Impact of therapeutic ionising radiation on extracellular matrix proteins with application to breast tissue

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2023

Ren Jie Tuieng

School of Biological Sciences

Contents

Word count: 35,943

List of Tables

List of Figures

Abbreviations

AA	Amino acid
ADAMTS	A disintegin and metalloproteinase with thrombospondin motifs
AFM	Atomic force microscopy
ANOVA	Analysis of variance
AN-PAGE	Acidic native polyacrylamide gel electrophoresis
BN-PAGE	Blue native polyacrylamide gel electrophoresis
CD	Circular dichroism
COL1A1	Collagen I alpha 1 chain
cFN	Cellular fibronectin
CMP	Collagen mimetic peptides
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DSF	Differential scanning fluorimetry
DTT	Dithiothreitol
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EL	EGF-like domain
EMILIN	Elastin microfibril interfacers
FA	Formic acid
fECM	*In vitro* fibroblast derived extracellular matrix
FN	Fibronectin
GO	Gene ontology
Gy	Gray (SI unit for radiation dose, J/kg)
HER-2	Human epidermal growth factor receptor-2
IAM	Iodoacetamide
LC-MS/MS	Liquid chromatography with tandem mass spectrometry
LE	Laminin EGF-like domain
LINAC	Linear accelerator
LOX	Lysyl oxidase
LTBP	Latent TGFβ-binding protein
MMP	Matrix metalloproteinases
MOPS	3-morpholinopropane-1-sulfonic acid
MS1	Mass spectrometry precursor ion
OCT	Optimal cutting temperature compound
PBS	Phosphate-buffered saline
pFN	Plasma fibronectin
PSM	Peptide spectrum match
P/S	Penicillin/streptomycin
ROS	Reactive oxygen species
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sv	Sieverts (unit for effective radiation dose)
TBS	Tris buffered saline
tECM	*Ex vivo* tissue extracellular matrix

TGFβ	Transforming growth factor beta
TIMP	Tissue inhibitor of metalloproteinase
TNBC	Triple negative breast cancer
UVR	Ultraviolet radiation
VEGF	Vascular endothelial growth factor
βig-H3	TGFβ-induced glycoprotein-H3

Abstract

Breast cancer is the most common cancer for women around the world and advances in its treatment, such as in radiation therapy, have significantly improved prognosis and survival rates. Quality of life of survivors are, however, impacted by late side effects of treatment, like fibrosis, which is prevalent in breast radiotherapy patients. Current knowledge on the physiopathology of radiation-induced late side effects remains incomplete. Research had focused on the cellular impact of radiation while in contrast, the extracellular matrix (ECM) had been overlooked despite their importance in governing cellular behaviour and outcomes.

To improve the understanding of roles that ECM proteins may play in mediating radiation-induced side effects, this PhD aimed to test the hypothesis that breast associated ECM proteins are profoundly altered by therapeutic doses of x-rays, using biomolecular techniques and a mass spectrometry-based analysis—peptide location fingerprinting (PLF). PLF is highly sensitive to the changes in proteolytic susceptibility of large ECM assemblies, allowing detection of regional structural changes within ECM proteins.

Purified human collagen I and plasma fibronectin were first chosen as exemplar ECM proteins given their distinct composition and molecular structure. Collagen I and fibronectin responded differently to therapeutic x-rays; the primary structure of x-ray exposed collagen was fragmented with minimal impact to its triple helical structure, while irradiated fibronectin had structurally altered binding sites, resulting in increased binding to collagen.

Many ECM components, including fibrillar collagens, exist as higher-order assemblies in tissue. To examine if the packing of fibrillar collagen I could impact its response to therapeutic x-rays, three increasingly complex forms of rat collagen were investigated: solubilised monomeric collagen, reconstituted collagen gel, and ex vivo tendon collagen. While solubilised collagen was fragmented by therapeutic x-rays, fibrillar collagen (gels and tendons) were not. However, proteolytic susceptibility (as analysed by PLF), and therefore structure, was altered in gel and tendon. The ultrastructure of collagen I appears to be crucial in determining its peptide fingerprint's response to therapeutic x-ray exposure.

Finally, the impact of therapeutic x-ray doses in complex ECM-rich proteomes (*in vitro* fibroblast derived ECM (fECM) with ex vivo tissue ECM (tECM)) were analysed using PLF to screen for ECM proteins compromised by therapeutic x-rays. Mass spectrometry revealed similar collagen composition in both fECM and tECM. Collagen I was also found to be structurally similar in both matrices, but not for FN and basement membrane proteins. However, therapeutic x-rays was found to alter the proteolytic susceptibility of these ECM proteins regardless of their source, which may implicate their biological functions *in vivo*.

In conclusion, this thesis successfully characterised the impact of therapeutic x-rays on crucial ECM proteins and demonstrated the capabilities of PLF as a multiscale proteomics tool. The work also identified a body of ECM proteins found in tissues, including the breast stroma, which are vulnerable to therapeutic x-rays. Further work will be needed to establish if these proteins are also affected by x-ray exposure *in vivo* and what the consequences might be for tissue physiology and the development of radiotherapy related pathologies.

Thesis Outline

This thesis is organised into six main chapters. Chapter 1 encompasses the introduction to the motivation behind working on therapeutic x-rays, its clinical importance for breast cancer treatment, and why the extracellular matrix (ECM) was chosen as our focus. A background of current literature on electromagnetic radiation and its biological impact was given, and studies showing the impact of radiation on ECM proteins were highlighted. A brief overview of breast tissue and its ECM components were also mentioned.

Chapter 2 comprises the methodologies used in all experiments throughout the thesis. This includes how the material were sourced, biochemical methods like gel electrophoresis, protein characterisation techniques like light scattering and circular dichroism. Atomic force microscopy and mass spectrometry methods were also described. The peptide location fingerprinting method and modifications were also illustrated.

Chapter 3 to 5 contains the thesis main findings and results, with chapter 3 starting off with the work on the simple purified collagen I and fibronectin to illustrate isolated molecular responses of ECM proteins with differing molecular structure. Chapter 4 brings the focus back to collagen I but in different hierarchical structures. Therapeutic x-ray responses were compared between solubilised collagen I, reconstituted collagen gel, and rat tail tendons. Finally, chapter 5 explores the responses of therapeutic x-ray exposure on *in vitro* fibroblast derived ECM with *ex vivo* breast tissue using PLF to screen for ECM proteins vulnerable to radiation damage.

Chapter 6, the final chapter, concludes the thesis and provides the limitations of the techniques and work presented. Future work that could derive from this study was also discussed.

Declaration

The author declares that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

Acknowledgements

I want to begin by expressing my gratitude to my supervisors, Professor Michael Sherratt, Professor Sarah Cartmell, and Professor Cliona Kirwan, for their unwavering support throughout my PhD journey, and also for their faith in selecting me (a physics undergraduate) for a cell matrix biology research project. Thank you, Mike, for being so supportive of the experiments I propose and always being available for discussions whenever things go wrong. The freedom you have given me to explore and take the PhD in my own direction has helped me learn how to think independently, but also made me realise there are many things I do not know and that I still have to work on. Sarah, thank you for the wealth of advice you have given me not just for the project, but for the academic life and beyond. I really appreciate the hours you have given me despite having to juggle between academia and Retendon. You are, and will continue to be, an inspiration. Cliona, thank you for always being so encouraging and reassuring even when things were bad. You've taught me to focus on the silver lining and to take life less seriously – I am really lucky to have the opportunity to learn from such a multidisciplinary surgeon/clinician scientist.

I would also like to thank lab members, both in the Sherratt and Cartmell group, for their support and companionship. In particular, Callum and Zeyad, who were instrumental in giving me a social life outside the lab (thanks for the occasional evening pubs) and would always check in on me to ensure I was still breathing. I am also grateful to Nathan, who was always willing to share his technical expertise to anyone in the lab. And finally, Josie, whose constant presence in the lab (even on weekends) spurred me to work harder.

I would also like to acknowledge the help from the core facilities in the University of Manchester, especially Dr Stacey Warwood, Dr Ronan O'Cualain, Dr David Knight, James Allsey, and Emma-jayne Keevil from the BioMS core facility for their technical support and training in mass spectrometry. Also Dr Nigel Hodson, for the going above and beyond in helping me with the AFM and even brainstorming ideas with me. Not forgetting Dr Thomas Jowitt and Natalie Terry from the Biomolecular analysis core facility for the help in CD and MALS, and Dr Stuart Cain and Dr Rana Dajani for their help and generous donations of reagents in my (failed) attempt to purify fibronectin. I would also like to thank Dr David Holmes for teaching me the crucial methods of obtaining collagen fibrils from rat tendons and Dr Aleksandr Mironov for generously allowing me to use the facilities at the EM laboratory. I would also like to thank the staff at the MCRC biobank and Biological Services Facility at Manchester for the help with collection of the breast tissue and rat tendon samples and last but not least, the Singapore Nuclear Research and Safety Initiative for sponsoring this PhD.

I would like to dedicate this thesis to my family back home in Singapore, Lao ba, Lao ma, and Lao jie, for their constant care and support despite being so far away. I also want to express my appreciation to my parents, who sacrificed their own education to provide a living so that I could have the privilege of pursuing my aspirations, giving me the chance to be where I am today.

Finally, I would like to thank my partner in life, Ya-Ting, for making precious memories with me in Manchester, and for tolerating the 8-hour time difference while we are apart. Thank you for the countless times you had to stay awake just to have a phone call even after a tiring day at work. I will be forever grateful for your care, support, and sacrifice throughout my PhD journey.

List of publications and conference presentations

Publications

1. **Tuieng, R.J.**; Cartmell, S.H.; Kirwan, C.C.; Sherratt, M.J. The Effects of Ionising and Non-Ionising Electromagnetic Radiation on Extracellular Matrix Proteins. Cells 2021, 10, 3041. https://doi.org/10.3390/cells10113041

Part of the introduction (Chapter 1) has been taken directly from the above publication in which the Author was the first author responsible for the majority of text and the first draft [1]. The specific sections in this thesis that were included in the published work include: Section 1.2, 1.3, 1.4, and 1.5.

Conference presentations

1. **Tuieng, R.J.**; Cartmell, S.H.; Kirwan, C.C.; Sherratt, M.J. Structural remodelling of collagen I and fibronectin by therapeutic x-ray doses, British Society for Matrix Biology conference at Liverpool, “The Matrix in Development”, Sept 2022

1 Introduction

Breast cancer is the leading cancer in females around the world, with an estimated 53,889 new cases in the UK alone for 2020 (more than double those due to the next leading cancer – lung (25000))[2] (Figure 1.1). A key treatment modality for breast cancer is radiotherapy, which utilises ionising radiation (usually photons in the MeV range) to target and kill cancer cells through deoxyribonucleic acid (DNA) damage via the production of reactive oxygen species (ROS). Whilst proven highly effective as a non-invasive treatment that has helped increased survival rates and decreased recurrence rates in breast cancer [3], breast radiotherapy inadvertently damages normal tissues surrounding the tumour site, specifically the heart and lung [4–6]. This can induce unwanted side effects, including fibrosis and secondary cancers, in those tissues, which reduces overall quality of life for patients. This issue is pertinent not only in breast radiotherapy, but also radiation therapy employed in the treatment of other cancers. For example, gastrointestinal toxicity can occur for prostate cancer radiotherapy [7], and osteoradionecrosis can result from radiation therapy of head and neck cancers [8].

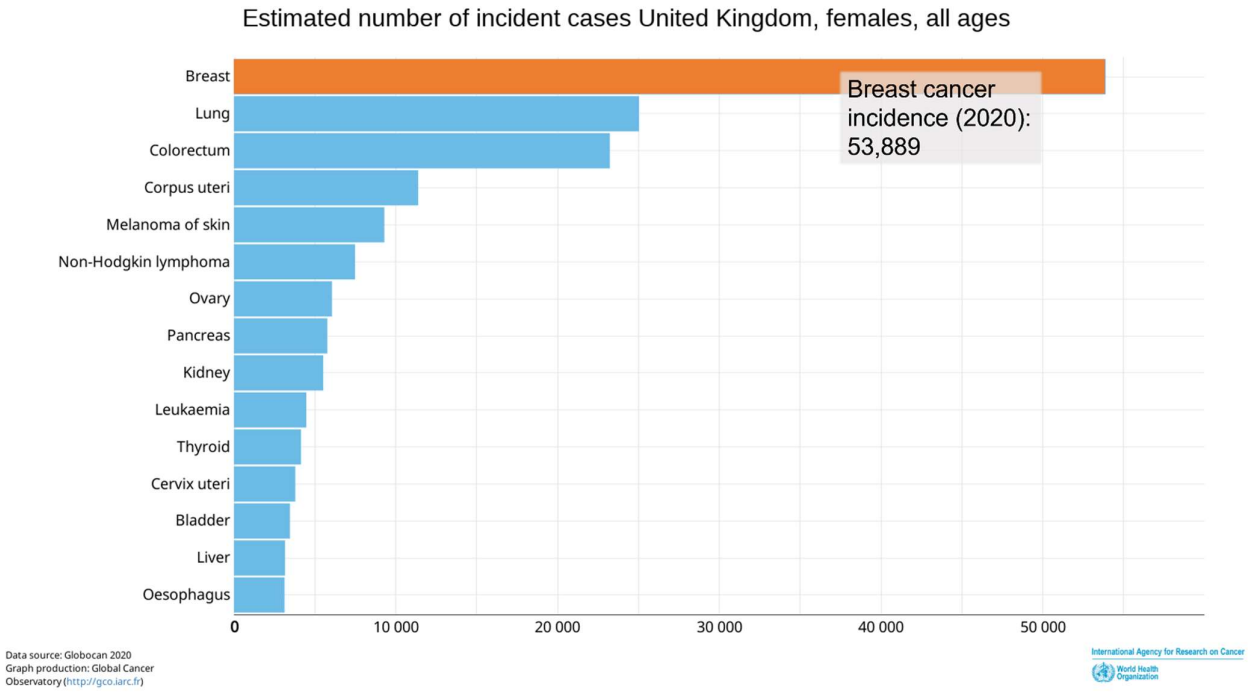


Figure 1.1. Estimated cancer incidence for females of all ages in the UK in 2020. Breast cancer leads all other cancers as the highest cancer incidence for females in the world and in the UK, with approximately 53,889 new incidents in the UK and more than 2.1 million worldwide [1]. With radiation therapy as a key treatment modality for breast cancer, there is a need to improve the post-treatment quality of life of patients by minimising the occurrence of side effects. Figure adapted from Global Cancer Observatory (http://gco.iarc.fr), accessed on 27 Feb 2023.

The pathology of acute high radiation exposure has been well understood, where high doses of ionising radiation cause extensive DNA and protein damage, driving a cell to apoptosis. However, it has become increasingly clear that lower, clinically relevant radiation doses may have sub-lethal effects that are complex, difficult to eliminate and delayed (persisting over long periods of time) [8–11]. These outcomes of radiation therapy are termed late side effects, which occur months or even years post-treatment, and remains least well-studied and understood as most investigations into the detrimental side effects of radiation on biological tissues have largely focused on cellular damage, and in particular, the sensitivity of DNA [12,13]. To understand the consequences of therapeutic radiation exposure and hence to potentially prevent or reverse the damage, it is necessary to characterise the interactions of ionising radiation with not only cells but also with their complex and dynamic extracellular environment.

Recent advances in understanding the extracellular environment have elucidated the importance of the extracellular matrix (ECM) in cellular behaviour. The ECM provides not only mechanical support but also acts as a storage for biochemical factors, with multiple cellular binding sites that translate mechanical cues to alter cell behaviour and even cell fate [14–19]. Current literature, however, does not provide a complete understanding of ECM proteins' response to ionising radiation, and whether they play a crucial role in mediating side effects of therapeutic ionising radiation exposure is still up for debate. This is in part due to how molecular mechanisms of radiation damage to ECM proteins is poorly defined. This thesis thus aims to improve the understanding of radiation damage to ECM proteins by utilising multiple biochemical and biophysical methods to explore if therapeutically relevant doses of ionising radiation can damage ECM proteins, specifically those found in breast and other radiation susceptible organs and tissues. Further, this thesis seeks to interrogate how those ECM proteins may respond differently in different states and experimental systems, which may improve our molecular understanding of radiation damage to proteins beyond the ECM.

This thesis is organised into five chapters. The introduction (chapter 1), provides a brief overview of breast cancer and ionising radiation and its biological effects, particularly in breast tissues exposed to therapeutic X-rays [1]. Chapter 2 describes the methodologies used in the subsequent sections to characterise the impact of radiation damage on the ECM.

Chapter 3 to 5 presents three different studies that aim to elucidate different aspects of the mechanism governing ECM response to ionising radiation exposure:

Chapter 3 – To compare therapeutic ionising radiation response between two structurally different ECM proteins, collagen I (fibrillar) and fibronectin (globular).

Chapter 4 – To investigate the importance of ultrastructure of collagen I, using rat tail collagen I in different forms, in mediating its response to therapeutic ionising radiation.

Chapter 5 – To evaluate the radiation response of fibroblast derived ECM proteins (*in vitro*) and tissue ECM from biopsies.

The final chapter discusses the key insights with reference to the peer reviewed literature, considers the strengths and limitations of the studies conducted and suggests potential future studies and outlook.

1.1 Breast cancer

1.1.1 Breast tissue

The breast is made up of three main tissue types: stromal tissue, which comprises mostly of collagen I, adipose tissue, made primarily of adipocytes, and glandular tissue, which houses the mammary glands (Figure 1.2).

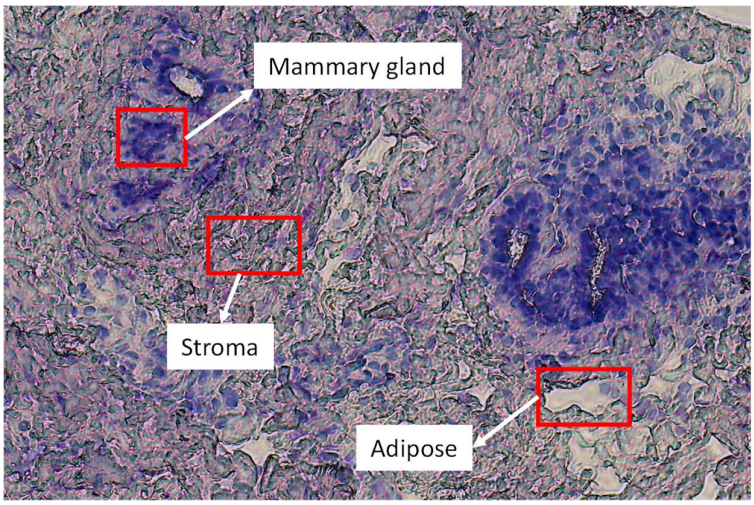


Figure 1.2. Hematoxylin and eosin (H&E) stain of a histological section of breast tissue. The breast tissue contains mammary glands, stromal tissue, and adipose tissue, which are indicated on the figure in the red boxes.

The primary functional architecture of breast tissue is the mammary gland, which contains epithelial 'trees' surrounded by stroma and adipose tissue [20,21]. The stroma comprises mainly collagens and an elastin fibre network that provides the mechanical structure and integrity of the tissue [22,23]. The breast epithelium comprises a spherical alveoli with a tubular ductal network, which contains a layer of apical luminal cells that make milk in lactation, and a layer of basal myoepithelial cells that contracts around the alveoli to squeeze milk into the ducts and to the nipple. A thin, continuous extracellular matrix (ECM) network, also called the basement membrane, surrounds all of the breast epithelium and acts as a barrier between the epithelium and the stroma.

1.1.2 Types of breast cancer

Breast cancer comprises four major classifications based on presence or absence of progesterone/oestrogen receptors in the tumour cells. The classifications include luminal A and B [24,25], human epidermal growth factor receptor-2 (HER-2) positive [26], and triple negative breast cancer (TNBC) [27]. In general, luminal A/B breast cancer have both hormone receptors are present, but luminal A lacks HER-2 expression, while luminal B may have HER-2 expression. HER-2 positive breast cancer lacks both hormone receptors but expresses HER-2, and finally triple negative breast cancer lacks both the hormone receptors and HER-2.

1.1.3 Breast cancer treatment

Breast cancer treatment can be divided into surgery, systemic therapy, and radiotherapy. Surgery is often employed for treatment of early breast cancer, which may be breast conserving surgery [28] or mastectomy [29]. In some cases, this is preceded by neoadjuvant systemic therapy such as chemotherapy (more commonly used in TNBC) [30,31], anti-Her2 therapy or endocrine therapy [32], either for tumour downstaging to facilitate smaller resection, or with the aim of earlier systemic treatment [33]. Where appropriate, surgery is followed by adjuvant treatments, again comprising chemotherapy, anti-Her2 therapy, or endocrine therapy (in oestrogen receptor positive disease) [34]. Adjuvant radiotherapy, which aims to reduce the chance of recurrence, is largely routine following breast conserving surgery and is also given following mastectomy for larger or lymph node positive tumours [35,36].

For hormone receptor positive breast cancers (Luminal A/B), endocrine therapy is often suggested, which works by preventing oestrogen from getting to tumour cells where it would

promote unwanted tumour growth [32]. Oestrogen/progesterone production can be limited by drug suppression or through ovarian ablation [37,38]. Alternatively, selective oestrogen receptor modulators could also be used to competitively bind to oestrogen receptors in tumour cells [39], such as Herceptin, which is a blocker for HER2+ breast cancer [40,41]. TNBC however, since lacking in specific receptors, is hard to target with blockers and hence requires chemotherapy [42].

The UK National Institute for Health and Care Excellence recommends radiation therapy for most types of breast cancer [43] as it has been shown to significantly reduce recurrence of all breast cancer forms [44] (though suggested radiation doses and treatment regime can vary between patients). There is increasing focus on radiation therapy due its tissue sparing property and non-invasiveness. The impact of ionising radiation on breast tissues has been extensively studied with a particular emphasis on the efficacy of radiotherapy in eliminating cancer cells and reducing recurrence [3,35,45]. However normal tissues in the breast and surrounding organs (including heart and lungs) are also subjected to a significant dose of ionising radiation [4–6], which can induce side effects which impact on a patient's long-term quality of life. Short-term side effects (or acute side effects), such as radiation erythema [46] and radiation fatigue [47], typically appear hours or days after radiotherapy treatment. In contrast, late side effects, such as radiation fibrosis [9], typically surface after several months or even years [48]. Prolonging patient lifespan at the expense of quality of life is not ideal – there is a need to reduce or reverse radiation damage to normal tissues to mitigate these side effects. This requires further understanding of the underlying mechanism of radiation impact on tissues.

1.2 Ionising radiation (X-rays/gamma rays)

The most common form of ionising radiation used for radiation therapy arises from the electromagnetic (EM) spectrum, in which radiation exists as alternating electric and magnetic waves that propagate energy. Particle radiation is another form of ionising radiation consisting of accelerated particles such as electrons and protons. While equally important in cancer therapeutics, proton/ion beam radiation therapy is out of scope for this thesis and is better covered by this comprehensive review/book by Solov'yov [49]. Moving back to the EM spectrum, ionising radiation here is comprised of X-rays (0.01 nm < typical λ (wavelength) < 10 nm) and gamma rays (typical λ < 0.001 nm) (Figure 1.3), where x-rays are generated by

processes outside the nucleus, typically from electrons and gamma rays are produced from the atomic nucleus. Both are sufficiently energetic to induce electronic transitions of most biomolecules [50]. Such as transition involves an electron absorbing the radiation and transiting into a higher electronic state, becoming less bounded to the nucleus and therefore more reactive [51]. The electron can escape the coulomb attraction of the nucleus if adequate amount of energy is imparted by the radiation, resulting in ionisation of the molecule.

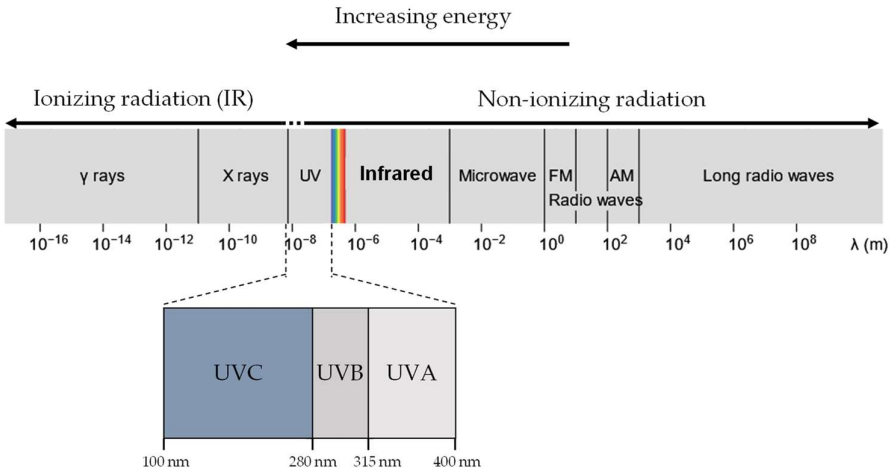


Figure 1.3. Ionising radiation (x-rays and gamma rays) lie on the high energy/low wavelength region of the electromagnetic spectrum. Non-ionising radiation includes ultraviolet radiation (UVR) (UV-A and UV-B) and subsequent lower energy EM radiation.

Human exposure to ionising radiation may arise from naturally occurring radon gases and cosmic radiation, which provide a constant background of ionising radiation that, on average, adds up to 2.4 mSv a year [52] (Sv, or sieverts, is a measure of effective radiation dose and represents the risk of radiation induced carcinogenesis [53]). On the other hand, man-made sources of ionising radiation, such as in mammography, would commonly only expose the patient to a dose of 0.36 mSv per screening [53,54]. Radiation therapy, however, lies in the range of 100-400 mSv [55] for whole body effective dose, or 40-60Gy (Gy, or gray, is a measure of radiation energy absorbed, i.e. J/kg) for the dose imparted to the cancer tissue [44,56,57], and the treatment is typically fractionated into 5-15 smaller doses over a few weeks [58].

The efficacy of radiotherapy lies in the ability of ionising radiation to penetrate biological tissues, allowing non-invasive targeting and killing of aberrant cells by causing irreparable DNA damage. Historically, radiotherapy utilised naturally occurring sources such as Co-60, which emits 1.2 MeV gamma rays. Modern external beam radiotherapy treatment regimens use linear accelerators (LINACs) to accelerate electrons towards a metal target to produce

ionising radiation [59], with exposures up to doses of 50Gy for breast cancer radiotherapy patients [60]. Other forms of radiotherapy include Brachytherapy, where a radioactive source is placed within the patient near the tumour (commonly prostate cancer) site [61]. Inadvertent exposure of healthy tissues along the irradiation path can lead to detrimental side effects, including radiation fibrosis [9] and secondary cancers [62]. While there are newer radiotherapy machines utilising proton or heavy ion beams to reduce exposures to healthy tissue by exploiting the Bragg peak [63], these treatment options are less widely available and are often reserved for paediatric patients [64]. X-ray/gamma ray radiotherapy remains the foremost therapeutic option, and hence, the impact of these radiation exposures on healthy tissues is a key biological and medical issue.

Photons of ionising radiation are energetic enough to ionise most molecules and atoms [50], potentially leading to the disruption of intermolecular bonds [65]. An abundance of water molecules in biological systems results in a large percentage of ionising radiation being absorbed by water in a process called water radiolysis [66], producing multiple ROS species (Figure 1.4). Water radiolysis induces the formation of not only hydrogen peroxide, and the superoxide anion [67], but also an abundance of highly reactive hydroxyl radicals [68].

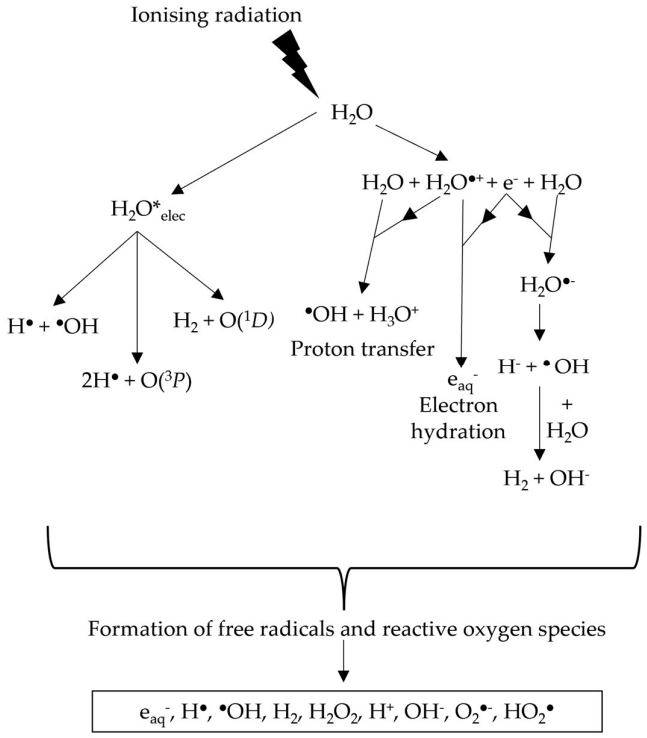


Figure 1.4. Ionising radiation produces a range of ROS and, more crucially, the hydroxyl radical through water radiolysis. This results in a larger concentration of hydroxyl radicals produced during ionising radiation irradiation compared to UVR due to the abundance of water molecules. Information from Figure (b) was sourced from Meesungnoen J. et al. [467].

1.3 Biological consequences of exposure to ionising radiation

The biological effects of ionising radiation can be crudely split into two categories: deterministic and stochastic [46]. Deterministic effects are often apparent only when tissues receive high doses of ionising radiation beyond a threshold level [69], while stochastic effects are probabilistic without any threshold dose [70].

1.3.1 Acute total body exposure to ionising radiation

A wide array of deterministic effects can be observed in humans exposed to acute doses of radiation to the whole body, such as in nuclear fallouts [71] or from the atomic bomb [72]. Whole body exposure had also been used clinically for suppressing the immune system, such as in preventing bone marrow rejection after transplant, or for low grade lymphoma [73].

In whole body exposure, lower doses between 1-7Gy primarily affects the hematopoietic system, with a significant decrease in population of blood progenitor stem cell and blood cell count [74]. At higher doses (>8Gy), radiosensitive organs such as the gastrointestinal and respiratory system are implicated [73,75,76].

1.3.2 Organ exposure to ionising radiation

In a controlled, clinical setting, ionising radiation can be exposed to organs and the effects on each of those organs can differ even at the same radiation dose. The radiosensitivity of organs can be measured based on the risk of carcinogenesis from exposure, and is highest in the bone marrow, stomach, lungs, colon, and breast [53].

Some organ specific examples are as follows:

For the gastrointestinal (GI) system, exposure to 8-14Gy of clinical radiation can result in acute GI syndrome, such as intestinal crypts formation, mucosal barrier breakdown, haemorrhage, vomiting or diarrhoea [77]. Skin, the organ exposed during most radiation therapy can suffer from radiation burns, erythema, and fibrosis [78,79]. Lungs exposed to clinical levels of ionising radiation can lead pneumonitis, inflammation and fibrosis of lung tissue, thus detrimentally affecting pulmonary volume [80,81]. For a more comprehensive review of the effects of ionising radiation on other organs which is outside of the scope of this thesis, the reader is directed to the following review by Wang and Tepper [82].

1.3.3 Cellular impact of ionising radiation

Outcomes of organs exposed to ionising radiation depends largely on how those cells respond to radiation damage and is thus influenced by the type of cell population and the microenvironment within the organs. However, the principal cellular damage arising from ionising radiation is largely indiscriminate. The damage cells receive from ionising radiation can be direct (radiation is absorbed by the biomolecule) or indirect (radiation is absorbed by water molecules, producing ROS that goes on to be absorbed by the biomolecule). Regardless, the primary targets of both are DNA and proteins, which disrupts cellular homeostasis.

1.3.3.1 DNA damage from ionising radiation

Most well studied cellular impact is damage to DNA in cells, which influences cellular behaviour, such as migration or protein translation and synthesis. Damage to DNA comes from induction of oxidation via deprotonation or electron removal, producing photolesions such as 8-oxoguanine [83]. Hydroxyl radicals produced from water radiolysis can also disrupt the bonds in the sugar backbone of DNA, resulting in single-strand breaks (SSBs) in the DNA [84,85]. As ionising radiation is highly energetic, electrons ejected from radical formation could potentially cause further radiolysis of nearby water molecules, resulting in a high density of hydroxyl radicals [68,86], increasing the probability of SSB occurring close enough in a DNA strand (within 10 base pairs) to promote the formation of double-stranded breaks (DSBs) [13,87]. DSBs are potentially highly cytotoxic due to the risk of failed repair, such as in non-homologous end joining (NHEJ) or homologous recombination, resulting in gene mutations [88,89], clastogenic effects [90], teratogenesis [91] and carcinogenesis [87].

1.3.3.2 Protein damage from ionising radiation

Ionising radiation-induced water radiolysis can also cause significant ROS-mediated damage to proteins through disruption of peptide bonds, protein oxidation [92], and lipid peroxidation [93], adversely altering protein structure and function [94–96]. The direct impact of ionising radiation on proteins can be observed during X-ray diffraction studies of protein crystals, where cryogenic temperatures reduce the effects of radicals produced by the solvent [97]. These studies demonstrate that disulphide bonds and carboxyl groups are most susceptible to localised radiation damage [98,99]. However, this damage may not be evenly distributed throughout the protein [100]. For example, Weik et al. (2000) have shown that the specific disulphide bond between Cys-254 and Cys-265 residues for *Torpedo californica*

acetylcholinesterase, as well as the disulphide bond between Cys-6 and Cys-127 for hen egg white lysozyme, are most susceptible to radiation damage. Radiation damage may also localise at active sites in proteins [99,101,102] such as for bacteriorhodopsin [103], DNA photolyase [104], malate dehydrogenases [105], and carbonic anhydrase [106]. This damage localisation has been hypothesised to be mediated either by the presence of metal ions, which have high proton numbers and hence more electrons for photo-absorption to propagate subsequent ionisation events [107], or by the relative accessibility of exposed active sites to ROS [99].

1.3.3.3 Cellular response to ionising radiation damage

In response to the damage due to photodynamically produced ROS, cells can initiate repair mechanisms, including nucleotide excision, to remove photolesions in the DNA [108]. Enzymes in the cell can repair reversibly oxidised proteins, such as cystine, which can be reduced back to cysteine by the thioredoxin reductase system [109], or may break down irreversibly oxidised proteins, typically products of hydroxylation and carbonylation processes [110,111]. In addition, ROS scavengers, such as superoxide dismutase, help restore the ROS balance in the intracellular and extracellular spaces by converting the superoxide anion to hydrogen peroxide [112,113], which is then converted to water and oxygen by catalase and glutathione peroxidase 3 to prevent the formation of hydroxyl radicals [114,115]. To repair DNA damage specific to ionising radiation, cells utilise base excision repair (BER) for oxidised nucleotides, such as 8-oxoguanine [116,117], while NHEJ and homologous recombination repair (HRR) is activated to remove DSBs [118–120].

1.3.4 Cellular response to ionising radiation damage in breast tissue

In breast tissue, mammary fibroblasts secrete and regulate ECM proteins in the stromal microenvironment, which provides structural support for mammary glands. Human mammary fibroblasts exposed to physiological doses of ionising radiation adopt a senescent-associated secretory phenotype (SASP), enhancing the secretion of ECM-degrading proteases promoting epithelial cell invasiveness and growth in 3D culture [121–123]. Key secreted proteases include MMP2 and MT1-MMP1, which drive not only ECM degradation but also epithelial cell migration in the basement membrane by exposing a cryptic site in laminin-332 for cell receptors to bind to [124,125]. Mammary epithelial cells exposed to ionising radiation were also found to become more susceptible to epithelial-mesenchymal transition when

exposed to TGF-β1 [126], which can lead to carcinogenesis and metastasis [127]. Ionising radiation has also been shown to activate latent TGF-β1 in the ECM [128], which binds to fibroblasts, triggering their differentiation into myofibroblasts [129].

Ionising radiation can also mediate the release of the growth factors due to ROS-mediated proteolytic cleavage of ECM components [130]. The basement membrane, a key ECM structure that provides architectural support and acts as a barrier to the mammary gland, can act as a source of matrikines and growth factors, such as the insulin-like growth factor (IGF) [131], which are often sequestered in the ECM. Paquette et al. have shown that reconstituted basement membrane containing these growth factors, when irradiated with ionising radiation, enhanced the invasiveness of breast cancer cells (MDA-MD-231) [130]. The release of other growth factors such as TGF-β1, which is commonly localised in the ECM [132], can also stimulate upregulation of MMPs (e.g., MMP-2, MT1-MMP) in fibroblasts or cancer cells to remodel the ECM [133].

1.4 The extracellular matrix as a target of radiation damage

Whilst the impact of radiation exposure on cells and cellular components is well characterised, ECM-radiation interactions and the downstream biological consequences are not well understood. Crucially, damage to ECM may mediate long-term radiation effects as a consequence of the long half-life and limited repair of key ECM components [134]. The synthesis of many ECM proteins is usually highest during development and diminishes over time [135,136]. Elastin, for example, may persist over the human lifetime [137], whilst dermal and cartilage collagens have half-lives of 15 and 95 years, respectively, [138–140]. The slow replacement of damaged ECM proteins would allow changes in, for example, mechanistic signals from ECM to persist, which can lead to long-term complications.

To improve our understanding of the repercussions of ionising and non-ionising radiation damage to the ECM, there is a need to investigate the impact of radiation on three key mechanisms through which ECM influences cells, namely molecular structural changes, mechanical changes, and biochemical changes.

1.4.1 Radiation and ECM Mechanical Properties

Ionising radiation is capable of inducing molecular changes in large ECM proteins, many of which govern the mechanical properties of tissue. These ECM proteins include collagens,

elastin, and laminins [141], which form interconnecting networks or fibrillar structures to allow cells to bind to. Changes to their tertiary and quaternary structures, such as by exposure to ionising radiation, could impact the function of these proteins, thereby altering the mechanical properties of the ECM [142,143]. In addition, structural damage to cell-adhesive proteins such as fibronectin could diminish cell–ECM interactions [18]. This implies that radiation exposure may compromise mechanosensing pathways. Altered tissue stiffness and elasticity may trigger different cellular responses including: initiating epithelial to mesenchymal transition in cancer cells [144,145], triggering senescence in fibroblasts [146,147], determining the fate of differentiating mesenchymal stem cells [148] and even enhancing replication of glioma cells [149]. Determining the mechanical effects of radiation exposure on complex extracellular matrices may provide a better picture of biological radiation response by helping to differentiate between the direct and indirect responses of cells to radiation.

1.4.2 Radiation and ECM Biochemistry

Radiation can also alter the biochemistry of the cellular environment by triggering the release of growth factors that are sequestered in the ECM. Paquette et al. [130] had shown that ionising radiation exposure (20Gy, Co-60) of Matrigel, which are made from reconstituted basement membranes, released pro-invasive growth factors that enhanced invasion of MDA-MB-231 cells. A plausible mechanism for the release of these factors could be attributed to radiation-induced structural changes to key ECM proteins, such as fibronectin [150–152], which binds to a variety of sequestered growth factors, including insulin-like growth factors (IGFs), fibroblast growth factors (FGFs), TGF-β1 and vascular endothelial growth factors (VEGFs) [131,153]. These factors serve as important signals to alter cell behaviour typically via integrins binding [154], MMP-mediated ECM degradation [155] or in wound healing [16]. Radiation damage to fibronectin and other similar ECM proteins may diminish their ability to bind to growth factors, thus increasing the availability of these factors [131] in the extracellular space. Abnormal levels of such growth factors would be taken up by cells, potentially triggering unwanted proliferation and migration due to FGFs [156] or ECM deposition due to TGF-β1 [157].

In addition to growth factors, radiation is also hypothesised to be able to introduce biologically active peptides in the extracellular environment through the fragmentation of

ECM proteins. These peptides, often referred to as 'matrikines', may be derived from abundant ECM proteins, such as collagen I and IV [158] or elastin [159], and are able to influence cellular behaviour just like growth factors. Whilst there is experimental evidence for the generation of matrikines by MMPs [160] there is a lack of evidence for the direct induction of matrikines by radiation. However, the ability of both non-ionising and ionising radiation to produce ROS that can fragment ECM proteins makes the possibility of radiation-produced matrikines (albeit with less specificity than MMPs), an interesting phenomenon to explore. In all, undertaking these biochemical studies may help explain certain non-local radiation effects, such as bystander effects, where local mechanical influences are not applicable.

1.5 Challenges of studying the ECM and the current state of knowledge

Studying the ECM is critical for furthering our understanding of radiation damage, but ECM proteins can be challenging to characterise due to their insolubility, necessitating the use of strong dissociative reagents which may affect protein structure [161,162]. Secondly, studying the ECM from tissues often requires decellularisation to prevent cellular influence, during which the ECM may be damaged and altered by chemicals used to remove the cells. Various models and experimental systems have been used in recent papers to study the ECM under UV and ionising radiation, but due to their limitations, these systems produce results which can be hard to interpret in relation to other literature. Summarised below are the four general categories of approach for studying the ECM under radiation exposure found in current literature, namely: (1) purified proteins; (2) decellularised cultures; (3) *ex vivo*; and (4) *in vivo*.

Purified protein experiments (Table 1.1), which is a bottom-up approach to target specific ECM proteins, is advantageous in its ability to isolate the effects of radiation on the protein of interest. For example, an increasingly common group of purified ECM "proteins" (better termed as peptides) are the collagen mimetic peptides (CMPs). These peptides contain multiple repeats of the tri-peptide sequence (Gly-X-X'), a key motif in collagen fibrils [65], in hopes of mimicking collagen properties while being more conducive for biophysical and biochemical methods, such as crystallographic studies [163–165]. Being able to control the specific composition of these CMPs also allows isolation of individual contribution of each type of amino acid in the overall stability of the collagen structure to radiation exposure. Using CMPs, Schwob et. al. (2017) were able to show preferential cleavage of the Gly-Pro peptide bond by absorption of near x-ray energies (slightly above UVC) [96]. Other studies have also

purified whole collagen, such as from rat tendons, before exposing them to radiation. This was done in a study by Miller et al. (2018) [142] where collagen from rat tendons were reconstituted into 3D collagen gels and were able to show reduction in stiffness of these gels with 10-63Gy of γ-ray exposure, thus shedding light on possible mechanical alterations of collagen in tissue after irradiation. Purified proteins are thus useful for investigating the chemical and structural changes of individual proteins that might occur during irradiation. However, the results are difficult to extrapolate to the protein in actual tissues and the potential downstream effects *in vivo* as the experimental system is not representative of the protein's natural environment or state in the ECM. Nonetheless, such a reductionist approach can provide a fundamental picture of the molecular mechanisms that occur for the proteins in question during radiation exposure and is a first step to predicting or understanding complex phenomena in tissues.

Table 1.1. Selected studies utilising purified ECM proteins for radiation damage experiments are useful to elucidate molecular mechanisms of radiation responses for individual ECM proteins. Most studies show that purified collagens in solution exposed to ionising radiation induces significant structural change and peptide bond cleavages albeit at non-physiological doses.

Dose	Method	Ref.	Results
Co-60 γ-ray at 1.289kGy/hr, 5kGy–50kGy	Lyophilised collagen from rat tail tendon irradiated and tested for solubility and melting temperature	[166]	Irradiated samples were, in general, more than twice as soluble as non-irradiated in 0.02 M acetic acid, 6 M lithium chloride and 6 M urea. Melting temperature reduces with increasing dose.
γ-ray (1 MeV), 50kGy–500kGy	Grounded collagen irradiated in dry/wet (5%/80% moisture) state in the presence and absence of oxygen/nitrogen	[167]	Solubility unchanged when irradiated wet due to cross-linking, and solubility increased when irradiated dry. Significant molecular changes likely due to the breakage of peptide bonds. Degradation of Tyr; Hyp/Pro; Asp sensitive to oxygen/nitrogen.
Near X-ray (13.8–22.1 eV)	Isolated collagen mimetic peptides, photon absorption in gas phase + mass spectrometry	[96]	Gly-Pro peptide bonds are more susceptible to cleavage, collagen triple helix stabilised by hydroxyproline.

Decellularised cultures (Table 1.2) involve taking tissue samples from living organisms and removing the cells from the tissue, leaving behind the ECM scaffold and proteins for cell culture applications [168,169]. Utilising ECM from tissue provides the advantage of good *in vivo* representation, with ECM structures, growth factors and binding ligands largely intact for radiation studies. There is evidence of X-rays, in the kGy range, altering the stiffness of these ECM scaffolds [170,171]. Behaviour of cells seeded onto X-ray irradiated ECM scaffold was also altered with increased proliferation [168] or poor adhesion [172]. These studies show X-rays' ability to induce changes in mechanical properties of biologically relevant ECM scaffolds and that cellular behaviour alteration may in part be attributed to an indirect contribution from ECM damaged by radiation exposure. While decellularised cultures enables us to explore the interplay between different components of the ECM, as well as post-irradiation cellular remodelling of the ECM, the abundance of ECM components in the culture can also be detrimental when attempting to identify the cause of downstream effects. Furthermore, the variability of ECM proteins in different organisms, or even in different regions of the same organism, can make experiments difficult to replicate. Lastly, the decellularisation process can also alter ECM protein's ultrastructure during the chemical removal of cells or during the sterilisation process [171], making it difficult to control for unwanted changes.

Table 1.2. Decellularised tissues exemplify a highly representative ECMs that can mimic key aspects of in vivo responses to radiation damage. Studies using these systems show X-rays' ability to induce changes in mechanical properties of ECM scaffolds, and that X-ray exposure can affect subsequent cell responses to the ECM.

Dose	**Method**	**Ref.**	**Results**
γ-ray (wavelength unspecified, 5000Gy)	Rabbit kidney decellularised	[170]	Reduced tensile strength and young's modulus with gamma ray.
Co-60 γ-ray, 25kGy	Gamma irradiation of decellularised cornea	[171]	Increased stiffness/tensile strength, reduced elongation at break after irradiation, due to fragmented collagen cross-linking.
Cs-137 γ-ray, 1k–10kGy	Decellularised whole porcine kidney	[172]	3kGy resulted in more than 50% loss in collagen content. Human renal cortical tubular epithelium (RCTE) cells reseeded and resulted in poor adhesion/growth.
Cs-137 γ-ray, 20Gy	Murine mammary fat pads decellularised and made into hydrogels.	[168]	Increased proliferation for murine TNBC reseeded on irradiated hydrogel.

Ex vivo systems (Table 1.3) refer to tissues that are extracted from organisms and experimented on (usually in less than 24 hours) with minimum alteration. A key advantage of *ex vivo* experiments is their ability to provide clinically relevant insights into complex tissues, making them useful for determining end point consequences of radiation effects. Radiation studies on *ex vivo* tissues have shown changes in mechanical strength, even at radiotherapeutic doses of X-rays [142,173,174], although they were not consistent between different tissue types. These systems are often complicated to analyse as they contain both ECM and cells that can influence the remodelling of the ECM after irradiation (Table 1.4). Furthermore, the results are hard to generalise as the tissues used are made of specific cells and ECM environments, which may only be applicable to that organism.

Table 1.3. Ex vivo experiments utilise complex model systems that give biologically relevant consequences of radiation effects. Studies show that radiotherapeutic doses of X-rays (around 50Gy) can alter the mechanical properties of ex-vivo samples.

Dose	Method	Ref.	Results
Cs-137 γ-ray, 10–63Gy	Mammary tumours (MMTV-PyMT transgenic mice) immediately irradiated and frozen before tested for compression	[142]	Significantly reduced tensile and compression modulus after 60Gy irradiation (fractionated and single dose).
6–10 MeV clinical (LINAC) X-rays, 30-56Gy	Biopsy from radiation therapy treated breast cancer patients. Irradiated/non irradiated samples from the same patient 10-96 months after treatment	[173]	No observable change in elastic fibres/collagen, but stiffness is higher for irradiated regions.
21 KeV synchrotron x-rays, 50–35000Gy	Lumbar vertebrae excised and removed of soft tissue. Wrapped in saline-soaked gauze	[175]	Monotonic strength (one direction) decreased at 17000Gy and above. Increase in non-enzymatic cross-links at a lower dose (50–1000Gy) by analysing AGEs. Crosslinks do not have a significant impact on vertebral strength.
6 MeV clinical (LINAC) X-rays, 10–100Gy	Bovine pericardial tissue (collagen), Bovine ligamentum nuchae (elastin)	[174]	For pericardial tissue, elastic modulus increased for small strain and decreased at larger strain after irradiation. Elastin has significantly reduced elastic modulus after irradiation.

For *in vivo* systems (Table 1.4), mouse models are often used to observe longer-term tissue responses to radiation. Such experiments are useful to account for the various effectors of radiation response, allowing full interaction between different mediators and are helpful for observing long-term effects such as secondary malignancies, fibrosis, and metastasis. In addition, such models can elucidate if irradiated regions may exert a local or systemic influence on, for example, the immune responses [176]. *In vivo* studies show that irradiated animals experience ECM remodelling, which is likely to be mediated by MMPs. However, therein lies the challenge of relating animal models back to humans, as genetic differences could invalidate the radiation response elucidated in these models [176]. In addition, it is difficult to determine if the radiation outcomes are associated with the acute effects of radiation on ECM proteins or with long-term remodelling from cellular expression of MMPs.

Table 1.4. In vivo models allow for observations of long-term radiation responses not only in the targeted area but also surrounding tissues or organs for bystander studies. In vivo studies showed that both UVR and ionising radiation-exposed animals experience ECM remodelling as a consequence of protease action.

Dose	Method	Ref	Results
Co-60 γ-ray, 2–22Gy in fractions of 2Gy/day	White, outbred rats, irradiated in bladder and rectum. For 2Gy, rats were harvested 1 day/1 week/1month after irradiation. Higher doses harvested after 1 day.	[177]	One-month post-2Gy irradiation showed thickening of collagen fibres and tight, parallel packing for the bladder and rectum. One day post-irradiation for higher dose observed the same effects with the severity dependent on dose. Skin most sensitive showing similar damage at 8Gy.
300kVp X-rays (30–60Gy) for local, Cs-137 γ-ray (6–10Gy) for whole body	C57BL/6 mice with smad3 gene knockout	[178]	Smad3 knockout mice have less TGF-β1 expression, less inflammation, less myofibroblasts after radiation
Co-60 γ-ray, 2–40Gy, 1.7Gy/min	2-month-old, white wild type outbred rats, ~ 200g, harvested 1 day/1 week/1,2,3 months after irradiation for rat's tail tendon	[179]	Differential scanning calorimetry showed a dose-dependent increase in denaturing temperature 24 h after irradiation, but dose-independent after 1 week. Negligible change was observed for tertiary/secondary structures using second harmonic generation/cross-polarisation optical coherence tomography

While studies conducted across the four types of model systems were useful in elucidating the outcomes of ECM proteins to radiation exposure *in vivo*, there remains a lack of consensus and thus, a gap in understanding the underlying mechanisms, which this thesis aims to address.

1.6 Hypotheses and aims

Clinical outcomes of exposure of breast tissue to ionising radiation, such as the fibrosis or the hardening of breast tissues, have been well documented and detrimentally impact patients' quality of life. However, this is not limited to breast tissues (which is the focus of this study); it is important to acknowledge that x-ray exposure can affect other organs, resulting in adverse clinical consequences. For example, in breast radiotherapy, excessive ECM accretion may occur in the lung leading to fibrosis [180]; In prostate radiotherapy, ECM degradation often precedes radiation proctitis [181]; For lung radiotherapy, pneumonitis often develops with aberrant ECM deposition [81]; In glioblastoma radiotherapy, the ECM is found to increase the invasiveness of glioblastoma cells, possibly contributing towards the high relapse of glioma patients after radiation therapy [182]. To mitigate these undesirable complications post-radiotherapy, there is a significant necessity to explore the underlying mechanisms that elicit such responses, and the role ECM may play in those mechanisms following therapeutic radiation exposure.

Although the impact of X-ray exposure on breast cells is well characterised, tissue function and physiology are determined by a complex interplay between both the cellular and the extracellular (ECM) components. However, to date, the impact of ionising radiation on ECM proteins remains poorly defined. The current literature (as seen in section 1.5) lacks consensus in the impact of radiation on ECM, partly due to a wide range of radiation doses and energies used. Furthermore, each study was often confined to addressing a single type of experimental system, thus lacking the continuity required for a complete picture of the impact of radiation on ECM proteins at multiple scales. To address this gap in knowledge, it is crucial to characterise the interaction of clinical X-ray doses with appropriate biological samples using methodologies which are sufficiently sensitive to detect low, clinically relevant radiation dose effects across multiple experimental systems.

This thesis, hence, aimed to address these gaps in order to improve our molecular understanding of ionising radiation damage to ECM proteins (both purified suspensions and complex tissue-derived ECM mixtures) while using consistent, clinically relevant doses and energy of X-rays (50-100Gy, 300kVp). X-ray dose of 50Gy was chosen as it represents the typical total dose to the tumor bed for radiation therapy of many cancer patients [58,183,184], while 100Gy was chosen to investigate possible dose dependent effects.

The impact of X-ray exposure was characterised using complimentary biophysical and biochemical characterisation including the newly developed approach of peptide location fingerprinting (PLF), which has been previously been shown to detect small changes to structure and enzyme susceptibility in large ECM molecules, such as fibrillin-1 and collagen VI [185]. Furthermore, PLF can be applied to proteomic analysis of both purified ECM suspensions and complex tissues due to the flexibility of mass spectrometry.

Using the principles mentioned, this thesis aims to address three hypotheses:

1. Therapeutic X-ray doses differentially affect the structure of proteins dependent on their amino acid composition.

Collagen I is the most abundant protein in the human body and a key component of breast stroma, while fibronectin is a large, ubiquitous glycoprotein which performs key roles in cellular binding and signalling. In contrast to collagen I, fibronectin is rich in oxidation sensitive amino acid residues and hence would be expected to suffer more radiation-induced damage due to the radiolytic production of ROS. Using a consistent panel of biochemical and biophysical techniques the impact of X-ray exposure on the structure of purified collagen I and fibronectin is compared.

2. Exposure to therapeutic X-rays doses affects collagen I structure in an assembly-dependent manner.

In common with other ECM components (principally of the elastic fibre system), collagen I is present in tissues as large multi-component macro-molecular assemblies. Here the aim is to determine if native, macro-molecular assemblies of ECM proteins, specifically collagen I (fibrils), are affected by x-ray exposure and if so, whether assembly into fibrils affects relative susceptibility. While other ECM proteins, such as elastin, are equally important in biomechanical function as collagen I, they are much harder to isolate and purify, and their structure is less distinct (unlike the triple helical structure of collagens). Thus collagen I has been chosen as the focus of this study. It is suspected that the accessibility of water molecules many cause localised differences in ROS concentrations in the suspension of collagen monomers compared with the interior of densely packed collagen fibrils, which may influence the outcome to therapeutic x-ray exposure.

3. Mass spectrometry data can identify key targets of X-rays damage in complex proteomes.

The final aim is to compare the therapeutic radiation response of ECM proteins derived from *in vitro* fibroblast cultures with *ex vivo* breast tissue samples. In addition to investigating collagen I and fibronectin, a global search (or shotgun proteomics) was conducted using mass spectrometry to screen for damaged ECM proteins flagged by PLF and compared between the two sample types. Identifying these targets is a key step in understanding the impact of low-dose ionising radiation on tissue biology and hence breast radiotherapy associated pathologies.

2 Materials and methods

2.1 Purified proteins, cell lines, and breast tissue sources

Purified human collagen I (from human placenta) were sourced from Sigma-Aldrich (UK) (catalogue number: CC050); purified plasma fibronectin (from blood plasma) was purchased from Abcam (Cambridge, UK) (catalogue number: ab80021). Solubilised rat tail collagen I was purchased from Gibco (UK) (catalogue number: A1048301).

Immortalised human mammary fibroblast (HMFU-19) for *in vitro* ECM production were kindly gifted by Alis Hales from the Gilmore lab at the University of Manchester. HMFU-19 are normal primary human mammary fibroblasts immortalised by human telomerase reverse transcriptase (hTERT) and a temperature-sensitive mutant of the simian virus 40 large-tumour antigen (U19tsA58) [186].

Rat tails tendons were dissected from young, male Sprague Dawley rats. The care and use of all rats in this study was carried out in accordance with the UK Home Office regulations, UK Animals (Scientific Procedures) Act of 1986 under the Home Office Licence (#70/8858 or I045CA465). Rats were culled by CO_2 inhalation and tails were immediately dissected and stored at -80°C wrapped in phosphate buffer (pH 7.4) saturated gauze.

Ex vivo frozen normal breast tissue samples were obtained from biobank (Project number: 22_MISH_02, Sample serial number: H000884T1Na, patient age at surgery: 64, patient diagnosis: Grade II Infiltrating Ductal Carcinoma) under the Manchester Cancer Research Centre (MCRC) general ethics approval.

A summary table of all materials used, and the sample size is included in the table below (Table 2.1).

Table 2.1. Table of summary for all samples and materials used in the study and the sample size.

