Expression And Post-Translational Modification Of Intermediate Filament Proteins In Colonic Mucosa Of Patients With Ulcerative Colitis: Role In Pathogenesis Of Colitis Associated Colorectal Neoplasia And Dysplasia

Dr Debabrata Majumdar

Supervisors:

Dr Bernard Corfe,

Dr Caroline Evans,

Professor Alan J Lobo

Registration no: 090258888

ACKNOWLEDGEMENTS

First of all, I would like to thank my supervisors Professor Alan Lobo, Dr Bernard Corfe and Dr Caroline Evans for their constant encouragement, support and guidance to help me complete the project. Although patient recruitment, the experimental work and data collection and analyses had been done on schedule, difficult personal circumstances, and a new job as a Consultant Gastroenterologist could have easily derailed the submission of the thesis on time. For this I am immensely thankful to my supervisors for having provided me with a timeframe for completion and giving up their valuable time to provide input towards completing the write-up.

I would also like to express my sincere gratitude to Bardhan Research and Education Trust for their grant support which enabled me to undertake the laboratory experiments.

As a clinician, this was the first time I was exposed to the harsh realities of basic science research. Without the help and guidance of the laboratory staff at the Department of Oncology, it would not have been possible for me to learn the important skills. Thank you to all, in particular Ms. Joanna Chowdry, Dr. Caroline Evans and Mrs. Jenny Globe. There are many others whose 'tips & tricks' made a big difference. I am indebted to my colleagues Miss. Ria Rosser (for having someone to discuss and share frustrations when things were not going right) and Ms. Alexandra Marsh (for the IHC data).

An important part of this study included recruitment of over 100 patients and controls and obtaining colonic biopsies at colonoscopy. This would not have been possible without the support and help from the endoscopy team at Royal Hallamshire Hospital.

Finally, without the constant support (and the occasional prodding) of my wife, Swati, and my two boys Patrick and James, this project would not have been completed on time.

CONTENTS

Publications 9
Presentations 10
Aims 12
Objectives 13
List of abbreviations 14
Abstract 17
1. Introduction 20
1.1. Ulcerative colitis 21
1.1.1. Disease distribution 21
1.1.2. Epidemiology 22
1.1.3. Aetiopathogenesis 22
1.1.4. Natural history of ulcerative colitis 23
1.2. Cancer in ulcerative colitis 24
1.2.1. Risk factors for colitis associated cancer 25
1.2.1.1. Duration of disease 25
1.2.1.2. Anatomical extent of colitis 25
1.2.1.3. Family history of colon cancer 26
1.2.1.4. Role of inflammation 26
1.2.1.5. Primary sclerosing cholangitis 27
1.2.2. Surveillance for dysplasia and colorectal cancers in ulcerative colitis 28
1.2.2.1. Limitations of current surveillance strategies 29
1.2.2.2. Need for pan-colonic biomarkers 29
1.3. Proteomics 31
1.3.1. Proteome and proteomics 31
1.3.1.1. Proteomic techniques 33
1.3.1.2. Protein separation techniques 34
1.3.1.2.1. One-dimensional gel electrophoresis mass spectrometry 34
1.3.1.2.2. Two-dimensional gel electrophoresis mass spectrometry 35
1.3.1.2.3. Shotgun proteomics(LC-MS/MS) 35
1.3.2. Proteomic studies in inflammatory bowel diseases 36
1.3.2.1. Studies investigating proteomic markers in IBD pathogenesis 36
1.3.2.2. Proteomics to identify diagnostic biomarkers in IBD 38

1.3.2.3. Markers of inflammation induced colon cancer pathogenesis 39
1.3.2.4. Markers of disease activity and response to treatment 42
1.3.2.5. Conclusion 44
1.4. Keratin and Intermediate filaments 46
1.4.1. Isolation and solubilisation of keratins 46
1.4.2. Post-translational modifications of keratins 46
1.4.3. Functions of keratin in normal epithelium 48
1.4.4. Keratins and inflammatory bowel disease 49
1.4.5. Keratins in colitis associated colorectal cancer 52
1.5. Hypothesis of the study 53
2. Patients and methods 54
2.1. Patient recruitment 55
2.1.1. Inclusion and exclusion criteria 55
2.1.1.1. Inclusion criteria 55
2.1.1.2. Exclusion criteria 55
2.1.2. Patient groups 56
2.1.3. Initial assessment of patients 57
2.1.4. Colonoscopy technique 57
2.1.5. Flexible sigmoidoscopy technique 57
2.1.6. Biopsy protocol 58
2.1.7. Categorising patients into groups 59
2.1.8. Pathological evaluation 59
2.1.9. Summary of investigation in the study 60
2.2. Reagents and chemicals 61
2.2.1. Antibodies for western immunoblotting 61
2.2.2. Reagents for protein assay 62
2.2.3. Reagents for cell culture 62
2.2.4. Reagents for WB and gel analysis 62
2.2.5. Reagents for mass spectrometry 64
2.2.6. Buffer solutions 65
2.2.6.1. Phosphate buffer saline 65
2.2.6.2. Running buffer 65
2.2.6.3. Wash buffer 66
2.2.6.4. Low detergent buffer 67
2.2.6.5. High detergent buffer 68
2.2.6.6. IF protein solubilising buffer 69

2.2.6.7. Buffer for strong cation exchange 69
2.2.6.8. Buffer for HPLC 70
2.2.6.9. Other buffers 70
2.2.7. SDS PAGE solution for preparing gels 71
2.3. Scientific equipment 72
2.4. Endoscopic equipment 73
2.5. Developing a technique for isolation of intermediate filament proteins from colorectal biopsies followed by parallel analyses of the isolated proteins using western immunoblotting and mass spectrometry. 74
2.5.1. Cell culture technique 74
2.5.1.1. Obtaining cell cultures 74
2.5.1.2. Splitting cell cultures 75
2.5.2. Intermediate filament protein extraction and solubilisation 75
2.5.2.1. Intermediate filament protein isolation from MCF-7 cancer cell lines 76
2.5.2.2. Isolation of intermediate filament proteins from human colorectal biopsies 76
2.5.2.3. Solubilisation of insoluble cytoskeletal fraction (fraction 3) in guanidine hydrochloride (GuHCl) and urea 79
2.5.2.4. Assessing suitability of guanidine hydrochloride as a solvent for cytoskeletal proteins obtained from colorectal biopsies 82
2.5.2.5. Sonication to maximise protein solubilisation 82
2.5.3. Removal of substances (GuHCl and urea) from dissolved cytoskeletal proteins, which could potentially interfere with effective protein separation on SDS-PAGE gels 84
2.5.4. Demonstration of adequate band intensity on SDS-PAGE with both urea and GuHCl 84
2.5.5. Demonstration of post- translational modifications (PTMs) in the resolved samples by western immunoblotting 84
2.6. Analysing colorectal biopsies using mass spectrometry and orthogonal validation using western blotting. 88
2.6.1. Protein concentration assay 88
2.6.2. Preparing samples for analysis by mass spectroscopy 88
2.6.2.1. Reduction and alkylation of proteins 88
2.6.2.2. Trypsin digestion step 88
2.6.2.3. iTRAQ (8-plex) labelling 89

