Misfolded Forms of Hen Egg White Lysozyme

Sarah Elizabeth Barron

Thesis presented for the degree of Doctor of Philosophy

Department of Chemistry

University of Glasgow

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

Acknowledgments

I wish to dedicate this piece of work to my grandmother Mrs. Lilias Barron, who never got to write her thesis.

I would like to thank Alan Cooper for allowing me to do this research and for all his help and encouragement. There are a number of people who I must thank for helping to provide some of the data presented in this thesis: Andy Pitt of Strathclyde University for use of the electrospray mass spectrometry facility; Sharon Kelly and Nick Price of Stirling University for their circular dichroism work; Alex Last and Carol Robinson of Oxford University for their hydrogen exchange work and electrospray mass spectrometry. I must also thank the Crystallography group at Glasgow University who helped when I was having crystallography and chromatography problems, especially Jeremy Beauchamp. Thank you to Fiona McMonagle for her advice.

Thanks to everyone in the lab who have prevented me from going insane in times of great stress: Andrew Heron, Lindsay McDermott, Margaret Nutley, Samantha Rutherford and Tennie Videler.

I thank the EPSRC for funding this project for the past three years.

And of course thank you to my parents, who are responsible for my existence in the first place, and my sister Suzanne, who have been a constant source of support.

Abstract

Although thermal unfolding of globular proteins is in principle reversible, in practice this is rarely true due to various factors. In the case of Hen Egg White Lysozyme (HEWL) we have evidence that thermal cycling through the normal unfolding transition gives rise to possible misfolded monomeric forms of the protein.

Repeat Differential Scanning Calorimetry (DSC) scans of HEWL suggests the presence of misfolded forms of the protein, the amount of which increases with time or number of scans. This is indicated by peaks at a lower temperature than the main transition peak. Approximately three peaks are visible in this region. The total number of peaks suggest that they may be due to the 4 possible proline isomers of HEWL. The properties and structure of these misfolded forms have been investigated using various biophysical techniques.

Analysis has shown that the heat treated HEWL (HT HEWL) is very similar to native HEWL except that it shows a uncooperative unfolding transition when denatured using GdnHCl and it binds ANS, a hydrophobic probe. The fact that it binds ANS indicated that the HT HEWL had exposed hydrophobic patches and that this aspect of its behaviour would allow separation of the HT HEWL into the four misfolded species of which it is composed.

HT HEWL was successfully separated into four species named L1 - 4. DSC of these species showed that they were the species identified in the original DSC experiment. Again these separate species show similarities to native HEWL by electrospray mass spectrometry (ESMS), sodium dodecyl sulphide polyacrylamide gels (SDS PAGE) and circular dichroism (CD), the only identifiable differences are in the stability of these species towards denaturants and temperature, and a decrease in activity when compared to native HEWL.

L1 - 4 when reheated and when left for a time period are interconvertible. This is the main evidence that there is to date that these species are proline isomers. Interestingly, L2 appears to behave like native HEWL in that it has the same elution profile and thermal stability, but this behaviour does not extend to its activity or chemical stability.

The species identified by this work are possible proline isomers of HEWL, as suggested by their interconvertibility.

Contents

Figures

Tables

Abbreviations

2`CMP	2` cytidine monophosphate
ANS	anilinonapthalene sulfonate
APP	Alzheimers precursor protein
BSE	bovine spongiform encephalopathie
CD	circular dichroism
CFTR	cystic fibrosis transmembrane regulator
CI2	chymotrypsin inhibitor 2
CJD	Creutzfeld Jacob disease
CSP	cold shock protein
DSC	differential scanning calorimetry
EDTA	ethylenediaminetetraacetic acid
ESMS	electrospray mass spectrometry
GdnHCl	guanidinium hydrochloride
HEWL	hen egg white lysozyme
HIC	hydrophobic interaction chromatography
HPLC	high performance liquid chromatography
HT HEWL	heat treated hen egg white lysozyme
HSQC	heteronuclear single quantum correlation
IEF	isoelectic focussing
IR	infra red spectroscopy
ISP	isomeric specific proteolysis
Ksv	Stern Volmer constant
L1	fraction of separated heat treated lysozyme which corresponds to peak 1 on chromatogram
L2	fraction of separated heat treated lysozyme which corresponds to peak 2 on chromatogram
L3	fraction of separated heat treated lysozyme which corresponds to peak 3 on chromatogram
L4	fraction of separated heat treated lysozyme which corresponds to peak 4 on chromatogram
NMR	nuclear magnetic resonance
NOESY	nuclear overhauser enhancement spectroscopy

pI	isoelectric point
PDI	protein disulphide isomerase
PMSF	phenylmethylsulfonylfluoride
PPI	peptidyl prolyl isomerase
Rnase A	ribonuclease A
SDS	sodium dodecyl sulphide
TEMED	N`,N`,N`,N`, tetramethylethylenediamine
TGFβ	transforming growth factor β
Tm	transition midpoint
Us	slow folding
UV	ultra violet spectroscopy

Chapter 1: Introduction

Introduction

Protein misfolding is now recognized to be a major problem in biotechnology and medicine. Despite initial suggestions that protein structure is determined by amino acid sequence, it is clear that this is just an ideal picture - folding mistakes can, and do, occur and nature has evolved a complex "chaperone" system to help prevent or correct such errors. But when mistakes do slip through , the effects can be catastrophic - protein misfolding diseases (such as Alzheimer's, BSE, etc.) and aggregation/instability of recombinant proteins in biotechnology for example.

Early observations in this laboratory suggested that even simple globular proteins such as lysozyme might sometimes fold incorrectly after thermal denaturation. This thesis reports on efforts to isolate and characterize some of these misfolded forms.

Protein Aggregation and Chaperones

Although for the past few decades the sticky aggregate and solid material at the bottom of protein chemists test tubes was ignored, recently it has been noticed that this aggregated protein product is yet another interesting aspect of protein behaviour. Interest has grown in misfolding and aggregation of proteins as large scale production of proteins for pharmaceuticals etc. has increased and the discovery that this material has a more sinister role in some diseases.

Aggregation occurs when unfolded protein, or protein folding intermediates come together, possibly because of exposed hydrophobic areas (Mitraki & King 1989, Jaenicke 1996, Fink 1998, Bauer et al 2000). It is highly concentration dependent and therefore it has multi order kinetics.

The study of aggregation has become more important because of increased use of recombinant proteins (see *Misfolding in Biotechnology* section) and because it appears to be one of the features of many diseases (see *Misfolding in Disease* section). This problem is not seen in vivo as the cell has developed molecules which help prevent aggregation. These molecules can be split into two groups : the protein folding enzymes and the chaperones.

The enzymes are Peptidyl Prolyl Isomerase (PPI) and the cyclophilins which catalyse proline isomerization around peptidyl proline bonds (Fischer & Schmid 1990, Jaenicke 1995). Protein Disulphide Isomerase (PDI) can catalyse the formation of disulphide bonds; this was one of the first of the protein folding enzymes that was discovered and helped to explain the discrepancy in timescale of in vivo and in vitro Rnase A folding (Anfinsen 1973).

The chaperones recognize and stabilize partially folded intermediates during protein folding. They appear to have many functions which relate to their ability to bind to partially folded intermediates (Ranson et al 1995). Examples are the Heat Shock Proteins (hsp) hsp 70, hsp 90 and hsp 60 (GroEL/Gro ES); as the name suggests these proteins are released when the cell experiences heat shock, thereby preventing aggregation (Gething & Sambrook 1992).

Misfolding in biotechnology

Mass protein production in the biotechnology industry has been problematic, genes for proteins are grafted into the DNA of eukaryotes such as E.coli overexpressed and the protein harvested (Mitraki et al 1991). Differences in how proteins are folded and expressed has led to aggregation problems.

In prokaryotic cells proteins are folded post translationally (Gething 1997, Netzer 1997), i.e. the whole protein is expressed first then folded, whereas in eukaryotic cells the proteins are folded co-translationally as the protein is expressed. This is possible due to differences in protein size and organization into domains (Wetzel 1996).

This difference between cells causes problems when a prokaryotic protein is expressed in a eukaryotic cell, because it is longer and isn't folded as it is produced. The presence of large amounts of these partially folded proteins with exposed hydrophobic areas leads to aggregation and formation of inclusion bodies (Jaenicke 1995). These aggregates, unlike the aggregate formed in a test tube, can be solubilised using denaturants and oxidising agents. The process is costly to the companies involved and a mass production technique which avoids this step will be of great interest to biotechnology companies.

Misfolding in disease

The correct folding of a polypeptide into its unique conformation is essential for its proper function. Incorrect folding or misfolding (Taubes 1996, Prusiner 1997, Perrett 1998, Cohen 1999, Saunders & Nagy 2000) can lead to a protein that doesn't function properly and can cause disease.

The product of all genes are proteins. In diseases where the root cause is a genetic mutation, the product of the gene, i.e. the protein, is the agent that leads to the physical symptom of the particular disease. Generally in these cases the mutation changes the proteins folded structure which prevents the protein from functioning normally. The study of such "misfolded " proteins is therefore important, allowing an understanding of how the disease damages and could lead to a means to either treat or prevent the disease.

Sickle cell anaemia and emphysema (Stryer 1995) are diseases where a mutation has led to proteins that don't operate as normal. In sickle cell anaemia a mutation in the haemoglobin protein means that it doesn't take up oxygen as well as usual. The deoxygenated haemoglobin molecule forms a fibrous precipitate which deforms the red cells which become sickle shaped. A single amino acid replacement causes this problem, placing a hydrophobic amino acid on the exterior of the protein and increasing its propensity towards aggregation.

Emphysema is caused by a mutation in $\alpha 1$ anti trypsin protein (Stryer 1995). This protein protects tissues from digestion by elastase, a secretory product of neutrophils, by binding to it`s active site. In emphysema the elastase destroys the alveolar walls in the lungs. So the resulting deficiency of this protein caused by emphysema makes the lungs less elastic and makes it difficult for people to breathe.

Cystic fibrosis (Welsh & Smith 1995) transmembrane regulator (CFTR) protein forms a channel in cell membranes which allows chloride ion transport. The disease Cystic fibrosis is caused by deletion of one amino acid, a phenylalanine, at position 508 and by many other mutations. This leads to an intracellular trafficking defect and the mutant protein is not allowed out of the endoplasmic reticulum as it is folded incorrectly and is degraded.

Retinitis pigmentosa (Liu et. al. 1996, Garriga et. al. 1996) is caused by mutations in a protein called Rhodopsin. Rhodopsin forms a 7 transmembrane helix which binds 11 cis-retinal and

plays a part in retinal function and sight. The many single amino acid mutations that cause this disease lead to a misfolded protein which doesn't function properly failing to bind retinal.

Alzheimers disease (Wellcome News Supplement 1998, MedicineNet) is a growing problem as more of the population live longer, and will in the future cause a drain on the economy and nursing resources. Discovery of a cure or some sort of preventative measure for this disease is of great interest to pharmaceutical companies all over the world. It is a disease of great interest to scientists because it doesn't seem to have just one definite cause, and because so many factors appear to be involved in its development. It causes dementia, changes in personality, disorientation and loss of short term memory.

Alzheimers disease can be genetic. Four gene products have been found to be implicated in the disease; presenilin -1 and 2, Amyloid precursor protein and apolipoprotein E (Wisniewski 1995). It may also be caused by a frameshift mutations (Voge1998) or the onset of old age, its exact mechanism is not known but so far it seems to involve a large number of enzymes and other factors. One of the proteins that are involved is Amyloid Precursor Protein (APP); this is snipped by an enzyme, giving a fraction of about 42 amino acids long called β amyloid. This forms plaques with hyper phosphorylated tau protein which is normally associated with microtubules. Ubiquitin can also be found in these plaques. Other chemical changes are observed in conjunction with Alzheimers such as a decrease in the amount of acetyl choline.

The Transmissible Spongiform Encephalopathies (TSEs) (Prusiner 1997, Lanchester 1996, Almond & Pattison 1997) are caused by a proteinaceous agent called a prion. This is a protein which is ubiquitous and has unknown function. It can assume different secondary structures in certain stretches of sequence, the infectious prion protein is mostly β-sheet and is protease resistant, whereas the normal prion protein is mostly α-helical. The conversion of the normal form to the infectious form is thought to be caused by close contact with the β-sheet infectious protein and possibly another conversion agent known as protein X. This prion, which is found in nerve cells, is the transmission agent and is unlike any other disease agent currently known, as it has no DNA or RNA, is incredibly hardy and cannot be destroyed in a similar manner to any other infectious agent. These diseases include kuru, scrapie (in sheep), bovine spongiform encephalopatie (BSE), Creutzfeld Jacob disease (CJD).

Protein Stability

How a protein attains its final folded structure is a subject of much discussion. The forces which drive the conversion of a long chain of amino acids into a three dimensional jigsaw piece which fits neatly into the huge jigsaw puzzle that make up the biochemical processes that are part of life, are not thoroughly understood.

The difference in Gibbs free energy between a native protein and the unfolded molecule is very small (Jaenicke 2000). The folded molecule is held together by non-covalent forces which are created by the exclusion of water from the natively folded molecule. The resulting increase in entropy of the water or other solvent molecules is the driving force for protein folding (Fischer & Schmid 1990, Creighton 1991, Seckler & Jaenicke 1992, Honig 1999) . It is these non-covalent forces which contribute to the folded protein's stability. They make the folded structure cooperative, this is confirmed by the fact that most proteins undergo a two-state unfolding transition when exposed to unfolding conditions. Most proteins remain in their folded state until the conditions are very close to that under which it unfolds. They then go through the unfolding transition, the midpoint of which is when 50% of the molecules are folded and 50% are unfolded, to the unfolded state (Creighton 1983).

The unfolded protein exists as a distribution of microstates which relate to the primary sequence and the environment that the protein is in. Tanford (Tanford 1968) originally suggested that the unfolded protein existed in a random coil state, but this is rarely true. In a random coil there should be free rotation around the bonds adjacent to the α carbon but rotation is inhibited because of restrictive intramolecular interactions (Hammerström 2000). For example free rotation about proline residue α carbons is restricted because of the nature of the side chain, these help keep the unfolded molecule in a more compact state. These forces are balanced by Van der Waals interactions which will limit the closeness of approach between non-bonded atoms. The nature of the solvent also plays a part in the formation of the unfolded state. In a poor solvent the polymer chain will interact with itself, whereas a good solvent will encourage interactions between the solvent and polymer chain. Due to the chemical properties of the amino acid side chains these solvents are unlikely to be good for all amino acids in a polymer (Tanford 1968, Dill &Shortle 1991). A solvent which is good for all amino acid side chains is known as a theta solvent. Other factors affecting the unfolded molecule are the presence of disulfide bonds, if these are not broken in the unfolded state they will increase the

likelihood of residual structure because of the increased interactions between distant parts of the polypeptide chain.

There is a lot of experimental evidence that residual structure is present in unfolded proteins, even at high denaturant concentrations. For example NMR has been used (Bierzynski & Baldwin 1982) to characterise GuHCl denatured RnaseA at low pH and it was found that the α helix which relates to the C-peptide is present for a significant fraction of time. Other evidence of this occurence has been reviewed by Dill & Shortle (1991) and Brockwell et al (2000).

Kinetics of protein folding

The Levinthal (Levinthal 1968) paradox indicated that it was impossible for a protein to sample every possible folded structure it could until it found the right one, since this process would take a long time and did not fit in with the known times taken for protein folding. In the late 50s Anfinsen performed refolding experiments using fully denatured and reduced Ribonuclease A and discovered that after a couple of hours the protein refolded so as to be indistinguishable from the original protein (Anfinsen 1973, Jaenicke 1996). The conclusion drawn from this experiment is that the primary structure, i.e. the amino acid sequence, of a protein contains all the information required by a protein to fold and assume its fully functional final structure. Other investigations from a similar period also established the "thermodynamic hypothesis" which states that a protein's native structure in physiological conditions is the one with the lowest free energy. (Fischer & Schmid 1990, Creighton 1992).

Because of the discrepancy between experimental protein folding time and the timescale of protein folding expected if a protein sampled every possible conformation, it was presumed that protein folding was determined by a kinetic pathway. This has been best explained by the "folding funnel" theory.

Over the last decade a new theory of protein folding (Baldwin 1995, Dill 1999, Honig 1999) has been developed which has neatly found a way round the Levinthal paradox and has provided an alternative view to other folding theories. Three dimensional free energy diagrams (free energy is represented on the vertical axis, the other axes relate to conformational degrees of freedom), known as folding funnels have been used to describe the kinetic landscape a protein molecule traverses whilst folding (Review Dill and Chan 1997). The landscape

resembles a valley with a deep narrow well in the middle which contains the lowest free energy state or the global minima/native state, and the top of the valley the vast number of conformations that an unfolded molecule can have.

The unfolding molecules start at the very top of the funnel and their journey down the slope into the well and the native state determines their folding pathway. Generally an analogy of the molecule being like a skier is used to describe their route down the funnel. The time that a molecule takes to fold, it`s kinetics, reflects it`s route down the funnel, for example a protein that has multi-state kinetics will meet many bumps, hills and moguls (local minima) which it will have to find its way round before it can get to the bottom of the funnel. A protein molecule with lower state kinetics will have a much smoother landscape. A variance in the kinetics of populations of the folding mixture are due to the different routes traversed by these populations in their journey down the funnel. This indicates yet another feature of the funnel theory, that it is consistant with thermodynamic and kinetic control of folding (Tsai 1999).

This new perspective on folding has also led to an alteration in the meaning of some terms, for example the transtion state, which in chemical terms means the state with highest free energy and is the rate limiting barrier of all chemical reactions. Instead of this being one particular molecular confomation, the transition state now becomes a particular high feature on the folding landscape that must be surmounted before folding can go on to completion.

So far this theory is only a model for protein folding, data has been calculated for model proteins (Wolynes 1995) and not for any real proteins. It has been used to illustrate principles and to direct research. It has also expanded the earlier pathway folding principle and gives a greater importance to folding intermediates as discussed in this thesis.

Many other protein folding theories have been developed and research carried out to provide evidence for them. Logical deduction and experimentation (Privalov (Review) 1996, Sosnick 1994) has lead to the conclusion that protein folding probably occurs on a pathway or pathways that involve the formation of intermediates. These intermediates can be either equilibrium or kinetic as folding is under thermodynamic and equilibrium control, pushing the protein to the structure with the lowest free energy in the fastest time possible. Identification of kinetic intermediates is possible only if the rate constant for the folding of the intermediate into the native state protein is slower than that of folding of the unfolded protein into the native state protein. Unfortunately this is not always the case, as the intermediate can also have a rate constant which is faster than the rate of folding from a completely unfolded protein, thus

making identification and isolation almost impossible. These wouldn't be seen using a thermodynamic technique such as DSC instead stopped flow experiments are used to identify kinetic intermediates. These intermediates are thought to be states of decreasing free energy which help push the extended unfolded molecule into its native folded state, the last intermediate will be the one closest in structure to the native protein. They also provide local minima which help prevent off pathway reactions such as aggregation (Baldwin 1995, Privalov 1996, Schmid 1999) and should mean that there is rapid formation of native structure, and that rapid folding is de rigeur. But slow folding does occur and arises when the polypeptide chain corrects mistakes in its structure. Study of the intermediates of protein folding reactions is important as it will allow a total understanding of what forces drive folding (Jones et al 2000).

It is interesting to note that some proteins have been shown to fold without producing intermediate states. Experiments using cytochrome c (Sosnick 1994) have shown that different solution/folding conditions can lead to different folding pathways, in this case it is due to the ionization of the two histidine residues. In neutral conditions (pH 6.2) these residues become charged and will bind the haem ligand incorrectly, this gives a heterogeneous folding mixture with a slow folding step which corrects this misfolding. Even when an aliquot of the slow folding protein is put into the correct conditions the folding remains slow. On the fast folding pathway, at lower pH, the protein folds quickly because it has no structural corrections to make and can collapse from an extended form to a compact form quickly. This indicates that folding in conditions where mistakes/barriers that block folding can occur on the folding pathway, is slower because kinetic intermediates are produced, and that these intermediates aren't necessary and only occur in certain subsections of the folding mixture. Stopped flow folding and unfolding of cytochrome c under the above conditions identified (Bai 1999) the intermediate formed in the folding pathway as being on-pathway, i.e. that the intermediate formed before correct folding occurred, because the rate constant of its formation was slow and there was a lag phase before appearance of the natively folded protein.

Cold shock protein (CSP) (Schindler et al 1995), is a very thermodynamically unstable protein, which folds without formation of intermediates, instead it has a native like activated state. The existence of this activated state can be related to the difficulties in trying to form partially folded structure in small marginally stable proteins, when only a subset of nativelike interactions are available to stabilize it. Again this would suggest that perhaps intermediates are not a necessary part of the folding pathway, at least for some proteins.

There is increasing evidence for protein folding being a nucleation condensation type process. Evidence for this type of mechanism has been found by studying the folding of chymotrypsin inhibitor 2 (CI2) (Itzhaki et al 1995). This protein has no disulphides or cis proline bonds, making it a useful protein for folding studies as there should be little opportunity for misfolding. CI2 has a known rate determining transition state and a two state folding pathway. Various mutations of CI2 have been made, the mutations were chosen so as to delete a small amount of side chain; removing a defined interaction but without introducing a new one, this was then used to probe its folding. The kinetics of folding of these mutants is measured and any changes in rate constant is related to the mutation. The results of this work indicate that the folding of this protein occurs by a nucleation condensation type mechanism, whereby a nucleus of adjacent amino acids is formed which make interactions with distant amino acids forming a stable structure. In this case the α-helix is formed first driven by long range interactions between distant amino acids, then the remainder of the protein condenses around the helix, forming native interactions leading to overall stabilization. This mechanism depends on the stability of the separate substructures within the protein; if they are stable separately they won't necessarily fold in this manner as a concerted mechanism occurs where the separate substructures can`t exist alone.

Recent advancements have lead to the development of equipment which allows the measurement of the rate of submillisecond refolding and unfolding in proteins. For example temperature jump laser equipment (Greubele et al 1998) can be used to measure the kinetics of these processes by fluorescence /UV/IR, basically any method which can measure a change in a protein. Also methods of rapid solution mixing have been invented (Shastry et al 1998, Yeh at al 1998). A lot of this work has focused (Winkler & Gray 1998) on the folding of cytochrome c because of the heme group which is a useful probe. Evidence from these experiments has also indicated a nucleation condensation mechanism for cytochrome c; folding occurs with the formation of a partially folded core domain of a loosely packed helix, the source of which is non native tertiary interactions. The rest of the protein folds around this structure resulting in a shift of free energy to a favorable value for folding.

The disparity between experimental results i.e. that there is evidence for different types of folding mechanism in different proteins, show how difficult it will be to propose a general theory of protein folding.

Proline Isomerisation

Careful study of folding kinetics has shown that although usually quite rapid, subpopulations of protein molecules seem to fold more slowly than their companions (Schmid 1999). This has been shown, at least in part, to be due to the slow cis-trans isomerization of peptide bonds adjacent to proline residues in the unfolded polypeptide (see figure 1).

Unfolded protein molecules exist as a heterogeneous mixture (Schmid 1999), so the unfolded state is different for each molecule and changes over time. This is why folding can occur at different rates, as some of these species will fold with different kinetic rates. The slow refolding molecules can be caused by the presence of incorrect isomers of Xaa - Pro peptide bonds. Such incorrect isomers might usually be corrected (enzymatically or otherwise) during folding, but the possibility arises that occasionally non-native cis - trans proline isomers may get trapped or locked in to an incorrectly folded conformation.

These non-native proline isomers can lead to a deceleration in the protein's folding pathway, as these isomers must assume the correct conformation before protein folding can occur along the correct pathway. As when the peptide bond involves proline, the conversion from cis to trans is slow, 10 - 100 secs (Schmid 1993). Hence the slow steps along the folding pathway, and the relationship that has been observed between folding rate and number of prolines in a protein. This slowing down of the folding pathway is due to the Us (slow folding species) on the folding pathway. For example when a protein is released from the ribosome it has all its proline peptide bonds in the trans form (Seckler & Jaenicke 1992) as do most folded proteins, this can be quite significant.

The peptide bond, in general, has appreciable double bond character which prevents free rotation around the peptide bond. Isomerization can occur around the peptide bond but it is very slow. The bond can be described as being in either the cis or trans conformation, see figure 1.1. Generally peptide bonds not involving proline are trans, although a few cis non-proline peptide bonds have been identified in crystallographic studies (Stewart et al 1990). Peptide bonds involving Pro can exist as a mixture of cis and trans isomers in solution (Fischer 2000), unless there is some structural constraint similar to that in the folded molecule. However the difference in Gibbs free energy between the cis and trans conformation of a proline peptide bond, is less than that of a normal peptide bond and subsequently at equilibrium in an unfolded protein molecule, both conformers are significantly populated.

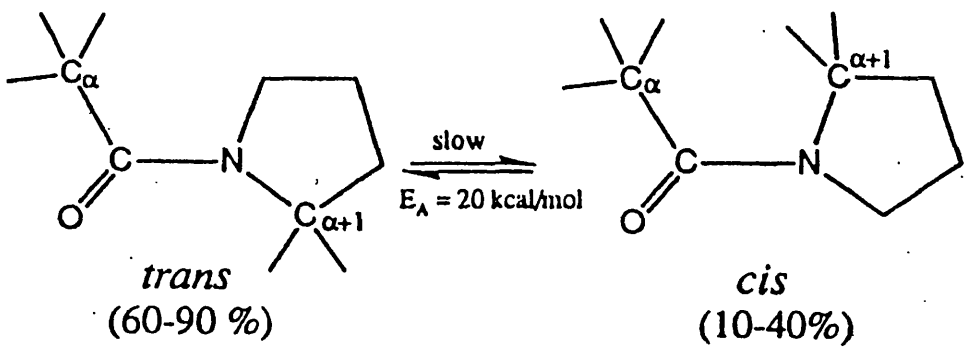


Figure 1.1 The two isomeric states of a peptidyl prolyl bond (cis ↔ trans). Taken from Schmid 1999.

Kinetic studies using NMR and stopped flow techniques have shown that proline isomerization is one of the rate limiting steps in folding of proteins (Walkenhorst et al 1997, Ikura et al. 1996). For example a mutant form of staphylcoccal nuclease (Baldwin 1995) where all proline residues have been replaced by alanines and glycines, folds very quickly with 2 unfolding intermediates and 1 folding intermediate, leading to a simplification of the kinetics of the wild type protein which contains 6 prolines. In spite of this the folding and unfolding of staphylococcal nuclease still has several phases. This has allowed the folding and unfolding of staphylococcal nuclease to be studied in detail.

To date most of the work on proline isomerisation has been limited to the area of identification of slow steps in protein folding and very little has been done to try and identify non-native proline conformations in proteins. They have been identified in some folded molecules (Langsetmo et al 1989, Ikura et al 1996, Svensson 1992), but it is likely that they are more prevalent as the potential energy difference between the cis and trans isomers is calculated to be about 3kJ/mol (Stewart et al. 1990) suggesting that the occurrence of a cis proline bond is about 30 - 40 %, but only very few cis proline bonds have been identified in crystal structures (Stewart et al. 1990). This deficit is likely to be due the refinement procedures used by crystallographers which tend to assume that all proline imide bonds are trans, and this is fixed into the refinement strategy. Generally cis prolines are only identified in high resolution protein structures (Weiss et al 1999).

In a few cases a non-native proline isomer has been found to exist in the folded protein. Ribonuclease A shows two folding periods, a slow folding phase and a fast folding phase, the slow folding phase has been identified as when proline isomerisation occurs around proline residue 93. (Schmid & Baldwin 1978). Ribonuclease A appears to be quite a useful protein for studying proline isomerisation as it has a tyrosine residue next to one of its proline residues, which can then be used as a probe in fluorescence folding experiments (Schmid 1981). Rnase A refolding experiments performed under conditions where the native protein conformation is favoured show only a slow folding phase indicating the presence of an intermediate with a non-native proline isomer, this structure has a compact folded structure which shields the tyrosine residue of interest from the external environment and is also capable of binding 2`CMP as well as native Rnase A does (Schmid 1981, Schmid & Blaschek 1981).

Brandts & Lin (Brandts & Lin 1986) have developed an enzymatic method which can be used to identify the amounts of cis and trans isomers of a protein. They discovered that certain proteases have

absolute specificity towards peptide bonds in the trans configuration, for example prolidase is specific for an X-Pro dipeptide in the trans configuration. Aminopeptidase P (APP) is an exopeptidase which is specific for N terminal X-Pro bonds in the trans configuration, and trypsin and proline specific endopeptidase are specific for bonds next to X-Pro when the X-Pro bond is trans. Therefore this isomeric specificity of the enzymes can be used to study the kinetics or thermodynamics of certain residues within a protein if hydrolysis is measured as a function of time at high enzyme concentrations and low temperatures, as this biases the equilibrium towards a high rate of hydrolysis (fast phases) and slow proline isomerization (slow phase).

This method involves hydrolysis of the protein by different enzymes, generally starting with a burst of pepsin hydrolysis to irreversibly unfold the protein then, if necessary, cleavage of the resultant peptides to make them a suitable substrate for the enzyme used in the stereo-specific cleavage. Hopefully the stereospecific cleavage results in a product which is easily assayed, preferably some free amino acid which can be measured by HPLC, otherwise the product may need to be hydrolysed again by another enzyme.

This method has been used by Lin & Brandts (Lin & Brandts 1983) to look at Rnase A isomerization, but appears to have been superseded by NMR methods and not used as a general method to assay proline isomerization. Using this method they discovered that native Rnase A has 100% of its Pro 93 residues in a cis configuration, oxidized Rnase A has 36% cis and unfolded Rnase A has 70% cis, these results from the isomer specific proteolysis method agree very well with the same measurements made using different methods. It is interesting to note that Rnase A is a very useful protein for this sort of method because it has the sequence Lys 91- Tyr 92- Pro 93, this sequence is a substrate for trypsin when the Tyr - Pro bond is trans. Trypsin is used to hydrolyse the Lys -Tyr bond, then the Tyr-Pro bond is hydrolysed by APP, resulting in free Tyr which can be easily assayed. This particular sequence has proved to be useful before, as the Tyr can be used as a fluorescence reporter of which configuration the Tyr-Pro is in.

The development of protein NMR has led to an increase in the amount of published papers identifying non-native proline isomers, for example Fox (Fox et al 1986) used 1D ^{1}H NMR experiments to study the changes in the interactions of proline with its neighbouring amino acid residues. These generally show minor peaks within the same chemical shift region as the expected major peaks from the natively folded protein (Truckses et al 1996, Alexandrescu et al 1989). These minor peaks correspond to native-like protein with non-native proline configurations. If the isomerisations are occurring on the NMR timescale, or can be made to by increasing temperature, 2D

^{1}H NMR hydrogen exchange experiments can be performed to provide information on the isomerisation process (Chazin et al 1989). More complicated NMR experiments can be used to calculate the structure and structural changes that occur when a non-native proline isomer is present. For example Chazin et al, have used the above experiments in conjunction with 2D Nuclear Overhauser Enhancement Spectroscopy (NOESY) to assign the resonance's and diagnose which proline residue is in a non-native configuration for Calbindin D_{9K}.

Even more complex experiments such as those performed by Yuan (Yuan et al 1998) have allowed computation of the structure of TGF-β binding protein like domain. NMR experiments were used to identify the existence of two stable isoforms of transforming growth factor β binding (TGF-β) protein like domain from human fibrillin, this has 6 proline residues, four of which adopt the trans conformation and the other two, Pro 22 and 52, are either cis or trans. These isomers do not interconvert on the millisecond or hour timescale and hence don't appear to interconvert directly from one isomer to the other. They used ^{1}H ^{15}N Heteronuclear Single Quantum Correlation Spectroscopy (HSQC) to study the TGF-β binding protein like domain, assigned the resonances and found major and minor resonances for some peaks, which were due to the folded form with a non-native proline configuration. The information gained from this experiment was enhanced with results of a D_2O NOESY experiment which then let them identify that only two prolines (Pro 22 and 52) were able to isomerise into the alternative configurations. The secondary structure of these isomers was then calculated using the NOEs and the populations of some of the resonances. Finally they came to the conclusion that the isomerization would alter the secondary structure of the protein, changing the type of structure that certain areas of the protein will assume, and the length of structural strands (Yuan et al. 1998). This sort of technique has been applied to other proteins to calculate the structure of non-native proline configuration folded proteins (Amodeo et al 1994) .

The presence of non-native proline conformations in folded proteins have been shown for other proteins such as Calbindin D_{9K} (Svensson et al. 1992, Chazin et al. 1989) by NMR and crystal structure, Staphyloccal Nuclease (Alexandreascu et al. 1989), Salmon Calcitonin (Amodeo et al. 1994) by NMR, and by the use of urea denaturing gels for Thioredoxin (Langestmo et al. 1989).

The proline amino acid appears to be an important part of many biologically active peptides and proteins (Vanhoof et al 1995). This is because of its ability to confer unique conformational constraints on the peptide backbone (Eyles et al 2000). It is found in biological

molecules such as the cytokines, growth factors, G coupled receptors and neuro and vaso active petides.

For example proline is found with high frequency in α helix transmembrane proteins (Vanhoof et al 1995). This is thought to be because an intrahelical proline will expose the carbonyl oxygen of neighbouring amino acids providing a cation binding site, this can play a part in transport systems such as proton pumps, Na/K channels and calcium pumps. They may also provide the conformational change that's required in protein structure for regulation of a transport channel.

Proline residues in peptides and proteins may protect against non-specific proteolytic degradation (Vanhoof et al 1995). Some pathogens take advantage of this by secreting proteases which are specific for Pro - X sequences. Diagnosis of the infection bacterial vaginosis is made by measuring the amount of proline aminopeptidase in vaginal secretions. IgA mediates immunity to infection at mucosal surfaces and has a Pro - X motif at the hinge between its Fab and Fac region on its heavy chain. Mucosal surface pathogens such as Neisseria Gonorrhoea and Neisseria Meningitis all secrete a protease which can cleave this particular sequence, thereby releasing intact Fab and Fc regions. It is interesting to note the production of these proteases relates to the virulence of these diseases.

The proline motif is also important as there are very few peptidases which can hydrolyse the proline peptide bond, and their activity is usually influenced by the isomerisation state of this bond (Brandts & Lin 1986, Vanhoof et al 1995). The cytokines and growth factor hormones have an X - Pro sequence at their amino terminus. Examples of these are interleukin 1β, interleukin 2 and erythropoietin. This sequence contributes to their biological activity and protects against nonspecific proteolytic degradation such as C terminal amidation, acetylation or N terminal cyclisation. Growth hormone and prolactin are protected against proteolytic degredation by having proline motifs at both ends of their chains. So far only 2 proline specific amino peptidases have been discovered, amino peptidase P and dipeptidyl peptidase IV. These have specificities towards N terminal prolines. This suggests that these proteins are under evolutionary pressure to retain these sequences, possibly because they have an important role in their biological activity.

Because the proline isomerization process is so very slow the cell has evolved particular enzymes to speed up the process. In particular peptidyl prolyl isomerase (PPI) and the

cyclophilins catalyse the 180° rotation around the C - N linkage of the peptide bond adjacent to proline. There is a lot of interest in these enzymes as it appears that they may be part of the body's immune system (Matouschek et al. 1995). The cyclophilins are an intracellular target for cyclosporin A, an immunosuppressant. Together they inhibit the calmodulin dependent phosphotase calcineurin and thereby prevent translocation of a subunit of T cell specific transcription factor from the cytosol to the nucleus. They are also part of the protein folding process in vivo. Drosophilla require a cyclophilin homologue to secrete Rh1 opsin in their photoreceptor cells. There is also a relationship between human cyclophilin and the HIV virus. The HIV-1 virus is the product of the *gag* and *pol* genes, which encode the poly-protein precursor of the actual virus and a protease. This protease is involved in maturation of the polyprotein to the the virus capsid. The *gag* polyprotein has been shown to specifically bind to Cyp18cy and Cyp23sec. These associations can be disrupted by addition of cyclosporin A, which inhibits the enzymatic action of the cyclophilins (Fischer 1994). Cyclosporin A is known to block HIV-1 replication, suggesting the involvement of cyclophilins in the HIV disease process. A prolyl isomerase has even been found to restore the function of phosphorylated tau protein which is found associated with plaques in Alzheimers disease (Lu et al 2000).

The effect of these peptidyl prolyl isomerases on small proline containing peptides and on the refolding and unfolding of proteins has been studied extensively, for review see Schmid 1993 and Fischer 1994. In particular the effect of PPIs on the refolding of Ribonuclease A has been studied. RnaseA is an interesting case as it has a tyrosine residue adjacent to one of its slow isomerising proline residues, and the tyrosine residue can be used as a spectroscopic probe. PPIase has been found to increase the rate of a minor slow phase in refolding of urea denatured Rnase A (Fischer & Bang 1985), but the Rnase A proved to be a poor substrate for the enzyme in comparison to a peptide. The refolding of RnaseA consists of two slow phases, referred to as the minor and the major slow phase, the major slow phase involves isomerzation of a proline residue in a native like intermediate and the minor one the isomerisation of proline 93. Proline 93 is highly exposed and the PPIase is able to catalyse this isomerisation, but not the one that involves the native like intermediate (Lang et al. 1987). Hence the PPIases are only able to catalyse isomerisation of highly exposed proline residues, and therefore it is unlikely that these enzymes are able to catalyse isomerizations of non-native proline configurations within folded protein molecules although currently very little work appears to have been done on this subject.

Hen Egg White Lysozyme

Hen egg white lysozyme (HEWL) is a 129 amino acid protein. Its function is as an enzyme which cleaves glycosidic bonds, such as those present in the cell walls of bacteria. Its structure was elucidated crystallographically (Phillips 1966) revealing that it had two domains, see figure 1.2, one is made up of 4 α-helices and one which is a triple stranded antiparallel β-sheet. There is a 3_{10} -helix in each domain and a short region of double stranded antiparallel β-sheet which links the two domains . The whole structure is stabilized by 4 disulphide bonds. The folding of HEWL in vitro has been studied extensively and all the results brought together and reviewed (Radford & Dobson 1995, Matagne & Dobson 1998), giving an overview of the folding pathway of HEWL.

The compact protein structure of HEWL is achieved within a few milliseconds of refolding, at this time the molecules have nativelike secondary structure but are heterogeneous. The heterogeneity is likely to have arisen from the coexistence of correctly folded and misfolded secondary structure or as a result of rapid collapse of the polypeptide chain. The latter forms are thought to be quite significant in the case of HEWL because of the 4 disulphide bonds.

NMR studies have shown that during HEWL refolding in general the α domain is formed first, within 65ms (Bai 1999), then the β domain (350ms) (Radford et al 1992).There is also some folding which occurs very slowly (20%) because of the presence of (incorrect) cis proline isomers. It is worth noting that lysozyme contains 2 proline residues, which are close together in the sequence (Pro 70, Pro 79), in a loop region of the protein. Kato and coworkers (Kato et al. 1982) first identified the nature of the slow folding step. Their work showed that the ratio of the rate constants for the slow step of conversion of one unfolded intermediate to another was similar to that of proline isomerisation, and showed that the time constants for conversion of one intermediate into the other was comparable to that of proline isomerisation in other proteins.

Kinetic folding studies of HEWL indicate that at lower temperatures a two phase folding pathway is present. Results from different experiments have been compared (Matagne et. Al. 1997), importantly the binding of a labelled inhibitor to the binding site of HEWL in refolding experiments (Itzhaki et al 1994, Kulkarni et al 1999) , giving an accurate measurement of the

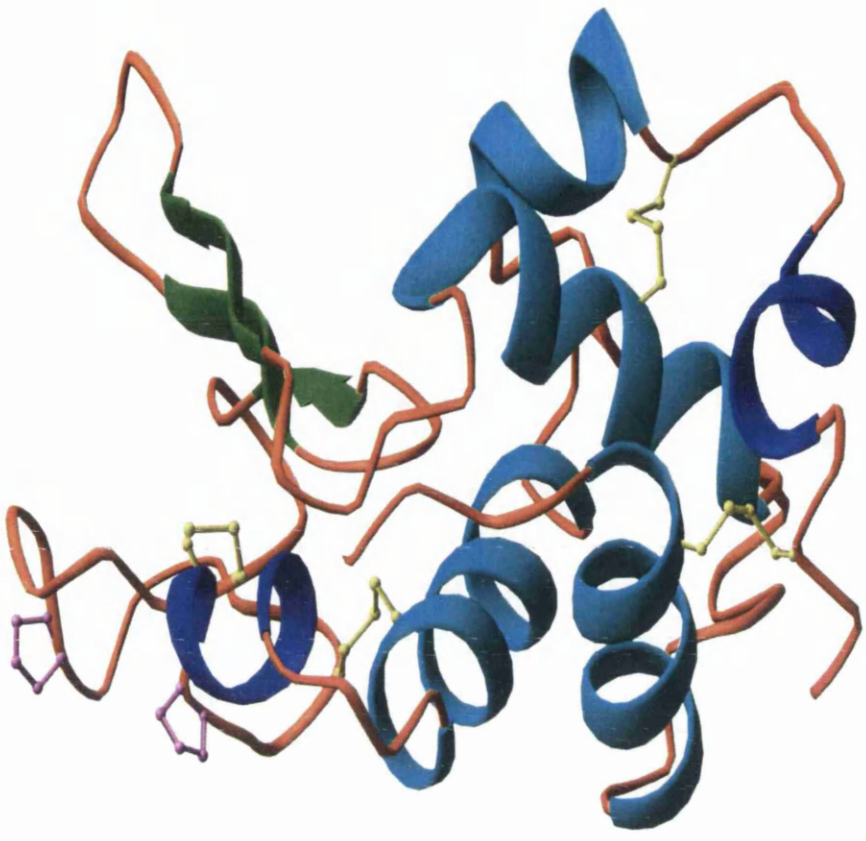


Figure 1.2: Cartoon diagram representing the structure of Hen Egg White Lysozyme, both proline residues are shown in magenta.

amount of natively folded HEWL. These have shown that 25% of the molecules fold on a fast pathway and the remaining 75% on a slow one. Hydrogen exchange indicates that the refolding population on the slow pathway only shows protection in the α domain and that the slow step is caused by the formation of the β domain. Binding of the labelled inhibitor suggests that the active site and the β domain develop at approximately the same time. The fast pathway has an intermediate with persistent native like structure in both domains, but doesn't bind the inhibitor till after another step. The timescale on which this intermediate is formed is similar to that found for the folding of small single domain proteins, indicating that in the absence of the need to reorganize bits of misfolded structure lysozyme can fold into a highly native like state with a rate constant similar to that of a smaller protein, and that subsequent steps in folding are limited by the need to reassemble the domains. The overall conclusion from this work is that the folding of larger protein occurs within domains.

