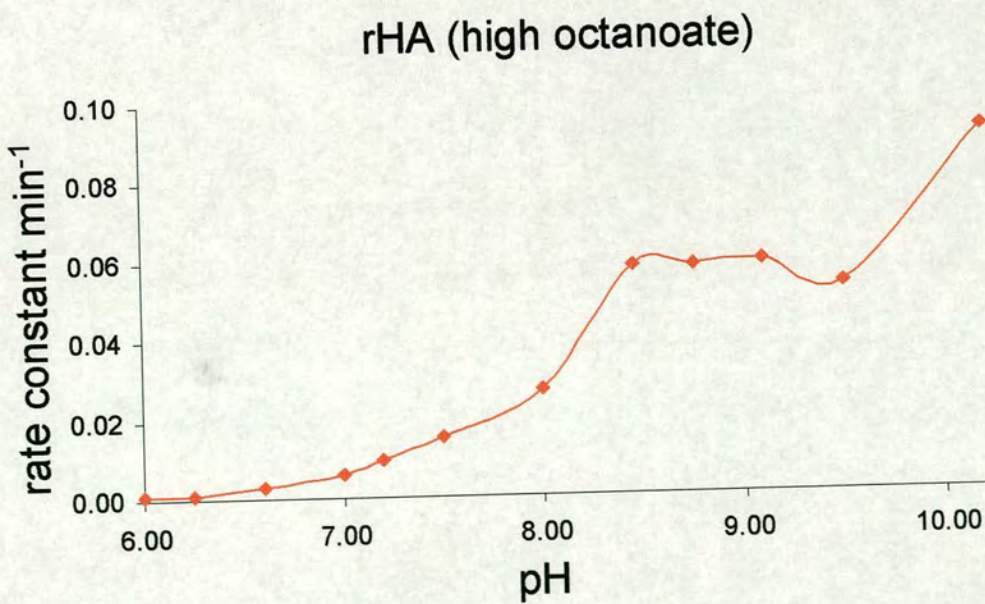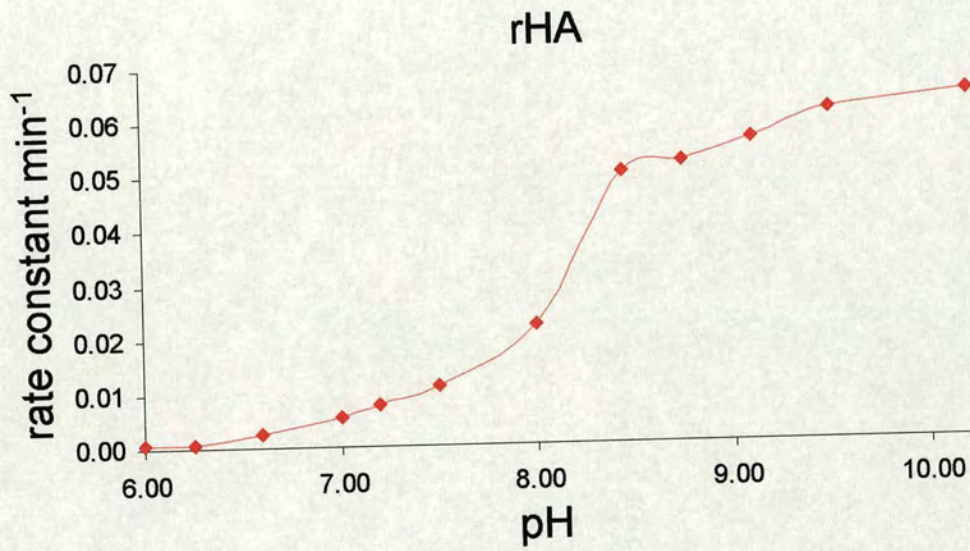
His³⁹Leu
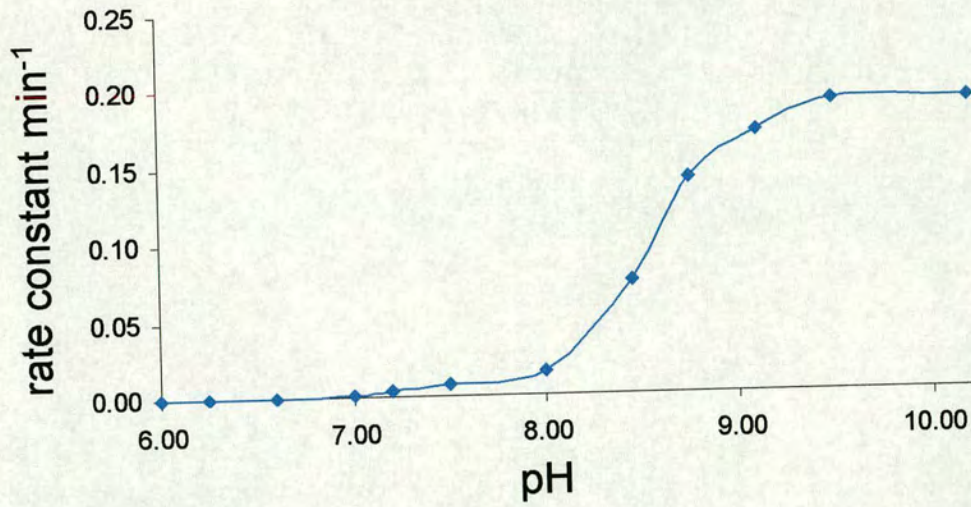
Tyr⁸⁴Phe