2.6.2.4. Strong cation exchange (SCX) 91
2.6.2.4.1. Fractionation of peptides 91
2.6.2.4.2. Sample desalting 91
2.6.2.4.3. Shotgun proteomic analysis 91
2.6.2.4.4. iTRAQ data analysis 93
2.7. Orthogonal validation using western immunoblotting. 95
2.7.1. Transferring of resolved proteins onto PVDF membranes and immunoblotting 95
2.7.2. Probing of membranes with primary and secondary antibodies and analysing band intensities 95 95
2.7.3. Re-probing PVDF membrane
2.7.4. Densitometric analysis of immunoblots and calculating acetyl and phosphorylation ratios 96
2.8. Setting up the study and obtaining ethical approval 97
2.8.1. Time line for ethics and recruitment of patients 97
2.8.2. Ethical assessment 99
2.8.3. Grant application 99
2.9. Recruitment of patients 99
2.9.1. Categorisation of patients 99
2.9.2. Clinical, histological and endoscopic data 100
2.9.2.1. Ulcerative colitis patients and controls recruited in the study 100
2.9.2.2. Categorisation of UC phenotype according to disease activity 103
2.9.2.3. Patients with distal active disease only 103
2.9.3. Categorisation of patients for proteomic study 104
3. Mass spectrometric analysis of intermediate filaments 107
3.1. Introduction 108
3.2. Experimental design 109
3.3. Results 110
3.3.1. iTRAQ results: protein expression profile in patient categories compared to controls 110
3.3.2. Longstanding colitis is associated with increased mucosal levels of 114
3.3.3. Keratins 8, 18 and 19 are reduced in primary sclerosing cholangitis and in dysplasia 116
3.3.4. Acute inflammation in the colorectal mucosa is associated with reduced levels of mucosal cytoskeletal proteins 118
3.3.5. Hierarchical cluster analysis of patient groups based on identified 123

peptides
3.4. Conclusion 125
4. Validation of iTRAQ results by western blotting and demonstration of post-translational modifications in keratins 127
4.1. Introduction 128
4.2. Experimental design 129
4.3. Results 129
4.3.1. Validation of iTRAQ results by western blotting and demonstration of post-translational modifications in keratins 129
4.3.2. Intermediate filament proteins show different levels of expression in health and in disease states 130
4.3.3. Keratin 8 exists in multiple isoforms in controls as well as patients with UC 132
4.3.4. Post-translational modification in Keratin 8 (phosphorylation and acetylation) occurs in colonic mucosa 133
4.3.4.1. Phosphorylation in Keratin 8 occurs in healthy colonic mucosa as well in mucosa of patients with ulcerative colitis 133
4.3.4.2. Acetylation of K8 is noted only in LSPC 135
4.3.5. Densitometry analysis of keratin 8 phosphorylation in comparison to total K8 levels identifies distinct phenotypic patterns. 135
4.3.6. Mucosal vimentin levels are altered in UC and an increased vimentin to keratin ratio may be a marker of elevated cancer risk 140
4.4. Conclusion 142
5. Discussion 144
5.1. Introduction 145
5.2. Developing a technique for intermediate filament isolation from biopsies and parallel analysis of IF proteins by western blotting and mass spectrometry 148
5.3. Implication of altered keratin levels in the mucosa: an apparent failure of restoration of keratins could predispose to pathogenesis of colitis-associated cancer 149
5.4. Altered vimentin to keratin ratio: a marker of aggressive disease?
5.5. Conclusion 150
6. Bibliography 151
7. Appendices 154
8. List of tables and figures 168
8.1. Tables 186

8.2. Figures 187
189

PUBLICATIONS

1. Corfe B M, Majumdar D, Assadsangabi A, Marsh AMR, Cross SS, Connolly JB, Evans CA, Lobo AJ. Inflammation decreases keratin level in ulcerative colitis; inadequate restoration associates with increased risk of colitis-associated cancer. BMJ Open Gastro. 2015 18;2(1).
2. Majumdar D, Tiernan JP, Lobo AJ, Evans CA, Corfe BM. Keratins in colorectal epithelial function and disease. Int J Exp Pathol. 2012;93(5):305-18.
3. Majumdar D, Rosser R, Havard S, Lobo AJ, Wright PC, Evans CA, Corfe BM. An integrated workflow for extraction and solubilisation of intermediate filaments from colorectal biopsies for proteomic analysis. Electrophoresis. 2012;33(13):1967-74.

PRESENTATIONS

1. Increased Mucosal Expression of Insoluble Keratins 8, 18 And 19 In Long-standing Ulcerative Colitis In Comparison To Recent-onset Ulcerative Colitis: Validation of Mass Spectrometry Data. **BSG poster presentation** (2015).
2. The fate of epithelial keratins in active ulcerative colitis. **BSG poster presentation** (2015).
3. Keratin 8 Expression Is Reduced in Active Ulcerative Colitis Relative to Proximal Inactive Mucosa. **American Gastroenterology Association, Digestive Diseases Week (2015).**
4. Does alteration in mucosal Keratin 8 levels, phosphorylation and relative ratio to vimentin associate with development of colitis-associated cancer? **BSG Oral presentation** (2013).
5. Keratin expression profile in active colitis and in colitis phenotypes associated with increased cancer risk: an immunohistochemical assessment. **BSG Oral presentation** (2013).
6. Altered keratin expression and phosphorylation in mucosa of patients with longstanding pancolitis and dysplasia: potential biomarker of disease progression? **UEGW** Amsterdam **UEG Travel Grant for Basic Science Research (**2012).
7. Dynamic alterations in mucosal vimentin levels in patients with long standing pancolitis and acute ulcerative colitis: potential implications in disease pathogenesis **UEGW** Amsterdam (2012).
8. Quantitative proteomic analysis of intermediate filament profile in

ulcerative colitis reveals increased levels of keratins 8, 18 and 19 in patients with longstanding pan colitis, which are reduced, with development of dysplasia. **American Gastroenterology Association, Digestive Diseases Week,** San Diego, USA/**Digestive Diseases Federation** Liverpool (2012).

9. Quantitative proteomics in ulcerative colitis reveals mucosal inflammation reduces levels of keratins in the insoluble fraction of the intermediate filament proteome. **American Gastroenterology Association, Digestive Diseases Week,** San Diego, USA/ **Digestive Diseases Federation,** Liverpool. **Poster of distinction** (2012).

AIMS

1. Investigate expression, alteration in levels and post-translational modifications (PTMs) in keratin 8, 18 and 19 in colonic mucosa of patients with ulcerative colitis as a potential aetiological factor in either the pathogenesis of inflammatory changes in the mucosa or in development of colitis associated colorectal cancers.
2. Recruit and investigate the following categories of patients.

    a. Patients with active colitis compared with inactive colitis and controls: to identify proteomic changes as a consequence of inflammation on the colonic mucosa.
    b. Patients with active colitis comparing samples from inflamed mucosa with those from normal (i.e. uninvolved) mucosa in the same patient (i.e. above the "cut-off" between inflamed and normal tissue): to identify any changes in the mucosa between the inflamed and normal colonic mucosa in the same patient
    c. Patients with longstanding colitis compared with newly diagnosed colitis (with same disease activity scores): to identify any proteomic changes in the colonic mucosa as a consequence of increased duration of the disease.
    d. Mucosa in patients with cancer or high-grade dysplasia (HGD) elsewhere in the colon (to check for any field effect): since dysplasia in the colonic mucosa is often multifocal, we intend to identify proteomic changes in mucosa away from the area of dysplasia
    e. Colitis patients with concurrent primary sclerosing cholangitis (PSC): to identify changes in the colonic mucosa in this high-risk group.
3. Develop method for analysis of insoluble keratins in the colorectal mucosa

OBJECTIVES

1. Obtain ethical approval for the study
2. Recruit newly diagnosed patients with UC
3. Recruit previously diagnosed patients with a history of UC at follow-up endoscopy (stratified by duration since diagnosis of colitis)
4. Recruit control patients from non-urgent routine endoscopy lists
5. Obtain biopsies from active and quiescent regions of the colon at colonoscopy or flexible sigmoidoscopy
6. Undertake biopsy fractionation and separation of intermediate filament proteins
7. Develop strategies for effective solubilisation of IF proteins to enable parallel analyses by mass spectrometry and immunoblotting
8. Undertake mass spectrometry of pooled samples categorised into distinct 8 groups of patients and controls
9. Validate proteomic targets using immunoblotting approaches
10. Analyse biopsies for alterations in keratin 8, 18 and 19 expression
11. Analyse alterations in K8 phosphorylation and acetylation by immunoblotting approaches