Aims of this thesis:

We have observed that repeat heat treatment of HEWL leads to a composite species which contains four conformers of HEWL. This thesis contains the results of the biophysical characterization of this species, the subsequent separation of this into four conformers and then their characterization. Finally I shall discuss what these results mean with respect to the postulation that the misfolding of hen egg white lysozyme has been caused by proline isomerisation.

The appendices contain the results of some experiments that were done that produced inconclusive results. Appendix I outlines the attempts to crystallise the HT HEWL and also contains the results of experiments where cyclophilin, a peptidyl prolyl isomerase, was added to the HT HEWL. Appendix II contains an overview of lysozyme aggregation studies, where creation of a β-fibrillar species was attempted.

Chapter 2 : Instrumentation

Differential Scanning Calorimetry (DSC)

DSC is a technique which can follow the heat energy changes in samples subjected to changes in temperature. This is one of the few techniques that can be used to study directly interactions within or between macromolecules and the energetics of unfolding of a protein or nucleic acids. This has been the main method used to look at the forces that hold proteins together (Cooper & Johnson 1994 a).

The actual calorimeter itself has 2 cells (Cooper & Johnson 1994 b), a sample cell and a reference cell (the sample cell will contain protein and buffer and the reference cell buffer alone). These are carefully insulated from the exterior atmosphere, the cells are connected to separate heaters, which compensate for any difference in the temperature of the cells and the surrounding outer heating jacket, see figure 2.1. What is measured in an experiment is the heat supplied to the sample cell to keep it at the same temperature as the control cell, this power difference is recorded as a function of temperature and relates to the difference in heat capacity.

The experiment is performed under 1 -2 atmospheres of nitrogen and the samples are initially degassed to prevent release of bubbles during the experiment which lead to errors in the heat capacity. Sample concentration is usually 2-3mg/ml in the case of these experiments. Usually a buffer baseline is measured before or after the main experiment with protein in it.

A typical DSC trace is shown in figure 2.2, this is an example of an equilibrium unfolding process which is caused by heating. It is described as an endothermic transition as heat is absorbed, and is an example of thermal denaturation of a protein in this case HEWL. There is one sharp, symmetrical peak which shows that the unfolding transition is highly cooperative, the midpoint of which is known as the T_m. At this point, assuming that the transition is two state, the molecules go from the folded to the unfolded state with no intermediates, so that at the temperature of the peak of the unfolding transition 50% of the molecules are unfolded and 50% folded. The baseline shift is due to the increase in heat capacity with temperature.

Analysis of the data from such an experiment is done using Microcal Origin™ computer program (Microcal) using the following equations to deconvolute the data.

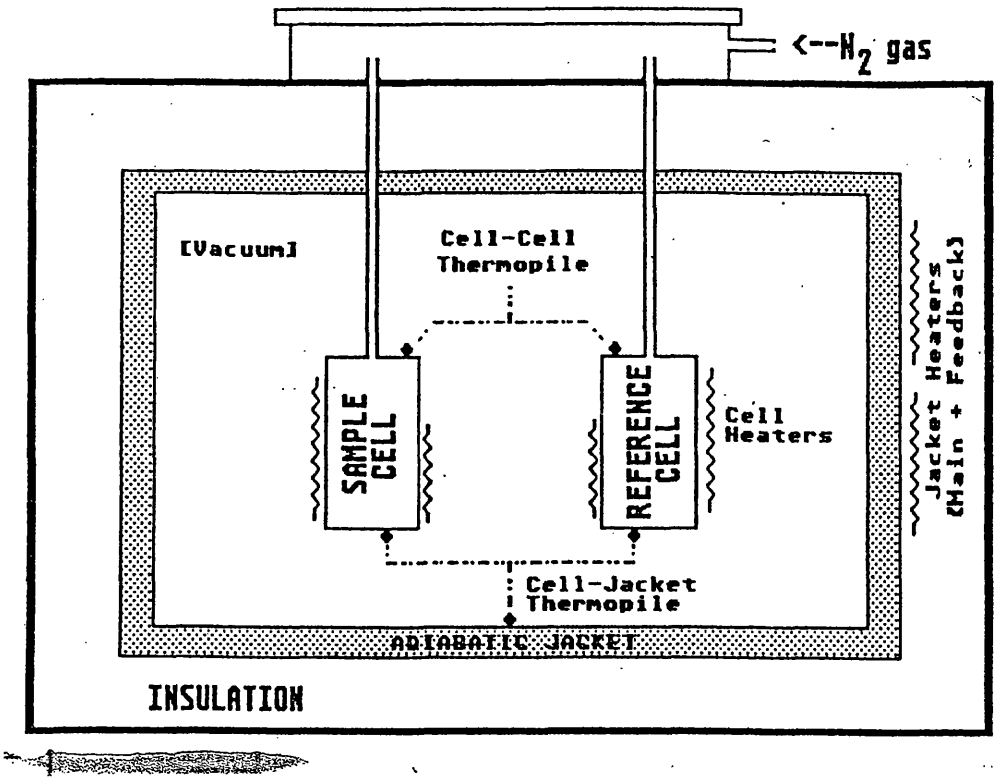


Figure 2.1 : Diagram of the organisation of the interior of a Differential Scanning Calorimeter, from Cooper & Johnson 1994 (a).

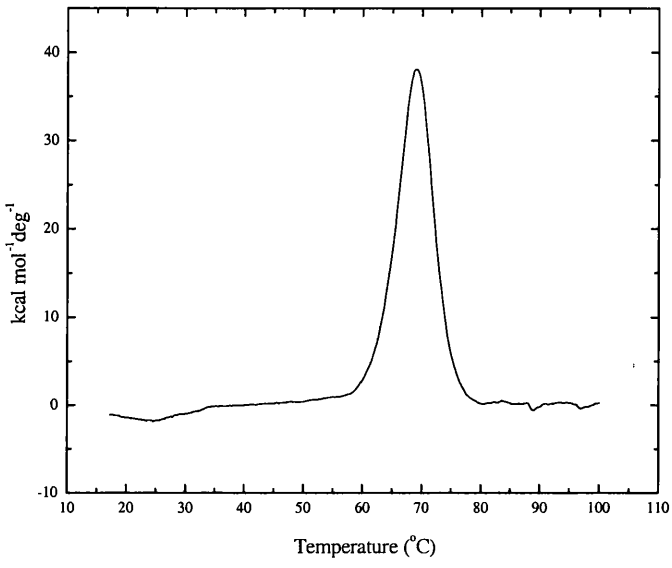


Figure 2.2 : DSC data for a typical two-state unfolding transition, 2mg/ml hen egg white lysozyme in pH 3.4, 0.1M citrate buffer. Midpoint of transition (Tm) occurs at 70°C under these conditions.

It is assumed that a protein is composed of a number of structural domains, each of which is involved independently in a transition between the folded and the unfolded forms (A = A`, B = B`, ...). The equilibrium constants are expressed as fractions ($K_A = f_{A`}/f_A$, $K_B = f_{B`}/f_B$, ...) and the calorimetric value expressed as ΔH, (ΔH_A, ΔH_B, ...). H_N is the enthalpy of the native folded form and is included as ΔH is measured relative to the folded form. The total molar enthalpy of the system is therefore:

$$H = H_N + f_{A`}\Delta H_A + f_{B`}\Delta H_B + \ldots \tag{1}$$

The temperature derivative of this equation is the total molar heat capacity:

$$C_P = C_{PN} + \left[f_{A`}\Delta C_{PA} + \Delta H_A \left( \frac{\delta f_{A`}}{\delta T} \right) \right] + \ldots \tag{2}$$

where C_{PN} is the molar heat capacity of the totally folded state, ΔC_{PA} is the change in heat capacity on the unfolding of the A domain and the bracketed term is repeated for each domain involved in unfolding. Since $f_A = 1 - f_{A`}$, then the fractional concentration of the unfolded species can be expressed in the following terms:

$$f_{A`} = \frac{K_A}{1 + K} \tag{3}$$

Differentiation of this gives :

$$\left(\frac{\delta f_{A`}}{\delta T} \right) = \left(\frac{K_A}{\left(1 + K_A\right)^2} \right) \left(\frac{\delta \ln K_A}{\delta T} \right) \tag{4}$$

The derivative on the left hand side is already known as the van`t Hoff equation :

$$\left(\frac{\delta \ln K_A}{\delta T} \right) = \frac{\Delta H_A^*}{RT^2} \tag{5}$$

and ΔH^*_A is the van`t Hoff enthalpy change for unfolding of cooperative unit A. Substitution of equations 3 and 5 into equation 2 gives the following equation :

$$C_P = C_{PN} + \left[\frac{K_A \Delta C_{PA}}{1 + K_A} + \frac{K_A \Delta H_A^* \Delta H_A}{\left(1 + K_A\right)^2 RT^2} \right] + \ldots \qquad (6)$$

This gives the general equation which can be applied to any type of protein unfolding as long as all parameters are evaluated at the same temperature T.

The model for Independent Two State Transitions including ΔC_P Effects

If the transition is assumed to be two state then all of the van`t Hoff enthalpy values above will become equal to the calorimetric enthalpy. If ΔC_{PN} can be expressed as a linear function of temperature ($\Delta C_{PN} = B_0 + B_1 T$), and is replaced in equation 6 by this linear function it gives the next equation :

$$\Delta C_P(T) = B_0 + B_1 T + \left[\frac{K_A(T) \Delta C_{PA}}{1 + K_A(T)} + \frac{K_A(T) \Delta H_A^* \Delta H_A(T)^2}{\left(1 + K_A(T)\right)^2 RT_2} \right] + \qquad (7)$$

all the temperature dependent parameters have the bracketed temperature term next to them.

The enthalpy of unfolding at a particular temperature ΔH_A (T) can be expressed in terms of its temperature independent value at the midpoint T_{mA} and the heat capacity change for the transition ΔC_{PA} :

$$\Delta H_A(T) = \Delta H_{mA} + \Delta C_{PA}\left(T - T_{mA}\right) \qquad (8)$$

then integrate equation 5 from T_{mA} where $K_A(T)$ is unity to a specific temerature T

$$K_A(T) = \exp\left\{ \frac{-\Delta H_{mA}}{RT}\left(1 - \frac{T}{T_{mA}}\right) - \frac{\Delta C_{PA}}{RT}\left(T - T_{mA} - T \ln \frac{T}{T_{mA}}\right) \right\} \qquad (9)$$

Equation 7 can be used to calculate , when substituted for each unfolding domain, the value of the systems heat capacity at any temperature T once the values for temperature independent parameters, B_0, B_1, T_{mA} etc, have been given.

The model for independent non-two-state transitions

This model is only applied to transitions with no ΔC_P, before curve fitting with this model, a progress baseline must be subtracted from the ΔC_P effects if they are present (this also sets ΔC_{PN} to zero at all temeratures). To treat non-two state transitions, the appropriate place to begin is equation 6 which still includes both calorimetric and van`t Hoff heat changes. Indicating the temperature -dependent parameters, this can be re-written as

$$C_P(T) = \left[\frac{K_A(T)\Delta H^*_{mA}\Delta H_{mA}}{\left(1+K_A(T)\right)^2 RT^2} \right] + \ldots \qquad (10)$$

The equilibrium constants are then calculated using the van`t Hoff heat.

$$K_A(T) = \exp\left\{ \frac{-\Delta H^*_{mA}}{RT}\left(1 - \frac{T}{T_{mA}}\right)\right\} \qquad (11)$$

The above equations are used for curve fitting and the parameter set is T_{mA}, ΔH_{mA}, ΔH^*_{mA}, T_{mB}, ΔH_{mB}, ΔH^*_{mB}, ...

Curve fitting can then be performed when the programme is provided with the number of transitions required to fit the heat capacity curve, and then puts in the T_m (the midpoint of the transition) of each of these, so Origin ™ can provide guesses for the other fitting parameters. With this information it can calculate the heat capacity at certain temperatures and then compare it with the experimentally observed values. An iterative process is then used to find the best fit of the calculated heat capacity to the experimental heat capacity.

Therefore DSC can provide us with a measure of the enthalpy of unfolding of a protein at a particular temperature and information on the unfolding transition.

Generally for this thesis DSC was used to visualise the misfolding process and as a means to identify the misfolded protein.

Fluorescence Spectroscopy

Fluorescence can occur when a molecule absorbs light energy and is excited up to higher electronic energy states from the ground state. This energy can be emitted in two ways, it can be given back out as heat or as radiation. In the case of fluorescence some of the energy is reemitted as radiation, the rest is dissipated during transitions from the highest excited state back down to lower excited states, this is shown in figure 2.3. Because the fluorescence energy emitted is less than is put in, the wavelength of the emitted radiation is greater than that of the radiation used to excite the molecule.

Fluorescent molecules have delocalised electrons usually associated with conjugated double bonds. In the case of proteins the relevant fluorophores are the aromatic amino acids tryptophan and tyrosine. Fluorescence is a highly sensitive technique and only small amounts of the fluorophore are usually required. The fluorescence intensity and wavelength of the fluorophore relates to its surrounding environment, and can therefore be used as a structural probe. If the fluorophore is in a more polar environment the fluorescence peak becomes red shifted. Physically this due to a decrease in the separation between the excited state and the ground state so the excitation becomes of lower energy as does the resulting fluorescence. In the corresponding case of a less polar environment the fluorescence peak becomes blue shifted as the energy levels move further apart, so more energy is required to get from the ground state level to the excited state, so more energy is given out in the resulting fluorescence.

The following are the types of fluorescence experiments which are mentioned in this thesis.

Emission spectra, can be measured when the sample is irradiated with light at or close to the absorption maxima of the sample molecule. The resulting spectrum is measured as a function of emission wavelength and the maximum wavelength noted. This is a steady state fluorescence measurement and measures the molecule`s intrinsic fluorescence.

Fluorescence quenching experiments using succinimide were also performed. A quencher is a molecule which can decrease the fluorescence of other molecules. A fluorescence quenching reaction normally involves physical contact between the quenching molecule (e.g. O_2 , I^-, acrylamide, succinimide (Eftink & Ghiron 1984)) and an excited indole ring (aromatic amino acid excited at appropriate wavelength in the fluorimeter, usually tryptophan). There are two different types of quenching molecule; a static quencher is a molecule which forms a

complex with the fluorophore and decreases its fluorescence. A dynamic quencher collides with the fluorophore to decrease its fluorescence.

Tryptophan fluorescence is typically used for protein studies as its side chains are very sensitive to the polarity of its environment. A surface tryptophan exposed to a polar solvent has a different emission maximum than a buried tryptophan, the more buried the tryptophan the lower the maximum. The ease with which a fluorophore is quenched depends on its relative exposure to the quencher. Aromatic side chains on the surface of a protein are quenched by diffusion controlled encounters with quenchers, but large quenchers aren't as efficient as O_2 at quenching internal residues as they are too large to enter the proteins interior .

$$M^{*}+Q \underset{k_{-d}}{\overset{k_d}{\leftrightarrow}} (M^{*}....Q) \overset{k_i}{\rightarrow} M+Q+\Delta$$

The above equation (Eftink & Ghiron 1976) describes the reaction between M* ,the excited indole ring, and Q, the quenching molecule. (M*....Q) is the complex formed by the diffusional encounter between the 2 molecules which has rate constant k_d. The excited state then collapses producing heat with rate constant k_i. This quenching process can be described by the Stern-Volmer equation

$$F_0/F = 1 + K_{sv}[Q]$$

where F_0 is the initial fluorescence intensity, F the fluorescence intensity after addition of succinimide and K_{SV} is the Stern-Volmer constant. Therefore, K_{SV} can be determined by finding the gradient of a graph of F_0/F against [Q].

The one problem with using tryptophan as the probe for quenching is that if there is more than one in the molecule the analysis of the quenching reaction can become very difficult. This difficulty is due to the heterogeneity of tryptophan accessibility due to variation in the quenching rate constant and the fluorescence lifetime of the individual tryptophan residues. In the case of HEWL there appears to be a high degree of heterogeneity in the accessibility of the tryptophan residues, experimentation has shown that the quenching rate constant for HEWL corresponds to the quenching mechanism involving the solvent exposed tryptophan residues and not the buried residues (Eftink & Ghiron 1984).

Extrinsic fluorescent probes are molecules which bind to protein molecules and whose fluorescence changes with their environment. For example, 8-anilinonapthalene sulphonate (ANS) is a hydrophobic probe. It binds to small exposed hydrophobic patches on the surfaces of proteins (Ptitsyn et al 1990, Boothe et al 1997). This probe is fluorescent in a hydrophobic environment. It does not bind to folded or unfolded molecules but will bind to folding intermediates, protein subunit interfaces and substrate binding sites and is therefore a useful probe for following protein folding, enzyme reactions and subunit interactions.

Nile red (Sackett & Wolff 1987) is an uncharged hydrophobic molecule whose fluorescence is strongly influenced by its environment. If the molecule is placed in a highly polar environment the fluorescence intensity, with excitation at 550nm, increases greatly giving a blue shift and vice versa. This ability can be used as a method to probe the formation of hydrophobic surfaces.

In the latter two experiments only the fluorescence of the probes is monitored, i.e. the extrinsic fluorescence, in the other experiments the fluorescence of the protein is looked at. This is known as the intrinsic fluorescence.

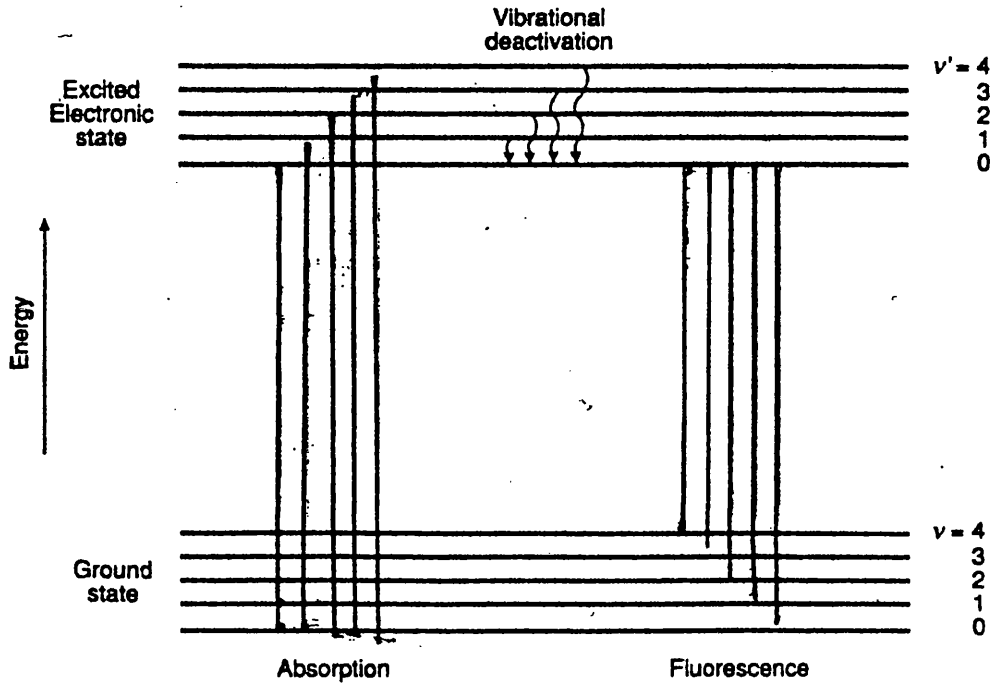


Figure 2.3 : Diagram showing absorption and fluorescence electronic transitions. Excitation is from the zero vibrational level in the ground state to various higher vibrational levels in the excited state. Fluorescence occurs when there is a transition from the zero vibrational level in the excited state to the various vibrational ground states, from d'Albis & Gratzer 1970.

Ultra-Violet (UV) Spectroscopy

UV spectroscopy depends on the ability of the chromophore to absorb light. The chromophore in the case of proteins is the aromatic amino acids and disulphide linkages (Drake 1994 a). The Beer Lambert Law ;

$$c = \frac{A}{\varepsilon \quad l}$$

is the relationship between A - absorption, and c- concentration, l is cell path length and ε is the extinction coefficient which is the measure of the probability of the transition from ground state to higher state. The extinction coefficient is the sum of extinction coefficients of each of the aromatic amino acids and cysteine residues (Gill & von Hippel 1989).

Circular Dichroism CD

Circular dichroism measures the difference in absorption of left and right handed circularly polarised light. This difference is small and cannot be looked at using usual ultraviolet or visible methods (Drake 1994b, Kelly & Price 1997).

The presence of an optically active/chiral sample leads to preferential absorption during one of the polarization periods giving a variation of the transmitted light at a specific wavelength or a CD spectra of a sample.

CD gives insights into the structure of a folded molecule and can also be used to gain information about what happens to a protein molecule during its unfolding or folding transition.

Measurements taken in the wavelength range 250-170nm, the far UV, gives information about the secondary structure. The interaction of light with the optically active molecule or chromophore leads to a π - π^* and an n - π^* transition, the chromophore in this range is the amide group. The interactions between amide groups in a folded protein molecule, for example hydrogen bonding in an α-helix, gives the molecule the necessary chirality for CD activity. Far UV can also be used to monitor changes in conformation and local environment caused by addition of agents, such as ligands or cofactors, into the sample solution. It is

regularly used to monitor protein unfolding caused by the addition of denaturants (Kelly & Price 1997).

The near UV is measured from 310 - 240nm, transitions in this range are due to the positions that the aromatic amino acid side chains assume when the protein folds into its tertiary structure, putting these side chains into chiral environments. The disulfide linkages can also effect the CD in this range.

Gel Electrophoresis

This is a method for separating proteins (Laemmli 1970) and estimating their approximate molecular weight. Molecules, such as proteins, which carry an overall net charge can be separated as they move in an electric field. If they move through some sort of matrix, such as an acrylamide gel, this behaves as a sieve which can be used to enhance separation, where small proteins move through the gel quickly and large proteins move through it slowly.

The proteins are diluted in sample buffer which contains SDS (Sodium Dodecyl Sulfate) a protein denaturing agent, so the samples are denatured when they are loaded onto the gel. The SDS binds to the denatured protein giving it an overall negative charge which overrides the negative charge of the protein giving a charge density which is independent of the size and composition of the protein, consequently the mobility through the gel matrix is determined by molecular mass.

Isoelectric Focussing (IEF)

This is an electrophoretic method similar to the one mentioned above which separates proteins by differences in isoelectric point (pI), instead of by size. The electrophoresis gels used here have amphoteric molecules fractionated according to their pI's along a continuous pH gradient. The gradient is created by the passage of current through a mixture of amphoteric compound with closely spaced pI, encompassing a given pH range (Righetti 1989). The surface charge on the molecule changes as it moves along the pH gradient until it reaches the point where it has zero net charge. This pH is the isoelectric point of the molecule.

Electrospray Mass Spectrometry (ESMS)

This technique can be used to measure the molecular weight of a sample, to determine the sequence of a protein or to characterise sites of post translational modification. The technique requires only small amounts, typically sub picomole concentrations, and maintains structural integrity in the gas phase by virtue of being a "soft ionisation" technique. (Johnstone & Rose 1996, Robinson & Radford 1995)

ESMS separates and measures the mass of a molecule by looking at mass to charge ratio (m/z), therefore the molecule or molecules being examined have to be ionized. This is achieved by the addition or subtraction of atoms or molecules leading to the generation of charged species. In positive ionisation addition of a proton gives the molecule a 1+ charge, so when the charge (z) equals 1 the mass (m) is the mass of the molecular ion. Such molecular ions have an internal energy which can lead to fragmentation of the molecule and production of fragment ions. This is often exploited, especially in tandem mass spectrometers (MS/MS) to elucidate structural information. Because of the initial range in molecular ions, after a short time many ions are present. Hence the number of peaks on a mass spectrum. The amount of ions depends on the individual rates of formation and decomposition of the ions, and on the initial energy imparted to the molecule in the first place.

A mass spectrum shows the mass to charge ratio against the ion abundance, the ion abundance is a normalised record, the largest peak is taken as the base peak and the relative heights of the other peaks are compared against this.

Electrospray works by producing a fine mist of charged solvent droplets containing the protein of interest, sequential evaporation of the solvent leads to the release of the protein ions. This is known as ion evaporation and is one of the main ionization processes in ESMS. The electrospray is performed at room temperature and atmospheric pressure, using solvents such as methanol/water with methanoic acid as a source of protons. Subsequently, the ions are drawn into the vacuum of the mass spectrometer by a potential gradient for detection.

The sample solution passes through the capillary to the ion source, where a large electric field is applied disrupting the emerging liquid surface and providing the droplet spray. This process can be assisted by a pneumatic spray where a constant gas flow can help shear droplets from the liquid stream (ion spray), which allows a larger flow rate. Electrospray removes solvent by combining a flow of dry warm gas (nitrogen) over the sample and by the

provision of a heated ion source. This process is known as desolvation. The charged droplets shrink, concentrating the charge density and when the repulsive forces become nearly the same as the surface tension the droplet explodes, making smaller droplets. This process is outlined in figure 2.4 . Eventually when the droplets get to a critical size the protein ions are released into the gas phase. The ions can be separated by m/z ratio by magnetic field or electric field. In the case of quadrupolar mass spectrometry, the ions are separated whilst they travel through a chamber which contains 4 parallel electrodes two of which are positively charged and two negatively charged. The molecular ion trajectory through this system can be described as being stable or unstable and is achieved by rapid switching of pairs of electrodes between DC and Rf voltage.

The ions that get through the quadrupole have a stable trajectory and are the ions that are seen by the detector and that result in the mass spectrum. Measurement of a protein`s mass by mass spectrometry is only possible because they can carry multiple charges, as their large mass puts them out of the spectrometer`s detector range. However, because they can be protonated at different sites giving a m/z (mass/ charge) ratio, this can then allow their mass to be calculated using mass spectrometry. The resulting spectrum will show a series of ions differing by 1 charge, these can be compared against each other in a set of simultaneous equations and the overall molecular weight calculated by a process called deconvolution. Given reasonable quadrupole calibration, accuracy in the order of 0.01 - 0.1 % is achieveable for such average mass molecules.

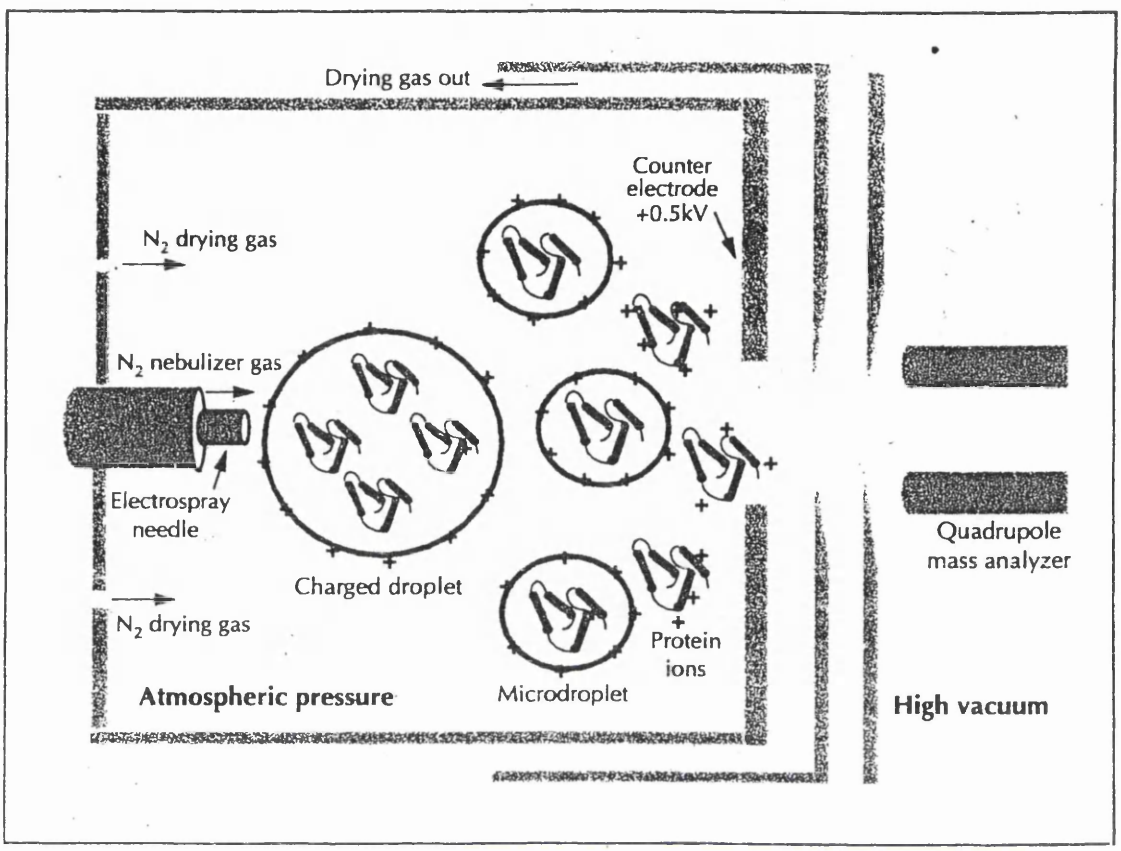


Figure 2.4 : Diagram outlining the electrospray process. The protein solution is injected into the mass spectrometer through an electrospray needle using a nebulizer-assisted spray. The solvent evaporates forming highly charged microdroplets. These release the protein ions into the gaseous phase after further drying. These ions are then drawn through the electrospray source into the mass analyzer. Taken from Robinson & Radford 1995.

Electrospray Mass Spectrometry : Hydrogen Exchange Experiments

Proteins in solution are in a state of constant conformational fluctuation and this means that some of the amino acids that make up the protein are able to react with reagents in the solution. In particular the protons of amino acid side chains are able to exchange with protons in the solution. If the solvent that the protein is in contains deuterons (^{2}H/D) these can exchange with the exchangeable hydrogens on protein side chains and vice versa. Mass spectrometry can be used to detect the resulting change in molecular weight which allows calculation of the number of exchangeable residues. This number is variable depending on the folded state of the protein (Last & Robinson 1999). Also the charged state of the protein detected in the mass spectrometer can be related to how native-like the protein fold is (Katta & Chait 1991). Mass spectrometry has been used for these experiments as, unlike NMR, it will not give an averaged picture of the protein structure, and thus allows identification of populations of proteins with different masses. It also requires a lot less sample.

Exchangeable protons are found in the peptide backbone amides and certain side chain residues. The rate and amount of exchange relates to the exposure of these amides and side chains. Protons which are involved in hydrogen bonding and very buried residues exchange very slowly.

How hydrogen exchange in the interior of proteins occurs is not fully understood but it is thought that local unfolding or " breathing " is a possibility (Woodward et al 1982). This is used in the classical Linderstrom-Lang interpretation of hydrogen exchange :

$$folded \underset{k_{-1}}{\overset{k_1}{\Leftrightarrow}} open \xrightarrow{k_{ex}} hydrogen\ exchanged$$

The rate of exchange then relates to the amount of time the structure is open. There are two possible mechanisms for this; the EX_1 mechanism, which occurs when the rate of exchange is greater than the rate of refolding of the open structure $k_{ex} > k_{-1}$. In electrospray experiments this type of exchange would show 2 peaks, which relate to the fully exchanged form and the protected form. As exchange is occurring faster than refolding, the protected fully folded form and the deuterated fully folded form are seen. The opposite case where $k_{ex} < k_{-1}$ is referred to as the EX_2 mechanism. In electrospray experiments this type of mechanism gives one peak, as refolding is happening at a rate faster than deuterium exchange the fully protected form is only seen. For example in a study of the equilibrium folding behaviour of

HEWL using hydrogen exchange and mass spectrometry (Miranker et.al 1993), where the rate of exchange was is slower than the rate of interconversion between the folded and unfolded forms 1 peak is seen at the beginning of exchange which by the end of exchange time has decreased in mass. Most proteins appear to demonstrate the EX_2 mechanism.

This method has been used to look at the folding of protein in an attempt to identify intermediates (Miranker et al 1993, Englander 2000), also to look at differences in exchange between folded and unfolded species (Chung et al 1997) and to look at conformational changes in proteins (Katta & Chait 1991). All of these studies indicate that it may be of some use in the misfolding phenomena studied here.

High Performance Liquid Chromatography HPLC

Chromatography is a molecular separation method, the term high performance relates to the equipment used (Scopes 1994). Separation is facilitated by an interaction between the molecule that is to be separated and the stationary phase of the column being used. The stationary phase is a porous support matrix with chemical groups bound to it, this support matrix must be relatively inert so only the groups bound to it cause the separation. An HPLC unit is just a means of pumping the eluent onto and through the column. It provides the high pressure needed to force the eluent/mobile phase through the column`s porous stationary phase. The eluent is the means of adsorbing the protein of interest onto the column, this is then desorbed or eluted off the column by adding another eluent as a gradient or by maintaining the solvent strength throughout the run (isocratic flow). Separation can be improved by manipulating either the mobile or stationary phase, as well as by using a longer column.

The main methods of chromatography are

- Ion Exchange : separation by charge.
- Hydrophobic Interaction : Hydrophobic Interaction.
- Affinity : separation by specific binding
- Size Exclusion : separates by molecular weight.

- Reverse Phase : Hydrophobic Interaction.

Because of the way the HPLC operates, caution is required with anything that is pumped through it. For example any protein that is loaded on to the column has to be centrifuged to remove any particulate matter and eluents must be regularly degassed, also the column needs to be washed through often to prevent any salt crystallizing out and destroying the column integrity.

The following is a description of the chromatography method mainly mentioned in this thesis.

Hydrophobic Interaction Chromatography (HIC)

This method of chromatography (Pharmacia a) requires the protein to be loaded onto the column in a high salt concentration (e.g. 1.5M Ammonium Sulphate) solution. This leads to a hydrophobic interaction between the protein and the solid phase medium. The bound protein is then eluted off with a buffer containing no salt. The amount of salt used is important, as it is imperative that the protein is eluted on in a salt concentration which is less than the concentration that causes salting out.

Chapter 3 : Materials and Methods

General Solutions

Buffers

All aqueous solutions used ultrapure deionised water. (Elga or Millipore)

0.1M citrate buffer pH 3.4 : 0.1M citric acid (Fisher Scientific, UK Ltd.) and 0.1M tri-sodium citrate dihydrate (Fisher Scientific, UK Ltd.) were mixed to a pH of 3.4.

Hydrophobic Interaction Chromatography buffer : 0.1M Phosphate pH 7: 0.1M solution of disodium hydrogen phosphate (Fisher Scientific, UK Ltd.) and sodium dihydrogen phosphate (BDH Laboratory Supplies) were mixed to a pH of 7.To 500ml of the above solution 1.5M Ammonium Sulphate (BDH Laboratory Supplies) was added to provide the salt solution for HPLC.

Ion Exchange Buffer : 0.1M buffer as above, elution buffer was 0.1M pH7 buffer with 1.5M NaCl (Fisher).

Trypsin Digestion : 0.1M Tris-HCl and Tris-base (Sigma) were mixed to a pH 7.5.

SDS polyacrylamide gels

2x sample buffer: 5g SDS (Schwarz/Mann Biotech), 5ml 1M Tris pH7.5, 100mM PMSF (Sigma) in isopropanol, 100mM EDTA (Sigma), 10ml glycerol, 2% bromophenol blue (BDH). Make up to volume with deionised water.

Running buffer stock solution : 1440g glycine (Sigma), 300g Tris-base (Sigma),50g SDS per 10L. Make up to volume with deionised water. This solution is then diluted 1:10 to make up to correct concentration.

Destain : 25% methanol, 10% acetic acid, 1% glycerol.

Coomassie Blue stain: 0.1% Coomassie Blue R-250 (Sigma) in destain.

Running gel : 15% : 2.35ml dH_2O. 2.5ml 1.5M Tris-HCl (Sigma) pH8.4, 0.1ml 10% SDS, 5ml acrylamide/bis-acrylamide solution (30% w/v - 0.8 w/v, ratio 37.5:1) (Scotlab) 0.05ml 10% APS (Sigma), 0.005ml TEMED for 2 stacks.

20% : 2.35ml dH_2O. 2.5ml 1.5M Tris-HCl (Sigma) pH8.4, 0.1ml 10% SDS, 5ml acrylamide/bis-acrylamide solution (30% w/v - 0.8 w/v, ration 37.5:1) (Scotlab) 0.05ml 10% APS (Sigma), 0.005ml TEMED for 2 stacks.

Stacking gel: 2.44ml dH_2O. 1ml 1.5M Tris-HCl (Sigma) pH8.4, 0.04ml 10% SDS, 0.52ml acrylamide/bis-acrylamide solution (30% w/v - 0.8 w/v, ratio 37.5:1) (Scotlab) 0.02ml 10% APS (Sigma), 0.004ml TEMED for 2 stacks.

Lysozyme Assay

β-N-Acetylhexosaminidase (Sigma) and p-Nitrophenyl Penta-N-Acetyl-β-Chitopentaoside (Seikagaku Kogyo Co. Ltd.).

Protein Analysis

Protein concentration measurement

UV absorbance measurements were made using a Shimadzu UV-160A double beam spectrophotometer, generally using two matched quartz cuvettes with a path length of 10mm. The baseline was corrected using cuvettes containing buffer alone. The extinction coefficient used to calculate lysozyme concentration was 38940 mol^{-1} L^{-1} cm^{-1}.

Gel analysis

The procedures for sodium dodecyl sulphate (SDS)-polyacrylamide gels were similar to those described by Laemmli (Laemmli 1970). 0.75mm gels were prepared in a Hoefer Mighty Small ™ SE245 Dual Gel Caster (Hoefer Scientific Instruments, San Francisco) and proteins electrophoresed through a 15% gel, unless otherwise stated, using a discontinuous buffering system. Lyophilised low molecular weight markers came from Amersham Pharmacia

Biotech. These were dissolved in 200µl of deionised water. All protein samples were diluted with an equal volume of 2x sample buffer, denatured by boiling for 4 minutes, unless otherwise stated, then electrophoresed at 30mA, 150V per 2 gels. Protein bands were visualised by staining with Coomassie Brilliant Blue R250 for 45 minutes, then destaining the gels until the desired background was achieved. Gels were carefully covered in cellophane and then dried using a BioRad vacuum drier.

Isoelectric focusing (IEF) gels

Isoelectric focusing was performed using a Pharmacia PhastSystem™. IEF gels pI 3 - 9 Phast Gel™ (Pharmacia b) which contain Pharmalyte™ carrier ampholytes were purchased from Pharmacia. The gel was run at 2000V, 5mA, 7W and 16°C (programme 7 on this particular machine) . The gel was loaded onto the PhastSystems electrodes plastic side down and the protective cover removed to reveal the gel. 6µl samples are put onto a well stamp which was previously covered in film, the samples are taken up by capillary action into a comb. The comb is inserted into a holder above the gels which during the programme is lowered onto the gel allowing the protein to run into the IEF gel.

Staining was performed using the fast Coomassie stain shown in **appendix III**.

Method for heat treating HEWL

This procedure was developed to simulate the effects of repetitive DSC scans. A solution of approximately 2mg/ml concentration HEWL in 0.1M, pH3.4 citrate buffer, was decanted into test tubes. These were sealed using Nescofilm and placed in a beaker of water on a hot plate, this was then heated to 100°C. The power supply was then turned off and the test tubes were left in the water on top of the hotplate for 5 minutes, then they were placed in a bucket of ice for thirty minutes. This procedure was repeated 5 times.

Differential Scanning Calorimetry

Initial experiments were performed using a MicroCal MC-2 DSC fitted with EM electronics and N2a DSC nanovoltmeter. Subsequent work used a MicroCal MCS and, for later experiments on separated peaks L1-4, the more sensitive VP-DSC was used. All machines were computer controlled. In the first two machines protein concentrations of approximately 2mg/ml (0.1mM) were used. In the VP machine concentrations of less than 0.5mg/ml (0.025mM) were used.

All samples were degassed for approximately two minutes in a vacuum dessicator attached to a water aspirator, during which the samples were stirred using a magnetic stirrer and bar. Samples were introduced into the DSC cells using a syringe with teflon tubing attached in the case of the MicroCal MC-2 DSC, and by using a Hamilton syringe for the other two. In all cases the samples were put under pressure of 1-2 atmospheres (15 - 30 psi) N_2 or air in the DSC cell.

The experiments were performed using the following DSC conditions: a scan rate of 60°C/hr, with a filtering period of 15/16 seconds, (data collection occurred every 15/16 seconds). During most experiments the data were collected over the range 10 - 100 or 20 -100°C. If multiple scans were performed there was a 30 minute cooling period between each scan.

The first measurement taken in every set of DSC experiments was a buffer baseline, where buffer is placed in the sample cell instead of a protein sample. This was done to check the performance of DSC and for use as a baseline in data manipulation.

All the DSC thermograms in this thesis have been concentration normalised using Microcal Origin™ software.

VP DSC experiments involving the separated peaks (L1-4) required desalting using a PD 10 column (Amersham Pharmacia Biotech.) and extensive dialysis into pH 3.4 citrate buffer. Both pH and concentration of samples was measured after dialysis.

Electrospray Mass Spectrometry

Mass spectrometry for determination of molecular weight was carried out using facilities at Strathclyde University with help from Dr. A. Pitt. Mass spectrometry was performed on a VG quadropole mass spectrometer fitted with a pneumatically assisted electrospray (ion spray) source and controlled via the MassLynx version 1.6 software. (Micromass UK Ltd., Wythenshawe, Manchester). Capillary voltages were set at 3.6kV, extraction cone voltages were between 20-30 V and the focusing cone voltage was offset by 10 V. The source temperature was set at 75°C, the nebulising gas flow between 15-20L/h and the drying gas flow at 250-350L/h. Lens stack voltages were adjusted to give the maximum ion currents. The mass range from 1000m/z was observed for 3-$3^1/_2$ minutes, halting data collection when the signal weakened. The instrument was calibrated over the correct mass range immediately before the experiment using horse heart myoglobin (Sigma Chemical Co,. Poole, Dorset, UK). Protein samples were prepared by dialysis into sterile filtered water (Elga Option 3 Water Purifier) at 2 mg/ml concentration (approximately 20pmol/μl) and then diluted 1:1 using the carrier solvent (1:1 acetonitrile and water with 1% formic acid), this solution was then centrifuged for 2 minutes at 5000g and 20-50 μl was injected directly into the carrier stream. Analysis of the spectra was performed using the MassLynx program.

