Proteomic analysis of tendons at different risk of injury.

Karen Sanders

December 2016

Table of Contents

Acknowledgments

This work was supported by the Wellcome trust, which allowed me to undertake this research in a yearlong project. This year has given me many opportunities, which I am very appreciative of.

With special thanks to my supervisor Professor Peter Clegg, who provided much needed encouragement, feedback and patience whenever it was required. Furthermore for the helpful discussions during the preparation of this thesis and all the time spent reading, correcting and re reading. Also to Dr Mandy Peffers for your patience, support, countless ideas, enthusiasm and motivation. Thank you too Professor Robert Beynon for allowing me to join your team at the Centre for Proteome Research and of course for your honesty and assistance any time I required it.

The Mass Spectrometry Imaging investigations were undertaken with The M41 division of Mass Spectrometry at the University of Maastricht, this was under the supervision of Professor Ron Heeren . Particular thanks goes to Dr Berta Cillero-Pastor, Dr Gert B. Eijkel and Dr Anne Bruinen who took time out of their busy schedules to welcome me into their laboratory and provided a fantastic learning experience.

With thanks to all my colleagues in Musculoskeletal Biology I within the institute of ageing and chronic disease. I have enjoyed working with each and every one of you. My family for all of their support and the education they provided me with, which has allowed me to be where I am today.

Lastly to Neil, for your unwavering support, encouragement and the countless hours spent in Coffee Shops while I wrote.

Abbreviations

AGC	Automatic gain control
AMBIC	Ammonium bicarbonate
ANOVA	One-way analysis of variance
AT	Achilles Tendon
ATT	Anterior Tibialis Tendon
CDET	Commen digital extensor tendon
CHCA	α-Cyano-4-hydroxycinnamic acid
COMP	Cartilage oligomeric matrix protein
Da	Daltons
DA	Discriminant Analysis
DTT	Dithiothreitol
ECM	Extracellular Matrix
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionisation
ESWT	Extracorporeal Shock Wave Therapy
FA	Formic acid
FASP	Filter aided sample preparation
FDR	Falso Discovery Rate
FTMS	Fourier Transform Mass Spectrometry
GAG	Glycosaminoglycan
GnHCL	Guanidine hydrochloride
H	Heavy
H&E	Hematoxylin-eosin
H_2O	Water
HCD	Higher energy collisional dissociation
HCL	Hydrogen chloride
HPLC	High Performance Liquid Chromatography
IAA	Iodoacetamide
IFM	Interfascicular matrix
IL	Interleukin
ITMS	Ion Trap Mass Spectroscopy
ITO	Indium tin oxide
L	Light
LC	Liquid chromatography
LTQ	Linear ion trap mass spectrometer
m/z	Mass to charge ratio
MALDI	Matrix Assisted Laser desorption/ionization
MEGA-8	*N*-Octanoyl-*N*-methylglucamin
MeOH	Methanol
MMP	Matrix metalloproteinase
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Mass spectrometry/Mass spectrometry

MSI	Mass Spectrometry Imaging.
OCT	Optimum cutting temperature
OcGlu	n-Octyl β-D-glucopyranoside
PANTHER	Protein annotation through evolutionary relationship
PCA	Principal component analysis
PMF	Peptide mass fingerprinting
PT	Posterior Tibialis
PTMs	Post translational modifications
RIA	Relative Isotope Abundance
S/N	Signal to noise ratio
SA	Sinapinic acid
SDFT	Superficial Digital Flexor Tendon
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SF	Surfactant
SIMS	Secondary ion mass spectrometry
SLRPS	Small leucine-rich proteoglycans
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TFA	Trifluoroacetic acid

Abstract

Tendon injury is a common problem in both human and veterinary medicine, which impacts on the both the quality of life of the individual in question but also a financial implication on the supporting health system. Previous research into the pathogenesis of tendon injury and also the healing and treatment options available are limited. This one year project aimed to advance our knowledge in proteomics of human and equine tendons by developing proteomic methodologies for the analysis of the normal protein composition.

In order to build a proteomic tendon library three different protein extraction protocols were considered, urea, guanidine and Rapigest. Results revealed urea had a poor extraction ability and that guanidine and Rapigest had a better success rate. Further work to combine these to agents would be worthwhile in the future.

Further to this to this analysing spatial distribution of these proteins was then investigated in equine tissue. This work made some progress in optimizing a protocol, finding trypsin digestion worked better when incubated for 16 hours and also when the digestion occurred with an ammonium bicarbonate buffer.

The fourth chapter investigated whether it is possible to detect in vivo changes in the proteomic composition of dialysis fluid subsequent to shockwave therapy of human Achilles tendon. This study revealed that current methods were sensitive enough to ascertain the protein composition but further work is needed to ascertain what significant protein changes occur in response to therapy.

Finally tendon protein turnover was explored in an in-vivo model. Results confirmed proteins within the non collagenous matrix had a higher rate of protein turnover compared to components of the collagenous matrix. All four projects have been the first of their kind so optimisiation and a 'look and see' approach was required, which has laid foundations for several subsequent studies.

Chapter 1, Literature Review

Introduction

Tendons are highly organised connective tissue structures that have essential roles in the biomechanical function of the musculoskeletal system. These tissues primarily transfer tensile strength from muscle to bone, and associated with this transfer, are varying degrees of energy storage depending upon the individual tendon. Tendons absorb and then return energy reducing the metabolic cost of locomotion by a 'spring board' effect. As importantly, these tissues also have a 'positional' role to provide structural stability to the joint and limb. Tendon injury is a common problem in both human and veterinary medicine. Injuries are characterised by a combination of pain, swelling and impaired tendon performance resulting in disability and a loss in productivity. This not only impacts on the quality of life of the individual but also has an economic burden to the health care system. Epidemiological studies have emphasised the significance of tendon injury in both humans and the performance horse. In humans the prevalence of Achilles tendon (AT) injury in the general population under 45 years of age, was reported to be 3%, which increased to 15-56% in elite athletes depending upon the type of sport undertaken (Knobloch *et al*, 2008; Kujala *et al*, 2005). In the same study 291 elite runners were monitored over an 8-month period, it was documented that 56.6% of the athletes had an Achilles tendon injury at some point during this time frame (Knobloch *et al*, 2008). A large retrospective study in Thoroughbred racehorses, found the prevalence of superficial digital flexor tendon (SDFT) tendonitis to be 11.1% during a 12-month time span (Kasashima *et al,* 2004). In another study SDFT pathology was detected by ultrasonography in 24% of National Hunt racehorses over two racing seasons *(*Avella *et al,* 2009*)*.

In each of the above studies, it was also noted that there was an increased risk of injury linked to increasing age (Thorpe *et al* 2015). Injury is believed to occur due to an accumulation of microdamage over a number of loading cycles rather than an acute injury (Riley, 2008; Thorpe *et al,* 2015). In the equine patient injuries are often localised to the central core of the mid metacarpal region (Pool and Meagher 1990; Stromberg and Tufvesson 1969; Webbon 1977).

As well as a cause of pain and morbidity, healing of these injuries are usually slow and often mechanical properties do not return to normal after periods of prolonged rehabilitation (Jung *et al*, 2009).

To date, no treatment methods have been proven in the re-establishment of normal tendon structure and function due to poor understanding of the underlying pathology (Ackermann, 2015). The prevalence and significance of tendinopathies warrants research into this area. The development of new treatment strategies for tendon repair can only be addressed by further understanding of tendon function, structure, composition, and basic mechanisms of healing.

Animal models are often used in tendon research to allow the examination of biological processes that occur. This can range from normal homeostasis through to tendon pathology and repair. Using animal models is not without drawbacks as differences between species means that direct comparisons cannot be made, and cannot truly replicate conditions seen in the human patient. Small animal models do not replicate the same conditions seen in humans as different lifestyles mean that the same stresses and strains do not apply. Also their basic anatomical structure is different, in mice the structure of the interfascicular region is different (Kohrs *et al*, 2011). The use of horses as models is quite common and justifiable as not only does research on this tissue benefit the species but the equine SDFT is also an excellent model of exercise-induced AT injury in humans. The SDFT and AT are functionally and clinically equivalent plus lifestyle factors and longevity of horses are comparable (Patterson-Kane and Rich, 2014). The relatively easy access to equine tissues has facilitated important advances in tendon research. Examination of human tissue tends to occur during end stage pathology, therefore using an animal model provides a host of opportunities to learn more about the normal tendon and early changes that develop during pathology.

A vast amount of knowledge so far has been gained from *ex vivo* analysis

however further advances in clinical tests may, one day, allow the use of *in vivo* testing which would be a huge breakthrough in this field. This literature review will discuss the major developments in tendon research and highlight the gaps in the current knowledge that need to be investigated.

Tendon function

Despite the common features that all tendons hold, there is a remarkable variation in the morphological, molecular and mechanical properties between different tendons within the body (Birch, 2007). Broadly speaking tendons can be divided into two functional groups; those that primarily have a positional role and those that also act as energy stores. It is widely documented that the rate of injury to these two functional types varies: with energy storing tendons, being injury prone in contrast to their positional counterparts (Thorpe, 2014). Examples of energy storing tendons are the SDFT in the horse and the AT in humans. Positional tendon examples would be the common digital extensor tendon (CDET) in the Horse or the anterior tibialis tendon (ATT) in humans.

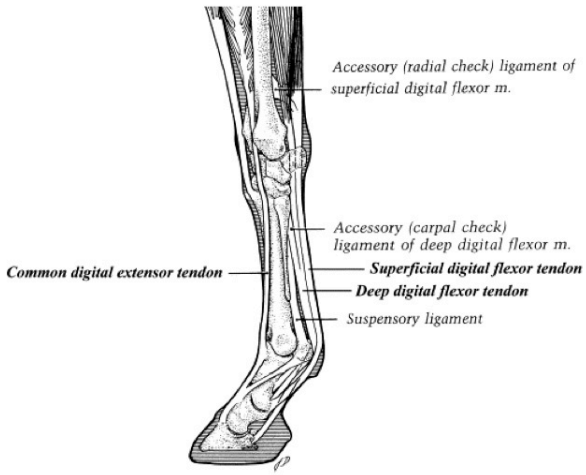


Figure 1. Diagram showing anatomy of the equine forelimb and the location of the energy storing SDFT and the positional CDET. Reproduced from Thorpe *et al*, 2013.

In order to understand how these differences are achieved it is necessary to have an understanding of the structure and the mechanical properties of the two tendon types. Research in recent years has come a long way in explaining how tendons succeed in their individual roles. Energy storing tendons such as the SDFT improve the efficiency of locomotion by their ability to stretch and recoil under physiological loads to ensure efficient return of stored energy (Birch,

2007) with one *in vivo* study measuring strains of up to 16% in the SDFT during galloping (Stephens *et al,* 1989). High levels of extension and subsequent recoil can increase the efficiency of locomotion by as much as 36% during high speed locomotion (Briewener, 1998). In comparison a positional tendon needs to be relatively inextensible in order to allow the correct placing of the limb and also to transfer force from muscle to bone. In contrast to the SDFT *in vivo* studies have calculated the maximum strains of the CDET to be approximately 2.5% (Birch *et al,* 2008; Batson *et al,* 2003 and Brown *et al,* 2003).

Basic tendon composition and structure

In order to meet its functional requirements, tendon tissue has a unique anatomical structural hierarchy. The collagen fibers are highly aligned in a longitudinal manner to create subunits, comprising of, fibrils, fascicles and fibers (Kastelic *et al*, 1978)(Figure 2).

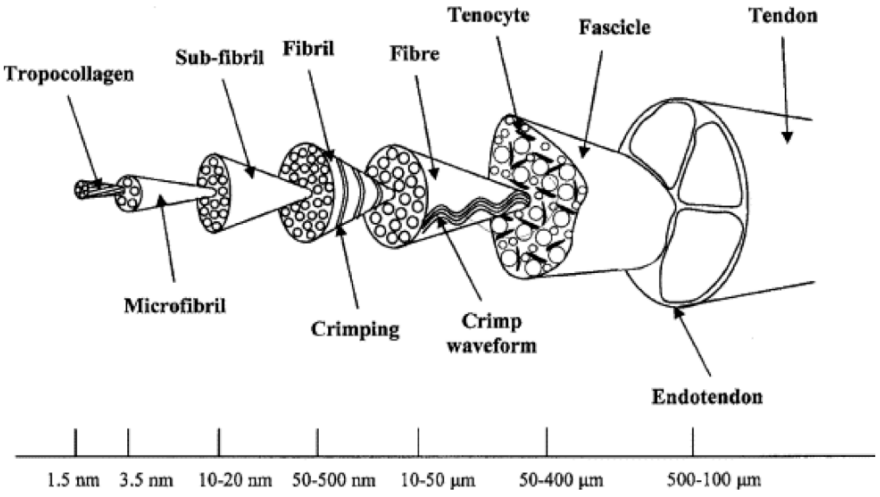


Figure 2: Schematic illustrating the structure of tendon tissue. Reproduced from Kastelic *et al*, 1978.

The division of the tendon tissue into individual smaller components ensures that minor damage does not necessarily spread to the entire tendon. (Kjaer, 2014). Investigations looking into tendon composition have ascertained that water makes up 55-70% of the wet weight (Kjaer, 2004). The primary molecular components are collagens, which make up 60-85% dry weight (Thorpe *et al*, 2013). Approximately 95% of this is type I collagen, with small amounts of collagen types III, V, XII and XIV (Thorpe *et al*, 2013; Birch *et al,* 1999). The collagen fibers are highly aligned in a longitudinal manner to create subunits,

comprising of, fibrils, fibers, and the largest of which is the fascicle (Kastelic *et al*, 1978). The fascicle can vary in diameter typically ranging from 50 to 400µm. The role of collagen is well characterised and contributes to the high tensile strength that this tissue requires to function (Silver *et al,* 2003; Franchi *et al,* 2008; Kjaer *et al,* 2009).

As to be expected there are differences in the compositional components of the energy storing tendons and their positional counterparts. Work has shown that the positional CDET has a significantly higher collagen content (80.4% ± 1.3) compared to the SDFT (75.8% ± 1.5). The size of the fascicles also differ, fascicles dissected from the SDFT were significantly smaller in cross sectional area (0.12 ± 0.006mm^2) than fascicles obtained from the CDET (0.16 ± 0.09mm^2) (Birch et al, 2013). The fascicles in the CDET appeared to be more tightly packed than those in the SDFT (Birch *et al,* 2013).

A connective tissue membrane, named the epitenon extends into and around the tendon covering the entire structure and separates the fascicules. Where this membrane separates the fascicules, it is known as the endotenon or interfascicular matrix (IFM)(Figure 3).

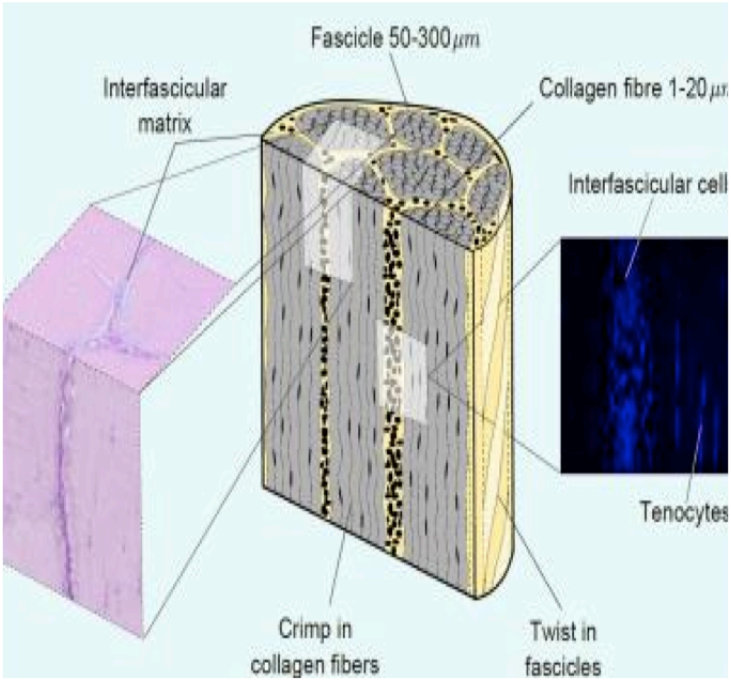


Figure 3. Schematic and histological section of a bovine flexor tendon, revealing the collagen rich fascicles, termed fascicular matrix (FM) and the surrounding endotendon, termed interfascicular matrix (IFM). Reproduced from Speisz *et al*, 2015.

The IFM is primarily composed of macromolecules subdivided into three groups proteoglycans, glycoproteins and collagens, (Collagen types III and VI are present in low concentration). The remainder of the weight is composed of cellular proteins and the fibrous protein elastin (Birch, 2007; Speisz *et al*, 2015; Thorpe *et al*, 2015).

Proteoglycans are composed of a core protein to which one or more glycosaminoglycan chains attached. Two groups of proteoglycan's have been demonstrated in the tendon; small leucine-rich proteins (SLRPS) such as decorin, biglycan, fibromodulin and lumican and large aggregating proteoglycans such as aggrecan, lubrican and versican (Rees *et al,* 2000; Yoon & Halper 2005). To date the function of each SLRPS is unclear, however they are thought to influence the mechanical and structural properties in a tendon specific manner, by regulating fibrillogenesis and matrix organisation (Gordon *et al*, 2015). Studies involving decorin and biglycan knockout mice showed altered dynamic and failure loads of the Achilles tendon, specifically this study detected a decrease in the organisation of the collagen fibres in these mice (Gordon *et al*, 2015).

The modular proteoglycans function to hold water so act in a spacing and lubricating role for the tendon (Thorpe *et al,* 2013). Cartilage oligomeric matrix protein (COMP) is an example of a glycoprotein, which resides predominately within the fascicles. Its precise function is unknown but thought to be involved in collagen fibrillogenesis and tendon growth (Thorpe *et al*, 2013). Elastin, as its name suggest is a highly elastic protein capable of extending by more than 100% of its original length, other relevant properties are a high fatigue resistance and the capacity for energy storage (Aaron & Gosline, 1981). It has been reported to be present in concentrations in the range of 1–10% of the tendon dry weight (Silver *et al,* 2003; Korol *et al,* 2007; Gosline *et al,* 2002). The matrix is maintained by small population of tenocytes, which are responsible for remodelling the matrix as they synthesise both the matrix molecules and also the enzymes responsible for degradation. Within the extracellular matrix tenoblasts and tencytes represent about 90-95% of the tendon cells, they either reside between the collagen fibres within the fascicles or are grouped within the

interfascicular matrix. The remaining 5-10% of cells comprise of chrondrocytes, vascular cells, synovial cells and smooth muscle cells (Ackermann, 2015).

The paratenon is loose connective tissue surrounding the whole tendon and allows the structure to glide against adjacent tissue. This structure contains mostly fibroblastic cells, which are responsible for the production of the extracellular matrix (Benjamin et al, 2008).

Tendon mechanics

As energy storing tendons must withstand great physiological loads, studies have attempted to investigate the mechanics behind this additional function. Recent research has established a distinctly specialised matrix as the protein profile differs between tensile and compressive regions of the tendon. For instance a region that experiences compressive loads has a greater content of non-collagenous proteins such as aggrecan, and lubricin (Batson *et al,* 2003; Jarvinen *et al,* 2003). Functions of these proteins are thought to increase tendon water content providing an increased stiffness and a resistance to compression (Rees *et al,* 2000). The levels of COMP have been recorded to be much higher in structures subjected to higher strains, one study recorded levels of 22.1 ± 6.1μg/mg in the SDFT compared to 1.9 ± 0.6μg/mg in the CDET (Birch *et al*, 2013). It has also been demonstrated that the matrix changes during development, maturation, ageing and injury.

Another study looked into the alignment and orientation of the fascicles. It was established that the extension and recoil mechanisms differed between the SDFT and CDET fascicles. In the CDET fibre sliding was the dominant functional mechanism. In comparison, SFDT fascicles were thought to contain a helical component, which relaxed under strains then recoiled. This allows a greater recovery following physiological loads and a lower level of energy loss (Thorpe *et al*, 2013).

The SDFT needs to have a high elasticity (low stiffness) to stretch and recoil, in comparison, the CDET needs to have a low elasticity (high stiffness) to transfer forces quickly and efficiently. Mechanical testing studies in both tendon types

have scrutinised the different properties of both the whole tendon and the material from which the tendon is composed, findings included that fascicles from energy storing tendons were stiffer than their positional counterparts and it was differences in the interfascicular region which allowed for the specialisation of the two tendon types. As stiffness of the fascicular interface in the SDFT was lower compared to the CDET (Thorpe *et al*, 2012). Figure 1, demonstrates results acquired.

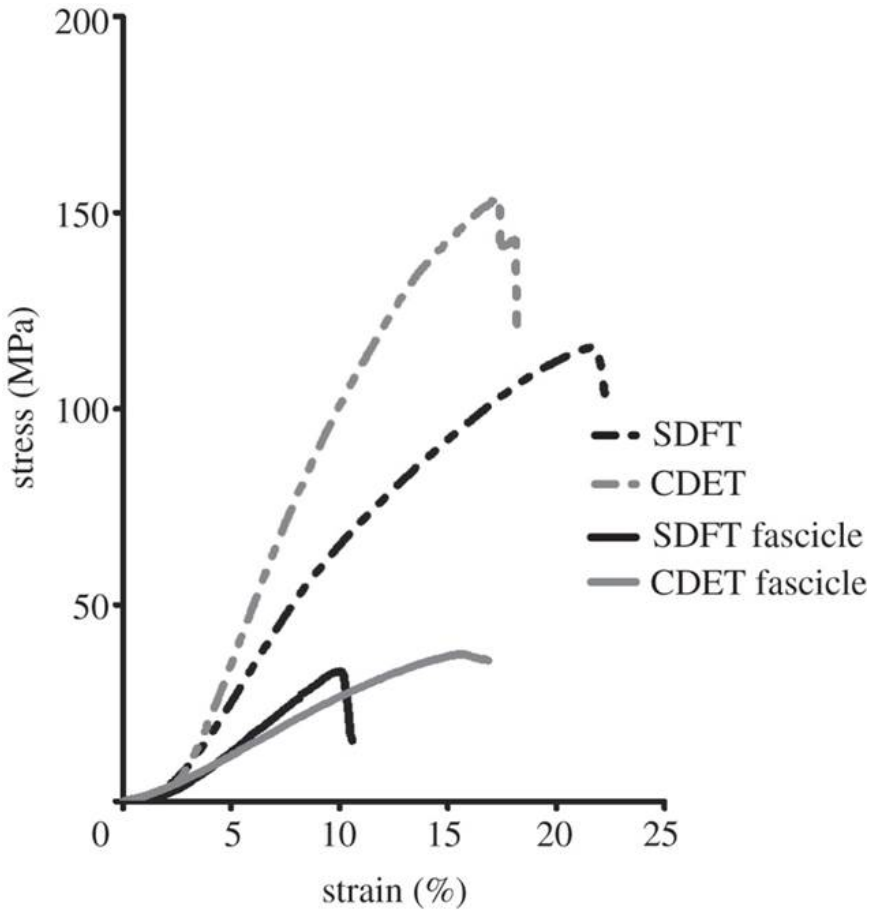


Figure 1. A representative tendon and fascicle stress–strain curves in the two tendon types, Used with permission from Thorpe *et al*, 2012. Illustrating differences in the extensibility of the two tendon types. The SDFT failed at higher strains compared to the CDET demonstrating increased extensibility. However the opposite was seen when the fascicles were tested. This is due to differences in the interfascicular matrix, in the SDFT the matrix was found to be less stiff than the CDET allowing a greater extensibility.

As mentioned previously, it is widely known that the incidence of tendon injury increases with increasing age; the same mechanical testing studies were repeated in young and old sample groups. This time there was significantly less fascicular sliding in aged SDFT tendons demonstrating increased interfascicular matrix stiffness. It is possible this could then lead to the fascicles being loaded at an earlier point during tendon extension (Thorpe *et al,* 2013). In CDET samples there was no change in the fascicular sliding between young and old samples.

The differences demonstrated between the two tendon types could be a basis of increased injury in energy storing tendons (Thorpe *et al,* 2013).

From the information gathered from the literature available it can be surmised that the increase in tendon injury in older individuals is associated with ageing of the IFM. Further study into the IFM structure and composition is the next phase of research as this region could potentially be a therapeutic target or used in conditioning and rehabilitation interventions.

Tendon ageing and injury

Several studies have assessed the effect of ageing on tendon composition. Reported changes in the matrix composition include, increased type III collagen at the core of the tendon, although the total collagen content was no different between old and young SDFT samples (Thorpe *et al,* 2015). A difference has also been found in the amount of partially degraded collagen, increasing amounts with increasing age within the SDFT (Thorpe *et al,* 2010). It is possible that an inability to remove partially degraded collagen from the matrix leads to a reduced mechanical competence that could account for the increased stiffness of the IFM in energy storing tendons. Proteomic studies into the effects of ageing have found distinct proteomic profiles in young and old tendon samples. The main findings were that in old tendon there was a reduction in fibromodulin, mimecan and asporin, which are all SLRPs. These proteoglycans have been shown to regulate collagen fibrillogenesis and fibril diameter. The heat shock protein 27, histones and integrins also decreased in ageing. Proteins that had increased with age were keratins, gap junction proteins and interfilament proteins (Peffers *et al,* 2014).

It is well documented that tendon injury occurs after an accumulation of microdamage over a number of loading cycles rather than an acute injury (Riley, 2008; Thorpe *et al,* 2015). It is thought that the biological response to tendon injury can be divided into three overlapping stages, the first stage involves an influx of inflammatory cells, growth factors and endothelial chemoattractants that enable the formation of a fibrin clot which aims to provide a temporary

stiffness to the area. The second stage begins two days after the initial injury, this phase is directed by macrophages and tenocytes. Macrophages switch from a phagocytic role to starting the repairing process by releasing growth factors and directing cell recruitment. Tenocytes start to deposit collagen III which acts as a temporary matrix. The final stage of tendon injury tends to occur 1-2 months after the initial injury and can last more than a year. In this stage collagen I synthesis is more prominent and the extracellular matrix becomes more aligned (Voleti *et al,* 2012). From this research is it evident that there are proteomic changes depending upon what stage of injury the tendon is in.

Macroscopic examination of injured tendons can show the following characteristics: collagen fibril disorganisation, increased size of the non collagenous extracellular matrix, hypercellularity and neovascularisation (Nourissat *et al,* 2015). Evidence of degeneration has been found at post mortem when the patient exhibited no clinical signs of tendon injury (Birch, 1998). Multiple degenerative changes in the extracellular matrix have been detected in injured tendons. Equine SDFT tendons with visual signs of tendinopathy ie abnormal colouration were analysed, changes repeatedly detected were increased levels of GAG and type III collagen in comparison to normal tendons (Birch, 1998). An increase in heat shock proteins has also been observed in tendinopathy (Millar *et al,* 2012).

Matrix metalloproteinases (MMPs) belong to a family of degradative enzymes responsible for matrix breakdown. It thought that some of these enzymes are required for the general health and maintenance of the tendon, as they remove damaged matrix and remodel scar tissue in order to try and regain its previous functions (Riley, 2005). This is supported by the observation that some drugs, which are known to affect the activity of MMP can induce lesions with patients being treated for other conditions (Corps *et al*, 2002). However some MMPs are thought to be damaging if their production is uncontrolled. A study which investigated the rate of collagen remodelling in human supraspinatus tendons found a relatively high turnover mediated by MMP 2, 3 and 1. However if the

tendon was ruptured there was an increase in expression and production of collagenase (MMP-1) and a down-regulation of stromelysin (MMP-3) (Clegg *et al,* 2007; Riley *et al,* 2002). In conclusion it is important to fully characterise these profiles of enzyme activity in order to produce future drug therapies.

Tendon tissue has little intrinsic regenerative capacity and so healing is characterized by the formation of fibrovascular scar tissue (Galatz et al, 2015). This is likely to be weaker than the original tissue and often a re-injury occurs (Corr *et al,* 2009; Marr *et al,* 1993). Understanding the mechanisms involved in tissue turnover and response to injury is essential in addressing pathologies in ageing, injury and disease.

Tendon turnover

Protein turnover is essential for the general health and maintenance of tissue. In tendons, it is also essential in the response to exercise, mechanical strain and injury. As mentioned already tendon healing can be problematic. In an effort to learn more, it is useful to consider the protein turnover rate within the tissue. An interesting exploration of turnover rates was undertaken in human Achilles tendon, on the basis of an increase in ^{14}C levels due to nuclear bomb tests in 1955-1963. The levels of ^{14}C were measured several decades later and showed that the core of the tendon had undergone very limited tissue turnover (Heinemeier *et al*, 2013). This was in stark contrast to skeletal muscle tissue which held no evidence of ^{14}C suggesting a much higher turnover rate.

Another turnover investigation used the racemization of aspartic acid to calculate the age of the collagenous and non-collagenous components of the two tendon types. Results revealed that the mean collagen half-life in the SDFT was found to be 197.53 (±18.23) years and this significantly increased with the horse's age, the CDET had a collagen half-life of 34.03 (±3.39) years. In comparison the half life of the non collagenous protein was 2 years in the SDFT and 3.5 years in the CDET. It was hypothesised that the energy storing tendons were protected from remodelling in order to try and preserve its structural integrity and mechanical functions. This was further emphasised when increased

collagen degradation markers were observed in the CDET compared with the SDFT (Thorpe *et al,* 2010).

The breakdown and synthesis of proteins is crucial for cellular homeostasis. Breakdown allows the clearance of abnormal or damaged proteins in order for regeneration of tissue. Protein turnover rates vary widely and are thought to decrease with age. One method of measuring protein turnover rates is to use stable isotope labeling. Mice are fed a diet containing a 'heavy' isotope that is stable incorporated into amino acids and thus labels newly synthesized proteins. Proteins with higher turnover rates will reveal a higher abundance of the heavy isotope. The ratio between light and heavy labeled peptides can be analysed by mass spectrometry.

Over the last decade significant advances in the field of tendon research have been achieved however there is still much to learn especially regarding the protein composition and turnover of tendon.

Introduction to Proteomics

Proteomics is the large-scale study of the structure and function of proteins. The expression and abundance of all proteins often mirror physiological processes at the cellular level. Protein analysis encompasses a variety of different aspects, including protein identification, protein spatial distribution, analysis of post-translational modifications (PTMs), protein-protein interactions, absolute and relative quantification of target proteins and the rate of protein turnover (Chen and Yates, 2007). These features are fundamental in the investigation and understanding of biological processes and disease. The identification of proteins plus determining how much of each protein is present, either as an absolute value or relative to the same protein in another sample provides information that allows us to understand the biological activities occurring. This information can help form theories into the pathogenesis of disease, further down the line this analysis then allows the development of screening tests for certain diseases.
Previous successes have been the discovery of biomarkers that can help to establish a diagnosis or a prognosis of a disease or by developing a targeted

therapy. For example, one study that undertook proteomic screening of urine, found that in bikunin loss was a potential bladder carcinoma marker (Tsui *et al*, 2010). Research across the world is continually aiming to increase our knowledge in the normal and diseased biological processes with the aim of improving healthcare. This section discusses why proteomic analysis is useful for the current investigation and the theory behind the techniques used.

