Intermediate Filament Protein Assembly
A Proteomic Approach

Thesis submitted in partial fulfilment of the requirements

for the degree of

Doctor of Philosophy

at the

University of Canterbury

by

Louise Nancy Paton

2005

Table of Contents

Intermediate Filament Protein Assembly
A Proteomic Approach

Intermediate filament proteins (IFPs) form the main structural elements of a wool fibre. The IFPs of wool are comprised of two families; the acidic type I family and the neutral-basic type II family. During follicle development, one type I and one type II IFP develop into an obligate heteropolymer, which, through a series of associations with other heteropolymers, forms an intermediate filament.

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) methods have been used to provide high-resolution separation of wool IFPs. Improvements in the method for maintaining reducing conditions and chaotrope constitution, combined with low %T polyacrylamide gels, allowed the high-resolution separation of the two keratin IFP families and their individual family members. The IFPs were separated to produce a clearly defined spot pattern, with numerous discrete minor spots not previously observed.

Genomic studies have reported that there are eight genes which produce eight abundant IFPs in wool. It was hypothesised that the large number of additional spots seen on a 2D-PAGE gel was due to post-translational modification (PTM) of the protein. Several common PTMs of proteins produce charge heterogeneity, including phosphorylation and glycosylation. However, analysis of wool IFPs by 2D-PAGE techniques and mass spectrometry revealed no evidence of phosphorylation or glycosylation modifications.

Conformational equilibria as a cause of protein charge heterogeneity has recently been reported. Investigations with both the type I and type II IFPs have shown that when single protein spots from a 2D-PAGE separation are eluted, re-focused and re-electrophoresed, several spots are formed on both the acidic and basic side of the original spot. The cause of this heterogeneity is thought to be a conformational equilibrium between several different forms of the same protein in the rehydration solution used for the first dimension. This technique allowed the accurate assignment of IFPs resolved by 2D-PAGE to protein families.

Fractionation methods to separate the IFPs and intermediate filament associated proteins (IFAPs) were successfully developed. Further fractionation into the type I and type II IFPs was achieved along with partial success at isolating individual spots. *In vitro* assembly experiments with the different IFP families gives important information about the strength of different protein pairings. To date there are no reproducible, efficient, *in vitro* assembly conditions for keratinised wool IFPs. A comprehensive study to investigate assembly conditions for keratinised wool IFPs was undertaken.

Assembly of filaments from IFPs was achieved after a partial digestion with chymotrypsin. Filaments were formed that varied in diameter from 10 to 40 nm, showing that higher ordered structures were being formed. This demonstrates that IFPs can be successfully assembled *in vitro* to form filamentous structures that may be able to be manipulated for biomaterial uses.

Abbreviations

% T	% of total monomer in solution (acrylamide + bis), w/v
% C	% of cross-linker from total cross-linker and monomer, w/w
1DE	One-dimensional electrophoresis
2DE	Two-dimensional electrophoresis
2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
Å	Ångström
A	Ampere
A_{340}	Absorbance at 340 nm
APS	Ammonium peroxodisulphate
cDNA	Circular deoxyribonucleic acid
CE	Collision energy
CHAPS	3-[(3-Cholamidopropyl)-dimethylammonio]-1-propanesulfonate
Da	Dalton
DCM	Dichloromethane
DTT	Dithiothreitol
EDTA	Ethylenediamine-tetraacetic acid
EBS	Epidermolysis bullosa simplex
ER	Endoplasmic reticulum
EtOH	Ethanol
g	Grams
GFAP	Glial fibrillary acidic protein
GMP	Glycomacropeptide
HAc	Acetic acid
HCl	Hydrochloric acid
HGTP	High glycine tyrosine protein
HPL	Human pancreatic lipase
HSP	High sulfur protein
IAM	Iodoacetamide
IEF	Isoelectric focusing
IF	Intermediate filament

IFAP	Intermediate filament associated protein
IFP	Intermediate filament protein
IMAC	Immobilised metal affinity chromatography
IPA	Isopropanol
IPG	Immobilised pH gradient
IPG-IEF	Immobilized pH gradient - isoelectric focusing
KRT	Keratin
KRTAP	Keratin associated protein
L	Litre
m	Metre
MALDI-TOF	Matrix assisted laser desorption ionisation-time of flight
MeOH	Methanol
mRNA	Messenger ribonucleic acid
ms/ms	Tandem mass spectrometry
MWCO	Molecular weight cut off
m/z (amu)	Mass to charge ratio
NEM	N-Ethylmaleimide
NMR	Nuclear magnetic resonance
O-GlcNAc	O-linked N-acetyl glucosamine
PAGE	Polyacrylamide gel electrophoresis
PDI	Protein disulfide isomerase
pI	Isoelectric point
PITC	Phenylisothiocyanate
pK_a	Acid dissociation constant
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
Q1	First quadrupole
Q2	Second quadrupole
rpm	Revolutions per minute
SCBP	Stratum corneum basic protein
SCMK	S-Carboxymethyl keratins
SB3-10	*N*-Dimethyl-3-ammonio-1-propane-sulfonate
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis

TBP	Tributylphosphine
TBST	Tris buffered saline solution containing Tween™20
TCEP	Tris (2-carboxyethyl) phosphine hydrochloride
TEM	Transmission electron microscopy
TEMED	NNN'N'-Tetramethylethylenediamine
TFMS	Trifluromethanesulfonic acid
ThT	Thioflavin T
TLCK	N-Tosyl-L-lysine-chloromethyl ketone
TPCK	Tosyl phenylalanyl chloromethyl ketone
Tris	Tris(hydroxymethly)methylamine
UHSP	Ultra high sulfur protein
ULF	Unit-length filaments
V	Volts
W	Watts

Acknowledgements

I would like to start by thanking Carl, my husband, for his love and support right through my postgraduate studies. Also to my Mum and Dad and my ever-increasing family, who are a constant reminder to me of what the truly important things in life are. Special thanks go to my parents for supporting me through my studies in all sorts of ways.

I was encouraged by Warren to go into PhD studies and he has pushed me to do my best right through. He has done an excellent job in securing funding for the proteomic group at Canesis, resulting in one of the best-equipped proteomic labs in NZ. It has been a privilege to work with such a fantastic lab set-up. The Fundamental Sciences group (and the old Platform Sciences group) at Canesis has been a great source of support with everyone willing to help and offer advice. The group would listen to me grumble when things weren't working, or I needed to let off some steam and were happy for my successes. Here's to more morning teas at Mrs O's! Special thanks to James for TEM assistance and Nigel and Scott for help with the mass spectrometry.

On the University side of things, Juliet has been a fantastic motivator and is constantly enthusiastic (especially when I needed it the most). I'm glad that I had the opportunity to have Juliet as my supervisor and to be part of the 'purple bubble' philosophy. Thanks to the 'purple bubble' group members, past and present, who helped make me feel like a real student for at least one hour of every week.

Finally, I must thank Canesis for supporting me throughout my PhD and FRST for giving me a Bright Future scholarship.

Chapter One

Introduction

This thesis sets out to characterise the intermediate filament (IF) properties of wool and characterise their assembly properties *in vitro* with a view to designing novel biomaterials. The IFs are the main structural elements of a wool fibre [1]. The IFs are made up of many monomeric intermediate filament proteins (IFPs), which associate with one another to form IFs. The classification, identification and assembly properties of wool IFPs have been partially studied, but are yet to be fully understood. Many of the important properties of the wool fibre are determined by the composition of its proteins, and how they interact.

1.1 Intermediate filaments in eukaryotic cells

Intermediate filaments are a superfamily of 10 nm fibres, ubiquitous in multicellular eukaryotes [2]. IFs were named due to their diameter; they are intermediate in size between microtubules (~25 nm) and microfilaments (~5-7 nm) [3-5]. The IFs differ from the microtubules and microfilaments in a number of ways. Several classes of cell-specific IFs exist, whereas microtubules and microfilaments are widely distributed. The IFs exist mainly in the polymerised form and the microtubules and microfilaments exist in a dynamic equilibrium with a large soluble pool of monomers [5]. There also appears to be no energy needed for the formation of IFs *in vivo*, which suggests relatively simple assembly dynamics compared to the complex formation of the two other structural systems [5]. The subunits of microfilaments and microtubules are globular proteins, which form a 'string', with the IFs forming rope-like polymers that are able to be denatured then renatured to form filaments indistinguishable from native filaments.

Initially, the IFs were considered static structures that merely kept cells rigid [6]. It has now been shown that IFs are able to form dynamic networks *in vivo*, which are able to assemble and disassemble within the cell and interact with other cellular organelles [3]. Intermediate filaments perform many functions within cells. Specifically, they connect the spot desmosomes of epithelial cells and stabilise the epithelium; form the major structural proteins of skin and hair; form a scaffold that holds the Z discs and myofibrils in place in muscle cells; and give strength and rigidity to nerve axons.

In higher vertebrates, there are six major classes of IFs. The IFs are grouped relative to their homologies within their α-helical domains. The types I and II come from the 'hard' and 'soft' α-keratins, the type IIIs include vimentin, desmin, and glial fibrillary acidic protein (GFAP), the type IVs are expressed in neuronal tissues, the type Vs are the nuclear lamins and a single type VI has been identified [7]. All of the IFPs are located in the cytoplasm except for the type Vs, which are located within the nucleus [8].

The type I and II IFPs, the keratins, are the most complex group of IFPs. The type I IFPs are slightly acidic (pI = 4-6) and the type II IFPs are neutral/basic (pI = 6-8). Assembled keratin filaments have distinct physical properties, suggesting that differential expression and specific keratin pairing may tailor networks to suit tissue-specific structural requirements of tensile strength, flexibility and dynamics [2].

In humans, six different pathologies have been identified in IFs which affect IF structure and/or function *in vivo* [3]. Mutations, which cause amino acid substitutions, limit the normal dynamic behaviour of IFs and can cause severe disruption to an entire tissue. Without a proper IF network, cells become fragile and prone to breakage [2]. The most widely known IF disease is epidermolysis bullosa simplex (EBS). It is a rare genetic skin disease which effects 1 in 50 000 people [2]. Typically, it causes blistering due to cytolysis within the basal layer, with the most severe forms showing at birth as blistering of the epidermis and oral mucosa. Point mutations have been identified in EBS patients; the most severe forms of the disease involve

mutations which occur in the domains that are involved in multiple contact points within IFs. Assays using bacterially expressed proteins with point mutations show a correlation between the degree to which IF assembly is disturbed and the severity of the disease [2].

1.2 Evolution of IFPs

All IFPs discovered so far share a common structure, and are thought to have a common evolutionary origin. The first firm evidence that hard α-keratins belonged to the IF family came from a study by Dowling, Parry and Sparrow [9]. They determined the primary sequence of two wool keratins and compared them to desmin and vimentin sequences. The sequences showed marked homology and provided evidence for all IFPs having similar structures and a common evolutionary origin.

Immunological, peptide map data and polypeptide gel patterns show a very high similarity in IFPs between bovine, ovine and human hairs. This suggests that there was a very high degree of conservation of keratin IFPs during the evolution of higher mammals [10]. This high degree of similarity is also shown in nail IFPs, suggesting that the eight trichocyte keratin IFPs are involved in the production of other epidermal appendages [10].

In the human genome there are at least 62 cytoplasmic IFP genes [11]. The teleost fish genome has at least 42 cytoplasmic IFP genes [11]. Genome comparisons between human and teleost fish show that there are some unexpected differences [11]. Teleost fish have a sizable excess of type I IFP genes over type II IFP genes. Four of the type I IFP genes map closely to six of the type II IFP genes, which suggests that a single gene cluster of the type II IFP genes, like the one found in mammals, is unlikely. This suggests that the keratin gene clustering in mammals evolved after the fish ancestry separated from the lineage, which led to higher vertebrates [11]. Further studies on amphibian and avian genomic comparisons will determine when the gene clustering was acquired during vertebrate evolution [11].

It has been postulated that all eukaryotes might have lamins and that the putative primordial IF gene encoded a lamin (a type V IF). Lamins form a fibrous meshwork on the inner surface of the nuclear membrane. This structure seems to provide a framework for the nucleus, and may facilitate chromatin organisation [2].

1.3 Wool as a source of IFs

The major components of the wool fibre are proteins [1]. Approximately 60% of the cortical cells in wool are made up of IFs, together with its surrounding matrix proteins [1]. Intermediate filaments in the hard α-keratins (a subgroup of the type I and II IFPs) are highly ordered and are amenable to analysis by X-ray diffraction, which provides information about structural features [12]. Wool is a plentiful source of hard α-keratin, with NZ producing more than 175 000 tonnes per year [13]. Methods for the extraction of proteins from the wool fibre have been developed over approximately 50 years [14]. Recently, the company Keratec Ltd (Lincoln, NZ) has been developed and commercialised. This is a large-scale plant, which extracts individual protein components from wool for use as biomaterials in the areas of cosmetics, medicine, and nanotechnology. Fundamental research into the structure and function of the wool hard α-keratins can be used to further develop these innovative products, in addition to providing information regarding their structure, function and assembly *in vitro* and *in vivo*.

1.4 Intermediate filament protein structure

All IFPs have a tripartite structure, with a central rod domain flanked by the head and tail domains (Figure 1.1) [2, 15]. The head and tail domains are rich in cysteine residues which are able to form extensive disulfide bonding cross-links with the intermediate filament associated proteins (IFAPs) [16].

Figure 1.1
Tripartite structure of an IFP. Non-helical N (head) and C (tail) terminal regions are separated by α-helical rod domains.

Initial work by Crewther and Dowling, using partial pronase digestion, suggested that the keratin proteins were made up of three units of α-helix linked by sections of nonhelical chain [17]. More recently, it was inferred using standard secondary structure prediction methods, that the rod domain contains four segments named 1A, 1B, 2A and 2B (Figure 1.2) [6, 7].

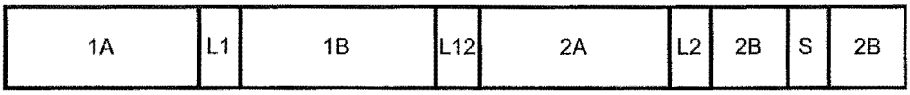


Figure 1.2
Schematic drawing of rod domain segments in an IFP. The S in segment 2B represents the stutter region [3].

Each segment has a heptad repeat motif $(abcdefg)_n$ [6, 7] with the *a* and *d* residues being hydrophobic [2] and the *e* residues being predominantly basic and *g* residues being predominantly acidic [18]. In vimentin, nearly all of the *a* and *d* residues are occupied by leucine and valine [19]. These residues are involved in the bonding of IFPs by hydrophobic and ionic interactions [3].

The four rod segments are separated by three linker regions named L1, L12 and L2 [6, 7] (Figure 1.2). The linker regions allow the rod to be flexible and bend at the links. The L1 linker varies in size (from eight to 16 residues [20, 21]) and sequence within IFs, and is the most flexible rod domain region [20]. The linker L2 has a highly conserved length and sequence [7]. The L1 linker has been modelled, and it has been suggested that it may act as a hinge, when segment 1A is unwound, allowing the head domains to function over a

wide lateral range, meaning they can interact with cellular moieties and function more efficiently [20, 21].

In the middle of the 2B region there is a discontinuity in the heptad repeat motif. This region is commonly called the 'stutter' and is caused by the deletion of three amino acid residues in an otherwise perfect heptad substructure [3] (Figure 1.2). The stutter is highly conserved. A crystal structure of the vimentin 2B region shows that local unwinding of the coiled-coil in the 2B domain compensates for the stutter, and has a global effect on dimer structure (Figure 1.3) [19].

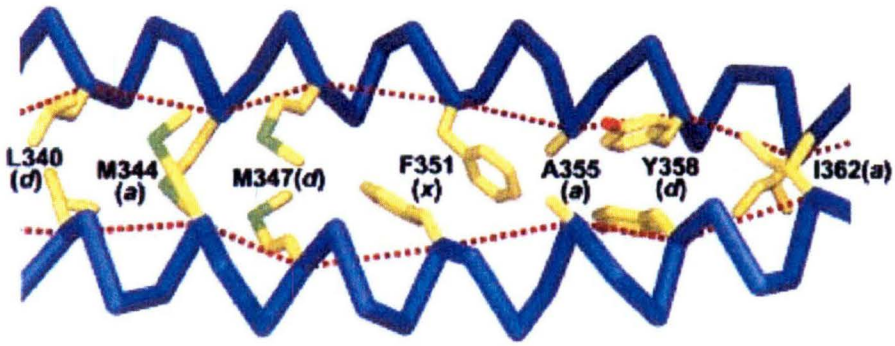


Figure 1.3
Hydrophobic core organisation near the stutter in a dimer. The letters in parentheses after the residue number indicate the heptad position [19].

Removal of the stutter, by inserting three amino acids, stops assembly at the unit-length fragment (ULF) stage, suggesting that the stutter has structural and/or functional significance [22]. The stutter contributes to the relative azimuthal orientation of functionally relevant groups in the N and C terminal regions. Distortion of these orientations may adversely affect the correct intermolecular interactions which occur during the annealing steps to form IFs from ULFs [19]. Genetic linkage analysis methods have shown that, in a mild form of EBS, a single substitution in the 2B region near the stutter changes a leucine to a proline residue. This substitution would be expected to affect the structural integrity of the keratin IFP and its assembly into an IF.

This demonstrates the functional relationship between the observed mutation and the disease phenotype [3].

In epidermal keratins, the N-terminal domain contains the regions E1, V1 and H1. The C-terminal domain has the E2, V2 and H2 domains. The V1 and V2 domains are variable regions that are rich in glycine and serine residues [3]. E1 and E2 are short basic regions of sequence at the extremities of the chain [3]. H1 (which is found in both the type I and II IFP chains [6]) flanks the beginning of the rod domain [21], H2 (only found in type II IFP chains [6]) flanks the end of the rod domain [21]. The H1 domain is one of the most conserved regions in all the IFPs [22]. It is likely to exist as a turn-β-strand-Ω-loop-α-helix which leads into the 1A region [22]. The importance of the H1 domain is shown by mutational studies whereby a single substitution in the sequence can lead to disease. Diseases caused by a mutation in the H1 domain are blistering skin diseases, namely, epidermolytic hyperkeratosis and the Weber-Cockayne form of EBS. The substitutions alter the keratin chain structure and cause a cascading effect on keratin IF structure. This affects the dynamic behaviour of the IFs leading to cytolysis, vacuolisation and failure of tissue integrity causing blistering [3].

Cross-linking modelling studies have shown that the H1 and H2 domains are in alignment with the ends of the rod domains and/or the L2 segment [22]. Using synthetic peptides, it has been shown that the H1 and possibly the H2 domains play an important role in aligning nearest neighbour molecules [22]. The H1 domains are proposed to be the first regions to interact during dimer formation, orientating the domains and allowing coiled-coil formation. The two chains are then fixed in alignment, which facilitates hydrophobic interactions between the two dimers [23].

The H2 domain is predicted to form a turn-β-strand-turn conformation [22]. Studies have shown that when the H2 domain is removed, the IFs are unable to assemble [24]. This may explain the importance of heterodimer formation in keratins, as the H2 domain occurs only in the type II IFP chain [22].

1.5 Intermediate filament secondary structure

Intermediate filaments from wool were originally studied by Astbury over 70 years ago using X-ray diffraction to look at their structure [25]. The orientation pattern showed a well defined secondary structure. Subsequent analysis of the X-ray diffraction data showed a strong α-helical pattern (Figure 1.4). From these data, Crick suggested that the α-helices could form a super-helix or a coiled-coil [26]. Crick also hypothesised that if the side-chains on the α-helix were thought of as knobs on the surface of the cylindrical helix, then the surface was made up of knobs alternating with holes. The holes were areas where the knobs from the neighbouring α-helix would fit [26]. Sequenced fragments of wool IFPs have been shown to contain helix favouring residues [27]. Partial pronase digestion of wool IFPs produced a fragment that was shown to be highly helical using optical rotary dispersion measurements [17]. X-ray diffraction data of vimentin fragments have been analysed and the 1A and 2B crystals were shown to be α-helical (Figure 1.3) [19]. Confirmation of this structure in solution came when circular dichroism also showed vimentin fragments to be α-helical [28].

Secondary structure prediction techniques using the amino acid sequence of a hard α-keratin revealed that the regions at either end of the chains are not α-helical and contain many potential β-bends [29].

1.6 Type I and II IFP coiled-coil dimers

Using X-ray diffraction data, Fraser, MacRae and Suzuki suggested two packing arrangements for the coiled-coil ropes [30]. The first was a two-stranded rope and the second a three-stranded rope. Later that same year, work on the assembly of bovine epidermal filaments concluded that the filaments were made up of a three-chained unit structure [31]. The first evidence of a four chain structure came when Ahmadi *et al.* used cross-linking with dimethyl suberimidate, followed by denaturation and polyacrylamide gel electrophoresis, to show that the protein molecule consisted of four protein chains [32]. Subsequent experiments using partial

chymotryptic digestion of wool IFs produced a four-chain structure, which was proposed to consist of a pair of two-stranded coiled-coil ropes [33]. This isolation of a four-chain structure suggested that the smallest association of molecules was a dimer rather than a trimer [33]. In 1986, ultracentrifugation techniques, electron microscopy and cross-linking confirmed that IFs are composed of two discrete polypeptide forms [34]. Theoretical estimates suggested the dimers were oriented in parallel and probably in a coiled-coil configuration (Figure 1.5).

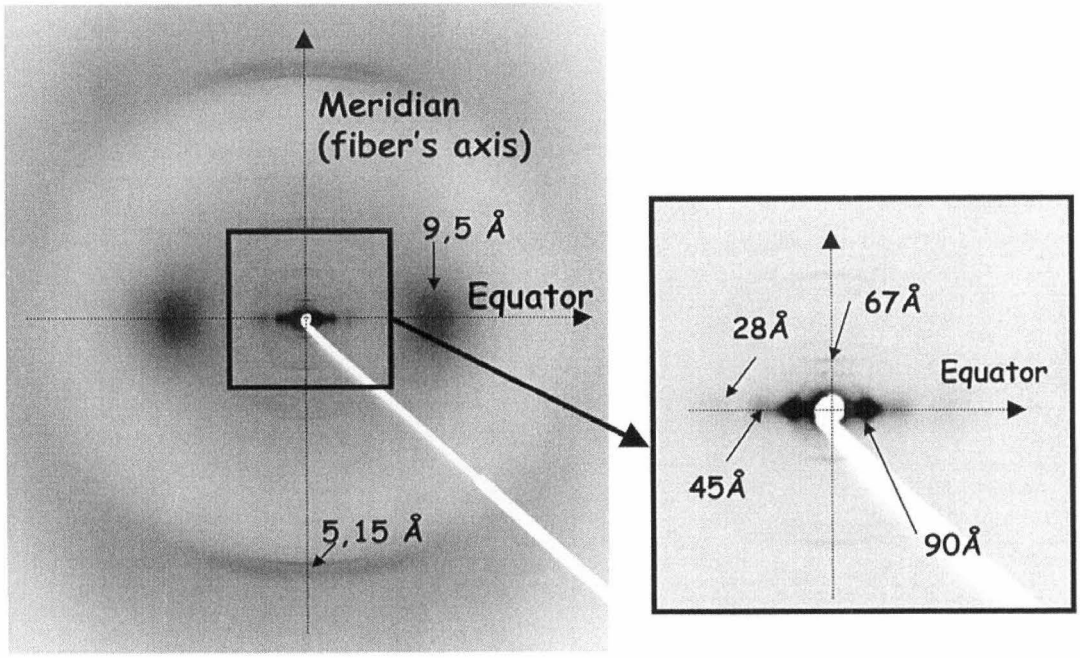


Figure 1.4
X-ray scattering pattern of a bundle of human hair obtained using synchrotron radiation and recorded on a flat detector perpendicular to the X-ray beam. The hairs' axes are vertical [35].

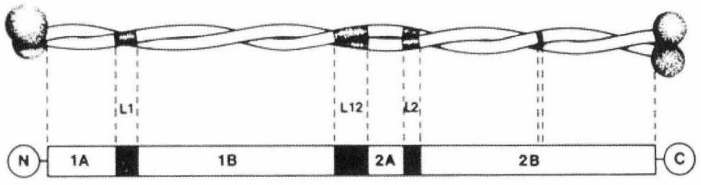


Figure 1.5
Parallel, coiled-coil conformation of an IFP dimer [36, 37].

These dimers had a tendency to form double dimers (tetramers). A refined model of two parallel 1A helices has been constructed and fitted onto a chain of the GCN4 leucine zipper, which is known to have two-stranded coiled-coil geometry. The superimposed image shows only a small number of stereochemical conflicts, most of which can be eliminated by adjusting side-chain conformation (Figure 1.6). The modelling suggests that a coiled-coil formation within the 1A segment is possible with an average coil radius of 5.0 Å and a pitch length of 165 Å [20].

Further secondary structure prediction techniques, using the amino acid sequence of a hard α-keratin, predicted that the ionic interactions which occur between the rod domains of two different keratin IFPs are optimised when the coiled-coil is aligned parallel and in-register [29]. Recently, site-directed spin labelling-electron paramagnetic resonance has shown, in real-time, that IF monomers are aligned as in-parallel and in-register dimers [38].

For some IFPs a 'divide and conquer' approach to determine the structure of the dimer has been adopted. Vimentin proteins cannot be used for crystallisation due to the tendency of the proteins to form filaments. By producing overlapping segments of vimentin, the crystal structure of each segment may potentially be determined, and the fragments then modelled into a complete structure (Figure 1.3) [28]. Six vimentin fragments, ranging from 39 to 84 amino acids, have yielded macroscopic crystals and X-ray diffraction data have been collected [28]. Three of these fragments have been analysed and published, and the molecular organisation of the 1A and 2B segments has been established. The 1A segment showed α-helical geometry and the 2B segment revealed a double-stranded coiled-coil, which unwound to accommodate the stutter. This approach to IF architecture and function revealed data for the first time at the atomic level and is entirely consistent with the models postulated to date [19].

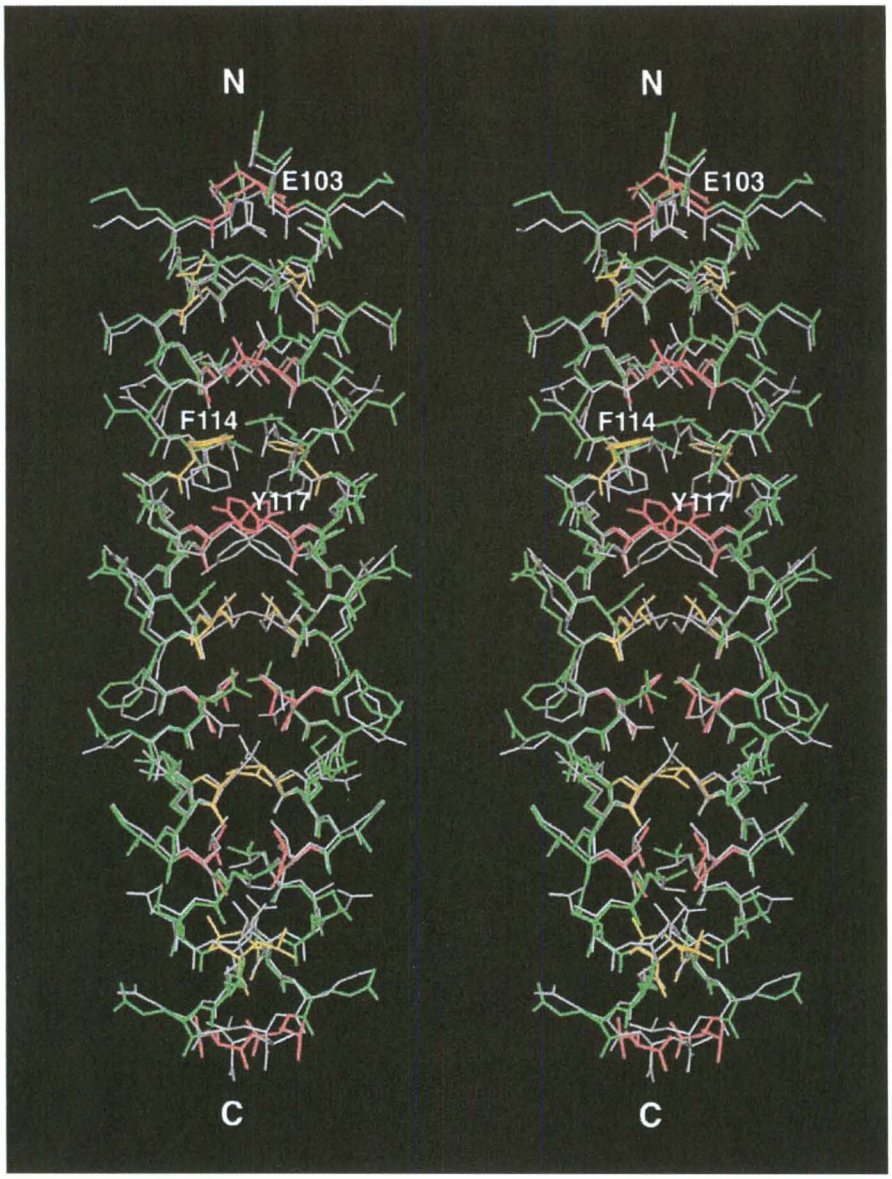


Figure 1.6
Stereo view of the constructed 1A dimer. The final refined model (shown in bright colours) is superimposed onto the initial model (grey) that had been produced by a rigid-body fitting of two crystallographically determined monomeric 1A helices onto the GCN4 leucine zipper. The residues in *a* and *d* positions of the heptad repeat motif are coloured orange and red, respectively [20].

In keratins, the dimers are made up of one type I IFP and one type II IFP. Individual cytokeratin proteins cannot form IFs [39]. It has been shown that, in solution, heterodimers are much more stable than homodimers and therefore are the favoured form [40]. However, cross-linking, high-resolution polyacrylamide gel electrophoresis (PAGE) and blotting with specific K10

antibodies have shown that keratins are able to form homodimers and homotetramers. These homooligomers were unable to form IFs. If the homodimers were dialysed in urea solutions, they would rearrange into heterodimers, which were capable of forming IFs [40]. Using site-specific mutagenesis, a single cystine cross-link has been introduced into keratins 8 and 18. These mutants could form homodimers that were only capable of forming IFs if they rearranged to form heterodimers [41].

In the rod domain, the *a* and *d* hydrophobic residues of the heptad repeat wind around the α-helix of the IFP creating a hydrophobic seal and allowing coiling between two IFPs [2]. The dimers are held together mainly by strong hydrogen bonds [42]. Hydrophobic and non-ionic interactions stabilise the coiled-coil dimer. The hydrophobic *a* and *d* residues in the monomer stabilise the interactions between the two protein chains [6] and interchain ionic interactions occur between the *e* and *g* residues (Figure 1.7) [3, 18].

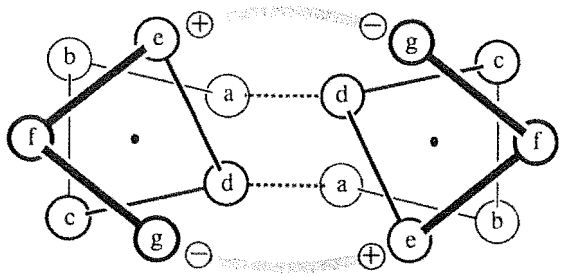


Figure 1.7
The stability of the coiled-coil dimer comes from the *a* and *d* hydrophobic interactions and the interchain ionic interactions between the *e* and *g* residues [36].

The keratin IFPs are expressed sequentially during hair cortical cell differentiation [16]. Different keratin pairs have varying physical properties, which suggests that the differential expression of keratin pairs may be tailored to suit specific structural requirements [2, 43]. Different pairings form IFs with various tensile strengths, flexibilities and dynamics [2]. Using filter

binding assays it was shown that all of the type I IFPs were able to bind to all of the type II IFPs [44]. The assays also showed that the wool proteins could interact with cow snout IFPs, demonstrating that the 'hard' wool proteins were able to recognise the 'soft' epidermal proteins [44]. This shows that even though sequential protein expression occurs in the follicle, *in vitro* assembly of all combinations of keratin IFPs is possible.

1.7 Tetramers

The correct alignment of the molecules past the dimer stage is the rate limiting step for keratin IF formation *in vitro* [6]. Cross-linking data initially revealed three modes of alignment A_{11}, A_{12}, and A_{22} [22]. In the A_{11} alignment, the two antiparallel dimers are staggered so that their 1B segments are aligned. The A_{22} alignment has the 2B segments aligned and in the A_{12} alignment the two dimers are aligned in-register [6]. Further work on vimentin revealed the A_{CN} alignment where there is a head to tail overlap, the 2B rod domain segment of one molecule overlaps with the beginning of the 1A domain segment of a similarly directed molecule (Figure 1.8) [7].

Using these alignment models it becomes apparent that several key sequences overlap each other five times per unit molecule length of 46 nm [6]. Five of these sequence regions (H1, the beginning of 1A, L2, the end of 2B and H2) represent the most highly conserved sequences in the entire IF family [6].

The alignment of hard α-keratins have been studied in detail. After hard α-keratin IFs are formed, terminal differentiation occurs. In the upper areas of the wool follicle, the environment changes from a reducing one into an oxidising one. The process is not fully understood, but it is known that the keratin proteins are keratinised (stabilised) by the conversion of sulfhydryl groups to disulfide bonds [45]. This may influence the interactions between proteins [46]. Cross-linking data have shown that the 1A segment interacts with the head domain [12]. In oxidised hard α-keratin, modelling studies have suggested that the head domain folds back and interacts with the rod

domain. The head domains of IFs are very basic and the rod domains and tail are acidic [19]. The possible role of this fold may be to stabilise the 1A segments [21]. During assembly or when modified (e.g. phosphorylation) the head domain may no longer associate with the 1A segments. This could lead to destabilisation of the 1A segment, causing the two strands to separate. The flexible L1 linker would then act as a hinge. A model for the 1A region suggests that it may exist as a twisted β-sheet which wraps around the 1A segment; this would lead to stabilisation of the 1A region (Figure 1.9) [21].

Detailed analysis of the sequence and structure of the 1A and L1 regions have revealed some interesting features. Across all chain types, there are more conserved residues in 1A than in any other region. There are more leucine residues in position *d*, and more hydrophobic residues in positions *a* and *d*. There are also major differences in the charged residues in the inner *e* and *g* positions and the outer *b*, *c* and *f* regions. Segment 1A is highly hydrophobic, suggesting a role in facilitating aggregation: the ends are also well sealed. The L1 linker is the most elongated and flexible linker region. There are common features between the sequences of the L1 linker in the type I, II, III and IV IFP chains. The highly conserved nature of both the 1A and L1 regions suggests an important role in IF aggregation and stabilisation.

1.8 Protofilaments versus ULFs

Electron microscopy studies on unravelled reconstituted epidermal keratins demonstrated that 4.5 nm protofibrils were the building blocks of the filaments. These protofibrils could be further unravelled to form 2 nm protofilaments. The dimensions of the protofibril were consistent with a pair of protofilaments, which were tetramers that had associated end on end [47]. The unravelling experiments showed that there was a 5.4 nm axial repeat [47]. A schematic model was determined that was made up of tetramers joined end to end to form the 2 nm protofilament; a multi-stranded helix of protofilaments made the 4.5 nm protofibril and four of these protofibrils made a 10 nm filament [47].

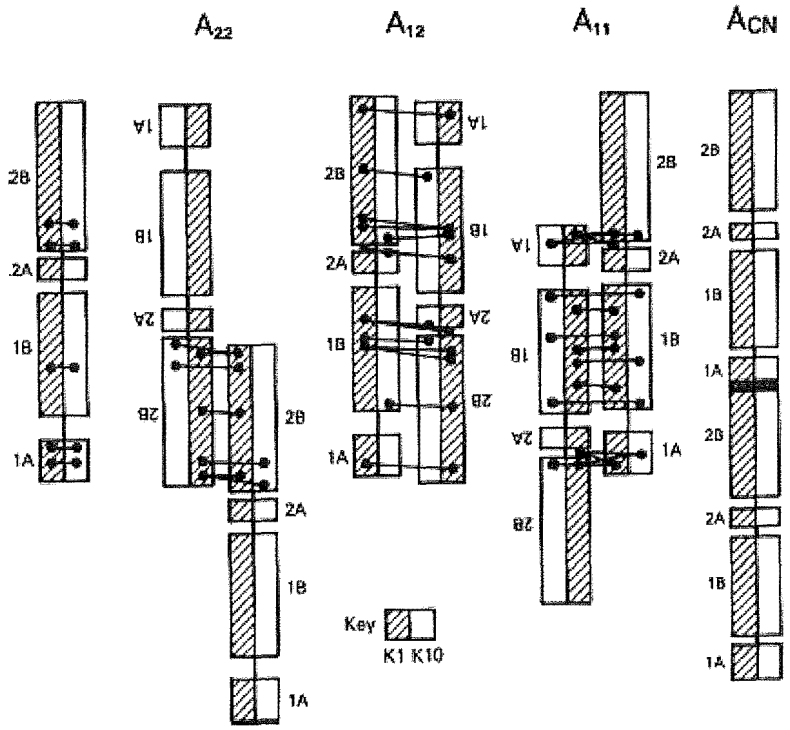


Figure 1.8
Disulfosuccinimidyl tartrate has been used to induce covalent crosslinks between lysine residues that are spacially adjacent. The links have been characterised and the residues involved have been specified by protein sequencing. The link alignments show four different modes; the A_{11} mode of alignment, the 1B segments are overlapped in antiparallel molecules, the A_{22} mode of alignment, the 2B segments are overlapped in antiparallel molecules, the A_{12} mode of alignment, the molecules are almost completely overlapped and anitparallel and the A_{CN} mode of alignment, there is a small head to tail overlap of parallel molecules [4].

Extensive *in vitro* assembly studies on vimentin show that the lateral association of dimers and tetramers into ULFs is preferred over the longitudinal association into protofilaments or protofibrils [48]. 'Stop-flow' studies on vimentin have shown that one second after filament initiation, assembled intermediates can be seen. During the next few seconds, ULFs were annealed longitudinally to generate IFs in excess of 300 nm long. During the next couple of minutes, the filaments compacted to form long smooth walled filaments [49]. It was concluded that experiments that use unravelling of IFs to give protofibrils do not represent true intermediates in the process of IF assembly [50]. Vimentin cross-linking analysis shows that the A_{11} alignment is predominant in ULFs. The alignment modes A_{12} and A_{22}

were shown in mature IFs, suggesting that rearrangements occur during longitudinal annealing and radial compaction to form IFs [51].

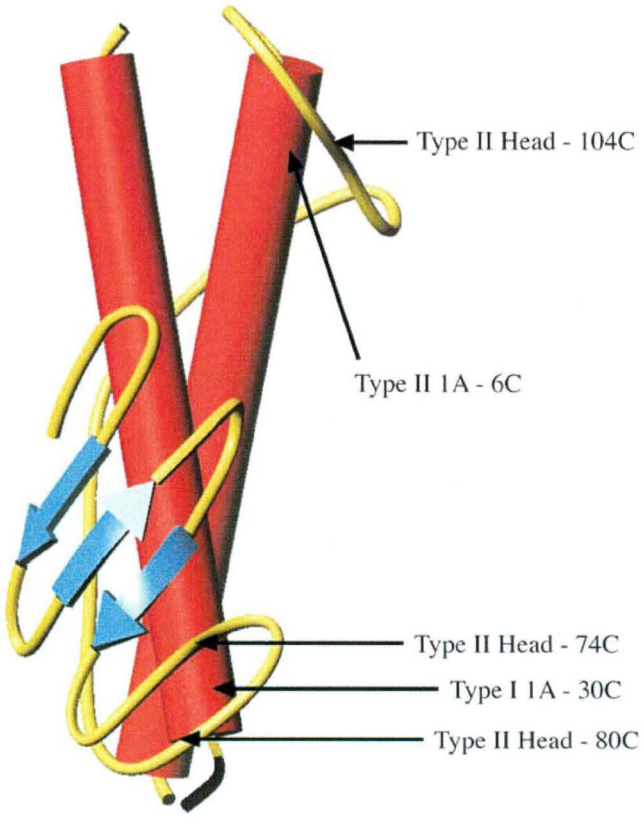


Figure 1.9
Prerefinement model for key elements of the type II IFP head domain in the presence of the adjacent 1A coiled-coil segment. Disulfide bonds can be induced between head domain residue 80 and type I 1A residue 30, and head domain residue 104 and type II 1A residue 6 [21].

1.9 IF structure

Detailed cross-linking studies on vimentin have determined the axial length and the alignment of neighbouring molecules [7]. The studies have shown that when compared to keratin, vimentin has different axial repeats and the alignments of neighbouring molecules are different. The vimentin molecule is shorter than the keratin molecule (43.9 nm compared to 46.2 nm) due to the linker sequences being more extended. Keratin and vimentin share the same alignment modes, but the dimensions of these alignments are different (Figure 1.10). The amount of stagger in the alignments differs, leading to a displacement of two to four heptads between vimentin and keratin.

This explains why mixed vimentin and keratin molecules are incapable of forming IF *in vivo* or *in vitro* [7]. This also suggests that the axial parameters of all type I and II IFPs will be the same, as they appear to co-assemble with any other keratin [39].

When the surface lattice model is wrapped into a tube, the molecules do not match precisely [3]. A seam is formed which coils around IFs, but, due to connections that do not lie in the most favourable mode of alignment, this produces a phase discontinuity. In the past, IFs were considered as static structures; however, it has since been shown that the structures are more dynamic. IFs can exchange individual protein chains everywhere along their length [3]. Microinjection experiments have shown that foreign IFPs are rapidly integrated into the existing IF network [2] (Figure 1.11). A type I IFP that is microinjected could mobilise type II IFPs from the small soluble pools in the cell, alternatively it could take a type II IFP from a preexisting pool of heterodimers. The integration of these proteins may occur at the seam site [3].

There is limited information available about the transverse architecture of IFs. Scanning transmission electron microscopy shows that native hard α-keratin filaments have 32 IFP chains in cross-section [52]. The radial density profile has been debated for many years. Pauling and Corey in 1953 suggested that the α-keratins may consist of a seven-strand cable in parallel orientation [53]. The two possibilities commonly considered for the profile are a ring-core profile or a uniform profile. There have been several suggestions for the ring-core profiles of IFs including: nine outer and two inner, nine outer and one inner, seven outer and one inner, eight outer and zero inner and seven outer and zero inner. Many of these models were developed in the 1970s prior to the findings that the coiled-coils are two-stranded and not three-stranded [35].

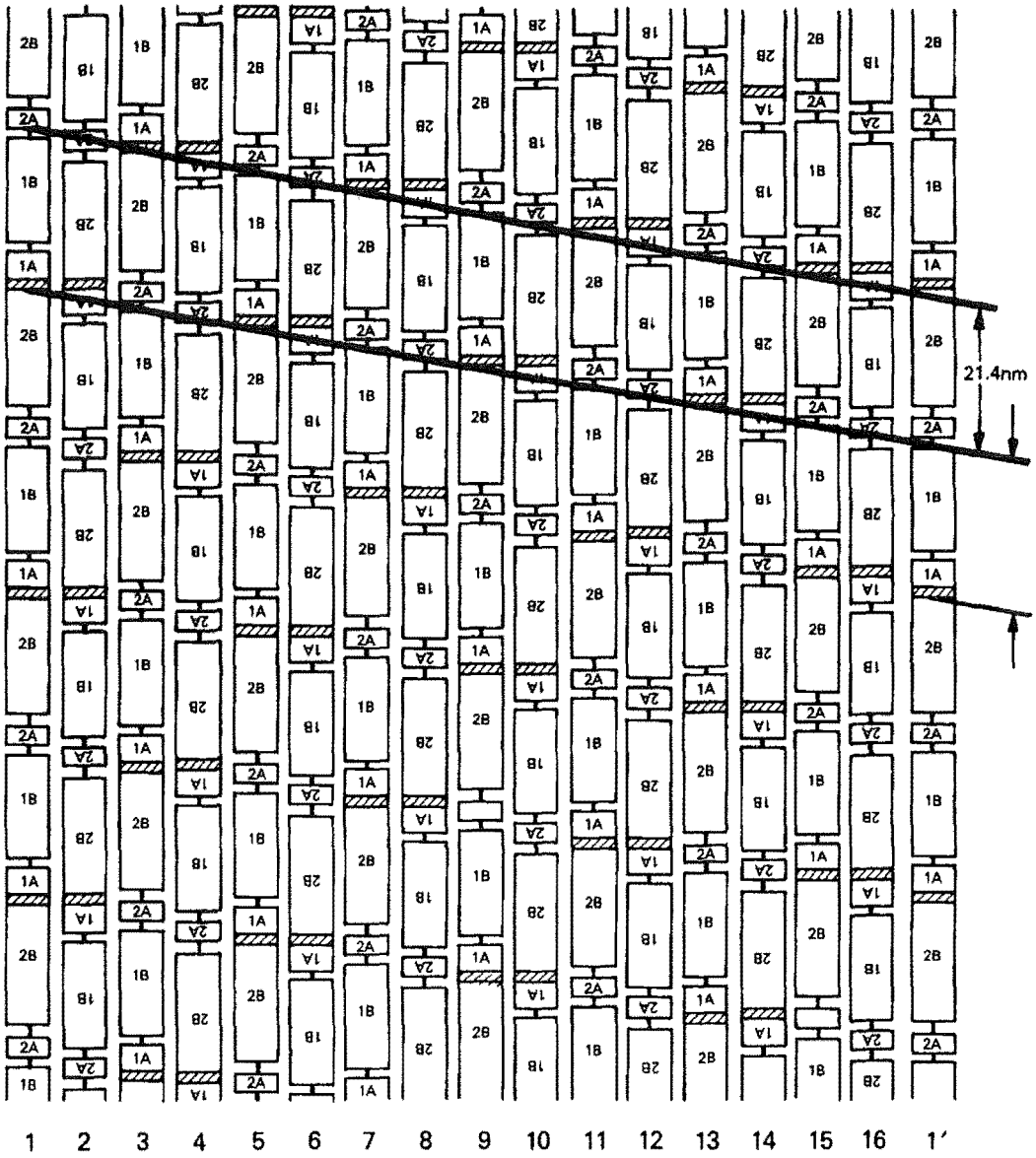


Figure 1.10
Two-dimensional surface lattice model for vimentin IF derived from cross-linking data. The helix has an axial repeat of 21.4 nm, keratin has an axial repeat of 22.6 nm [7].

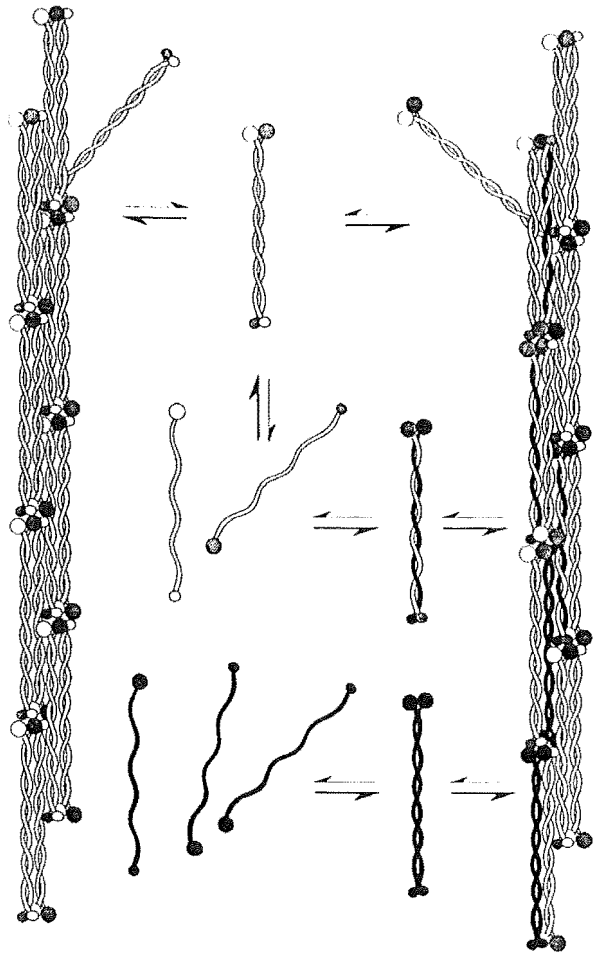


Figure 1.11
IFs can dissociate into monomeric proteins. When biotinylated proteins are microinjected into a cell they can associate with the monomeric proteins to form dimers which are incorporated into the existing IF. Type I IFPs have white terminal domains and type II IFPs have black terminal domains. Biotinylated type I IFPs have black terminal and rod domains [54].

Comparisons between chemically fixed and cryo-fixed samples are not easy. Intermediate filaments from dehydrated resin-embedded samples appear to have IFs clustered closer together and diminished interfilament distances when compared to fully hydrated native samples [55]. Cryo-fixed samples are the method of choice when trying to determine native filament structure. Recent cryo-electron micrographs of near-native IFs show an ~3 nm area of low density around the filament axis [52]. Analysis of the X-ray scattering equatorial profile of human hair in comparison to simulated profiles from atomic models has given information about the architecture of IFs under *in*

vivo conditions (Figure 1.12). Many models were tested using variations in the numbers of inner and outer tetramers. Modelling of the IFs as hollow cylinders produced results which were incompatible with the X-ray diffraction data [52]. Calculated equatorial profiles showed that the radial density across the IF is nearly uniform (if it contained a core, its radius would be less than ~7 Å) [35].

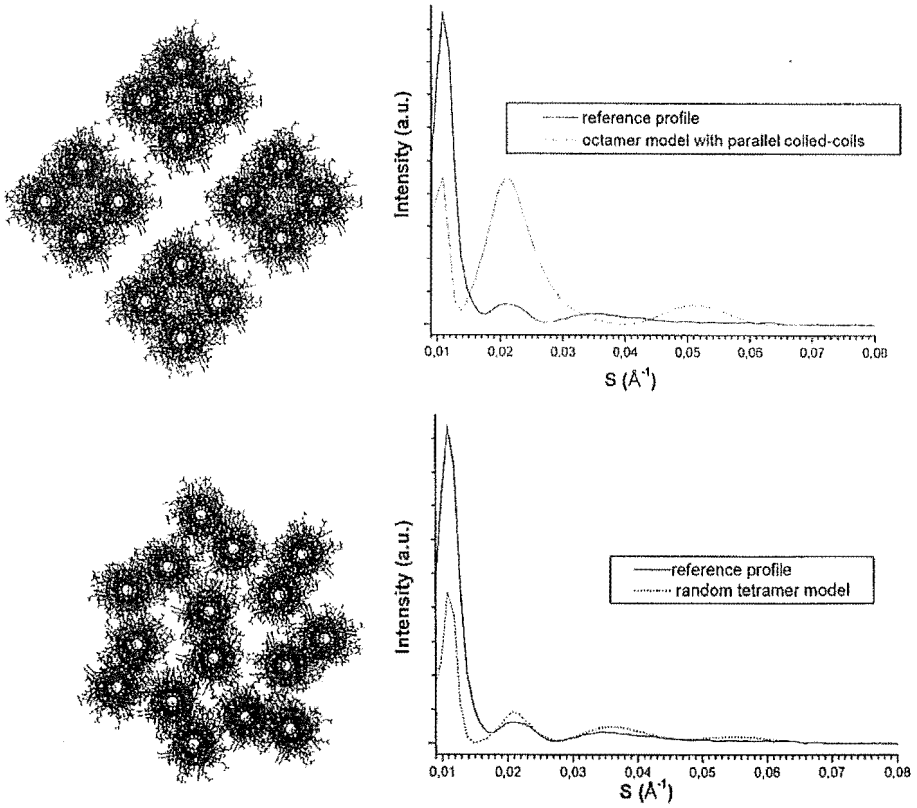


Figure 1.12
The top diagram shows an octamer-based transverse architecture with parallel coiled-coils, and its corresponding calculated X-ray scattering equatorial profiles. The simulated profile yielded is very different from the experimental profile. The bottom diagram shows a randomly ordered tetramer-based herringbone packing and its corresponding X-ray scattering equatorial profiles. The simulated profile is almost identical to the experimental profile [35].

The analysis of the X-ray scattering equatorial profile of human hair showed that the coiled-coils are straight and not supercoiled into oligomers, the coiled coils are probably assembled into oligomers and the positions and orientations of the oligomers do not show any regularity [35].

The longitudinal axial alignment of IFs has also been determined. Cross-linking studies on reduced hard α-keratin IFs have shown that the model derived does not agree with the model determined from X-ray diffraction data. However, when reduced IFs are oxidised there is a rearrangement of the structure [12]. This molecular shift, of approximately 15 residues, maximises the number of stabilising disulfide bonds. The A_{11} alignment mode is shifted, resulting in an increase in the net longitudinal axial repeat. This molecular shift results in a cross-linking model that fits with the model determined from X-ray diffraction data [12].

1.10 Intermediate filament macrofibrils

Filaments were first observed directly using high-resolution transmission electron microscopy. It revealed the IFs, which were 7-8 nm in diameter, embedded in a matrix that contained proteins high in tyrosine and cysteine residues [3]. The IFs, embedded in matrix proteins, formed macrofibrils, which were approximately 0.2-0.5 μm in diameter [56]. The macrofibrils associated to form the contents of cortical cells, which are spindle-shaped, and approximately 100 μm long and 3 μm wide. Transmission electron microscopy (TEM) shows three types of cortical cells, the paracortex, orthocortex and the mesocortex (Figure 1.13).

The paracortical cells always occur on the inside of the crimp wave of high crimp, fine wool and the orthocortex cells are always on the outside of the crimp wave [57]. In the paracortex, the macrofibrils are large, fused, irregular and the IF packing is quasi-hexagonal [58]. The mesocortex has defined macrofibrils with predominantly hexagonal packed IFs that are parallel to the fibre axis [58]. In the orthocortex, the macrofibrils are smaller, cylindrical, separated by intermacrofibrillar material, and are often inclined to the fibre axis which increases with radius [58]. Different combinations of IFPs that are used to form IFs may have an effect on the cortical cell type distribution and abundance, leading to differences in wool fibre properties [59].

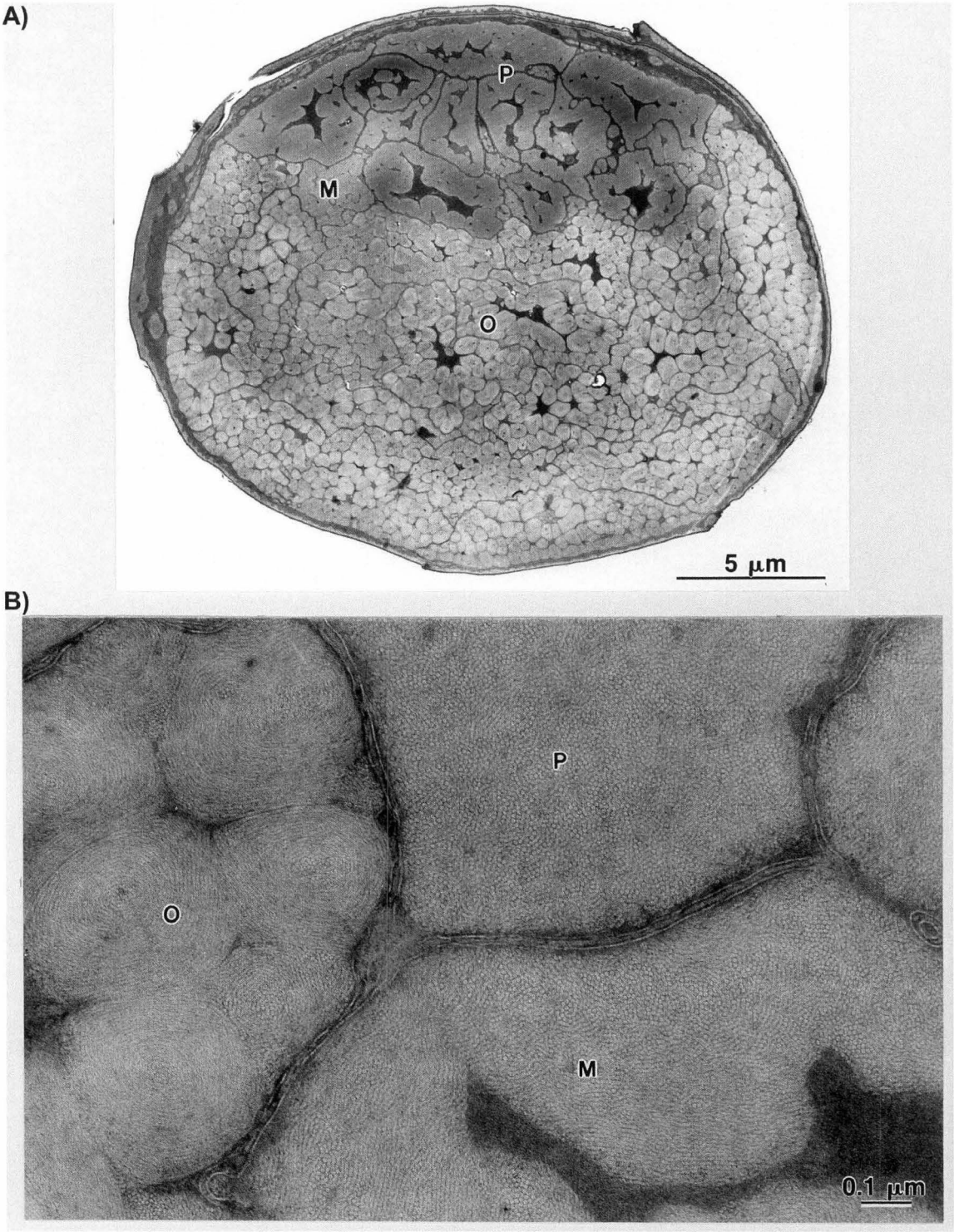


Figure 1.13
A) Cross-section of a wool fibre showing the three different types of cells: P, paracortex, O, orthocortex and M, mesocortex. B) Higher magnification image showing the different packing modes in the cortical cell types.

1.11 Role of the head and tail domains

Head and tail domains seem to have an effect on the thermodynamic interactions between subunits. Modelling studies have been used to suggest roles for the head and tail domains. The models suggested that the head domains may play a role in interfilamentous interactions, whereas the tail domains appear to control intrafilamentous associations [21].

Mutational studies have been used to determine the dispensability of regions of the head and tail domains. Mutational studies, where variable amounts of the head domain was deleted, showed that the length of deletion correlated with the degree of tetramer destabilisation [23]. Mutants without the entire head domain were only able to form soluble oligomers [23]. If a headless mutant was assembled with its native partner, lateral association was affected giving filaments with variable diameter and branch points [23]. It was concluded that their head domain and especially the H1 domain play an important role in stabilisation, alignment and elongation [23]. The head regions are probably also extremely important in interactions between IFs.

Keratin 18 site-directed mutagenesis has shown that the head region is essential for the elongation reaction [24]. Specific vimentin deletion mutants show that the amino acid residues at the beginning of the head domain are critically important for IF polymerisation and the residues 97-103 at the C-terminal end of the head domain are also required [60]. The head domains appear to control end-to-end lateral associations in IFs [2].

The tail domains appear to control lateral associations [2]. Vimentin mutants, lacking a tail domain, formed ULFs, but the diameters of the filaments were considerably larger than the wild-type [49]. Hatzfeld and Weber produced tailless keratin 18 by site-directed mutagenesis (the mutant still had six additional residues at the C-terminal end of the rod). They found that there was no requirement for a tail domain for filament formation *in vitro* [24]. The part of the tail domain closest to the rod may have an essential role is assembly, but mutations which didn't affect these amino acids were able to form filaments [24]. It was concluded that either the tail domains were not

required for assembly, or the presence of the non-mutant assembly partner was sufficient to compensate for the lack of a tail in the mutant [24].

Almost all IFs have a consensus sequence of YRKLLEGEE at the C-terminal end of their central rod domain [61]. This consensus sequence controls filament diameter during assembly, and is crucial for tetramer formation [61]. It may exert this control by steric hindrance or interaction with the rod domain [61]. All the studies on the requirements for head and tail domains in IFs highlight the importance of specific, conserved sequences in IFPs.

1.12 Post-translational modification (PTM)

The study of PTMs in IFs is important from regulatory/mechanistic and functional perspectives [62]. Modifications to IFPs preferentially occur in the head and tail domains [63, 64]. These domains contain the most structural diversity and the tissue-specific functions are likely to be modulated in these regions [64]. Soft keratins have been shown, by radioactively labelling cell cultures, to undergo many PTMs including phosphorylation, acetylation and glycosylation [65]. From past studies, it is thought that the modification by phosphorylation and glycosylation is likely to regulate keratin function [62]. Since IFs in live cells are being continuously restructured, the active exchange of subunits for remodelling requires a mechanism involving PTM [66].

1.12.1 Phosphorylation

Phosphorylation often acts as a molecular switch controlling many cellular functions such as metabolic signal transduction and cell division [67, 68]. All IFPs studied so far have shown serine/threonine phosphorylation, with serine being predominant [64]. Phosphorylation of epidermal keratins has been associated with solubility, reorganisation, protection against cell stress, cell signalling, mitosis, apoptosis and cell compartment specific roles [62, 69] Epidermal keratins undergo a very high phosphorylation turnover [69]. Treatment of epithelial keratins with phosphatase inhibitors leads to an increase in phosphorylation, accompanied by solubilisation, filament

reorganisation and collapse [70]. The epithelial keratins have different phosphorylation levels depending on the biological state of the cell [71].

Studies on vimentin IFs have shown site-specific phosphorylation induces the disassembly of the filaments *in vitro* [72]. The vimentin polymers form soluble tetrameric oligomers upon phosphorylation [66]. Further studies on vimentin and desmin showed that not only did phosphorylation induce disassembly, but it also inhibited polymerisation of filaments. Once the filaments were dephosphorylated, they regained their ability to form IFs [72]. Reversible enzymatic phosphorylation-dephosphorylation of desmin and vimentin may modulate the state of polymerisation and their reversible disassembly in various cell functions [72]. Comparisons between phosphorylated/unphosphorylated and phosphorylation-deficient vimentin indicates that the equilibrium between vimentin polymers and depolymerised subunits and the turnover of subunit exchange are regulated by kinase-phosphatase equilibria [66]. These studies all suggest that the role of phosphorylation is primarily to drive disassembly [66].

Phosphorylation of epithelial cell keratins K5, 6 and 8 leads to solubilisation and a migration shift on 1DE gels. In the same set of experiments, it was also shown that phosphorylation acted as an on/off switch during mitosis [70]. Mitosis-mediated reorganisation of vimentin filaments is associated with an increase in phosphorylation [2]. The nuclear lamin complex is disassembled and reassembled at cytokinesis as the chromosomes condense [73]. This process is controlled by a series of specific protein kinases that cause disassembly into soluble tetramers. At metaphase, the lamins are once again phosphorylated [73].

Neurofilament proteins are heavily phosphorylated at serine residues as the proteins move from cell body to axon. The phosphorylation in neurofilaments does not appear to regulate assembly, but rather it modulates the surface properties of the filaments, affecting functional interactions [2].

1.12.2 Glycosylation

The O-linked N-acetylglucosamine (O-GlcNAc) modification appears to be as abundant and important as phosphorylation in the regulation of many cellular processes [74-76]. Glycosylation may act in a similar way to phosphorylation, as it is a dynamic modification which may have a regulatory role [77]. All of the glycosylated IFPs identified so far also occur as phosphorylated proteins [75].

Chou *et al.* [65] examined the glycosylation of two soft keratins, K8 and K18. Cells were labelled metabolically with [^{3}H] glucosamine. This labelling was not inhibited by tunicamycin. Tunicamycin blocks N-linked glycosylation by blocking donor lipid-linked oligosaccharide formation [65]. This suggests that the glycosylation of K8 and K18 was not N-linked [65]. β-Elimination produced N-acetyllactosaminitol, which showed that the cytokeratins contained single O-GlcNAc residues. Chemical analysis showed that K8 had 1.5 and K18 contained two molecules of O-GlcNAc per protein molecule.

In the soft keratins that contain O-GlcNAc, (K8, 13 and 18 have been identified to date) it has been noted that phosphate and O-GlcNAc appear to occur on distinct hydroxyl moieties [74]. The relationship between O-GlcNAc and O-phosphorylation seems to be highly complex [75]. The turnover rate of the carbohydrate moiety was found to be much faster than the protein turnover rate, which suggests that CK8/18 glycosylation could be acting as a control switch. O-GlcNAc proteins are commonly involved in the assembly of multi-protein complexes [74]. The glycosylation of neurofilaments seems to play a direct role in neurofilament assembly [74]. Assembly experiments using K18 did not show an obvious role for K18 glycosylation in assembly. K18 glycosylation mutants were made by mutating serine and threonine to alanine and expressing the mutants in a baculovirus-insect cell system. Assembled IFs formed with mutated K18 were indistinguishable from wild type filaments [78].

Roles for glycosylation could include modulation of phosphorylation, subcellular localisation, regulating protein turnover [62], facilitating protein-

protein interactions [78] and modulating DNA binding and transactivation [79]. Glycosylation may play a role in transporting macromolecules between the nucleus and the cytoplasm [80].