Sample	Details	Methods	Sample size (n)
Purified human collagen I	Sigma-Aldrich (UK), CC050	SDS-PAGE Native-PAGE CD PLF	3 1 1 5
Purified human plasma fibronectin	Abcam (Cambridge, UK), ab80021	SDS-PAGE Native-PAGE Light scattering PLF	3 2 3 5
Rat tail collagen I (Solubilised/Gel)	Gibco (UK), A1048301	SDS-PAGE Native-PAGE CD PLF	3 3 1 5
Rat tail tendon	Dissected from young, male Sprague Dawley rats	SDS-PAGE Native-PAGE CD PLF	3 3 1 5
Human mammary fibroblast ECM	HMFU-19, passage 20-22	PLF	5
Human breast biopsy	Frozen normal tissue (at least 3cm from tumor site) Biobank project: 22_MISH_02 Sample number: H000884T1Na Patient age: 64 Diagnosis: Grade II Infiltrating Ductal Carcinoma	PLF	1

2.2 Reconstituted rat tail collagen gels

All solutions were kept on ice during the procedure (unless otherwise stated) to prevent premature fibrillogenesis of collagen solutions. Protocol was adapted from manufacturer's instructions (MAN0007327). Firstly, solubilised rat tail collagen (Gibco, UK) at 3mg/mL was diluted to 1mg/mL with PBS (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4) (Sigma-Aldrich, UK). 37.5µL of 1M NaOH was added to the mix per 1mL of 3mg/mL solubilised rat tail collagen used and mixed thoroughly by pipetting slowly without introducing air bubbles. pH was verified to be neutral (pH 7) by pipetting 40µL of mix onto a Whatman pH indicator paper strip (Cytiva, Amersham, UK) and allowing it to equilibrate for 1 minute. pH was adjusted with additional 1M NaOH if required. Neutralised collagen solution was distributed into 0.5mL Lo-bind Eppendorf tubes and placed in a humidified incubator for 1 hour at 37°C to speed up fibrillogenesis. Following incubation, the collagen gel was visually inspected for increased turbidity for successful gel formation, before storing at 4°C for up to a week.

2.3 Sample preparation prior to x-ray exposure

2.3.1 Complex ECM production and decellularisation with *in vitro* immortalised human mammary fibroblast (HMFU)

Immortalised human mammary fibroblast (HMFU-19) [186] were cultured at passage 20-22 and grown in RPMI-1460 media, with 1% Glutamax and 10% FBS. Cells were cultured in low wall, 50mm µ-Dish (Ibidi, Gräfelfing, Germany) and in 6 well plates. Cells were seeded at 10^5 per dish/well and grown for 2-3 days until confluent. Media was changed every 2 days and cells were allowed to deposit matrix until 9 days post-confluence. Cells were subsequently removed by incubating with PBS- (137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4) (Sigma-Aldrich, UK) containing 0.5mL of 20mM NH_4OH, 0.005% Triton-X 100 until no cells were visible. Additional 1mL of PBS- was added to each well and allowed to incubate overnight at 4°C. Plates/wells were carefully washed the next day with PBS- twice and PBS+ (PBS with 0.9 mM $CaCl_2$ and 0.5mM $MgCl_2$) twice to remove detergent and remaining cells. The leftover matrix was kept in PBS+ supplemented with 1% Penicillin/Streptomycin (P/S) at 4°C (for up to 5 days) until needed.

2.3.2 Rat tail tendon extraction

Frozen rat tails dissected from young, male Sprague Dawley rats were left to defrost at room temperature for 10 minutes, following which the skin was removed by making a vertical cut with a scalpel along the tail and carefully peeling it off. Two horizontal cuts, approximately 2cm apart, were made near the base of the tail to expose tendon fascicles. Fascicles were gently extracted using forceps while minimising extension and were immediately washed in a tank of phosphate buffer (PBS) before transferring to a clean Eppendorf tube with PBS containing protease inhibitor cocktail (Promega, Hampshire, UK) and stored at 4°C.

2.3.3 *Ex vivo* breast tissue

Ex vivo breast tissue samples were cryosectioned prior to x-ray irradiation. Firstly, frozen tissue was embedded in Tissue-tek optimal cutting temperature (OCT) compound (Sakura Finetek, Torrance, CA, USA) and left in the cryostat to stabilise at -33°C for at least 20 minutes. Subsequently, the tissue was sectioned to a nominal thickness of 20µm. 20µm was determined to contain sufficient tissue sample for mass spectrometry while retaining similar tissue composition (by H&E staining) across five sequential sections (Appendix 7.3). These five sequential sections were collected on individual Epredia Superfrost™ microscope slide (ThermoFisher, UK) to form a group. Each slide contained three cryosections from different groups but of the same sequential position in their respective group (Figure 2.1). Slides were kept at -20°C and transported on dry ice until x-ray exposure.

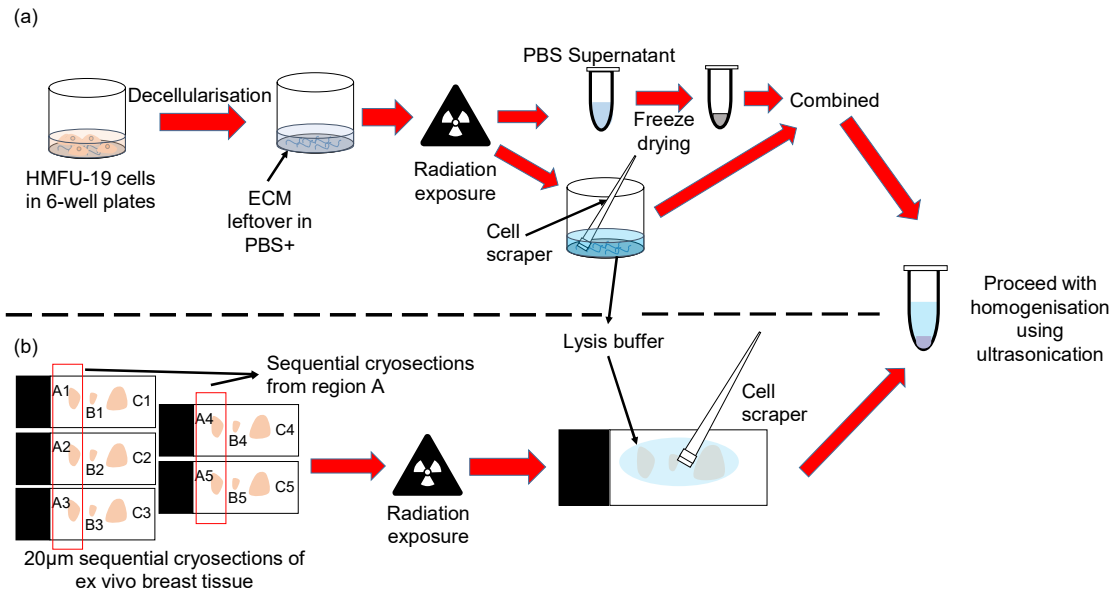


Figure 2.1. Post-irradiation sample extraction and homogenisation steps for in vitro HMFU-19 cell culture and ex vivo breast tissues from biobank. (a) For HMFU cell culture, cells were first removed by addition of extraction buffer containing Triton X-100 and NH_4OH. The leftover ECM in PBS was exposed to radiation, following which the supernatant was collected and freeze-dried while the ECM was scraped off from the wells into SDS-containing lysis buffer. The lyophilised supernatant was resuspended in the same lysis buffer and combined with the ECM fraction before being homogenised with ultrasonication.

(b) Ex vivo breast tissue samples were first cryosectioned into sequential sections of 20µm thickness (E.g. In figure, A1 to A5 are sequential sections from a region of tissue. B1 to B5 are sequential sections from a different region of the same tissue.) Five sequential sections were obtained for each region of the tissue, to a total of 9 regions, with 3 regions combined into a slide. This gave a total of 3 sets of 5 = 15 slides. Within each set of slides, the first and last slides were kept for H&E staining, while the 3 remaining slides were randomly chosen to be exposed to 0, 50, or 100Gy of radiation. After radiation exposure, OCT was removed by a quick ethanol wash, and subsequently the sections were scrapped off the slides with a cell scrapper and addition of lysis buffer. The detached sections were then ultrasonicated.

2.4 Ionising (X-ray) irradiation

Samples were irradiated on the C1X3 Self-contained cabinet irradiator (Xstrahl, Camberley, UK), fitted with a 0.7mm copper filter and placed on top of a rotating Perspex platform. The dose rate was previously monitored at a setting of 300kVp tube voltage, 10mA tube current and was calculated to give a dose rate of 2.2Gy/min to a water body (calibrated yearly) with a half value layer of 2.3mm aluminium. Samples were irradiated to 50Gy or 100Gy in 22min 16s and 45min 27s respectively at room temperature. All samples were placed at room temperature throughout the irradiation process (including control samples) until the last sample has completed to minimise possible variation between samples due to sample degradation.

Purified protein, rat tendons, and rat tail collagen reconstituted gels were irradiated while in their respective Eppendorf tubes (thickness of plastic of Eppendorf tube estimated to have negligible dose effect of about 5% reduction). *In vitro* HMFU-19 derived ECM was irradiated

while in 6-well plates with lids removed. *Ex vivo* cryosectioned tissue samples were irradiated while on the glass slides.

Fractionation of doses was not employed to investigate the direct impact of radiation on ECM proteins without the influence of cellular remodelling. Without cells, ECM components are unlikely to respond differently to radiation doses given in multiple fractions or in a single dose.

2.5 Sample extraction and homogenisation after x-ray exposure

After x-ray irradiation, all samples (except solubilised proteins) were extracted and homogenised for subsequent biochemical analysis as follows.

2.5.1 Collagen gel

An equal volume of 20mM acetic acid to the volume of gel was added to the Eppendorf tube and gently vortexed to lift the gel off the bottom of the tube. The mix was then left at 4°C overnight. The solution was then mixed thoroughly by pipetting to ensure no remaining gel was left.

2.5.2 Rat tendon fascicles

Extraction of collagen from rat tail tendon fascicles were based on protocol in previously published work by Holmes et. al. (2010) [187]. Rat tail tendon fascicles were diced and snap frozen in liquid N_2. Next, they were placed under a cooled metallic cylinder (also cooled with liquid N_2) and crushed rapidly with a hammer. The crushed tendons were suspended with a dounce homogenizer in 50mM Tris-HCl, 50 mM Ethylenediaminetetraacetic acid (EDTA), 150 mM NaCl, 100 mM sucrose (pH 7.4), supplemented with protease inhibitor cocktail (Promega, Hampshire, UK). Samples were homogenised with ultrasonication using Covaris LE220+ (Covaris LLC, Woburn, MA, USA) at 500W peak power, 20% duty factor (100W average) for 5 minutes and concentration was subsequently determined by Direct Detect™ (Merck Millipore, Darmstadt, Germany) assay before storage at -20°C.

2.5.3 HMFU-19 derived matrix

Previously prepared decellularised fibroblast (HMFU-19) derived ECM, post-exposure to ionising radiation, was processed for LC-MS/MS analysis as follows (Figure 2.1). The supernatant (PBS+ containing P/S) were collected into tubes and 400µL of lysis buffer is added to the decellularised ECM in 6-well plates. Following which, the ECM was removed with the buffer and collected into an Eppendorf tube by using a mini cell scraper (VWR, Radnor, PA,

USA). Supernatant was frozen down in -20°C freezer and lyophilized in a freeze dryer overnight. Subsequently, 100µL of lysis buffer was added to the lyophilised supernatant and combined with the ECM samples to a total volume of 500uL. The samples were then homogenised with ultrasonication using Covaris LE220+ (Covaris LLC, Woburn, MA, USA) at 500W peak power, 20% duty factor (100W average) for 5 minutes. Protein concentration was subsequently determined by Direct Detect™ (Merck Millipore, Darmstadt, Germany) assay before storage at -20°C.

2.5.4 *Ex vivo* breast tissue samples

After x-ray irradiation, the breast tissue cryosections were first processed to remove the OCT compound as follows. Slides were dipped in 70% EtOH for 30s, 100% EtOH for 30s, followed by ten quick dips in two different containers of deionised H_2O (five dips each). The slides were then given the same EtOH treatment (70% EtOH 30s, 100% EtOH 30s) and left to dry for 30 minutes at room temperature.

Each cryosection was re-hydrated with approximately 200µL of lysis buffer (5% SDS, 50 mM triethylammonium bicarbonate (TEAB), pH 7.5), and carefully loosened from the slides using a mini cell scraper (Figure 2.1). The tissue and the buffer on the slides were then transferred over to a microTUBE-500 (Covaris LLC, Woburn, MA, USA) by careful pipetting. Additional lysis buffer was added to the slides to collect as much cellular material leftover as possible and subsequently combined into the microTUBE-500. Total volume was topped up with lysis buffer to 500µL. Ultrasonication using the Covaris LE220+ (Covaris LLC, Woburn, MA, USA) at 500W peak power, 20% duty factor (100W average) for 5 minutes. Protein concentration was subsequently determined by Direct Detect™ (Merck Millipore, Darmstadt, Germany) assay, before storage at -20°C.

2.6 Biochemical techniques (Gel electrophoresis)

2.6.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was conducted in XCell surelock mini-cell electrophoresis system (ThermoFisher, UK) with 4-12% Bis-tris NuPAGE gels (ThermoFisher, UK), in 3-morpholinopropane-1-sulfonic acid (MOPS) SDS running buffer (ThermoFisher, UK) at 200V for 60 min according to manufacturer's instructions (unless otherwise stated). Gels were stained with ReadyBlue

Protein Gel stain (Sigma-Aldrich, UK) for coomassie stains, or Pierce silver staining for Mass spectrometry kit (ThermoFisher, UK) for silver stain.

2.6.2 Acidic native-PAGE

Acidic native-PAGE (AN-PAGE) protocol was adapted from the methodologies published by Ornstein [188] and Lebendiker [189]. Gels were cast at 5%/7% using formulations in Table 2.2. The running buffer was 35mM β-alanine, 14mM acetic acid at pH 4.3. Gels were run at 100V for 8 hours and subsequently 200V for 16 hours at 4°C.

Table 2.2. Composition of separating and stacking gel for acidic-Native PAGE adapted from Ornstein and Lebendiker.

Acidic-Native PAGE gel recipe

	Separating gel (7% acrylamide)	**Stacking gel**
30% Bis/Acrylamide	6.4mL	1.0mL
50% Glycerol	6.0mL	
dH20	6.6mL	6.4mL
0.25M Acetic Acid – KOH, pH 4.3		2.5mL
1.5M Acetic Acid – KOH, pH 4.3	6.7mL	
10% APS	320µL	100µL
TEMED	40µL	10µL

2.7 Atomic force microscopy (AFM)

2.7.1 Atomic force microscope mechanical testing

Mechanical testing on the AFM was conducted using the Peakforce Quantitaive Nanomechanics (QNM) mode in fluid on the BioScope Catalyst AFM (Bruker, Santa Barbara, CA, USA). Briefly, Peakforce QNM is a recently developed AFM mode designed to minimise sample and probe damage by limiting the force experienced by the probe while scanning and mapping the sample [190]. A spherical probe of 5µm diameter with cantilever spring constant of 0.14-0.16N/m was used (CONT-Silicon-SPM-Sensor with colloidal particle, sQUBE, Bickenbach, Germany). Prior to the start of each experiment, the probe deflection sensitivity and spring constant were calibrated. The deflection sensitivity was set by ramping/indenting onto an empty glass slide at least three times and calibrating on the approach curve. Once the deflection sensitivity was set, the probe was withdrawn from the surface and thermal

tuning was conducted to obtain the spring constant of the probe [191]. Following calibration, indentation of the sample was conducted by indenting the probe to a depth of 50nm into the sample at a rate of 6µm/s to obtain the force curve (Figure 2.2).

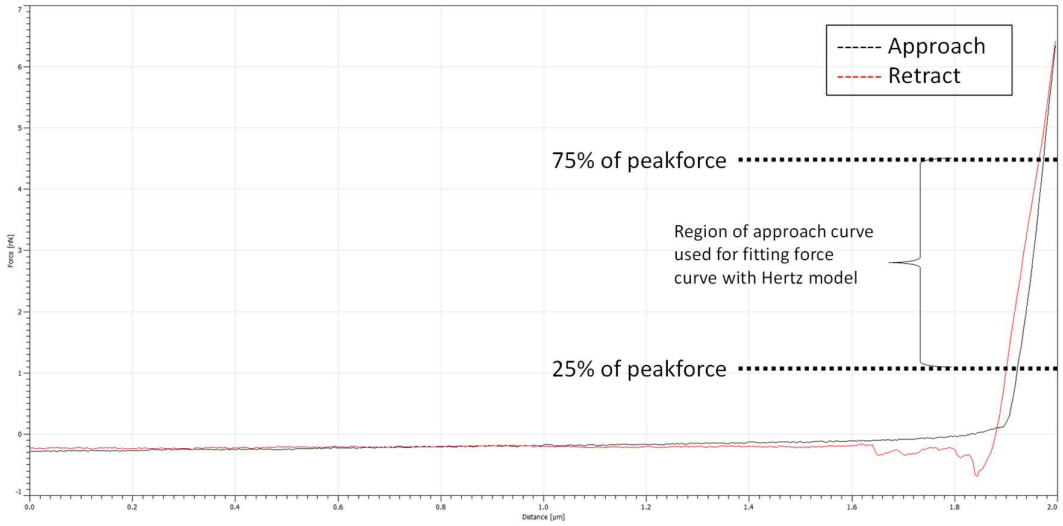


Figure 2.2. An example force curve obtained using the AFM Peakforce QNM mode from indentation of a collagen gel with a spherical probe. The force curve was further analysed by fitting the approach curve in the region between 75% and 25% of the peak force with the Hertzian model to obtain the reduced modulus.

Approximately 225 force curves were obtained in a region of 15µm x 15µm spaced 1µm apart, with a total of three regions tested per sample. For each force curve, the approach curve was used to estimate the reduced modulus of the sample based on the Hertzian model (spherical) by fitting the curve within the region bounded between 75% and 25% of the peak force. The hertzian model assumes the force curve is derived from the collision of two perfectly elastic spheres, which, in the case for the AFM, the material is assumed to be a sphere with infinite radius [192]. The fitting equation used (according to the manufacturer's manual) for the Hertzian spherical model is:

$$F^{\frac{2}{3}} = \delta \left(\frac{4}{3} E^* \sqrt{R}\right)^{\frac{2}{3}}$$

Where F is the force from the force curve, R is the radius of the spherical tip, δ is the indentation depth, and E^* is the reduced modulus to be obtained. The reduced modulus, which accounts for the stiffness of both the cantilever and the indentation material, can be related to the Young's modulus (E) by the following equation:

$$E^* = \frac{E}{1 - v^2}$$

Where v is the poisson's ratio, which is typically estimated to be around 0.2-0.5. For this study, the results are kept as reduced modulus. The data collected was then filtered to exclude data with an r^2 value below 0.9, and subsequently for values that lie beyond two standard deviations of the original raw dataset.

2.7.2 Atomic force microscope imaging

AFM imaging was conducted using the ScanAsyst Air mode on the BioScope Catalyst AFM (Bruker, Santa Barbara, CA, USA). The ScanAsyst-Air probe (Bruker, Santa Barbara, CA, USA) with cantilever spring constant of 0.4-0.8N/m was used. 5µm to 10µm images were taken at 512 x 512 or 1024 x 1024 resolution respectively at a rate of 0.5Hz per line. The peak force error images were processed using Gwyddion 2.63 [193] to obtain their 2D-fourier transformed images.

2.8 Solid phase enzyme-linked immunosorbent assay for collagen-fibronectin binding

This assay was conducted to test for protein-protein binding between purified human collagen I and human plasma fibronectin and generally follows previously published protocol by Biesiadecki [194] with modifications to blocking and antibody solutions. In brief, human collagen I (ab7533, Abcam, Cambridge, UK) was diluted to 2.5µg/mL using PBS- (Sigma-Aldrich, UK) and added to 96-well plates at 50uL/well for overnight coating at 4°C. Plates were washed three times (0 min, 2 min, 3min incubations) with 25mM Tris, 2.7mM KCl, 137mM NaCl, 0.05% Tween-20, pH 7.4 (Prepared from 20X TBS (ThermoFisher, UK) and Tween-20 (Sigma-Aldrich, UK)) (TBST) then blocked with 200uL/well of 5% skim milk (SERVA Electrophoresis GmbH, Germany) in TBST at room temperature (RT) for 1.5 hours. Following the same wash procedure, plasma FN (ab80021, Abcam), which was diluted to a range of 20 - 0.05µg/mL in 20mM HEPES, 150mM NaCl, was added at 50uL/well and incubated overnight at 4°C. Subsequently, the fibronectin solution was removed from the plates by washing with TBST, before addition of FN mouse monoclonal 1° antibody (66042-1-Ig, Proteintech Europe, UK) at 1:2000 with 100uL/well for overnight incubation at 4°C. The primary antibody was washed out with a stricter wash procedure (5X wash: 0 min, 2 min, 2 min, 3 min, 3 min incubations), before addition of the 2° antibody (HRP-conjugated Affinipure Goat anti-rabbit IgG(H+L), SA00001-2, Proteintech Europe, UK) at 1:2000, 100uL/well, for 1.5 hours incubation at RT. Finally, following a strict 5X wash with TBST, 100µL of 3,3',5,5'-tetramethylbenzidine

(TMB) substrate (ab171522, Abcam, Cambridge, UK) is added per well and incubated for 40 minutes. The reaction was stopped by adding equal volume (100uL) of 1M HCl to the wells. The absorbance was recorded at 450 nm with Multiskan FC (ThermoFisher, UK).

2.9 Protein structure and thermal stability characterisation

2.9.1 Differential scanning fluorimetry and light scattering

Protein melting temperature (T_m) and aggregation temperature (T_{agg}) of purified proteins were determined using differential scanning fluorimetry (DSF) and light scattering, both static and dynamic, on the UNcle protein stability screening platform (Unchained labs, Malvern, UK) based on manufacturer's guidelines. 8.5µL of 0.5mg/mL purified protein samples were added to each well on the Uni cassette (Unchained labs, Malvern, UK). For T_m/ T_{agg} experiments, samples were incubated at 15°C for 180s, following which the samples were exposed to 266nm and 473nm laser and the static light scattering (SLS) spectrum was measured. Subsequently, the temperature of the samples was raised by 1°C and held for 60s before repeating the same measurements, until a temperature of 90°C was reached. The change in intensity of scattered light at 266nm and 473nm were monitored separately and was used to determine protein aggregation (T_{agg}) (266nm for smaller molecules, 473nm for larger molecules) [195], while the shift in barycentric mean (BCM) of tryptophan fluorescence (305 - 355nm) was used as indication of the extent of tryptophan exposure to surrounding hydrophilic solvent molecules and used to determine the melting temperature (T_m) [196].

2.9.2 Circular dichroism

Collagen I triple helical structure was investigated using circular dichroism (CD) conducted on the J-810 spectropolarimeter (Jasco, Heckmondwike, UK) according to manufacturer's guidelines. Collagen samples were prepared to a concentration of 0.5mg/mL. Purified collagen I samples (solubilised rat tail collagen/reconstituted rat tail collagen gel) were diluted in 20mM acetic acid to obtain the required concentration. Rat collagen I from homogenised rat tail tendons fascicles were first dialysed into 20mM acetic acid using Slide-A-Lyzer mini dialysis device (ThermoFisher, UK) overnight at 4°C with a single change of 20mM acetic acid after the first 2 hours.

160µL of sample was pipetted into a demountable quartz cuvette (Jasco, Heckmondwike, UK) with a path length of 0.5mm. The CD spectrum was obtained between 250 – 195nm

wavelengths at a resolution of 0.2nm, at lowest temperature possible (10°C) to avoid thermal denaturation. The exposure time was set to 0.5 s, with 1nm bandwidth and the spectrum was averaged over three to five accumulations. To investigate thermal stability of protein secondary structure, heating experiments were also conducted. Samples were first incubated at 20°C for 5min, before measuring the CD signal at a chosen specific wavelength (222nm for collagen to observe triple helical structure) every 0.1°C from 20°C to 50°C at a heating rate of 20°C/hour.

2.10 Liquid chromatography – tandem mass spectrometry (LC-MS/MS) sample preparation and analysis

LC-MS/MS was conducted to examine proteome composition as well as peptide location fingerprinting (PLF) analysis to investigate changes in proteolytic susceptibility of proteins, from which localised structural changes in proteins were inferred [185,197]. The overall workflow is described in the figure below (Figure 2.3). Preparation of solubilised/homogenised samples prior to LC-MS/MS injection follows previously established protocol from the BioMS core facility at University of Manchester.

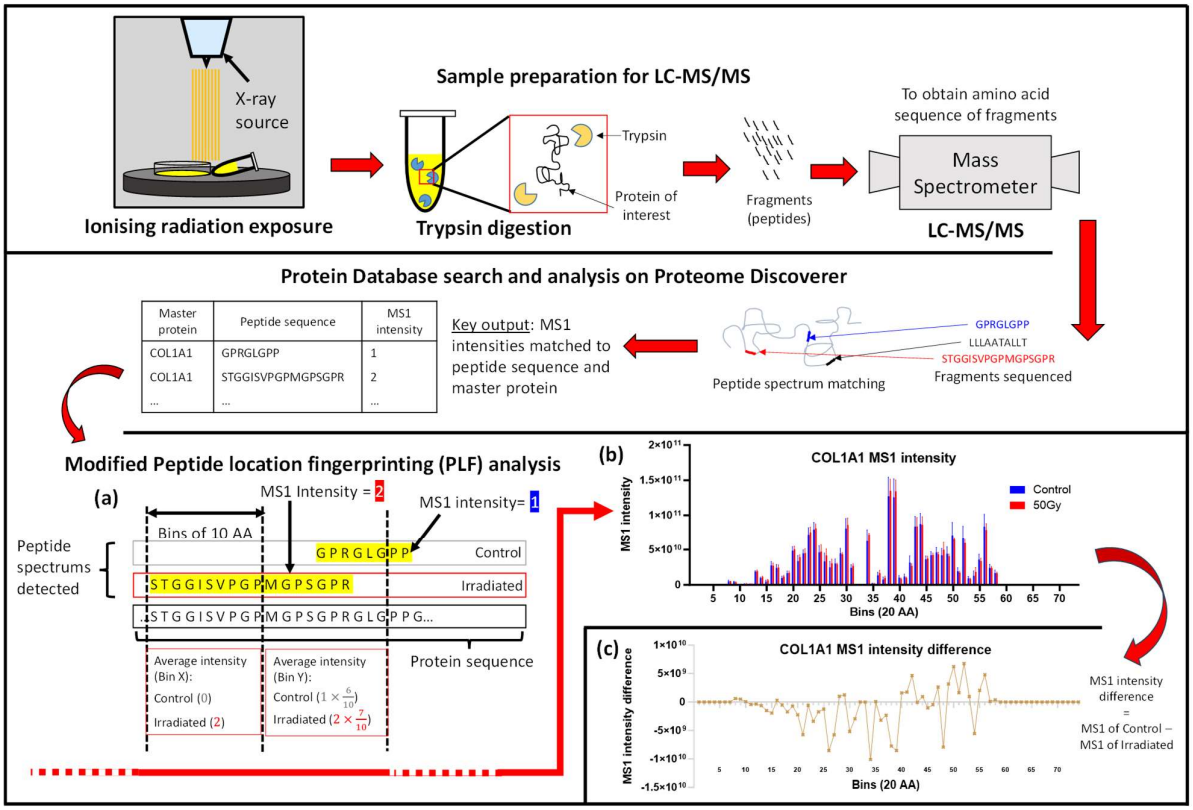


Figure 2.3. Workflow of LC-MS/MS protein database search to peptide location fingerprinting analysis (PLF), starting from top left. After exposure to ionising radiation, samples were prepared accordingly (depending on sample type) and digested with trypsin to produce peptide fragments. Peptides were then identified through LC-MS/MS and protein database search using Proteome Discoverer. The MS1 intensity of all peptides and the respective amino acid sequences were exported into excel and processed for PLF analysis using python. (a) Briefly, the peptides are aligned according to the sequence of their respective master protein. Following which the master protein sequence is segmented into arbitrarily chosen sized bins (10 amino acids (AA) for the example at bottom left). Based on the overlap of the peptides at each bin, the MS1 intensity was averaged at each bin across the whole protein sequence to obtain a graph (b) of the peptide fingerprint of the protein. (c) By subtracting the MS1 intensity of irradiated sample from the control sample, we obtain a graph of MS1 intensity difference.

2.10.1 Reduction, alkylation, and trypsin digestion

10µg of purified protein samples (solubilised rat collagen, plasma fibronectin, reconstituted rat collagen gel) or 100µg of homogenised complex protein mix (Rat tail tendon, HMFU-19 derived ECM, *ex vivo* breast tissue) were diluted to 130µL (purified protein) or 500µL (complex protein mix) containing 5% SDS, 50 mM triethylammonium bicarbonate (TEAB), pH 7.5, prior to reduction with 5mM Dithiothreitol (DTT) and alkylation with 15mM Iodoacetamide (IAM). Next, phosphoric acid was added to a final concentration of 1.2% to acidify the samples, following which binding buffer (100mM TEAB, 90% Methanol, pH 7.1) was added in ratio of 6:1 (binding buffer to volume of samples) to precipitate the proteins. Precipitated proteins were then concentrated by an S-trap column (ProtiFi LLC, Fairport, NY, USA), and the column was washed with binding buffer to remove impurities and excess DTT/IAM before addition of

1:10 trypsin in 20µL of 50mM TEAB (based on the ratio 1:10 trypsin to estimated amount of sample, i.e. 1µg trypsin for 10µg sample, 10µg trypsin for 100µg sample), and left overnight in a 37°C incubator (16-24 hours).

2.10.2 Peptide desalting

Following digestion with trypsin, the resultant peptide suspension was sequentially eluted with 50mM TEAB, followed by 0.1% (v/v) aqueous Formic acid (FA), and finally 0.1% (v/v) FA in 30% (v/v) acetonitrile to give a final 5% (v/v) acetonitrile concentration in the combined elute. The samples were further desalted using Oligo R3 beads added to 96-well Corning FiltrEX desalt filter plates. Samples were loaded into wells with primed beads and washed with aqueous 0.1% formic acid to remove impurities and salt. Subsequently, peptides were eluted with 0.1% formic acid in 30% (v/v) acetonitrile and dried with Heto speed vacuum concentrator centrifuge (ThermoFisher, UK) without heating for 90 minutes. Vacuum-dried peptide samples were then stored at 4°C and resuspended in 5% (v/v) acetonitrile in 0.1% (v/v) FA prior to LC-MS/MS.

2.10.3 Mass spectrometry

LC-MS/MS was performed by the BioMS core facility at the University of Manchester with previously published protocol. Briefly, LC-MS/MS was conducted on UltiMate® 3000 Rapid Separation LC system (RSLC) (Dionex Corporation, Sunnyvale, CA, USA) coupled to either an Orbitrap Exploris 480 (ThermoFisher, UK) (all samples except *ex vivo* breast tissue) or Orbitrap Fusion Lumos (ThermoFisher, UK) (*ex vivo* breast tissue) with buffer A as 0.1% formic acid in water and buffer B as 0.1% formic acid in acetonitrile. Resuspended peptides were first injected an analytical column (nanoEase M/Z Peptide CSH C18 Column, 130Å, 1.7 µm, 75 µm X 250 mm) (Waters Corporation, Milford, MA, USA) at a flow rate of 300nl/min. Peptides were separated over a multistage gradient as follows: 1% B to 6% B over 2 min, 6% B to 18% B over 44 min, 18% B to 29% B over 7 min and 29% B to 65% B over 1 min. MS2 data were acquired in a data-dependant manner over a scan range of 300 to 1750Th.

2.10.4 Protein identification and quantification with Proteome Discoverer

Protein identification was implemented using Proteome Discoverer 2.5 (ThermoFisher, UK) with a pre-optimised workflow consisting of spectrum search using SEQUEST within the Proteome Discoverer software with the following parameters:

Table 2.3. Search parameters for SEQUEST in Proteome discoverer

Species/ Database	For human collagen I, human plasma fibronectin, HMFU, breast tissue: Homo sapiens/swissprot, 2022 Database retrieved from UniProtKB release 2022_01 For rat tail collagen I (solubilised/gel/tendons): Rattus norvegicus/swissprot, 2022 Database retrieved from UniProtKB release 2022_01
Modifications	Fixed: Carbamidomethylation of cysteine residues Variable: Oxidation of lysine/methionine/Proline/Arginine
Error tolerance	For Orbitrap/Exploris 480: Peptide tolerance – 10 ppm MS/MS tolerance – 0.2 Da For Orbitrap Lumos/Ion trap: Peptide tolerance – 5 ppm MS/MS tolerance – 0.3 Da For both, False discovery rate (FDR) for peptides were set at 1%

Within the ProteomeDiscoverer software, peptides identified were filtered to remove those with poor pre-cursor ion (MS1) intensity quantification or lack thereof ('QuanInfo' is not 'None' or 'not reliable'). Only peptides that were unique to a single protein were accounted for ('Master protein' is '1'). The peptide data crucial for analysis were normalised abundances, which were used for MS1 intensity quantification. For the protein data, total peptide spectrum matches (PSMs) for each protein were needed for abundance approximations in chapter 5. All data were exported to an excel file for post-processing on Python.

2.10.5 Detecting protein structural changes with peptide location fingerprinting

A python program was developed from ground up for peptide fingerprinting (PLF) analysis based on published work by Eckersley et. al. (2020) [185] (Figure 2.3), with added modifications. Briefly, peptide sequences, their respective MS1 intensities, and master protein accession were obtained from the Proteome Discoverer exported search data. Peptides were sorted into their respective master protein, and the peptide's MS1 intensity is added onto the amino acids in their respective protein's sequence. After all peptides were located and their intensities accounted for, each protein is arbitrarily segmented into 20 amino acid-sized bins (remaining amino acid that does not fit into the last bin is combined into the last bin). Within each bin, the MS1 intensity of the amino acids are averaged amongst themselves to obtain the MS1 intensity for that specific bin and specific protein. The MS1 intensity of the bins across the entire protein sequence is termed the peptide fingerprint of the protein. The peptide fingerprint data is obtained from python and input into GraphPad Prism 9.0 (GraphPad Software, Boston, MA, USA) for testing statistical significance. Python code can be accessed on Github: (https://github.com/RJTuieng/MS1_PLF.git) (also attached in Appendix 7.6)

2.10.6 Enrichment analysis using STRING webtool

To test for enrichment of biological pathways or protein categories amongst the proteins detected by peptide location fingerprinting analysis, the list of proteins were entered into the STRING webtool [198] (accessed from https://string-db.org/ on 20 June 2023). Search was conducted using the 'Multiple proteins' option by inputting the list of uniport accession obtained from MS. Proteins with less than 45 unique peptide spectrum matches (PSMs) were excluded from subsequent steps.

2.11 Statistical analyses

Statistical testing of data were conducted on Graphpad Prism 9.0 (GraphPad Software, Boston, MA, USA) for SDS-PAGE/Native-PAGE/PLF. Unpaired student's t-test were used for comparing differences between SDS-PAGE/Native-PAGE band intensities. Two-way repeated measures ANOVA with with Geisser-Greenhouse correction and Bonferroni-corrected multiple comparisons test was conducted for PLF analysis. Data are presented as mean with standard deviation as error bars. Level of significance is defined as $p < 0.05$ unless otherwise specified.

For STRING enrichment analysis, the statistical background group of proteins were set to be the complete set of proteins identified by LC-MS/MS for the particular sample type. Enrichment analysis was computed using a hypergeometric test with Benjamin and Hochberg correction for multiple testing [199]. Gene ontology categories and UniProtKB associated keywords with false discovery rate of less than 5% were labelled as significantly enriched.

3 Differential response of Collagen I and Fibronectin to therapeutic ionising radiation

3.1 Introduction

Although it is well established that breast cells are affected by therapeutic does of X-rays [200], the impact of ionising radiation on ECM components remains poorly defined. The complexity of tissue ECMs make it difficult to study and identify key targets of radiation *in vivo*. Therefore, two abundant ECM proteins in breast stroma were selected as potential targets of therapeutic X-ray doses. Collagen I is the major structural protein in breast stroma [201] whilst fibronectin (FN) facilitates cell-ECM interactions [18,202]. In addition to their important biological functions, they have distinct amino acid compositions, which may influence their sensitivity to radiation damage as it has previously been found that proteins with higher compositions of oxidation sensitive amino acids were more susceptible to UV radiation [152,203]. Furthermore, collagen I and FN differ in secondary and tertiary structures (fibrillar for collagen I, and globular for FN). Protein structure can alter the type and number of amino acid side chains that are exposed to the environment and hence oxidative damage through water radiolysis [204,205]. Studying these two proteins may thus provide valuable insights into how protein sequence and structure can impact the outcome of radiation exposure on biomolecules.

3.1.1 Collagen I

Collagens make up 30% of our body's dry weight [206], making it the most abundant protein in the body. Collagens are comprised of three amino acid chains characterised by a repeating -Gly-X-Y- (X, Y being other amino acids) motif along the primary sequence, forming a triple helical structure. Amongst fibrillar collagens, collagen I is the most common, comprising 80-90% of all types of collagens, and being the most abundant protein in skin, tendons and even bone.

The collagen I monomer consists of two α_1 chains and a single α_2 chain. They are first transcribed as individual chains, after which they spontaneously form the triple helix within the cell. They are then secreted into the extracellular space, where their N- and C- terminal propeptides are cleaved off by metalloproteinases to allow self-association, and thereby, fibril formation to occur. Collagen I fibrils provide tissues with a mechanically resilient scaffold to

maintain structural integrity, while also serving as cellular binding sites [207,208]. Being the underlying foundation for tissue strength, collagen I is consequently a key player in overall mechanical stiffness of tissue and alterations to collagen content are thought to play critical roles in mechanosensitive pathways that can alter cell behaviour or cell fate [145,146].

3.1.2 Fibronectin

Fibronectin (FN) is a dimeric glycoprotein consisting of multiple repeating units (namely FN I, FN II, and FN III) along its primary sequence. There are two isoforms: cellular FN (cFN) is secreted by fibroblasts, and plasma FN (pFN) by hepatocytes. Cellular and plasma FN are produced via alternative splicing of FN mRNA. In pFN there is a lack of EDA and EDB domains and the interconnecting region (ICSIII) domain exists only on one of the pFN subunits in a dimer. Circulating in blood plasma, pFN plays a role in mediating early wound healing responses. pFN is therefore a major component of the fibrin clot formed during wound healing, and is crucial for allowing cells to bind to the fibrin matrix to facilitate downstream repair processes [209,210]. Cellular FN (cFN) comprises a mixture of alternatively spliced isoforms that contain either EDA domain, EDB domain, or both, with the ICSIII domain in both FN units of the dimer [211]. Fibroblasts secrete cFN into the ECM of the breast where it self-associates to form fibrils [212] and mediates cell-ECM interactions via integrins, most notably α5β1 [213] through the RGD sequence. This tissue bound isoform can also bind to the ECM through collagen I [214] and plays a role in collagen I fibrillogenesis [215,216]. Its interaction with collagen and other ECM proteins allow cFN to play a key role in transducing mechanical cues from the ECM to the cells through the integrin-cytoskeleton network [202,217].

3.1.3 Molecular impact of ionising radiation on ECM proteins

Given the abundance and functional importance of collagen I and FN in the breast stroma, it is important to characterise their susceptibility to ionising radiation. Changes to protein structure, especially to such large, multi-functional ECM proteins, may trigger a wide cascade of altered binding affinities to cells or other proteins in the ECM [218–220]. However, current literature lacks studies that probe the ECM without cells, concentrating instead on radiation-induced outcomes associated with altered cell behaviour.

3.2 Hypothesis

To avoid conflating cellular and acellular mechanisms of radiation-induced degradation of ECM proteins, this study aimed to investigate the impact of X-ray exposure on purified, solubilised forms of collagen I and pFN to test the hypothesis that therapeutic x-ray doses would differentially affect the structure of proteins based on their amino acid composition and native protein shape. To address this, the study was further split into three sub-hypotheses:

1) Collagen I and FN are differentially susceptible to therapeutic doses of X-ray radiation
2) X-ray exposure will induce region-specific structural changes in both collagen I and pFN.
3) X-ray-induced remodelling of molecular structure will impact FN/collagen I binding.

Purified human collagen I and plasma FN (pFN) sourced from Abcam (Native human collagen I, Ab7533; Native human plasma fibronectin, Ab80021) were used for all experiments. pFN was chosen as it has less structural variation (i.e. it comprised of mainly dimeric fibronectin) and has high purity. In comparison, cellular FN (cFN) sourced from manufacturers comprised a mixture of monomeric and dimeric FN, and contained significant amount of contaminating proteins that were co-purified. All samples were exposed to 50Gy and 100Gy of x-ray dose from the Xstrahl CIX3 x-ray cabinet.

To test the first hypothesis, SDS- and Native-PAGE were used to detect changes in electrophoretic mobility and structure (n=3, technical replicates, unless otherwise stated). Further investigation on changes to thermal stability was carried out using differential scanning fluorimetry (n=3), light scattering (n=3), and circular dichroism (n=3). For the second and third hypotheses, LC-MS/MS was conducted and data were examined with peptide location fingerprinting (PLF) analysis (n=5) [185,197].

3.3 Results

3.3.1 Ionising radiation disrupts collagen I primary structure while triple helical structure and thermal stability appear unaffected

Following SDS-PAGE, x-ray exposed collagen I had increased background staining between the α monomer, β dimer, and γ trimer bands, and this increase in background was observed in both 50Gy and 100Gy exposed samples (n=3) in a dose dependant manner (Figure 3.1a). This indicates possible oxidation and breakage of the α_1 and α_2 strands due to exposure to ionising radiation, forming smaller peptide strands that constitutes the increased background stain. The non-specific smearing suggests that X-ray induced protein fragmentation is random and hence that there are likely no specific intermediates formed from the effects of X-rays on the primary structure. The background-corrected mean intensity of the regions with increased background staining were obtained using ImageJ. In between the α monomer and β dimer bands (fragment region 1, Fr1) a 22% increase in mean intensity was observed at 50Gy (p=0.06) and 31% increase at 100Gy (p=0.06). In the region below the α monomeric chains (fragment region 2, Fr2), the increase in mean intensity was greater with a 180% increase for the 50Gy exposure (p=0.02) and 415% increase at 100Gy (p=0.008) (highlighted in red brackets). This suggests that the fragmentation of collagen I by therapeutic x-rays is dose dependant and increases with increasing dose.

SDS-PAGE denatures the collagen from its triple helical state and separates the monomer into individual alpha chains, thus do not allow detection of any x-ray induced changes to the collagen native structure. Hence, native-PAGE was also conducted in parallel to characterise any possible changes in native structure and protein folding of collagen post irradiation. Due to isoelectric point of collagen, acidic native PAGE (AN-PAGE) was conducted for collagen I. Electrophoretic mobility of native collagen monomers did not appear to be altered by radiation as observed on native-PAGE (Figure 3.1b). Collagen I produced two bands on AN-PAGE, neither of which showed observable difference in intensity between irradiated and control samples (n=1) (Figure 3.1b). This suggest that collagen monomers remain intact (as triple helical assemblies of three alpha chains) at doses of up to 100Gy. However, by running a 2nd dimension SDS-PAGE on excised lanes from AN-PAGE (Figure 3.1b), both the lower and upper bands were determined to contain both α_1 and α_2 strands of collagen I. The 2nd dimension SDS-PAGE exhibited similar smearing/fragmentation signature from the 1-

dimension SDS-PAGE, but only in the upper band and not the lower band. In addition, all three samples had smearing of α_1 chains along the native gel between the two bands.

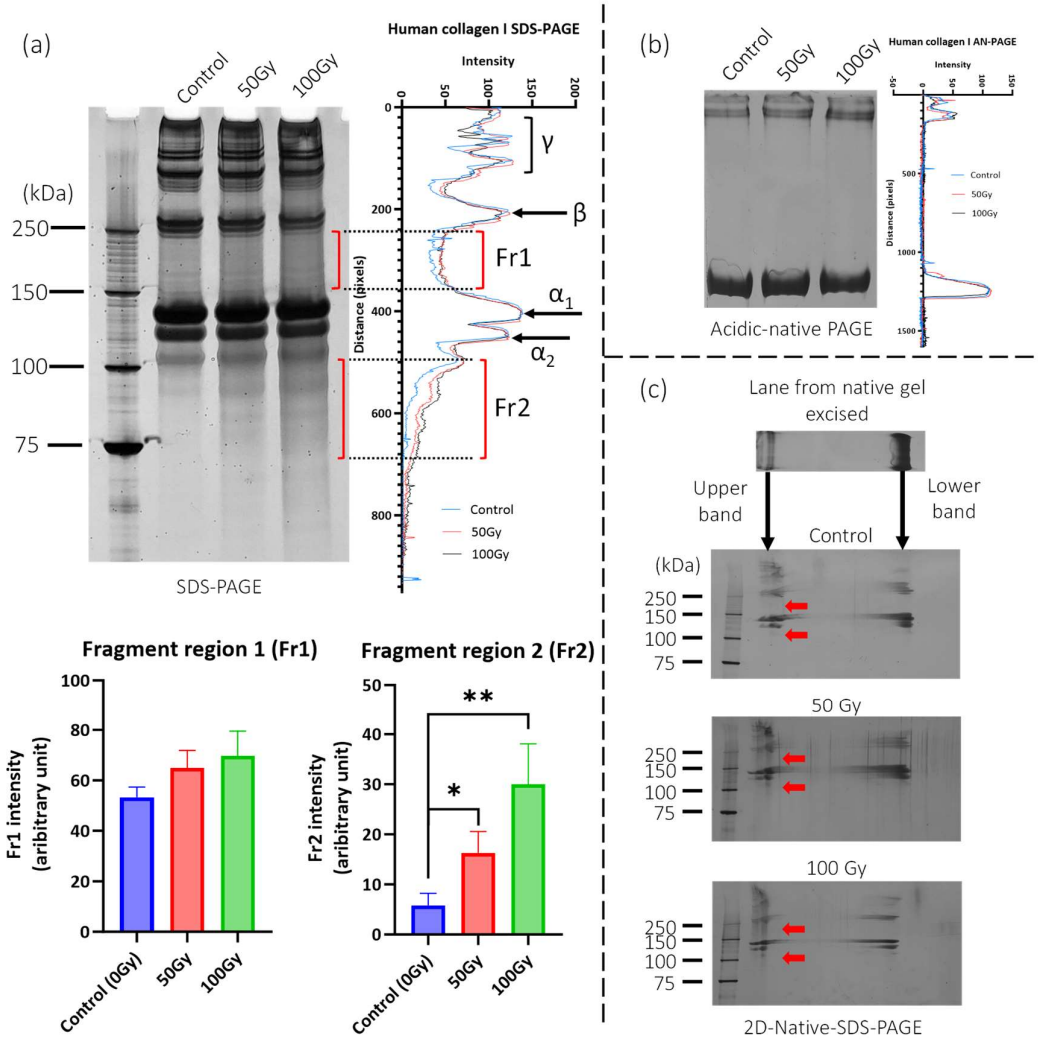


Figure 3.1. SDS-PAGE and acidic native-PAGE of human collagen I exposed to therapeutic x-ray doses. (a) SDS-PAGE for Collagen I showed 50Gy of therapeutic X-rays is sufficient to induce significant fragmentation of the primary backbone of collagen α1 and α2 strands, exhibited by a 22% increase in background-corrected mean staining intensity between the β dimer and γ trimer bands (Fr1) (p=0.06) and a 180% increase in staining intensity in the lower molecular weight region (Fr2) (p=0.02) (highlighted in red brackets). This effect seems to be dose dependant and is greater for the 100Gy exposed sample, with Fr1 appearing 31% higher in mean intensity (p=0.06) and 415% higher for Fr2 (p=0.008). Statistically significant differences were observed (using t-test) in Fr2 while for Fr1 the p values were only close to significant.

(b) Acidic native-PAGE (AN-PAGE) for human collagen I showed no discernible difference between control and irradiated samples (50/100Gy) (n=1). Native structure seems unaffected by therapeutic radiation. (c) In subsequent 2D-Native-SDS-PAGE, smearing observed from 1D-SDS-PAGE data seem to appear only on the upper bands of the native gel (regions indicated by red arrows), implying possible contribution of fragments to aggregation of the protein.

To confirm the AN-PAGE observation that the monomeric assembly of collagen I was unaffected by therapeutic x-rays, circular dichroism (CD) was also used to characterise the impact of X-ray exposure on the collagen triple helical structure. The CD peak value at 222nm was used as proxy for the abundance of triple helical structure within the sample [221–223]. There was no observable difference in peak value at 222nm between control and radiation exposed samples (n=2) (Figure 3.2a), indicating that the triple helical structure of collagen was unaffected by therapeutic doses of X-rays. The thermal stability of the triple helix of collagen I was also investigated using the CD peak at 222nm (n=1). Upon heating the samples (Figure 3.2b), two melting points were observed where the CD signal dipped: 36°C (T_{m1}) and 43°C (T_{m2}). The melting points were obtained by taking the first derivative of the melting curve and finding the position of the peak values [224]. The CD value was observed to start off slightly higher for the irradiated samples than the control, but subsequently dropped to the same level after the first melting point near 36°C (T_{m1}). All samples underwent a sharp decrease in CD signal at approximately 43°C (T_{m2}), indicating denaturation and melting of the triple helical structure. The absence of observable changes to the melting points of collagen I imply therapeutic x-rays did not compromise the thermal stability of the monomeric collagen.

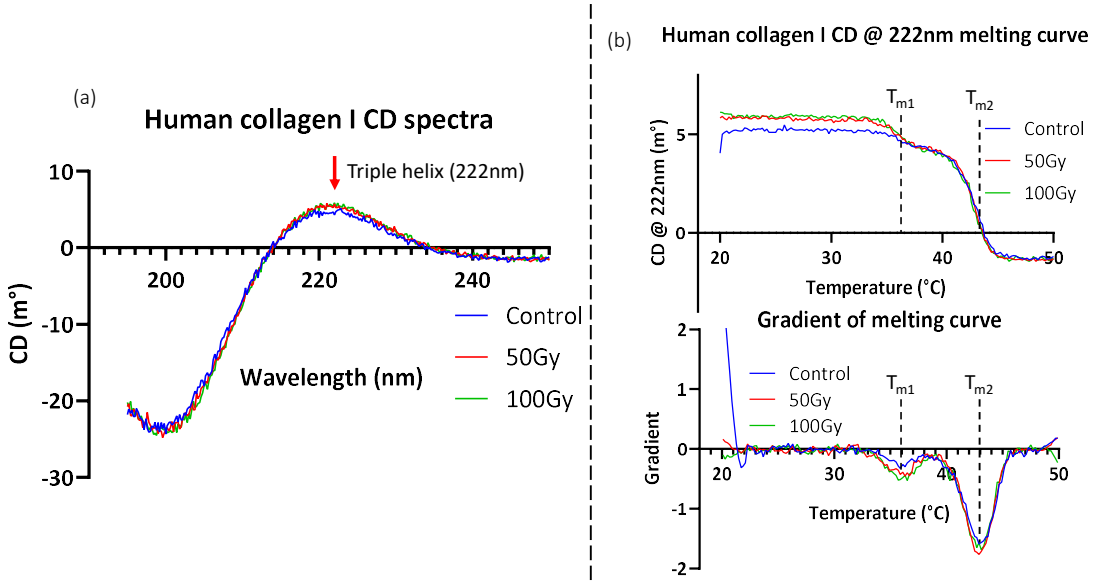


Figure 3.2. CD spectrum and melting curve of human collagen I. (a) Therapeutic x-ray had no observable effect on the CD spectrum and the 222nm triple helical structure peak for collagen I at 50/100Gy exposure (n=3). This suggests that the triple helical structure was unlikely compromised. (d) A melting curve was obtained by measuring the CD value at 222nm over a temperature ramp from 20 to 50°C (n=1). The melting points were obtained by taking a derivative of the melting curve and determining the position of the peaks. The CD value was initially higher for the irradiated samples than the control but dropped to the same level near 36°C (T_{m1}). All samples underwent a sharp decrease in CD signal at approximately 43°C (T_{m2}), indicating denaturation and melting of the triple helical structure. No difference in melting temperature was observed between control and radiation exposed samples.

Overall, SDS-PAGE showed that therapeutic x-rays fragment the primary structure of monomeric collagen I, but CD and AN-PAGE suggests that it did not have measurable impact on the triple helical structure nor the thermal stability of monomeric collagen I

3.3.2 Therapeutic ionising radiation altered regional susceptibility to tryptic digestion of collagen I

LC-MS/MS was performed for non-irradiated and irradiated collagen I and the data analysed with peptide location fingerprinting (PLF) (n=5). PLF is an analytical tool that uses mass spectrometry data to detect regional changes in proteolytic (trypsin) susceptibility of large ECM proteins [197,225]. Here, PLF was implemented in-house using a python program developed from scratch with added modifications (see section on methods for more details). Briefly, two improvements were made to the PLF workflow: 1. MS1 peptide intensities were used instead of peptide spectrum counts, which is a more accurate representation of the peptide amounts [226]; 2. The MS1 intensities for each peptide were averaged across the arbitrarily chosen sized bins while taking into account the overlap of peptides in each bin, instead of crude counting of peptides at each bin that it overlaps into without averaging. To better illustrate the process of analysis, the peptide fingerprint of collagen α_1 chain is shown (Figure 3.3) (Peptide fingerprint for both collagen α chains and pFN can be found in Appendix 7.1) . This fingerprint represents the regional sum of MS1 intensities of all peptides identified from LC-MS/MS data mapped onto the protein sequence, which has been arbitrarily split into bins of 20 amino acids (Figure 3.3). The size of 20 amino acids was chosen as it is close enough to the maximum peptide limit for the search parameters (5kDa, approximately 45 amino acids assuming average mass of 110Da) to minimise peptide overlaps into multiple bins, while at the same time provide sufficient resolution for the various domain positions along the protein sequence.

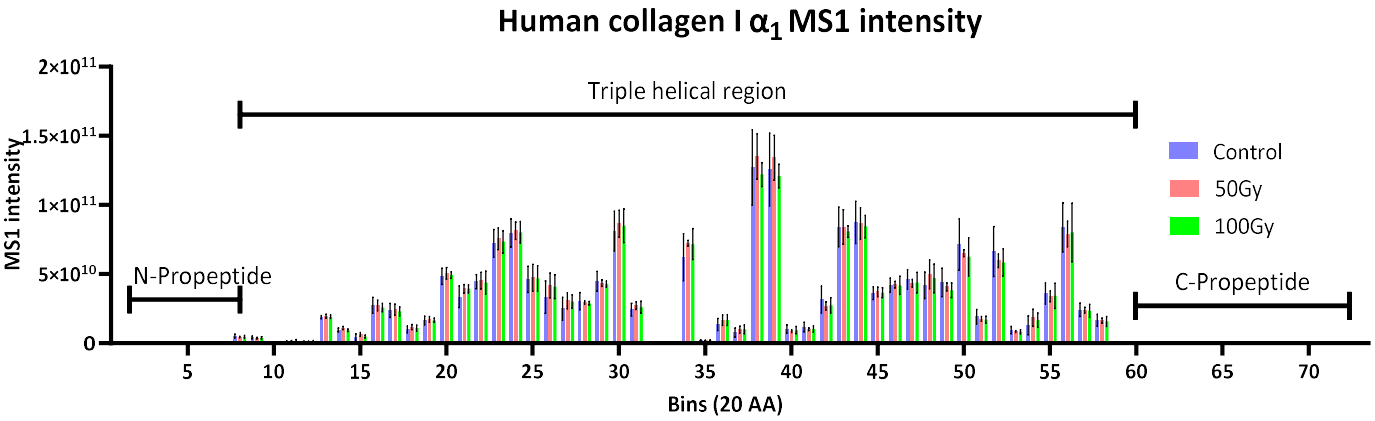


Figure 3.3. Peptide fingerprint for collagen I α_1 chain. Mass spectrometry MS1 intensity of peptides were summed across the primary sequence of collagen I α_1 chain. The protein was arbitrarily subdivided into bins of 20 amino acids, and the MS1 intensity of peptides falling within those regions are weighted according to their overlap and summed. Each bin contains three bars representing MS1 intensities for control, 50Gy, and 100Gy samples respectively (n=5 each). The pattern of intensities across the bins are unique to the protein and sample, hence termed a peptide fingerprint.

Peptide fingerprints of control samples were compared separately with fingerprints of 50Gy and 100Gy x-ray exposed samples by calculating the difference in MS1 intensities between control and the x-ray exposed sample at each bin. From this, a plot of MS1 intensity difference was obtained (Figure 3.4). Regions where the MS1 intensity difference was significantly different ($p<0.05$) are marked red and deemed as 'flagged' by PLF analysis.

PLF analysis of collagen I chains show that difference in peptide MS1 intensity appear to correlate well across the protein for both 50Gy and 100Gy samples (values of MS1 intensity difference and the relative intensities between bins are similar). Only small exceptions were detected such as in bins 38-39 for α_1 and bins 37-38 for α_2. This implies the effects of radiation on collagen I peptide fingerprint could be relatively consistent. α_2 chains also had greater number of regions flagged by PLF analysis than the α_1 chain. Although there were more flagged regions at the lower dose of 50Gy than 100Gy. Crucially however, the PLF analysis demonstrates, as did SDS-PAGE analysis, that therapeutic X-ray doses can affect the structure of collagen I alpha chains (hypothesis 1), and that these changes are localised and repeatable (hypothesis 2).