Mass Spectrometry

There are many commonly used technologies to study proteins. Mass Spectrometry (MS) in particular is commonly used and more recently the use of mass spectrometry imaging (MSI) has been developed. Simply put, MS is an analytical technique that helps identify the composition of a given sample. The identification of lipids, peptides and proteins is achieved by electron ionisation and separation of the ions according to their mass to charge ratio. The relative abundance of each ion type is recorded and displayed as a spectrum. Figure 1 demonstrates a simplified schematic of the process

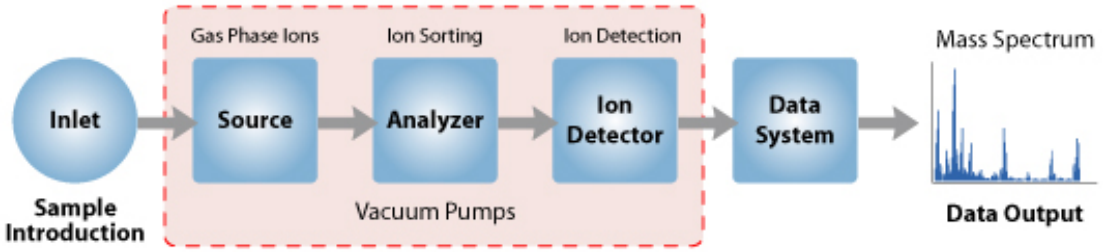


Figure 1. Schematic of Mass Spectrometer (http://www.premierbiosoft.com. Accessed on the 22.02.2016).

Frequently used MS methods are generally centered on liquid chromatography coupled to different electrospray ionization (ESI) MS/MS techniques, mainly being ESI and matrix-assisted laser desorption ionization (MALDI). The different types and abundances of proteins present in a given sample are recorded and analysed accordingly.

By using MS as an analytical tool to determine the chemical composition of samples, the mass spectra of normal vs diseased samples can be analysed and compared to create profiles of a disease. Using this technology, with the expectation that one day these tests could be transferred to in a clinical laboratory setting and used as a diagnostic test (Verrills *et al,* 2006).

However there are still many improvements to be made before this transfer to the clinical laboratory. Validated and approved clinical tests are required to be cost effective and robust. These samples create an enormous amount of data that needs to be processed and analysed so improvements in the handling of this data is required to enable cost effectiveness in clinical practice.

Many terms are used in proteomics to explain the strategies used, an understanding of these terms is required to understand the literature, some of these will be explained in the next section.

Mass spectrometry analysis

MS can characterise and sequence proteins via two approaches, a 'top down' approach ionizes intact proteins, which are then introduced to a mass detector and analysed. A 'bottom up' approach uses a protease to enzymatically digest the proteins into peptides, digestion can be carried out *in-solution* or after electrophoretic separation via SDS-PAGE termed *in-gel* digestion. The peptides are then ionised, separated according to mass to charge ratio and analysed. When the characteristic pattern of peptides is used for the identification of a protein, this is called peptide mass fingerprinting (PMF). Both the relative abundance of peptides from each sample, and protein identifications can be simultaneously obtained. Shotgun proteomics is a term used to describe bottom up proteomics in identifying proteins in complex mixtures using a combination of high performance liquid chromatography combined with mass spectrometry.

The study of complex samples (whole tissue or biofluids) is often achieved by 'bottom up' proteomics where the proteins are solubilised and then digested enzymatically, or chemically, into peptides. The term shot gun proteomics is

often used to describe a 'look and see' approach and is often achieved by utilising the combination of tandem MS and database searching. This can have limitations such as loss of information due to limited protein sequence coverage.

As biological samples are complex mixtures of multiple proteins and molecules MS analysis can present problems. Ionisation techniques work well when the sample contains an approximately equal amount of constituents, however in biological samples different proteins tend to be present in broadly differing amounts. When a mixture is ionised the more abundant molecules can suppress the signals generated from the less abundant ones. Samples can create very complex spectra that can be difficult to interpret (Verrills *et al,* 2006). This can be further exacerbated if the sample has been previously digested as this creates a large number of peptide products.

Mass Spectrometry Imaging

Mass spectrometry imaging combines the specificity of mass spectrometry analysis with microscopic imaging capabilities without the need for labeling. Detecting multiple analytes in parallel while preserving their natural location enables more information to be gleamed from their exact roles and functions. It is already established that the distribution of proteins and biomolecules is intrinsically related to their role, they can have multiple roles depending on their location (Simpson *et al,* 2000). MS imaging so far has played a large part in pharmaceutical research as it can give information regarding the pharmokinetics, pharmacodynamics and transport properties of drugs which is useful during early development of treatment protocols (Cobice *et al,* 2015). The visualisation of the distribution of compounds can answer important questions such as blood brain barrier penetration (Cobice *et al,* 2015). In the field of tendon biology, the differences between the fascicular matrix and interfascicular matrix could be examined. Matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI) and secondary ion MS (SIMS) are employed for this task. The basics of which are discussed below.

Background of MALDI Imaging

In 1997, Caprioli et al. demonstrated the use of MALDI MSI for the ion imaging of peptides and proteins in intact tissue sections, where the spatial arrangement of specific compounds were determined in an untargeted manner. Since then there have been significant improvements with instrumentation and data processing. MALDI-MSI is now a developing technique used to detect and record intact peptide and protein distributions of the mid to low molecular proteome with the detectable mass range being between 500m/z to over 100,000m/z (Kaletas, 2009). This technique has several other benefits such as a high sensitivity, large tolerance for salts, label free detection, the interpretation of data is simple as the majority of ions are singly protonated (Chungtai and Heeren, 2010).

The process of analysis is relatively simple, fresh frozen samples are thaw mounted onto glass slides and then a small organic compound, the matrix, is applied mixed in with a solvent solution. As the solvent evaporates, co-crystallisation of the analyte and the matrix compound occurs. As the analyte is incorporated into the matrix crystal, the molecules are separated from each other, thus minimizing the incidence of fragmentation (Kaletas *et al,* 2009). A laser beam is then aimed at the sample and the matrix molecules absorb the energy, this allows desorption and ionisation of the analyte to occur. Once ionised, the ions are accelerated at a fixed potential and become separated according to their mass/charge ratio. Images are created by plotting the intensities of a given ion on a map of the tissue. MALDI MSI has a lateral resolution of 10-100µm (Seeley and Caprioli, 2008). Spectra are recorded in an orderly fashion over the tissue by moving the target plate containing the sample underneath a fixed laser beam. Figure 2 demonstrates a schematic diagram of this process.

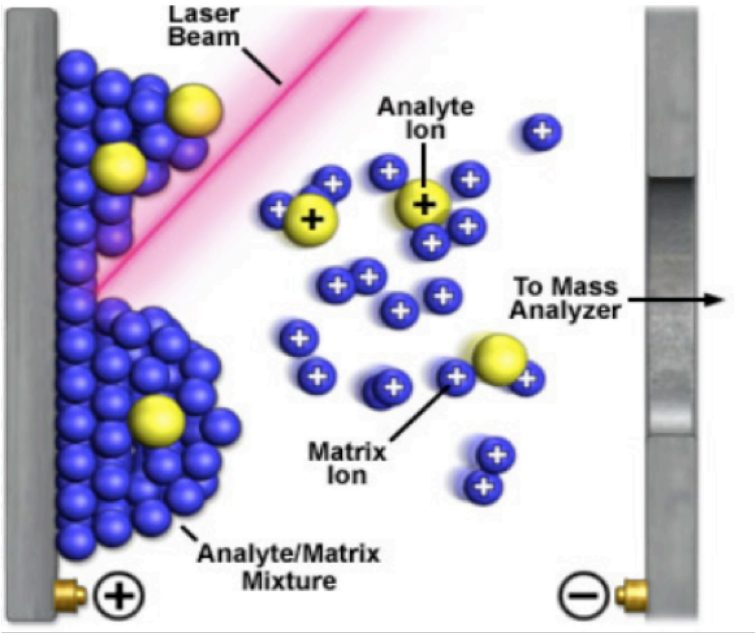


Figure 2. Schematic of MALDI Imaging process (Chugtai and Heeren, 2010).

Background to SIMS Imaging

SIMS is another desorption and ionisation technique which allows for a very high spatial resolution imaging (<1um) the limits of this technique is that the detectable mass range is limited to 1000da or less (Seeley and Caprioli, 2008). SIMS, in contrast to MALDI, is performed without any sample pre-treatment, it is able to identify most components of biological systems of masses below m/z 2000, the highest yields therefore are obtained for small molecules such as lipids and small metabolites. This technique uses a primary ion beam directed at the sample surface, this beam, which can be focused as sharply as 50nm, produces secondary ions. The primary ion beam is significantly higher in energy compared to the laser beam used in MALDI imaging and as a result of this, fragmentation of the service molecules can occur (Chugtai and Heeren, 2010). The secondary ions ejected at each beam position are collected and the intensities of each molecule recorded. This can then be used to create a composition map of the sample. A significant advantage of this method is that extensive sample preparation is not required. The tissue is loaded onto an indium tin oxide (ITO) coated glass slide without any further washing steps. Figure 3 demonstrates a simplified schematic of the process. A disadvantage to SIMS is the extensive fragmentation caused by

the primary beam can lead to difficulty in the interpretation of data (Chugtai and Heeren, 2010).

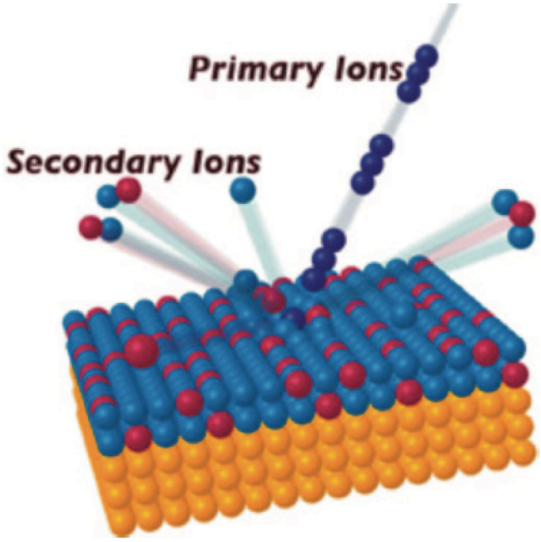


Figure 3 demonstrates a simplified schematic of SIMS (Chugtai and Heeren, 2010).

The images produced by MALDI MSI and SIMS are complementary to each other as a broad range of molecules can be detected from lipids and small metabolites to peptides and proteins.

By correlating a dual imaging approach that combines MALDI-MSI to obtain a low resolution molecular map of the peptides within tendon tissue. I intend to employ gas-phase ion mobility separation to enhance the number of peptides identified. The resulting images could then be used to direct subsequent microscopic SIMS imaging to examine small molecule distributions in areas of interest at higher spatial resolution. This multimodal approach allows multiplexed visualisation of unlabeled analytes at cellular and subcellular levels increasing our understanding of tendon functional structure and the changes in ageing. By using SIMS as well as MALDI-MSI small molecules present in the IFM that require higher spatial resolution will be localised. Mass spectrometry imaging has not been employed in this tissue before and so for MALDI-MSI a protocol for sample preparation must be generated and optimised.

Aims and objectives for study

Referring back to the start of this chapter, tendon injury is a major problem in both human and veterinary medicine. This not only causes an impact on the quality of life of the patient but also has cost implications to go alongside this. This year long study attempts to touch on three areas where there is no published data in order to increase our understanding of tendon biology.

To achieve an increased understanding of tendon biology it is crucial to determine the composition of the tendon tissue and how this differs between, not only, the fascicular matrix and interfascicular matrix but also how the composition varies between positional and energy storing tendons. This thesis aims to develop proteomic methodologies, which will enable future experiments to collect this data. Initially a comparative study using three different methods of protein extraction in two tendon types, in both human and equine species, with the aim of increasing coverage of the tendon proteome. To extend our understanding further, it is helpful to not only know the content of proteins but also the location within the tendon structure. Using mass spectrometry imaging, allows the spatial detection of proteins producing a molecular map of unlabeled peptides, lipids and small molecules at cellular and subcellular levels. The second task was to devise an appropriate protocol for equine tendon sample preparation for MALDI imaging. We hypothesise that there will be significant protein differences between the interfascicular and fascicular matrix. By extending the skills learnt in the previous two studies we then investigated the possibility of detecting *in vivo* changes in the proteomic composition of dialysis fluid subsequent to shock-wave therapy of human Achilles tendon. We hypothesised that if detection of peptides was possible in this media then there should be an increase in proteins expressed following treatment. Lastly protein turnover rates in mouse-tail tendon tissue were researched using the stable isotope of lysine. We hypothesized based on previous research available that there would be a higher rate of turnover within the non-collagenous matrix in comparison to the collagenous matrix. Bringing all of this data and research together provides a platform for further questions to be answered.

Chapter 2, Comparison of Protein Extraction Methods in Tendons for Mass Spectrometry Analysis

Using proteomic techniques to define tendon protein profiles and patterns is a relatively unexplored field. With the use of mass spectrometry the initial goal in this study was to build a reduced complexity database of diagnostic peptide ions and fragments, this can then be used to generate quantitative profiles of protein abundance in different tendon samples. The ability to detect proteins and variations in their abundance will not only give information regarding the biological processes that are occurring but also these changes can be correlated with disease detection and progression.

Biological samples are complex mixtures of proteins and in order to optimise the data sets following LC MS/MS analysis, sample preparation must be carefully considered. There are large variations in protein complexity within these samples that can result in the most abundant proteins masking less abundant ones. To increase coverage and sensitivity of data gained, techniques are employed to reduce the complexity of the sample while keeping them compatible with MS analysis. Alongside this, tendon tissue contains many protein types, *i.e.* cellular, membrane proteins and those contained within the extracellular matrix. Optimising extraction of all types needs to be considered. It is also critical to have a method that is reproducible in order to detect changes in the abundance of proteins between samples.

As it is the detection and quantification of peptides rather than proteins that is undertaken in LC MS/MS, a 'bottom up approach' is employed. Providing effective denaturisation of the protein will permit the digestion enzymes better access to the protease specific cleavage sites. Some believe that an ideal protocol would achieve a complete digestion of the protein as this could result in a better reproducibility while improving the overall coverage of the tendon proteome (Chen et al, 2007).

Various protocols for protein denaturation and extraction have been described in the literature such as the use of heat, sonication, solvents, surfactant and chaotropic agents (Proc *et al,* 2010). Simplicity, cost effectiveness and compatibility with down stream mass spectrometry are all crucial properties for a protein extraction protocol (Proc *et al,* 2010). To increase our understanding of tendon biology we must first optimise a method to extract proteins to analyse the tissue effectively. Having knowledge of the biological processes that occur within healthy, injured and old tendons will provide a basis for future advances in diagnostics and treatment plans.

Aims and objectives

This is a comparative study using three different methods of protein extraction in two tendon types, with the aim of increasing coverage of the tendon proteome. The ability to produce robust and effective guidelines in sample preparation will be invaluable for future studies in tendon research.

Materials and Methods

All reagents were obtained from Sigma-Aldrich unless otherwise stated.

Sample collection

Sample collection, handling and dissection of human tissue was undertaken by Queen Mary University London. Samples were harvested from limbs amputated during surgical procedures to treat sarcomas at the Royal National Orthopaedic Hospital, Stanmore. Tissue collection was carried out through the Stanmore Musculoskeletal Bio-Bank, which has ethical approval from the Cambridgeshire 1 Research Ethics Committee (REC reference 09/H0304/78) to collect tissue for research into musculoskeletal conditions. All patients gave consent for their tissue to be used for musculoskeletal-related research. Local research-and development approval for this project was given by the UCL/UCLH/RF Joint Research Office (reference number11/0464).

Equine samples received were obtained from a commercial equine abattoir so ethical approval was not required. No previous medical histories for the horses were available. Forelimbs, distal to the carpus were collected from half to full

Thoroughbred horses. A sample of both the SDFT and CDET were dissected from the mid metacarpal region. There were three donors with a mean ±SD age of 12.3±1.15 years. Only tendons that had no evidence of previous tendon injury at post mortem examination were included in this study. Human samples were obtained from cadavers from three age groups; a young male (9yrs old) a middle aged man (56yrs old) and an elderly female (96yrs old). Samples of the Achilles, Anterior Tibialis and Posterior Tibial tendons were collected. Each individual sample had three replicates; each replicate was assigned to a different extraction technique. Ethical permission was granted for this study by the NRES committee North east – Sunderland (Ref: 14/NE/0154).

Protein Extraction

For equine tendon samples, the mean±SD wet tissue sample size was 13.4±5.09mg. For human tendon samples, the mean wet tissue sample size was 5.2mg, S.D. 1.09. Deglycoslation was achieved in all samples by the addition of 200 µl of 100 mM Tris acetate, protease inhibitors (Complete Protease Inhibitors, EDTA-free, Roche Applied Science), and 0.1 unit of chondroitinase ABC, pH 8.0 to each Eppendorf and incubated at 37˚C for 6 hours. The supernatant was removed after centrifugation at 13,000 x g for 5 minutes. The collagenous and non-collagenous matrix components of the tendon were then separated by one of three methods: -

Method one – 4M Guanidine Protein Extraction

Protein extraction was achieved by the addition of 0.5 ml of guanidine extraction buffer (4 M guanidine hydrochloride (GnHCl), 65 mM dithiothreitol (DTT), and 50 mM sodium acetate, pH 5.8). Each sample was sonicated on ice (3 pulses of 10s each at 40% output) on a ultrasonic processor (Sonoplus HD 2070, Bandelin) and the extraction performed at 4^0C for 48hours on an orbital shaker (ThermoScientific Max Q). The supernatant was removed after centrifugation at 15,000 x g @4˚C for 15 minutes. The final insoluble fraction was kept aside for further analysis of the non collagenous matrix components.

Method two – 7M Urea Protein Extraction

Protein extraction was achieved by the addition of 0.5ml of urea extraction buffer (7M urea, 0.15M Sodium Chloride, 50mM Tris hydrochloride and protease

inhibitors (Complete Protease Inhibitors, EDTA-free, Roche Applied Science) pH 6.5). A Sonicator (Sonoplus HD 2070, Bandelin) was used in an attempt to maximize protein solubilization. Each sample was sonicated on ice (3 pulses of 10s each at 40% output) the extraction was then performed at 4˚C for 48hours on an orbital shaker (ThermoScientific Max Q). The supernatant was removed after centrifugation at 15,000 x g @4˚C for 15 minutes. The final insoluble fraction was kept aside for further analysis of the non collagenous matrix components.

Method Three – Rapigest SF Protein Extraction

50µl of 25mM Ammonium bicarbonate was added to each sample, which was then sonicated on ice (3 pulses of 10s each at 40% output). Subsequently 1mg Rapigest SF (Waters, UK) was added and the solution incubated @60˚C for 60 minutes. Within this time period samples were vortexed on two occasions. The supernatant was removed after centrifugation at 14,000 x g @4˚C for 10 minutes. The final insoluble fraction was kept aside for further analysis of the non collagenous matrix components.

Protein Assay

Ethanol precipitation was performed on aliquot's of the supernatant prior to estimation of protein concentration for both urea and guanidine samples. Samples were resolubilised with 7M Urea, 2M Thiourea, 30mM Tris, pH 8) then analysed by the Bradford assay read at 660nm using Coomassie Plus™ protein assay reagent (Thermo Scientific, Rockford, IL)

One-dimensional Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Silver Staining

To remove interfering substances GdnHCL and Urea soluble extracts were precipitated using Strataclean Resin (Stratagene R, Hycor Biomedical Ltd, Edinburgh, UK). All samples were subsequently separated in single dimensions based on their molecular weight by one dimensional SDS-PAGE to compare quantitative/qualitative differences in the protein profiles. 1ug of protein was loaded in equal with reducing buffer (2-mercaptoethanol) onto 4-12% acrylamide Bis-Tris NuPAGE gels (Invitrogen). Before loading samples were

incubated with reducing buffer for @ 90˚C for 5 minutes and then placed on ice to return sample to room temperature. Proteins were visualised by silver staining (Thermo Scientific) according to the manufacturer's instructions.

Protein Reduction/Alkylation and In-Solution Trypsin Digestion

Reduction of disulphide bonds and alkylation of free cysteines was achieved by incubation with 10mM DTT (10min at 60˚C) and then 50mM iodoacetamide (30min at room temperature in the dark) respectively. The soluble fraction was removed after centrifugation for 15 min at 13,000 × *g* at 4˚C and co precipitated with 1µg of trypsin (Promega) at 37˚C overnight. The concentration of trypsin was chosen as known workable concentration which was frequently used in the groups laboratory. To avoid auto digestion, the reaction was stopped with the addition of 1% trifloroacetic acid (TFA). To check digestion had occurred samples were again combined in equal volumes with reducing buffer (2-mercaptoethanol) and run onto 4-12% acrylamide Bis-Tris NuPAGE gels (Invitrogen), gels were run for 50 minutes at 200V. Silver staining (Thermo Scientific) was used to check for complete protein digestion.

Analysis of non collagenous matrix components

The final insoluble fraction from each of the extraction methods were incubated in 0.1M acetic acid containing pepsin at 100 µg/ml and shaken overnight at 4 °C. Samples were centriguged at 13,000 x g for 15 minutes, the supernatant was removed, lyophilized, re-suspended in water, re-lyophilized and then heated in 20µl of Laemmli buffer containing 50mM DTT for 5 min at 95 °C. The material was resolved using 3–8% acrylamide Tris acetate gels (Invitrogen) gels were run for 65minutes at 150V then silver stained.

Protein Identification

Samples were desalted and purified using C_{18} resin in the form of a ZipTip® (Merck Millipore, Rockland, MA, USA). Analysis was performed by LC MS/MS. (Orbitrap) using aliquots of tryptic peptides equivalent to 200ng of tendon

protein. Protein identification was achieved by searching through MASCOT Unihorse and Unihuman databases and by comparing samples using Progenesis™ QI software (version 4, Waters). Proteins with 2 or more unique peptide sequences were included in the final analysis.

Results

Results compare both the equine and human groups separately

1D SDS-PAGE

Results can be found in Figure 1 and 2. 10ng of each sample was loaded onto the gels. A quantitative comparisons of protein extracted using these three protocols cannot be made as silver staining was performed until the molecular standard lane was 'over-developed' to allow maximum staining of proteins extracted. Unfortunately the results obtained from the pepsin protocol were unable to be assessed due to technical artefacts with the gels.

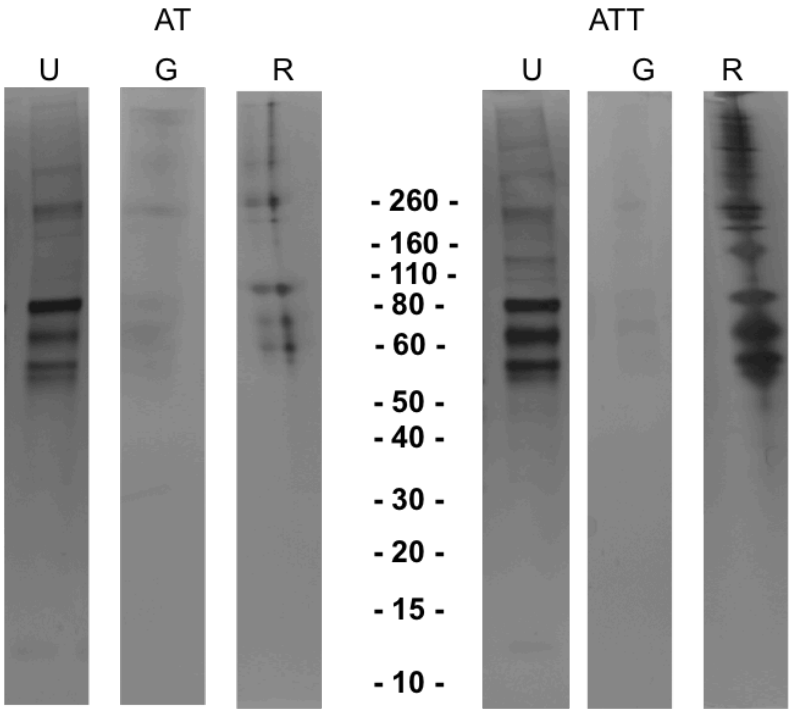


Figure 1. 1D-SDS-PAGE of Urea (U) Guanidine (G) and Rapigest SF (R) Human Tendon Samples collected from 9 year old male. Molecular weight

markers expressed in KDa. The reduced staining of the guanidine gel was due to under staining. Quantitative/Qualitative assumptions cannot be made, as staining time varied. Gels demonstrate that the gross protein profiles from these two different tendon types were similar.

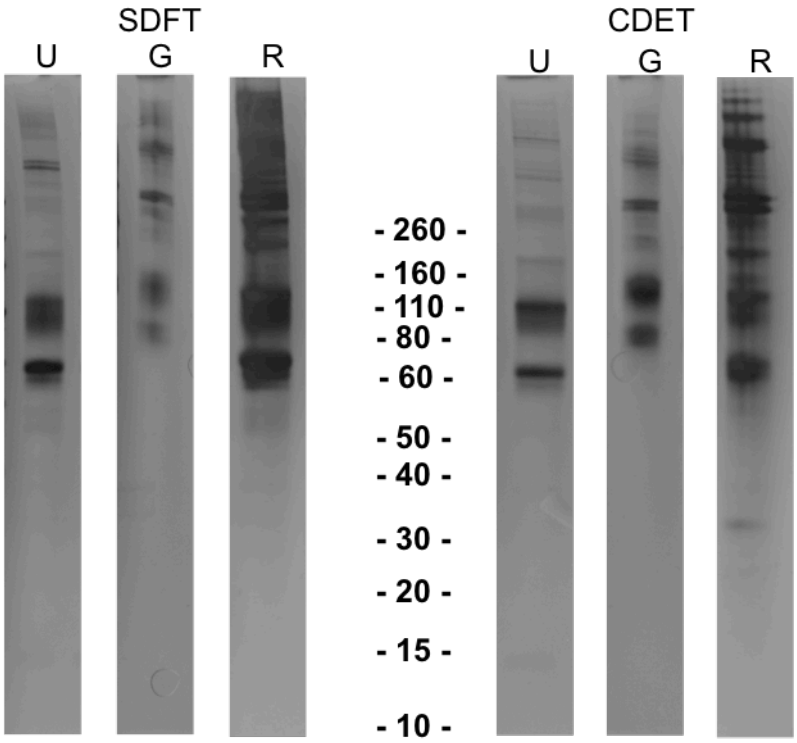


Figure 2. 1D-SDS-PAGE of Urea (U) Guanidine (G) Rapigest SF (R) and Equine Tendon samples from SDFT and CDET samples from the same donor. Molecular weight markers in KDa. Quantitative/Qualitative assumptions cannot be made, as staining time varied. More bands are detected with rapigest extraction

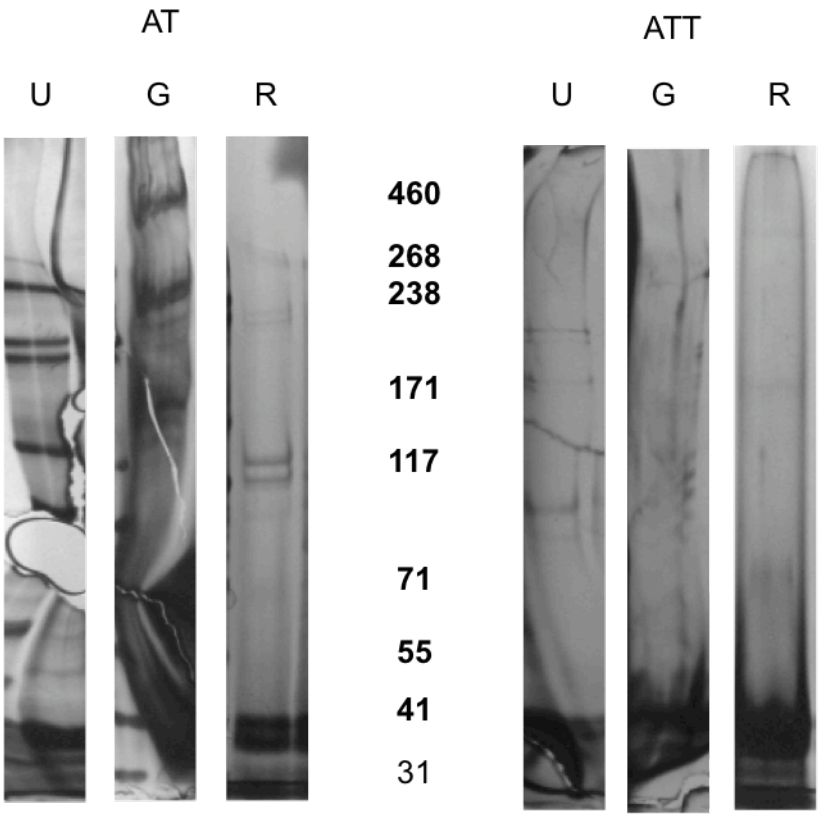


Figure 3. Pepsin released material from the entire insoluble fraction was resolved by Tris acetate 3-8% NuPage. Urea (U) Guanidine (G) Rapigest SF (R) human samples of AT and ATT from the same 9 year old donor. Molecular weight markers KDa. Artefacts due to a manufacturing fault within the gels meant that Quantitative/Qualitative assumptions could not be made. Possibly this fault occurred from the gels being to dry within their packaging.

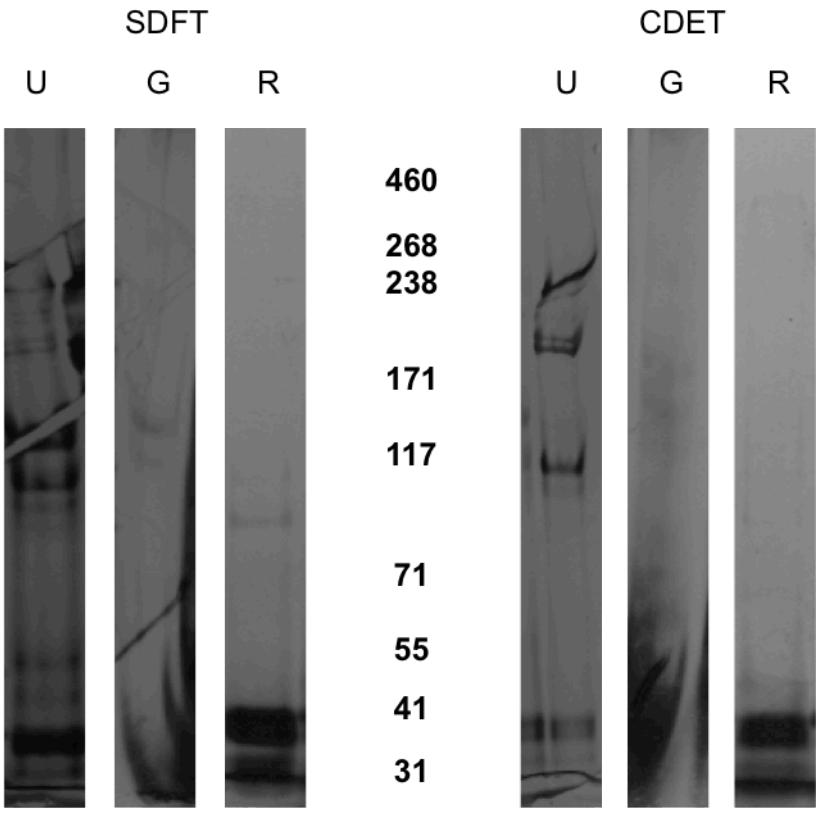


Figure 4. Pepsin released material from the entire insoluble fraction was resolved by Tris acetate 3-8% NuPage. Urea (U) Guanidine (G) Rapigest SF (R) Equine Tendon samples from SDFT and CDET samples from the same donor. Molecular weight markers KDa. Artefacts due to a manufactering fault within the gels meant that Quantitative/Qualitative assumptions could not be made.