Hydrogen Exchange Mass Spectrometry

These experiments were performed (with great patience) at Oxford Centre for Molecular Sciences by Alex Last using Micromass Platform II spectrometer with a nanoflow electrospray ionisation source. Gold plated borosilicate needles were produced in house. Capillary voltage was 1.20kV, cone voltage 65V. The high and low mass resolution was 15. The ion energy was set at 0.9V and the ion energy ramp set at 0V. The multiplier was set at 720V. The analyser vacuum pressure was $4.4x10^{-5}$ mBar and the backing pressure $5.7x10^{-1}$ m Bar. The cycle time was 4.5 secs., the scan duration 3.95 secs., and the interscan delay 0.55 secs. . The mass over the range 1000 - 2000 m/z was monitored. The spectrometer was calibrated using HEWL. Data were deconvoluted using MassLynx version 5 software. (Micromass UK Ltd., Wythenshawe, Manchester).

The number of exchangeable residues were calculated by Mr. A. Last by looking at the sequence of the protein. Each residue has one exchangeable site apart from proline and

disulphide linked cysteines. Protons which make the residue positively charged are discarded. Two extra exchangeable protons are added on for the terminii (Last & Robinson 1999)

Protein samples with a concentration of approximately 100μM were used in a solution of pH 5 d-H_2O/formic acid. 1μL of the protein sample was added to 2μL 0.02μM Human Lysozyme (internal standard) and 7μL H_2O. 6μL of this sample was diluted into 50μL D_2O and sprayed immediately (sample concentration approximately 1.1μM). Samples were then removed every 2 minutes (for the first 30 minutes) and sprayed, this process was then repeated every 6 minutes (for the second 30 minute period), then every 12 minutes (for the third thirty minute period) and finally every 50 minutes for the next 2 hour period. The samples were sprayed until the spectra deteriorated.

The data were processed using Mass Lynx version 2.3, then smoothed and centered. Three peaks were picked as representative and then the mass calculated for each of these at each time point and the difference calculated, giving the amount of exchange.

Fluorescence measurements

Emission spectra were obtained using a Spex Fluoromax Fluorimeter (Spex Industries, Edison, NJ) using 2ml samples in a silica cuvette. Protein concentration in all experiments showed an UV absorption of less than 0.1at 295nm. Raman scattering caused by solvent was corrected for by subtraction of buffer baseline.

A) Quenching Experiments

A stock solution of 2.5M recrystallised succinimide (Sigma Aldrich, Poole Dorset) was made up and 20μl aliquots added to 2ml of a dilute protein solution (the control was 2mL buffer) and the fluorescence emission maximum noted (λex 295nm) until 200μL in total had been added. The experiment was also performed on free tryptophan as another comparison to show the quenching propensity of totally exposed tryptophans.

Data analysis was performed using the method of Eftink and Ghiron (Eftink & Ghiron 1976) and Microcal Origin™ software.

B) Hydrophobic Probe Experiments

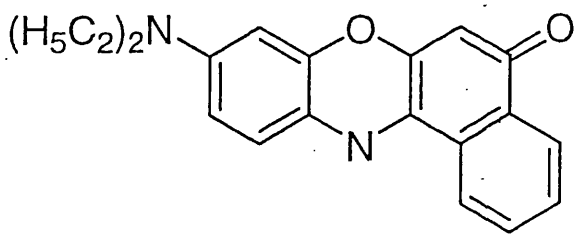
* 10μL of a fresh 20μM solution of ANS, 8-anilino-1-napthalene sulfonate (for structure see figure 3.1), (Aldrich Chemical Company, Poole Dorset) was added to 2mL protein solution with an A_{280} of 0.03. ANS binding is noticeable as a blue shift and increase in fluorescence intensity after fluorescence excitation at 390nm (Boothe et al. 1997).

*A solution of 0.25mM Nile red (for structure see figure 3.1), (Sigma Aldrich, Poole Dorset) in DMSO was made. This was added to the protein solution giving a final concentration of 1μM Nile red. The protein sample was then stirred, and allowed to stand for 15 mins before measurement of the fluorescence spectrum. The fluorescence emission spectrum was measured with excitation at 550nm. This experiment was performed a few times to compare the ability of Nile red to bind to HT HEWL and untreated HEWL (Sackett & Wolff 1987).

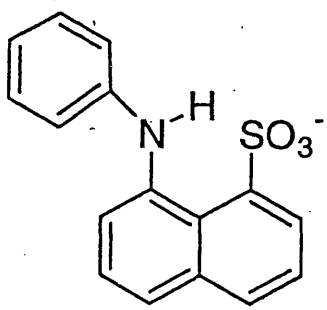
CD Experiments

The CD work was performed by Dr. Sharon Kelly at Stirling University. For all samples CD spectra were measured using a Jasco J-600 spectropolarimeter with a water bath circulator attached to thermostated cell holders to keep the cells at a constant temperature, in this case 25°C. Protein samples were 2mg/ml in the case of HT HEWL and native HEWL and approximately 1mg/ml in the case of the separated peaks. All samples were dialyzed into pH 3.4 20mM citrate buffer to lower the concentration of citrate buffer, which absorbs at below 200nm, i.e. the far UV region. Far UV CD (180 - 240nm) was measured using 0.02 or 0.05 cm cells, with 1mg/ml protein. Near UV CD (240 - 290nm) was measured using 1cm cells at 2mg/ml protein concentration.

Guanidinium hydrochloride (GdnHCl) unfolding experiments were performed by adding aliquots of protein to GdnHCl of known concentration. GdnHCl concentration was measured using the refractometer procedure described by Nozaki (Nozaki 1972). The circular dichroism was monitored at 225nm giving an indication in the changes in secondary structure induced by addition of denaturant.



(i) Nile Red



(ii) 1-Anilino 8 napthalene sulfonate

Figure 3.1 : The structure of Nile Red and 1-anilino 8 napthalene sulfonate.

Ammonium Sulphate Precipitation

This method was used after the heat treatment of the protein to concentrate samples for separation on the HIC column (Bollag et al 1996). Solid ammonium sulphate was slowly added to the heat treated lysozyme solution to give a final concentration of 85% by weight. The heat treated lysozyme was kept on ice with gentle stirring during the addition. Once all the ammonium sulphate had been added and dissolved the mixture was decanted into centrifuge tubes and spun at 10000g for 10 minutes. The supernatant was discarded and the pellet resuspended into buffer. This was centrifuged again for 10 minutes at 13000g to remove any excess undissolved protein.

HPLC :

Ion Exchange Chromatography

An Amersham Pharmacia Biotech Mono Q cation exchange column and Bio-Rad HPLC system were used. Loading buffer was pH 7, 0.1M phosphate buffer, and elution buffer 1M sodium chloride in 0.1M pH 7 phosphate buffer. Sample was 250µL, 2mg/ml heat treated lysozyme.

Hydrophobic Interaction Chromatography (HIC)

Preliminary investigations to find the best HIC medium for separation of heat treated lysozyme into the different misfolded forms was carried out using an HiTrap HIC test kit purchased from Amersham Phamacia Biotech and a Bio-Rad HPLC system. The best method was scaled up by packing a 25ml Omnifit glass chromatography column with Phenyl Sepharose® High Performance media. Packing was done by placing approximately 50 ml of the sepharose gel into a beaker with 450ml water and 250µl Tween 20 (Polyoxyethylenesorbitan monolaurate, BDH) which is added to decrease the surface tension of the slurry allowing the gel to pack more evenly. This was poured into the column with its bottom end closed, allowing gravity to pull the water through, and retaining the column packing. When the column was filled and the matrix had settled down evenly it was attached to the HPLC and 15% isopropanol was pumped through with increasing pressure, until the

column had bedded down. Then the top piece was screwed on, making sure no bubbles were introduced and the column was ready to use. The column volume was approximately 17ml.

The HIC/HPLC method that was routinely used is detailed in **appendix IV** , the column was equilibrated first with 0.1M pH 7 phosphate buffer and then with 1.5M ammonium sulphate in 0.1M pH7 phosphate buffer. The buffer was degassed and filtered to remove any particulates using a Nalgene® 250ml 0.25μM filter . Protein was injected onto the column via a 5ml loop purchased from Bio-Rad. Solid ammonium sulphate was added to the sample to give a final concentration of 10% by weight. The samples were then thoroughly centrifuged to remove any aggregate or any other particulate material.

Treatment of separated peaks

The separated peaks were concentrated by using Amicon Centriplus ™ (Millipore UK) centrifugal concentrators with a 15ml volume and 3000 Da. molecular weight cut off. They were then desalted using a PD-10 gravitational desalting column, from Amersham Phamacia Biotech, prepacked with Sephadex® G - 25 M stationary phase. The column was equilibrated with 25ml of the desired buffer. A maximum of 2.5ml of the protein solution was applied to the column and allowed to flow through the matrix. It was then eluted off using 3.5 mls of the same buffer, the protein elutes off in the first 2mls (approximately). This protein solution was then dialyzed into the required buffer using a Pierce Slide-A-Lyzer® dialysis cassette, molecular weight cut off of 3500Da, the dialysis buffer changed a minimum of 3 times. Any further concentration was done using Amicon® Centricon® centrifugal concentrator with a 3000Da molecular weight cut off.

HEWL activity assay

(Nanjo, Sakai & Usui 1988) To an approximately 50 μg sample of HEWL in 1.25ml pH 5 citrate buffer, 150μl of 0.09mM (1mg/ml) para - Nitrophenyl Penta - N - Acetyl -β- Chitopentaoside (note that before use this solution needs to be spun down to pellet out any excess substrate) and 100μl, 0.2units of β - N - Acetyl hexosaminidase were added. This solution was then incubated in the UV spectrophotometer at 37°C for 15 minutes, after this

time 1.5ml of 1M sodium carbonate was added to increase the pH and make the para-nitrophenol product turn yellow. The UV was then read at 405nm, and used to calculate the concentration of free para-nitrophenol (ε_{450nm} 18000 mol^{-1} L^{-1} cm^{-1}). This is then related to the activity using the following equation :

$$\frac{[\text{p-nitrophenol}]}{7.5\ [\text{HEWL}]}$$

giving the activity in number of moles of p-nitrophenol released per minute.

Trypsin digestion

(Seielstad et al 1995, Hubbard 1998, Price and Johnson 1990) Samples were desalted and then dialyzed into pH 7.5, 0.1M tris/HCl buffer. The concentration of the protein was measured and samples with a concentration of approximately 0.5mg/ml were used. 2 x 150µl of the samples were brought up to temperature in a 37°C waterbath and then to one set of samples 50µg of trypsin (100µg - 1mg) was added. The other samples were used as controls.

At different time points over 24 hours including t=0, 20µl of the sample was removed. 20µl of reducing sample buffer; which was SDS PAGE sample buffer (see SDS PAGE) with 10% β-mercaptoethanol added, was added to the 20µl of sample. The samples were then boiled for 10 minutes and frozen at -20°C. The sample was subsequently run on an SDS PAGE gel, the gel was stopped when the dye front is approximately 1 cm away from the bottom of the gel. The gel was stained and dried as before.

Chapter 4 : Results and Discussion

Section A : Preparation and Gross Properties of Thermally Misfolded Lysozyme

Although thermal unfolding of globular proteins is in principle reversible, in practice this is rarely true due to various factors. Repeat thermal cycling of HEWL at pH 3.4 in the differential scanning calorimeter (DSC) yields four apparent folding isomers, see fig.4.1. These can be identified as the 3 peaks which are at a lower temperature than the main HEWL peak. Our working hypothesis is that these peaks are caused by misfolding and more specifically by proline isomerisation, where somehow the HEWL molecules have refolded with the proline residues in the wrong conformation. Proline is a small ring amino acid and its connecting amide bond can be either cis or trans and must be in the right conformation before the protein can fold correctly. HEWL has all of its prolines in the trans conformation normally, but if enough energy is supplied the 85kJ/mol activation energy barrier can be overcome and a proline can change its conformation.

In this first section of results the characterization of the heat treated lysozyme will be discussed. The term heat treated lysozyme (HT HEWL) here relates to native HEWL which has been heated from 20 - 100 °C, with 30 minutes cooling in between each heat treatment. This process is repeated 5 times.

This project was stimulated by an initial observation by Cooper (Cooper 1999) that multiple differential scanning calorimeter scans of HEWL (3.12mg/ml in pH 3.4 glycine/HCl buffer) gave a trace that showed four peaks as shown in figure 4.1. A normal HEWL thermogram under these conditions shows one symmetrical peak at approximately 60°C figure 4.2, this thermogram has three extra peaks at temperatures lower than the main transition peak. HEWL is generally thought of as a fully reversible refolding protein and because of this behaviour is used as a model protein. These three peaks are postulated to be caused by misfolding during the cooling periods between DSC scans whilst the protein is refolding. The misfolding is possibly due to proline isomerization as this explains the appearance of the 4 peaks. HEWL has 2 proline residues so in theory can have a total of 4 proline isomers.

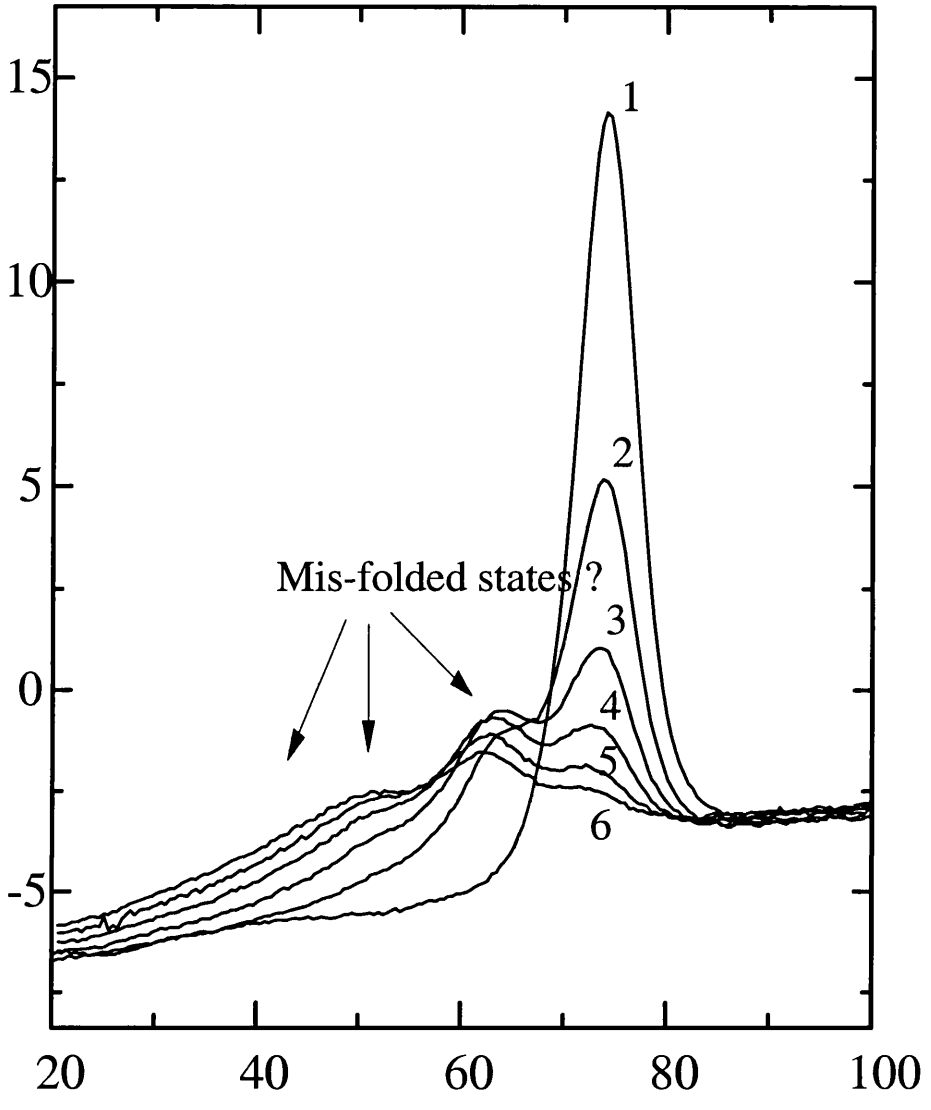


Figure 4.1 Repeat Differential Scanning Calorimeter (DSC) scans of lysozyme in pH 3.4 0.1 M glycine / HCl buffer, lysozyme concentration of 3.12mg/ml, scan rate 60°C/hr. This suggests the possible accumulation of misfolded forms of lysozyme as indicated by the appearance of peaks at lower Tm than lysozyme. Reproduced with the permission of Prof. A. Cooper.

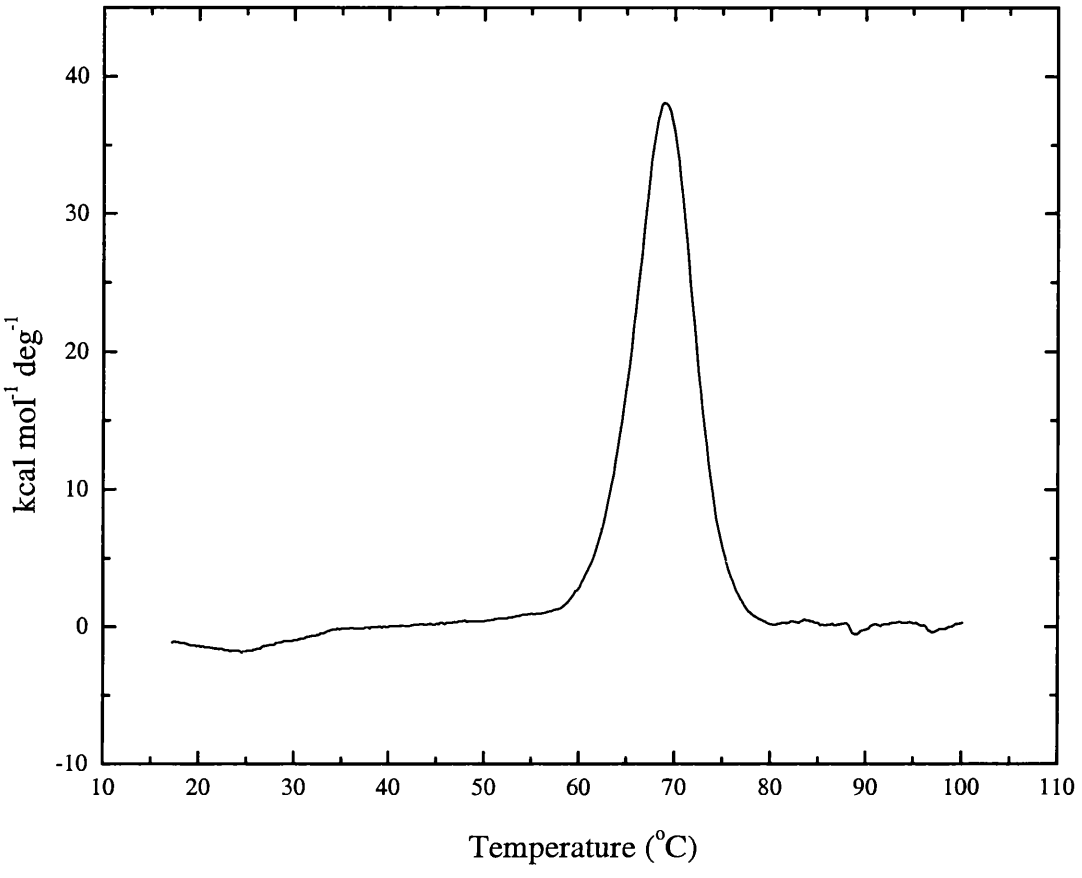


Figure 4.2 : Single DSC scan of Hen Egg White Lysozyme : 2mg/ml lysozyme, pH 3.4 0.1 M citrate buffer, scan rate 60°C/hr. This is an example of a typical symmetrical unfolding transition of a protein, the Tm or midpoint of unfolding temperature is 71.6°C in this case.

An initial concern was that this behaviour might be due to buffer salts present in Sigma lysozyme. The preliminary experiment was repeated using Sigma HEWL which had been extensively dialyzed against d-H_2O and then freeze dried. The resulting thermogram also had 4 peaks as shown in figure 4.3, indicating that the behaviour shown in the thermogram was not due to low molecular weight impurities in the HEWL used. Figure 4.4 shows each separate scan from this experiment and the appropriate unfolding model fit. Scan a shows one peak, Scan b can be fit with three unfolding peaks, Scan c can be fit with four peaks but the small peak at 60°C only fits in this thermogram. In Scan c the peak at approximately 65°C has increased in size between Scan b and c. Scan d can be fit with four peaks, the fourth one appearing in this fourth scan at 45°C, the second fit peak at 65°C has increased in size meanwhile the peak at 75°C, the main HEWL transition has decreased in size. This pattern is repeated in the final scan Scan d. These results allow us to conclude that four species are made by heat treating HEWL, that they all have a Tm at a lower temperature than HEWL and that the amount of native HEWL has decreased during the heat treatment.

Also, to ensure that the appearance of the extra peaks wasn't caused by one particular buffer, the experiment was performed using citrate buffer as well as glycine buffer at pH 3.4; fig. 4.5 shows the resulting thermogram. Again this showed 4 peaks like the other experiments. These experiments indicate that the "misfolding" occurs in both buffers giving thermal transition peaks at a temperature lower than the main transition, therefore this is not due to the presence of buffer salts in the Sigma HEWL nor is it a specific buffer induced effect.

The original DSC experiment was repeated using a variety of different lysozyme concentrations, 3/4/8/11 mg/ml HEWL in pH 3.4 citrate buffer. This was carried out to determine if the misfolding was a concentration dependent effect. The resulting thermograms, fig. 4.6, show that the same pattern of misfolding occurs at different concentrations. Also, although the thermograms show no apparent aggregation, above 5mg/ml the solution that came out of the DSC was cloudy and couldn`t be used in any further experiments. This indicated that increasing the concentration of HEWL in an attempt to produce more of the HT HEWL was not a viable option.

To produce the amounts of HT HEWL required for experimentation it was necessary to find a way to heat treat the HEWL other than in the DSC, as the DSC would only give 2ml of 2mg/ml protein at a time. For the majority of the experiments the HEWL (2mg/ml in pH 3.4

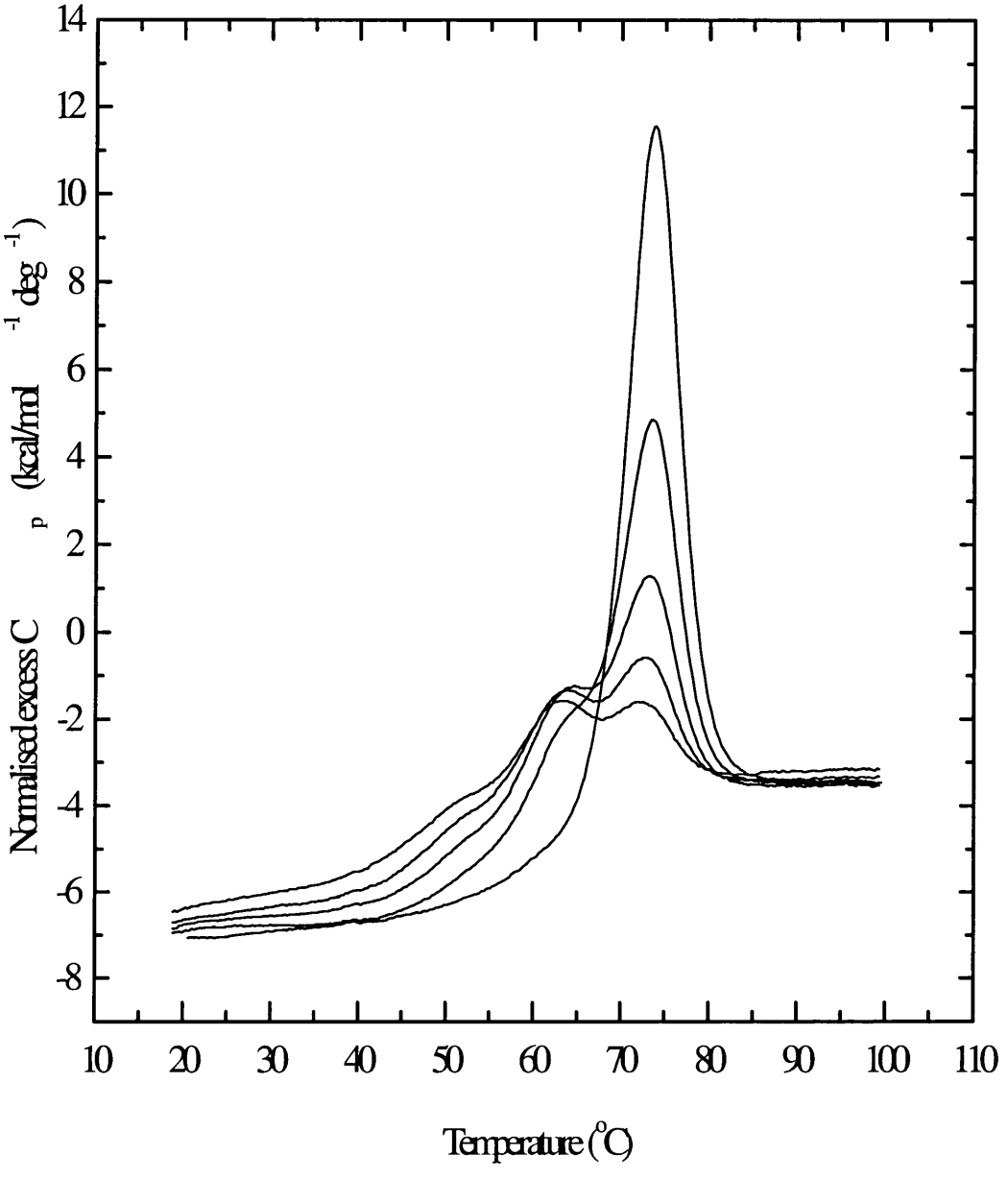
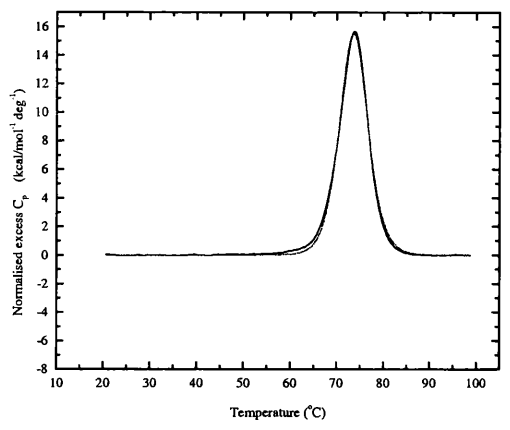
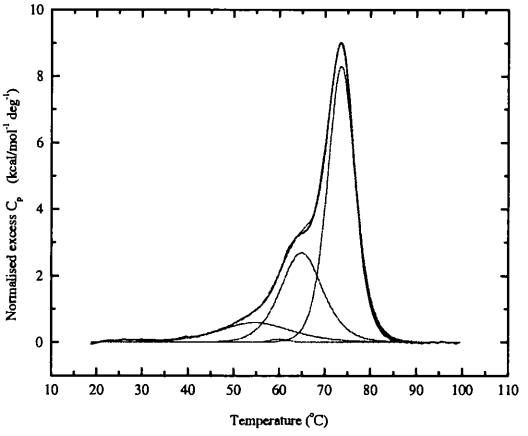


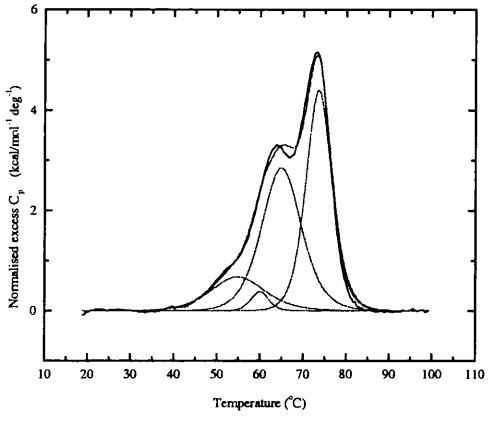
Figure 4.3 : Repeat DSC scans of dialysed and freeze dried HEW Lysozyme, 2mg/ml lysozyme in pH 3.4 0.1M glycine/HCl buffer, scan rate 60°C/hr. Evidence that the extra peaks that accumulate during the repeat heat treatment process are not caused by buffer salts which are present in Sigma HEW Lysozyme, these salts were removed in this case by dialysis and freeze drying.



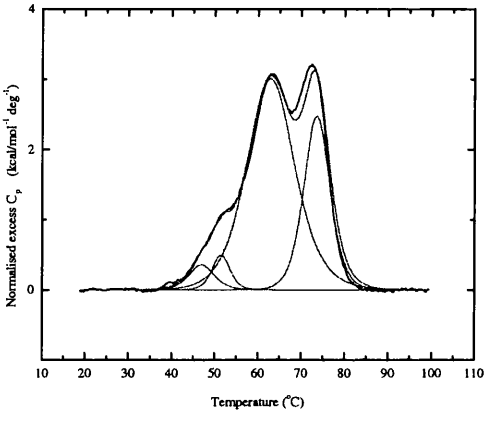
(a)



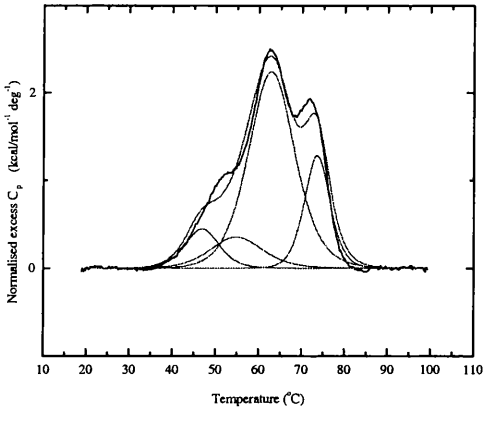
(b)



(c)



(d)



(e)

Figure 4.4 : Repeat DSC scans of HEWL, each thermogram shows a single scan, which has been fitted with either a two state unfolding model in the case of Scan a or a multi-state unfolding model in the other thermograms. Thermogram (a) Scan 1, (b) Scan 2, (c) Scan 3, (d) Scan 4, (e) Scan 5.

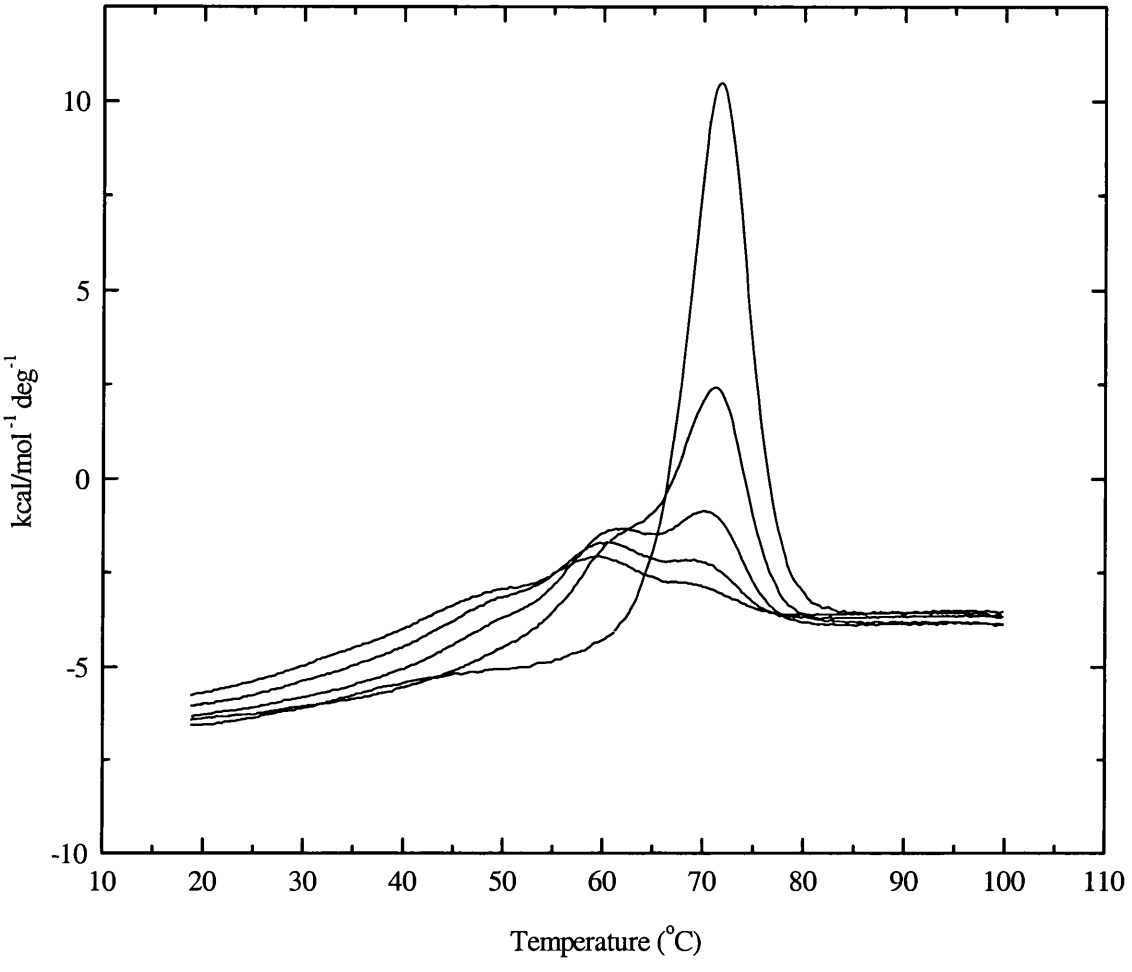


Figure 4.5 : Repeat DSC scans of HEWL, 1.92mg/ml lysozyme in pH 3.4 0.1M citrate buffer, scan rate 60°C/hr. Evidence that the extra peaks that accumulate during the repeat heat treatment process are not a side effect of using a particular buffer.

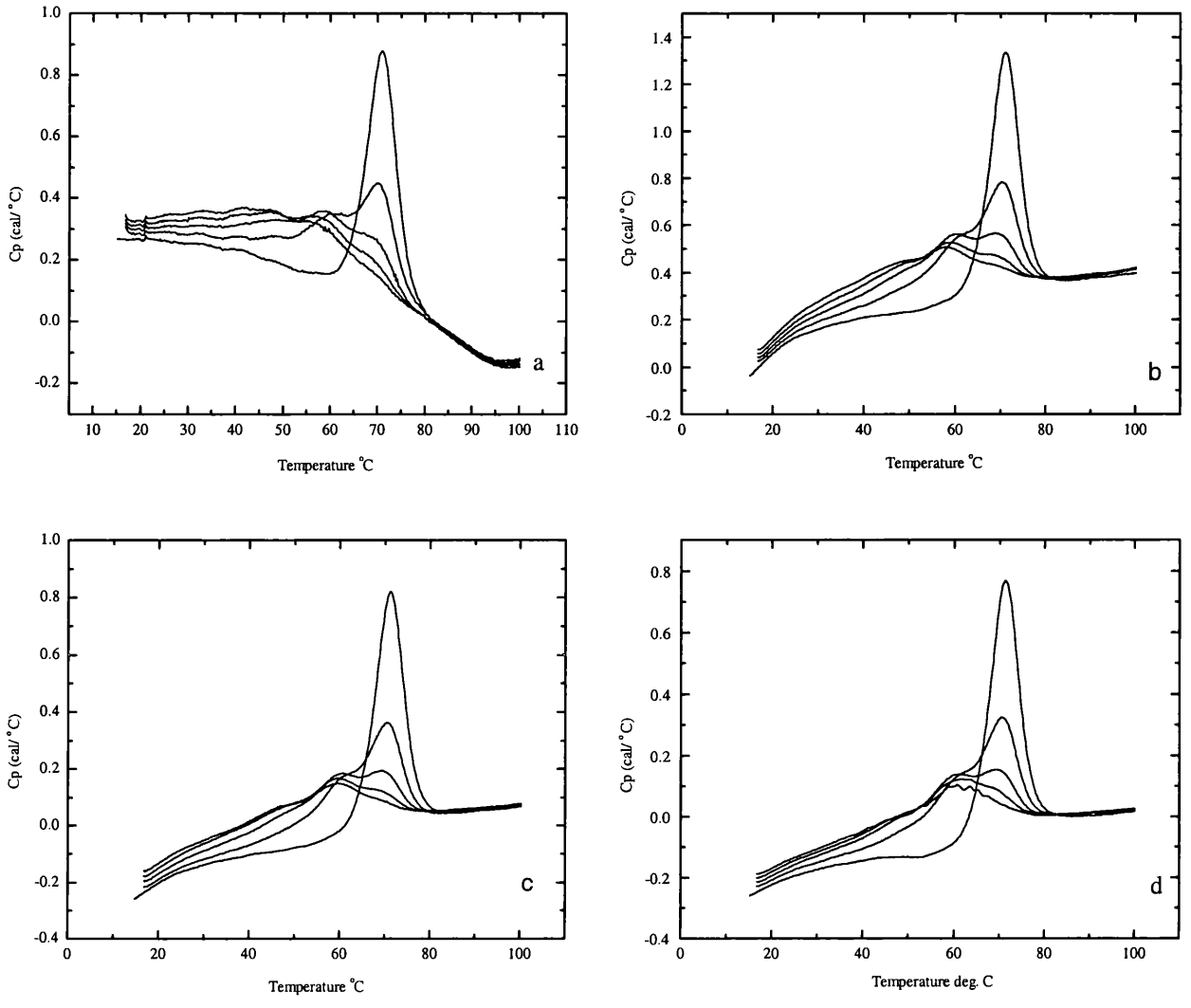


Figure 4.6 : Repeat DSC scans of HEWLysozyme, at varying concentrations of lysozyme in pH 3.4 0.1M citrate buffer, scan rate 60°/hr, to discover what effect changing protein concentration had on the production of the misfolded forms. Thermogram labelled (a) the lysozyme concentration was 3mg/ml, (b) 4mg/ml, (c) 8mg/ml, (d) 11 mg/ml. All data has been normalised for concentration. Although there is no DSC apparent aggregation, all of the samples removed from the calorimeters cells were cloudy and had a greater tendency towards aggregation in the long term.

0.1M citrate buffer) was heat treated by repeatedly heating in a water bath from 20 - 100°C and then cooled for 30 minutes in an ice bucket, mimicking the DSC.

The specific activity of the HT HEWL was compared to that of native HEWL using the p-Nitrophenyl-penta-n-acetyl-β-chitopentaoside assay (Nanjo & Sakai 1988). The enzyme activity of the HT HEWL was 1.045 (+/- 0.035) mols PNP min^{-1} mol HEWL $^{-1}$ compared to 1.01 (+/- 0.014) mols PNP min^{-1} mol HEWL $^{-1}$ for native HEWL. This showed that the HT HEWL had the same specific activity as native HEWL, suggesting that the misfolding caused by heat treatment had not affected the HEWL enzyme activity.

Figure 4.7 is an SDS PAGE gel comparing the molecular weights of the HT HEWL and native HEWL, this was done to see if the heat treatment process had cleaved or chemically altered the HEWL in any way. This shows that the HT HEWL ran in the same position as native HEWL, so the molecular weight of the two species are broadly similar. This indicates that no major chemical reaction, such as chain cleavage has occurred, altering the molecular weight of the HT HEWL. However, later experiments with reduced and denatured SDS PAGE gels indicated that there were some cleaved protein (peptides) present (fig. 4.36).

The similarity in MWt has been supported by ESMS, which indicated that the molecular weight of HT HEWL is the same as HEWL (fig. 4.8), although in the HT HEWL spectra there is a peak next to the HT HEWL peak with a M/Z ratio which identifies it as an HT HEWL species with a phosphate group associated with it. Also nanoflow quadropolar MS has shown HT HEWL species present with a M/Z ratio that identifies them as being HT HEWL with and without an extra water group associated to it. This could be due to the removal of a water molecule between two amino acid side chains. This can be seen in fig. 4.25 as the two peaks at the side of each HEWL with the molecular weights +/- 18.