1.13 Intermediate filament proteins as potential building blocks for biomaterials

The biocompatible materials market is estimated to reach nearly $11.9 billion by 2008 [81]. Polymers are the most popular biomaterial component with sales in 2003 accounting for almost 88% of the entire biomaterial market [82]. Polymers are ideal biomaterials as they are cheap compared to other biomaterials such as alloys and ceramics, can be manipulated to accommodate whatever trait the manufacturer needs and can exhibit toughness, but also maintain the elasticity of a plastic [82].

There are many potential uses for biomaterials. Protein-based materials can be fabricated in a wide variety of forms, such as solid films and coatings for biomedical implants, porous sponges for wound dressings, fibrous materials in woven and non-woven form and liquid components and hydrogels that solidify when administered into tissue [83]. Tissue engineering of skin (a soft keratin) is being studied so that the polymers developed could be used as a transfer dressing on which to grow and transfer a patient's skin cells back to a wound bed such as a burn or an ulcer [84].

Protein polymers are ideal for potential nanotechnology uses, as the building blocks for assembly can easily be manipulated [85]. To be medically useful, a biomaterial must display traits of biocompatibility, meaning that it does not elicit a negative immune response. Biomaterials must also display biodurability, meaning that they resist wear or corrosion [82]. Wool protein is an ideal biomaterial as it is natural, can be obtained with no harm to the animal, has been shown not to elicit any adverse biological response, and is very durable due to the high number of disulfide bonds that can be induced between the proteins.

Keratec Ltd uses a novel process which dissolves and then isolates fractions of keratin proteins from NZ wool, while preserving the functional characteristics - reconstitution of these biopolymer materials enables the capture of the distinct and highly functional characteristics of the original proteins [86]. Keratec Ltd proteins are currently used to produce: fine fibres for use in high value textile markets; adhesives and resins which are environmentally safe; cosmetics, where keratins can be used to impart superior properties in skin, hair and nail care formulations; medical materials where new biopolymer materials can create highly innovative biomaterial products, for example, surgical screws used in bone surgery [87].

Future uses of Keratec proteins could include materials for use in wound healing, bone regeneration, tissue scaffolds for tissue engineering, and keratin protein coatings for medical devices. To fully exploit the *in vivo* qualities displayed by IFPs, *in vitro*, a greater understanding of their structure, PTM and assembly properties is required. This thesis aims to add to that understanding. In the long term, assembled IFs may be used as biomaterials. The proteins could be carefully selected and assembled to produce polymers which have specific properties for specific uses.

1.14 Aims

The aims of this project were:

1. to develop a system for high-resolution separation of wool IFPs,
2. to determine the cause of the charge heterogeneity in the IFPs and
3. to find conditions for the successful reassembly of wool IFPs *in vitro*.

The results will be used to develop new structural knowledge of keratin proteins as an essential prerequisite to understanding how IFPs interact to form stable higher ordered structures. This understanding of the relationships between the IFPs and the higher ordered structures will assist in the development of keratin-derived biopolymers.

1.15 Thesis outline

The first section of work (Chapter Two) focused on the separation of wool IFPs so further analysis could be undertaken. Individual proteins spots were separated on two-dimensional electrophoresis (2DE) gels so they could be further characterised and identified. The methods used to extract and separate the proteins are applicable to other intractable protein systems.

The second section of work (Chapters Three, Four and Five) concentrated on determining the cause of charge heterogeneity in the IFPs. Several possible causes of charge heterogeneity were investigated. The results were compared to what is already known about the soft keratins to determine whether there were important similarities between the two systems.

The final section of work (Chapter Six) comprised a preliminary study on the self-assembly of wool IFP *in vitro*. The long-term goal of this work is to develop IFP assemblies, which can be used as biomaterials and are able to be freely manipulated for specific uses.

1.16 References

1. Plowman JE. Proteomic Database of Wool Components. *Journal of Chromatography B* **2003**;787:63-76
2. Fuchs E, Weber K. Intermediate Filaments: Structure, Dynamics, Function, and Disease. *Annual Review of Biochemistry* **1994**;63:345-382
3. Parry DAD, Steinert PM. Heidelberg, Germany: Springer-Verlag, **1995**:1-183
4. Parry DAD, Steinert PM. Intermediate Filaments: Architecture, Assembly, Dynamics and Polymorphism. *Quarterly Reviews of Biophysics* **1999**;32:99-187
5. Darnell J, Lodish H, Baltimore D. Chapter 22: Actin, Myosin, and Intermediate Filaments. In: *Molecular Cell Biology*. NY, USA: Scientific American Books Inc, **1990**:894-902

6. Steinert PM. Structure, Function, and Dynamics of Keratin Intermediate Filaments. *The Society for Investigative Dermatology, Inc* **1993**:729-734
7. Steinert PM, Marekov LN, Parry DAD. Diversity of Intermediate Filament Structure. *Journal of Biological Chemistry* **1993**;268:24916-24925
8. Gu L, Troncoso JC, Wade JB, Monterio MJ. *In vitro* Assembly Properties of Mutant and Chimeric Intermediate Filament Proteins: Insight into the Function of Sequences in the Rod and End Domains of IF. *Experimental Cell Research* **2004**;298:249-261
9. Dowling LM, Parry DAD, Sparrow LG. Structural Homology Between Hard Alpha-Keratin and the Intermediate Filament Proteins Desmin and Vimentin. *Bioscience Reports* **1983**;3:73-78
10. Heid HW, Werner E, Franke WW. The Complement of Native Alpha-Keratin Polypeptides of Hair-Forming Cells: A Subset of Eight Polypeptides that Differ from Epithelial Cytokeratins. *Differentiation* **1986**;32:101-119
11. Zimek A, Stick R, Weber K. Genes Coding for Intermediate Filament Proteins: Common Features and Unexpected Differences in the Genomes of Humans and the Teleost Fish *Fugu rubripes*. *Journal of Cell Science* **2003**;116:2295-2302
12. Wang H, Parry DAD, Jones LN, Idler WW, Marekov LN, Steinert PM. *In vitro* Assembly and Structure of Trichocyte Keratin Intermediate Filaments: A Novel Role for Stabilization by Disulfide Bonding. *Journal of Cell Biology* **2000**;151:1459-1468
13. www.woolsnz.com.
14. Gillespie JM. The Fractionation of S-Carboxymethyl Kerateine 2 From Wool. *Australian Journal of Biological Science* **1957**;10:105-117
15. Coulombe PA, Ma L, Yamada S, Wawersik M. Intermediate Filaments at a Glance. *Journal of Cell Science* **2001**;114:4345-4347
16. Powell BC, Rogers GE. The Role of Keratin Proteins and Their Genes in the Growth, Structure and Properties of Hair. In: Jollès P,

Zahn H, Höcker H, eds. *Formation and Structure of Human Hair.* Basel/Switzerland: Birkhäuser Verlag, **1997**:59-148

17. Crewther WG, Dowling LM. The Preparation and Properties of Large Peptides from the Helical Regions of the Low-Sulfur Proteins of Wool. *Applied Polymer Symposium* **1971**;18:1-20
18. Mothes E, Shoeman RL, Traub P. Effect of Pepstatin A on Structure and Polymerisation of Intermediate Filament Subunit Proteins *in vitro*. *Journal of Structural Biology* **1991**;106:64-72
19. Strelkov SV, Herrmann H, Geisler N, Wedig T, Zimbelmann R, Aebi U, Burkhard P. Conserved Segments 1A and 2B of the Intermediate Filament Dimer: their Atomic Structures and Role in Filament Assembly. *The EMBO Journal* **2002**;21:1255-1266
20. Smith TA, Strelkov SV, Burkhard P, Aebi U, Parry DAD. Sequence Comparrisons of Intermediate Filament Chains: Evidence of a Unique Functional/Structural Role for Coiled-Coil Segment 1A and Linker L1. *Journal of Structural Biology* **2002**;137:128-145
21. Parry DAD, Marekov LN, Steinert PM, Smith TA. A Role for the 1A and L1 Rod Domain Segments in Head Domain Organisation and Function of Intermediate Filaments: Structural Analysis of Trichocyte Keratin. *Journal of Structural Biology* **2002**;137:97-108
22. Steinert PM, Parry DAD. The Conserved H1 Domain of the Type II Keratin 1 Chain Plays an Essential Role in the Alignment of Nearest Neighbor Molecules in Mouse and Human Keratin 1/Keratin 10 Intermediate Filaments at the Two- to Four-Molecule Level of Structure. *Journal of Biological Chemistry* **1993**;268:2878-2887
23. Hatzfeld M, Burba M. Function of Type I and Type II Keratin Head Domains: Their Role in Dimer, Tetramer and Filament Formation. *Journal of Cell Science* **1994**;107:1959-1972
24. Hatzfeld M, Weber K. Tailless Keratins Assemble into Regular Intermediate Filaments *in vitro*. *Journal of Cell Science* **1990**;97:317-324
25. Fraser RDB, MacRae TP, Rogers GE. Keratins. In: Kugelmass IN, ed. Springfield: Charles C Thomas, **1972**
26. Crick FHC. Is Alpha-Keratin a Coiled Coil? *Nature* **1952**;170:882-883

27. Crewther WG, Dowling LM, Gough KH, Inglis AS, McKern NM, Sparrow LG, Woods EF. The Low-Sulphur Proteins of Wool: Studies on Their Classification, Characterisation, Primary and Secondary Structure. In: *Proceedings of the 5th International Wool Textile Research Conference*. Aachen, Germany, **1975**:233-242

28. Strelkov SV, Herrmann H, Geisler N, Lustig A, Ivaninskii S, Zimbelmann R, Burkhard P, Aebi U. Divide-and-Conquer Crystallographic Approach Towards an Atomic Structure of Intermediate Filaments. *Journal of Molecular Biology* **2001**;306:773-781

29. Dowling LM, Crewther WG, Parry DAD. Secondary Structure of Component 8c-1 of Alpha-Keratin. *Biochemical Journal* **1986**;236:705-712

30. Fraser RDB, MacRae TP, Suzuki E. Structure of the Alpha-Keratin Microfibril. *Journal of Molecular Biology* **1976**;108:435-452

31. Steinert PM, Idler WW, Zimmerman SB. Self-Assembly of Bovine Epidermal Keratin Filaments *in vitro*. *Journal of Molecular Biology* **1976**;108:547-567

32. Ahmadi B, Boston NM, Dobb MG, Speakman PT. Possible Four-Chain Repeating Unit in the Microfibril of Wool. In: Parry DAD, Creamer LK, eds. *Fibrous Proteins: Scientific, Industrial and Medical Aspects*. New York, USA: Academic Press, **1980**:161-166

33. Woods EF, Gruen LC. Structural Studies on the Microfibrillar Proteins of Wool: Characterisation of the Alpha-Helix-Rich Particle Produced by Chymotrypsin Digestion. *Australian Journal of Biological Science* **1981**;34:515-526

34. Quinlan RA, Hatzfeld M, Franke WW. Characterisation of Dimer Subunits of Intermediate Filament Proteins. *Journal of Molecular Biology* **1986**;192:337-349

35. Rafik ME, Doucet J, Briki F. The Intermediate Filament Architecture as Determined by X-Ray Diffraction Modeling of Hard Alpha-Keratin. *Biophysical Journal* **2004**;86:3893-3904

36. Conway JF, Parry DAD. Structural Features in the Heptad Substructure and Longer Range Repeats of Two-Stranded Alpha-

Fibrous Proteins. *International Journal of Biological Macromolecules* **1990**;12:328-334

37. Parry DAD. Primary and Secondary Structure of IF Protein Chains and Modes of Molecular Aggregation. In: Goldman RD, Steinert PM, eds. *Cellular and Molecular Biology of Intermediate Filaments*. New York, USA: Plenum, **1990**
38. Hess JF, Voss JC, Fitzgerald PG. Real-time Observation of Coiled-Coil Domains and Subunit Assembly in Intermediate Filaments. *Journal of Biological Chemistry* **2002**;277:35516-35522
39. Hatzfeld M, Franke WW. Pair Formation and Promiscuity of Cytokeratins: Formation *in vitro* of Heterotypic Complexes and Intermediate-Sized Filaments by Homologous and Heterologous Recombinations of Purified Polypeptides. *Journal of Cell Biology* **1985**;101:1826-1841
40. Steinert PM. The Two-Chain Coiled-Coil Molecule of Native Epidermal Keratin Intermediate Filaments is a Type I-Type II Heterodimer. *Journal of Biological Chemistry* **1990**;265:8766-8774
41. Hatzfeld M, Weber K. The Coiled Coil of *in vitro* Assembled Keratin Filaments is a Heterodimer of Type I and II Keratins: Use of Site-Specific Mutagenesis and Recombinant Protein Expression. *Journal of Cell Biology* **1990**;110:1199-1210
42. Hatzfeld M, Maier G, Franke WW. Cytokeratin Domains Involved in Heterotypic Complex Formation Determined by *in vitro* Binding Assays. *Journal of Molecular Biology* **1987**;197:237-255
43. Yamada S, Wirtz D, Coulombe PA. Pairwise Assembly Determines the Intrinsic Potential for Self-Organization and Mechanical Properties of Keratin Filaments. *Molecular Biology of the Cell* **2002**;13:382-391
44. Herrling J, Sparrow LG. Interactions of Intermediate Filament Proteins from Wool. *International Journal of Biological Macromolecules* **1991**;13:115-119
45. Orwin DFG. The Cytology and Cytochemistry of the Wool Follicle. *International Review of Cytology* **1979**;60:331-374

46. Hofmann I, Winter H, Mücke N, Langowski J, Schweizer J. The *in vitro* Assembly of Hair Follicle Keratins: Comparison of Cortex and Companion Layer Keratins. *Biological Chemistry* **2002**;383:1373-1381

47. Aebi U, Fowler WE, Rew P, Sun T-T. The Fibrillar Substructure of Keratin Filaments Unraveled. *Journal of Cell Biology* **1983**;97:1131-1143

48. Herrmann H, Wedig T, Porter RM, Lane EB, Aebi U. Characterisation of Early Assembly Intermediates of Recombinant Human Keratins. *Journal of Structural Biology* **2002**;137:82-96

49. Herrmann H, Häner M, Brettel M, Müller SA, Goldie KN, Fedtke B, Lustig A, Franke WW, Aebi U. Structure and Assembly Properties of the Intermediate Filament Protein Vimentin: The Role of its Head, Rod and Tail Domains. *Journal of Molecular Biology* **1996**;264:933-953

50. Herrmann H, Aebi U. Intermediate Filament Assembly: Fibrillogenesis is Driven by Decisive Dimer-Dimer Interactions. *Current Opinion in Structural Biology* **1998**;8:177-185

51. Mücke N, Wedig T, Bürer A, Marekov LN, Steinert PM, Langowski J, Aebi U, Herrmann H. Molecular and Biophysical Characterization of Assembly-Starter Units of Human Vimentin. *Journal of Molecular Biology* **2004**;340:97-114

52. Watts NR, Jones LN, Cheng N, Wall JS, Parry DAD, Steven AC. Cryo-Electron Microscopy of Trichocyte (Hard Alpha-Keratin) Intermediate Filaments Reveals a Low-Density Core. *Journal of Structural Biology* **2002**;137:109-118

53. Pauling L, Corey RB. Compound Helical Configurations of Polypeptide Chains: Structure of Proteins of the Alpha-Keratin Type. *Nature* **1953**;171:59-61

54. Miller RK, Khoun S, Goldman RD. Dynamics of Keratin Assembly: Exogenous Type I Keratin Rapidly Associates with Type II Keratin *in vivo*. *Journal of Cell Biology* **1993**;122:123-135

55. Norlén L, Al-Amoudi A. Stratum Corneum Keratin Structure, Function, and Formation: The Cubic Rod-Packing and Membrane

Templating Model. *Journal of Investigative Dermatology* **2004**;123:715-732

56. Kreplak L, Franbourg A, Briki F, Dallé D, Doucet J. A New Deformation Model of Hard Alpha-Keratin at the Nanometer Scale: Implications for Hard Alpha-Keratin Intermediate Filament Mechanical Properties. *Biophysical Journal* **2002**;82:2265-2274
57. Fraser RDB, Rogers GE, Parry DAD. Nucleation and Growth of Macrofibrils in Trichocyte (Hard-Alpha) Keratins. *Journal of Structural Biology* **2003**;143:85-93
58. Bryson WG, Mastronarde DN, Caldwell JP, Nelson W, Woods JL. High Voltage Microscopical Imaging of Macrofibril Ultrastructure Reveals the Three-Dimensional Spatial Arrangement of Intermediate Filaments in Romney Wool Cortical Cells - A Causative Factor in Fibre Curvature. In: *The 10th International Wool Textile Research Conference*, **2000**:1-8
59. Bryson WG, Parbhu AN, Mastronarde DN, Liu H, Caldwell JP, Lal R. New Structural Research Suggests a Three-Component Model for Describing Wool Fiber Curvature and Bending Stiffness. In: *Proceedings of the 30th Textile Research Symposium*. Mt Fuji, Japan, **2001**:227-236
60. Shoeman RL, Hartig R, Berthel M, Traub P. Deletion Mutagenesis of the Amino-Terminal Head Domain of Vimentin Reveals Dispensability of Large Internal Regions for Intermediate Filament Assembly and Stability. *Experimental Cell Research* **2002**;279:344-353
61. Herrmann H, Strelkov SV, Feja B, Rogers KR, Brettel M, Lustig A, Häner M, Parry DAD, Steinert PM, Burkhard P, Aebi U. The Intermediate Filament Protein Consensus Motif of Helix 2B: Its Atomic Structure and Contribution to Assembly. *Journal of Molecular Biology* **2000**;298:817-832
62. Omary MB, Ku N-O, Liao J, Price D. Keratin Modifications and Solubility Properties in Epithelial Cells and *in vitro*. In: Herrmann, Harris, eds. *Subcellular Biochemistry*. New York: Plenum Press, **1998**:105-140

63. Steinert PM, Wantz ML, Idler WW. *O*-Phosphoserine Content of Intermediate Filament Subunits. *Biochemistry* **1982**;21:177-183

64. Ku N-O, Omary MB. Expression, Glycosylation, and Phosphorylation of Human Keratins 8 and 18 in Insect cells. *Experimental Cell Research* **1994**;211:24-35

65. Chou C-F, Smith AJ, Omary MB. Characterization and Dynamics of O-Linked Glycosylation of Human Cytokeratin 8 and 18. *Journal of Biological Chemistry* **1992**;267:3901-3906

66. Eriksson JE, He T, Trejo-Skalli AV, Ann-Sofi H-B, Hellman J, Chou Y-H, Goldman RD. Specific *in vivo* Phosphorylation Sites Determine the Assembly Dynamics of Vimentin Intermediate Filaments. *Journal of Cell Science* **2004**;117:919-932

67. Sickmann A, Meyer HE. Phosphoamino Acid Analysis. *Proteomics* **2001**;1:200-206

68. Coulombe PA, Omary MB. 'Hard' and 'Soft' Principles Defining the Structure, Function and Regulation of Keratin Intermediate Filaments. *Current Opinion in Structural Biology* **2002**;14:110-122

69. Paramio JM. A Role for Phosphorylation in the Dynamics of Keratin Intermediate Filaments. *European Journal of Cell Biology* **1999**;78:33-43

70. Toivola DM, Zhou Q, English LS, Omary MB. Type II Keratins are Phosphorylated on a Unique Motif During Stress and Mitosis in Tissues and Cultured Cells. *Molecular Biology of the Cell* **2002**;13:1857-1870

71. Liao J, Ku N-O, Omary MB. Two-Dimensional Gel Analysis of Glandular Keratin Intermediate Filament Phosphorylation. *Electrophoresis* **1996**;17:1671-1676

72. Inagaki M, Gonda Y, Matsuyama M, Nishizawa K, Nishi Y, Sato C. Intermediate Filament Reconstruction *in vitro*. The Role of Phosphorylation on the Assembly-Disassembly of Desmin. *Journal of Biological Chemistry* **1988**;263:5970-5978

73. Steinert PM, Idler WW, Cabral F, Gottesman MM, Goldman RD. *In vitro* Assembly of Homopolymer and Copolymer Filaments from Intermediate Filament Subunits of Muscle and Fibroblastic Cells.

Proceedings of the National Academy of Science USA **1981**;78:3692-3696

74. Hart GW, Greis KD, Dong D, Blomberg MA, Chou T-Y, Jiang M-S, Roquemore EP, Snow DM, Kreppel LK, Cole RN, Comer FI, Arnold CS, Hayes BK. O-linked N-acetylglucosamine: The "yin-yang" of Ser/Thr Phosphorylation? *Glycoimmunology* **1995**:115-123
75. Comer FI, Hart GW. *O*-Glycosylation of Nuclear and Cytosolic Proteins. *Journal of Biological Chemistry* **2000**;275:29179-29182
76. Jung E, Veuthey A-L, Gasteiger E, Bairoch A. Annotation of Glycoproteins in the SWISS-PROT Database. *Proteomics* **2001**;1:262-266
77. Greis KD, Hayes BK, Comer FI, Kirk M, Barnes S, Lowary TL, Hart GW. Selective Detection and Site-Analysis of O-GlcNAc- Modified Glycopeptides by ß-Elimination and Tandem Electrospray Mass Spectrometry. *Analytical Biochemistry* **1996**;234:38-49
78. Ku N-O, Omary MB. Identification and Mutational Analysis of the Glysosylation Sites of Human Keratin 18. *Journal of Biological Chemistry* **1995**;270:11820-11827
79. O'Donnell N. Intracellular Glycosylation and Development. *Biochimica et Biophysica Acta* **2002**;1573:336-345
80. Snow CM, Senior A, Gerace L. Monoclonal Antibodies Identify a Group of Nuclear Pore Complex Glycoproteins. *Journal of Cell Biology* **1987**;104:1143-1156
81. www.bccresearch.com/press.
82. www.devicelink.com/mddi/archive/05/05/024.html.
83. www.devicelink.com/mddi/archive/98/04/005.html.
84. www.chemsoc.org/chembytes/ezine/2000/france_jun00.htm.
85. Waterhouse SH, Gerrard JA. Amyloid Fibrils in Bionanotechnology. *Australian Journal of Chemistry* **2004**;57:519-523
86. www.woolequities.co.nz/keratec.htm.
87. www.Keratec.co.nz.

Chapter Two

Two-Dimensional Gel Electrophoresis of Wool Intermediate Filament Proteins

The goal of this section of work was to separate the wool IFPs so that individual family members could be characterised. This was an important first step towards understanding the way that IFPs interact with each other and with other proteins. The separation and characterisation of wool IFPs could lead to breakthroughs in the manipulation of the proteins for biomaterial uses.

2.1 Introduction

The largest and most complex group of IFPs are the keratins [1, 2]. The keratins are divided into two subgroups: the soft and the hard α-keratins. The keratin IFPs are further split into the type I and type II groups. The type I IFPs are acidic and range in size from 392-416 amino acid residues [2]. The type II IFPs are neutral/basic and range in size from 479-506 amino acid residues [2]. Type I and type II IFPs have 30% sequence identity [3].

The proteins of hair and wool are products of several gene families, each having a number of closely related members [2]. In the wool follicle, eight wool IFPs are synthesised from different mRNAs and are probably all encoded by different genes [4, 5]. Mature wool polypeptides are genuine translation products and are not derived from precursor forms [5]. Studies looking at the IFPs as they progress up the hair shaft show that the proteins are not cleaved or degraded [5].

Chemically, the collection of different wool proteins has proven extremely difficult to study. Hard α-keratins form some of the most stable protein-

protein interactions known in nature [1]. This stability is important in nature as the hard α-keratins function as protective layers (e.g. wool, hair, hoof and horn). Keratins are insoluble in the usual protein solvents and are resistant to attack by proteolytic enzymes [6, 7].

Very little keratin protein can be extracted until a high proportion of the disulfide bonds have been broken [6]. These disulfide bonds impose a high level of cross-linking in the mature wool and allow tight associations with other proteins [5]. Even after the disulfide bonds of cystine have been reduced and then blocked to prevent protein-protein interaction *via* reoxidation, the proteins are still resistant to extraction [8]. Once extracted, keratin proteins have a tendency to aggregate, making further study difficult [6]. Some soft keratins require solubilisation in 10 M urea prior to electrophoresis [9].

Early work [10-15] to separate wool proteins on electrophoretic gels separated the protein into the two main groupings, the low-sulfur IFPs, and the IFAPs (matrix proteins), the latter made up of two classes, the high-sulfur proteins and the high glycine tyrosine proteins (Figure 2.1).

O'Donnell and Thompson [16] initially classified the IFPs based on starch gel electrophoresis patterns. The proteins were numbered, based on their mobility, from 1-8. Subsequent gel filtration showed that components 1-4 were likely to be aggregates of component 7 and maybe component 8 [16].

Urea PAGE separated the IFPs into the major bands numbered 5, 7 and 8. Sodium dodecyl sulfate (SDS)-PAGE further separated band 7 into three sub-components, 7a, 7b and 7c. Component 8 separated into five sub-components, 8a, 8b, 8c-1, 8c-2 and 8c-3 [3]. Component 8c-3 was later found to be an artifact of storage in alkaline solutions [17].

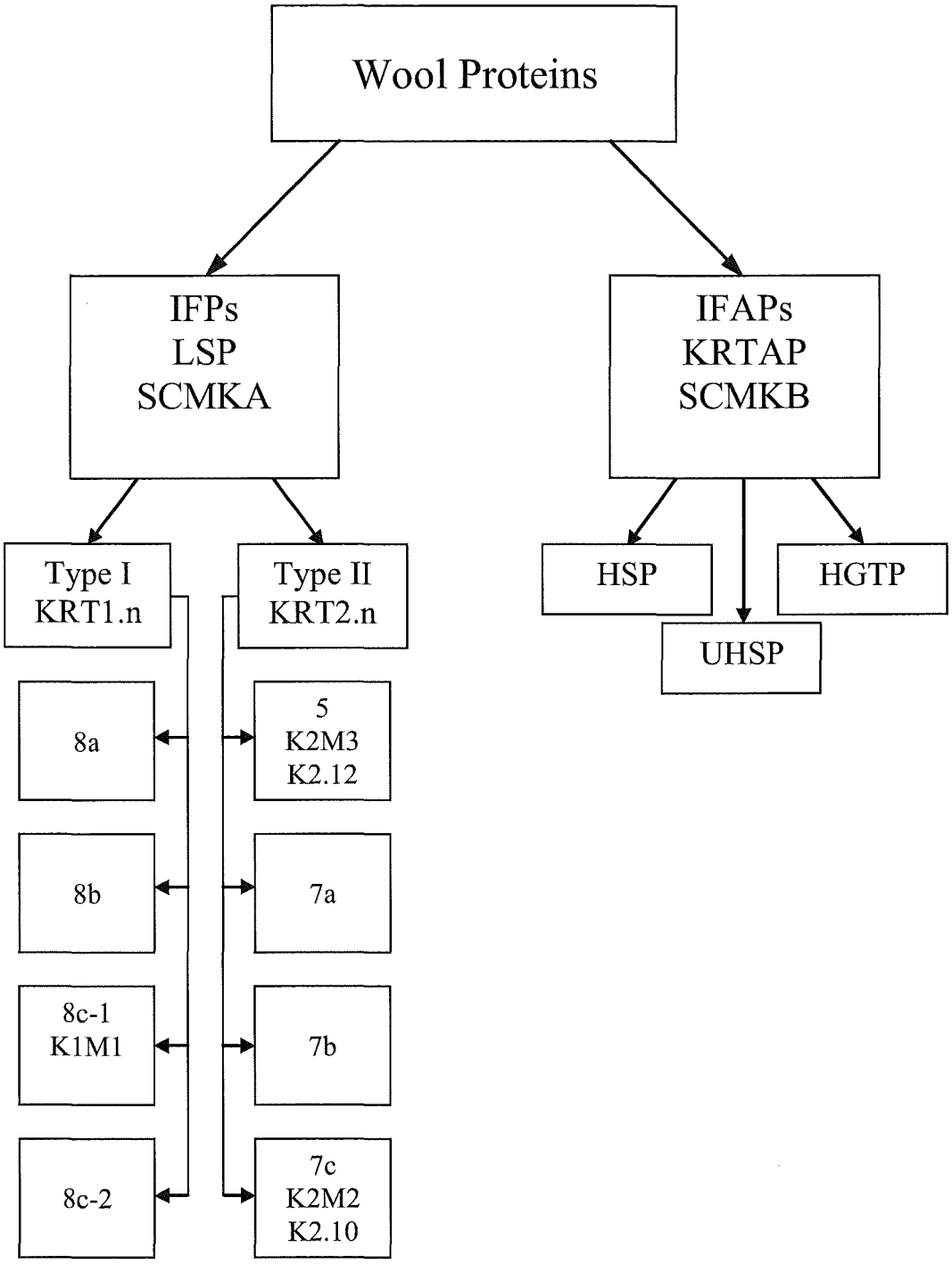


Figure 2.1
Nomenclature of wool proteins [2, 16, 18].

Abbreviations
IFP - intermediate filament protein, LSP - low sulfur protein, SCMK - *S*-carboxymethylcysteine kerateine, IFAP - intermediate filament associated protein, KRTAP - keratin associated protein, KRT - keratin, HSP - high sulfur protein, UHSP - ultra high sulfur protein, HGTP - high glycine tyrosine protein.

Alkaline PAGE in the first dimension followed by SDS-PAGE in the second dimension produced fluorographs of wool proteins that had been labelled with iodo (2-^{14}C) acetate, which showed eight low-sulfur protein spots corresponding with the low-protein components 5, 7a, 7b, 7c, 8a, 8b, 8c-1, 8c-2 [12]. Amino acid sequences have been determined for components 8c-1, 7c, and 5 [19-21]. A cDNA clone bank was prepared from wool follicle RNA, which was then hybridised with [^{32}P]cDNA. Clones that hybridised strongly with the radioactive cDNA were characterised by nucleotide sequencing. An almost complete sequence of component 8a has been determined but not published [10].

Marshall and Blagrove used isoelectric focusing (IEF) in a polyacrylamide gel containing 8 M urea to separate wool IFPs [22]. They demonstrated charge heterogeneity within the proteins focused. One-dimensional (1D) PAGE gels of wool IFPs have also shown that the low-sulfur group contains eight major proteins, with apparent molecular weights of 53 200-75 200 Da [11]. These proteins did not differ between sheep or between samples taken from different areas on the same sheep. Herbert and Woods [13] used immobilized pH gradient-isoelectric focusing (IPG-IEF) combined with SDS-PAGE to separate over 50 protein spots from *S*-carboxymethyl kerateines (SCMK). Resolution of the 2DE gels using IPG-IEF was greater than any previously obtained using acid or alkaline non-equilibrium PAGE in the first dimension. In 1997, improvements in 2D-PAGE allowed Herbert *et al.* [14] to separate the type II IFPs into a string of proteins of similar molecular weight but differing in isoelectric point: the type I IFPs were separated out at a lower molecular weight, and isoelectric point, and were more closely spaced. Using amino acid analysis, they were unable to differentiate between K2M2 and K2M3 due to the high sequence identity (type I IFPs 92%, type II IFPs 77%).

Two-dimensional electrophoresis and peptide mass fingerprinting has led to the identification, by comparison to the known sequences, of 14 of the type I and II IFP spots [15]. However, resolution of the IFPs, especially the type Is, was not ideal, with many spots overlapping neighbouring spots [15].

Thus, high-resolution separation of IFPs is necessary to allow detailed investigations into the composition of the IFs and the structural analysis of their proteins, and to understand their assembly properties. In this chapter, changes to the 2D-PAGE method are described, which improve the separation of both the type I and type II IFPs. These improvements to separation will allow more detailed studies into IFP composition, including characterisation of PTMs. Compositional studies will facilitate a better understanding of how IFPs assemble, both *in vivo* and *in vitro*.

2.2 Results & Discussion

2.2.1 Previous methods

Using past methods to separate wool proteins by 2DE [13, 23, 24], the type I IFPs focus at around pI 4.7 into a group of spots that are hard to distinguish. They appear to resolve vertically into about two rows of spots (Figure 2.2). The type II IFPs focus between the pIs of about 5.3 to 6.4. The more basic proteins are very faint and resolve at a higher molecular weight than the more acidic type II IFPs (Figure 2.2). The protein spots appear very faint for the amount of protein loaded onto the first dimension IPG strip, and there are several regions of horizontal and vertical smearing. Changes were made to the 2DE method to try to improve the resolution of the IFP spots to make subsequent analysis easier.

2.2.2 Reduction of smearing

Most of the changes that were made to the 2DE method have been made to alleviate the problem of smearing on the gels. There are many causes of smearing on a 2DE gel, including precipitation and aggregation of the proteins, deposited debris, short equilibration times and oxidation of proteins [25].

Horizontal streaks in the 2DE spot pattern can be caused by debris in the sample [25]. All protein solutions were centrifuged at 14 000 rpm for 10

minutes to remove insoluble material, before being used to rehydrate the IPG strip.

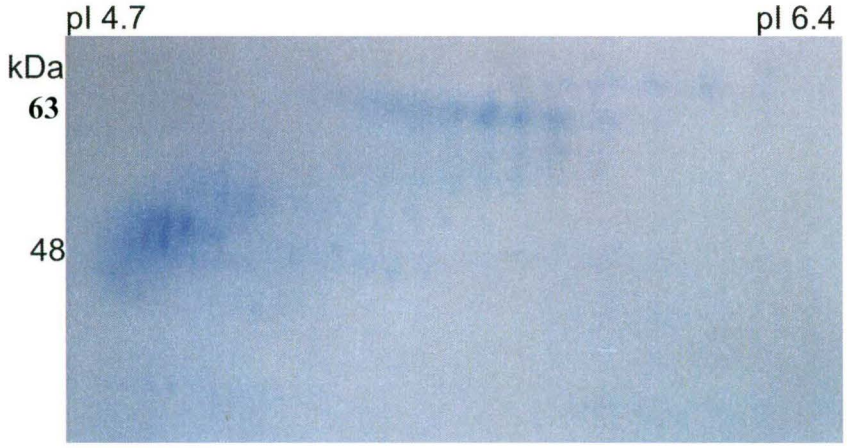


Figure 2.2
2DE gel of wool IFPs. No reductant or degassing was used after the extraction stage. Chaotrope constitution in rehydration buffer was 8 M urea with no thiourea. Equilibration was a single 15 minute step with 0.05 M Tris, 6 M urea, 20% glycerol, 2% SDS, pH 8.8. Proteins (250 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8.5% T gel using a continuous buffer system. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of duplicate experiments.

2.2.3 Choice of reductants

During electrophoresis of wool IFPs, it is very important to maintain reducing conditions to stop reformation of inter- and intra-chain disulfide bonds. Horizontal streaking in the 2DE pattern can be due to the reoxidation of sulfhydryl groups and an accompanying loss of solubility, especially for proteins that interact *via* disulfide bonds [25]. Herbert *et al.* [26] recommended the use of tributyl phosphine (TBP) to overcome this effect. TBP is a non-ionic reducing agent that improves solubility of keratin proteins and should not migrate during IEF [27]. It has been used in 2D-PAGE of wool proteins resulting in decreased horizontal streaking [26]. However, more recently, TBP has been shown to be already consumed during the rehydration of the IPG strips [28]. Tributyl phosphine has been suggested to

be a poor reductant for IEF due to its low solubility and poor stability in rehydration solution [29]. Previously, ß-mercaptoethanol has been used as a reducing agent in IEF rehydration solution. Dithiothreitol (DTT) surpassed ß-mercaptoethanol as an IEF reductant, as ß-mercaptoethanol is required in high concentrations and impurities may result in artefacts [29]. Görg *et al.*, working with mouse liver and bovine vitreous proteins, found that 0.4% DTT in the rehydration solution was the optimal concentration to reduce streaking, as higher amounts of DTT resulted in distorted protein patterns [25, 30]. Dithiothreitol was chosen as the best reductant for use in this work on 2DE of wool proteins. It was used in the rehydration solution at a concentration of 0.4%, which helped to reduce horizontal streaking, and increased solubility of the IFPs.

Dithiothreitol can be depleted during IEF in alkaline conditions. Dithiothreitol is negatively charged at alkaline pH and will therefore migrate towards the anode during IEF, which will deplete the cathodic end of the gel of DTT [27]. Dithiothreitol in water was added to the cathodic wick, so any DTT that was carried towards the anode could be replenished [30]. This was added at the same concentration as the DTT in the rehydration solution. Degassing the paraffin oil prior to focusing was another way to minimise the chance of protein reoxidation during isoelectric focusing [31].

2.2.4 Chaotropes

The chaotropic constitution of the rehydration solution is very important. Chaotropic agents allow proteins to unfold and expose their hydrophobic cores by changing their hydrogen bond structure in the solution, and decreasing the energy penalty for contact of the hydrophobic residues with the solution [27]. Urea is a neutral chaotrope and is always in the 2DE sample solution. It minimises protein aggregation by decreasing intra- and inter-protein hydrogen bonds and hydrophobicity interactions, whilst maximising protein interaction with the buffer system [32].

Sample solubilization has been improved by the use of additional chaotropes such as thiourea [33]. In 1997, Rabilloud *et al.* [34] discovered that the

presence of 2 M thiourea was a key factor for increasing protein solubility in membrane and nuclear proteins. The effect of thiourea concentration of 1-3 M was evaluated. Below 2 M, protein solubility did not improve significantly, whereas thiourea concentrations above 2 M resulted in streaking [34]. The optimal conditions for using thiourea are 2 M thiourea in 5-7 M urea. Thiourea is poorly soluble in water and requires a high concentration of urea for solubility [27]. As well as contributing to disrupting hydrogen bonds, 2 M thiourea produces the same amount of sulfhydryl reducing groups as 0.1 M DTT [35].

2.2.5 Equilibration

Post-IEF equilibration of the IPG strip is very important to ensure proteins are denatured during the second dimension. Sodium dodecyl sulfate is an anionic detergent, which binds to proteins in a mass ratio of 1.4:1 [29]. This gives the proteins a negative charge. To fully bind SDS, it is essential that the proteins are unfolded and all disulfide bonds are broken [26]. Keratin proteins have many disulfide bonds; therefore incorporation of a reductant into the equilibration solution should resolubilise any proteins that may have cross-linked in the IPG strip [26]. This was achieved by a 15 minute equilibration with a solution containing 1% DTT and 2% SDS.

The length of the equilibration time is very important. Focused proteins bind strongly to the IPG gel matrix, therefore prolonged equilibration time is important to desorb these proteins and allow interaction with SDS to get improved protein transfer from the first to the second dimension [25, 30]. A second equilibration step with iodoacetamide (IAM) is important to remove excess DTT from the IPG strip [36].

Urea in the equilibration solution reduces the effects of electroendosmosis by increasing the viscosity of the buffer, which improves transfer from the first to the second dimension. Thiourea, however, can not be added to the equilibration solution as it inhibits protein binding of SDS, leading to poor transfer of protein from the first to the second dimension [37].

2.2.6 Improved isoelectric separation of wool IFPs

The improvements made to the rehydration, IEF and equilibration steps all led to improved solubility and transfer of the wool IFPs from the first to the second dimension. The changes made to the methodology also greatly improved the first dimension resolution.

The type I IFPs have separated out into many horizontal rows of spots. There are four main horizontal rows containing the most abundant proteins. There are several minor spots below and to the basic side of the major type I IFPs. Many of these proteins have been resolved into discrete spots (Figure 2.3).

The type II IFPs have separated out at a higher molecular weight and pI. There are two distinct rows of major type II IFPs, the more basic proteins having a higher apparent molecular weight than the more acidic ones. There are many minor rows of spots below the acidic half of the type II IFPs (Figure 2.4).

When compared to a gel run under conditions which have been used in the past (Figure 2.2) a marked improvement in the resolution can be seen. The amount of protein loaded onto each IPG gel was the same (250 µg); however, using the improved method the gel appears to have much less protein lost during the 2DE process. Many spots are resolved that were not seen using past methods.

2.2.7 Improved second dimension separation of wool IFPs

Once the first dimension running conditions had been optimised, the second dimension conditions were altered to get the best vertical electrophoretic separation possible. Changes were made to the running buffers, stacking gels and run times. Gels that were run with the discontinuous buffer system showed good separation of both the type I and type II IFP families (Figure 2.5). However, the resolution when using a continuous buffer system was slightly better than the discontinuous buffer system. The continuous buffer

system has the added advantage of being able to use the same buffer for both the anode and cathode.

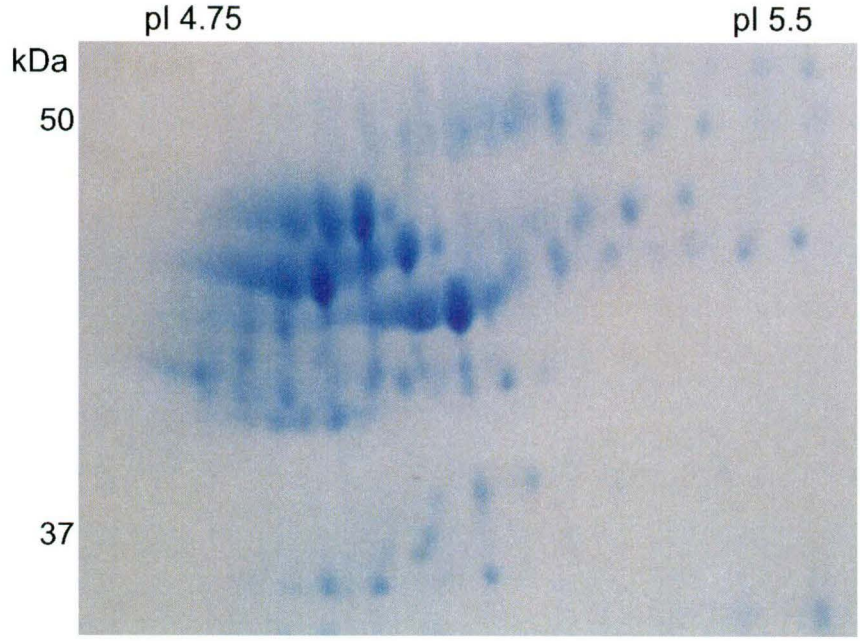


Figure 2.3
2DE separation of type I IFPs using extraction method 1. Proteins (250 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8.5% T gel using a continuous buffer system. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of triplicate experiments.

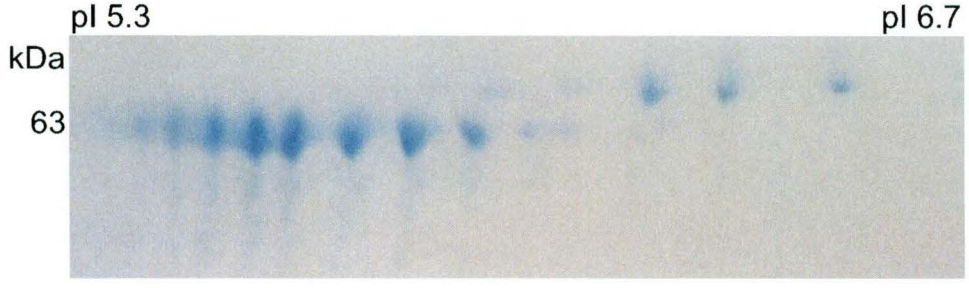


Figure 2.4
2DE separationl of wool type II IFPs using extraction method 1. Proteins (250 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8.5% T gel using a continuous buffer system. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of triplicate experiments.

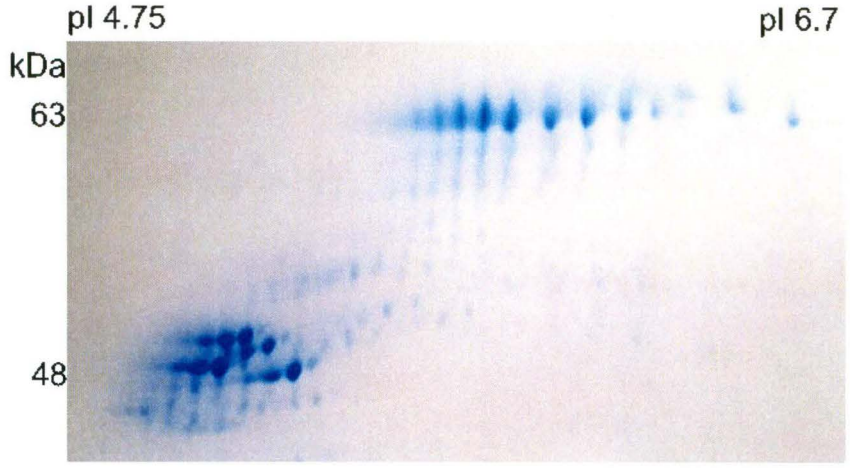


Figure 2.5
2DE separation of wool IFPs using a discontinuous buffer system. Extraction Method One was used. Proteins (250 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8.5% T gel. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of duplicate experiments.

Overnight second dimension runs were initially used to separate IFPs [23]. Separation of the IFPs was adequate with such runs (Figure 2.6). Changing to a faster one-day run appeared to improve resolution of the spot rows. Faster second dimension runs generally produce less diffuse spots, which results in better resolution, sensitivity and better enzyme kinetics for in-gel tryptic digestions prior to mass spectrometry analysis [38].

Stacking gels have been used in gel electrophoresis to concentrate the protein sample before entering the separating gel. Stacking gels did not improve the vertical resolution of the IFPs (Figure 2.7). In 2DE, when using IPG strips a stacking gel is generally considered to be unnecessary, the IPG strip already contains concentrated bands of protein in a low % T matrix, and acts as a small stacking gel [39]. There are several problems associated with the use of stacking gels. The edge between the stacking gel and the separating gel may contain unpolymerised acrylamide. This can cause spurious alkylation of proteins [38]. Additionally, proteins may become trapped between the stacking gel and the separating gel. Advantages of not using a stacking gel include the convenience of not having to cast a stacking

gel before each gel run, and the reduced chance of variability between gel casting.

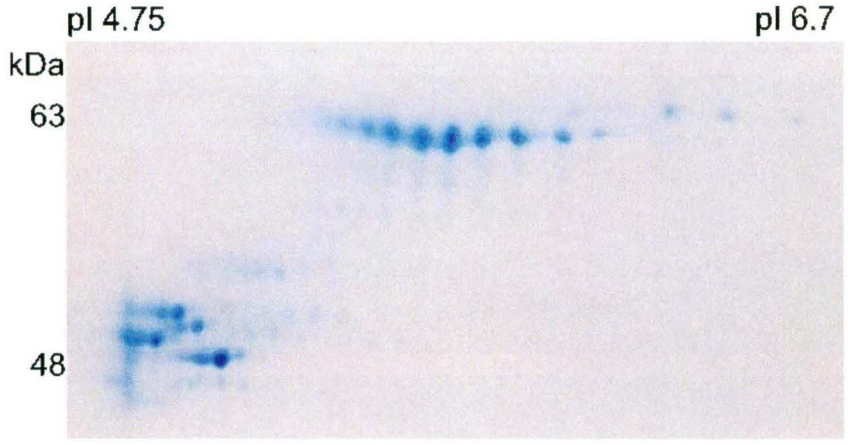


Figure 2.6
2DE separation of wool IFPs using an overnight second dimension run. Extraction Method One was used. Proteins (250 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8.5% T gel. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of duplicate experiments.

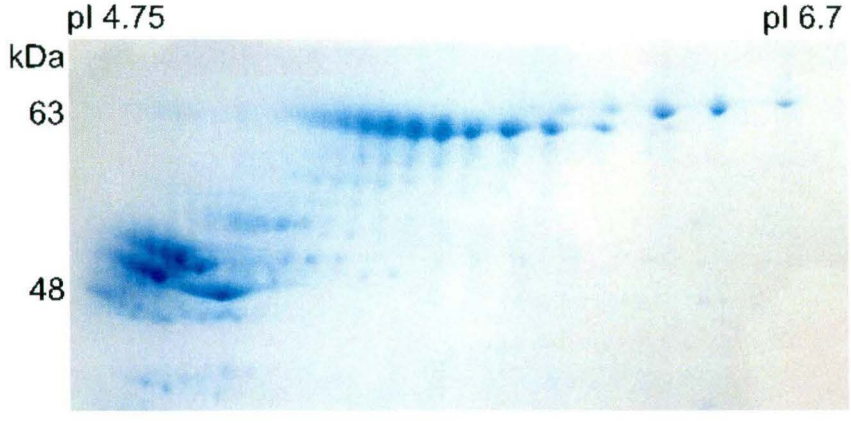


Figure 2.7
2DE separation of wool IFPs using a 4% T stacking gel cast above the separating gel. Extraction Method One was used. Proteins (500 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8% T gel. Gel was stained with Coomassie brilliant blue G-250. The gel is representative of duplicate experiments.

2.2.8 Improved extraction of wool IFPs

Once optimal conditions for second dimension separation were determined, an alternative wool protein extraction method was developed to simplify the sample preparation steps prior to the first dimension separation.

The aim of any protein extraction method is to disrupt all molecular interactions to ensure that each spot on the resulting gel represents a single protein, without creating artifacts [39]. The standard wool protein extraction method was originally based on the O'Farrell method, with the main components being urea and reductant [40]. This method is relatively gentle, which means it will not necessarily disrupt all protein complexes. All extraction methods should be as simple as possible to increase reproducibility: they also need to minimise the risk of modification during sample preparation [39].

Recently Harder *et al.* ran a series of experiments comparing extraction methods for yeast cell proteins [41]. They used three different extraction methods, and discovered that disruption in the presence of SDS instead of urea reduced the amount of protein degradation seen. The cell lysis buffer composition was altered to include thiourea and the concentration of 3-[(3-cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS) was increased. This resulted in an increase of detectable spots of more than 50% [41].

This method of extraction was adapted for use with wool proteins. In the initial SDS sample buffer, a reductant was included to promote cleavage of disulfide bonds. The amount of Tris was halved to be more compatible with subsequent IEF and the pH was increased to gain better reduction of disulfide bonds [42]. The sonication power was increased from 60 W to 150 W to improve disruption of the cells. Other changes made to the Harder *et al.* method included increasing the % of DTT in the lysis buffer to maintain the cysteines in their reduced state, and increasing the time of shaking in the lysis buffer from one hour to four hours.

This extraction method, combined with the improvements made to the first dimension isoelectric focusing and the second dimension molecular weight separation, gave the best resolution of wool IFPs seen to date (Figure 2.8). Both the type I and type II IFPs are separated in the horizontal and vertical dimensions. Individual spots are easily discerned, which makes subsequent mass spectrometric analysis easier.

The improved method for extraction eliminates the use of dialysis and freeze-drying between the extraction and rehydration steps. There is the potential for loss of protein in both these steps. The problem of trying to redissolve protein after freeze-drying has also been eliminated. Removing both of these steps from the protocol allowed the extraction and rehydration for first dimension to be achieved in one day. This saved three days in comparison to the original method [13, 23, 24].

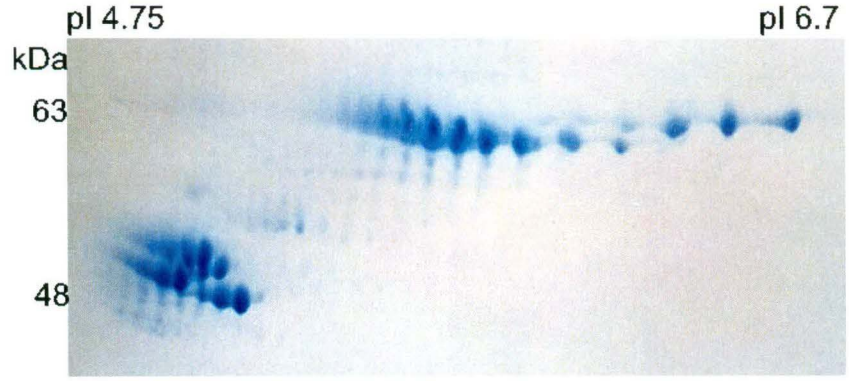


Figure 2.8
2DE separation of wool IFPs using Extraction Method Two. The new extraction method involves sonication, boiling, and shaking of the wool in sample and lysis buffers prior to rehydration. Proteins (470 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on an 8.5% T gel using a continuous buffer system. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of triplicate experiments.

2.2.9 Spot labelling design

To assist with identification of the major protein spots of both the type I and type II IFPs a numbering system was designed. The spots were numbered from the acidic end of the gel, from highest molecular weight to lowest. The major type I IFP spots were numbered 1-11 (Figure 2.9) and the major type II IFPs were numbered 12-24 (Figure 2.10).

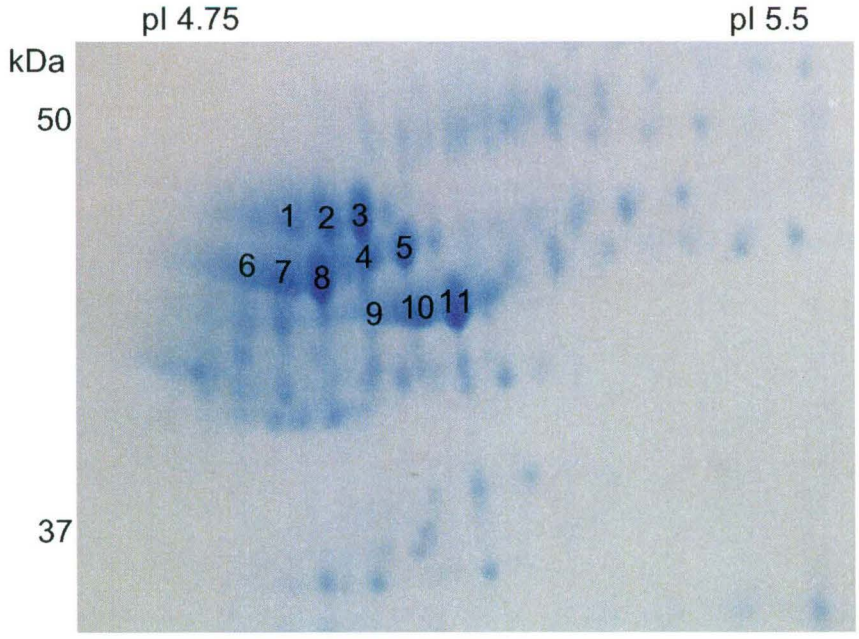


Figure 2.9
Major wool type I IFP spots labelled 1-11

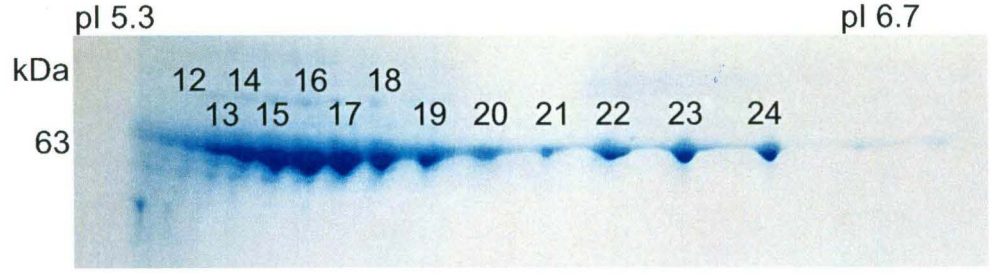


Figure 2.10
Major wool type II IFP spots labelled 12-24

2.3 Summary

Changes have been made to the way wool proteins are extracted and separated on 2D-PAGE, which allowed the high-resolution separation of the two keratin IFP classes and their individual family members. The wool IFPs have been separated by 2DE to produce a spot pattern with clearly defined spots and many minor spots not seen using past methods. Further work to analyse the composition of each of the protein spots has been made much easier by being able to separate the IFPs into discrete spots. Experiments in following chapters analyse the individual spot rows to discern if PTMs are responsible for the differences in pI.

2.4 References

1. Fuchs E, Weber K. Intermediate Filaments: Structure, Dynamics, Function, and Disease. *Annual Review of Biochemistry* **1994**;63:345-382
2. Powell BC, Rogers GE. The Role of Keratin Proteins and Their Genes in the Growth, Structure and Properties of Hair. In: Jollès P, Zahn H, Höcker H, eds. *Formation and Structure of Human Hair*. Basel/Switzerland: Birkhäuser Verlag, **1997**:59-148
3. Crewther WG, Dowling LM, Gough KH, Inglis AS, McKern NM, Sparrow LG, Woods EF. The Low-Sulphur Proteins of Wool: Studies on Their Classification, Characterisation, Primary and Secondary Structure. In: *Proceedings of the 5th International Wool Textile Research Conference*. Aachen, Germany, **1975**:233-242
4. Powell BC, Beltrame JS. Characterisation of a Hair (Wool) Keratin Intermediate Filament Gene Domain. *Journal of Investigative Dermatology* **1994**;102:171-177
5. Heid HW, Werner E, Franke WW. The Complement of Native Alpha-Keratin Polypeptides of Hair-Forming Cells: A Subset of Eight Polypeptides that Differ from Epithelial Cytokeratins. *Differentiation* **1986**;32:101-119
6. Fraser RDB, MacRae TP, Rogers GE. *Keratins*. Springfield, Illinois, USA: Charles C Thomas, **1972**:1-304

7. Goddard DR, Michaelis L. A Study on Keratin. *Journal of Biological Chemistry* **1934**;106:605-614
8. Parry DAD, Steinert PM. Heidelberg, Germany: Springer-Verlag, **1995**:1-183
9. Achtstaetter T, Hatzfeld M, Quinlan RA, Parmelee DC, Franke WW. Separation of Cytokeratin Polypeptides by Gel Electrophoresis and Chromatographic Techniques and Their Identification by Immunoblotting. *Methods in Enzymology* **1986**;134:355-371
10. Plowman JE. Proteomic Database of Wool Components. *Journal of Chromatography B* **2003**;787:63-76
11. Woods JL, Orwin DFG. Wool Proteins of New Zealand Romney Sheep. *Australian Journal of Biological Science* **1987**;40:1-14
12. Marshall RC. Analysis of the Proteins from Single Wool Fibers by Two-Dimensional Polyacrylamide Gel Electrophoresis. *Textile Research Journal* **1981**;51:106-108
13. Herbert BR, Woods JL. Immobilised pH Gradient Isoelectric Focusing of Wool Proteins. *Electrophoresis* **1994**;15:972-976
14. Herbert BR, Molloy MP, Yan JX, Gooley AA, Bryson WG, Williams KL. Characterisation of Wool Intermediate Filament Proteins Separated by Micropreparative Two-Dimensional Electrophoresis. *Electrophoresis* **1997**;18:568-572
15. Plowman JE, Bryson WG, Flanagan LM, Jordan TW. Problems Associated with the Identification of Proteins in Homologous Families: The Wool Keratin Family as a Case Study. *Analytical Biochemistry* **2002**;300:221-229
16. O'Donnell IJ, Thompson EOP. Studies on Reduced Wool. IV. The Isolation of a Major Component. *Australian Journal of Biological Science* **1964**;17:973-989
17. Crewther WG, Dowling LM, Gough KH, Marshall RC, Sparrow LG. The Microfibrillar Proteins of Alpha-Keratin. In: Creamer LK, ed. *Fibrous Proteins: Scientific, Industrial and Medical Aspects*. London: Academic Press Inc. (London) Ltd., **1980**:151-159

18. Harrap BS, Gillespie JM. A Further Study on the Extraction of Reduced Proteins from Wool. *Australian Journal of Biological Science* **1963**;16:542-557
19. Dowling LM, Crewther WG, Parry DAD. Secondary Structure of Component 8c-1 of Alpha-Keratin. *Biochemical Journal* **1986**;236:705-712
20. Sparrow L, Robinson C, Caine J, McMahon D, Strike P. Type II Intermediate-Filament Proteins from Wool. The Amino Acid Sequence of Component 5 and Comparison with Component 7c. *Biochemical Journal* **1992**;282:291-297
21. Wilson BW, Edwards KJ, Sleigh MJ, Byrne CR, Ward KA. Complete Sequence of a Type-I Microfibrillar Wool Keratin Gene. *Gene* **1988**;73:21-31
22. Marshall RC, Blagrove RJ. Successful Isoelectric Focusing of Wool Low-Sulphur Proteins. *Journal of Chromatography* **1979**;172:351-356
23. Plowman JE, Bryson WG, Jordan TW. Application of Proteomics for Determining Protein Markers for Wool Quality Traits. *Electrophoresis* **2000**;21:1899-1906
24. Plowman JE, Flanagan LM, Paton LN, Fitzgerald AC, Joyce NI, Bryson WG. The Effect of Oxidation or Alkylation on the Separation of Wool Keratin Proteins by Two-Dimensional Gel Electrophoresis. *Proteomics* **2003**;3:942-950
25. Görg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W. The Current State of Two-Dimensional Electrophoresis with Immobilized pH Gradients. *Electrophoresis* **2000**;21:1037-1053
26. Herbert BR, Molloy MP, Gooley AA, Walsh BJ, Bryson WG, Williams KL. Improved Protein Solubility in Two-Dimensional Electrophoresis Using Tributyl Phosphine as Reducing Agent. *Electrophoresis* **1998**;19:845-851
27. Herbert BR. Advances in Protein Solubilisation for Two-Dimensional Electrophoresis. *Electrophoresis* **1999**;20:660-663
28. Berkelman T. Use of a Novel Reductant for Improved 2-D Separations of Basic Proteins. In: *Proteomics Forum*. Munich, **2001**

29. Berkelman T, Steastedt T. *2D Electrophoresis Using Immobilized pH Gradients. Principles and Methods*, AB ed. Uppsala, Amersham: Amersham Pharmacia Biotech, **1998**:1-100

30. Görg A, Boguth G, Obermaier C, Posch A, Weiss W. Two-Dimensional Polyacrylamide Gel Electrophoresis with Immobilized pH Gradients in the First Dimension (IPG-Dalt): The State of the Art and the Controversy of Vertical *versus* Horizontal Systems. *Electrophoresis* **1995**;16:1079-1086

31. Görg A, Obermaier C, Boguth G, Weiss W. Recent Developments in Two-Dimensional Gel Electrophoresis with Immobilized pH Gradients: Wide pH Gradients up to pH 12, Longer Separation Distances and Simplified Procedures. *Electrophoresis* **1999**;20:712-717

32. Sanchez J-C. Sample Preparation and Solubilization: Crucial Steps Preceding the Two-Dimensional Gel Electrophoresis Process. In: *ABRF*. San Diego, **1998**:1-16

33. Molloy MO, Herbert BR, Walsh BJ, Tyler MI, Traini M, Sanchez J-C, Hochstrasser DF, Williams KL, Gooley AA. Extraction of Membrane Proteins by Differential Solubilization for Separation Using Two-Dimensional Gel Electrophoresis. *Electrophoresis* **1998**;19:837-844

34. Rabilloud T, Adessi C, Giraudel A, Lunardi J. Improvement of the Solubilization of Proteins in Two-Dimensional Electrophoresis with Immobilized pH Gradients. *Electrophoresis* **1997**;18:307-316

35. Rabilloud T. Use of Thiourea to Increase the Solubility of Membrane Proteins in Two-Dimensional Electrophoresis. *Electrophoresis* **1998**;19:758-760

36. Yan JX, Sanchez J-C, Rouge V, Williams KL, Hochstrasser DF. Modified Immobilized pH Gradient Gel Strip Equilibration Procedure in SWISS-2DPAGE Protocols. *Electrophoresis* **1999**;20:723-726

37. Galvani M, Rovatti L, Hamdan M, Herbert BR, Righetti PG. Protein Alkylation in the Presence/Absence of Thiourea in Proteome Analysis: A Matrix Assisted Laser Desorption/Ionization-Time Of Flight-Mass Spectrometry Investigation. *Electrophoresis* **2001**;22:2066-2074

38. Weistermeiner R, Naven T. *Proteomics in Practice.* Freiburg: Wiley-VCH Verlag-GmbH Weinheim, **2002**:316

39. Görg A, Weiss W, Dunn MJ. Current Two-Dimensional Electrophoresis Technology for Proteomics. *Proteomics* **2004**;4:3665-3685

40. O'Farrell PH. High Resolution Two-Dimensional Electrophoresis of Proteins. *Journal of Biological Chemistry* **1975**;250:4007-4021

41. Harder A, Wildgruber R, Nawrocki A, Fey SJ, Larsen PM, Görg A. Comparrison of Yeast Cell Protein Solubilization Procedures for Two-Dimensional Electrophoresis. *Electrophoresis* **1999**;20:826-829

42. Galvani M, Hamdan M, Herbert BR, Righetti PG. Alkylation Kinetics of Proteins in Preparation for Two-Dimensional Maps: A Matrix Assisted Laser Desorption/Ionization-Mass Spectrometry Investigation. *Electrophoresis* **2001**;22:2058-2065

Chapter Three

Proteomic Investigations into Phosphorylation as a Post-Translational Modification of Wool Intermediate Filament Proteins

The goal of this chapter was to determine whether the charge heterogeneity of the wool IFPs separated by 2DE was related to phosphorylation of the proteins. A multiplexed approach was taken to determine the phosphorylation state of the proteins.

3.1 Introduction

Many IFs have been shown to be phosphorylated at different stages of their cell cycle (see 1.12.1 for a review). When wool IFPs are separated by 2D-PAGE, both the type I IFPs and the type II IFPs separate out into rows of spots. The type I IFPs range from about pI 4.7 to 5.2 and the type II IFPs from about 5.0 to 6.7. The cause of this charge heterogeneity is unknown. Several common PTMs of proteins could cause charge heterogeneity including phosphorylation, glycosylation, and deamidation [1].