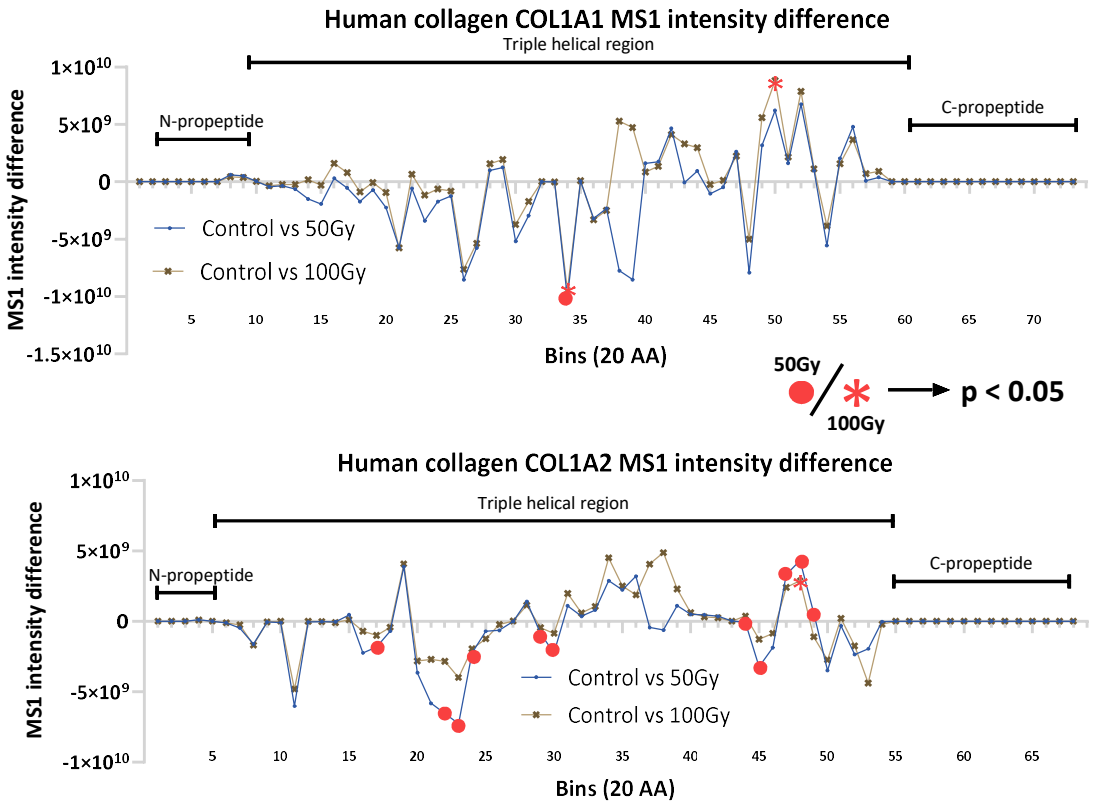


Figure 3.4. Difference in peptide MS1 intensities along collagen I α_1 (COL1A1) (top) and α_2 (COL1A2) (bottom) chains compared between control and treated groups (50Gy and 100Gy). The collagen I sequence is divided into equal bins of 20 amino acids and bins where peptide intensities were significantly altered ($p<0.05$) are marked in red (circle for 50Gy, star for 100Gy). Positive values indicate peptide intensities were greater in control, while negative values indicate it was greater in treated samples. 50Gy exposed collagen I α_2 chain appeared to have greater number of regions where there was significantly altered peptide intensities compared to the α_1 chain.

3.3.3 The primary and tertiary structures of plasma fibronectin are resistant to ionising radiation when analysed by gel electrophoresis and differential scanning fluorimetry

In contrast to collagen I, pFN did not exhibit any significant changes in its electrophoretic mobility or band intensity post-radiation exposure up to 100Gy (Figure 3.5a, b). Background intensity and monomer and dimer band intensities remained unchanged (n=3) ($p<0.05$) by SDS PAGE, indicating there was no observable fragmentation of the primary structure for pFN or induced aggregation. Blue native-PAGE (BN-PAGE), which was utilised to characterise possible changes to the native tertiary structure of pFN, did not find any significant changes to pFN from exposure to therapeutic x-rays (Figure 3.5c). The main band in native-page was further confirmed to be the dimer by its apparent molecular weight in 2D-SDS PAGE (Figure 3.5d). No differences were observed between control and irradiated samples for all bands, implying there were no significant changes to the pFN self-interaction or molecular weight following ionising radiation exposure.

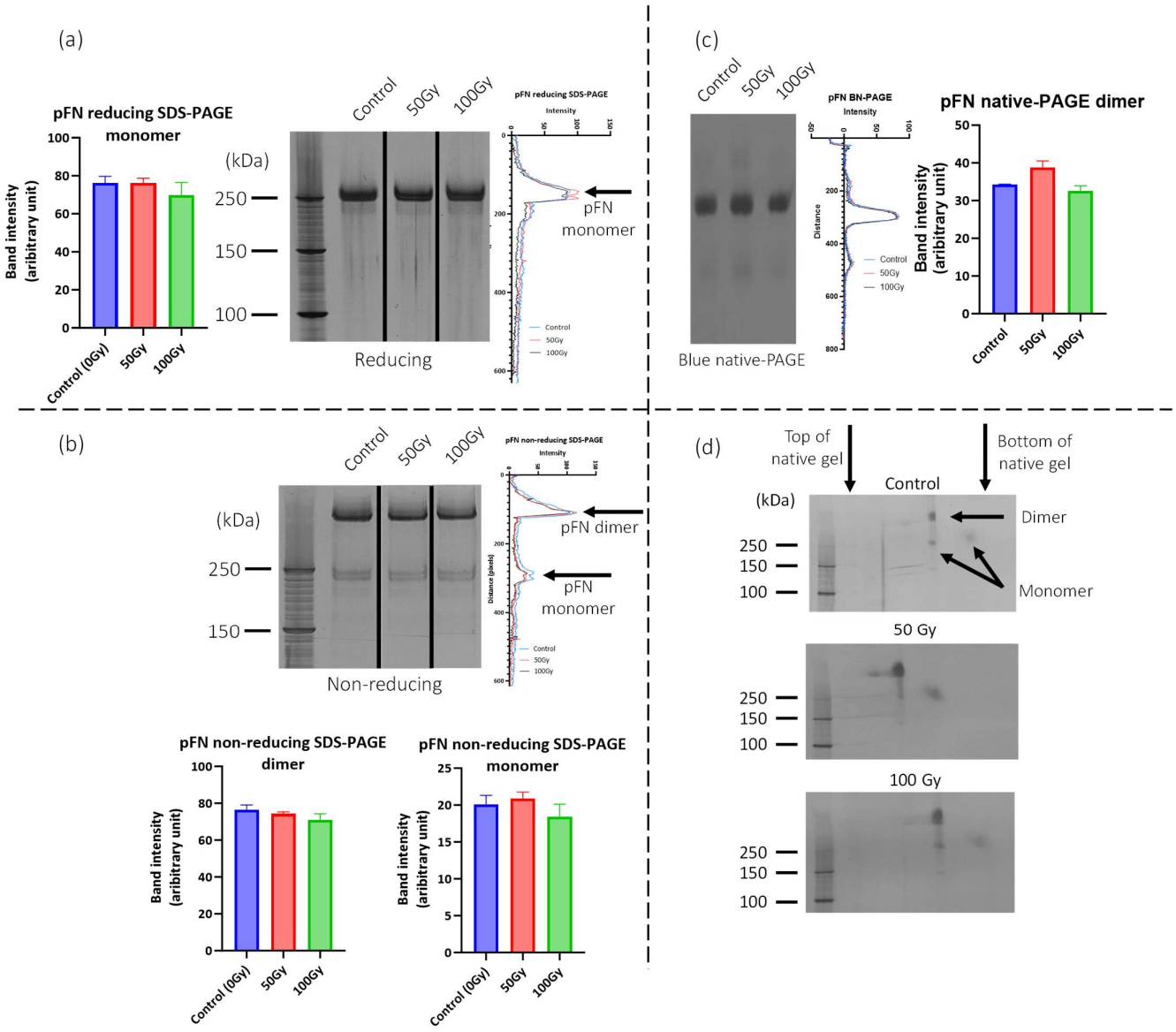


Figure 3.5. SDS-PAGE and blue native-PAGE for pFN exposed to therapeutic x-rays. (a) Both reducing and (b) non-reducing SDS-PAGE for of pFN indicated that 100Gy of therapeutic x-rays has no significant effect on the electrophoretic mobility of fibronectin dimer or monomer (n=3).

(c) Blue native-PAGE (BN-PAGE) for human plasma fibronectin (pFN) indicated that exposure of up to 100Gy of therapeutic X-rays has no significant effect on the overall native dimeric structure or its electrophoretic mobility (n=3). (d) 2D-BN-SDS-PAGE for plasma fibronectin confirmed that the high intensity band observed in BN-PAGE is indeed the dimer with a molecular weight of ~500kDa.

Differential scanning fluorimetry was unable to detect any alteration to the melting point (T_m) and aggregation temperature (T_{agg}) of pFN. T_m was investigated by observing the barycentric mean of the emissions from tryptophan residues (BCM) as a proxy for the melting of pFN structure [196] (Figure 3.6a). By taking a first derivative of the BCM curve (Figure 3.6b), the melting point (T_m) can be located at the position of highest gradient, which corresponded to 59.3 ± 0.6°C for control, 59.3 ± 0.6°C for 50Gy, and 59 ± 1°C for 100Gy x-ray exposed pFN (n=3). The aggregation temperature (T_{agg}) was probed by studying the scattered intensity of 266nm wavelength laser incident on the sample, which was used to correlate with particle size in sample [195] (Figure 3.6c). Increased scattering indicated increased mean particle size and thus, aggregation. T_{agg} was determined by the onset of aggregation, which was found to be 59 ± 0.3°C for control, 59.3 ± 0.6°C for 50Gy and 60 ± 1°C for 100Gy x-ray exposed pFN (n=3). The 100Gy samples appear to aggregate at a slower rate seen by the gentler increase in 266nm scattering compared to control and 50Gy. However, no statistically significant change in T_m nor T_{agg} was found for x-ray exposed pFN, suggesting that overall thermal stability of the protein was not compromised.

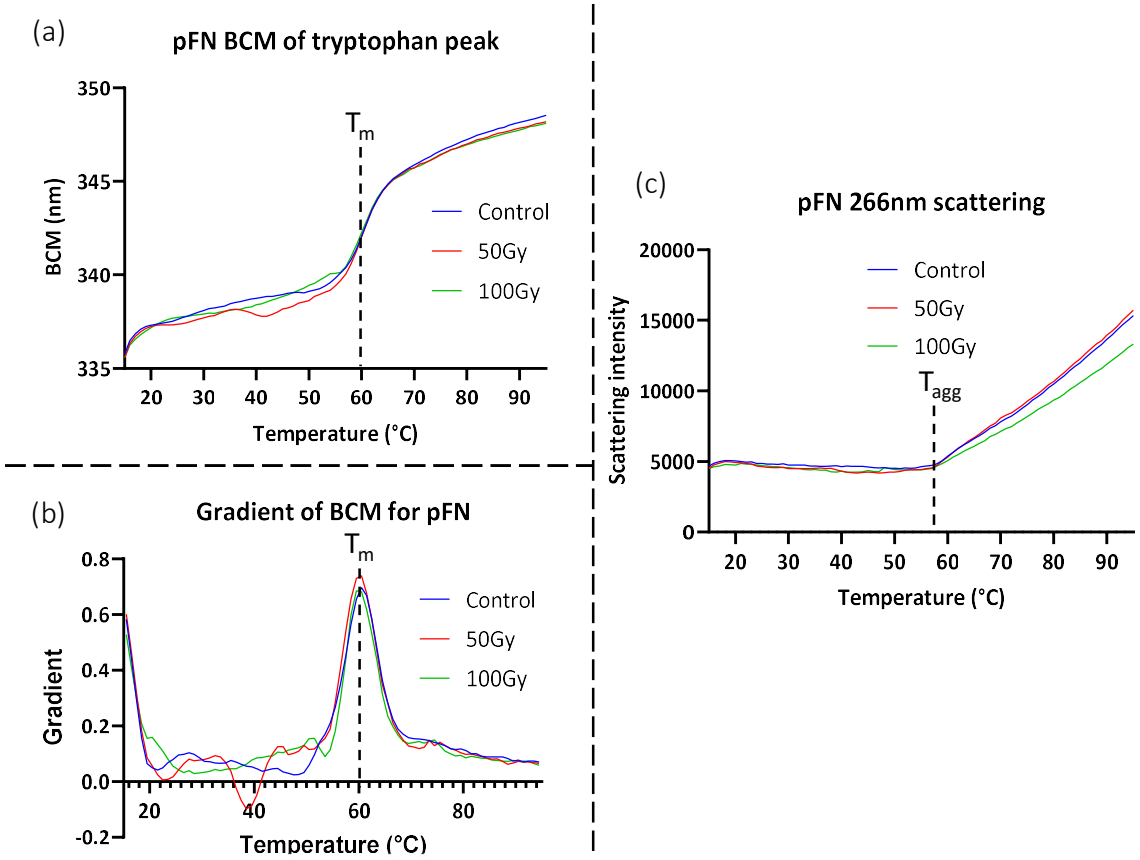


Figure 3.6. Differential scanning fluorimetry was conducted to probe the melting (T_m) and aggregation (T_{agg}) temperature of pFN. (a) The barycentric mean of the emissions from tryptophan residues (BCM) was used to detect melting of the protein structure. (b) The first derivative of the BCM curve was used to locate the maximum gradient, inferred to be the melting temperature, and was determined to be 59.3 ± 0.6°C for control, 59.3 ± 0.6°C for 50Gy, and 59 ± 1°C for 100Gy x-ray exposed pFN (n=3). (c) To obtain T_{agg}, light scattering at 266nm wavelength was used and the increase in scattering intensity correlates with increased particle size in sample, thus aggregation. T_{agg} was determined by the onset of aggregation, which was found to be 59 ± 0.3°C for control, 59.3 ± 0.6°C for 50Gy and 60 ± 1°C for 100Gy x-ray exposed pFN (n=3). Post-irradiated pFN did not have statistically significantly different T_m nor T_{agg} from control samples, suggesting that overall thermal stability of the protein was not compromised.

3.3.4 Therapeutic ionising radiation altered regional susceptibility to tryptic digestion of plasma fibronectin

Although X-ray exposure had no discernible effect on the ultrastructure of pFN as analysed by gel electrophoresis or differential scanning fluorimetry, PLF analysis of LC-MS/MS data indicates that x-ray irradiation impacts tryptic digestion (and possibly localised structure) of pFN in a dose dependent manner (Figure 3.7). Exposure to 50Gy resulted in a single region, between amino acids (AA) 920 to 940, where the fingerprint differed significantly ($p<0.05$) from that of the control sample (n=5). Exposure to 100Gy, however, resulted in multiple (seven) remodelling events including in regions which coincide with pFN binding sites for, fibrin, integrin (cell attachment), and collagen. This supports the hypotheses (hypothesis 1 and 2), showing that pFN, while resistant to fragmentation (unlike collagen), also appears to have structural changes when exposed to x-rays.

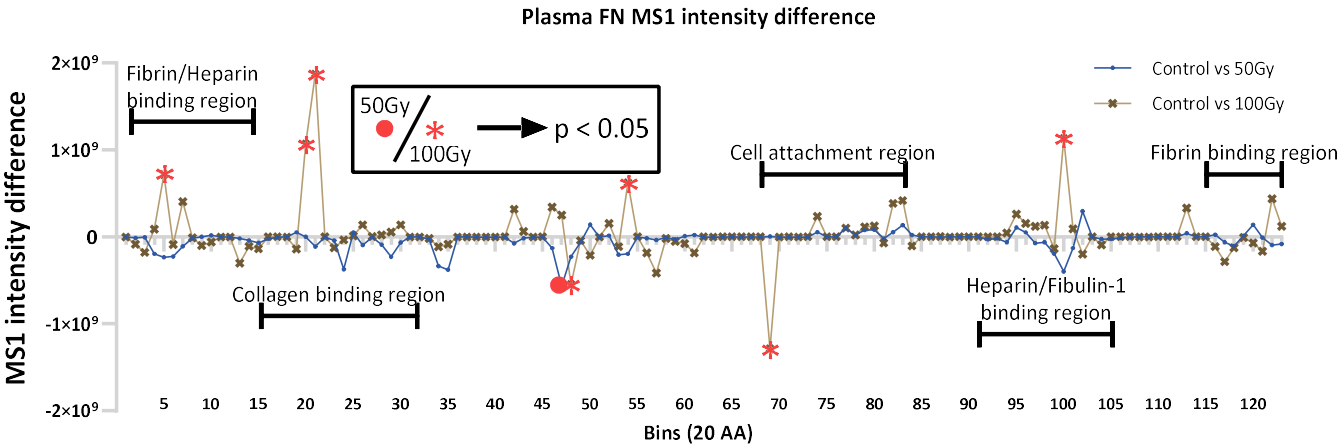


Figure 3.7. Difference in peptide MS1 intensities along pFN sequence between control and treated groups (50Gy and 100Gy). pFN sequence is divided into equal bins of 20 amino acids and bins where peptide intensities were significantly altered are marked in red (circle for 50Gy, star for 100Gy). Positive values indicate peptide intensities were greater in control, while negative values indicate it was greater in treated samples. Greater number of regions were flagged in 100Gy exposed pFN, and those regions are associated with fibrin, heparin, and collagen binding sites.

3.3.5 Presence of oxidation sensitive amino acid does not correspond with PLF data

Given that previous reports have shown that suscpetibiliy to UVR and oxidative damage may be related to the presence of reactive oxygen species (ROS) susceptible amino acids, namely tryptophan, tyrosine, methionine, cysteine, and histidine, similar differences in amino acid composition was suspected to drive localised remodelling in x-ray (and therfore ROS) exposed proteins [152,185,203]. To test that hypothesis, the composition of each bin was analysed for the percentage make-up of these ROS susceptible amino acids (Figure 3.8). However, there were no observable correlation between the percentage of these amino acids and the likelihood of those regions being flagged by PLF.

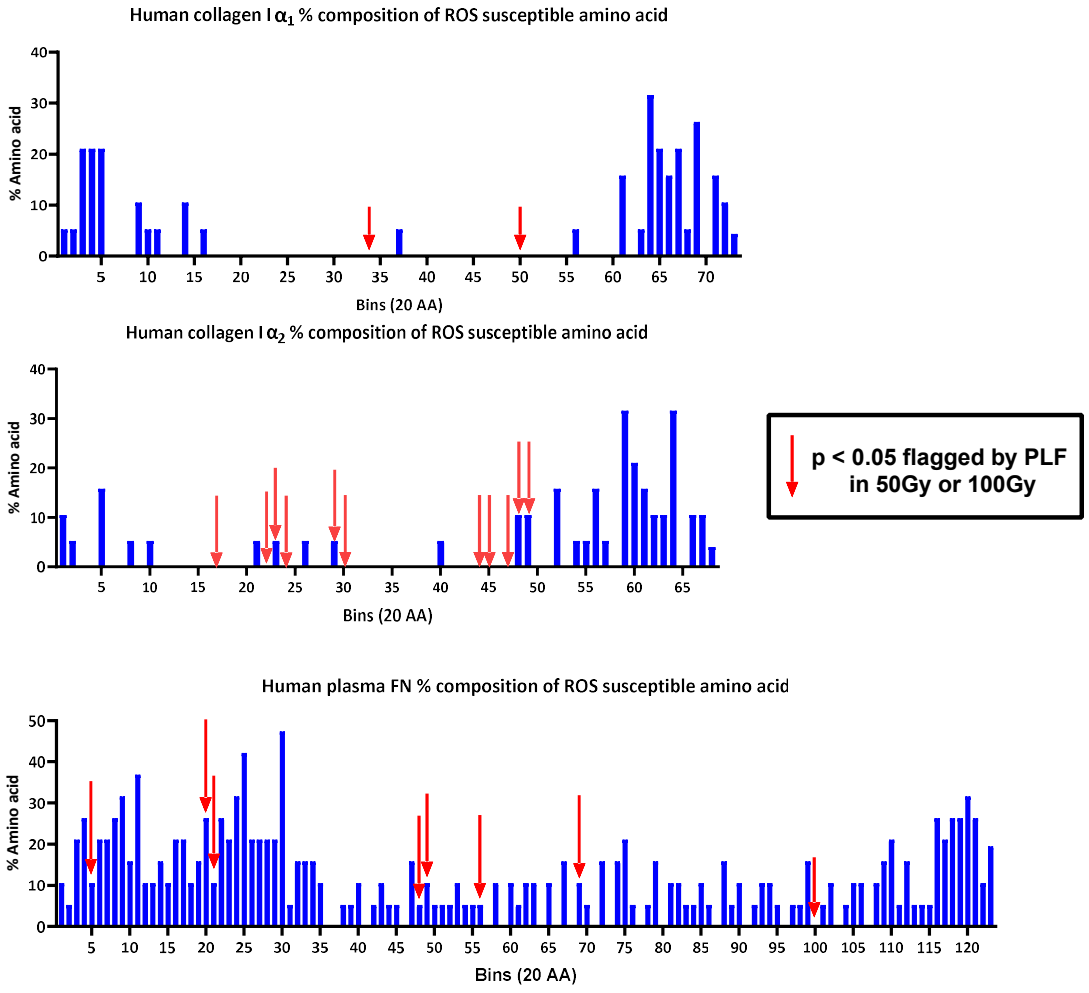


Figure 3.8. The composition of each bin (that were arbitrarily chosen) in PLF analysis was analysed for the percentage of reactive oxygen species (ROS) susceptible amino acids. The regions flagged in PLF analysis in both 50Gy and 100Gy results are indicated in the graphs. No observable correlation was found between the percentage of ROS susceptible amino acids within the bins and the likelihood for PLF to detect significant alteration of peptide intensities at those regions.

3.3.6 Therapeutic radiation may increase plasma fibronectin's binding affinity to collagen I

The apparent structural remodelling within the collagen binding region of pFN by x-ray exposure provided an opportunity to test the third hypothesis, that x-ray induced remodelling of molecular structure will impact binding of FN with collagen I. PLF data of pFN highlighted a significant increase in tryptic peptide yields from the collagen binding region of pFN exposed to 100Gy of ionising radiation, implying that the region may be structurally compromised thus resulting in increased proteolytic accessibility to those regions. Its binding affinity to collagen I, which is likely sensitive to the conformation of the binding sites, could thus be altered. To test the hypothesis, a solid-phase ELISA assay was conducted, with unirradiated collagen I as the substrate coated on 96-well plates and 50Gy or 100Gy irradiated pFN as the binding partner (Figure 3.9). A monoclonal FN antibody paired with horseradish peroxidase conjugated secondary antibody was used to detect the amount of pFN bound to the collagen substrate.

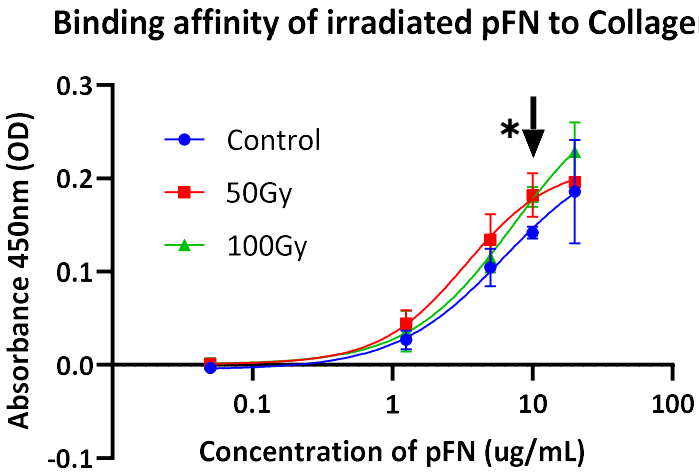


Figure 3.9. Solid-phase ELISA of pFN with collagen I substrate show that exposure of pFN to 100Gy of ionising radiation had increased binding affinity to collagen I, with 29% and 27% increase in absorbance for the 50Gy (p=0.04) and 100Gy (p=0.006) respectively at 10µg/mL. A consistent increased binding rate was also observed for both 50Gy and 100Gy exposed pFN across 1-10µg/mL concentrations although they were not statistically significant. Curve fitting was done with GraphPad Prism 9.0 software's built-in sigmoidal 4-parameter logistic curve (4PL) with R^2 values of 0.91 (Control), 0.96 (50Gy) and 0.97 (100Gy).

The ELISA results displayed a 29% and 27% increase in absorbance for the 50Gy (p=0.04) and 100Gy (p=0.006) exposed pFN respectively at the coating concentration of 10µg/mL (n=3). , This indicates a statistically significant increase in the amount of pFN bound to the collagen substrate for radiation exposed pFN at 10µg/mL. Whilst not significant at other concentrations, irradiated pFN (both 50Gy and 100Gy) consistently bound to collagen I with higher affinity than unirradiated pFN.

3.4 Discussion

3.4.1 Collagen I and plasma fibronectin exhibit differential susceptibility to therapeutic x-rays

3.4.1.1 Impact on collagen

The conventional biochemical techniques used in this study found that collagen I primary structure was profoundly altered. The smearing observed on the SDS-PAGE from radiation exposed collagen I represents a stochastic fragmentation of the collagen I primary structure, similar to that produced chemically by oxidative mechanisms [227]. This implies that collagen I's primary structure could be compromised at therapeutic levels of x-rays (50Gy). In contrast, investigations on its native structure by native-page and CD revealed that the triple helical structure appear radioresistant even at twice the therapeutic dose of x-rays. Taken together, the result indicates that after x-ray exposure, collagen chains peptide backbone fragments, but remains in the triple helical state.

To investigate the thermodynamic stability of collagen triple helix, a thermal ramp was conducted while monitoring the CD spectra of the collagen suspension, in which two melting points (36°C and 43°C) were observed. Previously published work by Mu et. al. (2007) observed the same phenomenon and attributed the first 'shoulder' transition at the lower temperature (36°C) to the disassociation of a small group of fibrils formed amongst the monomers [228]. Here, x-ray exposed samples appear to have slightly higher initial CD values but dropped to similar values to control samples immediately after the first transition. This may only be possible if therapeutic x-rays induced a small amount of fibrillar structure formation, of which were then degraded by the heating process. However, the mechanism of how therapeutic x-rays might induce collagen fibril formation (if there is any) is still not known and no other literature has reported to investigate x-rays impact on collagen fibrillogenesis. Additionally, the increase in CD values may simply be a technical error due to the low number of repeats (n=2). The data presented here is insufficient to provide irrefutable evidence of fibrillogenesis and further work needs to be conducted to confirm this phenomenon.

The mass spectrometry and PLF analyses could provide some possible mechanistic insights on how the collagen monomer is affected by therapeutic x-rays. The regions flagged by PLF is indication that proteolytic susceptibility in those regions had been altered and could imply changes to regional conformation. Analysing the PLF results from collagen I show that there are greater number of regions affected in α_2 chains than α_1, implying that proteolytic

accessibility to α_2 was preferentially altered. This may be due to the inherent resistance of the α_2 to proteolytic cleavage [229], which confers greater potential for altered proteolytic susceptibility. The higher proteolytic resistance of α_2 chain in a native collagen monomer may be due to the wrapping of two α_1 chains it, possibly providing some protection against proteolysis by reducing the accessibility of trypsin cleavage sites in a native collagen monomer [229,230]. After x-ray exposure, however, the fragmentation of the collagen primary structure (as was observed by SDS-PAGE), may destabilise the triple helix during preparation for mass spectrometry, where an elevated temperature was used while performing trypsin digestion (37°C) [222,231]. This destabilisation may explain why 8 out of 11 flagged regions indicated increased peptide yields (negative MS1 peptide intensity differences) after 50Gy x-ray exposure for α_2 chain, suggesting that proteolytic susceptibility of α_2 has generally increased.

Overall, the work on collagen I revealed that collagen I peptide backbone is susceptible to oxidative cleavage by therapeutic x-rays, and this may destabilise the collagen triple helix at *in vivo* temperature, exposing the α_2 chain to proteolytic cleavage.

3.4.1.2 Impact on of X-ray exposure pFN

Investigating pFN's response to therapeutic x-rays using similar biochemical techniques employed for collagen I, no significant changes were detected in its structure, electrophoretic mobility, nor thermal stability even at twice the therapeutic dose of x-rays. Comparing this with the results from collagen I, this supports the hypothesis that collagen I and pFN respond very differently on exposure to therapeutic x-rays. pFN's relative resistance to fragmentation may be attributed to its globular shape, which allows ROS susceptible hydrophobic residues, which are often buried within the protein, to avoid hydroxyl radicals produced from water radiolysis [105,205]. Oxidation of amino acid side chains that are on the surface of the protein could be less prone to drastically alter the overall protein shape, compared with oxidation of amino acids within hydrophobic regions [232,233]. In contrast, the elongated nature of collagen I monomers provides high solvent accessibility to its protein backbone and thus allows hydroxyl radicals to readily access and oxidise amino acids or the peptide backbone along the collagen chains. Hence, collagen I seems to be more readily damaged by ionising radiation compared to pFN. This however, appears to contradict previous studies on ultraviolet radiation (UVR), whereby aggregation was observed for FN exposed to 500mJ/cm^2

of UVB while collagen I was insensitive to the same amount of UVB [152] . Given the significantly higher energies of x-rays (about 100 times of UVB) and radiation dose used in this study compared to UVR, this was unexpected at first glance. However, considering that UVB has much lower probability of being absorbed by water [234] than 300keV x-rays [235] (approximately 30 times less by rough calculations, see appendix 7.4), there is a higher chance of UVB penetrating into the centre of globular domains in pFN, where it could be subsequently absorbed by chromophores. Oxidation of these chromophores within the hydrophobic core would have a more profound effect on the folding of the protein [233,236].

In contrast, x-rays are more likely to be preferentially absorbed by water molecules surround the protein (approximately 99.88% of all absorbed photons would be absorbed by water, see appendix 7.4 for calculations), thus producing hydroxyl radicals in the solvent surrounding the protein and oxidising amino acids at the surface. In fact, synchrotron x-rays have been used as a method for mapping the surfaces of proteins by looking oxidised side chains using water radiolysis [237,238].

The results from the peptide location fingerprinting (PLF) analysis further support the idea that the surface amino acids in pFN are likely damaged by therapeutic ionising radiation. The regions highlighted by PLF analysis of pFN encompassed numerous binding regions, such as for fibrin, heparin, fibulin and collagen, specifically after 100Gy of therapeutic x-ray exposure. These protein binding sites are often at the surface of the protein [218,239], and thus reinforces the idea that therapeutic x-rays are mainly indirectly oxidising and damaging amino acids on the surface, without drastically altering the overall protein folding of pFN nor causing fragmentation of the peptide backbone. Given that binding sites are sensitive to local conformation and chemistry, if PLF truly indicated an alteration in regional structure, then the impact of therapeutic radiation on pFN should encompass a change in binding affinity of pFN with other ECM proteins.

The data from the ELISA binding test support the PLF observations that structural changes in the collagen binding region affect function. This data seems promising but raises a conundrum – the MS1 difference flagged by PLF was strongly positive at the collagen binding region, indicating reduced proteolytic cleavage at the collagen binding region in radiation exposed pFN. With the absence of fragmentation of pFN detectable by SDS-PAGE, the reduction in peptide intensities may be attributed to crosslink formation, which should have reduced

collagen's ability to bind as well. How does this lead to increased collagen I binding to pFN? Crystallographic studies by Erat el. al. (2010) on FN's collagen binding region revealed that it comprises two regions (FNI_6-$FNII_{1-2}FNI_7$, and FNI_{8-9}) that are possibly acting synergistically [240], while work by Kubow et. al. (2015) showed that mechanical stretching of the FN molecule can abrogate collagen-FN binding, and thus postulated that the physical separation of the domains may be crucial for successful binding between collagen and FN [241]. The observed increased binding of x-ray exposed pFN to collagen could hence be mediated by improved stability in the distance between the domains by radiation induced crosslinking of residues between the domains. When compared to the native state of pFN where thermal motion (or Brownian motion) can perturb and vary the distances between the domains, the crosslinked binding site may hence facilitate improved collagen binding to FN. Further work needs to be done to verify the claim and better understand the mechanism underlying it, and to further test other binding proteins like fibrin and heparin to compare if their binding was similarly altered like for collagen.

3.4.2 Biological consequences of therapeutic x-rays on collagen I and fibronectin

While collagen I primary structure was compromised and fragmented, molecular ultrastructure of collagen I remain relatively unchanged, as seen in native-PAGE and CD. However, experiments here were conducted at low temperatures (below 10°C) to prevent protein degradation. *In vivo*, the higher body temperature would destabilise the collagen triple helix, especially for collagen monomers that have yet to form fibrils [222]. Studies have hypothesised there is a constant and reversible "micro-unfolding" of collagen triple helix *in vivo* at body temperatures [242–244], which, in x-ray exposed and fragmented collagen, may cause irreversible unfolding instead once fragments leave the triple helix. The release of such fragments can act as matrikines to influence cellular behaviour beyond its local region [158,245,246]. One such example is the proline-glycine-proline (PGP) fragment, which is an abundant peptide arrangement scattered throughout the sequence of collagen I α_1 and α_2 chains, and has been shown to act as a chemoattractant to neutrophils, possibly playing a role in neutrophilic inflammation in cystic fibrosis [247]. In addition to radiation-induced fragmentation, destabilisation of the collagen triple helix can also enhance MMP-1 mediated degradation of collagen I [160,248,249], thereby increasing matrikine production, and possibly reducing mechanical integrity of collagen fibrils. This could further influence

mechanosensitive pathways, such as cellular migration rates [250], or influence cell behaviour due to change in stiffness sensed by the cells [148,251].

The results for pFN suggests that therapeutic x-rays may impact the binding affinity of fibronectin to other proteins. Here, ELISA was used to demonstrate a statistically significant increase in collagen binding to pFN. *In vivo*, the effects of this increased binding is speculated to negatively impact collagen I fibrillogenesis [252]. On one hand, transitory FN binding to collagen I has been found to facilitate initiation of collagen I fibrillogenesis [216], while at the same time, high binding affinity of FN to collagen III adversely affects collagen III fibrillogenesis [253]. It was hence argued FN's transient and relatively weak binding to collagen I is essential to facilitate proper nucleation and subsequent fibrillogenesis. Increased binding affinity to collagen I may disrupt this delicate balance and tip FN binding to become adversarial to collagen I fibrillogenesis, which may be detrimental to the tissue mechanics and structural integrity over time.

While PLF is a powerful tool to interrogate changes to the intricate, local structure within a protein's overall shape, it lacks the ability to pinpoint the actual location of radiation damage within the protein. The change in proteolytic degradation in a protein region can be caused by radiation damage to regions far away in the primary sequence but spatially nearby when folded, such as in globular proteins like FN, or damage to a secondary chain bound together in a multimer, such as in collagen I. This makes predicting the location of x-ray damage to proteins non-trivial; it does not simply correlate with regions of highly abundant ROS sensitive amino acids seen from our results. Further work is needed to incorporate three-dimensional configuration of proteins to visually investigate not only regions flagged by PLF, but other close-by regions that may impact trypsin's accessibility as well. This may be done by incorporating protein structure databases, such as AlphaFold, which is a contemporary machine learning-enabled predictive tool for whole protein structures.

This study is also limited by the idealistic nature of purified samples. Given that the collagen I used in this study is pepsin treated and acid-solubilised, the radiation impact observed may not be replicated *in vivo* where it is insoluble, crosslinked, and bound to other proteins. Furthermore, pFN in its soluble form is not fully representative of its structure in its active form when recruited into the fibrin clot, where it would undergo conformational changes into an extended form due to binding and crosslinking with fibrin [209,254]. Given the change in

configuration of pFN when bound [255], the targeted regions of therapeutic x-rays could be completely distinct to those identified here in this study. The following chapter thus attempts to examine the effects of therapeutic x-rays on collagen I in the fibrillar state, which is more representative of the state of collagen I *in vivo.*

3.5 Conclusion

In conclusion, the results here have shown collagen I and FN, both crucial ECM proteins that have different structural profile, are both affected by exposure to X-rays although these effects require a more sensitive method – mass spectrometry – to detect and analyse. Collagen I chains, both α_1 and α_2, fragment while FN do not. Although the ultrastructure and thermal stability of both collagen I and FN appear unchanged, peptide location fingerprinting analysis was able to detect regional structural changes in both ECM molecules. Consequently, therapeutic x-rays exposure may destabilise collagen I at body temperatures or may enhance degradation of collagen I by tissue proteases. Exposure of pFN to therapeutic x-ray doses may alter its binding to ECM proteins and cells.

4 Therapeutic x-ray doses affects structure of native, fibrillar collagen I

4.1 Introduction

4.1.1 Protein structure and radiation damage

Whilst it is well established that individual amino acid residues differ in their relative susceptibility to oxidative damage [256], the impact of higher order protein structure (2°- 4°) on ROS susceptibility is difficult to predict. Numerous ECM proteins are often arranged into large, macro-molecular assemblies or multimers, such as in elastic fibres and collagen fibrils [257,258]. Such hierarchical assembly may contribute to local resistance or susceptibility to radiation damage within domains [204]. For example, regions of a protein which are protected from the aqueous environment due to protein folding (such as hydrophobic domains) are less likely to be affected by ROS produced from water radiolysis [237]. The opposite is also likely true whereby more hydrophilic amino acids or those at the surface of protein domains are likely to be damage by water radiolysis-induced ROS [259]. As was shown in the previous chapter, globular plasma fibronectin (pFN) exposed to therapeutic x-rays domains were preferentially damaged near binding sites, which are often at the surface of proteins. In contrast, impact of therapeutic x-ray on solubilised monomeric collagen I appear to be randomly distributed throughout the protein sequence. However, collagen I *in vivo* has greater structural complexity [207,231,260], making the solubilised collagen study limited in scope and biological significance. Hence, to study how the ultrastructure of collagen I may alter its response to therapeutic x-rays, three different forms of rat collagen I are employed: acid-solubilised triple helical monomers (hereafter referred to as 'solubilised collagen'), randomly aligned fibrillar collagen in reconstituted gel under neutral pH (hereafter referred to as 'gel'), and linearly aligned fibres in *ex vivo* rat-tail tendons (hereafter referred to as 'tendon'). In this chapter, rat collagen I was chosen due to sample availability of *ex vivo* tendons and the lower cost associated in producing collagen gels due to the high concentrations needed. The differences in collagen structure between these three forms are illustrated in the section below.

4.1.2 Collagen I assembly

4.1.2.1 Solubilised monomeric collagen I

Collagen I, in its monomeric form, comprises three polypeptide chains of two kinds – α_1 and α_2. Similar to human collagen described in the previous section, the polypeptide chains of rat tail collagen I contains repeating -Gly-X-Y- units along the triple helical regions (Figure 4.1). The monomer consists of two α_1 chains and one α_2 chain ($[\alpha_1]_2[\alpha_2]_1$). Collagen I remains in this monomeric state when in low pH [261,262]. Solubilised monomeric collagen I can be obtained through incubating collagen-rich tissues in acetic acid or through pepsin digestion (at low pH) to remove the C-terminus telopeptides [263–265]. In our study, the solubilised collagen purchased from Gibco (A1048301) was extracted by pepsin digestion and kept in low pH (20mM acetic acid) to maintain its monomeric form.

4.1.2.2 Collagen I gel

When pH of the collagen monomers' environment is raised to pH 6.0-7.2, self-assembly is energetically favoured and the monomers interact through hydrogen bonding to form fibrils [208,262]. Collagen monomers self-associates into a quasi-hexagonal configuration, forming the collagen fibril that exhibits a characteristic 65-67nm D-banding [260,266] (Figure 4.1).

In vitro, this assembly process is difficult to control, resulting in randomly arranged fibrils of varying diameters. For this study, collagen gel was produced by diluting solubilised collagen with PBS and neutralising with NaOH, followed by incubating at 37°C for 2 hours to speed up the process of fibril assembly [267].

4.1.2.3 Collagen in rat tail tendon

In vivo, collagen fibrils can form into highly aligned fibres in tissues such as in tendon and (to a lesser extent breast), due to the strict cellular regulation and production of collagen and its fibrillogenesis [268]. To outline this process briefly, the cells first produce and deposit procollagens in the extracellular space; these procollagens are similar to the monomeric collagens but further contain globular domains on the C- and N- terminal that facilitates formation of triple helices but prevents fibril assembly. Once in a triple helical form, the procollagen C- and N-proteinases cleave the C- and N-propeptides respectively, leaving telopeptides on either ends for initiation of assembly of collagen fibrils. Lysyl oxidases (LOX) then covalently crosslinks adjacent monomers through lysine and hydroxylysines within the

telopeptides. LOX plays a crucial role in helping the collagen fibrils form highly aligned and uniform fibres in tendons, which come together to form fascicles [269,270]. While rat tail tendons may not be completely representative of the structure of fibrillar collagen I in other animal or human tissues [271], it was chosen as our model given its well-studied structure [266,269,272,273], high collagen I content [274], and close homology with human collagen I (α_1: 92.77%, α_2: 91.14% similarity in sequence).

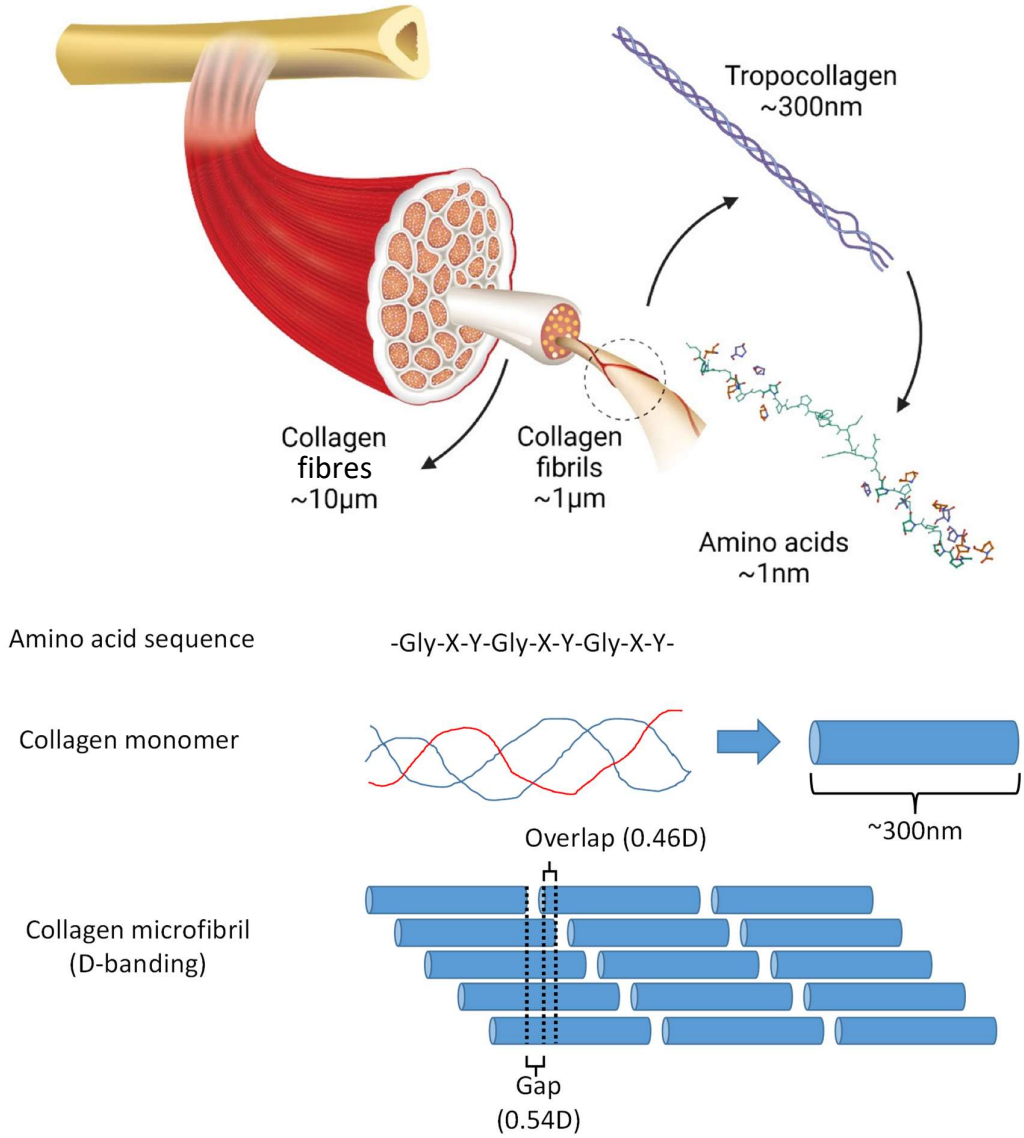


Figure 4.1. Collagen I structure at different scales. The fundamental amino acid sequence of a collagen chain consists of Gly-X-Y repeats, where X/Y represents other amino acids, commonly hydroxyproline or proline. The collagen monomer comprises three collagen chains, which is made of two α_1 and one α_2 chain. This monomer arranges itself in a quasi-hexagonal manner to form a microfibril with characteristic D-banding of approximately 67nm due to the overlap and gaps of staggered collagen monomers alignment in a fibril. Many fibrils come together to form collagen fibres, in which crosslinking by lysyl oxidase can occur and are essential in tendons. Top figure adapted from Amirrah, I.N. (2022) [419].

4.2 Aims and hypothesis

The different forms of rat collagen I, namely the solubilised monomer, reconstituted gel, and tendons, are made of similar chemical composition, differing only in physical properties and molecular structure. By exposing the three systems to the same radiation dose and using consistent biomolecular techniques applicable to all three forms, this study aims to investigate how increasingly complex rat collagen I forms respond differently to therapeutic x-rays doses. This would help address the second hypothesis of the thesis – that the effects of therapeutic x-rays is dependent on the state of assembly of collagen I. This is split into the following sub-hypotheses:

1. Therapeutic X-ray doses will affect the structure of collagen I regardless of assembly state.

However:

2. The nature of X-ray induced damage will depend on assembly state.

To test these hypotheses the first aim was to characterise the structure adopted by collagen I in reconstituted gel and native tendon samples. Neutralising the pH of solubilised collagen should induce self-assembly of fibrils with similar structures to the fibrils found in native tendon [257,265]. The presence of fibrillar collagen in gel and tendon was confirmed using atomic force microscopy (AFM). To investigate the impact of therapeutic x-rays on solubilised, gel, and tendon collagen, biochemical techniques were employed to characterise different possible aspects of radiation damage. SDS-PAGE was used to observe fragmentation of protein backbone while acidic native-PAGE (AN-PAGE) and circular dichroism (CD) was conducted to investigate changes in native protein structure. Finally, LC-MS/MS (with peptide location fingerprinting (PLF) analysis) is utilised to look for changes to the collagen peptide fingerprint profile and therefore infer possible changes to local molecular structure. A general workflow of processing the various samples is depicted in the figure below (Figure 4.2)

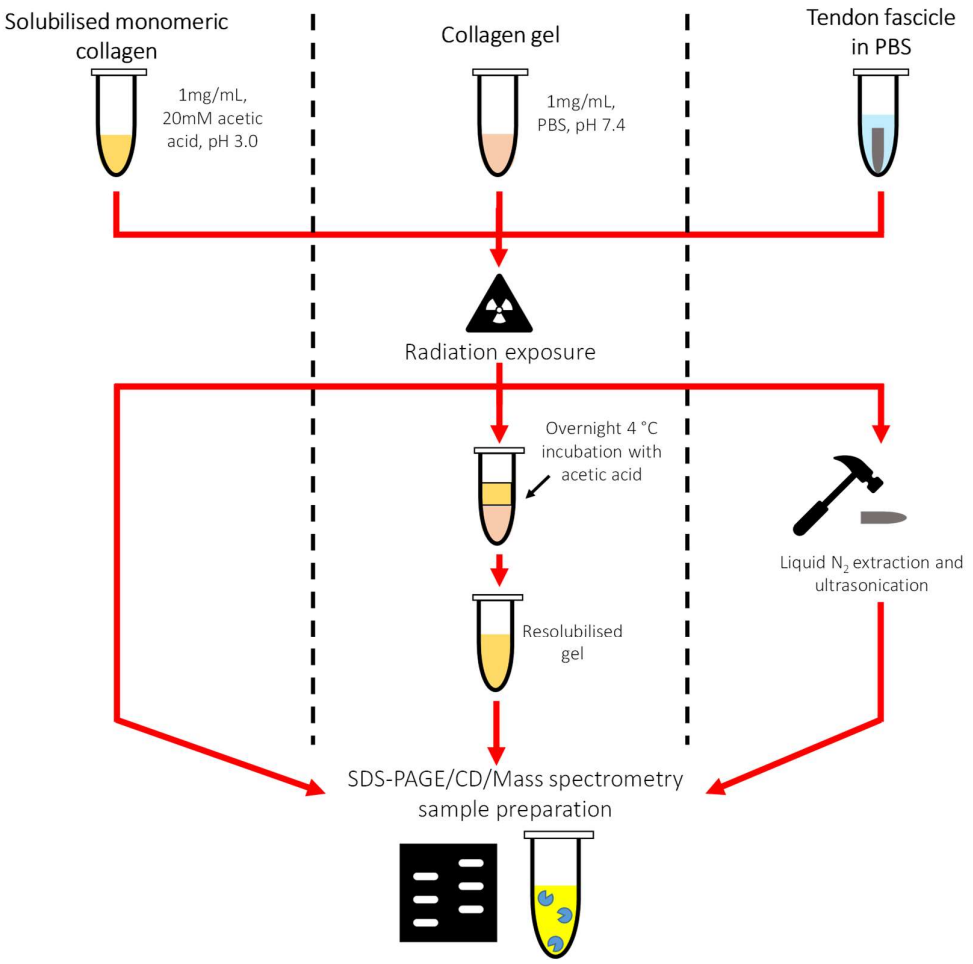


Figure 4.2. Experimental workflow for different collagen forms. Solubilised monomeric collagen was irradiated in acetic acid with no further processing. Collagen gels were made from neutralising solubilised collagen. Following x-ray exposure, gels were incubated overnight in acetic acid to allow for re-solubilisation. Tendon fascicles were exposed to therapeutic x-rays in PBS+, following which collagen was extracted from the tendons by liquid N_2 crushing, resuspending with a dounce homogeniser, then ultrasonicated. Following this workflow, all samples were processed in the same manner for SDS-PAGE and mass spectrometry.

4.3 Results

4.3.1 Atomic force microscopy revealed randomly aligned fibrillar collagen in reconstituted gels and tendons

AFM imaging of collagen gel and tendons revealed characteristic collagen fibril dark and light patterning, which is attributed to the D-banding in both the collagen gel and in tendons (Figure 4.3). 2D fast Fourier transform (2D-FFT) was performed on the AFM images to confirm the presence of D-banding. This technique, in brief, decomposes the original image into a summation of sinusoidal functions, whose frequencies are mapped onto phase space – each point in phase space represents a certain frequency of a sinusoidal wave function, and the intensity of the pixel at that point represents its amplitude. The frequency increases as you move away radially from the centre (zero at centre).

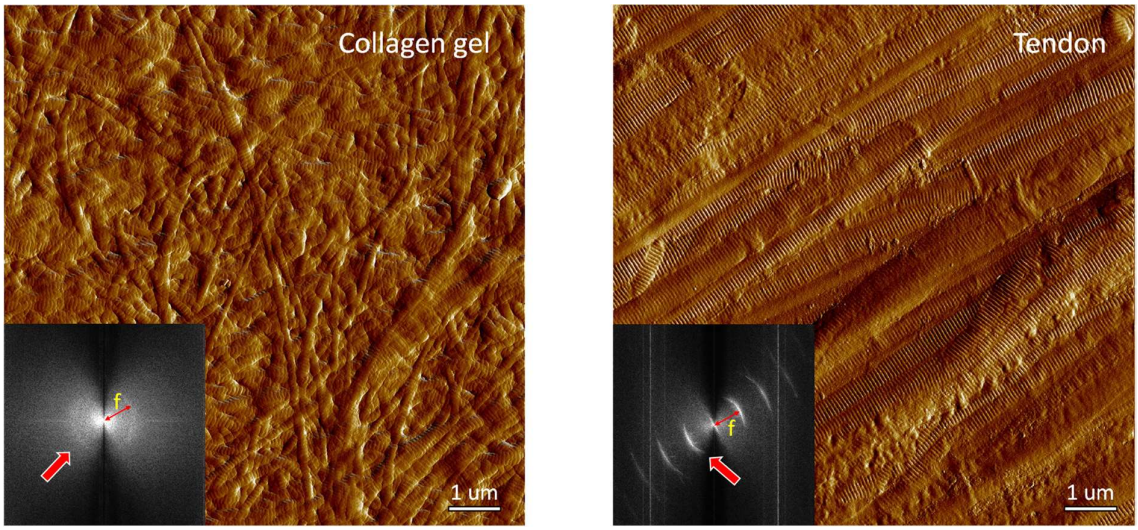


Figure 4.3. AFM peak error images of collagen gels and tendons with the corresponding 2D fast fourier transform (2D-FFT) images (insert). Fibrillar collagen was confirmed in gel samples by the observing and measuring the characteristic D-banding of collagen fibrils in FFT of original images (red arrows). Frequency was measured by taking distance from centre to edge of banding pattern, f, in triplicates. Fibrils in collagen gel appear randomly aligned while those in rat tail tendons were aligned.

For collagen gel, a circular imprint was observed in the FFT image, indicating a presence of a periodic pattern in our original image with a frequency 15.7 ± 0.4μm^{-1} (measured radius from centre in triplicates). Inverting this, the spacing of this periodic pattern corresponded to 64 ± 2nm. The same measurements were performed for FFT image of tendons, where directional stripes were observed instead, to obtain a frequency of 15.2 ± 0.1μm^{-1} and a periodicity of 65.7 ± 0.5nm. Both periodicities obtained were within the range for the D-banding of collagen fibrils found in literature [271], thus provides good evidence of the presence of fibrils in gel and tendons. The relatively circular pattern in gel implies that the D-bands have no specific

direction and is randomly aligned, while in tendons, the stripes in the FFT image were directional implying that the D-bands were highly aligned.

4.3.2 X-ray exposure induced fragmentation of primary structure for solubilised collagen I, but not collagen in gel or tendons

Reducing SDS-PAGE for all rat tail collagen I exhibited three regions of protein bands (as was observed for human collagen from the previous chapter). These were the α monomers, β dimers, and γ trimers (Figure 4.4). All three bands were present in solubilised, gel and tendon collagen, although the γ region for tendons appeared to encompass a more complex mixture of populations of high molecular weight proteins. With exposure to 50Gy and 100Gy of therapeutic x-rays, solubilised collagen exhibited a 170% (50Gy, p=0.004) and 210% (100Gy, p=0.002) increase in staining intensity respectively in the region between the β and α bands (Fr1). Similarly, there was an also increase in intensity of 170% (50Gy, p=0.007) and 270% (100Gy, p=0.04) in the region below the α bands (Fr2) (Figure 4.4a). This implied possible fragmentation of the primary structure of α_1 and α_2 chains. This fragmentation signature was not observed in x-ray exposed collagen gel and tendons (Figure 4.4b, c).

SDS-PAGE suggests that solubilised monomeric collagen but not fibrillar collagen (like in gel and tendons) are susceptible to therapeutic x-rays. That said, SDS-PAGE is disruptive to the native triple helical structure of collagen. Thus, native-PAGE and CD was further employed to investigate x-ray impact on the native structure of collagen.

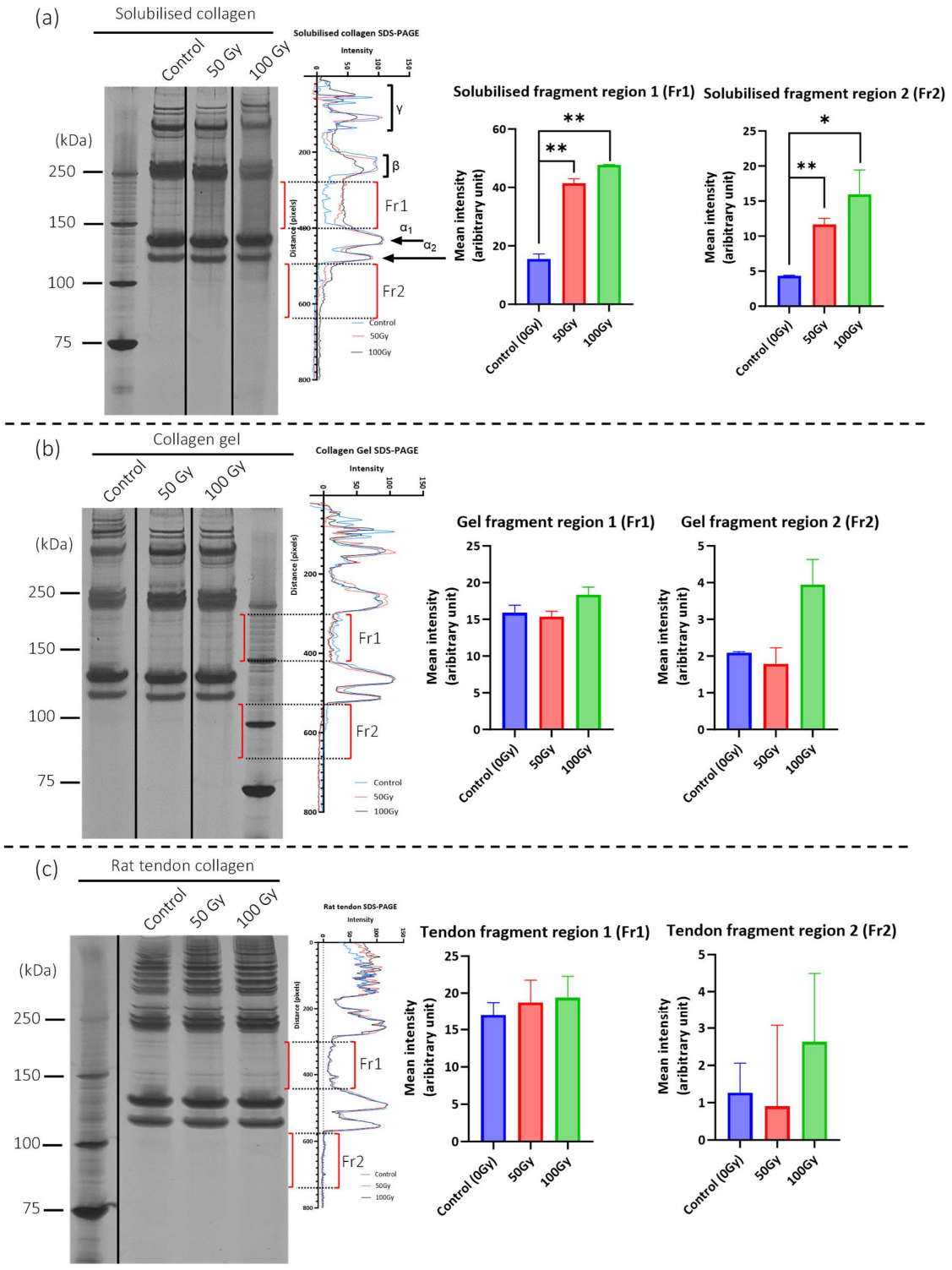


Figure 4.4. SDS-PAGE for rat collagen exposed to therapeutic x-rays (50Gy and 100Gy) in three different forms: solubilised, gel, and in ex vivo tendons (n=3). An intensity profile for each lane was obtained for all samples and an average was taken amongst replicates. Additionally, the mean intensity between the β and α bands (Fragment region 1, Fr1) and below the α bands (Fragment region 2, Fr2) were obtained and compared between control and x-ray exposed samples. (a) In solubilised collagen, a 170% (50Gy, p=0.004) and 210% (100Gy, p=0.002) increase in staining intensity was observed in Fr1. Similarly, there was an increase in intensity of 170% (50Gy, p=0.007) and 270% (100Gy, p=0.04) for Fr2. This smearing implied therapeutic x-rays induced fragmentation of the primary backbone of collagen α_1 and α_2 chains in solubilised collagen. (b/c) However, in gels and tendons, no statistically significant changes in intensities were observed in neither Fr1 nor Fr2.