LC MS/MS

Initial analysis came from comparing the complexity of the chromatograms generated from LC MS/MS. A typical chromatogram of each extraction method in samples from the same donor is shown for comparison in figure 3. Label-free relative protein quantification was then undertaken using Progenesis™ LC-MS software. To increase confidence in the results acquired, the average number of significant proteins was adjusted to 1% false discovery rate (FDR) on the MASCOT search, this reduces the proportion of false positive results (Pawitan *et al,* 2005). These results can be found in Figure 4. Although the guanidine extraction method appeared to have identified the greatest number of significant proteins, there was no statistical difference numbers of proteins identified between guanidine and Rapigest SF samples. The urea extraction group

produced the lowest number of significant proteins. To illustrate the relationship between the three protocols a Venn diagram was produced for both equine and human samples. Data was searched within MASCOT Unihuman and Unihorse databases. Figure 4 demonstrates Venn Diagrams, the Venn diagram illustrating human samples guanidine and Rapigest SF methods had 155 of the same proteins in common, guanidine and urea methods had 62 of the same proteins in common and Rapigest SF and Urea methods had 7 of the same proteins in common. 132 of the proteins detected were common to all extraction techniques. In equine samples guanidine and Rapigest SF methods had 25 of the same proteins in common, guanidine and urea methods had 61 of the same proteins in common and Rapigest SF and urea methods had 6 of the same proteins in common. There were 115 proteins detected in common to all extraction techniques. Following this initial data analysis bioinformatics tools PANTHER (protein annotation through evolutionary relationship) and STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) were employed to analyse the functional protein classifications and protein-protein interactions respectively (Mi *et al*, 2015; Franceschini *et al,* 2013). PANTHER is a classification system that provides ontological information about the significant proteins identified. STRING provides information on known and predicted protein-protein interactions. Before the analysis with the bioinformatics software, to increase confidence in the identification of proteins, data was further filtered to include proteins with at least two unique peptides and a FDR 1% in this data set. The urea extraction protocol did not achieve any results that were within these strict criteria, hence were not included in figures 5 and 6. Figure 5a-d reveals STRING analysis for each group; clustering was performed at a medium confidence level of ≥ 0.4. Figure 6 reveals PANTHER analysis of proteins classified according to their location, this aims to establish differences between the types of protein extracted, for instance cellular versus extracellular proteins. Within the human data, the Rapigest SF extraction protocol achieved a higher percentage of extracellular matrix components, 53%, compared to the guanidine extraction protocol, which was 36%. Within the equine data set the percentage of extracellular matrix components to cellular components between the two

protocols were similar. As there was a small sample size no comparisons could be made between subsets; in particular, how protein profiles may vary between young and old samples and between different tendon types as there was only one sample in each group.

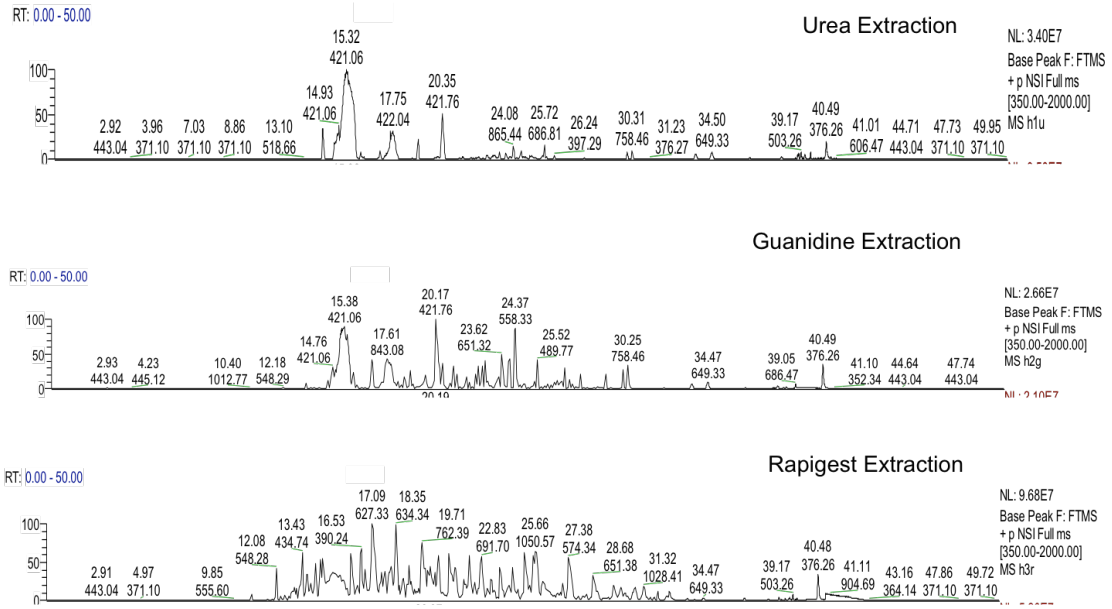


Figure 3 – Comparison chromatograms of all Extraction Methods in a 9yr Achilles Tendon. The same donor was used so a direct comparison can be made. The complexity of the chromatogram differs between each method. The most complex chromatogram is from the Rapigest SF extraction method and the least complex chromatogram results from the urea extraction method.

Human Venn Diagram

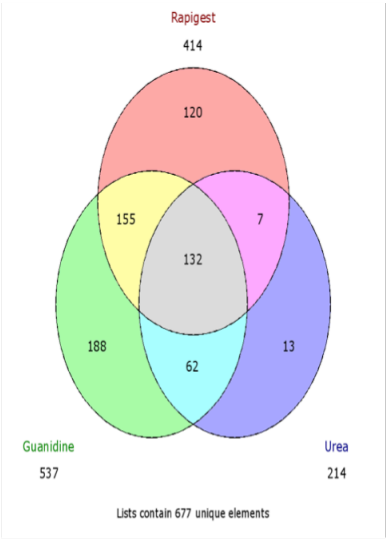


Figure 4a.

Equine Venn Diagram.

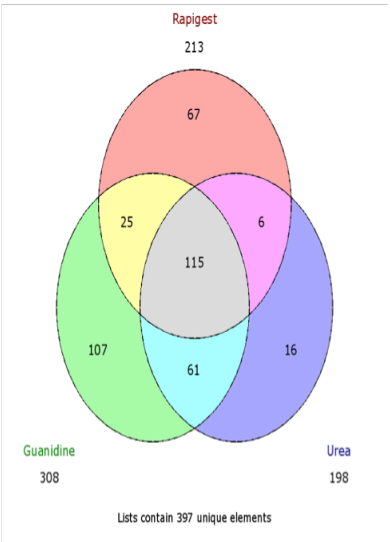


Figure 4b.

Figures 4a. Venn Diagram of Human proteins. Figure 4b. Venn diagram of equine proteins. Shows the average number of significant proteins after adjustment to 1% FDR in MASCOT search against Unihuman and Unihorse databases respectively. Results revealed that the guanidine method extracted the highest amount of unique proteins, and urea the least amount in both human and equine samples.

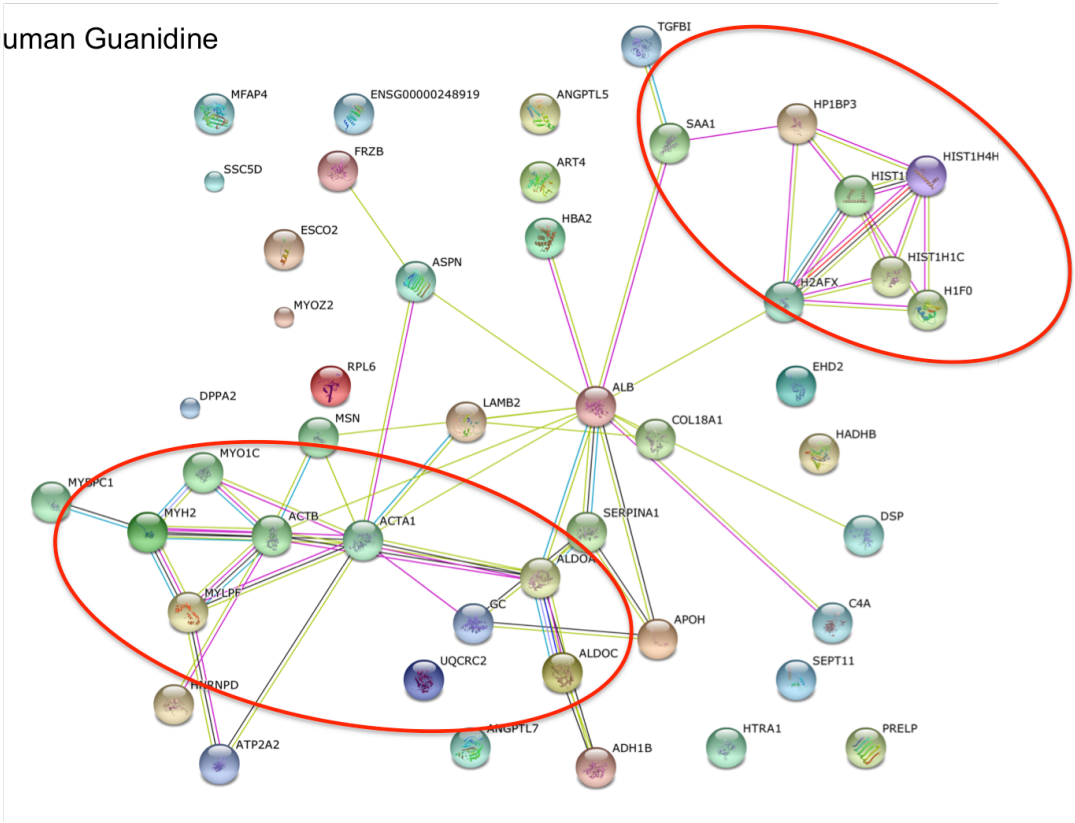


Figure 5a-d. Protein-protein interaction network visualised using the bioinformatics tool STRING. Proteins were filtered to include at least two unique peptides and a FDR<0.01. These proteins were specific to the guanidine protocol. A red ellipse on each figure encloses the most stable and connected interactions within each data set. Human guanidine data revealed clusters involving intracellular proteins such as histones and ribosomal proteins.

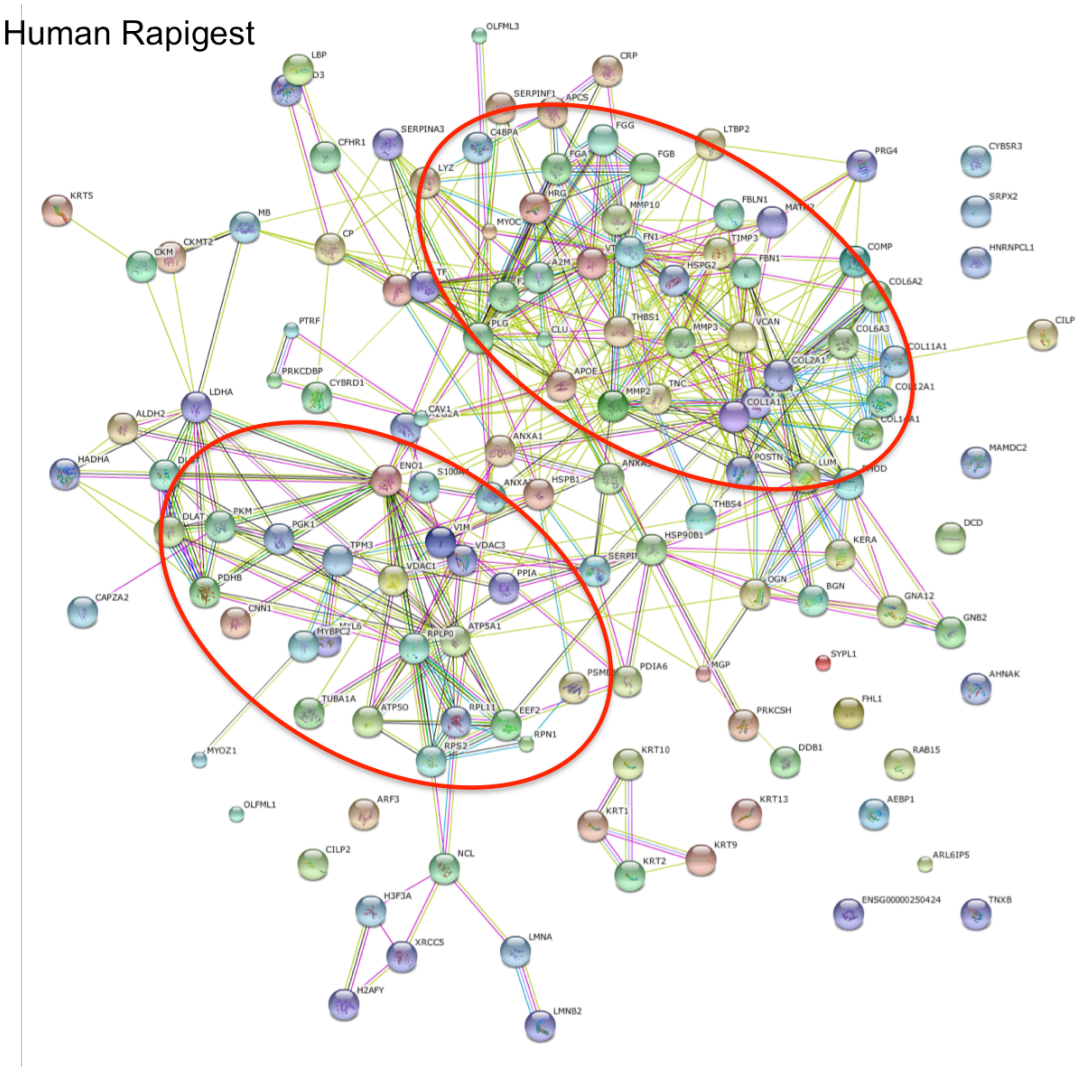


Figure 5b. This interaction network revealed proteins extracted in human samples via Rapigest SF. Clusters can be seen to involve extracellular proteins such as collagens, lumican and fibronectin.

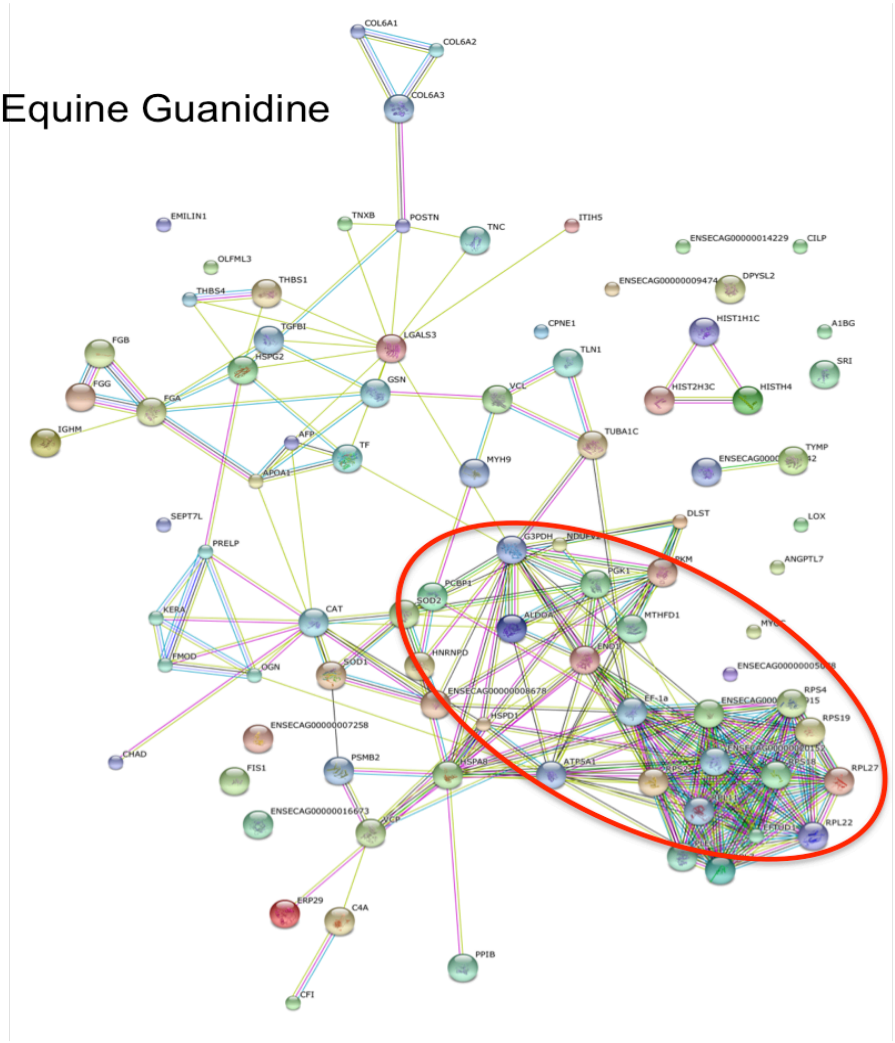


Figure 5c. This interaction network revealed proteins extracted in equine samples via guanidine, this data clusters heavily with intracellular proteins such as histones and ribosomal proteins.

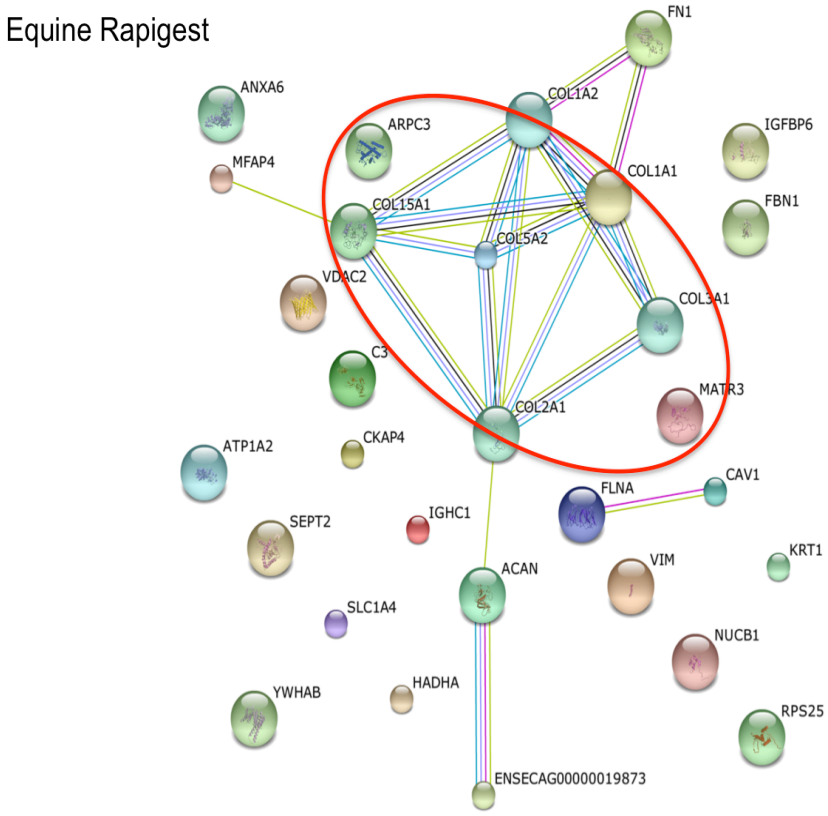


Figure 5d. This interaction network revealed proteins extracted in equine samples via Rapigest SF. Interactions between collagen extracellular proteins are seen here.

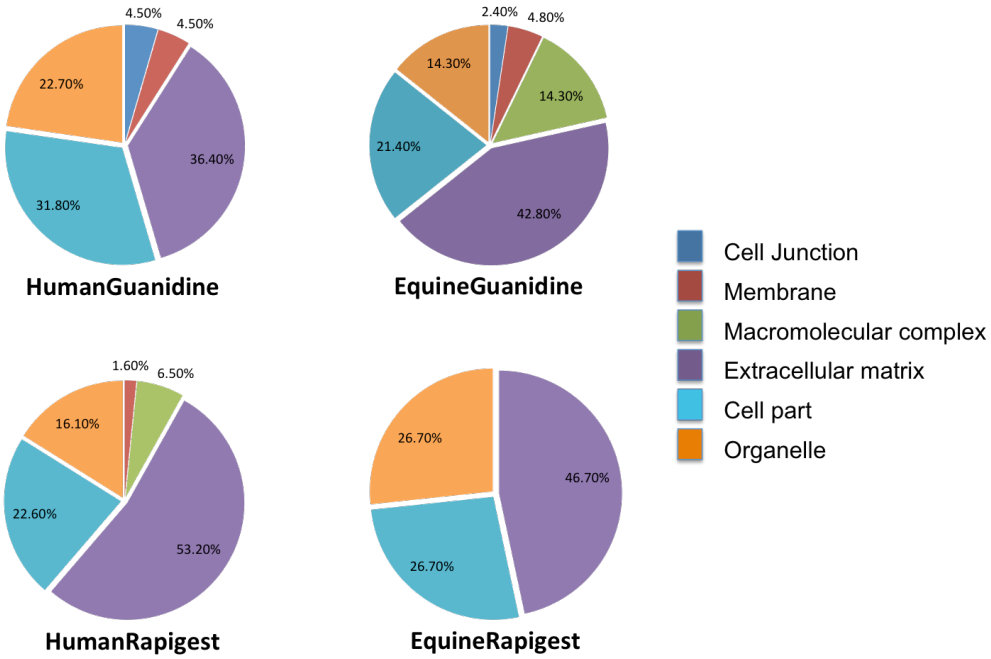


Figure 6. Using the gene ontology tool PANTHER, proteins were classified according to their location. Differentially expressed proteins with at least two unique peptides and a FDR 1% were included in this data set. Profiles of the guanidine and Rapigest SF datasets were species independent.

Discussion

The efficiency of enzymatic digestion can vary from protein to protein; cleavage site access may be inaccessible to enzymes depending upon the structural conformation of the protein. Denaturisation is the disruption of the internal bonds that give the proteins their quaternary, tertiary and secondary structures. This is advantageous prior to proteolysis, as the digestion enzyme has improved contact with cleavage sites, leading to a more efficient digestion.

During this study two different methods of protein denaturation were investigated by comparing two chaotropic agents and one detergent. There are various protein extraction techniques described in the literature (Norrgran *et al*, 2009; Peffers *et al*, 2014; Wilson *et al*, 2010) however, to date, no comparison study has been undertaken which compares these methods in tendon tissue.

Three agents were chosen, guandidine and urea, which are chaotrophic agents and Rapigest SF which is a detergent.

Protocols were devised to compare the solubility of proteins alongside their abundance between conditions, methods also need to be compatible with techniques of analysis and detection. For this tissue, which is largely composed of extracellular matrix, we expect there to be mostly collagen proteins, however to learn about the biological processes within the tissue we need to devise methods to detect both the non collagenous components of this matrix, such as proteoglycans, and also to detect cellular proteins. Due to the fibrous composition of tendon tissue, this tissue is highly resistant to extraction (Peffers *et al,* 2014; Wilson *et al,* 2010). Extrapolation of methods from previous studies in musculoskeletal tissues was deemed an appropriate starting point. Previous studies have used the chaotropic agent guanidine within an extraction buffer, with success, for protein extraction in tendons and cartilage (Peffers, *et al*, 2014. Wilson *et al*, 2010). Urea was chosen as an alternative chaotropic agent to trial. Drawbacks to these agents are that, at the concentration required to denature proteins they will also reduce proteolytic activity of the digestion enzyme so samples must be cleaned up and concentrations of these agents diluted before digestion occurs (Proc *et al*, 2010). Regarding the use of urea, when exposed to heat, degradation into ammonium cyanate can occur which leads to the carbamylation of free primary amines in the sample being analysed which needs to be avoided (Proc *et al,* 2010). Although heat alone is a powerful denaturant, both the guanidine and urea extraction protocols avoided its use for reasons described above, along side this, protein degradation could occur which needs to be avoided. A previous comparison study revealed that the use of heat alongside guanidine gave no benefit to the end results. Interestingly, in the same study, heat improved the efficiency of extraction using urea (Proc *et al*, 2010).

Finally the surfactant Rapigest SF was chosen for the third protocol. Rapigest SF is the brand name for the anionic detergent sodium 3-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl)methoxy]-1-propanesulfonate. This agent acts as a denaturant by solubilising and unfolding the substrate protein and has been shown to have several advantages. These benefits include: - not interfering with mass spectrometry analysis, does not suppress peptide ionisation and does not modify the proteins and peptides (Yu *et al*, 2003).

For all three methods, samples were pre-treated with chondroitinase ABC for deglycosylation. Glycosylation is a post-translational modification in proteins that makes the identification of glycoproteins challenging during mass spectrometry analysis. As discussed in the introduction, proteoglycans are one of the principal components of the non-collagenous matrix in tendon tissue. Proteoglycans are comprised of a core protein plus one or more glycosaminoglycan (GAG) side chains; previous studies have used chrondroitinase ABC to effectively remove these side chains (Fessel and Snedeker, 2011).

The effectiveness of each method was ultimately assessed on the number of identified peptides reported following LC MS/MS analysis. Originally, before filtering, Rapigest SF samples revealed higher signal intensities. Using Progenesis QI software to filter the data and threshold the identified proteins to a 1% FDR identification rate with at least a two-fold change in expression, the guanidine HCL extraction yielded the greatest number of unique identified proteins and urea the lowest. To increase confidence levels for the identification of proteins, the data was further filtered to only include proteins with at least two unique peptides and a FDR less than 1%. Applying these more rigorous acceptance thresholds reduced the number of proteins identified following urea extraction to zero. Due to the poor performance of urea extraction, subsequent bioinformatic analysis were only performed on the guanidine and Rapigest SF samples.

Proteomic profiles of the guanidine and Rapigest SF datasets appeared to follow the same pattern despite them being from two different species. Guanidine samples had a higher amount of intracellular proteins such as ribosomal proteins and histones. Rapigest SF samples tended to have a roughly equal amount of intracellular and extracellular matrix proteins, such as collagens and proteoglycans. Further study with a higher number of biological repeats would be needed to confirm whether this was a true representation.

Possible explanations for my results are as follows: - both guanidine and urea are classed as chaotrophic agents, meaning that they disrupt the non covalent bonds within the proteins tertiary structure, it is possible that the poor extraction efficiency of urea can be explained by a lower ability to solubilise the lipid bilayer of membranes and subsequently cell contents are not released. This is in stark comparison with guanidine, which showed a high percentage of cellular components. Another possible speculation is that urea causes a higher proportion of hydrophobic residues during the unfolding process, which then lead to an increase in protein aggregation and a subsequent reduction in identification hits.

There are limitations to the above study. Statistical analysis could not be performed due to the small sample size. Further work would include replicates in order to test repeatability. Also due to the small sample size no comparisons could be made between subsets; in particular, how protein profiles may vary between young and old samples and between different tendon types. Following these preliminary results further work could also expand on these findings to see whether combining methods could increase peptide recovery and overall coverage of the tendon proteome. For instance it would be interesting to combine extraction using guanidine and Rapigest SF. If protein extraction was performed by Rapigest SF, the supernatant removed and set aside, then on the same pellet continue with the guanidine extraction protocol, once completed the two protocol supernatants could be combined and analysed with LC MS/MS. Different concentrations of Rapigest could also be trialled to see whether the same effects could be achieved with lower amounts.

Once an optimal extraction protocol has been devised, it would be beneficial to look at different age groups as previous studies have found differences in protein profiles. One study found a reduction in the levels of several SLRPs and an increase in intermediate filament proteins were found as a result of ageing (Peffers *et al,* 2014). These proteins are involved in matrix organisation and regulation of cell tension. In support of this, another study detected a reduced

amount of neopeptides in aged IFM possibly indicating reduced turnover (Thorpe *et al,* 2016) however the protein composition of both the IFM and FM did not show large differences (Thorpe *et al,* 2016). If we can enhance the process of extracting proteins from tendon tissue this can only lead to better discoveries in the field of tendon proteomics as by establishing compositional changes between age groups, this will provide insights into tissue mechanics and potentially target areas for diagnostic/therapeutic methods

Chapter 3, Mass Spectrometry Imaging (IMS) of tendon tissue.

Introduction

Proteomic analysis of tendons is a field that has received limited attention. Gaining a clearer insight into the composition of tendons will in time have great consequences for cell based therapies and tissue engineering strategies. Using Liquid chromatography Mass spectrometry/Mass spectrometry (LC-MS/MS) provides label free analysis of our biological samples. However LC-MS/MS analysis ignores the spatial distribution of molecules, visualisation of the protein distribution allows the investigation of differentially expressed proteins in these areas (Cobice *et al,* 2015). The introduction of MS imaging allows questions regarding distribution patterns/ functions of molecules to be answered. Matrix assisted laser desorption/ionization (MALDI) mass spectrometry allows the spatial analysis of molecules <100,000Da whereas secondary ion mass spectrometry (SIMS) provides analysis of the elements and lower molecular mass compounds, <1000Da (Todd et al., 2001). The application of these methods together can attain a considerable amount of information on the biological functions of the tendon tissue and how alterations may occur during ageing, disease and between different types of tendon.

Since the invention of MALDI-MS several crucial preparation steps of the biological samples have been identified to optimise data collection. Developing protocols that have a high level of uniformity and reproducibility are vital in the development of future clinical tests and therapies. In a complex sample such as tendon tissue the identification of noteworthy features can be hidden beneath higher abundance molecules. There are many high abundance proteins, which may contain post-translational modifications, that can interfere with a sample affecting the analysis. Sample clean up steps are recommended to reduce the concentration of interfering substances. As there is no published data on the imaging of tendons it was necessary to optimise and adapt protocols previously used in other tissue types. This chapter will detail the major details of sample preparation according to the current literature; these measures can then be used as a starting point for the investigation of a novel tissue sample.

Aims and objectives

To devise an appropriate protocol for tendon sample preparation for MALDI imaging. Following this to employ a multimodal approach for the visualisation of unlabeled peptides, lipids and small molecules at cellular and subcellular levels of tendons at different risk of injury and ages.

Materials and Methods

All reagents were obtained from Sigma-Aldrich unless otherwise stated. To minimise non-protein contamination, high quality reagents HPLC grade solvents and all glass ware were pre-rinsed with high quality water. Buffers were made on the day. Simplified flow diagram of methods used can be viewed in figure 3.