Isoelectric focussing has also been used to compare HT and native HEWL, in order to exclude the possibility of deamidation (Ahern & Klibanov 1985) (fig. 4.9). If deamidation had occured the pI of the protein would have been affected. Deamidation is the conversion of an amide side chain to a carboxylate group, this effects the overall charge that the proten carries and hence the pI. This method separates proteins by isoelectric point (pI), which

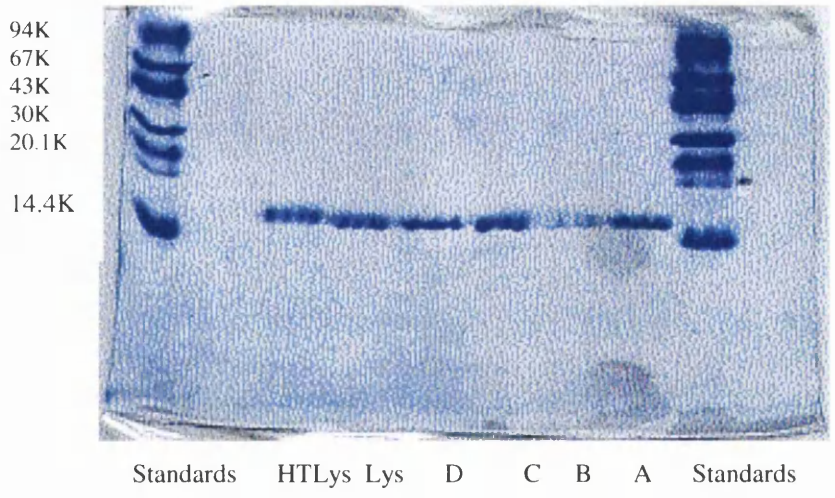
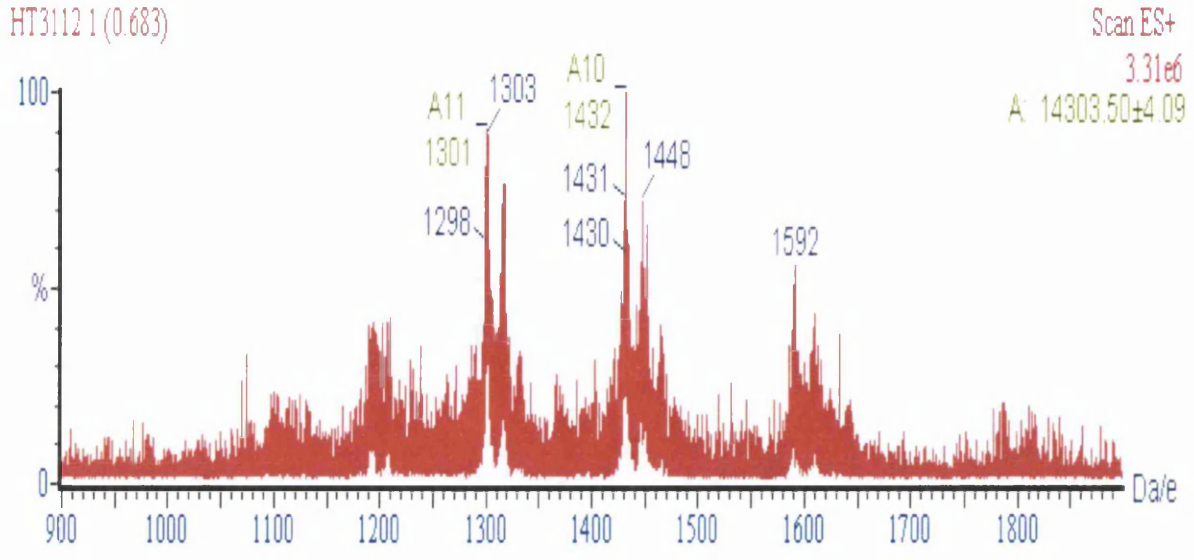
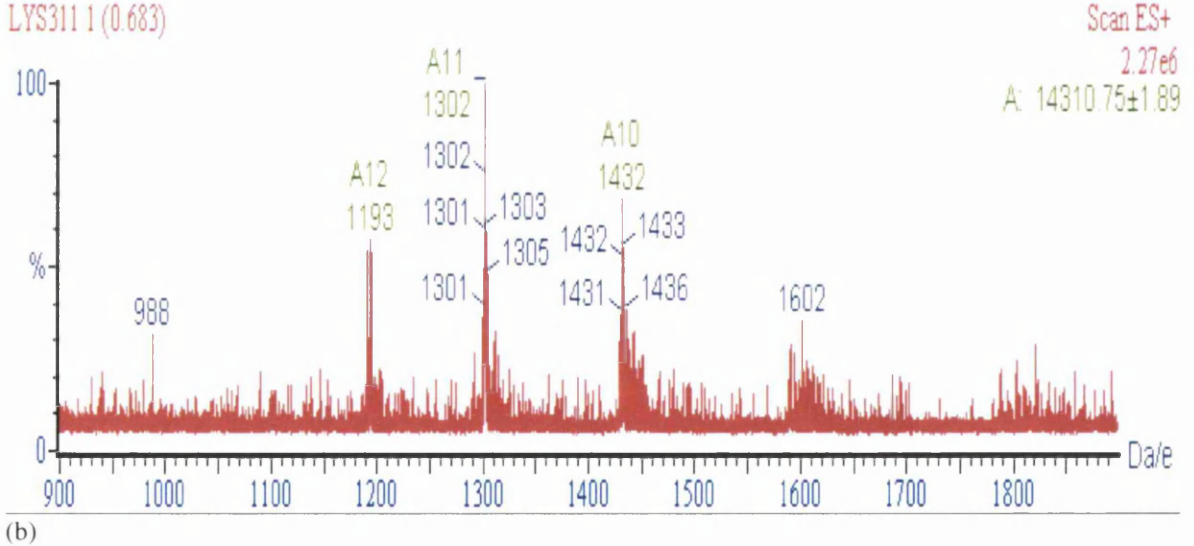


Figure 4.7 : SDS PAGE of native lysozyme, heat treated lysozyme and various mixtures of the two. HTLys - heat treated lysozyme, Lys - lysozyme, A 10 µl each of HTLys and native Lys, B 5 µl of each of HTLys and native Lys and 10 µl of buffer, C 15 µl of heat treated lysozyme and 5 µl of lysozyme and D 15 µl of lysozyme and 5 µl of heat treated lysozyme. In all case the proteins run to the same extent on the gel, and mixtures of the two proteins didn`t lead to resolution or separation of the two species.



(a)



(b)

Figure 4.8 : Electrospray mass spectra of HT HEWL (a), and HEWL (b). The calculated molecular weight of HEWL and HT HEWL using these mass spectra agree to within experimental error, thus the molecular weight of these two species is the same.

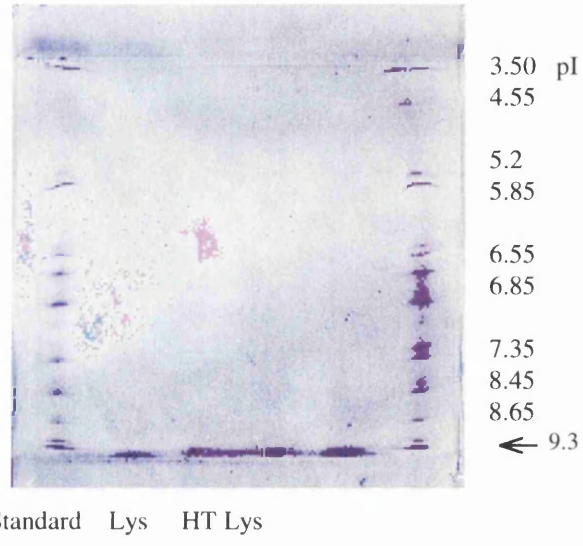


Figure 4.9 : IEF gel of native and heat treated lysozyme. Lys - lysozyme, HT Lys - heat treated lysozyme. The proteins both run to the same extent on this pI 3-9 gel.

would be altered if any chemical changes had occurred. Since HEWL has an isoelectric point of around 11(Worthington Biochem. 2000), and the IEF gels stop at a pI of 9 only a large difference in pI between the two species would be noted using this method. The samples of HT and native HEWL both ran to the end of the gel, see fig. 4.9, suggesting that there is no significant difference between their pIs.

Circular dichroism experiments were carried out to look for differences in the secondary/tertiary structure of the HT and native HEWL. In fig. 4.10 the far UV CD spectrum of HT and native HEWL is shown. The far UV CD is a measure of the secondary structure of a protein i.e. α-helices and β-sheet. The spectrum shown in fig. 4.10 indicates that the two species have similar secondary structure. The near UV CD spectrum of native and HT HEWL is in fig. 4.11. This method indicates the amount of tertiary structure of a protein. These spectra show no substantial difference in the tertiary structure between HT and native HEWL. However, a later near UV CD spectrum of HT HEWL indicates that there may be a small change in the contribution of aromatic amino acids towards the secondary structure. This is indicated by the crossover in spectra between native and HT HEWL (fig. 4.28), suggesting that there is a difference in how they are packed between the two species, although this could relate to a difference between the two batches of HT HEWL. If any conformational changes have occured during the heat treatment they are too small to be picked up by CD.

The unfolding of the HT HEWL in GdnHCl was monitored using CD. This was compared to the unfolding of HEWL, in order to discover how the misfolding caused by the heat treatment has affected the folding stability or cooperativity of the HEWL. The results of this experiment are illustrated in fig. 4.12. This showed that the cooperativity normally shown by HEWL when unfolding, which is due to interactions between parts of structure in HEWL, had been destroyed by the heat treatment. The HEWL unfolding profile at low concentrations of GdnHCL (0-2M) indicates that 100% of the molecules are folded. As the concentration of GdnHCl reaches 4-5M all molecules are unfolded, the midpoint between these two is where 50% of the molecules are folded and 50 % unfolded. HT HEWL and HEWL have different unfolding curves presumably because the HT HEWL is a mixture of different isomers each with different stabilities. Each isomer possibly has a different unfolding point, therefore unfolding will be initiated by different concentrations of GdnHCl, hence resulting in an attenuated unfolding profile.

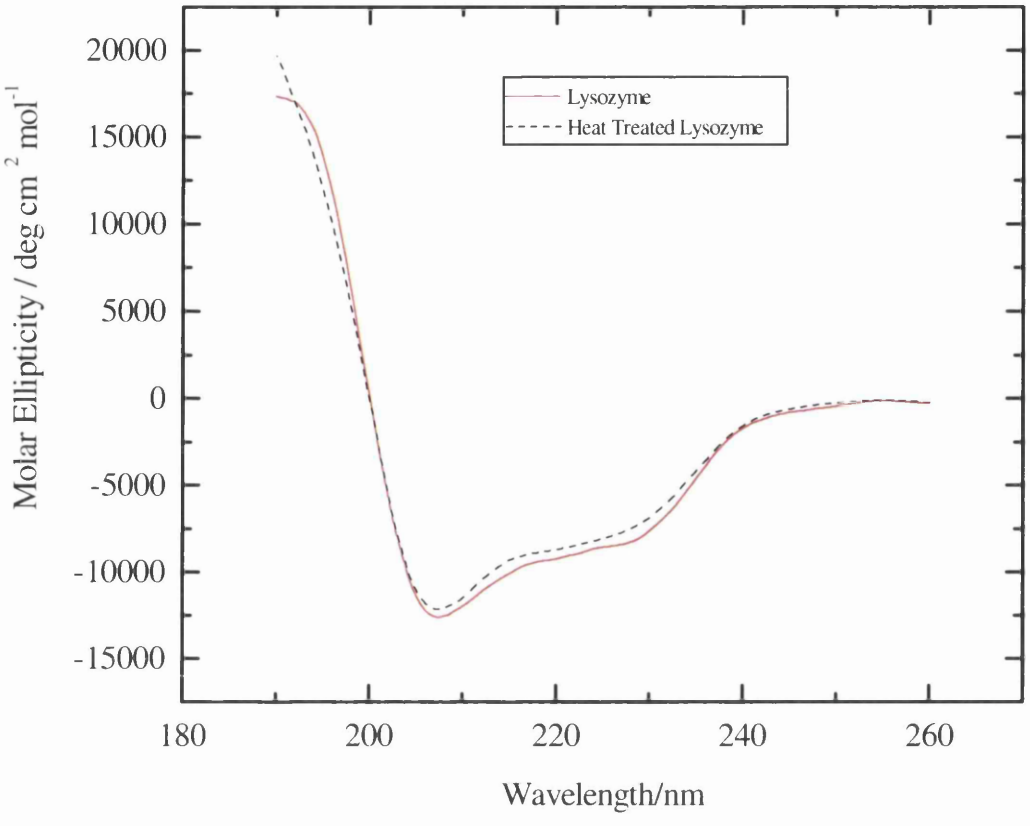


Figure 4.10: Far UV CD spectra of heat treated lysozyme, (broken line) : native lysozyme (solid line). This method gives an measure of the secondary structure of the protein. The shape of the spectrum is characteristic of a mostly α helical protein. In this case the shape and the height of the two spectra are similar, suggesting that there is no difference in secondary structure.

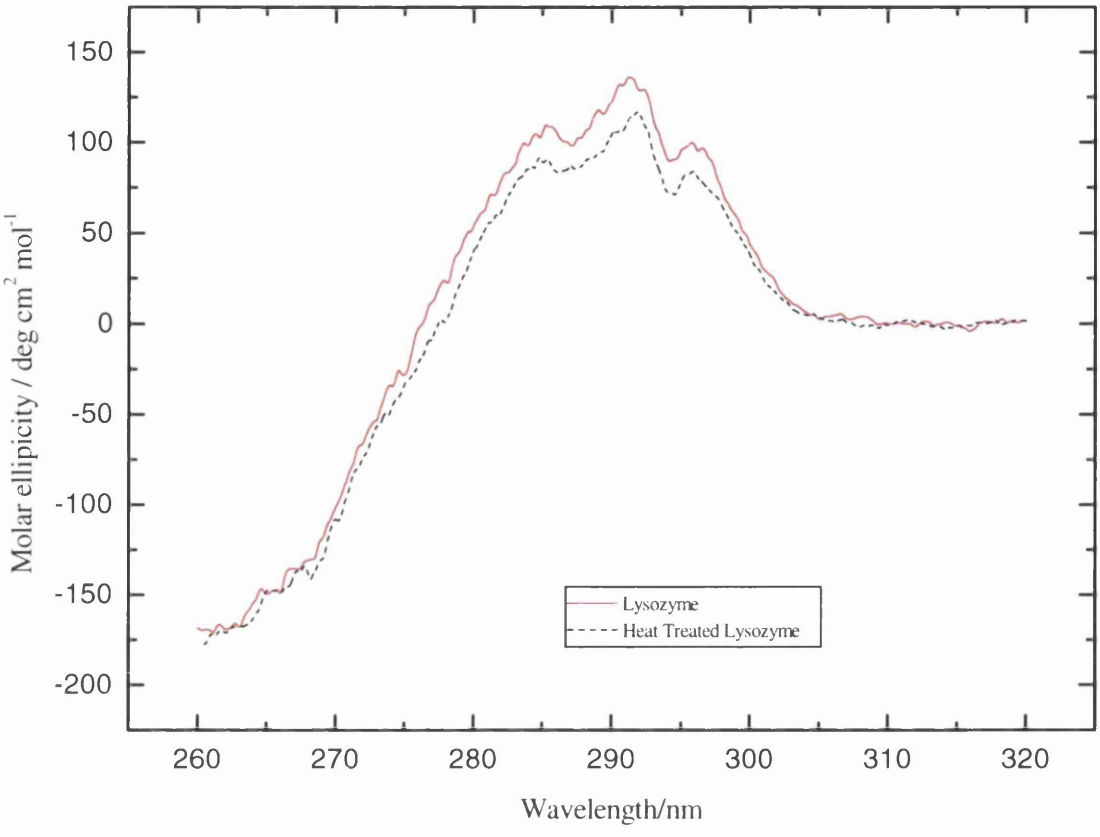


Figure 4.11: Near UV CD spectra of heat treated lysozyme, (broken line) : native lysozyme, (solid line). This method gives an measure of the tertiary structure of the protein, in this case the tertiary structure of both species is similar.

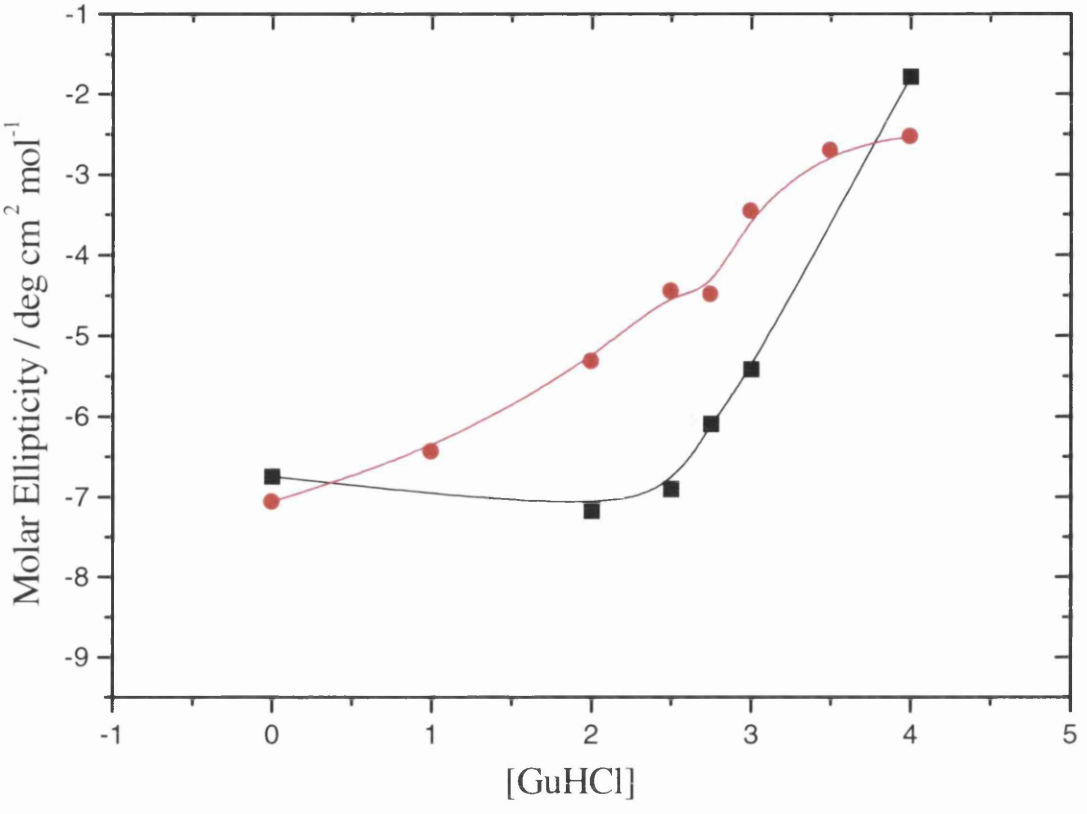


Figure 4.12: Guanidium hydrochloride induced unfolding of native lysozyme and heat treated lysozyme monitored by looking at the far UV CD spectra at 225nm. Native lysozyme unfolding curve marked by squares (■), the heat treated lysozyme by circles (●). Lysozyme shows a typical two state unfolding profile, whereas the heat treated lysozyme appears to have lower stability towards GuHCl.

The fluorescence spectra of HT HEWL and HEWL are compared in figure 4.13. These show that the fluorescence emission maximum (excitation at 295nm) of HT HEWL is 342nm, and HEWL maximum is at 340nm. The difference in maximum is not significant as it is only 0.04%. The greatest difference between the two spectra is the intensity.The difference in intensity may reflect a difference in the exposure of the tryptophan residues of the two proteins, particularly that the HT HEWL may have one or more tryptophans more exposed than native lysozyme hence the increase in intensity and also the red shift in fluorescence maximum, although the shift may not be significant.

The increase in intensity may also be because the HT HEWL is composed of more than one species, resulting in the addition of the fluorescence spectra of each species.

Fluorescence studies were used to compare HT and native HEWL. The graph shown in figure 4.14 is a plot of F_0/F (F_0 - fluorescence of protein with no succinimide present. F - fluorescence of protein in the presence of succinimide.) against succinimide concentration for native HEWL, HT HEWL and tryptophan. The gradient of these lines is the Stern-Volmer constant (Ksv) of each of these proteins, providing a measurement of the accessibility of tryptophan residues to succinimide as a quenching agent. Succinimide quenching of the heat treated protein compared to the wild type protein seems to show that the HT HEWL has a similar Stern Volmer constant to HEWL indicating that the quencher was equally able to come into contact with the tryptophan residues of the heat treated protein as those of the wild type form. This suggests that the misfolding has not affected the accessibility of the HEWL tryptophan residues to succinimide.

Nile red binding (Sackett & Wolff 1987) was used to assess any differences in accessibility to hydrophobic areas within the protein, as shown in fig. 4.15. The fluorescence emission spectra of HT HEWL with nile red had a greater intensity and was slightly red shifted, this suggests that HT HEWL has more hydrophobic areas accessible to nile red than HEWL.

The same experiment was repeated using ANS (Ptitsyn et al 1990) and the results plotted in fig. 4.16. ANS is another hydrophobic probe, and is a more popular method for assessing the accessibility of hydrophobic patches. The results of this experiment were exactly the same as when nile red was used. Figure 4.16 shows that ANS binds to beta lactoglobulin, which is a very hydrophobic protein, and is used as a standard in these experiments. The binding to beta lactoglobulin (fig. 4.16) resulted in a large increase in fluorescence intensity and a red shift in

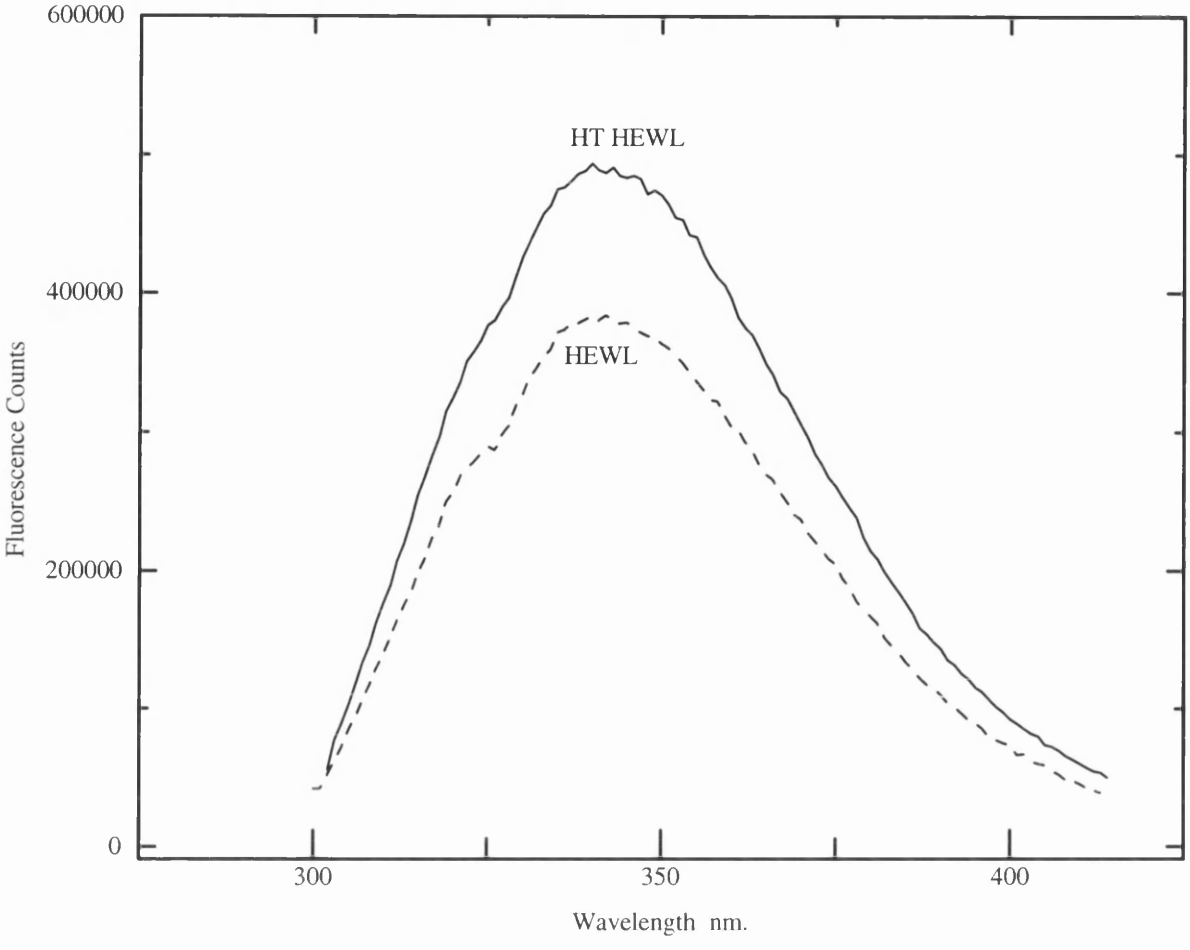


Figure 4.13 : The fluorescence emission spectra of HT HEWL and HEWL, excitation at 290nm.

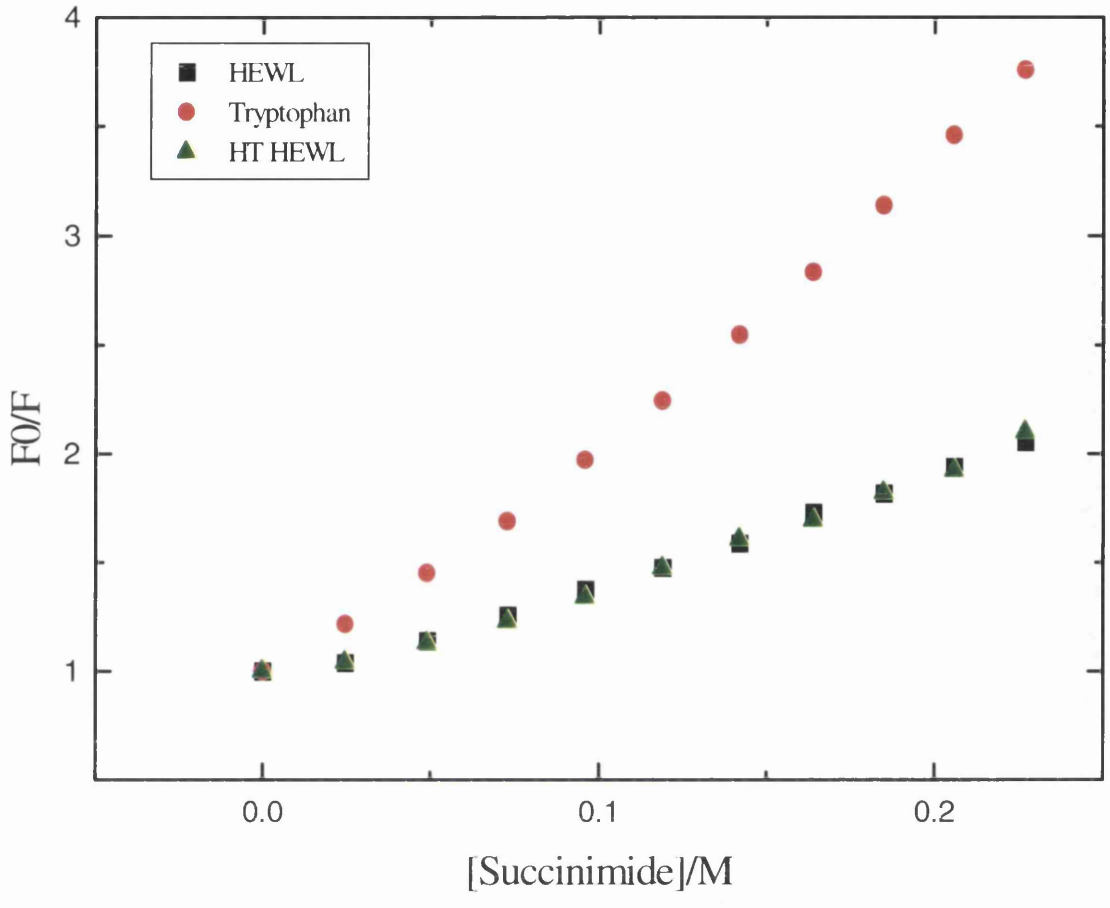


Figure 4.14: Succinimide quenching of available tryptophan residues in native lysozyme and heat treated lysozyme. Native lysozyme marked by filled in squares (■), the heat treated lysozyme by open triangles (▲) and free tryptophan by filled in circles (●).The gradient of the resulting lines when F0/F is plotted against succinimide concentration for native lysozyme and heat treated lysozyme is the same within the margin of error. Ksv for HEWL is 4.83 +/- 0.13, Ksv for HT HEWL is 4.93 +/- 0.19. These values are less than that of free tryptophan, which has a Ksv of 12.27 +/- 0.34.

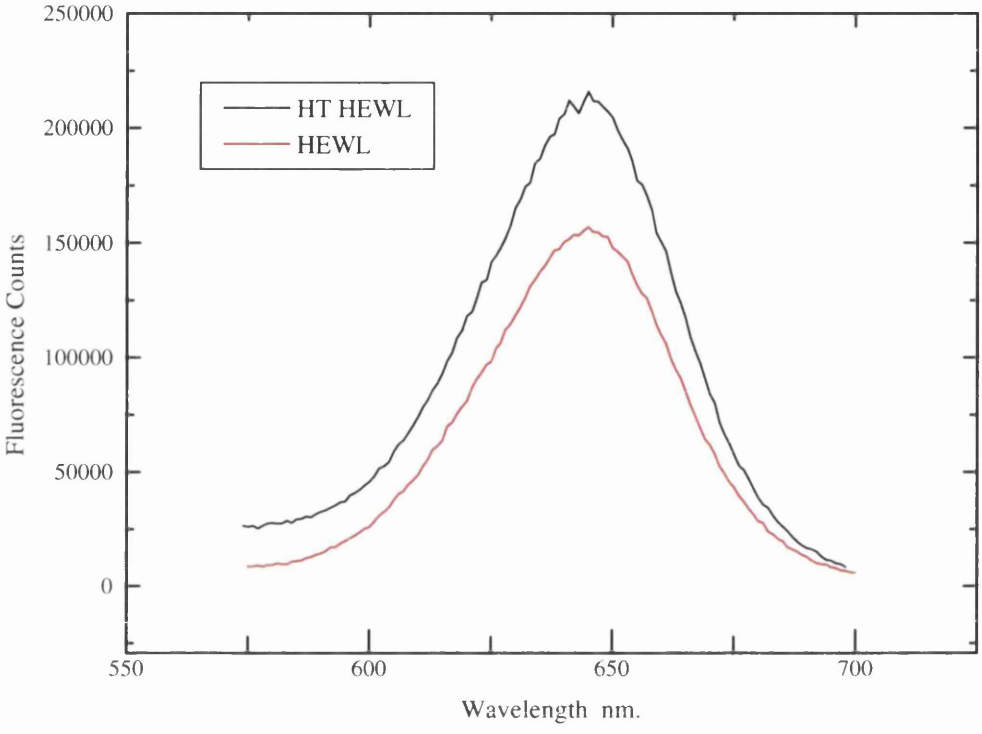


Figure 4.15: Nile Red binding to native lysozyme and heat treated lysozyme, measured by changes in the dye`s intrinsic fluorescence. Nile red binding to native lysozyme indicated by the solid line, to the heat treated lysozyme by the dashed line. Nile red appears to show a great affinity towards the heat treated lysozyme in comparison to native lysozyme.

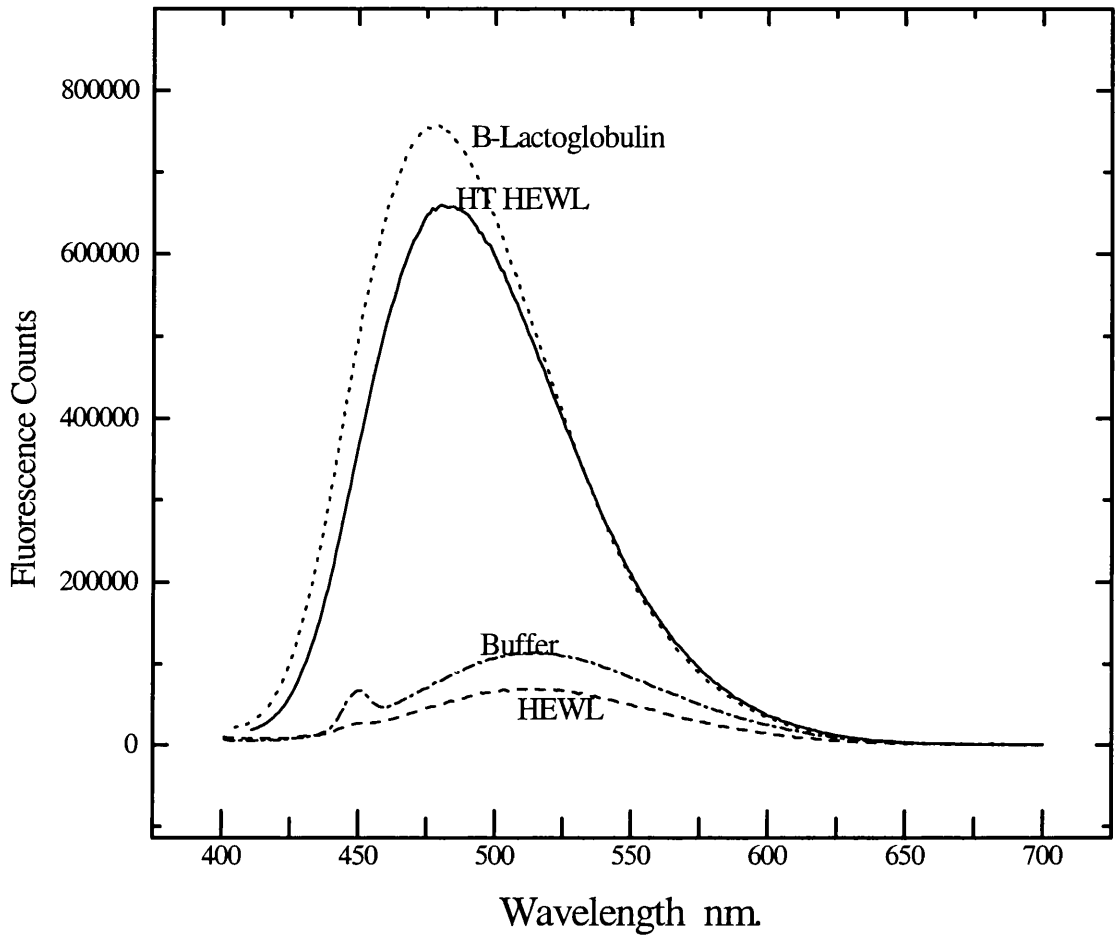


Figure 4.16: ANS binding to native lysozyme and heat treated lysozyme, measured by changes in ANS`s intrinsic fluorescence. ANS binding to native lysozyme, to the heat treated lysozyme and the standard β - lactoglobulin is shown. Also ANS binding to the pH 3.4 0.1M citrate buffer is displayed. ANS appears to show a great affinity towards the heat treated lysozyme in comparison to native lysozyme.

emission peak to lower wavelength. A similar effect is seen with ANS binding to HT HEWL, but not with native HEWL, which has no exposed hydrophobic sites. ANS has a similar affinity towards HT HEWL and the standard beta lactoglobulin. These experiments suggest that misfolding has exposed hydrophobic sites on HT HEWL which can bind probes of the size of nile red and ANS. It also allows suggests that hydrophobic interaction chromatography may be used to separate the misfolded forms of HEWL.

In summary : the experiments described so far on the crude mixture suggest that repeated thermal cycling of lysozyme through the folding/unfolding transition results in a mixture of conformers that have similar structural properties. SDS-PAGE, mass spectrometry, CD, fluorescence, enzyme activity assays, etc, indicate little or no gross structural chemical change in the proteins as a result of this treatment. Nevertheless, the different species show marked differences in thermal stability and response to denaturant, together with change in ability to bind hydrophobic probes, consistent with more subtle conformational changes.

This is also consistent with the working hypothesis that the effects are due (mainly) to trapping of "incorrect" cis-trans proline isomers in an otherwise more-or-less correctly folded protein.

There are other possible reasons for the behaviour of HEWL in the previous experiments such as aggregation, deamidation or disulfide rearrangement. This will be considered in the general discussion.

The next stage of the investigation involves separation of the individual isomers from the mixture in order to probe the proline isomerization hypothesis in more detail.

Section B : Isolation of the Misfolded Conformers

The separation of the heat treated lysozyme into the separate species identified by DSC was first attempted using a Mono Q ion-exchange column, using salt gradient to elute off the bound protein. This column separates proteins by charge. The column matrix has a positive charge which binds negatively charged proteins, so less negatively charged proteins will be eluted off first. This proved to be unsuccessful, indicating that the four species identified by DSC do not differ in charge under the conditions used for this experiment.

After discovering that HT HEWL bound ANS, indicating that the HT HEWL had exposed hydrophobic sites, separation of the HT HEWL was attempted using a Hydrophobic Interaction Chromatography (HIC) column.

Various conditions and different types of HIC columns were used, and eventually it was discovered that a phenyl sepharose column gave a reproducible chromatogram with 4 peaks (see fig. 4.17 (a)). Separation occurred when the heat treated protein was loaded onto the column at high ammonium sulphate concentration (1.5M) and the bound protein eluted off during a decreasing ammonium sulphate gradient. Elution of native HEWL off of the HIC column gave just a single peak (fig. 4.17 (b)), and appears to have the same elution pattern as the second peak of the HT HEWL chromatogram, fig 4.17(a).

The method was scaled up by packing a longer column with the chromatography medium and the heat treated protein was concentrated by ammonium sulphate precipitation. It should be noted however that the ammonium sulphate precipitation gave extra peaks on the chromatogram, but only the peaks that corresponded with the original chromatogram were collected. These extra peaks are the result of the ammonium sulphate precipitation process, which may have slightly altered the conformation of the protein thus exposing extra hydrophobic sites, giving them slightly different elution profiles from the main peaks.

The percentage area of the whole chromatogram that each peak takes up was calculated, pk1-L1 25%, pk2-L2 12.41%, pk3-L3 22.3% and pk4 -L4 8.8%, so peaks 1 and 3 make up the majority of the species created when the protein is heat treated and peaks 2 and 4 the minority. The separate species are referred to as L1 - 4. After separation, the peaks were reloaded onto the column after different time periods to see if the separate peaks would reisomerise when separated. It is known that that proline isomerization is a slow process (10 –100 secs (Schmid 1993)).

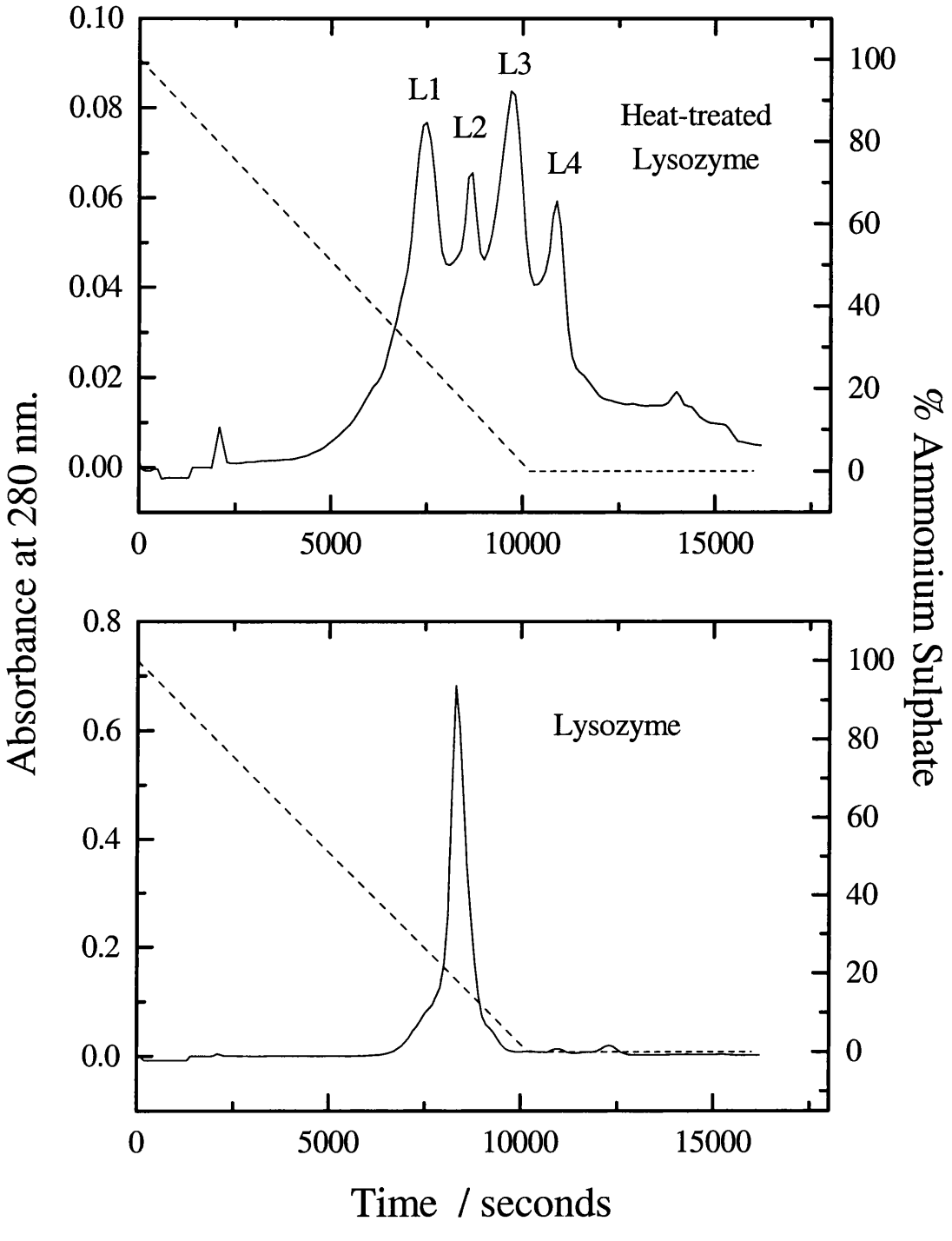


Figure 4.17: Hydrophobic interaction chromatography of heat treated and native lysozyme. The column was loaded with 1.5M ammonium sulphate in 0.1M pH 7 phosphate buffer and eluted using the same phosphate buffer. (a) Separation of heat treated lysozyme into 4 separate peaks. Four peaks were in the original DSC experiment. (b) Native lysozyme run down the HIC column under the same conditions at which separation occured. This appears to indicate that native lysozyme has the same elution profile as the second peak of the separation of HT HEWL

This experiment was designed to determine whether any of these factors would affect the elution profile of the separate peaks.

Figure 4.18 shows chromatograms of the four peaks which were reloaded onto the column within 24 hours of separation. These show that the peaks ran at approximately the same time and percentage of buffer B, as when they were originally separated. This indicated that within this time period no isomerisation had occurred. This experiment was repeated after 1 and 2 weeks, during which time the separate peaks had been stored at 4°C. In both these instances the peaks did not rerun as a single peak (see fig. 4.19 and 4.20), extra peaks appeared which corresponded, when compared to the original chromatogram, to the other three peaks implying that some sort of interconversion process has taken place. So over this time period some sort of interconversion process is occurring, where each separate peak is interconverting into the other peaks. The timescale for this interconversion is slow as would be expected at 4°C, which is consistent with proline isomerisation. It should be noted that the material which is eluted /washed off the column at low to 0% concentration of ammonium sulphate is aggregated or severely misfolded protein as it would have the greatest affinity to the column matrix.

The separated samples L1 to 4 were desalted and then dialyzed into pH 3.4 citrate buffer so that they could be heat treated as in the original experiment to see if the heat treatment would lead to the same pattern of peaks on the FPLC/HIC. The resulting chromatograms from this experiment are shown in figure 4.21. Peak 1 shows two peaks after reheating, these correspond to peaks 1 and 3. Peak 3 also shows the same pattern of peaks. Peak 4 gives three peaks which relate to peak 1,3 and 4. It should be noted that peak 2 when reheated gives four peaks. Peak 2 (before heat treatment) has the same elution profile as native HEWL.

In all cases the separate peaks after heating have reisomerised giving other peaks. This is a very interesting result indicating that the separate peaks are interconvertible isomers rather than just a static misfolded structure and that the interconversion is more rapid at higher temperature. This is consistant with the theory that the misfolding has been caused by proline isomerisation.