4.3.3 Triple helix and shape of tertiary structure of collagen I remains unchanged

Acidic native-PAGE could not detect any change in electrophoretic mobility of collagen I, suggesting that the triple helix remains intact after therapeutic x-rays exposure in all forms of collagen I (n=2) (Figure 4.5). Smearing was observed for x-ray exposed solubilised collagen but not in the control sample. For gel and tendons, both control and x-ray exposed samples appear to exhibit this smearing pattern. The smearing pattern could indicate presence of small groups of monomers aggregating/self-associating. In the case for gel and tendons, it could be crosslinked collagen monomers that were not fully dissociated.

CD was also used to investigate the impact of irradiation on the triple helical structure of collagen I by characterising changes to the 222nm absorption peak. No observable change was seen for all forms of rat tail collagen after exposure to either 50Gy or 100Gy of therapeutic x-rays.

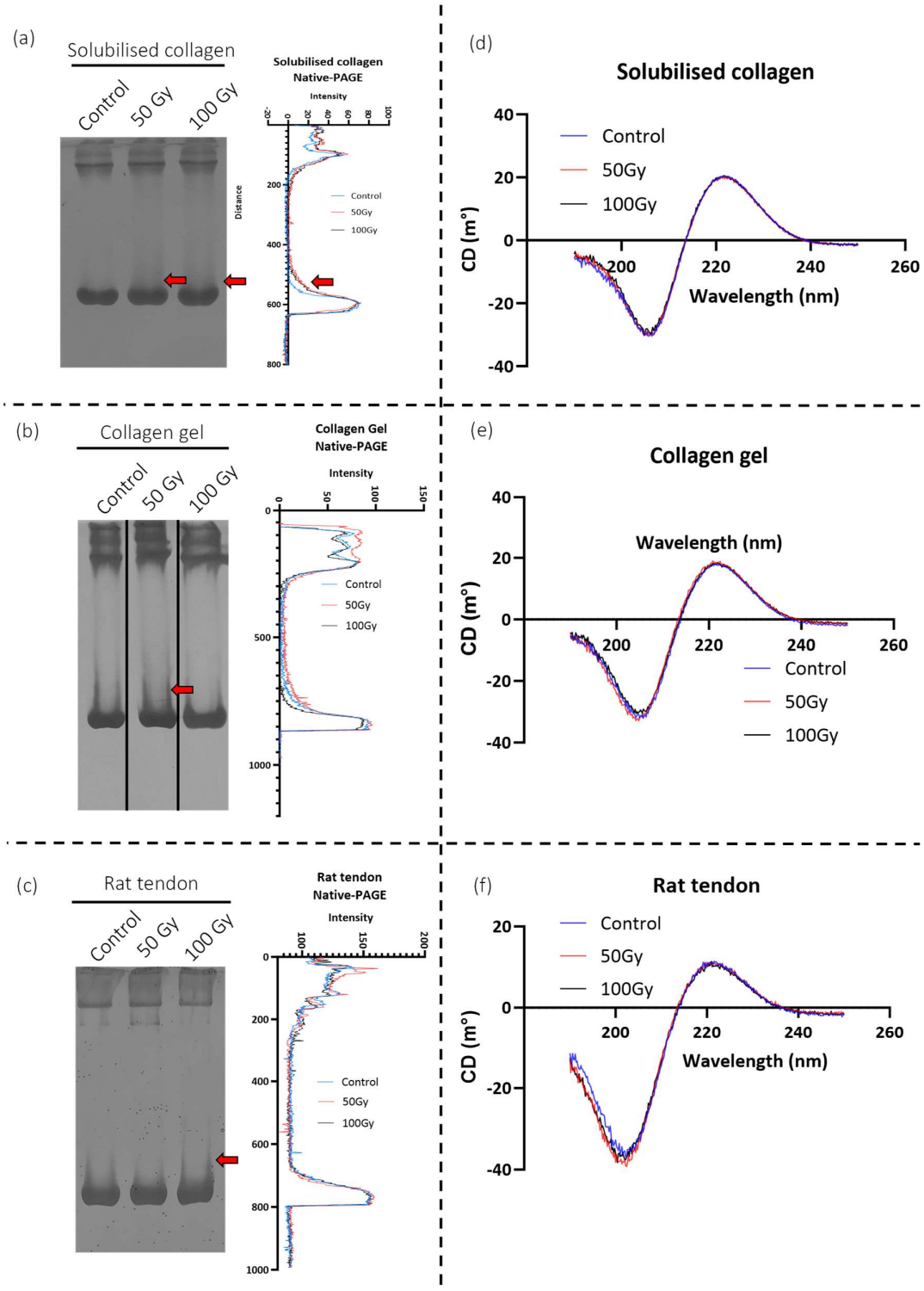


Figure 4.5. Acidic native-PAGE and circular dichroism spectrum for radiation exposed rat collagen I in different forms. (a/b/c) Acidic native-PAGE for all three forms of rat collagen I (solubilised (a), gel (b), and ex vivo tendons (c)) exposed to 50Gy and 100Gy of therapeutic x-rays (n=2). No change in electrophoretic mobility was observed, implying no significant change in overall structure of collagen monomers. Smearing was observed for x-ray exposed solubilised collagen but not in the control sample (highlighted by red arrows). For gel and tendons, both control and x-ray exposed samples appear to exhibit this smearing pattern.

(d/e/f) Circular dichroism spectrum where the 222 nm peak corresponds to the triple helical structure of collagen, remain unchanged by x-ray exposure in all three collagen forms. This implies the triple helix and tertiary structure of collagen is insensitive to therapeutic radiation exposure up to 100Gy.

4.3.4 Peptide location fingerprinting identified more structurally affected regions in gel and tendons than solubilised collagen

LC-MS/MS was employed to analyse regional changes to proteolytic susceptibility of collagen I in the solubilised, gel, and tendon samples. Similar to chapter 3, the peptide MS1 intensities obtained from mass spectrometry were mapped onto the original collagen I sequence, producing an intensity pattern that termed a peptide fingerprint. This highly reproducible fingerprint is used for peptide location fingerprinting (PLF) analysis by comparing the peptide intensity at each region between control and radiation exposed samples. Regions where the difference in peptide intensities were statistically significant ($p<0.05$) indicated changes to proteolytic susceptibility and are flagged for the α_1 (Figure 4.6) and α_2 chains (Figure 4.7). A greater number of flagged regions were detected for gel (16) and tendon (20) samples compared to solubilised collagen (3) in both collagen alpha chains after exposure to 50/100Gy of radiation (Table 4.1). Additionally, at 50Gy exposure, the α_1 chain appear to have consistently higher number of flagged regions than α_2. All 100Gy samples had fewer marked regions than the 50Gy samples within their respective collagen forms.

In all, PLF data suggests that therapeutic x-rays have profound impact on the proteolytic susceptibility of all three forms of collagen, although the impact seems to be more apparent in gel and tendons than in solubilised collagen.

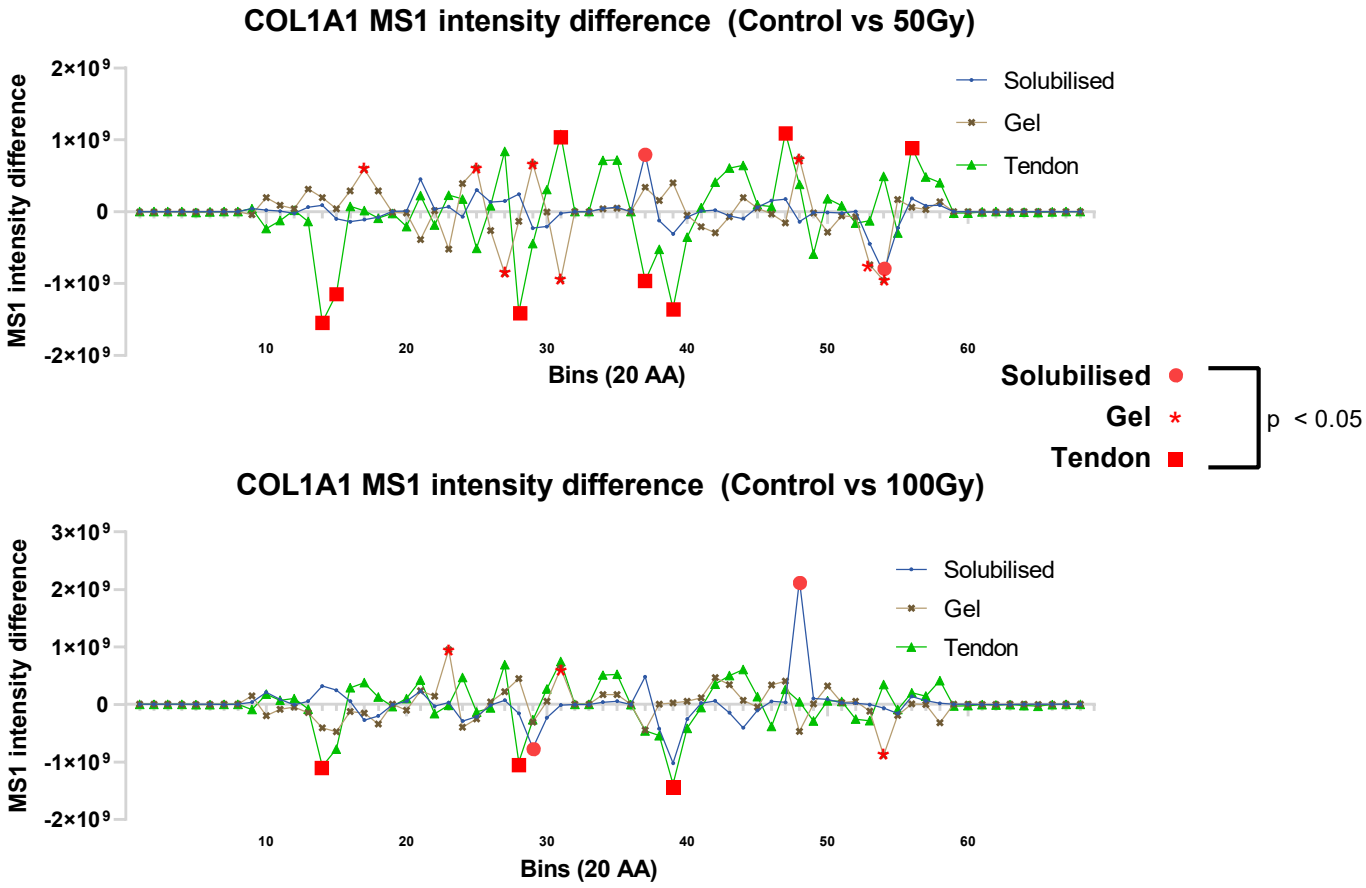


Figure 4.6. MS1 intensity difference of peptide fingerprint for collagen I α1 chain between control and 50Gy (top) or 100Gy (bottom) for all three forms of collagen I. Regions with significantly altered MS1 intensities ($p<0.05$) were marked with a corresponding red icon.

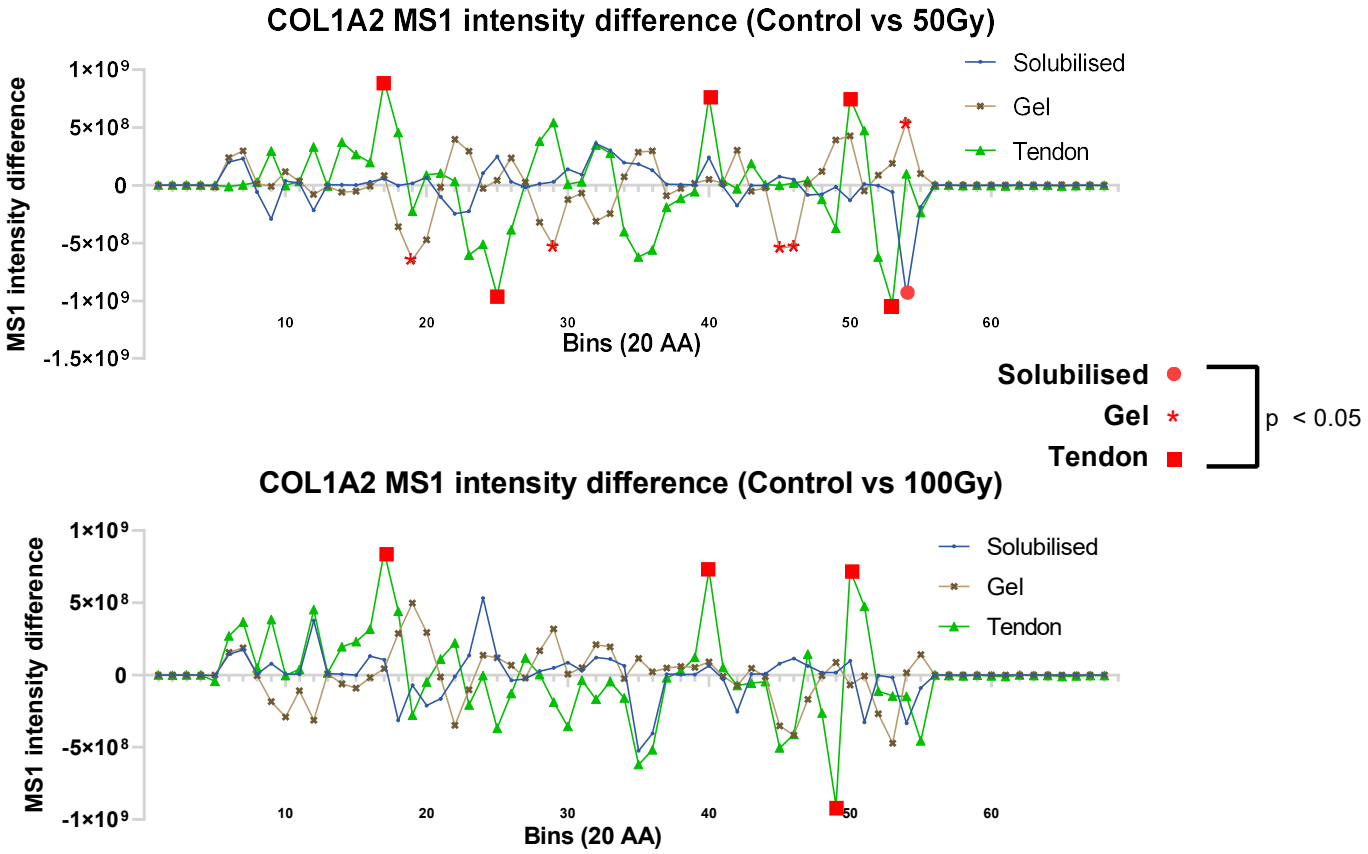


Figure 4.7. MS1 intensity difference of peptide fingerprint for collagen I α2 chain between control and 50Gy (top) or 100Gy (bottom) for all three forms of collagen I. Regions with significantly altered MS1 intensities ($p<0.05$) were marked with a corresponding red icon.

Table 4.1. Number of regions with statistically significant altered peptide intensities flagged by peptide location fingerprinting. There are drastically more in gels and tendons compared to solubilised collagen. Further, the number of regions seem to decrease with higher dose (100Gy) compared to 50Gy.

	Solubilised		**Gel**		**Tendons**	
Dose	**COL1A1**	**COL1A2**	**COL1A1**	**COL1A2**	**COL1A1**	**COL1A2**
50Gy	2	1	8	5	8	5
100Gy	0	0	3	0	3	4

4.3.5 Peptide fingerprint of solubilised collagen is distinct from gel and tendons

Further investigation into the peptide fingerprint of control, unirradiated samples for all three collagen forms revealed that the solubilised collagen produced a distinct peptide fingerprint that differed from gel and tendons, while the gel and tendons had very similar fingerprints (raw data in appendix 7.2). To quantify this, the fingerprint data of control samples for all three collagen forms were min-max normalised. Bins in which % MS1 intensity of solubilised collagen was significantly distinct from both gel and tendon (t-test, $p<0.05$), either significantly higher (S > G/T, black arrows) or lower (S < G/T red arrows), are highlighted (Figure 4.8).

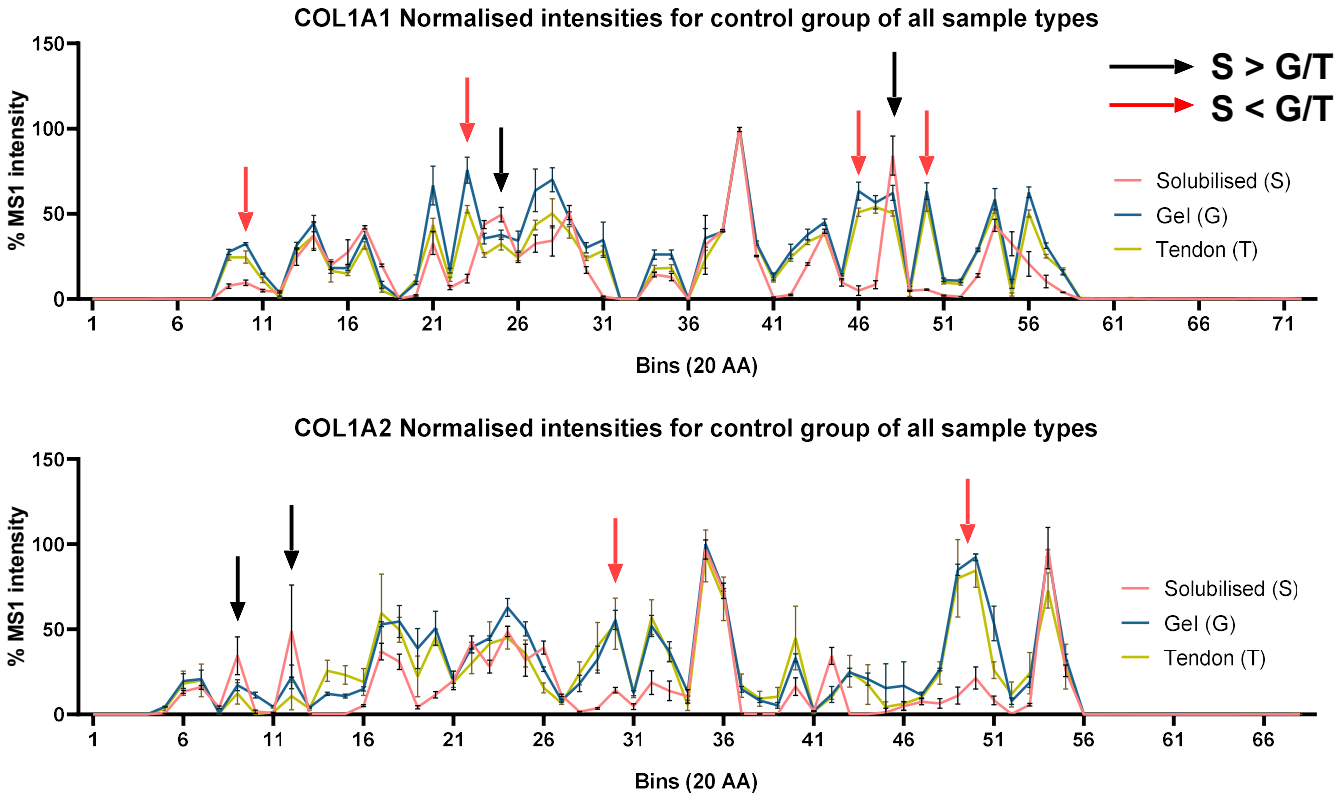
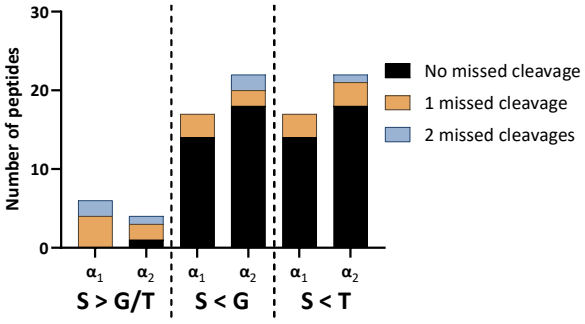


Figure 4.8. Peptide fingerprint for control group of the three collagen forms are normalised to the maximum ion intensity for each replicate (percentage of maximum ion intensity). While the fingerprint pattern for gel and tendons were very similar, the fingerprint of solubilised collagen significantly differed from both the gel and tendons. Regions with significant differences ($p<0.05$) between solubilised collagen and both gel and tendons were flagged for further investigation (arrows). Black arrows indicated where solubilised collagen had higher intensities while red arrows indicated regions of where solubilised collagen had lower intensities than both gel and tendon. Arrows on graph indicate selected examples and is not exhaustive – full comprehensive regions/bins are tabulated in Figure 4.9.

At each of those regions where there was significant difference between % MS1 intensity of solubilised collagen to gel/tendons, the peptide sequence contributing the highest MS1 intensity was noted down and listed in the table below (Figure 4.9). A pattern was discovered at those regions – bins where solubilised collagen had greater peptide intensity than gel or tendons (S>G/T), had a large proportion of peptides with 1 or 2 missed tryptic cleavage sites (or missed cleavages) (7/8 peptides, 87.5%). Vice versa, regions where solubilised collagen had lower peptide intensity than gel or tendons (S<G/T) were characterised by a large proportion of peptides with no missed cleavage (32/39, 82%) (Figure 4.9).

Classification of peptides in significantly different bins

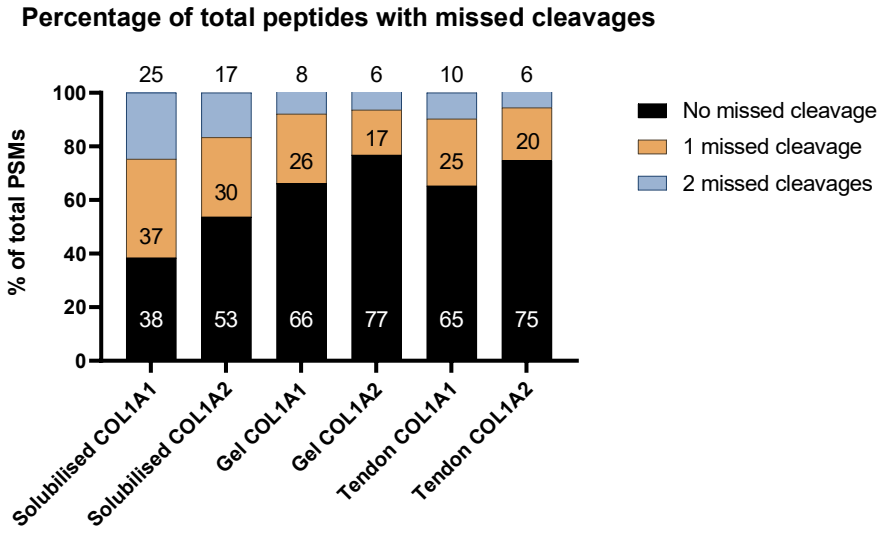


COL1A1	Bin	Peptide with greatest MS1 intensity contribution	Missed cleavages
Solubilised > Gel/tendons	16	GLPGERGRPGPPGSAGAR	1
	18	GNDGAVGAAGPPGPTGPTGPPGFPGAAGAKGEAGPQGAR	1
	24	GARGEPGPSGLPGPPGERGGPGSR	2
	25	GFPGADGVAGPKGPAGER	1
Gel/Tendon > Solubilised	9	SAGVSVPGPMGPSGPR	0
	10	GLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPR	0
	23	GEPGPAGVQGPPGPAGEEGKR	1
	27	GLTGSPGSPGPDGK	0
	28	TGPPGPAGQDGRPGPAGPPGAR	0
	31	DGEAGAQGAPGPAGPAGER	0
	41	GEPGPPGPAGFAGPPGADGQPGAK	0
	42	GDAGPPGPAGPAGPPGPIGNVGAPGPK	0
	43	GAAGPPGATGFPGAAGR	0
	46	GETGPAGRPGEVGPPGPPGPAGEK	0
	47	GSPGADGPAGSPGTPGPQGIAGQR	0
	50	GPPGPMGPPGLAGPPGESGR	0
	53	NGDRGETGPAGPAGPIGPAGAR	1
	54	NGDRGETGPAGPAGPIGPAGAR	1
	56	GFSGLQGPPGSPGSPGEQGPSGASGPAGPR	0
	57	GFSGLQGPPGSPGSPGEQGPSGASGPAGPR	0
	58	DGLNGLPGPIGPPGPR	0

COL1A2	Bin	Peptide with greatest MS1 intensity contribution	Missed cleavages
Solubilised > Gel/tendons	9	GVVGPQGAR	0
	12	GRVGAPGPAGAR	1
	26	GPSGDPGKPGEKGHPGLAGAR	1
	42	EGIRGPRGDQGPVGR	2
Gel/Tendon > Solubilised	14	GELGPVGNPGPAGPAGPR	0
	15	GELGPVGNPGPAGPAGPR	0
	16	GATGLPGVAGAPGLPGPR	0
	17	GATGLPGVAGAPGLPGPR	0
	18	GIPGPVGAAGATGPR	0
	19	GETGNKGEPGSAGAQGPPGPSGEEGKR	2
	20	GSPGEPGSAGPAGPPGLR	0
	23	GPNGDAGRPGEPGLMGPR	0
	28	GEQGPAGPPGFQGLPGPSGTAGEVGKPGER	0
	29	GEQGPAGPPGFQGLPGPSGTAGEVGKPGER	0
	30	GLPGEFGLPGPAGPR	0
	32	GPSGAPGPDGNKGEAGAVGAPGSAGASGPGGLPGER	1
	33	GPSGAPGPDGNKGEAGAVGAPGSAGASGPGGLPGER	1
	37	GEVGPAGPNGFAGPAGSAGQPGAKGEKGTK (GEVGPAGPNGFAGPAGSAGQPGAKGEK)	2 (1)
	40	GDGGPPGMTGFPGAAGR	0
	43	TGEIGASGPPGFAGEK	0
	44	GPSGEPGTTGPPGTAGPQGLLGAPGILGLPGSR	0
	48	GYPGNIGPTGAAGAPGPHGSVGPAGK	0
	49	GYPGNIGPTGAAGAPGPHGSVGPAGK	0
	50	GEPGPAGSVGPVGAVGPR	0
	51	GPSGPQGIR	0
	53	GHNGLQGLPGLAGLHGDQGAPGPVGPAGPR	0

Figure 4.9. Comparison of peptides that contributed to highest proportion of ion intensities within bins that were significantly different between the control, unirradiated samples of solubilised collagen and both gel/tendons. For bins where solubilised collagen had higher normalised intensities (S>G/T), 7 out of 8 peptides (87.5%) were found to contain 1 or 2 missed cleavages. On the contrary, in bins where solubilised collagen had lower normalised intensities (S<G/T), only 7 out of 39 peptides (18%) had missed cleavage sites, while the rest had no missed cleavage sites.

To assess if this pattern of missed cleavages is persistent throughout the entire protein length for the different collagen forms, comparison was made between the complete group of identified peptides from solubilised, gel, and tendon collagen by sorting the peptides based on the number of missed cleavages. For each peptide, their peptide spectrum matches (PSMs) across all replicates and treatment groups was used as a measure of the overall peptide contribution. Within each collagen form, the total PSM and the raw PSM of peptides with 0, 1, or 2 missed cleavages was normalised into a percentage to compare between different collagen forms (Figure 4.10).



	Solubilised		Gel		Tendons	
	COL1A1	**COL1A2**	**COL1A1**	**COL1A2**	**COL1A1**	**COL1A2**
0 missed cleavages	4918 **(38.4%)**	3843 **(53.7%)**	19132 **(66.2%)**	13974 **(76.8%)**	17637 **(65.3%)**	13305 **(74.8%)**
1 missed cleavage	4716 **(36.9%)**	2123 **(29.7%)**	7483 **(25.9%)**	3063 **(16.8%)**	6753 **(25.0%)**	3492 **(19.6%)**
2 missed cleavages	3163 **(29.7%)**	1189 **(16.6%)**	2282 **(7.9%)**	1159 **(6.4%)**	2629 **(9.7%)**	992 **(5.6%)**

Figure 4.10. Total peptide spectrum matches (PSMs) for both α1 (COL1A1) and α2 (COL1A2) chains categorised into peptides with 0, 1, or 2 missed cleavages. Solubilised collagen had a higher number of peptides with 2 missed cleavage sites for both α1 and α2 compared to gel and tendons. In contrast, the number of peptides with 0 or 1 missed cleavage sites were significantly lower in solubilised collagen than in gel and tendons. Table below shows the exact number of PSMs and the corresponding percentage (based on the total number of PSMs for the respective collagen chain) detected by the mass spectrometer for each collagen form categorised by the number of missed cleavages.

Solubilised collagen were found to contain 66.6% and 46.3% of peptides with at least one missed cleavage for the α_1 and α_2 chains respectively, which is much higher than those of gel (α_1: 33.8%, α_2: 23.2%) and tendon (α_1: 34.7%, α_2: 25.2%) collagen chains. Consequently, the percentage of peptides without any missed cleavages were lower for the solubilised collagen chains at 38.4% for α_1 and 53.7% for α_2, compared to gel (α_1: 66.2%, α_2: 76.8%) and tendon collagen (α_1: 65.3%, α_2: 74.8%).

Overall, solubilised collagen appears to yield a greater proportion of peptides with missed cleavages while contrastingly gel and tendon collagens generate a larger fraction of peptides without missed cleavages. This difference in peptide populations between the solubilised collagen and gel and tendon collagen suggests that the peptide fingerprint variations are correlated with the disparities in effectiveness of tryptic digestion, specifically the generation of peptides with or without missed cleavages.

4.3.6 Collagen gel and tendon stiffness undetectable change by AFM

As LC-MS/MS analysis combined with PLF demonstrated that fibrillar collagen (in gel and tendons) undergoes some structural remodelling as a consequence of x-rays exposure, an attempt was made to investigate possible alterations in the mechanical integrity of collagen gel and tendons by AFM (Figure 4.11) (n=3). For each sample, three 15µm x 15µm regions were randomly chosen to be mechanically tested. At each of those regions, 225 evenly distributed force curves were obtained before and after x-ray exposure. Substantial variation in stiffness was observed for the control sample within the same area probed despite lack of x-ray exposure. As the variation in control samples were of similar magnitudes to the variations x-ray exposed samples, the results are inconclusive.

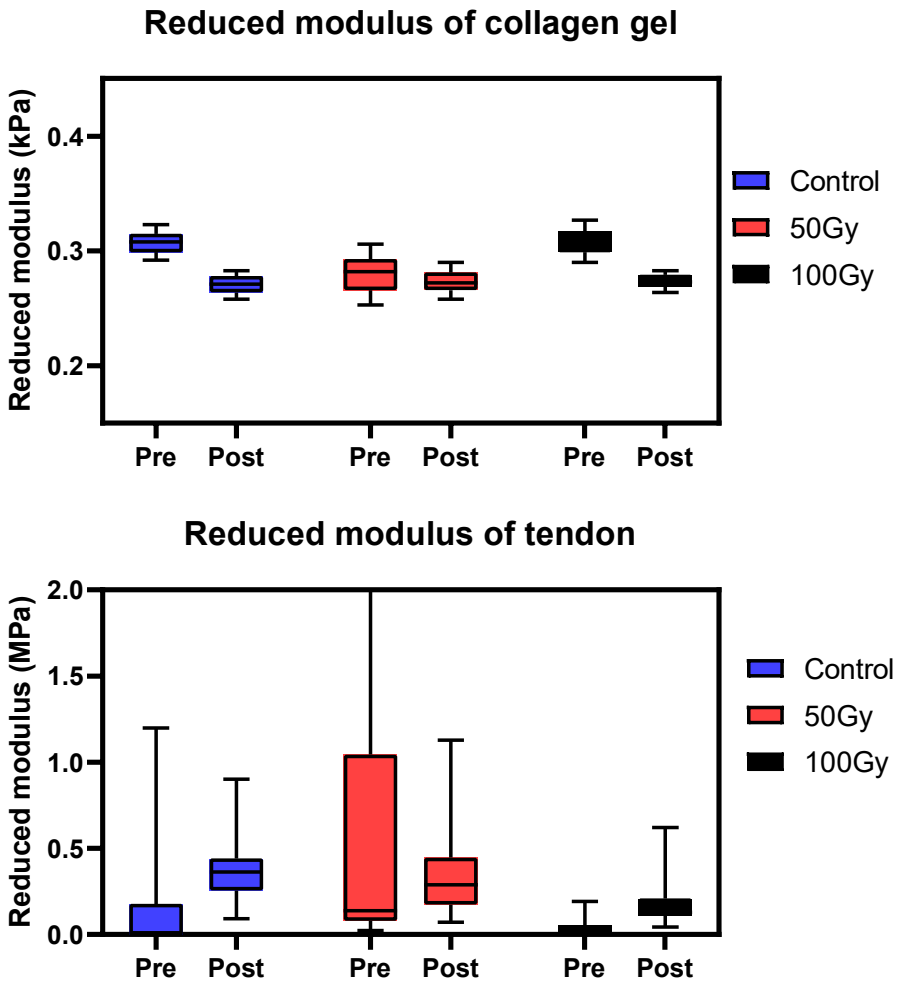


Figure 4.11. Reduced modulus of collagen gels and tendons obtained from atomic force microscopy (AFM) as proxy for relative stiffness. Due to the large variation in stiffness (reduced modulus) of the control samples, results remain inconclusive on whether exposure to 50/100Gy of therapeutic radiation alters the microscopic stiffness of both collagen gel and tendons.

4.4 Discussion

4.4.1 Differential response of solubilised, gel, and tendon collagen to therapeutic x-rays

In this study, solubilised monomeric collagen, reconstituted collagen gel, and collagen in rat tail tendons, all of which contain the same rat collagen I sequence but with different structural complexities, were investigated for their responses to therapeutic x-rays using numerous biochemical methods. SDS-PAGE of solubilised collagen shows clear fragmentation of the protein primary backbone at 50Gy and 100Gy of x-ray exposure. At the same doses, gel and tendon collagen, which have been confirmed to exhibit similar fibrillar structure using atomic force microscopy, do not. This result is surprising considering that the gel is made of the same collagen source and are of identical concentration to the solubilised collagen. This implies the fibrillar structure, which is a key difference between gels and solubilised collagen, could be the crucial mediating factor for fragmentation resistance that was observed in gels and tendons. A possible explanation for the resistance of gel and tendon collagen to fragmentation could be the compact nature of collagen fibrils. The fibrils, which contain a high number of collagen monomers in a very small volume of space [257], may result in less exposure of collagen I chains to hydroxyl radicals produced by water radiolysis. Simulation studies have identified that the tight packing of collagen monomers in fibrils forces water molecules to settle into the gap and overlap regions in the D-bands [275,276], thus reducing the overall number of water molecules within the fibril compared to those surrounding a solubilised collagen monomer. This reduces the chance of producing hydroxyl radicals from water molecules within the fibrils and subsequently compromising the peptide backbone [238,259]. Hence, the number of fragmented collagen monomers would be minimal in fibrillar collagen. Instead, there could be an increased chance for hydroxyl radical production outside of the fibrils, which would be disproportionately absorb by the group of monomers at the surface of the fibrils [205], perhaps producing fragments so small that it becomes undetectable by gel electrophoresis. Further work is needed to verify the fibrillar damage, possibly by using AFM or electron microscopy as a means to observe changes to the D-band structure, which may be altered if x-rays are indeed targeting those surface collagen chains [167].

The impact of therapeutic x-rays on the native structure of solubilised, gel, and tendon collagen I was also further studied with CD and acidic native-PAGE. No detectable change was

observed for the CD spectrum for all three forms of collagen, indicating that the triple helical structure was not compromised by therapeutic x-rays. In contrast, streaking of bands can be observed in native-PAGE for gel, tendon, and radiation exposed solubilised collagen, which are often associated with aggregated protein [277]. Protein aggregation for acidic native-PAGE can occur during the stacking phase of the gel, where the local concentration of the protein is dramatically increased as proteins are concentrated into a tight band [278,279]. The streaks in gel and tendon collagen were consistent amongst both control and treated samples, indicating they were unlikely induced by therapeutic x-rays. Additionally, the collagen in tendons and gel may not have been effectively solubilised from its fibrillar state since no detergents were used for the native-PAGE. Combined with the low salt concentration of the buffer, this may be insufficient to keep the collagen I solubilised when stacked, hence causing it to precipitate and result in smears in the separating gel. In the solubilised collagen, however, those streaks are only apparent in radiation exposed samples, implying they could be related to therapeutic x-ray effects on solubilised collagen. We postulate that the smear observed may not be aggregated protein but rather slightly fragmented collagen monomers with poorer electrophoretic mobility. Radiation-induced fragmentation of collagen chains (seen in SDS-PAGE) may cause partial unwinding of the triple helix in those regions, which, in contrast to the compact and fully wound triple helix, would exhibit lower electrophoretic mobility [280]. Another possibility would be the presence of x-ray induced collagen fibrillogenesis that was seen with human collagen in the previous section. However, in contrast to previous results in human collagen, no thermal ramp was conducted with CD so there is currently a lack of evidence to support this.

Using biochemical techniques, no changes were detected from exposure to therapeutic x-rays in gel and tendon collagen I. However, PLF was able to identify significant changes to proteolytic susceptibility in all three forms of collagen I after x-ray exposure. Remarkably, a greater number of regions with altered proteolysis were found in gel and tendon collagen, instead of solubilised collagen. This was unexpected considering that the observation of fragmentation, which was expected to alter trypsin's accessibility to digestion sites [281], occurred only for solubilised collagen. It could be argued that solubilised collagen's monomeric structure is too simplistic for any change in proteolytic accessibility to be detected, with its elongated structure very much open to proteolytic attack [282].

Additionally, digesting the samples with trypsin overnight may have abolished any slight differences in proteolysis due to the samples being digested to completion [283]. Using a limited proteolysis protocol, such as that suggested by Hubbard (1998), may be more optimal for detection of minute changes in proteolytic susceptibility for solubilised collagen samples [282].

In contrast to solubilised collagen, PLF detected numerous regions in gel and tendon collagen where proteolytic susceptibility was altered. Both gel and tendon collagen were observed to have greatest number of flagged regions at 50Gy of x-ray exposure, which were subsequently reduced at 100Gy. This phenomenon was also observed for the solubilised collagen, although less substantially, and similarly in human collagen (Chapter 3). The same reasoning could be argued – that there is possibly an increase in crosslinking events at 100Gy, offsetting the fragmentation-induced alterations to proteolysis susceptibility at the lower dose of 50Gy. Although, this reduction in flagged regions was not observed for the 100Gy exposed α_2 chain in tendons. It is suspected that presence of collagen associated proteins in tendons could bind to and cause increased resistance of collagen I to higher doses of x-rays. Examples of such proteins are decorin and biglycan, two proteoglycans which are crucial for tendon homeostasis [284], which were identified to be present in tendon samples through our mass spectrometry data. Their presence were shown to be important in maintaining mechanical strength of tendons by acting as a link between collagen microfibrils [284–286]. It is postulated that their presence in tendons connecting between microfibrils could help compact collagen fibrils [287,288], thereby increasing resistance to proteolysis. Collagen fibrils in tendons also have inherent crosslinking conferred by lysyl oxidase while *in vivo* [289,290] that may also increase its resistance to proteolytic degradation. While counterintuitive, the inherent resistance to trypsin may help preserve or even enhance the detection of subtle proteolytic changes induced by therapeutic x-rays. The overnight digestion protocol implemented here would elevate the proteolytic differences as the digestion proceeds, with less risk of reaching completion due to its enzymatic resistance, thereby avoiding “over-digestion” as seen in solubilised collagen.

4.4.2 Differences in peptide fingerprint profile of solubilised, gel, and tendon collagen

From comparing the mass spectrometry peptide fingerprint profiles, the solubilised collagen fingerprint unexpectedly stood out from the profile of gel and tendon collagens. This result

draw parallels to previous work by Eckersley (2018), whom showed that collagen VI α_3 chain from human eye and human skin, despite having the same amino acid composition, had distinct peptide fingerprints, likely due to differences in bead morphology [225]. In the same manner, the differences observed in peptide fingerprint between solubilised collagen and collagen gel could be attributed to the additional fibrillar structure in the gel. Most fascinating (to the author) was how the gel peptide fingerprint matched the tendon fingerprint very closely, despite lacking crosslinking, cellular influence, and other associated collagen proteins. This suggests the peptide fingerprint of collagen I may be strongly governed by the fibrillar structure alone, while the other factors mentioned may provide subtler shifts or refinements to the peptide fingerprint. Another crucial observation when comparing the peptide fingerprint was a pattern of higher missed cleavages for regions where solubilised collagen had greater MS1 peptide intensity than gel and tendons, and no missed cleavages when MS1 peptide intensity was lesser. As previously discussed, solubilised collagen likely has greater proteolysis susceptibility and thus greater probability of over-digestion and producing non-tryptic peptides [291–293]. As our search parameters excludes non-tryptic peptides, the reduction of MS1 peptide intensities in regions without missed cleavages in solubilised collagen could simply be due to an increase in non-tryptic peptides. On the other hand, a greater number of peptides with missed cleavages were identified in solubilised collagen. The presence of missed cleavages could be a sign of poor binding capability and hence digestion efficiency of trypsin in those regions, that could be attributed to the occurrence of incompatible amino acids, such as proline, lysine, arginine, glutamate, or aspartate [294,295] in specific amino acid positions surrounding the cleavage site. This effect may be compounded with increased proteolytic resistance in fibrillar collagens, resulting in little to no proteolysis in those regions, thus producing lower MS1 peptide intensities for gel and tendons in regions with missed cleavages.

Overall, the comparison of peptide fingerprint profiles between solubilised, gel, and tendon collagens using mass spectrometry highlights the importance of protein structure in influencing proteolytic susceptibility. This is especially crucial when using PLF analysis to probe the effects of therapeutic x-rays. A sample that has completely different protein structure compared to when it is *in vivo* (such as solubilised collagen) may result in less clinically relevant observations or may require re-optimisation of protocols to better suit their

needs. The findings here also revealed that the peptide fingerprint, being able to distinguish between monomeric and fibrillar collagen, may prove useful for studying structural similarity between proteins of different sources or in different conditions. This may be valuable in drug discovery [296], where protein structure is crucial in maintaining the ability for drugs to bind to their targets [297]. PLF may thus be utilised for validating the structural stability of drugs by looking at changes to peptide fingerprint profiles.

4.4.3 Functional consequence of therapeutic x-ray exposure

To attempt to link the proteolytic changes observed in x-ray exposed collagen gel and tendons to their biological and functional consequence, mechanical studies using the atomic force microscope (AFM) were conducted to test for possible changes in stiffness. Miller et. al. (2018) had previously shown that collagen gels exposed to similar doses of therapeutic x-rays (63Gy) had a reduction in bulk tensile and compression modulus [142]. However, a global stiffness change may not necessitate a local change, especially at the micron scale where cells reside in [298–300]. To validate x-ray induced alteration in stiffness at the cellular level, the AFM was hence chosen to investigate the mechanics of collagen gels and tendons. The stiffness obtained for collagen gel and tendons were similar in magnitude to those obtained in literature, which varies from 0.1 to 0.2kPa for the gel [301] and 2.0 to 6.0 MPa for tendons [302]. The large difference could possibly be attributed to the difference in collagen fibril concentration, which is much greater in tendons. The experimental data comparing between radiation exposed and control samples were, however, inconclusive for both collagen gels and tendons and did not support the drastic decrease in bulk collagen stiffness as shown by Miller. The inconclusiveness of the results stems from the large variability of control samples. For the collagen gel, this may be attributed to the degradation of samples, poor re-alignment of the sample under the AFM microscope, or probe contamination [303]. In tendons the variability may be attributed to the cryosectioning process; the tendon's collagen fibre orientation during sectioning, either parallel or perpendicular to the slides, greatly influences the mechanical property that is detected with the AFM. When parallel, the measured property would be the longitudinal tensile strength, while if perpendicular, the transverse tensile strength would be measured [304,305]. The difficulty in manipulating the tendons results in cryosections of fibres in a mix of parallel and perpendicular orientations, influencing and greatly conflating the tensile strength detected on the AFM.

The limitations of the AFM may be reduced if we could obtain a portable source of ionising radiation to expose the samples while it remains on the AFM. This can alleviate sample re-alignment issues and avoid sample degradation during transportation. Tendon samples may also be tested on the AFM without cryosectioning and instead simply attaching the tendon samples on glass slides to prevent disorientation of collagen fibres [306].

4.5 Conclusion

In conclusion, the results here confirmed that therapeutic x-rays profoundly affect collagen I differently in different forms. Monomeric solubilised collagen is fragmented by therapeutic x-rays, while fibrillar collagen in gel and tendon do not. Instead, the proteolytic susceptibility of collagen alpha chains in collagen fibrils is altered, implicating possible changes to their ultrastructure. Remarkably, solubilised collagen peptide fingerprint from PLF analysis was highly distinguishable from gel and tendon collagen, which is likely attributed to the fibrillar structure. This may allow PLF to be a useful tool for applications wherein protein structure validation is important such as in drug discovery.

5 Effects of ionising radiation on *in vitro* fibroblast-derived ECM vs *ex vivo* breast tissue ECM

5.1 Introduction

This thesis has been focused on investigating if ECM components of breast tissue are compromised by therapeutic ionising radiation to understand whether ECM proteins may contribute to unwanted side effects of radiotherapy. Earlier chapters focused primarily on purified ECM protein samples or tendon (which has a relatively simple composition compared to breast), which are useful for isolating the impact of ionising radiation on specific proteins and simplify analysis (Chapters 3, 4). However, these biological samples fail to capture the proteome complexity of a tissue, in which ECM components interact with each other. These interactions can alter the structure of these ECM proteins [307], which could modify their sensitivity or resistance to x-ray damage. That said, experimenting on *ex vivo* tissue samples can be expensive and biological variations would be hard to control without a large sample size [308]. Utilising an *in vitro*, cell culture-based model that is able to produce complex ECM proteins for experimentation may help bridge between the simple, purified protein system to the complex, expensive, and often hard to get patient tissue samples. Hence, this study seeks to answer if an *in vitro*, fibroblast-derived cell matrix can be used as a model system for the breast stroma *in vivo*, and more specifically, for therapeutic x-ray responses.

5.1.1 Breast tissue and the extracellular matrix

The breast is made up of three main tissue types: stromal tissue, which comprises mostly of collagen I, adipose tissue, made primarily of adipocytes, and glandular tissue, which houses the mammary glands (Figure 5.1). These mammary glands are made of myoepithelial cells and are separated from the stromal tissue by the basement membrane, which is a subtype of ECM. There is interest in studying the stromal ECM, such as collagen I, fibronectin, and elastic fibre proteins as they provide the tissue with structural integrity, and also proteins of the basement membrane which includes collagen IV, laminin, perlecan, and nidogens [309]. The basement membrane is of particular interest due to its importance in acting as a barrier between epithelial cells and the bulk stroma of the breast tissue [201,310]. Breaching of the basement membrane by epithelial cells or alterations to the basement membrane mechanical integrity and/or stiffness has been associated with breast cancer progression and metastases

[311–313]. Below, the importance of the main components of the basement membrane is briefly highlighted.

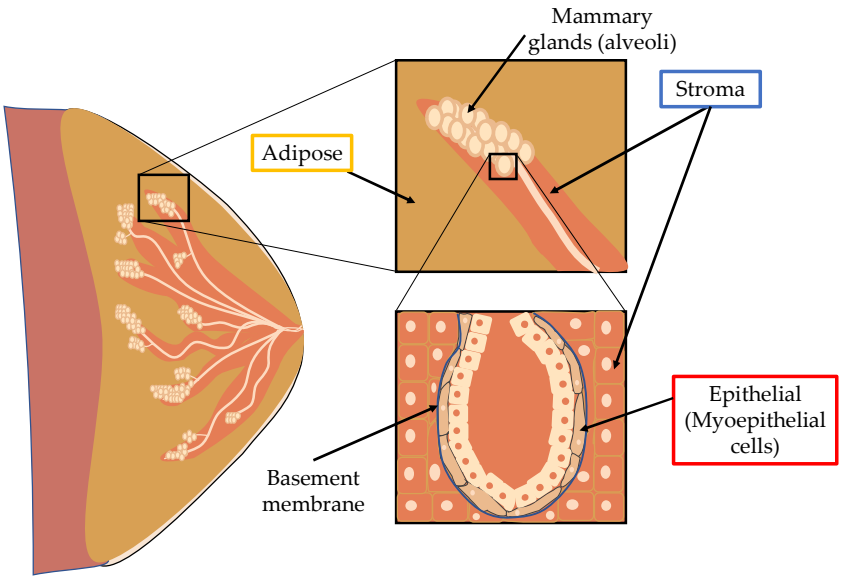


Figure 5.1. Breast tissue comprises the stromal tissue, adipose tissue, and glandular tissue. The stroma consists primarily of collagen I, while adipose tissue is made up of adipocytes. Glandular tissue is separated from the stroma by a layer of ECM known as the basement membrane, and within, the mammary glands are lined with myoepithelial cells. Figure reproduced from Tuieng et. al. [1] *with permission.*

Collagen IV is a non-fibrillar collagen that forms the network structure of the basement membrane and has six distinct α chains forming three different heterotrimers, the most common being $[(\alpha_1)_2]\ \alpha_2$. The heterotrimers interact through the C-terminal NC1 domain to form a hexamer. The hexamers are then crosslinked together at the N-terminal 7S domain to form a 'chicken-wire' mesh network. This network provides a scaffold for other ECM proteins and cells to bind to [314,315].

Laminin is another network-forming ECM protein in the basement membrane that forms independently of collagen IV [316,317] and similarly acts as a scaffold. Laminin is usually comprised of three chains (α, β, γ) with at least 15 different isoforms made from different combinations of α, β, and γ chains, the most common being laminin-111 (Laminin-111 implies $\alpha_1\beta_1\gamma_1$). Another isoform commonly associated with breast tissue and breast cancer is laminin-511 [317]. The trimeric laminin typically forms a cross-like structure, with three short extended arms contributed by N-terminals of each α, β, γ chains, and a longer body made of a triple helical coil of all three chains. The laminin network forms primarily by linking together the short arms of multiple trimeric laminins (three-arm interaction model) [318].

Another key basement membrane protein is the basement-membrane specific heparan sulphate proteoglycan (HSPG), or perlecan. Perlecan is a multifunctional protein and interacts with numerous ECM and cell membrane proteins, of which a comprehensive review is required to cover [319]. In the basement membrane, perlecan function as a matrix stabiliser, binding to both collagen IV and laminin, thus connecting the two structural networks. Furthermore, perlecan has also been found to bind to endothelial cells through the α2β1 integrin, hence could also facilitate cell-ECM mechanosensitive pathways [320].

Similarly, nidogen-1 also serves as a mediator of cell-ECM communication in the basement membrane through binding with laminin, specifically the domain III of the laminin γ chain [321,322], and perlecan. This has been postulated to allow nidogen-1 modulation of the interaction between laminin and integrins, and hence with cells. Additionally, it also connects between the collagen IV and laminin networks as its G2 domains can interact with the triple helical region of collagen IV [323,324].

While these crucial ECM components of the breast stroma has been thought to be primarily synthesised by mammary fibroblast [325,326], there is a lack of literature studying the matrix produced by an *in vitro* mammary fibroblast culture in comparison to the matrix in *ex vivo* breast tissue.

5.1.2 Extracellular matrix derived from *in vitro* cell culture

Cells grown in culture also produces a complex extracellular matrix as they adapt and grow despite being in an unnatural environment such as in plasticware [327,328]. Fibroblasts specifically are particularly efficient in producing these matrices as they are the main source of ECM deposition and remodelling in tissues [329]. Recently these matrices have been increasingly adopted as an alternative cell substrate as they are able to better mimic the cellular environment *in vivo* [328,330]. The matrix is produced by allowing fibroblasts to populate and deposit the matrix for a few weeks, before isolating the matrix by subsequent decellularization (detailed methodology in chapter 2). This makes it a convenient source of a biologically complex matrix. Studies of the proteomic profile of such cell culture-derived ECM from primary human lung fibroblasts identified major ECM proteins, including collagens I, III, IV, VI, laminin, nidogen and perlecan (not exhaustive) [331]. Matrix components of primary human dermal fibroblast cultures were also found to contain versican, laminin, fibulin-2, fibronectin, and thrombospondin-1 (not exhaustive) [225]. Evidently, the ECMs produced by

fibroblast cell cultures appear to produce a highly relevant and complex repertoire of ECM proteins.

5.1.3 *In vitro* VS *ex vivo* model systems

Whilst *in vitro* synthesis of ECM components by cultured cells may be convenient and reproducible, experiments on *ex vivo* tissue (in this case, breast tissue) are more likely to yield clinically relevant results across a diverse range of ECM components. That said, given the dynamic nature of tissue composition, it is difficult to generalise or extrapolate the results obtained from a small group of tissue donors due to biological variability, thus a large sample size is often needed [332]. This is problematic as it is often difficult and expensive to obtain tissue from patients [308]. Additionally, tissue samples typically originate from patients that may be diseased and are seeking treatment or surgery; normal tissues from healthy patients are hard to come by [333].

In vitro cell cultures help alleviate a large part of the complexity and biological variations. Cell cultures, which are often grown in tightly controlled environments, are likely to produce ECM that is more consistent, making experiments more reproducible [334,335]. Despite being less clinically relevant, the ECM extracted from cell cultures, as shown in literature, still contain ECM proteins that may be of interest and can provide significant insights [225,331]. However, before such insights can be drawn, there is a need to ascertain the biological similarity of such *in vitro* cell culture-derived ECM with the tissue of interest, as different fibroblasts secrete a specific combination of matrix proteins distinct for that tissue type [329,336].

5.2 Aims and hypothesis

There is currently (to the best of the author's knowledge) no published literature quantifying the ECM proteins produced by normal human mammary fibroblasts in cell culture and how well it encapsulates the ECM proteins of *ex vivo* breast tissue. Establishing such a comparison would validate future work conducted on the ECM produced by these mammary fibroblasts. Hence, this section of the thesis aims to provide a quantitative comparison on the proteome profile of *in vitro* immortalised human mammary fibroblast (HMFU-19 [186])-derived ECM (fibroblast ECM, or fECM) and *ex vivo* breast tissue ECM (tissue ECM, or tECM). Subsequently, this section aims to utilise fECM and tECM to investigate and compare the therapeutic x-ray damage signature of collagen I, FN, and associated basement membrane ECM proteins, specifically by looking at their changes to proteolytic susceptibility using peptide location

fingerprinting (PLF) analysis. This may help to address the final hypothesis presented at the start of the thesis, that mass spectrometry (PLF analysis) is able to identify therapeutic x-ray targets from a complex proteome derived from *in vitro* fibroblast cultures and *ex vivo* tissues. In order to achieve this, the following sub-hypotheses are postulated, that:

1. *In vitro* immortalised mammary fibroblast-derived ECM (fECM) would contain comparable ECM proteome to that from *ex vivo* breast tissue (tECM).
2. The protease susceptibility of some ECM proteins (as assessed by PLF) would differ between fECM and tECM,
3. Radiation induced changes observed ECM proteins would be sufficiently similar between fECM and tECM to serve as model for therapeutic x-ray studies

To test hypothesis 1 and 2, the protein composition of fECM and tECM was assessed using mass spectrometry proteomics with a specific focus on ECM proteins (Collagens, proteoglycans, glycoproteins, and secreted factors) as labelled in the human matrix proteome database from MatrisomeDB [337]. A bioinformatic tool, STRING [198], was used to group identified proteins into their functional and biological relevance to investigate overall functional similarities between the ECM proteins detected in fECM and tECM. PLF was used to quantify structural differences, inferred from variations in proteolytic susceptibility of major ECM proteins (such as collagen I and FN), between fECM and tECM.

The third hypothesis was examined by, again, applying PLF to analyse the protease susceptibility of collagen I and FN, the two most abundant ECM proteins, before and after exposure to therapeutic levels of ionising radiation (50Gy and 100Gy). STRING was subsequently used for investigating possible functional enrichment amongst all (intracellular and extracellular) proteins flagged by PLF in fECM and tECM.

The workflow for the LC-MS/MS data analysis using PLF and STRING is illustrated in the figure below (Figure 5.2).