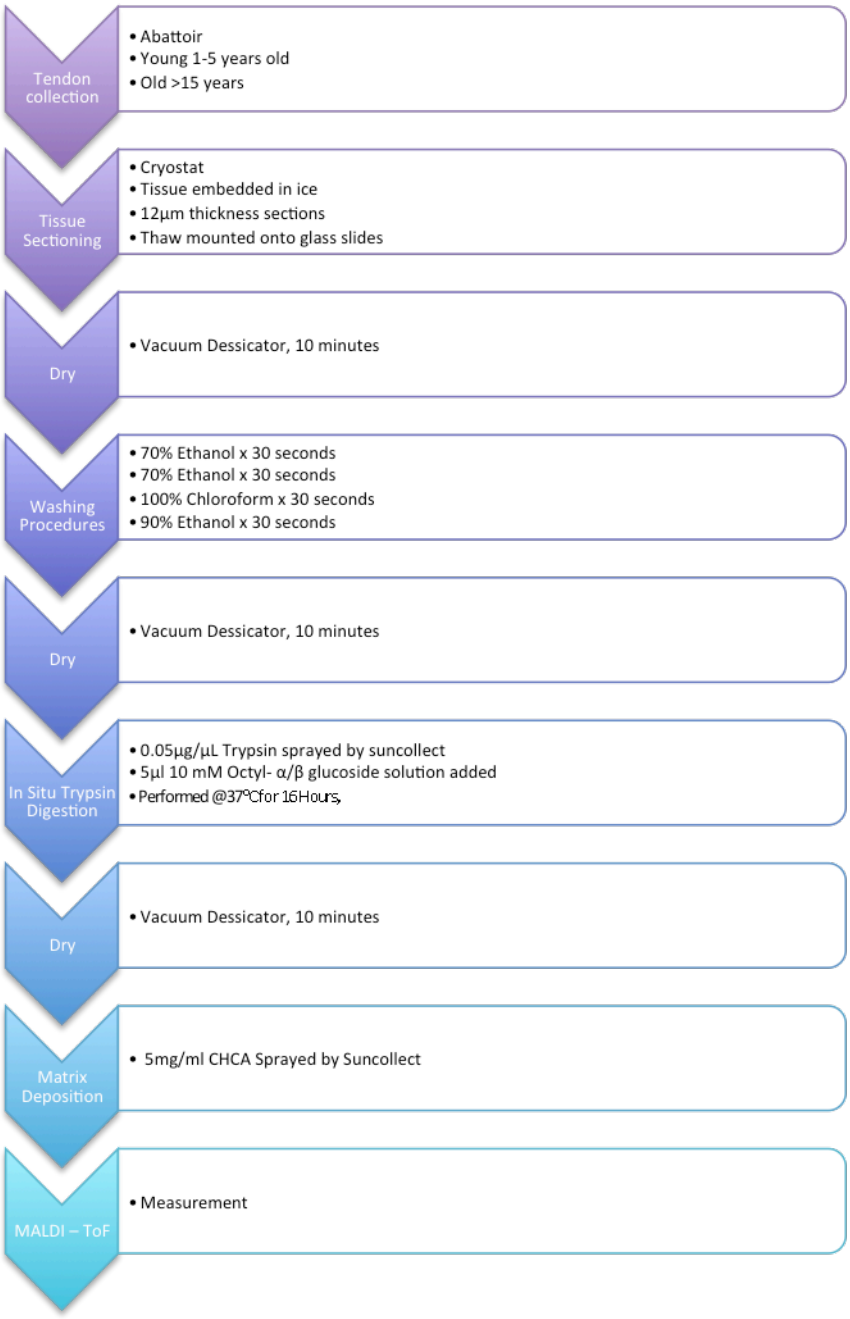


Figure 3. Flow diagram of the methods used during the preparation of tendon tissue for MALDI imaging.

Sample collection

Ethical approval was not required for this study as samples were obtained from a commercial equine abattoir. No previous medical histories for the horses were available. Forelimbs, distal to the carpus were collected from Thoroughbred or Thoroughbred crosses. A sample of both the SDFT and CDET were dissected in duplicate from the mid metacarpal region from young skeletally mature

(<4years) and old horses (>15years). Only tendons that had no evidence of previous tendon injury at post mortem examination were included in this study. Tendon sections were then either snap frozen in liquid nitrogen and stored at −80 °C until further analysis or stored in 4% paraformaldehyde solution for future histological analysis.

Sample preparation for SIMS

Tissue, which had previously been stored at -80˚C, was transferred to the Cryotstat machine (Leica CM1860UV); the stage holding and blade temperature were set to -21˚C. Tendon samples were attached onto a sample holder with ice. Thickness of a sample from each category were cut to 12 µm and thaw mounted onto indium tin oxide (ITO) coated high conductivity glass slides (Delta Technologies, CO, USA) and frozen at - 20˚C for later use.

SIMS Experiment

Samples were transferred to a desiccator box, defrosted and dried in a vacuum desiccator to prevent the molecular de-localisation. SIMS experiments were performed in positive and negative ion modes on a Physical Electronics (Chanhassen, MN, USA) TRIFT II secondary ion mass spectrometer with a 30 keV Bi_3^{2+} primary ion beam. Areas of 8.4 x 6.8 mm were analysed using the mosaic mode of 21 x 17 tiles, each 400 x 400 µm in size containing 256 x 256 pixels per tile (lateral resolution is 1.56 µm per pixel). The acquisition time was 24 seconds per tile.

SIMS Data Interpretation

Raw data were converted to Matlab format and then spatial binning, spectral binning and peak picking were performed. Principal component analysis (PCA) was used to look for spectral similarities and differences between the samples using Maastricht University in house built ChemoneTricks Toolbox for MATLAB version 2014b

Sample preparation for MALDI Imaging

Sample cutting

Tissue which had previously been stored at -80˚C was transferred to the Cryotstat machine (Leica CM1860UV). Tendon samples were carefully mounted onto a sample holder and embedded in ice. The use of optimum cutting temperature (OCT) polymer was avoided to prevent suppressed ion formation and intensity during MALDI MSI analysis. Thickness of a sample from each category was cut to 12 µm and thaw mounted onto glass slides. After sectioning the samples were stored at - 20˚C for later use. For optimisation purposes SDFT sample from a horse >15 years old was used.

Tissue washing

Washing steps prior to on-tissue trypsin digestion were undertaken to remove lipids, salts and other environmental factors that would reduce the efficiency of enzymatic digestion. This was achieved by carefully washing the sample twice with 70% ethanol for 30 seconds. Successively the sample was washed with 100% chloroform for 30 seconds and then 95% Ethanol for 30 seconds. These steps act to remove lipids that interfere at the low mass range. Samples were then dried in a vacuum desiccator for 10 minutes.

On-tissue Tryptic digestion

Trypsin solution 5ug/ml was reconstituted in either deionised water or 25mM Ammonium Bicarbonate (AMBIC), both contained 10mM n-Octyl β-D-glucopyranoside (OcGlu) Tissue was then coated with a homogeneous layer of trypsin solution with the SunCollect™ automatic sprayer (SunChrom, Friedrichsdorf, Germany) or the trypsin was manually applied. Suncollect settings used can be found in Table 1. Whilst spraying was carried out, continued observation was necessary to ensure spraying was consistent with settings. Digestion was performed either over 4 hour or 16 hours. To prevent samples from drying out, they were maintained in a wet/humid environment. This was achieved by incubation in a parafilmed container at 37˚C in a saturated atmosphere of MeOH/H_2O (1:1 v/v). Following digestion, samples were again dried in a vacuum desiccator for 10 minutes.

SunCollect settings for enzyme application	
Vial X	0.4mm
Vial Y	2mm
Z Value	40mm
Z offset	0mm
Number of layers	8
Flow rate Layer 1	10µL/min
Flow rate Layer 2	15µL/min
Flow rate Layer 3	20µL/min
Flow rate Layer >4	25µL/min
Speed X	Low (4)
Speed Y	Medium (1)

Table 2. Suncollect settings for enzyme application.

Matrix application

The matrix solution, α-Cyano-4-hydroxycinnamic acid (CHCA) (5mg/ml diluted in 50% acetonitrile and 50% water containing 0.2% trifluoroacetic acid, was sprayed on top of the tissue section with SunCollect™ automatic sprayer (SunChrom, Friedrichsdorf, Germany). Two spraying conditions were trialed with this application, Suncollect settings can be found in Table 2 and 3. Manual application was also tested.

SunCollect settings for CHCA Matrix Application 'Dryer'	
Vial X	0.4mm
Vial Y	2mm
Z Value	30mm
Z offset	0mm
Number of layers	8
Flow rate Layer 1	15µL/min
Flow rate Layer 2	15µL/min
Flow rate Layer 3	20µL/min
Flow rate Layer >4	20µL/min
Speed X	Low (4)
Speed Y	Medium (1)

Table 3. SunCollect Settings for Dryer Matrix Application.

SunCollect settings for CHCA Matrix Application 'wetter'	
Vial X	0.4mm
Vial Y	2mm
Z Value	40mm
Z offset	0mm
Number of layers	8
Flow rate Layer 1	10µL/min
Flow rate Layer 2	15µL/min
Flow rate Layer 3	20µL/min
Flow rate Layer >4	25µL/min
Speed X	Low (4)
Speed Y	Medium (1)

Table 4. SunCollect Settings for Wetter Matrix Application.

Experiments run

The initial phase in this study was to trial a method that has been successfully used in other tissue types. This experiment compared young and old SDFT sections to a non-digested control sample. The SunCollect spray was used for both enzymatic application and the matrix application. Suncollect settings for the matrix application were set to 'dry' parameters.

In order to improve the spectra obtained, a series of experiments were undertaken to evaluate the optimal sample preparation protocol. To achieve optimisation of the *in situ* digestion, different techniques were tested. The experiments can be viewed in table 1.

Sample Number	Trysin + Ammonium Bicarbonate	Sample Number	Trypsin + ddH_20
1	Manual Trypsin application Manual Matrix application	4	Manual Trypsin application Manual Matrix application
2	Manual Trypsin application Suncollect Matrix application	5	Manual Trypsin application Suncollect Matrix application
3	Suncollect Trypsin application Suncollect Matrix application	6	Suncollect Trypsin application Suncollect Matrix application

Table 1. Six experiments performed to optimise trypsin reconstitution. These experiments also determined differences between manual and Suncollect application of both enzyme and matrix.

MALDI Imaging

Digital optical scans of tissue sections were taken prior to MALDI-IMS experiments using an HP scanner. The digital images were then imported into the MALDI imaging software HDImaging V4.1 (Waters Corporation, Manchester, UK) which was utilised to select the area to be sampled. The horizontal and vertical step sizes between each sample spot were set at 100 µm. Instrument calibration was performed using a red phosphorous standard, in positive ionisation mode between *m/z* 100 to *m/z* 5000. Data was acquired in sensitivity mode, positive polarity and in the 100 to 3500 Da mass range. The quadrupole profile was set to 900Da and other MS setting were set to the manufacturers default values.

Data Analysis

As full analysis of data is very time consuming a broad overview of the spectra generated from different conditions were compared and contrasted via the biomap (Novartis, Switzerland) software. This was to enable a modification of the final protocol without prolonged amounts of time performing in depth data analysis.

For full analysis, acquired mass spectra were combined into a single image-file and then processed via Masslynx (Waters), mMass [30] and BioMap (Novartis, Switzlerland) software. Spectra were processed in MassLynxTM (Waters Corporation) by means of peak smoothing and baseline subtraction. Ion images were generated with the Biomap 3.8.0.4 software (Novartis Pharma AG, Basel, Switzerland). The intensity of all m/z channels were normalised, autoscaled and peak picked. The peak picking was performed on peaks with threshold S/N (i.e. signal to noise ratio) ≥2. Following this Principal Component Analysis (PCA) and Discriminant Analysis (DA) were applied to reveal and visualise the variability within the data.

Results

The first experiment assessed old SDFT sections to a non-digested control sample, results revealed similar spectra and are displayed in figure 4.

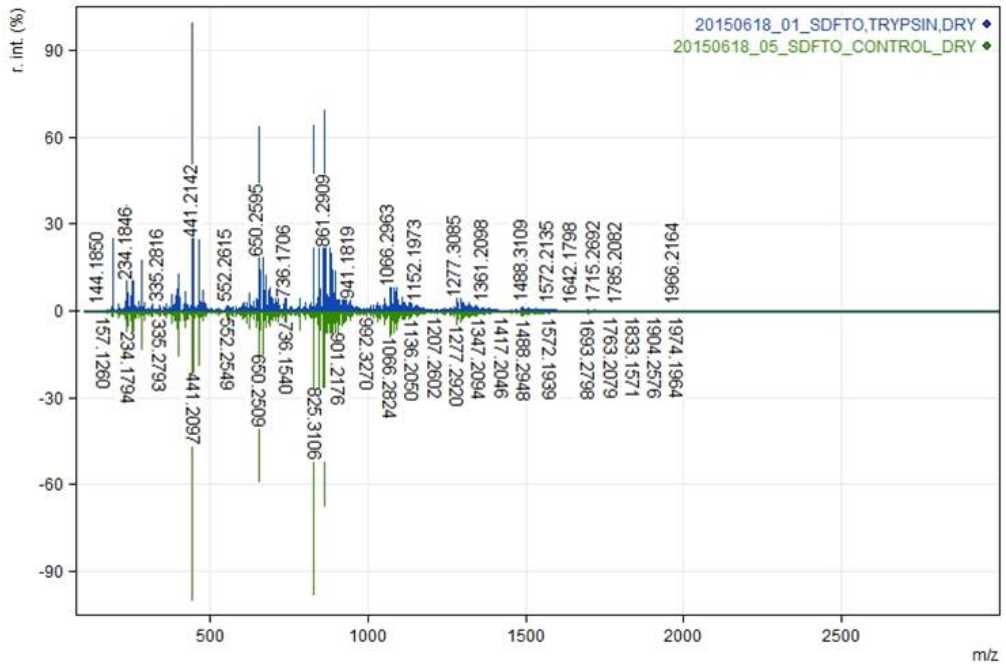
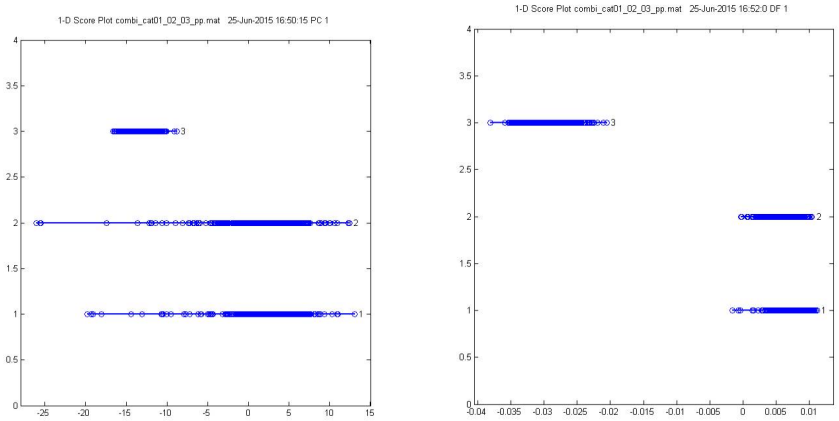


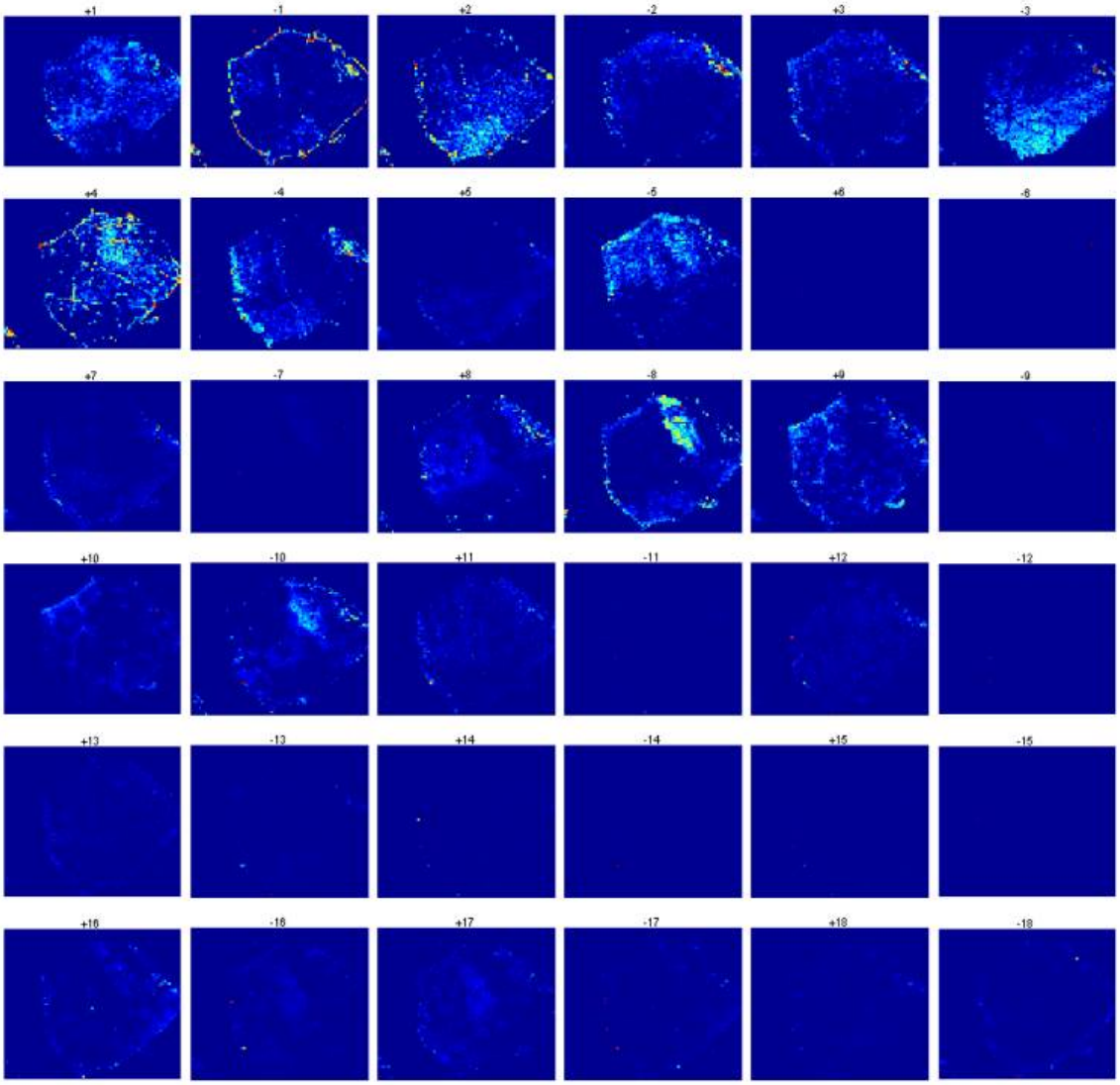
Figure 4. Using biomap software. Top spectrum represents in situ trypsin digestion with Dry Matrix settings compared to control sample. The two spectra are similar.

Secondly a comparison was made between the old and young subjects. Principal component analysis (PCA) and discriminant analysis (DA) were applied to see whether the two samples revealed any molecular differences based on the tryptic peaks that had been generated. Results can be seen in figure 5.



(A) **(B)**

Figure 5. Sample 1 = young SFDT. Sample 2 = Old SDFT and Sample 3 = control i.e. non digested SDFT (A) PCA plot, this plot aims to illustrate any variation between the undigested control and the digested SDFT old and young samples. From this plot, samples 1 and 2, show a great degree of variation in points exhibiting the digestion process as there are a larger number of peptides in these samples. However the difference between the two samples shows no significant differences. (B) Discriminant analysis, this plot aims to characterize or separate the two groups. Here samples 1 and 2 reveal no differences between old and young samples but a difference is seen between the digested and non digested control.



(C)

5 (C) MALDI Image SDFT Young – Mosaic image showing the first 18 principal components (PCs). The different PCs show the intra-tissue heterogeneity of the SDFT group.

The PCA and DA plots reveal that digestion has occurred however there was not a significant difference between the age groups.

Enzyme application optimisation

Trypsin reconstituted in 25mM AMBIC was compared to trypsin reconstituted in deionized H_20 (figure 6) Manual application was compared to SunCollect application (figure 7). The optimal digestion time was also studied; 4 hour compared to 16 hour digestion times (figure 8)

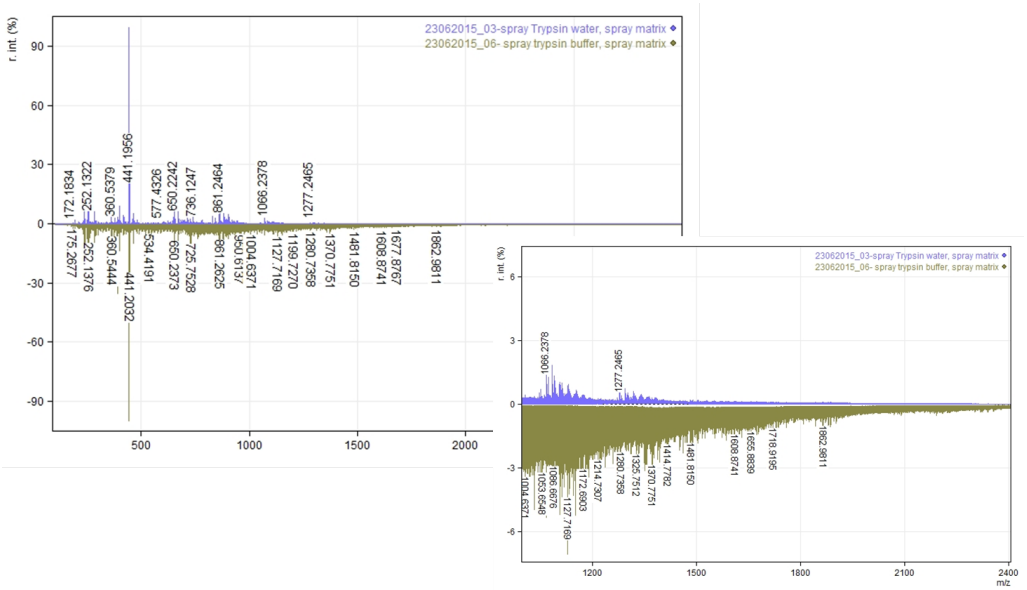


Figure 6. Comparison of spectra when trypsin was reconstituted up in 25mM ammonium bicarbonate (buffer) compared to the use of ddH_2O. The top spectra reveals the highest intensity and an increased number of peaks showing that this is the superior method. Matrix was applied by the Suncollect machine on the wet settings for both.

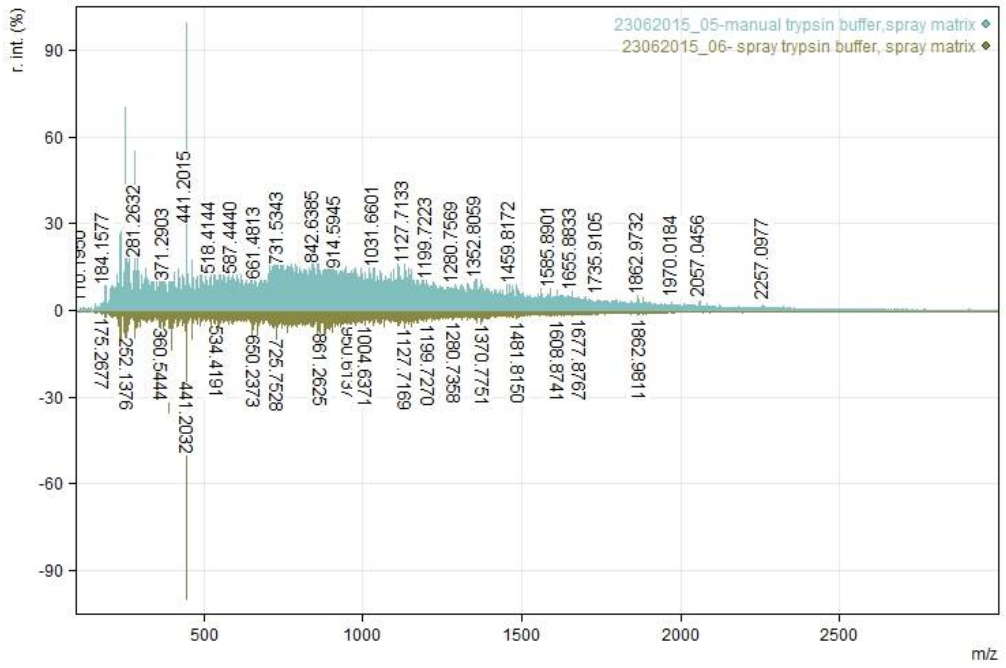


Figure 7. Manual vs sprayed trypsin in buffer. Matrix was applied using the SunCollect on the wet settings. Top spectrum represents manual application compared to Suncollect application. Higher intensity peaks are seen throughout when trypsin was applied manually.

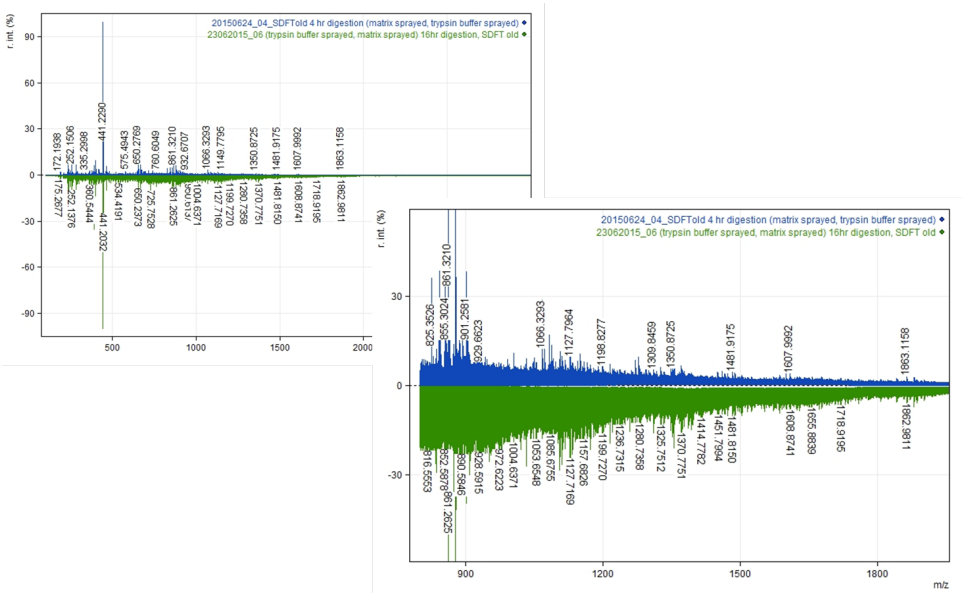


Figure 8. Comparison of digestion times. Top spectrum represents in situ trypsin 4 hour digestion (reconstituted in buffer) compared to 16 hour digestion. The longer, 16 hour digestion, reveals the highest peak intensities.

Matrix application optimisation

Suncollect settings were modified to ascertain whether this improved co-crystallization and gave an increased number of protein identifications. Dryer vs wetter settings are compared in figure 9.

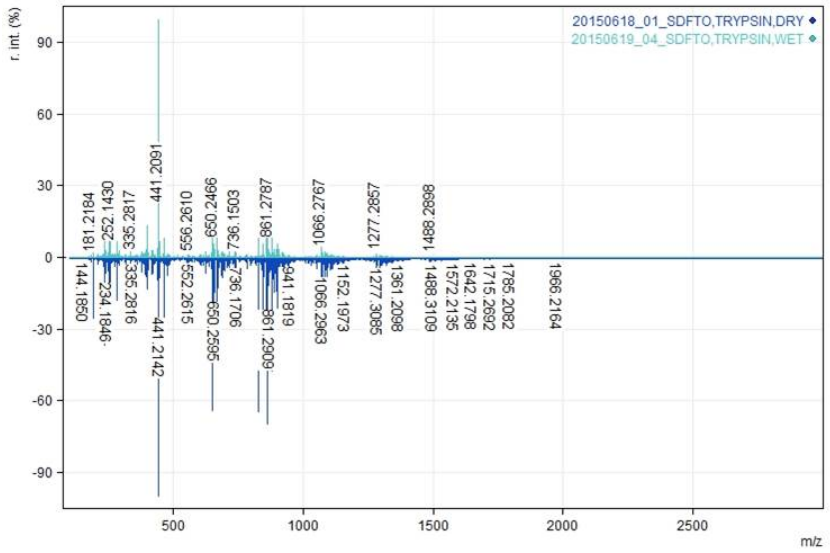


Figure 9. Comparison of SunCollect matrix application settings. Top spectrum represents 'dryer' settings in comparison to the 'wetter' settings. In this

comparison the dryer settings looked marginally better, as hits were identified in 1500-2000 m/z region.

SIMS Analysis

All samples were measured in both positive and negative ion modes. Most common peaks 23 (sodium), 39 (potassium), 104, 86, 184 (all originating from phosphocholine, a fragment of phospholipids), 369 (cholesterol). In negative mode, 79 is the highest, corresponding to phosphate, but also 281 and 255 which are free fatty acids. Which is demonstrated by figure 11. No significant differences between the four groups were observed.

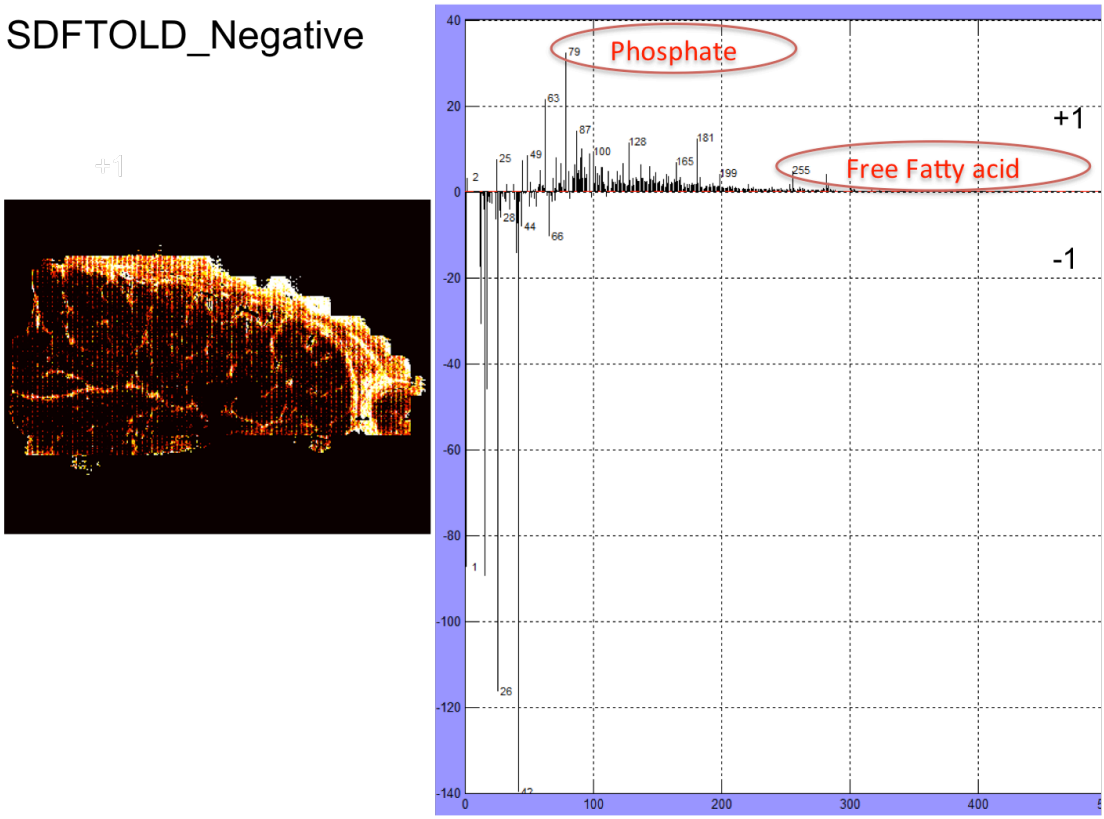


Figure 10. Negative ion mode SIMS image and corresponding spectra of old energy storing tendons (SDFT). Spectra generated came from the interfascicular matrix region. No significant differences were detected between samples.

Adjacent sections were prepared with haematoxylin-eosin (H&E) staining for optical microscopy, this enabled the verification of where the interfascicular matrix was located and revealed artifacts in the tissue sections which assisted with accurate data analysis. CDET young sample showed slightly different spectra; most likely due

to artifacts within the tissue caused by the freeze thaw process. However looking over all four groups, the peaks detected revealed no significant difference.

Discussion

In this study we investigated the use MALDI and SIMS as a novel method for the proteomic analysis of tendon tissue of different ages and types. Results demonstrated for both techniques that sample preparation protocols have yet to be fully optimised.