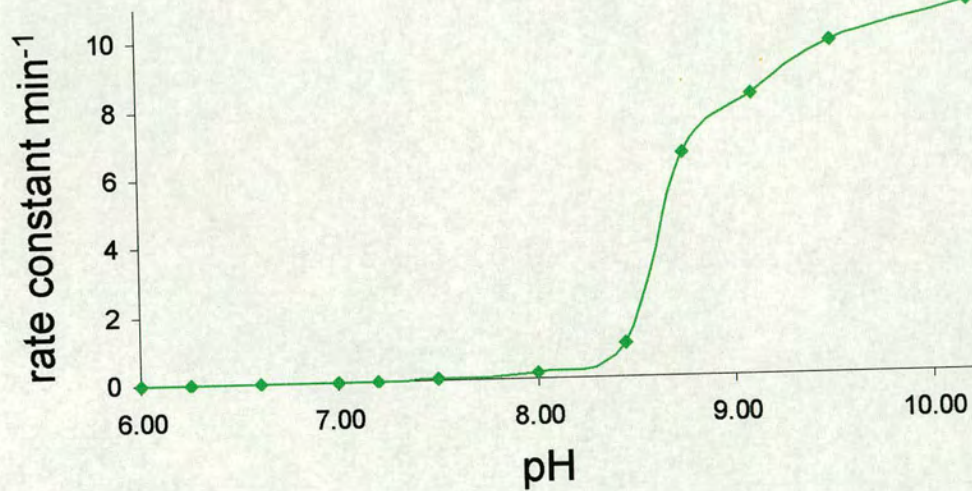
The 4 plots show the effects of pH on the pseudo-first order rate constants in the redox reaction between DTNB and each of the albumins.