List of abbreviations

1DE	One dimensional electrophoresis
2DE	Two dimensional electrophoresis
A1BG	Alpha-1B-glycoprotein
A2GL	Alpha-2-glycoprotein
ACT	Distal active colitis
AMBP	α-1-microglobulin
AOM	Azoxymethane
APOA1	Apolipoprotein A-I (),
APOE	Apolipoprotein E
ASA	Amino salicylic acid
BSG	British Society of Gastroenterology
C1R	Complement C1R
CAC	Colitis associated cancer
CD	Crohn’s disease
CI	Confidence interval
CLUS	Clusterin
CMGB	Chromogranin B
CO4B	Complement C4-B
COC3	Complement C3
CON	Healthy controls
CPS	Carbamoyl phosphate synthase
CRC	Colorectal cancer
DAI	Disease activity index
DIGE	Difference gel electrophoresis
DR	Pancolitis with dysplasia with biopsies taken from the rectum (DR)
DSS	Dextran sodium sulphate
DT	Biopsies obtained from dysplastic/neoplastic lesions (DT) in patients with UC
EGTA	Ethylene glycol tetra acetic acid
ESI	Electrospray ionisation
FIBA	fibrinogen alpha chain

GuHCl	Guanidine hydrochloride
HDAC	Histone deacetylase
HDB	High detergent buffer
HGD	High grade dysplasia
HNP	Human neutrophil peptide
HPLC	High-performance liquid chromatography
HSP	Heat shock protein
IBD	Inflammatory bowel disease
ICAT	Isotope-coded affinity tags
IF	Intermediate filament
IHC	Immunohistochemistry
IL	Interleukin
INACT	Inactive proximal colonic segment in those with distal colitis
ITRAQ	Isobaric tagging for relative and absolute quantification
K	Keratin
LDB	Low detergent buffer
LGD	Low grade dysplasia
LMW	Low molecular weight
LSPC	Long-standing pancolitis
LTQ	Linear Trap Quadrupole
MALDI	Matrix-assisted laser desorption/ionization
MDR	Multi drug resistance
MMTS	Methyl methanethiosulfonate
PMF	Peptide mass fingerprinting
PLMN	Plasminogen
PSC	Primary sclerosing colitis
PTM	Post translational modifications
PVDF	Polyvinylidene difluoride
ROC	Receiver operating characteristic
ROUC	Recent onset ulcerative colitis
RPLC	Reversed-phase liquid chromatography

RPMI	Roswell Park Memorial Institute medium
SCX	Strong cation exchange
SDS PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SELDI	Surface-enhanced laser desorption/ionization
SILAC	Stable isotope labeling by amino acids in cell culture
SPP-24	Secreted phosphoprotein 24
TBST	Tris-Buffered Saline with Tween 20
TCEP-HCl	Tris (2-carboxyethyl) phosphine hydrochloride
TEAB	Triethyl ammonium bicarbonate buffer
TFA	Trifluroacetic acid
TLR	Toll like receptor
TNF	Tumour necrosis factor
TNFR	TNFR
TOF	Time of flight
TRAP	TNF receptor associated protein
TRFE	Serotransferrin
TTHY	Transthyretin
UC	Ulcerative colitis
UDCA	Ursodeoxycholic acid
VIM	Vimentin
VTDB	Vitamin D-binding protein

ABSTRACT

Background

Ulcerative colitis(UC) is a chronic inflammatory disease of the colon associated with increased cancer risk of colitis associated cancer(CAC). Keratins(K) are intermediate filament(IF) proteins, and are key component of the cellular cytoskeleton. K8 null mice develop colitis; a subset of IBD patients have mutation in K8 gene. Keratins play a role in cell-death signalling pathways; epithelial cells lacking K8 and K18 are more sensitive to TNF-mediated apoptosis.

Methods

Colonic biopsies obtained from UC patients (and controls) were grouped into eight categories based on risk of cancer and presence of mucosal inflammation: quiescent recent onset (<5 years) UC (ROUC); quiescent long-standing pancolitis (20–40 years) (LSPC); UC with primary sclerosing cholangitis (PSC); active colitis (ACT) and un-inflamed proximal colonic mucosa (INACT) in the same patient; pancolitis with dysplasia-both dysplastic lesions (DT) and distal rectal mucosa (DR). An iTRAQ and western immunoblotting compatible extraction and solubilisation protocol for insoluble IF proteins was developed. Labelled peptides from pooled patient groups were analysed by quantitative proteomics. Results noted were validated by western immunoblotting.

Results

52 proteins were identified, 32(61.5%) were matched by 2 or more peptides. Acute inflammation was associated with reduced K8, K18, K19 and vimentin ($p<0.05$) compared to controls and un-inflamed mucosa; reduced levels were also seen in DT and DR. LSPC relative to controls or ROUC showed increased levels of IF proteins (K8, K18, K19 and vimentin, $p<0.05$). Multiple forms of K8 forms were identified on immunoblotting; in aggressive phenotypes relative K8 phosphorylation (K8pS23) was reduced along with an increase in vimentin:K8 ratio. Acute inflammation reduces K8 levels and

phosphorylation; such changes are restored in longstanding quiescent disease LSPC but not in ROUC (despite clinical and endoscopic remission).

Conclusion

Alteration in mucosal levels of IF proteins (keratin and vimentin) may play a role in pathogenesis of colitis associated cancers.

CHAPTER 1

INTRODUCTION

1.1 Ulcerative colitis

Ulcerative colitis (UC) is a chronic, inflammatory disease bowel disease (IBD) affecting the large bowel (colon) often with a relapsing and remitting clinical course [1]. The main symptoms of the condition include diarrhoea, often bloody diarrhoea with passage of mucus. Other symptoms which can be variably present include abdominal pain, tenesmus or fever [2].

1.1.1 Disease distribution

The distribution of the disease in the colon is quite variable and is summarised in table 1.1. The distribution of disease in the colon can progress or regress with time. The risk of proximal extension of proctitis (rectal involvement with colitis) is estimated to be between 41-54% over a 10-year period. Regression of the disease too may be occur; rates varying between 1.6% to 71% over 10 years have been noted [1].

Table 1.1 *Montreal classification of ulcerative colitis based on disease extent in the colon. Adapted from* [1]

Extent	Nomenclature	Anatomy
E1	Ulcerative proctitis	Involvement limited to the rectum (that is, proximal extent of inflammation is distal to the recto-sigmoid junction)
E2	Left sided UC (distal UC)	Involvement limited to a proportion of the colo-rectum distal to the splenic flexure
E3	Extensive UC (pancolitis)	Involvement extends proximal to the splenic flexure

1.1.2 Epidemiology

The incidence rate of UC in Europe varies from 1.5 to 20.3 cases per 100,000 person-years [3]. Its age-sex adjusted point prevalence was noted in a general practice database based study from Northern England at 243 per 100,000 persons (CI:217.4±269.4) [4]. Although it is usually recognized as a disease of younger people, with a peak incidence between 10 to 40 years, it can affect any age. In fact, 15% of patients are noted to be above 60 years of age at diagnosis [2].