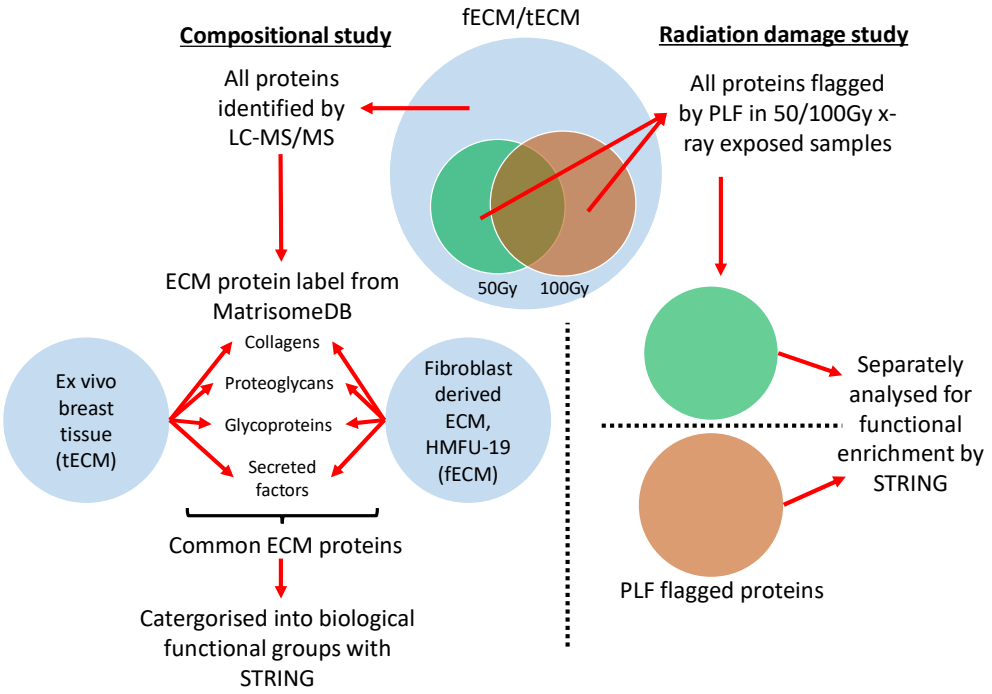


Figure 5.2. Workflow of the analysis of LC-MS/MS data using PLF and STRING [198]. *The total protein list identified in fECM and tECM were first separately organised into various ECM protein categories based on MatrisomeDB and comparison was made between fECM and tECM. Next, the group of common ECM proteins present in both fECM and tECM were subsequently entered into the STRING database and functional groups with highest representations from the group of common proteins were noted.*

For proteins flagged in PLF analysis, functional enrichment analysis was conducted using STRING database to identify possible functional/biological groups of proteins whose proteolytic susceptibility was more likely to be significantly altered by therapeutic radiation.

5.3 Results

5.3.1 *In vitro* immortalised fibroblast and *ex vivo* tissue samples contain similar collagen and glycoprotein composition but have distinct secreted factors and glycoprotein proteomes

A comparison was made between the ECM proteins identified from LC-MS/MS data in fECM and tECM, focussing on the four main categories of ECM proteins, namely collagens, glycoproteins, proteoglycans, and secreted factors. For each group, the proteins detected were compared between fECM and tECM and the number of common or distinct proteins are illustrated in the figure below (Figure 5.3).

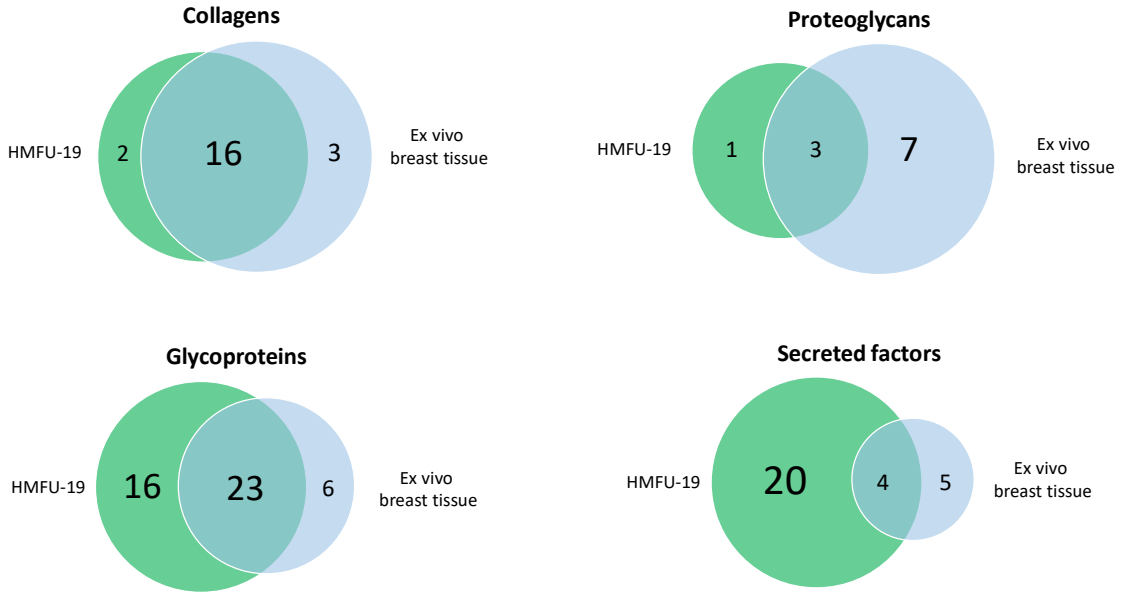


Figure 5.3. Compositional difference in ECM proteins obtained from in vitro culture of immortal human mammary fibroblasts (HMFU-19) (fECM) with ex vivo normal breast tissue (tECM). A significant portion of collagens are seen in both in vitro and ex vivo samples, making the collagen profile similar. On the other hand, secreted factors and glycoproteins were distinct between fECM and tECM; Significantly greater number of secreted factors were identified in fECM than tECM.

In the family of collagens, fECM and tECM shared a large pool of important collagens that are highly abundant, including fibrillar collagens I, III, V, as well as network-forming and basement membrane associated collagens VI, XVIII. These collagens contribute a large part to the mechano-environment influence on cellular behaviour. However, we note that the collagen pool for tissue was slightly more complex. For example, the collagen IV α_5 chain was only found in tECM and not in fECM.

In contrast to collagens, the group of glycoproteins identified were largely different between the two ECMs. There were significantly greater number of glycoproteins identified in fECM

and most notably, there was presence of latent-transforming growth factor-beta binding proteins (LTBP) 1-4 in fECM but not in tECM. Similarly, a large proportion of secreted factors were uniquely produced and sequestered in fECM. A list of biologically significant proteins identified exclusively to either fECM or tECM is compiled in the table below (Table 5.1)

Table 5.1. Biologically relevant ECM proteins exclusively identified in fibroblast derived ECM (fECM) or ex vivo tissue (tECM) and their respective biological roles and significance.

Source		Biological significance
	Collagens	
Fibroblast-derived ECM	Collagen XVI α_1 chain	Fibril associated collagen with interrupted triple helices (FACIT), Part of Fibrillin-1 microfibrils in skin [338], may play a role in anchoring fibrillin microfibrils to basement membrane [339]
Ex vivo tissue	Collagen II α_1 chain	Secreted by chondrocytes, major component in cartilage tissue [340]
	Collagen V α_3 chain	Less abundant chain in collagen V, associated with fibrillogenesis of collagen I and III [341].
	Collagen VI α_6 chain	Similar to alpha-3 chain of collagen VI, but more abundant in endomysium/perimysium of skeletal muscles [342]. Collagen VI binds to many ECM components including fibronectin, collagen IV [343,344].
	Glycoproteins	
Fibroblast-derived ECM	Latent-transforming growth factor-β binding protein (LTBP) 1-4	LTBP associates with ECM components fibrillin microfibrils and fibronectin [345,346]. Binds to and maintains TGF-β1 latency. LTBP-1 may play a role in TGF- β1 integrin-mediated activation [132,347]
	Tenascin-C	Binds FN [348], perlecan, also growth factors EGF, bFGF, PDGF
	Thrombospondin-2	MMP-2 regulation, can reduce cell adhesion when MMP-2 not regulated [349]
	Cellular communication network factor 1 (CCN1)	Wound healing/integrin binding [350,351]
	Fibrillin-2	Fibrillin microfibrils, elastin deposition [138]
Ex vivo tissue	Tenascin-X	Architectural support for collagen and elastin fibres, modulates cell adhesion [352,353]
	Fibrinogen α/β/γ chains	Circulates in blood, forms fibrin clot after cleavage by thrombin [354,355]
	Elastin	Elastic fibre for elasticity of tissue [356]
	Proteoglycans	
Fibroblast-derived ECM	Hyaluronan and proteoglycan link protein-1	Binds to hyaluronan and stabilises proteoglycan aggregates [357]
Ex vivo tissue	Decorin	Collagen fibrillogenesis [358], cell adhesion [359], TGF-β1 binding [360]
	Fibromodulin	Collagen fibrillogenesis [361], enhance cross-linking by lysyl oxidase [362]
	Lumican	Accelerate collagen fibril formation [358], may promote fibrosis [363]
	Prolargin	links basement membrane to connective tissues by binding to perlecan [364], angiogenic regulator and cancer associated fibroblast tumour suppressor [365]
	Podocan	Binds collagen I, inhibit cell proliferation/migration [366]
	Secreted factors	
Fibroblast-derived ECM	FGF-2/7	Mediate cell proliferation, migration, differentiation [367]
	TGF-β1	Multifunctional cytokine, associated with fibrosis [368,369]
	Protein wnt-5a/5b	Ligands for binding to frizzled proteins, activate/inhibit wnt signalling [370,371]
	Growth differentiation factor 15	Secreted ligand of TGF-β family of proteins that regulate gene expression through SMAD [372]
	Inhibin/Activin β-E chain	TGF-β family [157,373,374]
Ex vivo tissue	Protein S100-A4	Expressed exclusively by fibroblast, expressed during injury/disease, calcium binding. Increased expression in tumour stromal cells. Inhibit FGF2-induced FGF receptor signalling pathway [375]
	Secreted frizzled-related protein 1	Modulate wnt signalling, can induce angiogenic response [376,377]

Besides looking at protein composition of the ECM, the ECM dynamics is also an important factor to consider in comparing the biological similarity between the *in vitro* and *ex vivo* systems [378]. A crucial part of ECM dynamics involves remodelling of the ECM by extracellular proteinases, as they regulate the proteolysis of key structural matrix proteins. A major class of proteinases is the matrix metalloproteinases (MMPs). Proteomics data from LC-MS/MS seemed to suggest that fECM and tECM contains very different classes and quantities of MMPs. In fECM, major MMPs were present including MMP-1, -2, -3, and -14. In comparison, only MMP-2 was detected in tECM. However, both fECM and tECM comprised tissue inhibitor of metalloproteinase 3 (TIMP3), which is an important regulator of ECM composition through MMP inhibition [379].

5.3.1.1 Common ECM proteins

While differences were observed in the family of glycoproteins and in the profile of MMPs, still, a total of 48 ECM proteins were identified to be common amongst fECM and tECM (Figure 5.4). Within those 48 common ECM proteins comprised crucial basement membrane proteins, namely, collagen IV, laminin, perlecan, and nidogen-1.

Peptide spectrum matches (PSMs) were used as an estimate of protein abundance (Figure 5.4). Collagens were found to be most abundant in both fECM and tECM by total PSM count, specifically collagen I, III, V, and VI. Fibrillin-1 was also highly abundant in both ECMs. However, a few crucial glycoproteins were found in higher abundance within fECM than tECM, including thrombospondin-1, TGFβ-induced glycoprotein-H3 (βig-h3), perlecan, agrin, elastin microfibril interfacer-1 (EMILIN-1), and fibronectin.

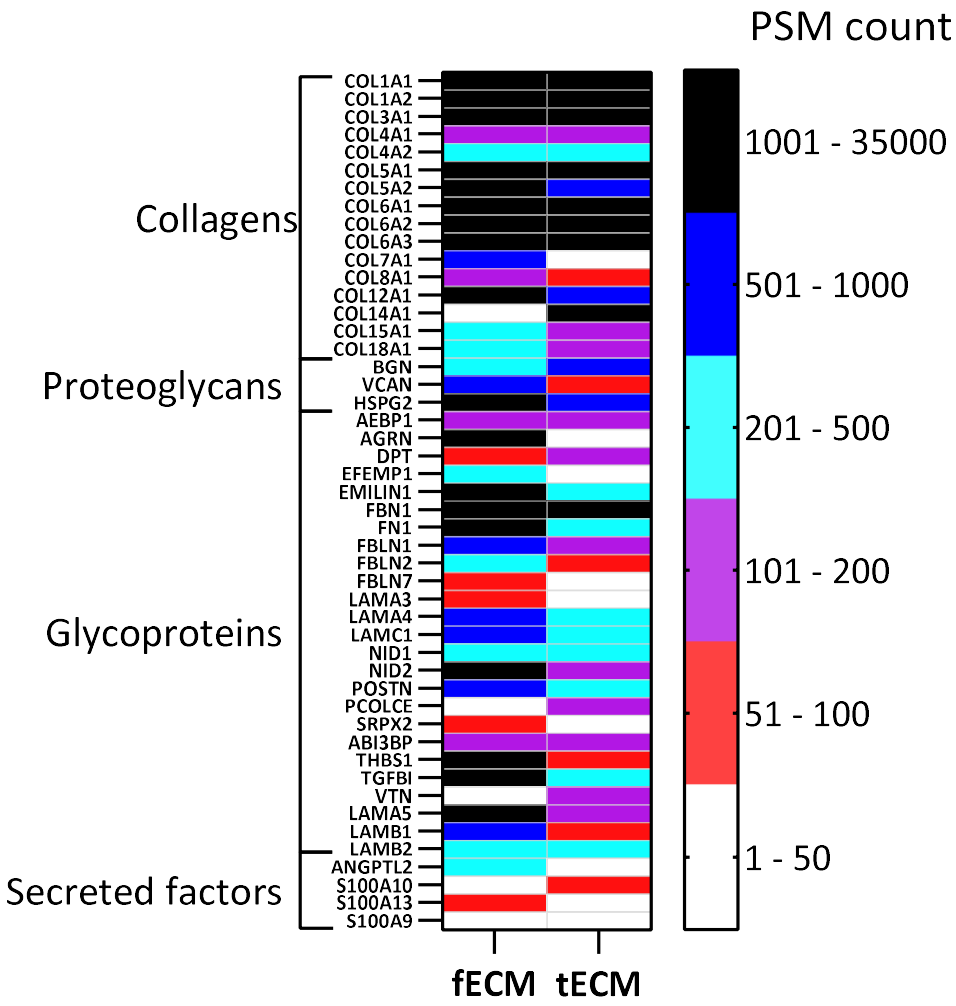


Figure 5.4. 48 common ECM proteins identified with mass spectrometry in both fECM and tECM and their respective abundances. Total count of peptide spectrum matches (PSM) for each protein was used as proxy for the level of abundance of protein in the respective ECM.

To explore the association amongst the common ECM proteins found in both fECM and tECM, a search was conducted on the STRING database and associated protein groups were highlighted. In total, 22 (out of 48) common ECM proteins were found to be associated with the basement membrane (Gene ontology (GO):0005604), while 25 proteins are associated with the organisation of the ECM (GO:0030198). The STRING interaction pathways of the proteins are illustrated below (Figure 5.5). Additionally, there were a small but significant group of elastic fibre associated proteins (GO:0030023, Extracellular matrix constituent conferring elasticity), which are EMILIN-1, Fibrillin-1, and Fibulin-2, that were also found in both fECM and tECM (interaction pathway too small hence not shown). In all, the ECM proteins found in fECM appear to encompass crucial ECM constituents with biologically important functions that are also represented in tECM.

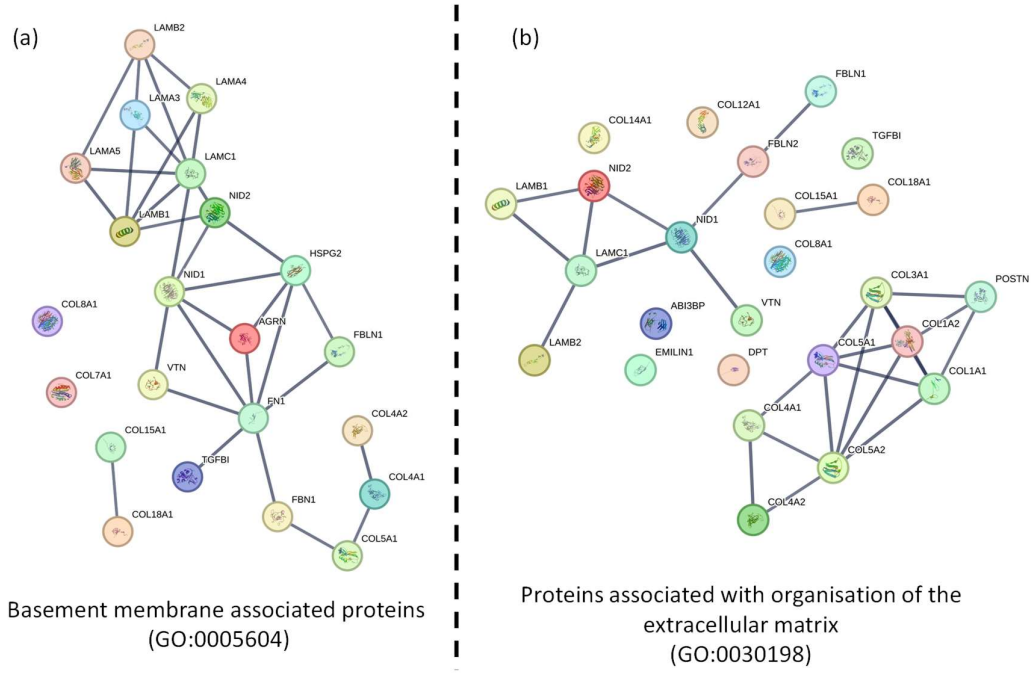


Figure 5.5. Gene ontology terms associated with common proteins found in fECM and tECM. Out of a total of 48 ECM proteins identified by PLF analysis in both fibroblast-derived (fECM) and ex vivo tissue ECM (tECM), (a) 22 proteins are associated with the basement membrane (GO:0005604), and (b) 25 proteins are associated with the organisation of the ECM (GO:0030198) based on annotation by the gene ontology database. The connecting lines represents an interaction between connected proteins based on STRING's confidence score greater than 0.9 [468].

5.3.2 Functional enrichment analysis of all proteins flagged by PLF analysis reveal therapeutic ionising radiation may preferentially alter structure of extracellular matrix proteins

To test if proteins from fECM and tECM exhibit differential responses to therapeutic ionising radiation, PLF was used to analyse all (intra- and extracellular) proteins from both fECM and tECM, comparing between control and samples exposed to therapeutic ionising radiation (50Gy and 100Gy). Proteins (either intra- or extracellular) with PSM > 45 (arbitrarily chosen to minimise false positives from low PSM) and had at least one region of significantly altered MS1 peptide intensity were identified, and was considered 'flagged' by PLF.

A total of 1097 and 738 proteins with PSM > 45 were detected for fECM and tECM respectively (Figure 5.6a). In fECM, 206 proteins were flagged by PLF from 50Gy exposure, and 245 proteins were flagged for the 100Gy group, of which 103 proteins were common between both groups. For tECM, 72 proteins were flagged in the 50Gy group and 378 proteins for the 100Gy group, with 63 common proteins. STRING webtool was also utilised to analyse in bulk the entire group of intracellular and extracellular proteins flagged by PLF to test if they were enriched for proteins associated with biological ontology terms [198] (Figure 5.6b).

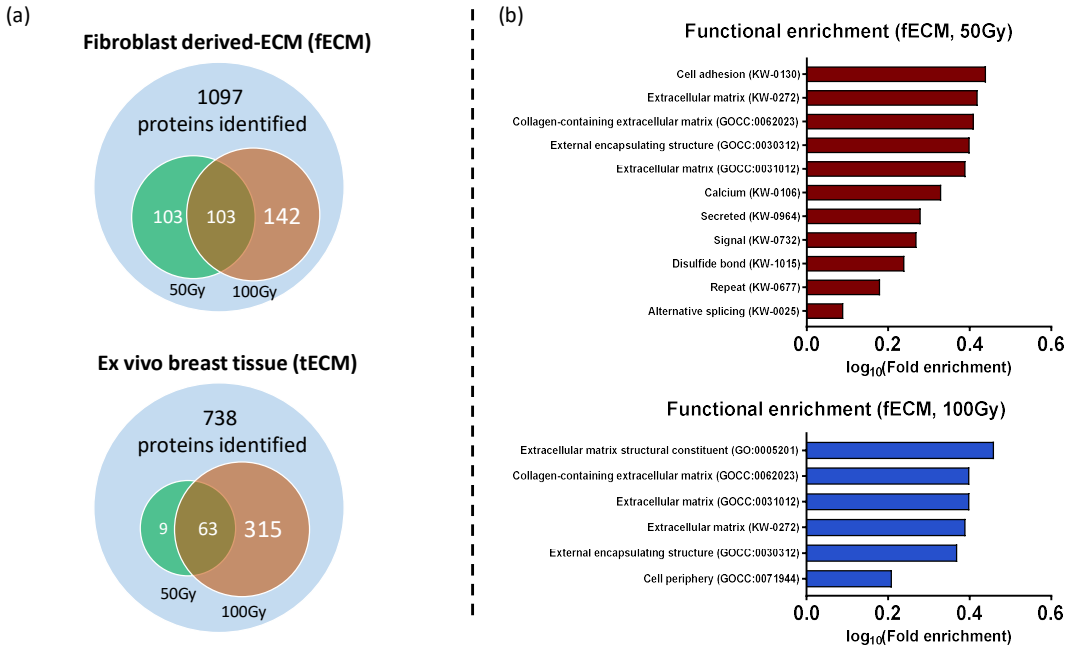


Figure 5.6. Functional enrichment analysis of proteins deemed structurally altered by therapeutic x-rays (as flagged by PLF) in both fECM and tECM. (a) The total number of intra- and extracellular proteins identified (PSM > 45) were 1097 and 738 for fibroblast-derived ECM (fECM) and ex vivo breast tissue (tECM) respectively. The proteins that were flagged by PLF to have at least one region with significantly altered peptide intensity after exposure to ionising radiation (50Gy and 100Gy) were tallied.

(b) Functional enrichment analysis was conducted using the STRING webtool [198] *on the group of all (intra- and extracellular) proteins flagged by PLF for fECM and tECM separately, and for each radiation dose. The background gene/protein dataset was set to be the full protein list (PSM > 45) obtained from LC-MS/MS for the respective set of samples, and a false discovery rate of 0.05 was set. Functional enrichment was found for the group of proteins flagged by PLF in fECM at 50Gy and 100Gy radiation exposure. The enrichment encompassed proteins associated with cell adhesion and the ECM at 50Gy, while at 100Gy the top four enriched proteins annotations were all related to the ECM.*

Only flagged proteins from fECM appear to have functional enrichment while tECM did not, hence only terms associated with fECM are displayed. The lack of functional enrichment for tECM implies that the group of flagged proteins in tECM did not have specific groups of proteins that were over-represented – remodelling by therapeutic x-rays did not seem to specifically target any protein group. In contrast, the group of proteins flagged by PLF in fECM exposed to 50Gy of therapeutic x-rays appear to be functionally enriched for proteins that play a role in cellular adhesion (UniprotKB keywords database (KW):KW-0130) and are associated with the ECM (KW-0272, Gene ontology cellular component (GOCC) database:0062023). Similarly, in the group of proteins flagged in 100Gy exposed fECM, the top 4 functional enrichments were all associated with the ECM.

5.3.3 Similarity in control peptide fingerprint of collagen I chains observed between fibroblast-derived ECM and *ex vivo* tissue ECM

Whilst similar ECM proteins were detected in both fECM and tECM, it is also clear that protein function, particularly for large ECM proteins, is determined by their molecular structure and maturation. Compared with the newly synthesised fECM from *in vitro* cell culture, the tECM will likely be predominantly composed of mature proteins in large assemblies, which may be structurally distinct to those from fECM. Before investigating and comparing the impact of radiation on these proteins between fECM and tECM, there is a need to ascertain the structural similarity of these proteins *in situ*. Using the peptide fingerprint obtained from LC-MS/MS, the structural state on individual proteins was inferred and compared between the two different systems based on their susceptibility to tryptic digestion. Here, collagen I and FN were chosen to focus on given their importance in the ECM and that they were well studied in previous sections.

To compare the peptide fingerprint between fECM and tECM, while minimising the effects from different protein abundance, the data was min-max normalised and subsequently plotted as a percentage of the maximum peptide intensity (Figure 5.7). Data for solubilised collagen I and solubilised human plasma fibronectin (pFN) from Chapter 3 were compiled together for further evaluation.

For collagen I α_1 and α_2 chains, there was remarkable correlation between the peptide fingerprint obtained from fECM and tECM. Barring a few select regions that had different percentage MS1 intensities (%MS1) (annotated in figure by red (fECM>tECM) or black (fECM<tECM) arrows), the overall fingerprint of collagen chains appears to match between the two types of ECM. Solubilised collagen stood out and was generally higher in %MS1, such as between bins 24-31 and 44-51 in the α_1 chain, implying greater proteolytic susceptibility of solubilised collagen. Notably, it was observed in certain bins (those annotated with arrows) that %MS1 between fECM and solubilised collagen coincided better than with tECM, such as bins 39, 40, and 57 for α_1 and bins 42, 43, 48, and 49 for the α_2. This implies that those regions in the collagen of fECM were similar in proteolytic susceptibility to purified, solubilised collagen.

In contrast, the FN fingerprint was different between fECM and tECM. The FN fingerprint from fECM had better coverage of the entire protein sequence and were consistently higher in

percentage MS1 intensity throughout most of the protein. The presence of peptides in the extra domain A (EDA) (between bin 86-91, AA 1727-1814) and V region (bin 104-110, AA 2083-2202) implies that the FN in fECM is cellular FN [211]. Compared with the peptide fingerprint of pFN, there were also very few similarities with either fECM or tECM, indicating that FN had highly dissimilar proteolytic susceptibility in its solubilised state compared to when it is deposited in the ECM.

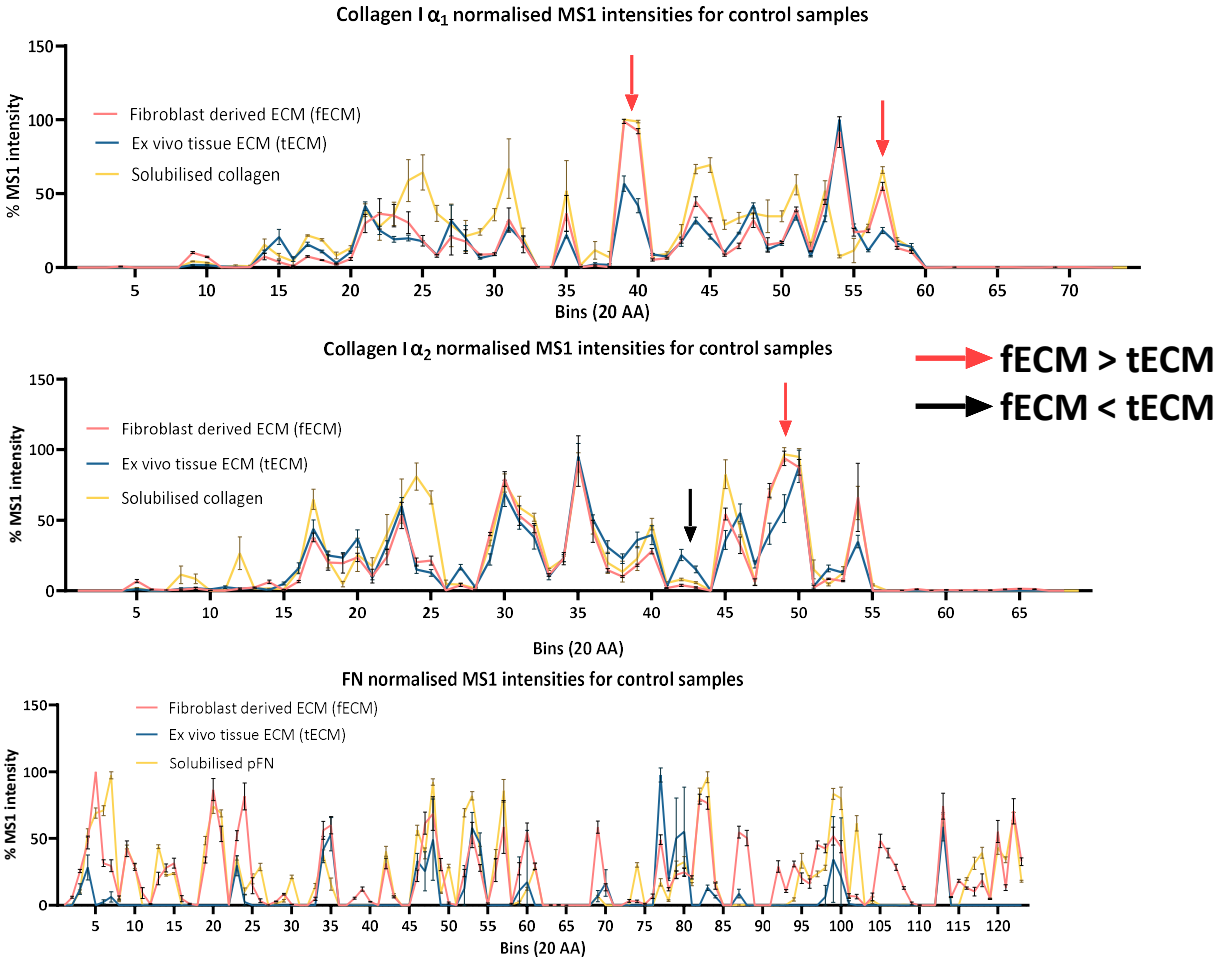


Figure 5.7. Peptide fingerprint of control groups of collagen I α_1/α_2 chains and fibronectin from fibroblast derived ECM (fECM), ex vivo tissue ECM (tECM), and purified solutions (from chapter 3) normalised to the maximum MS1 intensity among bins. While collagen I chains had very high similarity in fingerprint profile between fECM and tECM samples (slight differences highlighted with red (fECM>tECM) or black (fECM<tECM) arrows, not exhaustive), the fingerprint profile for fibronectin was highly distinct with HMFU having much higher intensities and more abundant peptide yields across the entire protein sequence. Certain regions in collagen where fECM differed from tECM (arrows), fECM appear to correspond better with solubilised collagen, indicating similarity in proteolytic susceptibility of those regions.

In all, the peptide fingerprint for collagen suggests the proteolytic susceptibility and possibly ultrastructure of collagen chains is closer in similarity between fECM and tECM than solubilised collagen. Proteolytic susceptibility of FN from fECM appeared to be distinct to solubilised pFN. However, it is difficult to conclude for tECM due to the lack of proper peptide coverage and hence a robust peptide fingerprint in tECM.

5.3.4 Alteration of collagen I and fibronectin peptide fingerprint from exposure to therapeutic x-rays differs between fibroblast-derived ECM and *ex vivo* tissue ECM.

After validating the similarities and differences in proteolytic susceptibility of collagen I and FN, PLF analysis was conducted to compare between the control and x-ray exposed group for fECM and tECM separately. Collagen I and FN from both fECM and tECM were found to contain at least 1 region with significantly altered ($p<0.05$) proteolytic susceptibility after exposure to therapeutic x-rays. To further investigate where those specific regions were located in the protein, the data was plotted as MS1 intensity difference between control and radiation exposed samples (50Gy: Figure 5.8, 100Gy: Figure 5.9).

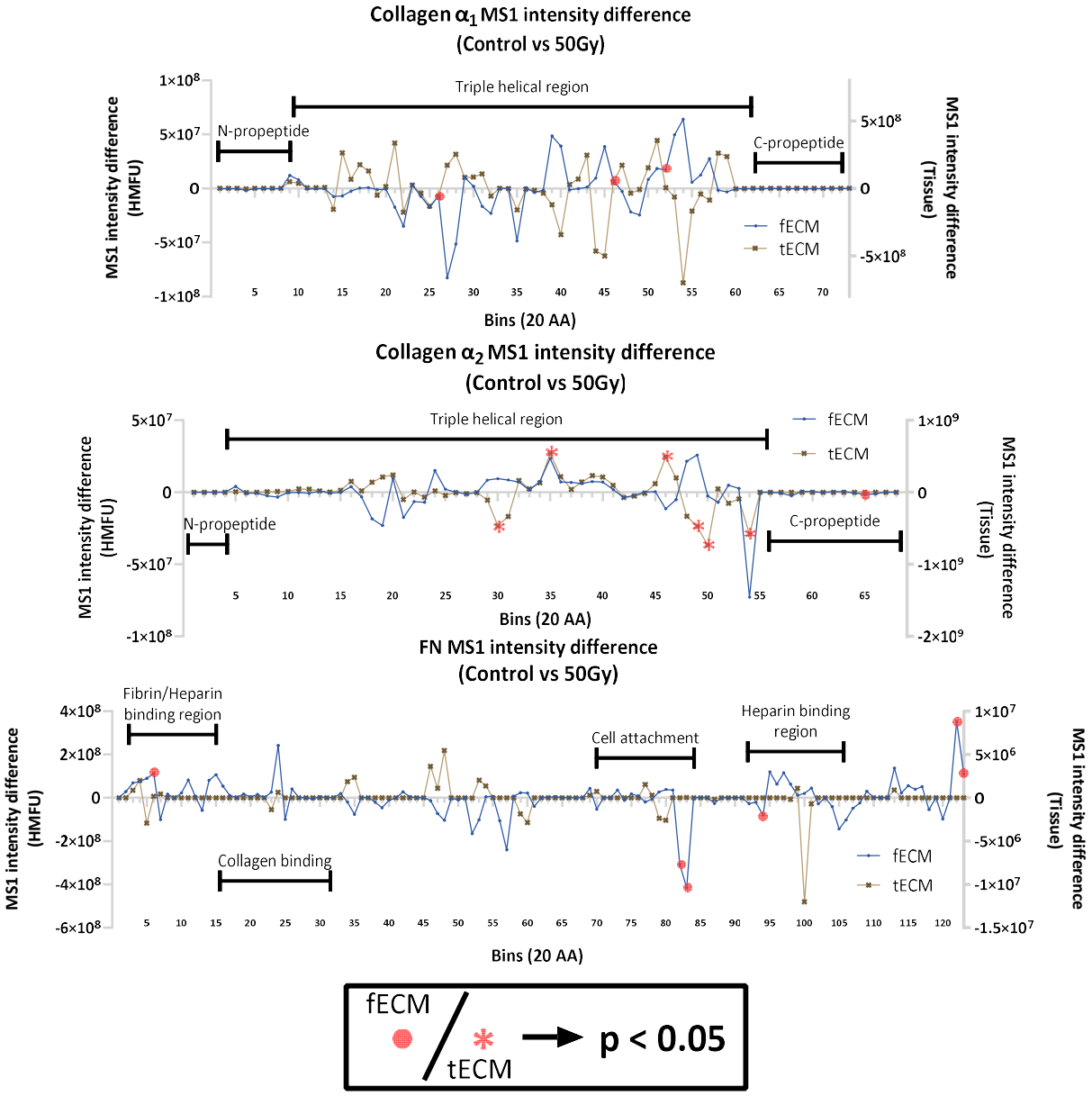


Figure 5.8. Peptide location fingerprinting (PLF) analysis was implemented on LC-MS/MS data of collagen I α_1/α_2 chains and fibronectin for fibroblast derived ECM (fECM) and ex vivo tissue ECM (tECM) exposed to 50Gy of therapeutic x-rays. Regions where the difference in peptide MS1 intensities were statistically significant implies alteration in proteolytic susceptibility and were highlighted in red. Collagen I α_1 appears preferentially damaged compared to α_2 in fECM, while it is the opposite in tECM based on the number of regions flagged. For FN, there were no correlation in regions flagged for PLF between the two samples, and fECM appear to be implicated in more regions and appear to affect similar binding regions of interest such as heparin/fibrin, and cell attachment areas as those found in plasma FN studies (Chapter 3).

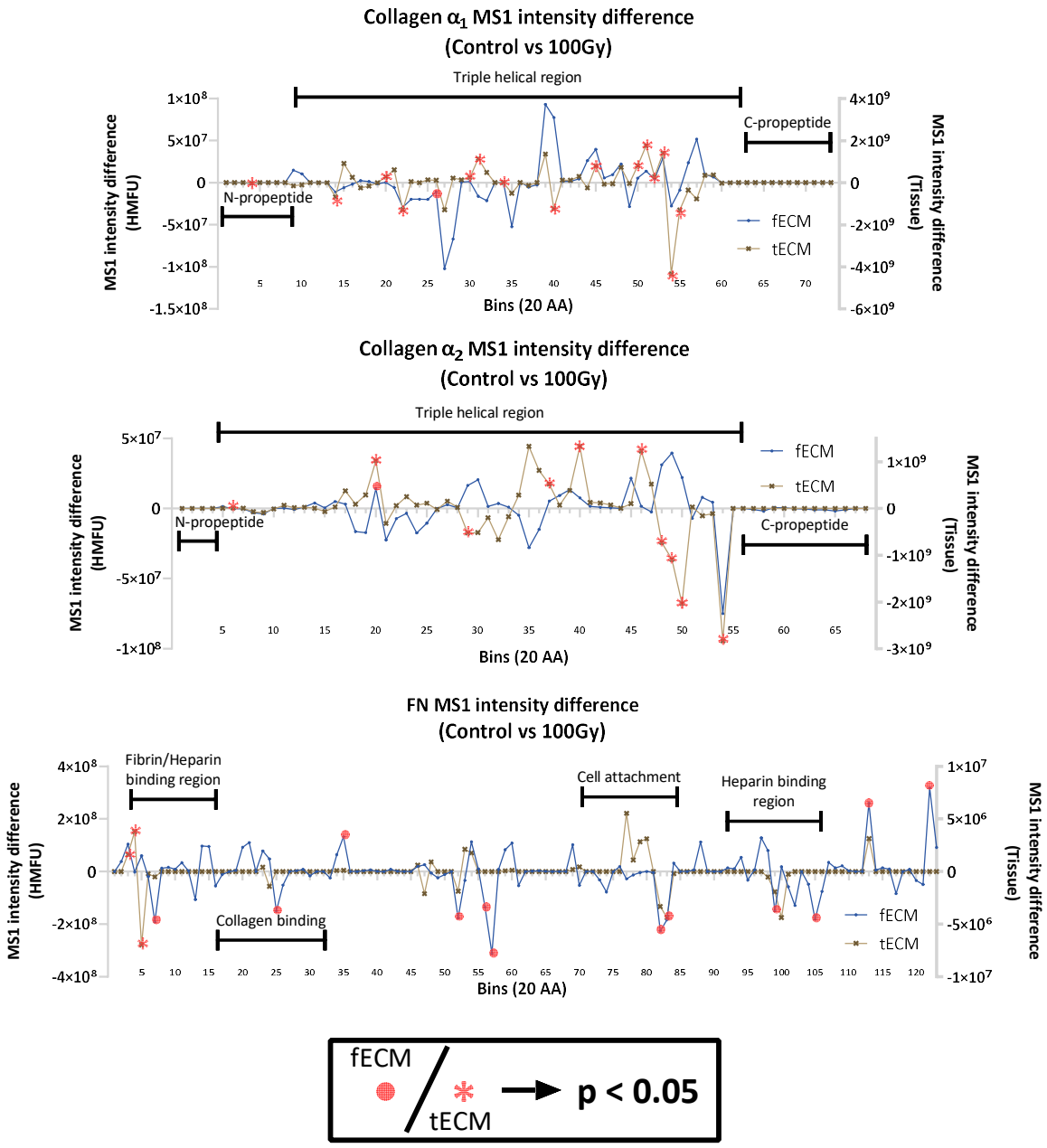


Figure 5.9. Peptide location fingerprinting (PLF) analysis was implemented on LC-MS/MS data of collagen I α_1/α_2 chains and fibronectin for fibroblast derived ECM (fECM) and ex vivo tissue ECM (tECM) exposed to 100Gy of therapeutic x-rays. Regions where the difference in peptide MS1 intensities were statistically significant implies alteration in proteolytic susceptibility and were highlighted in red. In both collagen I chains, tECM contained more regions that were highlighted than fECM. Furthermore, collagen chains from fECM appear to have greater flagged regions compared with 50Gy, unlike tECM where he number of flagged regions reduced compared to 50Gy.

For FN, both fECM and tECM had an increase in number of regions flagged in 100Gy compared to 50Gy, possibly hinting at some dose dependency. Still, fECM had a much higher number of flagged regions than tECM. Collagen binding region appears to be affected in fECM but not tECM after 100Gy exposure of therapeutic x-rays.

In collagen I, x-ray exposed α_1 chain had greater number of altered regions than α_2 for fECM, but is reversed in tECM, especially for 50Gy. At 100Gy, there were fewer altered regions than 50Gy for fECM but tECM exhibited an increase in number of regions, with α_1 exhibiting a larger increase than α_2. The comparison of number of flagged regions between fECM and tECM for the collagen chains is compiled in the table below (Table 5.2). The only region where both x-rays induced changes in peptide fingerprint matched for fECM and tECM collagen was the α_2 chain at bin 20 (aa. 380-400).

Table 5.2.Peptide location fingerprinting analysis using LC-MS/MS data on collagen I α_1/α_2 chains and the number of regions with statistically significantly altered MS1 intensities. At 50Gy of therapeutic x-ray exposure, collagen I α_1 appears preferentially damaged compared to α_2 in fECM, while the opposite is true in tECM. At 100Gy, fECM had a decrease in number of flagged regions while tECM drastically increased, and α_1 chain had significantly more regions flagged than α_2.

	PLF flagged regions in Fibroblast derived ECM (fECM)		**PLF flagged regions in *Ex vivo* tissue (tECM)**	
Dose	COL1A1	COL1A2	COL1A1	COL1A2
50Gy	3	1	0	6
100Gy	1	1	15	10

PLF analysis for FN in fECM highlighted a larger number of regions with significantly altered proteolytic susceptibility in response to therapeutic radiation (50Gy: Figure 5.8, 100Gy: Figure 5.9). At 50Gy, affected regions were observed around the fibrin/heparin binding sites and cell attachment regions, while at 100Gy this extended even to the collagen binding site. The peptide fingerprint of FN from tECM appeared more radioresistant to structural changes from radiation exposure, with no regions affected at 50Gy and a small group of regions around the heparin/fibrin binding site at 100Gy.

Overall, the changes to proteolytic susceptibility of collagen I and FN from therapeutic x-ray exposure seems to be distinct between fECM and tECM. While collagen chains in both fECM and tECM had flagged regions from PLF (indicating possible proteolytic changes), there were no clear consistency in number of regions affected nor the positions of those regions. FN appear more susceptible to therapeutic x-ray in fECM, particularly at 100Gy, compared to

tECM. However, this may be associated with the poor peptide fingerprint coverage of FN in tECM. This suggests that structural responses of collagen I and FN to therapeutic x-rays in *ex vivo* tissue may not be replicable using *in vitro* fibroblast derived ECM.

5.3.5 Basement membrane associated ECM proteins may also be structurally affected by therapeutic ionising radiation

Besides collagen I and FN, basement membrane proteins also play crucial roles in the mammary ECM, thus warranting further examination into their responses to therapeutic x-ray exposure. As LC-MS/MS proteomics data revealed that both fECM and tECM contained major basement membrane proteins (collagen IV, laminin, nidogen-1, and perlecan), PLF was used to further investigate their differences in peptide fingerprints between fECM and tECM, and how their proteolytic susceptibility is affected by therapeutic x-rays.

5.3.5.1 Peptide fingerprint profile of proteins are more conserved for fibroblast-derived ECM than ex vivo tissue ECM

Similar to the comparison made for collagen I and FN, PLF was utilised to ascertain the similarity or differences in structure of basement membrane proteins identified in both fECM and tECM by comparing the peptide fingerprint of non-irradiated control samples (Figure 5.10).

In general, proteins' peptide fingerprints from fECM had higher MS1 intensities and better coverage compared to tECM. This was prevalent in numerous other ECM proteins, including collagen VI, XII, biglycan, periostin (data not shown). Amongst the basement membrane proteins, the difference in peptide coverage was most distinct between perlecan and laminin subunits, which may indicate a higher protein abundance or better proteolysis of these ECM proteins in fECM.

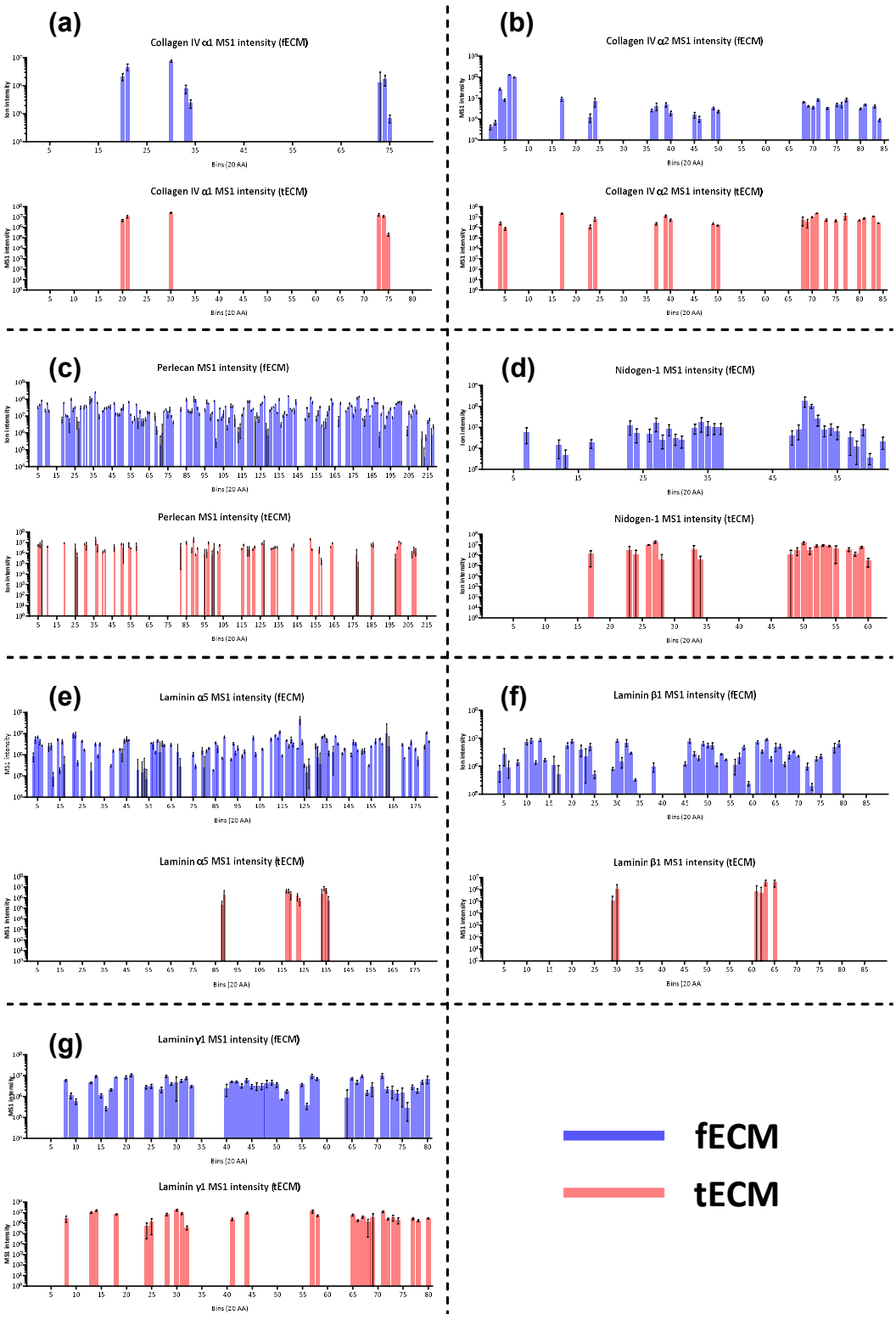


Figure 5.10. Peptide fingerprint profile of control (unirradiated) samples of other key basement membrane ECM proteins such as (a) collagen IV α1 and (b) α2 chains, (c) perlecan, and (d) Nidogen-1, (e-g) Laminin subunits α_5, β_1, and γ_1, illustrate that LC-MS/MS of fibroblast derived ECM (fECM)(blue) provided better peptide coverage of these basement membrane proteins than ex vivo tissue (tECM) (red).

5.3.5.2 Collagen IV

The proteomics data suggested that collagen IV from both fECM and tECM were mostly α_1 and α_2 chains, implying that the most abundant trimer is likely that of $[(\alpha_1)_2]$ α_2. Hence, PLF analysis focused only on these two collagen IV chains (Figure 5.11). PLF data was sparse due to the low coverage of both collagen IV chains. However, it was observed that the α_2 chain for tECM had significantly altered regions within and surrounding the C-terminal NC1 domain after exposure to 100Gy of therapeutic ionising radiation.

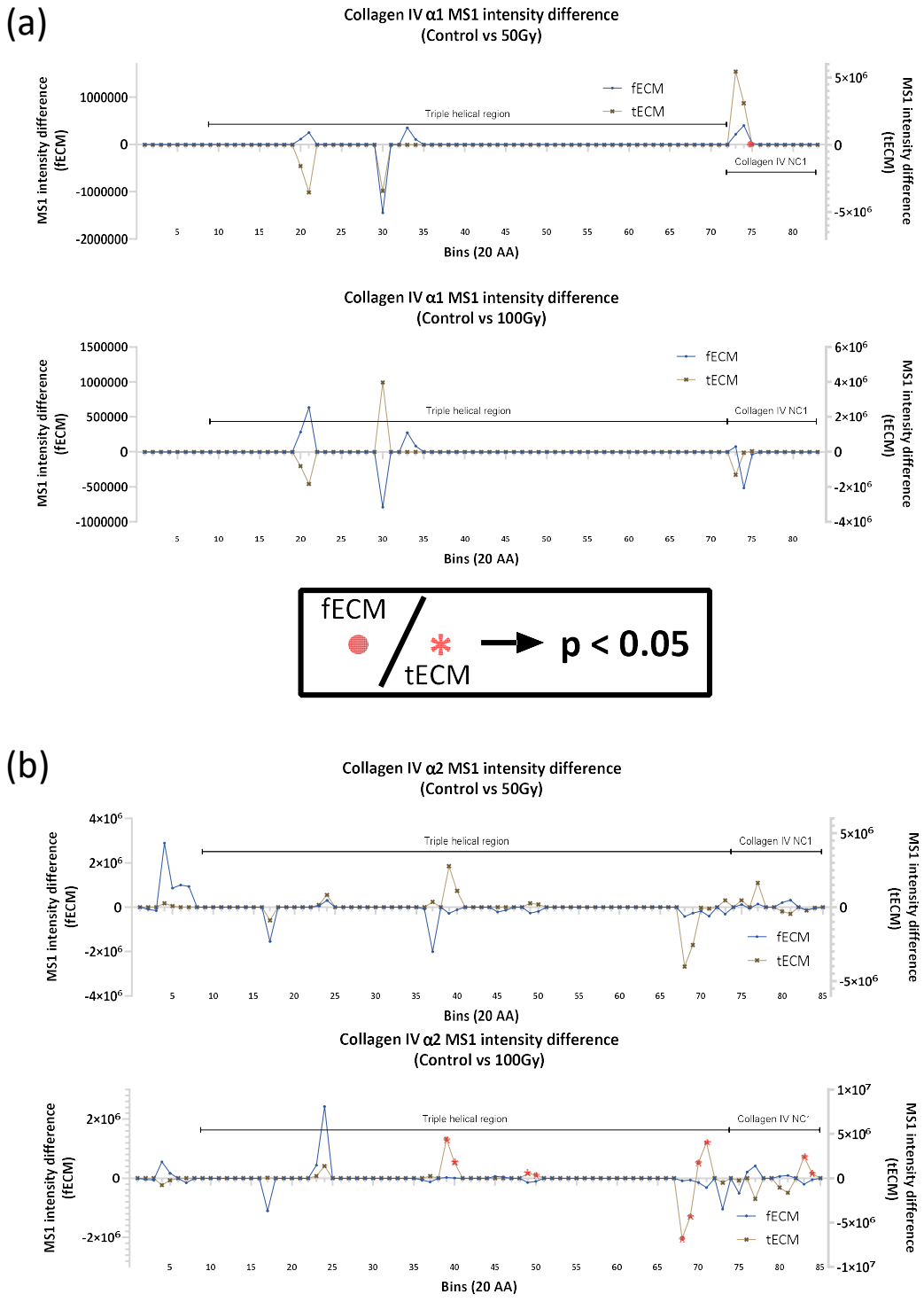


Figure 5.11. PLF analysis for basement membrane protein collagen IV chains α_1 and α_2. (a) α_1 chain of collagen IV did not appear to have significant changes to its proteolytic susceptibility after therapeutic x-ray exposure of up to 100Gy. (b) Conversely, α_2 chain of tECM appear to have significantly altered proteolytic susceptibility after 100Gy of x-ray irradiation. Regardless, proteolysis of collagen IV chains of fECM remain unaffected.

5.3.5.3 Laminin-511

The presence of laminin subunit α_5 suggested that a high probability of laminin-511, which comprises the subunits α_5, β_1, γ_1, is present in both fECM and tECM. Laminin-511 is particularly biologically relevant due to its presence in basement membranes and its expression has been shown to correspond with breast cancer and metastases [15,317]. Therefore, here we chose to analyse the respective subunits from laminin-511 (Figure 5.12a, b, Figure 5.13a).

The α_5 subunit of laminin appear to have most regions affected by 100Gy of ionising radiation exposure compared to the other subunits. Additionally, it seemed to only be coming from tECM samples. The regions seemingly affected included the laminin B domain, domain I and II, and the laminin G-like domains.

5.3.5.4 Nidogen-1

PLF data of Nidogen-1 exposed to 50Gy of ionising radiation suggested that in fECM, most regions where there was coverage had significantly altered peptide intensities (Figure 5.13b), while at 100Gy this effect was abolished. This seemed unlikely and could be attributed to a normalisation error (further discussed in section below). On the other hand, tECM data only had a specific region within the G2 β-barrel domain that was flagged at 50Gy. Furthermore, this coincided with the 100Gy data, where the same domain was affected in tECM, albeit in a slightly different region.

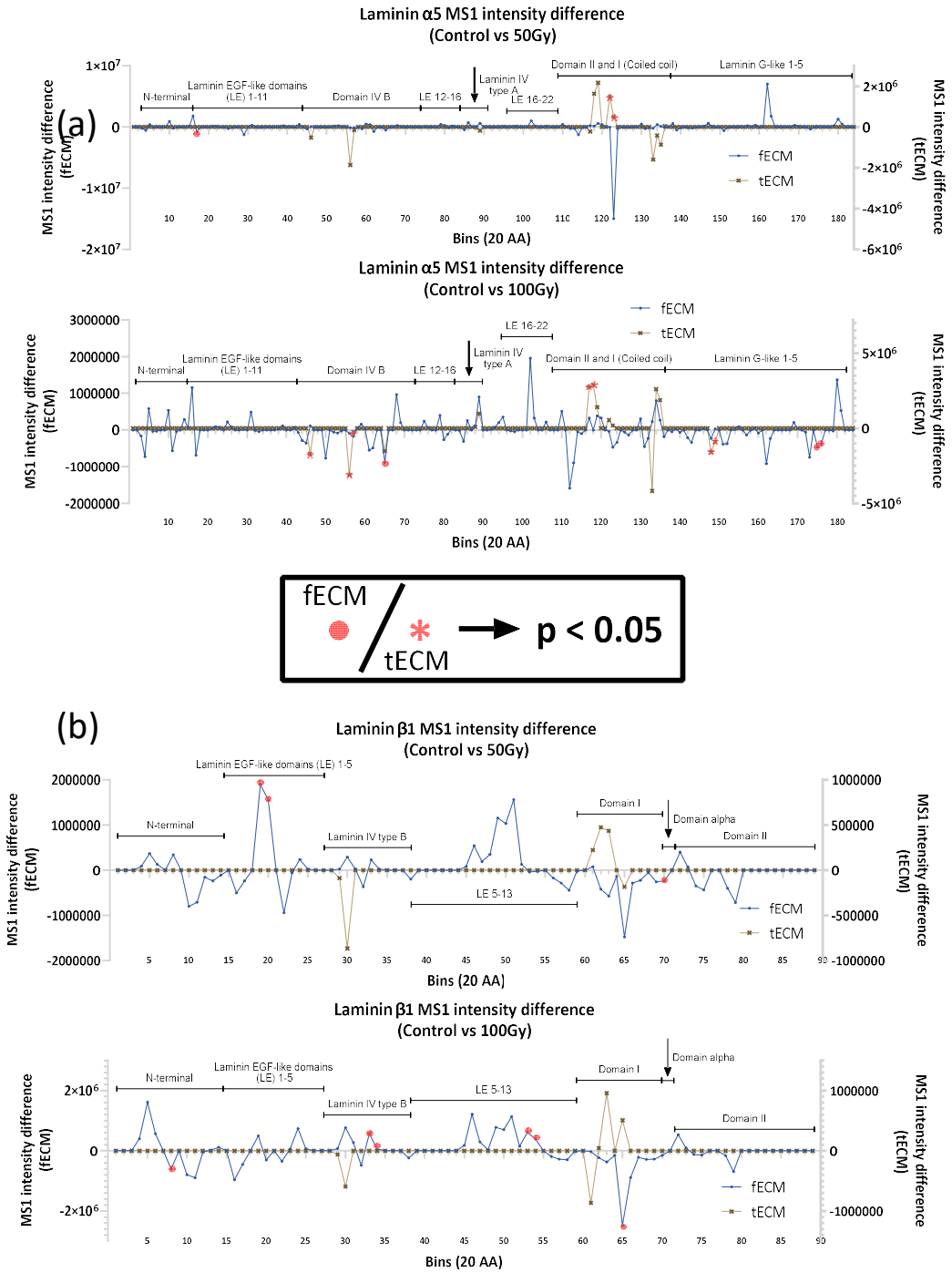


Figure 5.12. PLF analysis of laminin α_5 and β_1. PLF analysis was conducted for laminin subunits α_5, β_1, and γ_1, as they make up laminin-511 which plays important roles in forming a network in the basement membrane. (a) The α_5 chain in tECM, appear to have changes to its proteolytic susceptibility around domain I and II after exposure to 50Gy therapeutic x-rays. At 100Gy, domain IV B and Laminin G-like domains were also affected. The α_5 chain from fECM seemed to be affected only at the higher 100Gy dose of irradiation. (b) β_1 chains from fECM exhibited changes to proteolytic susceptibility in the EGF and alpha domains at 50Gy, while transiting to the laminin IV type B domain at 100Gy. In contrast, tECM did not have any flagged regions.