Appendix 3

Recombinant Human Albumin Crystallisation Experiment 1

60% PEG 400 pH 5.0	64% PEG 400 pH 5.0	68% PEG 400 pH 5.0	72% PEG 400 pH 5.0	76% PEG 400 pH 5.0	80% PEG 400 pH 5.0
60% PEG 400 pH 5.2	64% PEG 400 pH 5.2	68% PEG 400 pH 5.2	72% PEG 400 pH 5.2	76% PEG 400 pH 5.2	80% PEG 400 pH 5.2
60% PEG 400 pH 5.4	64% PEG 400 pH 5.4	68% PEG 400 pH 5.4	72% PEG 400 pH 5.4	76% PEG 400 pH 5.4	80% PEG 400 pH 5.4
60% PEG 400 pH 5.6	64% PEG 400 pH 5.6	68% PEG 400 pH 5.6	72% PEG 400 pH 5.6	76% PEG 400 pH 5.6	80% PEG 400 pH 5.6

All wells contained polyethylene glycol (PEG) 400 at the concentrations indicated, 50 mM potassium phosphate at the pH values indicated and 5 mM sodium azide.

Appendix 4

Recombinant Human Albumin Crystallisation Experiment 2

20% PEG 4000 pH 7.0	20% PEG 4000 pH 7.2	20% PEG 4000 pH 7.4	20% PEG 4000 pH 7.6	20% PEG 4000 pH 7.8	20% PEG 4000 pH 8.0
22% PEG 4000 pH 7.0	22% PEG 4000 pH 7.2	22% PEG 4000 pH 7.4	22% PEG 4000 pH 7.6	22% PEG 4000 pH 7.8	22% PEG 4000 pH 8.0
24% PEG 4000 pH 7.0	24% PEG 4000 pH 7.2	24% PEG 4000 pH 7.4	24% PEG 4000 pH 7.6	24% PEG 4000 pH 7.8	24% PEG 4000 pH 8.0
26% PEG 4000 pH 7.0	26% PEG 4000 pH 7.2	26% PEG 4000 pH 7.4	26% PEG 4000 pH 7.6	26% PEG 4000 pH 7.8	26% PEG 4000 pH 8.0

All wells contained polyethylene glycol (PEG) 4000 at the concentrations indicated, 50 mM potassium phosphate at the pH values indicated and 5 mM sodium azide.

Appendix 5

Recombinant Human Albumin Crystallisation Experiment 3

40% PEG 400 pH 6.8	40% PEG 400 pH 6.8	40% PEG 400 pH 6.8	40% PEG 400 pH 6.8	40% PEG 400 pH 6.8	40% PEG 400 pH 6.8
40% PEG 400 pH 5.5	40% PEG 400 pH 5.5	40% PEG 400 pH 5.5	40% PEG 400 pH 5.5	40% PEG 400 pH 5.5	40% PEG 400 pH 5.5

The 12 wells contained 40% polyethylene glycol (PEG) 400 and 50 mM potassium phosphate at the pH values indicated.

Appendix 6

Recombinant Human Albumin Crystallisation Experiment 4

25% PEG 3350 pH 7.5	25% PEG 3350 pH 7.5	25% PEG 3350 pH 7.5	25% PEG 3350 pH 7.5	25% PEG 3350 pH 7.5	25% PEG 3350 pH 7.5
26% PEG 3350 pH 7.5	26% PEG 3350 pH 7.5	26% PEG 3350 pH 7.5	26% PEG 3350 pH 7.5	26% PEG 3350 pH 7.5	26% PEG 3350 pH 7.5
27% PEG 3350 pH 7.5	27% PEG 3350 pH 7.5	27% PEG 3350 pH 7.5	27% PEG 3350 pH 7.5	27% PEG 3350 pH 7.5	27% PEG 3350 pH 7.5
28% PEG 3350 pH 7.5	28% PEG 3350 pH 7.5	28% PEG 3350 pH 7.5	28% PEG 3350 pH 7.5	28% PEG 3350 pH 7.5	28% PEG 3350 pH 7.5

The wells contained polyethylene glycol (PEG) 400 at the concentrations indicated and 50 mM potassium phosphate, pH 7.5.

Appendix 7

Summary of Solutions and Results of the Crystal Screen (Exp. 5).