In the soft keratins the polypeptides appear as a series of isoelectric variants when separated by 2DE [2]. Steinert *et al.* used [^{32}P] labelled epidermal cells to show that phosphoserine could be detected on the IFPs using thin layer chromatography and autoradiography [3]. The labelled epidermal filaments were subjected to limited chymotryptic digestion, then separated by chromatography. The α-helical eluted peaks showed no labelling whereas the non α-helical eluted peaks were labelled. The conclusion was that the phosphoserine was located in the non α-helical regions of the IFs [3]. It has

been suggested by Herbert *et al.* [4] that the IFPs of wool are phosphorylated. To demonstrate this, wool IFPs were dephosphorylated with alkaline phosphatase. Two-dimensional electrophoresis gels run with dephosphorylated wool protein showed that in some instances the type II IFPs decreased in intensity at the acidic end of the row and new spots appeared at a more alkaline pI and at a higher molecular weight. This experiment was followed up by Herbert *et al.* [5] in 1997 by amino acid analysis of the IFPs that showed very low levels of phosphoserine and phosphotyrosine and trace amounts of phosphothreonine in some type II IFPs.

More recently, Nakamura *et al.* [6] used Western blot analysis with antibodies against phosphoserine, phosphothreonine and phosphotyrosine to show that the proteins extracted from human hair had both serine and threonine phosphorylations. There was no reactivity with phosphotyrosine antibodies. They suggested a role for phosphorylation in modulating the organisation and structure of microfibrils.

3.2 Characterisation of phosphoproteins

Characterisation of phosphoproteins requires a combination of specific and sensitive analytical techniques [7]. There are many established methods for detecting phosphorylations in protein samples. Prior to the development of proteomic techniques, the most common method for detecting phosphorylations involved [^{32}P] labelling and autoradiography. This is only a useful technique if the sample in question is able to incorporate the radioactive label prior to electrophoretic separation [8].

Amino acid analysis was used in the past to detect phosphorylated amino acids. O-Phosphates (serine, threonine, tyrosine) are stable under acid conditions and therefore can be separated using ion-exchange HPLC after partial acid hydrolysis [8]. This technique allows the detection of phosphorylations but does not indicate where the modifications occur in the

sequence. Other past methods of separation include thin-layer 2D separation and capillary electrophoresis [8].

Gel-based techniques have been used as a means of separation of phosphoproteins before detection using immunoblotting and mass spectrometry. Immunodetection for phosphorylations with antibodies is problematic due to the limited specificity and cross-reactivity of the antibodies [9, 10]. Antibodies against phosphoserine and phosphothreonine usually detect a sequence motif as the antibodies are not specific enough to detect a single phosphorylated serine or threonine as this epitope is too small [8, 11]. Generally, phosphotyrosine antibodies show the best results when detecting phosphorylations.

Enzymatic dephosphorylation followed by 2DE has recently been used as a tool to distinguish between phosphorylated proteins [12]. The protein sample is divided into two aliquots, one is dephosphorylated using a phosphatase and the other is not treated with the enzyme. The two aliquots are then subjected to 2DE and the differences in spot positions are mapped. The spots that show a shift to a more basic pI can then be analysed by mass spectrometry.

Mass spectrometry has become the method of choice for identification of phosphorylation sites [13]. Phosphate locations can be detected using a combination of precursor ion scanning, peptide mass fingerprinting and product ion scanning. The loss of 79 Da (HPO_3^{2-}) and 98 Da (H_3PO_4) indicate which peptides are phosphorylated [8]. Mass spectrometry allows the determination of the type and exact sites of modification in the sequence. Modifications to the mass spectrometry techniques used, such as the method of ionisation, immobilised metal ion affinity chromatography (IMAC), alkaline phosphatase treatment and alternative enzymes, are being developed to improve the detection of small amounts of phosphate [7, 14, 15].

Recently, fluorescent stains have been developed that are very specific for detecting phosphorylations. Most of the stains are mass spectrometry

compatible, the spots that are identified as possibly containing a phosphorylation are easily transferred from the gel and into the mass spectrometer. Specifically stained gels can also be subsequently stained with whole protein stains to act as a control gel. This reduces inaccuracies due to differences in gel running.

3.3 Results and Discussion

To determine whether the keratinised IFPs from wool were post-translationally modified by phosphorylation, several different proteomic techniques were employed, including on-gel staining, enzymatic dephosphorylation, immunoblotting and mass spectrometry.

3.3.1 Alkali-labile phosphoprotein stain

This method of detection of phosphorylated proteins on polyacrylamide gels relies on the alkaline lability of phosphoserine and phosphothreonine residues [16]. The phosphate is initially released, after running the gel, by alkaline hydrolysis, before precipitation with calcium. An insoluble phospho-molybdate complex is then formed which is stained by a basic dye [16]. Figure 3.1 shows a 2DE gel, which has been stained with the alkali-labile phosphoprotein stain. Proteins show up as white spots against the pink background. The positive control lane shows phosvitin as a dark pink spot. The negative control, soybean trypsin inhibitor, stained as white bands (not visible in Figure 3.1 due to its low molecular weight). All of the IFPs were detected on the gel as proteins that contained no phosphorylations.

3.3.2 Dephosphorylation

Lambda protein phosphatase was chosen as the enzyme for dephosphorylation as it will dephosphorylate all known phosphoamino acid residues [12]. α-Casein was dephosphorylated as a control protein. The major α-casein spot shifted to a more basic pI, suggesting that the dephosphorylation conditions were successful (data not shown). In contrast, two-dimensional electrophoresis gels of wool IFPs after phosphatase treatment showed no compression of the rows of spots, or a change in

molecular weight of any of the proteins (Figure 3.2). This suggests that either the proteins have not been dephosphorylated by the conditions employed, or that phosphorylation is not the cause of the pI heterogeneity of the IFPs.

Figure 3.1
Alkali-labile phosphoprotein staining of wool IFPs on a 2DE gel. A pH 4-7 IPG strip containing 1 mg of focused protein was run in the second dimension on a 10-20% T criterion gel followed by detection with an alkali labile phosphoprotein stain. Phosvitin is shown at the left hand side of the gel as a positive control. The gel is representative of triplicate experiments.

Previous experiments by Herbert *et al.* [4] on wool IFPs showed changes in spot patterns of the type II IFPs after dephosphorylation and separation by 2D-PAGE with the first dimension separation using pH 3-10 IPG strips. The original spots decreased in intensity and new spots at a higher molecular weight and pI appeared. The new spots were presumed to be the dephosphorylated spots. However, when dephosphorylated wool IFPs were run on pH 4-7 IPG strips no shift in pI was observed. The higher molecular weight of dephosphorylated proteins on the pH 3-10 IPG strip was explained by possible differences in the SDS binding of these proteins after dephosphorylation. No explanations were given for why SDS would bind differently if the protein had been dephosphorylated. If phosphorylations were removed from a protein, one would expect a decrease in the molecular weight of 80 Da per phosphate. The loss of a single phosphate group would not be detected on a gel, however, if several phosphates were lost, a slight decrease in molecular weight would be seen. Evidence of this is shown by SDS-PAGE of a phosphorylated precursor of stratum corneum basic protein

(SCBP) [17]. This protein shows a higher apparent molecular weight than the non-phosphorylated SCBP, even though the amino acid sequences are virtually the same. It is possible that different protein samples will contain different amounts of phosphate. When Steinert *et al.* [3] analysed bovine IF subunits for phosphoserine content some showed phosphorylation and some did not. The reason for this variance was not clear [3].

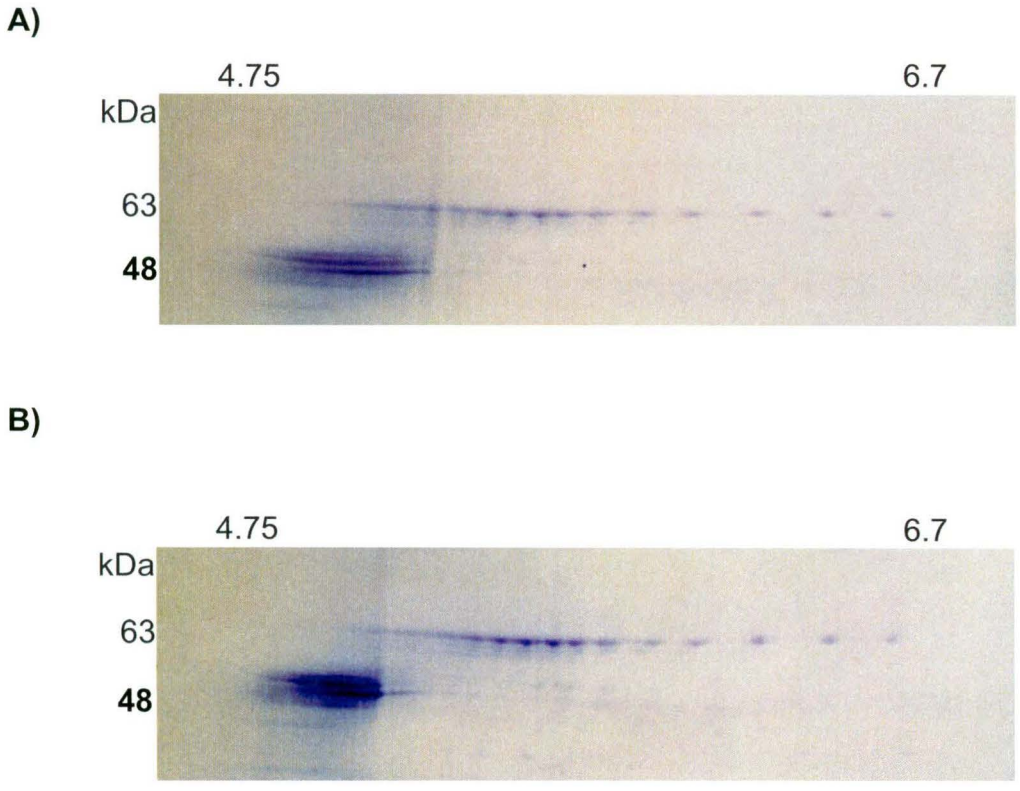


Figure 3.2
Enzymatic dephosphorylation of wool IFPs. A) A pH 4-7 IPG strip containing 400 µg of focused dephosphorylated protein and B) a pH 4-7 IPG strip containing 400 µg of focused protein prior to dephosphorylation were run in the second dimension on a 10-20% T criterion gels. The gels were stained with Coomassie brilliant blue R-250. The gels are representative of duplicate experiments.

As demonstrated in Chapter Two, with optimum 2D-PAGE conditions the type II IFPs can be separated vertically, with the proteins with a more alkaline pI migrating at a slower rate than the more acidic type II IFPs. This may explain the differences in molecular weight seem by Herbert *et al.* [4] during

his dephosphorylation experiment, as the 2DE gels showed poor reproducibility and lacked resolution in both the first and second dimensions.

Herbert *et al.* [5] followed up their experiments with amino acid analysis of several type II IFPs. They found small amounts of phosphorylation in the proteins but failed to use spiking experiments with standard phosphorylated amino acids to prove that the peaks that were being assigned to phosphorylated amino acids were indeed eluting at the same time as the standards. The small peaks that were found are likely to have been artifactual.

3.3.3 Phosphoserine Immunoblotting

Serine was chosen as the most likely amino acid of wool IFPs to be phosphorylated based on previous studies of IFPs [5, 18-22]. Standard immunoblotting conditions were employed using an anti-phosphoserine primary antibody. Figure 3.3 shows the film of the 2DE gel blot after chemiluminescent detection. Three of the IF type II IFPs have reacted with the antibody (arrows in Figure 3.3), and the area where the type I IFPs are has also reacted (circled in Figure 3.3). Figure 3.4 shows the standard lane from the chemiluminescent blot. The only protein that has reacted is lysozyme at 14 400 kDa. Lysozyme is one of the negative controls for phosphoprotein detection. This suggests that there is non-specific binding between the bound proteins and the primary or secondary antibodies. This non-specific binding could have been caused by contamination of the primary, or secondary antibodies, or the monoclonal antibodies are cross-reacting with the SDS denatured proteins [8-10, 23].

The results shown here with monoclonal antibodies show the importance of running control proteins when using immunodetection techniques. The results of Nakamura *et al.* [6], showing phosphoserine and phosphothreonine antibodies reacting with hair proteins, may be misleading due to the lack of control proteins to rule out cross reactivity of the antibodies with the hair proteins. They found that the type I IFPs reacted more than the type II IFPs, which correlates with the results in this work. They also found that the matrix

proteins cross-reacted with the phosphoserine antibody, which has not been observed before.

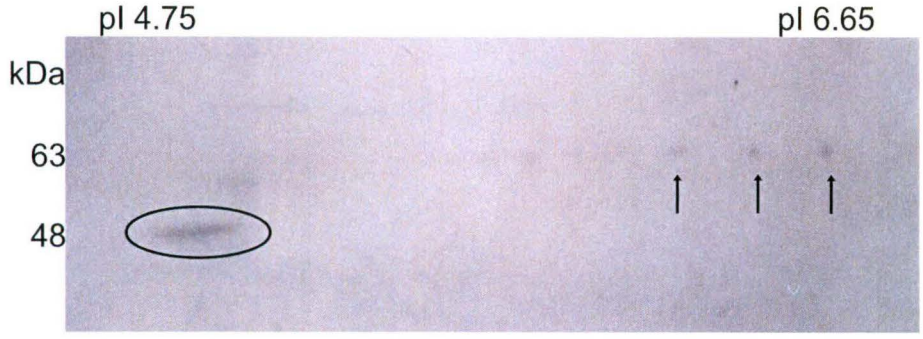


Figure 3.3
Wool IFPs after 2DE followed by transfer to Polyvinylidene fluoride (PVDF) and chemiluminescent detection. Arrows show Type II IFPs that have reacted with the antibodies. Circled area shows the Type I IFP region, which has reacted with the antibodies. The immunoblot is representative of triplicate experiments.

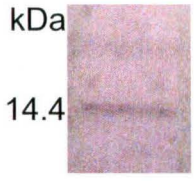


Figure 3.4
PeppermintStick™ phosphoprotein molecular weight standards after 1DE followed by transfer to PVDF and chemiluminescent detection run in conjunction with the 2DE gel in Figure 3.3. The standards contain two phosphorylated proteins and four non-phosphorylated proteins. The only protein to react with the antibodies was lysozyme, which is one of the negative control proteins.

The apparent random reactivity of the antibodies with the wool IFPs means that no conclusions about the phosphorylation states of the proteins can be drawn. The most likely amino acid candidate for phosphorylation was serine, with threonine being detected in small amounts in epidermal keratins and the amount of phosphotyrosine is either very small or nil [3]. Further

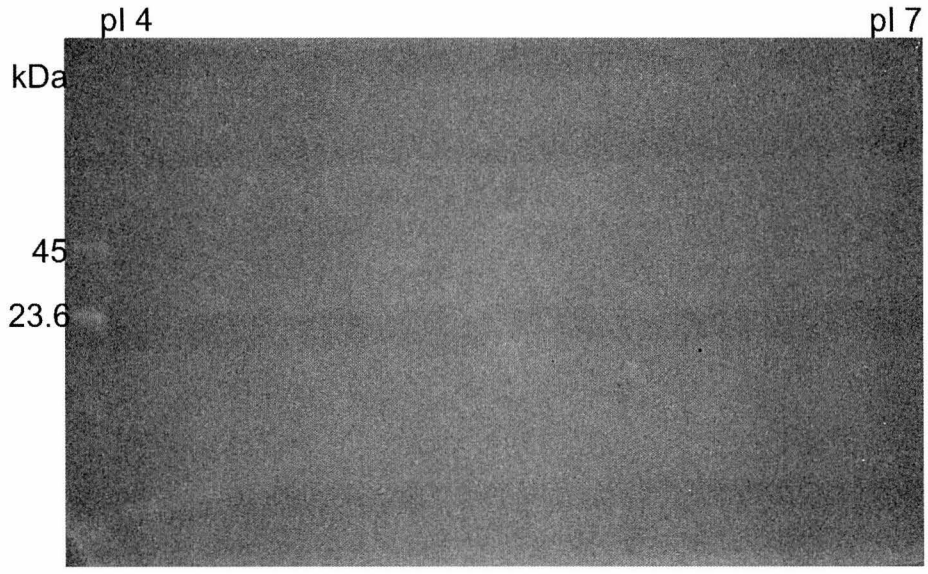
immunodetection using the wool IFPs was ruled out as the phosphoserine immunoblotting demonstrated the ambiguous nature of phosphorylation detection in the wool proteins with antibodies, and there was a low likelihood of detecting phosphothreonine or phosphotyrosine based on previous studies.

3.3.4 Pro-Q® Diamond phosphoprotein gel stain

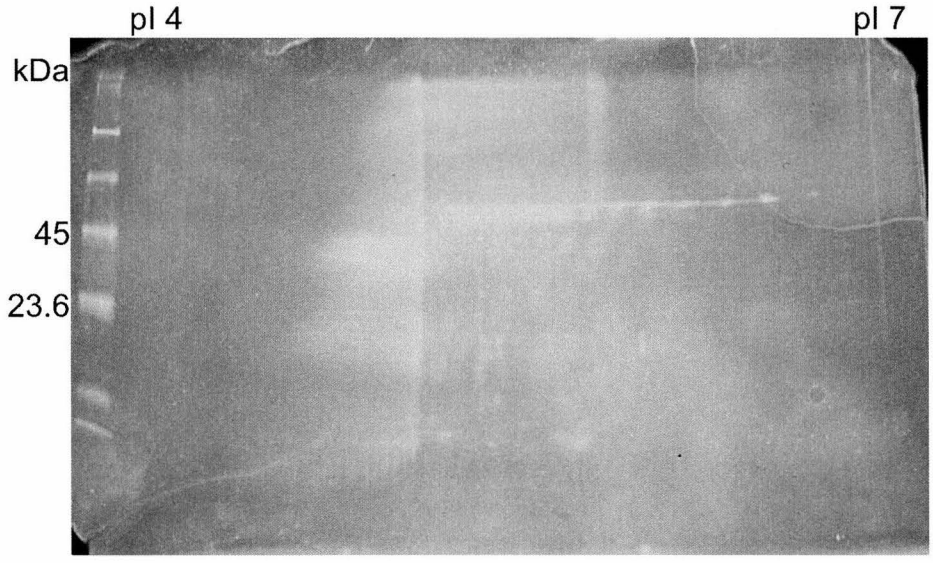
The Pro-Q® Diamond phosphoprotein gel stain allows selective detection of phosphorylated proteins in polyacrylamide gels [24]. The stain will detect any phosphate groups attached to serine, threonine or tyrosine residues. The staining technique has a very low detection level of 1-16 ng, depending on the degree of phosphorylation.

When the 2DE gels of IFPs from wool were incubated with Pro-Q® Diamond phosphoprotein stain, none of the IFPs were stained (Figure 3.5 A). The PeppermintStick™ phosphoprotein molecular weight markers showed two stained proteins at the molecular weights of 45 kDa and 23.6 kDa. These correspond to the two positive control proteins, ovalbumin and β-casein. This confirms that the staining technique was selective for phosphorylated proteins, and did not stain any non-phosphorylated proteins. Figure 3.5 B shows the same gel from Figure 3.5 A after it had been stained by the whole protein stain, Sypro® Ruby. All of the IFP and matrix proteins have been stained, as well as all of the positive and negative control proteins, showing that the proteins were loaded onto the gel at a concentration that should be sufficient to detect any phosphorylation. A Coomassie brilliant blue G-250 stained gel with equivalent loading of IFPs (Figure 3.5 C) shows that the sensitivity of the Sypro® Ruby stain is slightly less than the Coomassie brilliant blue G-250 stain, due to the methods used to visualise and document the fluorescent Sypro® Ruby stain.

A)



B)



C)

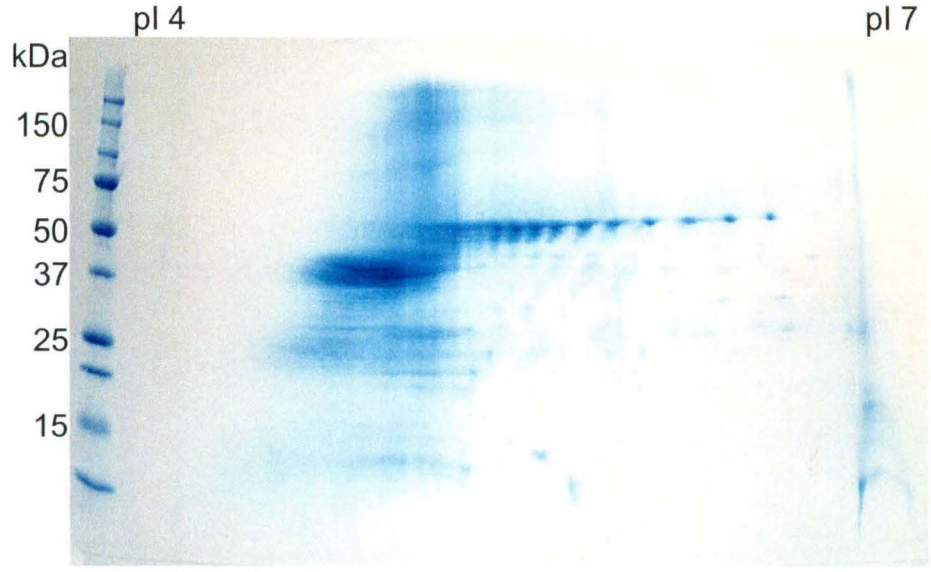


Figure 3.5
A) Pro-Q® Diamond phosphoprotein staining of wool IFPs on a 2DE gel. A pH 4-7 IPG strip containing 100 µg of focused protein was run in the second dimension on a 10-20% T criterion gel followed by detection with Pro-Q® Diamond phosphoprotein stain. Positive and negative control proteins were run in a well to the left of the gel (PeppermintStick™ phosphoprotein molecular weight standards). B) The same gel from A after it has been stained with Sypro® Ruby. All of the proteins on the gel have been stained including all of the control proteins. C) A comparative gel stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

3.3.5 Mass Spectrometry

3.3.5.1 Method validation with α-casein

Mass spectrometry was used to determine the location of any phosphorylations, rather than amino acid analysis, as it has a lower detection limit and can give additional information, such as the exact amino acid residue that has been modified [8]. Initially, α-casein was used as a model phosphorylated protein to establish the detection limit of the mass spectrometer.

Peptide mass fingerprint scans were initially matched to the database sequence of α-casein to determine that the digestion had worked well and

that the peptides could be easily detected in the mass spectrometer (Figure 3.6).

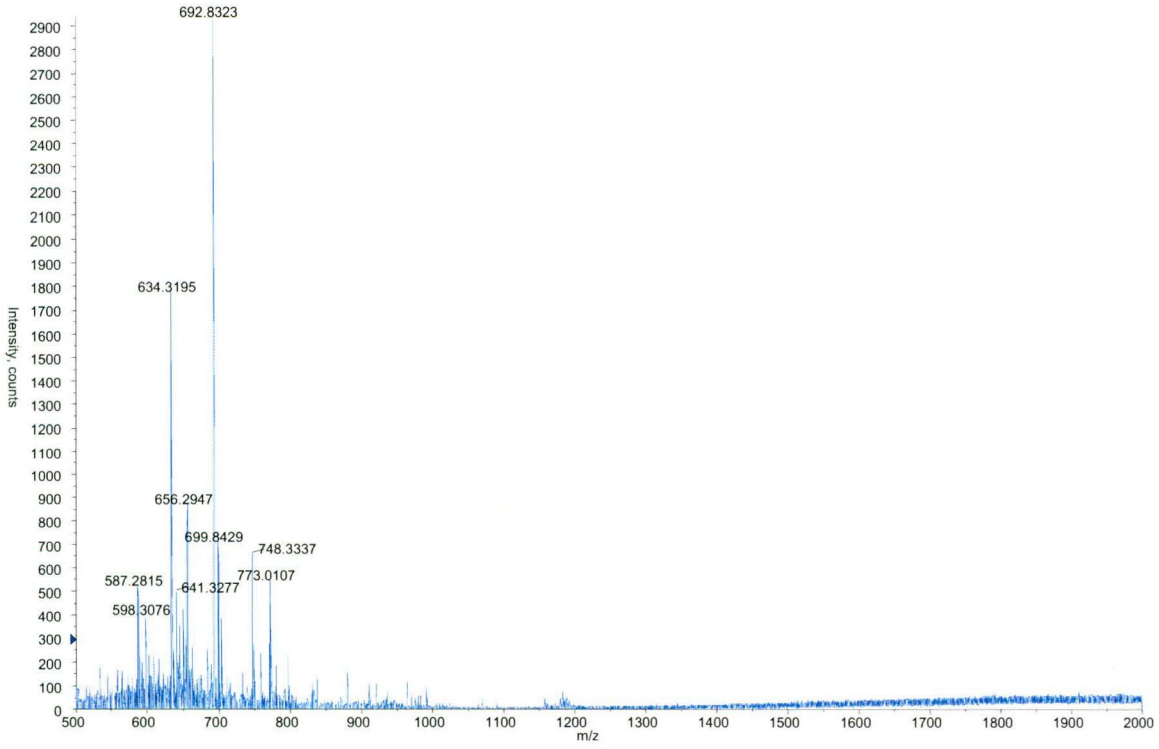


Figure 3.6
Peptide mass fingerprint of α-casein digested with trypsin. 102 nM was analysed in positive scanning mode. The spectrum is representative of triplicate experiments.

The peptide mass fingerprint data matched to an α_{s2}-casein sequence (20% sequence coverage) (Appendix One A1.1).

The precursor ion scan, detecting the loss of 79 Da, showed many potentially phosphorylated masses (Figure 3.7) (Appendix One A1.2). Closer inspection ruled out some singularly charged masses as not originating from peptides.

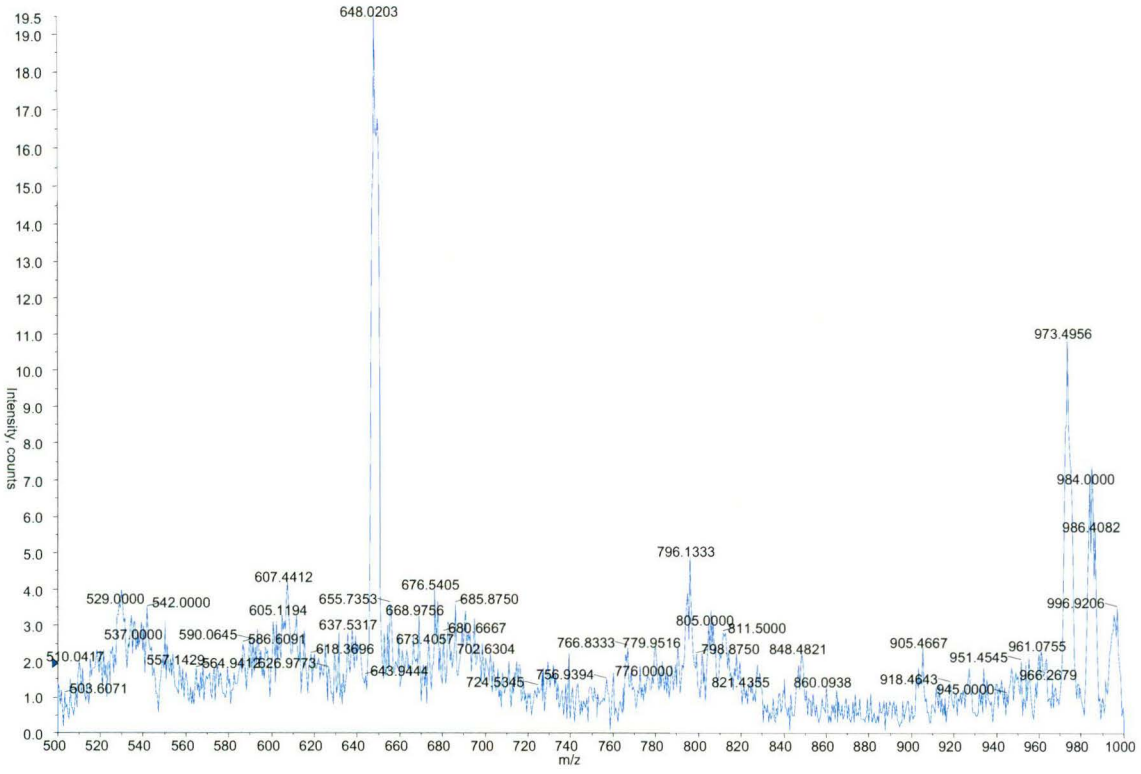


Figure 3.7
Precursor ion scan for the loss of 79 Da (HPO_3^{2-}) from α-casein. 102 nM was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

Once peptides were identified from the precursor ion scan they were subjected to product ion scanning to determine their sequence (Appendix One A1.3). An example of a product ion scan is shown in Figure 3.8. The neutral loss of H_3PO_4 (98 Da) from the doubly charged peak 730 gives the 681 peak. The doubly charged peptide was then confirmed by *de novo* sequencing to be QLST[pS]EENSK.

The sensitivity limit for the detection of phosphopeptides was tested using a tryptic digest of α-casein from a 2DE gel. At 200 fmol of total digest, all the major phosphopeptides could be seen. At 20 fmol the most abundant phosphopeptides could still be detected. From these data, we can assume that the amount of wool protein loaded onto 2DE gels (commonly in the mg range) is in large excess compared to the amount required for positive phosphorylation detection in the mass spectrometer. When working with a pure phosphopeptide, commonly less than 100 fmol is needed for the

localisation as assignment of phosphorylated amino acids [8]. In mixtures with unphosphorylated peptides, the ionisation rate of the phosphopeptide will be impaired and more starting peptide material is needed [8]. Phosphorylation has been detected in myosin I heavy-chain kinase, which was separated by gel electrophoresis. A single gel slice containing 3 pmol of protein was sufficient to identify 12 phosphorylation sites [25].

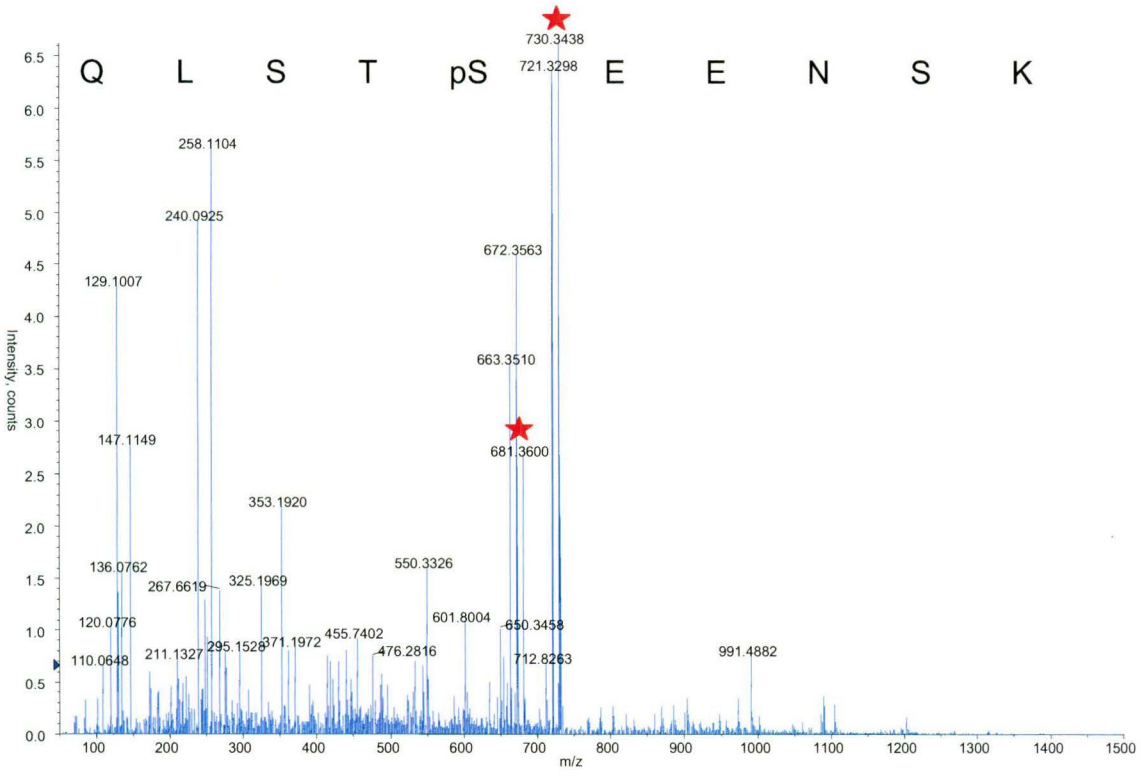


Figure 3.8
Product ion scan of the doubly charged peptide at m/z 730^{+2}. 102 nM was analysed in positive scanning mode. The doubly charged peptide was confirmed by *de novo* sequencing to be QLST[pS]EENSK. The neutral loss of H_3PO_4 (98 Da) from the doubly charged peak 730 gives the 681 peak (annotated with red stars). The spectrum is representative of triplicate experiments.

Peptide mass fingerprinting, precursor ion scanning and production ion scanning were trialled as a combined mass spectrometry approach to locating phosphorylation modifications in a control protein. Phosphorylation

modifications were able to be located at a level that would easily detect similar modifications in wool IFPs separated by 2DE.

3.3.5.2 Intermediate filament proteins

Precursor ion scanning was used on the wool IFPs separated by 2DE and digested with trypsin (Figure 3.9) (Appendix One A1.4). Subsequent product ion scanning and peptide mass fingerprint comparisons showed that approximately half the peaks detected using precursor ion scanning did not result from peptides (Appendix One A1.5 and A1.6). These peaks could have come from contaminants in the solvents used prior to mass spectrometry analysis, as when high concentrations of acid solutions are pipetted they can become contaminated with polymers from the pipette tips. Peptide sequences that were detected in the precursor ion scan were subjected to product ion scanning but none showed the loss of phosphate moieties.

A problem associated with mass spectrometry of phosphopeptides is the suppression effect, which occurs when the acidic phosphopeptides are present in crude and complex mixtures [8]. Another common problem is the hydrophilic nature of the modification, which means that when using reversed-phase chromatography for peptide purification the phosphopeptides may elute in the void volume [7, 26]. However, further experimentation with other techniques to increase the likelihood of detecting phosphopeptides, such as MALDI-TOF ionisation, IMAC [13], alkaline phosphatase treatment [15] and alternative enzymes [14] did not reveal any phosphorylation modifications on the wool IFPs (data not shown).

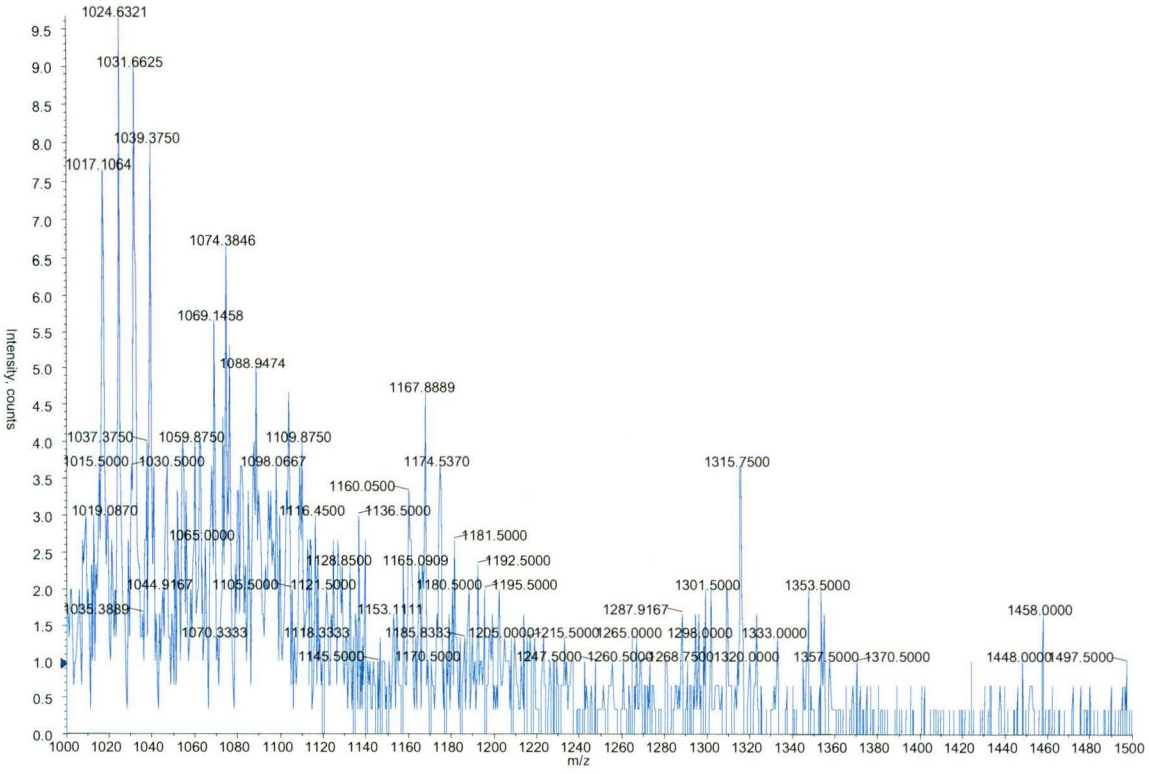


Figure 3.9

Precursor ion scan for the loss of 79 Da (HPO_3^{2-}) from a wool IFP. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

Despite the variety of analytical techniques employed, no evidence was found for wool IFP phosphorylation in the keratinised wool fibre. This contradicts previous work by Herbert *et al.* [5] and Nakamura *et al.* [6]. There is a close relationship in other IFPs between phosphorylation and disassembly [27]. Many IFPs including vimentin [21], desmin [22], epithelial IFs [18] and nuclear lamins [28, 29] are disassembled when they are phosphorylated. It would therefore follow that the IFPs of wool may be phosphorylated in the follicle before final assembly into the wool fibre. In the follicle, the IFPs may be dephosphorylated by phosphatases in final preparation for assembly into filaments that will form the wool fibre. Once the IFs are assembled and keratinised, the cells die and become static structures which would no longer be able to exchange phosphate groups. It has been shown in histochemical studies that there are phosphatase enzymes in the wool follicle bulb [30].

This possible loss of phosphate before terminal differentiation would correlate with studies on stratum corneum (the outer most layer of cells of the epidermis). The precursor to SCBP is highly phosphorylated (15-20 mol of phosphate per mol of protein). The mature SCBP has no phosphate associated with it, suggesting that phosphatases may be involved in forming terminally differentiated stratum corneum [17]. Other functions for phosphorylation, such as modulation of surface properties, which affect functional interactions as shown in neurofilament proteins, would be no longer required in fully keratinised proteins [29].

During IF polymerisation, many of the phosphorylation sites in the head domain become inaccessible, suggesting that the head domain undergoes conformational changes during IF formation [31]. Structural studies on oxidised hard α-keratin support this hypothesis. As discussed in Chapter One, modelling studies have shown that the head domain folds back and interacts with the rod domain. When modified by phosphorylation the head domains may no longer associate with the 1A segments, leading to destabilisation of the 1A segment causing the two strands to separate.

Future work on IFPs may involve extracting the proteins from the follicle before they are assembled. Phosphorylation studies on these proteins may prove that the IFPs are phosphorylated before they are assembled into filaments, and at what stage of development dephosphorylation occurs.

3.4 Summary

Several specific and sensitive techniques have been used to determine whether the IFPs are phosphorylated in mature wool. None of the techniques used found any of the IFPs to be phosphorylated. It is possible that the IFPs are phosphorylated within the follicle and are dephosphorylated prior to assembly and keratinization.

3.5 References

1. Sickmann A, Marcus K, Schäfer H, Butt-Dörje E, Lehr S, Herkner A, Suer S, Bahr I, Meyer HE. Identification of Post-Translationally Modified Proteins in Proteome Studies. *Electrophoresis* **2001**;22:1669-1676
2. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R. The Catalog of Human Cytokeratins: Patterns of Expression in Normal Epithelia, Tumors and Cultured Cells. *Cell* **1982**;31:11-24
3. Steinert PM, Wantz ML, Idler WW. *O*-Phosphoserine Content of Intermediate Filament Subunits. *Biochemistry* **1982**;21:177-183
4. Herbert BR, Chapman ALP, Rankin DA. Investigation of Wool Protein Heterogeneity Using Two-Dimensional Electrophoresis with Immobilised pH Gradients. *Electrophoresis* **1996**;17:239-243
5. Herbert BR, Molloy MO, Yan JX, Gooley AA, Bryson WG, Williams KL. Characterisation of Wool Intermediate Filament Proteins Separated by Micropreparative Two-Dimensional Electrophoresis. *Electrophoresis* **1997**;18:568-572
6. Nakamura A, Arimoto M, Takeuchi K, Fujii T. A Rapid Extraction Procedure of Human Hair Proteins and Identification of Phosphorylated Species. *Biological Pharmaceutical Bulletin* **2002**;25:569-572
7. Stensballe A, Andersen S, Jensen ON. Characterization of Phosphoproteins from Electrophoretic Gels by Nanoscale Fe(III) Affinity Chromatography with Off-Line Mass Spectrometry Analysis. *Proteomics* **2001**;1:207-222
8. Sickmann A, Meyer HE. Phosphoamino Acid Analysis. *Proteomics* **2001**;1:200-206
9. Yan JX, Packer NH, Tonella L, Ou K, Wilkins MR, Sanchez J-C, Gooley AA, Hochstrasser DF, Williams KL. High Sample Throughput Phosphoamino Acid Analysis of Proteins Separated by One- and Two-Dimensional Electrophoresis. *Journal of Chromatography A* **1997**;764:201-210

10. Kaufmann H, Bailey JE, Fussenegger M. Use of Antibodies for Detection of Phosphorylated Proteins Separated by Two-Dimensional Gel Electrophoresis. *Proteomics* **2001**;1:194-199
11. Steinberg TH, Pretty on top K, Berggren KN, Kemper C, Jones L, Diwu Z, Haugland RP, Patton WF. Rapid and Simple Single Nanogram Detection of Glycoproteins in Polyacrylamide Gels on Electroblots. *Proteomics* **2001**;1:841-855
12. Yamagata A, Kristensen DB, Takeda Y, Miyamoto Y, Okada K, Inamatsu M, Yoshizato K. Mapping of Phosphorylated Proteins on Two-Dimensional Polyacrylamide Gels Using Protein Phosphatase. *Proteomics* **2002**;2:1267-1276
13. Posewitz M, Tempst P. Immoblilzed Gallium (III) Affinity Chromatography of Phosphopeptides. *Analytical Chemistry* **1999**;71:2883-2892
14. Schlosser A, Bodem J, Bossemeyer D, Lehmann WD. Identification of Protein Phosphorylation Sites by Combination of Elastase Digestion, Immobilized Metal Affinity Chromatography, and Quadrupole-Time Of Flight Tandem Mass Spectrometry. *Proteomics* **2002**;2:911-918
15. Larsen MR, Sørensen GL, Fey SJ, Peter R. Phospho-Proteomics: Evaluation of the use of Enzymatic De-Phosphorylation and Differential Mass Spectrometric Peptide Mass Mapping for Site Specific Phosphorylation Assignment in Proteins Separated by Gel Electrophoresis. *Proteomics* **2001**;1:223-238
16. Debruyne I. Staining of Alkali-Labile Phosphoproteins and Alkaline Phosphatases on Polyacrylamide Gels. *Analytical Biochemistry* **1983**;133:110-115
17. Lonsdale-Eccles JD, Haugen JA, Dale BA. A Phosphorylated Keratohyalin-Derived Precursor of Epidermal Stratum Corneum Basic Protein. *Journal of Biological Chemistry* **1980**;255:2235-2238
18. Omary MB, Ku N-O, Liao J, Price D. Keratin Modifications and Solubility Properties in Epithelial Cells and *in vitro*. In: Harris, ed. *Subcellular Biochemistry*. New York: Plenum Press, **1998**:105-140

19. Ku N-O, Omary MB. Expression, Glycosylation, and Phosphorylation of Human Keratins 8 and 18 in Insect Cells. *Experimental Cell Research* **1994**;211:24-35

20. Chou C-F, Riopel CL, Rott LS, Omary MB. A Significant Soluble Keratin Fraction in 'Simple' Epithelial Cells. *Journal of Cell Science* **1993**;105:433-444

21. Inagaki M, Nishi Y, Nishizawa K, Matsuyama M, Sato C. Site-Specific Phosphorylation Induces Disassembly of Vimentin Filaments *in vitro*. *Nature* **1987**;328:649-652

22. Inagaki M, Gonda Y, Matsuyama M, Nishizawa K, Nishi Y, Sato C. Intermediate Filament Reconstruction *in vitro*. The Role of Phosphorylation on the Assembly-Disassembly of Desmin. *Journal of Biological Chemistry* **1988**;263:5970-5978

23. Bio-Rad. Protein Blotting - A Guide to Transfer and Detection.

24. Steinberg TH, Agnew BJ, Gee KR, Leung W-Y, Goodman T, Schulenberg B, Hendrickson J, Beechem JM, Haugland RP, Patton WF. Global Quantitative Phosphoprotein Analysis Using Multiplexed Proteomics Technology. *Proteomics* **2003**;3:1128-1144

25. Zhang X, Herring CJ, Romano PR, Szczepanowska J, Brzeska H, Hinnebusch AG, Qin J. Identification of Phosphorylated Sites in Proteins Separated by Polyacrylamide Gel Electrophoresis. *Analytical Biochemistry* **1998**;70:2050-2059

26. Immler D, Gremm D, Kirsch D, Spengler B, Presek P, Meyer HE. Identification of Phosphorylated Proteins from Thrombin-Activated Human Platelets by Two-Dimensional Gel Electrophoresis by Electrospray Ionization-Tandem Mass Spectrometry (ESI-MS/MS) and Liquid Chromatography-Electrospray Ionization-Mass Spectrometry (LC-ESI-MS). *Electrophoresis* **1998**;19:1015-1023

27. Paramio JM. A Role for Phosphorylation in the Dynamics of Keratin Intermediate Filaments. *European Journal of Cell Biology* **1999**;78:33-43

28. Steinert PM. Structure, Function, and Dynamics of Keratin Intermediate Filaments. *The Society for Investigative Dermatology, Inc* **1993**:729-734

29. Fuchs E, Weber K. Intermediate Filaments: Structure, Dynamics, Function, and Disease. *Annual Review of Biochemistry* **1994**;63:345-382

30. Orwin DFG. The Cytology and Cytochemistry of the Wool Follicle. *International Review of Cytology* **1979**;60:331-374

31. Shoeman RL, Hartig R, Berthel M, Traub P. Deletion Mutagenesis of the Amino-Terminal Head Domain of Vimentin Reveals Dispensability of Large Internal Regions for Intermediate Filament Assembly and Stability. *Experimental Cell Research* **2002**;279:344-353

Chapter Four

Proteomic Investigations into Glycosylation as a Post-Translational Modification of Wool Intermediate Filament Proteins.

The goal of this chapter was to determine whether the charge heterogeneity of the wool IFPs separated by 2DE was related to glycosylation of the proteins. As with phosphorylation analysis (Chapter Three) a multiplexed approach was taken to determine the glycosylation state of the proteins.

4.1 Introduction

The results from Chapter Three imply that the charge heterogeneity seen when wool IFPs are separated by 2DE was not caused by phosphorylation. Another common PTM of IFs is glycosylation [1-4]. Glycosylation could explain the charge heterogeneity seen when wool IFPs are subjected to 2DE, as glycoforms usually separate as 'trains' of spots in the first dimension of 2DE [5]. Past studies on soft keratins have discovered a single O-GlcNAc modification is common (See Chapter One for a review). There appear to have been no studies done on the glycosylation state of hard α-keratins.

Histochemical studies on the wool follicle have demonstrated that plasma membrane proteins at the base of the wool follicle show polysaccharide material at the surface thought to be glycoproteins [6]. Further up the follicle, where more protein synthesis occurs, the level of polysaccharide material is greatly decreased. No further polysaccharide material is detected before the zone of keratinisation [6]. This suggests there may be a role for glycosylation in the formation of mature IFs.

4.2 Techniques for characterising glycosylated proteins

Techniques for characterising glycosylations are limited in hard α-keratin fibres, as it is difficult to perform studies that involve the metabolic labelling of cells. To label keratinised wool proteins, the sheep would need to be fed a diet containing radioactively labelled amino acids that could be incorporated into the growing fibre. Cultured cells are often labelled with [^{3}H] or [^{14}C] using galactosyltransferase before proteolysis and high pressure liquid chromatography (HPLC). Fractions are then screened for radioactivity. Once a radioactive fraction is found, the peptide is isolated and identified by Edman sequencing [1, 7, 8]. This procedure is very tedious and requires a lot of glycopeptide starting material [9].

Polyacrylamide gel-based electrophoretic techniques provide a useful means to analyse glycosylation. Two-dimensional electrophoresis gels have become the preferred method to provide fast purification and concentration of samples for analysis [5, 7, 10]. The initial step in glycosylation analysis is often to detect carbohydrates directly on the proteins separated by electrophoresis [11]. Electrophoresis has advantages over other methods as it allows separation and detection of glycosylated proteins in complex samples [11]. Often the effect of glycosylation can be readily apparent on 2DE gels, as carbohydrates can greatly affect the molecular weight of a protein [5]. As with phosphorylation analysis (Chapter Three) fluorescent stains selective for glycosylations have recently been developed [7] and improved [12]. This type of staining is very sensitive and is the most sensitive gel-based method for the detection of glycosylated proteins [12].

Enzymatic release of glycosylations has also been used to determine which proteins are glycosylated [13]. However, a study on the filamentous fungi *Trichoderma reesei* determined that chemical deglycosylation with trifluromethanesulfonic acid (TFMS) was more efficient than enzymatic methods [14]. When the fungal proteins were deglycosylated, the 2DE gel patterns showed reduced complexity. A study on the deglycosylation of hemonectin (a bone marrow protein) also found that deglycosylation with TFMS was more effective at removing all glycosylations when compared to

enzymatic deglycosylation [15]. They found that degradation of the polypeptide backbone was not a problem with this method of deglycosylation [15].

Mass spectrometry is one of the most useful methods to analyse the structures of carbohydrates [16]. Improvements to mass spectrometry techniques to detect glycosylated peptides are constantly being made [5, 9, 16-25]. Product ion scanning can detect sites of glycosylation at low levels [9]. Improvements in the software tools to predict possible modifications and determine the structures of modifications have allowed the exact determination of complex glycosylations easier [26, 27].

4.3 Results and Discussion

To determine whether the keratinised IFPs from wool were post-translationally modified by glycosylation, several different proteomic techniques were employed, including on-gel staining, chemical deglycosylation, immunoblotting and mass spectrometry.

4.3.1 Thymol-sulfuric acid stain

The thymol-sulfuric acid stains the carbohydrates by hydrolysing glycoside bonds, dehydrating the sugars, followed by condensation of the furfural with a phenol to yield quinoidal pigments. The reaction is sensitive (the detection limit was approximately 0.05 μg of carbohydrate) and general for pentoses, hexoses, uronic acids and all of their polymers [28]. The staining is relatively quick, and is insensitive to protein and non-carbohydrate lipid [28]. Post-electrophoretic staining of wool IFPs on 2DE gels showed no glycosylation staining (Figure 4.1 A). The positive control (horse radish peroxidase) loaded onto the left-hand side of the gel, showed a red-brown band at 44 kDa.

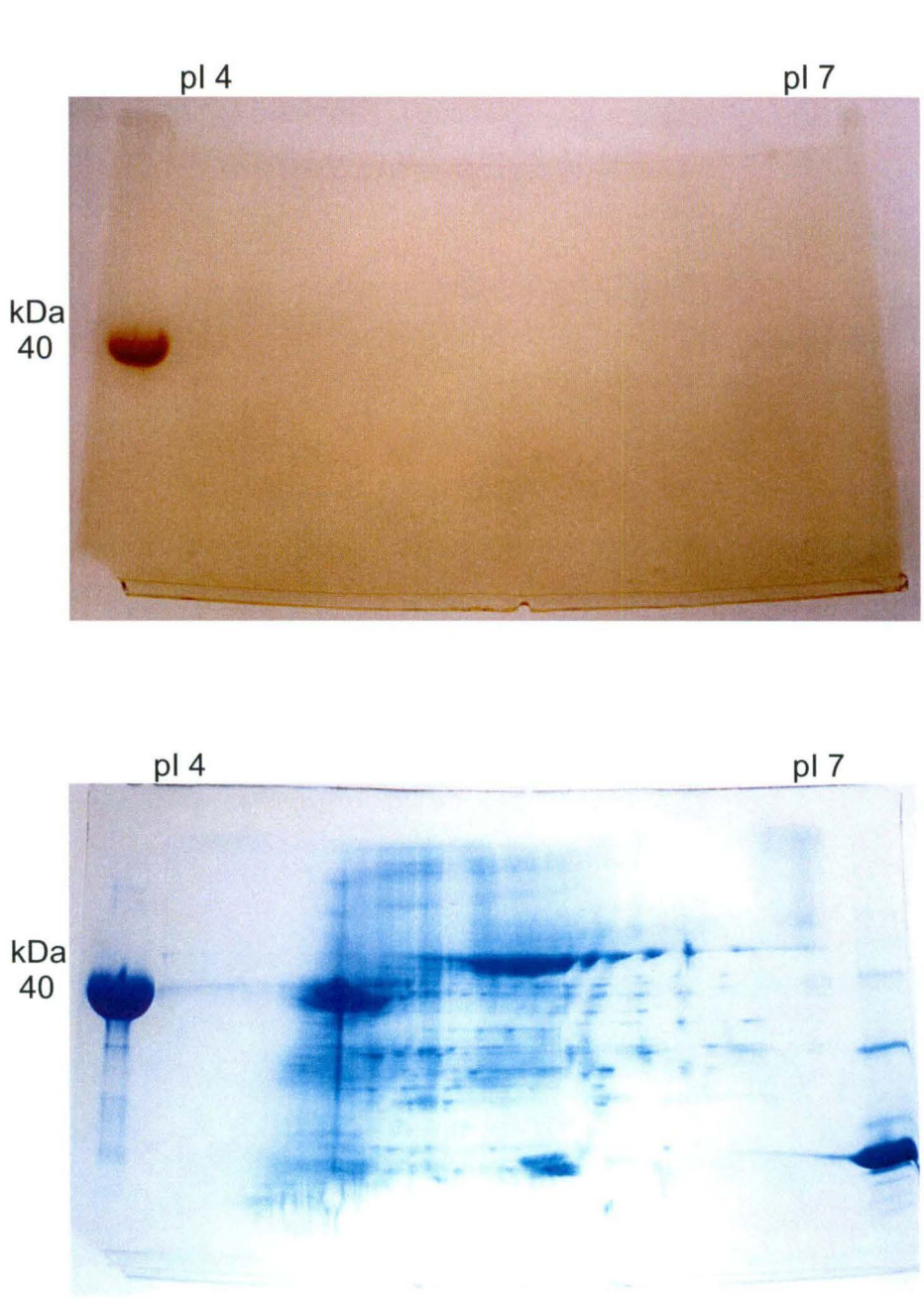


Figure 4.1
A) Thymol-sulphuric acid staining of wool IFPs on a 2DE gel. A pH 4-7 IPG strip containing 1 mg of focused protein was run in the second dimension on a 10-20% T criterion gel. Horse radish peroxidase (133 µg) was loaded at the left-hand side of the strip as a positive control and lysozyme (133 µg) was loaded on the right-hand side of the strip as a negative control. B) Equivalent gel stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

The negative control (lysozyme) was loaded on the right-hand side of the gel and showed no staining, suggesting the staining technique was specific for

glycosylated protein. Figure 4.1 B shows an equivalent gel stained with Coomassie brilliant blue G-250. It demonstrates that the large amount of wool protein, which was loaded onto the first dimension IPG strip, had caused overloading and smearing. Therefore, there was more than adequate protein on the gel for the detection of glycosylations.

4.3.2 Periodic acid-Schiff's staining

As the thymol-sulfuric acid stain revealed no glycosylations, a more sensitive staining technique was employed. New gels were run, which were then stained with the periodic acid-Schiff's stain, in case the level of glycosylation was below the detection limit of the thymol-sulfuric acid stain. As with the previous experiment, duplicate gels were Coomassie brilliant blue G-250 stained to confirm the presence of IFPs on the gels.

A periodic acid-Schiff's stain was used that can detect as little as 25-100 ng of glycoprotein on a gel [11]. The staining of glycoproteins involves the periodate oxidation of *cis*-diols to form aldehydes. The aldehydes then react with amine-containing molecules to form Schiff bases [11]. Initial results showed that the IFPs were giving a positive result; however, the negative control was also staining (Figure 4.2).

The staining procedure was revised and repeated (Method B). The oxidation step to produce aldehydes had to be tightly controlled to reduce the risk of oxidation of amino acids such as cysteine, methionine, serine, threonine and tyrosine [29]. The oxidation reaction is dependent on the concentration of periodate, the pH, temperature and light exposure [11, 30]. Controlling the oxidation conditions can reduce non-specific staining and allows detection of weakly glycosylated proteins. Overoxidation can be avoided most easily by increasing the pH, lowering the temperature, or reducing the oxidation time, as was done here [30].

Figure 4.2
Periodic acid-Schiff's staining of wool IFPs on a 2DE gel. A pH 4-7 IPG strip containing 1 mg of focused protein was run in the second dimension on a 10-20% T criterion gel. Horse radish peroxidase (133 µg) was loaded at the left-hand side of the strip as a positive control, and soybean trypsin inhibitor (133 µg) was loaded on the right-hand side of the strip as a negative control. The gel is representative of duplicate experiments.

The second periodic acid-Schiff's stain shows the positive control as a dark magenta band with no other bands or spots staining on the gel (Figure 4.3 A). The control gel (Figure 4.3 B) shows that the IFPs, and the control proteins, are able to be detected easily with a general protein stain. This suggests that if the IFPs of wool are glycosylated then it is at a very low level, below the detection limit (~150 ng [7]) of the periodic acid-Schiff's staining method.

4.3.3 Chemical deglycosylation

Romney wool protein and a standard glycosylated protein, fetuin, were deglycosylated using a method developed by Hakimuddin and Bahl [31]. Fetuin has three asparagine-linked triantennary complex glycosylations and three serine/threonine-linked glycosylations [31]. A large attached carbohydrate on a protein may increase the apparent mass, as determined on a polyacrylamide gel, by as much as 50% [5].

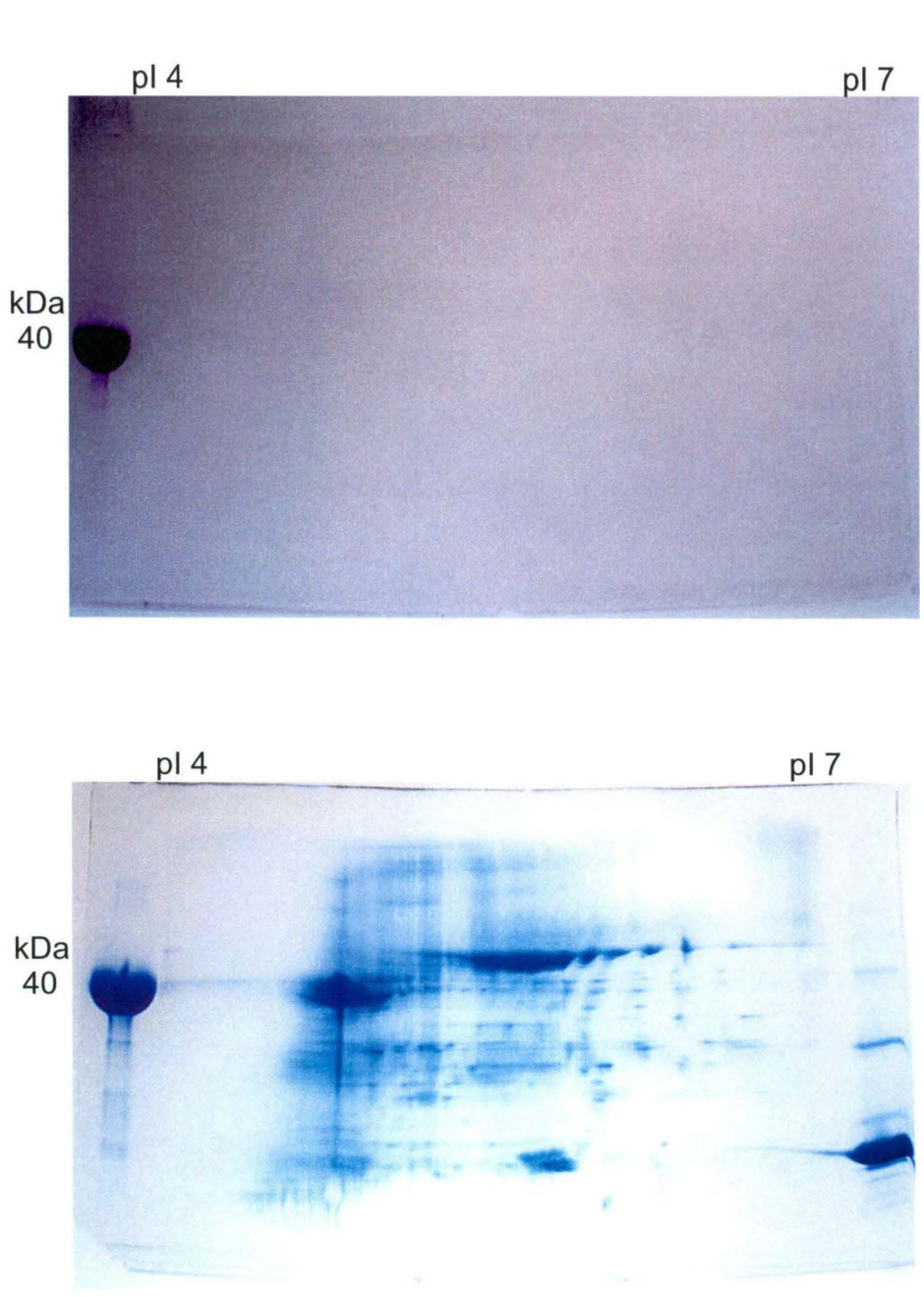


Figure 4.3
Modified periodic acid-Schiff's staining of wool IFPs on a 2DE gel. A pH 4-7 IPG strip containing 1 mg of focused Romney wool protein was run in the second dimension on a 10-20% T criterion gel. Horse radish peroxidase (47 µg) was loaded at the left-hand side of the strip as a positive control and lysozyme (47 µg) was loaded on the right-hand side of the strip as a negative control. B) Equivalent gel stained with Coomassie brilliant blue G-250. The gels are representative of triplicate experiments.

When subjected to TFMS chemical deglycosylation, approximately 90% of the carbohydrate is cleaved [31]. After deglycosylation of fetuin, a mass shift to a lower molecular weight can be seen on a 1DE gel. Figure 4.4 shows deglycosylated fetuin in lane 12 and fetuin before deglycosylation in lanes seven, nine and 11. There are several impurities in the fetuin sample but using data from previous studies the fetuin band can be identified [31]. There is a clear difference in the masses before and after deglycosylation (approximately 13 kDa) suggesting that the deglycosylation reaction worked successfully.

Lanes two, four and six show the wool protein after deglycosylation, and lanes one, three and five show wool protein before deglycosylation. There appears to be no mass difference between the proteins before and after deglycosylation. This result suggests that if the wool IFPs are glycosylated, the carbohydrates attached must be small and have no effect on the apparent mass of the protein when run on a 1DE gel. It is possible the deglycosylation method used was not able to remove the type of carbohydrates attached. The soft keratins are known to contain single O-GlcNAc (204 Da) (Chapter One). The loss of a single O-GlcNAc from the protein via chemical deglycosylation would not distinguish it from the native protein mass on a 1DE gel. However, if several O-GlcNAc modifications were lost, the difference in mass on a 1DE gel would easily be seen.

Glycoforms commonly separate into rows of horizontal spots when separated by 2DE [5]. After undergoing chemical deglycosylation, IFPs still separated out into rows of spots on a 2DE gel (Figure 4.5 A and B). This suggests that glycosylation is not responsible for the heterogeneity of the pIs of the IFPs. It is possible that either another type of PTM is the cause of the heterogeneity, or that the gel electrophoretic process may be contributing to the variations in pI seen on the 2DE gels.