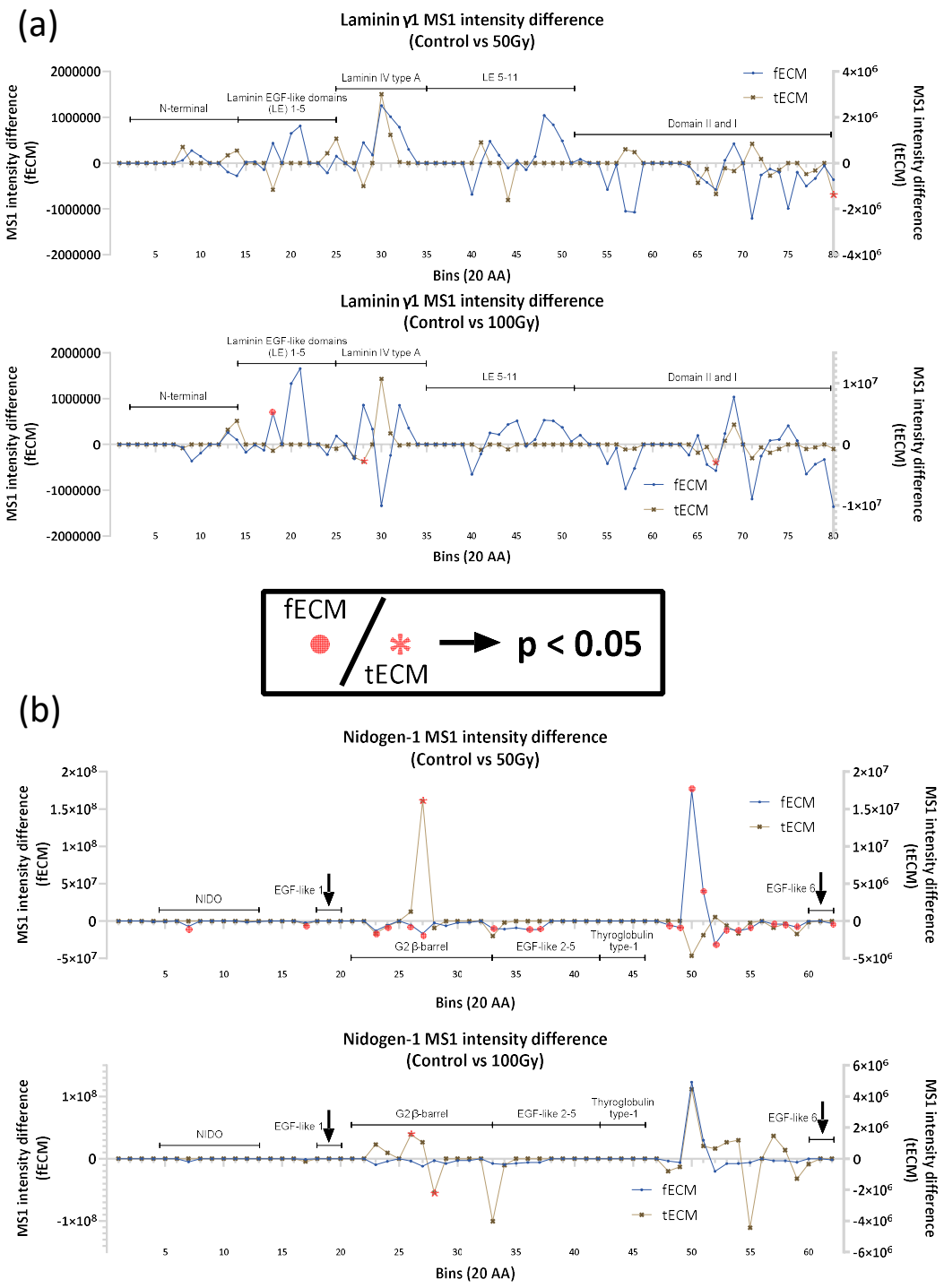


Figure 5.13. PLF analysis of laminin γ_1 and nidogen-1. PLF analysis was conducted for laminin subunits α_5, β_1, and γ_1, as they make up laminin-511 which plays important roles in forming a network in the basement membrane. (a) The laminin γ_1 chain in tECM, appear to be affected in the regions of domain II and I in both 50Gy and 100Gy, while the laminin IV type A domain was affected only at 100Gy. fECM only had a single EGF-like domain affected at 100Gy x-ray exposure.

(b) Nidogen-1 appear to have significant number of flagged regions for the fECM sample at 50Gy across the whole protein sequence but no flagged regions at 100Gy. In contrast, nidogen-1 from tECM only appear to have flagged regions in the G2-β barrel domains that is consistent in both 50Gy and 100Gy x-ray irradiation.

5.3.5.5 Perlecan

PLF analysis of perlecan (Figure 5.14) showed that exposure to 50Gy of therapeutic ionising radiation primarily altered proteolytic susceptibility around the C-terminal domains, including Ig-like C2 domains for fECM, and the C-terminal domain V containing the laminin G-like domains for tECM. At 100Gy, the effects appear to be greater, where the central Ig-like C2 domain had a significant increase in number of flagged regions. The C-terminal domain V remain affected for both fECM and tECM, particularly in laminin G-like 2.

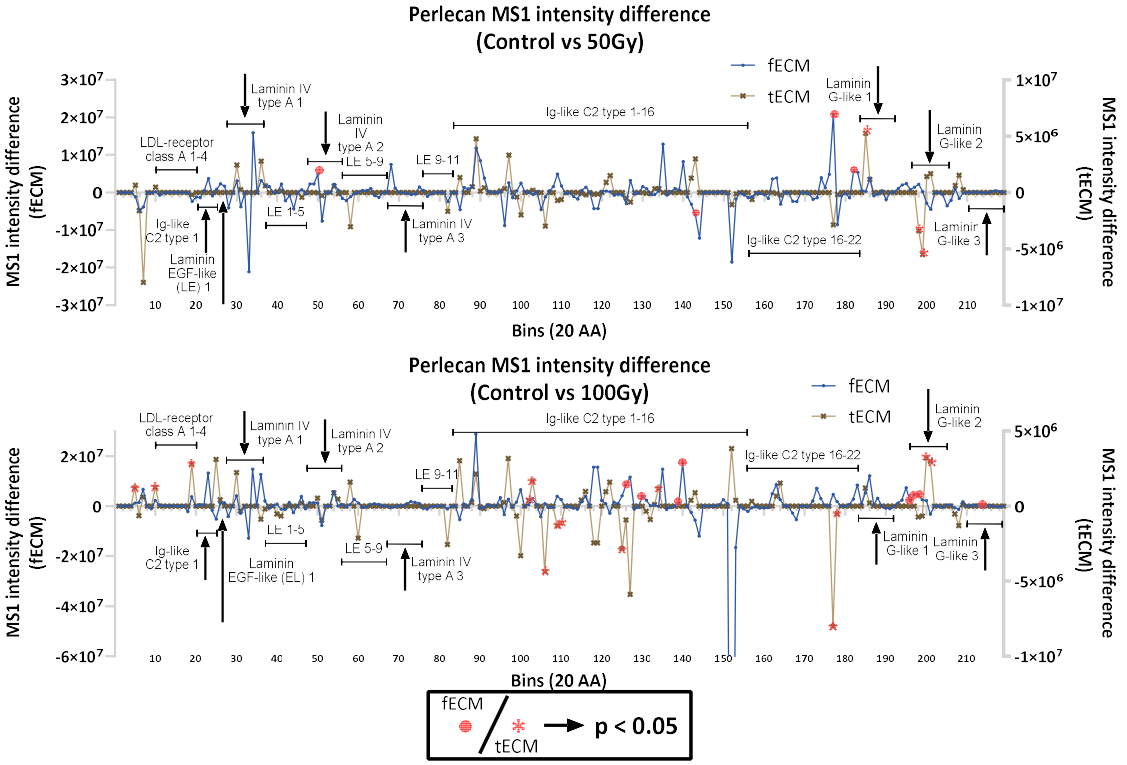


Figure 5.14. PLF analysis for Perlecan, which is an important basement membrane proteoglycan that interacts with collagen IV and laminin. It also binds to integrins, thus serving as a cell-ECM signalling mediator and mechanotransducer. PLF shows that at 50Gy, C-terminal domains like laminin G-like domains and Ig-like C2 domains were primarily affected for both fECM and tECM. At 100Gy, the central Ig-like C2 domain had significant number of regions in both fECM and tECM tissue indicative of changes to proteolytic susceptibility. The Laminin G-like domains were also affected in C-terminal domain V (endorepellin), particularly Laminin G-like 2.

The data here shows basement membrane proteins are susceptible to therapeutic x-rays doses and appear to have altered proteolytic susceptibility in both fECM and tECM. While fECM protein coverage was significantly better than tECM, both fECM and tECM exhibited regions of altered proteolysis, which may be attributed to structural alterations by therapeutic x-rays.

5.4 Discussion

5.4.1 Compositional similarities and differences

This study has quantified and compared the ECM proteins generated by *in vitro* immortalised human mammary fibroblast (HMFU-19) (fECM) with ECM proteins extracted from *ex vivo* breast tissue samples (tECM) using mass spectrometry-based proteomics. To determine if fECM can be utilised as a representative ECM of tECM, the group of proteins identified within major ECM classes was compared between tECM and fECM. Collagens were found to have the highest percentage of common proteins between fECM and tECM, and both also contained collagen-associated proteins that help support cell-ECM communication including periostin [380,381], βig-h3 [382], vitronectin [383], EMILIN-1 [384,385], dermatopontin [386,387]. Given that interactions with other ECM proteins can induce conformational changes [307,388–390], the presence of these key collagen-interacting proteins is particularly important to ensure that the structure of collagens being studied is comparable to when it is *in vivo*. The abundance of those collagen-associated proteins found in both fECM and tECM may indicate similar binding conditions and hence structure for collagens. This was further supported by PLF analysis, where the peptide fingerprint profile obtained was used to indirectly infer structural similarity [225]. In particular, both alpha chains of collagen I have remarkably similar peptide fingerprint patterns in fECM and tECM. This may be evidence of structural similarity of collagen I in fECM and tECM, and thus their biological relevance even when derived *in vitro*.

Compared with solubilised collagen I, specific regions of the peptide fingerprint where collagen I from fECM differs from tECM appear to correspond with high similarity between fECM and solubilised collagen I. This was surprising given that collagen I in fECM and solubilised collagen I are highly dissimilar in structural state (fibrillar vs monomeric – in which chapter 4 had shown induces distinct peptide fingerprints between solubilised rat collagen with fibrillar rat collagen). Knowing that fECM comprises freshly made collagen matrix while the collagen in tECM has probably matured for years, the extrinsic ageing of the tissue may be the underlying cause of altered proteolytic susceptibility of those regions [197]. Matiss et. al. (2020) had shown how extrinsically aged human skin tissue produces different peptide fingerprints of ECM proteins from those that were photo-protected, which may be attributed to photo-induced crosslinking from the absorption of UV radiation [391,392]. Chronological

or intrinsic ageing of tissue may also play a role in altering collagen structure and hence its peptide fingerprint through remodelling of the ECM [393] or the accumulation of advanced glycosylation end products [394]. Regrettably, the work by Matiss et. al. (2020) was unable to obtain reliable data for collagen I due to sample variability, thus, it is not yet certain if these differences in peptide fingerprints observed between fECM and tECM is primarily from ageing of the tissue or due to other protein/cellular influences that may also be at play in tECM. To confirm that those distinct regions are not due to the fibrillar state of collagen I, a study with reconstituted human collagen I gel (similar to the comparison between solubilised rat collagen I and rat collagen gel in chapter 4) may help clarify the underlying cause of those differences. It would also be beneficial to study those specific regions and how they might play a role in ageing or specialised protein binding that only occur in aged tissues.

While it would be tempting to correlate the similarity of peptide fingerprints between fECM and tECM collagen with a high similarity in protein ultrastructure, our work on collagen I rat tail (Chapter 4) has shown that a comparable peptide fingerprint profile may only signify a certain level of structural similarity; Tendon collagen illustrated a greater uniformity in alignment, and larger fibrils compared to those in gel, even though they produced peptide fingerprints of high resemblance. Therefore, while PLF may be used as a quick screening tool to reject ECM proteins with drastically different structure based on distinct peptide fingerprints, matched peptide fingerprints need to be interpreted with caution and may have less power in determining how identical the protein structures are. Other proteases may also be considered for use in conjunction with trypsin to produce a different fingerprint profile, such as using elastase in previous work by Eckersley (2020) [185], that would provide further validation of the structural similarity at different cleavage sites.

Taken together, the compositional analysis by STRING and structural analysis by PLF provide strong indication of the immortalised HMFU-19 fibroblasts' ability to construct an intricate collagen network in its ECM, closely mimicking the complexity and interplay between different collagen associated components in *ex vivo* tissue ECM, thereby resulting in similar collagen molecular ultrastructure.

In addition to collagens, another key finding that supports that fECM may be a representative model of breast tissue ECM is the presence of basement membrane proteins such as laminin, nidogen-1 and perlecan. Their presence indicates that there could be formation of a

functioning basement membrane within the *in vitro* culture [309]. However, PLF analysis highlighted stark differences between the peptide fingerprint of these proteins, with fECM having much better coverage than tECM. This could be attributed to poor solubility and thus reduced proteolytic efficacy of those proteins in tECM [395,396]. The diminished solubility could have resulted from crosslinking [397] or extrinsic ageing [350], which are likely only present in tECM proteins where they had matured over a longer period of time. Such processes, which is absent in fECM, would thus alter the proteolytic susceptibility and hence the peptide fingerprint of tECM basement membrane proteins substantially [185,197]. The abundance level of the protein may also be crucial in affecting the peptide fingerprint. Our results for FN exemplifies how protein abundance could result in distinct peptide fingerprints, and is likely exacerbated by the data-dependant acquisition mode used in our LC-MS/MS sampling, in which low abundance protein peptides are under-sampled [398,399]. More work is required to determine if the underlying reason for such differences is due to poor network formation (which can be confirmed by immunohistochemistry), or the level of protein abundance (which can be confirmed by western blotting).

While the collagens in fECM and tECM were largely similar, there were also certain unique collagens only found in tECM that may be biologically relevant. For example, presence of a small amount of collagen IV α_5 chain was found only in tECM and not fECM. This could imply that while tECM collagen IV network is predominantly ($[\alpha_1]_2$ α_2) like fECM, it might be comprised of other trimers, such as $\alpha_3\alpha_4\alpha_5$ or ($[\alpha_5]_2$ α_6). Although the biological functions of the other trimeric isoforms of collagen IV are not yet well established in the breast tissue, they have been implicated in diseases, such as Goodpasture's disease [400] and Alport syndrome [401]. If its importance in breast tissue is revealed in future work, this may impair the credibility of fECM as a biological representation of breast tissue ECM. Another collagen found in tECM and not in fECM is collagen II, which is secreted mainly by chondrocytes. The presence of collagen II was not expected as it is often mainly found in cartilage tissue. This could be associated with chondrogenic differentiation of mesenchymal stem cells in the tissue, and could be an indicator of abnormal cellular behaviour [402]. Such irregular variations were not observed in fECM, which highlights the benefits of utilising in vitro cell cultures for replicability of experiments.

Besides collagens, the profile of glycoproteins is also distinct between fECM and tECM. One particularly crucial difference is the absence of latent transforming growth factor beta binding protein (LTBP) in tECM. LTBP is biologically important as it plays a key role in TGF-β1 sequestering or regulation, and TGF-β1 has been implicated in many cellular functions and fibrotic diseases [132,369,403]. LTBP is also able to associate with other ECM components like fibrillin microfibrils and fibronectin [345,346], in addition to binding with TGF-β1 and its propeptide to form a large latent complex [404], thus allowing it to sequester TGF-β1 in the matrix. Furthermore, it has been thought to assist with mechano-activation of TGF-β1 activation by binding to cells through integrins, thus transducing cellular forces to deform and activate the growth factor [132,347]. The absence of LTBP in tECM may indicate an absence of TGF-β1 sequestering or regulation, unlike in fECM where it is highly abundant. This is further supported by the fact that TGF-β1 was only detected within fECM and not tECM. The presence of TGF-β1 and LTBP in fECM may thus imply a fibrotic or wound healing ECM phenotype, which would likely differ from a homeostatic tissue ECM like in our tECM biopsy.

Overall, this suggests that while collagens in fECM and tECM share similarities in both structure and compositions, other components of the ECM that may influence biological behaviour can differ. Depending on the research question, fECM may not be as suitable, particularly if re-seeding cells on matrix. The higher abundance of growth factors available in fECM (transforming growth factor (TGF), VEGF, epidermal growth factor (EGF), etc.) may cause substantial differences in cell behaviour compared to tECM [405,406]. Further investigations could look at how the matrix composition of fECM may change under a longer culture period, which may perhaps aid in transition from the growth/wound healing phase into homeostasis that would better represent *in vivo* tissue ECM.

Besides compositions of ECM proteins, proteases in the ECM are important as they govern the remodelling of the ECM, which can release ECM fragments or sequestered growth factors as cytokines and hence affect cellular behaviours [16,407–409]. In fECM, numerous basement membrane degrading MMPs were identified, including MMP-1, -2, -3, and MMP-14. MMP-1 acts as a collagenase and cleaves many types of collagens including collagen I. In addition, it also acts as an activator of MMP-2 and MMP-9. MMP-2, which is present in both fECM and tECM, is able to digest collagen IV present in basement membrane [410]. Similarly, MMP-3 can degrade collagen IV but can additionally affect glycoproteins and proteoglycans like

fibronectin [409], laminin, and perlecan [411]. MMP-14 is a membrane-type MMP that cleaves collagen I while also being able to activate MMP-2 and TGF-β1, and has been attributed to mediate cell migration through ECM degradation [412,413]. The presence of these basement membrane degrading MMPs in fECM may indicate active remodelling processes in our *in vitro* culture, supporting the idea that the fECM is in a wound healing state [414]. In contrast, only MMP-2 was detected in tECM. In addition, TIMP3, which is an important inhibitor of a wide range of substrates, including most MMPs, several a disintegrin and metalloproteinases (ADAMs) and ADAM with thrombospondin motifs (ADAMTSs), was identified in both fECM and tECM [379].

Overall, the presence of a wide variety of MMPs and TIMP3 in fECM suggests healthy regulation and remodelling of ECM components [415] and may represent a wound healing phenotypic matrix. In tECM, the low abundance of MMPs, alongside the presence of TIMP3, signifies a more homeostatic environment [414,416]. However, we note that while we may have detected the presence of those MMPs in fECM, they might not be activated [416]. Enzyme activity assays may need to be conducted to ascertain activity of MMPs and that actively remodelling is occurring in fECM.

5.4.2 Radiation damage to fECM and tECM proteins

PLF was utilised to screen for all (intra- and extracellular) proteins that could be affected by therapeutic x-rays. Using the functional enrichment analysis in STRING database, ECM proteins in fECM were found to be functionally enriched in the group of proteins with altered proteolytic susceptibility. An optimistic interpretation of this result is that therapeutic x-rays are targeting ECM proteins preferentially. This may have implications in redefining the key players of tissue side effects in radiation therapy, which have mainly focused on cellular damage [82], and reiterates the call for increased focus on understanding the biological implications of radiation damage to ECM proteins [1]. However, there are several concerns – this assumes that the PLF analysis is completely unbiased for all types of proteins, which is untrue. There is an overall greater chance for ECM proteins to be flagged by PLF due to false positives. ECM proteins, which are often larger, have more amino acids and hence bins. False positives are likely more prevalent in these large ECM proteins, thus having a higher chance of being flagged. Furthermore, PLF is biased towards large proteins with structural complexity that would be able to affect protease degradation. Smaller proteins may simply be digested

to completion, leaving no room for subtle changes in protease susceptibility to be detected by PLF. While a simple cutoff for a minimum number of flagged regions may be implemented to weed out "one-hit wonders" (proteins with only a single flagged region), this may omit proteins where radiation damage may have largely focused on a single region of the protein. An example from our data could be seen in 50Gy exposed Nidogen-1 from tECM, where a single region in the G2-β barrel domain had a reduction of 87% ± 16% (p=0.001) in MS1 intensity after 50Gy exposure and may be biologically relevant due to its large effect size. Then again, the point in which an effect size becomes biologically relevant is also subjective, and comparison of effect sizes between different regions is not straight-forward considering that the impact on MS1 intensity encompasses not just peptide abundance, but also peptide charge, solubility, and ionisation efficiency [226,417,418]. As such, the author advises caution when interpreting the results of the functional enrichment without further statistical optimisation and considerations. Still, PLF was beneficial in helping to filter out a pool of potential radiation sensitive candidate proteins, in which further investigation may choose to focus on.

Amongst the pool of candidate proteins were collagen I and FN. Thus, the specific response of collagen I and fibronectin in both fECM and tECM to therapeutic x-rays were further investigated given that they were previously studied in purified solution and were found to have significantly altered proteolytic susceptibility (Chapter 3).

The number of regions PLF flagged for collagen was not consistent between fECM and tECM and did not appear to follow the same pattern with increasing x-ray dose nor does it appear to preferentially impact the α_1 or α_2 chain. It appears that while the collagen chain has very similar peptide fingerprint profiles, the underlying stability or sensitivity to radiation seem to differ. This could be associated with the innate differences as discussed above, where tECM contains crosslinking [419], in conjunction with the effects of ageing [166] and general protein interaction differences that, whilst they may not impact the overall peptide fingerprint significantly, may transform the radiation responses of the protein. Nonetheless, the results show that therapeutic dose of x-rays at 100Gy seem to noticeably impact the proteolytic susceptibility of collagens in tECM. In our work from collagen gels (where there were only collagen), there was no observation of such a drastic increase in proteolytic susceptibility at 100Gy, despite having fibrillar structure similar to *ex vivo* tendons (Chapter 4). This makes the

most likely contributing factor to be the other associated proteins binding to collagen in tECM. The exact contribution of collagen binding partners, such as decorin [420], fibronectin [216], and periostin [381], to the radiation response need to be further explored. This would yield better understanding of how ECM proteins interactions affect proteolytic changes in therapeutic x-ray irradiation.

Key basement membrane proteins were, surprisingly, all flagged by PLF to be affected by therapeutic x-rays in both fECM and tECM. This included collagen IV, perlecan, nidogen, laminin. Collagen IV α_2 chain had a significantly altered region around the C-terminal NC1 domain after exposure to 100Gy of therapeutic x-rays. This may implicate collagen IV network formation as this domain has been previously identified to be important in initiating triple helical association between the α chains [421], and is necessary for initiating the formation of the collagen IV network [315,422]. Change in proteolytic susceptibility, which may imply structural modifications, thus could impact the formation of new collagen IV networks [423].

Amongst the subunits of laminin, α_5 appear to have the most regions with altered proteolytic susceptibility after 100Gy of therapeutic x-ray exposure. Laminin IV type B (L4b) domain, domain I and II (laminin coiled coil domain, LCC), and the laminin G-like (LG) domains were found to be affected in the α_5 chain. Intriguingly, the L4b domain in laminin β_1 subunit also had changes to proteolytic susceptibility, but their biological functions are still not well known [424]. The affected LG domains are important for heparin and α-dystroglycan binding [425,426], particularly in the LG 4-5 domains where PLF highlighted in the α_5 chain after 100Gy of x-ray exposure. This may be biologically important as laminin binding to α-dystroglycan has been implicated in muscular dystrophy and is crucial for normal muscle function [427].

In the PLF analysis for 50Gy exposed nidogen-1 in fECM, 21 out of 32 regions were found to have significantly altered peptide intensities, while at 100Gy, this effect was completely abolished (i.e. no regions were flagged). This was highly unusual which warranted further investigation of its peptide fingerprint. Expectedly, it was found that 50Gy MS1 intensities were consistently higher than control in 30 out of a total of 32 regions with detectable MS1 intensities (Appendix 7.5), which could arise from a normalisation error. As the data was normalised based on the means, the exceptionally low MS1 intensity at bin 50 (AA 980-1000) (~10x smaller than control/100Gy) for the 50Gy samples could skew the normalisation factor for the data to be slightly higher. As such, the flagged regions are likely false positives for the

50Gy data for fECM. On the other hand, tECM data only had a specific region within the G2 β-barrel domain that was flagged at 50Gy that further coincided with the 100Gy data. This gave much higher confidence of true proteolytic alteration at this specific region. This particular region was found to bind strongly to perlecan [323,428], making it an appealing target for future investigations into its binding affinity with perlecan with exposure to therapeutic x-rays.

X-ray exposure on perlecan revealed that 50Gy was sufficient to alter proteolytic susceptibility around the C-terminal domains, including Ig-like C2 domains for fECM, and the C-terminal domain V containing the laminin G-like domains for tECM. The radiation impact on the laminin G-like domains, or domain V, in perlecan is particularly intriguing given that it can be endogenously cleaved by MMP to produce endorepellin, which is a matrikine [429]. Endorepellin can bind with cell integrins (such as $\alpha_2\beta_1$ integrin on endothelial cells) and interact with growth factors like vascular endothelial growth factor to promote tissue repair [429,430]. Given that PLF was able to detect alteration to the proteolytic susceptibility of the matrikine domain, it would be interesting to see if x-ray exposure could modify the production of these MMP mediated peptide fragments, and, given their highly potent ability to influence cell behaviour, how they may be connected to fibrosis.

5.5 Conclusion

This study successfully quantified and compared the ECM proteins generated by *in vitro* immortalised human mammary fibroblast (HMFU-19) (fECM) with ECM proteins extracted from *ex vivo* breast tissue samples (tECM) using LC-MS/MS proteomics. Comparison made between fECM and tECM in the various classes of ECM protein found that neither were considerably more diverse in ECM proteins, and both contain similar collagen and collagen-associated protein families. The evidence supports the hypothesis that fECM may provide a sufficiently complex collagen-focused ECM comparable to *ex vivo* tissue tECM. That said, the composition of ECM proteins in fECM appear to mimic a wound healing phenotype while tECM leaned towards a more homeostatic environment. PLF confirmed the second hypothesis, that there were inherent differences in proteolytic susceptibility of crucial ECM proteins, including FN and basement membrane proteins, as detected in the significantly different peptide fingerprint profiles between fECM and tECM. Finally, PLF also proved valuable as a screening tool to detect potential radiation sensitive proteins through

proteolytic changes of proteins identified in LC-MS/MS shotgun proteomics. In addition to collagen and FN, proteolytic susceptibility of key basement membranes were also altered with therapeutic x-ray exposure, with numerous biologically important domains found to be impacted. However, these proteins and their affected domains appear to be distinct between fECM and tECM. Future experiments could narrow down on those proteins and domains to further explore the impact of x-rays on their binding affinities and biological repercussions.

6 Discussion and future perspectives

6.1 Summary of thesis

Breast radiation therapy is widely used and the pathology of its side effects to normal breast tissues are still not well understood. The ECM had been implicated in regulating many biological processes but have not been well studied for its responses to therapeutic radiation and how they may contribute to the side effects of radiation therapy.

The overarching aim of the thesis was to investigate and characterise the impact of therapeutic x-ray on extracellular matrix proteins using biochemical techniques. In conjunction, the novel LC-MS/MS analytical tool (peptide location fingerprinting (PLF)), previously developed by Eckersley in the lab, was further improved. Firstly, peptide spectrum counting was replaced with the precursor ion MS1 intensity as a better proxy for peptide abundance [226]. Secondly, peptide overlaps between multiple bins were better accounted for by doing a weighted sum of the peptides that overlap into each bin. These improvements not only help enhance the sensitivity at which PLF is able to detect localised changes to peptide abundance, but also enable bin sizes to be arbitrarily chosen without impact on the accuracy and detail of the analysis.

To address the aim of the thesis, three experimental studies were conceived, each hoping to shed light on a different aspect of the molecular impact of therapeutic radiation on ECM proteins. The first study focused on understanding how therapeutic x-ray impact isolated and solubilised ECM proteins collagen I and plasma fibronectin (pFN), which are highly abundant and present in many tissues. Mechanistic insights were drawn into how therapeutic x-rays impact these ECM proteins differentially based on their structural and amino acid compositional differences. The study was able to show profound x-ray induced fragmentation in the elongated, solubilised monomeric collagen I that were characteristic of reactive oxygen species-mediated protein fragmentation [227]. In contrast, the impact of therapeutic x-rays on the globular shaped pFN was more subtle; X-ray exposed pFN had localised changes to structure which appeared to be concentrated on binding sites. Detection of therapeutic x-ray impact on pFN was only possible through PLF by observing differences in the peptide yields from proteolytic degradation before and after x-ray exposure. The alteration of proteolytic susceptibility of pFN was further found to have functional consequences with an increase of

collagen I binding affinity to irradiated FN, which was postulated to impact collagen fibrillogenesis [241,253].

While profound x-ray remodelling was observed in purified collagen I, the artificial, solubilised state of purified collagen makes it difficult to draw relevant biological conclusions. Hence, the second study aimed to test if similar remodelling could still be detected in collagen I exposed to therapeutic x-rays in the fibrillar state: that is, in collagen gel and *ex vivo* tendon. The results show that fragmentation that was observed in solubilised collagen were absent in gel and tendons, which implied the fibrillar structure may have increased collagen I resistance to fragmentation by therapeutic x-rays. However, collagen I chains in both solubilised and fibrillar collagen were all prone to X-ray induced changes in tryptic proteolysis susceptibility, indicating possible structural changes in localised regions of collagen I. Tendons were also found to exhibit greater number of regional alterations, which may be in part due to their complex hierarchical fibrillar structure formed with other tendon-associated proteins that is not present in solubilised collagen and reconstituted collagen gels. This raises further questions on how collagen I within a more complex environment, such as the breast tumour microenvironment, may be remodelled by therapeutic x-rays. The presence of other cellular components, such as proteases or MMPs, that can also remodel collagen is prevalent in these tissues [431–433], and its interplay with the impact of therapeutic x-rays on collagen may have biologically relevant consequences.

Therefore, the third and final study aimed to investigate the radiation impact in two complex ECM proteomes, namely *in vitro* mammary fibroblast derived ECM (fECM) and *ex vivo* breast tissue (tECM). Compositional studies of the two proteomes revealed closely related collagen and collagen-associated protein families as well as crucial components of the basement membrane present in both ECM. However, the ECM produced indicate differences in cellular biological states – cells in fECM appear to be in a wound healing state with active ECM remodelling, while tECM seem to be a less active, mature ECM with cells under homeostasis. This can be seen by the presence of growth factors, higher abundance of fibronectin, and appearance of numerous ECM remodelling MMPs in fECM [16,207,257,414,434] that were not found in tECM. PLF was utilised to screen and detect innate proteolytic, and possibly structural, differences between the same ECM proteins from fECM and tECM while also detecting the impact of x-ray exposure across all proteins. The widespread alteration of

proteolytic susceptibility in ECM proteins from therapeutic x-ray exposure allude to the importance of ECM proteins' role in the side effects of radiation therapy. This study also highlighted several ECM proteins, including basement membrane proteins like collagen IV, laminins, perlecan, and nidogen-1, and specific regions in their proteins sequence in which they may be impacted by therapeutic x-rays. Although, instead of a single damaged ECM protein governing the radiation response of tissues, it is postulated that a collective of radiation-exposed, structurally-altered ECM proteins may contribute their own small, but crucial piece to the overall physiology of radiation, either through matrikines production [429], a change in mechanical stiffness [142], or through alteration in protein binding affinities [435].

Overall, the three studies have shown clearly that ECM proteins, both solubilised and in a complex proteome, are profoundly altered by therapeutic x-rays and it can play crucial roles in mediating long-term radiation side effects. The work on collagen I also highlighted the importance of ultrastructure in influencing the outcome of large, multi-component ECM proteins' exposure to therapeutic x-rays. The capabilities of PLF were also showcased in its ability to work well with complex mixtures, thereby eliminating the need for costly and inefficient protein purification for biochemical characterisation techniques. While the aims of the thesis were met, the data presented here is not without limitations.

6.2 Limitations and future work

To address the limitations of the work presented in this thesis, further work is required and are elaborated below.

In the purified protein work, the concentration of protein when exposed to radiation was arbitrarily chosen to be 1mg/mL. The effect of concentration on the extent of radiation damage is has not been considered in this study. Given that radiation dose is determined by energy absorbed per unit mass (or volume since the density is approximate the same ~ water), the amount of energy absorbed per molecule changes with different concentrations, which would in turn affect the degree of change in each protein molecule. Preliminary results show that at lower concentrations, collagen I was found to be more extensively fragmentated when exposed to the same amount of x-ray dose (Figure 6.1). This would be interesting to further investigate as concentration of collagen differs in different tissues [436,437] and may thus respond differently even when exposed to the same amount of therapeutic x-rays. Further work can utilise spatial proteomics (laser-capture microdissection with LC-MS/MS [382,438,439]) to look at differences in proteolytic changes between high ECM concentration regions (such as the breast stroma) to lower concentration regions (such as the lumen of mammary glands).

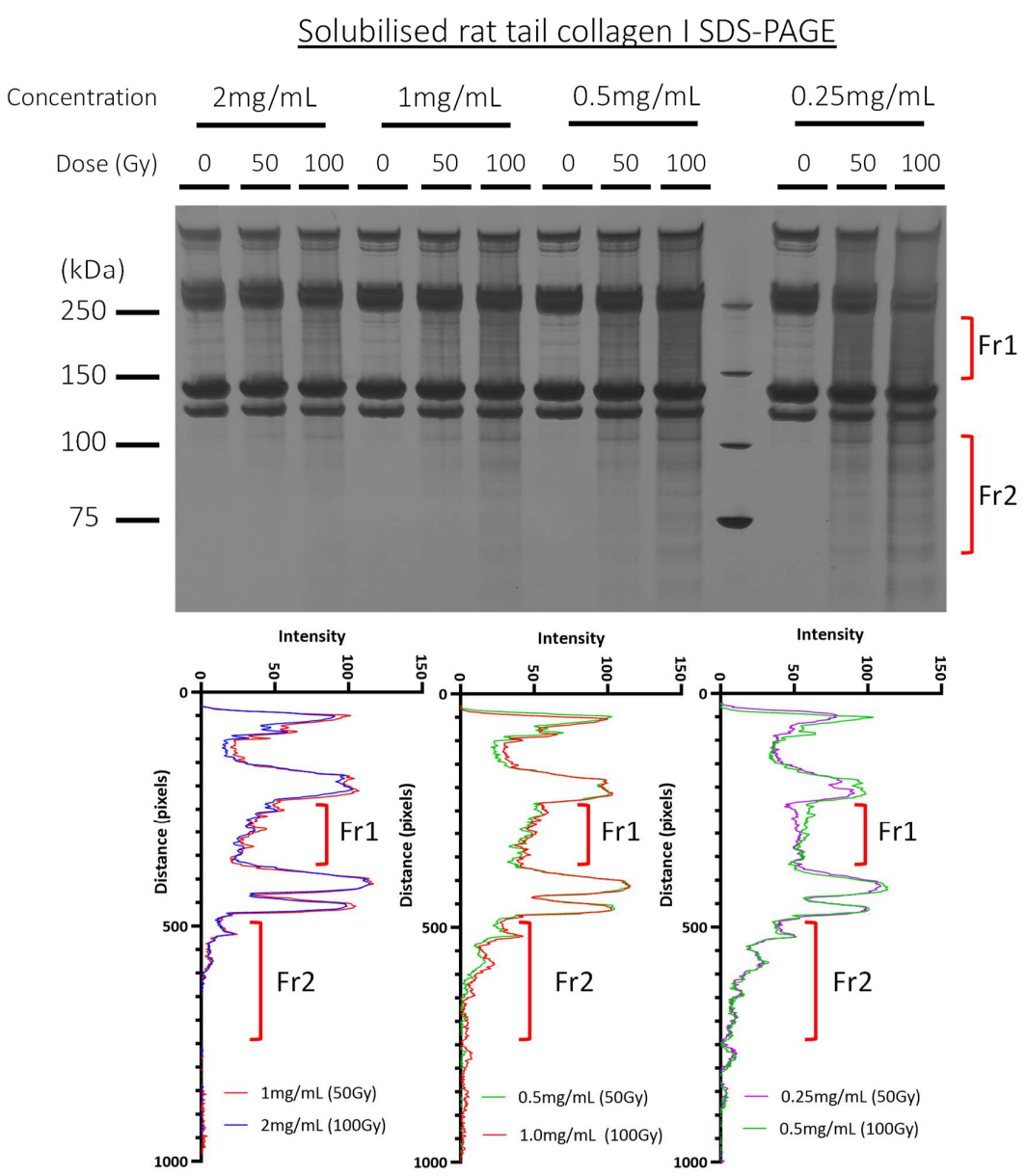


Figure 6.1. SDS-PAGE of solubilised rat tail collagen I exposed to therapeutic x-ray doses of 50Gy and 100Gy at varying concentrations diluted in 50mM acetic acid. Decreased concentration resulted in increased fragmentation detected by the increased smearing and background staining at regions between β and α bands (Fr1), as well as below the α bands (Fr2). The intensity profile appears to be consistently reproduced in a dose-concentration ratio dependant manner. Three comparisons were made that had the dose-concentration ratio. (Left) 1mg/mL exposed to 50Gy produced the same fragmentation profile as 2mg/mL exposed to 100Gy. (Centre) Similarly, 0.5mg/mL exposed to 50Gy returned the same graph as 1.0mg/mL exposed to 100Gy.

Using an ELISA assay, it was shown in chapter 3 that x-ray exposed plasma fibronectin appear to bind with greater affinity to collagen I. However, an underlying assumption was the antibody binding affinity to fibronectin was not affected by therapeutic x-rays, which may not hold if there were structural changes in those regions. While controls were carefully implemented in the work done to test for variations in antibody binding (using direct pFN binding to the plastic substrate without collagen), saturation of absorbance signal was reached much quicker for these wells (1-2 minutes) compared to the other wells (40 minutes) despite using the lowest concentration of pFN. This makes the final comparison of the controls unreliable due to oversaturation which would conceal the slight differences (if any) in antibody binding to x-ray exposed pFN. Further work could be conducted to optimise separately an experiment to determine if there are changes to antibody binding, or to repeat the pFN-Collagen binding experiments using the surface plasmon resonance technique [440,441], which eliminates the need for antibodies while improving sensitivity.

While PLF analysis assumed that MS1 peptide changes were due to changes in proteolytic susceptibility, there could also be chemical changes associated with the peptides that are not captured. Chemical changes could impact the ionisation efficiency of peptides, which would alter the efficiency of detection of those peptides and hence its MS1 peptide intensity [417,442]. Future work can focus on optimising a time course experiment for tryptic digestion times. There would inspire greater confidence in regions where differences (between control and x-ray treated groups) in MS1 peptide intensities increase with digestion time as this would signify those differences are indeed influenced by proteolytic susceptibility. PLF can also be utilised in conjunction with other specialised mass spectrometry approaches such as cross-linking [443] or hydroxyl radical footprinting [259] mass spectrometry to provide a more holistic structural information on the proteins of interest.

There is a need to determine if current statistics for PLF is sufficiently robust to prevent false positives. The Bonferroni correction applied across treatment groups may not be sufficient due to the high number of bins per protein, especially larger proteins [444,445]. However, applying the correction for by considering all bins within each protein as a comparison would exponentially increase the correction size, resulting in overly conservative results and an issue with false negatives instead [446]. The work here had decided to adopt the former, thus requiring careful analysis and further validation with other biochemical methods. Further

work can be done to improve the statistical power for PLF analysis and explore how to correct for false positives, or to take into account practical significance (effect sizes) when considering regions of significant MS1 peptide intensity alteration [447].

Our work on collagen I has focused on alteration on individual α chains and their biological impact. However, the collagen I fibril *in vivo*, as argued by Zhu et.al. (2018) [448], contains three layers of complex binding motifs governed by the primary sequence of the collagen α strands, post-translationally modified residues on the amino acids of the chain, and finally the staggering of the α_1 and α_2 chains along the triple helix. These three layers of complexity allow different combinations of possible binding motifs across chains which could be altered with changes to proteolytic susceptibility. The study here is unable to provide direct evidence at which level the binding is subjected to most alteration *in vivo*. Further work is needed for looking at a combination of proteases or protein-protein interactions that work at the three different levels to elucidate which is most affected. For example, one could use the activity of MMP-1 to determine if its specific cleavage sites were compromised by therapeutic x-rays, given that the cleavage sites needs to be partially unwound for initiation of MMP-mediated collagen degradation [248,412].

Many other ECM proteins were highlighted to be impacted by therapeutic x-rays in the compositional study comparing the *in vitro* ECM with tissue ECM could be further explored with additional resources. These proteins and interactions include laminin-511, whose network formation and integrity of the trimeric structure is crucial for cell-ECM communication [317]. Another ECM protein of interest is perlecan and studying how therapeutic x-rays can impact its binding affinities to different components of the ECM would be crucial given that it has numerous binding partners [319], and to also explore possible matrikine production by therapeutic x-ray exposure [429]. Similarly, fibronectin, which also binds to many ECM and cellular proteins, could be further studied for its binding affinity to tenascin-C [449], heparin [450] and fibrin [434] post x-ray irradiation.

Besides looking at individual ECM proteins, other tissues that are prone to radiation damage could be further explored. Examples would be heart or lung tissues, which as mentioned in the introduction, are particularly sensitive to radiation and can often be exposed in radiation therapy to the breast [4,47,80]. Work on such tissues would complement the dataset obtained here for breast tissue, given the similarity in composition of the stroma with

collagens and elastic fibres as major components and them playing important mechanical roles in lung and heart tissues [451–456]. This would improve the confidence of the results obtained through this dataset and help extend the impact of this study beyond breast tissue and into cardio-oncology. Additionally, the mechanics of these tissue could be explored with optimisation of the AFM by addressing some of the challenges faced here in sample realignment and degradation. Further AFM imaging of these tissues may also highlight some of the underlying collagen ultrastructure, like shown in previous work by Graham et. al. [457]. Elucidating visual changes, if any, of the collagen molecular structure by ionising radiation exposure would also help further our understanding of the mechanism in which collagen is damaged in those complex tissues.

Finally, while fibroblasts were specifically chosen for the *in vitro* cell culture work in chapter 5, production of ECM *in vivo* is contributed by other cells as well, including adipocytes [458,459], chondrocytes [460], and macrophages [461]. The crosstalk between these cells can further influence the type of ECM being produced. As seem from work by Witherel et. al. (2021), adding pre-conditioned media of M2 macrophages appear to enhance ECM secretion of dermal fibroblasts and promote a fibrotic ECM phenotype [462]. Future work can examine ways to improve the *in vitro* ECM model by co-culturing cells or using organ-on-a-chip methods to combine cells [463,464] and recreate an ECM model that could be more representative of tissues.

6.2.1 Conclusion

In summary, the work presented here had addressed the aims and hypotheses set out at the start of the PhD. By combining conventional biochemical techniques with the updated PLF analysis tool, the damage of therapeutic x-rays on collagen I, fibronectin, and numerous ECM biomolecules were investigated and revealed. PLF was able to locate specific regions of FN with altered proteolytic susceptibility from therapeutic x-ray exposure. Work has confirmed the change in proteolytic susceptibility in FN binding domains may be associated with altered binding functions. Experimenting on different scales of the collagen I protein also helped identified the importance of protein structure in governing the outcome of therapeutic x-ray exposure, as well as the interpretation of PLF data. The proteome composition of *in vitro* mammary fibroblast derived ECM was also validated and compared with *ex vivo* breast tissue. PLF was useful for validating structural similarity and to screen bioinformatically for proteins

altered by therapeutic x-rays. While PLF has proved to be a powerful tool in investigating subtle proteomic changes, there are limitations and various factors to consider when interpreting results from PLF. Overcoming the challenges of PLF, while uncovering the biological/functional consequences of ECM proteins identified in this thesis to be altered by therapeutic radiation, would further expand our understanding of ECM proteins in the pathophysiology of breast radiotherapy side effects.

7 Appendix

7.1 Peptide fingerprint of human collagen and fibronectin

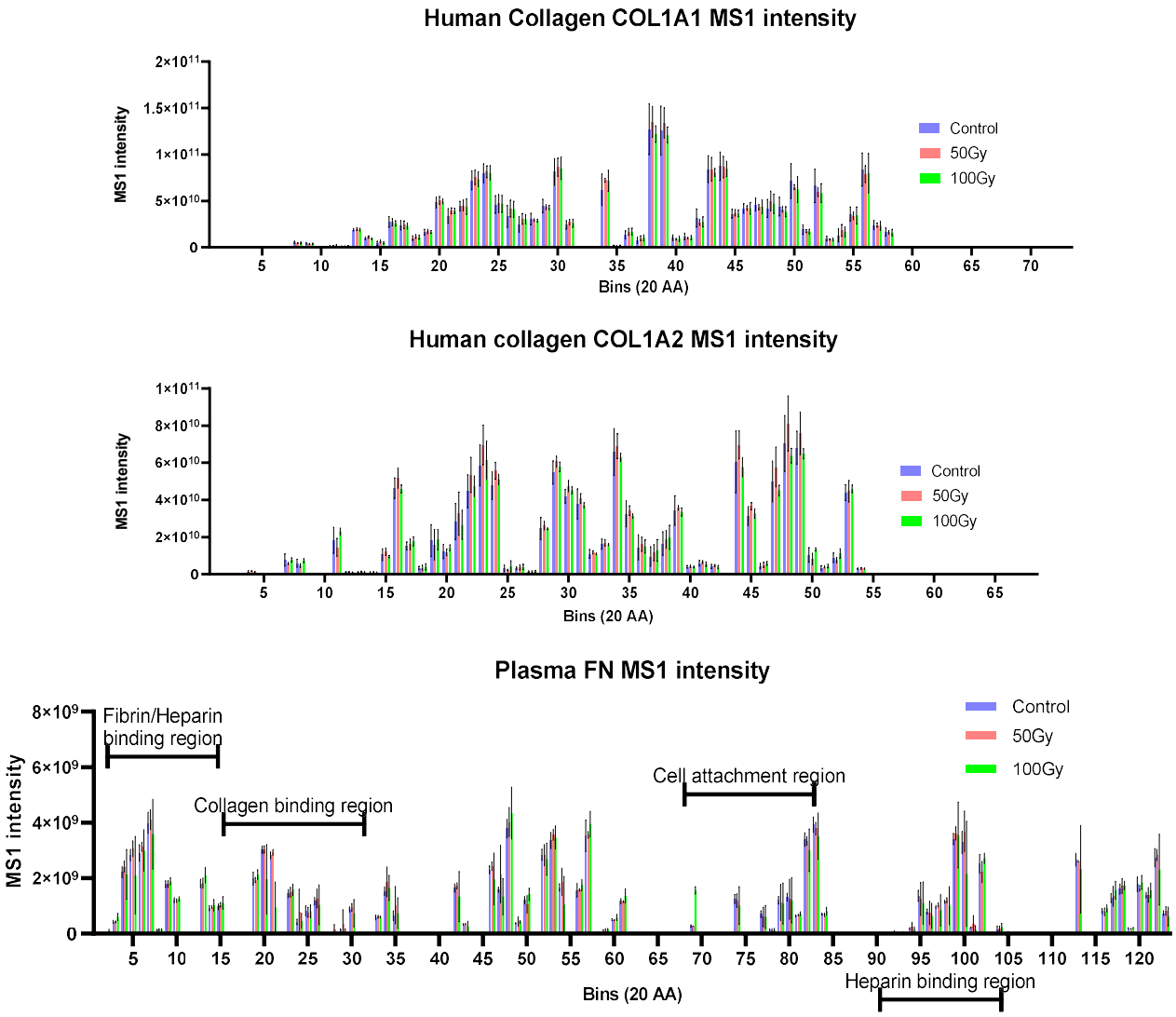


Figure 7.1. Peptide fingerprint for human collagen I α_1/α_2 chain and human plasma fibronectin. Mass spectrometry MS1 intensity of peptides were summed across the primary sequence of the protein. The protein was arbitrarily subdivided into bins of 20 amino acids, and the MS1 intensity of peptides falling within those regions are weighted according to their overlap and summed. Each bin contains three bars representing MS1 intensities for control, 50Gy, and 100Gy samples respectively (n=5 each). The pattern of intensities across the bins are unique to the protein and sample, hence termed a peptide fingerprint.

7.2 Peptide fingerprint of rat collagen I in solubilised, gel, and tendon collagen

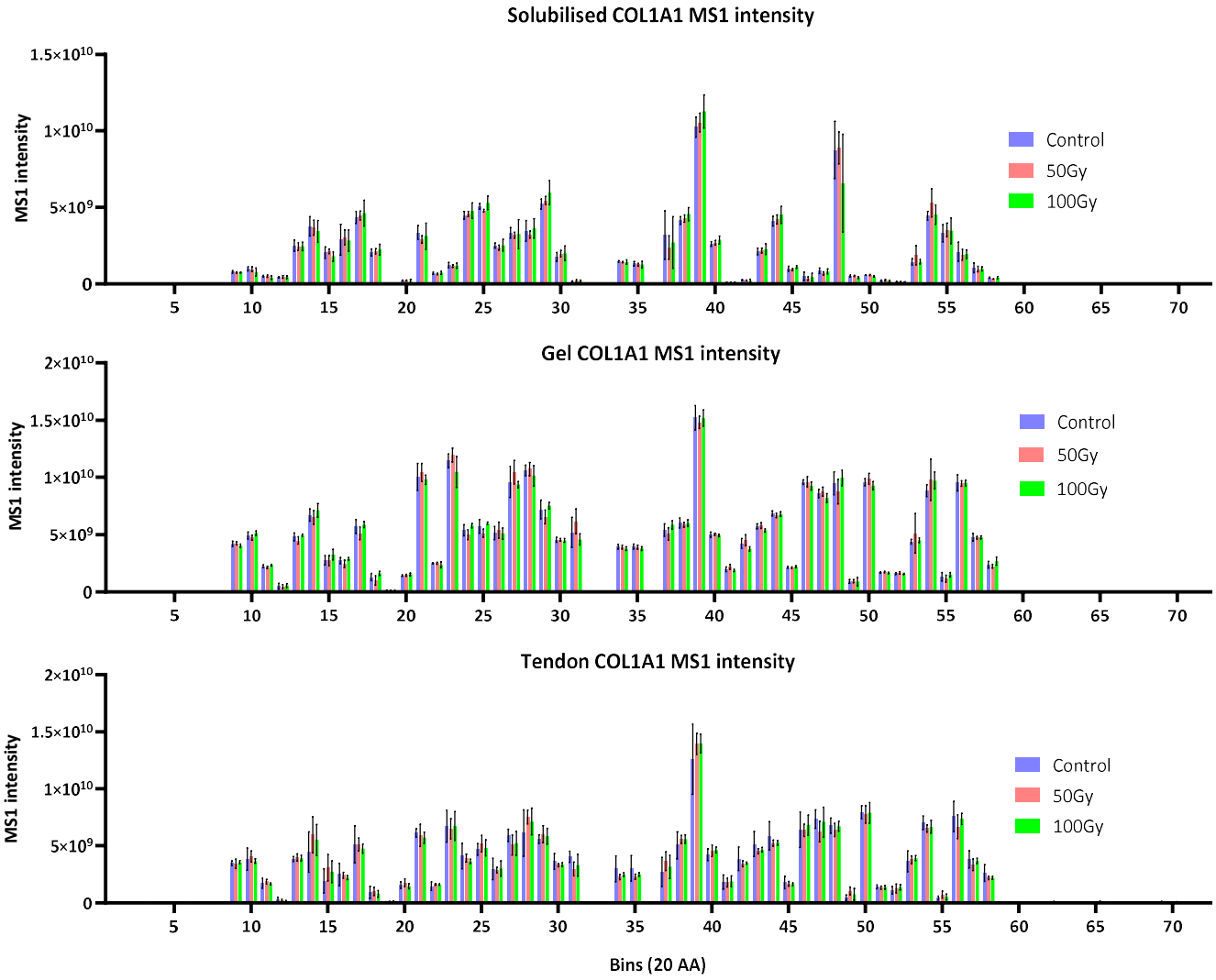


Figure 7.2. Peptide fingerprint obtained by LC-MS/MS of rat collagen I α_1 chain in solubilised, gel, and tendon collagen I.

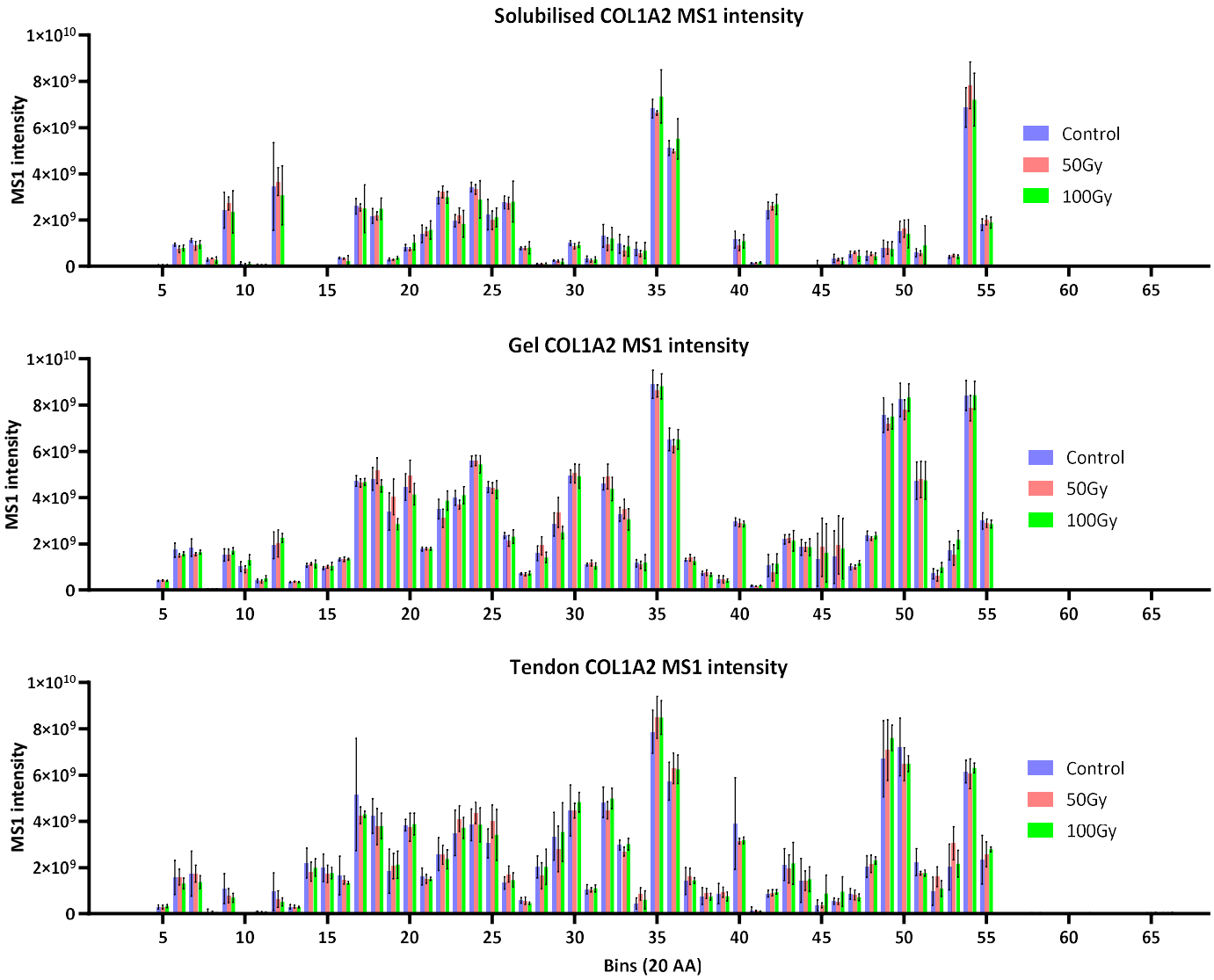


Figure 7.3. Peptide fingerprint obtained by LC-MS/MS of rat collagen I α_2 chain in solubilised, gel, and tendon collagen I.

7.3 H&E staining of serial cryosections of *ex vivo* breast tissue

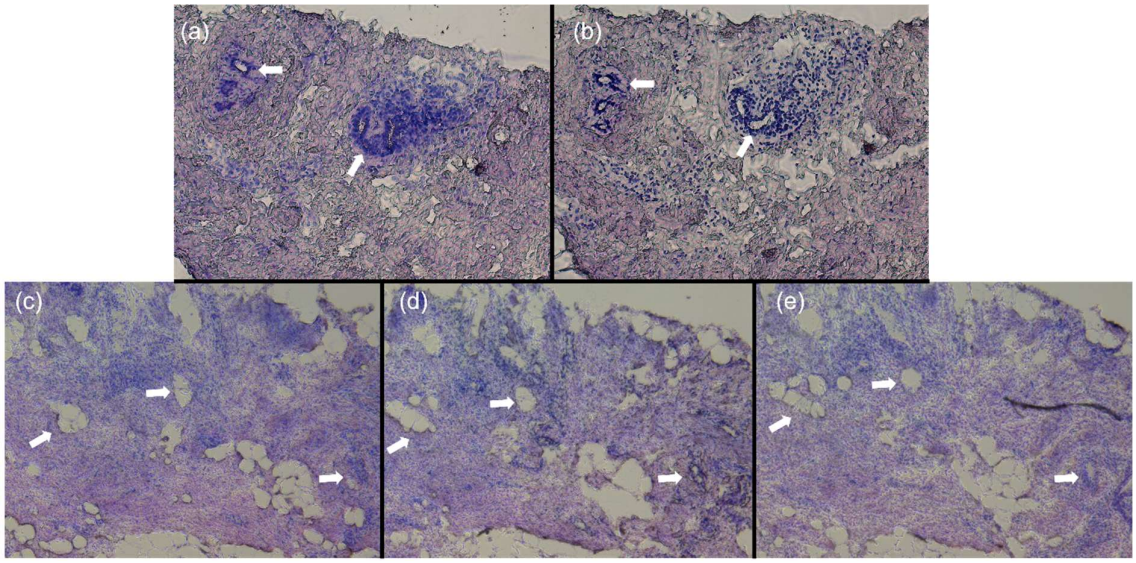


Figure 7.4. H&E stained cryosections of ex vivo breast tissue (biobank). (a/b) Comparison between original cryosection (a) and a sequential section treated with extraction buffer (NH_4OH and Triton X-100) overnight at 4°C and washed with PBS+ (b). Cells were still observed in treated slide (b) seen in dark blue staining of nucleus from H&E stain around the lobular structures.