1.1.3 Aetiopathogenesis

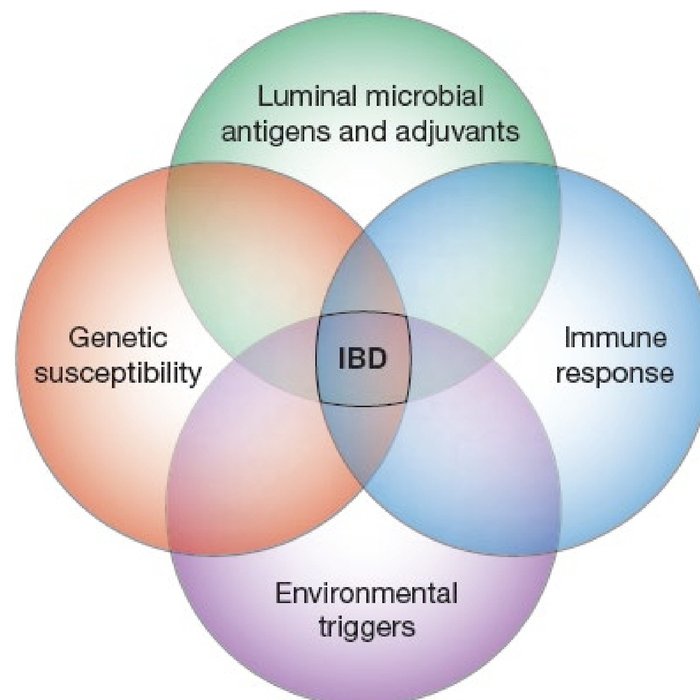
The aetiology of UC remains unknown and no single agent or distinct single mechanism has been implicated in the pathogenesis of ulcerative colitis [5]; it is rather believed that several factors may contribute. UC may develop in genetically susceptible individuals as a consequence of disturbance in the homeostasis between the normal commensal intestinal bacterial microflora and the host immune system. These processes result in activation of different immunological and non-immunological responses involving a variety of cell types [6, 7]. Various environmental factors are believed to contribute. This theory involves four diverse components: genetic susceptibility, environmental triggers and an altered immune response to luminal microbial antigens, which interact for the condition to become clinically apparent (figure 1.1) [8].

Environmental factors which may be significant include cigarette smoking and previous appendectomy; smokers have a significantly reduced risk of UC [9]. Appendectomy too is associated with almost 70% reduction in risk of developing UC [10]. It is felt that such environmental factors may trigger off the disease in susceptible individuals. There is an activation of the innate (macrophage, neutrophil) and acquired (T and B cell) immune responses and loss of tolerance to enteric commensal bacteria [11].

Several genomic loci have been identified in patients which increase the susceptibility to IBD in general; these may however represent only a small

subset of all genes involved in the pathogenesis of IBD [12]. The genes implicated in the pathogenesis of IBD are believed to modulate innate immune responses and barrier function of the colonic mucosa. *MDR1* is one such gene associated with ulcerative colitis [13]. Multi drug resistance gene 1 (*MDR1)* deleted mice have been noted to develop colitis [13-15].

Figure 1.1 *Various factors(viz. host immune response, environmental factors and gastrointestinal microbes), which interact and contribute to the pathogenesis of ulcerative colitis in genetically susceptible individuals. Adapted from [8]*



1.1.4 Natural history of the disease

The natural history of UC is unpredictable and characterised by disease flares (exacerbations) and quiescent periods (remission). Most patients present with mild to moderate disease; 10% will have severe disease at presentation [16, 17]. About 50% of patients experience a flare up of their symptoms in any year [2]. It is believed that approximately 15% of patients with UC will have a severe flare up of their disease at some point. A smaller number have a

frequently relapsing or chronic, continuous disease [2]. Around 10-40% of patients may have to undergo colectomy [16, 17]. In addition, a small number of patients with UC will develop colon cancers (see below).

1.2 Cancer in ulcerative colitis

Patients with long standing ulcerative colitis have an increased risk of developing colorectal cancer (CRC). IBD as a risk factor in the development of colorectal cancer was first described by Crohn and Rosenberg in 1925 [18]. Although CRC in IBD constitutes only 1% of all cases of colorectal cancers, patients with IBD represent one of the groups with highest risk for developing this complication [19].

Dysplasia (defined as unequivocal neoplasia of the epithelium confined to the basement membrane without invasion into the lamina propria) represents the pre-malignant phase of the condition [20]. Dysplasia in UC is usually classified as per Riddel *et al* [20] (Table 1.2).

***Table 1.2**: Classification of dysplasia in ulcerative colitis* [20]

1	**Negative for dysplasia**
2	Indefinite for dysplasia
3	Low grade dysplasia
4	High grade dysplasia

In patients who harbour low grade dysplasia and undergo immediate colectomy, 19% will already have concurrent CRC or high grade dysplasia

(HGD) [21]. HGD carries a 43% risk of synchronous malignancy and is considered to be an indication for immediate colectomy [21]

1.2.1 Risk factors for colitis associated cancers

Various factors have now been shown to be associated with an increased risk of colorectal cancer in UC. These are elucidated below.

1.2.1.1 Duration of colitis

Overall the most important factor identified across most studies is the duration of the disease. In a meta-analysis by Eaden *et al*, the cumulative CRC incidence was noted at 2% at 10 years, 8% at 20 years and 18% at 30 years [22]. The prospective surveillance programme at St Marks Hospital over 30 years showed that cumulative incidence of CRC by colitis duration was 2.5% at 20 years, 7.6% at 30 years, and 10.8% at 40 years [23]. Since the risk of CRC increases appreciably after 8 to 10 years of disease, British and American Gastroenterology Societies advocate surveillance colonoscopies to commence at this stage [24-26]. A retrospective study from The Netherlands has shown that the diagnosis of colorectal cancer is delayed or missed in a large number of patients (17–28%) when surveillance is conducted according to standard guidelines [27]. Despite that, it is generally accepted that commencing surveillance prior to 8 years of disease increases the cost of a surveillance program without any significant benefit [28].

1.2.1.2 Anatomical extent of disease in the colon

Extent of the disease in the colon is another important risk factor. Overall there is little or no increased risk of CRC in patients who have proctitis (UC limited to the rectum) or proctosigmoiditis (colitis involving the rectum and sigmoid colon); in patients with left sided disease, the risk is considered as intermediate [29, 30]. A population based cohort study from Sweden in patients with UC assessed by barium enema, showed a standard incidence ratio of 1.7 [95% confidence interval (CI) 0.8-3.2] in patients with proctitis; this was 2.8 (95% CI 1.6-4.4) in patients with left sided disease and in those with pancolitis was 14.8 (95% CI 11.4-18.9) [30]. Histological extent of disease in

the colon may be more important. A review of histology specimens in 30 patients who had undergone colectomy, showed that colitis associated neoplasia can occur in areas of the colon not grossly involved with colitis [31].

Heuschen *et al* evaluated the significance of backwash ileitis (the presence of mucosal inflammation in the terminal ileum in patients with pancolitis) in UC as a risk factor for CRC [32]. They prospectively analysed 590 consecutive patients with UC who had a restorative proctocolectomy. Although backwash ileitis was noted to increase the risk of colon cancer by 3 fold this has not been replicated in subsequent studies [33]. The increased risk noted in patients with backwash ileitis may just be a consequence of the fact that these patients have pancolitis.

1.2.1.3 Family history of colon cancer

A positive family history of sporadic CRC is associated with an increased risk of CRC in patients with UC. A population based cohort study from Sweden, found that a positive family history of CRC was associated with a more than 2-fold increased risk of CRC in patients with IBD (adjusted RR, 2.5; 95% CI, 1.4 – 4.4) [34]. IBD patients who had a first-degree relative diagnosed with CRC before the age of 50 were also noted in this study to have much higher risk (RR 9.2, 95% CI 3.7-23) [34]. In a recent case–control study of patients with pancolitis, family history of CRC was associated with an increased risk [HR 3.37 (1.02 – 11.14)] [35]. A family history of sporadic colorectal cancer was again noted as an independent risk factor for cancer in ulcerative colitis (odds ratio, 2.33; 95% CI, 1.06–5.14; P = 0.03) in a case-control study from the Mayo Clinic [36].