The first objective of this study was to determine the most efficient sample preparation for MALDI imaging of tendon tissue. The data was processed pre-processed and visualized using Masslynx (Waters), mMass and biomap (Novartis, Switzerland) software. Firstly a comparison, using biomap software, was made between an *In situ* trypsin digestion of an SDFT old tendon and the non-digested control. Results revealed similar spectra so different conditions were trialed in the attempt to reveal the peptides we were looking for.

For MALDI-IMS investigations revealed:

1) Trypsin reconstituted in 25mM Ammonium bicarbonate was seen to generate more peaks compared to deionised H_2O.
2) Although manual application of trypsin yielded better results than the SunCollect application for comparable results the robotic sprayer is superior as the repeatability was considered to be more consistent. As spectra seen revealed increases in intensity of peaks identified.
3) The optimal trypsin digestion time was found to be 16 hours.

During this study, no differences were detected between tendon types and age groups. It is likely that methodologies were not fully refined however it is also a possibility that results obtained are true findings. Further work needs to be undertaken to optimise sample preparation before full analysis between sample types can be undertaken. Potential areas of improvement are in situ tissue digestion, matrix application and data collection.

Determining methods to maximize the efficiency of this *in-situ* tissue digestion should increase the amount of peptides visualised. By performing on-tissue digestion more protein identifications may be made, while preserving the spatial distribution of proteins. Currently the most commonly used enzyme is trypsin, it generates peptides compatible with MS by cleaving mainly at the carboxyl side of the amino acids lysine or arginine. Studies have been undertaken to improve in situ proteolysis and research has revealed that digestion is improved by the addition of a detergent (Patel et al, 2015). In this study n-octyl glucoside was used to aid digestion of hydrophobic proteins as these proteins can be resistant to enzymatic digestion due to inaccessible cleavage sites. An alternative detergent that could be trial is *N*-Octanoyl-*N*-methylglucamin (MEGA-8) as this has also been shown to achieve efficient proteolysis and is compatible with mass spectrometry. Another possibility is to include the addition of Rapigest SF (Waters Corporation, Manchester, UK) to the *in situ* digest, this has previously been used with success in trypsin digestions. One study demonstrated that when RapiGest was used in conjunction with trypsin digestion an increased number of peptide peaks were generated in comparison to experiments with trypsin digestion alone was used (Patel et al, 2015). Previous work has also demonstrated that when using this surfactant there is no interference with MS analysis. In the previous chapter Chrondroitinase ABC was used to remove glycosaminoglycan (GAG) side sides as they are difficult to analyse and can cause ion suppression of other molecules. One of the reasons this occurs is due to the negative charge that these side chains carry (Thanawiroon *et al,* 2004). A pretreatment with chrondroitinase ABC could be considered. Finally an alternative enzyme such as collagenase could be considered, as tendons are primarily composed of collagens this may enable a greater number of peptides to be generated. Proteins would then be identified on MS/MS fragmentation pattern.

The type of matrix chosen is influenced by the size of the analyte. Commonly sinapinic acid (SA) is used for the analysis of larger peptides and proteins (>10kDa) (Meetani and Voorhees, 2005) α-Cyano-4-hydroxycinnamic acid CHCA

is often used for the detection of smaller peptides (Chughtai and Heeren, 2010). By substituting CHCA with SA, whole proteins would be detected which would eliminate the need for a digestion step.

Data was acquired in sensitivity mode, positive polarity and in the 100 to 3500 Da mass range. The majority of peaks were observed below m/z 1000, there was a substantial amount of these peaks that originated from the matrix molecules themselves. Improvements could result from changing the mass range in which data is acquired, e.g. in the 900-3500Da mass range. This could reduce background spectra and possibly allow a detection of peptides in the higher mass range.

For SIMS Analysis, Mass spectrums presented were focused on the interfascicular matrix. The abundant peaks were fragments of phospholipids, which represent a major class of lipids and are a component of the cellular membrane. Other peaks detected were of fatty acids and cholesterol, again these are components of the cellular membrane. These results correspond with a recently published study comparing proteomic profiles of the IFM and fascicular matrix (FM) it was demonstrated that the IFM contained more proteins and protein fragments (neopeptides) in comparison with the FM (Thorpe *et al,* 2016). Using label free quantification this study showed that the majority of proteins with higher abundance in the IFM were cellular proteins compared to the FM which contained a higher abundance of ECM proteins (Thorpe *et al,* 2016).

There was reduced signal on the CDET young sample potentially due to artifacts created during the freezing process. It has been described that different parts of the tissue can cool at different rates, this can lead to the formation of ice crystals that then lead to fragmentation of the sample to be analysed (Chughtai and Heeren, 2010). During this investigation a protocol of snap freezing the sample in liquid nitrogen was used in order to try and prevent this occurring. This not only prevents the sample from biological degradation, it also preserves tissue morphology. Extra samples were taken for validation of MS by histological

staining, so that the morphological information could be correlated between the two techniques, this proved that there was fragmentation within the tissue and removal of background spectra was easier when SIMS images were correlated with H and E stained images of tissue.

A brief comparison of the four sample groups revealed little differences between the spectra. As there was only one sample representing each group a general conclusion cannot be made. It is possible that the results seen in this study are a true impression. However, looking at previous studies, one investigation revealed a decrease mRNA expression of proteoglycan 4, which was confirmed by immunostaining (Kostrominova and Brooks, 2013). In conjunction with this, other research groups have detected differences in the proteomic profiles between these types of sample groups (Peffers, 2014), it would be advisable to alter sample preparation and to look again.

To reduce fragmentation of the sample would be difficult as all attempts were made to avoid this, however, an increase in sample numbers in each group would give more accurate overview. Another option could be to employ metal assisted SIMS by depositing a gold coating to the sample tissue with the hope of increasing ion yields of higher mass molecules (McDonnel and Heeren, 2006).

Further work is necessary in this area and suggestions have been made. Another avenue that could be explored is the use of laser microdissection to selectively isolate the region of choice ie the interfascicular matrix, then using methods discussed in chapter 2 to perform Liquid chromatography mass spectrometry/mass spectrometry LC MS/MS. This study has already been undertaken and results were discussed previously (Thorpe *et al*, 2016) however future analysis via MALDI–IMS or SIMS could also be performed on isolated sections once optimum sample preparation methods are defined.

Chapter 4, Dialysis

Tendinopathy is a broad term use to describe a tendon disorder characterised by a combination of pain, swelling and impaired performance. Tendinopathies are often associated with overuse tendon injuries and pose a major problem for the human and equine patient. There is some debate concerning the role of inflammation in tendon injury.

Some investigations have indicated little or no inflammation present in tendons exposed to overuse (Andres and Murrell, 2008; Khan *et al*, 1999; Wilson and Best, 2005). This could be a true finding or could be inaccurate due to the time of patient presentation, poorly sensitive techniques or the definition of inflammation based purely on cellular components (Dakin *et al*, 2013). Whilst inflammation is necessary after injury, if this process persists, fibrosis is thought to occur which can add to the risk of subsequent reinjury. The development of effective protocols in managing the injured tendon, and also preventing reinjury, is dependant upon the full understanding of tendon pathology and the role of inflammation. Many treatment options include the use of anti-inflammatories, firstly to manage the associated pain but also to reduce the harmful effects inflammation causes however there are known adverse consequences (Shier *et al,* 1996). Therefore this treatment option may or may not be the most effective choice of treatment. Further research is needed into the mechanisms of disease pathogenesis, therapeutic interventions and their efficiencies.

There are a number of alternative conservative treatment options available to the physician each having various efficiencies, one of which is extracorporeal shock wave therapy (ESWT). During this treatment, acoustic shockwaves are passed through the skin to the affected area. The mechanism by which this treatment works has not been fully determined, but it is thought to accelerate the healing process by increasing blood flow to the area; by having a direct suppressive effect on nociceptors, hence causing an degree of denerving and by activating interstitial and extracellular biological responses (Notarnicola and Moretti, 2012).

Repeated mechanical loading or injury to the tendon causes an upregulation of MMP's which lead to degradation of the ECM which in time can lead to a recurring injury and chronic tendinopathy. It has been demonstrated by one investigation that after shockwave stimulation there was a down regulation of cytokines and MMPs (Han *et al,* 2009). It has also been hypothesized that the mechanical stimulus which is provided by this therapy might aid in tendon remodeling by stimulating the inflammatory and catabolic processes that are required for removing any damaged matric constituents (Waugh *et al*, 2015). A study which explored the effects of ESWT to normal tendinous tissue found that there was a disorganization of the matrix structure and changes were observed in the amount of degraded collagen present. An upregulation of COL1 expression was noted 6 weeks after the treatment which was hypothesized to be indicative for repair (Bosch *et al,* 2009). Possible side effects that have been mentioned in the literature are pain during the procedure, reddening of the skin and bruising (Haake *et al*, 2002).

This technique has also been used successfully in the fields of urology, to disintegrate uroliths, and in orthopaedics to induce tissue repair and regeneration (McClure and Dorfmüller, 2003). Until recently the use of ESWT in tendinopathies was controversial, however a recent review of ESWT found this procedure to be a successful alternative conservative treatment in tendinopathies (Mani-Babu *et al,* 2014).

Proteomic analysis by mass spectrometry has the ability to reveal both the different types and abundances of proteins in a sample. From the author's knowledge, to date, it has been mainly *ex vivo* investigations that have been undertaken to characterize the tendon proteome. A significant development would be the ability to detect *in vivo* changes in the proteomic composition, leading to clearer and more direct understanding of the efficiency of therapies. Biological samples can vary in complexity, and minimization of interfering substances through sample handling and cleaning can drastically improve the resolved peptide spectra. During this investigation, samples of microdialysis media from the peritendinous space of the Achilles tendon pre and post

shockwave therapy were used to establish whether the *in vivo* identification and quantification of proteins was possible.

Aims and objectives

To determine whether current methods for proteomic analysis using mass spectrometry are sensitive enough to ascertain the protein composition of dialysis fluid following ESWT in human tendinopathy. To determine any alterations in differentially expressed proteins in pre and post shockwave dialysis samples.

Materials and Methods

All reagents were obtained from Sigma-Aldrich unless otherwise stated.

Sample collection

Sample collection and handling was undertaken by Queen Mary University London. The full details of the experimental protocol, ethical permissions and methods used can be found in Waugh et al, 2015. Thirteen participants with healthy tendons with no previous history of tendinopathy (7 men, 6 women, aged 25.7 ± 7.0 years). Six patient participants (6 men, aged 39.0 ± 14.9 years) with established (*i.e.* symptomatic for > 6 months) mid-portion Achilles tendinopathy, diagnosed by a suitably qualified healthcare professional. Patients were excluded from participation if they had received steroid, platelet- rich plasma or high-volume image-guided injections in the previous 6 months, or had previous tendon surgery. Each patient was screened for a history of systemic inflammatory conditions, anti-inflammatory medication prescriptions.

A microdialysis catheter was placed under strict aseptic technique. within the peritendinous space (plantar to the Achilles tendon). For healthy participants, the skin either side of the free Achilles tendon was anaesthetised using subcutaneously injected Lidocaine (0.4-0.7 mL, 20 mg/mL), approximately 25 mm proximal to the calcaneal bone. For patient participants, the skin either side of the site with most pain and/or tendon thickening on palpation, depending on proximity to bony structures. The area was perfused with sterile lactated ringers solution. Due to a possible immediate biological response from the trauma of placing the catheter a 60-minute wait period was introduced prior to dialysate

collection. Collection then occurred every 30 minutes and samples pooled for two hours, ESWT was administered to each tendon mid-portion (mean treatment size of 1-2 cm^2) in a single session using a British kite-marked device (Swiss DolorClast Classic, Electro Medical Systems, Nyon Switzerland) and radial hand piece. Treatment was given perpendicular to the longitudinal tendon axis without anaesthesia and consisted of 2500 impulses administered at 8 Hz. The total energy delivered was 160 mJ/mm. After ESWT treatment, dialysate was collected every 60 min for the following 4 h (Waugh *et al,* 2015). This study included healthy patients and patients diagnosed with established (*i.e.* symptomatic for greater than 6 months) Achilles tendinopathy diagnosed by a suitably qualified healthcare professional. What sample group, our samples belonged to is unknown.

Methods

Six microdialysis samples were obtained from human patients, donor signalment and history was unknown. Two samples were pre shockwave therapy, two samples were post shockwave therapy and two samples were from unknown groups. The initial protein concentrations in all samples were measured and recorded with the NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Sample volume was limited, so in-solution tryptic digestion was undertaken following loading onto StrataClean™ resin (Agilent, USA). This enabled digestion of equal protein amounts of the samples and the removal of potential interfering substances (to the in-line chromatography). The in-solution trypsin digestion was performed with the prior addition of 1% Rapigest SF solution (Waters, UK) (incubated at 80˚C for 10 minutes) to improve the efficiency of digestion. Reduction of disulphide bonds and alkylation of free cysteines was achieved by incubation with 10 mM DTT (10min at 60˚C) and then 50mM iodoacetamide (30min at room temperature in the dark) respectively. Samples were digested overnight using 1µg of trypsin (Promega) at 37°C. To avoid auto digestion, the reaction was stopped with the addition of 1% trifloroacetic acid. Figure 1 demonstrates schematic of methods used. One-dimensional SDS-PAGE to compare quantitative/qualitative differences in the protein profiles could not be performed, as sample size was limited.

Mass Spectrometry methods

Trypsin-digested peptides were separated using an Ultimate 3000 RSLC (Thermo Scientific) nanoflow LC system. Using an ESI Easy Spray source at 50°C ~ 0.5 µg protein was loaded with a constant flow of 5 µl/min onto an Acclaim PepMap100 nanoViper C18 trap column (100 µm inner-diameter, 2cm; Thermo Scientific). After trap enrichment, peptides were eluted onto an Acclaim PepMap RSLC nanoViper, C18 column (75 µm, 50 cm; ThermoScientific) with a linear gradient of 2–40% solvent B (80% acetonitrile with 0.08% formic acid) over 65 min with a constant flow of 300 nl/min. The HPLC system was coupled to a linear ion trap Orbitrap hybrid mass spectrometer (LTQ-Orbitrap Velos, Thermo Scientific) via a nanoelectrospray ion source (Thermo Scientific). The spray voltage was set to 1.6 kV, and the temperature of the heated capillary was set to 250 °C. Full-scan MS survey spectra (m/z 335–1800) in profile mode were acquired in the Orbitrap with a resolution of 60,000 after accumulation of 1,000,000 ions. The fifteen most intense peptide ions from the preview scan in the Orbitrap were fragmented by collision-induced dissociation (normalised collision energy, 35%; activation Q, 0.250; and activation time, 10 ms) in the LTQ after the accumulation of 10,000 ions. Maximal filling times were 1,000 ms for the full scans and 150 ms for the MS/MS scans. Precursor ion charge state screening was enabled, and all unassigned charge states as well as singly charged species were rejected. The lock mass option was enabled for survey scans to improve mass accuracy (Olsen *et al,* 2004). Data were acquired using the Xcalibur software.

Data analysis

The Orbitrap Velos raw files were loaded into Proteome Discoverer (Version 1.4.0.288) and searched against Mascot Search Engine (Version 2.3.2).

Enzyme specificity was set to that of trypsin, allowing for cleavage N-terminal to proline residues and between aspartic acid and proline residues. Other parameters used were: (i) variable modifications, methionine oxidation, methionine dioxidation, protein N-acetylation, gln → pyro-glu; (ii) fixed modifications, cysteine carbamidomethylation; (iii) database: IPI-HUMAN_20130120); (iv) MS/MS tolerance: FTMS- 10ppm, ITMS- 0.6 Da; (v)

maximum peptide length, 6; (vi) maximum missed cleavages, 2; and (vii) false discovery rate, 1%.

Label free quantification

For label free quantification the raw files of the acquired data were analysed by the Progenesis QI™ software (Version 4, Nonlinear Dynamics). After selection of the reference sample, the retention times of the other samples were aligned. These were then exported and searched in UniHuman database with MASCOT server (Version 2.3.01). The following search parameters were used: peptide mass tolerance of 10ppm, fragment mass tolerance of 0.6Da, one missed cleavage, fixed modification, carbamidomethylation; variable modifications, methionine oxidation; a false discovery rate of 1% was used before the search results were re-imported back into Progenesis QI™. Statistical analysis was performed on all detected features using transformed normalised abundances for one-way analysis of variance (ANOVA). All peptides (with Mascot score >36 and $p < 0.05$) of an identified protein were included, and the protein p value (one-way analysis of variance) was then performed on the sum of the normalized abundances for all runs. Adjusted analysis of variance values of $p < 0.05$ and additionally regulation of >2-fold or <0.5-fold were regarded as significant.

Removal of lost labeled samples

To ensure the highest confidence in the data, the initial analysis was carried out excluding the two samples with lost labels. Unfortunately, within both the pre and post dialysis groups the sample variation was too great to allow effective processing with Progenesis software, when only the 4 labeled samples were anlysed. PEAKS software was also trialed but the combination of large sample variation and low sample number prevented any significant expression changes from being determined. This is a discovery project with the main aim of determining whether current methods for proteomic analysis using mass spectrometry are sensitive enough to ascertain the protein composition of dialysis fluid. Therefore analysis was continued using the data of all 6 samples and the unlabeled samples were speculatively grouped by how they clustered in principal components analysis.

Protein identification and ontology

The resulting significant protein lists were then entered into bioinformatics tools PANTHER (protein annotation through evolutionary relationship) and STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) to analyse the functional protein classifications and protein-protein interactions respectively (Mi et al, 2005) (Franceschini et al, 2013). PANTHER is a classification system that provides ontological information about the significant proteins identified. STRING provides information on known and predicted protein-protein interactions. Tendon tissue has a large proportion of extracellular matrix (ECM) compared to other tissues, so protein hits were entered into the Matrisome Database which provides live cross-referencing to gene and protein databases for every ECM and ECM-associated genes

Results

Initial protein concentrations were measured and recorded with the NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific). Samples 1 and 6 were pre dialysis; samples 2 and 5 were post dialysis. Sample size was too small to perform a protein assay. Results are displayed in Table 1.

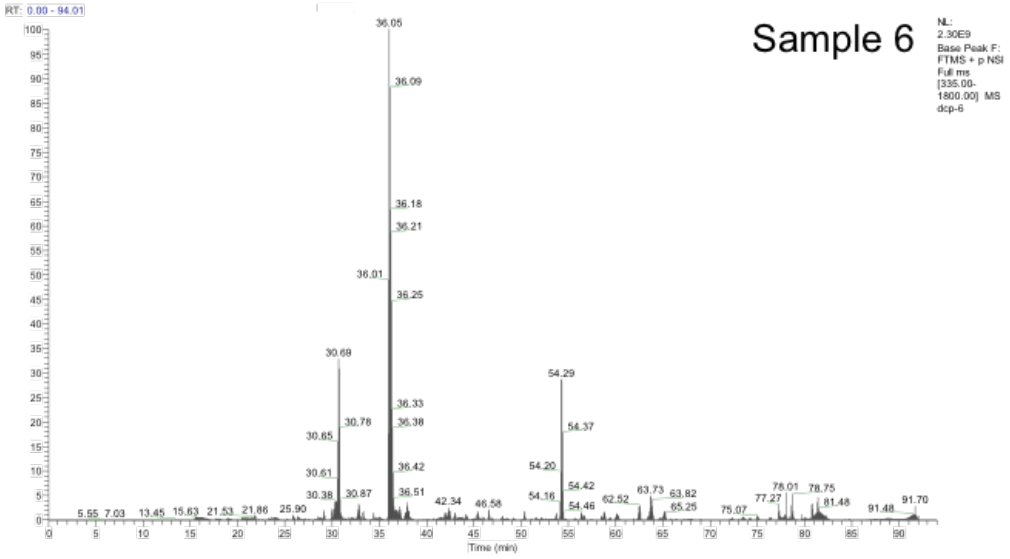
Sample number	Protein Concentration ug/ul
1	0.02
2	0.7
3	0.49
4	0.48
5	0.6
6	0.08

Table 1. Protein quantification in dialysis fluid assayed using NanoDrop Spectrophotometer.

Due to a problem with labeling, sample numbers 3 and 4 could not be identified, and were from unknown groups. From initial protein concentration measurements it could be predicted, but not conclusively, that both of these samples belonged in the post-ESWT dialysis group. After in solution trypsin digestion, equal amounts of proteins were loaded onto the LTQ-Orbitrap-Velos mass spectrometer. Total ion currents were comparable and a definite difference

in complexity exists between the two groups. Chromatograms from samples are displayed in Figure 2.

Pre shockwave therapy



Post shockwave therapy.

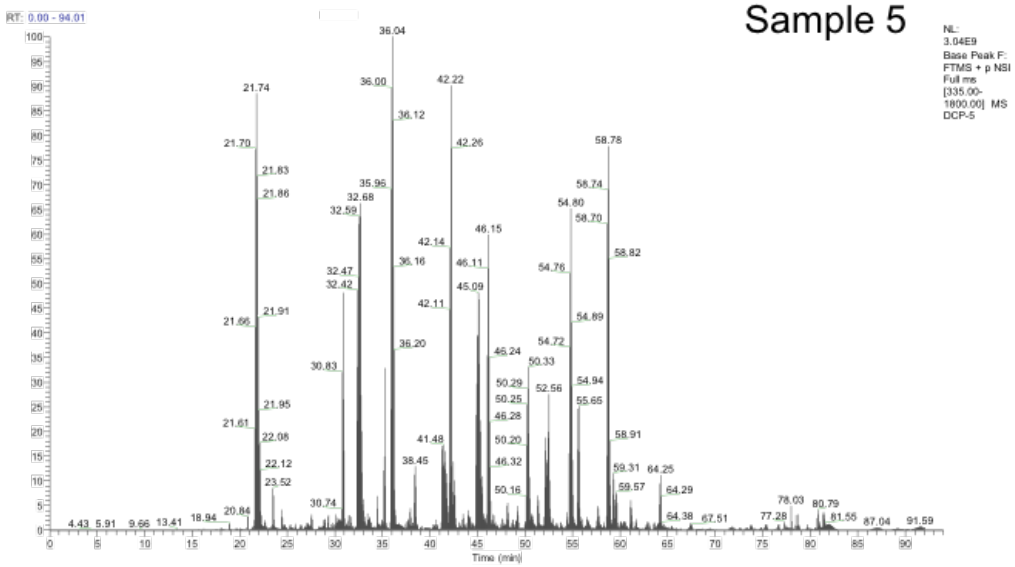


Figure 2 Comparison of chromatograms from liquid chromatography tandem mass spectrometry (LC-MS/MS) in the pre and post shockwave samples. An increase in chromatogram complexity is evident in post therapy samples.

Identification of proteins

In order to identify what proteins were present in our samples, raw data from MS/MS runs were imported into Progenesis QI. This software performed a multitude of tasks, and the main findings from this analysis are shown as follows. Principal Component Analysis (PCA) was employed to check for consistent behavior within the two experimental conditions: the pre and post ESWT samples. Figure 3 reveals there was a clear difference between the pre and post shockwave therapy groups, from this analysis, unknown samples 3 and 4 naturally fell into the pre and post shockwave therapy category respectively. This was in contrast with the protein concentration measurements in which both samples looked to be in the post ESWT group.

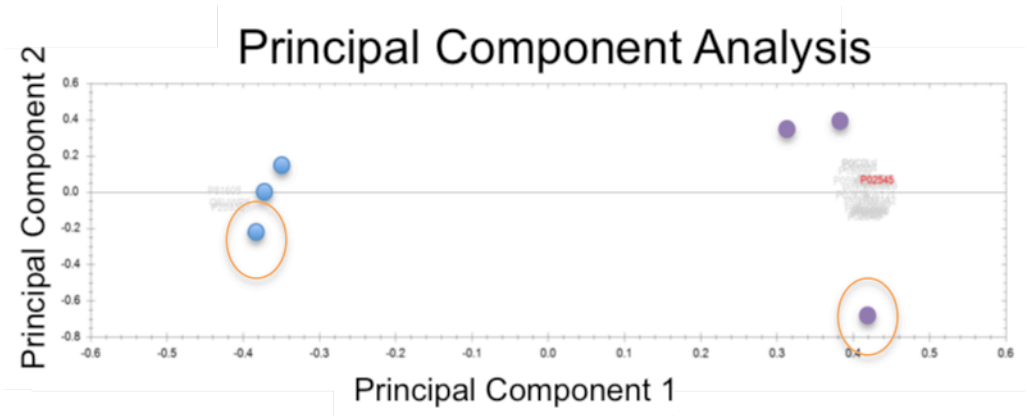


Figure 3. Principal component analysis (PCA) of filtered proteins P<0.01 and two fold change in expression were identified in MASCOT and reimported back into ProgenesisQI. PCA was used to search for correlations between the two groups. This approach clearly discriminated between pre and post shockwave therapy. The two unlabelled samples are circled.

Protein identification and ontology

Protein identifications included ECM proteins such as collagens and proteoglycans, and cellular proteins. After filtering, in total there were 93 quantifiable proteins identified, of these 80 proteins had a p<0.05 and 2 fold change difference as a consequence of treatment; 61 higher and 19 lower in post shock wave therapy. The 80 significant proteins were processed through the matrisome database, of these, 18 were recognised and are detailed in Table 1.

Functions of these proteins included roles in the clotting cascade and roles relating to cell adhesion.

Genome	Gene Symbol	Name of protein	Matrisome division	Category
Human	FGA	fibrinogen alpha chain	Core matrisome	ECM Glycoproteins
Human	FGB	fibrinogen beta chain	Core matrisome	ECM Glycoproteins
Human	FGG	fibrinogen gamma chain	Core matrisome	ECM Glycoproteins
Human	FN1	fibronectin 1	Core matrisome	ECM Glycoproteins
Human	COL1A1	collagen, type I, alpha 1	Core matrisome	Collagens
Human	CLEC3B	C-type lectin domain family 3, member B	Matrisome-associated	ECM-affiliated Proteins
Human	LGALS1	lectin, galactoside-binding, soluble, 1	Matrisome-associated	ECM-affiliated Proteins
Human	A2M	alpha-2-macroglobulin	Matrisome-associated	ECM Regulators
Human	CSTB	cystatin B (stefin B)	Matrisome-associated	ECM Regulators
Human	F12	coagulation factor XII (Hageman factor)	Matrisome-associated	ECM Regulators
Human	F2	coagulation factor II (thrombin)	Matrisome-associated	ECM Regulators
Human	ITIH4	inter-alpha (globulin) inhibitor H4	Matrisome-associated	ECM Regulators
Human	PLG	plasminogen	Matrisome-associated	ECM Regulators
Human	FLG	filaggrin	Matrisome-associated	Secreted Factors
Human	FLG2	filaggrin family member 2	Matrisome-associated	Secreted Factors
Human	HRNR	hornerin	Matrisome-associated	Secreted Factors
Human	PF4	platelet factor 4	Matrisome-associated	Secreted Factors
Human	S100A9	S100 calcium binding protein A9	Matrisome-associated	Secreted Factors

Table 1. The 80 proteins detected with p<0.05 and 2 fold change difference were processed through the matrisome database, 18 were recognised and are displayed. Functions of these proteins ranged from roles in the clotting cascade to roles in cell adhesion.

In the following analysis, to increase confidence in the identification of proteins, data was filtered to include proteins with at least two unique peptides and a FDR<0.01 in this data set. The following set of analysis uses this data set: -

Table 2 lists the top 10 increased proteins found by MASCOT to be differentially expressed in the post shockwave therapy samples. The top protein, peroxiredoxin-6, plays a role in protection against oxidative injury. The other proteins found had roles in the immune system and cell transport.

Figure 4 demonstrates protein classification according to their biological process using PANTHER. Results reveal higher proportion of cellular proteins and a response to external stimulus is evident. Figure 5 demonstrates STRING analysis on the same data set. This provides information on known and predicted protein-protein interactions. As this data is from post shockwave therapy it is

reassuring to find that 41% of these proteins are known to be involved in a response to an external stimulus (figure 5).

Accession	Peptide count	Anova (p)	Max fold change	Highest mean condition	Description	Protein type	Location
P30041	2	0.0029	Infinity	post	Peroxiredoxin-6	Redox regulator	Cellular
P02774	2	0.0088	51.26	post	Vitamin D-binding protein	Transport	Extracellular space
P01861	2	0.0014	33.8	post	Ig gamma-4 chain C region	Immune protein	Secreted
P15090	2	0.0092	22.77	post	Fatty acid-binding protein, adipocyte	Transport	Cellular
P02766	3	0.0046	16.18	post	Transthyretin	Transport	Cellular and secreted
P01023	4	0.0096	13.77	post	Alpha-2-macroglobulin	Proteinase inhibitor	Secreted
P01834	2	0.005	12.27	post	Ig kappa chain C region	Immune protein	Cellular
P01024	13	0.0028	9.12	post	Complement C3	Immune protein	Secreted
P02753	4	0.0014	8.64	post	Retinol-binding protein 4	Transport	Secreted
P0C0L4	6	0.0038	5.95	post	Complement C4-A	Immune protein	Secreted

Table1. Top 10 differentially expressed proteins following shock wave therapy (those with at least two unique peptides and a FDR<0.01 were included in this data set). Top protein Peroxiredoxin-6 plays a role in protection against oxidative injury.

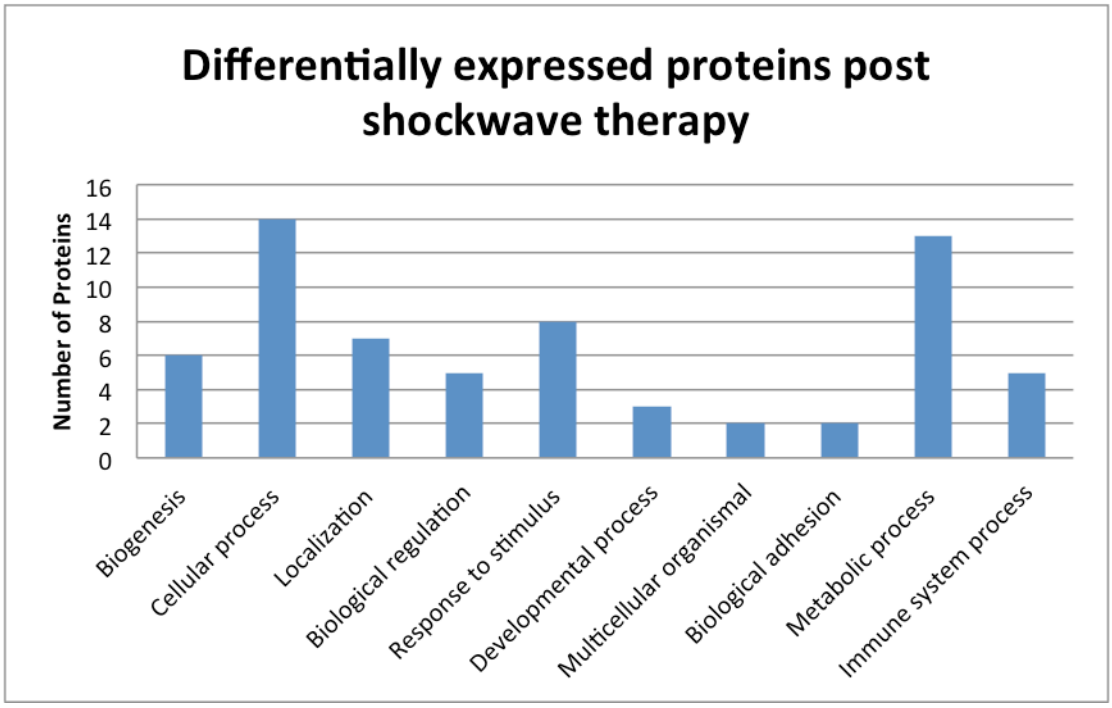


Figure 4. Using the gene ontology tool PANTHER, differentially expressed proteins from post ESWT samples were classified according to their biological process.