Solution No.	Buffer	Precipitant	Additional component	HD	SD
1	0.1 M Na acetate trihydrate pH 4.6	30% 2-methyl-2,4 pentanediol	0.02 M calcium chloride dihydrate	P	P
2	0.1 M Na acetate trihydrate pH 4.6	30% PEG 4000	0.2 M ammonium acetate	P	P
3	0.1 M Na acetate trihydrate pH 4.6	25% PEG 4000	0.2 M ammonium sulphate	P	P
4	0.1 M Na acetate trihydrate pH 4.6	2.0 M sodium formate	None	N	N
5	0.1 M Na acetate trihydrate pH 4.6	2.0 M ammonium sulphate	None	P	P
6	0.1 M Na acetate trihydrate pH 4.6	8% PEG 4000	None	N	N
7	0.1 M tri-sodium citrate dihydrate pH 5.6	30% PEG 4000	0.2 M ammonium acetate	P	P
8	0.1 M tri-sodium citrate dihydrate pH 5.6	30% 2-methyl-2,4-pentanediol	0.2 M ammonium acetate	N	N
9	0.1 M tri-sodium citrate dihydrate pH 5.6	20% 2-propanol, 20% PEG 4000	None	N	P
10	0.1 M Na citrate pH 5.6	1.0 M ammonium dihydrogen phosphate	None	N	N
11	0.1 M Na acetate trihydrate pH 4.6	20% 2-propanol	0.2 M calcium chloride dihydrate	N	N
12	0.1 M Na cacodylate pH 6.5	1.4 M Na acetate trihydrate	None	N	N
13	0.1 M Na cacodylate pH 6.5	30% 2-propanol	0.2 M tri-sodium citrate dihydrate	N	N
14	0.1 M Na cacodylate pH 6.5	30% PEG 8000	0.2 M ammoniium sulphate	P	P
15	0.1 M Na cacodylate pH 6.5	20% PEG 8000	0.2 M magnesium acetate tetrahydrate	P	P
16	0.1 M Na cacodylate pH 6.5	30% 2-methyl-2,4-pentanediol	0.2 M magnesium acetate tetrahydrate	N	N

17	0.1 M imidazole pH 6.5	1.0 M sodium acetate trihydrate	None	N	N
18	0.1 M Na cacodylate pH 6.5	30% PEG 8000	0.2 M sodium acetate trihydrate	P	P
19	0.1 M Na cacodylate pH 6.5	18% PEG 8000	0.2 M zinc acetate dihydrate	P	P
20	0.1 M Na cacodylate pH 6.5	18% PEG 8000	0.2 M calcium acetate hydrate	N	N
21	0.1 M Na hepes pH 7.5	30% 2-methyl-2,4-pentanediol	0.2 M tri-sodium citrate dihydrate	N	N
22	0.1 M Na hepes pH 7.5	30% 2-propanol	0.2 M magnesium chloride hexahydrate	N	N
23	0.1 M Na hepes pH 7.5	28% PEG 400	0.2 M calcium chloride dihydrate	N	N
24	0.1 M Na hepes pH 7.5	30% PEG 400	0.2 M magnesium chloride hexahydrate	N	N
25	0.1 M Na hepes pH 7.5	20% 2-propanol	0.2 M tri-sodium citrate dihydrate	N	N
26	0.1 M Na hepes pH 7.5	0.8 M K, Na tartrate tetrahydrate	None	N	N
27	0.1 M Na hepes pH 7.5	1.5 M lithium sulphate monohydrate	None	N	N
28	0.1 M Na hepes pH 7.5 and 0.8M K dihydrogen phosphate monohyd.	0.8 M Na dihydrogen phosphate	None	N	N
29	0.1 M Na hepes pH 7.5	1.4 M tri-sodium citrate dihydrate	None	N	N
30	0.1 M Na hepes pH 7.5	2% PEG 400, 2.0 M amm sulphate	None	N	N
31	0.1 M Na hepes pH 7.5	10% 2-propanol, 20% PEG 4000	None	P	P
32	0.1 M tris HCl pH 8.5	2.0 M ammoniium sulphate	,None	N	N
33	0.1 M tris HCl pH 8.5	30% PEG 4000	0.2 M magnesium chloride hexahydrate	P	P
34	0.1 M tris HCl pH 8.5	30% PEG 400	0.2 M tri-sodum citrate dihydrate	N	A
35	0.1 M tris HCl pH 8.5	30% PEG 4000	0.2 M lithium sulphate monohydrate	P	P