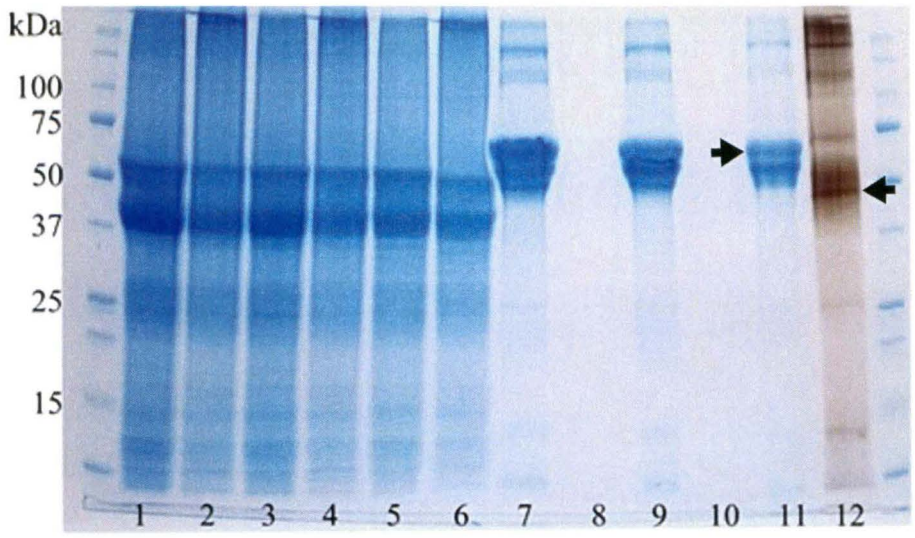
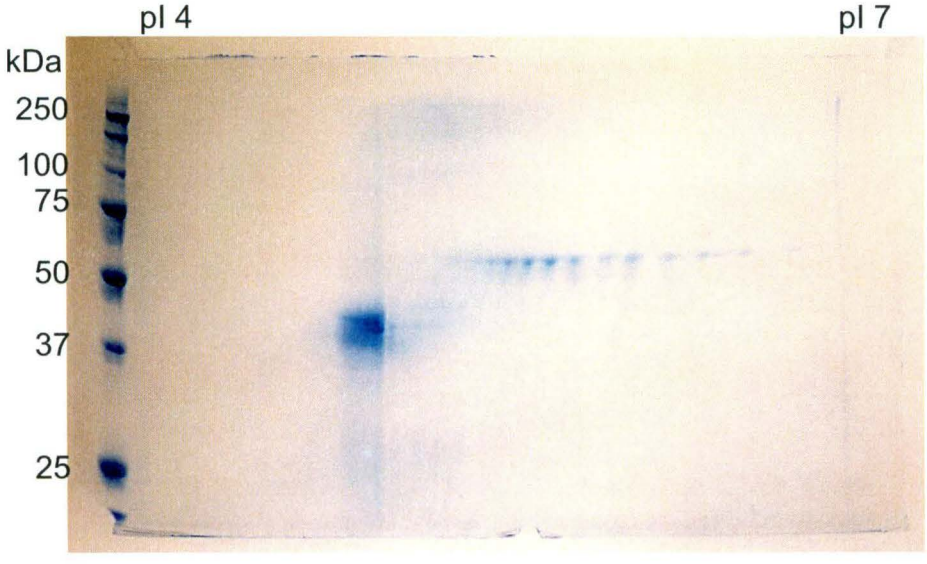


Figure 4.4 A 1DE gel showing wool IFPs and fetuin before and after chemical deglycosylation with TFMS. Lanes two, four and six, wool protein after deglycosylation. Lanes one, three and five, show wool protein before deglycosylation. Lane 12, shows fetuin after deglycosylation. Lanes seven, nine and 11, show fetuin before deglycosylation. Arrow in lane 11 indicates the fetuin band before deglycosylation . Arrow in lane 12 indicates the fetuin band after deglycosylation. The 1DE gel was stained with Coomassie brilliant blue G-250. Lane 12 was subsequently stained with the Blum silver stain. The gel is representative of duplicate experiments.

A)



B)

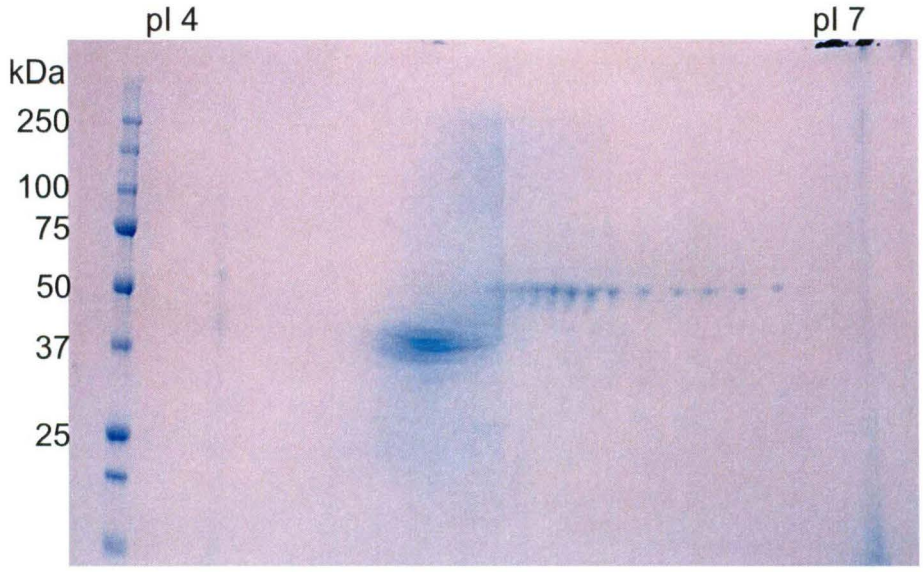


Figure 4.5
A) 2DE gel of wool IFPs after chemical deglycosylation with TFMS. A pH 4-7 IPG strip containing focused protein was run in the second dimension on a 10% T criterion gel. B) 2DE gel of wool protein before deglycosylation. A pH 4-7 IPG strip containing focused protein was run in the second dimension on a 10-20% T criterion gel. Gels were stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

4.3.4 O-GlcNAc Immunoblotting

The antibody (MA1-076) was chosen as the most likely to detect O-GlcNAc glycosylation in IFPs, based on previous studies [1-3]. All other IFPs that have been shown to be glycosylated have had this modification [1-3]. O-GlcNAc is a small glycosylation, which would not show a measurable change in the apparent mass of a protein separated by 2DE if a single molecule was attached. All of the IFPs showed reactivity (Figure 4.6); however, some of the negative control proteins also showed reactivity (Figure 4.7). Nonspecific reactions between bound proteins and probes can be caused by contaminated primary or secondary antibodies. Monoclonal antibodies can also react non-specifically with SDS denatured proteins [32].

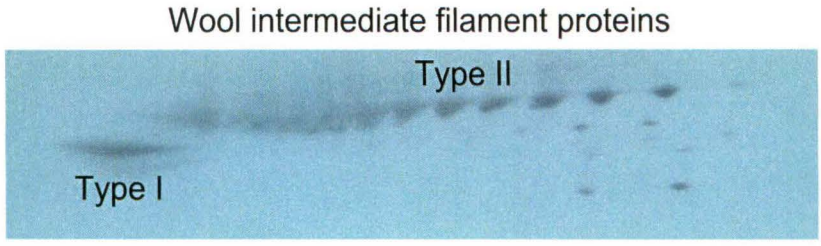


Figure 4.6
Wool IFPs after transfer to PVDF and chemiluminescent detection of antibody labelled proteins. The immunoblot is representative of triplicate experiments.

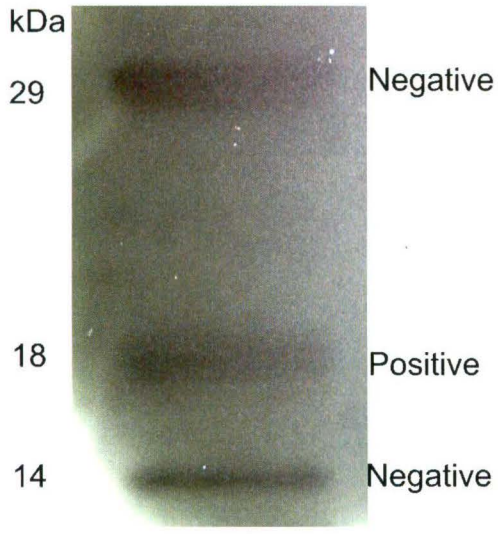


Figure 4.7
Chemiluminescent detection of CandyCane™ glycoprotein molecular weight standards, after 1DE separation and transfer to a PVDF membrane. The standards contain both positive and negative control proteins for glycosylation. Two negative control proteins at 29 kDa and 14 kDa show a positive reaction to the antibody, suggesting a lack of selectivity by the antibody.

4.3.5 Pro-Q® Emerald 300 glycoprotein gel stain

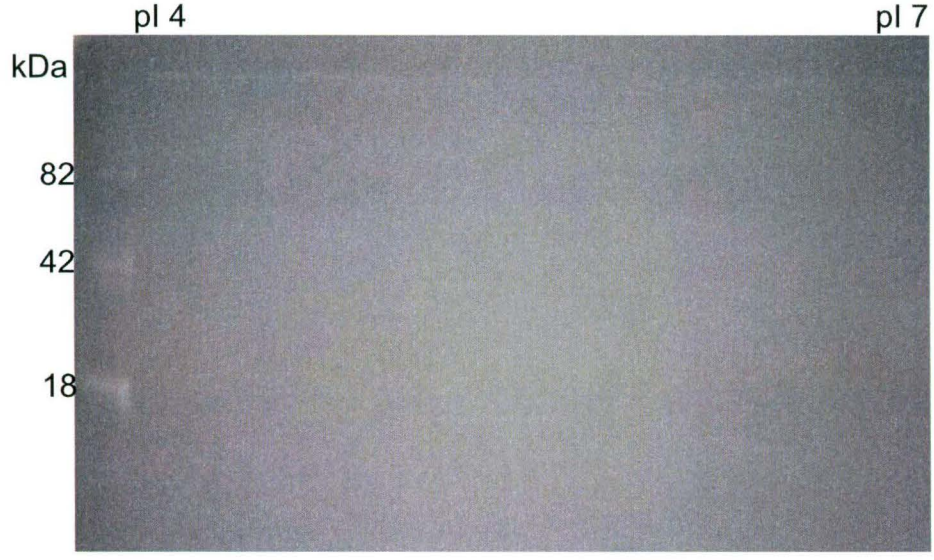
The Pro-Q® Emerald 300 glycoprotein gel stain is approximately 150 – 250 times more sensitive than periodic acid-Schiff's staining [7]. Depending on the degree of glycosylation, as little as 300 pg of glycoprotein can be detected on a gel [12]. Wool protein run on a 2DE gel showed no fluorescence under UV illumination (Figure 4.8, A). CandyCane™ molecular

weight markers, which have alternating glycosylated and nonglycosylated proteins, showed staining specifically at the glycosylated protein molecular weights (Figure 4.8 A).

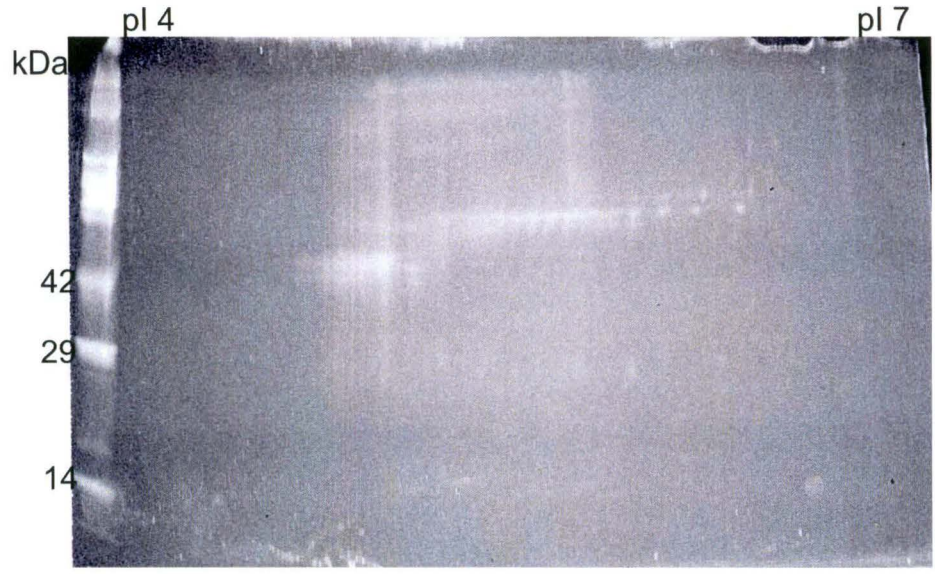
An advantage of the Pro-Q® Emerald glycoprotein stain is that after staining to detect glycosylated proteins, the same gel can be stained with a general protein stain. Sypro® Ruby staining of the gel showed all the IFPs were present on the gel at a level which was able to be stained (Figure 4.8 B). The negative glycosylation controls were also stained (Figure 4.8 B, left hand lane). A Coomassie brilliant blue G-250 stained gel with equivalent loading of IFPs (Figure 4.8 C) shows that the sensitivity of the Sypro® Ruby stain is slightly less than the Coomassie brilliant blue G-250 stain, due to the methods used to visualise and document the fluorescent Sypro® Ruby stain.

All of the attempts to detect glycosylated IFPs on PAGE gels gave results that suggest the wool IFPs are not glycosylated.

A)



B)



C)

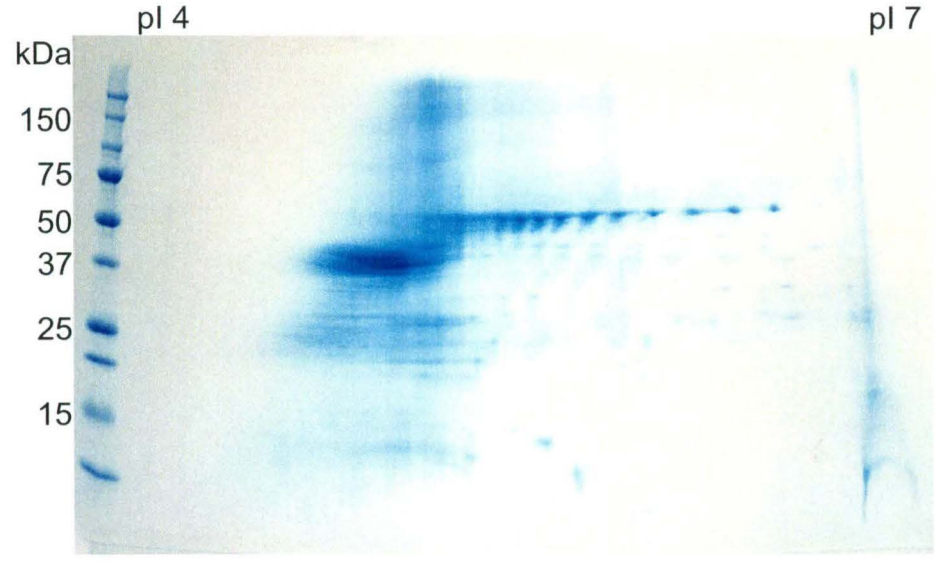


Figure 4.8
A) Pro-Q® Emerald 300 glycoprotein gel staining of a 2DE gel with 100 µg of wool protein. Protein was focused on a pH 4-7 IPG strip before being run on a 10–20% T criterion gel. Left hand lane shows CandyCane™ molecular weight markers. B) The same gel stained with Sypro® Ruby. C) A comparative gel stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

4.3.6 Mass Spectrometry

Tandem mass spectrometry (MS/MS) methods, where a parent ion is selected in the first quadrupole (Q1) by its mass to charge ratio (m/z) then fragmented in the second quadrupole (Q2) into ions that reveal the parent ion composition, are useful for the characterisation of peptides modified by PTMs [18, 25].

4.3.6.1 Method validation with Glycomacropeptide (GMP)

Glycomacropeptide variant B is known to be glycosylated on three neighbouring threonines positioned at 131, 133 and 135 [33]. The protein was digested with chymotrypsin in solution at a ratio of chymotrypsin to protein of 1:60. A peptide mass fingerprint scan of a 16 hour digestion is shown (Figure 4.9) (Appendix Two shows a 4 hour digestion peptide mass fingerprint scan, A2.1). This concentration of protein, and ratio of protease to protein, gave a spectrum with high abundance peaks. The peptide mass fingerprint data give information about which peaks represent peptides which aids further precursor ion and product ion scanning analysis.

Precursor scans based on oxonium ions are useful for locating glycosylated peptides (Table 4.1) [23].

Glycomacropeptide precursor ion scans for the loss of 204 Da (N-acetylhexosamine cation) indicated the peptides in the m/z range 1000-1340 (Figure 4.10) were potentially glycosylated (Appendix Two, A2.2 shows the peak list).

The precursor ion scan was low in abundance, but it was possible to select oxonium ion containing peptides above the background. Once peptides were identified from the precursor scan, they were subjected to product ion scanning to determine their amino acid sequence (Appendix Two A2.3) [19]. An example of a product ion scan is shown in Figure 4.11. The ion fragments (Table 4.1) representing $HexNAc^+$ (m/z 204), $NeuAc^+$ (m/z 292) and $HexHecNAc^+$ (m/z 366) could all be positively identified, demonstrating that the peptide was glycosylated. The sequence of this peptide was

manually interpreted by *de novo* sequencing to be ^{147}EASPEVIESPPEINTV^{162}. This demonstrated that this technique was able to detect the loss of glycosylated fragment ions from peptides.

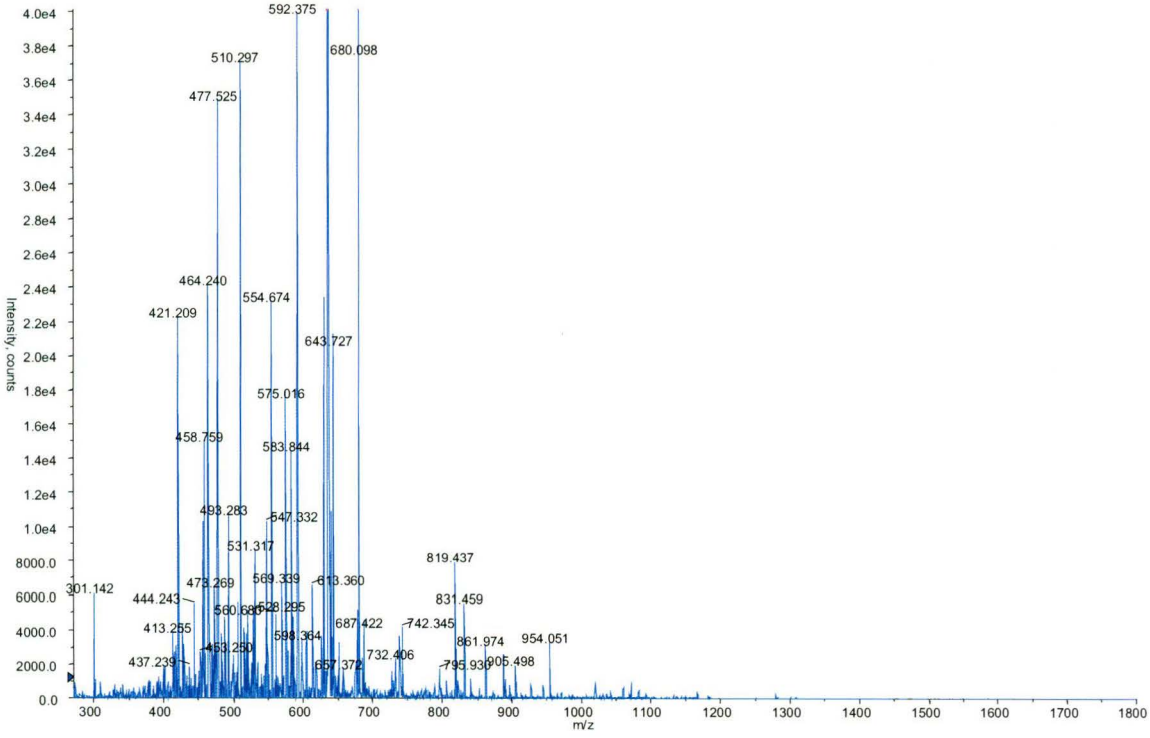


Figure 4.9
Peptide mass fingerprint of GMP variant B at a concentration of 14 µM digested with chymotrypsin for 16 hours. The spectrum is representative of triplicate experiments.

Ion Fragment	m/z	Abbreviated code
Hexose	163.060	Hex^+
N-Acetylhexosamine	204.084	$HexNAc^+$
N-Acetyl-neuraminic acid	292.095	$NeuAc^+$
Hexoylhexosamine	366.139	$HexHexNAc^+$
N-Acetyl-neuraminic-hexoylhexosamine	657.235	$NeuAc\text{-}HexHexNAc^+$

Table 4.1
Ion fragments for identification of common glycopeptides [21].

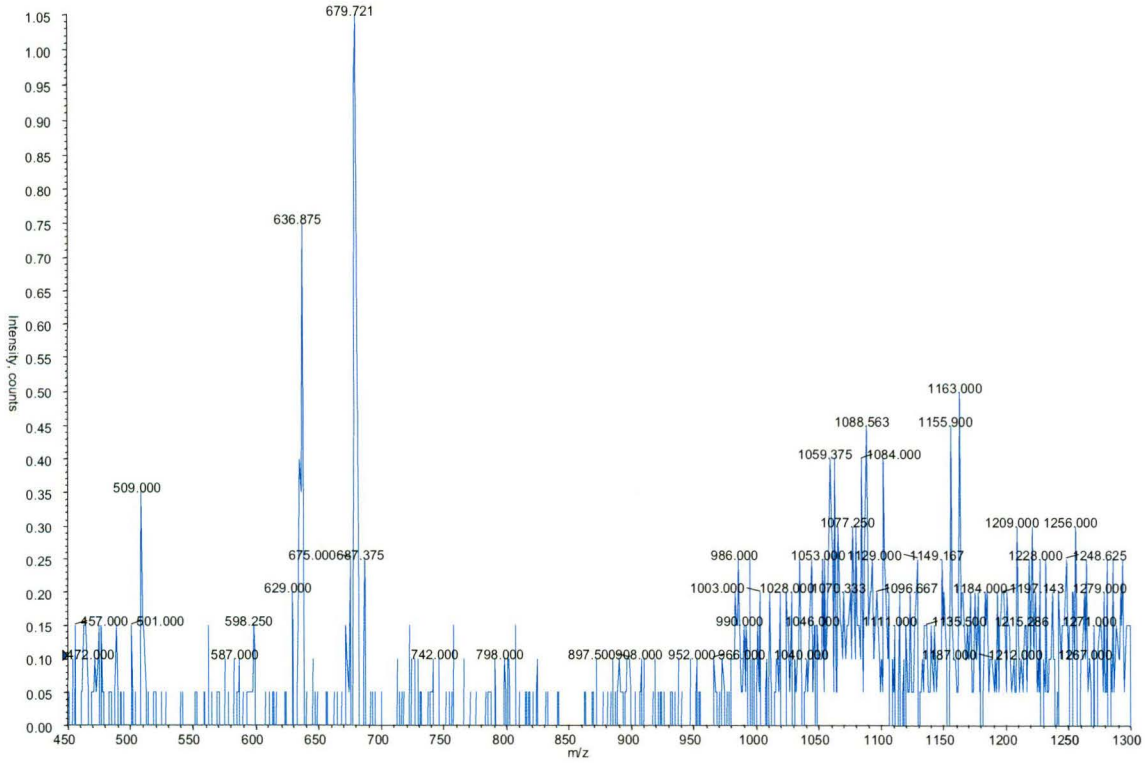


Figure 4.10
Precursor scan for the loss of 204 Da (N-acetylhexosamine) from GMP variant B at a concentration of 14 μM. The spectrum is representative of triplicate experiments.

The sensitivity of the mass spectrometer was maximised by optimising the collision energy (CE) [21]. The abundance of the ion fragment at m/z 204 ($HexNAc^+$) was determined as the CE of the mass spectrometer was altered. A CE of 34, when using argon as the collision gas, was optimal for production of m/z 204. Further analysis of the product data also suggested that the m/z 274 [NeuAc - H_2O]$^+$ and 292 [$NeuAc^+$] ions will also be potentially useful for precursor scanning based on the GMP spectra. Optimal CE is important when analysing glycans, as they require lower CEs to cleave the glycosidic bonds compared to other PTMs [21].

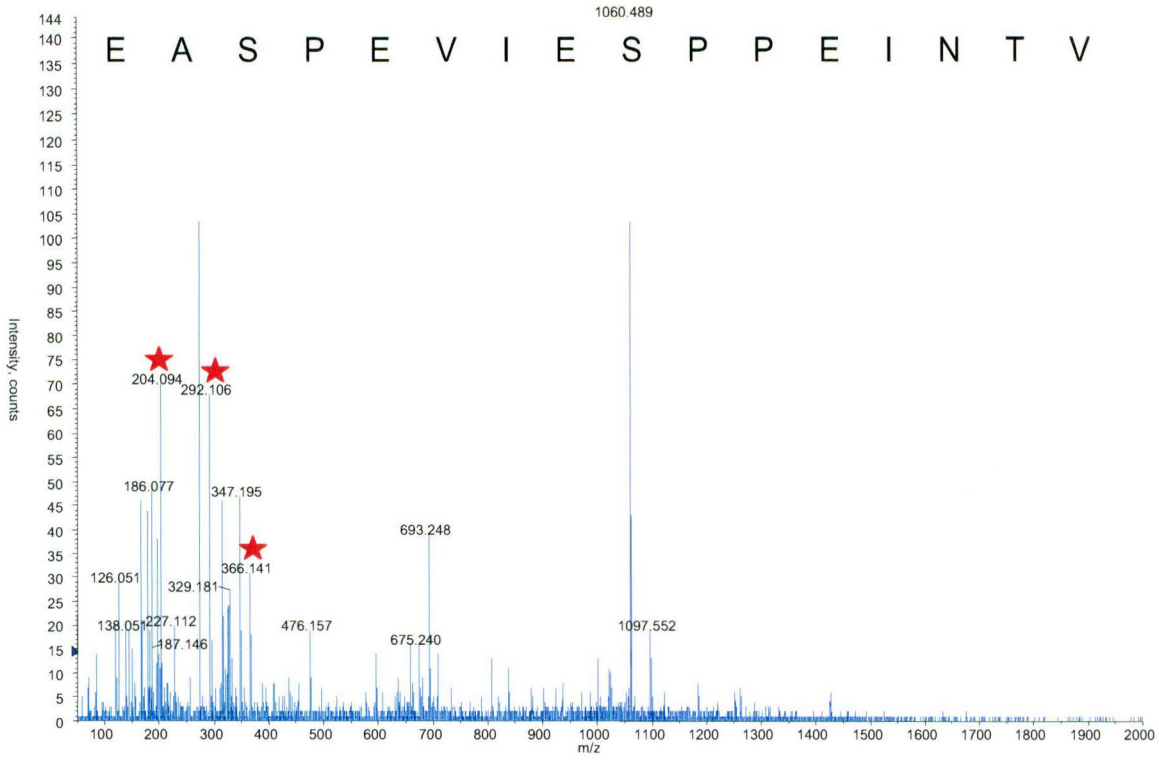


Figure 4.11
Product ion scan of the doubly charged peptide at m/z 1060^{2+}. The spectrum clearly shows the ion fragments at m/z 204, 292 and 366, which are characteristic for the loss of glycosylation modifications from the peptide (annotated with red stars). The spectrum is representative of triplicate experiments.

Precursor ion scanning aided in isolating a glycopeptide from a relatively complex peptide mix of GMP variant B at a concentration of 14 nmole. This concentration would seem the minimum for glycopeptide identification using the precursor ion scan approach. The GMP reference material enabled optimisation of peptide fragmentation by fine-tuning of the collision parameters at Q2 of the Q-STAR instrument in precursor ion scanning mode. A similar approach to finding glycosylations was then applied to 2DE separated wool IFPs.

4.3.6.2 IFPs

Intermediate filament protein trypsin digests, with protein concentrations well above the detection limit of the mass spectrometer, were scanned for the 204.087 precursor ion using the methods developed with GMP. No potentially glycosylated peptides were indicated in any precursor ion scans, suggesting that the IFPs do not contain O-GlcNAc PTMs (Figure 4.12), (Appendix Two A2.4). Four different IFP spots were analysed in triplicate. These protein spots were all excised from 2DE gels that large amounts of protein loaded, to maximise the chance of detecting glycosylation modifications.

2DE gels of wool proteins separated using optimised methodology (Chapter Two) show individual protein spots with little smearing between proteins. Glycosylated proteins often focus into spots with extensive smearing between them [5, 14]. This corroborates the suggestion that the wool IFPs are not highly glycosylated (Sections 4.3.1 to 4.3.5).

In the soft keratins, glycosylation acts to regulate many different cellular processes [2, 3]. There is significant evidence for a role as a nuclear targeting signal *via* regulation of serine and threonine phosphorylation sites [1]. Glycosylation is a highly dynamic PTM, with the rate of turnover being much greater than the rate of turnover of the polypeptide backbone [1]. Glycosylation is a reversible modification which can act as an on-off switch for many biological processes [34]. Wool, which is made up of hard α-keratins, is composed of dead cells. It is possible that the wool IFPs are glycosylated in the follicle prior to assembly, while they are still within living cells, and are deglycosylated before they form filaments.

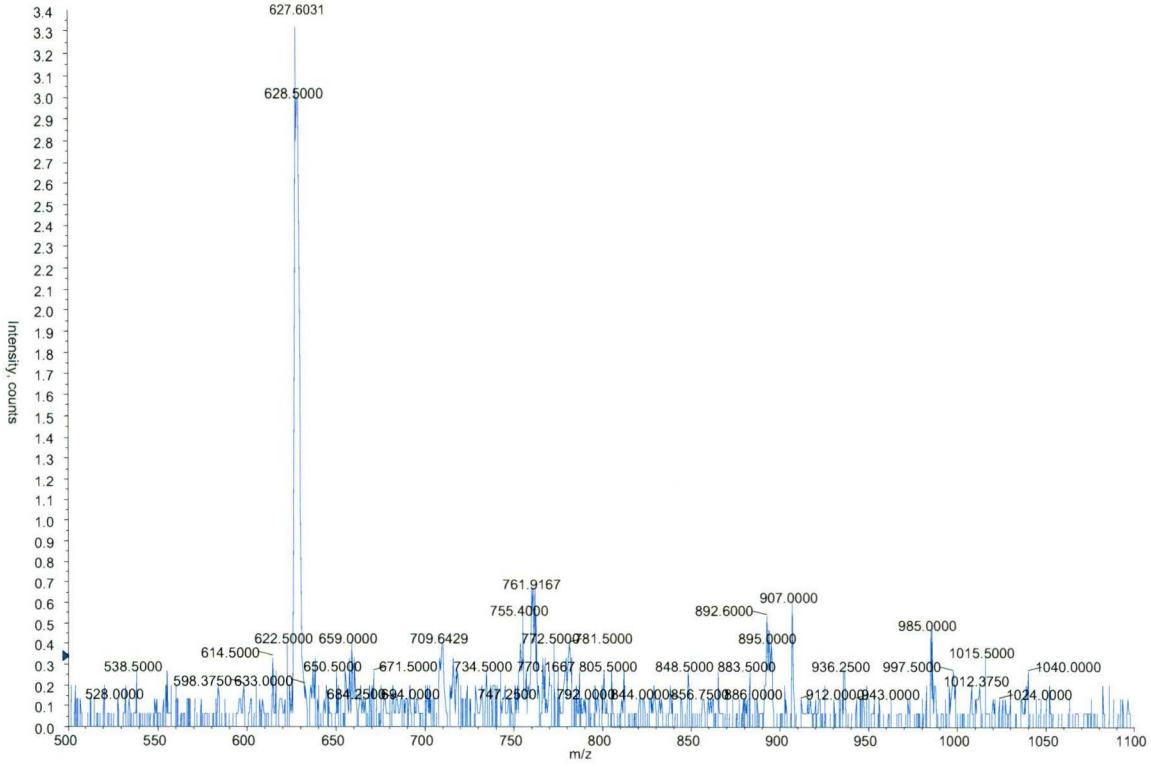


Figure 4.12
Precursor ion scan for the loss of 204 Da (N-acetylhexosamine) from a wool IFP. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

There may be a relationship between phosphorylation and glycosylation in the wool follicle, in which glycosylation controls phosphorylation. Initially, in the developing wool follicle, phosphatase-kinase pathways may regulate the solubility and rearrangement of wool IFPs [6]. When the proteins start to assemble into IFs glycosylation may act to control the disassembly action of phosphorylation. This postulates a role for glycosylation in the control of phosphorylation pathways and assisting in macromolecule assembly in the wool follicle. The requirement for controlling phosphatase-kinase activity would no longer be required once the IFs had reached the stage of terminal differentiation. The modifications would thus be removed from the proteins prior to terminal differentiation.

4.4 Summary

Several different gel protein staining methods with specificity for glycoprotein detection have been used to investigate the glycosylation of wool IFPs. The gel stains all showed that none of the IFPs were glycosylated. Deglycosylation of these proteins showed no shift in molecular weight or a decrease in the pI heterogeneity of the IFPs when isoelectrically focused. Chemiluminescent detection after immunoblotting gave no information about the glycosylation of the IFPs as there were many non-specific reactions between the proteins and the probes, shown by the detection of negative control proteins. Mass spectrometry precursor ion scanning was successful at detecting and identifying standard glycosylated peptides but found no O-GlcNAc in IFP peptides. These results all suggest that the IFPs of wool are either not glycosylated, or are glycosylated at a level not detectable by the methods used. It is likely that the wool IFPs are glycosylated in the wool follicle and are deglycosylated prior to terminal differentiation.

4.5 References

1. Chou C-F, Smith AJ, Omary MB. Characterization and Dynamics of O-Linked Glycosylation of Human Cytokeratin 8 and 18. *Journal of Biological Chemistry* **1992**;267:3901-3906
2. Comer FI, Hart GW. *O*-Glycosylation of Nuclear and Cytosolic Proteins. *Journal of Biological Chemistry* **2000**;275:29179-29182
3. Hart GW, Greis KD, Dong D, Blomberg MA, Chou T-Y, Jiang M-S, Roquemore EP, Snow DM, Kreppel LK, Cole RN, Comer FI, Arnold CS, Hayes BK. O-linked N-acetylglucosamine: The "yin-yang" of Ser/Thr Phosphorylation? *Glycoimmunology* **1995**:115-123
4. Ku N-O, Omary MB. Identification and Mutational Analysis of the Glysosylation Sites of Human Keratin 18. *Journal of Biological Chemistry* **1995**;270:11820-11827
5. Packer NH, Lawson MA, Jardine DR, Sanchez J-C, Gooley AA. Analyzing Glycoproteins Separated by Two-Dimensional Gel Electrophoresis. *Electrophoresis* **1998**;19:981-988

6. Orwin DFG. The Cytology and Cytochemistry of the Wool Follicle. *International Review of Cytology* **1979**;60:331-374

7. Steinberg TH, Pretty on top K, Berggren KN, Kemper C, Jones L, Diwu Z, Haugland RP, Patton WF. Rapid and Simple Single Nanogram Detection of Glycoproteins in Polyacrylamide Gels on Electroblots. *Proteomics* **2001**;1:841-855

8. Chalkley RJ, Burlingame AL. Identification of Novel Sites of O-*N*-Acetylglucosamine Modification of Serum Response Factor Using Quadrupole Time-Of-Flight Mass Spectrometry. *Molecular and Cellular Proteomics* **2003**;2:182-190

9. Greis KD, Hayes BK, Comer FI, Kirk M, Barnes S, Lowary TL, Hart GW. Selective Detection and Site-Analysis of O-GlcNAc- Modified Glycopeptides by ß-Elimination and Tandem Electrospray Mass Spectrometry. *Analytical Biochemistry* **1996**;234:38-49

10. Küster B, Krogh TN, Mørtz E, Harvey DJ. Glycosylation Analysis of Gel-separated Proteins. *Proteomics* **2001**;1:350-361

11. Bouchez-Manhiout I, Doyen C, Laurière M. Accurate Detection of Both Glycoproteins and Total Proteins on Blots: Control of Side Reactions Occurring after Periodate Oxidation of Proteins. *Electrophoresis* **1999**;20:1412-1417

12. Hart C, Schulenberg B, Steinberg TH, Lueng W-Y, Patton WF. Detection of Glycoproteins in Polyacrylamide Gels and on Electroblots Using Pro-Q Emerald 488 Dye, a Fluorescent Periodate Schiff-Base Stain. *Electrophoresis* **2003**;24:588-598

13. Wilson NL, Schulz BL, Karlsson NG, Packer NH. Sequential Analysis of N- and O-Linked Glycosylation of 2D-PAGE Separated Glycoptoreins. *Journal of Proteome Research* **2002**;1:521-529

14. Fryksdale BG, Jerdrzejewski PT, Wong DL, Miller BS. Impact of Deglycosylation Methods on Two-Dimensional Gel Electrophoresis and Matrix Assisted Laser Desorption/Ionization-Time Of Flight-Mass Spectrometry for Proteomic Analysis. *Electrophoresis* **2002**;23:2184-2193

15. Sullenbarger BA, Petitt MS, Chong P, Long MW, Wicha MS. Murine Granulocytic Cell Adhesion to Bone Marrow Hemonectin is Mediated by Mannose and Galactose. *Blood* **1995**;86:135-140

16. Kawasaki N, Ohta M, Hyuga S, Hashimoto O, Hayakawa T. Analysis of Carbohydrate Heterogeneity in a Glycoprotein Using Liquid Chromatography/Mass Spectrometry and Liquid Chromatography with Tandem Mass Spectrometry. *Analytical Biochemistry* **1999**;269:297-303

17. Mo W, Sakamoto H, Nishukawa A, Kagi N, Langridge JI, Shimonishi Y, Takao T. Structural Characterization of Chemically Derivatized Oligosaccharides by Nanoflow Electrospray Ionization Mass Spectrometry. *Analytical Chemistry* **1999**;71:4100-4106

18. Haynes PA, Aebersold R. Simultaneous Detection and Identification of O-GlcNAc-Modified Glycoproteins Using Liquid Chromatography-Tandem Mass Spectrometry. *Analytical Biochemistry* **2000**;72:5402-5410

19. Ritchie MA, Gill AC, J DM, Lilley K. Precursor Ion Scanning for Detection and Structural Characterization of Heterogeneous Glycopeptide Mixtures. *American Society for Mass Spectrometry* **2002**;13:1065-1077

20. Wells L, Vosseller K, Cole RN, Cronshaw JM, Matunis MJ, Hart GW. Mapping Sites of O-GlcNAc Modification Using Affinity Tags for Serine and Threonine Post-Translational Modifications. *Molecular and Cellular Proteomics* **2002**;1:791-804

21. Jebanathirajah J, Steen H, Roepstorff P. Using Optimized Collision Energies and High Resolution, High Accuracy Fragment Ion Selection to Improve Glycopeptide Detection by Precursor Ion Scanning. *Journal of the American Society for Mass Spectrometry* **2003**;14:777-784

22. Harvey DJ. Identification of Protein-Bound Carbohydrates by Mass Spectrometry. *Proteomics* **2001**;1:311-328

23. Chalkley RJ, Burlingame AL. Identification of GlcNAcylation Sites of Peptides and Alpha-Crystallin Using Q-TOF Mass Spectrometry. *American Society for Mass Spectrometry* **2001**;12:1106-1113

24. Mills PB, Mills K, Johnson AW, Clayton PT, Winchester BG. Analysis by Matrix Assisted Laser Desorption/Ionisation-Time Of Flight Mass Spectrometry of the Post-Translational Modifications of Alpha1-Antitrypsin Isoforms Separated by Two-Dimensional Polyacrylamide Gel Electrophoresis. *Proteomics* **2001**;1:776-786

25. Dell A, Morris HR. Glycoprotein Structure Determination by Mass Spectrometry. *Science* **2001**;291:2351-2356

26. Cooper CA, Gasteiger E, Packer NH. GlycoMod - A Software Tool for Determining Glycosylation Compositions from Mass Spectrometric Data. *Proteomics* **2001**;1:340-349

27. Jung E, Veuthey A-L, Gasteiger E, Bairoch A. Annotation of Glycoproteins in the SWISS-PROT Database. *Proteomics* **2001**;1:262-266

28. Racusen D. Glycoprotein Detection in Polyacrylamide Gel with Thymol and Sulfuric Acid. *Analytical Biochemistry* **1979**;99:474-476

29. Clamp JR, Hough L. The Periodate Oxidation of Amino Acids with Reference to Studies on Glycoproteins. *Biochemical Journal* **1965**;94:17-24

30. Wolfe CAC, Hage DS. Studies on the Rate and Control of Antibody Oxidation by Periodate. *Analytical Biochemistry* **1995**;231:123-130

31. Hakimuddin T, Sojar T, Bahl OP. Chemical Deglycosylation of Glycoproteins. *Methods in Enzymology* **1987**;27:341-350

32. Bio-Rad. Protein Blotting - A Guide to Transfer and Detection.

33. Kanamori M, Kawaguchi N, Ibuki F, Doi H. Attachment Sites of Carbohydrate Moieties to Peptide Chain of Bovine K-Casein from Normal Milk. *Agricultural and Biological Chemistry* **1980**;44:1855-1861

34. Yan SCB, Grinnell BW, Wold F. Post-Translational Modifications of Proteins: Some Problems Left to Solve. *Trends in Biochemical Science* **1989**;14:264-268

Chapter Five

Investigations into Charge Heterogeneity of Wool Intermediate Filament Proteins

5.1 Introduction

Eight genes form the proteins that produce the type I and type II IFP spot pattern [1]. However, when run on 2D-PAGE, the IFPs separate out into approximately 24 major spots (Chapter Two). Isoelectric heterogeneity is generally considered an indication of PTM of a protein [2]. The most common PTMs of IFPs are phosphorylation and glycosylation [3]. In Chapters Three and Four both of these modifications were investigated and were not found in keratinised wool proteins.

5.2 Common causes of isoelectric heterogeneity in proteins

A review of past papers on isoelectric heterogeneity, that was not caused by PTM, suggested several possibilities. Conformational isomers, which exist in a state of equilibrium, may be responsible for charge heterogeneity of proteins when separated by denaturing IEF, as denatured proteins may adopt varied degrees of unfolding, resulting in differences in exposure of amino acid residues [4].

Isoelectric focusing can be a valuable technique when trying to show conformational change of proteins. Miled *et al.* [5] have used IEF as a tool to differentiate between closed and open forms of human pancreatic lipase (HPL). In most lipases, the active site is protected from the surrounding environment by loops and helices. In HPL, a lid consisting of 23 amino acid residues covers the active site when it is in its closed conformation. When the lid opens, the number of exposed negative residues increases and positive charges are masked, resulting in a decrease in the pI at which the

lipase focuses on an IPG gel. This demonstrates how conformational change in a protein can result in a change in the pI seen on a gel.

Microheterogeneity is common in enzymes [6]. Enzymatic conformers have the same sequence and molecular weight but differ in conformation, which results in different pIs [7]. The conformational differences may mean there is a change in exposed charges, which will show variations in electrophoretic mobility and behaviour on cation- and anion-exchange resins [7]. The theory that stable conformational variants are formed assumes that an amino acid sequence can give a three-dimensional structure with a set of stable conformations [6].

Isoelectric focusing has been used to show that some trains of spots on gels are caused by conformational isomerism. N-Terminal fragments of a *Porphyromonas gingivalis* outer membrane protein showed trains of spots when run on 2DE gels. When individual spots were eluted and re-electrophoresed, the complete set of spots was reproduced. Almost 100% of the amino acid sequences of the spots were identical. They concluded that the heterogeneity was related to conformational equilibria [8]. Experimental manipulations to try to reduce the heterogeneity, such as boiling in SDS and detergent, did not reduce the pI heterogeneity.

Purified pullulanase from spinach has also been shown to form discrete bands when separated by IEF [9]. When an individual band was electroeluted and refocused the whole set of isoforms was reproduced. Sometimes the band that had been eluted was represented the most, but the other bands were always present. They termed the isomers "conformers" and reasoned that the differences occurred due to more open or closed molecular shapes.

In 2001, Lutter *et al.* [10] showed that recombinant mistletoe lectin separated by 2D-PAGE produced series of spots independent of urea concentration, heat treatment or cysteine alkylating reagents. Mass spectrometry gave very high sequence coverage and showed a lack of chemical, PTM, or non-

enzymatic deamidation. When individual spots were resubjected to 2DE, the initial spot pattern was reformed. The observed charge variants were suggested to be caused by conformational equilibrium and not by microheterogeneity in the primary structure of the protein.

Similar studies on *Methylococcus capsulatus* polypeptides showed that trains of spots could be reformed, independent of alkylation prior to the first dimension [2]. Folding isomers were determined to be caused by the separation procedure and not by covalent PTM. Covalent PTM as a result of carbamylation or deamidation was considered; however, it was reasoned that covalent modifications would produce a stable product and the modification process would not be readily reversible. The re-establishment of spots on both the acidic and basic side of the original spot also ruled out any covalent modification, as covalent modifications would cause a change in pI in one direction only.

These re-running studies have shown no effect of urea on the charge heterogeneity of proteins. However, in 1973 Ui [11] showed that urea could cause a reduction in the heterogeneity of human serum albumin and caused a shift to a higher pI. He concluded that human serum albumin exists as a broad distribution of molecular species, due to the variation in the number of groups buried within the interior of the molecule, and that urea denaturation normalised these groups by exposing them. Ui also studied concanavalin A, and concluded that the different pIs could be explained by increasing amounts of side chain groups with a high pKa value being buried within the core of the molecule [11].

5.3 Possible causes of isoelectric heterogeneity in wool IFPs

Urea denaturation studies in wool have provided evidence that, in 6 M urea, IFPs have a random coil conformation when the secondary structure is determined using optical rotary dispersion [12]. The same study showed that wool IFPs have the ability to retain a 50% α-helical conformation after reduction, extraction and alkylation [12]. Other studied proteins, such as

ribonuclease, bovine serium albumin and insulin, have little or no α-helical structure after disruption to their disulfide bonds [12]. Denaturation may be very important when determining structural differences in wool IFPs using IEF. If several forms of a partially denatured α-helix structure still remained after denaturation, then the conformational isomers would be expected to have differing exposures of acidic and basic groups, which would show on a 2DE gel as isoelectric heterogeneity.

Heterogeneity of wool IFPs during IEF could reflect different charge states due to variations in the binding of the ampholytes to the proteins. Williamson *et al.* [13] discovered that the more acidic bands of bovine plasma albumin were formed by noncovalent binding to an unidentified component of the ampholine carrier ampholytes. In this work, ampholyte binding was tested to determine whether ampholytes were the cause of charge heterogeneity in wool IFPs. The possibility that conformational isomers were causing the heterogeneity was then tested.

This chapter presents work to determine what causes the charge heterogeneity seen during IEF of the IFPs, and whether the charge heterogeneity is related to urea denaturation, disulfide bonding and conformation.

5.4 Results & Discussion

5.4.1 Spot labelling design

To assist with identification of the major protein spots of both the type I and type II IFPs, a numbering system was designed. The spots were numbered from the acidic end of the gel, from highest molecular weight to lowest. The major type I IFP spots were numbered 1-11 (Figure 5.1) and the major type II IFPs were numbered 12-24 (Figure 5.2).

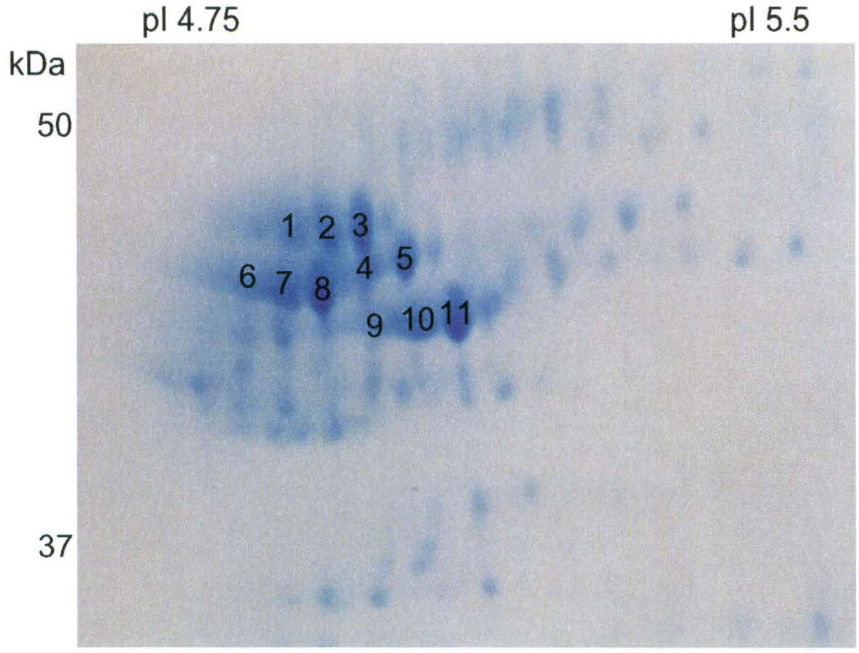


Figure 5.1
Major type I IFP spots labelled 1-11

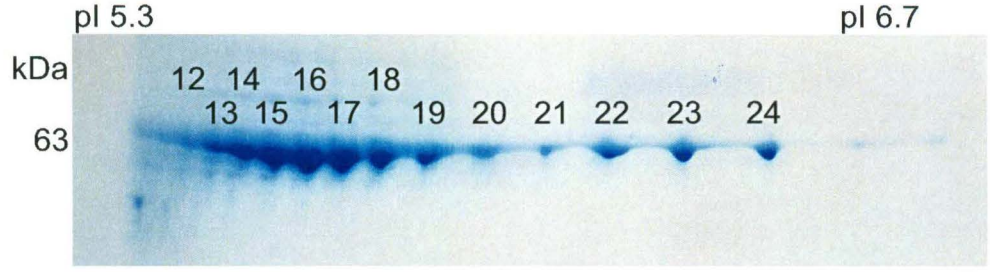


Figure 5.2
Major type II IFP spots labelled 12-24

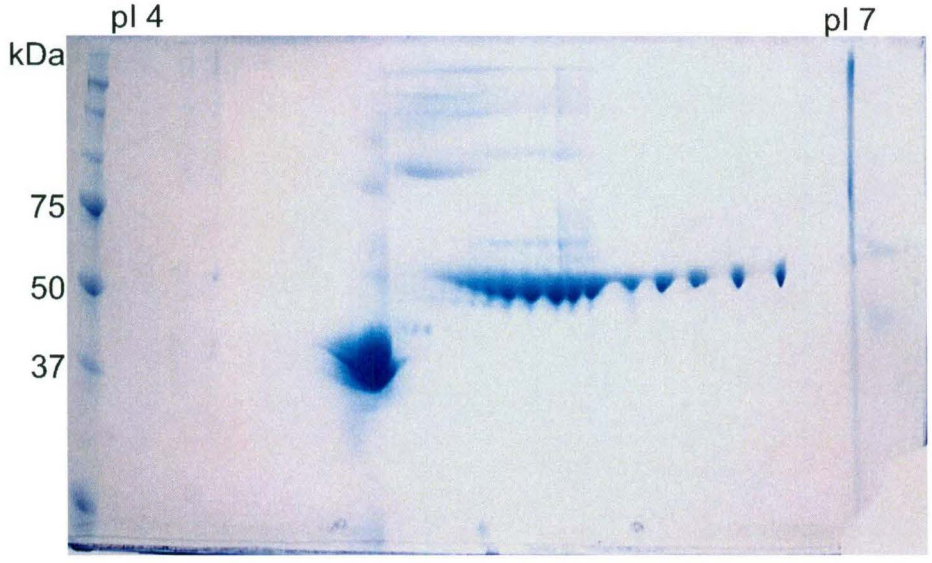
5.4.2 IEF rehydration without Pharmalyte™

A past paper on allomorphism and microheterogeneity of proteins [13] suggested that charge heterogeneity of wool IFPs could be due to the proteins binding to varying amounts of ampholytes. Pharmalytes™ are carrier ampholytes that are co-polymerised with glycine, glycylglycine, amines and epichlorohydrin. They contain numerous ampholytes, which have a high buffering capacity. Pharmalytes™ are used in the rehydration solution for IEF as they form an extremely stable linear pH gradient, exhibiting even conductivity across the gel.

Gels that were run with no Pharmalyte™3-10 in the rehydration solution showed no differences in the focusing pattern when compared to gels run in the presence of Pharmalyte™3-10 (Figure 5.3). The binding of ampholytes to proteins appears to be a problem in the alkaline ranges of IPGs, and in IPG gels which are not properly washed before use [14]. The majority of IPG gels used in wool protein separation use acidic pH ranges (most commonly pH 4-7). The experiments done here, with wool proteins, use commercial IPG strips, which were well washed before use.

Figure 5.3 demonstrates that removing the Pharmalyte™3-10 from the rehydration solution did not reduce the charge heterogeneity of the IFPs. This suggests that carrier ampholytes do not have any effect on the IEF pattern seen when focusing IFPs.

A)



B)

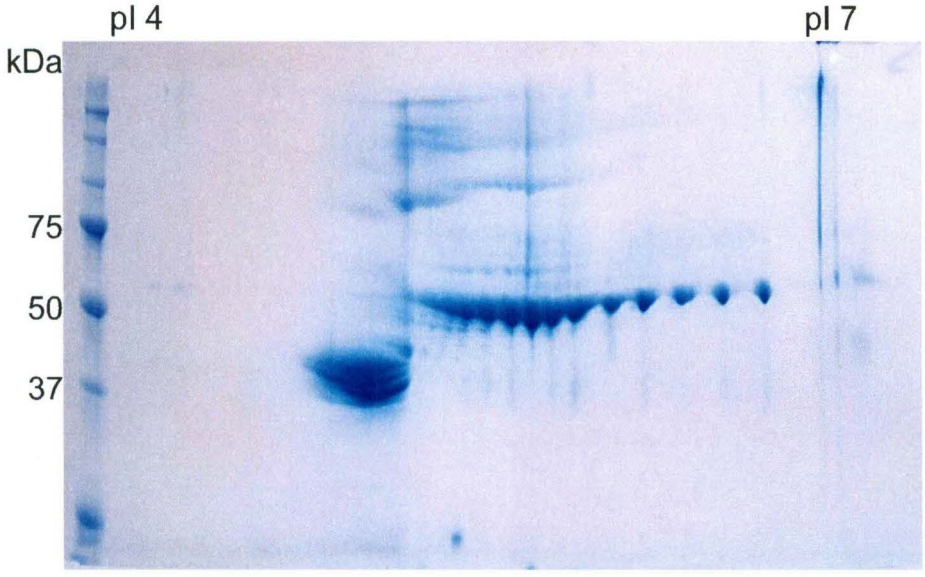


Figure 5.3
A) 2DE separation of wool IFPs without Pharmalyte™ 3-10 in the rehydration solution. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10% T gel. B) 2DE separation of IFPs with Pharmalyte™ 3-10 in the rehydration solution. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10% T gel. The gels were stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

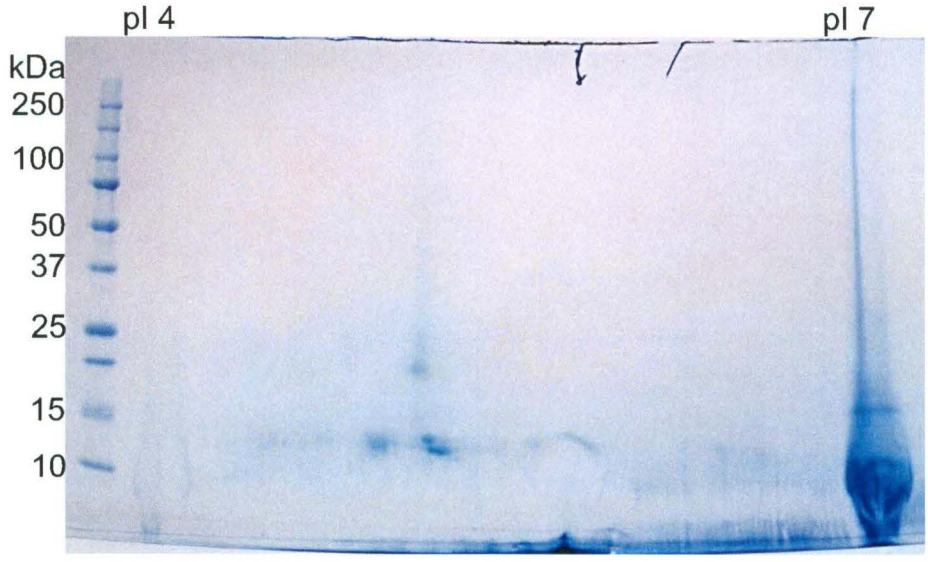
5.4.3 Varying urea concentrations in the rehydration solution

Isoelectric focusing in rehydration solution with low concentrations of urea showed no or very pale IFP spots (Figure 5.4 A, and B). Isoelectric focusing in rehydration solution containing high concentrations of urea showed the rows of spots usually seen when IFPs are separated by 2DE (Figure 5.4 C). The rows of spots are not as well resolved as the separation seen in Chapter Two, (Figure 2.3 and 2.4) as thiourea, a chaotrope, was not included in the rehydration solution when the effect of urea was tested.

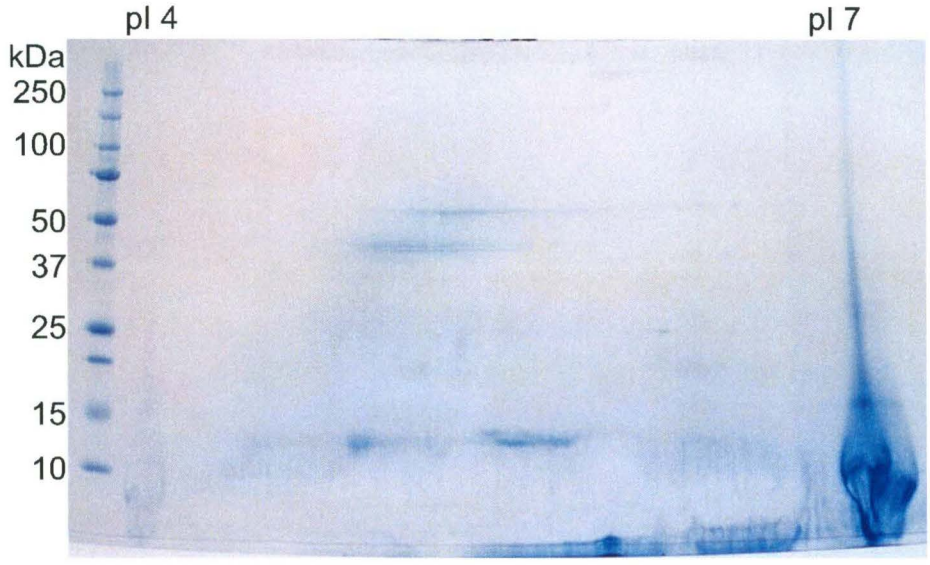
Low amounts of urea do not sufficiently solubilise the IFPs for IEF. However, when higher concentrations of urea are used, the focusing pattern remains the same as when using medium concentrations of urea. This suggests that the charge heterogeneity of the IFPs is not related to the amount of urea

denaturant used in the rehydration solution. In Chapter Two, the importance of chaotropic agents in the rehydration solution was discussed (2.2.4). In high concentrations of urea it would be expected that all hydrogen bonds and hydrophobic interactions would be minimised.

A)



B)



C)

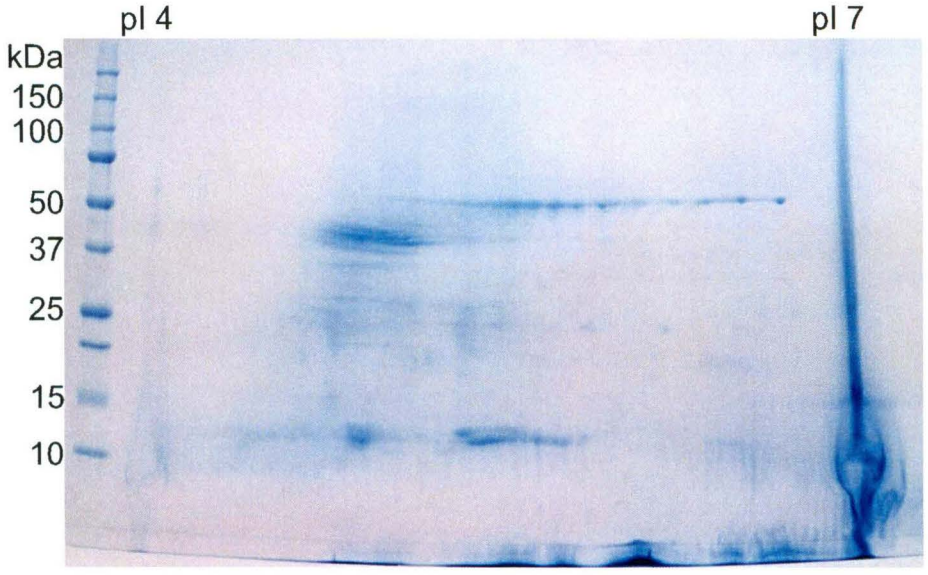


Figure 5.4
2DE separation of wool IFPs with A) 2 M, B) 5 M, and C) 8 M urea in the rehydration solution. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10-20% T gel. The gels were stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

5.4.4 Mass Spectrometry

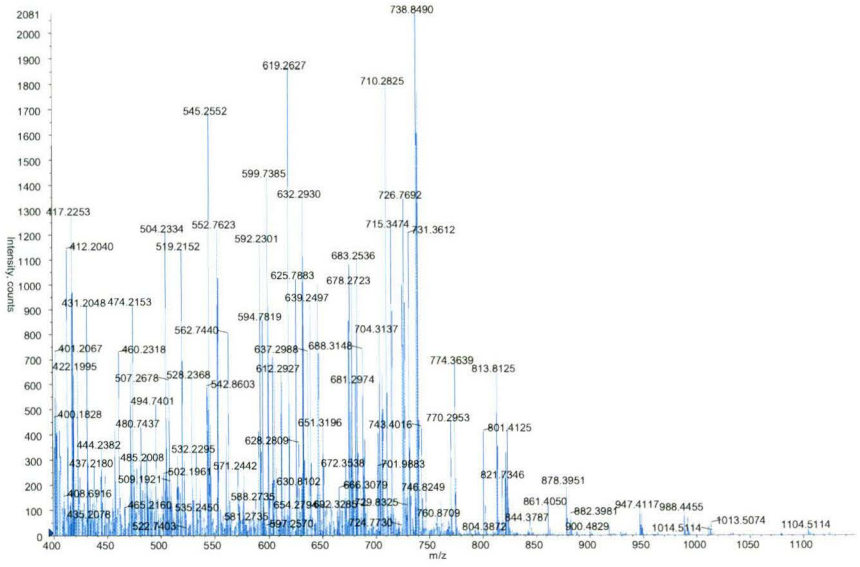
Two neighbouring type I IFP spots (spots 10 and 11) and 2 neighbouring type II IFP spots (spots 23 and 24) were excised, digested and analysed by peptide mass fingerprinting to determine whether their sequences were the same. Examples of the peptide mass fingerprint spectra are shown in Figure 5.5.

The mass spectrometry results suggest that the proteins are identical but don't exclude the possibility that PTMs are causing the pI shift. When the two type I IFPs peak lists were compared, by matching the peak lists generated from the peptide mass fingerprint data, 124 peaks matched and 15 did not match (89.2% match). When the type II IFPs peak lists were compared 342 peaks matched and 46 did not match (88.1% match) (Appendix Three, A3.1). This suggests that there is a very high chance that the two proteins in each case have identical sequences. Past mass

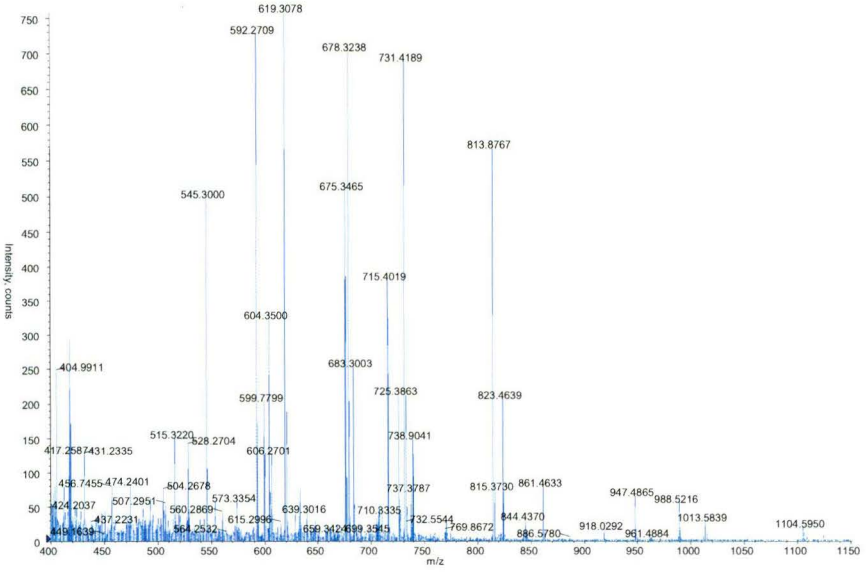
spectrometry on wool IFPs has used %coverage (a comparison of the peptide mass fingerprint data to a protein sequence in a database) as low as 17.7% to identify the protein [15]. 100% sequence identity would be required to confirm that the proteins have the same sequences. However, many of the tryptic peptides formed during digestion are either too long or too short to be detected [16]. Different proteases can be combined to give optimal sequence coverage of proteins. Each protein results in a unique set of peptide masses after cleavage with a specific protease [17]. Several digestions with different proteases can produce a series of overlapping sequences, which can be aligned and combined into longer sequences or even the entire protein sequence [17, 18]. Currently no computer software is available for this task [17].

Another problem with using mass spectrometry to gain 100% sequence coverage occurs due to the different hydrophobicities of the peptides produced [19]. Hydrophobic compounds are very insoluble, they may require specific solvents to keep them in solution during mass spectrometry analysis [19].