1.2.1.4 Role of inflammation

The role of inflammation in the pathogenesis and progression of colorectal and other cancers is well established [37, 38]. Once a tumour is initiated there is suppression of anti-tumour immunity; pro-angiogenesis and chronic stimulation of cellular proliferation promotes tumourigenesis [39].This has also been corroborated in animal (mouse) models of colitis. Mice administered dextran sulfate sodium (DSS) which causes colonic inflammation have shown

changes of dysplasia in the colon followed by development of colorectal cancer [40]. The pro-inflammatory cytokines TNF∝(tumour necrosis factor ∝), IL6 and IL 23 have been shown to contribute to the neoplastic changes [41-46].

Severity of inflammation in the colon is now recognised as an independent risk factor for colorectal cancer. In a case-control study by Rutter *et al*, 68 patients with long standing ulcerative colitis were assessed endoscopically. On univariate analysis, a significant correlation was noted between the colonoscopic (odds ratio, 2.5; $P = 0.001$) and histological (odds ratio, 5.1; $P < 0.001$) inflammation scores and risk of colorectal cancer. Multivariate analysis, showed the association of histological inflammation score as being significant (odds ratio, 4.7; $P < 0.001$) [47].

In another cohort study of 418 patients with UC, the severity of inflammation on histology in biopsy specimens was assessed using a 4-point scale. Progression to neoplasia was analysed using proportional hazards models with the inflammation. Of the study patients, 15 were noted to progress to advanced neoplasia (high grade dysplasia or colorectal cancer); 65 patients progressed to low-grade dysplasia, high-grade dysplasia, or colorectal cancer. Both univariate and multivariate analysis showed significant relationships between inflammation on histology and subsequent progression to advanced neoplasia (hazard ratio, 3.0; 95% CI, 1.4 – 6.3) [48].

Post-inflammatory pseudo polyps, a commonly noted benign finding at endoscopy that reflects underlying inflammation and regeneration has been shown to predict neoplastic changes. Patients with post-inflammatory polyps were noted by Rutter *et al* to have twice the risk of colorectal cancers (OR 2.29; 95% CI: 1.28–4.11) [49]. Similar findings were also noted in another study (OR, 2.5; 95% CI: 1.4–4.6) [50]

1.2.1.5 Primary Sclerosing Cholangitis

Primary sclerosing cholangitis (PSC) is a rare chronic fibrotic disease of the biliary tract of unknown aetiology. About 80% of patients with PSC have concomitant IBD [51, 52]. Association of PSC in patients with UC has been

shown to confer an increased risk of colorectal cancer. In a population based study from Sweden, the incidence of CRC was noted at 10%, 33% and 40% respectively at 10, 20 and 30 years respectively [53]. A meta analysis of 11 studies conducted by Soetikno *et al* noted an approximate 4-fold increased risk of colorectal neoplasia in patients with PSC and UC (OR 4.79, 95% CI 3.58-6.41) [54]. A recent Dutch study of 211 patients with PSC, of whom 60% had concomitant IBD showed that the 10-year and 20-year risks for CRC were 14% and 31% respectively. This was significantly higher than for patients without IBD (2% and 2% respectively (P = 0.008)) [55].

In patients with PSC and UC who had undergone orthotopic liver transplantation (for curative treatment of their biliary tract disease), the risk of CRC still remains high [56, 57]. Patients with UC and PSC should undergo annual colonoscopy commencing at the time of diagnosis of PSC, and continue indefinitely, even following liver transplantation [28]. Ursodeoxycholic acid (UDCA) (which is the most extensively studied medical therapy of this condition) use in patients with concomitant PSC and UC too has been shown to reduce the incidence of CRC and dysplasia [58]. There is good evidence that UDCA can induce histone hypoacetylation and senescence in colon cancer cells [59].

1.2.2 Surveillance for dysplasia and colorectal cancers in ulcerative colitis

Unlike sporadic CRC, which begins as dysplastic adenomas, usually in the form of polypoid lesions, dysplasia in IBD can arise in apparently normal appearing mucosa at endoscopy. Progression of neoplastic changes in UC usually occurs in a stepwise manner of histological changes. It commences from negative to indefinite for dysplasia and eventually progresses to low-grade dysplasia, high-grade dysplasia (HGD) and ultimately cancer [20] (Table 3). Dysplasia and cancers in colitis are often multifocal.

Current guidelines advocate routine surveillance colonoscopy and biopsies as the cornerstone of prevention [2, 26, 60]. Although colonoscopic surveillance in patients with UC has never been investigated by a randomised clinical trial,

this is still considered the gold standard for identification of histological dysplasia as markers of premalignant changes. The American Society for Gastrointestinal Endoscopy guidelines suggests quadrantic non-targeted biopsies for every 10 cm of colon, equating to 20–40 biopsies per colon [25]. Rubin *et al* noted that 33 biopsy specimens are needed to detect dysplasia with 90% probability [61]. A recent update to the British guidelines has highlighted chromoendoscopy and targeted biopsy as possible alternatives to multiple random biopsies [26].

1.2.2.1 Limitations of current surveillance strategies

The strategy of multiple colonic biopsies however has its limitations, perhaps most notably non-adherence to accepted guidelines [62]. A survey showed that 57% of British gastroenterologists took 10 or less biopsies during each colonoscopy, and only 2% of endoscopists took more than 20 biopsies [63], perhaps reflecting financial and time constraints both at colonoscopy and in the pathology department, as also a perceived lack of benefit[64]. In addition, there is also a potential for sampling error associated with non-targeted biopsies to look for flat dysplasia [65]. Tissue obtained from the colon at colonoscopic surveillance probably only samples less than 0.1% of the colonic mucosa[66]. The inter- and intra-pathologist-dependent interpretation of dysplasia in biopsy samples is another potential limitation [61, 62, 67]. Finally, the evidence demonstrating a benefit from surveillance colonoscopy is also limited. In particular, there is no evidence of prolongation of survival. Although cancers tend to be detected at an earlier stage with better prognosis, this could be accounted for by a ‘lead-time’ bias [68].

1.2.2.2 Need for a pan-colonic biomarker

In view of the perceived shortcomings of the current endoscopic strategies, alternate technologies are required to enable early identification and diagnoses of neoplastic changes in colitis, as also enable a better understanding of the pathogenesis of such changes. Biomarkers are specific proteins or molecules that act as surrogate markers of disease states [69]. In recent years as a consequence better understanding of protein biology and

their interactions, there is now an increased emphasis on identifying such biomarker molecules. Since CRC and dysplasia in colitis can develop insidiously and in apparently normal mucosa, biomarkers might enable us in future to identify such neoplastic changes at an early stage. Colitis associated CRC is usually preceded by and co-exists with dysplastic changes elsewhere in the bowel. There is therefore a potential to explore biomarkers that could be exploited to improve cancer detection and therefore surveillance.

Neoplastic changes in colitis have been shown to be associated with genomic abnormalities throughout the whole colon, including in the non-dysplastic mucosa. This is referred to as molecular field defect or genomic field defect [70-73]. Chen *et al* investigated patients with long standing UC, 11 of whom had dysplasia/cancer (termed progressors); the other group of 10 patients without any dysplasia (termed non-progressors) [71]. Polymerase chain reaction (PCR) based DNA fingerprinting was used to measure quantitatively genomic instability. UC progressors had marked genomic instability, which was noted in both their dysplastic and non-dysplastic colonic mucosa. These changes were not evident in the most of the non-progressors. The degree of genomic instability in non-dysplastic mucosa was similar to that noted in the dysplastic/cancerous mucosa from the same patient, suggesting that these changes are widespread. These results suggest that UC patients who develop dysplasia or cancer have an underlying process of genomic instability in their colonic mucosa. In another study using dual colour fluorescence *in situ* hybridization with centromere probes and locus-specific arm probes for chromosomes 8, 11, 17, and 18, chromosomal instability was shown to be present throughout the colon of UC patients with HGD or cancer [70].Since genomic field defects occur in the colonic mucosa, a pan colonic biomarker could potentially avoid the pitfalls of sampling error associated with random colonic mucosal biopsy.