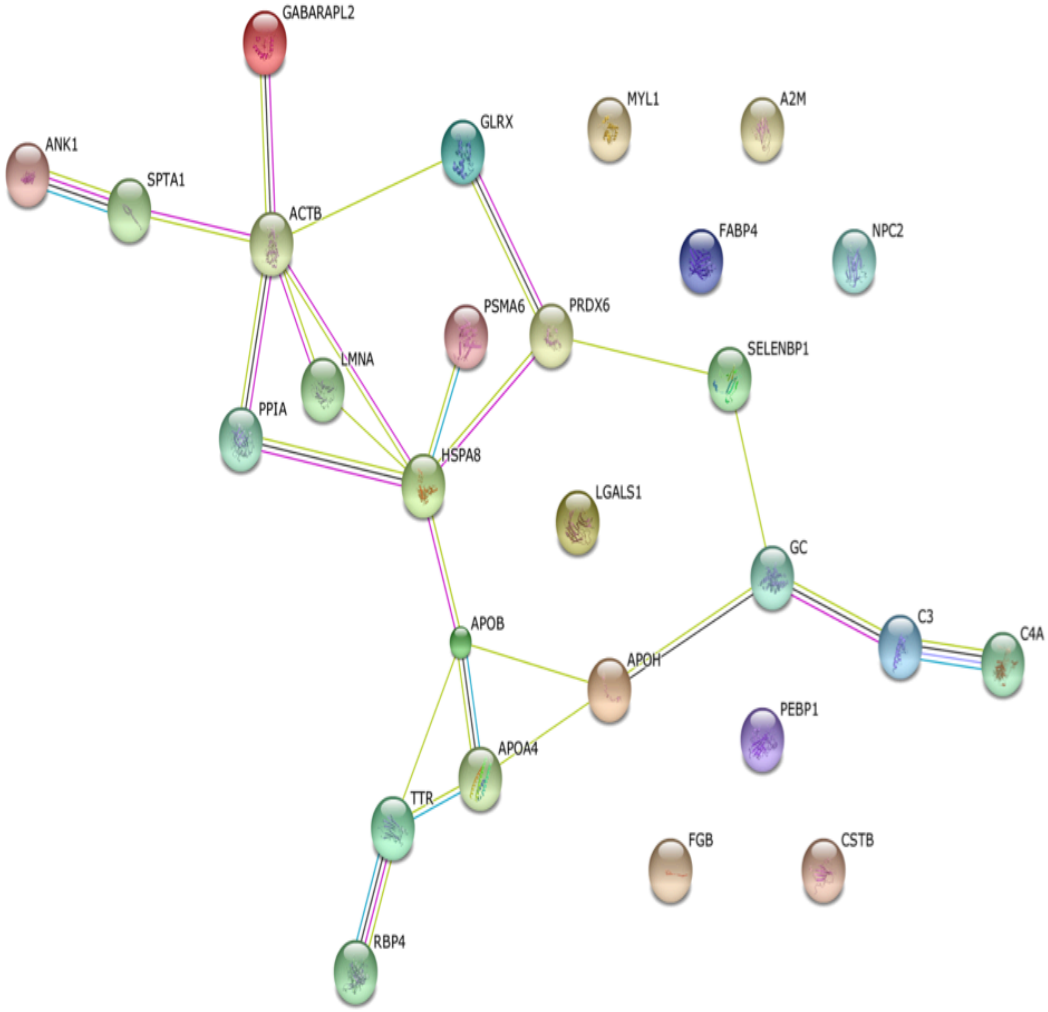


Figure 5. Protein-protein interaction network visualised by the bioinformatics tool STRING using differentially expressed proteins with at least two unique peptides and a FDR<0.01 in post EWST samples. Clustered proteins seen are normally involved in response to an external stimulus

There were 19 proteins that were downregulated in the post ESWT section, biological processes analysed via STRING analysis revealed that these proteins were involved in the establishment of skin barriers, skin development water homeostasis and the regulation of body fluid levels. Figure 6 demonstrates the Protein-protein interaction network generated.

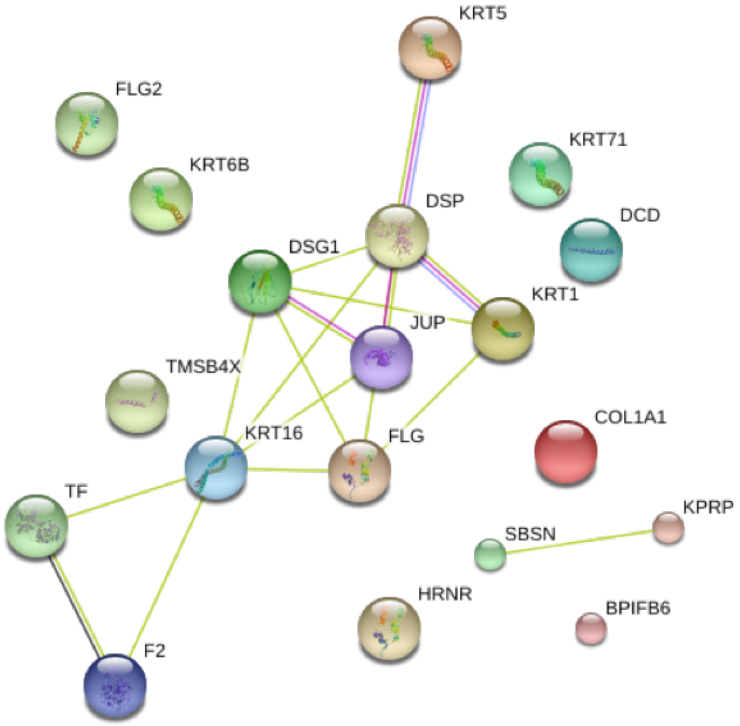


Figure 6. Protein-protein interaction network visualised by the bioinformatics tool STRING of down regulated proteins. Using differentially expressed proteins with at least two unique peptides and a FDR<0.01 in post EWST samples. Clustered proteins seen are normally involved in skin development

Discussion

The main aim of the current study was to establish if current methods for proteomic analysis, using mass spectrometry, are sensitive enough to ascertain the protein composition of dialysis fluid following ESWT in human tendinopathy. If this was achieved then the next step was to analyse any alterations that were observed between the two sample sets. Prior to discussing results acquired during this investigation, it is interesting to review previous findings within this area of research. Firstly this investigation is unique, in that, to the author's knowledge, no study using mass spectrometry has been performed on these types of samples before. So the following discussion targets related research.

As mentioned earlier histopathological changes of tendinopathy have been recorded to include collagen fibril disorganisation, increased size of the non-collagenous extracellular matrix, hypercellularity and neovascularization (Nourissat *et al,* 2015). However, the molecular driver responsible for these changes is unknown. Proteomic analysis of healthy and injured superficial digital flexor tendons of Thoroughbred horses revealed increased matrix degeneration and fragmentation in injured tendons compared to healthy samples. A large number of neopeptides were identified for proteoglycans and collagens in injured tendon (Peffers et al, 2014). Another study detected the pro-inflammatory cytokines IL-1α, IL1-β, TNF-α and IFN-γ in injured equine SDFT tissue by immunohistochemistry, the same cytokines were not detected in normal tendon samples (Hosaka et al, 2002).

The samples used during this study came from the investigation performed by Waugh et al, samples were analysed by cytometric bead array and zymography. It was determined that interleukins 6, 8 and the zymogen forms of matrix metalloproteinase (MMP) 2 and 9 were increased following ESWT. The authors provisionally concluded that this treatment aids tendon remodeling by promoting the inflammatory and catabolic processes that are responsible for removing damaged matrix constituents (Waugh et al, 2015). During this investigation we did not identify any of these proteins. This could have been due to inadequate protein loading, the sensitivity of the instrument or ion suppression due to other peptide components.

Data gained from this investigation is promising. It was established that the *in-vivo* identification and quantification of proteins was possible from samples taken from the peritendinous region following dialysis. This is a promising outcome, as the ability to perform in vivo analysis will allow assessment of the efficiency of treatments and give a real time analysis of how biological processes are occurring. However this experiment should be repeated with a larger sample size to confirm the conclusions drawn.

From using the data analysis available, we were able to detect changes between the pre and post ESWT samples. An upregulation of proteins was seen in post shockwave therapy samples, corresponding with previous investigations results (Waugh *et al,* 2015). Types of protein that were up regulated were involved in oxidative injury, transport and the immune response. Many of the proteins that were higher in the post ESWT samples are abundant proteins in serum. This could signify serum leakage following treatment.

Limitations in this investigation included small sample size and poor labeling of two of the samples. Therefore the data determining differences between treatments should not be relied on, as there is uncertainty in the exact provenenace of these two unlabeled samples. However the PCA plot grouped them into labeled and unlabeled groups. Also no information was available on sample phenotype such as age/sex/healthy/patient. Knowing information on sample phenotype would have provided additional information on this aspect of tendon biology.

This unreliability of the data should not detract from the central aim of this chapter, which was to determine whether proteomic methodologies could be used to profile microdialysis fluids. This could easily be improved upon with larger sample sizes to enable a more accurate analysis of protein changes. This small study has proved that continuing with proteomic analysis of dialysis fluid is worthwhile. A larger sample sizes would increase the accuracy of results and hopefully identify additional evidence on what biological processes are generated for healing. There was no identification of MMP proteins, which may just be due to the small sample numbers, due to the levels of abundance and sampling handling alternatively this could be a real finding. By repeating the investigation with increased sample numbers this could be examined further. Alongside this, results could be validated by using a different technique such as western blot analysis. This could used to confirm the proteomic data of interest generated by mass spectrometry. An alternative approach for a sensitive and specific quantification of targeted proteins would be to employ multiple reaction monitoring (MRM). This technique uses a directed MS/MS approach, which then

reinjects a selection of peptides to supplement the identifications acquired (Ademowo *et al,* 2014). This process enhances the lower detection limit for peptides and so would be a valuable tool with discovery of biomarkers in, for instance, tendinopathy or during the healing process.

The other factor to consider when analyzing these results is that the samples were collected from the peritendinous region via saline dialysis. This, of course, will give a different view of the biological process compared to taking a sample of the tendon itself and performing LC MS/MS. However more invasive sampling such as a biopsy, in a tissue, which is known to have poor healing abilities, is far from ideal and alongside this, using this method, would make the identification of less abundant proteins seen in this study difficult to detect due to the high level of extracellular matrix seen in tendon tissue.

This type of analysis could be used in the future to help in the understanding of tendinopathy, monitor the effectiveness of treatments, and identify the mechanisms that are responsible for the clinical success of ESWT.

Chapter 5, Tendon protein turnover

Introduction

In this thesis so far, we have aimed to develop our understanding of the biological function and processes occurring in tendon tissue by protein identification and quantification. However another important issue is the rate at which these proteins are synthesised and degraded. The study of protein turnover has long been realised as an extremely important feature, as a dysfunction in protein turnover can underlie certain pathological phenotypes. This dysfunction, can occur by various means, such as an accumulation of damaged proteins from reduced quality control, or from misfolding or a reduction in the amount of particular required proteins (Kim *et al*, 2012). To have a dynamic understanding of metabolism in normal tissue, allows us to understand changes that occur during different physiological and environmental conditions. With increased knowledge of what is occurring in pathologies, this should lead to the creation of new biotechnologies for diagnosis and treatment, as well as development of novel approaches to prevent diseases.

Turnover rates are specific to the individual protein and can vary widely from minutes to years (Price et al, 2010). Previous research of tendon tissue turnover has been described within the first chapter. Briefly, studies so far have indicated that tendon tissue undergoes very limited turnover in comparison to other tissues types (Heinemeier *et al*, 2013) and there are documented turnover differences in the collagenous and non-collagenous components of tendon tissue. For instance, the collagenous components of the equine superficial digital flexor tendon are relatively inert 197.53 (±18.23) years compared to the non collagenous components, which had a turnover rate of 2 years for (Thorpe *et al,* 2010).

However, data on this subject is limited and needs further work to validate and develop previous data. The investigation of 'proteome dynamics' has been undertaken by numerous methods, which are reviewed in the paper by Beynon, 2004. The main approach to measure the flux of amino acids through the protein

pool is by employing a metabolic tracer. Stable isotope labeling of mammals to examine the rate of protein synthesis and protein degradation is a well recognised technique.

The feeding of an essential stable isotope amino acid such as $^{13}C_6$ lysine to experimental animals, such as mice, allows the exploration of processes within metabolic pathways; this heavier form isotope allows the differentiation between preexisting and newly synthesized proteins by analysis with mass spectrometry. This method is not thought to cause harm to the treated animals, consequently data can be achieved in a safe and ethical manner (Doherty and Beynon, 2006). As mentioned previously, biological samples can vary in complexity, and minimization of this through sample handling can drastically improve the resolved peptide spectra. During this investigation filter assisted sample preparation (FASP) methods were employed to increase the reliability and accuracy of results by increasing the quality and consistency of the final samples.

Aim and objectives

To determine if dietary supplementation of the essential, stable, labeled amino acid lysine can be used to ascertain turnover rates in tendon proteins. To determine the potential to compare, collagenous and non-collagenous protein turnover.

Materials and methods

Ethics statement

The procedures involving animals were carried out in accordance with the UK Animals (Scientific Procedures) Act 1986, and authorised on personal and project licenses issued by the UK Home Office.

Materials

All reagents were obtained from Sigma-Aldrich unless otherwise stated.

Sample collection

Sample collection was undertaken by John Waters, Mammalian Behaviour & Evolution Group in the Institute of Integrative Biology, The Unversity of

Liverpool.

Male adult C57 BL/6 mice were housed individually in polypropylene cages (48 x 15 x 13 cm North Kent Plastic Cages Ltd UK) containing substrate (Corn Cob Absorb 10/14 substrate) and paper wool nest material. Food and water were provided ad libitum (LabDiet 5002 Certified Rodent diet purina Mills St Louis MO, USA) Animals were kept in an environment with no stressors, temperature was set to 21 +/- 1˚C and a reversed photo period (light 17h and dark 7h). All mice were acclimatised to a reconstituted semi-synthetic diet which was prepared, in house, based on their usual diet plus supplementation of $^{12}C_6$ Lysine at a quantity equal to the natural lysine content of the diet (1.18% w/w)) for 7 days prior to the start of the experiment. The 'light' diet was then replaced with an identical diet apart from the substitution of the crystalline amino acid by $^{13}C_6$ lysine at a relative isotope abundance (RIA) of 0.5. Animal's weight, food and water intake was recorded throughout. An outline of when mice were introduced into the study is displayed in table 1. At the end of the study all animals were humanely killed. The tail was transected at its base and the skin removed. Lengths of tendon were removed with careful attempts to avoid the collection of muscle. The tissue was washed in chilled saline and then stored in -80^{o}C until needed.

Time point (day)	MOUSE NUMBER INTRODUCED										
30	M1										
29	M1										
28	M1										
27	M1										
26	M1										
25	M1										
24	M1										
23	M1										
22	M1	M2									
21	M1	M2									
20	M1	M2									
19	M1	M2									
18	M1	M2									
17	M1	M2	M3								
16	M1	M2	M3								
15	M1	M2	M3								
14	M1	M2	M3								
13	M1	M2	M3								
12	M1	M2	M3	M4							
11	M1	M2	M3	M4							
10	M1	M2	M3	M4							
9	M1	M2	M3	M4	M5						
8	M1	M2	M3	M4	M5						
7	M1	M2	M3	M4	M5						
6	M1	M2	M3	M4	M5	M6					
5	M1	M2	M3	M4	M5	M6					
4	M1	M2	M3	M4	M5	M6	M7				
3	M1	M2	M3	M4	M5	M6	M7	M8			
2	M1	M2	M3	M4	M5	M6	M7	M8	M9		
1	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	
0	Culled	Culled	Culled	Culled	Culled	Culled	Culled	Culled	Culled	Culled	M11 (control) Culled

Table 1. Outline of study, each mouse was introduced to the diet at different time points between day 0, M11 and day 30, M1. During this investigation only two time points were processed which are highlighted above. The other samples have been kept for future analysis.

Methods

Protein extraction

The wet weight of each tail section was recorded. Protein extraction was achieved by the addition of 0.5 ml of guanidine extraction buffer (4 M guanidine hydrochloride (GdnHCl), 65 mM dithiothreitol (DTT), and 50 mM sodium acetate, pH 5.8) and the extraction performed at 4^0C for 48hours on an orbital shaker (ThermoScientific Max Q). The supernatant was removed after centrifugation at 15,000 x g @4^oC for 15 minutes. The final insoluble fraction was retained for future analysis of the non collagenous matrix components.

Protein assay

Ethanol precipitation was performed on aliquots of the supernatant prior to estimation of protein concentration. Samples were resolubilised with 7M Urea, 2M Thiourea, 30mM Tris, pH 8) then analysed by the Bradford assay read at

660nm using Coomassie Plus™ protein assay reagent (Thermo Scientific, Rockford, IL)

One-dimensional Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) and Coomassie blue staining

To remove interfering substances, GdnHCL soluble extracts were precipitated using Strataclean Resin Filtration (Stratagene R, Hycor Biomedical Ltd, Edinburgh, UK). All samples were subsequently separated in single dimensions based on their molecular weight by one dimensional SDS-PAGE to compare quantitative/qualitative differences in the protein profiles. 25ug of protein was loaded with reducing buffer (2-mercaptoethanol) onto 4-12% acrylamide Bis-Tris NuPAGE gels (Invitrogen). Before loading samples were incubated with reducing buffer for @ 90^0C for 5 minutes and then placed on ice. Proteins were visualised by Coomassie Blue staining (Thermo Scientific) according to the manufacturer's instruction.

Filter aided sample preparation (FASP)

50 µg GdnHCL soluble extract was loaded onto the filter unit, (VivaCon 500 10kDa MWCO HydroSart Filter units (VN01H01) from Sartorius), which had been pretreated with 100 µL of 1% (v/v) Formic acid (FA), and centrifuged at 14 000g for 15 min. The protein extracts were washed through with 200µl of urea buffer (8M Urea, 100mM Tris pH 8.5 (using HCL)) centrifuge for 15 minutes at 12,500rpm x 3. Reduction of disulphide bonds was achieved by the addition of 100µL 8mM Dithiothreitol (DDT) solution to the filter unit, mixed by gentle vortex and incubate in the dark at 56°C for 15 min. The unit was then centrifuged for 10 min at 12,500rpm. To remove the DTT, samples were washed through with 200µl of urea buffer centrifuge for 15 minutes at 12,500rpm x 3. Alkylation of free cysteine residues was undertaken by incubation with 100 of 50mM iodoacetamide (IAA), mixed by gentle vortexing and incubated in the dark at 20 min at room temperature. Again the sample was centrifugation at 12,500rpm for 10 min. To remove IAA the sample was washed with 100µl of urea buffer centrifuge for 15 minutes at 12,500rpm x 3. Taking care that the entire buffer had passed through by the end cycle. Buffer exchange occurred by adding 100µl

of 50mM Ammonium Bicarbonate (AMBIC) to the filter and then centrifuged for 10 min at 12,500rpm x 3, taking care that all buffer had passed through by the end cycle.

For the digestion step, a new collection low bind tube was obtained, which had been pre washed with 500ul 1% FA. Trypsin (Promega) 0.4µg was added in AMBIC solution, gently vortexed and digested overnight at 37^{o}C in a moist chamber. It was then centrifuged at 12,500rpm for 10 min, and the flow through collected. A further addition of 40µl of ammonium bicarbonate was added to the filter and spun at 12,500rpm for 10 min. The flow through was collected and pooled. To avoid auto digestion, the reaction was stopped with the addition of 10%Trifloroacetic acid solution. The tendon FASP sample was diluted 20-fold in 0.1%(v/v) TFA/3%(v/v) acetonitrile and run on a 50min LC-MS gradient. Figure 1 demonstrates a simplified schematic of this process.

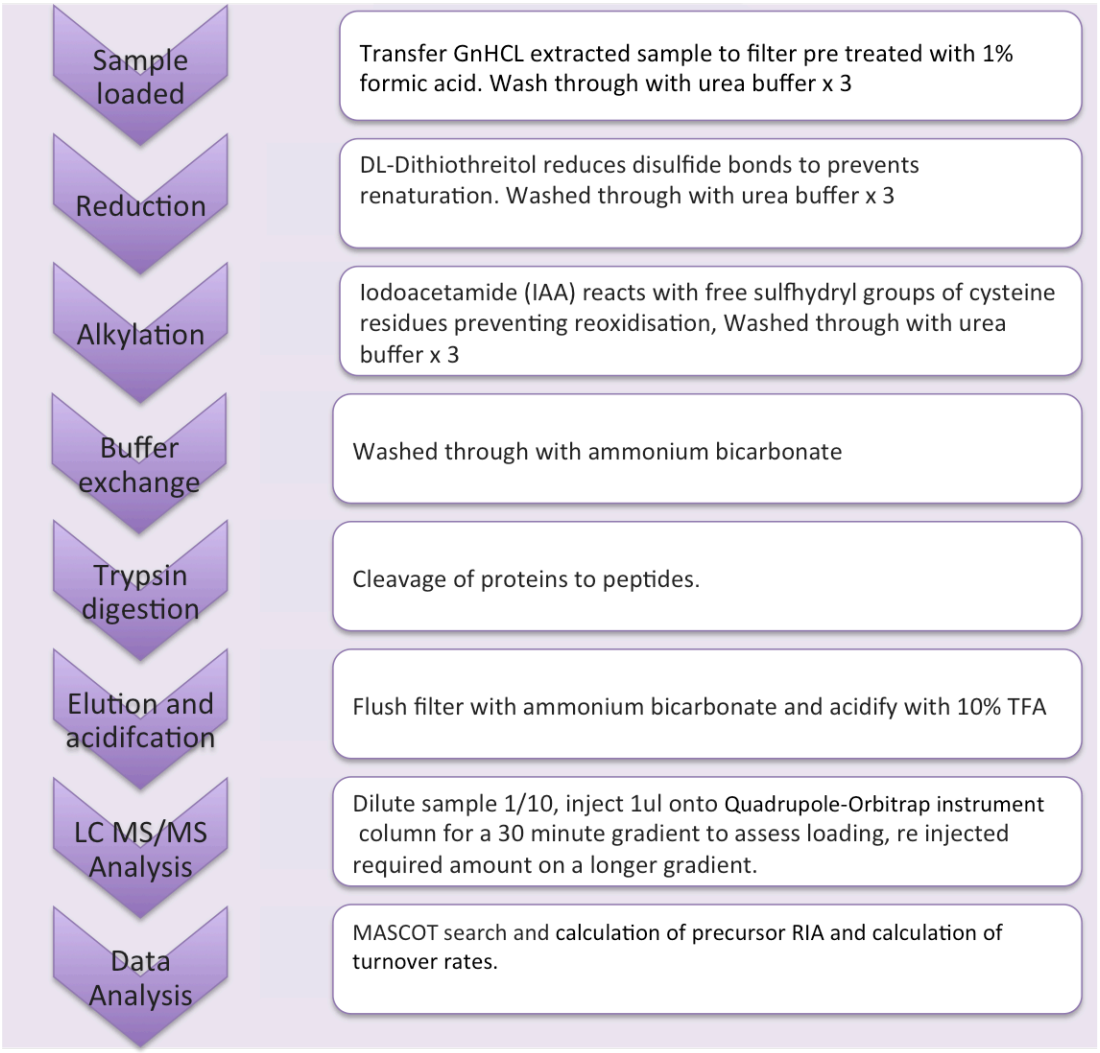


Figure 1. Schematic of FASP protocol.

LC Separation

All peptide separations were carried out using an Ultimate 3000 nano system (Dionex/Thermo Fisher Scientific). For each analysis the sample was loaded onto a trap column (Acclaim PepMap 100, 2cm x 75mm inner diameter, C_{18}, 3mm, 100Å) at 5mL/min with an aqueous solution containing 0.1%(v/v) TFA and 2%(v/v) acetonitrile. After 3min, the trap column was set in-line with an analytical column (Easy-Spray PepMap® RSLC 50cm x 75mm inner diameter, C_{18}, 2mm, 100Å) (Thermo Fisher). Peptide elution was performed by applying a mixture of solvents A and B. Solvent A was HPLC grade water with 0.1%(v/v) formic acid, and solvent B was HPLC grade acetonitrile 80%(v/v) with 0.1%(v/v) formic. Separations were performed by applying a linear gradient of 3.8% to 40% solvent B over 30 min at 300nL/min followed by a washing step (5min at

99% solvent B) and an equilibration step (10 min at 3.8% solvent B).

Analyses on a Quadrupole-Orbitrap instrument

The Q Exactive instrument was operated in data dependent positive (ESI+) mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (*m/z* 300-2000) were acquired in the Orbitrap with 70,000 resolution (*m/*z 200) after accumulation of ions to $1x10^6$ target value based on predictive automatic gain control (AGC) values from the previous full scan. Dynamic exclusion was set to 20s. The 10 most intense multiply charged ions ($z \geq 2$) were sequentially isolated and fragmented in the octopole collision cell by higher energy collisional dissociation (HCD) with a fixed injection time of 100ms and 35,000 resolution. Typical mass spectrometric conditions were as follows: spray voltage, 1.9kV, no sheath or auxillary gas flow; heated capillary temperature, 275°C; normalised HCD collision energy 30%. The MS/MS ion selection threshold was set to 2×10^4 counts. A 2 m/z isolation width was set.

Calculation of relative isotope abundance (RIA) and turnover rates

Thermo raw MS data files from both samples were analysed using the Progenesis QI™ software (Version 4, Nonlinear Dynamics, Newcastle, UK) These were then exported and searched in UniMouse database with MASCOT server (Version 2.3.01). The following search parameters were used: peptide mass tolerance of 10ppm, fragment mass tolerance of 0.6Da, one missed cleavage, fixed modification, carbamidomethylation; variable modifications, methionine oxidation and $^{13}C_6$ lysine; a false discovery rate of 1% was used before the search results were re-imported back into Progenesis QI™.

Statistical analysis was performed on all detected features using transformed normalized abundances for one-way analysis of variance (ANOVA). All peptides (with Mascot score >36 and $p < 0.05$) of an identified protein were included, and the protein p value (one-way analysis of variance) was then performed on the sum of the normalized abundances for all runs. Adjusted analysis of variance values of $p < 0.05$ and additionally regulation of >2-fold or <0.5-fold were regarded as significant.

Tissue at the start of the experiment will have a RIA of zero, the label is then introduced into the diet and incorporation into the protein pool is monitored over time. The flux of isotopic labels through the protein pool varies between protein types and should be determined by calculations based on Heavy + Light isotopic ratios

There were two time points available for analysis (time 0 and time 30 days). Extracted ion chromatograms were performed for heavy labeled peptides using Xcalibur 2.0 software (Thermo -Electron, Hemel Hempstead, UK). Using mascot, the total peptide intensity was used to determine the relative isotopic abundance. To allow the determination of RIA the following calculation was used

$$\mathbf{RIA = H/[H+L]}$$

(H= heavy isotope relative abundance L=unlabeled isotope abundance)

Figure 2 demonstrates the possible outcome that could occur from this experiment.

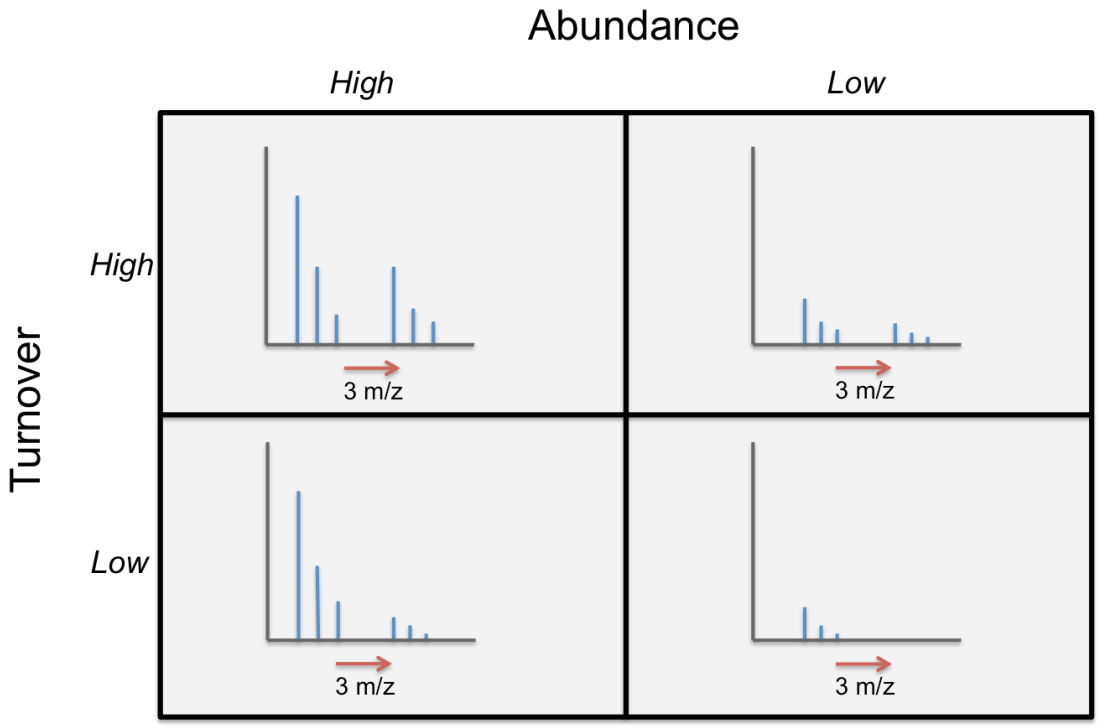


Figure 2 demonstrates a schematic diagram of potential outcomes, i.e. a high abundance and high turnover protein will have a spectrum revealing high intensity peaks which are separated by 3m/z for a double charged ion. As a 50:50 heavy:light diet was used to increase palatibility, this meant that that the highest possible Relative Isotopic Abundance is 0.5, heavy isotope peaks will always be no more than 50% of the light form of the peptide.

Results

Using the GnHCL protein extracts samples were run on 1D SDS PAGE. Results can be found in figure 3 and demonstrate that protein profiles are very similar between the different time points and the control. The majority of bands are of a higher mass range, which could be due to the sample being heavily composed of collagen proteins. Subsequent to LC MS/MS analysis raw data was filtered using the MASCOT search engine for the two time points. A list of top ten proteins based on mascot score are listed in tables 2 and 3.

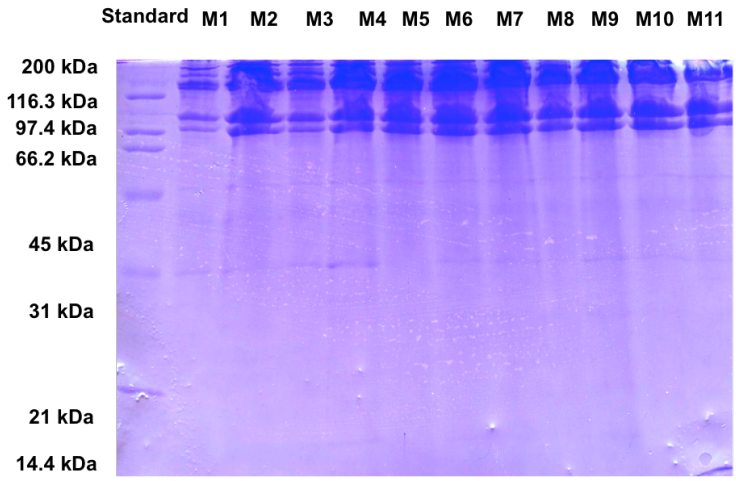


Figure 3. Coomassie stained 1D-SDS-PAGE of all guanidine soluble tendon protein samples from each time point (M1-M10) and control (M11). The majority of proteins are found in the higher molecular weight range suggesting that the sample primarily consists of collagen proteins. The gross protein profiles for the different time points and the control are similar.