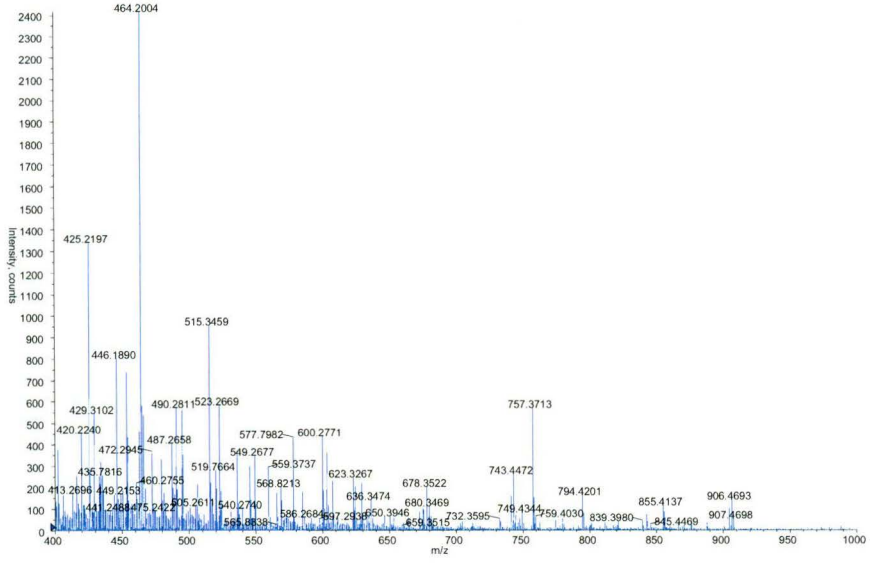
A)



B)


Intensity, counts
m/z
619.3078
592.2709
678.3238
731.4189
813.8767
545.3000
675.3465
715.4019
604.3500
683.3003
404.9911
599.7799
725.3863
823.4639
515.3220
528.2704
738.9041
606.2701
417.2587
431.2335
456.7455
474.2401
504.2678
737.3787
815.3730
861.4633
947.4865
988.5216
424.2037
507.2951
573.3354
639.3016
710.8335
844.4370
1013.5839
437.2231
560.2869
615.2996
732.5544
769.8672
918.0292
1104.5950
886.5780
961.4884

C)


Intensity, counts
m/z
464.2004
425.2197
515.3459
446.1890
429.3102
490.2811
523.2669
757.3713
420.2240
577.7982
600.2771
487.2658
472.2945
549.2677
559.3737
519.7664
623.3267
743.4472
435.7816
460.2755
568.8213
678.3522
794.4201
636.3474
680.3469
855.4137
906.4693
749.4344
759.4030
839.3980
845.4469
907.4698

D)

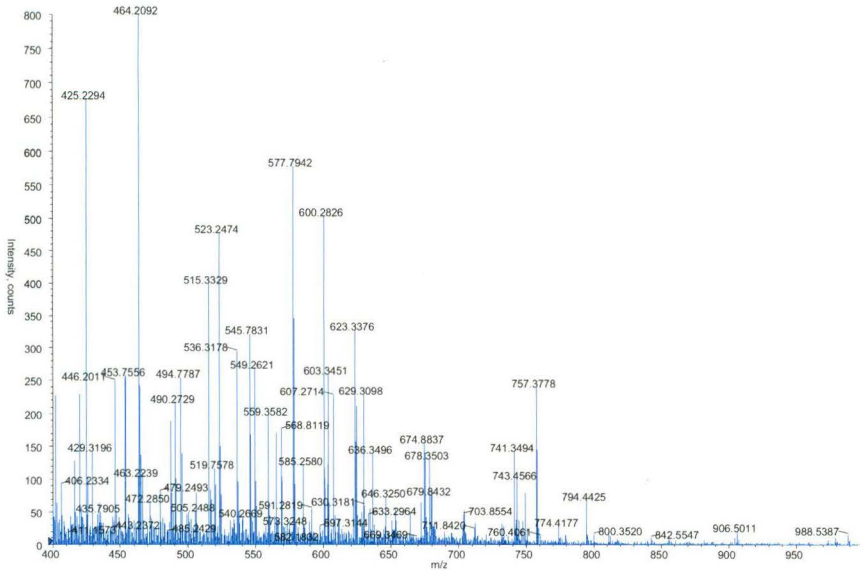


Figure 5.5
Peptide mass fingerprints of spot numbers A) 24, B) 23, C) 10 and D) 11 excised from a 2DE gel. The spectra are representative of duplicate experiments.

To help with interpretation of mass spectra, peptides can be tagged using protein-chemical means [17]. Methyl groups can be added to the free carboxyl groups in the side chains of aspartic acid and glutamic acid to yield an increase of mass by 14 Da [20]. Tyrosine residues can be modified by iodination, which will give a mass increase of 126 Da per tyrosine [20]. Procedures are being developed to increase or decrease ion intensities. Spectra complexity can be reduced by knocking out a complete ion-series. Ion-series intensities can be increased by attaching positively charged groups or masking negative charges. Conversion of the C-terminal lysine to homoarginine can be used to increase the ionisation of the y-ion-series [17].

Techniques to improve the sequence coverage of the IFPs were not employed here. The intransigent nature of wool proteins made it unlikely that sequence coverage would be improved. Gel-based techniques have recently been developed that can determine whether protein spots are covalently modified [2]. This technique can also determine whether the spots separated by 2DE into rows had the same sequence.

5.4.5 Re-running

Re-running the type I IFP spots showed that when one spot from each row of the type I IFPs was eluted, re-run and pooled with single spots from the other rows, the original spot pattern of the major proteins in the type I IFP area was reformed (Figure 5.6).

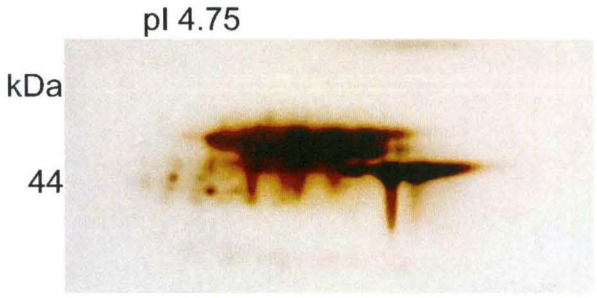


Figure 5.6
Spots numbered 3, 4, 8 and 11 were eluted, pooled and re-run to reform the type I IFP pattern. Gels were stained with the Blum silver stain. The gel is representative of triplicate experiments.

Four type II IFP spots can be eluted, pooled and re-run to form the original type II IFP spot pattern (Figure 5.7).

Re-running of individual type II IFP spots showed that trains of spots could be reformed from a single spot (Figure 5.8).

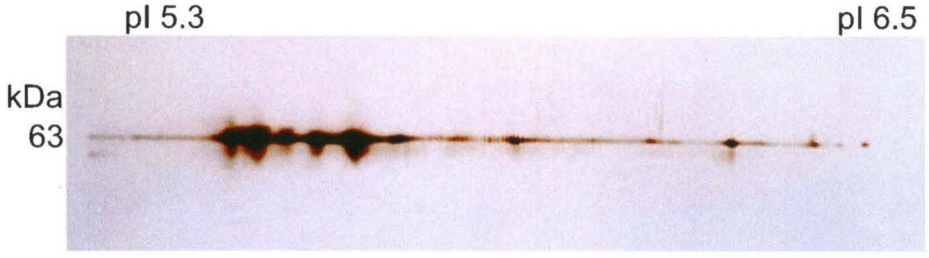


Figure 5.7
Spots numbered 15, 18, 21 and 24 were eluted, pooled and re-run to reform the type II IFP pattern. Gels were stained with the Blum silver stain. The gel is representative of triplicate experiments.

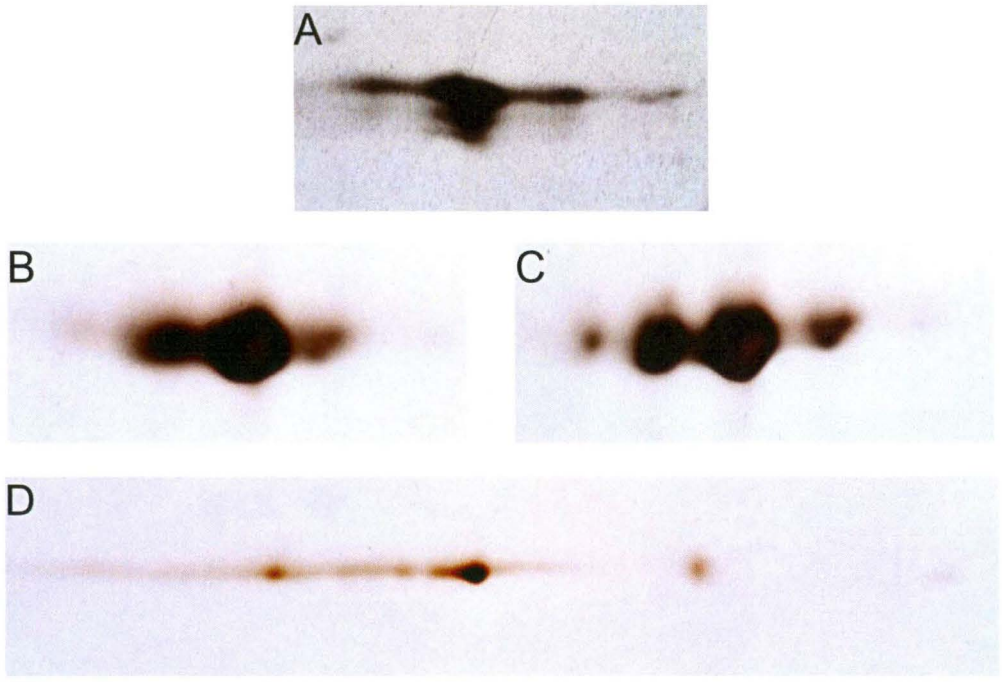


Figure 5.8
A) Spot number 15 re-run to form spots 14, 15, 16 and 17 B) Spot number 16 re-run to form spots 15, 16 and 17. C) Spot number 17 re-run to form spots 15, 16, 17 and 18. D) Spot number 24 re-run to form spots 24, 23 and 22. Gels were stained with the Blum silver stain. The gels are representative of triplicate experiments.

Individual spots that were cut out were carefully excised within their borders to eliminate any contamination from neighbouring spots. Spot contamination from excising too close to the edges of spots would be expected to produce the spot directly beside the original. However, spots that were eluted from the most acidic end of the rows were able to form spots at the most basic end of the rows and *vice versa*. This demonstrates that the original spot was able to form spots that were not directly beside the original, ruling out excision contamination as a potential cause of multiple spots. The type II IFP spots are spread over a much wider pI range than the type I IFP spots, making excision of individual spots much easier. At the basic end of the gels, spots were separated by as much as 15 mm, which substantially reduced the possibility that excision contamination could be causing spot heterogeneity.

Contaminated spots were easy to recognise. Figure 5.9 shows a spot, which may contain two separate proteins that have been excised and eluted as

one. When they are re-run, they form spots that have a slightly different molecular weight. This could mean that the spot that was originally excised was actually made up of two spots from two different families. The high-resolution separation of IFP type II IFPs seen in Chapter Two (Figure 2.8), show that there are possibly three rows of vertically separated proteins at the acidic end and one row of proteins separated vertically at the basic end. To allow better resolution of the spot families at the acidic end of the type II IFPs, improved vertical resolution is required.

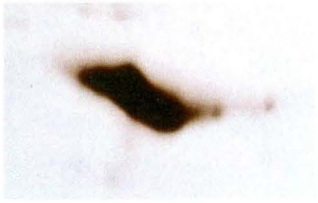


Figure 5.9
Spot number 14 re-run to form spots 13, 14 and 15. The gel was stained with the Blum silver stain. The gel is representative of triplicate experiments.

When the spots were re-run, the spot that had been cut out and eluted tended to form a larger spot than the others, suggesting that equilibrium between the different conformations had not been reached. After spots had been eluted, they were placed in rehydration solution. During this time in rehydration solution, the proteins would have been reforming the original spot pattern. Proteins were unable to be left in rehydration solution to see if complete equilibrium could be reached, as the reducing component of the solution becomes less effective the longer it is in solution, creating artifactual results [21, 22]. The reducing component can not be replenished during rehydration, as high concentrations in the rehydration solution lead to smearing on 2DE gels [23].

Spot number 2 (Figure 5.10, A) was able to reform spots on both the acidic and basic side of the spot that was originally eluted. This demonstrates that a covalent modification could not be responsible for the charge

heterogeneity. Covalent modifications, which were able to have an effect on the charge of the protein, would only move the gel spot in one direction on the gel. For example, if a protein was phosphorylated, the gel spots would migrate to a more acidic pI with each addition of a phosphate group. All of the internal type II IFP spots (13-23) that were re-run were able to produce spots on both the acidic and basic side of the original spot, showing that covalent modifications are not responsible for the charge heterogeneity seen in the type II IFPs (Figure 5.8).

When the type I IFPs are re-run, they clearly form four rows of spots that are vertically separated. All of the type II IFPs that are re-run on small format gels appear to resolve at approximately the same molecular weight. High-resolution separation of the type II IFPs (Figure 2.4) shows there is a clear vertical separation between spots at the basic end of the type II IFP row. However, on small format gels, the resolution is not high enough to separate out the spots vertically. This resulted in the two rows of spots that are separated on high-resolution gels being merged into one row of spots on a small format gel. When spots 20 and 21 from small format gels were excised, they contained more than one spot. When they were re-run, one row of spots was produced for each protein.

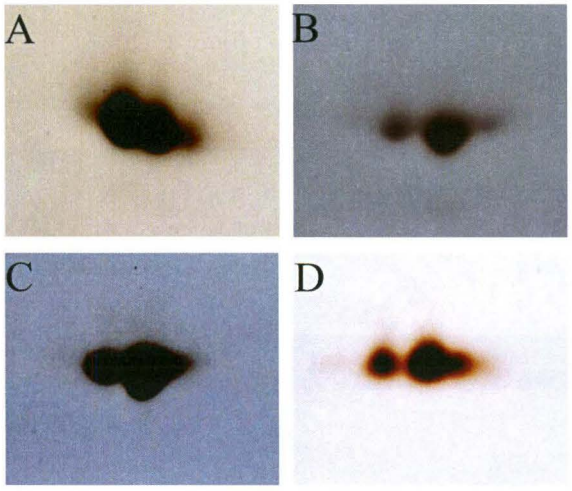


Figure 5.10
A) Spot number 2 re-run to form spots 1,2 and 3. B) Spot number 5 re-run to form spots 4 and 5. C) Spot number 7 re-run to form spots 6,7 and 8. D) Spot number 11 re-run to form spots 9,10 and 11. Gels were stained with the Blum silver stain. The gels are representative of triplicate experiments.

The 2DE gel separation combined with the re-running results suggest that there are four different type I IFPs, which is in agreement with genetic data. Four genes form the proteins that produce the type I IFP spot pattern. There were clearly four rows of spots with each row having one protein represented. Using the original electrophoretic separations of wool IFPs (Chapter Two Introduction) the wool type I IFPs can be placed into the original classification system based on their molecular weights. Spots 1-3 represent isoforms of 8c-1, spots 4-5 isoforms 8c-2, spots 6-8 isoforms 8b and spots 9-11 isoforms 8a (Table 5.1).

Using the original classification system and the differences seen in molecular weight on a 2DE gel, the three spots families at the acidic end of the type II IFPs can be assigned from highest molecular weight to lowest as 7a, 7c and 7b. The type II IFP spots at the basic end of the row represent the family originally labelled 5 (Figure 5.11).

Original classification	Spots
8c1	1-3
8c2	4 + 5
8b	6-8
8a	9-11

Table 5.1
Type I IFPs classified based on the original electrophoretic separations [24, 25].

Other techniques commonly used to study protein folding such as, ultraviolet difference spectroscopy, circular dichroism, fluorescence and nuclear magnetic resonance (NMR), require a stable, pure, protein sample [6, 26]. The results from the re-running work show that the different conformations of IFPs interconvert readily, making them not stable enough for use with the usual techniques. Even if the spots were stable, gaining enough protein out

of single gel spots would be challenging. Therefore, an alternative method was required to corroborate the theory that each train of spots represented a single protein.

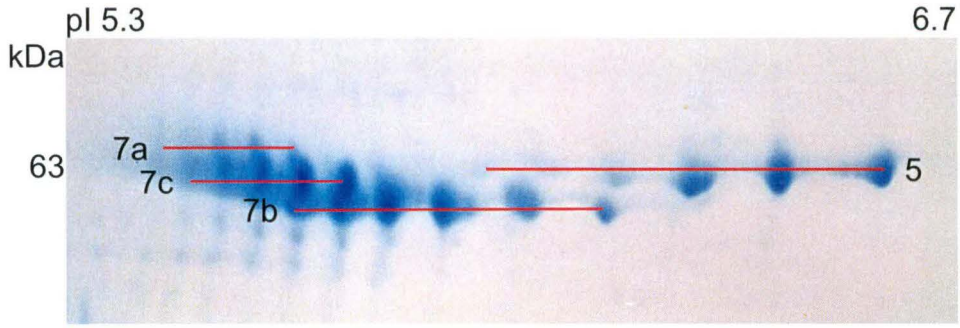


Figure 5.11
Assignment of the type II IFPs to families based on the original classification system [24, 25].

5.4.6 Alkylation time course

An alkylation time course was used to determine whether the charge heterogeneity of the IFPs was associated with variation in disulfide bonding within the monomer. Variations in disulfide bonding could lead to different residues being exposed, altering the pI of the protein. In bovine plasma albumin, different permutations in disulfide pairings have been shown to be a source of microheterogeneity [27].

Three different alkylating agents were tried, IAM, N-ethylmaleimide (NEM), and acrylamide [14, 28-30]. Long alkylation times led to reduced charge heterogeneity in the type II IFPs (Figure 5.12). IAM gave the most successful results and will be discussed here; the other alkylating agents gave spot patterns consistent with those seen with IAM (Appendix Three, A3.2, NEM and acrylamide alkylation time course figures).

At 6 hours (Figure 5.12 E) there was a slight compression in the type II IFP rows of spots and by 24 hours the type II IFPs had compressed towards the acidic side, to the extent that they separated out in approximately half the

distance that they did before alkylation (Figure 5.12 F). There was a large decrease in the spot resolution, with the protein spots at 24 hours forming a continuous row with no individual spots being identifiable.

The type I IFP spots did not show any large changes. After 10 minutes, the whole type I IFP area moved to a slightly more acidic pI, but no further change was observed with longer incubation times. The resolution of the type I IFP spots improved after 10 minutes alkylation, but then became progressively more smeared after 10 minutes. The resolution was very poor at alkylation times above two hours, making individual spots hard to discern (Figure 5.12). The alkylation time-course results could mean that the heterogeneity of the type II IFPs is partially due to scrambled disulfide bonds within the monomer protein. This theory, however, cannot be used to explain the heterogeneity of the type I IFPs, as alkylation appeared to have no effect on their pI spread. Another possibility is that the type II IFPs are more susceptible to over-alkylation, other amino acids such as histidine, lysine and methionine may become alkylated when all cysteines have been alkylated and other amino acids are then exposed to alkylating reagents for long periods of time [14, 30]. In order to eliminate this possibility, the Ellman's assay was used to determine when all of the cysteines in the IFPs had been alkylated.

5.4.7 Ellman's assay

The Ellman's assay showed that, with all the alkylating agents used, virtually all of the free cysteine had been alkylated after 10 minutes (Figure 5.13). This confirms a mass spectrometry study on the effect of alkylation on wool proteins [30]. After 10 minutes alkylation, the major protein species were fully alkylated at cysteine. It was suggested that side reactions, such as alkylation of lysine, histidine and methionine, could be responsible for the observed pI shift and loss of resolution of the IFPs with increasing alkylation times [30].

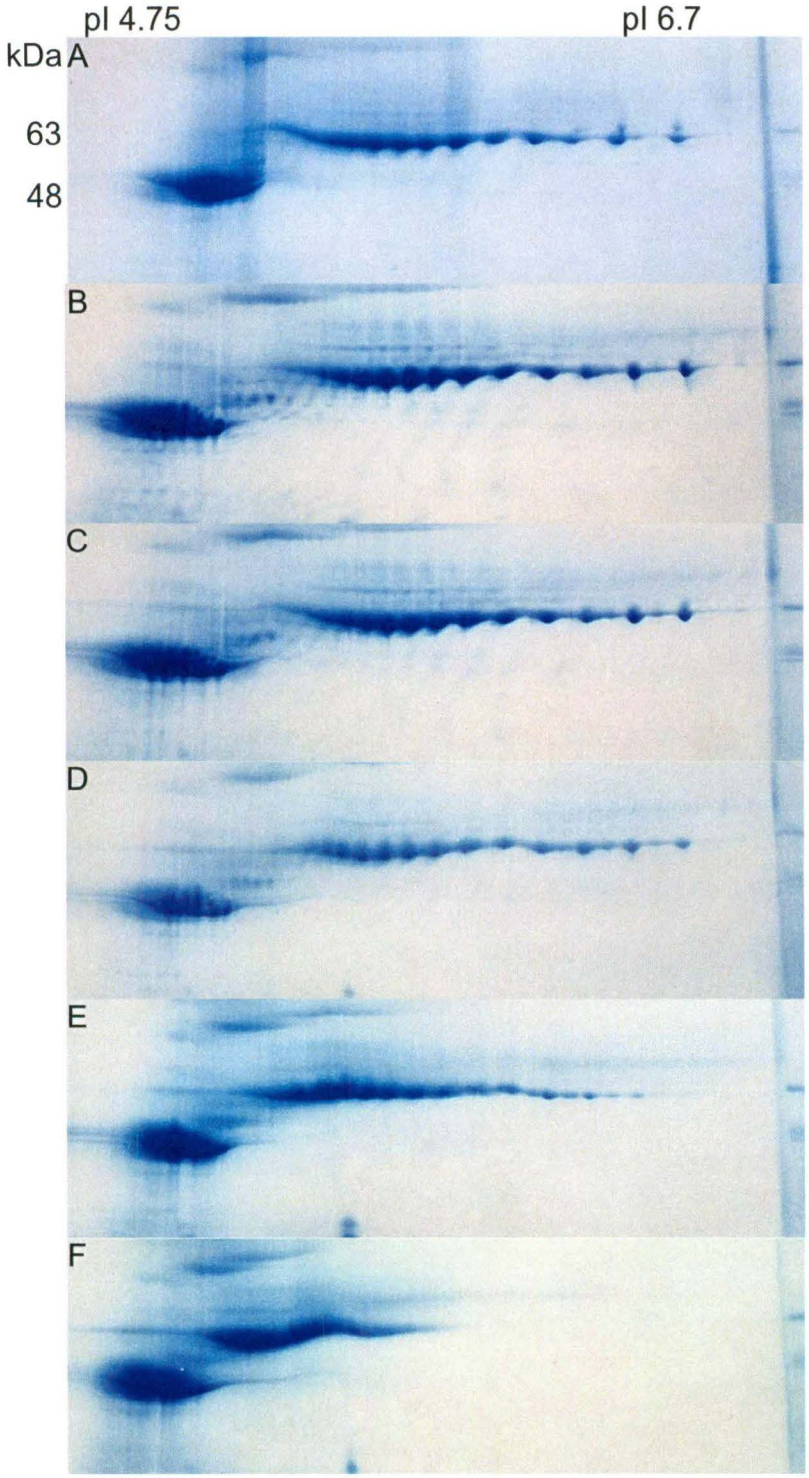


Figure 5.12
2DE separation of wool IFPs after alkylation with IAM for A) 0 minutes, B) 10 minutes, C) 1 hour, D) 2 hours, E) 6 hours and F) 24 hours. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10-20% T gel. Gels were stained with Coomassie brilliant blue G-250. The gels are representative of triplicate experiments.

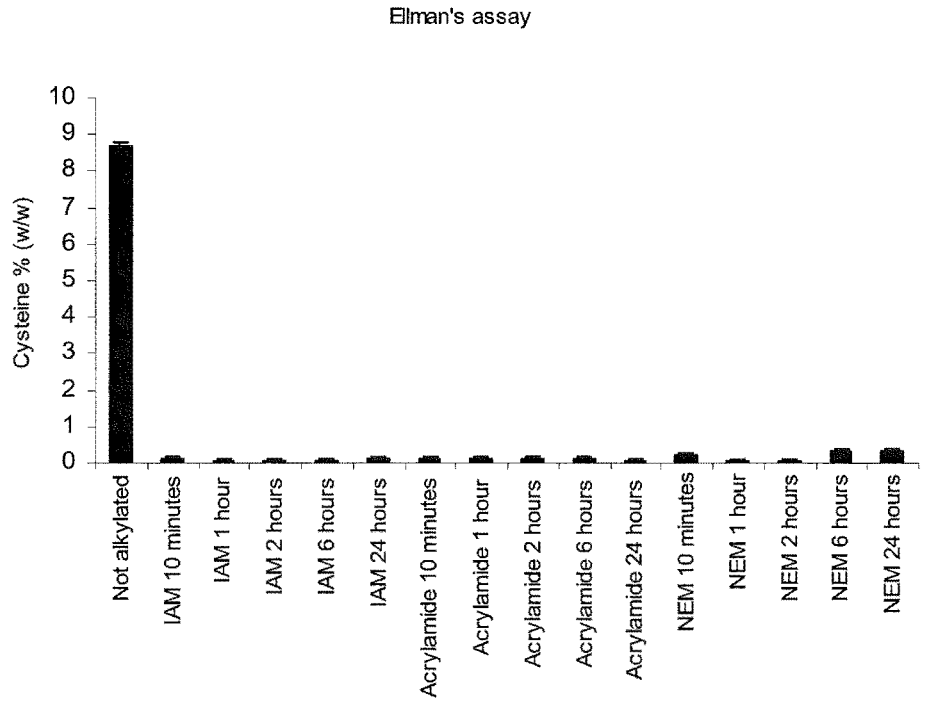


Figure 5.13
Amount of cysteine% (w/w) after alkylation with IAM, acrylamide or NEM for between 0 and 24 hours. Values are the mean of duplicate experiments. Error bars represent the standard deviation from duplicate experiments.

This assay ascertained that the charge heterogeneity could not be caused by variations in disulfide bonding within the monomer IFPs. Amino acid analysis was carried out on the IAM alkylated samples to establish whether other amino acids were being alkylated during the long alkylation times.

5.4.8 Amino acid analysis

Other amino acids that are susceptible to alkylation include lysine, histidine and methionine [30]. Amino acid analysis of their relative amino acid concentrations (Figure 5.14) shows that there was no methionine or histidine detected, even in the control samples. This was expected, as amino acid analysis of Merino wool has shown that histidine accounts for around 0.6% of the amino acid content and methionine only 0.4% [31].

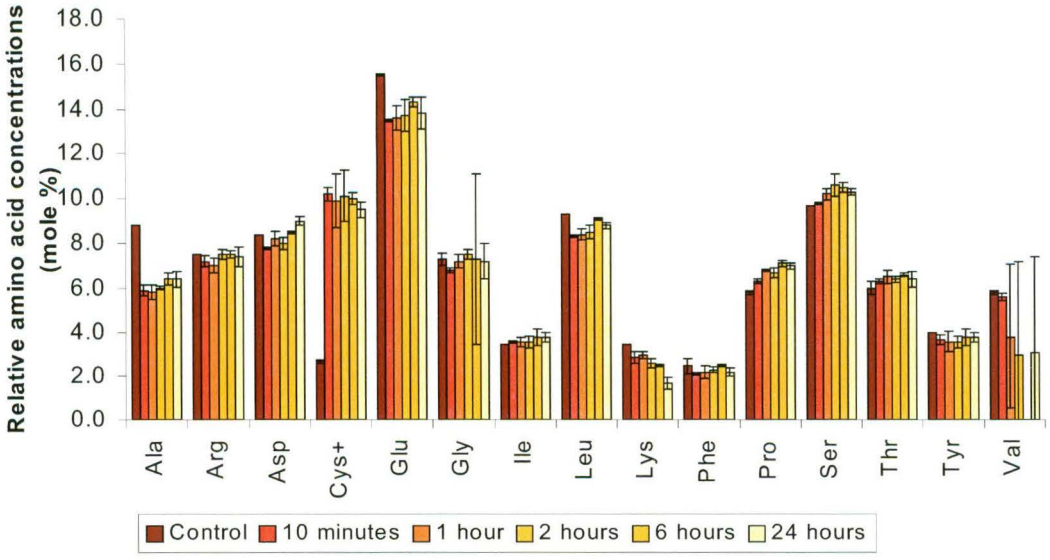


Figure 5.14
Relative amino acid concentrations of wool proteins after an alkylation time-course with IAM. Cys^+ is the combined values of IAM alkylated cysteine, cysteic acid and cysteine. Values are the mean of triplicate experiments. Error bars represent the standard deviation from triplicate experiments.

The preferred alkylation site of attack for alkylating reagents is the thiol group [14]. The reactivity of thiol groups of proteins is two to three orders of magnitude higher than other potential sites of alkylation [14]. The second most reactive, after the thiol, is the ε-amino group of lysine. Once all the thiol groups had been alkylated, the side chain of lysine was alkylated [14]. There appeared to be a reduction in the amount of lysine detected with increasing alkylation times, especially at 24 hours. There was a significant reduction in lysine detected after 6 hours of alkylation with IAM. This suggests that the shift in pI and loss of resolution of the IFPs, when alkylated for long periods, then run on 2DE gels, is due to overalkylation, which results in the lysine residues becoming alkylated. This suggests that when the lysine residues were alkylated, they were unable to maintain their folded formations. Presumably because the lysines were involved in stabilising the folded state of particular conformers causing the pI shift of the protein on a 2DE gel.

The charge heterogeneity of wool IFPs appears to be due to differences in conformation of the monomer proteins on the gels. Different molecular shapes of the IFPs in rehydration solution leads to the burying and exposing

of charged groups. The charge differences caused by alternative arrangements of residues leads to a series of spots on a gel. Each spot must be a semi-stable product, as in high-resolution gels very little smearing is seen between gel spots.

Charge differences are also seen in the wool HGTPs. The wool HGTPs contain several families of proteins. Rogers [32] has suggested that the numbers of different polypeptide chains may be a conformational anomaly. Studies used a cDNA clone that was isolated from a cDNA library of follicle mRNA. The genomic clone was Southern blotted to reveal a single gene, even at low stringency. When the library was screened, the number of recombinants distinguished by the clone was much smaller than expected, as there were approximately 20 variants expected. The conclusion was that either cross-hybridisation was not working or the HGTPs were not as complex as suggested by the protein data [32].

Genetic data shows that eight genes are responsible for producing the IFPs [1]. Four genes form the proteins that produce the type I IFP spot pattern and four genes form the proteins that produce the type II IFP spot pattern [1]. The high-resolution 2DE separation combined with the re-running results has enabled the assignment of the type I IFPs into individual protein families and has allowed the tentative assignment of the type II IFPs.

The different conformations could be related to the "swinging head" hypothesis, which suggests that in the keratinised wool fibre the IFP's head domain folds back and interacts with the rod domain [33]. Models also suggest that the head domain may wrap around the rod domain forming strong interactions [34]. When the proteins are extracted and denatured, variations in the association of the terminal groups to the rod domain may give different conformations (Figure 5.15).

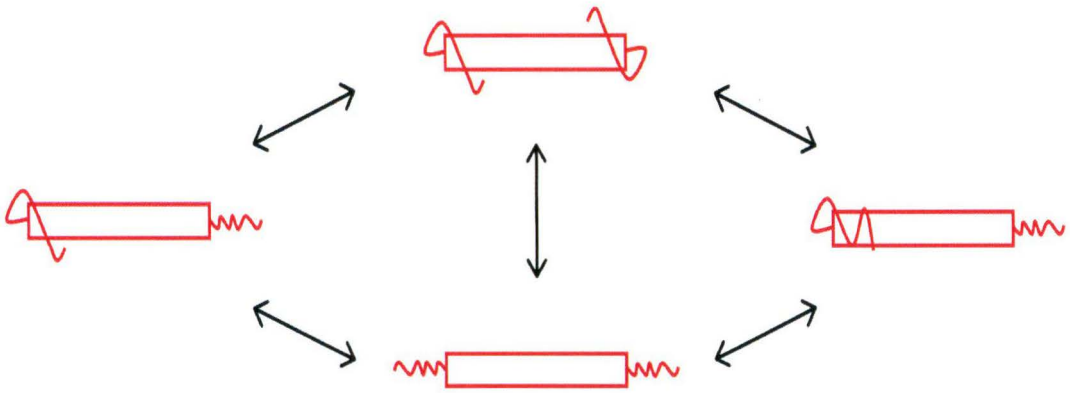


Figure 5.15
Diagrammatic representation of possible conformations of an IFP.

This hypothesis suggests that IEF conditions are not fully denaturing, as a random coil form has clearly not been achieved. Controlled alkylation of the lysine residues followed by mass spectrometry may reveal which lysines are involved in masking the charges from other amino acid residues. Little can be done to increase the denaturing conditions when running IEF other than those already tried (chaotrope concentration, SDS during extraction, boiling and alkylation). Isoelectric focusing is sensitive to the use of non-zwitterionic or charged chemicals, therefore, future work to increase the denaturing capacity using zwitterionic and neutral chemicals is required.

5.5 Summary

Two-dimensional gel electrophoresis analysis has excluded the possibility that the spot heterogeneity of the IFPs is caused by binding of ampholytes to the proteins. Using high concentrations of urea in the rehydration solution did not reduce the spot heterogeneity. Amino acid analysis, mass spectrometry and Ellman's assay results support the hypothesis that the proteins have the same sequence but vary in isoelectric charge due to differences in exposure of charged residues on the molecular surface. The cause of IFP charge heterogeneity is concluded from these results to be a conformational equilibrium between several different forms of the same protein in the rehydration solution used for the first dimension. This contrasts to earlier work on the heterogeneity of wool IFPs by Ben Herbert *et al.* [35,

36]. The results from this chapter showing conformational equilibrium in the wool IFPs suggest that earlier results showing that the wool IFPs were phosphorylated were artifactual.

5.6 References

1. Heid HW, Werner E, Franke WW. The Complement of Native Alpha-Keratin Polypeptides of Hair-Forming Cells: A Subset of Eight Polypeptides that Differ from Epithelial Cytokeratins. *Differentiation* **1986**;32:101-119
2. Berven FS, Karlsen OA, Murrell JC, Jenson HB. Multiple Polypeptide Forms Observed in Two-Dimensional Gels of *Methylococcus capsulatus* (Bath) Polypeptides are Generated During the Separation Procedure. *Electrophoresis* **2003**;24:757-761
3. Chou C-F, Smith AJ, Omary MB. Characterization and Dynamics of O-Linked Glycosylation of Human Cytokeratin 8 and 18. *Journal of Biological Chemistry* **1992**;267:3901-3906
4. Chang J-J, Märkl W, Lai P-H. Analysis of the Extent of Unfolding of Denatured Insulin-Like Growth Factor. *Protein Science* **1999**;8:1463-1468
5. Miled N, Riviere M, Cavalier JF, Buono L, Verger R. Discrimination Between Closed and Open Forms of Lipases Using Electrophoretic Techniques. *Analytical Biochemistry* **2005**;338:171-178
6. Epstein CJ, Schechter AN. An Approch to the Problem of Conformational Isozymes. In: Vesell ES, ed. *Multiple Molecular Forms of Enzymes*. Proceedings of The New York Academy of Sciences, **1968**:85-101
7. Kaplan NO. Nature of Multiple Molecular Forms of Enzymes. In: Vesell ES, ed. *Multiple Molecular Forms of Enzymes*. Proceedings of The New York Academy of Sciences, **1968**:382-399
8. Veith PD, Talbo GH, Slakeski N, Reynolds EC. Identification of a Novel Heterodimeric Outer Membrane Protein of *Porphyromonas gingivalis* by Two-Dimensional Gel Electrophoresis and Peptide Mass

Fingerprinting. *European Journal of Biochemistry* **2001**;268:4748-4757

9. Henker A, Schindler I, Renz A, Beck E. Protein Heterogeneity of Spinach Pullulanase Results from the Coexistence of Interconvertible Isomeric Forms of the Monomeric Enzyme. *Biochemical Journal* **1998**;331:929-935
10. Lutter P, Meyer HE, Langer M, Wittholn K, Dormeyer W, Sickmann A, Blüggel M. Investigation of Charge Variants of rViscumin by Two-Dimensional Gel Electrophoresis and Mass Spectrometry. *Electrophoresis* **2001**;22:2888-2897
11. Ui N. Conformational Studies on Proteins by Isoelectric Focusing. *Annals New York Academy of Sciences* **1973**;209:198-209
12. Harrap BS. The Conformation of a Soluble Wool Keratin Derivative. *Australian Journal of Biological Science* **1962**;16:231-240
13. Williamson AR, Salaman MR, Kreth HW. Microheterogeneity and Allomorphism of Proteins. *Annals New York Academy of Sciences* **1973**;209:210-224
14. Galvani M, Hamdan M, Herbert BR, Righetti PG. Alkylation Kinetics of Proteins in Preparation for Two-Dimensional Maps: A Matrix Assisted Laser Desorption/Ionization-Mass Spectrometry Investigation. *Electrophoresis* **2001**;22:2058-2065
15. Plowman JE, Bryson WG, Flanagan LM, Jordan TW. Problems Associated with the Identification of Proteins in Homologous Families: The Wool Keratin Family as a Case Study. *Analytical Biochemistry* **2002**;300:221-229
16. Westermeier R, Naven T. *Proteomics in Practice - A Laboratory Manual of Proteome Analysis.* Weinheim, Germany: Wiley-VCH Verlag-GmbH, **2002**:316
17. Reinders J, Lewandrowski U, Moebius J, Wagner Y, Sickmann A. Challenges in Mass Spectrometry-Based Proteomics. *Proteomics* **2004**;4:3686-3703
18. Doucette A, Liang L. Investigation of the Applicability of a Sequential Digestion Protocol Using Trypsin and Leucine Aminopeptidase M for

Protein Identification by Matrix-Assisted Laser Desorption/Ionisation-Time Of Flight Mass Spectrometry. *Proteomics* **2001**;1:987-1000

19. Szambó PT, Kele Z. Electrospray Mass Spectrometry of Hydrophobic Compounds Using Dimethyl Sulfoxide and Dimethylformamide as Solvents. *Rapid Communications in Mass Spectrometry* **2001**;15:2415-2419
20. Patterson SD, Aebersold R. Mass Spectrometric Approaches for the Identification of Gel-Separated Proteins. *Electrophoresis* **1995**;16:1791-1814
21. Cleland WW. *Dithiothreitol, A New Protective Reagent for SH Groups*. CA, USA: Calbiochem-Novabiochem Corporation, **1963**
22. Herbert BR, Galvani M, Hamdan M, Olivieri E, MacCarthy J, Pederson S, Righetti PG. Reduction and Alkylation of Proteins in Preparation of Two-Dimensional Map Analysis: Why, When, and How? *Electrophoresis* **2001**;22:2046-2057
23. Görg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W. The Current State of Two-Dimensional Electrophoresis with Immobilized pH Gradients. *Electrophoresis* **2000**;21:1037-1053
24. Crewther WG, Dowling LM, Gough KH, Inglis AS, McKern NM, Sparrow LG, Woods EF. The Low-Sulphur Proteins of Wool: Studies on Their Classification, Characterisation, Primary and Secondary Structure. In: *Proceedings of the 5th International Wool Textile Research Conference*. Aachen, Germany, **1975**:233-242
25. Marshall RC, Blagrove RJ. Successful Isoelectric Focusing of Wool Low-Sulphur Proteins. *Journal of Chromatography* **1979**;172:351-356
26. Pace CN. Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods in Enzymology* **1986**;131:266-280
27. Sogami M, Peterson HA, Foster JF. The Microheterogeneity of Plasma Albumins. V. Permutations in Disulfide Pairings as a Probable Source of Microheterogeneity in Bovine Albumin. *Biochemistry* **1969**;8:49-58

28. Sechi S, Chait BT. Modification of Cysteine Residues by Alkylation. A Tool in Peptide Mapping and Protein Identification. *Analytical Biochemistry* **1998**;70:5150-5158

29. Mineki R, Taka H, Fujimura T, Shindo N, Murayama K. *In situ* Alkylation with Acrylamide for Identification of Cysteinyl Residues in Proteins During One- and Two-Dimensional Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis. *Proteomics* **2002**;2:1672-1681

30. Plowman JE, Flanagan LM, Paton LN, Fitzgerald AC, Joyce NI, Bryson WG. The Effect of Oxidation or Alkylation on the Separation of Wool Keratin Proteins by Two-Dimensional Gel Electrophoresis. *Proteomics* **2003**;3:942-950

31. Maclaren JA, Milligan B. *Wool Science: The Chemical Reactivity of the Wool Fibre*. Marrickville, Australia: Science Press, **1981**

32. Rogers GE. Genes for Hair and Avian Keratins. *Annals New York Academy of Sciences* **1985**;455:403-425

33. Smith TA, Strelkov SV, Burkhard P, Aebi U, Parry DAD. Sequence Comparisons of Intermediate Filament Chains: Evidence of a Unique Functional/Structural Role for Coiled-Coil Segment 1A and Linker L1. *Journal of Structural Biology* **2002**;137:128-145

34. Parry DAD, Marekov LN, Steinert PM, Smith TA. A Role for the 1A and L1 Rod Domain Segments in Head Domain Organisation and Function of Intermediate Filaments: Structural Analysis of Trichocyte Keratin. *Journal of Structural Biology* **2002**;137:97-108

35. Herbert BR, Molloy MO, Yan JX, Gooley AA, Bryson WG, Williams KL. Characterisation of Wool Intermediate Filament Proteins Separated by Micropreparative Two-Dimensional Electrophoresis. *Electrophoresis* **1997**;18:568-572

36. Herbert BR, Chapman ALP, Rankin DA. Investigation of Wool Protein Heterogeneity Using Two-Dimensional Electrophoresis with Immobilised pH Gradients. *Electrophoresis* **1996**;17:239-243

Chapter Six

Wool Intermediate Filament Assembly

The goals of the previous chapters were to gain an efficient high-resolution separation system and to determine the cause of the charge heterogeneity of the IFPs, as a prerequisite for understanding the *in vitro* assembly of wool IFPs. One of the goals of this chapter was to develop a method for the *in vitro* assembly of IFs from wool. Currently, there is no standard method to reassemble IFs that are extracted from a keratinised source (i.e. not from bacterial expression) [1] and thus direct extraction from wool was employed. Another goal was to develop fractionation methods so that once optimal assembly conditions were found, individual proteins could be used in assembly experiments.

6.1 Introduction

One of the important defining characteristics of IFs is their ability to form filaments [2]. To appreciate IFP function, it is necessary to understand how they polymerise [3]. Studies into how IFs assemble *in vitro* give significant information about how they assemble *in vivo*. *In vitro* studies have been able to give information about the stability of protein pairings. Hatzfeld and Franke [4] have shown that it is possible to reassemble IFs from combinations of different cytokeratins and even from different species, such as human, cattle and rat. However, there was a difference in stability between different pairings. Using urea 'melting' experiments, a specific melting curve could be attained for each pair. The melting curves for *in vitro* and *in vivo* samples were almost identical. Other studies have also shown that some mixtures of polypeptides polymerise into filaments more efficiently than others [5].

Wool protein is an ideal starting material for biomaterial manufacture, as it exhibits several characteristics necessary for biomaterials. It is natural, can be obtained with no harm to the animal, has been shown not to elicit any adverse biological response, and is very durable due to the high number of disulfide bonds that can be induced to form between the proteins. NZ is the second highest producer of wool worldwide; consequently, it is plentiful and cheap when compared to other biomaterial sources [6]. Keratec Ltd currently uses NZ wool to produce biopolymer materials for a wide range of uses [7]. The controlled assembly of wool IFPs may produce biomaterials with superior properties that could be manipulated depending on the required usage.

6.2 Previous work on hard α-keratin assembly

There has been very little work done on the assembly of hard α-keratins. An early paper by Thomas *et al.* [2] attempted to reconstitute wool IFs. Sulfitolysis was used to extract wool proteins in the S-sulfo form before using zinc acetate to fractionate out the IFPs. Dialysis at low temperatures and Tris-HCl concentrations, along with reducing urea concentrations in the dialysis buffers gave material that appeared filamentous. However, the filaments did not have the same long smooth walled filamentous morphology shown in other soft keratin IF assembly papers [4, 8, 9]. Subsequent researchers have been unable to repeat this work [1]. These results demonstrate that *in vitro* reconstitution of wool IFs cannot be achieved as easily as epidermal IF reconstitution [2].

It has been shown recently that recombinant human hair keratins were unable to form IFs using Tris-HCl concentrations between 5 and 50 mM [10]. However, when vimentin (Type III IF) assembly conditions were employed (NaCl 100-160 mM) 9-10 nm wide and 500-1 000 nm long filaments were formed [10].

Wang *et al.* [1] found that the bacterially expressed hair IFs optimal assembly conditions were pH 7.5-8 and ionic strengths between 150 and 200 mM.

Intermediate filaments that were extracted from adult mouse hair would not form IFs under the optimal assembly conditions used for expressed IFs [1]. This suggests that there is some difference in the bacterially expressed proteins and the native proteins. Optimal assembly conditions for mature wool hard α-keratins need to be studied so the results can be used to assemble high value products from low value wool.

Assembly of the bacterially expressed hair IFs using 5% DTT as the reducing agent gave yields between 30-50% and short IF particles. Yields above 90% were obtained when TCEP was used as the reducing agent and buffers were degassed and nitrogen restored. Under physiological conditions, only the reducing agents TCEP and DTT allowed efficient assembly of bacterially expressed hair follicle keratins [11]. The reducing agents may be acting to prevent oxidative cross-linking during early stages of assembly. In hard α-keratins, high concentrations of thiol agents (much higher than those necessary for epidermal assembly) are required to minimise oxidative cross-linking [12]. The maintenance of a reducing environment during assembly of hard α-keratins is therefore essential [1].

The importance of reducing conditions has been shown in other IF systems. Bovine epidermal keratin filaments would only produce long filaments when the thiol groups of the proteins were kept in a fully reduced form [9]. Interchain disulfide bonds could inhibit assembly by causing steric hindrance [9]. At later stages of assembly, disulfide bonds become structurally important. In the wool follicle, the sulfhydryl groups are converted to disulfide bonds by oxidative enzymes during keratinisation [13]. The cysteines in the rod domains of the IFPs help stabilise the molecules in the oxidised form [14].

Assembly experiments often involve conditions very different from those occurring in the cell [15]. Mutagenesis studies on vimentin showed that some mutants were able to form filaments *in vivo* but were unable to form filaments *in vitro* [16]. Hard α-keratins assembled *in vivo* have tapered ends,

whereas *in vitro* assembled IFs have blunt ends. This may indicate differences in the assembly pathways [17]. Refolding of proteins *in vitro* may be initiated by the collapse of hydrophobic regions, the formation of secondary structures and the formation of covalent bonds e.g. disulfide bonds to add stability [15].

Shoeman *et al.* [16] suggested that cellular proteins may interact with IFs *in vivo* enabling polymerisation of mutant IFs. Chaperones may be used to help proteins assemble *in vivo*. Misfolding and aggregation is a common problem when assembling proteins *in vitro* [15]. Intermediate filaments do not require chaperones *in vitro* to assemble; however, this does not exclude the possibility that chaperones are used *in vivo* to make assembly more efficient [18]. Protein disulfide isomerase (PDI) is a chaperone which may be involved in IF assembly *in vivo*. It facilitates the correct formation of disulfide bonds by reshuffling incorrect disulfide pairings [15]. Sulfhydryl oxidases also contribute to correct disulfide bonding [1, 13]. Other important *in vivo* factors may include the contribution from the intermediate filament associated proteins (high sulfur proteins and the high glycine tyrosine proteins, refer to Figure 2.1) to give correct alignment of IFs [1]. IFs isolated from follicles have been shown to be decorated with associated proteins [17]. The IFAPs may influence the stability and assembly of hard α-keratins [11].

Assembly of IFs *in vitro* is critically dependent on the conditions used, e.g. dialysis versus dilution, ionic strength and pH [19]. In desmin, increased temperature and time produced longer filaments [20]. As dialysis time increased, the diameter of the filaments decreased, showing that more highly ordered packing of subunits was occurring. The use of non-physiological cations caused an increase in the IF diameter.

The effect that cations have on IF assembly is variable. Polymerisation of bovine epidermal keratin filaments has been shown to not be influenced by the presence of $CaCl_2$, NaCl, KCl or $MgCl_2$ [5]. However, divalent cations, at identical ionic strengths to univalent cations, are more effective promoters of desmin assembly [20]. Assembly of epidermal keratin fibres of newborn rat

is altered by the addition of divalent cations [21]. Wider filaments were formed after the addition of $ZnCl_2$, $CuSO_4$, and $HgCl_2$ when compared to filaments formed with $CaCl_2$ or $MgCl_2$. At 1 mM concentration, Zn^{2+}, Cu^{2+} or Hg^{2+} caused precipitation, whereas at a lower concentration (0.3 mM) they caused gelation [21]. Aebi *et al.* [22] found that Ca^{2+} above ~5 mM formed human epidermal filaments that looked like folded ribbons, and that polylysine above ~0.01 mM gave compact highly aligned bundles of filaments. Chelating agents are often added to initial assembly buffers to remove divalent cations [1].

It is possible that the divalent cations form cross-links between keratin molecules and alter solubility [21]. Nelson and Traub [23] suggested from their studies on vimentin that di- and polyvalent cations may stabilise IFs and may help inter-connect them into parallel arrays or bundles. Gu *et al.* [24] believe that divalent cations may stabilise protein-protein interactions. Others believe that cations may perform a regulatory role [21]. IFs are often found close to the endoplasmic reticulum (ER) where high levels of calcium are stored. It is possible that high levels of calcium could regulate IF assembly [24].

Other conditions likely to effect IF assembly include pH. Glial fibrillary acidic protein (GFAP), a type III IFP, is more likely to aggregate at lowered pH and vimentin, another type III IFP, will aggregate at higher pH [25]. At high pH, the very basic head domains of IFs may form an unproductive conformation that may be reactivated by high salt concentrations [26].

6.3 Methods to monitor IFP assembly

Conditions less favourable for IFP assembly, such as low ion concentrations and elevated pH, can be used to study reconstitution steps [9]. Transient electric birefringence has been used to follow assembly and calculate particle length values [27].

One of the techniques used to study IF assembly is transmission electron microscopy (TEM). Assembled filaments are negatively stained on hydrophilic grids. A good indicator of complete IF assembly is the appearance of smooth IFs observed using TEM [24]. IF assembly occurs in two steps, the initial formation of ULFs from tetramers/octamers, then annealing and compaction of the ULFs into long IFs [28]. Before annealing and compaction the IFs may have a rough appearance. IFs will have a diameter ranging from 7 to 15 nm [19, 29].

6.4 Results & Discussion

6.4.1 IFP/IFAP fractionation

One of the initial aims of the assembly studies was to fractionate the IFPs from the IFAPs, and to attempt to isolate individual IFP family members. Assembly of specific protein pairings may give filamentous biomaterials with differing physical properties [30].

The first chemical fractionation method used was based on a fractionation method developed by Thomas *et al.* [31]. Using acetic acid fractionation, the IFPs are theoretically separated from the IFAPs in a simple precipitation step. Using this technique, the IFAPs appear to be depleted but are not completely removed (Figure 6.1).

In 1998, Kon *et al.* [32] developed a method to fractionate hair components to study the degree of damage due to perming. Using different concentrations of reductant, they were able to separate and quantify the IFPs and IFAPs. When this method was used here for fractionating wool proteins, there was little difference between the two fractions (Figure 6.2). The initial IFAP fraction (Fraction 1) contained large amounts of IFPs. The IFP fraction (Fraction 2) did appear to be slightly enriched in IFPs but still contained IFAPs.

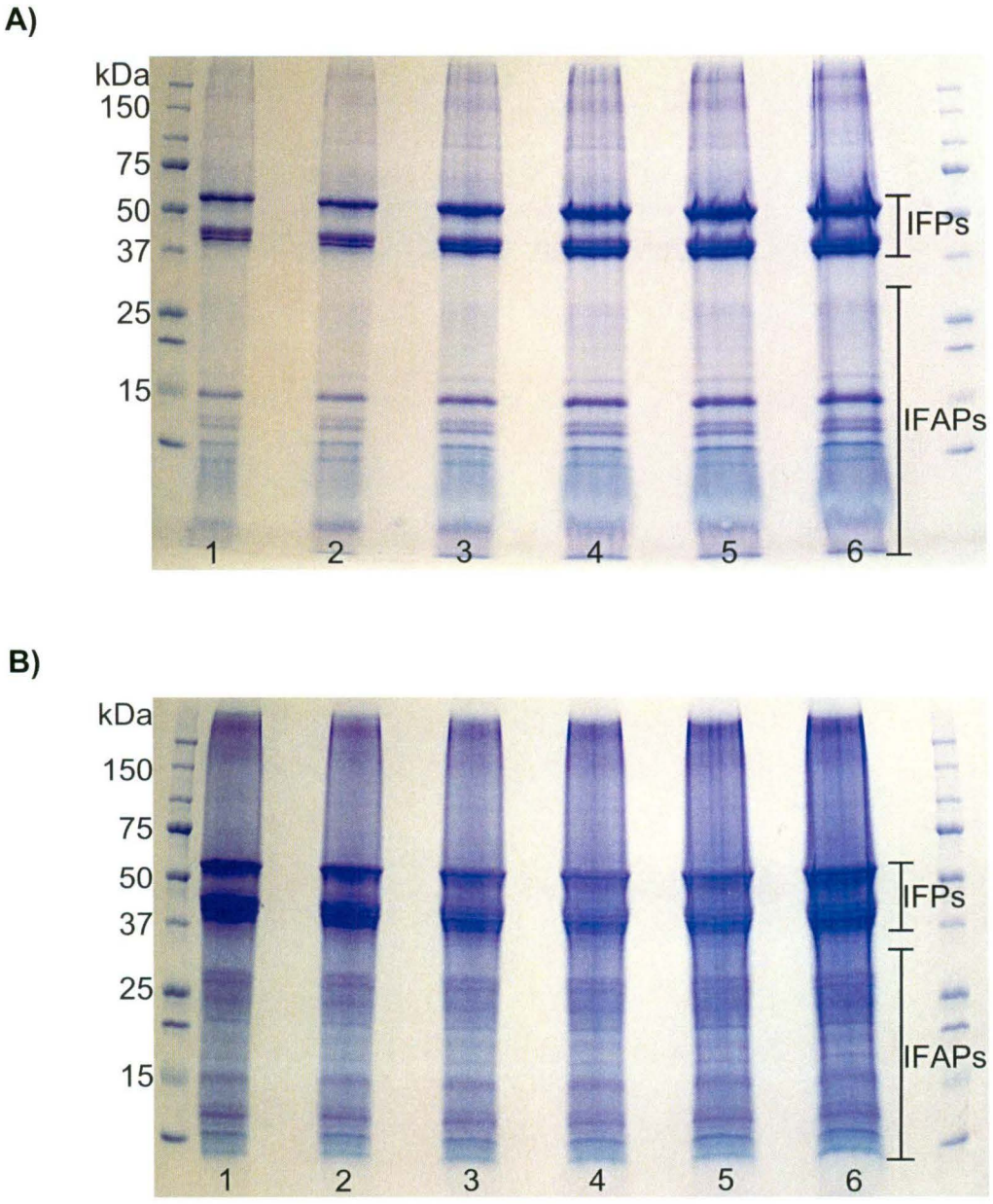


Figure 6.1
A) 1DE gel showing fractionation of IFPs from IFAPs using an acid precipitation. B) 1DE gel of wool proteins prior to fractionation. Lane loadings are; 1-33 µg, 2-47 µg, 3-60 µg, 4-73 µg, 5-87 µg, 6-100 µg. The gels were stained with Coomassie brilliant blue R-250. The gels are representative of duplicate experiments.

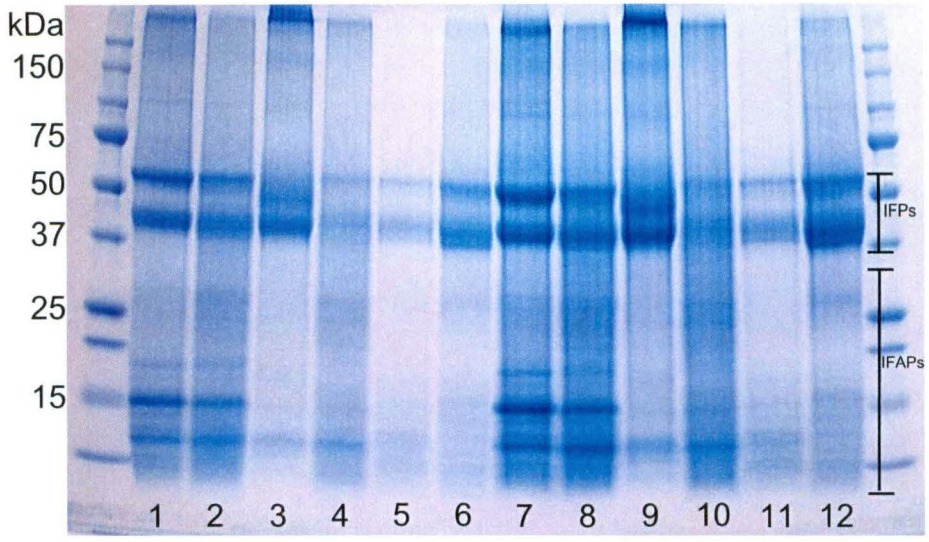


Figure 6.2
1DE gel showing fractionation of wool proteins using the method of Kon *et al.* [32] method. Lanes 1, 2, 7 and 8 show whole wool extractions. Lanes 5 and 11 show fraction 1 (IFAP extraction) and lanes 6 and 12 show fraction 2 (IFP extraction). Lanes 3, 4, 9 and 10 show human hair extractions. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of duplicate experiments.

The final chemical extraction method used here was a sulfitolysis extraction followed by an isoelectric precipitation with zinc acetate. This method was used by Thomas *et al.* [2] to separate the wool proteins into classes prior to their assembly studies. Using this method, the majority of the IFPs were separated from the IFAPs (Figure 6.3). There was a small amount of contamination from the high sulfur proteins (HSPs) (around 25 kDa), and no stained high glycine tyrosine proteins (HGTPs) (around 15 kDa).

Thus fractionation of the wool proteins into IFPs and IFAPs was achieved by the sulfitolysis extraction/fractionation method. This allowed assembly studies to be performed on just the IFPs.

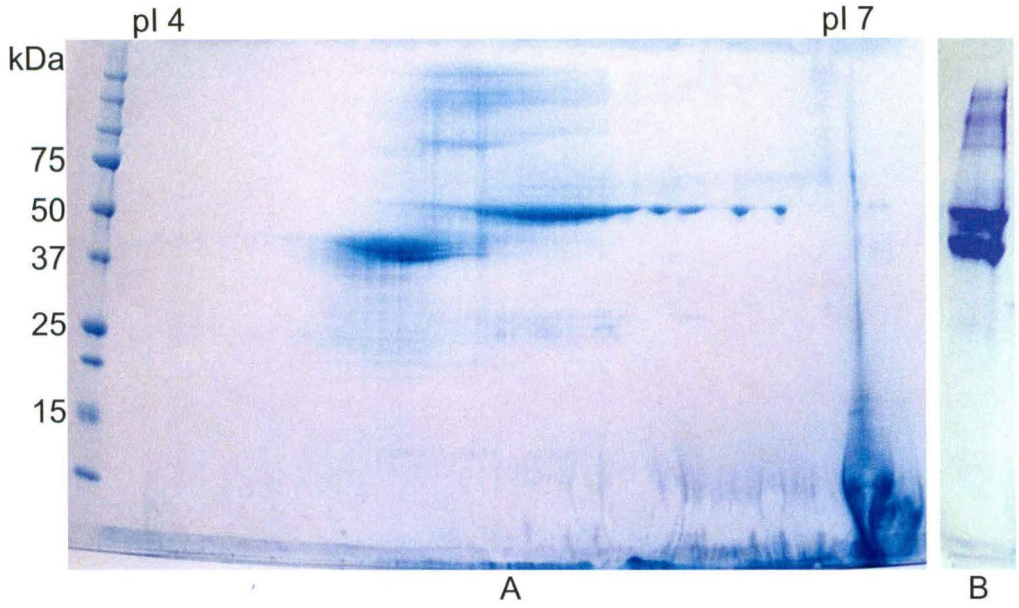


Figure 6.3
A) 2DE separation of wool IFPs after sulfitolysis extraction/fractionation. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10 - 20% T gel. The gel was stained with Coomassie brilliant blue G-250. B) 1DE separation of wool IFPs after sulfitolysis extraction/fractionation. The lane was loaded was 47 µg of protein. The gel was stained with Coomassie brilliant blue R-250. The gels are representative of triplicate experiments.

6.4.2 Type I/type II IFP fractionation

Once IFP fractionation had been achieved, methods to isolate the type I IFPs from the type II IFPs were developed. The fractionation method to separate the type I IFPs from the type II IFPs used large format, thick 1DE gels, so a large amount of protein could be separated. Approximately 60 mg of protein, containing both IFPs and IFAPs, was loaded into the 1DE well. Approximately 18% of this protein was extracted out of the gel as IFPs. The staining method is very important when elution of the proteins off the gels after identification is required. Ideally, the proteins are visualised without subjecting the gel to long periods in acidic alcohol solutions. Standard staining methods involve fixing the gel in acidic alcohol solutions for long periods to make the proteins insoluble and therefore unable to migrate out of the gel. In this fractionation method, the gels were stained with either a reversible negative stain or a fast Coomassie R-250 stain [33-35]. In the reversible negative stain method, the area around the proteins is stained

leaving the protein areas clear. This means that the proteins are never fixed into the gel matrix [33, 34]. In the second method, the gels are only briefly exposed to the acidic alcohol solution and the gels are only stained until bands become visible. 1DE gels of the proteins after they had been eluted off the large format 1DE gel showed that the fractionation of the proteins into the type I IFPs or the type II IFPs had worked well (Figure 6.4).

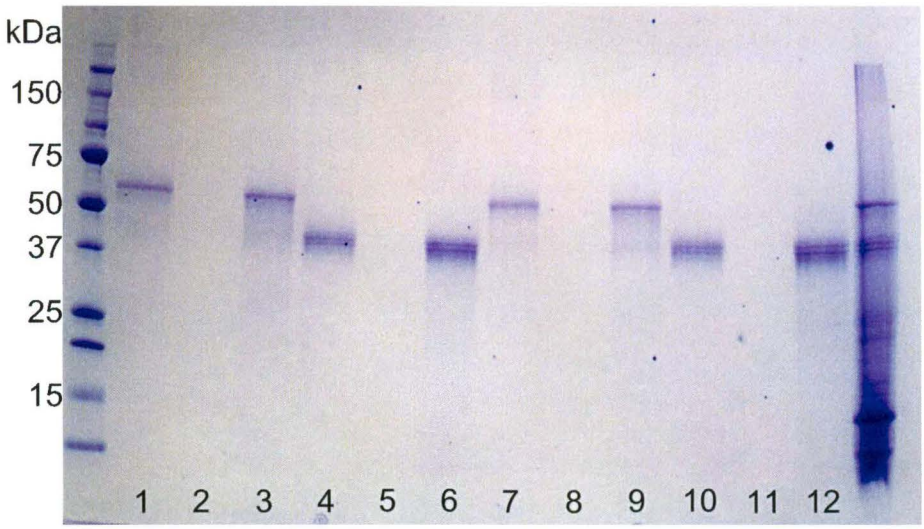


Figure 6.4
1DE gel showing fractionation of type I and type II wool IFPs using a preparative 1DE gel. Lanes 1 and 7 show type I IFPs and lanes 4 and 10 show type II IFPs that have been extracted out of the gel after Coomassie R-250 staining. Lanes 2, 5, 8 and 11 were loaded with material extracted out of a gel after Coomassie brilliant blue G-250 staining. Lanes 3 and 9 show type I IFPs and lanes 6 and 12 show type II IFPs that have been extracted out of the gel after reversible negative zinc stain. The gel was stained with Coomassie brilliant blue G-250. The gel is representative of duplicate experiments.

Proteins that were stained with Coomassie brilliant blue G-250 were not successfully extracted out of the gel (Figure 6.4, lanes 2, 5, 8 and 11). This staining method involves an overnight fixing step in acid and alcohol followed by several days staining in acid and alcohol. After such harsh treatment, the proteins were presumed to be insoluble and unable to migrate out of the polyacrylamide matrix. This demonstrates the importance of using gentle

staining methods if extraction of proteins out of the gel is required after staining.

Since 1DE had successfully separated the type I and type II IFPs, isoelectric fractionation using the ZOOM® IEF Fractionator was investigated to determine whether individual spots could be isolated.

The initial runs used in-house membrane discs, as commercially available discs in the range required were not available. Protein was focused in solution, using the ZOOM® IEF Fractionator. The proteins from each pI fractionation chamber were then isoelectrically focused using IPG strips to determine the accuracy of the ZOOM® IEF Fractionator was. The ZOOM® IEF Fractionator gave good separation in the more basic zones but gave poorer separation in the more acidic zones (Figure 6.5). The type I IFPs separate out between approximately pI 4.75 and 5.05, and the type II IFPs separate out between approximately pI 5.3 and 6.7. The type I IFPs have been separated from the type II IFPs (Figure 6.5, strip E). The type II IFPs have been separated into two fractions at the basic end (Figure 6.5, strips A and B) but the acidic end of the type II IFPs has not been fractionated well (Figure 6.5, strip C). The poor resolution could be due to using in-house membrane discs. The discs were very difficult to fit into the end of each chamber as they had to be cut to the correct shape with a scalpel. It is possible that some discs were not sitting correctly in the chambers and therefore weren't fractionating accurately. These problems may be overcome by increased use of the equipment to gain familiarity and possibly by casting the membrane discs onto thinner, harder material and cutting them out with a corer made to the correct shape.