It is increasingly recognised that gene expression does not correlate well at the level of protein expression or activation [74]. Information about DNA sequences or RNA expression alone cannot provide enough insight into the structure or interactions of various proteins [75]. Several other factors like

post-translational modifications (PTM) of proteins may be more important in controlling and regulating protein activation and stability. The emerging field of proteomics has developed in order to understand the expression, biological function, and regulation of the proteins [76, 77]. A range of diseases has been investigated using proteomic techniques, with the development of biomarkers for early cancer diagnosis, and only a few have focused on IBD (see below, section 1.3.2, page 36). Information obtained using proteomic techniques could potentially help in understanding complex biological processes occurring at a molecular level, in differing cell types, and their changes in a number of disease states [78]. It is currently felt that the new and emerging proteomic technologies may provide us with a better insight into critical changes that occur at different stages of IBD pathogenesis. They may also be used as novel biomarkers for diagnosis, assessing disease activity and in quantifying prognosis in different disease states.

1.3 Proteomics

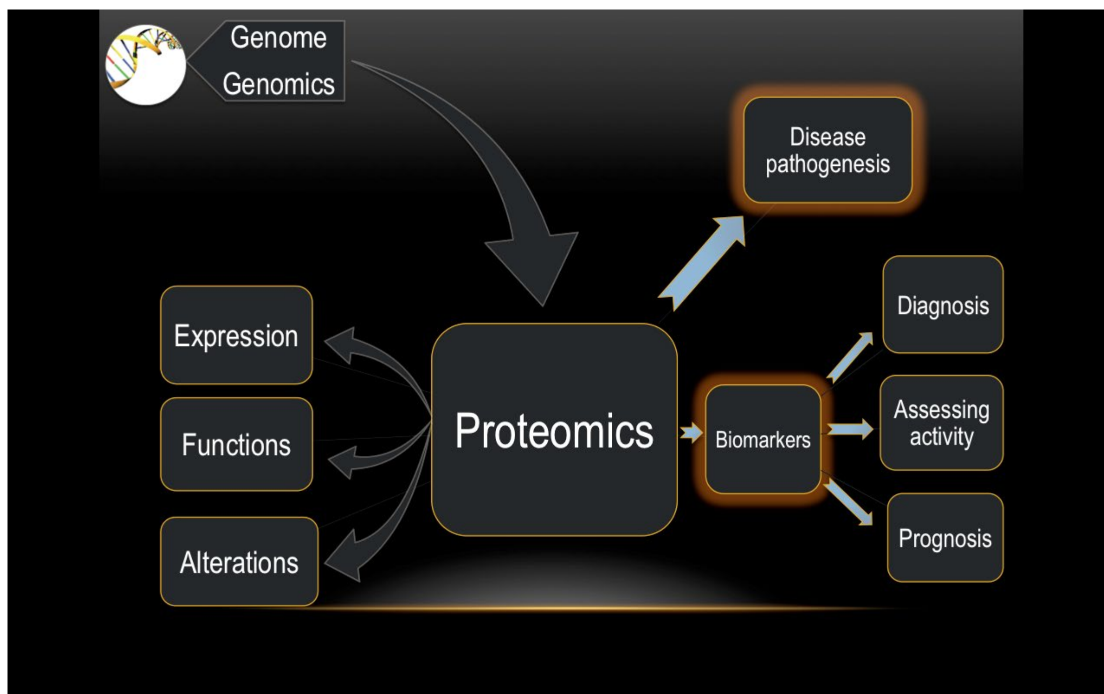
In recent years, 'Proteomics' the science of large-scale analysis of proteins in biological tissues has been developed as an alternative to traditional protein biochemistry, which tends to focus on one or few individual proteins. The available proteomic techniques, biology of proteins identified in inflammatory bowel disease and their alterations (and potential for use of proteomic technologies in diagnosis and management of this complex condition) are described below.

1.3.1 Proteome and proteomics

The term 'Proteome' the entire set of proteins encoded for by the genome, and 'Proteomics' (study of the proteome in biological tissues)(Figure 1.2) [79, 80] have been coined only fairly recently. Proteomics has emerged as an important tool in our understanding of various physiological and disease states. It is now known that in biological tissues, proteins constitute the main structural elements, as well as the messengers in signalling pathways [81]. It is increasingly being recognised that despite having mapped the human genome (and other plant and animal genomic sequences) and the protein

encoding genes in them, there is a marked discrepancy with the sheer number of proteins produced as well as in their biological functions [74]. A linear relationship between the proteome and the genes does not exist and a variety of other metabolic and regulatory pathways are involved [75]. Proteomics has enabled us to obtain a better understanding of the expression of proteins, their structure and alterations (isoforms and post translational modified forms) localisation in cellular tissue, interactions with other proteins (and signal pathways) as well as their physiological and pathophysiological functions [82, 83].

Figure 1.2 *The central role of proteomics in identifying expression, alteration and functions of proteins in biological tissues. This can enable identification of biomarkers, which could aid in diagnosing clinical conditions, assessing disease states and provide insight into disease prognosis.*



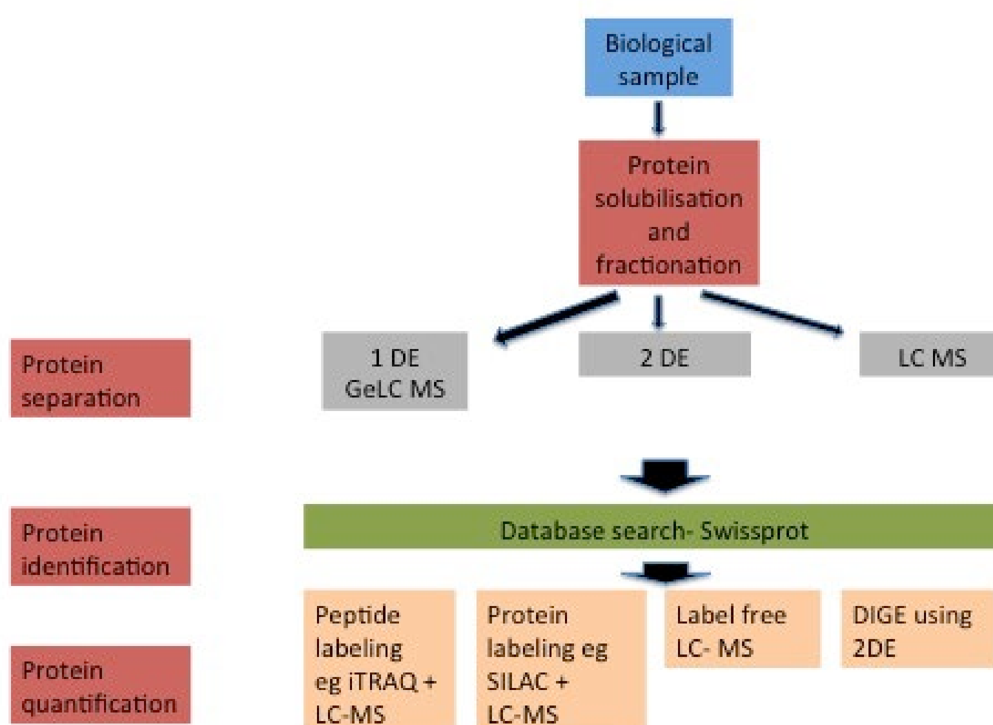
1.3.1.1 Proteomic techniques

The central technique in all proteomic strategies is tandem mass spectrometry (MS/MS), which enables a measure of the ratio of mass to charge for ionised constituents like peptides. Sample preparation and identifying the clinical context are also critically important in most proteomic experiments Essential strategies for analysing biological samples are summarised in figure 1.3.

A variety of mass spectrometry techniques including quantitative proteomic techniques [(isotope coded affinity tagging (ICAT), stable isotope labeling with amino acids in cell cultures (SILAC) and isobaric tag for relative and absolute quantitation (iTRAQ)], protein/antibody arrays, and multi-epitope-ligand cartograph have been developed for identifying and quantifying proteins [78]. Due to the complexity of biological tissues, a single technique may be

inappropriate for all [84]; in addition the role of sample preparation and subcellular fractionation acquires an additional significance.