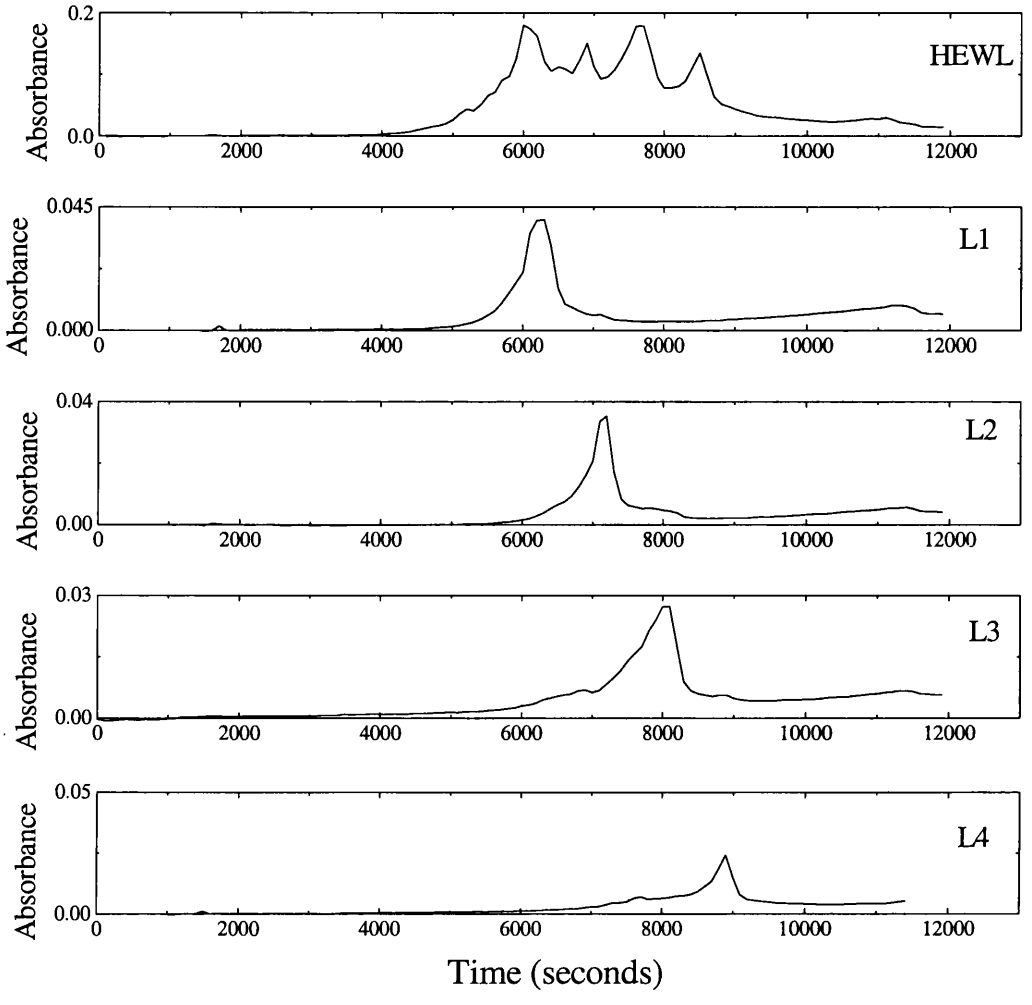


Figure 4.18 : The separated peaks rerun down the HIC column immediately after separation. (a) Heat treated lysozyme, (b) peak 1-L1, (c) peak 2-L2, (d) peak 3-L3, (e) peak 4-L4. This shows that the interconversion process doesn`t occur on a fast timescale.

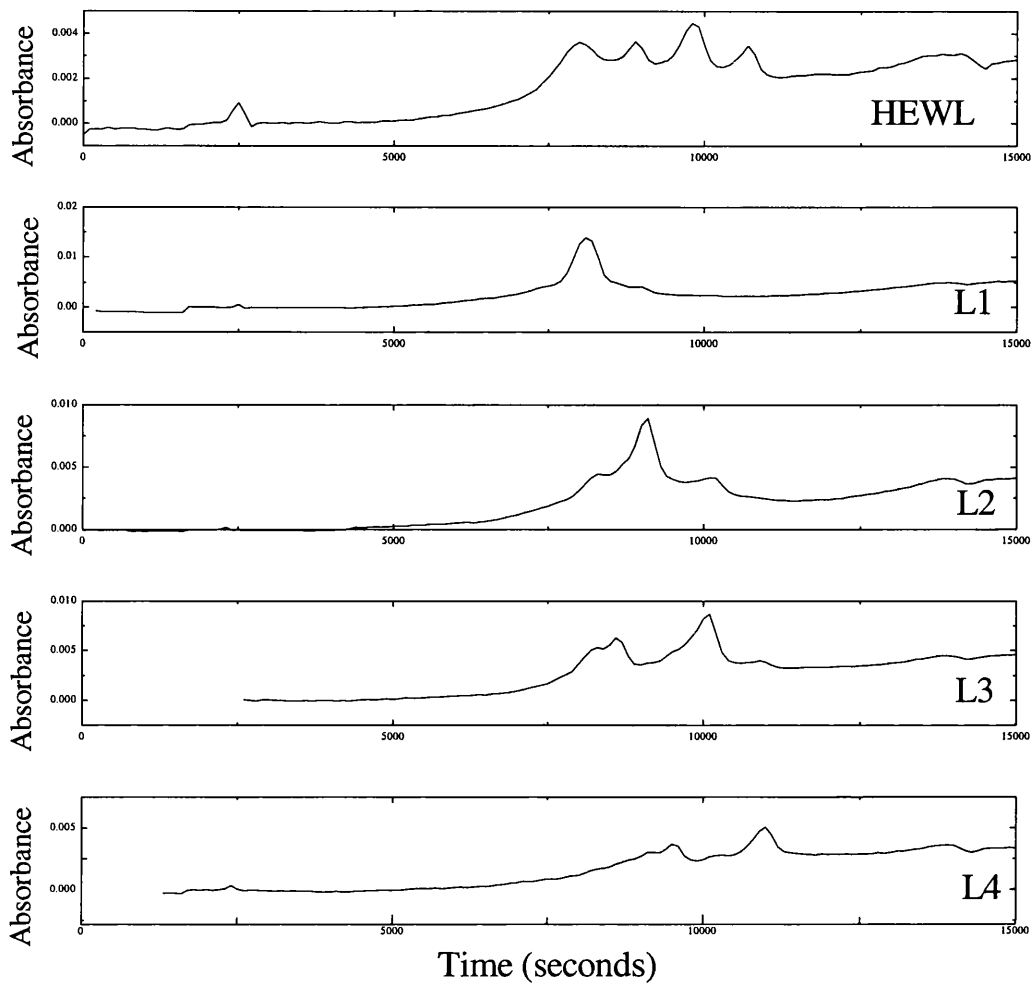


Figure 4.19: Hydrophobic interaction chromatography of the separated peaks after a week at 4°C. Again the column was loaded with 1.5M ammonium sulphate in 0.1M pH 7 phosphate buffer and eluted using the same phosphate buffer. (a) Separation of heat treated lysozyme into 4 separate peaks which relate to those seen in the original DSC experiment. (b) L1, (c) L2, (d) L3 and (e) L4 run down the HIC column under the same conditions at which separation occured.

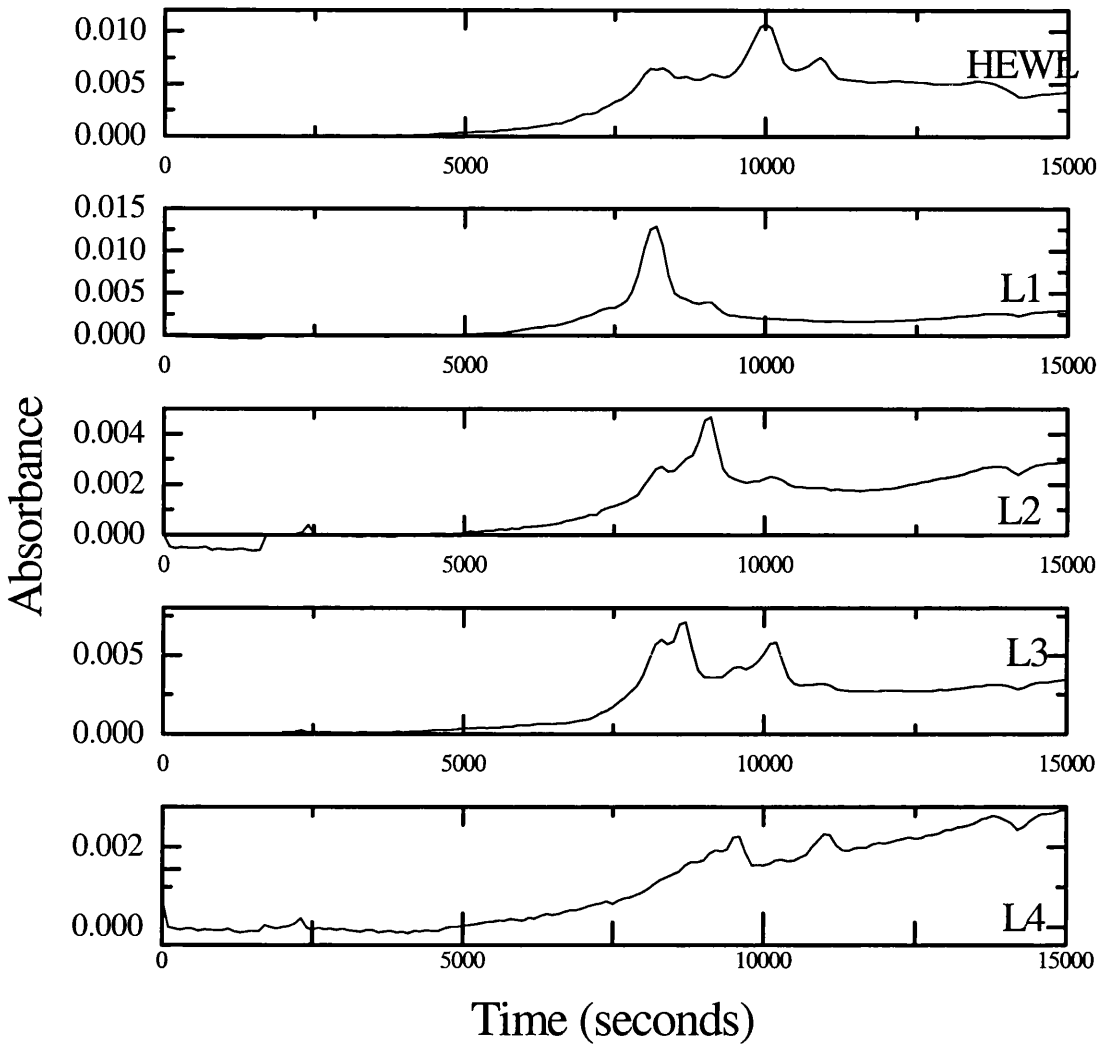


Figure 4.20: Hydrophobic interaction chromatography of the separated peaks after a fortnight at 4°C. Again the column was loaded with 1.5M ammonium sulphate in 0.1M pH 7 phosphate buffer and eluted using the same phosphate buffer. (a) Separation of heat treated lysozyme into 4 separate peaks which relate to those seen in the original DSC experiment. (b) L1, (c) L2, (d) L3 and (e) L4 run down the HIC column under the same conditions at which separation occured.

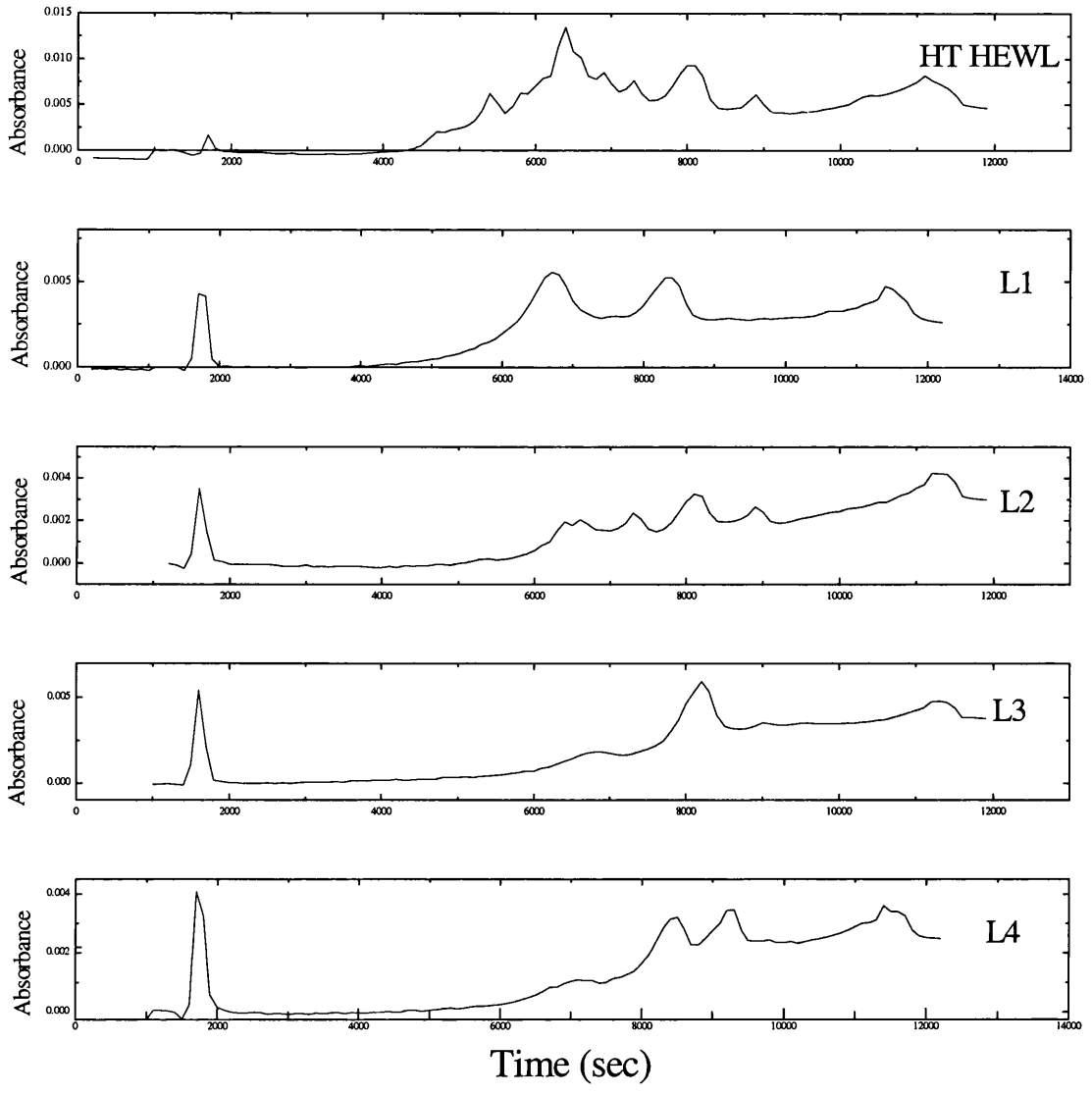


Figure 4.21: Hydrophobic interaction chromatography of the separated peaks after heat treatment. Separation of HT HEWL, HEWL, L1, L2, L3 and L4, into 4 separate peaks which relate to those seen in the original DSC experiment.

Experiments have been performed where the cooling period between each heat treatment was varied to see what affect this has on the area of the peaks produced when the resulting protein is eluted off of the column. Varying the cooling period may effect the amount of misfolding. The cooling period involved insertion of the test tube containing the protein, in the ice bucket until it felt cold to touch (approximately 1 minute) to 30 minutes, as in the original experiment. This appears to show that the maximum misfolding occurs after 10 minutes cooling in the case of L1, 3 and 4, and after 5 minutes cooling for L2.

In Summary: Separation of the HT HEWL into 4 separate species (L1- 4) was possible using an HIC column. The second of these four peaks has the same elution profile as HEWL, suggesting that pk2 (L2) is the native HEWL conformer. From this it can be concluded that there is a small conformational difference which leads to a difference in hydrophobicity between the four species.

The separate peaks appear to undergo an interconversion process if left for a period of weeks or reheated, indicating that the process that creates them is slow and reversible, like proline isomerisation.

In this section the separation of the heat treated lysozyme has been achieved, also the effects of time and heat treatment on these separate species has been studied. In the next section characterization of the separate species is discussed in attempt to identify any differences between them and native lysozyme.

Section C : Properties of Isolated Misfolded Forms

The original DSC observation that led to this work had shown that repeated heat and cool cycles had resulted in a thermogram with four peaks. To make certain that L1-4 , from the separation of the HT HEWL were the same species, it was important to look at the T_ms (midpoint unfolding temperature) of the separate peaks as this would be the only indication that the peaks actually related to what had been seen in the original experiment.

Because the concentrations of L1-4 were always low the new VP-DSC had to be used to look at these separated peaks. The samples were concentrated, desalted and dialyzed into pH 3.4 citrate buffer and their pH checked to make sure that they were in the same pH environment. Each sample was taken and was scanned separately in the DSC from 20 to 100 °C at 60°C/hr. All of the peaks had a discrete T_m, see fig. 4.22 and table 1, which related to those of the original DSC thermogram. Also fig. 4.23 shows the unfolding model fits used to derive the Tm`s of each of the separate peaks. These fits appear to show that there is a very small amount of some other species present in L1 and L3, this may be due to some non-specific aggregate but is not considered any further in this thesis. Notably L2 had the same T_m as native HEWL. This is in agreement with the results from the chromatography experiment, where L2 had the same elution profile as native HEWL. The other three peaks all have a lower T_m than native HEWL.

The results of this experiment were reproducible within the error caused by differences in pH. The pH of the peaks in each experiment range approximately from 3.33 to 3.5, in this region the T_m of unfolding is very sensitive to pH, for every 0.1 pH unit there is an error in T_m of about 1-2 degrees.

SDS-PAGE gels, fig. 4.24, indicated that the four separated peaks, which are identified as L1-4, have a similar molecular weight to HEWL and therefore do not appear to be chemically modified or cleaved forms of HEWL. Reducing and denaturing gels (see fig 4.36) indicate the presence of peptides in the HT HEWL samples, but these are removed during the HIC chromatography.

Sample	Lysozyme	L1	L2	L3	L4
Tm (°C)	71.37 (pH 3.48)	58.79 (pH 3.38)	69.99 (pH 3.33)	57.08 (pH 3.36)	67.08 (pH 3.38)
Tm (°C)	74.89	61.59	72.51	59.16	66.98

Table 1 : Tm of the separated peaks. The separated peaks and native lysozyme from the same experiments were taken, desalted and then extensively dialysed into pH 3.4. 0.1M citrate buffer, the pH was checked and each sample was scanned in the VP-DSC from 20 - 100 °C. The Tm values vary within a range of +/- 3°C, because the Tm of lysozyme is very sensitive to pH in this range.

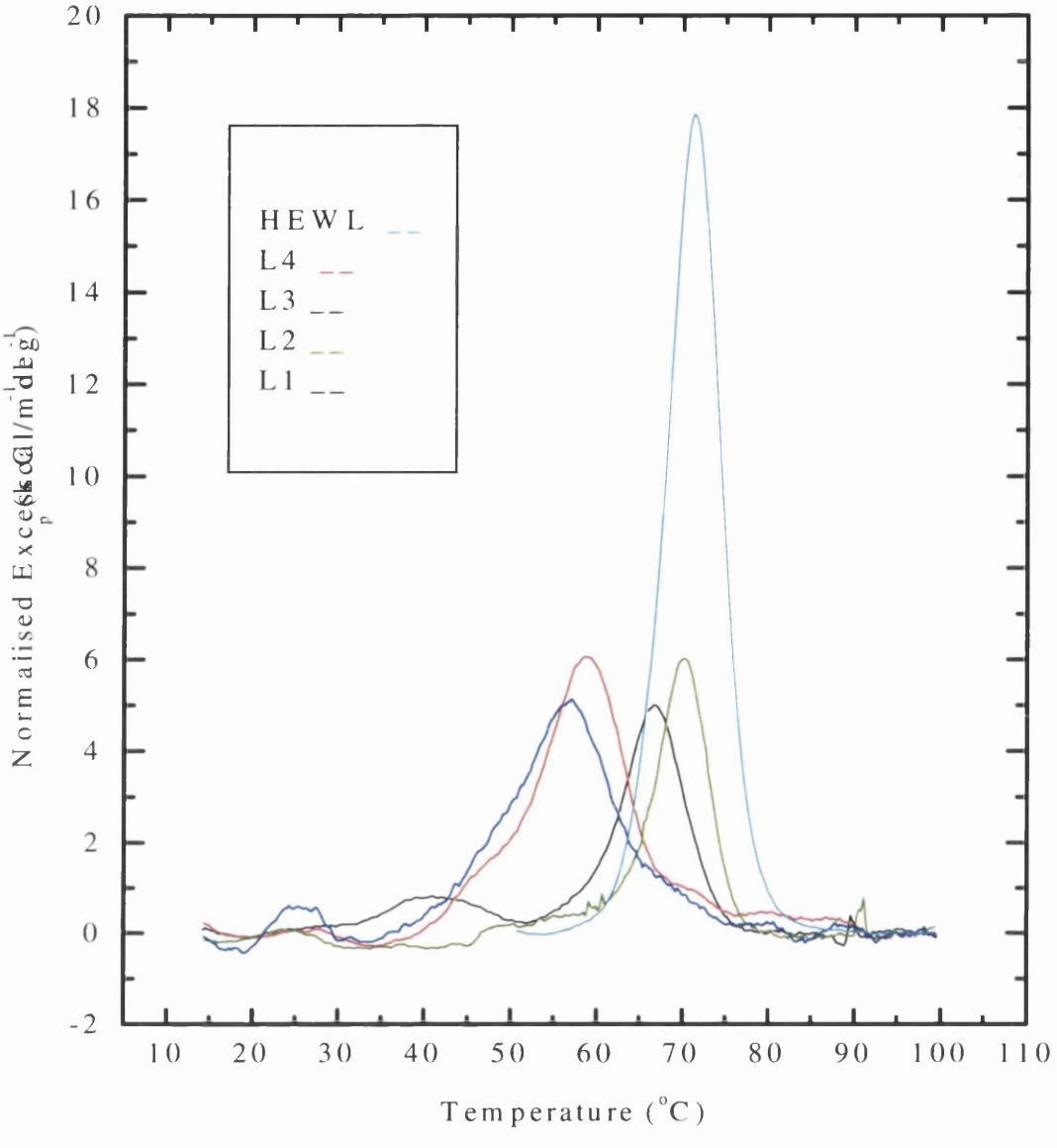


Figure 4.22: Differential scanning calorimeter scan of the separated peaks, after extensive dialysis into pH 3.4 citrate buffer.The thermogram shows the unfolding transtions of L1 - 4. Native HEWL and L2 have the same T_m. L1, 3 and 4 have a T_m lower than that of native HEWL. Each separate scan was fit to a two or multi -state unfolding model, see the following figure.

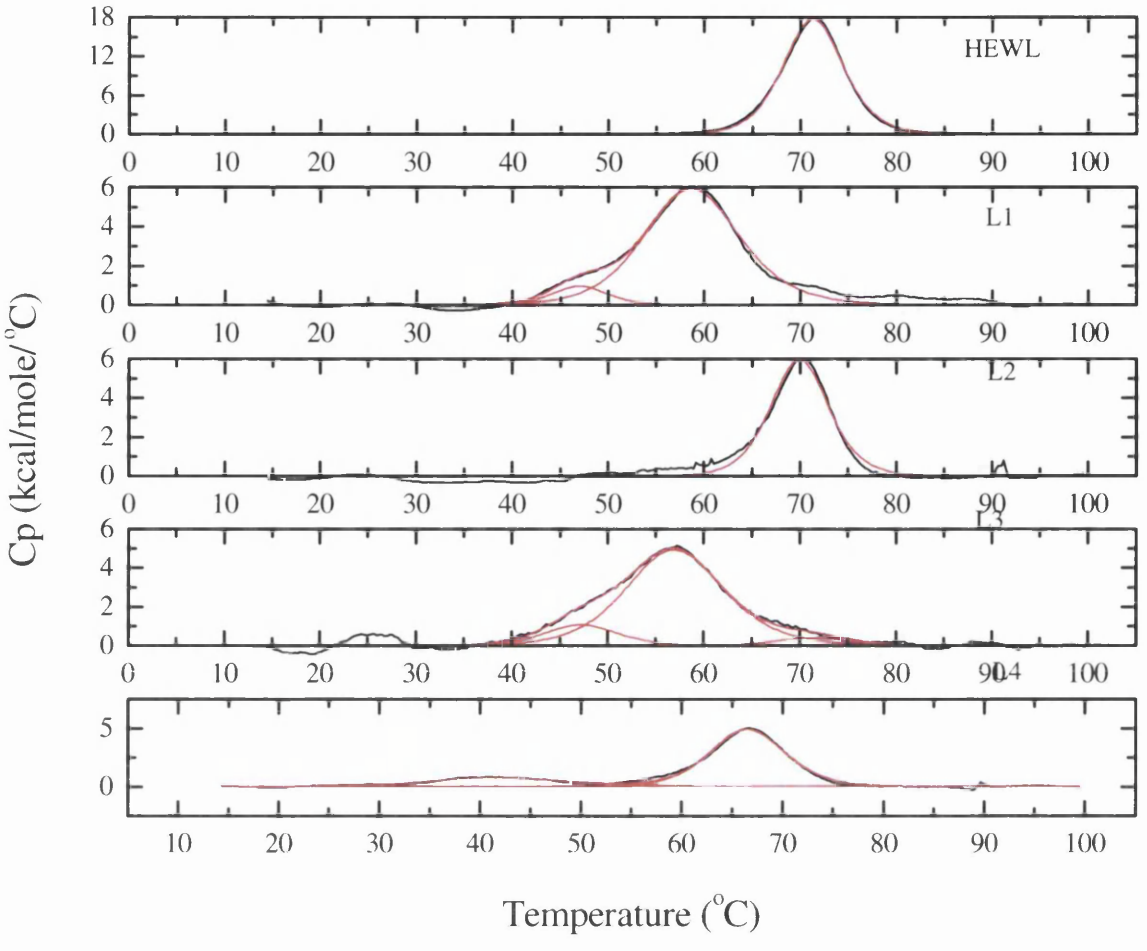


Figure 4.23 : The separate DSC scans with the appropriate unfolding model fit.

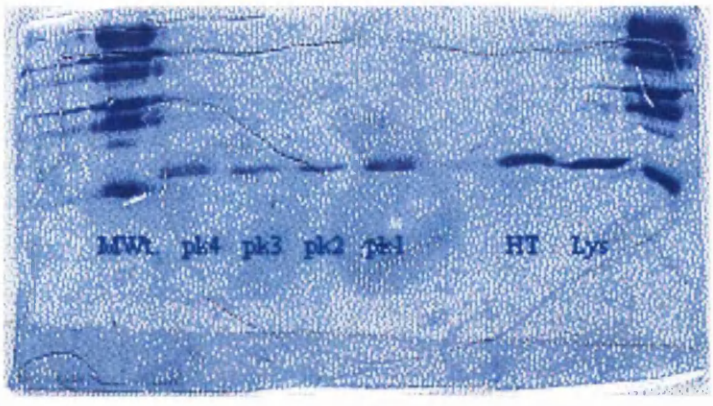


Figure 4.24 : SDS PAGE native gel of separated peaks, native and heat treated lysozyme. MWT, molecular weight markers, pk1 L1, pk2 L2, pk3 L3, pk4 L4, HT heat treated lysozyme and Lys, native lysozyme. All the bands run to the same position.

Electrospray mass spectrometry shows that all these species have the same molecular weight as native HEWL. Extra peaks were also present in the spectra. A small population of these species have a molecular weight of +/- 18, i.e. there is loss of a water molecule (figure 4.25 (d) peak at m/z =1299.8 & 1298.1, where z = 11^+). The loss of a water molecule is possibly due to reactions between, for example a lysine and glutamate residue producing a covalent bond (Personal Communication : Alex Last). Other peaks may be due to non-covalent addition of a Na^+ ion in solution (figure 4.25 (d) peak at m/z = 1302.9). The smaller peaks seen in figure 4.25 (d) m/z = 1306.3, 1308.7, 1310.3 correspond to the masses of [HEWL - $2H_2O$ + phosphate], [HEWL - $1H_2O$ + phosphate] and [HEWL + 1 phosphate] respectively[#]. The phosphate ions bind non-covalently. The even smaller peaks at 1315.6 and 1317.3 are due to the same non-covalent reaction but 2 phosphates are bound in this case. The peak at m/z = 1336.7 corresponds to human lysozyme, which was used as an internal standard for the mass spectrometry experiments performed at Oxford University. The two peaks 1345.6 and 1354.4 are due to the interaction between human lysozyme and one and two phosphates respectively.

Similar peaks are seen in each of the other spectra, and have the same origins as explained for L3.

Isoelectric focussing gels indicated, data not shown, that the 4 separated peaks, HT and native HEWL all run to the same pI. Again if there was a difference in overall charge caused by conversion of an amide group to a carboxylate group (deamidation) it would be seen by using IEF (Ahern & Klibanov 1985). The gels used in this experiment had a pI range of 3 - 9 and HEWL has a pI of approximately 11 (Worthington Biochem. 2000), so this indicates that none of the separated peaks have a pI which deviates significantly from that of native or the HT HEWL. There is no change in the charge carried by the species.

The fluorescence emission spectra of each of the separate species, HT HEWL and HEWL is compared in figure 4.26. Overall the peak maxima of each of the species is the same as HEWL, (340nm +/- 0.3%) to within 0.3%. So there is no significant difference in the fluorescence maximum of the species. There does, however, appear to be a difference in the

[#] This result was calculated as follows ; the m/z of HEWL was 1301.4, this was multiplied by the charge state, 11^+, giving the mass of 14345.1. From this was subtracted the molecular weight of a water molecule (18) and the molecular weight of one phosphate group (75) added on, giving 14390.4, this was then divided by 11^+ giving 1308.2.

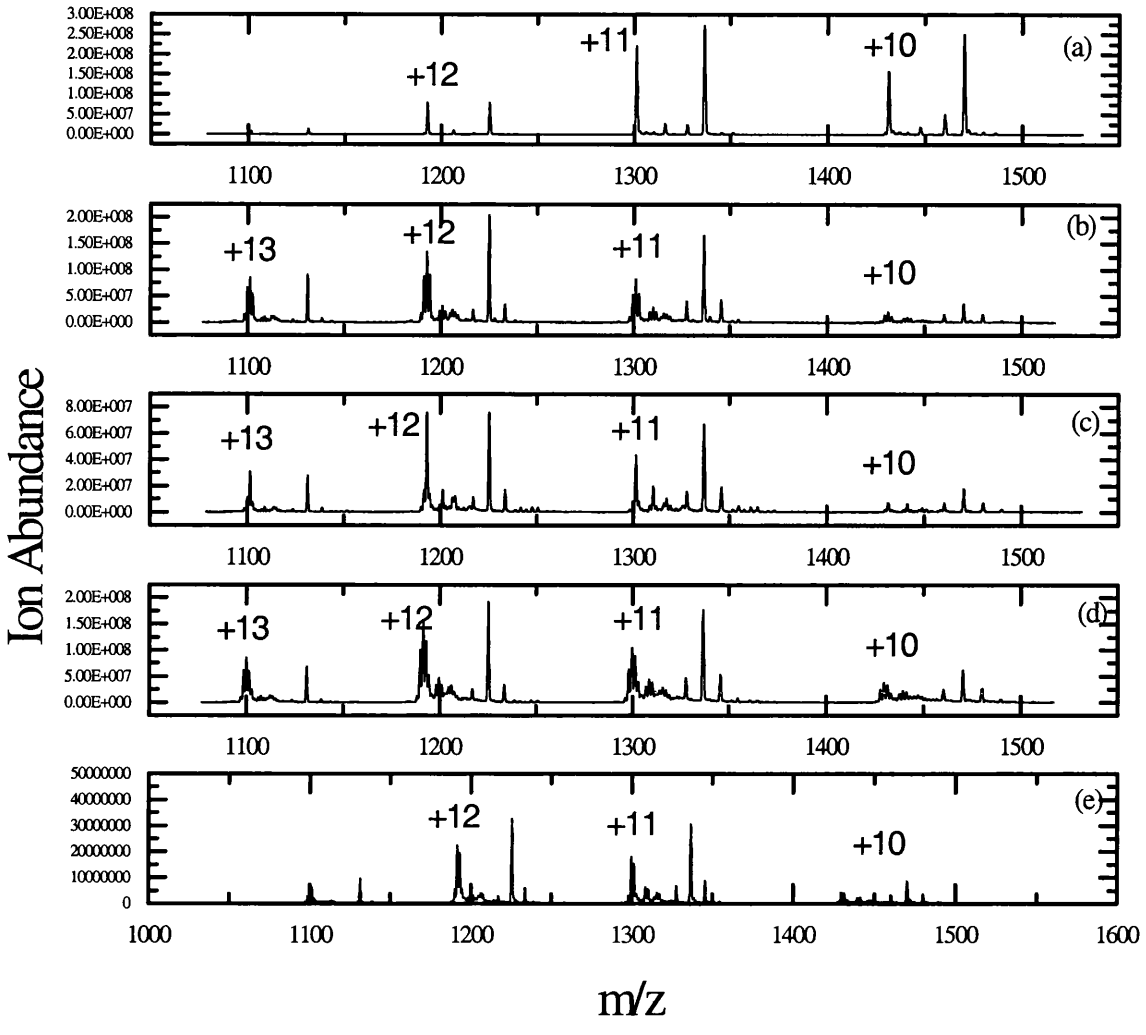


Figure 4.25 : Electrospray mass spectrum of HEWL (a), L1 (b), L2 (c), L3 (d), L4 (e). Each spectrum shows four sets of charge states, human lysozyme shows as peaks at an approximate m/z of 1132.1, 1225.4, 1336.6, 1470.2. HEWL shows at m/z of approximately 1101, 1191.6, 1301, 1430. The m/z of 1301.4 was followed during the hydrogen exchange experiments. The peaks measured in this experiment correspond to a charged state of +11.

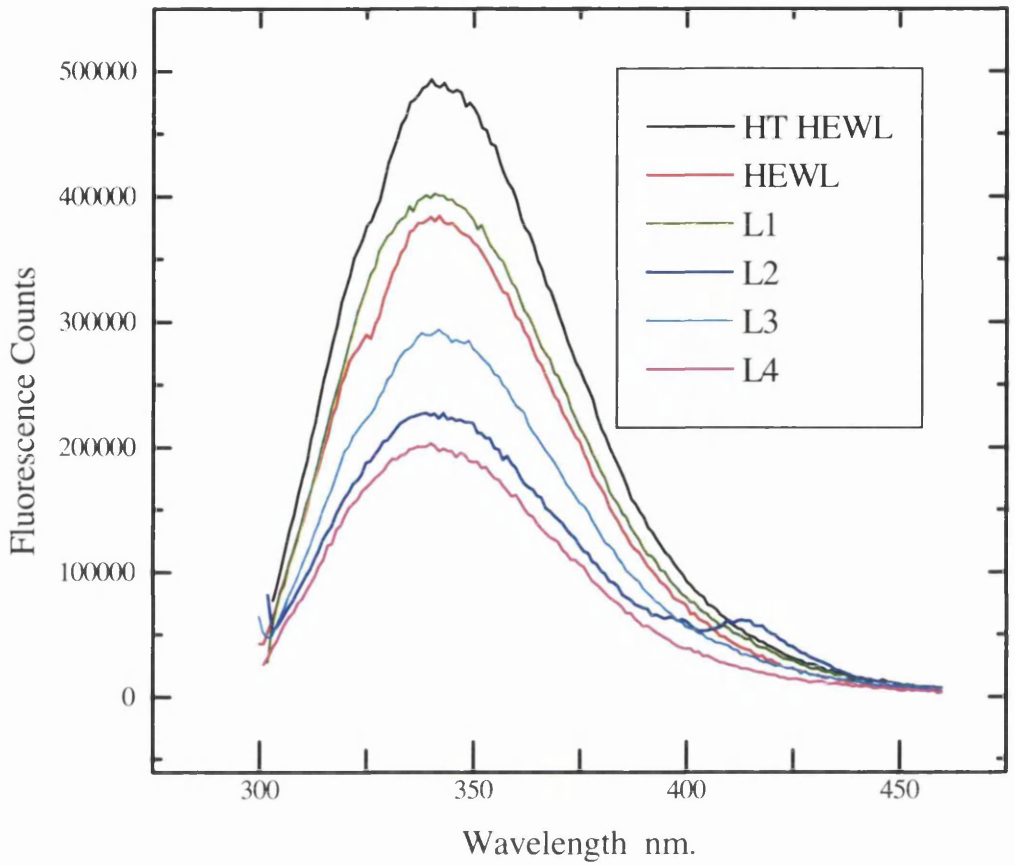


Figure 4.26 : The fluorescence emission spectra of HT HEWL, HEWL, L1, L2, L3 and L4, excitation at 290nm.

intensity of each of the species. The difference in intensity between HT HEWL and HEWL was discussed in the first section. L3, L2 and L4 all have a lower intensity than HEWL. The intensity of L3 is approximately 25% lower than HEWL, L2 41% lower and L4 48% lower. This decrease in intensity is likely to be due to a difference in the tryptophan exposure of L2 - 4. These species are likely to have one or more of their tryptophan residues less exposed than native lysozyme. This could be due to them being masked by some other amino acid side chain, hidden by some misfolded piece of structure or due to quenching by another aromatic amino acid side chain. Also there appears to be a slight red shift in the case of L3 and L2, this would concur with the conclusion that the some of the tryptophan residues are more exposed. This result reflects what is seen when hydrophobic interaction chromatography is used to separate HT HEWL into L1 - 4. L3 and L4 have a greater affinity for this column than native HEWL, which indicates they have more hydrophobic areas exposed or that they are more flexible.

The peak in the fluorescence spectrum of L1 has a greater intensity than native HEWL, suggesting that it has one or more of its tryptophan side chains more exposed than in native HEWL. Again this may be an effect of misfolding. This is mirrored in the results from the HIC separation of HT HEWL, L1 has an earlier retention time than HEWL.

The species L2, which has shown a lot of similarities to HEWL has an intensity that is 41% lower than that of HEWL, suggesting that it has tryptophan residues in a different environment from HEWL. This result does not agree with the HIC result, unlike L1, L3 and L4, as L2 has the same retention time and conditions as HEWL. Although it should be noted that, experimentally, L2 does not always have the same characteristics as HEWL.

Succinimide quenching experiments were performed using L1-4, and an example of the resulting graph is shown in figure 4.27. The data from the experiment shown and a replicate experiment was analysed using the data analysis tool in the Excel computer program. This 2-way analysis of variance showed that the method of analysis had no significant effect on the result obtained, as P, the probability of obtaining the results by chance, is greater than 0.05 (see table 2) at 0.0923. This therefore indicates that the accessibility of the tryptophan residues of all species to the succinimide quencher is similar. It is entirely possible that the succinimide molecule is too large a probe for use in this case and hence has resulted in all species having a similar Ksv.

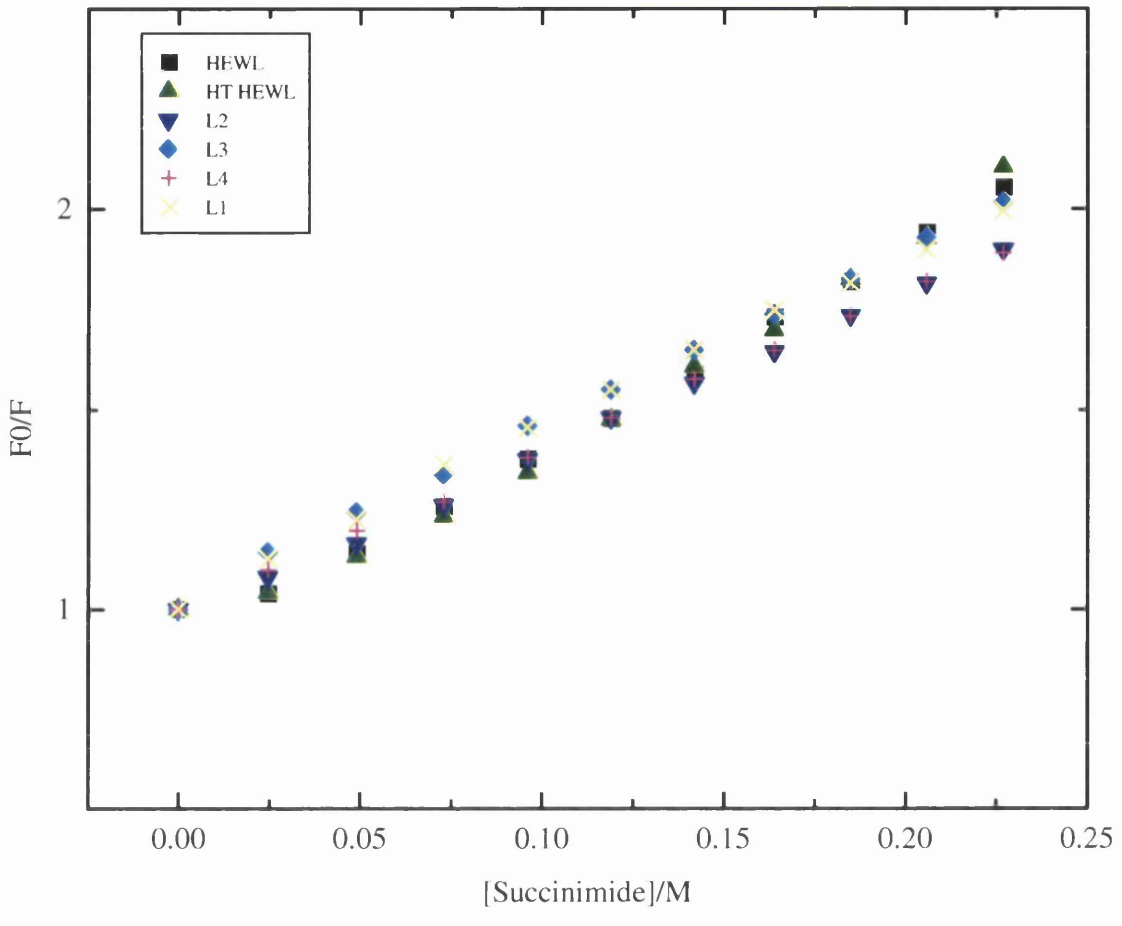


Figure 4.27 : Succinimide quenching of the separated peaks. The Ksv (Stern Volmer constant) of all the separate species are shown in table 2. The separate species have a lower Stern Volmer constant than native and HT HEWL.