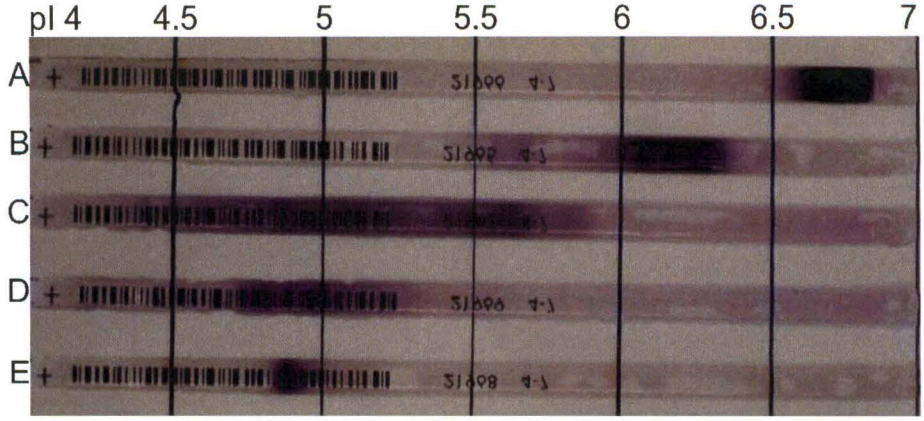


Figure 6.5
Isoelectric fractionation of wool proteins using the ZOOM® IEF Fractionator. Membranes used were: A) 6.5-7, B) 6-6.5, C) 5.5-6, D) 5-5.5, E) 4.5-5. The IPG strips were stained with acid violet. The IPG strips are representative of duplicate experiments.

To see if even smaller pI fractions could be separated a pI 6.1 membrane was made. The pI separation between individual spots of the type I IFPs is approximately 0.05 pI units. The pI separation between individual spots of the type II IFPs is approximately 0.15 pI units at the basic end and 0.06 pI units at the acidic end. The fraction, which had previously been separated into pI 6-6.5, was loaded between the membranes. When this was run, a very small fraction of protein (approximately 150 µg) could be fractionated and focused on an IPG strip (Figure 6.6). This shows that in solution IEF is capable of fractionating protein down to 0.1 pI units, which would easily separate individual type II IFPs at the basic end of the spot row. Even smaller pI separations may be possible.

The ZOOM® IEF Fractionator showed a varied ability to isoelectrically fractionate wool proteins. In the more basic pI range, excellent resolution was attained. However, fractionation resolution was poor at a more acidic pI. Future work with the fractionator may be able to resolve the more acidic fractions, allowing more precise assembly studies of individual subunits.

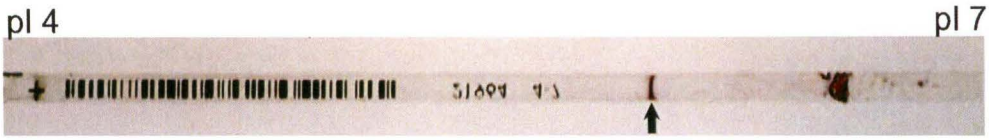


Figure 6.6
Isoelectric fractionation of wool proteins using the ZOOM® IEF Fractionator. Membranes used were pI 6 and 6.1. The arrow shows where the proteins fractionated in the pI 6-6.1 region focused on a pI 4-7 IPG strip. The IPG strip was stained with acid violet. The IPG strips are representative of duplicate experiments.

6.4.3 Assembly

Once the fractionation studies had been completed, attempts were made to assemble the IFPs using fractionated protein or protein which hadn't been subjected to fractionation (contained IFPs and IFAPs). All assembly studies used milligram amounts of starting material and were completed in duplicate, with at least 5 grid squares being observed using TEM. Initial assembly experiments used methods which had previously been used for both wool and hair proteins. The first method was based on the method that Thomas *et al.* [2] had used for wool IFPs. Briefly, wool proteins were fractionated by sulfitolysis prior to being dialysed into a series of buffers with reducing urea concentrations. No signs of assembly were seen when samples were looked at using TEM. The second method was based on the method of Wang *et al.* [1], which used bacterially expressed IFPs and a series of dialysis steps with solutions that had high concentrations of reducing agents and decreasing urea concentrations. Two different starting materials were used, one used type I and II IFPs that had been fractionated and eluted off a gel and the other used unfractionated protein which contained IFAPs. No signs of assembly were visible under TEM when the gel-eluted proteins were used. However, the unfractionated protein showed some small structures under TEM that were possibly protofilamentous (Figure 6.7).

From these results it was determined that the best chance of getting assembled IFPs was to start with protein which contained IFPs and IFAPs. A

systematic series of experiments was done to test variables that may have an effect on IFP assembly, based on previous literature (Table 6.1).

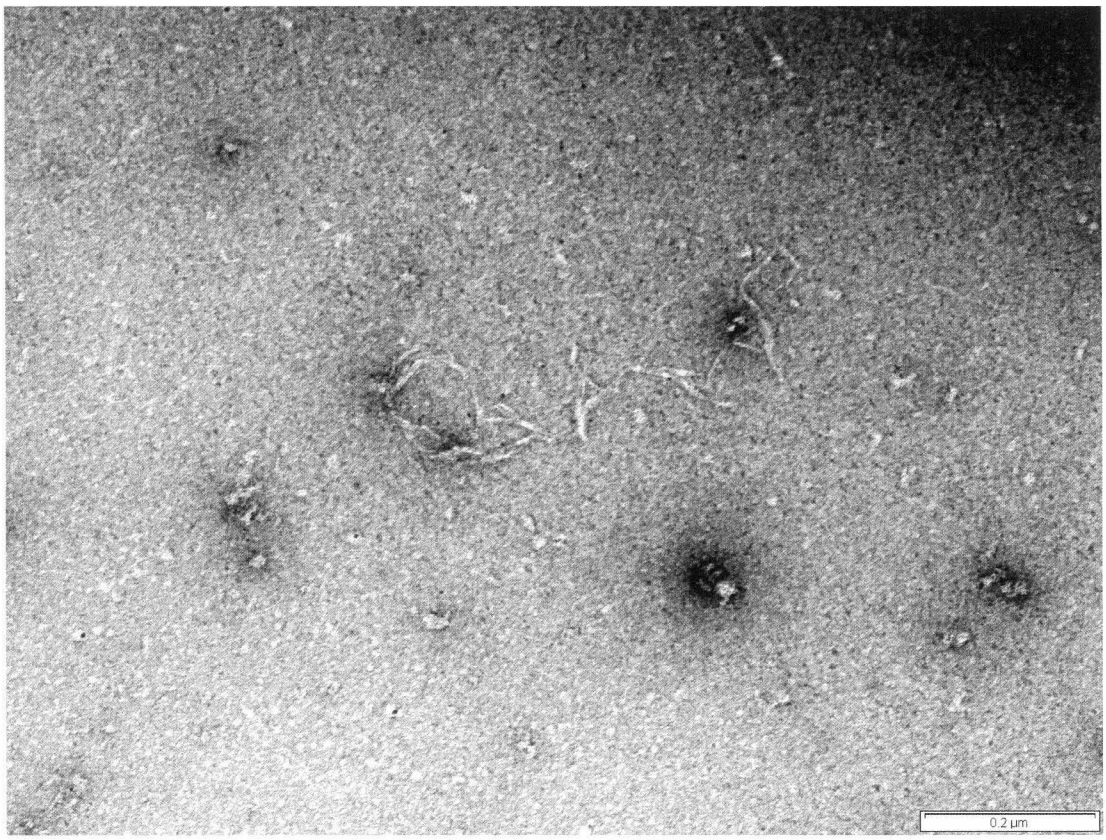


Figure 6.7
A possible protofilamentous structure formed when unfractionated proteins were subjected to the Wang [1] assembly conditions.

Variables tested were reductant concentration, pH, time, chaotrope, temperature, cations, alkylation, oxidation and a chaperone enzyme (PDI). With all variables examined, no signs of assembly were seen using TEM. Absorbance measurements have been used in the past to show IF assembly; the more the filaments assemble the higher the absorbance readings are [5, 20]. The A_{340} measurements showed variations in the absorbance of the final protein solutions (Figure 6.8), which could mean that under optimal assembly conditions the variables chosen may have an effect on assembly. Reduced absorbance measurements suggest that there were more monomer proteins in solution and less higher ordered structures. TEM did not show

any signs of assembly, therefore only small higher ordered structures may have been formed (e.g. dimers or tetramers) which were not visible under TEM.

Reductant	DTT was used as the reductant, instead of TCEP, at concentrations of 10 mM, 25 mM or 50 mM.
pH	Assembly was performed at pH 4, 5, 6, 7, 9 and pH 10.
Time	Solutions were changed once a day, every second day or once a week.
Guanidine-HCl	Guanidine-HCl was substituted for urea in all solutions.
Temperature	Assembly was performed at either low temperature (approximately 7°C) or high temperature (approximately 35°C).
Cations (Ca^{2+}, Mg^{2+}, Cu^{2+}, Ni^{2+})	5 mM of each cation was added to each assembly solution. No EDTA was added to the final solution.
Alkylation	Proteins were alkylated with 200 mM IAM prior to assembly.
Oxidised	No reductant was added to the final solution.
Protein disulfide isomerase (PDI)	Following overnight dialysis in the final solution, 8.5 units of PDI were added to each tube.

Table 6.1
Variations made to the assembly buffer conditions.

Three different starting materials were also tested to determine whether different extraction methods had an effect on the *in vitro* assembly characteristics. The three other starting materials were: Extraction Method Two unfractionated protein, an acid fractionated sample and a sulfitolysis fractionated sample. The unfractionated protein contained both IFPs and IFAPs in the usual concentrations that they are extracted in. The acid fractionated sample had slightly less IFAPs than an unfractionated sample (Figure 6.1) and the sulfitoylsis fractionated sample had very little IFAPs

(Figure 6.3). None of these starting materials gave any filaments when they were subjected to the Wang *et al.* [1] assembly conditions.

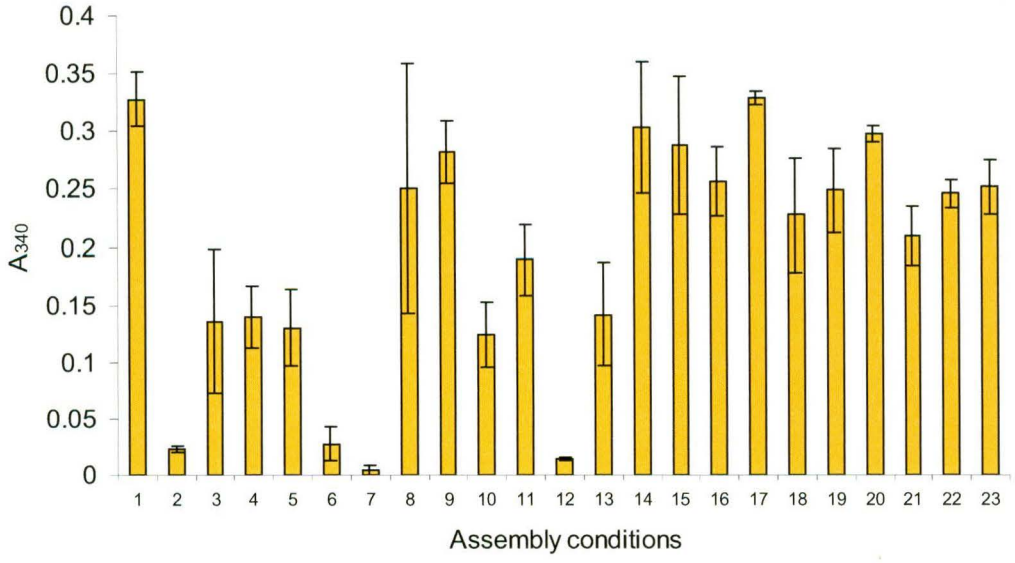


Figure 6.8
A_{340} measurements of the protein solutions after the final dialysis step was concluded. See Table 6.1 for details of the conditions used. Lane numbers represent; 1-standard Wang method, 2-7°C, 3-35°C, 4-IAM alkylated, 5-1 day dialysis steps, 6-2 day dialysis steps, 7-1 week dialysis steps, 8-Ca^{2+}, 9-Mg^{2+}, 10-Cu^{2+}, 11-Ni^{2+}, 12-Guanidine, 13-pH 4, 14-pH 5, 15-pH 6, 16-pH 7, 17-pH 9, 18-pH 10, 19-oxidised, 20-10 mM DTT, 21-25 mM DTT, 22-50 mM DTT, 23-PDI. The results are of triplicate experiments. The error bars represent the standard deviation.

In the final assembly experiment, a chymotrypsin digestion was done prior to assembly being attempted. In a past paper on wool IFs, Woods and Gruen [36] found that when the IFs were digested with chymotrypsin they produced a pair of double-stranded α-helices. The starting material for the digestion in this experiment was Extraction Method Two protein; either containing IFPs and IFAPs, or protein which had been acid fractionated.

Directly after the partial chymotrypsin digestion, a sample was taken and looked at under the TEM. The unfractionated protein showed no filamentous formations. However, the acid fractionated sample showed numerous short filamentous assemblies (Figure 6.9).

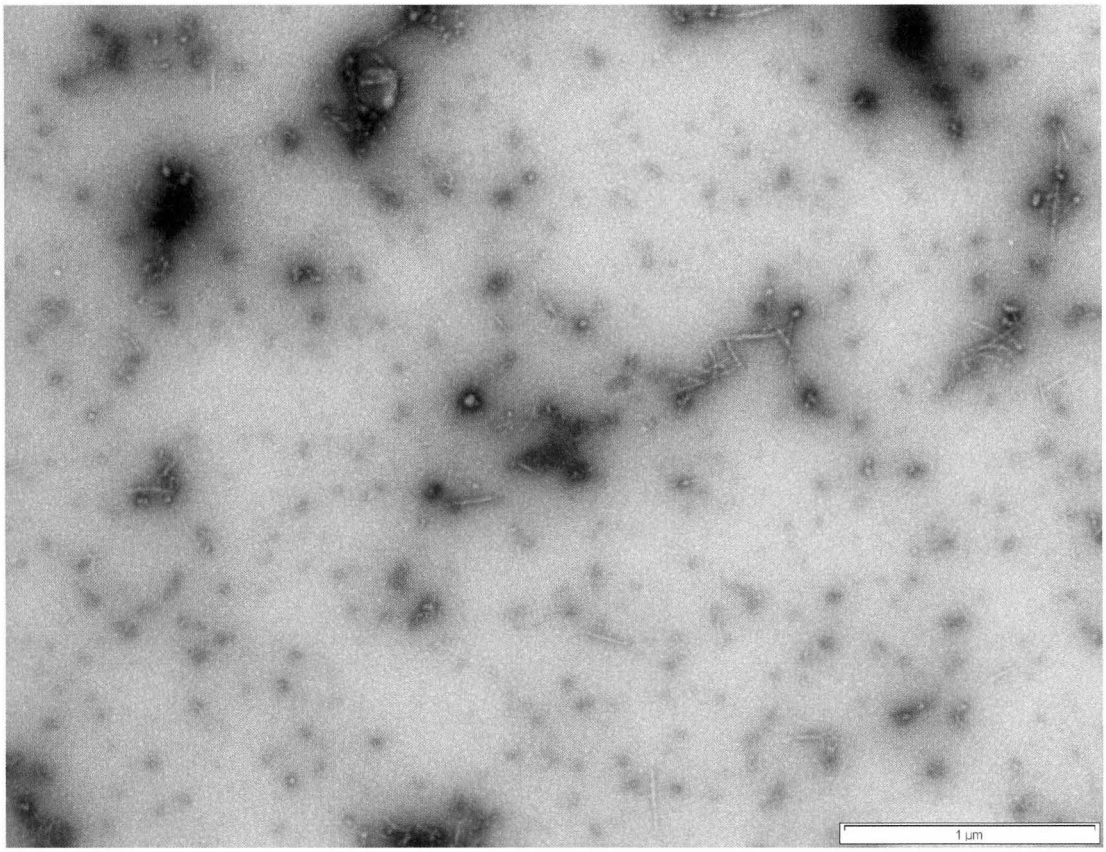


Figure 6.9
Filaments were formed from partially chymotrypsin digested acid fractionated protein.

The small filaments diameter was approximately 10 nm, with larger filaments with a diameter of approximately 40 nm also being present. This suggests that the basic unit was able to form higher ordered structures via further lateral assembly. Larger assembled structures show the filamentous nature of the assemblies (Figure 6.10 and 6.11). Some areas appear to be wrapping around others and some appear to be unravelled at their ends.

A control grid with starting protein material was examined to determine whether the results were artifactual. The control grid showed no filamentous material (Figure 6.12).

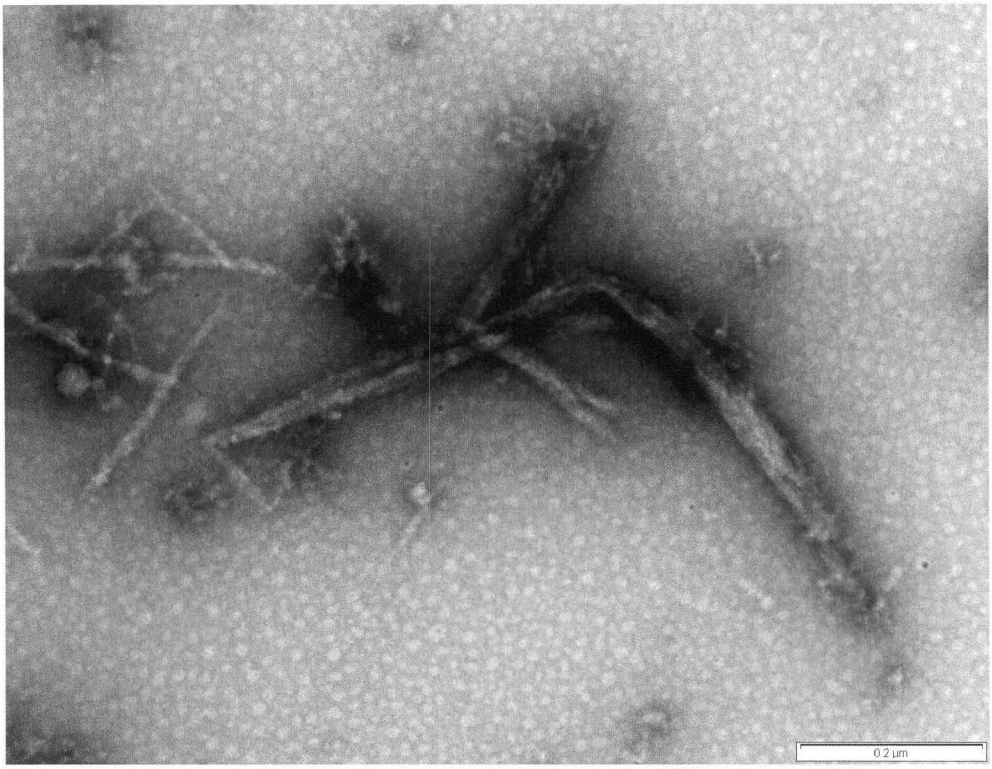


Figure 6.10
Filaments were formed from partially chymotrypsin digested acid fractionated protein.

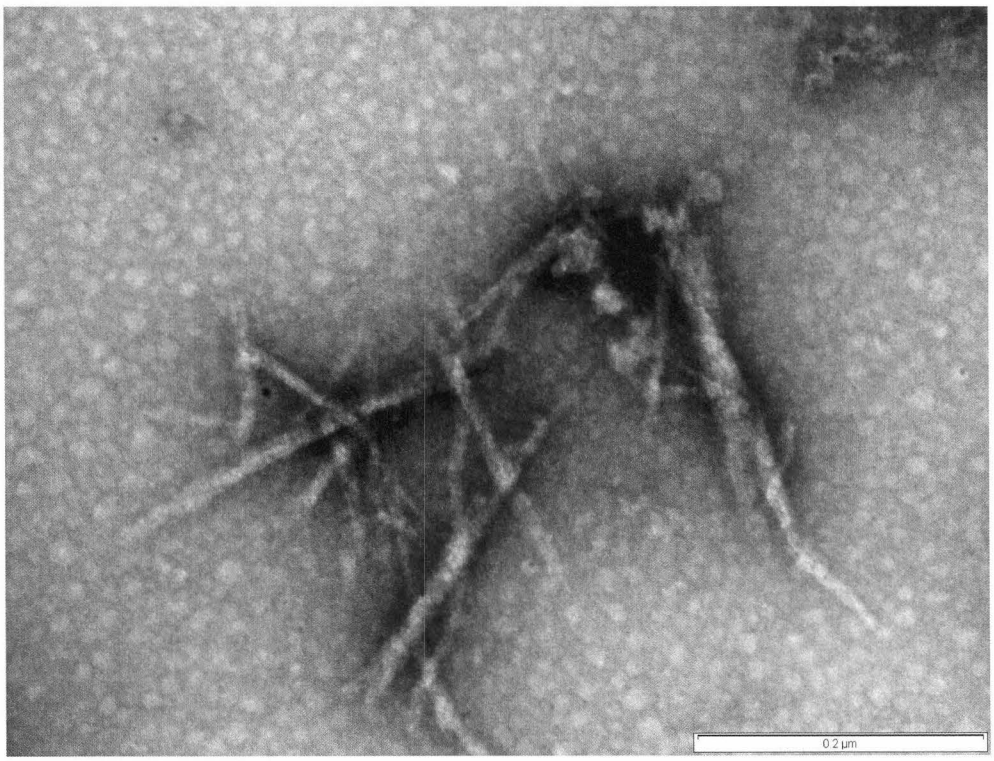


Figure 6.11
Filaments were formed from partially chymotrypsin digested acid fractionated protein.

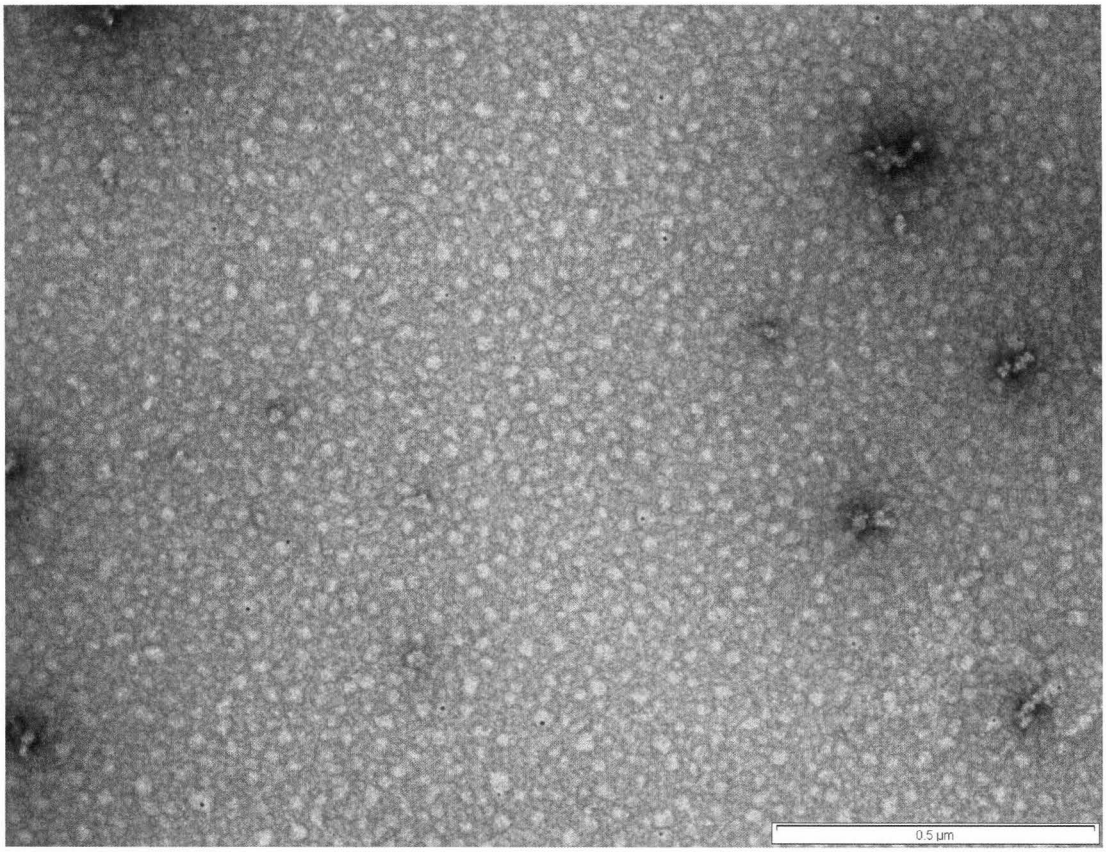
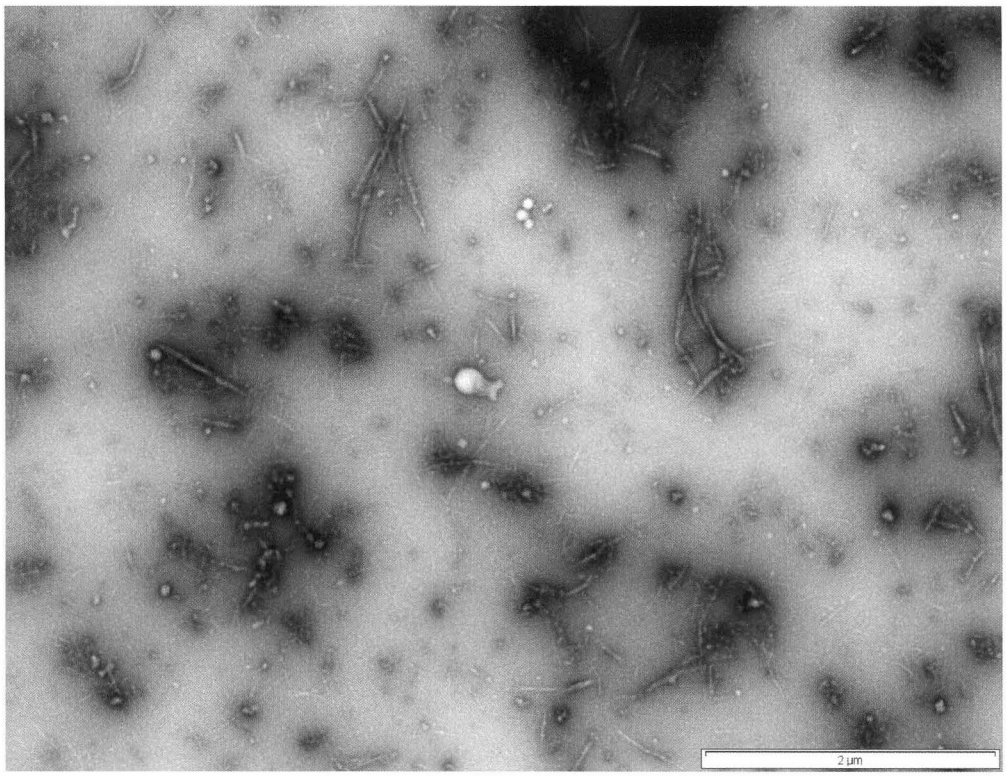


Figure 6.12
Control grid for filament formation. Protein which had not been partially chymotrypsin digested was viewed using TEM.

These filamentous assemblies were then dialysed using the Wang method [1]. The initial buffer for the Wang method contains urea, reductant, Tris and EDTA. This solution should disassemble any filaments that have been previously formed [37]. Inspection, after dialysis, with the TEM at low power (14 000 x), showed numerous long filamentous assemblies (Figure 6.13 A and B). Some of the filaments exceeded 1 µm in length. The longer dialysis appeared to produce many more long filaments (Figure 6.14). Filaments with a diameter of 10 nm were common. All the filaments showed a complex structure made up of smaller units appearing to associate. There were many filaments with a diameter of 40 nm, demonstrating the ability of four 10 nm filaments to form a higher ordered structure (Figure 6.15).

A)



B)

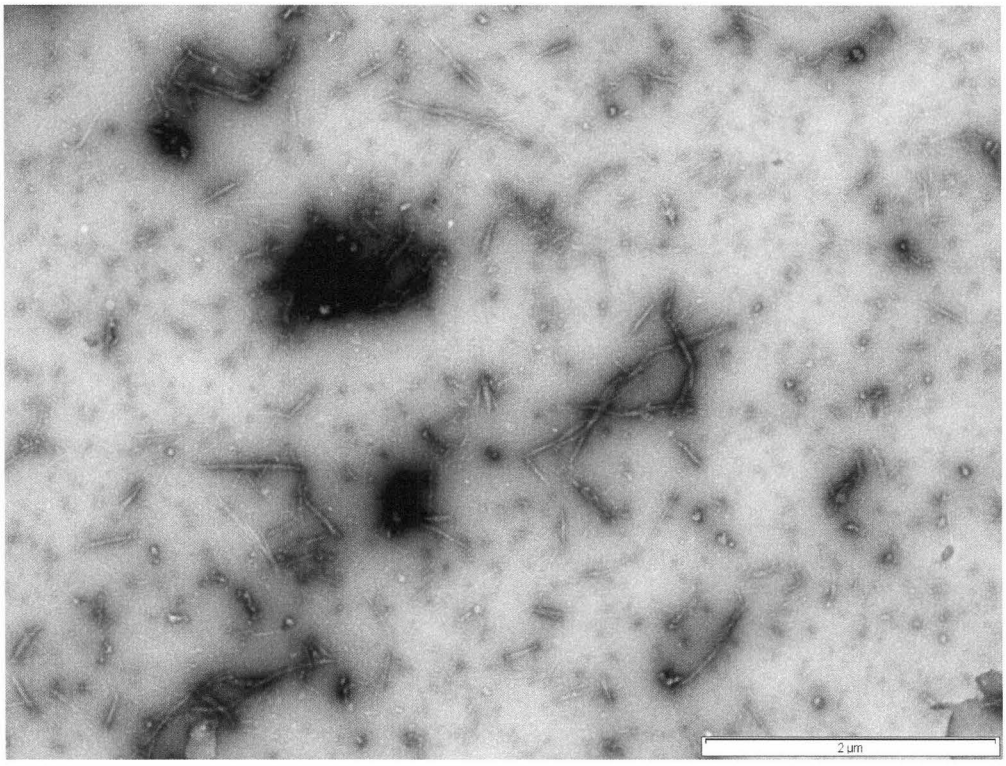


Figure 6.13
A) and B) Filamentous structures were formed when acid fractionated, partially chymotrypsin digested proteins were subjected to the Wang *et al.* [1] assembly conditions.

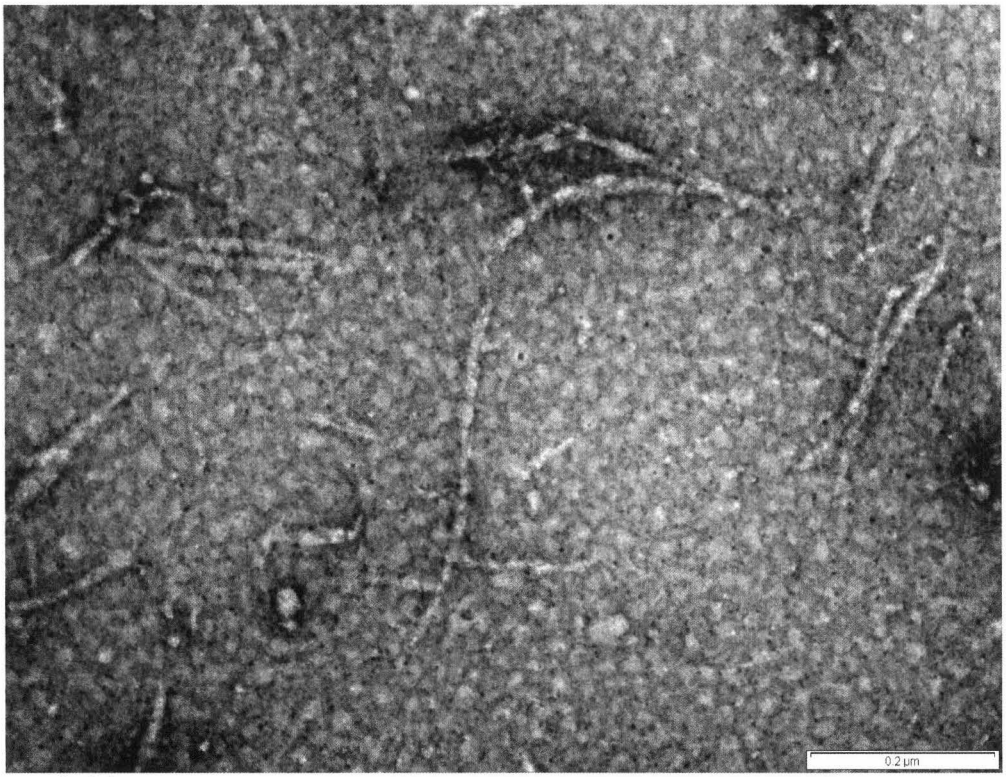


Figure 6.14
Long filaments were formed when acid fractionated, partially chymotrypsin digested proteins were subjected to the Wang *et al.* [1] assembly conditions.

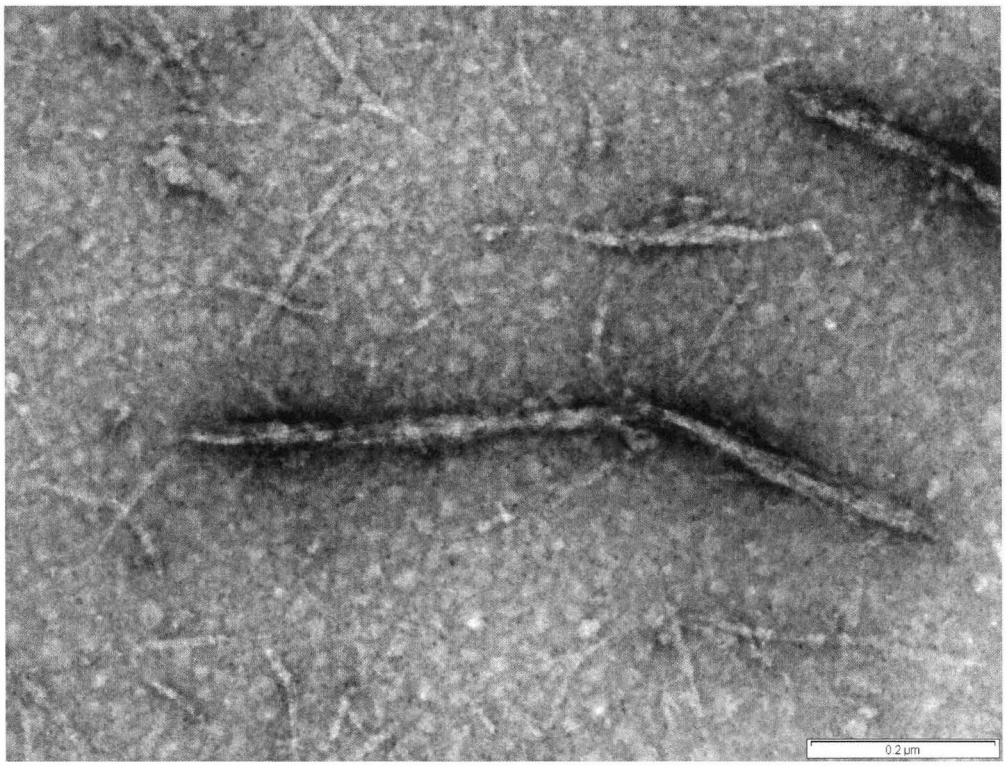


Figure 6.15
Wide filaments were formed when acid fractionated, partially chymotrypsin digested proteins were subjected to the Wang *et al.* [1] assembly conditions. Wide filaments had an approximate diameter of 40 nm.

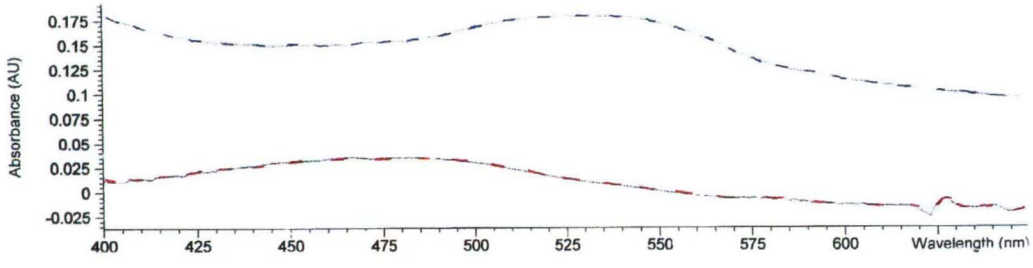
The starting material for the chymotrypsin digestion was monomeric protein (Figure 2.8). After partial digestion, the filaments are easily formed either by dilution into buffer or via dialysis. This suggests that the partial digestion has removed an area of the protein that was blocking assembly.

6.4.4 Characterisation of the assemblies

To determine whether the structures that had been formed had any β-sheet structure, two assays used for assessing amyloid formation were performed [38-40]. Using the Congo red assay, a shift in the spectrum to longer wavelengths, and a new point of maximal difference at 540 nm is classified as evidence for the presence of amyloid. Evidence for amyloid formation is seen with the thioflavin T assay as increase in the intensity. Both the assays showed that the starting material had more β-structure than the material after assembly (Figure 6.16). This suggests that the assembled structures have very little or no β-structure.

Gels that were run of the partially digested protein showed that the original rows of spots could be faintly seen (Figure 6.17). This suggests that some of the IFPs had not been digested, but that most of the proteins had been cleaved. The largest spot occurs at approximately 12 kDa and pI 4.5. The proteins appear to form a smear extending down from the type I IFP region to this protein spot. This protein spot may represent a portion of the rod domain of the IFPs.

A)



B)

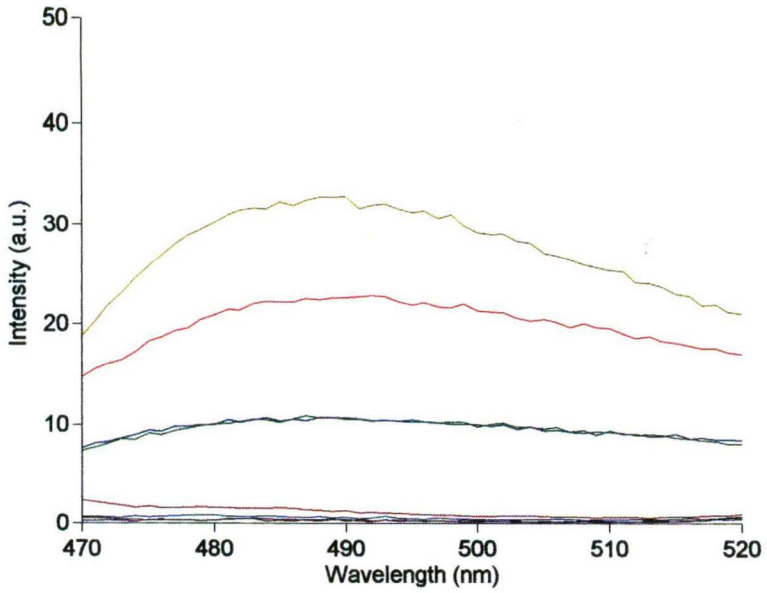


Figure 6.16
A) Spectrum from the Congo red assay. The blue line represents the protein starting material and the red line represents the protein after filament formation. B) Spectrum from the thioflavin T assay. The tan line represents insulin, a control protein with β-structure. The red line represents the protein starting material and the black line represents the protein after filament formation.

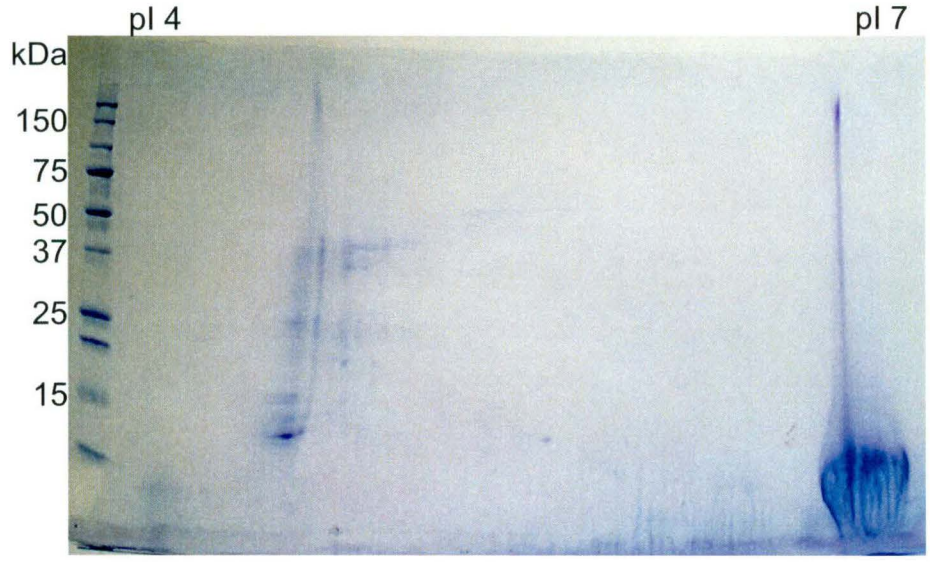


Figure 6.17
Partially chymotrypsin digested wool proteins (150 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10-20%T criterion gel. The gel was stained with Coomassie brilliant blue R-250. The gel is representative of duplicate experiments.

It is possible that the same pattern of bonding that were causing the isoforms investigated in Chapter Five, were causing the failure of IF assembly. The partial digestion may have led to a more linear form of the proteins that were assembly competent (Figure 6.18).

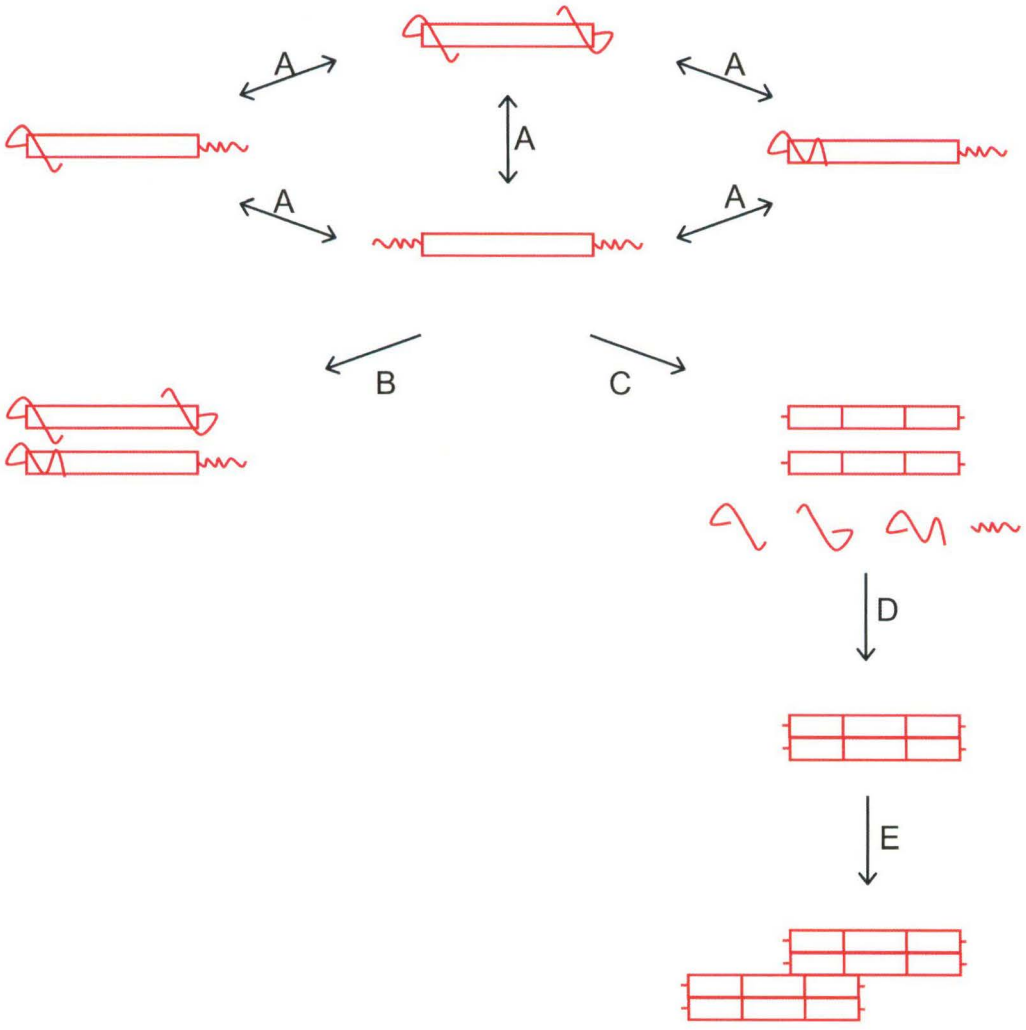


Figure 6.18
A) Conformational isomers of an IFP. B) Before digestion the terminal domains are in a conformation that blocks assembly. C) After partial digestion the IFPs are able to assemble into D) dimers, E) tetramers and higher ordered structures.

6.5 Summary

Fractionation methods to separate the IFP and IFAP have been successfully developed. Further fractionation into the type I and type II IFPs has been achieved along with partial success at isolating individual spots.

Assembly of filaments from IFPs was successful after a partial digestion. Filaments were formed that varied in diameter from 10 to 40 nm, showing that higher ordered structures were being formed.

This demonstrates that IFPs can be successfully assembled to form filamentous structures that will be able to be manipulated for biomaterial uses.

6.6 References

1. Wang H, Parry DAD, Jones LN, Idler WW, Marekov LN, Steinert PM. *In vitro* Assembly and Structure of Trichocyte Keratin Intermediate Filaments: A Novel Role for Stabilization by Disulfide Bonding. *Journal of Cell Biology* **2000**;151:1459-1468
2. Thomas H, Conrads A, Phan K-H, van de Löcht M, Zahn H. *In vitro* Reconstitution of Wool Intermediate Filaments. *International Journal of Biological Macromolecules* **1986**;8:258-264
3. Stuurman N, Sasse B, Fisher PA. Intermediate Filament Protein Polymerization: Molecular Analysis of *Drosophila* Nuclear Lamin Head-to-Tail Binding. *Journal of Structural Biology* **1996**;117:1-15
4. Hatzfeld M, Franke WW. Pair Formation and Promiscuity of Cytokeratins: Formation *in vitro* of Heterotypic Complexes and Intermediate-Sized Filaments by Homologous and Heterologous Recombinations of Purified Polypeptides. *Journal of Cell Biology* **1985**;101:1826-1841
5. Steinert PM, Idler WW, Zimmerman SB. Self-Assembly of Bovine Epidermal Keratin Filaments *in vitro*. *Journal of Molecular Biology* **1976**;108:547-567
6. http://www.woolequities.co.nz/keratec.htm.

7. www.Keratec.co.nz.
8. Steinert PM, Idler WW, Cabral F, Gottesman MM, Goldman RD. *In vitro* Assembly of Homopolymer and Copolymer Filaments from Intermediate Filament Subunits of Muscle and Fibroblastic Cells. *Proceedings of the National Academy of Science USA* **1981**;78:3692-3696
9. Franke WW, Schiller DL, Grund C. Protofilamentous and Annular Structures as Intermediates during Reconstitution of Cytokeratin Filaments *in vitro*. *Biology of the Cell* **1982**;46:257-268
10. Winter H, Hofmann I, Langbein L, Rogers MA, Schweizer J. A Splice Site Mutation in the Gene of the Human Type I Hair Keratin hHa1 Results in the Expression of a Tailless Keratin Isoform. *Journal of Biological Chemistry* **1997**;272:32345-32352
11. Hofmann I, Winter H, Mücke N, Langowski J, Schweizer J. The *in vitro* Assembly of Hair Follicle Keratins: Comparison of Cortex and Companion Layer Keratins. *Biological Chemistry* **2002**;383:1373-1381
12. Heid HW, Werner E, Franke WW. The Complement of Native Alpha-Keratin Polypeptides of Hair-Forming Cells: A Subset of Eight Polypeptides that Differ from Epithelial Cytokeratins. *Differentiation* **1986**;32:101-119
13. Orwin DFG. The Cytology and Cytochemistry of the Wool Follicle. *International Review of Cytology* **1979**;60:331-374
14. Steinert PM, Parry DAD. The Conserved H1 Domain of the Type II Keratin 1 Chain Plays an Essential Role in the Alignment of Nearest Neighbour Molecules in Mouse and Human Keratin 1/Keratin 10 Intermediate Filaments at the Two- to Four-molecule Level of Structure. *Journal of Biological Chemistry* **1993**;268:2878-2887
15. Gething M-J, Sambrook J. Protein Folding in the Cell. *Nature* **1992**;355:33-45
16. Shoeman RL, Hartig R, Berthel M, Traub P. Deletion Mutagenesis of the Amino-Terminal Head Domain of Vimentin Reveals Dispensability of Large Internal Regions for Intermediate Filament Assembly and Stability. *Experimental Cell Research* **2002**;279:344-353

17. Watts NR, Jones LN, Cheng N, Wall JS, Parry DAD, Steven AC. Cryo-Electron Microscopy of Trichocyte (Hard Alpha-Keratin) Intermediate Filaments Reveals a Low-Density Core. *Journal of Structural Biology* **2002**;137:109-118

18. Herrmann H, Eckelt A, Brettel M, Grund C, Franke WW. Temperature-Sensitive Intermediate Filament Assembly. *Journal of Molecular Biology* **1993**;234:99-113

19. Herrmann H, Häner M, Brettel M, Ku N-O, Aebi U. Characterisation of Distinct Early Assembly Units of Different Intermediate Filament Proteins. *Journal of Molecular Biology* **1999**;286:1403-1420

20. Stromer MH, Ritter MA, Pang S, Robson RM. Effect of Cations and Temperature on Kinetics of Desmin Assembly. *Biochemical Journal* **1987**;246:75-81

21. Fukuyama K, Murozuka T, Caldwell R, Epstein WL. Divalent Cation Stimulation of *in vitro* Fiber Assembly from Epidermal Keratin Protein. *Journal of Cell Science* **1978**;33:255-263

22. Aebi U, Fowler WE, Rew P, Sun T-T. The Fibrillar Substructure of Keratin Filaments Unraveled. *Journal of Cell Biology* **1983**;97:1131-1143

23. Nelson WJ, Traub P. Effect of the Ionic Environment on the Incorporation of the Intermediate Filament-Sized Filament Protein Vimentin into Residual Cell Structures Upon Treatment of Ehrlich Ascites Tumour Cells with Triton X-100. *Journal of Cell Science* **1982**;53:77-95

24. Gu L, Troncoso JC, Wade JB, Monterio MJ. *In vitro* Assembly Properties of Mutant and Chimeric Intermediate Filament Proteins: Insight into the Function of Sequences in the Rod and End Domains of IF. *Experimental Cell Research* **2004**;298:249-261

25. Kooijman M, van Amerongen H, Traub P, van Grondelle R, Bloemendal M. The Assembly State of the Intermediate Filament Proteins Desmin and Glial Fibrillary Acidic Protein at Low Ionic Strength. *FEBS Letters* **1995**;358:185-188

26. Herrmann H, Wedig T, Porter RM, Lane EB, Aebi U. Characterisation of Early Assembly Intermediates of Recombinant Human Keratins. *Journal of Structural Biology* **2002**;137:82-96

27. Herrmann H, Aebi U. Intermediate Filament Assembly: Fibrillogenesis is Driven by Decisive Dimer-Dimer Interactions. *Current Opinion in Structural Biology* **1998**;8:177-185

28. Herrmann H, Häner M, Brettel M, Müller SA, Goldie KN, Fedtke B, Lustig A, Franke WW, Aebi U. Structure and Assembly Properties of the Intermediate Filament Protein Vimentin: The Role of its Head, Rod and Tail Domains. *Journal of Molecular Biology* **1996**;264:933-953

29. Parry DAD, Steinert PM. Heidelberg, Germany: Springer-Verlag, **1995**:1-183

30. Yamada S, Wirtz D, Coulombe PA. Pairwise Assembly Determines the Intrinsic Potential for Self-Organization and Mechanical Properties of Keratin Filaments. *Molecular Biology of the Cell* **2002**;13:382-391

31. Thomas H, Greven R, Spei M. Experiments for the Isolation of the Matrix Proteins of Wool in Disulphide Form. *Melliand Textilberichte* **1983**;4:297-299

32. Kon R, Nakamura A, Hirabayashi N, Takeuchi K. Analysis of the Damaged Components of Permed Hair Using Biochemical Technique. *Journal of Cosmetic Science* **1998**;49:13-22

33. Castellanos-Serra L, Proenza W, Huerta V, Moritz RL, Simpson RJ. Proteome Analysis of Polyacrylamide Gel-Separated Proteins Visualized by Reversible Negative Staining using Imidazole-Zinc Salts. *Electrophoresis* **1999**;20:732-737

34. Castellanos-Serra L, Carlos F-P, Eugenio H, Huerta V. A Procedure for Protein Elution from Reverse-Stained Polyacrylamide Gels Applicable at the Low Picomole Level: An Alternative Route to the Preparation of Low Abundance Proteins for Microanalysis. *Electrophoresis* **1996**;17:1564-1572

35. Wong C, Sridhara S, Bardwell JCA, Jakob U. Heating Greatly Speeds Coomassie Blue Staining and Destaining. *BioTechniques* **2000**;28:426-432

36. Woods EF, Gruen LC. Structural Studies on the Microfibrillar Proteins of Wool: Characterisation of the Alpha-Helix-Rich Particle Produced by Chymotrypsin Digestion. *Australian Journal of Biological Science* **1981**;34:515-526

37. Harrap BS. The Conformation of a Soluble Wool Keratin Derivative. *Australian Journal of Biological Science* **1962**;16:231-240

38. Klunk WE, Pettegrew JW, Abraham DJ. Two Simple Methods for Quantifying Low Affinity Dye-Substrate Binding. *Journal of Histochemical Cytochemistry* **1989**;37:1273-1281

39. Klunk WE, Pettegrew JW, Abraham DJ. Quantititive Evaluation of Congo Red Binding to Amyloid-Like Proteins with a Beta-Pleated Sheet Conformation. *Journal of Histochemical Cytochemistry* **1989**;37:1393-1297

40. LeVine III H. Quantification of Beta-Sheet Amyloid Fibril Structures with Thioflavin T. *Methods in Enzymology* **1999**;309:274-284

Chapter Seven

Conclusions and Future Perspectives

7.1 Conclusions

The first aim of this work was to develop a simple method to extract wool IFPs from mature wool and separate them into their individual families. Electrophoresis allows the separation of thousands of proteins at once [1]. It is the method of choice when further proteomic techniques are used to analyse the structure and sequences of proteins. Previous work on wool IFPs have separated the proteins on 2DE gels, and many improvements have been made [2-7]. However, a constant problem with the gel resolution was the lack of clarity associated with the high level of insolubility of the proteins. Another problem was the lack of vertical resolution of the individual family members.

Changes were initially made to the methodology of wool protein 2DE separation to increase the solubility of the proteins. This increased solubility combined with the optimisation of the second dimension vertical separation, led to the high-resolution separation of the two keratin IFP classes and their individual family members. High-resolution separation meant that individual protein spots could be further analysed by mass spectrometry.

The extraction method used to obtain wool IFPs was then improved and simplified. Extraction and rehydration for the first dimension was carried out within one day, with no loss in extraction percentage or protein representation. This made the 2DE procedure significantly shorter, contributing to much higher throughput for analysis (Figure 7.1).

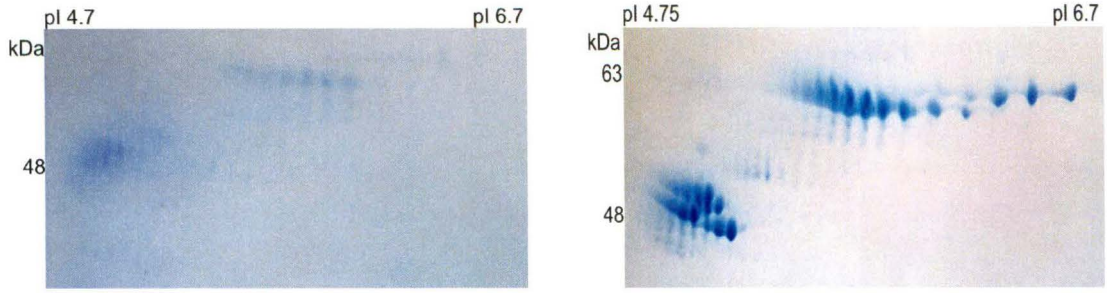


Figure 7.1
Left gel shows the resolution obtained when gels were run before the conditions were optimised. Right gel shows resolution obtained after the running and extraction conditions were optimised.

The high-resolution separation of the wool IFPs demonstrated a large amount of isoelectric heterogeneity, which had been shown before on wool protein gels [4-6]. This heterogeneity was presumed to be caused by PTM, as rows of spots were observed. In the soft keratins, PTMs have been characterised, including phosphorylation and glycosylation [8]. Almost all of the literature on PTMs of keratins involved the soft keratins. In the cytoskeleton of most cells, soft keratins are in a continual state of flux, whereas hard α-keratins are thought to be simply mechanical structures [9]. There may be similarities between hard and soft keratin development and differentiation but it was hypothesised that there may be important differences between the PTMs of the two groups. The assumption made in previous literature [5, 10], that results concerning the PTMs in the soft keratins could be extrapolated to the hard α-keratins, was therefore tested.

Several specific and sensitive techniques were used to determine whether the IFPs were post-translationally modified by phosphorylation or glycosylation in the mature wool. None of the techniques used found any of the IFPs to be modified. It is likely that the IFPs are phosphorylated and glycosylated within the follicle and that the modifications are removed prior to keratinisation. This demonstrates that there may be similarities in the soft keratins and the follicle hard α-keratins, but after keratinisation the hard α-keratins need to be considered as a distinct group.

The lack of evidence for phosphorylation and glycosylation in the mature wool IFPs suggested that something other than PTM was causing the charge heterogeneity of the proteins. Two-dimensional electrophoresis gel analysis excluded the possibility that the spot heterogeneity was caused by the binding of ampholytes to the proteins. High concentrations of urea in the rehydration solution did not reduce the spot heterogeneity. Individual gel spots were eluted and resubjected to 2DE analysis. All of the protein spots were able to reform all the members of their spot family. Amino acid analysis, mass spectrometry and Ellman's assay results support the hypothesis that the proteins have the same sequence but vary in isoelectric charge due to differences in exposure of charged residues on the molecular surface. Based on these results, it can be concluded that each protein spot was able to form a number of other spots by changing conformation and that the different conformations were in equilibrium under the electrophoretic conditions used. This contrasts with earlier work on the heterogeneity of wool IFPs by Ben Herbert *et al.* [5, 10], and suggests that earlier results showing that the wool IFPs are phosphorylated were artifactual. From the re-running results, the 24 main IFP spots could be characterised into the eight different families corresponding to the eight different genes that produce the wool IFPs [11].

The final section of work involved studying the assembly of wool IFPs. The initial studies involved developing a method to fractionate the wool protein classes. Fractionation methods to separate the IFP and IFAP were successfully developed. Further fractionation into the type I and type II IFPs was achieved, along with partial success at isolating individual spots. Once fractionation had been achieved, assembly studies were undertaken.

There was no standard method to reassemble extracted wool IFPs. One attempt had been made previously [12], but other researchers were not able to repeat their results [13]. Several different methods were tested and none produced any filamentous material. The final attempt at assembly of wool IFPs used partially chymotrypsin digested, crudely fractionated IFPs. It was hypothesised that the areas of the IFPs that were causing assembly to be

blocked were cleaved off by the digest treatment, allowing assembly to proceed. This method showed great success. Filaments were formed with a diameter of 10 nm, suggesting they were reassembled wool IFs. Assembled IFs have a diameter ranging from 7 to 15 nm (see Chapter Six, section 6.3). These 10 nm filaments appeared to associate even further to form 40 nm wide filaments (Figure 7.2). The filaments could be disassociated and then reformed into filaments, suggesting that the digestion had removed a part of the IFP that was causing assembly failure.

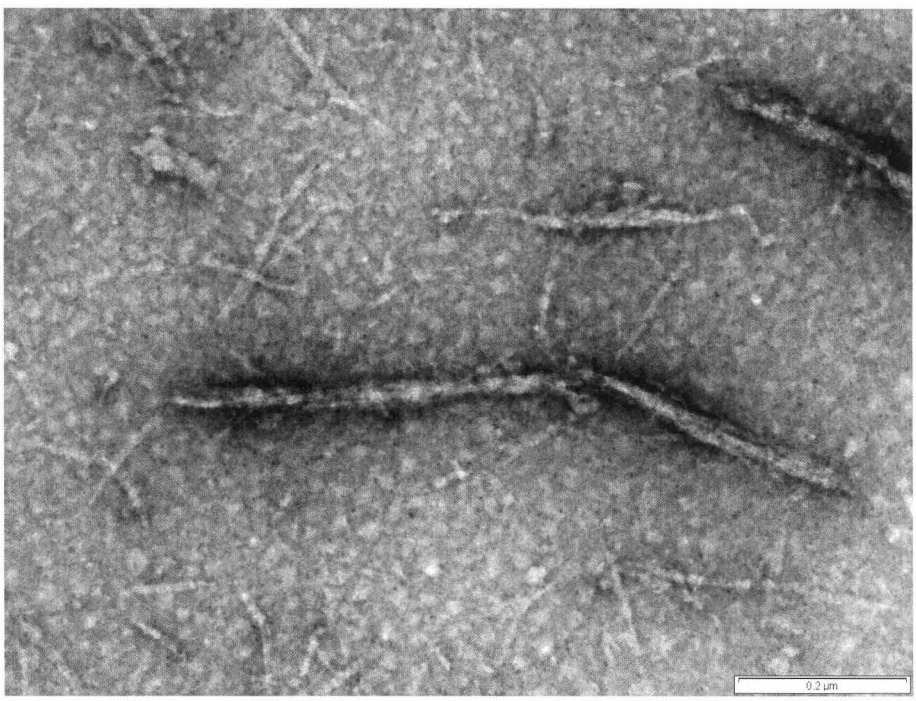


Figure 7.2
Wide filaments formed from acid fractionated, partially chymotrypsin digested proteins that were subjected to the Wang *et al.* [13] assembly conditions. Wide filaments had an approximate diameter of 40 nm.

7.2 Future Perspectives

There is still relatively little known about the genes that control IFs. Understanding the regulation of the gene families is an important first step towards understanding and controlling gene expression. *In situ* hybridisation studies have mapped the expression of major sheep IFPs [14]. However, many more molecular biology studies need to be undertaken before the

regulation of keratin genes will be fully understood. Currently, only four of the eight wool IFP sequences are known and a partial sequence has been identified. In order to confirm the assignment of the IFPs to families, the sequences need to be determined. Currently, mass spectrometry identification uses the four known sequences to assign the proteins to families [6]. It is likely that this has led to the misassignment of several of the IFPs. Once all the sequences have been determined, specific differences in protease fragments can be searched for and used as a tag to identify the IFP families.

Mass spectrometry techniques for the analysis of peptide fragments need to be developed to gain very high sequence coverage. Improvements could be made by derivatising specific amino acids that don't 'fly' very well in the mass spectrometer, and by using a combination of proteases to give differing cleavage patterns. Preparation methods used before mass spectrometry analysis need to be developed further as the wool IFPs are hydrophobic and are difficult to extract out of gels and maintain in a soluble form. Mass spectrometry of intact IFPs would give exact molecular weights of individual proteins, which could confirm family assignments. Techniques for whole protein analysis will be more difficult than peptide analysis, but the main problems will be the same: keeping the proteins soluble and extracting the proteins out of gels.

The results from the PTM work suggest that follicle studies should be undertaken to determine the modification state of the proteins before they are assembled. If modifications are found they could then be studied, to determine what role they play in assembly and at what point the modifications are removed.

Many more assembly studies need to be attempted. The proteins that have been assembled into filaments could be enzymatically phosphorylated to study the role of phosphorylation on the disassembly of hard α-keratin proteins. Glutaraldehyde is able to 'fix' proteins, and allows intermediate stages of assembly to be observed [15]. A timeline of phosphorylation

disassembly could then be determined. The proteins that have already been assembled could be separated using native gels to determine the range of higher ordered structures present and the ratios of these structures. The higher ordered structures could then be digested and analysed with the mass spectrometer. This would give information about the bonds holding the filaments together and which parts of the proteins are interacting. Many of the previous studies on soft keratins (e.g. cross-linking studies) could be done on the hard α-keratins to determine whether their properties are the same. Absorbance measurements on the assembly intermediates should give data to confirm the results seen using TEM.

The role that IFAPs play in assembly of IFPs is uncertain. In this study, only fractions that contained some IFAPs were able to form filaments. This suggests that a small amount of IFAP may be necessary for filament formation. In the wool follicle, it is known that some high sulfur proteins are synthesised concurrently with the IFPs [16]. It has also been shown that follicle IFs are decorated with associated proteins [17]. Assembly studies with specific additions of IFAPs could determine the importance of IFAPs and the quantity and specificity of proteins required.