Accession	Score	emPAI	Description	13C labelled peptides identified in MASCOT search
E9PWQ3	2028	0.93	Col6a3_MOUSE Protein Col6a3	0
Q60847	1292	0.72	COCA1_MOUSE Collagen alpha-1(XII) chain	0
Q01149	1101	0.7	CO1A2_MOUSE Collagen alpha-2(I) chain	0
P11087	1076	0.6	CO1A1_MOUSE Collagen alpha-1(I) chain	0
Q62507	603	1.92	COCH_MOUSE Cochlin	0
Q02788	568	0.86	CO6A2_MOUSE Collagen alpha-2(VI) chain	0
Q04857	568	0.69	CO6A1_MOUSE Collagen alpha-1(VI) chain	0
A2AQ53	550	0.35	FBN1_MOUSE Fibrillin-1	0
P20152	506	2.33	VIME_MOUSE Vimentin	0
Q9JK53	450	3.03	Prelp_MOUSE Prolargin	0

Table 2. Day 0, mouse tail tendon sample. Top 10 proteins identified by Mascot according to Mascot score. As expected, no 13C labeled peptides were identified.

Accession	Mascot Score	emPAI	Description	13C labelled peptides identified in MASCOT search
Q60847	2489	1.27	COCA1_MOUSE Collagen alpha-1(XII) chain	1
Q01149	1574	0.86	CO1A2_MOUSE Collagen alpha-2(I) chain	0
P11087	1563	0.78	CO1A1_MOUSE Collagen alpha-1(I) chain	0
Q61554	1469	0.62	FBN1_MOUSE Fibrillin-1	0
Q04857	1156	1.23	CO6A1_MOUSE Collagen alpha-1(VI) chain	0
Q62507	1064	4.47	COCH_MOUSE Cochlin	5
P28654	821	8.67	PGS2_MOUSE Decorin	10
P20152	784	4.47	VIME_MOUSE Vimentin	4
Q02788	740	1.2	CO6A2_MOUSE Collagen alpha-2(VI) chain	0
P50608	622	2.11	FMOD_MOUSE Fibromodulin	2

Table 3. Day 30, mouse tail tendon sample. Top 10 proteins identified by Mascot according to Mascot score. As expected, 13C labeled peptides were identified in mainly the non collagenous proteins. However one labeled peptide was detected in collagen alpha-1 (XII) Chain.

Analysis of labeled peptides was then undertaken. Looking at the top 10 significant proteins detected on day 0 (table 2) and day 30 (table 3), 13C labeled peptides were investigated by calculating the RIA and subsequently percentage turnover rate by comparing light to heavy ratios. Of the top 10 proteins, 4 proteins, Cochlin, Decorin, Vimectin and Fibromodulin revealed evidence of labeling. However collagen alpha-1 (XII) Chain also showed one peptide to be labeled, this is likely to be a false positive, significant labeling is deemed only to be present when all of the lysine cleaved peptides are labeled, which wasn't the case here. Figures 4 demonstrates one of the labeled peptides seen for Cochlin on day 30, figure 5 demonstrates the same peptide on day 0, proving no labeling has occurred. Figures 6-8 exhibit spectra for the remaining labeled proteins, Decorin, Vimectin and Fibromodulin. Table 4 reveals the mean RIA calculated from the labeled peptides, Cochlin had the highest turnover rate at 38% and vimentin the lowest at 24%.

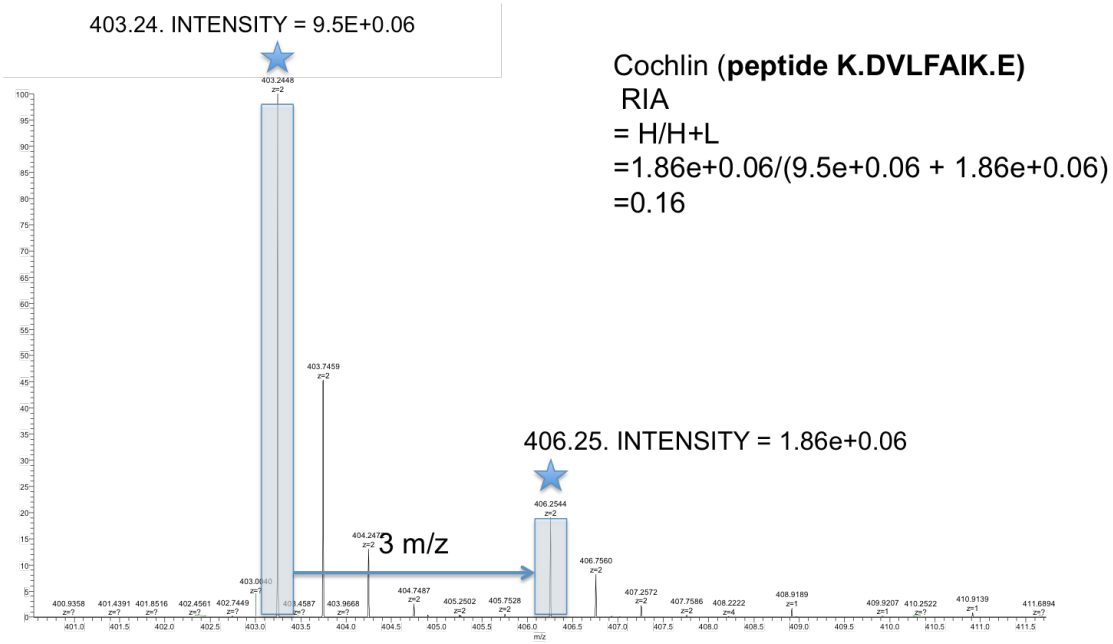


Figure 4. Spectra demonstrating $^{13}C_6$-lysine labeling of Cochlin peptide K.DVLFAIK.E at day 30, with doubly charged ions. Unlabeled peak is seen at m/z 403.24 and labeled peak is seen at m/z 406.25. RIA = 0.16 equating to a 32% turnover rate, for this peptide, taking RIA 0.5 to be 100% turnover.

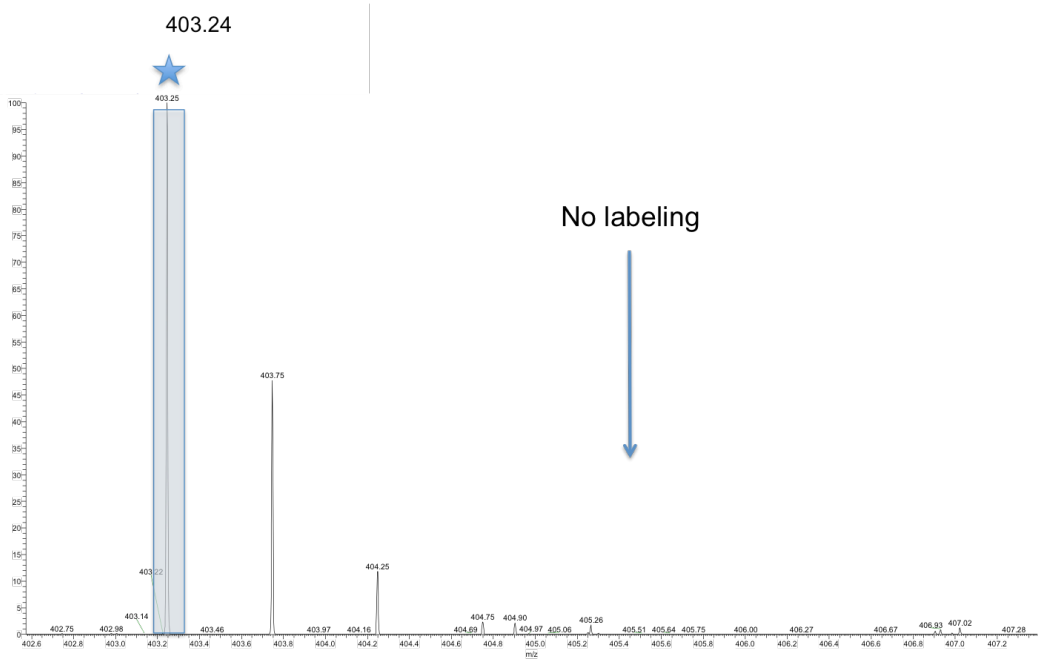


Figure 5. Spectra demonstrating Cochlin peptide K.DVLFAIK.E at day 0, with doubly charged ions. Unlabeled peak is seen at mass 472.76 and labeled peak is seen at mass 475.7.

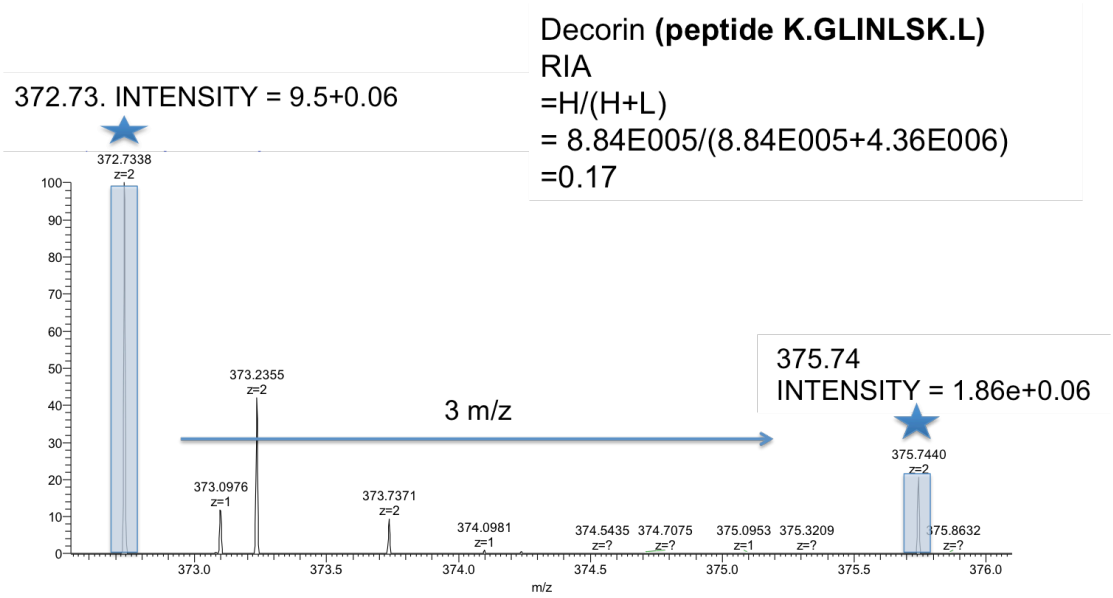


Figure 6. Spectra demonstrating $^{13}C_6$-lysine labeling of Decorin peptide K.GLINLSK.L at day 30, with doubly charged ions. Unlabeled peak is seen at m/z 372.73 and labeled peak is seen at m/z 375.74. RIA = 0.17 equating to a 34% turnover rate, for this peptide, taking RIA 0.5 to be 100% turnover.

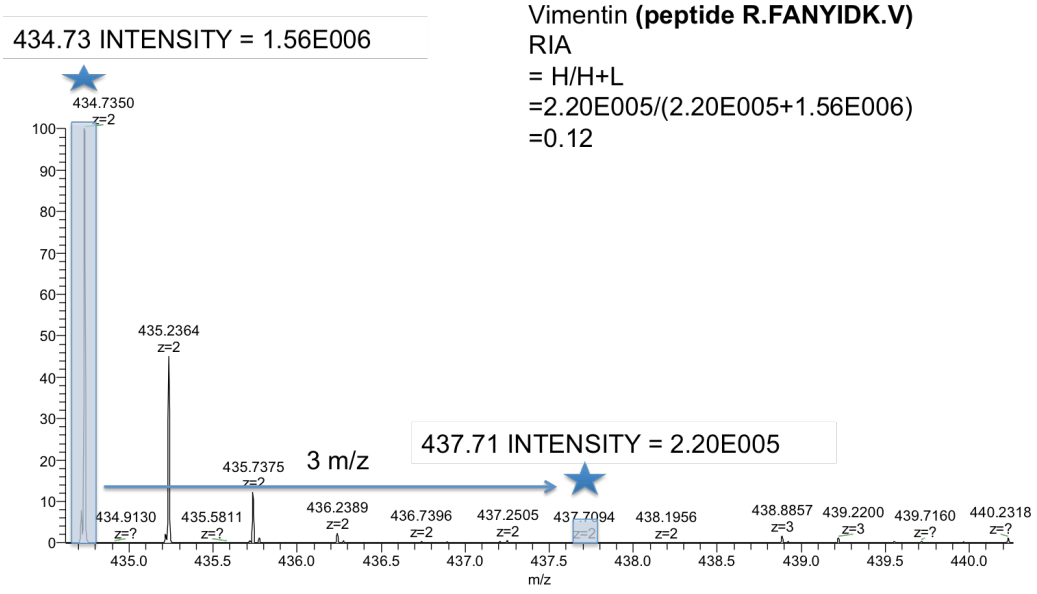


Figure 7. Spectra demonstrating $^{13}C_6$-lysine labeling of Vimentin peptide R.FANYIDK.V at day 30, with doubly charged ions. Unlabeled peak is seen at m/z 434.73 and labeled peak is seen at m/z 437.71. RIA = 0.12 equating to a 24% turnover rate, for this peptide, taking RIA 0.5 to be 100% turnover.

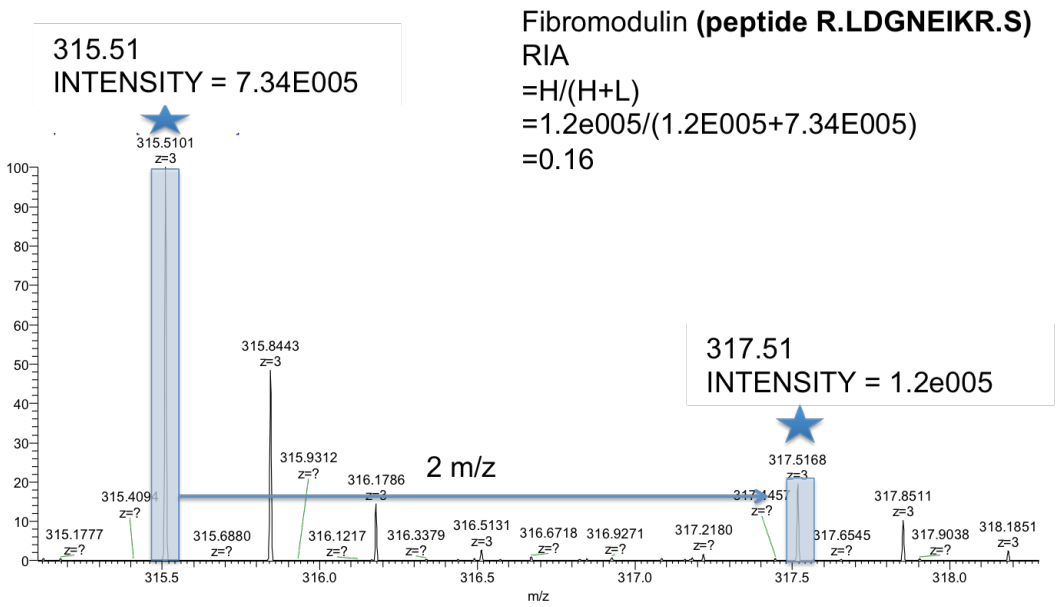


Figure 8. Spectra demonstrating $^{13}C_6$-lysine labeling of Fibromodulin peptide R.LDGNEIKR.S at day 30, with triply charged ions. Unlabeled peak is seen at m/z 315.51 and labeled peak is seen at m/z 317.51. RIA = 0.16 equating to a 32% turnover rate, for this peptide, taking RIA 0.5 to be 100% turnover.

Protein Description	Mean Relative Isotopic Abundance	Mean turnover rate % (RIA 0.5 to be 100% turnover)
COCH_MOUSE Cochlin	0.19	38
PGS2_MOUSE Decorin	0.17	34
VIME_MOUSE Vimentin	0.14	24
FMOD_MOUSE Fibromodulin	0.16	32

Table 4, mean RIA for each of the proteins displaying labeling and the resultant mean turnover rate. Cochlin had the highest turnover rate and vimentin the lowest.

Discussion

During this investigation, the dietary administration of the stable isotope lysine was used to assess protein turnover rates in mouse tendon tissue.

The main achievements of this small study were

- The identification of well known tendon tissue compositional components such as collagens and proteoglycans
- Labeling of the non collagenous component in 30 days. Tendon tissue is renowned for its poor healing ability and a slow turnover rate, in 30 days is was possible to label the non collagenous component, this study prove that further work would be worthwhile in this area.
- Confirmation that the collagenous component is relatively inert.
- Identification of higher turnover protein components of tendon tissue, which is consistent with previous studies.

The basis of this investigation was to measure the protein turnover in tendon tissue. Once accurate baseline turnover rates are achieved for the compositional components of healthy young tissue these could be compared to rates in injured, aged, exercised or from a chronic disease such as in diabetes. Further to this turnover rates could be calculated after treatments e.g. anti-inflammatory mediators or physiotherapy. Turnover rates could then be used to monitor tendon response to certain treatments. The end objective would be to use this information to then develop new approaches for the treatment of tendon diseases and injuries in patients.

A 50:50 heavy: light diet was chosen, as a completely heavy diet would be unpalatable, hence having a 50:50 diet increases palatability. Animals were firstly acclimated to a diet containing the same levels of non-labeled amino acid, and then transferred to the labeled diet at different time points. The essential amino acid lysine was used to measure protein turnover as the RIA is uncomplicated by de novo synthesis. Labeling with $^{13}C_6$ also removes complications seen with ^{15}N labeling due to the latters increased incidence of generating partially unlabeled peptides (Beynon and Pratt, 2005).

During this study, time was restricted, hence only two time points were included, day 30 and day 0. As a 50:50 heavy:light diet was used, this meant that that the highest possible RIA would be 0.5.

It was established that the non collagenous matrix, which consists of proteoglycans, had a higher rate of protein turnover compared to components of the collagenous matrix. Results are consistent with previous studies conducted in this area. (Heinemeier *et al*, 2013; Thorpe *et al,* 2010).

The proteins identified during this investigation largely correspond with previous work done in this area. As expected there were a high percentage of collagen proteins in both day 0 and day 30 samples. Non collagenous components in day 0 were cochlin, fibrillin, prolargin and vimentin. Day 30 non collagenous proteins identified were cochlin, fibrillin, decorin, vimentin and fibromodulin. For both samples some of these components were just as expected. Peffers et al, 2014 identified both decorin and fibromodulin, which are thought to be involved in fibrillogenesis. Studies that used decorin knockout mice, revealed a difference in the dynamic and failure loads of the Achilles tendon. This study also discovered a decrease in the organisation of the collagen fibres in these mice (Gordon *et al*, 2015). Another study demonstrated that an absence of fibromodulin also decreases collagen fibril organization and tendon strength (Kalamasjski *et al,* 2014). Fibrillin, a glycoprotein found in connective tissue has been isolated in tendon proteomic studies, one study identified fibrillin-2 only in the fascicular matrix (Thorpe *et al,* 2016). This protein is fundamental in providing a scaffold matrix for the deposition of elastin (Ritty *et al*, 2003). Prolargin is a class II SLRPs and has been detected in tendon tissue before, this protein, binds to type I collagen and is thought to anchor basement membranes to the connective tissue (Peffers et al, 2014). Vimentins are class-III intermediate filaments found in various non-epithelial cells, especially mesenchymal cells. It is also involved in the stabilisation of type I collagens mRNAs for CO1A1 and COL1A2 (www.uniprot.org) Vimentin has also been localized in the rat suprapatella, which is a fibrocartilage that resists the compression of the quadriceps tendon against the femur in the flexed knee

(Benjamin et al, 1991). Cochlin was the most unusual of proteins to be discovered here, to the authors knowledge this is the first work to identify the presence of Cochlin within tendon tissue. This protein plays a role in the control of cell shape and motility in the trabecular meshwork (www.uniprot.org). Its presence in these samples could be due to contamination or could be a genuine finding as the role cochlin has, is biologically plausible in tendon tissue..

One peptide from collagen alpha 1 (XII) chain revealed a labeled peptide, this, almost certainly is a false positive as all of the other lysine cleaved peptides were unlabeled. With this result removed, it was shown that there was no evidence of turnover identified in any of the collagenous proteins. This was expected, however previous studies have hypothesised that the IFM may contain a pool of more labile collagen (Thorpe *et al,* 2016), it would be worthwhile extending the sample numbers and the length of this study to see whether this hypothesis can be backed up. It would have also been useful to study a larger number of proteins in this study, much more information would have been gained from looking at the top 20 proteins however as this was the last study undertaken during the year, time was very restricted and so it was decided to just use the top 10 proteins which was unfortunate but necessary in order to finish the work in time.

In regards to the methods used in this study, this analysis of tendon tissue was actually performed alongside another study, which looked at turnover rates in different tissue types. Hence the methodology used here was slightly different to previous chapters, by doing this it will be possible to compare the dataset with other tissue types in the future. Guanidine was chosen as the extraction agent as this method has been used in previous published studies (Peffers *et al*, 2014; Wilson *et al*, 2010) therefore using this method was decided upon rather than expanding upon the work undertaken in chapter 2 and adding in the surfactant Rapigest.

The use of animal models for investigations into human disease and treatments is not without its drawbacks. Mouse tail tendon is not an ideal comparison model

for humans or horses, as their basic anatomical structure is different. This tendon does not have to withstand high strains and the interfascicular region is different (Kohrs *et al*, 2011). Labeling studies are expensive even in these tiny animals so to use a more relevant larger animal model, such as the horse, would be unrealistic. The rat model may be better as their basic anatomy is more comparable to humans and they can be used as a model of tendinosis which has been confirmed by previous studies, in contrast to the mouse model, meaning a superior overall comparison can be made (Warden, 2007)

As no study has been undertaken in this area before, the preparation of samples collected from other time points is warranted. Including more time points would allow a greater accuracy in the turnover calculations. Due to the known slow turnover rate of the collagenous part of tendon tissue it would be interesting to perform the study over a longer time period such as over 100 days. This will improve accuracy of results however collagens are predicted to have such a low turnover rate that the incorporation of the isotopic label would be predicted to occur over many years, so results specific to the collagenous matrix may not be too dissimilar to this study.

In conclusion, it has been established that the non collagenous matrix which consists of proteoglycans had a higher rate of protein turnover compared to components of the collagenous matrix. If repeated, as suggested, it will be possible to use results gained as a baseline of normal tendon turnover. This, in the future, could be compared to injured tissue to assess comparisons. Another possibility is to use this research to assess the tendon tissues response to certain treatments with the end objective in discovering new biotechnologies for the treatment of diseases/injuries in humans and animals.

Chapter 6, General Discussion

Tendon injury is an ever-increasing problem to the individual and the health service that provides for them. It has been documented that injuries to the Achilles tendon are becoming common in athletes and sedentary individuals alike (Uquillas et al, 2015). The same is true in our equine patients; tendon injury is a significant cause of morbidity and performance loss.

Any type of tendon injury can be a challenge to treat and so further work is warranted into the mechanisms that are occurring in tendon injury, degeneration and healing. Although a seemingly far off vision, advances that enable the recovery of tissue function would revolutionise current therapeutics models. Further methodologies to identify tendon injury at an early stage could lead to advances in injury diagnosis and prevention.

Over the past year we have aimed to gain further understanding into the field of tendon biology, with an end sight of contributing to the developments made by other research groups in recent years. Our objective was to increase understanding of tendon biology by developing proteomic methodologies, which will enable future experiments to determine the composition of the tendon tissue and how this varies between tendon types and within the tendon tissue itself. As the reader knows, this thesis was split into four preliminary investigations and here we discuss limitations and problems encountered for each one.

The protein extraction investigation was a comparative study investigating two types of protein denaturants, the chaotropic agents, guanidine and urea and the detergent Rapigest. Experiments were performed in two tendon types, positional vs. energy storing and in two species, human vs. equine. The aim was to maximize proteins removed from the tendon fascicles to allow the accurate definition of the tendon proteome via mass spectrometry.

The digestion process varies from protein to protein due to differences in primary and secondary structures. Difficulty in digestion can mean that there is a

reduction in the access to cleavage sites for the digestion enzyme. So agents that disrupt the protein structure can increase the efficiency of digestion. Currently there are no published studies that have compared efficiencies of protein extraction methods in tendon tissue. However we can study other musculoskeletal tissues for general comparisons. Wilson et al performed a protein extraction comparison study in articular cartilage that, like tendon, has a dominance of poorly soluble matrix components. This study investigated proteomic differences within juvenile cartilage and neocartilage, the most applicable aspect of this study to our investigation is that they compared the efficiencies of the chaotropic agent guanidine and saline in these tissues. Results revealed that both conditions had specific proteins that were significantly enriched. Mature cartilage had lower levels of identifiable likely due to the tightly interacting ECM network. Guanidine samples generated a higher yield of proteins in mature cartilage. From this we can devise that guanidine has the denaturing ability in this poorly soluble tissue therefore should be a suitable agent for protein extraction in tendon tissue. Subsequent to this study, a proteomic study in tendon tissue used the chaotropic agent guanidine with success (Peffers et al, 2014). For this investigation guanidine was a definite contender to trial. Rapigest is an agent, which has had success within proteomics due to its compatibility mass spectrometry so would offer a simple method, if effective for tendon tissue.

Although there were no statistically significant results, noticeable trends were identified which enabled broad conclusions to be drawn. For instance protein extraction using the chaotropic agent urea, yielded the lowest number of significant protein hits in both human and equine samples and guanidine HCL yielded highest. Using the PANTHER database, guanidine also yielded the highest amount of intracellular proteins, possibly as this agent is able to disrupt cellular membrane proteins in a more effective manor, which is consistent with other studies findings where guanidine was thought to disrupt the outer membrane of cells causing permeabilization (Naglak and Wang, 1990). Rapigest SF samples tended to have a roughly equal amount of intracellular and

extracellular matrix proteins, revealing this has the ability to denature some of the most resistant proteins.

This study offered preliminary data that offers foundations for new investigations, such as, repeating with a larger sample size and combining denaturants to see whether this permits greater peptide recovery. For instance using a protocol which firstly extracts proteins with guanidine and then using rapigest on the resulting insoluble pellet. This may increase the protein yield further as well as less waste from both a sample collection and a financial point of view. Alternatively different experiements could be run depending on the proteins of interest. It is known from this investigation and from Wilson et al that different conditions can yeild different unique proteins. Two experiments could be devised, one with a focus of extracting cellular proteins and another with a focus of extracting extracellular matrix proteins and the results read alongside each other. This would allow a fuller picture to be visualised and in time advance the proteomic extracellular matrix world by paving the way for differential analysis of normal and pathological tendon tissue

From previous studies, it can be surmised that the increase in tendon injury in older individuals is associated with ageing of the interfascicular matrix (Thorpe *et al,* 2013). Further study into the IFM structure and composition is warranted, as this region could potentially be a diagnostic or therapeutic target. A dual imaging approach that combined matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) and secondary ion mass spectrometry (SIMS) of young, old, energy storing and positional tendon tissue was evaluated. The aim was to obtain molecular maps of peptides, lipid and small molecules and to compare the spatial variation between groups. This was a challenging task as there was no previous published data on equine tendon imaging to offer known workable methods. As before, looking to previous investigations in other musculoskeletal tissues is helpful. Starting in a laboratory, which had limited experience in the direct imaging of tissues using MALDI, meant progress was slow and first attempts failed as all spectra collected were that of the matrix. Fortunately a three-week visit to the University of Maastricht under the

supervision of Professor Ron Heeren provided expert collaboration and assistance on the subject. This laboratory previously had performed MSI on cartilage and synovial tissue (Cillero-Paster *et al*, 2014; Peffers *et al*, 2014). So this provided a starting point for protocols for sample preparation, but different conditions needed to be trailed for tendon tissue, each step of the protocol used had to be trialled and progress was made. Results revealed that Trypsin reconstituted in 25mM Ammonium bicarbonate was seen to generate more peaks compared to dionised H_2O. Previous studies have highlighted the importance of the buffer solution used in the dissolution of trysin. The use of water is thought not to be ideal for efficient tryptic digestion (Cillero-Paster and Heeren, 2014). The reason for this could be because the ideal pH range for tryptic digestion is 7.5-8.5 (Worthington, 2011), therefore using a buffer solution which provides this is optimal. Scober et al showed that using ammounium bicarbonate as the trypsin buffer had a greater compatibility compared to Tris-Hydrochloride, as this buffer had a greater compatibility with the subsequent matrix application. This buffer would also be better for the downstream MS analysis (Proc et al, 2010).

During this study the optimal trypsin digestion time was found to be 16 hours. This overnight digestion time, previously has been used largely due to the convenience. A study conducted in plasma proteins revealed similar results to this study where 9 hour or 16 hour digestion period were superior to 4 hours (Proc et al, 2010).

The biggest limitation during this study was the time available, overall I detected no differences between the four sample groups being analysed using both MALDI and SIMS. However, as only one sample of each group was evaluated, results should be reviewed with this in mind, future studies should be undertaken with a larger sample size. It is possible that this representation is a true finding but based on previous immunohistochemistry results it is worth persevering with sample preparation optimisation. However, progress was achieved with in situ trypsin digestion.

For the final two studies of this thesis we were collecting data rather than developing methods of analysis. Both of them produced promising data. Chapter 4 investigated whether the *in vivo* identification and quantification of proteins was possible from samples of microdialysis media, collected from the peritendinous space of the Achilles tendon pre and post shockwave therapy. This was an innovative investigation as most tendon studies, so far, have been limited to *in vitro* systems. Although enormously useful, there are limitations with in vitro analysis, as the real time characterization of dynamic protein variations, which fluctuate in different biological conditions such as diseases or ageing, cannot be achieved.

Previous investigations using microdialysis fluid have been successful, for instance, comparative studies into peritoneal dialysate before and after the occurrence of a peritonitis episode by mass spectrometry. This study observed 10 proteins, which were significantly differentially expressed in the dialysate before and after peritonitis. The functional categories that these proteins were from included binding/transport, blood coagulation/hemostasis, and acute phase/immune response (Tyan *et al,* 2013). Cerebral microdialysis media has also been used, with success, to investigate brain metabolism after traumatic brain injury. This study revealed a differential proteome that indicated cellular destruction during the acute phase of illness (Lakshmanan, 2010). Using the above two studies as a small example, proteomic analysis of dialysis fluid has been performed successfully for analysis for other organs, however this technique has never been employed in the field of tendon biology. As described in Chapter 1, tendon tissue is primarily composed of extracellular matrix with a small population of tenocytes (Waggett *et al,* 1998). So it was quite possible that the detection of proteins within microdialysis media wouldn't be as easy.

The fact that this study was able to detect peptides in both pre and post ESWT sample and in addition to this, an upregulation of proteins was seen in post shockwave therapy samples provides us with a great opportunity to explore this subject further. Many of the proteins that were higher in the post ESWT samples

are abundant proteins in serum. This could signify serum leakage following treatment.