Figure 1.3 *Proteomic analysis of biological samples commences with sample preparation. This might involve steps like protein depletion (to remove the most abundant protein eg. while analysing serum proteins). Following trypsin digestion, proteins are separated and identified using protein databases like Swissprot/Uniprot. Relative quantification of peptides is possible using techniques described* [78].



1.3.1.2 Protein separation techniques

There are in general three distinct techniques for protein separation from biological samples.

1.3.1.2.1 One dimensional electrophoresis

In 1DE, proteins are initially separated by standard gel electrophoresis. Subsequently the entire lane or identified bands in the gel in a defined

molecular weight distribution are excised into gel slices followed by trypsin digestion. The digested samples are then analysed by LC-MS. Although a time consuming process, an advantage of this technique is that the SDS-PAGE gel provides an indication of the molecular weights of the resolved protein samples, which can provide information about protein modifications [78]. In addition, solubilisation of proteins in SDS results in maximal yield of resolved proteins [78].

1.3.1.2.2 Two dimensional electrophoresis

In this technique, proteins are resolved initially according to their charge using specific immobilised pH gradient strips(IPG), by a process known as iso-electric focusing(IEF) followed by separation based on their molecular weight in the second dimension using standard SDS-PAGE gels. The resolved proteins are identified as discrete spots on gels following staining with specific stains like silver stain, Coomassie blue etc. The identified gel spots are then excised followed by trypsin digestion and subsequently analysed by either tandem MS (MS/MS) or peptide mass fingerprinting (PMF). One of the greatest advantages of using this technique lies in the ability to identify post-translational modifications in complex biological mixtures [85]. Nevertheless, solubilisation of protein mixtures remains a limitation particularly in the context of compatibility with IEF. In addition there are difficulties in separation of very basic or high or low molecular weight proteins [78].

1.3.1.2.3 Shotgun proteomics (LC-MS/MS)

A gel-free approach towards identifying peptides can be achieved by shotgun proteomic technique or bottom-up proteomics. In this technique the following initial digestion, the solubilised proteins are separated by liquid chromatography. Using chromatography columns, peptides are separated (eluted) on the basis of their relative hydrophobicity. The proteomic workflow for identifying peptides initially eluted by liquid chromatography is termed as LC-MS/MS, with electrospray ionisation (ESI) usually being the preferred MS technique.

1.3.2 Proteomic studies in IBD

A variety of studies have investigated proteomic changes in IBD. These are summarised below.

1.3.2.1 Studies investigating proteomic markers in IBD pathogenesis

In a small study involving 4 patients with active UC, Hseih *et al* investigated proteomic changes in actively inflamed colonic mucosa and compared the changes noted with un-inflamed mucosa of the same patients, patients with infectious colitis as well as healthy controls[86]. 40 spots were differentially expressed between diseased and normal colonic mucosa. Several down-regulated and up-regulated proteins were noted, the former included mitochondrial proteins (heat-shock protein 90, heat-shock protein 60, H1-transporting two-sector ATPase, prohibitin, mitochondrial malate dehydrogenase, thioredoxin peroxidase, voltage-dependent anion-selective channel protein 1, and thiol-specific antioxidant) as well as proteins involved in energy generation (ATP5B, MDH2, triosephosphate isomerase) cellular antioxidants (PRDX1, PRDX2, SELENBP1) and stress proteins (HSPD1, HSPA9B, PRDX1, PRDX2, PHB, VDAC1). Up-regulated proteins included NFAT C1, vimentin and TRA1. Five-protein deregulation was specific for patients with UC (HSPD1, vimentin, TRA1, PHB, and galectin-3). The protein profile in the diseased colonic mucosa compared to un-involved mucosa in UC patients was relatively similar, although a specific protein, PHB, a component of mitochondrial inner membrane was down-expressed in the inactive segment only. On IHC, a difference in subcellular distribution of Galectin-3 was noted between colonocytes in normal mucosa (distributed in nuclei and cytoplasm) and diseased mucosa (in secreted mucus). These changes suggest specific mitochondrial dysfunction in the colonic mucosa in UC patients, changes that tended to occur even in the un-involved mucosa.

Fogt *et al* investigated proteomic changes in the mucosa and submucosa from colectomy specimens of 5 patients with longstanding UC (3 with pancolitis, one each with subtotal colitis and sigmoid colitis) [87]. Using 2D electrophoresis coupled with LC-MS/MS, they investigated the proteomic

profile of the inflamed segments of the colon and identified 7 protein spots with proteins associated with inflammation and tissue repair, viz caldesmon, protocadherin, α-1 anti trypsin, and tetratricopeptide repeat domains in the inflamed tissues. Fibroblastic proliferation in the colitis stroma could account for increased levels of caldesmon. In addition, mutated forms of desmin, verified by western blotting were noted in all 5 cases in the involved mucosa. The significance of these findings is unclear, and may constitute a primary or secondary event in pathogenesis of UC.

Effects of inflammation on proteome in mice have also been investigated using 2D- difference gel electrophoresis (2D-DIGE) coupled with MALDI-TOF MS analysis of colonic mucosa in DSS colitis mice [88]. Several proteins are down regulated in the inflamed mucosa, including vimentin, peroxiredoxin-6, disulphide isomerase A3. Up-regulated proteins included serpin b1a, HMG CoA synthase 2.

Proteomic profile in intestinal epithelial cells (IEC) of Crohn's patients and healthy mucosa in patients with cancer and diverticulitis was investigated by Nanni *et al* [89]. In addition, they investigated colon adenocarcinoma cell lines HT29 Cl.16E with or without treatment with IFN-γ to simulate mucosal inflammation in Crohn's disease. Following fractionation of proteins into four compartments- membrane, cytosolic, nuclear and cytoskeletal, a label-free MS comparative proteomic strategy was used to identify altered protein expression in the cell lines as well as IEC. IFN-γ treated cells demonstrated up-regulated cytosolic protein, adenosylhomocysteinase, a protein involved in regulation of methylation with an important role in in inflammation in epithelial cells[89]. Several other proteins were upregulated in IEC of Crohn's patients including cytosolic protein TRYA1 (associated with altered paracellular permeability in the intestine) and HSP70 (involved in intestinal epithelial protection).

Functional proteomic approaches to identify proteomic changes in cultured colonic epithelial cells after treatment with cytokines have also been studied [90]. Human adenocarcinoma cells DLD-1 before and after treatment with interferon Ƴ, IL-1β and IL-6 were noted to express several proteins including

tryptophanyl-tRNA synthatase, indoleamine 2,3 dioxygenase (IDO), lipocortin 1, JKTBP1 (a nuclear ribonucleoprotein), IFP 35 (interferon induced 35 kDa protein) and proteasome LMP2 were identified. Using colonic epithelial cells isolated from colonic crypts from patients with IBD and diverticulitis, IDO expression was found increased by western blotting in 85% of IBD patients; levels however were undetectable in either diverticulitis or in normal colonic mucosa. This suggests role of specific enzymes, involved in tryptophan metabolism in the pathogenesis of IBD.

In another study of macrophages with serum exosomes isolated from DSS induced mice identified phosphorylation of p38 and ERK and production of TNF α. Subsequent proteomic analysis identified 56 differentially expressed proteins, majority of which were acute-phase proteins and immunoglobulins and involved in macrophage activation [91].

1.3.2.2 Proteomics to identify diagnostic biomarkers in IBD

Serum sample analyses to identify biomarkers of Crohn's disease have been evaluated[92]. As serum contains many abundant proteins (see above), an ultrafiltration technique to enrich LMW proteins up to 10kDa was employed. Despite enrichment, subsequent LC-MS/MS analysis only identified several up-regulated proteins in CD patients. Fibrinopeptide A, a fibrinogen by-product, complement 3 protein and its fragment C3f are increased as is apolipoprotein A-IV, the later often considered as an independent predictor of disease activity in patients with IBD [93].