7.3 IFPs as biomaterials

Currently IFPs are used by Keratec Ltd as an additive in hair conditioning formulations. Keratec IFP™ allows the protein to form coherent films on the cuticle and provides moisturising and softening benefits. It provides anti-ageing efficacy, maintains hair manageability and youthfulness, sacrificially responds to external environmental aggression, has natural anti-oxidant activity and maintains hair fibre strength. Given that the results from this thesis suggests that the IFPs are not phosphorylated, and thus the degree of disassembly and incorporation of extraneous IFPs into the hair is likely to be extremely limited, the mechanism by which these beneficial properties are found to occur is intriguing.

Pure IFPs may be produced in bacterial expression systems [13] to produce high quantities of individual proteins. This system would be able to produce specific proteins for use in novel biomaterials. Product development for Keratec Ltd could involve making IFs from different combinations of IFPs, which would have different properties e.g. soft flexible assemblies could be used for film formation and harder combinations for scaffolds. Alternatively, the use of enzymes to phosphorylate the Keratec IFP™ may facilitate more facile incorporation into the hair.

In summary, the wool fibre has been studied for over 70 years, but there are still large areas of research that need to be undertaken. Many diverse experiments, ranging from fundamental work on the genes and assembly stages to applied development studies on the formation of biomaterials, are still to be achieved.

7. 4 References

1. Westermeier R, Naven T. *Proteomics in Practice - A Laboratory Manual of Proteome Analysis*. Weinheim, Germany: Wiley-VCH Verlag-GmbH, **2002**:316
2. Marshall RC. Analysis of the Proteins from Single Wool Fibers by Two-Dimensional Polyacrylamide Gel Electrophoresis. *Textile Research Journal* **1981**;51:106-108
3. Woods JL, Orwin DFG. Wool Proteins of New Zealand Romney Sheep. *Australian Journal of Biological Science* **1987**;40:1-14
4. Herbert BR, Woods JL. Immobilised pH Gradient Isoelectric Focusing of Wool Proteins. *Electrophoresis* **1994**;15:972-976
5. Herbert BR, Molloy MO, Yan JX, Gooley AA, Bryson WG, Williams KL. Characterisation of Wool Intermediate Filament Proteins Separated by Micropreparative Two-Dimensional Electrophoresis. *Electrophoresis* **1997**;18:568-572
6. Plowman JE, Bryson WG, Flanagan LM, Jordan TW. Problems Associated with the Identification of Proteins in Homologous

Families: The Wool Keratin Family as a Case Study. *Analytical Biochemistry* **2002**;300:221-229

7. Plowman JE. Proteomic Database of Wool Components. *Journal of Chromatography B* **2003**;787:63-76
8. Chou C-F, Smith AJ, Omary MB. Characterization and Dynamics of O-Linked Glycosylation of Human Cytokeratin 8 and 18. *Journal of Biological Chemistry* **1992**;267:3901-3906
9. Parry DAD, Steinert PM. *Intermediate filament structure*. Heidelberg: Springer-Verlag, **1995**:183
10. Herbert BR, Chapman ALP, Rankin DA. Investigation of Wool Protein Heterogeneity Using Two-Dimensional Electrophoresis with Immobilised pH Gradients. *Electrophoresis* **1996**;17:239-243
11. Heid HW, Werner E, Franke WW. The Complement of Native Alpha-Keratin Polypeptides of Hair-Forming Cells: A Subset of Eight Polypeptides that Differ from Epithelial Cytokeratins. *Differentiation* **1986**;32:101-119
12. Thomas H, Conrads A, Phan K-H, van de Löcht M, Zahn H. *In vitro* Reconstitution of Wool Intermediate Filaments. *International Journal of Biological Macromolecules* **1986**;8:258-264
13. Wang H, Parry DAD, Jones LN, Idler WW, Marekov LN, Steinert PM. *In vitro* Assembly and Structure of Trichocyte Keratin Intermediate Filaments: A Novel Role for Stabilization by Disulfide Bonding. *Journal of Cell Biology* **2000**;151:1459-1468
14. Powell BC, Beltrame JS. Characterisation of a Hair (Wool) Keratin Intermediate Filament Gene Domain. *Journal of Investigative Dermatology* **1994**;102:171-177
15. Herrmann H, Haner M, Brettel M, Ku N-O, Aebi U. Characterisation of Distinct Early Assembly Units of Different Intermediate Filament Proteins. *Journal of Molecular Biology* **1999**;286:1403-1420
16. Orwin DFG. The Cytology and Cytochemistry of the Wool Follicle. *International Review of Cytology* **1979**;60:331-374
17. Watts NR, Jones LN, Cheng N, Wall JS, Parry DAD, Steven AC. Cryo-Electron Microscopy of Trichocyte (Hard Alpha-Keratin)

Intermediate filaments Reveals a Low-Density Core. *Journal of Structural Biology* **2002**;137:109-118

Chapter Eight

Materials and Methods

8.1 Materials

Teric GN9 was purchased from Orica (Christchurch, N.Z). Dichloromethane (DCM) and *o*-phosphoric acid were purchased from Biolab (Clayton, Australia). Tris (2-carboxyethyl) phosphine hydrochloride (TCEP) was purchased from Fluka (Buchs, Switzerland). Pefabloc SC was purchased from Roche Diagnostics (Mannheim, Germany). Spectra/Por molecular weight cut off (MWCO) 3500 membrane was purchased from Spectrum (Houston, TX, USA). Thiourea, iodoacetamide (IAM), CHAPS, periodic acid, Tween™ 20, monoclonal anti-phospho serine clone PSR-45, anti-mouse igG peroxidase antibody, sodium acetate, Schiff's reagent, bovine albumin fraction V, chymotrypsin TLCK, NEM, sodium nitroprusside, phenylisothiocyanate (PITC), DL-norleucine, sodium tetrathionate, NNN'N'-tetramethylethylenediamine (TEMED), L-lysine, L-arginine, guanidine hydrochloride, protein disulfide isomerase (bovine pancreas), α-chymotrypsin, trypsin inhibitor and glutaraldehyde were purchased from Sigma (St. Louis, MO, USA). Pharmalyte™ 3-10, hyperfilm ECL, agarose NA, immobiline™ drystrips and immobiline buffers were purchased from GE Healthcare (Uppsala, Sweden). Orange G was purchased from George T Gurr LTD (London, England). ReadyStrip IPG Strips, immun-blot™ PVDF membrane, precision plus protein standards, Coomassie® brilliant blue R-250, Coomassie® brilliant blue G-250, criterion Tris-HCl gels, 40% acrylamide/bis solution, 37.5:1 (2.6%C) acrylamide 99.9% and N,N`-methylene-bis-acrylamide were purchased from Bio-Rad Laboratories (Hercules, CA, USA). SnakeSkin® pleated dialysis tubing, phosphoprotein control set, Super Signal® west femto, Ellman's reagent and cysteine-HCl were from Pierce (Rockford, IL, USA). 'Odina' paraffin oil was purchased

from Shell NZ Ltd (Wellington, NZ). Methanol (MeOH), trichloroacetic acid, ethanol (EtOH), trifluoromethanesulfonic acid (TFMS), pyridine, calcium chloride, zinc sulfate and sodium acetate trihydrate were purchased from Merck (Darmstadt, Germany). Thymol was purchased from the Aldrich Chemical Company (Milwaukee, USA). MA1-076 antibody was purchased from Affinity Bioreagents (CO, USA). The Pro-Q® Diamond phosphoprotein staining kit, PeppermintStick™ phosphoprotein molecular weight standards, Pro-Q® Emerald glycoprotein staining kit, CandyCane™ molecular weight markers and Sypro® Ruby stain were from Molecular probes (Eugene, Oregon, USA). Acetonitrile and ethylenediamine-tetraacetic acid (EDTA) were purchased from Mallinckrodt (NJ, USA). Polaroid® 667 film was from Polaroid Corporation (UK). Glyco-macropeptide variants A/B and B were supplied by Kate Palmano, Fonterra (Palmerston North, NZ). GBX developer and replenisher and GBX fixer and replenisher were purchased from Kodak (VIC, Australia). Dimethyl formamide was from Asia Pacific Speciality Chemicals Ltd (NSW, Australia). POROS®20R2 preparative chromatography media was purchased from Applied Biosystems (Foster City, CA, USA). Carbograph 120/400 mesh preparative media was from Alltech Asociated Inc. (Deerfield, IL, USA). Trypsin TPCK was from Promega™ (Madison, WI, USA). 1 mL syringes purchased from Terumo (Philippines). Nickel chloride was from May and Baker (Dagenham, England). All other chemicals were purchased from BDH (Poole, Dorset, England). NANOpure water with resistance higher than 18 MΩ was used throughout the experiments.

8.2 Preparation of wool protein extracts

All experiments used protein extracted from the same fleece to avoid inter-fleece variation. Mid-side wool samples from Romney fleece were de-tipped with diamond-edged scissors then scoured in a tea strainer with 0.15% Teric GN9 for 2 minutes at 60°C, 0.15% Teric GN9 for 2 minutes at 40°C, water for 2 minutes at 40°C, water for 2 minutes at 60°C, then left to air-dry overnight. The following morning, the dried wool sample was washed twice in DCM for 30 seconds, twice in EtOH for 30 seconds, twice in water for 30 seconds, then air-dried overnight.

Dried scoured wool (2 g) was cut into small sections with diamond-edged scissors then freeze-crushed to a fine powder using liquid nitrogen in a mortar and pestle. The crushed wool sample was dried under vacuum and over phosphorus pentoxide for two days.

Extraction Method One: dried wool powder (1 g) was extracted overnight by vigorously shaking in a Trident shaker (Canesis Developments LTD) at room temperature in 100 mL of extraction solution consisting of 8 M urea, 0.05 M Tris-base, 50 mM TCEP and 10 mM Pefabloc SC (a protease inhibitor) at approximately pH 4. Extracted protein was dialysed in Spectra/Por MWCO 3500 membrane against five changes of water and then freeze-dried.

Extraction Method Two: 8 mL of SDS sample buffer (1% w/v SDS, 50 mM Tris-HCl pH 8.8, 50 mM TCEP at 95°C) was added to 200 mg of wool contained in a vial. The vial was placed in a beaker filled with ice-cold water and sonicated at 150 W (6 times for 1 second with a 9.9 second wait between sonications). The vial was placed in a boiling water bath for 5 minutes before being briefly cooled in an ice-bath. 20 mL of lysis buffer (2 M thiourea, 7 M urea, 4% CHAPS, 2% DTT, 2% Pharmalyte 3-10™) was added. The vial was shaken for 4 hours before it was centrifuged at 14 500 rpm for 10 minutes. The supernatant was stored frozen in aliquots.

Quantitative analysis: protein, which had been extracted but not freeze-dried, was analysed for protein concentration by a gravimetric method. Briefly, the residue left after extraction was placed in filter paper in a funnel and rinsed with water 6 times before being dried at 37°C overnight. Residues were conditioned at room temperature for another 24 hours before being weighed, and extraction percentage calculated (Table 8.1).

8.3 Electrophoresis

All experiments were conducted in duplicate or triplicate.

Extraction Method	Percentage of protein extracted (w/w)
Extraction Method One	75.88% ± 13.29
Extraction Method Two	74.06% ± 1.09

Table 8.1
Extraction percentages of the two different extraction methods used. Values were calculated using quantitative analysis. Values are the result of triplicate experiments.

8.3.1 Rehydration

Wool proteins were solubilised by vortexing freeze-dried protein (Extraction Method One) in 350 µL [1] of an oxygen-free nitrogen flushed rehydration solution consisting of 6 M urea, 2 M thiourea, 0.5% Pharmalyte 3-10™, 0.004% Orange G, 4% CHAPS, 0.4% DTT at approximately pH 7. Extracted protein solution (Extraction Method Two) was diluted with rehydration solution. In experiments excluding Pharmalyte 3-10™ no pharmalyte was added to the rehydration solution. Experiments where the effect of urea in the rehydration solution was studied had no thiourea in the rehydration solution and urea was added to the rehydration solution at a concentration of 4 M, 6 M or 10 M.

All protein-rehydration solutions were centrifuged at 14 500 rpm for 10 minutes to remove cellular debris. IPG strips were placed in a plastic tissue culture pipette then rehydrated by pipetting the rehydration solution on top of the IPG strip. The plastic tissue culture pipettes containing the IPG strip and rehydration solution were flushed with oxygen-free nitrogen then sealed for overnight rehydration at room temperature. Alternatively, rehydration trays were used to contain the IPG strips and rehydration solution. The trays were placed in a sealed chamber with 2 open taps then flushed with nitrogen for 10 minutes before closing the taps for overnight rehydration.

After rehydration, all of the protein-rehydration solution had been absorbed by the IPG strip. IPG strips were rinsed with paraffin oil briefly to remove any crystallised urea.

8.3.2 IEF

IEF was carried out using a Pharmacia Multiphor® II electrophoresis system with a Dry Strip Kit. Paraffin oil used to cover the focusing strips, was degassed for at least 30 minutes prior to use [2]. Power was supplied from an EC6000P power supply (E-C Apparatus Corporation, Holbrook, NY) and cooling water at 20°C [3] was supplied by a Pharmacia Multitemp II, thermostatic circulator.

The anode electrode wick was soaked with water, then excess liquid blotted off with filter paper, the cathode wick was soaked with a 0.4% DTT solution before also being blotted and then applied to the electrodes. The settings used for IEF were: 50 V for 1 hr, 500 V for 1 hr, 1000 V for 1 hr and 6000 V to the steady state. After IEF, the strips were stored at -80°C on glass plates and covered with plastic wrap.

8.3.3 Equilibration

Prior to separation in the second dimension, IPG strips were equilibrated in a solution composed of 0.05 M Tris, 6 M urea, 20% glycerol, 2% SDS, 1% DTT pH 8.8 for 15 minutes, drained on filter paper, then equilibrated for a further 15 minutes in the same solution with 4% IAM instead of the DTT, to remove excess DTT.

8.3.4 Large format gels

After draining excess equilibration solution onto filter paper, each IPG strip was sealed onto a second dimension slab gel using 1% agarose in 1:4 diluted gel buffer. Second dimension gels were run using a BioRad Protean IIXL slab gel electrophoresis tank. Gels were run using either a continuous buffer system (192 mM glycine, 25 mM Tris, 0.1% SDS, at pH approximately 8.6) or a discontinuous buffer system (anode buffer: 375 mM Tris, 0.1% SDS pH 8.8 adjusted with hydrochloric acid (HCl), Cathode buffer: 192 mM Glycine, 0.1% SDS pH 8.3 adjusted with Tris.). Gels were cooled by connecting the tank's cooling core to a thermostatic circulator set to 8°C. Gels were run at 32 mA per gel for 50 minutes then at 48 mA per gel until the bromophenol blue front was at the bottom of the gel. Overnight separations

were run at 4 mA per gel for 2 hours, followed by 20 mA per gel until the next morning.

Vertically run second dimension gels were initially run at low milliAmps to reduce the effect of electroendosmosis [4]. Electroendosmosis occurs when the IPG gel strips carboxylic acid groups become negatively charged. The buffers in the IPG strip are immobilized and therefore can't move towards the anode; this is compensated for by a counter flow of H_3O^+ ions towards the cathode. This flow of counter ions can carry solubilized substances with it. Reducing the initial current during a second dimension run minimises the counter flow of ions [4].

Gels were 2 mm-thick, 20 x 20 cm, 8.5% T, 2.5% C, gels cast with a stacker had a 3 cm, 4% T stacking gel cast on top of the separating gel.

8.3.5 Medium format criterion gels

After draining excess equilibration solution onto filter paper the IPG strips were loaded into an IPG comb of a criterion Tris-HCl gel. For 1DE gels, protein was dissolved in sample buffer (6 M urea, 2% SDS, 0.062 M Tris-HCl pH 6.8, 6% β-mercaptoethanol, 10% glycerol, 0.0001% bromophenol blue) and boiled for 4 minutes before being loaded into a well in a Tris-HCl gel. Gels were run using Tris/glycine/SDS buffer (192 mM glycine, 25 mM Tris base, 0.1% SDS pH 8.3 unadjusted), at 200 V until the bromophenol blue front left the gel.

8.4 Gel stains

8.4.1 Coomassie brilliant blue G-250 stain

After electrophoresis, gels were fixed in 2% phosphoric acid, 50% MeOH overnight. Gels were washed 3 times with water for 30 minutes then incubated in 2% phosphoric acid, 34% MeOH, 17% ammonium sulphate for 1 hour. 0.66 g/L of Coomassie brilliant blue G-250 was added direct to the containers and the gels stained for 5 days. To clear the background, gels were washed with several changes of water [5].

8.4.2 Reversible negative zinc stain [6, 7]

Gels were rinsed with water for 30 seconds then equilibrated in 0.2 M imidazole, 0.1% SDS for 15 minutes. The gels were then developed in 0.2 M zinc sulfate until the gel background became deep white leaving the protein spots transparent (less than 1 minute). Staining was stopped by rinsing the gel in several changes of water.

8.4.3 Blum silver stain [8]

Gels were fixed in 50% MeOH, 12% acetic acid (HAc), 0.05% formaldehyde for at least 1 hour. After washing gels with 50% EtOH 3 times for 20 minutes gels were pretreated with sodium thiosulfate (0.2 g/L) for 1 minute. The gels were rinsed 3 times with water before impregnation with silver nitrate (2 g/L), 0.075% formaldehyde for 20 minutes. Gels were developed with sodium carbonate (60 g/L), 0.05% formaldehyde, sodium thiosulfate (4 mg/L) for 10 minutes, after twice washing with water for 20 seconds. Gels were washed with water twice for 2 minutes then the stain development was stopped with 50% MeOH, 12% HAc.

8.4.4 Coomassie brilliant blue R-250 stain [9]

The gels were microwaved in solution A (0.05% Coomassie brilliant blue R-250, 25% IPA, 10% HAc) for 90 seconds, then rocked for 5 minutes. After 2 water washes the gels were microwaved for 80 seconds in solution B (0.005% Coomassie brilliant blue R-250, 10% IPA, 10% HAc). The gels were once again washed twice in water then microwaved in solution C (0.002% Coomassie brilliant blue R-250, 10% HAc) for 80 seconds before 2 water washes. Solution D (10% HAc) was added and the gels were microwaved for 80 seconds before rocking for 5 minutes. Gels were then washed in water to reduce background staining.

8.4.5 Acid violet 17 stain [10]

The IPG gels were fixed in 20% trichloroacetic acid for 60 minutes, rinsed for 1 minute with 3% phosphoric and then stained with 0.1% acid violet (from a 1% acid violet stock solution which had been heated to 60°C while stirring) in 11% phosphoric acid for 10 minutes. The gels were destained with 3%

phosphoric acid until the background became clear. The gels were then washed with several changes of water before a final wash in 5% glycerol.

8.4.6 Alkali-labile phosphoprotein stain [11]

SDS-PAGE gels were fixed overnight with 25% IPA, 10% HAc. The next morning the solution was refreshed and washed for 1 hr. The gels were then transferred into fixative solution containing 0.5 M $CaCl_2$ for 20 minutes. The gels were rapidly rinsed with water to remove excess surface $CaCl_2$. Hot 0.5 M NaOH (80°C) was poured onto the gel and left to equilibrate. After 30 minutes, excess NaOH was removed by a brief wash with water. The gels were then equilibrated in 8.0 mM ammonium heptamolybdate, this solution was renewed after 20 minutes. After another 20 minutes, the gels were rapidly rinsed with water to remove excess ammonium heptamolybdate. During this step fresh colour reagent was prepared by adding 5 mL of rhodamine B to 40 mL 5 M HCl, 40 mL 34 mM ammonium heptamolybdate, 115 mL water, and stirred to produce a clear bright red solution. The gels were stained in colour reagent for 20-30 minutes then destained in 1 M HCl, which was refreshed after 1 hour.

8.4.7 Pro-Q® Diamond phosphoprotein gel stain

2-DE gels were fixed in 50% MeOH, 10% trichloroacetic acid overnight. After washing with water 2 times for 10 minutes, gels were placed in Pro-Q® Diamond phosphoprotein gel stain, and incubated in the dark for 2 hours. Background fluorescence was reduced and staining sensitivity was increased by incubating the gel in 20% acetonitrile, 50 mM sodium acetate, pH 4.0, for 1 hour in the dark. This step was repeated twice more. Gels were visualised by placing on a 300 nm UV transilluminator. Images were taken on Polaroid® 667 film using a Polaroid® camera fitted with a SYPRO® photographic filter. After detection of phosphorylated proteins, gels were stained with SYPRO® Ruby total protein

8.4.8 Thymol-sulfuric acid stain [12]

After electrophoresis, the gels were placed in a fixative solution of 25% isopropanol (IPA), 10% HAc, and placed on a rocker-action shaker; after 2

hours, the fixative solution was replaced with fresh fixative solution. The next morning fresh fixative solution was added which contained 0.2% thymol. The gels were then washed in 80% sulfuric acid, 20% EtOH at 35°C for 2.5 hours.

8.4.9 Periodic acid-Schiff's stain

A) Gels were fixed with 1% periodic acid and 3% HAc for 50 minutes. They were washed with water 6 times for 10 minutes each. Gels were stained with Schiff's reagent for 50 minutes in the dark. The gels were then washed with 0.5% sodium metabisulfite, 3 times for 3 minutes each, before washing with water. The gels were stored in 5% HAc [13].

B) After electrophoresis, gels were fixed with 50% MeOH overnight. In the morning, gels were washed with water for 20 minutes followed by a 15 minute wash in 2% periodate. The gels were washed twice with water before being placed in Schiff's reagent until the bands turned magenta. The gels were then washed twice with water and then in 2% sodium metabisulfite overnight. The following morning, the gels were washed with water until the water remained clear [14].

8.4.10 Pro-Q® Emerald 300 glycoprotein gel stain

2DE gels were placed in 50% MeOH, 5% HAc. This solution was refreshed before an overnight incubation. After washing the gels twice with 3% HAc for 20 minutes, they were placed in oxidising solution (0.04 M periodic acid, 3% HAc) for 30 minutes. The gels were then washed 3 times in 3% HAc for 20 minutes each wash, followed by staining with Pro-Q® Emerald 300 solution for 2 hours. After staining, the gels were washed twice in 3% HAc for 20 minutes each wash. Gels were visualised by placing on a 300 nm UV transilluminator. Images were taken on Polaroid® 667 film using a Polaroid® camera (Polaroid Corporation, USA) fitted with a Sypro® photographic filter. After detection of glycosylated proteins, gels were stained with Sypro® ruby total protein stain for 3 hours in the dark. After washing with a 10% MeOH, 7% HAc aqueous solution for 30 minutes, gels were viewed and photographed as before.

8.5 Enzymatic dephosphorylation

Romney wool protein was dephosphorylated using the method of Yamagata *et al.* [15]. Briefly, 20 mg of extracted wool protein was dissolved in lysis buffer (5 M urea, 2 M thiourea, 2% CHAPS, 2% *N*-dimethyl-3-ammonio-1-propane-sulfonate (SB3-10), 2% Pharmalyte ™ 3-10 and 1% DTT). 20 μL of this solution was mixed with 10 μL of 10% SDS. Water (945 μL), 20 mM $MnCl_2$ (0.5 μL) and λ protein phosphatase buffer (2 μL) were added sequentially and mixed before the addition of λ protein phosphatase (0.125 μL, 100 units). After gentle vortexing the mixture was incubated at 30°C overnight. α-casein was used as a control protein.

8.6 Chemical deglycosylation [16]

Romney wool protein (1 mg) and fetuin (1 mg) were dried overnight in 1 mL screw capped vials placed in a desiccator containing di-phosphorus pentaoxide. After placing the vials in a dry ice-EtOH bath, 150 μL of ice cold TFMS was added to each vial. The vials were flushed with nitrogen and sealed. After incubation with occasional shaking at 0°C for 2 hours, the vials were placed in a dry ice-EtOH bath and neutralized by gradual addition of 470 μL of 60% pyridine. The vials were cooled in dry ice for 5 minutes then placed in an ice-water bath for 15 minutes. 400 μL of 0.5% ammonium bicarbonate was added to each vial and mixed. Dialysis tubing was attached to the opening of the vials and the proteins were dialysed against 0.5% ammonium bicarbonate at 4°C.

8.7 Immunoblotting

Proteins separated by 1- and 2DE were transferred onto a PVDF membrane using a Trans-blot® electrophoretic cell (BioRad, Hercules, CA, USA) and blotting buffer consisting of 192 mM glycine, 25 mM Tris-base, 0.075% SDS. The membrane was then blocked for 1 hour in Tris buffered saline containing Tween™20 (TBST) (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Tween™20) containing 3% BSA. After draining, the blocking buffer primary

antibody (monoclonal anti-phospho serine clone PSR-45 or MA1-076) in fresh blocking buffer was added and the membrane incubated for 1 hour. After a brief wash with TBST, the membrane was washed six times with large amounts of TBST for 5 minutes each wash. The membrane was then incubated in secondary antibody (anti-mouse IgG peroxidase antibody) in blocking buffer for 1 hour. After a short wash and six high volume five minute washes in TBST, the membrane was incubated with Super Signal® West Femto for five minutes; after draining, the membrane was placed in a sheet protector and exposed to hyperfilm ECL for between 1 minute and 30 minutes. The films were developed using Kodak developer and fixer.

8.8 Mass Spectrometry

All mass spectrometry experiments were conducted in triplicate.

8.8.1 GMP digestion

Glycomacropeptide (GMP) variant B (300 µg/mL (14.06 pmole/µL)) was dissolved in 10 mM Tris-HCl, 10 mM calcium chloride buffer at pH 7.8. 1 mL was digested with 5 µg of α-chymotrypsin (1:60) at 30°C for 4 hours or 25°C for 16 hours. Digests were acidified with formic acid.

8.8.2 In-gel digestion

Excised gel protein spots from 2DE gels were cut into 1x1 mm pieces then placed into 1.5 mL plastic vials. The spots were de-stained twice with 100 mM ammonium bicarbonate in 40% acetonitrile for 30 minutes. Gel spots were washed twice with water then vacuum centrifuged to dehydrate. Gel spots were then reduced by covering the excised gel spots with 5 mM TCEP in 100 mM ammonium bicarbonate solution, for 45 minutes at 56°C. This wash was discarded, then gel pieces were washed twice with water. After vacuum centrifuging, the gel pieces were alkylated with 22 mM IAM in 100 mM ammonium bicarbonate. Samples were agitated with a vortex mixer in the dark for 30 minutes at room temperature. The alkylating reagent was discarded and gel pieces washed twice with 100 mM ammonium bicarbonate in 40% acetonitrile then twice with water. After discarding any residual

solution, the gel pieces were dehydrated with a vacuum centrifuge. Trypsin TPCK (35 µL of 12.5 ng/µL) was added prior to the addition of enough 50 mM ammonium bicarbonate in 30% DMF to cover the gel pieces. The alkylated proteins were digested for 16 hours at 37°C. The resulting peptide solution was then transferred to a 0.5 mL tube and the remaining gel pieces extracted twice with 100 µL volumes of 0.5% formic acid in 50% acetonitrile. Peptide extracts were then concentrated to near dryness (<50 µL) on a heating block at 30°C under a gentle stream of nitrogen. Prior to mass spectrometry peptides were re-suspended in 10-60 µL of 5% formic acid.

8.8.3 Peptide purification

Analysis by electro-spray ionisation-mass spectrometry is adversely affected by salts and buffers, therefore efficient methods for contaminant removal are required [17]. Peptides were desalted and concentrated with a combination of POROS®20R2 and carbograph media [18-20]. GELoader tips containing chromatography media were rinsed with 5% formic acid prior to sample addition. The trapped peptides were washed with 5% formic acid, before elution with 10-20 µL of 50% MeOH in water.

8.8.4 Instrumentation

All mass spectrometry was performed on a QSTAR Pulsar-i quadrupole time-of-flight tandem mass spectrometer (Applied Biosystems/PE Sciex, Toronto, Canada) equipped with a Protana™ nanoES source. Samples were introduced via nanocapillaries (Proxeon Biosystems, Odense, Denmark). The mass spectrometer was operated with AnalystQS™ software and data were interpreted and *de novo* sequenced with Bioanalyst 1.1™ or ProID™ (Applied Biosystems/PE Sciex, Toronto, Canada). Additional database analysis was with MASCOT from matrix science. Operation and analysis was carried out with the help of Nigel Joyce and Scott Bringans (Canesis Network Ltd).

8.9 Re-running

Following reversible negative zinc staining individual gel spots were excised with a scalpel. The gel spot was placed in a 1 mL syringe, which had 32 μm wire mesh placed in the bottom. The gel spot was crushed through the mesh into a 1.5 mL vial. 100 μL of elution buffer (25 mM Tris-HCl pH 8.3, 500 mM glycine) was used to flush the syringe of gel pieces. Crushed gel spots were vortexed for 10 minutes before another 100 μL of buffer was added via the syringe to flush out any remaining gel pieces. The vortexing and buffer additions were repeated twice before 1 mL of 80% ice-cold acetone was added to the vial. After a brief vortex to mix the solutions, the vial was incubated at -20°C for at least 2 hours. Vials were then centrifuged and the supernatant was pipetted off and discarded. Rehydration solution was added to each vial and 2D-PAGE was performed.

8.10 Alkylation time course

Dried wool powder (10 mg) was extracted overnight in vials by vigorously shaking in a Trident shaker (Canesis Developments LTD) at room temperature in 1 mL of extraction solution consisting of 8 M urea, 0.05 M Tris base pH 9.3, 50 mM DTT. Vials were centrifuged for 10 minutes at 14 500 rpm to pellet the wool residue. 500 μL of supernatant was pipetted into a new vial and excess alkylating reagent was added (1 M IAM, 200 mM NEM or 7 M acrylamide) in 2.3 M Tris buffer. After a quick vortex, vials were nitrogen flushed and incubated in the dark for 10 minutes, 1, 2, 6, or 24 hours. After incubation, 20 μL of ß-mercaptoethanol was added to quench the alkylation reaction. Following 10 minutes vortexing, a 5 μL aliquot was tested to verify that the alkylation reaction had been stopped. The test involved dissolving a few crystals of sodium nitroprusside in water, adding 5 μL of this solution to the aliquot and visualising a change to purple, indicating that there were free thiols present (i.e. there was an excess of ß-mercaptoethanol). Alkylated proteins were dialysed in their vials by cutting the lids off the vials and attaching SnakeSkin® dialysis tubing across the entry

of the vial. Inverted vials were placed in dialysis supports, and dialysed against 5 changes of water and then freeze-dried.

8.11 Amino acid analysis [21]

A 10 mg/mL solution of protein in 2% SDS was freeze-dried in a culture vessel then hydrolysed under nitrogen using 6 M HCl, 1% phenol, at 110°C for 24 hours. After re-drying, amino acids were derivatised with phenylisothiocyanate (PITC). The derivatised amino acids were run through a guard column (Widepore C18, 4 mm x 2 mm, Phenomenex, CA, USA) before being separated on a reversed-phase column (Pico•Tag®, 3.9 mm x 150 mm, Waters, Milford, MA, USA) at 36°C. Chromatographic separation was carried out using a ternary solution system. Buffer A was 0.14 M sodium acetate buffer (containing 0.05% triethylamine, 0.001% EDTA); the pH was adjusted to 5.9 with HAc. Buffer B was NANOpure water. Buffer C was 100% acetonitrile. The gradient profile used for amino acid separation is shown in Table 8.2. An internal standard, DL-norleucine was used in all experiments.

Time (minutes)	Flow (mL/minute)	A (%)	B(%)	C(%)
0.00	1.0	94.0	0.0	6.0
16.5	1.0	50.8	18.4	30.8
16.7	1.0	0.0	40.0	60.0
18.2	1.0	0.0	40.0	60.0
18.5	1.5	0.0	40.0	60.0
18.7	1.5	0.0	40.0	60.0
19.0	1.5	94.0	0.0	6.0
23.0	1.5	94.0	0.0	6.0
23.5	1.0	94.0	0.0	6.0

Table 8.2
Chromatographic gradient conditions for amino acid analysis.

8.12 Fractionation methods

All fractionation methods were conducted in duplicate or triplicate.

8.12.1 1DE preparative gels

Extracted wool protein (60 mg/ 9 mL) in 1DE sample buffer (6 M urea, 2% SDS, 0.062 M Tris-HCl pH 6.8, 6% β-mercaptoethanol, 10% glycerol, 0.0001% bromophenol blue) was loaded on top of a 4 mm thick 10% T gel after boiling for 4 minutes. Gels were run using a continuous buffer system (192 mM glycine, 25 mM Tris base, 0.1% SDS). Gels were cooled by connecting the tank's cooling core to a thermostatic circulator set to 8°C. Gels were run at 32 mA per gel for 50 minutes then at 48 mA per gel until the bromophenol blue front was at the bottom of the gel. Gels were stained using several methods. The first was a microwave Coomassie R-250 method [9]. The simplified protocol was used where the 2nd and 3rd steps of the full protocol are omitted. The gel was microwaved in 0.05% Coomassie brilliant blue R-250, 25% IPA, 10% HAc for 90 seconds, rocked for 5 minutes then rinsed 2 times with water. The gel was then microwaved in 10% HAc for 80 seconds, rocked for 5 minutes then rinsed 2 times with water. Gels were also stained with a reversible negative zinc stain (8.4.2) and Coomassie brilliant blue G-250 stain (8.4.1).

Zinc stained proteins were destained twice for 8 minutes with 50 mM Tris-HCl pH 8.3, 0.3 M glycine, 30% acetonitrile before being eluted. Proteins were eluted off the gel by excising the bands and placing in a tube with buffer (50 mM Tris-HCl pH 8.8, 2 mM EDTA, 5 mM TCEP, 0.2% SDS) then shaking overnight. Proteins were dialysed in SnakeSkin® dialysis tubing against 5 changes of water and then freeze-dried.

8.12.2 Acid fractionation [22]

Wool proteins were extracted using Extraction Method Two (8.2). Proteins were dialysed in SnakeSkin® dialysis tubing against 5 changes of water then acidified to pH 4 with HAc. After vigorous shaking the proteins were frozen for 5 hours. The precipitate was centrifuged then freeze-dried.

8.12.3 Kon fractionation [23]

Dried wool powder (1 mg) was placed in a vial with 1 mL of 25 mM Tris-HCl pH 8.3, 1% SDS, 2 M β-mercaptoethanol. The wool powder was extracted for 3 days at 50°C. The vial was centrifuged at 14 000 rpm for 10 minutes and the supernatant was stored at 4°C (Fraction 1). The residue was washed with 25 mM Tris-HCl pH 8.3 before 25 mM Tris-HCl pH 8.3, 1% SDS, 0.4 M β-mercaptoethanol was added. The residue was extracted for 3 days at 50°C. The vial was centrifuged at 14 000 rpm for 10 minutes and the supernatant was stored at 4°C (Fraction 2). 10 μL of fraction 1 or 2 was dissolved in 1DE sample buffer (8.12.1) and boiled for 4 minutes. Proteins were loaded onto a 1DE criterion Tris-HCl gel and were run using a continuous buffer (192 mM glycine, 25 mM Tris base, 0.1% SDS pH 8.3 unadjusted), at 200 V until the bromophenol blue front left the gel.

8.12.4 Sulfitolysis extraction/fractionation [24]

Dried wool powder (300 mg) was extracted overnight in tubes by vigorously shaking in a Trident shaker (Canesis Developments LTD) at room temperature in 30 mL of extraction solution consisting of 0.1 M Tris-HCl pH 9.5, 0.2 M sodium sulfite, 0.1 M sodium tetrathionate and 8 M urea. Tubes were centrifuged for 20 minutes at 13 400 rpm to pellet the wool residue. Proteins were then either precipitated with 80% ice-cold acetone, and pelleted, or they were fractionated into protein classes. The IFPs were precipitated with 1 M zinc acetate, centrifuged then suspended in 1% sodium citrate. After dialysis against 0.05 M sodium tetraborate, the proteins were once again precipitated with zinc acetate, dialysed against water and freeze-dried.

8.12.5 IEF fractionation

The ZOOM® IEF Fractionator (Invitrogen, CA, USA) was used to separate wool proteins based on pI. Membranes solutions [25] were made in 10 mL volumetrics. Immobiline buffers and Tris base were added (Table 8.3). Followed by 3.33 mL of 30% T, 8% C acrylamide. Volumetrics were made up to volume with water. TEMED (5 μL) and ammonium peroxodisulphate

(APS) (10 μL of a 40% solution) were added and mixed. Discs were cast by placing a 7 cm Whatman (Kent, England) glass microfibre filter in the lid of a petri dish, pouring over the membrane solution then placing the petri dish base on top of the lid. The discs were left at room temperature to polymerise overnight. The discs were stored in 20 mM sodium acetate, 2 mM sodium azide, and were washed with water 3 times 30 minutes before use.

Membrane pH	Immobiline buffer 3.6 (μL)	Immobiline buffer 4.6 (μL)	Immobiline buffer 6.2 (μL)	Immobiline buffer 7.0 (μL)	Immobiline buffer 9.3 (μL)	1 M Tris base (μL)	1 M Hac (μL)
4.5	353	154	271	34	99	63	0
5	310	229	235	65	190	47	0
5.5	268	305	199	95	280	32	0
6	225	381	163	126	371	16	0
6.1	216	396	155	132	289	13	0
6.5	182	456	126	157	461	0	0
7	139	532	90	188	551	0	9

Table 8.3
Formulae for making membranes for use in the ZOOM® IEF Fractionator.

Extracted freeze-dried protein (3 mg / mL) was dissolved in rehydration solution (8.3.1). 670 μL of this solution was added to each chamber of the ZOOM® IEF Fractionator. Each chamber was separated by a membrane disc. Cathode buffer was made from 10 x cathode buffer (20 mM lysine, 20 mM arginine) by adding 8.4 g of urea and 3.0 g of thiourea to 2 mL of 10 x cathode buffer and making up to 20 mL with water. Anode buffer was made from 50 x anode buffer (7 mM phosphoric acid) by adding 8.4 g of urea and 3.0 g of thiourea to 3.3 mL of 50 x anode buffer and making it up to 20 mL. The proteins were focused at 100 V for 20 minutes, 200 V for 80 minutes then 600 V for 80 minutes. Separate IPG strips were rehydrated with 214 μL of solution from each chamber, then focused (8.3.2) and stained with acid violet.

8.13 Assembly methods

8.13.1 Thomas *et al.* method [24]

Freeze-dried proteins were dissolved in Solution 1 (Table 8.4), at a concentration of 2 mg/mL. Filaments were reconstituted by dialysing against Solution 2 at 6°C overnight followed by Solution 3 at 6°C for 12-24 hours.

Solution 1	Solution 2	Solution 3
8 M urea	4 M urea	0 M urea
8% β-mercaptoethanol	25 mM β-mercaptoethanol	10 mM β-mercaptoethanol
50 mM Tris-HCl pH 7.5	10 mM Tris HCl pH 7.5	10 mM Tris HCl pH 7.5

Table 8.4
Solutions for use in the assembly of IFPs using the Thomas *et al* method [24].

8.13.2 Wang *et al.* method [26]

All solutions were degassed under vacuum and nitrogen restored before use. Proteins were dissolved in solution 1, and then dialysed for 1 hour each in Solution 2, 3, 4 then 4 hours in solution 5 (Table 8.5).

Solution 1	Solution 2	Solution 3	Solution 4	Solution 5
10 mM Tris-HCl pH 9	10 mM Tris-HCL pH 8	10 mM Tris-HCL pH 8	10 mM Tris-HCL pH 8	10 mM Tris-HCL pH 8
9.5 M urea	6 M urea	4 M urea	2.5 M urea	175 mM NaCl
2.5 mM EDTA	2.5 mM EDTA	2.5 mM EDTA	2.5 mM EDTA	2.5 mM EDTA
5 mM TCEP	5 mM TCEP	5 mM TCEP	5 mM TCEP	5 mM TCEP

Table 8.5
Solutions for use in the assembly of IFPs using the Wang *et al.* method [26].

8.13.3 Assembly buffer variables

All protein solutions were dissolved in Solution 1 of the Wang *et al* method [26] (approximately 2 mg/mL). Various changes to the Wang *et al* method

were made as shown in Table 6.1. Following the last step of dialysis, samples were taken for TEM analysis and their A_{340} was measured.

8.14 Electron microscopy [27]

Samples were applied to 200 mesh glow-discharged formvar carbon-coated copper grids (SPI supplies, PA, USA). 5 μL of protein solution was pipetted onto the grid and left to adhere for 20 seconds. The proteins were fixed with 3 μL of 0.8% glutaraldehyde before being washed twice with 3 μL of water. The proteins were negative stained with 8 μL of 0.7% uranyl acetate for 2 minutes before being air-dried. Samples were examined in a Morgagni 286D TEM (FEI company, OR, USA) operating at 80 kV and fitted with a 40 μm objective aperture.

8.15 Absorbance measurements [26, 28]

The A_{340} was measured in a CARY1 UV-visible spectrophotometer (Varian, Melbourne, Australia). The final buffer solution was used to zero the instrument. The A_{340} of the protein in Solution 1 before assembly was deducted from the A_{340} value.

8.16 Partial chymotrypsin digestion [29]

Proteins that had been extracted and fractionated using the acid fractionation were dissolved in 0.01 M sodium tetraborate (7.64 mg/2124 μL). 76 μL of 0.15 M $CaCl_2$ was added before 76 μL of chymotrypsin/trypsin inhibitor (2 mg chymotrypsin, 0.5 mg trypsin inhibitor in 10 mL), which had been at 37°C for 10 minutes, was added. The protein digestion solution was placed in a shaking water bath at 37°C for 5.5 hours and the pH was monitored to ensure a pH of 8.6 was constantly maintained. To terminate the reaction the pH was dropped to 4 with HAc. The solution was centrifuged for 10 minutes and the resulting precipitate was redissolved in 1.15 mL of 0.05 M sodium tetraborate. 42.9 mg of KCl was added to bring the solution to 0.5 M KCl.

The pH was once again reduced to 4 with HAc and centrifuged. The pellet was dissolved in 0.05 M sodium tetraborate and shaken overnight.

8.17 Assays

8.17.1 Ellman's assay

Cysteine standards were prepared (Table 8.6). 250 µL of each unknown/standard was added to separate tubes containing 2.5 mL of reaction buffer (0.1 M sodium phosphate, pH 8 containing 1 mM EDTA) and 50 µL of Ellman's reagent solution (4 mg of Ellman's reagent in 1 mL of reaction buffer). After mixing, the tubes were incubated at room temperature for 15 minutes before measuring the absorbance at 412 nm. Cysteine % (w/w) was calculated using the standard curve.

8.17.2 Congo Red Assay [30, 31]

A stock solution of Congo red (1 mM) was prepared in a phosphate buffer (0.01 M Na_2HPO_4, 0.0027 M KCl, 0.137 M NaCl, pH 7.4), containing 10% EtOH. This solution was filtered three times through Gelman extra-thick glass fibre filters before use, and stored at room temperature protected from light. The Congo red solution was diluted with phosphate buffer immediately prior to use in the assay to give a final Congo red concentration of approximately 10 µM, in the protein sample solution to be analysed. The protein concentration of the samples to be tested was between 10 – 20 µM. All samples were analysed in duplicate. For the assay, the Congo red solution and protein solution were mixed in a 1 mL plastic cuvette (Sarstedt, Germany) and allowed to stand at room temperature for 30 min prior to spectral analysis. The spectrum of the resulting solution was measured between 300 - 700 nm. The spectrum of phosphate buffer only, phosphate buffer with Congo red, and phosphate buffer with protein were measured as experimental controls. The spectral measurements were taken using a HP-8452A diode array spectrophotometer set in the wavelength-scanning mode to read from 300-700 nm.

Standard	Volume of Reaction buffer	Amount of Cysteine	Final Concentration
A	100 mL	26.3 mg	1.5 mM
B	5 mL	25 mL of Standard A	1.25 mM
C	10 mL	20 mL of Standard A	1.0 mM
D	15 mL	15 mL of Standard A	0.75 mM
E	20 mL	10 mL of Standard A	0.5 mM
F	25 mL	5 mL of Standard A	0.25 mM
G	30 mL	0 mL	0.0 mM (Blank)

Table 8.6
Cysteine standards preparation for the standard curve for the Ellman's assay.

8.17.3 Thioflavin T (ThT) Assay [32]

The protein sample to be analysed was added to Tris buffer (50 mM Tris, 100 mM NaCl, pH 7.5) containing ThT (5 or 10 μM) The final concentration of the protein in the 3 mL quartz cuvette was typically 20 μg/mL. The solution was mixed thoroughly and left to stand at room temperature for 3 min to allow binding between the dye and protein to equilibrate, prior to recording the emission spectrum. The spectrum of buffer only, buffer with ThT, and buffer with protein were measured as the experimental controls. Emission spectra for all samples were assayed in duplicate. The fluorescence spectrophotometer was programmed to measure the wavelength of excitation at 450 nm, and the wavelength of emission over a spectral range (typically from 470 - 540 nm) that included the wavelength of 482 nm. Both excitation and emission slits were set to 5 nm.

8.18 References

1. Sabounchi-Schütt F, Åström J, Olsson I, Eklund A, Grunewald J, Bjellqvist B. An Immobiline DryStrip Application Method Enabling High-Capacity Two-Dimensional Gel Electrophoresis. *Electrophoresis* **2000**;21:3649-3656
2. Görg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W. The Current State of Two-Dimensional Electrophoresis with Immobilized pH Gradients. *Electrophoresis* **2000**;21:1037-1053
3. Görg A, Boguth G, Obermaier C, Posch A, Weiss W. Two-Dimensional Polyacrylamide Gel Electrophoresis with Immobilized pH Gradients in the First Dimension (IPG-Dalt): The State of the Art and the Controversy of Vertical *versus* Horizontal Systems. *Electrophoresis* **1995**;16:1079-1086
4. Weistermeiner R, Naven T. *Proteomics in Practice*. Freiburg: Wiley-VCH Verlag-GmbH Weinheim, **2002**:316
5. Scheler C, Lamer S, Pan Z, Li X-P, Salnikow J, Jungblut P. Peptide Mass Fingerprint Sequence Coverage from Differently Stained Proteins on Two-Dimensional Electrophoresis Patterns by Matrix Assisted Laser Desorption/Ionization-Mass Spectrometry (MALDI-MS). *Electrophoresis* **1998**;19:918-927
6. Castellanos-Serra L, Proenza W, Huerta V, Moritz RL, Simpson RJ. Proteome Analysis of Polyacrylamide Gel-Separated Proteins Visualized by Reversible Negative Staining using Imidazole-Zinc Salts. *Electrophoresis* **1999**;20:732-737
7. Castellanos-Serra L, Carlos F-P, Eugenio H, Huerta V. A Procedure for Protein Elution from Reverse-Stained Polyacrylamide Gels Applicable at the Low Picomole Level: An Alternative Route to the Preparation of Low Abundance Proteins for Microanalysis. *Electrophoresis* **1996**;17:1564-1572
8. Blum H, Beier H, Gross HJ. Improved Silver Staining of Plant Proteins, RNA and DNA in Polyacrylamide Gels. *Electrophoresis* **1987**;8:93-99

9. Wong C, Sridhara S, Bardwell JCA, Jakob U. Heating Greatly Speeds Coomassie Blue Staining and Destaining. *BioTechniques* **2000**;28:426-432

10. Patestos NP, Fauth M, Radola BJ. Fast and Sensitive Protein Staining with Colloidal Acid Violet 17 Following Isoelectric Focusing in Carrier Ampholyte Generated and Immobilized pH Gradients. *Electrophoresis* **1988**;9:488-496

11. Debruyne I. Staining of Alkali-Labile Phosphoproteins and Alkaline Phosphatases on Polyacrylamide Gels. *Analytical Biochemistry* **1983**;133:110-115

12. Racusen D. Glycoprotein Detection in Polyacrylamide Gel with Thymol and Sulfuric Acid. *Analytical Biochemistry* **1979**;99:474-476

13. Beardmore T, Wetzel S, Burgess D, Charest PJ. Characterization of Seed Storage Proteins in *Populus* and their Homology with *Populus* Vegetative Storage Proteins. *Tree Physiology* **1996**;16:833-840

14. Jay GD, Culp DJ, Jahnke MR. Silver Staining of Extensively Glycosylated Proteins on Sodium Dodecyl Sulfate-Polyacrylamide Gels: Enhancement by Carbohydrate-Binding Dyes. *Analytical Biochemistry* **1990**;185:324-330

15. Yamagata A, Kristensen DB, Takeda Y, Miyamoto Y, Okada K, Inamatsu M, Yoshizato K. Mapping of Phosphorylated Proteins on Two-Dimensional Polyacrylamide Gels Using Protein Phosphatase. *Proteomics* **2002**;2:1267-1276

16. Hakimuddin T, Sojar T, Bahl OP. Chemical Deglycosylation of Glycoproteins. *Methods in Enzymology* **1987**;27:341-350

17. Harvey DJ. Identification of Protein-Bound Carbohydrates by Mass Spectrometry. *Proteomics* **2001**;1:311-328

18. Gobom J, Nordhoff E, Mirgorodskaya E, Ekmann R, Roepstorff P. Sample Purification and Preparation Technique Based on Nano-Scale Reversed-Phase Columns for the Sensitive Analysis of Complex Peptide Mixtures by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Journal of Mass Spectrometry* **1999**;34:105-116

19. Larsen MR, Cordwell SJ, Roepstorff P. Graphite Powder as an Alternative or Supplement to Reversed-Phase Material for Desalting and Concentration of Peptide Mixtures Prior to MALDI Mass Spectrometry. *Proteomics* **2002**;2:1277-1287

20. Kawasaki N, Ohta M, Hyuga S, Hashimoto O, Hayakawa T. Analysis of Carbohydrate Heterogeneity in a Glycoprotein Using Liquid Chromatography/Mass Spectrometry and Liquid Chromatography with Tandem Mass Spectrometry. *Analytical Biochemistry* **1999**;269:297-303

21. *Waters 616/626 LC System User's Guide*, Revision 1 ed. MA, USA: Millipore Corporation, **1994**

22. Thomas H, Greven R, Spei M. Experiments for the Isolation of the Matrix Proteins of Wool in Disulphide Form. *Melliand Textilberichte* **1983**;4:297-299

23. Kon R, Nakamura A, Hirabayashi N, Takeuchi K. Analysis of the Damaged Components of Permed Hair Using Biochemical Technique. *Journal of Cosmetic Science* **1998**;49:13-22

24. Thomas H, Conrads A, Phan K-H, van de Löcht M, Zahn H. *In vitro* Reconstitution of Wool Intermediate Filaments. *International Journal of Biological Macromolecules* **1986**;8:258-264

25. AmershamBiosciences. Isoelectric Membrane Formulas for IsoPrime® Purification of Proteins. In: *Protocol Guide #1 IsoPrime®*, **1997**

26. Wang H, Parry DAD, Jones LN, Idler WW, Marekov LN, Steinert PM. *In vitro* Assembly and Structure of Trichocyte Keratin Intermediate Filaments: A Novel Role for Stabilization by Disulfide Bonding. *Journal of Cell Biology* **2000**;151:1459-1468

27. Herrmann H, Häner M, Brettel M, Müller SA, Goldie KN, Fedtke B, Lustig A, Franke WW, Aebi U. Structure and Assembly Properties of the Intermediate Filament Protein Vimentin: The Role of its Head, Rod and Tail Domains. *Journal of Molecular Biology* **1996**;264:933-953

28. Stromer MH, Ritter MA, Pang S, Robson RM. Effect of Cations and Temperature on Kinetics of Desmin Assembly. *Biochemical Journal* **1987**;246:75-81

29. Crewther WG, Dowling LM. The Preparation and Properties of Large Peptides from the Helical Regions of the Low-Sulfur Proteins of Wool. *Applied Polymer Symposium* **1971**;18:1-20

30. Klunk WE, Pettegrew JW, Abraham DJ. Two Simple Methods for Quantifying Low Affinity Dye-Substrate Binding. *Journal of Histochemical Cytochemistry* **1989**;37:1273-1281

31. Klunk WE, Pettegrew JW, Abraham DJ. Quantititive Evaluation of Congo Red Binding to Amyloid-Like Proteins with a Beta-Pleated Sheet Conformation. *Journal of Histochemical Cytochemistry* **1989**;37:1393-1297

32. LeVine III H. Quantification of Beta-Sheet Amyloid Fibril Structures with Thioflavin T. *Methods in Enzymology* **1999**;309:274-284

Appendix One

A1.1 Peptide mass fingerprint match to α-casein

Matched peptides shown in **Bold Red**

1 MKLLILTCLV AVALARPKHP IKHQGLPQEV LNENLLR**FFV APFPEVFGK**E

51 KVNELSKDIG SESTEDQAME DIKQMEAESI SSSEEIVPNS VEQKHIQK**ED**

101 **VPSERYLGYL EQLLR**LKKYK VPQLEIVPNS AEERLHSMK**E GIHAQQK**EPM

151 IGVNQELAYF YPELFRQFYQ LDAYPSGAWY YVPLGTQYTD APSFSDIPNP

201 IGSENSEK**TT MPLW**

A1.2 Experimental and theoretical masses of peptides that matched to α-casein

Start-End	Mr (experimental)	Mr (calculated)	Sequence
38-49	1383.75	1383.72	FFVAPFPEVFGK
99-105	830.35	830.38	EDVPSER
106-115	1266.73	1266.70	YLGYLEQLLR
140-147	909.49	909.47	EGIHAQQK
209-214	747.37	747.36	TTMPLW

A1.3 Precursor ion scan peak list for the loss of 79 Da (% max intensity is the intensity (counts) of the peak as a percentage of the maximum peak intensity)

m/z	% max intensity	m/z	% max intensity	m/z	% max intensity
510	10.2564	613	11.7949	698.5	12.8205
515.5	10.7692	613.5	13.8462	700	10.7692
518	10.2564	615.5	14.8718	702.5	11.2821
518.5	11.7949	618	10.2564	703	10.2564
519	10.2564	618.5	11.2821	704	10.2564
520	11.7949	620	11.2821	711	10.2564
521.5	11.7949	621.5	10.2564	715	10.2564
523.5	10.2564	622	10.7692	729.5	10.2564
524	10.2564	625	12.8205	739	11.2821
525	11.7949	629.5	11.2821	766	11.2821
525.5	11.2821	630	10.2564	766.5	10.2564
526	12.3077	631.5	14.359	767	12.3077
527	11.2821	635.5	14.359	779.5	11.7949
528	16.9231	636	10.2564	780	12.3077
528.5	16.9231	637	11.2821	780.5	10.2564
529	18.4615	637.5	14.8718	783	10.2564

m/z	% max intensity	m/z	% max intensity	m/z	% max intensity
530.5	18.9744	639.5	13.8462	793.5	10.7692
531	15.8974	640	10.2564	794	15.3846
531.5	16.4103	640.5	12.8205	794.5	18.4615
532	11.7949	641	11.7949	795	20
533	12.8205	641.5	12.8205	795.5	13.3333
533.5	12.3077	642	10.7692	796	25.1282
534	15.3846	645	11.2821	796.5	18.4615
534.5	16.9231	645.5	11.2821	797	14.359
535	11.2821	646	24.6154	797.5	11.2821
535.5	15.8974	646.5	53.3333	799	11.2821
536	15.8974	647	74.8718	801.5	11.2821
536.5	11.7949	647.5	79.4872	804.5	11.7949
537	13.3333	648	100	805	15.3846
537.5	12.3077	648.5	84.1026	805.5	11.7949
538	14.359	649	83.5897	806	17.4359
538.5	12.8205	649.5	86.1538	806.5	12.8205
539	13.8462	650	80.5128	807	15.8974
539.5	15.8974	650.5	50.7692	807.5	11.2821
540	11.7949	651	21.5385	808	12.8205
540.5	15.3846	651.5	14.359	810	11.2821
541.5	13.8462	652	10.2564	810.5	10.7692
542	17.9487	653	14.359	811	11.7949
542.5	13.8462	654.5	15.3846	811.5	14.359
543	12.8205	655.5	18.4615	812	13.8462
543.5	11.7949	656	17.4359	812.5	14.359
544	12.3077	656.5	13.3333	813	12.8205
549	10.7692	658	10.2564	814	11.7949
550	14.359	659.5	14.359	818	11.2821
550.5	15.8974	661	11.7949	819.5	10.2564
551.5	12.8205	661.5	10.7692	848	10.2564
553	10.7692	663	11.7949	848.5	11.2821
554	11.2821	663.5	10.7692	905.5	12.3077
585.5	10.7692	666.5	13.8462	951.5	10.2564
586.5	12.8205	667	10.7692	954	10.2564
587	11.7949	667.5	10.2564	955.5	10.7692
589.5	10.2564	668	10.7692	960	11.2821
590	13.3333	669	16.9231	961	11.7949
590.5	12.3077	673	11.2821	961.5	10.7692
591.5	13.3333	673.5	12.8205	963	10.2564
592.5	12.3077	676	13.8462	963.5	10.7692
593.5	14.8718	676.5	20.5128	971	10.2564
594	12.3077	677	15.3846	971.5	20.5128
594.5	10.2564	677.5	13.8462	972	31.7949
595	11.7949	678	18.9744	972.5	43.0769
597	12.3077	679	12.3077	973	43.5897
597.5	11.2821	680.5	14.359	973.5	55.8974
598	11.7949	681	10.7692	974	43.0769
600	10.2564	682.5	15.8974	974.5	38.9744
600.5	14.359	683	15.3846	975	37.4359
601	15.8974	683.5	10.2564	975.5	32.3077
601.5	13.8462	684	11.7949	976	23.5897
602.5	15.8974	685	10.7692	976.5	10.2564
603	10.2564	685.5	12.3077	982.5	12.8205
603.5	14.359	686	18.4615	983	21.5385
604	12.3077	688	10.7692	983.5	26.6667
604.5	10.7692	688.5	11.7949	984	35.3846

m/z	% max intensity	m/z	% max intensity	m/z	% max intensity
605	16.9231	689	15.3846	984.5	25.641
605.5	14.8718	689.5	13.8462	985	37.9487
606.5	14.8718	690.5	15.8974	986	22.0513
607	20	691	17.4359	986.5	28.7179
607.5	21.5385	691.5	12.3077	987	17.4359
608	16.9231	692	14.359	987.5	10.7692
608.5	14.8718	692.5	13.3333	993.5	11.2821
609	11.2821	693	14.359	994	12.3077
609.5	11.2821	694	11.7949	994.5	14.359
610	13.8462	694.5	12.3077	995	14.359
610.5	14.359	695	16.4103	995.5	16.9231
611	14.8718	696	10.2564	996	15.8974
611.5	16.9231	696.5	12.8205	996.5	13.8462
612	11.2821	697	11.7949	997	17.9487
612.5	10.7692	697.5	10.7692	997.5	11.2821

A1.4 Product ion scans of α-casein

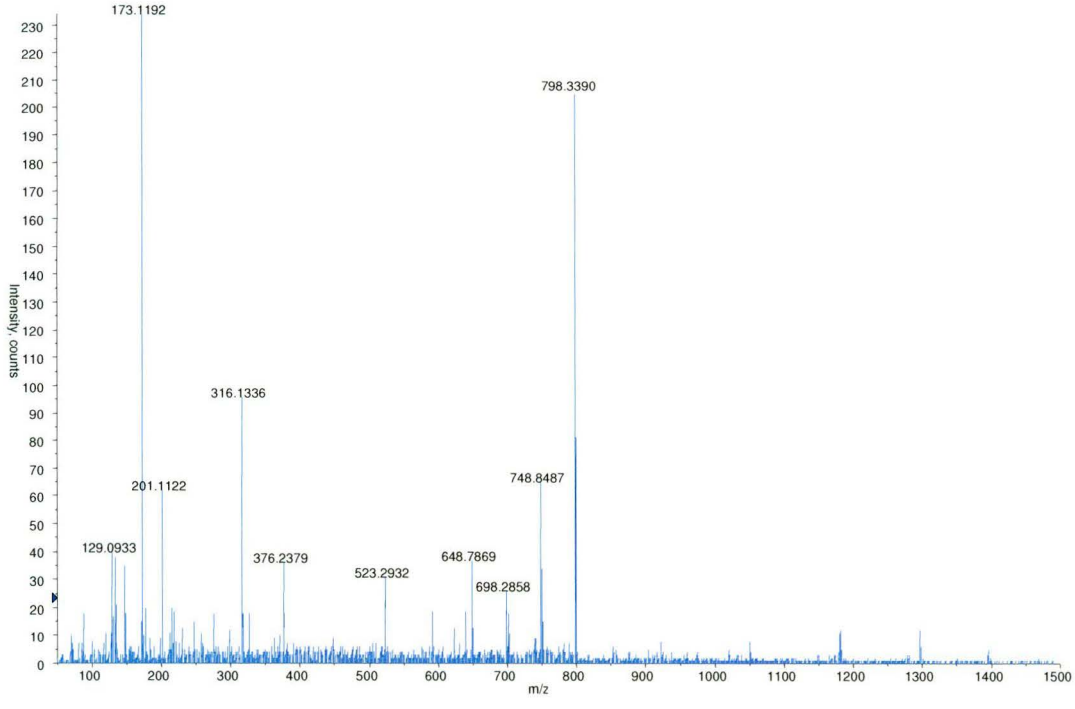


Figure A1.1
Product ion scan of the doubly charged peptide at m/z 798^{+2}. 102 nM was analysed in positive scanning mode. The doubly charged peptide was confirmed by *de novo* sequencing to be TVDME[pS]TEVFTKK. The spectrum is representative of triplicate experiments.

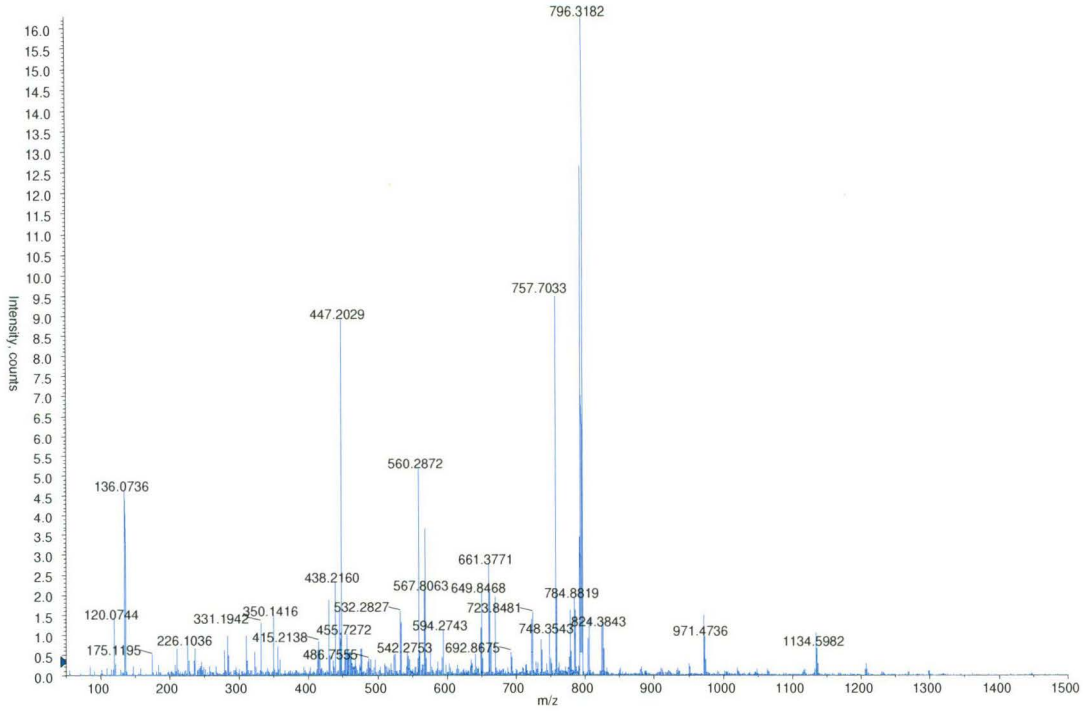


Figure A1.2
Product ion scan of the doubly charged peptide at m/z 796^{+2}. 102 nM was analysed in positive scanning mode. No phosphorylated amino acids were found. The spectrum is representative of triplicate experiments.