(c-e) Comparison of morphology and features of sequential cryosections of 20µm thickness. Major features including lobules and adipose tissue were consistent between sequential slices, indication that tissue composition was unlikely to alter significantly between 20µm thick sequential cryosection slices.

7.4 Estimating absorption of UV radiation and ionising radiation by fibronectin

Absorption of radiation by a material can be described by the following equation:

$$I_x = I_0 e^{-\frac{\mu_{en}}{\rho}x}$$

$$I_x \rightarrow Intensity\ at\ depth\ x\ in\ material;\ I_0 \rightarrow\ Incident\ intensity$$

$$\mu_{en} \rightarrow energy\ absorption\ coefficient;\ \rho \rightarrow density\ of\ material$$

$$x \rightarrow distance\ travelled\ in\ material$$

The value of linear absorption coefficient, or $\frac{\mu_{en}}{\rho}$, for UVR and 150keV x-rays is estimated to be approximately 0.1m^{-1} and 2.764m^{-1} respectively [234,235]. Here, the mean x-ray energy is assumed to be 150keV (half of peak tube voltage 300kVp). As samples were approximately microlitres in volume, we assume that samples are approximately 1cm (or 0.01m) in thickness. Using the above equation, the percentage intensity of UVR or x-rays that would be absorbed by water can be approximated to 0.1% and 2.73% respectively. The paper referenced in the section by Sherratt et. al. describes total incident UVB that was absorbed is 3% [152], implying that 2.9/3 *100 = 97% of all absorbed UVB is direct absorption by protein molecules (i.e. fibronectin).

For x-rays, a very approximate calculation can be made based on probability since they can be absorbed equally by water and by protein molecules. Assuming water molecules and protein molecules are homogenous, the only factor considered for x-ray to react with the protein directly is if it encounters one randomly based on chance, which can be computed by considering the total volume occupied by protein vs total volume occupied by water molecules.

Given the fibronectin concentration used was 1mg/mL, converting to mol/L: 10^{-3}g/mL * 1/(5.5*10^5) mol/g [fibronectin dimer is approximately 550kDa] = 1.82*10^{-6} mol/L. Comparatively, pure water has molarity 55.56 mol/L. From this, the ratio of number of water molecules to FN molecules is calculated to be 1: 3.3*10^{-8}. Now, there is a need to account for the volume occupied by FN compared to a water molecule. The size of water molecule is around 0.135nm in radius [465] and FN is 61nm in length (L), 2nm in diameter (d) for each FN monomer [466]. Assuming water molecules are spheres ($V = \frac{4}{3}\pi r^3$), and FN is a cylinder ($V =$

$\frac{\pi d^2}{4} L * 2$) (multiply by 2 for dimer) then the ratio of volumes (water molecule: FN molecule) can be found = 1 : $3.72*10^4$. Combining the ratio of volume to the ratio of number of molecules, the approximate chance of photon being absorbed by protein directly instead of water is 0.123%. This implies that 95.5% of x-rays would be first absorbed by water, thus it is likely that the effect of x-rays is mostly indirect.

7.5 Nidogen-1 fECM peptide fingerprint

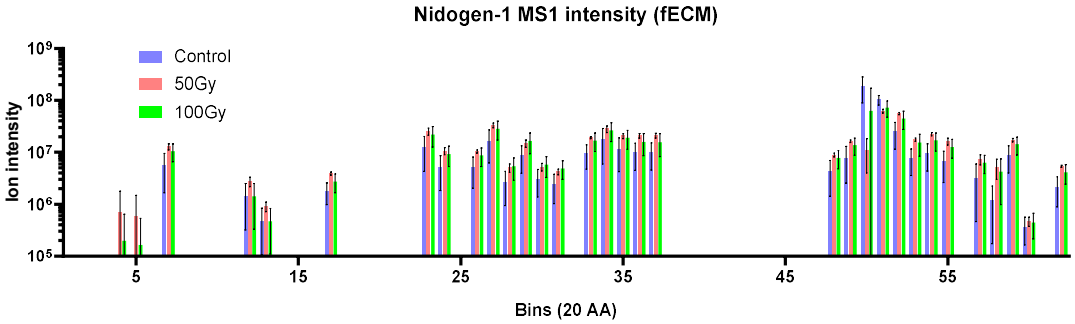


Figure 7.5. Peptide fingerprint of Nidogen-1 from in vitro fibroblast derived ECM (fECM). The 50Gy fingerprint was found to be consistently higher in MS1 intensities for 30 out of 32 total regions with identified peptides, which is unlikely due to therapeutic x-rays effects and could be associated with error in normalisation.

7.6 Python code for peptide location fingerprinting

```
"""
Created on Tue Mar 22 11:21:13 2022

@author: rj
"""
import xlsxwriter as xw
import tkinter,sys,itertools, warnings,win32com.client, matplotlib,os, io,
math,json
from Bio.SeqUtils import seq3,seq1
from Bio.SeqUtils.ProtParam import ProteinAnalysis
from tkinter.filedialog import askopenfilename
#from bioservices.uniprot import UniProt
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
from tqdm import tqdm
from Bio import SeqIO

def getSeq(protein_id):
    b=u.retrieve(protein_id,frmt='fasta')
    b = b.split('\n')[1:-1]
    protein_seq = ''.join(b)
    return protein_seq

def locatePeptides(seqData,locationPos):

    fragPos=list(seqData['pos'].dropna())
    t=[list(np.zeros(len(sampleList))) for _ in range(len(locationPos))]
    for j in range(len(fragPos)):
        for i in range(len(locationPos)):
            if fragPos[j][0]>=locationPos[i][0] and
fragPos[j][0]<=locationPos[i][1] and fragPos[j][1]>=locationPos[i][0] and
fragPos[j][1]<=locationPos[i][1]:
                t[i]=
list(np.sum([t[i],list(seqData[sampleList].iloc[j])],0))
            elif (fragPos[j][1]>=locationPos[i][0] and
fragPos[j][1]<=locationPos[i][1]) or(fragPos[j][0]>=locationPos[i][0] and
fragPos[j][0]<=locationPos[i][1]):
                t[i]=
list(np.sum([t[i],list(seqData[sampleList].iloc[j])],0))
    return t

def findArrayPos(df,pos):
    shape = np.shape(df)
    start=[int(np.floor((pos[0]-1)/shape[1])),int((pos[0]-1)%shape[1])]
    end = [int(np.floor((pos[1]-1)/shape[1])),int((pos[1]-1)%shape[1])]
    print(f'Start pos at array position [{start[0]},{start[1]}], end pos at
array position [{end[0]},{end[1]}]')
    return start, end

def sumNewArray(array,sumPos):
    newDF=np.zeros(len(sumPos))
    for k in range(len(sumPos)):
        newDF[k]=np.average(array[sumPos[k][0]:sumPos[k][1]+1],0)
    return newDF

def sortfunc(x): #For sorting posList
    y=np.float64(x.split("-")[0].split('[')[1])
    return y
```

```
def domainCleanUp(data):
    testlist=[]
    for x in range(len(data.split('/note='))-1):
        test1 = data.split('/note=')[x].split(';')[-2].split(' ')[-1]
        if '..' in test1:
            test = test1.split('..')
        else:
            continue
        if ':' in test[0]:
            test[0]=test[0].split(':')[1]

        position=[int(test[0]),int(test[1])]
        name= data.split('/note=')[x+1].split('";')[0].strip('"')
        testlist.append([position,name])
    return testlist

def convertUsefulFasta(protein_data):
    data_list=protein_data.split(' -n ')
    entry_name=data_list[0]
    protein_name=data_list[1]
    protein_seq=data_list[2]
    return entry_name, protein_name, protein_seq
################################################

fileLocation = os.getcwd()
# To open dialog to ask for file location
root=tkinter.Tk()
root.withdraw()
root.wm_attributes('-topmost', 1)
fileName= askopenfilename(parent=root)

database ="PythonStuff\\Database\\uniprot_sprot_rat_251022.xlsx"
# sprot Human database:
'PythonStuff\\Database\\uniprot_sprot_Human_2022.08.10-
09.13.40.25_Edited2.xlsx'
# sprot Rat database:
"PythonStuff\\Database\\uniprot_sprot_rat_251022.xlsx"
# sprot Mouse database:
"PythonStuff\\Database\\uniprot_sprot_trembl_mouse_131022.xlsx"
# sprot Chicken endatabase:
"PythonStuff\\Database\\uniprot_sprot_gallusgallus_220623.xlsx"
xls = pd.ExcelFile(database)
df = xls.parse(xls.sheet_names[0])
columns_want = ['Entry Name', 'Protein names','Sequence', 'Domain [CC]',
'Compositional bias',
'Domain [FT]', 'Motif', 'Region']

df_new=pd.DataFrame()
df_new=df[columns_want].apply(lambda x:' -n '.join(x.astype(str)),axis=1)
df_new.index=df.Entry
fasta_sequences=df_new.to_dict()

#Old database (FASTA)
#database='PythonStuff\\Database\\uniprot_sprot_Human_NoIsoform.fasta'
#fasta_sequences = SeqIO.index(database,'fasta',key_function=lambda x:
x.split('|')[1])

#Read csv file and arrange headings
book = pd.read_excel(fileName,header=0).fillna(0)
abun_headings=book.columns[book.columns.str.contains('Abundance:')]
```

```
norm_abun_headings=book.columns[book.columns.str.contains('(Normalized)')]
sampleList= [x.split(':')[1]+x.split(':')[2] for x in norm_abun_headings]
samplesize=len(sampleList)

#Remove oxidized proteins
#book=book[book['Modifications'].str.contains("1xOxidation"and
"\[C")==True]
#book=book[['Oxidation' in str(x) for x in book['Modifications']]]
#'X' in modifications is special: '×'
id_list=book['Master Protein Accessions'].unique()
book['Abundances']=book[abun_headings].apply(lambda x:
','.join(x.astype(str)),axis=1)
book['Abundances (Normalized)']=book[norm_abun_headings].apply(lambda x:
','.join(x.astype(str)),axis=1)
book['Pos']=book['Positions in Master Proteins'].str.split(' ').str[1]
book['Pos_Start']=book['Pos'].str.split('-').str[0].str.split('[').str[1]
book['Pos_End']=book['Pos'].str.split('-').str[1].str.split(']').str[0]
final_head=['Master Protein Accessions','Pos_Start','Pos_End','Abundances
(Normalized)','Abundances']
final_book=book[final_head]

#mode either bin or dom
mode='bin'
binSize=20
norm='y' #'y' for yes, 'n' for no
norm_mode='mean' #mean or median

if mode=='bin':
    outputName= f'Output_{mode}_{binSize}'
else:
    outputName= f'Output_{mode}'
if norm=='y':
    outputName= outputName+'_norm'

outputExcel=xw.Workbook(fileName.split('.xlsx')[0] + f'_{outputName}.xlsx',
{'constant_memory': True})
sheet=outputExcel.add_worksheet()
row=0
for k in tqdm(range(len(id_list))):
    col=3
    protein_id=id_list[k]
    try:
        entry_name, protein_name, protein_seq =
convertUsefulFasta(fasta_sequences[protein_id])
    except:
        print(f'{protein_id} not found in database!')
        continue
    xBook=final_book.loc[final_book['Master Protein
Accessions'].str.contains(protein_id,case=False).fillna(False)]

    abunDF=np.zeros([len(protein_seq),samplesize])

    for h in range(len(xBook)):
        #print(f'{len(x)} repeat for sequence {seqList[0]}
\n{x["Modifications"]}')
        #Get peptide location

        s=int(xBook['Pos_Start'].iloc[h])-1
        e=int(xBook['Pos_End'].iloc[h])-1 #Convert to zero indexing

        abunList=np.zeros(samplesize)
```

```
        #Abundances (Normalized)/# PSMs
        #q=xBook['Abundances (Normalized)'].iloc[h].split(',')
        q=xBook['Abundances (Normalized)'].iloc[h].split(',')
        for k in range(samplesize):
            if q[k]=='nan':
                q[k]='0'
        abunList= abunList+np.float64(q)
        abunDF[s:e+1]= abunDF[s:e+1]+abunList

        #Extra normalisation step using total abundance per protein (for
working with contaminated samples)
        with warnings.catch_warnings():
            warnings.simplefilter("ignore", category=RuntimeWarning)
            if norm_mode=='mean':
                normFactor=max(sum(abunDF))/sum(abunDF)
            if norm_mode=='median':

normFactor=max(np.median(abunDF,axis=0))/np.median(abunDF,axis=0)
        normFactor[normFactor==float('+inf')]=0 #Replace inf into 0
        normFactor=np.nan_to_num(normFactor,nan=1) #Replace nan into 1
    if mode=='dom':

        ###############  Get domain positions ############################
        dList=[]
        for k in range(3,8):
            dList2=domainCleanUp(fasta_sequences[protein_id].split(' -n
')[k])
            dList.append(dList2)

        dList = list(itertools.chain(*[x for x in dList if x!=[]]))

        def sort(x):
            return x[0][0]

        dList=sorted(dList,key=sort)
        domPos=[x[0] for x in dList] #1 indexing
        domPos=[[s-1,e-1] for s,e in domPos] #0 indexing
        domList=[x[1] for x in dList]
    ####################################################################
        pos = domPos
    elif mode=='bin':
        binPos=[]
        numBin=len(protein_seq)//binSize
        if numBin==0:
            pos=[0,len(protein_seq)-1]
        else:
            binPos.append([[x,x+binSize-1] for x in
range(0,numBin*binSize,binSize)])
            binPos=binPos[0]
            remainder=len(protein_seq)%binSize
            if remainder!=0:
                binPos[-1][1]=binPos[-1][1]+remainder
        pos=binPos

    #Write data into excel

    for g in range(samplesize):
        sheet.write(row,col-3,f'{protein_id}')
        sheet.write(row,col-2,f'{sampleList[g]}')
        if norm=='y':
```

```
            abun_row_data = sumNewArray(abunDF[:,g], pos)*normFactor[g] #If
additional normalisation needed
        if norm=='n':
            abun_row_data = sumNewArray(abunDF[:,g], pos)
        for m in range(len(pos)):
            try:
                sheet.write(row,col-1,abun_row_data[m])
            except:
                pass
            col+=1

        col=3 #reset column
        row+=1

outputExcel.close()


#%% Stats
print('Now starting stats...')

import pingouin as pg
import pandas as pd
import scipy
from statsmodels.stats.anova import AnovaRM

paired='n' #'y' for paired, 'n' for not
replicates_num=5
control_keyword='Control'
treatment_keyword='100Gy'

#fileLocation = os.getcwd()
# To open dialog to ask for file location
#root=tkinter.Tk()
#root.withdraw()
#root.wm_attributes('-topmost', 1)
#fileName= askopenfilename(parent=root)

df=pd.read_excel(fileName.split('.xlsx')[0] +
f'_{outputName}.xlsx',header=None)
ecm_data=pd.read_excel(f'{fileLocation}//PythonStuff//Database//matrisome_h
uman.xls',header=0)
#matrisome_human.xls (Homosapiens) or matrisome_mouse.xls (Mus musculus)

plfAnalysisExcel=xw.Workbook(fileName.split('.')[0] +
f'_PLFAnalysis_{control_keyword}VS{treatment_keyword}.xlsx')
sheet1=plfAnalysisExcel.add_worksheet()
headings=['Protein ID','Protein Name']
#for col,headings in enumerate(headings):
sheet1.write(0,0,'Protein ID')
sheet1.write(0,2,'Significant p value?')
sheet1.write(0,3,'Score')
sheet1.write(0,4,'Matrix protein?')
sheet1.write(0,5,'Matrix protein type')
proteinIDList=id_list
row=1

for k in tqdm(range(len(proteinIDList))):
    try:
        e,name,s=convertUsefulFasta(fasta_sequences[proteinIDList[k]])
    except:
        print(f'{proteinIDList[k]} not found in database!')
```

```
            continue
        col=7
        sheet1.write(row,0,proteinIDList[k])
        sheet1.write(row,1,name)
        #Use protein_id to find protein name because name is not fully unique


#sheet1.merge_range(row,1,row+5,1,df[df[0]==proteinIDList[k]][1].unique()[0
])

        #datalist=['pValue','cMean','cStd','tMean','tStd','meanDiff']
        datalist=['pValue','meanDiff','%meanDiff']
        for rowlist,head in enumerate(datalist):
            sheet1.write(row+rowlist,col-1,head)

        df_specific=df[df[0]==proteinIDList[k]].dropna(axis=1,how='all')

        test=df_specific.melt(id_vars=[0,1])
        bins=len(test['variable'].unique())
        test_num=bins-np.sum(np.sum(df_specific)==0) #remove zeros bins from
multiple comparison corrections
        hs_correction=0.05/(1-(0.95**(1/test_num))) #Holmes sidak correction
factor for p_value
        #Check if ECM protein
        try:
            ecm_index =
ecm_data[ecm_data.UniProt_IDs.str.contains(proteinIDList[k])].index[0]
            sheet1.write(row,4,'Y')
            sheet1.write(row,5,f'{ecm_data.iloc[ecm_index].Category}')
        except:
            sheet1.write(row,4,'N')
        #Do anova for each bin

        significant='N'
        score=0

        if paired=='y':
        #setup dataframe for repeated measure anova
            dfRM=df_specific[

df_specific[1].str.contains('|'.join([control_keyword,treatment_keyword]))]
            labels_df=pd.DataFrame({'SampleNumber':
np.tile(range(replicates_num), 2),
                                    'Dose': [x.split(',')[1] for x in
dfRM[1]]})
            dfRM.columns=dfRM.columns.astype(str)+'bin'

finalrm_df=pd.concat([labels_df,dfRM.iloc[:,2:].reset_index().drop(columns=
'index')],axis=1)

        for x in test['variable'].unique():

            if np.mean(test[test['variable']==x]['value'])==0: #If all values
zero then skip
                sheet1.write(row,col,'1')
                sheet1.write(row+1,col,'0')
                sheet1.write(row+2,col,'0')
                col+=1
                continue

            if paired=='n':
```

```
            #simple t-test
            val=test[test['variable']==x]
            cVal=val[val[1].str.contains(control_keyword)]['value']
            tVal=val[val[1].str.contains(treatment_keyword)]['value']
            with warnings.catch_warnings():
                warnings.simplefilter("ignore", category=RuntimeWarning)
                f,p=scipy.stats.ttest_ind(cVal,tVal)

            #p_corr=p
            p_corr=p*2 #Bonferroni correction -- two comparisons
controlvs50 and vs 100
            #p_corr=p*hs_correction #holm Sidak
            if p_corr>1:
                p_corr=1

        elif paired=='y':

            #repeated measure anova
            with warnings.catch_warnings():
                warnings.simplefilter("ignore", category=RuntimeWarning)

r=AnovaRM(finalrm_df,subject='SampleNumber',depvar=f'{x}bin',within=['Dose'
]).fit()
            p=r.anova_table['Pr > F'][0]
            p_corr=p
            #p_corr=p*len(test['variable'].unique())  #Bonferroni
correction
            if p_corr>1:
                p_corr=1

cVal=finalrm_df[finalrm_df['Dose'].str.contains(control_keyword)][f'{x}bin'
]

tVal=finalrm_df[finalrm_df['Dose'].str.contains(treatment_keyword)][f'{x}bi
n']

        cMean=np.mean(cVal)
        tMean=np.mean(tVal)
        meanDiff=cMean-tMean

        if meanDiff==0 or cMean==0:
            perc_meanDiff=0
        else:
            with warnings.catch_warnings():
                warnings.simplefilter("ignore", category=RuntimeWarning)
                perc_meanDiff=meanDiff/cMean*100

        if np.isnan(p_corr):
            p_corr=1

        #cStd=np.std(cVal)
        #tStd=np.std(tVal)

        sheet1.write(row,col,f'{p_corr}')
        #sheet1.write(row+1,col,f'{cMean}')
        #sheet1.write(row+2,col,f'{cStd}')
        #sheet1.write(row+3,col,f'{tMean}')
        #sheet1.write(row+4,col,f'{tStd}')
        sheet1.write(row+1,col,f'{meanDiff}')
        sheet1.write(row+2,col,f'{perc_meanDiff}') # % mean difference
        col+=1
```

```
        if p_corr<=0.05:
            significant='Y'
            score+=1
    sheet1.write(row,2,f'{significant}')
    if len(test['variable'].unique())==0:
        fin_score=0
    else:
        fin_score=score/len(test["variable"].unique())
    sheet1.write_number(row,3,fin_score) #score is weighted by number of
bins
    #row+=6
    row+=3

plfAnalysisExcel.close()
```

8 References

1. Tuieng RJ, Cartmell SH, Kirwan CC, Sherratt MJ. The effects of ionising and non-ionising electromagnetic radiation on extracellular matrix proteins. Cells. 2021;10(11):1–25.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021;71(3):209–49.

3. Yang TJ, Ho AY. Radiation Therapy in the Management of Breast Cancer. Surg Clin North Am. 2013 Apr;93(2):455–71.

4. Pignol J-P, Keller BM, Ravi A. Doses to internal organs for various breast radiation techniques - implications on the risk of secondary cancers and cardiomyopathy. Radiat Oncol. 2011 Dec 14;6(1):5.

5. Taylor C, Duane FK, Dodwell D, Gray R, Wang Z, Wang Y, et al. Estimating the Risks of Breast cancer radiotherapy: Evidence from modern radiation doses to the lungs and Heart and From previous randomized trials. J Clin Oncol. 2017;35(15):1641–9.

6. Henson KE, McGale P, Taylor C, Darby SC. Radiation-related mortality from heart disease and lung cancer more than 20 years after radiotherapy for breast cancer. Br J Cancer. 2013;108(1):179–82.

7. Biegała M, Hydzik A. Analysis of dose distribution in organs at risk in patients with prostate cancer treated with the intensity-modulated radiation therapy and arc technique. J Med Phys. 2016;41(3):198–204.

8. Lyons A, Ghazali N. Osteoradionecrosis of the jaws: current understanding of its pathophysiology and treatment. Br J Oral Maxillofac Surg. 2008;46(8):653–60.

9. Purkayastha A, Sharma N, Sarin A, Bhatnagar S, Chakravarty N, Mukundan H, et al. Radiation Fibrosis Syndrome: The Evergreen Menace of Radiation Therapy. Asia-Pacific J Oncol Nurs. 2019 Jul;6(3):238–45.

10. Weihermann AC, Lorencini M, Brohem CA, de Carvalho CM. Elastin structure and its involvement in skin photoageing. Int J Cosmet Sci. 2017;39(3):241–7.

11. Wlaschek M, Tantcheva-Poór I, Naderi L, Ma W, Schneider LA, Razi-Wolf Z, et al. Solar UV irradiation and dermal photoaging. J Photochem Photobiol B Biol. 2001 Oct;63(1–3):41–51.

12. Lomax ME, Folkes LK, O'Neill P. Biological consequences of radiation-induced DNA damage: Relevance to radiotherapy. Clin Oncol. 2013;25(10):578–85.

13. Yokoya A, Shikazono N, Fujii K, Urushibara A, Akamatsu K, Watanabe R. DNA damage induced by the direct effect of radiation. Radiat Phys Chem. 2008 Oct;77(10–12):1280–5.

14. Walker C, Mojares E, del Río Hernández A. Role of Extracellular Matrix in Development and Cancer Progression. Int J Mol Sci. 2018 Oct 4;19(10):3028.

15. Hu G, Li L, Xu W. Extracellular matrix in mammary gland development and breast cancer progression. Front Lab Med. 2017;1(1):36–9.

16. Schultz GS, Wysocki A. Interactions between extracellular matrix and growth factors in wound healing. Wound Repair Regen. 2009;17(2):153–62.

17. Kechagia JZ, Ivaska J, Roca-Cusachs P. Integrins as biomechanical sensors of the microenvironment. Nat Rev Mol Cell Biol. 2019;20(8):457–73.

18. Hubbard B, Buczek-Thomas JA, Nugent MA, Smith ML. Fibronectin Fiber Extension Decreases Cell Spreading and Migration. J Cell Physiol. 2016;231(8):1728–36.

19. Cheng A, Cain SA, Tian P, Baldwin AK, Uppanan P, Kielty CM, et al. Recombinant Extracellular Matrix Protein Fragments Support Human Embryonic Stem Cell Chondrogenesis. Tissue Eng - Part A.

2018;24(11–12):968–78.

20. McConnell JC, O'Connell O V., Brennan K, Weiping L, Howe M, Joseph L, et al. Increased peri-ductal collagen micro-organization may contribute to raised mammographic density. Breast Cancer Res. 2016;18(1):1–17.

21. Sherratt MJ, McConnell JC, Streuli CH. Raised mammographic density: Causative mechanisms and biological consequences. Breast Cancer Res. 2016;18(1):1–9.

22. Grant TM, Thompson MS, Urban J, Yu J. Elastic fibres are broadly distributed in tendon and highly localized around tenocytes. J Anat. 2013;222(6):573–9.

23. Toprak N, Aras I, Toktaş O, Yokuş A, Gündüz AM. Assessment of Stromal Elastin Fibers in Breast Cancer and Fibroadenomas: Is There a Correlation With Ultrasound Elastography Findings? Eur J Breast Heal. 2022;18(2):134–40.

24. Gao JJ, Swain SM. Luminal A Breast Cancer and Molecular Assays: A Review. Oncologist. 2018;23(5):556–65.

25. Ades F, Zardavas D, Bozovic-Spasojevic I, Pugliano L, Fumagalli D, De Azambuja E, et al. Luminal B breast cancer: Molecular characterization, clinical management, and future perspectives. J Clin Oncol. 2014;32(25):2794–803.

26. Loibl S, Gianni L. HER2-positive breast cancer. Lancet. 2017;389(10087):2415–29.

27. Bianchini G, Balko JM, Mayer IA, Sanders ME, Gianni L. Triple-negative breast cancer: Challenges and opportunities of a heterogeneous disease. Nat Rev Clin Oncol. 2016;13(11):674–90.

28. Kunkler IH, Williams LJ, Jack WJL, Cameron DA, Dixon JM. Breast-Conserving Surgery with or without Irradiation in Early Breast Cancer. N Engl J Med. 2023;388(7):585–94.

29. Kummerow KL, Du L, Penson DF, Shyr Y, Hooks MA. Nationwide trends in mastectomy for early-stage breast cancer. JAMA Surg. 2015;150(1):9–16.

30. Charfare H, Limongelli S, Purushotham AD. Neoadjuvant chemotherapy in breast cancer. Br J Surg. 2005;92(1):14–23.

31. von Minckwitz G, Martin M. Neoadjuvant treatments for triple-negative breast cancer (TNBC). Ann Oncol. 2012;23(SUPPL. 6):vi35–9.

32. Spring LM, Gupta A, Reynolds KL, Gadd MA, Ellisen LW, Isakoff SJ, et al. Neoadjuvant endocrine therapy for estrogen receptor-positive breast cancer a systematic review and meta-Analysis. JAMA Oncol. 2016;2(11):1477–86.

33. Zaborowski AM, Wong SM. Neoadjuvant systemic therapy for breast cancer. Br J Surg. 2023 Jun 12;110(7):765–72.

34. Makubate B, Donnan PT, Dewar JA, Thompson AM, McCowan C. Cohort study of adherence to adjuvant endocrine therapy, breast cancer recurrence and mortality. Br J Cancer. 2013;108(7):1515–24.

35. Darby S, McGale P, Correa C, Taylor C, Arriagada R, Clarke M, et al. Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: meta-analysis of individual patient data for 10 801 women in 17 randomised trials. Lancet. 2011 Nov;378(9804):1707–16.

36. Clarke M, Collins R, Darby S, Davies C, Elphinstone P, Evans V, et al. Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. Lancet. 2005 Dec;366(9503):2087–106.

37. Sánchez-Barceló EJ, Cos S, Mediavilla D, Martínez-Campa C, González A, Alonso-González C. Melatonin-estrogen interactions in breast cancer. J Pineal Res. 2005;38(4):217–22.

38. Nourmoussavi M, Pansegrau G, Popesku J, Hammond GL, Kwon JS, Carey MS. Ovarian ablation for

premenopausal breast cancer: A review of treatment considerations and the impact of premature menopause. Cancer Treat Rev. 2017;55:26–35.

39. Riggs BL, Hartmann LC. Selective Estrogen-Receptor Modulators — Mechanisms of Action and Application to Clinical Practice. Wood AJJ, editor. N Engl J Med. 2003 Feb 13;348(7):618–29.

40. Spector N, Xia W, El-Hariry I, Yarden Y, Bacus S. HER2 therapy. Small molecule HER-2 tyrosine kinase inhibitors. Breast Cancer Res. 2007;9(2):1–8.

41. Nahta R, Esteva FJ. Herceptin: Mechanisms of action and resistance. Cancer Lett. 2006;232(2):123–38.

42. Yin L, Duan JJ, Bian XW, Yu SC. Triple-negative breast cancer molecular subtyping and treatment progress. Breast Cancer Res. 2020;22(1):1–13.

43. National Institute for Health and Care Excellence. Early and locally advanced breast cancer: diagnosis and management. NG101 [Internet]. National Institute for Health and Care Excellence. 2023 [cited 2023 Sep 5]. Available from: www.nice.org.uk/guidance/ng101

44. Haussmann J, Corradini S, Nestle-Kraemling C, Bölke E, Njanang FJD, Tamaskovics B, et al. Recent advances in radiotherapy of breast cancer. Radiat Oncol. 2020 Dec 30;15(1):71.

45. Rutqvist LE, Rose C, Cavallin-Ståhl E. A systematic overview of radiation therapy effects in breast cancer. Acta Oncol (Madr). 2003;42(5–6):532–45.

46. Mettler FA. Medical effects and risks of exposure to ionising radiation. J Radiol Prot. 2012 Mar;32(1):N9–13.

47. LaRiviere MJ, Chao HH, Doucette A, Kegelman TP, Taunk NK, Freedman GM, et al. Factors Associated With Fatigue in Patients with Breast Cancer Undergoing External Beam Radiation Therapy. Pract Radiat Oncol. 2020;10(6):409–22.

48. Bentzen SM. Preventing or reducing late side effects of radiation therapy: Radiobiology meets molecular pathology. Nat Rev Cancer. 2006;6(9):702–13.

49. Solov'yov A V. Nanoscale insights into ion-beam cancer therapy. Nanoscale Insights into Ion-Beam Cancer Therapy. 2016. 1–498 p.

50. Reisz JA, Bansal N, Qian J, Zhao W, Furdui CM. Effects of ionizing radiation on biological molecules - mechanisms of damage and emerging methods of detection. Antioxidants Redox Signal. 2014;21(2):260–92.

51. ARMENTROUT PB. Chemistry of Excited Electronic States. Science (80-). 1991 Jan 11;251(4990):175–9.

52. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Sources and effects of ionizing radiation UNSCEAR 2008 report to the General Assembly with Scientific Annexes Volume 1. Vol. I. 2008. 156 p.

53. ICRP. The 2007 Recommendations of the International Commission on Radiological Protection. ICRP Publication 103. Vol. 37, Annals of the ICRP. 2007 Apr.

54. Young KC, Oduko JM. Radiation doses received in the United Kingdom breast screening programme in 2010 to 2012. Br J Radiol. 2016;89(1058).

55. Uselmann AJ, Thomadsen BR. On effective dose for radiotherapy based on doses to nontarget organs and tissues. Med Phys. 2015;42(2):977–82.

56. Nutting C, Dearnaley DP, Webb S. Intensity modulated radiation therapy: A clinical review. Br J Radiol. 2000;73(869):459–69.

57. Olaciregui-Ruiz I, Beddar S, Greer P, Jornet N, McCurdy B, Paiva-Fonseca G, et al. In vivo dosimetry in external beam photon radiotherapy: Requirements and future directions for research, development, and clinical practice. Phys Imaging Radiat Oncol. 2020;15(August):108–16.

58. Murray Brunt A, Haviland JS, Wheatley DA, Sydenham MA, Alhasso A, Bloomfield DJ, et al.

Hypofractionated breast radiotherapy for 1 week versus 3 weeks (FAST-Forward): 5-year efficacy and late normal tissue effects results from a multicentre, non-inferiority, randomised, phase 3 trial. Lancet. 2020;395(10237):1613–26.

59. Podgorsak EB. Radiation Oncology Physics: A Handbook for Teachers and Students. Chapter 6: External photon beams: Physical aspects. IAEA. International atomic energy agency; 2005. 161–217 p.

60. Koulis TA, Phan T, Olivotto IA. Hypofractionated whole breast radiotherapy: Current perspectives. Breast Cancer Targets Ther. 2015;7:363–70.

61. Yamada Y, Rogers L, Demanes DJ, Morton G, Prestidge BR, Pouliot J, et al. American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. Brachytherapy. 2012;11(1):20–32.

62. Yahyapour R, Salajegheh A, Safari A, Amini P, Rezaeyan A, Amraee A, et al. Radiation-induced Non-targeted Effect and Carcinogenesis; Implications in Clinical Radiotherapy. J Biomed Phys Eng. 2018;8(4):435–46.

63. Tommasino F, Durante M. Proton radiobiology. Cancers (Basel). 2015;7(1):353–81.

64. Mohan R, Grosshans D. Proton therapy – Present and future. Adv Drug Deliv Rev. 2017;109:26–44.

65. Lalande M, Schwob L, Vizcaino V, Chirot F, Dugourd P, Schlathölter T, et al. Direct Radiation Effects on the Structure and Stability of Collagen and Other Proteins. ChemBioChem. 2019;20(24):2972–80.

66. Feldberg RS, Carew JA. Water radiolysis products and nucleotide damage in γ-irradiated DNA. Int J Radiat Biol. 1981;40(1):11–7.

67. Hatano Y, Katsumura Y, Mozumder A, Meesungnoen J, Jay-Gerin J-P. Charged Particle and Photon Interactions with Matter. Hatano Y, Katsumura Y, Mozumder A, editors. Charged Particle and Photon Interactions with Matter Recent Advances, Applications, and Interfaces. CRC Press; 2010. 355–392 p.

68. Misawa M, Takahashi J. Generation of reactive oxygen species induced by gold nanoparticles under x-ray and UV Irradiations. Nanomedicine Nanotechnology, Biol Med. 2011;7(5):604–14.

69. Averbeck D, Salomaa S, Bouffler S, Ottolenghi A, Smyth V, Sabatier L. Progress in low dose health risk research: Novel effects and new concepts in low dose radiobiology. Mutat Res - Rev Mutat Res. 2018;776(January):46–69.

70. Baverstock K, Karotki A V. Towards a unifying theory of late stochastic effects of ionizing radiation. Mutat Res - Genet Toxicol Environ Mutagen. 2011;718(1–2):1–9.

71. Møller AP, Mousseau TA. Biological consequences of Chernobyl: 20 years on. Trends Ecol Evol. 2006;21(4):200–7.

72. Ozasa K, Grant EJ, Kodama K. Japanese legacy cohorts: The life span study atomic bomb survivor cohort and survivors' offspring. J Epidemiol. 2018;28(4):162–9.

73. Quast U. Whole body radiotherapy: A TBI-guideline. J Med Phys. 2006;31(1):5.

74. Hamasaki K, Imai K, Hayashi T, Nakachi K, Kusunoki Y. Radiation sensitivity and genomic instability in the hematopoietic system: Frequencies of micronucleated reticulocytes in whole-body X-irradiated BALB/c and C57BL/6 mice. Cancer Sci. 2007;98(12):1840–4.

75. Thomas O, Mahé MA, Campion L, Bourdin S, Milpied N, Brunet G, et al. Long-term complications of total body irradiation in adults. Int J Radiat Oncol Biol Phys. 2001 Jan;49(1):125–31.

76. Wills C, Cherian S, Yousef J, Wang K, Mackley HB. Total body irradiation: A practical review. Appl Radiat Oncol. 2016;(August):11–7.

77. Yu J. Intestinal stem cell injury and protection during cancer therapy. Transl Cancer Res. 2013;2(5):384–96.

78. Kole AJ, Kole L, Moran MS. Acute radiation dermatitis in breast cancer patients: Challenges and

solutions. Breast Cancer Targets Ther. 2017;9:313–23.

79. Ejaz A, Greenberger JS, Rubin PJ. Understanding the mechanism of radiation induced fibrosis and therapy options. Pharmacol Ther. 2019;204:107399.

80. Werner EM, Eggert MC, Bohnet S, Rades D. Prevalence and characteristics of pneumonitis following irradiation of breast cancer. Anticancer Res. 2019;39(11):6355–8.

81. Tsoutsou PG, Koukourakis MI. Radiation pneumonitis and fibrosis: Mechanisms underlying its pathogenesis and implications for future research. Int J Radiat Oncol Biol Phys. 2006;66(5):1281–93.

82. Wang K, Tepper JE. Radiation therapy-associated toxicity: Etiology, management, and prevention. CA Cancer J Clin. 2021;71(5):437–54.

83. Shtarkman IN, Gudkov S V., Chernikov A V., Bruskov VI. Effect of amino acids on X-ray-induced hydrogen peroxide and hydroxyl radical formation in water and 8-oxoguanine in DNA. Biochem. 2008;73(4):470–8.

84. Cadet J, Richard Wagner J. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. Cold Spring Harb Perspect Biol. 2013;5(2):1–18.

85. Fulford J, Nikjoo H, Goodhead DT, O'Neill P. Yields of SSB and DSB induced in DNA by AlK ultrasoft X-rays and α-particles: Comparison of experimental and simulated yields. Int J Radiat Biol. 2001;77(10):1053–66.

86. Ward JF. Some biochemical consequences of the spatial distribution of ionizing radiation-produced free radicals. Radiat Res. 1981;86(2):185–95.

87. Khanna KK, Jackson SP. DNA double-strand breaks: Signaling, repair and the cancer connection. Nat Genet. 2001;27(3):247–54.

88. Burma S, Chen BPC, Chen DJ. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. DNA Repair (Amst). 2006;5(9–10):1042–8.

89. Rothkamm K, Kühne M, Jeggo PA, Löbrich M. Radiation-induced genomic rearrangements formed by nonhomologous end-joining of DNA double-strand breaks. Cancer Res. 2001;61(10):3886–93.

90. Morgan WF. Non-targeted and delayed effects of exposure to ionizing radiation: I. Radiation-induced genomic instability and bystander effects in Vitro. Radiat Res. 2012;178(2):581–96.

91. De Santis M, Di Gianantonio E, Straface G, Cavaliere AF, Caruso A, Schiavon F, et al. Ionizing radiations in pregnancy and teratogenesis: A review of literature. Reprod Toxicol. 2005;20(3):323–9.

92. Pietraforte D, Paulicelli E, Patrono C, Gambardella L, Scorza G, Testa A, et al. Protein oxidative damage and redox imbalance induced by ionising radiation in CHO cells. Free Radic Res. 2018;52(4):465–79.

93. Prasad NR, Menon VP, Vasudev V, Pugalendi K V. Radioprotective effect of sesamol on γ-radiation induced DNA damage, lipid peroxidation and antioxidants levels in cultured human lymphocytes. Toxicology. 2005;209(3):225–35.

94. Guven M, Brem R, Macpherson P, Peacock M, Karran P. Oxidative Damage to RPA Limits the Nucleotide Excision Repair Capacity of Human Cells. J Invest Dermatol. 2015;135(11):2834–41.

95. Liu F, Lai S, Tong H, Lakey PSJ, Shiraiwa M, Weller MG, et al. Release of free amino acids upon oxidation of peptides and proteins by hydroxyl radicals. Anal Bioanal Chem. 2017;409(9):2411–20.

96. Schwob L, Lalande M, Rangama J, Egorov D, Hoekstra R, Pandey R, et al. Single-photon absorption of isolated collagen mimetic peptides and triple-helix models in the VUV-X energy range. Phys Chem Chem Phys. 2017;19(28):18321–9.

97. Henderson R. Cryo-protection of protein crystals against radiation damage in electron and X-ray diffraction. Proc R Soc London Ser B Biol Sci. 1990 Jul 23;241(1300):6–8.

98. Ravelli RBG, McSweeney SM. The "fingerprint" that X-rays can leave on structures. Structure.

2000;8(3):315–28.

99. Weik M, Ravelli RBG, Kryger G, McSweeney S, Raves ML, Harel M, et al. Specific chemical and structural damage to proteins produced by synchrotron radiation. Proc Natl Acad Sci U S A. 2000;97(2):623–8.

100. Burmeister WP. Structural changes in a cryo-cooled protein crystal owing to radiation damage. Acta Crystallogr Sect D Biol Crystallogr. 2000 Mar 1;56(3):328–41.

101. Dubnovitsky AP, Ravelli RBG, Popov AN, Papageorgiou AC. Strain relief at the active site of phosphoserine aminotransferase induced by radiation damage. Protein Sci. 2009 Jan 1;14(6):1498–507.

102. Garman EF, Weik M. Radiation damage in macromolecular crystallography. Methods Mol Biol. 2017;1607:467–89.

103. Matsui Y, Sakai K, Murakami M, Shiro Y, Adachi SI, Okumura H, et al. Specific damage induced by X-ray radiation and structural changes in the primary photoreaction of bacteriorhodopsin. J Mol Biol. 2002;324(3):469–81.

104. Mees A, Klar T, Gnau P, Hennecke U, Eker APM, Carell T, et al. Crystal structure of a photolyase bound to a CPD-like DNA lesion after in situ repair. Science (80-). 2004;306(5702):1789–93.

105. Fioravanti E, Vellieux FMD, Amara P, Madern D, Weik M. Specific radiation damage to acidic residues and its relation to their chemical and structural environment. J Synchrotron Radiat. 2007;14(1):84–91.

106. Sjöblom B, Polentarutti M, Djinović-Carugo K. Structural study of X-ray induced activation of carbonic anhydrase. Proc Natl Acad Sci U S A. 2009;106(26):10609–13.

107. Nass K, Foucar L, Barends TRM, Hartmann E, Botha S, Shoeman RL, et al. Indications of radiation damage in ferredoxin microcrystals using high-intensity X-FEL beams. J Synchrotron Radiat. 2015 Mar 1;22(2):225–38.

108. Schärer OD. Nucleotide excision repair in Eukaryotes. Cold Spring Harb Perspect Biol. 2013;5(10):1–19.

109. Holmgren A, Johansson C, Berndt C, Lönn ME, Hudemann C, Lillig CH. Thiol redox control via thioredoxin and glutaredoxin systems. Biochem Soc Trans. 2005;33(6):1375–7.

110. Davies KJA. Degradation of oxidized proteins by the 20S proteasome. Biochimie. 2001;83(3–4):301–10.

111. Fedorova M, Bollineni RC, Hoffmann R. Protein carbonylation as a major hallmark of oxidative damage: Update of analytical strategies. Mass Spectrom Rev. 2014;33(2):79–97.

112. Folz RJ, Crapo JD. Extracellular superoxide dismutase (sod3): Tissue-specific expression, genomic characterization, and computer-assisted sequence analysis of the human ec sod gene. Vol. 22, Genomics. 1994. p. 162–71.

113. Yao H, Arunachalam G, Hwang JW, Chung S, Sundar IK, Kinnula VL, et al. Extracellular superoxide dismutase protects against pulmonary emphysema by attenuating oxidative fragmentation of ECM. Proc Natl Acad Sci U S A. 2010;107(35):15571–6.

114. Schamberger AC, Schiller HB, Fernandez IE, Sterclova M, Heinzelmann K, Hennen E, et al. Glutathione peroxidase 3 localizes to the epithelial lining fluid and the extracellular matrix in interstitial lung disease. Sci Rep. 2016;6(June):1–15.

115. Thomas C, Mackey MM, Diaz AA, Cox DP. Hydroxyl radical is produced via the Fenton reaction in submitochondrial particles under oxidative stress: implications for diseases associated with iron accumulation. Redox Rep. 2009 Jun 19;14(3):102–8.

116. Caldecott KW. Single-strand break repair and genetic disease. Nat Rev Genet. 2008 Aug;9(8):619–31.

117. David SS, O'Shea VL, Kundu S. Base-excision repair of oxidative DNA damage. Nature. 2007;447(7147):941–50.

118. Lieber MR. The mechanism of DSB repair by the NHEJ. Annu Rev Biochem. 2011;79(3):181–211.

119. Natarajan AT, Palitti F. DNA repair and chromosomal alterations. Mutat Res - Genet Toxicol Environ Mutagen. 2008;657(1):3–7.

120. Takata M, Sasaki MS, Sonoda E, Morrison C, Hashimoto M, Utsumi H, et al. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. EMBO J. 1998;17(18):5497–508.

121. Liakou E, Mavrogonatou E, Pratsinis H, Rizou S, Evangelou K, Panagiotou PN, et al. Ionizing radiation-mediated premature senescence and paracrine interactions with cancer cells enhance the expression of syndecan 1 in human breast stromal fibroblasts: The role of TGF-β. Aging (Albany NY). 2016;8(8):1650–69.

122. Nguyen HQ, To NH, Zadigue P, Kerbrat S, De La Taille A, Le Gouvello S, et al. Ionizing radiation-induced cellular senescence promotes tissue fibrosis after radiotherapy. A review. Crit Rev Oncol Hematol. 2018;129(February):13–26.

123. Tsai KKC, Chuang EYY, Little JB, Yuan ZM. Cellular mechanisms for low-dose ionizing radiation-induced perturbation of the breast tissue microenvironment. Cancer Res. 2005;65(15):6734–44.

124. Giannelli G, Falk-Marzillier J, Schiraldi O, Stetler-Stevenson WG, Quaranta V. Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5. Science (80-). 1997;277(5323):225–8.

125. Koshikawa N, Giannelli G, Cirulli V, Miyazaki K, Quaranta V. Role of cell surface metalloprotease MT1-MMP in epithelial cell migration over laminin-5. J Cell Biol. 2000;148(3):615–24.

126. Andarawewa KL, Erickson AC, Chou WS, Costes S V., Gascard P, Mott JD, et al. Ionizing radiation predisposes nonmalignant human mammary epithelial cells to undergo transforming growth factor β-induced epithelial to mesenchymal transition. Cancer Res. 2007;67(18):8662–70.

127. Lima JF, Nofech-Mozes S, Bayani J, Bartlett JMS. EMT in breast carcinoma—A review. J Clin Med. 2016;5(7):1–14.

128. Ehrhart EJ, Segarini P, Tsang ML -S., Carroll AG, Barcellos-Hoff MH. Latent transforming growth factor β1 activation in situ: quantitative and functional evidence after low-dose γ-irradiation 1 . FASEB J. 1997;11(12):991–1002.

129. Herskind C, Rodemann HP. Spontaneous and radiation-induced differentiationof fibroblasts. Exp Gerontol. 2000;35(6–7):747–55.

130. Paquette B, Baptiste C, Therriault H, Arguin G, Plouffe B, Lemay R. In vitro irradiation of basement membrane enhances the invasiveness of breast cancer cells. Br J Cancer. 2007;97(11):1505–12.

131. Jones JI, Gockerman A, Busby WH, Camacho-Hubner C, Clemmons DR. Extracellular matrix contains insulin-like growth factor binding protein-5: Potentiation of the effects of IGF-I. J Cell Biol. 1993;121(3):679–87.

132. Hinz B. The extracellular matrix and transforming growth factor-β1: Tale of a strained relationship. Matrix Biol. 2015;47:54–65.

133. Juhl P, Bondesen S, Hawkins CL, Karsdal MA, Bay-Jensen AC, Davies MJ, et al. Dermal fibroblasts have different extracellular matrix profiles induced by TGF-β, PDGF and IL-6 in a model for skin fibrosis. Sci Rep. 2020;10(1):1–10.

134. Bonnans C, Chou J, Werb Z. Remodelling the extracellular matrix in development and disease. Nat Rev Mol Cell Biol. 2014;15(12):786–801.

135. Cole MA, Quan T, Voorhees JJ, Fisher GJ. Extracellular matrix regulation of fibroblast function: redefining our perspective on skin aging. J Cell Commun Signal. 2018 Mar 17;12(1):35–43.

136. Lu P, Takai K, Weaver VM, Werb Z. Extracellular Matrix Degradation and Remodeling in Development and Disease. Cold Spring Harb Perspect Biol. 2011 Dec 1;3(12):a005058–a005058.

137. Shapiro SD, Endicott SK, Province MA, Pierce JA, Campbell EJ. Marked longevity of human lung

parenchymal elastic fibers deduced from prevalence of D-aspartate and nuclear weapons-related radiocarbon. J Clin Invest. 1991;87(5):1828–34.

138. Kielty CM, Sherratt MJ, Marson A, Baldock C. Fibrillin microfibrils. Adv Protein Chem. 2005;70(04):405–36.

139. Sherratt MJ. Tissue elasticity and the ageing elastic fibre. Age (Omaha). 2009;31(4):305–25.

140. Verzijl N, DeGroot J, Thorpe SR, Bank RA, Shaw JN, Lyons TJ, et al. Effect of collagen turnover on the accumulation of advanced glycation end products. J Biol Chem. 2000;275(50):39027–31.

141. Padhi A, Nain AS. ECM in Differentiation: A Review of Matrix Structure, Composition and Mechanical Properties. Ann Biomed Eng. 2020;48(3):1071–89.

142. Miller JP, Borde BH, Bordeleau F, Zanotelli MR, LaValley DJ, Parker DJ, et al. Clinical doses of radiation reduce collagen matrix stiffness. APL Bioeng. 2018 Sep;2(3):031901.

143. Sionkowska A. Thermal denaturation of UV-irradiated wet rat tail tendon collagen. Int J Biol Macromol. 2005;35(3–4):145–9.

144. Dong Y, Zheng Q, Wang Z, Lin X, You Y, Wu S, et al. Higher matrix stiffness as an independent initiator triggers epithelial-mesenchymal transition and facilitates HCC metastasis. J Hematol Oncol. 2019;12(1):1–15.

145. Wei SC, Fattet L, Tsai JH, Guo Y, Pai VH, Majeski HE, et al. Matrix stiffness drives epithelial–mesenchymal transition and tumour metastasis through a TWIST1–G3BP2 mechanotransduction pathway. Nat Cell Biol. 2015 May 20;17(5):678–88.

146. Fisher GJ, Shao Y, He T, Qin Z, Perry D, Voorhees JJ, et al. Reduction of fibroblast size/mechanical force down-regulates TGF-β type II receptor: implications for human skin aging. Aging Cell. 2016 Feb 8;15(1):67–76.

147. Junker JPE, Kratz C, Tollbäck A, Kratz G. Mechanical tension stimulates the transdifferentiation of fibroblasts into myofibroblasts in human burn scars. Burns. 2008;34(7):942–6.

148. Park JS, Chu JS, Tsou AD, Diop R, Tang Z, Wang A, et al. The effect of matrix stiffness on the differentiation of mesenchymal stem cells in response to TGF-β. Biomaterials. 2011;32(16):3921–30.

149. Ulrich TA, De Juan Pardo EM, Kumar S. The mechanical rigidity of the extracellular matrix regulates the structure, motility, and proliferation of glioma cells. Cancer Res. 2009;69(10):4167–74.

150. Hibbert SA, Watson REB, Gibbs NK, Costello P, Baldock C, Weiss AS, et al. A potential role for endogenous proteins as sacrificial sunscreens and antioxidants in human tissues. Redox Biol. 2015;5:101–13.

151. Hibbert SA, Watson REB, Griffiths CEM, Gibbs NK, Sherratt MJ. Selective proteolysis by matrix metalloproteinases of photo-oxidised dermal extracellular matrix proteins. Cell Signal. 2019;54(August 2018):191–9.

152. Sherratt MJ, Bayley CP, Reilly SM, Gibbs NK, Griffiths CEM, Watson REB. Low-dose ultraviolet radiation selectively degrades chromophore-rich extracellular matrix components. J Pathol. 2010;222(1):32–40.

153. Taipale J, Keski-Oja J. Growth factors in the extracellular matrix. FASEB J. 1997 Jan;11(1):51–9.

154. Annes JP, Chen Y, Munger JS, Rifkin DB. Integrin αvβ6-mediated activation of latent TGF-β requires the latent TGF-β binding protein-1. J Cell Biol. 2004;165(5):723–34.

155. Hawinkels LJAC, Zuidwijk K, Verspaget HW, de Jonge-Muller ESM, Duijn W van, Ferreira V, et al. VEGF release by MMP-9 mediated heparan sulphate cleavage induces colorectal cancer angiogenesis. Eur J Cancer. 2008;44(13):1904–13.

156. Turner N, Grose R. Fibroblast growth factor signalling: From development to cancer. Nat Rev Cancer. 2010;10(2):116–29.

157. Walton KL, Johnson KE, Harrison CA. Targeting TGF-β mediated SMAD signaling for the prevention of fibrosis. Front Pharmacol. 2017;8(JUL).

158. Kisling A, Lust RM, Katwa LC. What is the role of peptide fragments of collagen I and IV in health and disease? Life Sci. 2019;228(April):30–4.

159. Duca L, Floquet N, Alix AJP, Haye B, Debelle L. Elastin as a matrikine. Crit Rev Oncol Hematol. 2004;49(3):235–44.

160. Wells JM, Gaggar A, Blalock JE. MMP generated matrikines. Matrix Biol. 2015 May;44–46(3):122–9.

161. Chang CW, Dalgliesh AJ, López JE, Griffiths LG. Cardiac extracellular matrix proteomics: Challenges, techniques, and clinical implications. Proteomics - Clin Appl. 2016;10(1):39–50.

162. Naba A. Ten Years of Extracellular Matrix Proteomics: Accomplishments, Challenges, and Future Perspectives. Mol Cell Proteomics. 2023;22(4):100528.

163. Xu Y, Kirchner M. Collagen Mimetic Peptides. Bioengineering. 2021 Jan 5;8(1):5.

164. Zheng H, Lu C, Lan J, Fan S, Nanda V, Xu F. How electrostatic networks modulate specificity and stability of collagen. Proc Natl Acad Sci U S A. 2018;115(24):6207–12.

165. Boudko SP, Engel J, Okuyama K, Mizuno K, Bächinger HP, Schumacher MA. Crystal structure of human type III collagen Gly991-Gly 1032 cystine knot-containing peptide shows both 7/2 and 10/3 triple helical symmetries. J Biol Chem. 2008;283(47):32580–9.

166. Svojtková E, Deyl Z, Adam M. Aging of connective tissue. Age dependency of collagen depolymerization by ionizing radiation. Exp Gerontol. 1973;8(3):157–64.

167. Bowes JH, Moss JA. The Effect of Gamma Radiation on Collagen. Radiat Res. 1962 Mar;16(3):211.

168. Alves SM, Zhu T, Shostak A, Rossen NS, Rafat M. Studying normal tissue radiation effects using extracellular matrix hydrogels. J Vis Exp. 2019;2019(149):1–8.

169. Giobbe GG, Crowley C, Luni C, Campinoti S, Khedr M, Kretzschmar K, et al. Extracellular matrix hydrogel derived from decellularized tissues enables endodermal organoid culture. Nat Commun. 2019;10(1).

170. Moradi L, Mohammadi Jobania B, Jafarnezhad-Ansariha F, Ghorbani F, Esmaeil-Pour R, Majidi Zolbina M, et al. Evaluation of different sterilization methods for decellularized kidney tissue. Tissue Cell. 2020;66(62):101396.

171. XIslam MM, Sharifi R, Mamodaly S, Islam R, Nahra D, Abusamra DB, et al. Effects of gamma radiation sterilization on the structural and biological properties of decellularized corneal xenografts. Acta Biomater. 2019;96:330–44.

172. Poornejad N, Nielsen JJ, Morris RJ, Gassman JR, Reynolds PR, Roeder BL, et al. Comparison of four decontamination treatments on porcine renal decellularized extracellular matrix structure, composition, and support of renal tubular epithelium cells. J Biomater Appl. 2016 Mar 20;30(8):1154–67.

173. Riekki, Arja Jukkola, Aarne Oikarin R. Radiation Therapy Induces Tenascin Expression and Angiogenesis in Human Skin. Acta Derm Venereol. 2001 Nov 1;81(5):329–33.

174. Mohamed F, Bradley DA, Winlove CP. Effects of ionizing radiation on extracellular matrix. Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip. 2007 Sep;580(1):566–9.

175. Pendleton MM, Emerzian SR, Liu J, Tang SY, O'Connell GD, Alwood JS, et al. Effects of ex vivo ionizing radiation on collagen structure and whole-bone mechanical properties of mouse vertebrae. Bone. 2019;128(June):115043.

176. Koontz BF, Verhaegen F, De Ruysscher D. Tumour and normal tissue radiobiology in mouse models: how close are mice to mini-humans? Br J Radiol. 2017 Jan;90(1069):20160441.

177. Kotova SL, Timashev PS, Belkova G V., Kochueva M V., Babak K V., Timofeeva VA, et al. Early Effects of Ionizing Radiation on the Collagen Hierarchical Structure of Bladder and Rectum Visualized by Atomic Force Microscopy. Microsc Microanal. 2018;24(1):38–48.

178. Flanders KC, Sullivan CD, Fujii M, Sowers A, Anzano MA, Arabshahi A, et al. Mice lacking Smad3 are protected against cutaneous injury induced by ionizing radiation. Am J Pathol. 2002;160(3):1057–68.

179. Maslennikova A, Kochueva M, Ignatieva N, Vitkin A, Zakharkina O, Kamensky V, et al. Effects of gamma irradiation on collagen damage and remodeling. Int J Radiat Biol. 2015 Mar 4;91(3):240–7.

180. Darby SC, Ewertz M, McGale P, Bennet AM, Blom-Goldman U, Brønnum D, et al. Risk of Ischemic Heart Disease in Women after Radiotherapy for Breast Cancer. N Engl J Med. 2013;368(11):987–98.

181. Takeuchi H, Kimura T, Okamoto K, Aoyagi E, Miyamoto H, Kaji M, et al. A mechanism for abnormal angiogenesis in human radiation proctitis: Analysis of expression profile for angiogenic factors. J Gastroenterol. 2012;47(1):56–64.