	Mean Ksv	Standard Deviation	Variance
HEWL	5.2385	0.4120	0.340
HT HEWL	4.9255	$9x10^{-4}$	$1.62x10^{-6}$
L1	3.8331	0..4995	0.4990
L2	4.1198	0.0779	0.012
L3	4.1217	0.2698	0.1455
L4	4.2028	0.2339	0.1095

Table 2 : Stern Volmer constants from succinimide quenching of HEWL, HT HEWL and the separate peaks. The results were analysed using Excel software. Variance was calculated using the following equation : $(dp)x(SD)^2$, where dp is the number of data points and SD the standard deviation. The P-value is 0.09227. The P-value is the probability of obtaining the results by chance and will be ≤ 0.05 if the results are obtained by chance.

The enzyme activity of the separate species (L1 - 4) was measured using a HEWL enzyme assay, see Material and Methods, as this would give a useful indication of how misfolding has affected the HEWL. The specific activity of these species, HT HEWL and native HEWL are tabulated in table 3. The results show that the separated peaks have a much lower specific activity (between 40 - 50 %) than that of native and HT HEWL; even in the case of L2 which appears to have some behaviour similar to that of native HEWL, the specific activity is approximately half that of native HEWL. This suggests that some sort of internal change has occurred during heat treatment which has somehow affected the binding site, obscuring it from the substrate or altering it and preventing it from functioning properly.

It is interesting to note that there appears to be a disparity between the results of the enzymatic activity assay and the results of other experiments. The mixture of the misfolded forms has a greater activity than any of the separate samples. The separate peaks appear to be the same as HEWL by CD, mass measurements, fluorescence and isoelectric point, but there are differences in their thermal and chemical stability. Unfortunately the results of these experiments can give no explanation for the disparity in enzyme activity between HT HEWL and the separate peaks.

Circular dichroism of each of the separate species shows that the secondary structure is very similar to native HEWL, (see fig. 4.28) although there appears to be a difference in the height of the peak, which is possibly due to difference in concentration. The tertiary structure of each of the separated peaks is shown in the near UV CD spectrum in figure 4.29, and this shows that in most cases the tertiary structure is similar. The spectra of HT HEWL and native HEWL show there is a very small change which is indicated by the fact the two lines crossover. There is also a difference in height of the aromatic amino acid absorption wavelength area (250 - 300nm), which could be caused by a slight difference in the packing of the aromatic amino acids in the tertiary structure of the HT HEWL. It is possible that the differences between these spectra are caused by differences in the batches of HT HEWL, as this was not seen in earlier CD spectra (fig. 4.12),which may have a different amounts of cleaved proteins (peptides) as indicated in reducing/denaturing gels (fig. 4.36), or because of unfolded/ drastically misfolded protein seen on the chromatograms at the end of the run (0% ammonium sulphate concentration). The difference in the spectra my also be due to there being different amounts of isomerisation present in the batches due to variations in heating time.

	Native Lysozyme	Heat Treated Lysozyme	L1	L2	L3	L4
Activity mol PNP min^{-1} mol $HEWL^{-1}$	1.02 +/- 0.01	1.02 +/- 0.025	0.52 +/- 0.015	0.64 +/- 0.05	0.68 +/- 0.05	0.65 +/- 0.005
Activity mol PNP min^{-1} mol $HEWL^{-1}$	1.00 +/- 0.01	1.07 +/- 0.025	0.55 +/- 0.015	0.65 +/- 0.05	0.67 +/- 0.05	0.65 +/- 0.005

Table 3 : Activities of native and heat treated lysozyme and the separated peaks as calculated using the p-nitrophenol penta n-acetyl β chitopentaoside assay as described in materials and methods. (PNP para-nitrophenol)

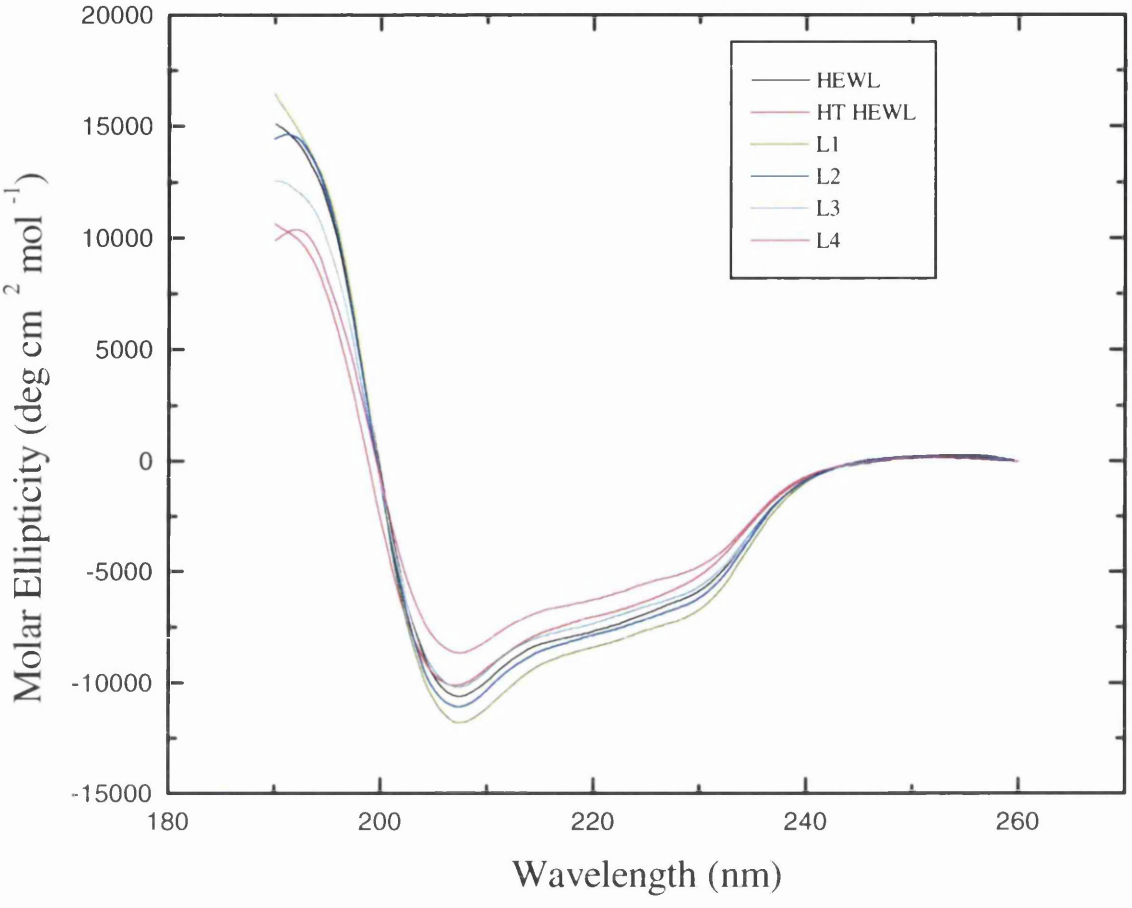


Figure 4.28: Far UV CD spectra of heat treated lysozyme, native lysozyme, L1, L2, L3 and L4 . This method gives an measure of the secondary structure of the protein, in this case the secondary structure of all species is similar.

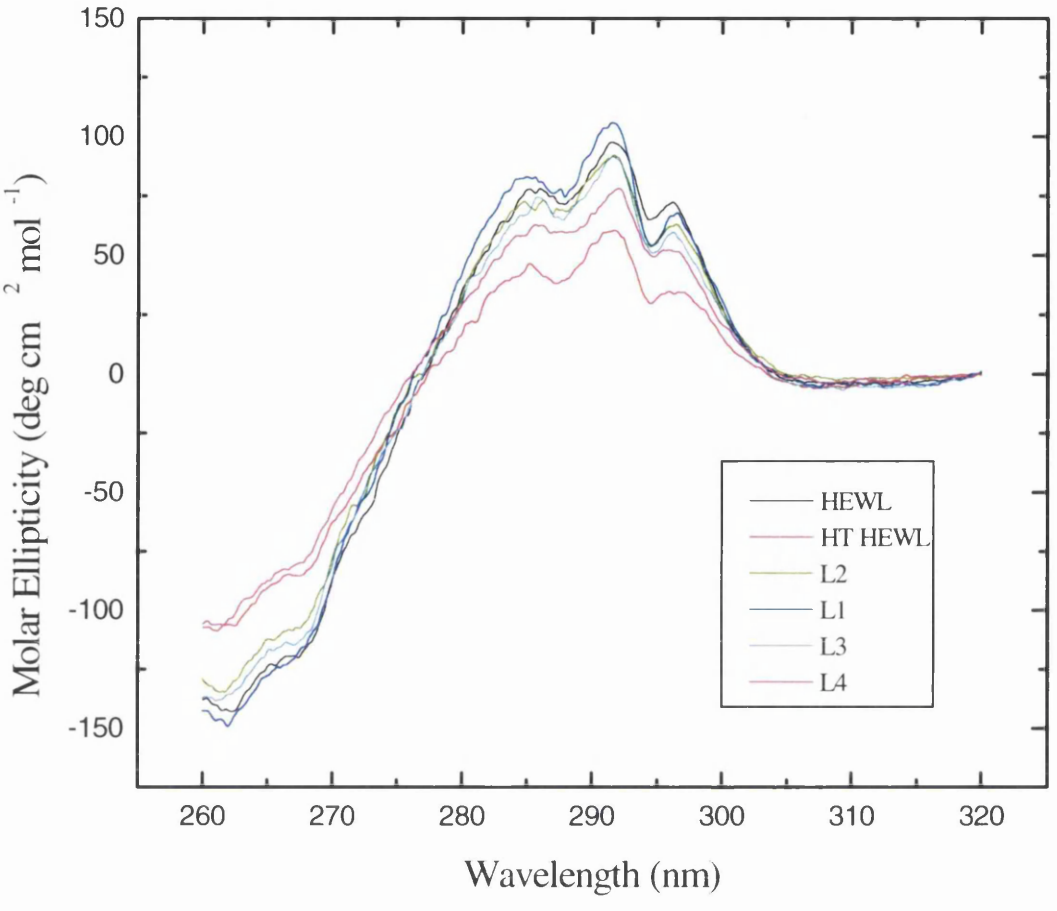


Figure 4.29: Near UV CD spectra of heat treated lysozyme, native lysozyme, L1, L2, L3 and L4. This method gives an measure of the tertiary structure of the protein. In this case the tertiary structure of all species is similar, except that the heat treated lysozyme which crosses the native lysozyme spectra suggesting a slight difference in structure.

In general the structure of the separate peaks appear to be similar to that of native HEWL. Therefore the heat treatment has not altered the structure of the protein significantly. In the case of HEWL the proline residues are situated in a loop region so it is possible that a change in their configuration would therefore not effect the folded conformation of the rest of the protein significantly.

The amount of secondary structure present in the separate species was estimated using the CONTIN (Provencher & Glockner 1981) procedure, the results are tabulated in table 4. The HT HEWL appears to contain slightly more β-sheet than the native HEWL, but this is possibly due to the presence of aggregated material which is seen as the material washed off of the HIC column at 0% ammonium sulphate concentration. Conformer L1 is estimated to have slightly more α-helix present, L2 has a similar secondary structure content to native HEWL, which is consistent with some of the other evidence that suggests it is native HEWL. The secondary structure content of L3 seems to be less than in native HEWL and L4 has been calculated to have less α-helical structure and more β-sheet. It is very important to note that these results are uncertain due to the nature of the CONTIN procedure, which relies on the CD spectra of 16 reference proteins of known structure. This has limited the accuracy of this procedure for estimating secondary structure as it does not take into account any type of structure not present in the 16 reference proteins, and also does not allow for disulphide bonds.

The unfolding of L1-4 caused by increasing guanidinium hydrochloride concentration and observed by measuring the change in CD in the secondary structure region, was carried out to study the cooperativity of unfolding of each of the separate species. The results of these experiments are shown in figure 4.30. For comparison the unfolding curves of native HEWL and HT HEWL have been included. Compared to native HEWL the separate peaks again do not show a 2 state unfolding profile (see Section A), this is indicative of an uncooperative unfolding transition. Each of the separate species, including L2, have a lower stability towards GdnHCl when compared to native HEWL.

These experiments give an indication of the proteins cooperativity, which relates to interactions within the protein molecule which help keep it in the folded conformation and that help drive the folding process to completion. This decrease in chemical stability is mirrored by a decrease in the thermal stability, and suggests that the misfolding has lead to disruption of some of the cooperative interactions within the protein.

	α-helix	β-sheet	Remainder
Native Lysozyme	24%	28%	48%
Heat Treated Lysozyme	24%	34%	42%
L1	30%	26%	45%
L2	25%	27%	48%
L3	22%	27%	51%
L4	18%	33%	49%

Table 4 : Results of secondary structure estimation of separate conformers L1- 4, heat treated and native lysozyme, calculated using the CONTIN procedure of Provencher & Glockner

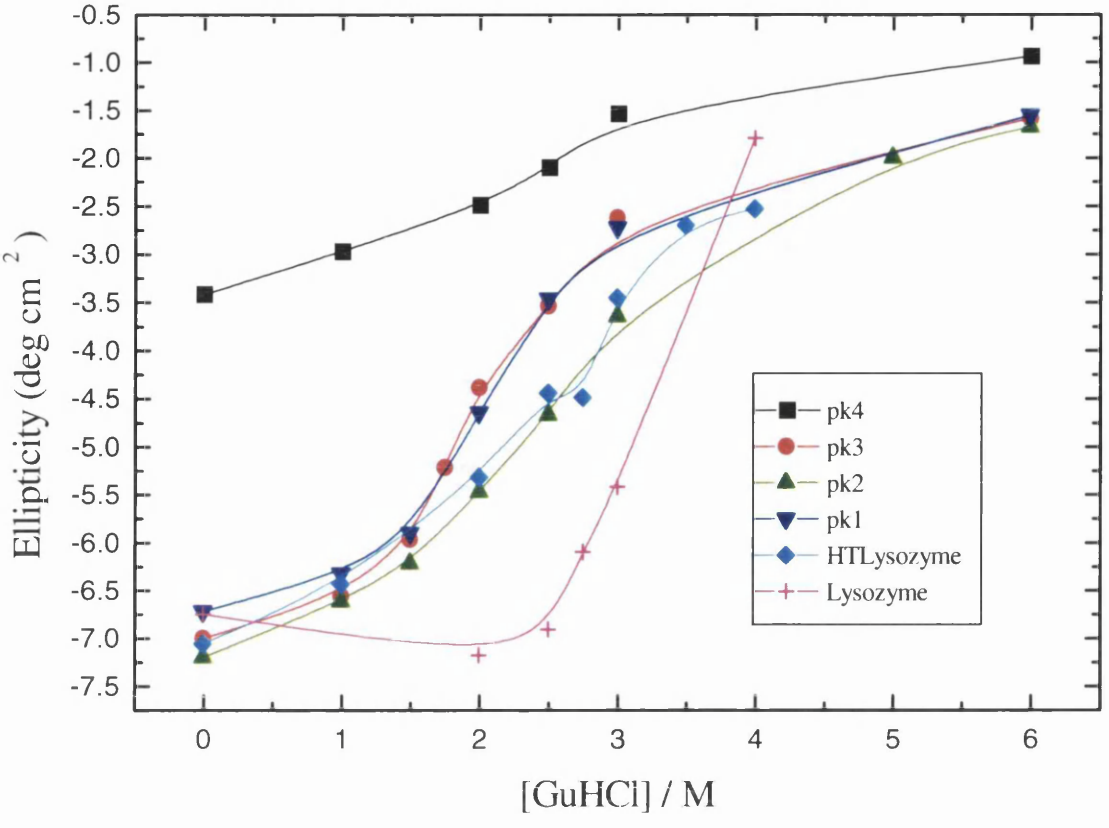


Figure 4.30 : Guanidinium hydrochloride induced unfolding of the separate peaks, heat treated and native lysozyme monitored by changes in far UV CD spectra. ◆ native lysozyme, + heat treated lysozyme, ■ L1, **O** L2, **V** L3 , Δ L4. All species have a lower stability towards GuHCl than native lysozyme. Data points were fitted to a linear plot.

The affinity of ANS towards each of the misfolded species was compared to that of native HEWL and HT HEWL. ANS is a hydrophobic probe and binds to exposed hydrophobic sites on proteins. It was this experiment that originally identified the possibility of using HIC to separate HT HEWL into the four separate species. The results from this experiment are shown in figure 4.31. ANS as previously, (see Section A) has a great affinity towards HT HEWL, and none or very little towards native HEWL. It also appears to bind to the separate peaks but with less affinity than that it has towards HT HEWL. This was unexpected as the misfolded species had been separated using HIC because HT HEWL had shown a great affinity for ANS, and it had been assumed that the separate species would show a similar affinity for ANS.

To explore the differences between the heat treated lysozyme and the separate peaks further, the separate peaks were mixed together in the proportions that they would be expected to be in, in the HT HEWL(L1 - 25%, L2 - 12.41%, L3 - 22.3%, L4 - 8.8%) and also in equal proportions (25% of each). This would help exclude the possibility that the 4 species were somehow "aggregating" giving a species with exposed hydrophobic sites. ANS was added to these mixtures to see if these would show the same result as when ANS binds to HT HEWL, see figure 4.32. Again ANS didn't have as great an affinity towards these mixture as to HT HEWL. The large increase in fluorescence, and the resulting blue shift that is seen when ANS binds to heat treated lysozyme, must have been due to some other species that is present in the HT HEWL and not in the mixtures of the separate peaks. It was concluded that the affinity for ANS that HT HEWL had shown was due to the very misfolded and degraded protein, whose presence is indicated by the increase in baseline, at the end of the chromatography run (0% ammonium sulphate) when separating HEWL, see figure 4.17(a). The protein material which comes off at the end of the HPLC run has a great affinity for the HIC column which is why it is eluted off the column at zero salt concentration. An experiment to look at the ANS binding of this protein was performed to positively identify these misfolded proteins and peptides as the cause of the increased ANS affinity. This showed (fig. 4.33) that ANS binding to this species resulted in a large increase in ANS fluorescence intensity and a blue shift in emission wavelength, which would explain the result of ANS binding to HEWL.

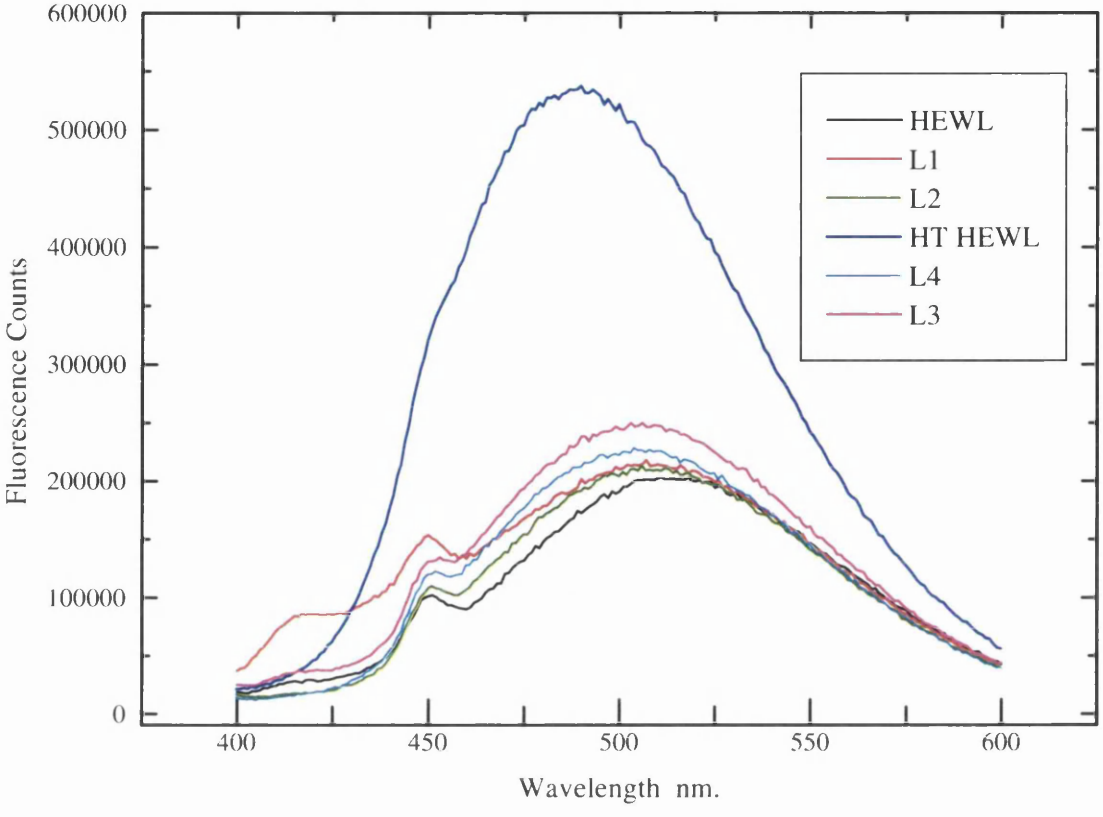


Figure 4.31 : ANS binding to the separate peaks, heat treated and native lysozyme. ANS has a great affinity towards heat treated lysozyme compared to native lysozyme and appears to have some affinity towards the separate peaks when compared to native lysozyme.

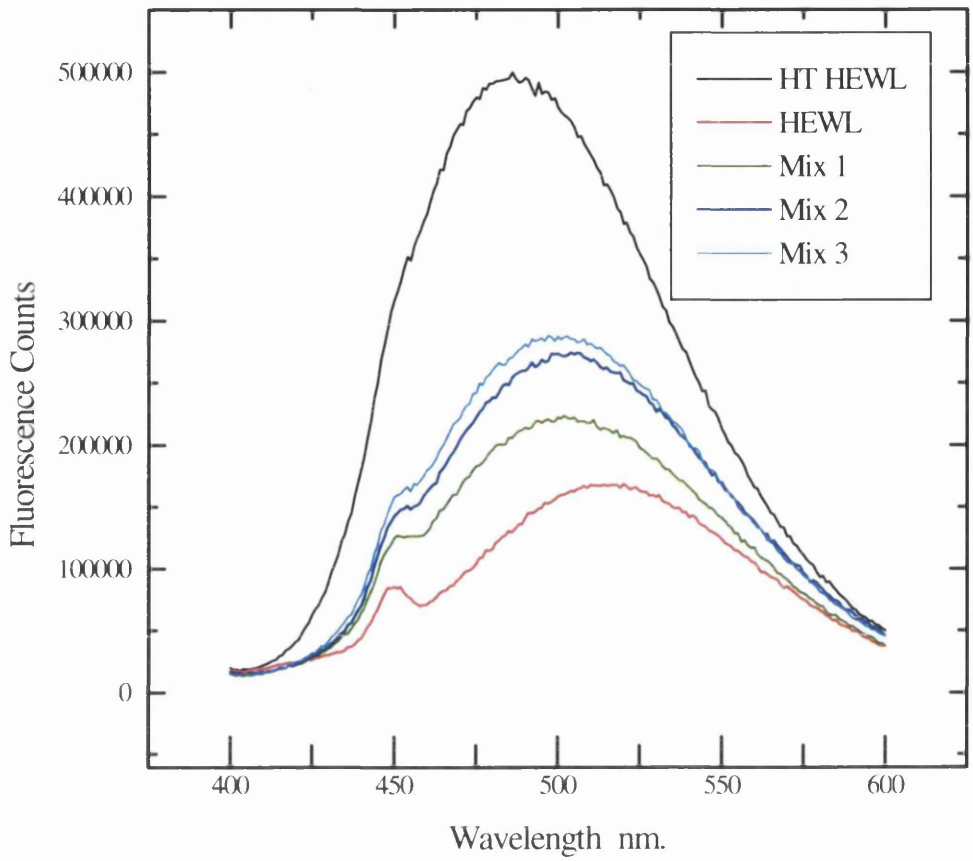


Figure 4.32 : ANS binding to mixtures of the separated peaks, heat treated and native lysozyme. Mix 1 refers to a mix of 36µl L1, 18µl L2, 20µl L3 and 28µl L4, roughly the amounts of each peak in a chromatogram of the heat treated protein, the area of the chromatogram calculated as the area where the baseline goes above 0 AU. Mix 2 refers to 46µl L1, 24µl L2, 54µl L3, and 36µl L4, again roughly the amounts of each peak in a chromatogram of the heat treated protein, the area of the chromatogram calculated as the area where the peaks start. Mix 3 equal amounts of each peak. In all cases the samples were diluted up to 3ml. ANS has a greater affinity for the mixtures than for native lysozyme but ANS has more affinity for the heat treated lysozyme.

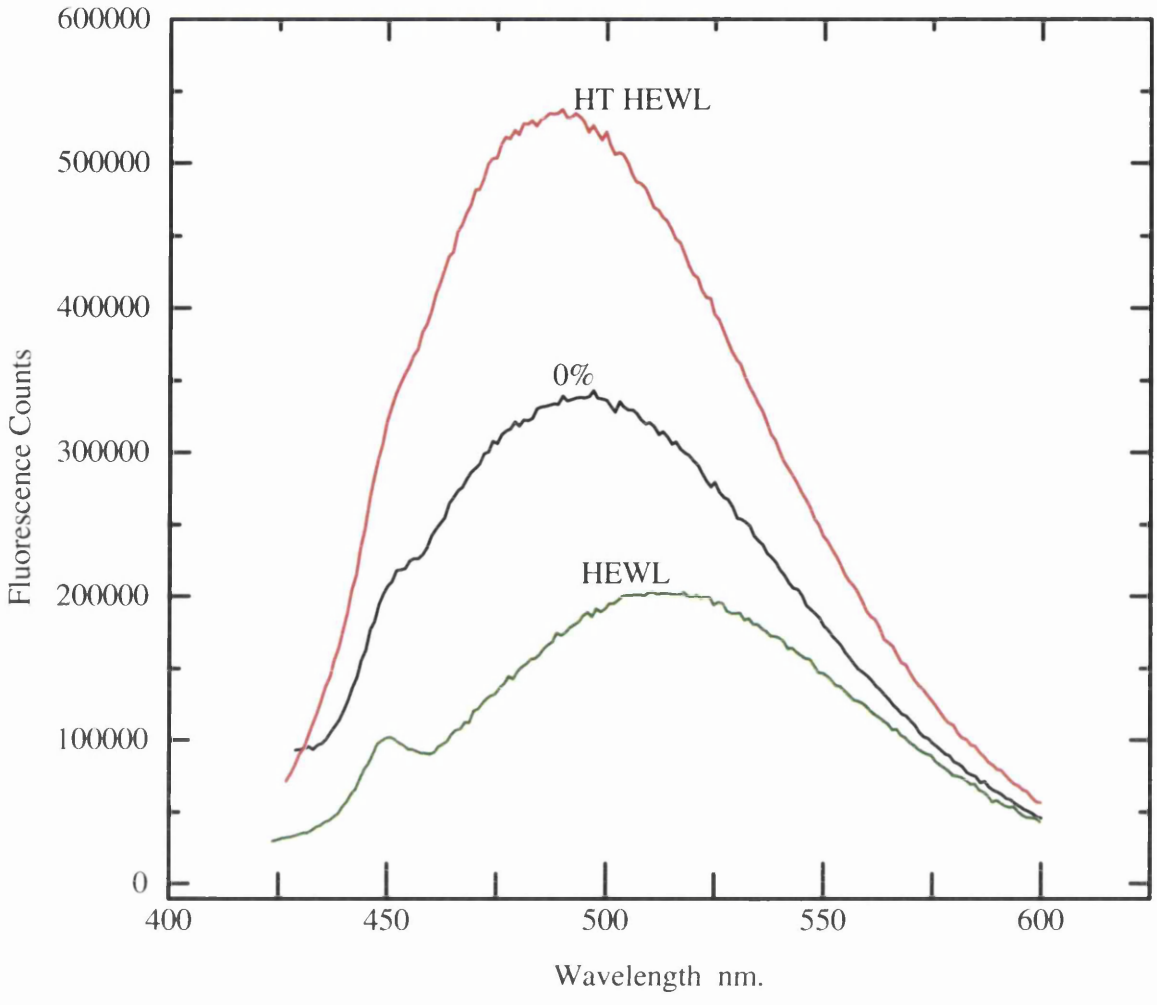


Figure 4.33 : ANS binding to the protein that is eluted off the HIC column at 0 salt concentration. Dashed line - heat treated lysozyme, dotted line - native lysozyme and solid line - the protein that is eluted off the HIC column at 0 salt concentration. The ANS has a greater affinity for this protein than native lysozyme and slightly greater affinity for the heat treated lysozyme.

In a further attempt to discover if the structure of these misfolded species was significantly different from that of native lysozyme, hydrogen exchange experiments were performed using mass spectrometry. Because deuterium exchange occurs between solvent molecules and exposed hydrogens this should provide better information on the structure of these misfolded species, as the deuterium ion is smaller than the other structural probes used previously.

These experiments were performed using peaks L1 - 4 and native HEWL, to study any change in the structure/ exposure of exchangeable hydrogens caused by the misfolding. Unfortunately HT HEWL could not be used as the mass spectrum of this species was too noisy due to the presence of aggregate and peptide, but experiments were possible on the isolated species. The samples were washed to remove any salt ions and dialyzed into water/formic acid at pH 5, then concentrated. The samples were then diluted into D_2O and human lysozyme added as an internal control. Human lysozyme is used as it has a different molecular weight (14,691) from HEWL. HEWL is assumed to have 255 exchange labile protons (Chung et al. 1997) and the human lysozyme 262, i.e. a difference of 7 labile protons. The number of exchangeable residues were calculated by Mr. A. Last, see Materials and Methods (Last & Robinson 1999).The change in molecular weight over exchange time was measured using a mass spectrometer.

The control for this experiment was human lysozyme and the difference in exchanged sites between this and HEWL was consistently 7, which agrees with the expected value. Figure 4.34(a) shows the change in number of protected sites over time, the two curves are similar in shape and the offset in the two curves is the difference in number of protected sites, which is 7 (see figure 4.35). When the results of the hydrogen exchange experiments performed using the separate species L1-4 are compared to the control experiment, L2 (fig. 4.34(c)) and L4 (fig. 4.34(d)) show a similar exchange pattern to HEWL. The difference in number of protected sites is also approximately 7 in both these cases, see figure 4.35. This suggests that these species are very similar to native HEWL, and that there is no difference in the accessibility of exchangeable sites between these species and native HEWL. In the case of L2 this is consistent with other results which indicate that this species is native HEWL, also it is interesting to note that L4 has the next highest T_m when compared to L2 \ native lysozyme.

The exchange pattern of L1 appears to be different from HEWL, see figure 4.34(b), the offset between the graph of L1 and human lysozyme is greater, and the overall difference in number

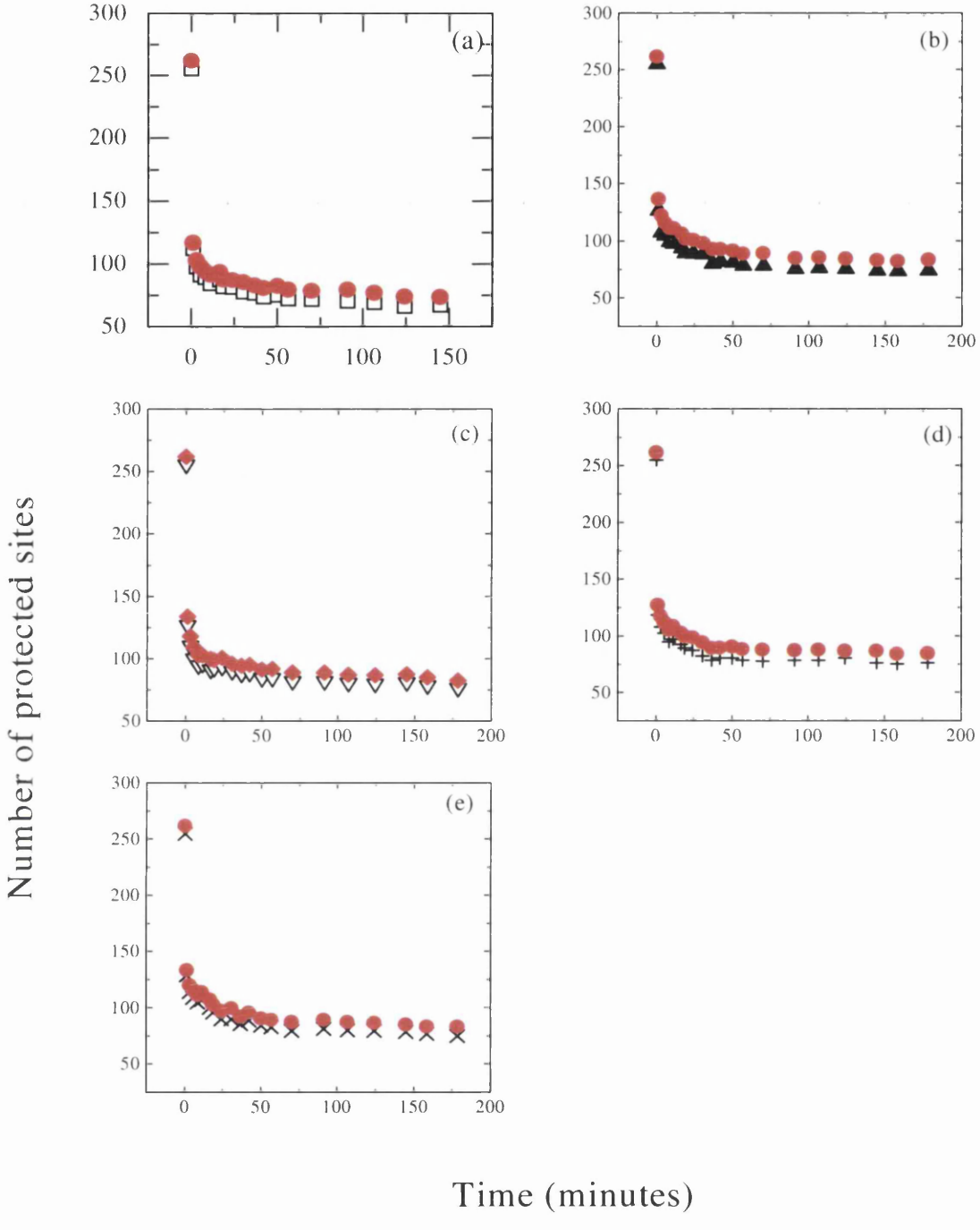


Figure 4.34 : Hydrogen exchange experiments, graphs show the change in number of exchangeable protons/ protected sites over hydrogen exchange time. Human Lysozyme has been used an internal standard in all these experiments and is denoted by the solid circle (●). Native HEWL (□) exchange shown in (a); L1 ▲ (b); L2 (c); L3 + (d) and L4 X (e).

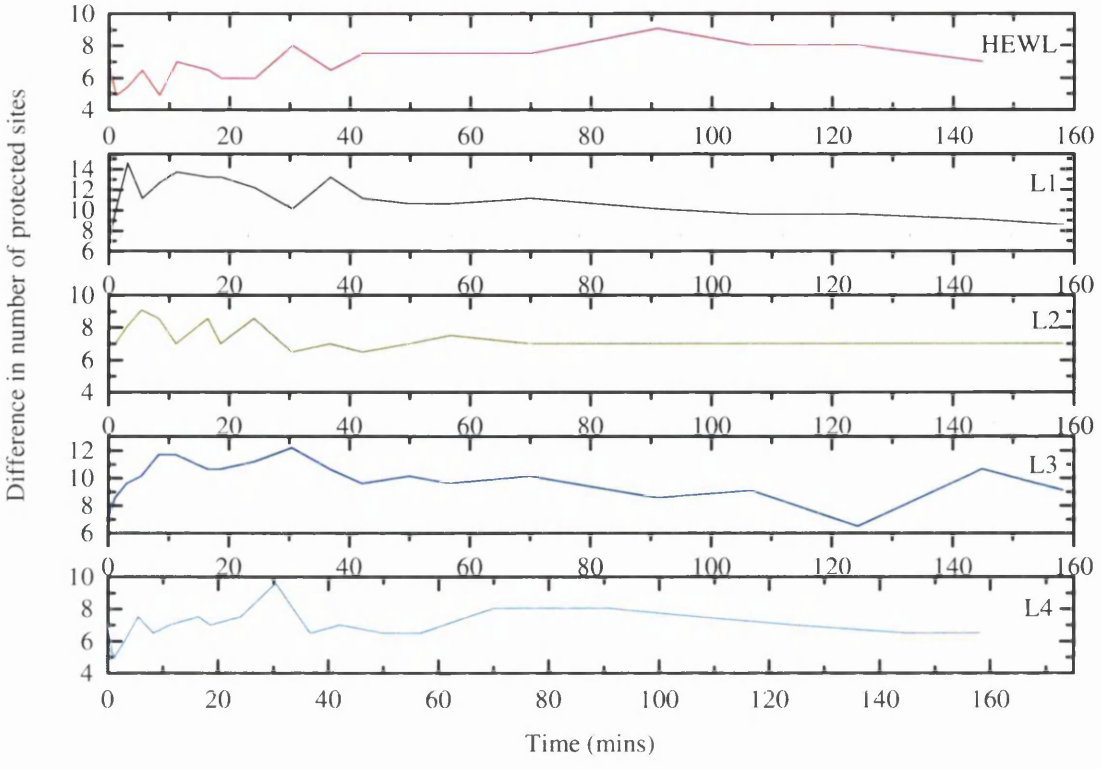


Figure 4.35 : Difference in the number of protected proton exchange sites between human lysozyme and each one of the separate species and native HEWL. The difference in number of protected sites between native HEWL and human lysozyme is approximately 7, which is what is expected. Note that L1 and L3 deviate most significantly from this value.

of protected sites deviates from the expected value of 7 (see figure 4.35 solid line) by as much as 7 residues. The measurement of the mass of a biomolecule is accurate to 0.01% of the biomolecules molecular weight (Ashcroft 2000), hen egg white lysozyme has a molecular weight of 14303K.Da therefore these measurements are accurate to 1Da. This would appear to indicate that there is some appreciable difference between HEWL and L1. Also L3 appears to show a different exchange pattern from native HEWL (4.34(d)), as with L1 the offset between the human lysozyme exchange graph and the L1 graph is slightly bigger than that between HEWL and human lysozyme. The difference in number of protected sites seems to vary by about 2 or 3 from the expected value of 7 (fig. 4.35). The difference in number of protected sites between L1 - 4 and HEWL when compared to human lysozyme, using the results from the exchange experiments, has been plotted as a graph in figure 4.35. This clearly illustrates the difference in number of protected sites between L1 and L3, when compared to the other species and HEWL. This result would seem to indicate that there is some difference in the accessibility of exchangeable protons of these species, and that a small, invisible by CD, change in conformation is allowing the deuterium ions access to a normally protected proton.

Native HEWL is resistant to trypsin digestion, so an interesting experiment was trypsin digestion of the separated peaks, to see if the alteration of the structure caused by heat treatment would affect this resistance. Any change in the structure of a protein which leads to a difference in the exposure of any lysine and arginine residues, compared to the native protein, will allow digestion to occur. Trypsin digestion was used to try and identify any changes in the structure of L1 - 4 compared to native HEWL. Native HEWL, HT HEWL and the separated peaks were incubated for 24 hours with trypsin and samples were removed at different time points, boiled and then frozen to prevent any further trypsin digestion and were subsequently analysed for cleavage products on denaturing and reducing SDS PAGE gels.

The first gel shown in fig. 4.36 shows the reduced and denatured gel of the protein samples without treatment with trypsin, or under the control experiment conditions, i.e. not incubated at 37°C for any time. This shows that the HT HEWL and the separate peaks appear to contain some small peptide fragments. These had not been noticed before because these gels have been run using reducing and denaturing conditions unlike previous gels which had been run using denaturing conditions. These fragments would also have been missed during the mass spectrometry experiments, as only high molecular weight species were being monitored. At zero time, figure 4.37, when the trypsin has had no time to act, all peaks are present on the

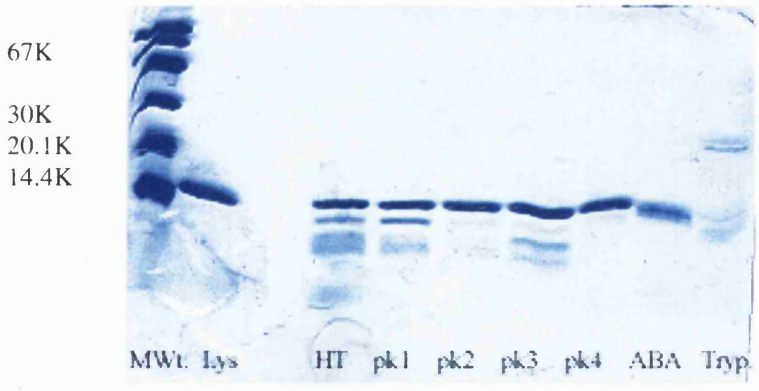
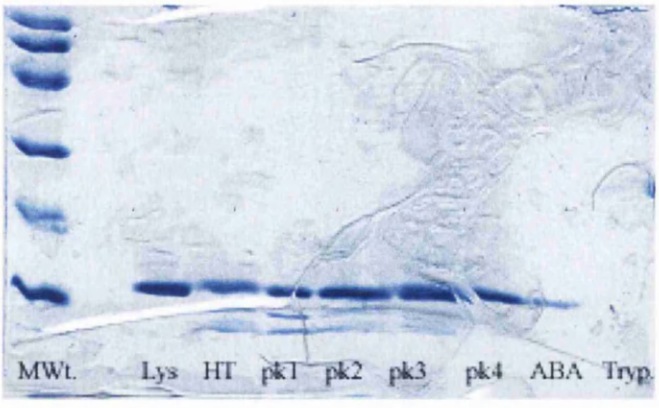
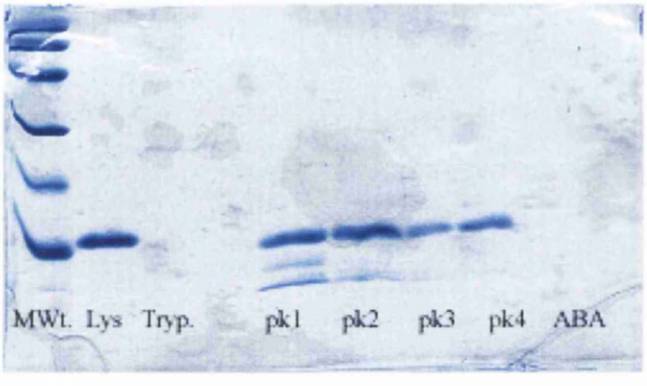


Figure 4.36: Reducing and denaturing gel of heat treated lysozyme (HT), native lysozyme (Lys) and the 4 separate peaks; pk1 - L1, pk2 - L2, pk3 - L3, pk4 - L4. All the samples run to the same extent on this gel. This also shows the presence of small peptide fragments in the samples which are caused by degredation of the protein chain at high temperatures.