During this investigation, samples were kindly received from a study undertaken by Waugh et al, this study detected cytokines IL-1β, IL-2, IL-6 and IL-8, with IL-6 and IL-8 becoming significantly elevated for four hours post ESWT. Alongside this Pro- forms of MMP-2 and -9 were increased after ESWT. It was proposed that ESWT might promote tendon remodeling by promoting inflammatory and catabolic processes that are associated with removing damaged matrix constituents. During our discovery mass spectrometry proteomic analysis, we did not identify any interleukins or MMPs, possibly due to inadequate protein loading, the sensitivity of the instrument or ion suppression due to other peptide components. In previous tendon Mass Spectrometry proteomic studies MMPs have been identified (Sereysky *et al,* 2013).

It is possible that MMPs and IL proteins were present but the amount within the sample was too low to be detected. However the proteins that were detected during this study, do concur with the Waugh et al study in that there was an upregulation of proteins involved in inflammatory pathways such as complements C3, C4 and Immunoglobulin chains,

Overall this investigation should be repeated with a larger sample size. This would give more information and hopefully statistically significant results. Advancing our knowledge of the biology of treatment efficiencies may accelerate the discovery of novel protein biomarkers for numerous treatment applications. The development of accurate *in vivo* analytical methods will steer research into the next stage of investigations hopefully bringing us a step closer to providing better medical care and treatments for injuries, diseases and aging.

Finally, an investigation into tendon turnover was carried out but time was limited and only two time points could be studied. It was established that the non collagenous matrix which consists of proteoglycans had a higher rate of

protein turnover compared to components of the collagenous matrix. Which gave similar conclusions to other research group's data (Thorpe *et al,* 2010). As hoped common tendon proteins were identified, mainly collagens and glycoproteins. Data can be improved by increasing time points and sample numbers, if this occurs then data can be analysed more scientifically with degradation rates and turnover curves (Hammond *et al,* 2016). During this study the top 10 proteins were discussed however other important proteins known to tendon were present. Thrombospondins are glycoproteins, which are involved in multiple roles of mediating cell-cell and cell-matrix interactions. Thrombospondin 4 has been detected in multiple tendon analysis studies (Frolova *et al,* 2014: Little *et al*, 2014: Mueller, 2015) therefore its detection could be a potential marker in this tissue. Mueller, 2015 also identified serpin in tendon tissue, this is a secreted protein, which inhibits angiogenesis and has been detected in previous tendon tissue studies (Farhat *et al,* 2012) and hence could be another potential tendon marker. Extending the research to look into the turnover rates for these proteins would be useful.

The long term aim for this part of the investigation would be to identify the proteomic components of tendon tissue, and to calculate the rate of turnover for each of these. Having a normal baseline would be invaluable as then further turnover studies could be conducted in this area where an intervention was given i.e. a treatment process, such as EWST or medicine e.g. an anti-inflammatory and then turnover of the protein recalculated.

Over the course of this yearlong research project, time was a limiting factor in all of these studies with plenty of further work to be undertaken on each project. The proteomic analysis of biological processes has the potential to provide extremely valuable information regarding the mechanisms of action that occur on a daily basis. The identification of proteins and their abundances can be used to indicate protein roles and interactions. Comparing sample types i.e. disease vs. healthy can give further insight into the pathophysiology that occurs, potentially providing ways in which treatments and therapies can be developed. Efforts

have already been made to investigate the tendon proteome; however there are many gaps in our knowledge and this thesis makes a start at addressing them, as there are many questions to be answered each chapter could be extended further. If time had permitted I certainly would have increased the number of proteins analysed in the tendon turnover study and also looked at other time points which would advance this work further. Then once this data set had been completed I could have progressed to adding in a treatment and assessing any differences in turnover that may occur.

In summery the general outline of this thesis was that the first two chapters of the thesis focused mainly on developing and optimising these techniques. As the ability to produce robust and effective guidelines in sample preparation will be invaluable for future studies in tendon research. Before results are acquired, the techniques in which they are gained must be developed and then optimised. For the final two studies of this thesis we were collecting data rather than developing methods of analysis, however the skills gained for protein extraction investigation were then extrapolated for analysis in both the dialysis chapter and protein turnover chapter. Time limitation was the main problem when performing experiments for MALDI and SIMS imaging. Hopefully with further work, data can be collected and it would be interesting to see whether differences are established, this could then be extended to injured samples and results compared with other studies such as *in vivo* dialysis study to see whether the same proteins are present or with older injuries how the peptide picture has changed. Protein turnover studies in tendon tissue will then allow information to be gleamed about the efficiency of treatments, for instance if a treatment increases the turnover of a known protein involved in a healing process.

We aimed to build a tendon proteomic library of proteins, to discover specific tendon protein markers and look for markers of ageing/disease and repair. If we learn about protein function, abundance, localization and turnover rates, this will all assist in the end objective in discovering new biotechnologies for the treatment of diseases/injuries in humans and animals.

References

Aaron, B. B., and Gosline, J. M. (1981). Elastin as a Random-Network Elastomer: A Mechanical and Optical Analysis of Single Elastin Fibers. Biopolymers *20*, 1247-1260.

Ackermann, P.W. (2015). Chapter 4 - Tendinopathy I: Understanding Epidemiology, Pathology, Healing, and Treatment. In: *Tendon Regeneration*, pp. 113-147. Boston: Academic Press.

Ademowo, O. S., Hernandez, B., Collins, E., Rooney, C., Fearon, U., Kuijk, A. W. van, Tak, P.-P. P.-P., Gerlag, D. M., FitzGerald, O., and Pennington, S. R. (2014). Discovery and confirmation of a protein biomarker panel with potential to predict response to biological therapy in psoriatic arthritis. Annals of the Rheumatic Diseases , 1-8.

Al-Abbad, H. and J. V. Simon (2013). "The effectiveness of extracorporeal shock wave therapy on chronic achilles tendinopathy: a systematic review." Foot Ankle Int 34(1): 33-41.

Andres, B. M. and G. A. Murrell (2008). "Treatment of tendinopathy: what works, what does not, and what is on the horizon." Clin Orthop Relat Res 466(7): 1539-1554.

Avella, C. S., E. R. Ely, K. L. Verheyen, J. S. Price, J. L. Wood and R. K. Smith (2009). "Ultrasonographic assessment of the superficial digital flexor tendons of National Hunt racehorses in training over two racing seasons." Equine Vet J 41(5): 449-454.

Batson, E. L., R. J. Paramour, T. J. Smith, H. L. Birch, J. C. Patterson-Kane and A. E. Goodship (2003). "Are the material properties and matrix composition of equine flexor and extensor tendons determined by their functions?" Equine Vet J 35(3): 314-318.

Benjamin, M., Tyers, R. N., and Ralphs, J. R. (1991). Age-related changes in tendon fibrocartilage. Journal of anatomy *179*, 127-136

Benjamin, M., Kaiser, E., and Milz, S. (2008). Structure-function relationships in tendons: A review. Journal of Anatomy *212*, 211-228.

Beynon, R. J. (2005). The dynamics of the proteome: Strategies for measuring protein turnover on a proteome-wide scale. Briefings in Functional Genomics and Proteomics *3*, 382-390.

Beynon, R. J., and Pratt, J. M. (2005). Metabolic Labeling of Proteins for Proteomics. Mol Cell Proteomics *4*, 857-872.

Biewener, A. A. (1998). "Muscle-tendon stresses and elastic energy storage during locomotion in the horse." Comp Biochem Physiol B Biochem Mol Biol 120(1): 73-87.

Birch, H. L. (2007). "Tendon matrix composition and turnover in relation to functional requirements." Int J Exp Pathol 88(4): 241-248.

Birch, H. L., J. V. Bailey, A. J. Bailey and A. E. Goodship (1999). "Age-related changes to the molecular and cellular components of equine flexor tendons." Equine Vet J 31(5): 391-396.

Birch, H. L., C. T. Thorpe and A. P. Rumian (2013). "Specialisation of extracellular matrix for function in tendons and ligaments." Muscles Ligaments Tendons J 3(1): 12-22.

Birch, H. L., S. Worboys, S. Eissa, B. Jackson, S. Strassburg and P. D. Clegg (2008). "Matrix metabolism rate differs in functionally distinct tendons." Matrix Biol 27(3): 182-189.

Bosch, G., Mos, M. de, Binsbergen, R. van, Schie, H. T. M. van, Lest, C. H. a van de, and Weeren, P. R. van (2009). The effect of focused extracorporeal shock wave therapy on collagen matrix and gene expression in normal tendons and ligaments. Equine veterinary journal *41*, 335-341.

Brown, N. A. T., C. E. Kawcak, C. W. McIlwraith and M. G. Pandy (2003). "Architectural properties of distal forelimb muscles in horses, Equus caballus." Journal of Morphology 258(1): 106-114.

Caprioli, R. M., T. B. Farmer and J. Gile (1997). "Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS." Anal Chem 69(23): 4751-4760.

Chen, E. I. and J. R. Yates, 3rd (2007). "Cancer proteomics by quantitative shotgun proteomics." Mol Oncol 1(2): 144-159.

Chughtai, K. and R. M. Heeren (2010). "Mass spectrometric imaging for biomedical tissue analysis." Chem Rev 110(5): 3237-3277.

Cillero-Pastor, B., G. B. Eijkel, F. J. Blanco and R. M. Heeren (2014). "Protein classification and distribution in osteoarthritic human synovial tissue by matrix-assisted laser desorption ionization mass spectrometry imaging." Anal Bioanal Chem.

Cillero-Pastor, B. and R. M. Heeren (2014). "Matrix-assisted laser desorption ionization mass spectrometry imaging for peptide and protein analyses: a critical review of on-tissue digestion." J Proteome Res 13(2): 325-335.

Clegg, P. D., S. Strassburg and R. K. Smith (2007). "Cell phenotypic variation in normal and damaged tendons." Int J Exp Pathol 88(4): 227-235.

Cobice, D. F., R. J. Goodwin, P. E. Andren, A. Nilsson, C. L. Mackay and R. Andrew (2015). "Future technology insight: mass spectrometry imaging as a tool in drug research and development." Br J Pharmacol 172(13): 3266-3283.

Corps, A. N., Curry, V. A., Harrall, R. L., Dutt, D., Hazleman, B. L., and Riley, G. P. (2003). Ciprofloxacin reduces the stimulation of prostaglandin E2 output by interleukin-1β in human tendon-derived cells. Rheumatology *42*, 1306-1310.

Corr, D. T., C. L. Gallant-Behm, N. G. Shrive and D. A. Hart (2009). "Biomechanical behavior of scar tissue and uninjured skin in a porcine model." Wound Repair Regen 17(2): 250-259.

Dakin, S.G., Dudhia, J and Smith, R.K.W (2013) "Science in brief: Resolving tendon inflammation. A new perspective." Equine Vet J. 45 (4):398-400

Doherty, M. K., and Beynon, R. J. (2006). Protein turnover on the scale of the proteome. Expert review of proteomics *3*, 97-110.

Fessel, G., and Snedeker, J. G. (2011). Equivalent stiffness after glycosaminoglycan depletion in tendon - an ultra-structural finite element model and corresponding experiments. Journal of Theoretical Biology *268*, 77-83.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J., Minguez, P., Bork, P., Von Mering, C., et al. (2013). STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Research *41*.

Franchi, M., M. Raspanti, C. Dell'Orbo, M. Quaranta, V. De Pasquale, V. Ottani and A. Ruggeri (2008). "Different crimp patterns in collagen fibrils relate to the subfibrillar arrangement." Connect Tissue Res 49(2): 85-91.

Frolova, E. G., Drazba, J., Krukovets, I., Kostenko, V., Blech, L., Harry, C., Vasanji, A., Drumm, C., Sul, P., Jenniskens, G. J., et al. (2014). Control of organization and function of muscle and tendon by thrombospondin-4. Matrix Biology *37*, 35-48.

Galatz, Leesa M., Louis Gerstenfeld, Ellen Heber-Katz, and Scott A. Rodeo. 2015. Tendon regeneration and scar formation: The concept of scarless healing. In *Journal of Orthopaedic Research*, 33:823-831.

Gonçalves, A. I., Costa-Almeida, R., Gershovich, P., Rodrigues, M. T., Reis, R. L., and Gomes, M. E. (2015). Cell Based Approaches for Tendon Regeneration (Elsevier Inc.)

Gordon, J. a, Freedman, B. R., Zuskov, a, Iozzo, R. V., Birk, D. E., and Soslowsky, L. J. (2015). Achilles tendons from decorin- and biglycan-null mouse models have

inferior mechanical and structural properties predicted by an image-based empirical damage model. Journal of biomechanics *48*, 2110-5.

Gosline, J., M. Lillie, E. Carrington, P. Guerette, C. Ortlepp and K. Savage (2002). "Elastic proteins: biological roles and mechanical properties." Philos Trans R Soc Lond B Biol Sci 357(1418): 121-132.

Haake, M., Böddeker, I. R., Decker, T., Buch, M., Vogel, M., Labek, G., Maier, M., Loew, M., Maier-Boerries, O., Fischer, J., et al. (2002). Side-effects of extracorporeal shock wave therapy (ESWT) in the treatment of tennis elbow. Archives of Orthopaedic and Trauma Surgery *122*, 222-228.

Hammond, D. E., Claydon, A. J., Simpson, D. M., Edward, D., Stockley, P., Hurst, J. L., and Beynon, R. J. (2016). Proteome dynamics: tissue variation in the kinetics of proteostasis in intact animals. Molecular & cellular proteomics. 15 (4) 1204-19.

Han, S. H., Lee, J. W., Guyton, G. P., Parks, B. G., Courneya, J.-P., and Schon, L. C. (2009). Effect of Extracorporeal Shock Wave Therapy on Cultured Tenocytes. Foot & Ankle International *30*, 93-98.

Heinemeier, K. M., Schjerling, P., Heinemeier, J., Magnusson, S. P., and Kjaer, M. (2013). Lack of tissue renewal in human adult Achilles tendon is revealed by nuclear bomb 14C. FASEB Journal *27*, 2074-2079.

Hosaka, Y., Kirisawa, R., Yamamoto, E., Ueda, H., Iwai, H., and Takehana, K. (2002). Localization of cytokines in tendinocytes of the superficial digital flexor tendon in the horse. The Journal of veterinary medical science *64*, 945-947.

Jarvinen, T. A., P. Kannus, T. L. Jarvinen, L. Jozsa, H. Kalimo and M. Jarvinen (2000). "Tenascin-C in the pathobiology and healing process of musculoskeletal tissue injury." Scand J Med Sci Sports 10(6): 376-382.

Jung, H. J., M. B. Fisher and S. L. Woo (2009). "Role of biomechanics in the understanding of normal, injured, and healing ligaments and tendons." Sports Med Arthrosc Rehabil Ther Technol 1(1): 9.

Kalamajski, S., Liu, C., Tillgren, V., Rubin, K., Oldberg, ??ke, Rai, J., Weis, M. A., and Eyre, D. R. (2014). Increased c-telopeptide cross-linking of tendon type i collagen in fibromodulin-deficient mice. Journal of Biological Chemistry *289*, 18873-18879.

Kaletas, B. K., I. M. van der Wiel, J. Stauber, C. Guzel, J. M. Kros, T. M. Luider and R. M. Heeren (2009). "Sample preparation issues for tissue imaging by imaging MS." Proteomics 9(10): 2622-2633.

Kasashima, Y., T. Takahashi, R. K. Smith, A. E. Goodship, A. Kuwano, T. Ueno and S. Hirano (2004). "Prevalence of superficial digital flexor tendonitis and

suspensory desmitis in Japanese Thoroughbred flat racehorses in 1999." Equine Vet J 36(4): 346-350.

Kastelic, J., A. Galeski and E. Baer (1978). "The multicomposite structure of tendon." Connect Tissue Res 6(1): 11-23.

Khan, K. M., Cook, J. L., Bonar, F., Harcourt, P., and Åstrom, M. (1999). Histopathology of common tendinopathies: Update and implications for clinical management. Sports Medicine *27*, 393-408.

Kim, T.-Y., Wang, D., Kim, A. K., Lau, E., Lin, A. J., Liem, D. A., Zhang, J., Zong, N. C., Lam, M. P. Y., and Ping, P. (2012). Metabolic labeling reveals proteome dynamics of mouse mitochondria. Molecular {&} Cellular Proteomics *11*, 1586-1594.

Kjaer, M. (2004). "Role of extracellular matrix in adaptation of tendon and skeletal muscle to mechanical loading." Physiol Rev 84(2): 649-698.

Kjaer, M. (2009). "The secrets of matrix mechanical loading." Scand J Med Sci Sports 19(4): 455-456.

Kjaer, M. and K. M. Heinemeier (2014). "Eccentric exercise: acute and chronic effects on healthy and diseased tendons." J Appl Physiol (1985) 116(11): 1435-1438.

Knobloch, K., U. Yoon and P. M. Vogt (2008). "Acute and overuse injuries correlated to hours of training in master running athletes." Foot Ankle Int 29(7): 671-676.

Kohrs, R. T., Zhao, C., Sun, Y. L., Jay, G. D., Zhang, L., Warman, M. L., An, K. N., and Amadio, P. C. (2011). Tendon fascicle gliding in wild type, heterozygous, and lubricin knockout mice. Journal of Orthopaedic Research *29*, 384-389.
Korol, R. M., H. M. Finlay, M. J. Josseau, A. R. Lucas and P. B. Canham (2007). "Fluorescence spectroscopy and birefringence of molecular changes in maturing rat tail tendon." J Biomed Opt 12(2): 024011.

Kostrominova, T. Y., and Brooks, S. V. (2013). Age-related changes in structure and extracellular matrix protein expression levels in rat tendons. Age (Dordrecht, Netherlands) *35*, 2203-14.

Kujala, U. M., S. Sarna and J. Kaprio (2005). "Cumulative incidence of achilles tendon rupture and tendinopathy in male former elite athletes." Clin J Sport Med 15(3): 133-135.

Lakshmanan, R., Loo, J. A., Drake, T., Leblanc, J., Ytterberg, A. J., McArthur, D. L., Etchepare, M., and Vespa, P. M. (2010). Metabolic crisis after traumatic brain injury is associated with a novel microdialysis proteome. Neurocritical Care *12*, 324-336.

Lemaire, R., J. C. Tabet, P. Ducoroy, J. B. Hendra, M. Salzet and I. Fournier (2006). "Solid ionic matrixes for direct tissue analysis and MALDI imaging." Anal Chem 78(3): 809-819.

Little, D., Thompson, J. W., Dubois, L. G., Ruch, D. S., Moseley, M. A., and Guilak, F. (2014). Proteomic differences between male and female anterior cruciate ligament and patellar tendon. PLoS ONE *9*.

Mani-Babu, S., D. Morrissey, C. Waugh, H. Screen and C. Barton (2015). "The effectiveness of extracorporeal shock wave therapy in lower limb tendinopathy: a systematic review." Am J Sports Med 43(3): 752-761.

Marr, C. M., S. Love, J. S. Boyd and Q. McKellar (1993). "Factors affecting the clinical outcome of injuries to the superficial digital flexor tendon in National Hunt and point-to-point racehorses." Vet Rec 132(19): 476-479.

McClure, S., and Dorfmüller, C. (2003). Extracorporeal shock wave therapy: Theory and equipment. Clinical Techniques in Equine Practice *2*, 348-357.

McDonnell, L. A. and R. M. Heeren (2007). "Imaging mass spectrometry." Mass Spectrom Rev 26(4): 606-643.

Meetani, M. A., and Voorhees, K. J. (2005). MALDI mass spectrometry analysis of high molecular weight proteins from whole bacterial cells: pretreatment of samples with surfactants. J Am Soc Mass Spectrom *16*, 1422-1426.

Mi, H., Lazareva-Ulitsky, B., Loo, R., Kejariwal, A., Vandergriff, J., Rabkin, S., Guo, N., Muruganujan, A., Doremieux, O., Campbell, M. J., et al. (2005). The PANTHER database of protein families, subfamilies, functions and pathways. Nucleic Acids Research *33*.

Muller, A. (2015) A systems biology approach to musculoskeletal tissue engineering:transcriptomic and proteomic analysis of cartilage and tendon cells. PhD thesis. University of Liverpool.

Naglak, T. J., and Wang, H. Y. (1990). Recovery of a foreign protein from the periplasm of Escherichia coli by chemical permeabilization. Enzyme and Microbial Technology *12*, 603-611.

Norrgran, J., T. L. Williams, A. R. Woolfitt, M. I. Solano, J. L. Pirkle and J. R. Barr (2009). "Optimization of digestion parameters for protein quantification." Anal Biochem 393(1): 48-55.

Notarnicola, A. and B. Moretti (2012). "The biological effects of extracorporeal shock wave therapy (eswt) on tendon tissue." Muscles Ligaments Tendons J 2(1): 33-37.

Nourissat, G., F. Berenbaum and D. Duprez (2015). "Tendon injury: from biology to tendon repair." Nat Rev Rheumatol 11(4): 223-233.

Patel, V. J., Thalassinos, K., Slade, S. E., Connolly, J. B., Crombie, A., Murrell, J. C., and Scrivens, J. H. (2009). A comparison of labeling and label-free mass spectrometry-based proteomics approaches. Journal of Proteome Research *8*, 3752-3759.

Patel, E., Clench, M. R., West, A., Marshall, P. S., Marshall, N., & Francese, S. (2015). Alternative Surfactants for Improved Efficiency of In Situ Tryptic Proteolysis of Fingermarks. *Journal of the American Society for Mass Spectrometry*, *26*(6), 862–872.

Patterson-Kane, J. C. and T. Rich (2014). "Achilles Tendon Injuries in Elite Athletes: Lessons in Pathophysiology from Their Equine Counterparts." ILAR Journal 55(1): 86-99.

Pawitan, Y., Michiels, S., Koscielny, S., Gusnanto, A., and Ploner, A. (2005). False discovery rate, sensitivity and sample size for microarray studies. Bioinformatics (Oxford, England) *21*, 3017-24.

Peffers, M. J., Thorpe, C. T., Collins, J. A., Eong, R., Wei, T. K. J., Screen, H. R. C., and Clegg, P. D. (2014). Proteomic analysis reveals age-related changes in tendon matrix composition, with age- and injury-specific matrix fragmentation. Journal of Biological Chemistry *289*, 25867-25878.

Peffers, M. J., B. McDermott, P. D. Clegg and C. M. Riggs (2015). "Comprehensive protein profiling of synovial fluid in osteoarthritis following protein equalization." Osteoarthritis Cartilage.

Pool, R. R. and D. M. Meagher (1990). "Pathologic findings and pathogenesis of racetrack injuries." Vet Clin North Am Equine Pract 6(1): 1-30.

Price, S. R., Du, J., Bailey, J. L., Mitch, W.E. (2001) Molecular mechanisms regulating protein turnover in muscle. Am J Kidney Dis 37. (1) S112-S114

Proc, J. L., Kuzyk, M. A., Hardie, D. B., Yang, J., Smith, D. S., Jackson, A. M., Borchers, C. H. (2010). A Quantitative Study of the Effects of Chaotropic Agents, Surfactants, and Solvents on the Digestion Efficiency of Human Plasma Proteins by Trypsin. *Journal of Proteome Research*, *9*(10), 5422–5437.

Rees, S. G., C. M. Dent and B. Caterson (2009). "Metabolism of proteoglycans in tendon." Scand J Med Sci Sports 19(4): 470-478.

Riley, G. (2008). "Tendinopathy--from basic science to treatment." Nat Clin Pract Rheumatol 4(2): 82-89.

Riley, G. P., V. Curry, J. DeGroot, B. van El, N. Verzijl, B. L. Hazleman and R. A. Bank (2002). "Matrix metalloproteinase activities and their relationship with collagen remodelling in tendon pathology." Matrix Biology 21(2): 185-195.

Ritty, T. M., Broekelmann, C. C., Mecham, R. P. (2003) Fibrillin-1 and −2 contain heparin-binding sites important for matrix deposition and that support cell attachment. Biochemical Journal. 375 (2) 424-432.

Seeley, E. H. and R. M. Caprioli (2008). "Imaging mass spectrometry: Towards clinical diagnostics." Proteomics Clin Appl 2(10-11): 1435-1443.

Seeley, E. H. and R. M. Caprioli (2008). "Molecular imaging of proteins in tissues by mass spectrometry." Proc Natl Acad Sci U S A 105(47): 18126-18131.

Sejersen, M. H. J., Frost, P., Hansen, T. B., Deutch, S. R., and Svendsen, S. W. (2015). Proteomics perspectives in rotator cuff research a systematic review of gene expression and protein composition in human tendinopathy. PLoS ONE *10*.

Sereysky, J. B., E. L. Flatow and N. Andarawis-Puri (2013). "Musculoskeletal regeneration and its implications for the treatment of tendinopathy." Int J Exp Pathol 94(4): 293-303.

Shrier, I., Matheson, G. O., and Kohl, H. W. (1996). Achilles tendonitis: are corticosteroid injections useful or harmful? Clinical journal of sport medicine : official journal of the Canadian Academy of Sport Medicine *6*, 245-50.

Silver, F. H., J. W. Freeman and G. P. Seehra (2003). "Collagen self-assembly and the development of tendon mechanical properties." Journal of Biomechanics 36(10): 1529-1553.

Simpson, J. C., Wellenreuther, R., Poustka, A., Pepperkok, R., and Wiemann, S. (2000). Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing. EMBO reports *1*, 287-92.

Spiesz, E. M., C. T. Thorpe, S. Chaudhry, G. P. Riley, H. L. Birch, P. D. Clegg and H. R. Screen (2015). "Tendon extracellular matrix damage, degradation and inflammation in response to in-vitro overload exercise." J Orthop Res.

Stephens, P. R., D. M. Nunamaker and D. M. Butterweck (1989). "Application of a Hall-effect transducer for measurement of tendon strains in horses." American Journal of Veterinary Research 50(7): 1089-1095.

Stromberg, B. and G. Tufvesson (1969). "Lesions of the superficial flexor tendon in race horses. A microangiographic and histopathologic study." Clin Orthop Relat Res 62: 113-123.

Thanawiroon, C., Rice, K.G., Toida, T and Linhardt, R.J. (2004) Liquid Chromatography/Mass Spectrometry Sequencing Approach for Highly Sulfated

Heparin-derived Oligosaccharides. J. Biol. Chem. 2004 279: 2608

Thorpe, C. T., H. L. Birch, P. D. Clegg and H. R. Screen (2013). "The role of the non-collagenous matrix in tendon function." Int J Exp Pathol 94(4): 248-259.

Thorpe, C. T., C. Klemt, G. P. Riley, H. L. Birch, P. D. Clegg and H. R. Screen (2013). "Helical sub-structures in energy-storing tendons provide a possible mechanism for efficient energy storage and return." Acta Biomater 9(8): 7948-7956.

Thorpe, C. T., B. T. McDermott, A. E. Goodship, P. D. Clegg and H. L. Birch (2015). "Ageing does not result in a decline in cell synthetic activity in an injury prone tendon." Scand J Med Sci Sports.

Thorpe, C. T., G. P. Riley, H. L. Birch, P. D. Clegg and H. R. Screen (2014). "Fascicles from energy-storing tendons show an age-specific response to cyclic fatigue loading." J R Soc Interface 11(92): 20131058.

Thorpe, C. T., Peffers, M. J., Simpson, D., Halliwell, E., Screen, H. R. C., and Clegg, P. D. (2016). Anatomical heterogeneity of tendon: Fascicular and interfascicular tendon compartments have distinct proteomic composition. Scientific reports *6*, 20455.

Thorpe, C. T., E. M. Spiesz, S. Chaudhry, H. R. Screen and P. D. Clegg (2015). "Science in brief: recent advances into understanding tendon function and injury risk." Equine Vet J 47(2): 137-140.

Thorpe, C. T., Streeter, I., Pinchbeck, G. L., Goodship, A. E., Clegg, P. D., and Birch, H. L. (2010). Aspartic acid racemization and collagen degradation markers reveal an accumulation of damage in tendon collagen that is enhanced with aging. Journal of Biological Chemistry *285*, 15674-15681.

Thorpe, C. T., C. P. Udeze, H. L. Birch, P. D. Clegg and H. R. Screen (2012). "Specialization of tendon mechanical properties results from interfascicular differences." J R Soc Interface 9(76): 3108-3117.

Thorpe, C. T., C. P. Udeze, H. L. Birch, P. D. Clegg and H. R. Screen (2013). "Capacity for sliding between tendon fascicles decreases with ageing in injury prone equine tendons: a possible mechanism for age-related tendinopathy?" Eur Cell Mater 25: 48-60.

Todd, P. J., Gregory Schaaff, T., Chaurand, P., and Caprioli, R. M. (2001). Organic ion imaging of biological tissue with secondary ion mass spectrometry and matrix-assisted laser desorption/ionization. Journal of Mass Spectrometry *36*, 355-369.

Tsui, K. H., P. Tang, C. Y. Lin, P. L. Chang, C. H. Chang and B. Y. Yung (2010). "Bikunin loss in urine as useful marker for bladder carcinoma." J Urol 183(1): 339-344.

Tyan, Y. C., Su, S. B., Ting, S. S., Wang, H. Y., and Liao, P. C. (2013). A comparative proteomics analysis of peritoneal dialysate before and after the occurrence of peritonitis episode by mass spectrometry. Clinica Chimica Acta *420*, 34-44.

Uquillas, C. a, Guss, M. S., Ryan, D. J., Jazrawi, L. M., and Strauss, E. J. (2015). Everything Achilles: Knowledge Update and Current Concepts in Management: AAOS Exhibit Selection. The Journal of Bone & Joint Surgery *97*, 1187-1195.

Verrills, N. M. (2006). "Clinical proteomics: present and future prospects." Clin Biochem Rev 27(2): 99-116.

Voleti, Pramod B., Mark R. Buckley, and Louis J. Soslowsky. 2012. Tendon Healing: Repair and Regeneration. *Annual Review of Biomedical Engineering* 14: 47-71.

Waggett, A. D., Ralphs, J. R., Kwan, A. P. L., Woodnutt, D., and Benjamin, M. (1998). Characterization of collagens and proteoglycans at the insertion of the human Achilles tendon. Matrix Biology *16*, 457-470.

Warden, S. J. (2007). Animal models for the study of tendinopathy. British journal of sports medicine *41*, 232-240.

Waugh, C. M., D. Morrissey, E. Jones, G. P. Riley, H. Langberg and H. R. Screen (2015). "In vivo biological response to extracorporeal shockwave therapy in human tendinopathy." Eur Cell Mater 29: 268-280; discussion 280.

Webbon, P. M. (1977). "A post mortem study of equine digital flexor tendons." Equine Vet J 9(2): 61-67.

Wilson, J. J. and T. M. Best (2005). "Common overuse tendon problems: A review and recommendations for treatment." Am Fam Physician 72(5): 811-818.

Wilson, R., Diseberg, A. F., Gordon, L., Zivkovic, S., Tatarczuch, L., Mackie, E. J., Gorman, J. J., and Bateman, J. F. (2010). Comprehensive profiling of cartilage extracellular matrix formation and maturation using sequential extraction and label-free quantitative proteomics. Molecular & cellular proteomics : MCP *9*, 1296-313.

Worthington, B. C. (2011). from http://www.worthington-biochem.com/try/default.html.

Yoon, J. H. and J. Halper (2005). "Tendon proteoglycans: biochemistry and function." J Musculoskelet Neuronal Interact 5(1): 22-34.

Yu YQ, Gilar M, Lee PJ, Bouvier ES, Gebler JC. (2003) Enzyme-friendly, mass spectrom- etry-compatible surfactant for in-solution enzymatic digestion of proteins. Anal. Chem. 75: 6023-6028.