182. Yoo KC, Suh Y, An Y, Lee HJ, Jeong YJ, Uddin N, et al. Proinvasive extracellular matrix remodeling in tumor microenvironment in response to radiation. Oncogene. 2018;37(24):3317–28.

183. Rancati T, Ceresoli GL, Gagliardi G, Schipani S, Cattaneo GM. Factors predicting radiation pneumonitis in lung cancer patients: A retrospective study. Radiother Oncol. 2003;67(3):275–83.

184. Haviland JS, Owen JR, Dewar JA, Agrawal RK, Barrett J, Barrett-Lee PJ, et al. The UK Standardisation of Breast Radiotherapy (START) trials of radiotherapy hypofractionation for treatment of early breast cancer: 10-year follow-up results of two randomised controlled trials. Lancet Oncol. 2013;14(11):1086–94.

185. Eckersley A, Ozols M, O'Cualain R, Keevill EJ, Foster A, Pilkington S, et al. Proteomic fingerprints of damage in extracellular matrix assemblies. Matrix Biol Plus. 2020;5:100027.

186. O'Hare MJ, Bond J, Clarke C, Takeuchi Y, Atherton AJ, Berry C, et al. Conditional immortalization of freshly isolated human mammary fibroblasts and endothelial cells. Proc Natl Acad Sci U S A. 2001;98(2):646–51.

187. Holmes DF, Tait A, Hodson NW, Sherratt MJ, Kadler KE. Growth of collagen fibril seeds from embryonic tendon: Fractured fibril ends nucleate new tip growth. J Mol Biol. 2010;399(1):9–16.

188. REISFELD RA, LEWIS UJ, WILLIAMS DE. Disk Electrophoresis of Basic Proteins and Peptides on Polyacrylamide Gels. Nature. 1962 Jul;195(4838):281–3.

189. Lebendiker M. ACIDIC-NATIVE GEL Protocol [Internet]. 2002 [cited 2022 Nov 12]. p. 0–1. Available from: http://wolfson.huji.ac.il/purification/Protocols/PAGE_Acidic.html

190. Pittenger B, Erina N, Su C. Quantitative Mechanical Property Mapping at the Nanoscale with PeakForce QNM. Appl Note Veeco Instruments Inc. 2010;1(1):1–11.

191. Belikov S, Alexander J, Wall C, Yermolenko I, Magonov S, Malovichko I. Thermal tune method for AFM oscillatory resonant imaging in air and liquid. Proc Am Control Conf. 2014;1009–14.

192. Kontomaris S V., Malamou A, Stylianou A. The Hertzian theory in AFM nanoindentation experiments regarding biological samples: Overcoming limitations in data processing. Micron. 2022;155(December 2021):103228.

193. Nečas D, Klapetek P. Gwyddion: An open-source software for SPM data analysis. Cent Eur J Phys. 2012;10(1):181–8.

194. Biesiadecki BJ, Jin JP. A high-throughput solid-phase microplate protein-binding assay to investigate interactions between myofilament proteins. J Biomed Biotechnol. 2011;2011.

195. Minton AP. Static light scattering from concentrated protein solutions, I: General theory for protein mixtures and application to self-associating proteins. Biophys J. 2007;93(4):1321–8.

196. Vivian JT, Callis PR. Mechanisms of tryptophan fluorescence shifts in proteins. Biophys J. 2001;80(5):2093–109.

197. Ozols M, Eckersley A, Mellody KT, Mallikarjun V, Warwood S, O'Cualain R, et al. Peptide location fingerprinting reveals modification-associated biomarker candidates of ageing in human tissue proteomes. Aging Cell. 2021;20(5):1–16.

198. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(D1):D638–46.

199. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013;41(D1):808–15.

200. Ravichandran A, Clegg J, Adams MN, Hampson M, Fielding A, Bray LJ. 3D Breast Tumor Models for Radiobiology Applications. Cancers (Basel). 2021;13(22):1–20.

201. Swamydas M, Eddy JM, Burg KJL, Dréau D. Matrix compositions and the development of breast acini and ducts in 3D cultures. Vitr Cell Dev Biol - Anim. 2010;46(8):673–84.

202. Johansson S, Svineng G, Wennerberg K, Armulik A, Lohikangas L. Fibronectin-integrin interactions. Front Biosci. 1997;2(11).

203. Ozols M, Eckersley A, Platt CI, Mcguinness CS, Hibbert SA, Revote J, et al. Predicting and validating protein degradation in proteomes using deep learning. bioRxiv. 2020;44(0):2020.11.29.402446.

204. Chang RL, Stanley JA, Robinson MC, Sher JW, Li Z, Chan YA, et al. Protein structure, amino acid composition and sequence determine proteome vulnerability to oxidation-induced damage. EMBO J. 2020;39(23):1–21.

205. Sharp JS, Becker JM, Hettich RL. Analysis of Protein Solvent Accessible Surfaces by Photochemical Oxidation and Mass Spectrometry. Anal Chem. 2004;76(3):672–83.

206. Deshmukh SN, Dive AM, Moharil R, Munde P. Enigmatic insight into collagen. J Oral Maxillofac Pathol. 2016;20(2):276–83.

207. Bella J, Hulmes DJS. Fibrillar collagens. Vol. 82, Subcellular Biochemistry. 2017. 457–490 p.

208. Muiznieks LD, Keeley FW. Molecular assembly and mechanical properties of the extracellular matrix: A fibrous protein perspective. Biochim Biophys Acta - Mol Basis Dis. 2013;1832(7):866–75.

209. Corbett SA, Lee L, Wilson CL, Schwarzbauer JE. Covalent cross-linking of fibronectin to fibrin is required for maximal cell adhesion to a fibronectin-fibrin matrix. J Biol Chem. 1997;272(40):24999–5005.

210. Knox P, Crooks S, Rimmer CS. Role of fibronectin in the migration of fibroblasts into plasma clots. J Cell Biol. 1986 Jun 1;102(6):2318–23.

211. Pankov R. Fibronectin at a glance. J Cell Sci. 2002 Oct 15;115(20):3861–3.

212. Mao Y, Schwarzbauer JE. Fibronectin fibrillogenesis, a cell-mediated matrix assembly process. Matrix Biol. 2005;24(6):389–99.

213. Schaffner F, Ray AM, Dontenwill M. Integrin α5β1, the fibronectin receptor, as a pertinent therapeutic target in solid tumors. Cancers (Basel). 2013;5(1):27–47.

214. Erat MC, Sladek B, Campbell ID, Vakonakis I. Structural analysis of collagen type I interactions with human fibronectin reveals a cooperative binding mode. J Biol Chem. 2013;288(24):17441–50.

215. Owens RJ, Baralle FE. Mapping the collagen-binding site of human fibronectin by expression in Escherichia coli. EMBO J. 1986;5(11):2825–30.

216. Paten JA, Martin CL, Wanis JT, Siadat SM, Figueroa-Navedo AM, Ruberti JW, et al. Molecular Interactions between Collagen and Fibronectin: A Reciprocal Relationship that Regulates De Novo

Fibrillogenesis. Chem. 2019;5(8):2126–45.

217. Van der Flier A, Sonnenberg A. Function and interactions of integrins. Cell Tissue Res. 2001;305(3):285–98.

218. Kennett EC, Chuang CY, Degendorfer G, Whitelock JM, Davies MJ. Mechanisms and consequences of oxidative damage to extracellular matrix. Biochem Soc Trans. 2011;39(5):1279–87.

219. Davis GE, Bayless KJ, Davis MJ, Meininger GA. Regulation of tissue injury responses by the exposure of matricryptic sites within extracellular matrix molecules. Am J Pathol. 2000;156(5):1489–98.

220. Joladarashi D, Salimath P V., Chilkunda ND. Diabetes results in structural alteration of chondroitin sulfate/dermatan sulfate in the rat kidney: Effects on the binding to extracellular matrix components. Glycobiology. 2011;21(7):960–72.

221. Hwang J, San BH, Turner NJ, White LJ, Faulk DM, Badylak SF, et al. Molecular assessment of collagen denaturation in decellularized tissues using a collagen hybridizing peptide. Acta Biomater. 2017;53:268–78.

222. Leikina E, Mertts M V., Kuznetsova N, Leikin S. Type I collagen is thermally unstable at body temperature. Proc Natl Acad Sci U S A. 2002;99(3):1314–8.

223. Drzewiecki KE, Grisham DR, Parmar AS, Nanda V, Shreiber DI. Circular Dichroism Spectroscopy of Collagen Fibrillogenesis: A New Use for an Old Technique. Biophys J. 2016;111(11):2377–86.

224. Rödiger S, Böhm A, Schimke I. Surface melting curve analysis with R. R J. 2013;5(2):37–52.

225. Eckersley A, Mellody KT, Pilkington S, Griffiths CEM, Watson REB, O'Cualain R, et al. Structural and compositional diversity of fibrillin microfibrils in human tissues. J Biol Chem. 2018;293(14):5117–33.

226. Wang X, Shen S, Rasam SS, Qu J. MS1 ion current-based quantitative proteomics: A promising solution for reliable analysis of large biological cohorts. Mass Spectrom Rev. 2019;38(6):461–82.

227. Vartio T. Regular fragmentation of hydrogen peroxide-treated fibronectin. J Biol Chem. 1989 Mar;264(8):4471–5.

228. Mu C, Li D, Lin W, Ding Y, Zhang G. Temperature induced denaturation of collagen in acidic solution. Biopolymers. 2007 Jul;86(4):282–7.

229. Chung L, Dinakarpandian D, Yoshida N, Layer-Fields JL, Fields GB, Visse R, et al. Collagenase unwinds triple-helical collagen prior to peptide bond hydrolysis. EMBO J. 2004;23(15):3020–30.

230. Orgel JPRO, San Antonio JD, Antipova O. Molecular and structural mapping of collagen fibril interactions. Connect Tissue Res. 2011;52(1):2–17.

231. Perumal S, Antipova O, Orgel JPRO. Collagen fibril architecture, domain organization, and triple-helical conformation govern its proteolysis. Proc Natl Acad Sci U S A. 2008;105(8):2824–9.

232. Bettinger J, Ghaemmaghami S. Methionine oxidation within the prion protein. Prion. 2020 Jan 1;14(1):193–205.

233. Nomura T, Kamada R, Ito I, Chuman Y, Shimohigashi Y, Sakaguchi K. Oxidation of methionine residue at hydrophobic core destabilitizes p53 tetrameric structure. Biopolymers. 2009;91(1):78–84.

234. Mason JD, Cone MT, Fry ES. Ultraviolet (250–550 nm) absorption spectrum of pure water. Appl Opt. 2016;55(25):7163.

235. Hubbell JH, Seltzer SM. X-Ray Mass Attenuation Coefficients [Internet]. 2004 [cited 2023 Aug 6]. p. NIST Standard Reference Database 126. Available from: https://www.nist.gov/pml/x-ray-mass-attenuation-coefficients

236. Lockhart C, Smith AK, Klimov DK. Methionine Oxidation Changes the Mechanism of Aβ Peptide Binding to the DMPC Bilayer. Sci Rep. 2019;9(1):1–12.

237. Kiselar JG, Maleknia SD, Sullivan M, Downard KM, Chance MR. Hydroxyl radical probe of protein surfaces using synchrotron X-ray radiolysis and mass spectrometry. Int J Radiat Biol. 2002 Jan 3;78(2):101–14.

238. Maleknia SD, Brenowitz M, Chance MR. Millisecond radiolytic modification of peptides by synchrotron X-rays identified by mass spectrometry. Anal Chem. 1999;71(18):3965–73.

239. Ringe D. What makes a binding site a binding site? Curr Opin Struct Biol. 1995;5(6):825–9.

240. Erat MC, Schwarz-Linek U, Pickford AR, Farndale RW, Campbell ID, Vakonakis I. Implications for collagen binding from the crystallographic structure of fibronectin 6FnI1-2FnII7FnI. J Biol Chem. 2010;285(44):33764–70.

241. Kubow KE, Vukmirovic R, Zhe L, Klotzsch E, Smith ML, Gourdon D, et al. Mechanical forces regulate the interactions of fibronectin and collagen i in extracellular matrix. Nat Commun. 2015;6.

242. Kadler KE, Hojima Y, Prockop DJ. Assembly of type I collagen fibrils de novo. Between 37 and 41°C the process is limited by micro-unfolding of monomers. J Biol Chem. 1988;263(21):10517–23.

243. Fields GB. Interstitial collagen catabolism. J Biol Chem. 2013;288(13):8785–93.

244. Persikov A V., Brodsky B. Unstable molecules form stable tissues. Proc Natl Acad Sci U S A. 2002;99(3):1101–3.

245. Ricard-Blum S, Salza R. Matricryptins and matrikines: Biologically active fragments of the extracellular matrix. Exp Dermatol. 2014;23(7):457–63.

246. Palmieri D, Camardella L, Ulivi V, Guasco G, Manduca P. Trimer carboxyl propeptide of collagen I produced by mature osteoblasts is chemotactic for endothelial cells. J Biol Chem. 2000;275(42):32658–63.

247. Gaggar A, Jackson PL, Noerager BD, O'Reilly PJ, McQuaid DB, Rowe SM, et al. A Novel Proteolytic Cascade Generates an Extracellular Matrix-Derived Chemoattractant in Chronic Neutrophilic Inflammation. J Immunol. 2008;180(8):5662–9.

248. Van Doren SR. Matrix metalloproteinase interactions with collagen and elastin. Matrix Biol. 2015 May;44–46:224–31.

249. Manka SW, Carafoli F, Visse R, Bihan D, Raynal N, Farndale RW, et al. Structural insights into triple-helical collagen cleavage by matrix metalloproteinase 1. Proc Natl Acad Sci U S A. 2012;109(31):12461–6.

250. Gjorevski N, Piotrowski AS, Varner VD, Nelson CM. Dynamic tensile forces drive collective cell migration through three-dimensional extracellular matrices. Sci Rep. 2015;5(February):11458.

251. Aragona M, Panciera T, Manfrin A, Giulitti S, Michielin F, Elvassore N, et al. A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. Cell. 2013;154(5):1047–59.

252. Kleinman HK, Wilkes CM, Martin GR. Interaction of fibronectin with collagen fibrils. Biochemistry. 1981 Apr 1;20(8):2325–30.

253. Speranza ML, Valentini G, Calligaro A. Influence of Fibronectin on the Fibrillogenesis of Type I and Type III Collagen. Top Catal. 1987 Jun;7(2):115–23.

254. To WS, Midwood KS. Plasma and cellular fibronectin: Distinct and independent functions during tissue repair. Fibrogenes Tissue Repair. 2011;4(1):1–17.

255. Maurer LM, Ma W, Mosher DF. Dynamic structure of plasma fibronectin. Crit Rev Biochem Mol Biol. 2016 Jul 3;51(4):213–27.

256. Stadtman ER, Levine RL. Free radical-mediated oxidation of free amino acids and amino acid residues in proteins. Amino Acids. 2003;25(3–4):207–18.

257. Shoulders MD, Raines RT. Collagen structure and stability. Annu Rev Biochem. 2009;78:929–58.

258. Godwin ARF, Dajani R, Zhang X, Thomson J, Holmes DF, Adamo CS, et al. Fibrillin microfibril structure identifies long-range effects of inherited pathogenic mutations affecting a key regulatory latent TGFβ-binding site. Nat Struct Mol Biol. 2023;30(5):608–18.

259. Biehn SE, Lindert S. Accurate protein structure prediction with hydroxyl radical protein footprinting data. Nat Commun. 2021;12(1):1–10.

260. Hulmes DJSS, Miller A. Quasi-hexagonal molecular packing in collagen fibrils. Nature. 1979 Dec;282(5741):878–80.

261. Harris JR, Soliakov A, Lewis RJ. In vitro fibrillogenesis of collagen type I in varying ionic and pH conditions. Micron. 2013;49:60–8.

262. Harris JR, Reiber A. Influence of saline and pH on collagen type I fibrillogenesis in vitro: Fibril polymorphism and colloidal gold labelling. Micron. 2007;38(5):513–21.

263. Rajan N, Habermehl J, Coté MF, Doillon CJ, Mantovani D. Preparation of ready-to-use, storable and reconstituted type I collagen from rat tail tendon for tissue engineering applications. Nat Protoc. 2007;1(6):2753–8.

264. Rubin AL, Drake MP, Davison PF, Pfahl D, Speakman PT, Schmitt FO. Effects of Pepsin Treatment on the Interaction Properties of Tropocollagen Macromolecules *. Biochemistry. 1965 Feb 1;4(2):181–90.

265. Zeugolis DI, Paul RG, Attenburrow G. Factors influencing the properties of reconstituted collagen fibers prior to self-assembly: Animal species and collagen extraction method. J Biomed Mater Res - Part A. 2008;86(4):892–904.

266. Orgel JPR., Miller A, Irving TC, Fischetti RF, Hammersley AP, Wess TJ. The In Situ Supermolecular Structure of Type I Collagen. Structure. 2001 Nov;9(11):1061–9.

267. Doyle AD. Generation of 3D collagen gels with controlled diverse architectures. Curr Protoc Cell Biol. 2016;2016(September):10.20.1-10.20.16.

268. Rossert J, de Crombrugghe B. Type I Collagen. In: Principles of Bone Biology. Elsevier; 2002. p. 189–XVIII.

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

270. Screen HRC, Lee DA, Bader DL, Shelton JC. An investigation into the effects of the hierarchical structure of tendon fascicles on micromechanical properties. Proc Inst Mech Eng Part H J Eng Med. 2004;218(2):109–19.

271. Fang M, Goldstein EL, Turner AS, Les CM, Orr BG, Fisher GJ, et al. Type i collagen D-spacing in fibril bundles of dermis, tendon, and bone: Bridging between nano- and micro-level tissue hierarchy. ACS Nano. 2012;6(11):9503–14.

272. Rowe RWD. The Structure of Rat Tail Tendon. Connect Tissue Res. 1985 Jan 7;14(1):9–20.

273. Rowe RWD. The Structure of Rat Tail Tendon Fascicles. Connect Tissue Res. 1985 Jan 7;14(1):21–30.

274. Rittié L. Type I collagen purification from rat tail tendons. Methods Mol Biol. 2017;1627:287–308.

275. Streeter I, De Leeuw NH. Atomistic modeling of collagen proteins in their fibrillar environment. J Phys Chem B. 2010;114(41):13263–70.

276. Bella J, Brodsky B, Berman HM. Hydration structure of a collagen peptide. Structure. 1995;3(9):893–906.

277. Kurien BT, Scofield RH. Common Artifacts and Mistakes Made in Electrophoresis. In: Methods in Molecular Biology. 2012. p. 633–40.

278. Wittig I, Karas M, Schägger H. High resolution clear native electrophoresis for in-gel functional assays and fluorescence studies of membrane protein complexes. Mol Cell Proteomics. 2007;6(7):1215–25.

279. Wittig I, Braun HP, Schägger H. Blue native PAGE. Nat Protoc. 2006;1(1):418–28.

280. O'Brien RW, White LR. Electrophoretic mobility of a spherical colloidal particle. J Chem Soc Faraday Trans 2 Mol Chem Phys. 1978;74:1607–26.

281. Fontana A, De Laureto PP, De Filippis V, Scaramella E, Zambonin M. Probing the partly folded states of proteins by limited proteolysis. Fold Des. 1997;2(2):17–26.

282. Hubbard SJ. The structural aspects of limited proteolysis of native proteins. Biochim Biophys Acta - Protein Struct Mol Enzymol. 1998;1382(2):191–206.

283. Hildonen S, Halvorsen TG, Reubsaet L. Why less is more when generating tryptic peptides in bottom-up proteomics. Proteomics. 2014;14(17–18):2031–41.

284. Robinson KA, Sun M, Barnum CE, Weiss SN, Huegel J, Shetye SS, et al. Decorin and biglycan are necessary for maintaining collagen fibril structure, fiber realignment, and mechanical properties of mature tendons. Matrix Biol. 2017;64(class I):81–93.

285. Redaelli A, Vesentini S, Soncini M, Vena P, Mantero S, Montevecchi FM. Possible role of decorin glycosaminoglycans in fibril to fibril force transfer in relative mature tendons - A computational study from molecular to microstructural level. J Biomech. 2003;36(10):1555–69.

286. Chen D, Smith LR, Khandekar G, Patel P, Yu CK, Zhang K, et al. Distinct effects of different matrix proteoglycans on collagen fibrillogenesis and cell-mediated collagen reorganization. Sci Rep. 2020;10(1):1–13.

287. Raspanti M, Viola M, Sonaggere M, Tira ME, Tenni R. Collagen fibril structure is affected by collagen concentration and decorin. Biomacromolecules. 2007;8(7):2087–91.

288. Elliott DM, Robinson PS, Gimbel JA, Sarver JJ, Abboud JA, Iozzo R V., et al. Effect of altered matrix proteins on quasilinear viscoelastic properties in transgenic mouse tail tendons. Ann Biomed Eng. 2003;31(5):599–605.

289. Herchenhan A, Uhlenbrock F, Eliasson P, Weis M, Eyre D, Kadler KE, et al. Lysyl oxidase activity is required for ordered collagen fibrillogenesis by tendon cells. J Biol Chem. 2015;290(26):16440–50.

290. Kagan HM, Li W. Lysyl oxidase: Properties, specificity, and biological roles inside and outside of the cell. J Cell Biochem. 2003;88(4):660–72.

291. Kolsrud H, Malerod H, Ray S, Reubsaet L, Lundanes E, Greibrokk T. A Critical Review of Trypsin Digestion for LC-MS Based Proteomics. Integr Proteomics. 2012;(October 2014).

292. Fang P, Liu M, Xue Y, Yao J, Zhang Y, Shen H, et al. Controlling nonspecific trypsin cleavages in LC-MS/MS-based shotgun proteomics using optimized experimental conditions. Analyst. 2015;140(22):7613–21.

293. Loziuk PL, Wang J, Li Q, Sederoff RR, Chiang VL, Muddiman DC. Understanding the role of proteolytic digestion on discovery and targeted proteomic measurements using liquid chromatography tandem mass spectrometry and design of experiments. J Proteome Res. 2013;12(12):5820–9.

294. Siepen JA, Keevil EJ, Knight D, Hubbard SJ. Prediction of missed cleavage sites in tryptic peptides aids protein identification in proteomics. J Proteome Res. 2007;6(1):399–408.

295. Thiede B, Lamer S, Mattow J, Siejak F, Dimmler C, Rudel T, et al. Analysis of missed cleavage sites, tryptophan oxidation and N-terminal pyroglutamylation after in-gel tryptic digestion. Rapid Commun Mass Spectrom. 2000;14(6):496–502.

296. Batool M, Ahmad B, Choi S. A structure-based drug discovery paradigm. Int J Mol Sci. 2019;20(11).

297. Śledź P, Caflisch A. Protein structure-based drug design: from docking to molecular dynamics. Curr

Opin Struct Biol. 2018;48:93–102.

298. Roeder BA, Kokini K, Voytik-Harbin SL. Fibril microstructure affects strain transmission within collagen extracellular matrices. J Biomech Eng. 2009;131(3):1–11.

299. Billiar KL. The Mechanical Environment of Cells in Collagen Gel Models: Global and Local Effects in Three-dimensional Biological Hydrogels. Stud Mechanobiol Tissue Eng Biomater. 2011;4(November 2010):201–45.

300. Pedersen JA, Swartz MA. Mechanobiology in the third dimension. Ann Biomed Eng. 2005;33(11):1469–90.

301. Van Helvert S, Friedl P. Strain Stiffening of Fibrillar Collagen during Individual and Collective Cell Migration Identified by AFM Nanoindentation. ACS Appl Mater Interfaces. 2016;8(34):21946–55.

302. Grant CA, Brockwell DJ, Radford SE, Thomson NH. Tuning the Elastic Modulus of Hydrated Collagen Fibrils. Biophys J. 2009 Dec;97(11):2985–92.

303. Wu X, Muthuchamy M, Reddy DS. Atomic force microscopy protocol for measurement of membrane plasticity and extracellular interactions in single neurons in epilepsy. Front Aging Neurosci. 2016;8(MAY):1–12.

304. Hayes A, Easton K, Devanaboyina PT, Wu JP, Kirk TB, Lloyd D. A review of methods to measure tendon dimensions. J Orthop Surg Res. 2019;14(1):1–12.

305. Lake SP, Miller KS, Elliott DM, Soslowsky LJ. Tensile properties and fiber alignment of human supraspinatus tendon in the transverse direction demonstrate inhomogeneity, nonlinearity, and regional isotropy. J Biomech. 2010;43(4):727–32.

306. Yoo L, Reed J, Shin A, Demer JL. Atomic force microscopy determination of Young's modulus of bovine extra-ocular tendon fiber bundles. J Biomech. 2014;47(8):1899–903.

307. Tobi D, Bahar I. Structural changes involved in protein binding correlate with intrinsic motions of proteins in the unbound state. Proc Natl Acad Sci U S A. 2005;102(52):18908–13.

308. Lawrence E, Sims J, Gander A, Garibaldi JM, Fuller B, Davidson B, et al. The Barriers and Motivators to Using Human Tissues for Research: The Views of UK-Based Biomedical Researchers. Biopreserv Biobank. 2020;18(4):266–73.

309. Randles MJ, Humphries MJ, Lennon R. Proteomic definitions of basement membrane composition in health and disease. Matrix Biol. 2017;57–58:12–28.

310. Kalluri R. Basement membranes: Structure, assembly and role in tumour angiogenesis. Nat Rev Cancer. 2003;3(6):422–33.

311. Pittayapruek P, Meephansan J, Prapapan O, Komine M, Ohtsuki M. Role of matrix metalloproteinases in Photoaging and photocarcinogenesis. Int J Mol Sci. 2016;17(6).

312. Chang TT, Thakar D, Weaver VM. Force-dependent breaching of the basement membrane. Matrix Biol. 2017;57–58:178–89.

313. Reuten R, Zendehroud S, Nicolau M, Fleischhauer L, Laitala A, Kiderlen S, et al. Basement membrane stiffness determines metastases formation. Nat Mater. 2021;20(6):892–903.

314. Mak KM, Mei R. Basement Membrane Type IV Collagen and Laminin: An Overview of Their Biology and Value as Fibrosis Biomarkers of Liver Disease. Anat Rec. 2017;300(8):1371–90.

315. Boudko SP, Danylevych N, Hudson BG, Pedchenko VK. Basement membrane collagen IV: Isolation of functional domains. 1st ed. Vol. 143, Methods in Cell Biology. Elsevier Inc.; 2018. 171–185 p.

316. Jayadev R, Sherwood DR. Basement membranes. Curr Biol. 2017;27(6):R207–11.

317. Pouliot N, Kusuma N. Laminin-511. Cell Adh Migr. 2013 Jan 27;7(1):142–9.

318. Hohenester E, Yurchenco PD. Laminins in basement membrane assembly. Cell Adhes Migr. 2013;7(1):56–63.

319. Hayes AJ, Farrugia BL, Biose IJ, Bix GJ, Melrose J. Perlecan, A Multi-Functional, Cell-Instructive, Matrix-Stabilizing Proteoglycan With Roles in Tissue Development Has Relevance to Connective Tissue Repair and Regeneration. Front Cell Dev Biol. 2022;10(April):1–23.

320. Brown JC, Sasaki T, Gohring W, Yamada Y, Timpl R. The C-terminal domain V of perlecan promotes pl integrin-mediated cell adhesion , binds heparin , nidogen. Eur J Biochem. 1997;250:39–46.

321. AUMAILLEY M, WIEDEMANN H, MANN K, TIMPL R. Binding of nidogen and the laminin-nidogen complex to basement membrane collagen type IV. Eur J Biochem. 1989 Sep;184(1):241–8.

322. Patel TR, Bernards C, Meier M, McEleney K, Winzor DJ, Koch M, et al. Structural elucidation of full-length nidogen and the laminin-nidogen complex in solution. Matrix Biol. 2014;33:60–7.

323. Hopf M, Göhring W, Ries A, Timpl R, Hohenester E. Crystal structure and mutational analysis of a perlecan-binding fragment of nidogen-1. Nat Struct Biol. 2001;8(7):634–40.

324. Fox JW, Mayer U, Nischt R, Aumailley M, Reinhardt D, Wiedemann H, et al. Recombinant nidogen consists of three globular domains and mediates binding of laminin to collagen type IV. EMBO J. 1991;10(11):3137–46.

325. Avagliano A, Fiume G, Ruocco MR, Martucci N, Vecchio E, Insabato L, et al. Influence of Fibroblasts on Mammary Gland Development, Breast Cancer Microenvironment Remodeling, and Cancer Cell Dissemination. Cancers (Basel). 2020 Jun 26;12(6):1697.

326. Unsworth A, Anderson R, Britt K. Stromal fibroblasts and the immune microenvironment: Partners in mammary gland biology and pathology? J Mammary Gland Biol Neoplasia. 2014;19(2):169–82.

327. Hedman K, Kurkinen M, Alitalo K, Vaheri A, Johansson S, Höök M. Isolation of the pericellular matrix of human fibroblast cultures. J Cell Biol. 1979 Apr 1;81(1):83–91.

328. Hoshiba T, Cho CS, Murakawa A, Okahata Y, Akaike T. The effect of natural extracellular matrix deposited on synthetic polymers on cultured primary hepatocytes. Biomaterials. 2006;27(26):4519–28.

329. Kendall RT, Feghali-Bostwick CA. Fibroblasts in fibrosis: Novel roles and mediators. Front Pharmacol. 2014;5 MAY(May):1–13.

330. Franco-Barraza J, Beacham DA, Amatangelo MD, Cukierman E. Preparation of extracellular matrices produced by cultured and primary fibroblasts. Curr Protoc Cell Biol. 2016;2016(June):10.9.1-10.9.34.

331. Merl-Pham J, Basak T, Knüppel L, Ramanujam D, Athanason M, Behr J, et al. Quantitative proteomic profiling of extracellular matrix and site-specific collagen post-translational modifications in an in vitro model of lung fibrosis. Matrix Biol Plus. 2019;1(2019):100005.

332. Kushner IK, Clair G, Purvine SO, Lee JY, Adkins JN, Payne SH. Individual Variability of Protein Expression in Human Tissues. J Proteome Res. 2018;17(11):3914–22.

333. Dickson D. Human tissue samples more difficult to obtain for academics. Nat Med. 2002;8(6):543.

334. Lavigne D, Guerrier L, Gueguen V, Michel JB, Boschetti E, Meilhac O, et al. Culture of human cells and synthesis of extracellular matrix on materials compatible with direct analysis by mass spectrometry. Analyst. 2010;135(3):503–11.

335. Hirsch C, Schildknecht S. In vitro research reproducibility: Keeping up high standards. Front Pharmacol. 2019;10(December):1–9.

336. Tracy LE, Minasian RA, Caterson EJ. Extracellular Matrix and Dermal Fibroblast Function in the Healing Wound. Adv Wound Care. 2016;5(3):119–36.

337. Shao X, Gomez CD, Kapoor N, Considine JM, Grams C, Gao Y, et al. MatrisomeDB 2.0: 2023 updates to the ECM-protein knowledge database. Nucleic Acids Res. 2023;51(1 D):D1519–30.

338. Kassner A, Hansen U, Miosge N, Reinhardt DP, Aigner T, Bruckner-Tuderman L, et al. Discrete integration of collagen XVI into tissue-specific collagen fibrils or beaded microfibrils. Matrix Biol. 2003;22(2):131–43.

339. Grässel S, Bauer RJ. Collagen XVI in health and disease. Matrix Biol. 2013;32(2):64–73.

340. Wu Z, Korntner SH, Mullen AM, Zeugolis DI. Collagen type II: From biosynthesis to advanced biomaterials for cartilage engineering. Biomater Biosyst. 2021;4(May).

341. Fichard A, Kleman JP, Ruggiero F. Another look at collagen V and XI molecules. Matrix Biol. 1995;14(7):515–31.

342. Sabatelli P, Gara SK, Grumati P, Urciuolo A, Gualandi F, Curci R, et al. Expression of the collagen VI α5 and α6 Chains in normal human skin and in skin of patients with collagen VI-related myopathies. J Invest Dermatol. 2011;131(1):99–107.

343. Carter WG. Transformation-dependent alterations in glycoproteins of the extracellular matrix of human fibroblasts. Characterization of GP250 and the collagen-like GP140. J Biol Chem. 1982;257(22):13805–15.

344. Kuo HJ, Maslen CL, Keene DR, Glanville RW. Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. J Biol Chem. 1997;272(42):26522–9.

345. Ono RN, Sengle G, Charbonneau NL, Calberg V, Bächinger HP, Sasaki T, et al. Latent transforming growth factor β-binding proteins and fibulins compete for fibrillin-1 and exhibit exquisite specificities in binding sites. J Biol Chem. 2009;284(25):16872–81.

346. Zilberberg L, Todorovic V, Dabovic B, Horiguchi M, Couroussé T, Sakai LY, et al. Specificity of latent TGF-β binding protein (LTBP) incorporation into matrix: Role of fibrillins and fibronectin. J Cell Physiol. 2012;227(12):3828–36.

347. Robertson IB, Horiguchi M, Zilberberg L, Dabovic B, Hadjiolova K, Rifkin DB. Latent TGF-β-binding proteins. Matrix Biol. 2015 Sep;47(1):44–53.

348. Chiquet-Ehrismann R, Kalla P, Pearson CA, Beck K, Chiquet M. Tenascin interferes with fibronectin action. Cell. 1988;53(3):383–90.

349. Yang Z, Strickland DK, Bornstein P. Extracellular Matrix Metalloproteinase 2 Levels are Regulated by the Low Density Lipoprotein-related Scavenger Receptor and Thrombospondin 2. J Biol Chem. 2001;276(11):8403–8.

350. Tu Y, Quan T. Oxidative stress and human skin connective tissue aging. Cosmetics. 2016;3(3):1–12.

351. Borkham-Kamphorst E, Schaffrath C, Van de Leur E, Haas U, Tihaa L, Meurer SK, et al. The anti-fibrotic effects of CCN1/CYR61 in primary portal myofibroblasts are mediated through induction of reactive oxygen species resulting in cellular senescence, apoptosis and attenuated TGF-β signaling. Biochim Biophys Acta - Mol Cell Res. 2014;1843(5):902–14.

352. Valcourt U, Alcaraz LB, Exposito JY, Lethias C, Bartholin L. Tenascin-X: Beyond the architectural function. Cell Adhes Migr. 2015;9(1–2):154–65.

353. Miller WL. Tenascin-X—Discovery and Early Research. Front Immunol. 2021;11(January):1–6.

354. Wu M, Ben Amar M. Growth and remodelling for profound circular wounds in skin. Biomech Model Mechanobiol. 2015;14(2):357–70.

355. Pieters M, Wolberg AS. Fibrinogen and fibrin: An illustrated review. Res Pract Thromb Haemost. 2019 Apr;3(2):161–72.

356. Martin BM, Ritchie AR, Toselli P, Franzblau C. Elastin synthesis and accumulation in irradiated smooth muscle cell cultures. Connect Tissue Res. 1992;28(3):181–9.

357. Evanko SP, Gooden MD, Kang I, Chan CK, Vernon RB, Wight TN. A Role for HAPLN1 During Phenotypic

Modulation of Human Lung Fibroblasts In Vitro. J Histochem Cytochem. 2020;68(11):797–811.

358. Neame PJ, Kay CJ, McQuillan DJ, Beales MP, Hassell JR. Independent modulation of collagen fibrillogenesis by decorin and lumican. Cell Mol Life Sci. 2000;57(5):859–63.

359. Ferdous Z, Peterson SB, Tseng H, Anderson DK, Iozzo R V., Grande-Allen KJ. A role for decorin in controlling proliferation, adhesion, and migration of murine embryonic fibroblasts. J Biomed Mater Res - Part A. 2010;93(2):419–28.

360. Yamaguchi Y, Mann DM, Ruoslahti E. Negative regulation of transforming growth factor-β by the proteoglycan decorin. Nature. 1990 Jul;346(6281):281–4.

361. Ezura Y, Chakravarti S, Oldberg A, Chervoneva I, Birk DE. Differential expression of lumican and fibromodulin regulate collagen fibrillogenesis in developing mouse tendons. J Cell Biol. 2000;151(4):779–87.

362. Kalamajski S, Bihan D, Bonna A, Rubin K, Farndale RW. Fibromodulin interacts with collagen cross-linking sites and activates lysyl oxidase. J Biol Chem. 2016;291(15):7951–60.

363. Krishnan A, Li X, Kao WY, Viker K, Butters K, Masuoka H, et al. Lumican, an extracellular matrix proteoglycan, is a novel requisite for hepatic fibrosis. Lab Investig. 2012;92(12):1712–25.

364. Bengtsson E, Mörgelin M, Sasaki T, Timpl R, Heinegård D, Aspberg A. The leucine-rich repeat protein PRELP binds perlecan and collagens and may function as a basement membrane anchor. J Biol Chem. 2002;277(17):15061–8.

365. Chiavarina B, Ronca R, Otaka Y, Sutton RB, Rezzola S, Yokobori T, et al. Fibroblast-derived prolargin is a tumor suppressor in hepatocellular carcinoma. Oncogene. 2022;41(10):1410–20.

366. Shimizu-Hirota R, Sasamura H, Kuroda M, Kobayashi E, Saruta T. Functional characterization of podocan, a member of a new class in the small leucine-rich repeat protein family. FEBS Lett. 2004;563(1–3):69–74.

367. Yun YR, Won JE, Jeon E, Lee S, Kang W, Jo H, et al. Fibroblast growth factors: Biology, function, and application for tissue regeneration. J Tissue Eng. 2010;1(1):1–18.

368. Martelossi Cebinelli GC, Paiva Trugilo K, Badaró Garcia S, Brajão de Oliveira K. TGF-β1 functional polymorphisms: A review. Eur Cytokine Netw. 2016;27(4):81–9.

369. Peng D, Fu M, Wang M, Wei Y, Wei X. Targeting TGF-β signal transduction for fibrosis and cancer therapy. Mol Cancer. 2022;21(1):1–20.

370. Suthon S, Perkins RS, Bryja V, Miranda-Carboni GA, Krum SA. WNT5B in Physiology and Disease. Front Cell Dev Biol. 2021;9(May):1–20.

371. Bueno MLP, Saad STO, Roversi FM. WNT5A in tumor development and progression: A comprehensive review. Biomed Pharmacother. 2022;155(July):113599.

372. Wischhusen J, Melero I, Fridman WH. Growth/Differentiation Factor-15 (GDF-15): From Biomarker to Novel Targetable Immune Checkpoint. Front Immunol. 2020;11(May).

373. Vale W, Wiater E, Gray P, Harrison C, Bilezikjian L, Choe S. Activins and inhibins and their signaling. Ann N Y Acad Sci. 2004;1038:142–7.

374. Xia Y, Schneyer AL. The biology of activin: recent advances in structure, regulation and function. J Endocrinol. 2009 Jul;202(1):1–12.

375. Fei F, Qu J, Li C, Wang X, Li Y, Zhang S. Role of metastasis-induced protein S100A4 in human non-tumor pathophysiologies. Cell Biosci. 2017;7(1):1–10.

376. Dufourcq P, Descamps B, Tojais NF, Leroux L, Oses P, Daret D, et al. Secreted Frizzled-Related Protein-1 Enhances Mesenchymal Stem Cell Function in Angiogenesis and Contributes to Neovessel Maturation. Stem Cells. 2008;26(11):2991–3001.

377. Üren A, Reichsman F, Anest V, Taylor WG, Muraiso K, Bottaro DP, et al. Secreted frizzled-related protein-1 binds directly to wingless and is a biphasic modulator of Wnt signaling. J Biol Chem. 2000;275(6):4374–82.

378. Hu M, Ling Z, Ren X. Extracellular matrix dynamics: tracking in biological systems and their implications. J Biol Eng. 2022;16(1):1–13.

379. Fan D, Kassiri Z. Biology of Tissue Inhibitor of Metalloproteinase 3 (TIMP3), and Its Therapeutic Implications in Cardiovascular Pathology. Front Physiol. 2020;11(June):1–16.

380. González-González L, Alonso J. Periostin: A matricellular protein with multiple functions in cancer development and progression. Front Oncol. 2018;8(JUN):1–15.

381. Dorafshan S, Razmi M, Safaei S, Gentilin E, Madjd Z, Ghods R. Periostin: biology and function in cancer. Cancer Cell Int. 2022;22(1):1–28.

382. Li J, Ma J, Zhang Q, Gong H, Gao D, Wang Y, et al. Spatially resolved proteomic map shows that extracellular matrix regulates epidermal growth. Nat Commun. 2022;13(1):1–16.

383. Schvartz I, Seger D, Shaltiel S. Vitronectin. Int J Biochem Cell Biol. 1999 May;31(5):539–44.

384. Colombatti A, Spessotto P, Doliana R, Mongiat M, Bressan GM, Esposito G. The EMILIN/multimerin family. Front Immunol. 2012;2(JAN):1–13.

385. Zanetti M, Braghetta P, Sabatelli P, Mura I, Doliana R, Colombatti A, et al. EMILIN-1 Deficiency Induces Elastogenesis and Vascular Cell Defects. Mol Cell Biol. 2004;24(2):638–50.

386. OKAMOTO O, FUJIWARA S, ABE M, SATO Y. Dermatopontin interacts with transforming growth factor β and enhances its biological activity. Biochem J. 1999 Feb 1;337(3):537–41.

387. Kato A, Okamoto O, Ishikawa K, Sumiyoshi H, Matsuo N, Yoshioka H, et al. Dermatopontin interacts with fibronectin, promotes fibronectin fibril formation, and enhances cell adhesion. J Biol Chem. 2011;286(17):14861–9.

388. Goh CS, Milburn D, Gerstein M. Conformational changes associated with protein-protein interactions. Curr Opin Struct Biol. 2004;14(1):104–9.

389. Weikl TR, Paul F. Conformational selection in protein binding and function. Protein Sci. 2014;23(11):1508–18.

390. Tandon H, de Brevern AG, Srinivasan N. Transient association between proteins elicits alteration of dynamics at sites far away from interfaces. Structure. 2021;29(4):371-384.e3.

391. Rittie L, Fisher GJ. Natural and Sun-Induced Aging of Human Skin. Cold Spring Harb Perspect Med. 2015 Jan 1;5(1):a015370–a015370.

392. Yamauchi M, Prisayanh P, Haque Z, Woodley DT. Collagen Cross-Linking in Sun-Exposed and Unexposed Sites of Aged Human Skin. J Invest Dermatol. 1991 Nov;97(5):937–40.

393. Langton AK, Alessi S, Hann M, Chien ALL, Kang S, Griffiths CEM, et al. Aging in Skin of Color: Disruption to Elastic Fiber Organization Is Detrimental to Skin's Biomechanical Function. J Invest Dermatol. 2019;139(4):779–88.

394. Gautieri A, Passini FS, Silván U, Guizar-Sicairos M, Carimati G, Volpi P, et al. Advanced glycation end-products: Mechanics of aged collagen from molecule to tissue. Matrix Biol. 2017;59:95–108.

395. Naba A, Clauser KR, Hoersch S, Liu H, Carr SA, Hynes RO. The matrisome: In silico definition and in vivo characterization by proteomics of normal and tumor extracellular matrices. Mol Cell Proteomics. 2012;11(4):1–18.

396. Melby JA, Roberts DS, Larson EJ, Brown KA, Bayne EF, Jin S, et al. Novel Strategies to Address the Challenges in Top-Down Proteomics. J Am Soc Mass Spectrom. 2021 Jun 2;32(6):1278–94.

397. Yuzhalin AE, Lim SY, Kutikhin AG, Gordon-Weeks AN. Dynamic matrisome: ECM remodeling factors

licensing cancer progression and metastasis. Biochim Biophys Acta - Rev Cancer. 2018;1870(2):207–28.

398. Johnson D, Boyes B, Fields T, Kopkin R, Orlando R. Optimization of data-dependent acquisition parameters for coupling high- speed separations with LC-MS/MS for protein identifications. J Biomol Tech. 2013;24(2):62–72.

399. Guo J, Huan T. Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography-Mass Spectrometry Based Untargeted Metabolomics. Anal Chem. 2020;92(12):8072–80.

400. Borza DB, Bondar O, Todd P, Sundaramoorthy M, Sado Y, Ninomiya Y, et al. Quaternary organization of the goodpasture autoantigen, the α3(IV) collagen chain. Sequestration of two cryptic autoepitopes by intraprotomer interactions with the α4 and α5 NC1 domains. J Biol Chem. 2002;277(42):40075–83.

401. Borza DB, Bondar O, Ninomiya Y, Sado Y, Naito I, Todd P, et al. The NC1 Domain of Collagen IV Encodes a Novel Network Composed of the α1, α2, α5, and α6 Chains in Smooth Muscle Basement Membranes. J Biol Chem. 2001;276(30):28532–40.

402. Kusafuka K, Muramatsu K, Kasami M, Kuriki K, Hirobe K, Hayashi I, et al. Cartilaginous features in matrix-producing carcinoma of the breast: Four cases report with histochemical and immunohistochemical analysis of matrix molecules. Mod Pathol. 2008;21(10):1282–92.

403. Verrecchia F, Mauviel A. Transforming growth factor-β signaling through the Smad pathway: Role in extracellular matrix gene expression and regulation. J Invest Dermatol. 2002;118(2):211–5.

404. Kanzaki T, Olofsson A, Morén A, Wernstedt C, Hellman U, Miyazono K, et al. TGF-β1 binding protein: A component of the large latent complex of TGF-β1 with multiple repeat sequences. Cell. 1990;61(6):1051–61.

405. Chute M, Aujla P, Jana S, Kassiri Z. The non-fibrillar side of fibrosis: Contribution of the basement membrane, proteoglycans, and glycoproteins to myocardial fibrosis. J Cardiovasc Dev Dis. 2019;6(4).

406. Belair DG, Le NN, Murphy WL. Design of growth factor sequestering biomaterials. Chem Commun. 2014;50(99):15651–68.

407. Wells JM, Gaggar A, Blalock JE. MMP generated matrikines. Matrix Biol. 2015;44–46:122–9.

408. Quan T, Qin Z, Xia W, Shao Y, Voorhees JJ, Fisher GJ. Matrix-degrading metalloproteinases in photoaging. J Investig Dermatology Symp Proc. 2009;14(1):20–4.

409. Zhang X, Chen CT, Bhargava M, Torzilli PA. A Comparative Study of Fibronectin Cleavage by MMP-1, -3, -13, and -14. Cartilage. 2012;3(3):267–77.

410. Monaco S, Sparano V, Gioia M, Sbardella D, Di Pierro D, Marini S, et al. Enzymatic processing of collagen IV by MMP-2 (gelatinase A) affects neutrophil migration and it is modulated by extracatalytic domains. Protein Sci. 2006;15(12):2805–15.

411. Whitelock JM, Murdoch AD, Iozzo R V., Underwood PA. The degradation of human endothelial cell-derived perlecan and release of bound basic fibroblast growth factor by stromelysin, collagenase, plasmin, and heparanases. J Biol Chem. 1996;271(17):10079–86.

412. Cui N, Hu M, Khalil RA. Biochemical and Biological Attributes of Matrix Metalloproteinases. 1st ed. Vol. 147, Progress in Molecular Biology and Translational Science. Elsevier Inc.; 2017. 1–73 p.

413. Itoh Y. MT1-MMP: A key regulator of cell migration in tissue. IUBMB Life. 2006;58(10):589–96.

414. Caley MP, Martins VLC, O'Toole EA. Metalloproteinases and Wound Healing. Adv Wound Care. 2015;4(4):225–34.

415. Jackson HW, Defamie V, Waterhouse P, Khokha R. TIMPs: Versatile extracellular regulators in cancer. Nat Rev Cancer. 2017;17(1):38–53.

416. Gaffney J, Solomonov I, Zehorai E, Sagi I. Multilevel regulation of matrix metalloproteinases in tissue

homeostasis indicates their molecular specificity in vivo. Matrix Biol. 2015;44–46(Ldl):191–9.

417. Kiontke A, Oliveira-Birkmeier A, Opitz A, Birkemeyer C. Electrospray ionization efficiency is dependent on different molecular descriptors with respect to solvent pH and instrumental configuration. PLoS One. 2016;11(12):1–16.

418. Schulze WX, Usadel B. Quantitation in mass-spectrometry-based proteomics. Annu Rev Plant Biol. 2010;61:491–516.

419. Amirrah IN, Lokanathan Y, Zulkiflee I, Wee MFMR, Motta A, Fauzi MB. A Comprehensive Review on Collagen Type I Development of Biomaterials for Tissue Engineering: From Biosynthesis to Bioscaffold. Biomedicines. 2022;10(9).

420. Nareyeck G, Seidler DG, Troyer D, Rauterberg J, Kresse H, Schönherr E. Differential interactions of decorin and decorin mutants with type I and type VI collagens. Eur J Biochem. 2004;271(16):3389–98.

421. Söder S, Pöschl E. The NC1 domain of human collagen IV is necessary to initiate triple helix formation. Biochem Biophys Res Commun. 2004;325(1):276–80.

422. Ries A, Engel J, Lustig A, Kühn K. The Function of the NC1 Domains in Type IV Collagen. J Biol Chem. 1995 Oct;270(40):23790–4.

423. Brown KL, Cummings CF, Vanacore RM, Hudson BG. Building collagen IV smart scaffolds on the outside of cells. Protein Sci. 2017;26(11):2151–61.

424. Moran T, Gat Y, Fass D. Laminin L4 domain structure resembles adhesion modules in ephrin receptor and other transmembrane glycoproteins. FEBS J. 2015;282(14):2746–57.

425. Harrison D, Hussain SA, Combs AC, Ervasti JM, Yurchenco PD, Hohenester E. Crystal structure and cell surface anchorage sites of laminin α1LG4-5. J Biol Chem. 2007;282(15):11573–81.

426. Wizemann H, Garbe JHO, Friedrich MVK, Timpl R, Sasaki T, Hohenester E. Distinct requirements for heparin and α-dystroglycan binding revealed by structure-based mutagenesis of the laminin α2 LG4-LG5 domain pair. J Mol Biol. 2003;332(3):635–42.

427. Hohenester E. Laminin G-like domains: dystroglycan-specific lectins. Curr Opin Struct Biol. 2019;56:56–63.

428. Kvansakul M, Hopf M, Ries A, Timpl R, Hohenester E. Structural basis for the high-affinity interaction of nidogen-1 with immunoglobulin-like domain 3 of perlecan. EMBO J. 2001;20(19):5342–6.

429. Goyal A, Pal N, Concannon M, Paul M, Doran M, Poluzzi C, et al. Endorepellin, the angiostatic module of perlecan, interacts with both the α2β1 integrin and vascular endothelial growth factor receptor 2 (VEGFR2): A dual receptor antagonism. J Biol Chem. 2011;286(29):25947–62.

430. Jung M, Lord MS, Cheng B, Lyons JG, Alkhouri H, Hughes JM, et al. Mast cells produce novel shorter forms of perlecan that contain functional endorepellin a role in angiogenesis and wound healing. J Biol Chem. 2013;288(5):3289–304.

431. Insua-Rodríguez J, Oskarsson T. The extracellular matrix in breast cancer. Vol. 97, Advanced Drug Delivery Reviews. Elsevier B.V.; 2016. p. 41–55.

432. Oskarsson T. Extracellular matrix components in breast cancer progression and metastasis. Breast. 2013;22(S2):S66–72.

433. Winkler J, Abisoye-Ogunniyan A, Metcalf KJ, Werb Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. Nat Commun. 2020;11(1):1–19.

434. Makogonenko E, Tsurupa G, Ingham K, Medved L. Interaction of fibrin(ogen) with fibronectin: Further characterization and localization of the fibronectin-binding site. Biochemistry. 2002;41(25):7907–13.

435. Lee SH, Cheng H, Yuan Y, Wu S. Regulation of ionizing radiation-induced adhesion of breast cancer cells to fibronectin by alpha5beta1 integrin. Radiat Res. 2014;181(6):650–8.

436. Weber L, Kirsch E, Muller P, Krieg T. Collagen type distribution and macromolecular organization of connective tissue in different layers of human skin. J Invest Dermatol. 1984;82(2):156–60.

437. Siddiqi NJ, Alhomida AS. Investigation into the distribution of total, free, peptide-bound, protein-bound, soluble- and insoluble-collagen hydroxyproline in various bovine tissues. J Biochem Mol Biol. 2003;36(2):154–8.

438. MacDonald ML, Favo D, Garver M, Sun Z, Arion D, Ding Y, et al. Laser capture microdissection–targeted mass spectrometry: a method for multiplexed protein quantification within individual layers of the cerebral cortex. Neuropsychopharmacology. 2019;44(4):743–8.

439. Espina V, Wulfkuhle JD, Calvert VS, VanMeter A, Zhou W, Coukos G, et al. Laser-capture microdissection. Nat Protoc. 2006;1(2):586–603.

440. Ingham KC, Brew SA, Migliorini M. Type I collagen contains at least 14 cryptic fibronectin binding sites of similar affinity. Arch Biochem Biophys. 2002;407(2):217–23.

441. Schasfoort RBM. Introduction to Surface Plasmon Resonance. In: Handbook of Surface Plasmon Resonance. The Royal Society of Chemistry; 2017. p. 1–26.

442. Oss M, Kruve A, Herodes K, Leito I. Electrospray ionization efficiency scale of organic compound. Anal Chem. 2010;82(7):2865–72.

443. O'Reilly FJ, Rappsilber J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. Nat Struct Mol Biol. 2018;25(11):1000–8.

444. Jafari M, Ansari-Pour N. Why, when and how to adjust your P values? Cell J. 2019;20(4):604–7.

445. Armstrong RA. When to use the Bonferroni correction. Ophthalmic Physiol Opt. 2014;34(5):502–8.

446. Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. Am J Hum Genet. 2007;81(6):1158–68.

447. Peeters MJ. Practical significance: Moving beyond statistical significance. Curr Pharm Teach Learn. 2016;8(1):83–9.

448. Zhu J, Hoop CL, Case DA, Baum J. Cryptic binding sites become accessible through surface reconstruction of the type I collagen fibril. Sci Rep. 2018;8(1):1–12.

449. van Obberghen-Schilling E, Tucker RP, Saupe F, Gasser I, Cseh B, Orend G. Fibronectin and tenascin-C: Accomplices in vascular morphogenesis during development and tumor growth. Int J Dev Biol. 2011;55(4–5):511–25.

450. Raitman I, Huang ML, Williams SA, Friedman B, Godula K, Schwarzbauer JE. Heparin-fibronectin interactions in the development of extracellular matrix insolubility. Matrix Biol. 2018;67:107–22.

451. Vesely I. The role of elastin in aortic valve mechanics. J Biomech. 1997 May;31(2):115–23.

452. Wang X, Ali M, Lacerda C. A Three-Dimensional Collagen-Elastin Scaffold for Heart Valve Tissue Engineering. Bioengineering. 2018 Aug 28;5(3):69.

453. Lee A, Hudson AR, Shiwarski DJ, Tashman JW, Hinton TJ, Yerneni S, et al. 3D bioprinting of collagen to rebuild components of the human heart. Science (80-). 2019 Aug 2;365(6452):482–7.

454. Weber KT, Sun Y, Tyagi SC, Cleutjens JPM. Collagen Network of the Myocardium: Function, Structural Remodeling and Regulatory Mechanisms. J Mol Cell Cardiol. 1994 Mar;26(3):279–92.

455. Thibeault DW, Mabry SM, Ekekezie II, Zhang X, Truog WE. Collagen Scaffolding During Development and Its Deformation With Chronic Lung Disease. Pediatrics. 2003 Apr 1;111(4):766–76.

456. Starcher BC. Lung Elastin and Matrix. Chest. 2000 May;117(5):229S-234S.

457. Graham HK, Hodson NW, Hoyland JA, Millward-Sadler SJ, Garrod D, Scothern A, et al. Tissue section AFM: In situ ultrastructural imaging of native biomolecules. Matrix Biol. 2010;29(4):254–60.

458. Divoux A, Clément K. Architecture and the extracellular matrix: The still unappreciated components of the adipose tissue. Obes Rev. 2011;12(5):494–503.

459. Mariman ECM, Wang P. Adipocyte extracellular matrix composition, dynamics and role in obesity. Cell Mol Life Sci. 2010;67(8):1277–92.

460. Chen JL, Duan L, Zhu W, Xiong J, Wang D. Extracellular matrix production in vitro in cartilage tissue engineering. J Transl Med. 2014;12(1):1–9.

461. Schnoor M, Cullen P, Lorkowski J, Stolle K, Robenek H, Troyer D, et al. Production of Type VI Collagen by Human Macrophages: A New Dimension in Macrophage Functional Heterogeneity . J Immunol. 2008;180(8):5707–19.

462. Witherel CE, Sao K, Brisson BK, Han B, Volk SW, Petrie RJ, et al. Regulation of extracellular matrix assembly and structure by hybrid M1/M2 macrophages. Biomaterials. 2021 Feb;269(1):120667.

463. Leung CM, de Haan P, Ronaldson-Bouchard K, Kim GA, Ko J, Rho HS, et al. A guide to the organ-on-a-chip. Nat Rev Methods Prim. 2022;2(1).

464. Bhatia SN, Ingber DE. Microfluidic organs-on-chips. Nat Biotechnol. 2014;32(8):760–72.

465. Huang Y, Zhang X, Ma Z, Li W, Zhou Y, Zhou J, et al. Size, separation, structural order, and mass density of molecules packing in water and ice. Sci Rep. 2013;3:1–5.

466. Erickson HP, Carrell N, McDonagh J. Fibronectin molecule visualized in electron microscopy: a long, thin, flexible strand. J Cell Biol. 1981 Dec 1;91(3):673–8.

467. Meesungnoen J, Jay-Gerin J-P. Charged Particle and Photon Interactions with Matter. Hatano Y, Katsumura Y, Mozumder A, editors. Charged Particle and Photon Interactions with Matter Recent Advances, Applications, and Interfaces. CRC Press; 2010. 355–392 p.

468. von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, et al. STRING: Known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res. 2005;33(DATABASE ISS.):433–7.