(a)



(b)

Figure 4.37 : Reducing and denaturing gel of heat treated lysozyme (HT), native lysozyme (Lys) and the 4 separate peaks; pk1 - L1, pk2 - L2, pk3 - L3, pk4 - L4. Gel (a) no trypsin has been added, (b) the samples have had trypsin added, were shaken and then the trypsin reaction was stopped (0 time). These show that over this time period trypsin has not digested any of the proteins nor does it appear to have digested any of the peptide fragments.

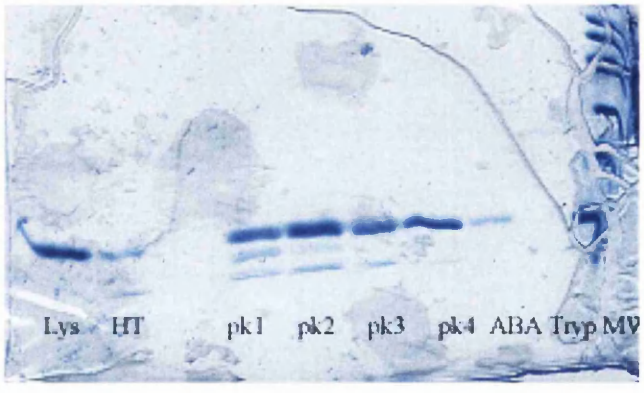
gel as are the bands at lower molecular weights, or the peptides. After 3 minutes the bands which correspond to the peptide fragments appear to have decreased in intensity, fig. 4.38, indicating that the trypsin has digested some of these peptides. Otherwise the bands which relate to the separate peaks, native and HT HEWL are still present in qualitatively similar amounts. Within 2 hours (figure 4.39) the peptide fragment bands have become very faint, and the main bands may have become less obvious too, although this may be due to experimental error or lower sample concentration due to sample boiling off when the samples are treated to stop trypsin digestion. A similar picture is seen after 4 hours fig. 4.40. At the end of the experiment, after 24 hours fig. 4.41, the trypsin has digested the small peptide fragments that appear to be present in the separate peaks and in HT HEWL, as none appear to be present in the gel. Also the separate peaks and the HT HEWL bands seem to be less obvious, indicating that some digestion has taken place. These experiments appear to indicate that the misfolded protein is more susceptible to trypsin digestion , so the conformation of the protein has changed so as to allow the trypsin enzyme access to its substrate sequence.

Summary

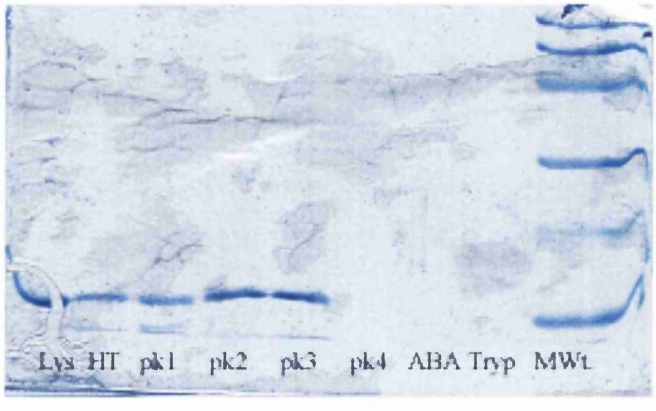
The separate species and HEWL have a similar molecular weight, isoelectric point and structure by CD, but most other techniques have indicated differences in the species, including GuHCl unfolding. Significantly, the separate species are more susceptible to trypsin digestion.

The thermal stabilities are different from native HEWL, as are the HIC retention times and the hydrogen exchange profile, with the exception of L2. Overall L2 appears to have a lot of similarities to native HEWL. It has not been possible to identify the underlying cause of the differences in these species, so we can call them non-native forms of HEWL.

If we assume that the process behind the production of these species is proline isomerisation, the results of the chromatography in conjunction with other results seem to suggest a proline isomerisation mechanism similar to the one described earlier. When HEWL is repeatedly heated from 20 - 100°C, chromatography and DSC indicate 4 species are created. L2 appears to have similarities to HEWL which would indicate that both prolines are in the trans configuration. When this peak is isolated and then reheated the same pattern of peaks is produced, L1-4 and L2 reduces in intensity. The other 3 species produced have different thermal and chemical stabilities when compared to native lysozyme, but have other

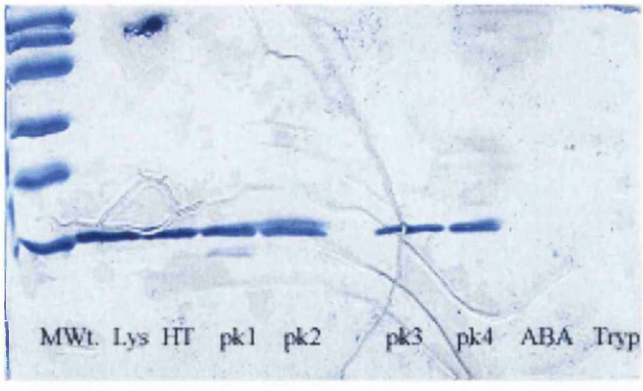


(a)

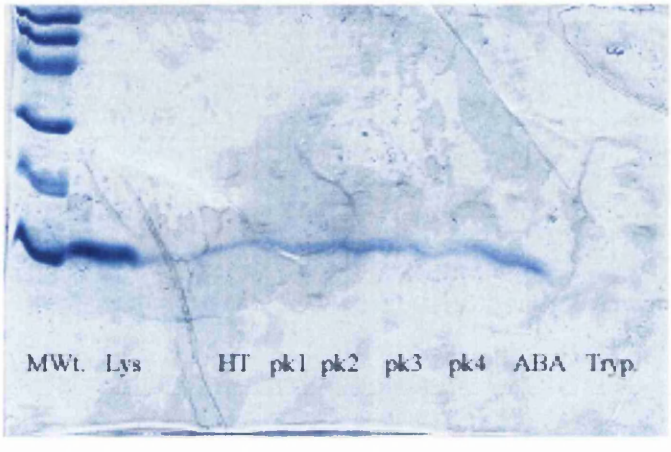


(b)

Figure 4.38 : Reducing and denaturing gel of heat treated lysozyme (HT), native lysozyme (Lys) and the 4 separate peaks; pk1 - L1, pk2 - L2, pk3 - L3, pk4 - L4. Gel (a) no trypsin has been added, (b) the samples have had trypsin added and were incubated for 3 minutes at 37°C. These show that over this time period trypsin has not digested any of the proteins, but that the intensity of the peptide fragment bands has decreased where trypsin has been added.

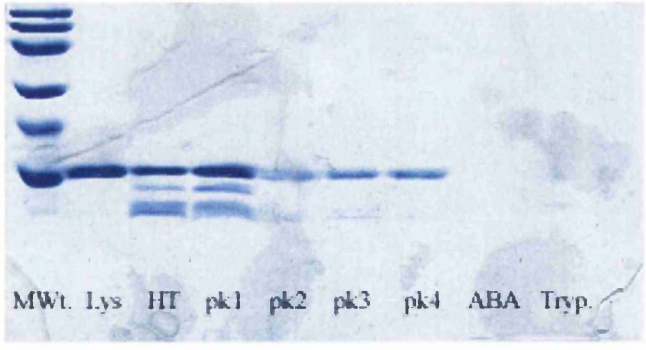


(a)

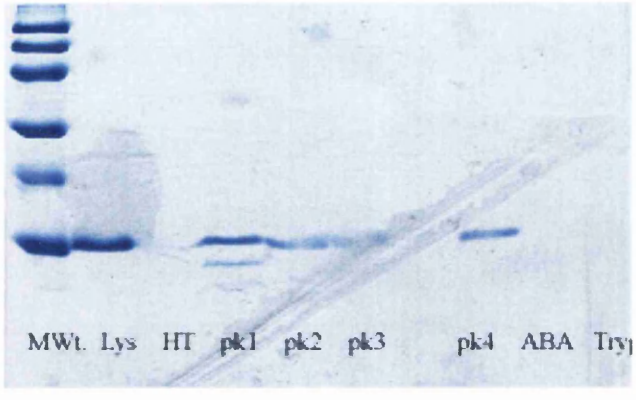


(b)

Figure 4.39 : Reducing and denaturing gel of heat treated lysozyme (HT), native lysozyme (Lys) and the 4 separate peaks; pk1 - L1, pk2 - L2, pk3 - L3, pk4 - L4. Gel (a) no trypsin has been added, (b) the samples have had trypsin added and were incubated for 2 hours at 37°C. These show that over this time period trypsin may have digested some of the proteins and that the intensity of the peptide fragment bands has decreased where trypsin has been added.

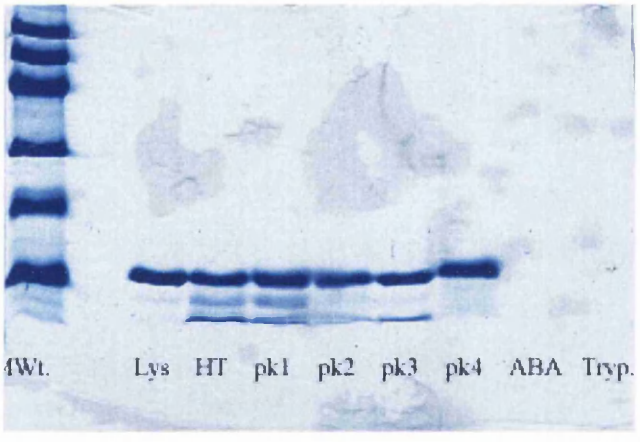


(a)

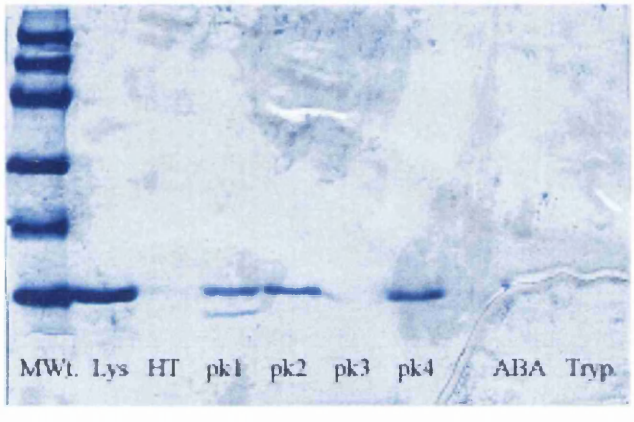


(b)

Figure 4.40 : Reducing and denaturing gel of heat treated lysozyme (HT), native lysozyme (Lys) and the 4 separate peaks; pk1 - L1, pk2 - L2, pk3 - L3, pk4 - L4. Gel (a) no trypsin has been added, (b) the samples have had trypsin added and were incubated for 4 hours at 37°C. These show that over this time period trypsin appears to have digested some of the proteins and that peptide fragment bands have disappeared where trypsin has been added.



(a)



(b)

Figure 4.41: Reducing and denaturing gel of heat treated lysozyme (HT), native lysozyme (Lys) and the 4 separate peaks; pk1 - L1, pk2 - L2, pk3 - L3, pk4 - L4. Gel (a) no trypsin has been added, (b) the samples have had trypsin added and were incubated for 24 hours at 37°C. These show that over this time period trypsin appears to have digested some of the proteins and that peptide fragment bands have disappeared where trypsin has been added.

characterisitics in common with HEWL. L1 and L3 when interconverted produce a pattern of peaks that only involve each other which would suggest that they have one proline in the trans configuration and the other cis, they also have a similar height and area. This leaves L4 which presumably would have both prolines in the unfavourable cis configuration, and experimentally gives L1 and 3 when reapplied to the HIC column, it is also a very small peak which would support it being an unfavourable isomer.

Chapter 5 : General Discussion

General Discussion

The results from this study suggest that heat treatment of HEWL creates four separate, but interconvertible, species (and some aggregate and peptides) which have similar structure to native HEWL but have some degree of flexibility. These species are also less active than native HEWL which may be a reflection of this flexibility.

Overall there is evidence that L2 is native HEWL, the T_m is the same as is the tryptophan exposure, the CD spectra are similar, the elution profile on an HIC column is the same, and the number of protected exchange sites is the same. Although other experiments show a disparity, i.e. CD unfolding experiments and the enzymatic activity. This may suggest a small change in the internal interactions holding the protein in its folded conformation.

The results of many of the experiments show disparities, for instance CD suggests that there is no change in the secondary or tertiary structure of the protein after heat treatment whilst hydrogen exchange experiments, guanidinium hydrochloride unfolding experiments and fluorescence experiments indicate that there may be a small difference in some aspects of the structure of the heat treated protein .

The CD experiments measure the interaction of circularly polarised light with amino acid side chains and the hydrogen exchange experiments measure the change in molecular weight after exchange of the amide hydrogen. The CD experiments suggest no major change in structure after heating as all of the separate species L1-4 have a similar spectrum. The hydrogen exchange shows a significant difference in the amide availability to deuteron exchange, suggesting that some of the amide groups are more exposed. These methods examine protein structure by looking at the amino acid side chains but look at different aspects of the side chain.

The guanidinium hydrochloride exchange experiments show that all of the four separate species are less stable towards denaturants, this agrees with the DSC experiments which show that all of them, except L2, have a lower thermal stability.

Fluorescence experiments were used to compare the tryptophan exposure of the separate species, the tryptophans are generally found in the interior of the protein as they are very hydrophobic. The fluorescence spectra indicate that there may be a small difference in the tryptophan exposure. The fluorescence quenching experiments, which measure tryptophan exposure, shows that the exposure of the tryptophans is the same for all species. The disagreement in results between these two

methods may be due to the difference in the way the tryptophan exposure is measured. The fluorescence experiment looks at it directly but the quenching experiment involves an interaction between the tryptophan and the quenching molecule, any differences between the results may be due to difficulties in the quenching agent gaining access to exposed tryptophan residues. The quenching agent may have been too large to gain access to the tryptophan residues.

Each of these techniques looks at a difference aspect of protein structure but all of them appear to suggest that there may be differences in the interactions that hold the protein together in its folded form, changing its thermal stability and accessibility to amino acids. Therefore the internal interactions which help stabilize the protein are more flexible than in native HEWL, so these species can be referred to as non-native forms of HEWL.

An alternative explanation for the disparities between experimental results is that because of the nature of these samples it was impossible to analyse one sample by all methods, so each experiment was performed using different HT HEWL samples. Therefore the differences in results may be due to different amounts of proline isomers being present in each of the samples. Each heat treatment will have created different amounts of proline isomers because conditions will have been different each time.

Another possible argument is that these four conformers are molten globule forms of HEWL.

The molten globule state is reported to be another possible folding intermediate (Ptitsyn et al 1995 and 1990, Creighton 1992, Denisov et al 1999). These are protein folding intermediates which have secondary structure but no tertiary interactions formed, and have a volume 20-30% greater than that of the native protein. Their unfolding enthalpy is the same as that of the unfolded protein, hence they are not visible by calorimetric methods. In some cases it is possible to form similarly structured species by putting the protein of interest into low pHs, low denaturant and salt concentrations thus making the study of the molten globule practicable.

This explanation is unlikely as these species have a visible DSC thermogram which a molten globule would not have. These species also have the same structure as the native HEWL by CD, a molten globule would not have as much structure.

Another explanation could be that the HEWL is aggregating this could be the reason for the extra peaks seen in the DSC thermogram, however it is unlikely that this is the case as aggregates usually unfold at higher rather than lower temperatures compared to the main transition. Also it is well

known that lysozyme tends to aggregate at pHs greater than 4 (Thomas et al 1996, Nesmelova & Fedetov 1998) and this experiment was performed at a more acidic pH. Study of the mechanism of enzyme inactivation at 100°C by Ahern & Klibanov (Ahern & Klibanov 1985) showed by size exclusion chromatography (SEC) that aggregation was not the cause of HEWL inactivation at pH 4, 6 or 8.

Zale and Klibanov (Zale and Klibanov 1986 , Ahern & Klibanov 1985) have shown that there are specific mechanisms that lead to enzyme deactivation; disulphide exchange, hydrolysis of peptide bonds and deamidation of gln/asn residues .

A chemical modification such as disulfide rearrangement or deamidation could be the reason behind this behaviour of HEWL. If intermolecular disulfide rearrangement had occurred this would have been seen in the non-reducing gels as a band at a multiple of the molecular weight of lysozyme (14303). Ahern & Klibanov (Ahern & Klibanov 1995), during their investigation of the causes of heat induced enzyme inactivation, found that reduction-oxidation of heat inactivated HEWL did not recover any of the lost enzyme activity, suggesting that intramolecular disulphide rearrangement is not responsible for the species made by heat treatment. Also disulfide rearrangement due to destruction of cystine residues tends to occur at basic pH (Zale & Klibanov 1986, Ahern & Klibanov 1995).

Deamidation would also have been visible on the isoelectric focusing gels as bands at a lower pI than HEWL, as deamidation leads to a decrease in isoelectric point (Ahern & Klibanov 1985). In the case of these experiments, even though the gels only ran to a pI of 9, if deamidation had occurred it is likely to have been seen in bands at lower pIs. Ahern & Klibanov (Ahern & Klibanov 1985) have shown that when HEWL is heated at a 100°C for a maximum of 220mins tri deamidation is seen. The trideamidated species would have a pI of 8. The HT HEWL was made by heating HEWL from 20 - 100°C, 5 times, then cooling which would be a comparable process.

They have also shown that at certain conditions in the case of HEWL (pH 8), that the formation of incorrectly folded and kinetically trapped structures contributed to its deactivation (Ahern & Klibanov 1985). In this thesis, under different conditions, we have shown that incorrectly folded structures have lead to a decrease in enzyme activity.

A small amount of peptide bond hydrolysis can be seen in the reducing and denaturing gels of HT HEWL and the separate species. Although there are some peptides present in the separated species there is not enough to significantly effect their behaviour.

It has been possible to compare some of the results of this study with published work on the subject of proline isomerization in other proteins.

Work done by Alexandrescu (Alexandrescu et al 1989)indicated that Staphyloccal Nuclease existed as 3 conformational isomers. They have identified resonances on the ^{1}H NMR spectra of staphyloccal nuclease as belonging to 3 folded forms of the protein. One of these was the native form, one a dimer and the third a non-native proline isomer. The amount of the non-native proline isomer could be increased by increasing temperature and lowering the pH. Acid pH is known to catalyse proline isomerisation (Schmid & Baldwin 1978) and increased temperature will provide the energy required to overcome the activation energy barrier. Therefore, the conditions which produced the proline isomer are similar to those used in our experiments which produce the misfolded form of HEWL (i.e. low pH (pH 3.4) and elevated temperatures). Therefore it is likely that the conditions used to produce the misfolded protein are conducive to proline isomerization.

Also the similarity in molecular weight between the HT HEWL and native HEWL is consistent with the notion that these different species are due to proline isomerisation, as this would in no way effect the molecular weight of a protein. This concurs with what was seen during the biochemical characterization of Calbindin D_{9K} (Chazin et al 1989). This protein molecule exists in two folded forms which differ by the isomerisation state of Pro 43, though no difference from native protein was identified by gel electrophoresis or isoelectric focusing.

The experiments which involved treating the separate species in such a manner as to increase the rate of proline isomerisation (by heating or by leaving to equilibrate) indicated that the interconversion process was slow. Again this would be consistent with proline isomerisation, especially at 4°C. Slow proline isomerisation has been identified before in the work of Yuan (Yuan et al 1998) who identified the existence of two isoforms of transforming growth factor β (TGF-β) protein like domain from human fibrillin. These isomers do not appear to interconvert on the millisecond or hour timescale, which is what is expected from other studies where isomerisation is studied in a protein refolding or unfolding situation. This lead to the conclusion that the isoforms they identified do not interconvert directly, though the interconversion may take place over a longer period of time than the NMR timescale.

The observed difference in stability between HT HEWL, L1 - 4 and native HEWL, is similar to observations made by Langestmo (Langestmo et al 1989). They showed that reduced thioredoxin appeared to have two folded forms with different stabilities, using urea gradient gels (Creighton 1980). The gels were performed at low temperature (2°C) so as to slow down proline isomerisation and at room temperature. When the reduced thioredoxin was run on the urea gels, either as unfolded or folded protein, it showed two transitions one at 4.2M and one at 5.5M urea at both temperatures. The experiments were repeated using a reduced thioredoxin variant, which had Pro-76 mutated to an alanine residue, this showed one transition at a lower urea concentration than oxidised thioredoxin. Therefore the reduced thioredoxin exists as two folded forms with Pro 76 in the cis or trans configuration with similar but distinguishable stabilities in urea, which interconvert slowly on the electrophoresis timescale hence they can be seen at both temperatures.

The separate species L1 - 4 and native HEWL have different enzyme activities. This is in contrast with the similarity in enzyme activity of native HEWL and HT HEWL. There are conflicting reports on the matter of substrate binding to the active site of a protein with a non-native proline isomer. Schimd & Blaschek (Schimd & Blaschek 1981) identified an intermediate on the folding pathway of Rnase A which had a non-native proline isomer and appeared to bind 2`CMP and remained enzymatically active. Later on Lin & Brandts (Lin & Brandts 1983) showed that Rnase A required all prolines to be in the correct configuration for enzymatic activity to be regained during folding. There is a possibility that this difference is related to the methods used to measure the isomerisation state of the protein; Schmid & Blaschek measured tyrosine fluorescence and Lin & Brandts used their Isomeric Specific Proteolysis (ISP) method.

In this case the activity of the separated species is less than that of native HEWL which would seem to agree more with Lin & Brandts results as Schmid & Blaschek showed their intermediates to have the same activity as native Rnase A.

There is no evidence which absolutely identifies these four species as being the proline isomers of HEWL, but their means of production, their behaviour and the manner in which they interconvert seems to be consistent with the theory that they are due to proline isomerisation and is similar to studies done by other people where proline isomerisation has been identified.

Looking at these results from the perspective of a biotechnology company the methodology used in this thesis could help towards understanding what type of processes cause misfolding in recombinant protein production. Problems that are known to arise with recombinant protein are aggregation and loss of activity (Fink 1998, Mitraki & King 1989) both of which are features that

the separate species show. Usually the activity of a protein is what is used to characterize any problems with the protein's production but the reason for this is not investigated, so it is interesting to postulate that proline isomerisation could be behind this too. Also the difference in T_m of the separate peaks relates to a problem where production of a glycosylated protein in a different organism led to aggregated protein which then was centrifuged to solubilise it. DSC of this solubilised protein indicated that it had a lower T_m than expected (Personal Communication Joseph Farchaus).

Future work

This section contains a discussion on possible future experiments that could positively identify these misfolded species as being proline isomers or not, as the case may be, and provide some information on the effects of manufacturing processes on misfolded proteins.

Chemical Analysis

A quantitative chemical analysis of the heat treated protein would determine how the processes of inactivation have contributed to the decrease in enzyme activity of the HT HEWL. This could be done using the methods used by Zale & Klibanov (Zale & Klibanov 1986) and Ahern & Klibanov (Ahern & Klibanov 1985), who studied the effect of heating enzymes such as HEWL and Ribonuclease A. The processes they tested for were disulfide rearrangement, deamidation and hydrolysis. Although some of the methods they used were applied during this research, a separate and more focused study using these methods alone would give a comprehensive view of the effect of heat treating lysozyme under these conditions.

Size Analysis

One of the possible reasons for this behaviour of HEWL is aggregation, as aggregation has been postulated to be one of the causes of loss of enzyme activity. Analysis of the separated species (L1-4) and the heat treated HEWL by size exclusion chromatography, light scattering (Ackland et al 1991, Bauer et al 2000) or ultracentrifugation (Varley et al 1997) to determine whether the separate species are aggregated forms of HEWL.

Nuclear Magnetic Resonance

NMR has proved to be one of the most useful techniques for identifying proline isomerisation, ^{1}H NMR experiments appears to be a good way of identifying conformational isomers. A typical spectrum of amino acids in the area near to the proline of interest will show major and minor resonances, the minor resonances will have a similar chemical shift to the major ones. The major peaks will be the expected shifts due to the native protein and the minor ones due to the isomers,

which will be native-like but distinct from the native form (Fox et al 1986, Alexandrescu et al 1989).

These preliminary experiments can then be expanded upon to look at the mechanism behind the isomerization, 2D hydrogen exchange NMR experiments can be used to study the conformational changes that occur between the isomers (Chazin et al 1989). More complex NMR experiments can then be used to determine the structure of isomers, this involves looking at the relationship between chemical shifts of different amino acid residues (Amodeo et al 1994, Yuan et al 1998).

These methods could be applied to the study of the species identified in this thesis. A useful experiment would involve looking at the ^{1}H NMR spectra of each of the separated peaks with native HEWL. Native HEWL would act as an internal standard for comparison, and that of the four separated peaks mixed together (not the heat treated lysozyme as it contains aggregated and severely misfolded protein) again with native HEWL. Also 2D hydrogen exchange NMR of the species could be instructive, this experiment would have to be performed at elevated temperatures to increase the rate of isomerisation, but would help identify the process whereby the separate peaks interconvert. Further experiments as mentioned above would also be interesting. Unfortunately there is a problem, the amounts of protein required for these experiments exceeds what can be made easily in the time period over which interconversion occurs. It is also possible that at these high concentrations that aggregation of the heat treated protein may occur. For this type of experimentation a scaled up method of production of the heat treated lyszoyme and the separated peaks would have to be developed.

Unfolding /Refolding Experiments

The kinetics of folding and unfolding of proteins can be monitored using fluorescence or UV spectroscopy, these experiments involved diluting aliquots of protein in appropriate solutions into buffer or in the latter case buffer containing denaturant and then watching the change in fluorescence or absorbance over time. If the kinetics of folding or unfolding are very fast stopped flow apparatus is used. This helps overcome the problem of missing kinetic phases which occur during the manual mixing time, although currently new methods are being developed all the time which overcome this delay.

It would be interesting to compare the kinetics of unfolding of native HEWL and the separate peaks. Any differences in kinetics may be due to proline isomerisation (Ibarra-Molero & Sanchez-Ruiz 1997). Also, if the correct conditions could be found, experiments monitoring the refolding of the separate peaks to native HEWL would provide valuable information. This refolding experiment may be difficult to monitor though, as there is no difference in absorbance or fluorescence between the separate peaks and native HEWL (Kato et al. 1981).

Isomeric Specific Proteolysis

This method was developed by Brandts & Lin (Brandts & Lin 1986) and involves the use of stereospecific proline bond enzymes to hydrolyse the protein, one of the resultant products of the hydrolysis is assayed. This can then be used to calculate the amount of cis or trans isomer. This has been successfully used to calculate the amount of trans (non-native) pro93 in RnaseA under different conditions (Lin & Brandts 1983). This method has been discussed in more detail earlier in the Introduction.

In the case of these non-native forms of HEWL, although it is dependent on the primary sequence of HEWL, the most likely sequence of treatment would be a short burst of pepsin hydrolysis to denature HEWL. This is followed by treatment with one of prolidase or APP, which are proline specific enzymes, depending on the peptides that pepsin hydrolysis produces. Otherwise these peptides would be treated with some other general enzyme prior to treatment with the proline specific enzymes, to provide a better substrate with an assayable product . The percentage of isomers in the protein compared to native HEWL, would be calculated by measuring one of the products of hydrolysis of the peptides by the proline specific enzymes.

Lysozyme Mutants

It may be useful to produce a mutant form of HEWL with one and both proline residue substituted with a small, non-imide amino acid. If this mutant then had the same structure as native HEWL (this would have to be studied also), the original DSC experiment could be performed using these mutants. Presumably if the effect we are seeing is caused by proline isomerization substitution of the proline residues will prevent it from occurring.

This type of strategy has been used extensively to study the slow phases in folding of stapphylcoccal nuclease (Ikura et al 1996, Evans et al 1987, Maki et al 1999, Truckses et al 1996).

The effect of freeze drying

The effect of freeze drying on proteins is not generally understood and since this is the usual method of preserving them for storage it would be helpful to find out how this treatment effects their behaviour.

Because the four separate peaks produced by heat treatment have lower enzyme activity than native HEWL, it is possible that the mechanism behind these differences is similar to the ones that cause decreased enzyme activity in mass produced protein. Therefore, if each of the four peaks is freeze dried separately, then run on the same type of HIC column as used in previous chromatography experiments, the effect of freeze drying on proteins could be studied.

Iodide quenching

Because the results of the succinimide quenching were at odds with what was determined from ANS binding and from the behaviour of the separate peaks on the HIC column, iodide quenching could be used to investigate more thoroughly the effect of heat treatment on the exposure of the tryptophan residues. This experiment wasn't carried out because of time constraints.

Appendix I : Supplemental Experiments

Crystallisation Trials

Crystallisation of the HT HEWL was attempted to see if it was possible to actually visualize any changes within HEWL`s structure, i.e. proline isomerisation. Trials were started by using conditions that HEWL crystals are known to grow under, method 34 of the Magic 50 and 25% PEG 4K, 250mM NaCl in 0.1M pH 4.5 sodium acetate buffer. When this was unsuccessful the conditions were altered by varying the amounts of salt and PEG in each, and by changing the pH of the buffer. Then the Magic 50, M and N screens and the footprint screen were tried. These efforts resulted in a lot of precipitate and the occasional twinned crystal. Crystals that did grow were identified as having the same unit cell dimensions as HEWL. It is possible that because the HT HEWL is made up of more than one isomer crystallisation could not occur (Giege & Ducruix). So these experiments were discontinued.

Cyclophilin Experiments

Cyclophilin is a peptidyl prolyl isomerase and is known to catalyse the folding and unfolding of proteins (Fischer & Bang 1985; Lang et al 1987). It increases the rate of any slow steps in folding /unfolding that involve a proline isomerisation (Mücke & Schmid 1992). Experiments were performed to see if this enzyme would catalyze isomerisation of any prolines that were potentially in the wrong conformation in the HT HEWL.

The cyclophilin and recombinant protein production facilities were kindly provided by Dr. Tony Page of Glasgow University (Page et al 1996). Preparation and production of recombinant B. Malayi cyclophilin fused with maltose binding protein was performed as described by the manufacturer of the plasmid that contained the cyclophilin gene (New England Biolabs 1993).

Catalytic amounts of cyclophilin, approximately 5.6μM per ml of lysozyme were added to 2mls, 2mg/ml heat treated lysozyme (HT HEWL). This solution was then incubated at 37°C in a warm room for 4/8/24 hours. Meanwhile controls with the same volume of buffer added to HT HEWL were also incubated. After the appropriate time period the samples were removed from the warm room and prepared for the DSC. A scan of the protein from 20 - 100°C at 60°C/hr scan was performed and the resulting thermogram was compared to those collected during the creation of the HT HEWL and the controls.

In the first instance these experiments were performed using HT HEWL which had been made at pH 3.4 and then dialyzed into pH 5 citrate buffer, but this led to aggregation. Preliminary experiments to discover an alternative method of making HT HEWL at the higher pH indicated that 3 scans in the DSC under the usual conditions created a small amount of misfolding similar to that seen at pH 3.4. Incubation of this with approximately 24μM cyclophilin for 4 hours gave the same sort of DSC thermogram as seen originally, except that the baseline was very noisy. In subsequent experiments the concentration of cyclophilin was increased giving a ratio of 1:25 cyclophilin to lysozyme and the time of incubation increased from 8 to 24 hours. In all cases the resulting DSC thermograms were very noisy due to the presence of aggregate and the DSC traces looked similar to the original HT HEWL thermograms. Increasing the concentration of cyclophilin and the time of incubation of cyclophilin with HT HEWL did not appear to effect the DSC trace which appeared to be similar to that of the HT HEWL. It should be noted however that the baseline of these thermograms was generally very noisy before the main transition, due to the presence of aggregate and therefore no definite conclusion could be drawn from these experiments.

After discovery of a method of separation of the HT HEWL into its constituent species the previous experiment was repeated and the cyclophilin treated species and controls were separated on the HIC column in the same manner as outlined in the HIC procedure to see if the cyclophilin would convert the heat treated lysozyme back into a single species. If the cyclophilin successfully isomerised any incorrect proline conformers, the sample would give a chromatogram similar to that of native HEWL. Again these experiments were inconclusive, the chromatogram of HT HEWL and HT HEWL with cyclophilin were exactly the same, as was the control of native HEWL, and native HEWL with cyclophilin. These results seem to indicate that the cyclophilin was not able to catalyse any isomerisations in these samples, presuming that the misfolded forms are caused by proline isomerisation.

There are other problems to be taken into consideration when interpreting the results from both these experiments, due to the conditions under which the experiments were performed. Cyclophilin is a thermally sensitive protein and its optimum pH is approximately 8. Unfortunately heat treated lysozyme aggregates at certain pH`s, including pH 8. So these experiments had to be performed at lower pH, because dialysis of the HT HEWL into higher pH conditions led to the creation of aggregate which was the protein of interest. The activity of cyclophilin is about 6 fold lower than at the optimum pH of pH 8 under these conditions (Lin et al 1988). So it is entirely possible that the cyclophilin was not functioning correctly

because of the solution conditions, or that because the heat treated protein seems to be folded approximately correctly (see Results and Discussion) that the cyclophilin could not gain access to the proline residues of interest.

Appendix II : Aggregation Experiments

Introduction

Because of the increasing interest in the prion diseases such as Bovine Spongiform Encephalopathies and Creutzfeld Jacob disease (Prusiner 1997, Lanchester 1996, Almond & Patterson 1997), where prion protein is converted from one conformation to another, it was relevant to see if this means of propagation was possible in other proteins. This conversion could be described as a result of "misfolding" and so was relevant to the other work done as part of this thesis.

The conversion from the normal prion protein to the infectious form, appears to occur when a normal non-infectious prion protein is exposed to the infectious form of the protein. The infectious form has a region of β sheet which in the non-infectious form is α helical. The β sheet form of the prion protein forms fibrillar plaques in the brain (a "protein aggregate") which can lead to cell death and the symptoms that are associated with these type of diseases. This plaque formation is also a symptom of many other neurodegenerative diseases like Alzheimers and Huntingtons (Wanker 2000), hence there is a need to discover the mechanism of in vivo plaque formation.

Because the α-helical form of the prion protein is presumably the most stable conformation, existence of a non-native infectious β-sheet form is not expected and contrary to all thinking on protein stability that has gone before. The conversion that the infectious form instigates presumably occurs when enough of this infectious form is present, as expected by equilibrium rules, by a nucleation type mechanism. If this is the case this type of aggregation should be more prevalent and could occur in proteins other than these plaque forming ones, as addition of excessive amounts of aggregated protein should induce aggregation under the right conditions.

This chapter describes experiments that have been performed to study the effects of adding protein aggregate to solutions of protein. The protein solutions were subject to conditions that have been proven to induce aggregation, to see if addition of aggregate will change the rate of aggregation. This would provide information on the mechanism of protein aggregation.

Material and Methods

Aggregation experiments

Concentration measurements were made using a Shimadzu double beam spectrometer, as were turbidity measurements (Parker & Dalgleish 1977, Jaenicke & Rudolph 1989). Turbidity was measured as the absorbance at 320nm.

Most involved incubation using a Techne waterbath with a Techne Tempunit®Tu-16D heating element, set to the desired temperature, and in the case of the experiments performed in the spectrophotometer, with an attached Techne RB-5 refrigerated circulating waterbath and Tempette ® TE-8D heating element.

Thioflavin T Binding

10µl of the aggregated suspension was added to 990µl of 0.3µM Thioflavin T made up in pH 7 phosphate buffer. Fluorescence measurements were performed with excitation at 450nm and emission monitored at 482nm. β-Fibres give a large increase in fluorescence intensity at this wavelength (Levine 1995, Wood et al 1996). As a standard, insulin β-fibres were made by heating a pH 2 (water/HCl), 1% insulin solution at 90 -100°C 3 - 4 times for 2 minutes with cooling in an acetone/ dry ice bath followed by defrosting under a warm running tap (Burke & Rougvie 1972). Insulin β-fibres should make a very viscous solution. The previous method gave a cloudy aggregate. A 1%, pH 2 insulin solution was heated at 90°C for 6 hours, giving a viscous solution, so this method was used for production of insulin fibrils.

Results

Preliminary experiments were performed to find the best conditions for causing aggregation.

HEWL at approx. 2mg/ml in pH 3.4 citrate buffer was placed in a waterbath at 60°C. Three other solutions were also exposed to the same conditions: one with 20µl aggregate made by heating HEWL in pH 5 buffer at 100°C for some time, one with aggregate made by heating HEWL in higher pH buffer, and the final one with aggregate made by heating HEWL from 20 - 100 °C at pH 5 in the DSC. The samples had completely aggregated within 11 days, and the concentration of the samples was measured by UV, after centrifugation to remove aggregate. The results of this experiment appear to show, see table 1 , that the concentration of most samples have remained the same, with some increase in the case of the sample with

Time/days	Solution (mg/ml)			
	Lysozyme	Lysozyme + DSC aggregate	Lysozyme + dialysis aggregate	Lysozyme + heated aggregate
0	1.6	1.6	1.6	1.6
11	1.66	2.3	1.67	1.5

(a)

Time/days	Solution (mg/ml)	
	Lysozyme	Lysozyme + heated aggregate
0	1.7	1.7
8	1.46	1.76
9	1.65	1.78

(b)

Table 1 : Rate of aggregation of lysozyme, and lysozyme with aggregate added.

aggregate made by repeat heat treatment in the DSC. The increase in concentration could be due to the aggregate breaking down or because of loss of water through condensation over the eleven days of the experiment. The time period of this experiment was too long to really be useful as a quick aggregation experiment, so conditions were changed allowing the experiment to be performed in one day.

The conditions for an aggregation experiment that would take around 1 hour were discovered to be 2mg/ml HEWL at 70°C solvated in pH 7 phosphate buffer. To 2mls of these solutions quantities of aggregate were added and 2mls were kept aside for use as a control. So it was possible to follow aggregation in the UV spectrophotometer. The experiment involved looking at the turbibidity of the mixture at 320nm. The UV spectrophotometer was connected to a circulating waterbath at approximately 70°C, and the temperature that the experiments were performed at was monitored as this would greatly effect the rate of aggregation.

The results of the three performed experiments are tabulated in figure 1. These results appear to show that after a time of approximately 150 minutes, the solution with aggregate in it had aggregated at a lower rate, than the solution without aggregate in it, suggesting that the presence of aggregate hindered aggregation. These results have possibly been affected by poor insulation around the spectrophotometer, so the experiment was repeated using a waterbath.

The next set of aggregation experiments involved heating the samples with aggregate (0, 2, 5, 10, 15, 20 μL) in them and the HEWL control in a waterbath at 70°C, approx. 2mg/ml HEWL in pH7 phosphate buffer. Samples were removed and their turbidity was read in the spectrophotometer. This enabled me to set up a larger range of samples with increasing amounts of aggregate in them to see what effect that had. This also meant that the temperature was almost constant compared to the experiments using the spectrophotometer, which had poor insulation. A graph showing the results of one of these experiments is shown in figure 2. The graph shows that addition of aggregate slows down the rate of aggregation. This result is consistent between experiments and appears to show that the addition of aggregate has no effect on rate of aggregation of the protein samples.

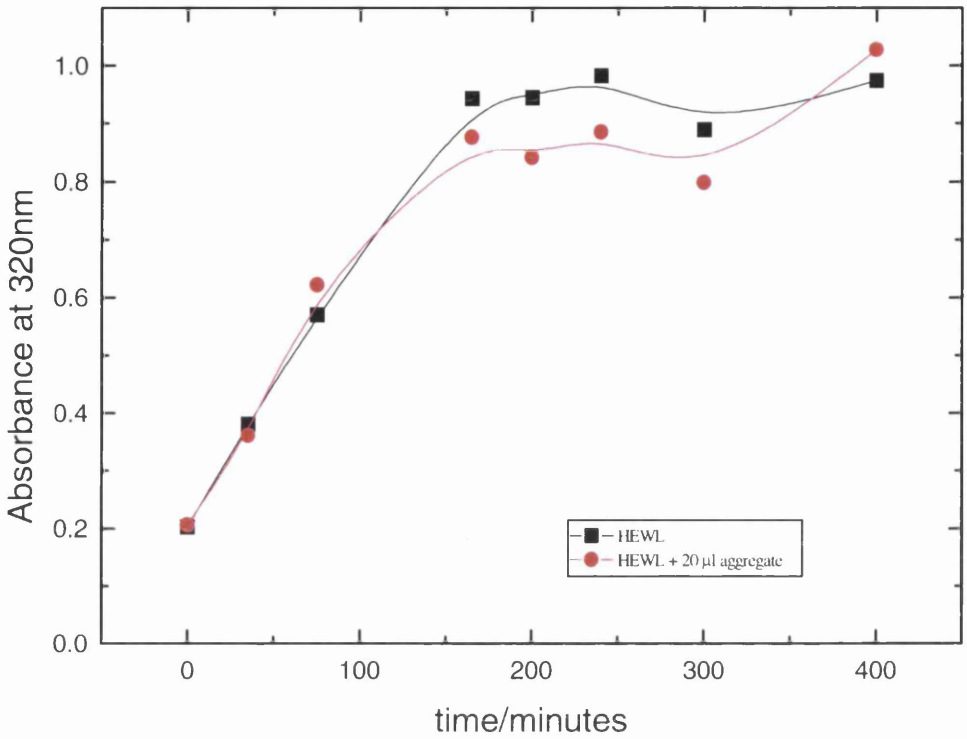


Figure 1: Graph of turbidity against heating time of a solution of lysozyme, ■ : lysozyme + 20µl aggregate, ●. This shows that the rate of aggregation does not appear to be affected by addition of aggregate.