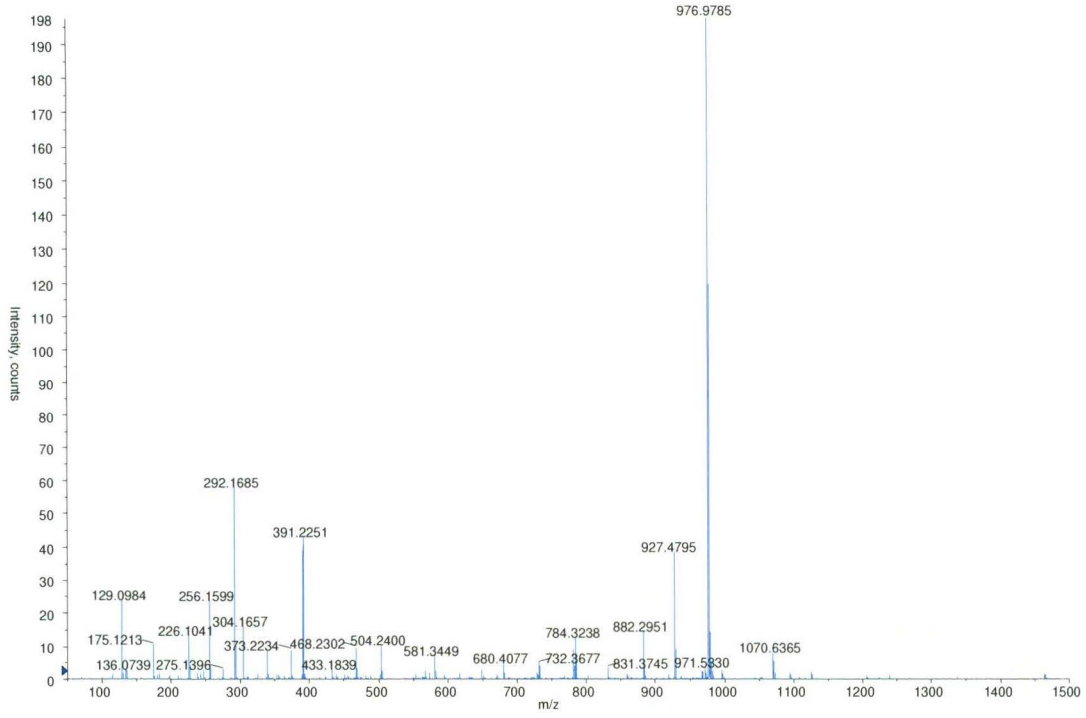


Figure A1.3
Product ion scan of the doubly charged peptide at m/z 976^{+2}. 102 nM was analysed in positive scanning mode. The doubly charged peptide was confirmed by *de novo* sequencing to be YVPQLELVPN[pS]A. The spectrum is representative of triplicate experiments.

A1.5 Intermediate filament protein precursor ion scans for the loss of 79 Da

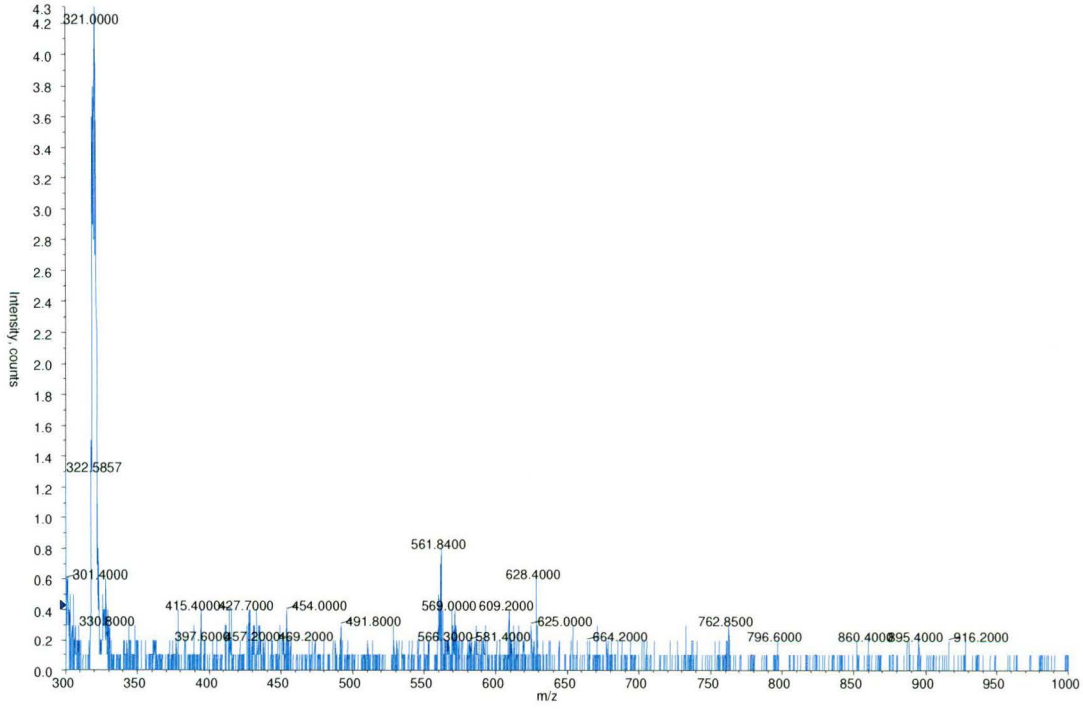


Figure A1.4
Precursor ion scan for the loss of 79 Da (PO_3^{2-}) from a wool IFP. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

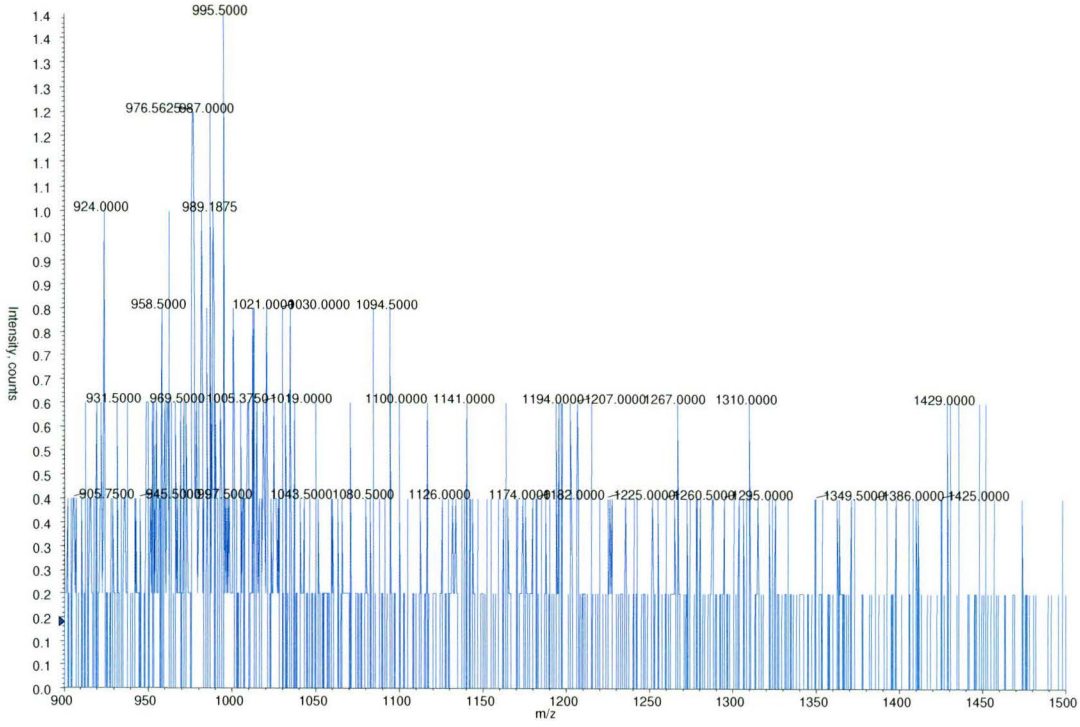


Figure A1.5
Precursor ion scan for the loss of 79 Da (PO_3^{2-}) from a wool IFP. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

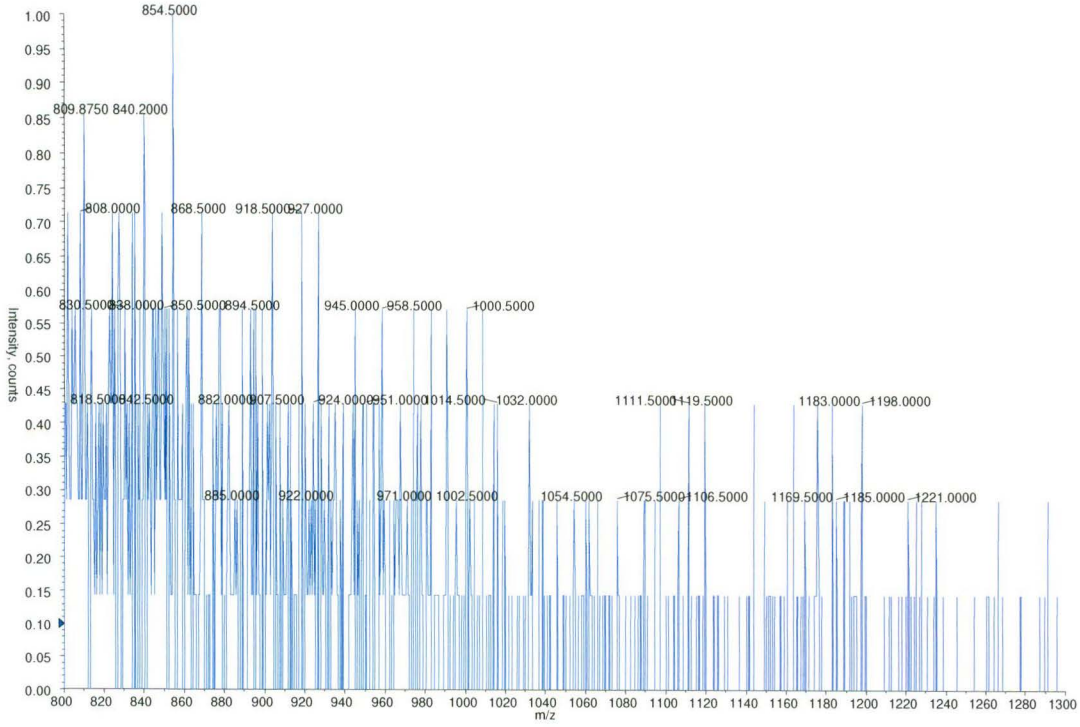


Figure A1.6
Precursor ion scan for the loss of 79 Da (PO_3^{2-}) from a wool IFP. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

A1.6 Product ion scans of wool IFPs

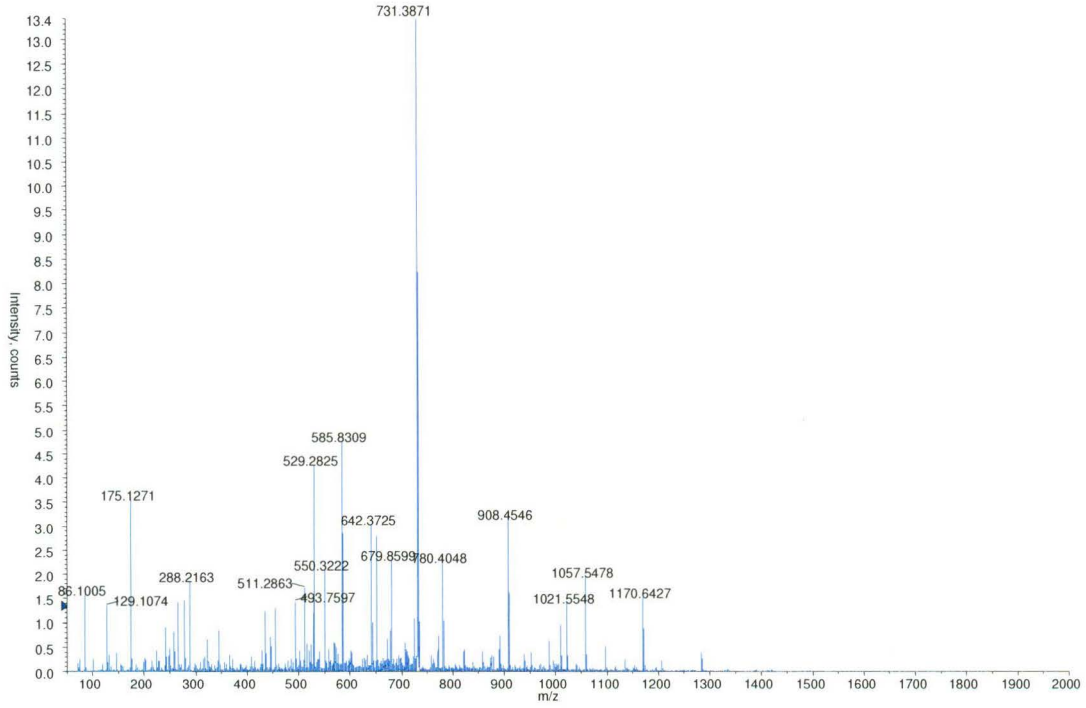


Figure A1.7
Product ion scan of the doubly charged peptide at m/z 731^{+2}. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

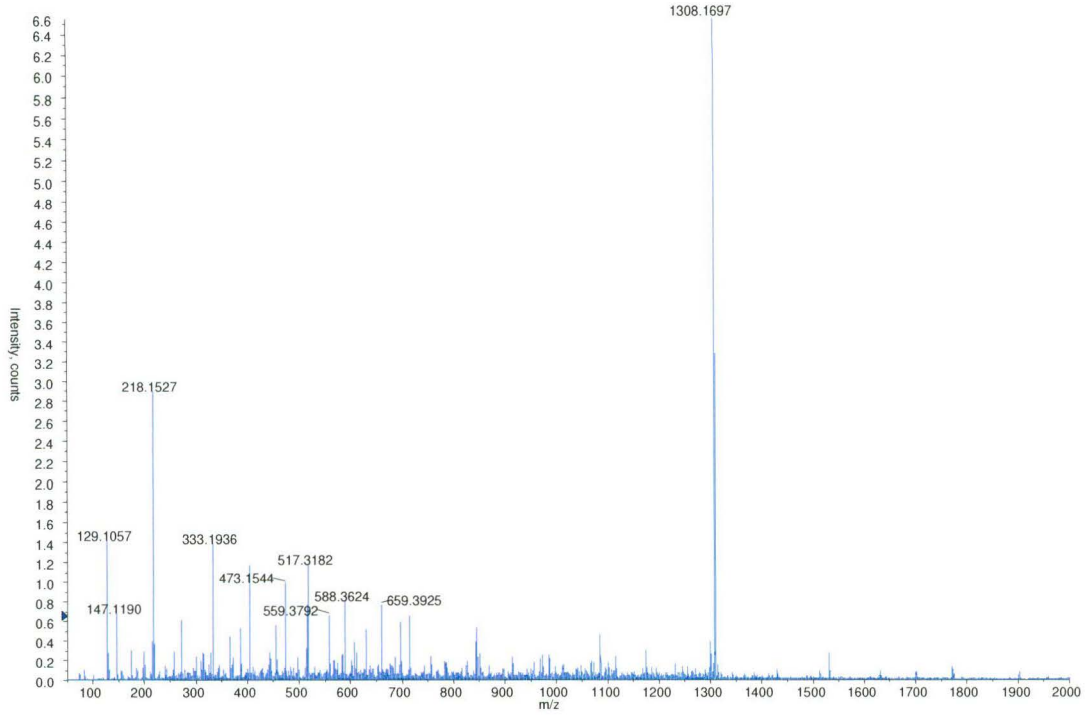


Figure A1.8
Product ion scan of the doubly charged peptide at m/z 1308^{+2}. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

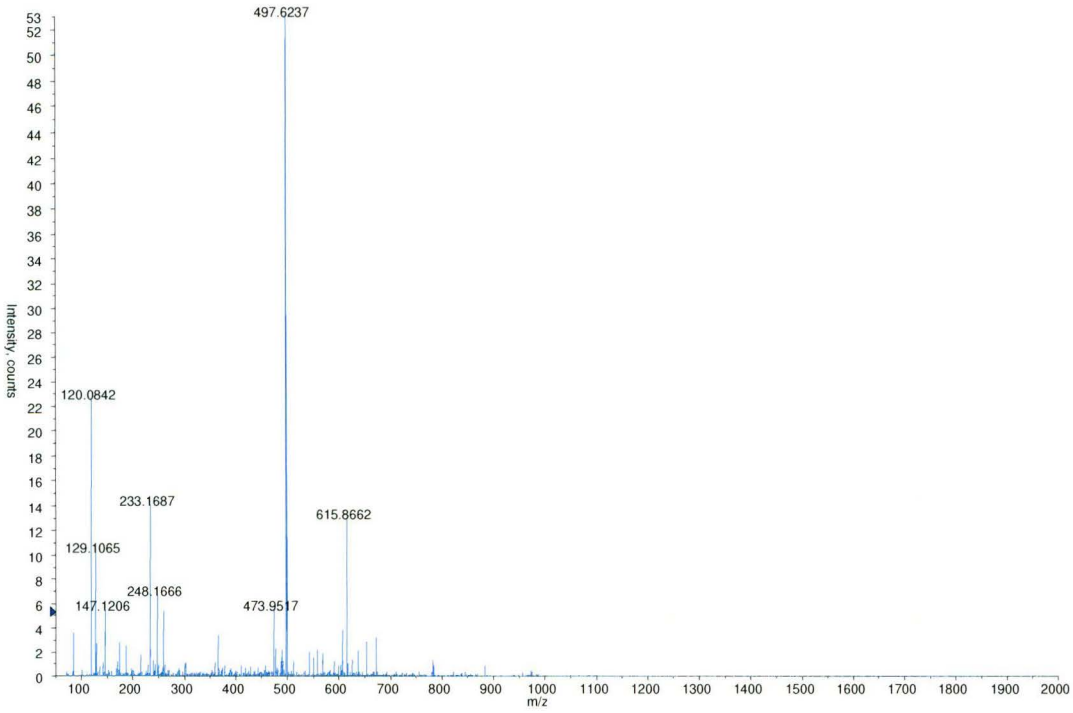


Figure A1.9
Product ion scan of the doubly charged peptide at m/z 497^{+2}. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

A1.7 Peptide mass fingerprint scans of wool IFPs

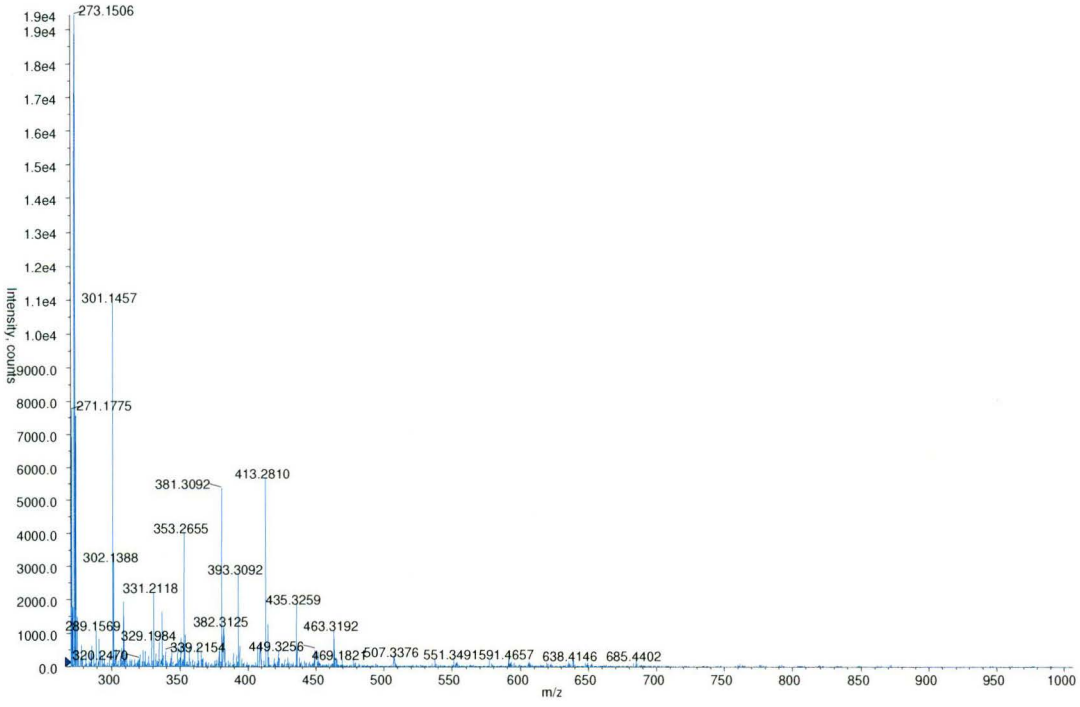


Figure A1.10
Peptide mass fingerprint of spot number 24 excised from a 2DE gel. A gel excised spot was analysed in negative scanning mode. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

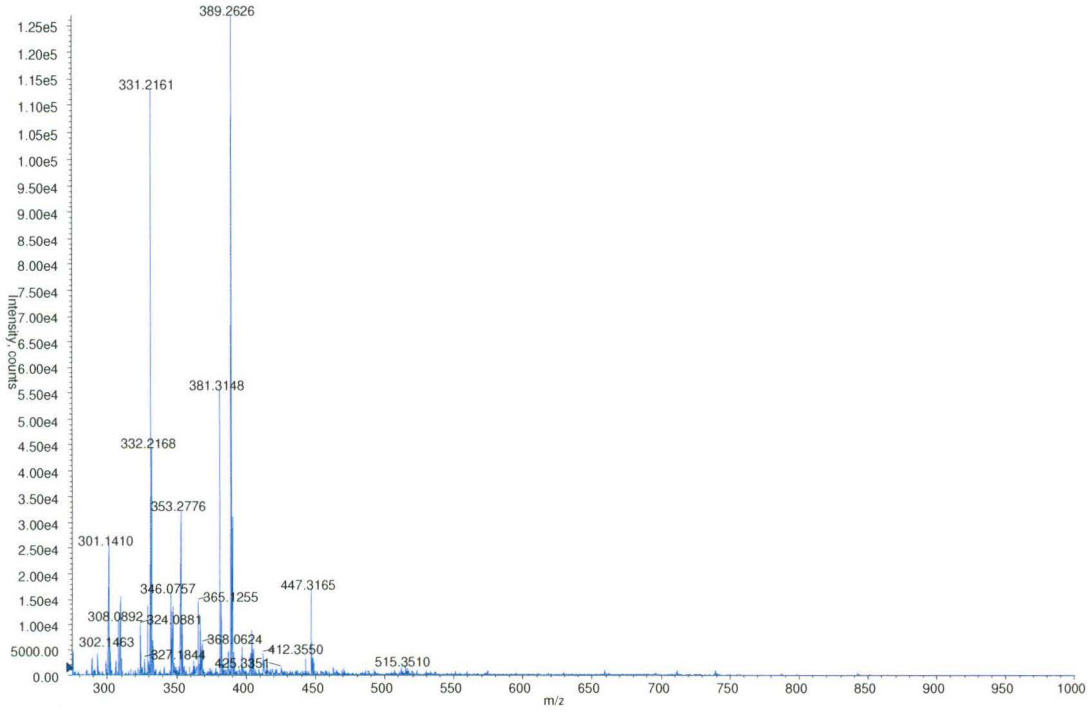


Figure A1.11
Peptide mass fingerprint of spot number 17 excised from a 2DE gel. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

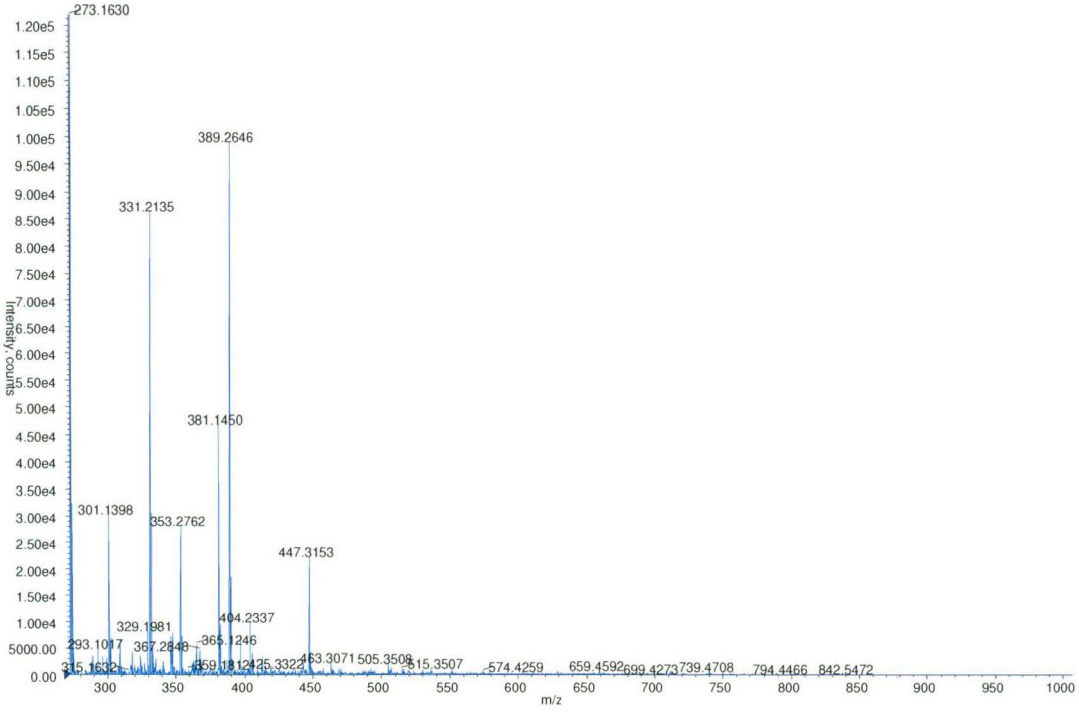


Figure A1.12
Peptide mass fingerprint of spot number 20 excised from a 2DE gel. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

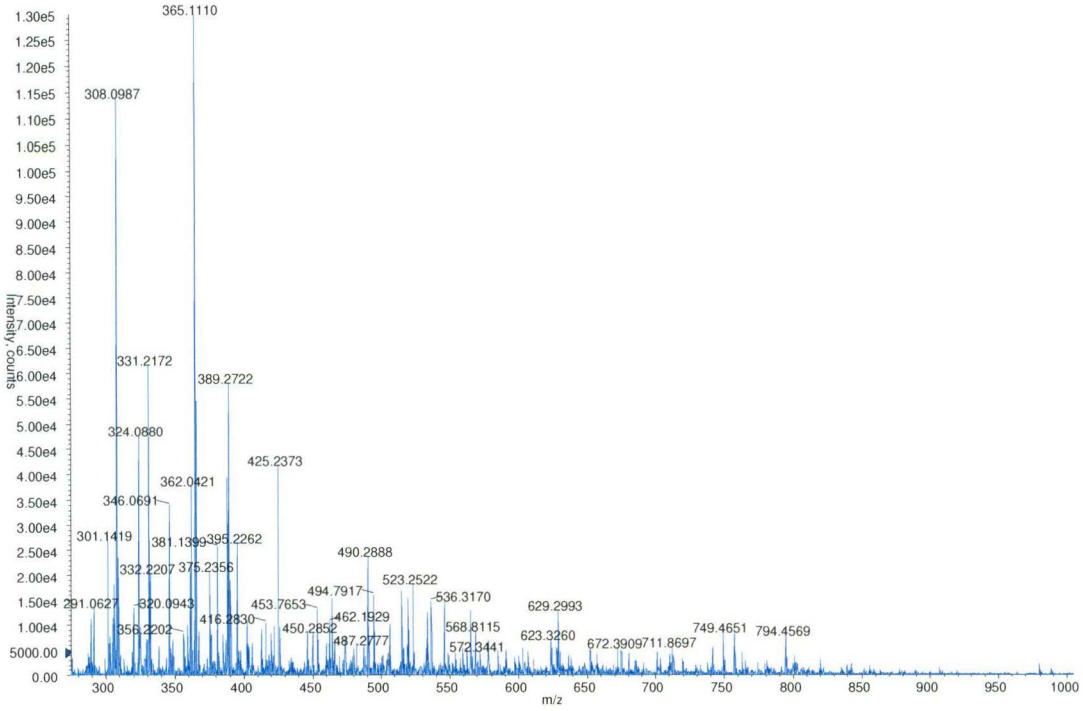


Figure A1.13
Peptide mass fingerprint of spot number 23 excised from a 2DE gel. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

Appendix Two

A2.1 Peptide mass fingerprint of a four hour digestion of GMP variant B

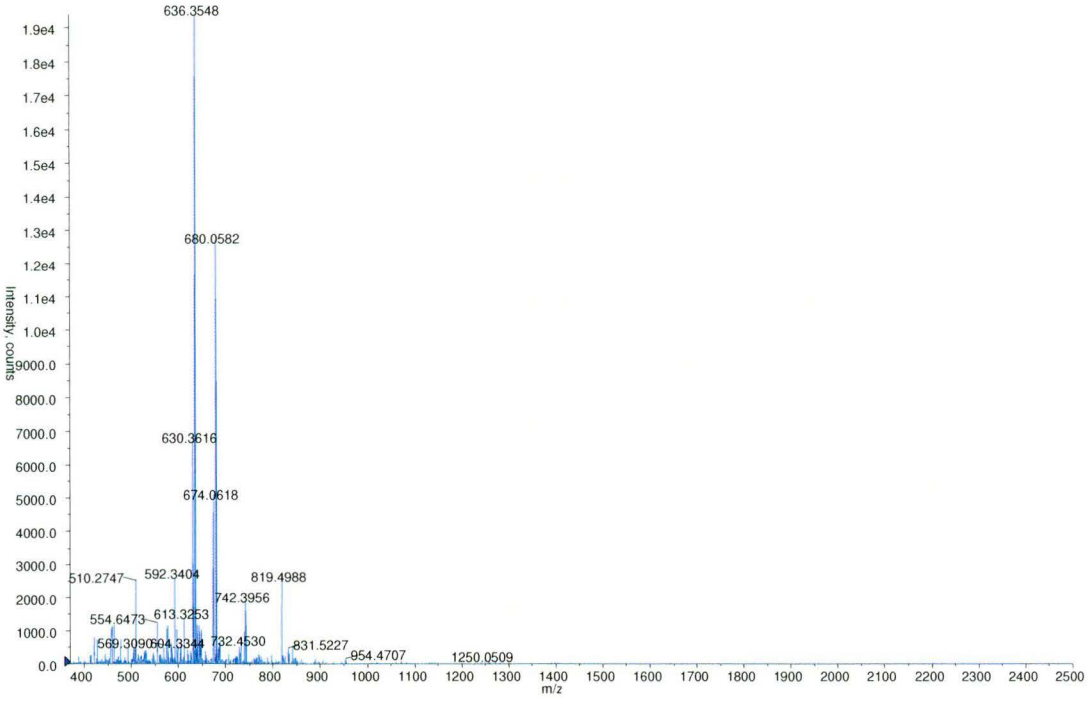


Figure A2.1
Peptide mass fingerprint of GMP variant B at a concentration of 14 μM digested with chymotrypsin for 4 hours. Spectrum is representative of triplicate experiments.

A2.2 Precursor ion scan peak list for the loss of 204 Da (% max intensity is the intensity (counts) of the peak as a percentage of the maximum peak intensity)

m/z	% max intensity	m/z	% max intensity	m/z	% max intensity
457	14.2857	1058	23.8095	1171	14.2857
463	14.2857	1059	38.0952	1172	19.0476
464	14.2857	1060	28.5714	1175	19.0476
465	14.2857	1061	23.8095	1177	14.2857
475	14.2857	1063	38.0952	1178	19.0476
477	14.2857	1066	28.5714	1180	14.2857
489	14.2857	1067	14.2857	1183	19.0476
501	14.2857	1068	14.2857	1184	14.2857
509	33.3333	1070	19.0476	1185	19.0476
510	14.2857	1071	14.2857	1193	19.0476
511	14.2857	1073	14.2857	1195	14.2857
562	14.2857	1074	14.2857	1196	14.2857
598	14.2857	1075	19.0476	1197	19.0476
629	19.0476	1077	28.5714	1198	19.0476
630	14.2857	1078	19.0476	1199	14.2857
634	14.2857	1080	28.5714	1201	19.0476
635	38.0952	1081	14.2857	1202	14.2857
636	33.3333	1082	14.2857	1203	14.2857
637	71.4286	1084	38.0952	1209	28.5714
638	14.2857	1086	14.2857	1210	14.2857
672	14.2857	1087	28.5714	1215	14.2857
675	23.8095	1088	38.0952	1216	14.2857
678	38.0952	1089	42.8571	1218	14.2857
679	90.4762	1090	14.2857	1219	23.8095
680	100	1091	19.0476	1221	28.5714
681	61.9048	1092	19.0476	1223	14.2857
682	14.2857	1093	23.8095	1224	19.0476
687	23.8095	1095	14.2857	1228	23.8095
688	19.0476	1096	14.2857	1232	23.8095
722	14.2857	1097	19.0476	1237	19.0476
758	14.2857	1102	38.0952	1243	19.0476
807	14.2857	1103	23.8095	1245	14.2857
983	19.0476	1104	19.0476	1246	14.2857
984	14.2857	1106	19.0476	1247	14.2857
985	14.2857	1111	14.2857	1248	19.0476
986	23.8095	1114	14.2857	1249	23.8095
987	14.2857	1115	19.0476	1251	14.2857
990	14.2857	1120	14.2857	1253	19.0476
992	14.2857	1123	19.0476	1254	19.0476
995	23.8095	1127	14.2857	1256	28.5714
1001	14.2857	1128	19.0476	1262	14.2857
1003	19.0476	1129	23.8095	1263	19.0476
1011	19.0476	1135	14.2857	1265	23.8095
1019	14.2857	1136	14.2857	1271	14.2857

m/z	% max intensity	m/z	% max intensity	m/z	% max intensity
1020	19.0476	1140	14.2857	1274	9.5238
1024	19.0476	1143	14.2857	1275	14.2857
1025	14.2857	1149	23.8095	1276	14.2857
1026	14.2857	1150	14.2857	1279	19.0476
1028	19.0476	1151	19.0476	1282	23.8095
1035	23.8095	1155	23.8095	1286	23.8095
1043	14.2857	1156	42.8571	1289	14.2857
1044	23.8095	1157	14.2857	1292	19.0476
1046	14.2857	1162	19.0476	1293	14.2857
1047	14.2857	1163	47.619	1294	23.8095
1049	14.2857	1164	14.2857	1297	14.2857
1053	23.8095	1165	19.0476	1298	14.2857
1055	23.8095	1166	19.0476	1299	14.2857
1057	14.2857	1169	14.2857	1300	14.2857

A2.3 Product ion scans of GMP variant B

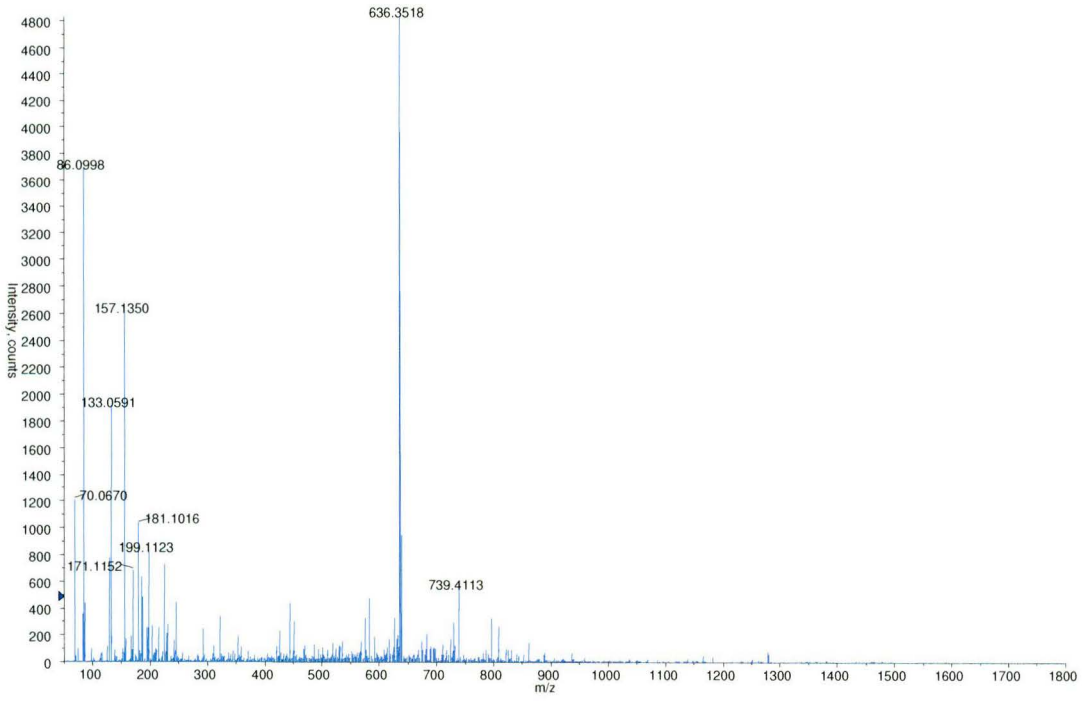


Figure A2.2
Product ion scan of the doubly charged peptide at m/z 636^{+2}. The spectrum is representative of triplicate experiments.

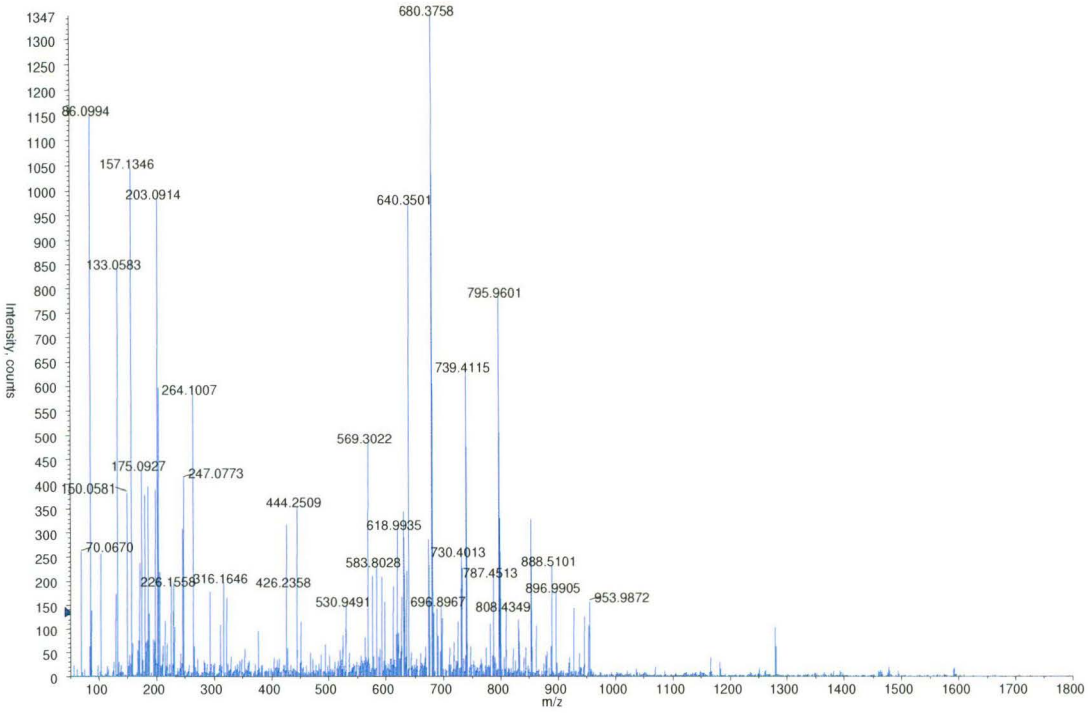


Figure A2.3
Product ion scan of the doubly charged peptide at m/z 680^{+2}. The spectrum is representative of triplicate experiments.

A2.4 Intermediate filament protein precursor ion scan for the loss of 204 Da

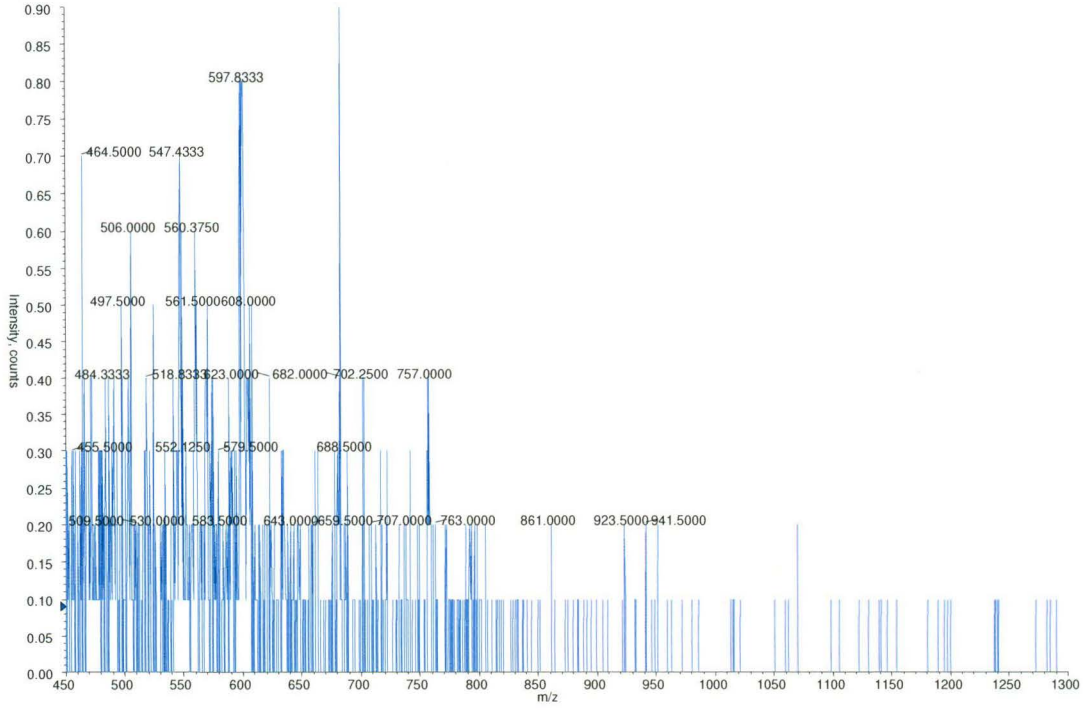


Figure A2.4
Precursor ion scan for the loss of 204 Da (N-acetylhexosamine) from a wool IFP. A gel excised spot was analysed in negative scanning mode. The spectrum is representative of triplicate experiments.

Appendix Three

A3.1 Peak list matching

A3.1.1 Peaks that matched – spots 10 and 11

m/z Spot 10	intensity	m/z Spot 11	intensity	m/z Spot 10	intensity	m/z Spot 11	intensity
400.1814	468	400.2314	55	552.7569	1225	552.7959	30
401.2079	727	401.2044	50	553.2497	1026	553.2679	59
402.2178	408	402.2502	71	560.2677	399	560.2865	42
403.2111	527	403.2436	80	573.2506	116	573.3346	57
404.1698	262	404.1934	49	592.2277	1171	592.271	728
404.959	163	404.9917	248	593.2265	336	593.2699	172
405.1924	419	405.1712	66	594.237	69	594.2805	33
406.1893	318	406.1952	33	599.7282	1432	599.7723	203
411.2015	78	411.2078	38	600.2306	908	600.2857	125
412.2149	1141	412.2122	79	604.3015	708	604.3461	323
413.2386	265	413.2178	43	605.3104	221	605.3441	72
415.1716	135	415.2054	44	606.2323	687	606.2661	127
417.2276	1281	417.2616	128	607.2318	267	607.2878	39
417.6922	968	417.7263	296	619.2571	1860	619.3033	757
418.1936	635	418.2278	172	620.2673	524	620.3025	189
419.2247	91	419.2133	30	621.2672	111	621.3136	31
420.2113	242	420.2549	37	628.2891	365	628.3028	39
422.1972	659	422.2318	51	632.2969	1349	632.3446	37
423.1965	170	423.2221	32	633.2841	1113	633.3094	75
424.1878	231	424.1859	48	639.2342	931	639.2713	41
425.1895	80	425.2244	30	647.2526	738	647.3018	31
427.1871	95	427.2131	29	675.2874	853	675.3398	509
428.22	132	428.2368	44	676.2959	1080	676.3484	387
430.1784	70	430.214	29	676.8063	493	676.8588	92
431.1964	912	431.232	127	677.3052	307	677.3577	71
442.2201	81	442.2194	31	678.2687	1003	678.3214	704
443.224	234	443.214	36	679.2678	294	679.3206	203
447.204	236	447.232	42	680.2677	118	680.309	36
450.7472	94	456.7411	79	683.2483	1103	683.3016	253
457.2078	219	457.2464	31	684.4027	74	684.3044	54
460.2354	725	460.236	32	688.3163	737	688.3584	29
470.2381	153	471.2264	36	704.3071	804	704.3389	34
472.2322	327	472.2821	33	706.2966	155	706.3524	29
473.2211	142	473.2518	31	706.83	643	706.8858	52
474.211	917	474.2515	82	707.328	501	707.3839	48
482.207	182	482.2289	34	710.2603	1801	710.3165	40

m/z Spot 10	intensity	m/z Spot 11	intensity	m/z Spot 10	intensity	m/z Spot 11	intensity
485.1982	297	485.2498	32	715.4751	175	715.4006	381
486.2006	320	486.213	47	715.8449	892	715.5676	34
487.1942	213	487.1968	30	716.346	394	716.4029	95
488.2183	152	488.2309	33	725.3371	998	725.383	213
491.2182	182	491.2705	33	725.8416	754	725.8875	160
493.2067	123	493.19	62	726.3342	279	726.3922	60
494.7429	522	494.7758	30	726.7669	1348	726.813	40
495.229	303	495.262	39	731.3543	1205	731.4128	689
500.233	158	500.2366	34	732.3555	451	732.4142	212
502.1997	241	502.2334	34	733.3333	159	733.4163	50
504.2203	1211	504.2542	75	737.3358	208	737.3708	72
505.2321	464	505.2461	45	737.8202	148	737.8916	57
507.2588	613	507.3031	54	738.9956	290	738.9095	149
509.1788	211	509.2032	31	739.3471	1607	739.3944	127
513.6967	57	513.7417	33	739.8201	1774	739.9038	70
514.2223	190	514.2573	29	740.3174	1332	740.3891	42
515.2846	322	515.3096	150	770.2918	455	770.3547	29
516.2872	193	516.3021	39	813.8099	649	813.8777	568
519.2094	1139	519.2347	52	814.306	431	814.3739	440
520.1955	695	520.2616	38	814.8024	207	814.8703	242
528.2303	625	528.2668	141	815.3243	66	815.3668	74
529.2454	187	529.282	46	823.4168	448	823.4857	207
539.2326	70	539.2806	32	824.4152	145	824.4842	72
543.2058	554	543.2646	33	844.367	58	844.4383	30
545.2551	1681	545.2934	498	861.4005	121	861.4736	79
546.2447	434	546.3039	106	947.4025	108	947.4856	66

A3.1.2 Peaks that did not match – spots 10 and 11

m/z Spot 11	intensity
406.9865	29
450.2171	45
479.2467	37
483.189	29
526.2598	37
527.2526	30
600.395	31
600.7884	36
715.9017	325
716.0449	29
731.5576	90
814.0049	92
814.5266	92
988.5527	56
1013.569	29

A3.1.3 Peaks that matched – spots 23 and 24

m/z Spot 23	intensity	m/z Spot 24	intensity	m/z Spot 23	intensity	m/z Spot 24	intensity
400.2225	135	400.2314	59	549.7699	120	549.7595	98
401.2134	113	401.2223	42	550.2615	88	550.251	59
402.2502	376	402.2591	228	551.2767	30	551.2348	13
403.2346	130	403.2615	65	552.2823	19	552.3033	20
404.2562	60	404.2472	34	555.2837	29	555.2522	15
405.2251	70	405.2161	39	556.293	34	556.314	18
406.2311	162	406.2222	92	557.2821	36	557.261	29
407.2205	95	407.2025	45	558.2932	25	558.2932	18
408.2019	65	408.2019	36	559.3685	291	559.3474	198
409.2207	77	409.2117	28	560.3815	61	560.3604	45
410.2227	52	410.1865	20	561.3108	35	561.3214	18
411.1987	68	411.1626	17	562.3573	25	562.3044	21
412.2303	57	412.2303	25	563.2989	19	563.2671	17
413.2631	176	413.2631	44	564.3048	32	564.2942	44
414.2246	94	414.2519	31	564.7922	17	564.7922	16
415.2327	84	415.1873	32	565.3329	176	565.3223	171

m/z Spot 23	intensity	m/z Spot 24	intensity	m/z Spot 23	intensity	m/z Spot 24	intensity
416.2693	253	416.2693	128	565.842	23	565.8207	39
417.2252	130	417.2616	50	566.3194	63	566.2982	36
418.2278	102	418.2187	44	567.2856	23	567.2856	16
419.2042	84	419.2407	31	568.295	18	568.2632	17
420.2183	455	420.2274	229	568.8267	201	568.8161	177
421.0689	87	421.0597	16	569.3267	140	569.3054	107
421.2153	122	421.2428	58	569.8376	16	569.8269	16
421.7646	122	421.7738	43	570.3061	48	570.3061	29
422.2685	100	422.2685	41	571.3609	53	571.3502	27
423.2313	80	423.2404	31	573.3346	60	573.3346	31
424.2043	59	424.2135	23	574.3495	56	574.3281	23
425.2152	1343	425.2244	676	575.3224	44	575.3117	38
426.2182	282	426.2274	104	576.2855	35	576.2962	20
427.2223	100	427.2315	24	577.2923	40	577.2816	19
428.2368	89	428.2645	28	577.796	430	577.7853	578
428.7722	84	428.7815	13	578.2999	229	578.2999	345
429.2156	75	429.2341	22	578.804	62	578.804	93
429.3172	548	429.3265	143	579.2977	40	579.2869	32
430.3065	126	430.3157	24	580.35	22	580.2856	13
431.2228	92	431.2506	28	581.2958	32	581.3281	27
432.2142	95	432.242	25	582.35	16	582.3069	17
433.2068	253	433.2439	47	583.2651	18	583.2974	14
433.736	219	433.736	29	584.321	20	584.2887	16
434.2098	317	434.2191	65	585.2592	180	585.27	122
435.2233	194	435.2326	46	586.2522	63	586.2846	26
435.7724	258	435.8003	50	587.2784	31	587.2568	16
436.2752	187	436.2845	34	588.3055	26	588.3163	17
437.2258	96	437.2258	33	589.3011	36	589.3011	36
438.2428	100	438.2335	24	590.3191	30	590.3408	21
439.2236	84	439.261	20	591.273	60	591.2947	56
440.2616	74	440.2991	15	592.2602	30	592.271	19
441.2634	93	441.2634	21	593.2808	27	593.3459	18
442.2663	70	442.2382	23	594.3131	32	594.3131	16
443.2422	109	443.2422	25	595.3137	17	595.3029	13
444.2098	167	444.2474	43	596.2498	20	596.2825	16
445.2349	115	445.2537	36	597.2848	50	597.3066	28
446.1858	802	446.1952	250	598.277	28	598.3097	18
447.1942	149	447.1942	52	599.27	32	599.3027	22
448.2322	119	448.2511	23	600.2748	439	600.2857	504
449.2241	171	449.2241	42	600.7775	187	600.7775	249
450.2644	187	450.2833	32	601.2804	101	601.3022	86
451.268	111	451.2585	21	602.3525	30	602.4072	13
452.2253	64	452.2348	14	603.3489	360	603.3489	260
453.2216	123	453.2596	20	604.3351	117	604.3242	65
453.7439	738	453.7629	257	605.3112	36	605.3222	18
454.2475	436	454.2475	128	606.3211	41	606.3211	24
454.7419	66	454.7514	18	607.2658	231	607.2658	227
455.2365	139	455.2651	26	608.2772	61	608.2662	45
456.1981	113	456.2171	47	609.3005	26	609.2675	20
457.2273	126	457.2941	30	610.3137	30	610.3247	20

m/z Spot 23	intensity	m/z Spot 24	intensity	m/z Spot 23	intensity	m/z Spot 24	intensity
458.2578	112	458.2482	19	610.8316	39	610.8426	34
459.2702	80	459.232	18	611.3276	58	611.3276	40
460.2838	216	460.2647	38	612.3094	22	612.2873	15
461.2507	148	461.2411	26	613.336	30	613.3471	25
462.2569	126	462.1993	26	614.3304	26	614.3083	13
463.2161	464	463.2353	104	615.4252	30	615.3699	16
464.1957	2417	464.2053	800	616.3105	26	616.2995	18
465.1954	583	465.2243	138	617.3184	22	617.2962	15
466.2444	541	466.2637	111	618.316	37	618.3382	22
467.2464	193	467.2271	38	619.2922	29	619.3144	18
468.2302	58	468.2205	17	620.2692	21	620.3247	15
469.2439	83	469.2246	17	622.27	47	622.27	69
470.2201	77	470.2685	29	623.3384	245	623.3384	328
471.2651	73	471.2457	19	623.8395	92	623.8395	158
472.2918	362	472.2918	66	623.962	22	623.9397	19
473.2906	117	473.2712	30	624.3297	203	624.3185	211
474.2807	94	474.2418	29	624.4523	45	624.4634	19
475.2232	93	475.2621	23	625.3218	56	625.3329	45
476.2542	67	476.2542	24	628.3252	65	628.314	50
477.2572	64	477.2084	14	628.8954	16	628.8059	34
478.2514	69	478.2709	28	629.3092	219	629.3092	229
479.2564	329	479.2662	82	630.3165	69	630.2941	61
480.2625	141	480.2527	35	630.3948	34	630.4508	15
481.2403	176	481.2501	41	631.3805	30	631.3469	21
482.2484	122	482.268	34	632.3446	23	632.3222	16
483.2381	72	483.2087	31	633.2645	73	633.2869	46
484.2876	59	484.2679	21	633.3879	33	633.4104	18
485.24	51	485.2105	20	634.2973	73	634.2973	55
486.272	68	486.272	35	635.2973	36	635.2748	23
487.2657	409	487.2559	189	636.3543	144	636.3431	140
487.7679	121	487.7777	70	636.4443	33	636.4556	23
488.3294	200	488.2703	80	637.3334	55	637.3672	31
489.2858	166	489.2759	53	638.3245	32	638.302	17
490.2727	575	490.2629	216	639.2939	23	639.339	23
490.7765	192	490.7666	44	640.4219	29	640.4445	17
491.231	189	491.231	101	641.3363	23	641.3363	15
491.399	18	491.3496	36	642.3645	20	642.3306	18
492.2397	61	492.2496	28	644.3892	27	644.3553	24
493.2494	51	492.3584	14	645.3518	33	645.3631	28
494.27	71	494.2601	24	646.3151	68	646.3264	73
494.7758	565	494.7758	254	647.3586	18	647.3245	21
495.2818	353	495.2818	140	648.3461	18	648.312	16
495.7782	56	495.7881	27	650.403	65	650.4143	37
496.2748	58	496.2748	22	650.9262	26	650.8921	23
497.2687	49	497.2488	21	652.372	24	652.3492	23
498.2438	82	498.2537	30	653.2607	57	653.2607	55
499.2596	95	499.2497	47	654.2527	20	654.2755	13
500.2865	109	500.2466	50	656.319	19	656.3075	14
501.2944	74	501.2645	35	657.3704	16	657.3361	21
502.3034	76	502.2834	40	659.3729	19	659.3729	13

m/z Spot 23	intensity	m/z Spot 24	intensity	m/z Spot 23	intensity	m/z Spot 24	intensity
503.2933	67	503.2533	27	660.3466	35	660.3351	17
504.3043	54	504.2542	25	664.3406	57	664.3406	50
505.2461	117	505.226	52	665.3294	19	665.341	16
506.249	214	506.239	80	666.3651	17	666.3881	16
506.7508	47	506.7508	19	672.3994	85	672.4109	66
507.2629	80	507.2428	29	673.3942	24	673.3363	14
508.2477	53	508.2879	19	674.8879	102	674.8763	156
509.2636	52	509.2636	20	675.0037	16	675.0037	18
511.2582	74	511.238	18	675.3861	99	675.3745	142
512.2468	37	512.2266	15	675.4904	26	675.5484	13
512.3478	16	512.3175	15	675.9077	27	675.8845	42
513.2364	33	513.2768	24	676.3368	28	676.3948	27
514.2775	46	514.2472	21	677.3926	20	677.3461	20
515.34	960	515.3197	399	678.3446	204	678.3563	131
516.3427	223	516.3325	85	678.4724	37	678.4492	16
517.3058	131	517.3261	69	679.3439	61	679.3323	48
518.3003	79	518.2698	30	679.8438	110	679.8321	76
519.2855	46	519.2551	28	680.3439	114	680.3439	83
519.7633	278	519.7429	117	680.7394	61	680.7394	17
520.2717	196	520.2616	91	681.0652	66	681.0885	28
521.2487	39	521.2894	17	681.3795	56	681.4144	28
522.2979	71	522.2877	28	682.3577	43	682.346	36
523.2462	587	523.236	478	683.3832	18	683.3482	13
523.7359	184	523.7359	151	689.3884	16	689.3416	19
524.277	118	524.2463	77	695.3611	16	695.3494	18
525.2679	29	525.2679	15	703.3806	31	703.3925	31
526.2291	30	526.2291	14	703.8775	21	703.8538	47
527.273	41	527.273	23	704.3863	41	704.3389	53
528.3283	39	528.2565	13	712.3504	30	712.3623	34
529.3436	21	529.3025	13	732.3539	49	732.3659	32
530.2982	25	530.2879	18	733.3679	39	733.3679	29
531.2845	90	531.2742	38	741.3359	159	741.3359	143
532.2923	47	532.282	27	742.3319	44	742.3562	39
533.2702	46	533.3011	24	743.438	260	743.4501	100
533.8264	38	533.8161	39	743.6082	27	743.5839	15
534.3211	63	534.3005	32	744.4475	70	744.4597	37
536.3228	371	536.3228	295	746.3591	35	746.3834	24
536.8186	127	536.829	97	747.334	55	747.334	36
537.325	79	537.294	51	748.334	21	748.3706	13
538.2661	46	538.2765	32	749.4324	84	749.4568	80
539.2702	33	539.2806	18	750.4338	31	750.446	26
540.2752	107	540.2648	43	757.3764	569	757.3764	240
540.3789	19	540.3789	14	757.4869	94	757.4869	33
540.7729	26	540.7729	18	757.8674	302	757.8797	145
541.2604	33	541.2707	19	758.3831	151	758.3831	44
542.288	60	542.2672	23	759.3905	62	759.4151	26
543.2854	30	543.3374	24	774.3827	48	774.4075	28
544.2941	40	544.3357	16	779.4417	54	779.4043	16
545.3142	30	545.2517	23	794.4033	163	794.441	68
545.7829	299	545.7933	321	795.4091	83	795.4468	16

546.2935	144	546.2831	168	800.347	23	800.347	16
546.7835	46	546.7731	43	811.4495	37	811.4622	15
547.2842	58	547.2946	36	812.4407	20	812.4788	13
548.2862	50	548.3175	19	906.4704	144	906.4973	21
549.2682	348	549.2577	269	988.5667	16	988.5247	16

A3.1.4 Peaks that did not match - spots 23 and 24

m/z Spot 23	intensity
402.3575	16
411.262	24
417.1523	13
427.1762	13
430.2232	26
447.298	22
473.2518	23
493.2692	18
495.4109	16
501.3743	14
519.8548	17
535.1977	18
535.7549	16
546.4082	18
548.3906	13
551.35	13
565.4071	36
569.4225	20
572.2673	21
585.3779	30
586.3926	15
591.7827	14
600.4168	76
600.4168	76
601.7944	13
605.388	15
623.4831	34
629.4547	35
651.279	14
651.8368	22
655.2911	15
658.3084	13
663.3524	18
674.3782	13

m/z Spot 23	intensity
680.8441	29
692.3422	13
694.3384	18
702.3284	13
704.4691	16
704.848	19
705.3453	17
711.8387	24
713.3624	13
714.3633	14
720.3592	16
736.366	13
760.4108	14

A3.2 N-Ethylmaleimide and acrylamide alkylation time courses

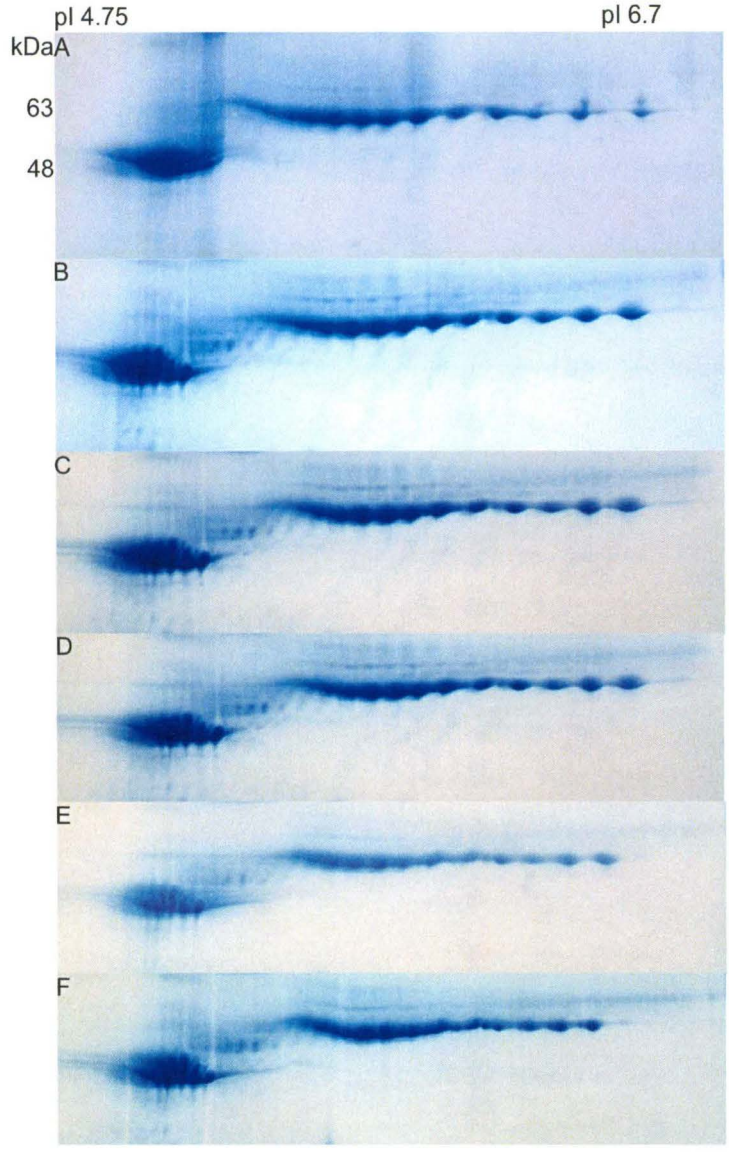


Figure A3.1
2DE separation of IFPs after alkylation with acrylamide for A) 0 minutes, B) 10 minutes, C) 1 hour, D) 2 hours, E) 6 hours and F) 24 hours. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10-20% T gel. The gels were stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments.

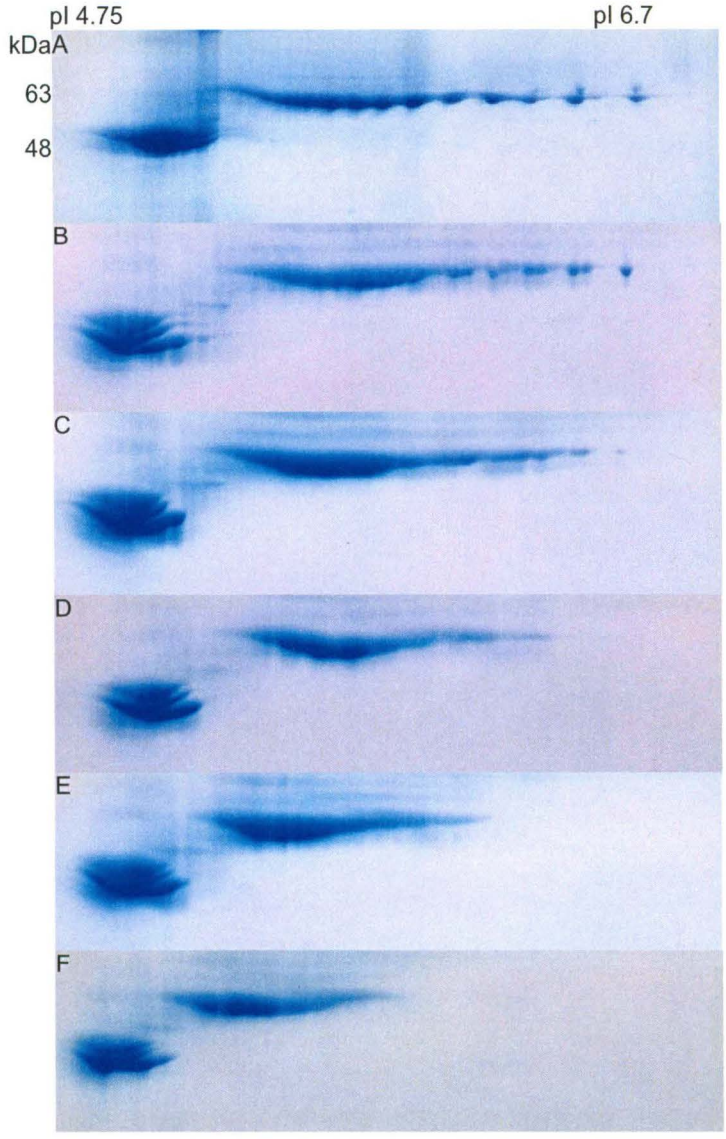


Figure A3.2
2DE separation of IFPs after alkylation with NEM for A) 0 minutes, B) 10 minutes, C) 1 hour, D) 2 hours, E) 6 hours and F) 24 hours. Proteins (200 µg) were separated in the 1st dimension on a pH 4-7 IPG strip. Second dimension separation was run on a 10% T gel. The gels were stained with Coomassie brilliant blue G-250. The gels are representative of duplicate experiments