36	0.1 M tris HCl pH 8.5	30% 2-propanol	0.2 M ammonium acetate	N N
37	0.1 M tris HCl pH 8.5	30% PEG 4000	0.2 M sodium acetate trihydrate	N P
38	0.1 M tris HCl pH 8.5	8% PEG 8000	None	N N
39	0.1 M tris HCl pH 8.5	2.0 M ammonium dihydrogen phosphate	None	N N
40	None	0.4 M K, Na tartrate tetrahydrate	None	N N
41	None	0.4 M ammonium dihydrogen phosphate	None	N N
42	None	30% PEG 8000	0.2 M ammonium sulphate	P P
43	None	30% PEG 4000	0.2 M ammonium sulphate	P P
44	None	2.0 M ammonium sulphate	None	P P
45	None	4.0 M sodium formate	None	N N
46	None	20% PEG 8000	0.05 M potassium dihydrogen phosphate	P P
47	None	30% PEG 1500	None	N P
48	None	0.2 M magnesium formate	None	N N
49	None	2% PEG 8000	1.0 M lithium sulphate monohydrate	N N
50	None	15% PEG 8000	0.5 M lithium sulphate monohydrate	N N

Abbreviations used: HD, hanging drop; SD, sitting drop; A, possible crystalline aggregate; N, no crystal; P, protein precipitated.

Appendix 8

Recombinant Human Albumin Crystallisation Experiment 6

25% PEG 400 pH 7.5	27% PEG 400 pH 7.5	29% PEG 400 pH 7.5	31% PEG 400 pH 7.5	33% PEG 400 pH 7.5	35% PEG 400 pH 7.5
25% PEG 400 pH 8.0	27% PEG 400 pH 8.0	29% PEG 400 pH 8.0	31% PEG 400 pH 8.0	33% PEG 400 pH 8.0	35% PEG 400 pH 8.0
25% PEG 400 pH 8.5	27% PEG 400 pH 8.5	29% PEG 400 pH 8.5	31% PEG 400 pH 8.5	33% PEG 400 pH 8.5	35% PEG 400 pH 8.5
25% PEG 400 pH 8.5	27% PEG 400 pH 8.5	29% PEG 400 pH 8.5	31% PEG 400 pH 8.5	33% PEG 400 pH 8.5	35% PEG 400 pH 8.5

The wells contained polyethylene glycol (PEG) 400 at the concentrations indicated, 50 mM potassium phosphate at the pH values indicated and 0.2 M tri-sodium citrate.

Appendix 9

Recombinant Human Albumin Crystallisation Experiment 7

20% PEG 400 pH 7.2	20% PEG 400 pH 7.4	20% PEG 400 pH 7.6	20% PEG 400 pH 7.2	20% PEG 400 pH 7.4	20% PEG 400 pH 7.6
22% PEG 400 pH 7.2	22% PEG 400 pH 7.4	22% PEG 400 pH 7.6	22% PEG 400 pH 7.2	22% PEG 400 pH 7.4	22% PEG 400 pH 7.6
24% PEG 400 pH 7.2	24% PEG 400 pH 7.4	24% PEG 400 pH 7.6	24% PEG 400 pH 7.2	24% PEG 400 pH 7.4	24% PEG 400 pH 7.6
26% PEG 400 pH 7.2	26% PEG 400 pH 7.4	26% PEG 400 pH 7.6	26% PEG 400 pH 7.2	26% PEG 400 pH 7.4	26% PEG 400 pH 7.6

The wells contained polyethylene glycol (PEG) 400 at the concentrations indicated and 50 mM potassium phosphate at the pH values indicated.