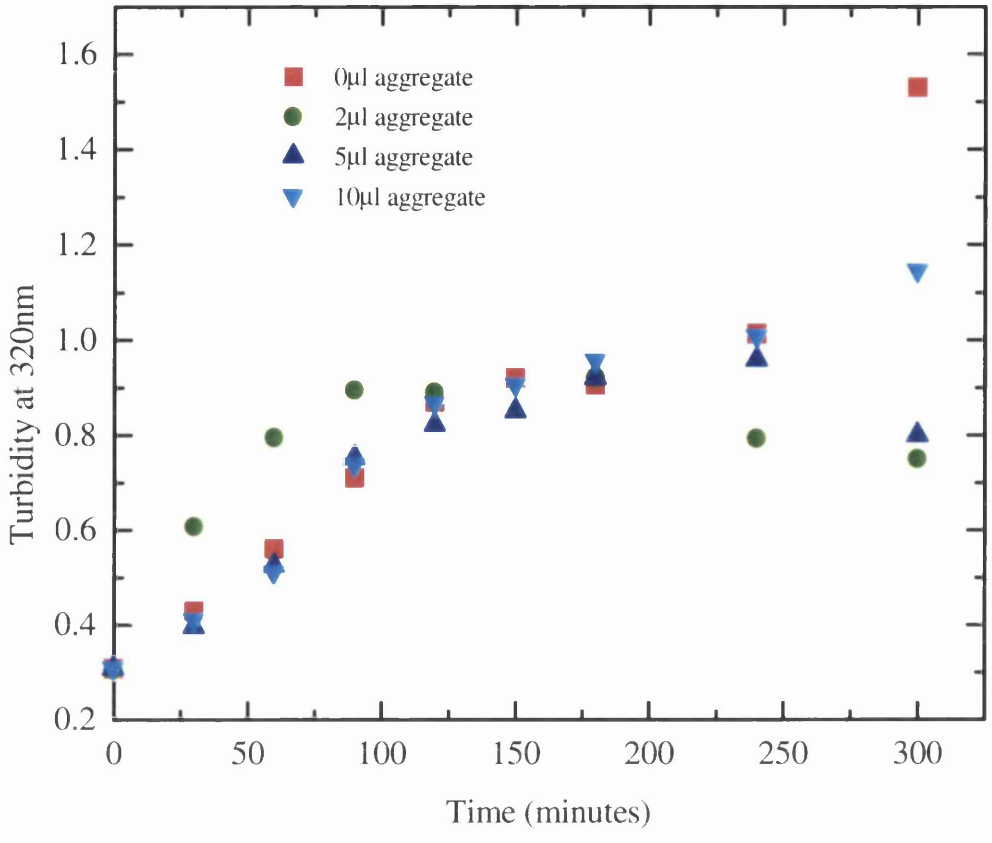


Figure 2: Graph of turbidity against heating time of a solution of lysozyme, and lysozyme plus varying amounts of aggregate. This shows that the rate of aggregation does not appear to be affected by addition of aggregate.

At this time more work on the subject of prion propagation was being published and this seemed to indicate that it would be more profitable to perform these experiments using aggregate that was fibrillar, as experiments performed by Harper & Lansbury (Harper & Lansbury 1997) (see over) indicated that non-specific aggregate would not have the desired effect.

"First, the aggregates formed at pH 5.8 bind more Congo Red on a molecular basis than do A β1-40 aggregates formed at pH 7.4. Second, the pH 7.4 A β1-40 aggregates are fibrillar, as shown by E.M., but do not exhibit turbidity at 405nm, whilst turbid aggregates formed at pH 5.8 do not appear to be fibrillar by E.M. (although they induce a significant enhancement in the Thioflavin T fluorescence spectrum). Third the aggregates grown at pH 5.8 are unable to seed polymerisation of A β1-40 at pH 7.4." **Harper and Lansbury, Seeding Amyloid Formation, Annu. Rev. Biochem. 1997, 66:385-407**

HEWL in different pH solutions varying from pH 1.87 - 8, as more basic pHs would lead to disulphide interchange, was heated at 80°C and the samples removed when aggregation had occured. For the samples at neutral pH this happened within 20 minutes, in an hour for the sample at pH 4, and for the other low pH samples aggregation didn`t occur after 20 hours. These samples were used to try and seed aggregation as fibrillar solutions are known to not have obvious aggregate in them (Wood et al 1996).

The thioflavin T binding of all these aggregates was looked at. Thioflavin T binds to fibrillar aggregates. In particular the fluorescence of the "aggregate" made at pH 1.87 had increased greatly, indicating that these in particular may be fibrillar. The fluorescence of this solution was compared to a standard fibrillar insulin solution. The insulin solution had fluorescence maximum of $2x10^6$, the lysozyme at pH 1.87 had a fluorescence maximum of $7x10^5$. It should be noted however, see the above quote, that thioflavin T is not the best method for identifying the existence of β fibrillar type structure. We have not been able to identify this aggregate as being β fibrillar using other methods such as congo red staining.

The previous aggregation experiments were repeated using this pH 1.87 HT HEWL aggregate, varying amounts of aggregate were added to 2mg/ml HEWL at pH 7 phosphate buffer, then incubated at 70°C for 5 hours and the turbidity measured at different time points. The results of this experiment are reproduced in figure 3. They appear to show that at higher concentrations (15 and 20 μL) of aggregate, slows down the rate of aggregation of the protein solution. This seems to suggest that whatever the structure of this aggregate it seems to

somehow slow down the aggregation process, and possibly this maybe indicates that the HEWL aggregate made at pH 1.87 is not fibrillar as non-fibrillar aggregate does not increase the rate of aggregation. It is also unlikely that this effect is due to the pH of the HEWL solution being changed significantly by the addition of the lower pH aggregate.

HEWL in different pH solutions varying from pH 1.87 - 8, as more basic pHs would lead to disulphide interchange, was heated at 80°C and the samples removed when aggregation had occured. For the samples at neutral pH this happened within 20 minutes, in an hour for the sample at pH 4, and for the other low pH samples aggregation didn`t occur after 20 hours. These samples were used to try and seed aggregation as fibrillar solutions are known to not have obvious aggregate in them (Wood et al 1996).

The thioflavin T binding of all these aggregates was looked at. Thioflavin T binds to fibrillar aggregates. In particular the fluorescence of the "aggregate" made at pH 1.87 had increased greatly, indicating that these in particular may be fibrillar. The fluorescence of this solution was compared to a standard fibrillar insulin solution. The insulin solution had fluorescence maximum of $2x10^6$, the lysozyme at pH 1.87 had a fluorescence maximum of $7x10^5$. It should be noted however, see the above quote, that thioflavin T is not the best method for identifying the existence of β fibrillar type structure. We have not been able to identify this aggregate as being β fibrillar using other methods such as congo red staining.

The previous aggregation experiments were repeated using this pH 1.87 HT HEWL aggregate, varying amounts of aggregate were added to 2mg/ml HEWL at pH 7 phosphate buffer, then incubated at 70°C for 5 hours and the turbidity measured at different time points. The results of this experiment are reproduced in figure 3. They appear to show that at higher concentrations (15 and 20 μL) of aggregate, slows down the rate of aggregation of the protein solution. This seems to suggest that whatever the structure of this aggregate it seems to somehow slow down the aggregation process, and possibly this maybe indicates that the HEWL aggregate made at pH 1.87 is not fibrillar as non-fibrillar aggregate does not increase the rate of aggregation. It is also unlikely that this effect is due to the pH of the HEWL solution being changed significantly by the addition of the lower pH aggregate.

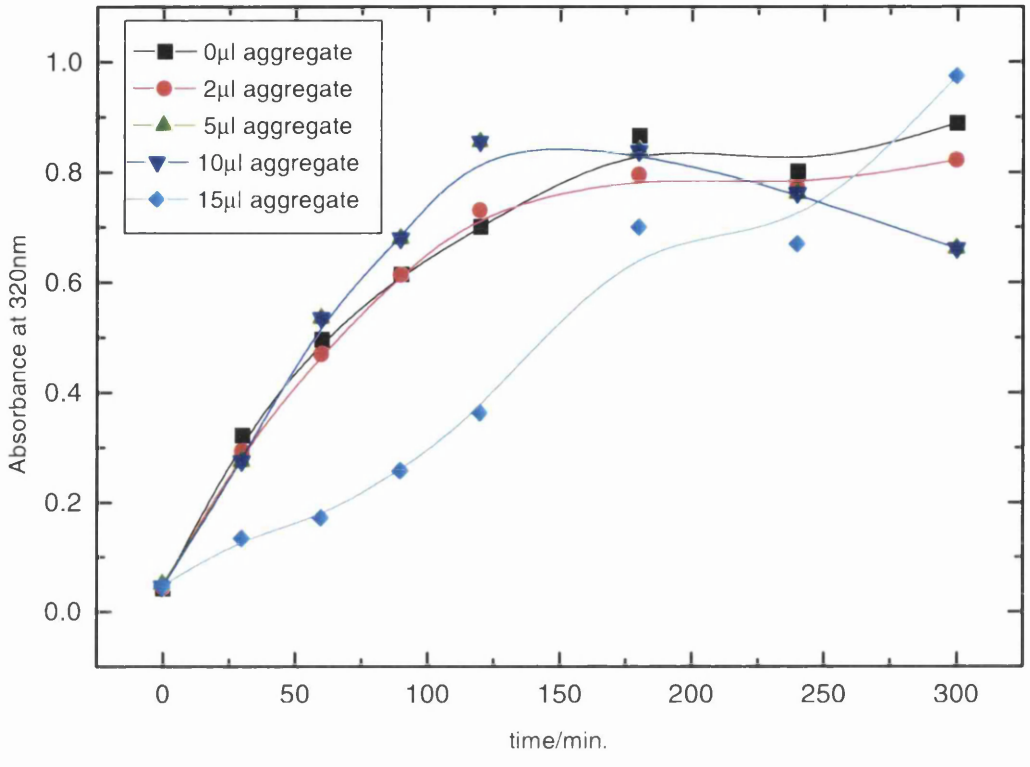


Figure 3: Graph of turbidity against heating time of a solution of lysozyme, lysozyme plus varying amounts of aggregate. All aggregate used in this experiment when added to Thioflavin T resulted in an increase in the dye molecule's fluorescence, suggesting that it was possibly β-fibrillar. This shows that when greater concentrations of aggregate was added to the lysozyme solution the rate of aggregation is decreased.

Discussion

The results of these experiments agreed with the work of other groups that non-specific protein aggregate does not increase the rate of aggregation (Wood et al 1996, Harper & Lansbury 1997). The use of aggregate which may have been β-fibrillar did not increase the rate of aggregation but instead decreased its rate. This may have been because the rate of non-specific aggregation was faster than that of β fibre formation. It is however entirely possible that the aggregate used in this experiment was not β-fibrillar. If this is the case this would suggest that addition of certain amounts of aggregate to a protein solution in conditions that will induce aggregation slows down the formation of protein aggregate. This could be a non-fibrillar version of an observation that addition of non-fibrillar aggregate to a protein solution with Alzheimers protein under fibre forming conditions slowed down fibre formation.

Conclusion

The pertinence of this work to the fibre formation argument is not currently obvious, as so far there has been no evidence that HEWL in particular is able to form an β-fibrillar aggregates. It is likely that this experiment has more relevance to a non-specific aggregation mechanism than a plaque formation mechanism. But this work has provided some interesting information on HEWL aggregation, and has shown that under these particular conditions β fibrillar lysozyme aggregate formation does not occur.

Appendix III : IEF Coomassie Blue Staining Protocol

IEF Coomassie Blue Staining Procedure

This technique file describes optimized methods for detecting proteins in PhastGel™ IEF and gradient media by coomassie staining. The methods are designed to give reproducible , fast resullts with good sensitivity and low background staining . Results are obtained within 25 - 45 minutes, depending on gel type and sensitivity required for the application.

The average sensitivity limit of this technique is estimated to be approximately 20 to 30 ng of protein per band. This is based on serial dilution studies of the Pharmacia Calibration Kit proteins.

1. Fix 20% trichloroacetic acid. This solution can be recycled 3 to 4 times.

Leave for 5 mins. At 20°C.

2. Wash : 30% methanol and 10% acetic acid in distilled water (3:1:6). This solution can be recycled 3 to 4 times. Leave for 2 mins. At 20 °C.

3. Stain : 0.02% PhastGel Blue R tablet in approximately 30% methanol and 10% acetic acid in distilled water and 0.1% (w/v) $CuSO_4$.

Stock solution : dissolve 1 PhastGel Blue R tablet in 80ml of distilled water by stirring for 5 - 10 minutes. Add 120ml of methanol and stir for 2 mins. This makes a 0.2% solution.

Final solution : mix 1 part of filtered stock solution with 9 parts of a methanol : acetic acid : distilled water (3:1:6) solution listed in the stock solution step. Add $CuSO_4$ to 0.1% (w/v). The $CuSO_4$ is added to decrease the background staining.

Make the final solution fresh the day you plan to use it. Do not recycle this solution.

Leave for 10 mins at 50°C.

4. Destain : 30% methanol and 10% acetic acid in distilled water (3:1:6), as in step 2 above. Leave for 10 mins at 50°C.

Appendix IV : HPLC/HIC Protocol

Method Name: hic17.4

Run Name: ht1721299

Run Operator: Sarah

Trace 1: UV Detector Trace 2: Conductivity Monitor Threshold: UV Detector

No Fraction Collection

Starting Chart Recorder ranges: UV = 0.1000, Conductivity = 5.0000

Run Date: 02-17-99

Total Run Time: 4:12.6 hr:min

14:32:58 Initializing...

14:32:58 Starting Set Zero Baseline.

14:32:58 Starting Isocratic Flow. Duration : 27:12, 100% Buffer B, 0% Buffer A flow: 1.25 ml/min

14:32:58 Isocratic Flow: Start gradient pump.

14:32:58 Running...

15:00:10 Starting Set Zero Baseline.

15:00:10 Starting Load/Inject Sample. Sample : Sample, Volume : 1.00 ml, flowrate : 1.00 ml/min

15:00:10 Sample Load: Move AV-73 to inject position.

15:00:10 Sample Load: Start gradient pump.

15:01:10 Starting Set Zero Baseline.

15:01:10 Sample Load: Move AV-73 to load position.

15:01:10 Starting Isocratic Flow. Duration : 27:12, 100% Buffer B, 0% Buffer A flow: 1.25 ml/min

15:01:10 Isocratic Flow: Start gradient pump.

15:28:22 Starting Linear Gradient. Duration: 02:50:00, %B grad: 100->0, flow: 1.00 ml/min

15:28:22 Linear Gradient: Start gradient pump.

18:18:20 Starting Isocratic Flow. Duration : 27:12, 0% Buffer B, 100% Buffer A flow: 1.25 ml/min

18:18:20 Isocratic Flow: Start gradient pump.

18:45:32 Run completed, finishing ...

18:45:34 Run finished.

147

Method Name: hic17.4

Step Number	Start (ml)	Step
1	0.0	Set UV baseline to 0.0
2	0.0	Isocratic Flow with 0% Buffer A, 100% Buffer B at 1.25 ml/min for 34.0 ml
3	34.0	Set UV baseline to 0.0
4	34.0	Static Loop: Inject 1.0 ml sample at 1.00 ml/min for 1.0 min
5	35.0	Set UV baseline to 0.0
6	35.0	Isocratic Flow with 0% Buffer A, 100% Buffer B at 1.25 ml/min for 34.0 ml
7	69.0	Linear Gradient with 100% to 0% Buffer B at 1.00 ml/min for 170.0 ml
8	239.0	Isocratic Flow with 100% Buffer A, 0% Buffer B at 1.25 ml/min for 34.0 ml
9	273.0	End of Protocol

References

d`Albis, A., Gratzer, W.B., (1970) in Companion to Biochemistry. Eds. Bull, A.T., Lagmado, J.R., Thomas, J.O., Tipton, K. F. Longmans, London, pg 166.

Ackland, C. E. , Berndt, W. G., Frezza, J. E., Landgraf, B. E., Pritchard, K.W. Ciardelli, T. L., (1991) Monitoring of protein conformation by high performance size exclusion chromatography and scanning diode array second derivative UV absorption spectroscopy. Jnl. Chrom. **540**, 187 - 98.

Ahern, T.J., Klibanov, A. M., (1985) The mechanism of irreversible enzyme inactivation at 100°C. Science **228**, 1280 - 1284.

Alexandrescu, A.T., Ulrich, E.L., Markley, J.L., (1989) Hydrogen-1 NMR evidence for three interconverting forms of staphylcoccal nuclease: Effects of mutations and solution conditions on their distribution. Biochemistry **28**, 204 -11.

Almond, J. Pattison, J. (1997) Human BSE. Nature 389, 437 - 438.

Amodeo, P., Morelli, M., A., C., Motta, A., (1994) Multiple conformations and proline cis-trans isomerization in salmon calcitonin: A combined nuclear magnetic resonance, distance geometry, and molecular mechanics study. Biochemistry **33**, 10754 - 10762.

Anfinsen, C. B., (1973) Principles that govern the folding of polypeptide chains. Science, Vol. **181** No 4096, 223 - 230.

Ashcroft, A. E., (2000) An Introduction to Mass Spectrometry. http://www.astbury.leeds.ac.uk/Facil/MStut/mstutorial.html

Bai, Y., (1999) Kinetic evidence for an on-pathway intermediate on the folding pathway of cytochrome c . Proc. Natl. Acad. Sci. USA **96**, 477-480.

Bai, Y., (2000) Kinetic evidence of an on-pathway intermediate in the folding of lysozyme. Protein Science **9**, 194 -6.

Baldwin, R. L., (1995) The nature of protein folding pathways. Journal of Biological NMR, **5**, 103 - 109.

Bauer, R., Carrotta, R., Rischel, C., Øgendal, L., (2000) Characterisation and isolation of intermediates in ß Lactoglobulin. Heat aggregation at high pH. Biophys. Jnl. **79**, 2 1030 - 8.

Bierzynski, A., Baldwin, R. L. (1982) Local secondary structure in Ribonuclease A denatured by guanidine hydrochloride near 1°C. Jnl. Mol. Biol. **162** 1 173 -186.

Bollag, D.M., Rozycki, M.D., Edelstein, S. J., (1996) In Protein Methods, Wiley-Liss, Inc. New York, Second Edition, pg 91-93.

Boothe, D.R., Sunde, M., Bellotti,V., Robinson, C.V., Hutchinson, W.L., Fraser, P.E., et al. (1997) Instability, unfolding and aggregation of human lysozyme variants underlying amyloid fibrilogenesis. Nature, **385**, 787-793.

Brandts, J., F., Lin, L., N., (1986) Proline isomerization studied with proteolytic enzymes. Met. In Enzym. Vol. **131**, 107 - 126.

Brockwell, D. J., Smith A.D. Radford, S.E., (2000) Protein folding mechanisms: new methods and emerging ideas. Curr. Opin. Struct. Biol. **10**, 16 - 25.

Burke, M. J., Rougvie, M. A., (1972) Cross β protein structures: Insulin fibrils. Biochemistry 11, 13, 2435 - 9.

Chazin, W. L., Kordel, J., Drakenburg, T., Thulin, E., Brodin, P., Grundström, T., Forsén, S., (1989) Proline isomerization leads to multiple folded conformations of of Calbindin D_{9K}: Direct evidence from two-dimensional ^1H NMR Spectroscopy. Proc. Natl. Acad. Sci. USA **86** 2195 - 2198.

Chung, E. W., Nettleton, E. J., Morgan, C. J., Groß, M., Miranker, A., Radford, S. E., Dobson, C, M., Robinson, C. V. (1997) Hydrogen excahnge properties of proteins in native and denatured states monitored by mass spectrometry and NMR. Protein Science **6** 1316 - 24.

Cohen, F. E., (1999) Protein misfolding and prion diseases. Jnl. Mol. Biol. **293**, 313-20.

Cooper, A. (1999) In Proteins: A Comprehensive Treatise. Ed. Allen, G. JAI Press Inc. Vol. 2 Chp. 6 Thermodynamics of Protein Folding and Stability.

Cooper, A., Johnson, C. M. (1994) In Methods in Molecular Biology,:Microscopy, Optical Spectroscopy, and Macroscopic Techniques. Ed. Jones,C., Mulloy, B., Thomas, A. H Humana Press Inc, Totowa NJ. Vol 22:Chp 10 Differential Scanning Calorimetry.

Cooper, A., Johnson, C. M. (1994) In Methods in Molecular Biology,:Microscopy, Optical Spectroscopy, and Macroscopic Techniques. Ed. Jones,C., Mulloy, B., Thomas, A. H Humana Press Inc, Totowa NJ. Vol 22:Chp 9 Introduction to Microcalorimetry and Biomolecular Energies.

Creighton, T., E. (1980) Kinetic studies of protein unfolding and refolding using urea gradient electrophoresis. J. Mol. Biol. **137**, 61 - 80.

Creighton, T., E. (1983) An empirical approach to protein conformational stability and flexibility, Biopolymers **22**, 49 - 58.

Creighton, T., E. (1991) Stability of folded conformations, Curr. Opin. Struct. Biol. **1**, 5 - 16.

Creighton, T. E., (Ed.) (1992) Protein Folding , W.H. Freeman and Company.

Creighton, T., E., (1996) Proteins : Structure, Fuction and Properties, fourth edition. W. H. Freeman and Company.

Denisov, V. P., Jonsson, B., Halle, B., (1999) Hydration of denatured and molten globule proteins. Nature: Structure Biology **6**, 253 -60.

Dill, K. A., (1999) Polymer principles and protein folding. Protein Science **8**, 1166 - 80.

Dill, K. A., Chan, H. S. (1997) From Levinthal to pathways to funnels. Nature Structural Biology **4**, 10 - 19.

Dill, K.A. Shortle, D. (1991) Denatured states of proteins. Annu. Rev. Biochem. **60**, 795 - 825.

Drake, A.F. (a) (1994) In Methods in Molecular Biology,:Microscopy, Optical Spectroscopy, and Macroscopic Techniques. Ed. Jones,C., Mulloy, B., Thomas, A. H Humana Press Inc, Totowa NJ. Vol 22:Chp 13 The Measurement of \electronic Absorption Spectra in the Ultraviolet and Visible.

Drake, A.F. b (1994) In Methods in Molecular Biology,:Microscopy, Optical Spectroscopy, and Macroscopic Techniques. Ed. Jones,C., Mulloy, B., Thomas, A. H Humana Press Inc, Totowa NJ. Vol 22:Chp 16 Circular Dichroism.

Eftink, M.R., Ghiron, C.A. (1976) Exposure of tryptophanyl residues in proteins. Quantitative determination by fluorescence quenching studies. Biochemistry, **15**, 672-680.

Eftink, M.R., Ghiron, C.A., (1984) Indole fluorescence quenching studies on proteins and model systems: Use of the inefficient quencher succinimide. Biochemistry, **23**, 3891-3899.

Englander, S. W., (2000) Protein folding intermediates and pathways studied by hydrogen exchange. Annu. Rev. Biophys. Biomol. Struct. **29**, 213 -238.

Evans, P., A., Dobson, C., M., Kautz, R., A., Hatfull, G., Fox, R., O., (1987) Proline isomerization in staphylococcal nuclease characterised by NMR and site directed mutagenesis. Nature Vol. **329**, 266 - 268.

Eyles, S. J., Gierasch, L. M., (2000) Multiple roles of prolyl residues in structure and folding. Jnl. Mol. Biol. **301** (3) 737 - 747.

Fink, A. L., (1998) Protein Aggregation: folding aggegates, inclusion bodies and amyloid. Folding and Design Vol. **3** 9-23.

Fischer, G., (1994) Peptidyl-prolyl isomerases and their effectors. Angew. Chem. Int. Ed. Engl. 33, 1415 - 1436.

Fischer, G., Bang, H., (1985) The refolding of urea denatured ribonuclease A is catalysed by peptidyl-prolyl cis -rans isomerase. BBA **828**, 39 - 42.

Fischer, G., Schmid, F. X., (1990) The mechanism of protein folding. Biochemistry, Vol. **29** No. 9, 2205 - 2212.

Fischer, G., (2000) Chemical aspects of peptide bond isomerization. Chem. Soc. Rev. **29** (2), 119 - 127.

Fox, R., O., Evans, P., A., Dobson, C., M., (1986) Multiple conformations of a protein demonstrated by magnetization transfer NMR spectroscopy. Nature **320**, 192 - 194.

Garriga, P., Khorana, H. G., Liu, X., (1996) Structure and function in rhodopsin. Proc. Natl. Acad. Sci. **93**, 4560 -4564.

Gething, M., (1997) The difference with prokaryotes. Nature, **388**, 329 - 331.

Gething, M., Sambrook, J., (1997) Protein folding in the cell. Nature **355**, 33 - 45.

Gill, S.C., von Hippel, P.H., (1989) Calculation of protein extinction coefficients from amino acid sequence data. Analytical Biochemistry, **182**, 319-326.

Greuble, M., Sabelko, J., Baliew, R., Elvin, J., (1998) Laser temperature jump induced protein refolding. Acc. Chem. Res. **31**, 699 -707.

Hammerström, P., Carlssom, U., (2000) Is the unfolded state the Rosetta stone of the protein folding problem. Biochem. Biophys. Res. Comm. **276**, 393 - 8.

Harper, J. D., Lansbury, P. T., (1997) Models of amyloid seeding in Alzheimers disease and scrapie: Mechanistic truths and physiological consequences of the time dependent solubility of amyloid proteins. Annu. Rev. Biochem. **66**, 385 - 407.

Honig, B., (1999) Protein Folding: From the Levinthal paradox to structure prediction. Jnl. Mol. Biol. **293**, 283 - 93.

Hubbard, S.J., (1998) The structural aspects of limited proteolysis of native proteins (Review). Biochemica Biophysica Acta **1382**, 191 -206.

Ibarra-Molero, B., Sanchez-Ruiz, J., M., (1997) Are there equilibrium intermediate states in the urea-induced unfolding of hen egg white lysozyme. Biochemistry **36**, 9616 - 9624.

Ikura, T., Tsurupa G. P., Kuwajima, K., (1996) Kinetic folding and cis/trans prolyl isomerization of staphylcoccal nuclease. A study by stopped flow absorption, stopped flow CD, and molecular dynamics simulations. Biochemistry **36**, 6529 - 6538.

Itzhaki,L.S., Otzen, D.E., Fersht,A., (1995) The structure of the transition state for folding of chymotrypsin inhibitor 2 analysed by protein engineering methods: Evidence for a nucleation condensation mechanism. Jnl.. Mol. Biol. **254**, 260-88.

Itzhaki, L. S., Evans, P. A., Dobson, C. M., Radford, S. E., (1994) Tertiary interactions in the folding pathway of HEWL: Kinetic states using fluorescent probes. Biochem. **33**, 5212 - 5220.

Jabs, A., Weiss, M. S., Hilgenfield, R., (1999) Non proline cis peptide bonds in proteins. Jnl. Mol. Biol. **286**, 291 -304.

Jaenicke, R., (2000) Stability and stabilization of globular proteins in solution. Jnl. Biotech. **79**, 193 – 203.

Jaenicke, R. , (1995) Folding and association versus misfolding and aggregation of proteins. Phil. Trans. Royal. Soc. Lon. B, **348**, 97 - 105.

Jaenicke, R. , (1996) How do proteins aquire their three dimensional structure and stability. Naturwissenschaftenaufsatze, **83**, 544 -554.

Jaenicke, R. , Rudolph, R. (1989) Protein Structure: a practical approach. Ed. Creighton, T. E., IRL Press at OUP. Chapter 9 Folding Proteins.

Johnstone R. A.W., Rose, M.E., (1996) Mass Spectrometry for Chemists and Biochemists, second edition, Cambridge University Press.

Jones, S., Reader J. S., Healy, M., Capaldi, A. P., Ashcroft, A. E., Kalvedra, A. P., Smith, D. A., Radford, S. E., (2000) Partially unfolded species populated during equilibrium denaturation of the ß sheet protein Y74W apo pseudoazurin. Biochemistry **39**, 5672 -5682.

Kato, S., Shimamoto, N., Utiyama, H., (1982) Identification and characterisation of the direct folding process of hen egg white lysozyme. Biochemistry Vol. **21**, 38-43.

Kato, S., Okamura, M., Shimamoto, N., Utiyama, H., (1981) Spectral evidence for a rapidly formed intermediate in the refolding kinetics of hen egg white lysozyme. Biochemistry **20**, 1080 - 1085.

Katta, V., Chait, B. T., (1991) Conformational changes inproteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. Rapid Communications in Mass Spectrometry **5** 214 - 217.

Kelly, S.M., Price, N.C., (1997) The application of circular dichroism to studies of protein folding and unfolding. B.B.A.,**1338**, 161-185.

Klunk, W. E., Pettigrew, J.W., Abraham, D. J., (1989) Quantitative evaluation of congo red binding to amyloid-like proteins with β-sheet conformation. Jnl. Histo. Cyto. **37**, 1273 - 1282.

Kulkarni, S. K., Ashcroft, A. E., Carey, M., Masselos, D., Robinsone, C. V., Radford, S. E., (1999) A near native state on the slow refolding pathway of hen lysozyme. Protein Science **8**, 35 -44.

Last, A. M., Robinson, C. V., (1999) Protein folding and interactions revealed by mass spectrometry. Curr. Opin. Struct. Biol. Vol. **3**, 5, 564 - 70.

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

Lanchester, J., (1996) A new kind of contagion. The New Yorker December 5, 70 - 81.

Lang, K., Schmid, F.X., Fischer, G., (1987) Catalysis of protein folding by prolyl isomerase. Nature **329**, 268 -270.

Langsetmo, K., Fuchs, J., Woodward, C., (1989) *Escherichia coli* thioredoxin folds into two compact forms of different stability to urea denaturation. Biochemistry **28**, 3211 - 20.

LeVine, H., (1995) Thiflavin T interaction with amyloid β sheet structures. Amyloid: Int. Jnl. Exp. Clin. Inv. **2**, 1 - 6.

Levinthal, C., (1968) Are there pathways for protein folding. J. Chim. Phys. **65**, 44 - 45.

Lin, L. N., Brandts, J. F., (1983) Determination of cis-trans proline isomerization by trypsin proteolysis: application to a model penta-peptide and to oxidised ribonuclease A. Biochemistry **22**, 553 - 559 and 559 - 563.

Lin, L. N., Hasumi, H., Brandts, J. F., (1988) Catalysis of proline isomerization during protein folding reactions. Biochim. Biophys. Acta **956**, 256 -266.

Liu, X., Garriga, P., Khorana, H. G., (1996) Structure and function in rhodopsin. Proc. Natl. Acad. Sci. **93**, 4554 -4559.

Lu, P., Wulf, G., Zhou, X. Z., Davies, P., Lu, K. (2000) The prolyl isomerase Pin 1 restores the function of Alzheimers associated phosphorlyated tau protein. Nature **399**, 784 - 8.

Maki, K., Ikura, T., Hayano, T., Takahashi, N., Kuwajima, K., (1999) Effects of proline mutations on the folding of staphylococcal nuclease. Biochemistry **38**, 2213 - 2223.

Matagne, A., Radford, S.E., Dobson, C. M. (1997) Fast and slow tracks in lysozyme folding; Insight into the roles of domains in the folding process. J. Mol. Biol. Vol. **267**, 1068-1074

Matagne, A., Dobson, C. M., (1998) The folding process of hen lysozyme: a perspective from the new view. Cell. Mol. Life Sci. **54**, 363 -71.

Matouschek, A., Rospert, S. Glick, B.S., Schatz, G., (1995) Cyclophilin catalyses protein folding in yeast mitochondria. Proc. Natl. Acad. Sci. USA **92**, 6319 - 6323.

MedicineNet Alzheimers Disease, http://www.medicinenet.com

Microcal Inc., DSC data analysis in Origin™.

Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T., Dobson, C. M., (1993) Detection of transient protein folding populations by mass spectrometry. Science **262**, 896 - 900.

Mitraki, A., King, J., (1989) Protein folding intermediates and inclusion body formation. Biotechnology **7** 690-697.

Mitraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J., King, J., (1991) Global suppression of protein folding defects and inclusion body formation. Science **253**, 54 - 58.

Mücke, M., Schmid, F. X., (1992) Enzymatic catalysis of prolyl isomerasation in an unfolding protein. Biochemistry **31**, 7848 - 54.

Nanjo, F., Sakai, K., Usui, T., (1988) p-Nitropheyl penta-N acetyl-β chitopentaoside as a novel synthetic substrate for the colourimetric assay of lysozyme. J. Biochem. , Vol. **104** , 255 - 258.

Netzer, W. J., Hartl, F. U. (1997) Recombination of protein domains facilitated by cotranslational folding in eukaryotes. Nature, **388**, 343 - 349.

New England Biolabs, (1993) Protein Fusion & Purification System : Expression and Purification of Protein from Cloned Genes, Instruction Manual.

Nozaki, Y., (1972) The preparation of Guanidinium hydrochloride. Methods in enzymology Vol. **26**, 43 - 50

Page, A., MacNiven, K., Hengartner, M. O., (1996) Cloning and biochemical characterisation of the cyclophilin homologues from the free living nematode caenorhabiditis elegans. Biochem. Jnl. **317**, 179 - 85.

Parker, T. G., Dalgleish, D. G., (1977) The use of light scattering and turbidity measurements to study the kinetics of extensively aggregating proteins: - $\alpha_s$casein. Biopolymers **16**, 2533 - 47.

Perrett, S., (1998) Mishapes and misfits: protein misfolding and disease. Chemistry and Industry, 18/05/98, 389 - 393.

Pharmacia a Hydrophobic Interaction Chromatography - Principles and methods.

Pharmacia b Phastsystem™ owners manual.

Phillips, D. C. , (1966) The three dimensional structure of an enzyme molecule. Sci. Amer. **215**, 65, 78 - 90.

Price, N.C., Johnson, C.M., (1990) In Proteolytic Enzymes - a practical approach, Second Ed., Ed. Reynon, R.J., Bond, J.S., IRL Press at OUP, Eynsham, England.: Chp 9. Proteinases as probes of conformation of soluble proteins.

Privalov, P. L., (1996) Intermediate states in protein folding. Jnl. Mol. Biol., **258**, 707 - 725.

Provencher, S. W., Glockner, J., (1981) Estimation of globular protein secondary structure from circular dichroism. Biochemistry **20**, 33 - 37.

Prusiner, S. B., (1997) Prion diseases and the BSE crisis. Science **278**, 245 - 251.

Ptitsyn, O.B., Pain, R.H., Semisotnov, G. V., Zerovnik, E., Razgulyaev, O. I. (1990) Evidence for a molten globule state as a general intermediate in protein folding. FEBS **262**, 20 - 24.

Ptitsyn, O.B., Bychkova, V. E., Uversky, V. N. (1995) Kinetic and Equilibrium Folding Intermediates. Phil. Trans. Biol. Sci. **348**, 1323, 35 - 41.

Radford, S. E., Dobson, C. M., Evans, P. A., (1992) The folding of HEWL involves partially structured intermediates and multiple pathways. Nature **358**, 302 -7.

Radford, S.E., Dobson, C. M. (1995) Insights into protein folding using physical techniques: studies of lysozyme and β-lactoglobulin. Phil. Trans. R. Soc. Lond. B **348**, 17-25.

Ranson, N. A., Dunster, N. J., Burston, S. G., Clarke, A. R., (1995) Chaperonins can catalyse the reversal of early aggregation steps when a protein misfolds. Jnl. Mol. Biol. **250**, 581 - 6.

Righetti, P. G., (1989) Isoelectric focusing of proteins in conventional and immobilized pH gradients. In Protein Structure a practical approach, Ed. T.E. Creighton. IRL Press at O.U.P., second edition, Chp 2.

Robinson, C.V., Radford, S.E., (1995) Weighing the evidence for structure : electrospray mass spectrometry of proteins. Structure, **3**, 861-865.

Sackett, D. L., Wolff, J., (1987) Nile red as a polarity sensitive fluorescent probe of hydrophobic protein surfaces. Ana. Biochem., **167** , 228 - 234.

Saunders, C.R., Nagy, J.K., (2000) Misfolding of membrane proteins in health and disease: the lady or the tiger. Curr. Opin. Struct. Biol. **10**, 438 -442.

Schindler; T.; Herrler; M.; Marahiel; M.A.; Schmid; F.X. (1995) Extremely rapid protein folding in the absence of intermediates. Nature Struct. Biol.; **2** No.8.

Schmid, F. X., Baldwin, R. L., (1978) Acid catalysis of the formation of the slow folding species of Rnase A: Evidence that the reaction is proline isomerization. Proc. Natl. Acad. Sci. USA, **75**, No. 10, 4784 - 4788.

Schmid, F. X., Blaschek, H., (1981) A native like intermediate on the ribonuclease A folding pathway.2/ Comparison of its properties to native Rnase A. Eur. Jnl. Biochem. **114**, 111 - 117.

Schmid, F.X., (1981) A native like intermediate on the ribonuclease A folding pathway 1/ Detection by tyrosine fluorescence changes. Eur. J. Biochem. **114**, 105 - 109.

Schmid, F.X., (1993) Prolyl iomerase: Enzymatic catalysis of slow protein folding reactions. Annu. Rev. Biophys. Biomol. Struct. **22**, 123 - 43.

Schmid, F. X., (1999) Protein Folding. In Proteins: A comprehensive treatise. Ed. Allen, G. JAI Press Inc. Vol. 2 Chp. 5 Protein Folding.

Scopes, R., K., (1994) Protein Purification principles and practice, Springer -Verlag New York, Inc., third edition, pg. 126 - 132.

Seckler, R., Jaenicke, R., (1992) Protein folding and protein refolding. FASEB Jnl **6**, 2545 - 2552.

Seielstad, D.A., Carlso, K.E., Kushner, P.J., Greene, G.L., Katzenellenbogen, J.A., (1995) Analysis of the Structural Core of the Human Estrogen Receptor Ligand Binding Domain by Selective Proteolysis/ Mass Spectrometric Analysis. Biochemistry **34**, 12605 - 12615.

Shastry, M. C. R., Sandler, J. M., Roder, H., (1998) Submillisecond folding events in cytochrome c. Acc. Chem. Res. **31,** 717 - 25.

Sosnick, T.R., Mayne, L. E., Hiller, R., Englander, S.W., (1994) The barriers in protein folding. Nature Struct. Biol., **1** No.3, 149-156.

Stewart, D. E., Sarker, A., Wampler, J.E., (1990) Occurence and role of cis peptide bonds in protein structures. J. Mol. Biol. **214**, 253 - 200.

Stryer, L. H., (1995), Biochemistry, W.H. Freeman and Company New York, Chp. 2, 7, 10.

Svensson, L.A.,Thulin, E., Forsen, S., (1992) Proline cis-trans isomers in calbindin D_{9K} observed by X-ray crystallography. J. Mol. Biol. **223**, 601 - 606.

Tanford, C. (1968) Protein Denaturation. Adv. Protein Chem. **23** 122 - 275.

Taubes, G., (1996) Misfolding the way to disease. Science **271**, 1493 - 1495.

Thomas, B. R., Vekilov, P. G., Rosenberger, F., (1996) Heterogeneity , determination and purification of commercial HEWL. Acta Cryst. **D52**, 776 - 84.

Truckses, D., M., Somoza, J., R., Prehoda, K., E., Miller, S., C., Markley, J., L., (1996) Coupling between trans/cis proline isomerisation and protein stability in staphylcoccal nuclease. Protein Science **5**, 1907 - 1916.

Tsai, C., Kumar, S., Ma, B., Nussinov, R. (1999) Review : Folding funnels , binding funnels and protein function. Protein Science, **8** , 1181 - 1190.

Vanhoof, G., Goossens, F., DeMeester, I., Hendriks, D., Scharpe, S., (1995) Proline motifs in peptides and their biological processing. FASEB Jnl. **9** 736 - 744.

Varley, P. J. (1994) In Methods in Molecular Biology,:Microscopy, Optical Spectroscopy, and Macroscopic Techniques. Ed. Jones,C., Mulloy, B., Thomas, A. H Humana Press Inc, Totowa NJ. Vol 22:Chp 15

Varley, P. G., Brown, A. J., Dawkers, H. C., Burns, N. R., (1997) A case study for the study and the use of sedimentation equilibrium analytical ultracentrifugation as a tool for biopharmaceutical development. Eur. Biophys. Jnl. **25**, 437 - 43.

Voge, G., (1998) Possible new cause of Alzheimers disease found. Science **279**, 174.

Walkenhorst; W.F.; Green; S.M.; Roder; H.; (1997) Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease. Biochemistry **36**, 5795-5805.

Wanker, E. E., (2000) Protein aggregation in Huntingtons and Parkinson`s disease: implications for therapy. Mol. Med. Today **6** (10), 387 -91.

Welsh, M. J., Smith A. E., (1995) Cystic Fibrosis Scientific American, December issue.

Wetzel, R., (1996) For protein misassembly its the "I" decade. Cell **86**, 699 -702.

Winkler, J.R; Gray, H.B.; (Ed.) (1998) Protein Folding (Special Issue) Acc. Chem. Res. Vol. **31** No. 11.

Wisniewski, T., Golabek, A. A., Kida, E., Wisniewski, K. E., Frangione, B., (1995) Conformational mimicry in Alzheimers disease. Am. Jnl. Of Path. **147** No. 2, 238 - 243.

Wolynes, P. G., Onuchic, J. N., Thirumalai, D. (1995) Navigating the folding routes. Science **267**, 1619 - 1620.

Wood, S. J., Maleef, B., Hart, T., Wetzel, R., (1996) Physical, morphological and functional difference between pH 5.8nand 7.4 aggregates of the Alzheimers amyloid peptide Aβ. J. Mol. Biol. **256**, 870 -877.

Woodward, C., Simon, I., Tuchsen, E. (1982) Hydrogen exchange and the dynaimc structure of proteins. Mol. Cell. Biochem. **48** 135 -160.

Worthington Biochemical Company (2000) Lysozyme. http://www.worthington-biochem.com/manual/L/LY.html

Yaron, A. Naider, F. (1993) Proline dependent structural and biological properties of peptides and proteins. Crit. Rev. Biochem. Mol. Biol. **28**, 31 - 81.

Yeh, S., Han, S., Rousseau. D. L., (1998) Cytochrome c folding and unfolding: A biphasic mechanism. Acc. Chem. Res. **31**, 727-30.

Yuan, X., Werner, J.M., Knott, V., Handford, P. A. Campbell, I. D., Downing K., (1998) Effects of proline cis-trans isomerization on TB domain secondary structure. Protein Science **7**, 2127 - 2135.

Zale, S. E., Klibanove, A. M., (1986) Why does Ribonuclease irreversibly inactivate at high temperatures. Biochemistry **25**, 5432 - 5444.