Serum proteomic profiling to diagnose IBD and discriminate UC from CD as well as other inflammatory states in the bowel has been investigated [94]. Using two types of chip arrays coupled with SELDI-TOF MS followed by a multivariate analysis model they identified potential discriminating acute phase protein biomarkers including PF4, Hpα2, FIBA and MRP8. Although a low sensitivity was noted in discriminating between active UC and CD, they may hold promise when combined with other biomarkers.

Similarly colonic biopsies with proteomic profiling of colonic mucosa and submucosa have been used to investigate effects of inflammation as well as

to distinguish between UC and CD [95]. Using a technique of histology-directed protein profiling, which enables contiguous sections of tissue sample to be analysed by MS and histology M'Koma *et al* identified 2 discrete significant peaks between inflamed UC and CD submucosa (m/z 8773 and 9245) and 3 significant discrete peaks discriminating between un-inflamed UC and CD submucosa (m/z 2778, 9232, and 9519), without any significant difference between inflamed UC and CD mucosa possibly in keeping with the distribution of inflammation in both the conditions. In addition, effect of inflammation in both UC and CD was noted by presence of significant peaks between inflamed and un-inflamed mucosa. Although this study did not delve into protein identification or include healthy controls / non-IBD inflamed colon, nevertheless the results are significant as it identified protein signatures which might enable differentiation of colitis in cases which appear indeterminate on routine histology.

In another proteomic study (using 2D gel electrophoresis and MALDI TOF/TOF tandem MS) circulating protein samples were investigated in interleukin-10 knockout (IL-$10^{-/-}$) mice, collected at day 30, 93 and 135 [96]. This mouse model, where alteration in colonic mucosa are predictable over time enabled a longitudinal assessment of proteomic changes. Fifteen proteins were differentially expressed in sera of mice with mid to late stages of inflammation, compared to non-inflamed IL-$10^{-/-}$ mice including hemopexin, transferrin, haptoglobin, contrapsin and SERPINA 1.

Another pilot proteomic study, in patients with stricturing Crohn's disease identified proteins involved in complement activation, fibrinolytic pathways, and lymphocyte adhesion [97].

1.3.2.3 Marker of inflammation induced colon cancer pathogenesis

Proteins associated with inflammation induced colonic tumourigenesis have been investigated in a dextran sodium sulfate (DSS) colitis model[98]. In rodents, oral DSS administration induces colonic inflammation, which is similar clinically and histologically to human UC. An "inflammation-dysplasia carcinoma sequence" of CRC development is also noted in such DSS colitis

model[98]. 6-weekold C57BL/6J mice were exposed to 15 cycles of DSS; colorectal tumours developed in 91.6% of the mice. 2-DE demonstrated more than 1.5-fold differential expression of 38 protein spots in density in colon tumours compared to normal colon. Using a MALDI-TOF MS, 27 proteins were identified, including GRP 94, HSC 70, enolase, prohibitin, and transgelin. While GRP 94 expression was increased in colon cancer tissue relative to normal colon, there was a 3.4-fold reduction of transgelin expression in colon tumours, which was confirmed by western blotting and immunohistochemistry. Further western blotting revealed loss of expression of transgelin in human colorectal cancer tissue. This loss of transgelin has thus been suggested as a candidate biomarker of colitis-associated colon cancer.

Chronic inflammation associated colorectal carcinogenesis has also been studied using an azoxymethane(AOM)/dextran sodium sulphate(DSS) mouse model[99]. AOM, a colonic carcinogen in combination to DSS has previously been shown to alter colonic mucosal gene expression[100]. Colonic tumours and non-tumour tissue were analysed by 2DE and MALDI-TOF MS identified 21 differentially expressed proteins in the cancerous tissue with more than 3 fold increase or decrease in the density in the cancerous tissue. Proteins showing increased expression were beta-tropomyosin, tropomyosin 1 alpha isoform b, S100 calcium binding protein A9, peptidylprolylisomerase A and an unknown protein. 16 proteins were shown to have reduced expression, including Car 1 protein, selenium binding protein 1, HMG CoA synthase, thioredoxin I, I Cys peroxiredoxin protein 2, Fcgbp protein, cytochrome c oxidase, subunit Va, ETHEI protein etc. These are proteins associated with metabolic functions, antioxidant systems and oxidative stress pathways as well as inflammation.

In contrast, in another study using UC associated cancer cell lines (UCCA-3, UCCA-21 and UCCA-24), a significantly higher expression of heat-shock protein of 47kDa (HSP47) was noted, in comparison to sporadic colorectal cancer cell lines [101]. HSP47 is an endoplasmic reticulum associated molecular chaperone and closely related to collagen in fibroblasts.

Using a label-free proteomics approach, May *et al* [102] identified potential biomarkers in non-dysplastic and normal-appearing rectal tissue from UC patients who have dysplasia or cancer and compared them with UC patients who were dysplasia or cancer-free. An LTQ-Orbitrap hybrid mass spectrometer coupled with nanoflow HPLC was used in this study. Of the differentially expressed proteins mitochondrial proteins, cytoskeletal proteins, RAS superfamily, and apoptosis related proteins were important protein clusters expressed in the non-dysplastic and dysplastic tissues of UC progressors. Mitochondrial dysregulation has thus been suggested as having a role in pathogenesis of UC associated cancer. In addition in view of role of cytoskeletal proteins in regulation of contact inhibition, apoptosis and anchorage-independent cell growth [103], alteration in cytoskeletal proteins actin may contribute to tumour progression. TRAP1 (TNF receptor associated protein 1) a mitochondrial heat shock protein was increased by 2.41 fold in rectal biopsies of those with dysplasia, and in dysplastic tissue (1.74 fold). CPS1 (carbamoyl phosphate synthase 1) was not noted by MS/MS in any of the non-dysplastic patients. In this study, IHC using antibodies to TRAP1 and CPS1 was undertaken. Whereas only one of five non-dysplastic samples had moderate staining, all the progressors demonstrated moderate to strong staining. CPS1 showed a statistically significant difference in IHC staining between the nonprogressor and progressor groups. ROC analysis suggested rectal CPS1 staining could be used to predict dysplasia or cancer in the colon with 87% sensitivity and 45% specificity. These results are thus significant as they could herald potential minimal invasive tests to identify high-risk individuals with dysplasia/colon cancer elsewhere in the colon.

Similarly iTRAQ based quantitative proteomic strategies too have been used to investigate pathogenesis and progression of colitis associated cancers [104]. In this study Brentnall *et al* compared proteomic changes in dysplastic and non-dysplastic mucosa from UC progressors (patients with UC patients having either HGD or cancer) to the colonic mucosa of UC non-progressors (patients with UC who are dysplasia-free) and healthy non-UC colon Using a 2.0-fold change in protein expression as cutoff value, they noted 60 differentially expressed proteins in UC non-progressors compared with

healthy colon, with 45 down-regulated and 15 up-regulated proteins. In UC progressors (compared with normal non-UC colon) the differentially expressed proteins were more than 90. Over expressed proteins in the colon increased with more aggressive disease. These included a variety of proteins including CPS1 (carbamoyl phosphate synthase 1), S100P, mitochondrial proteins as well as cytoskeletal proteins (including keratin) the later under-expressed in UC samples compared to healthy colon. S100 proteins (including S100P, S100A6, S100A11, and S100H), a family of calcium-binding proteins were differentially expressed in progressive disease, with highest expression noted in HGD. This study is significant as this for the first time demonstrated abnormality in the proteome of the non-dysplastic mucosa in patients with HGD/cancer. This suggests that genomic field defect noted previously [70, 71] in the rest of the colon also possibly extents to the proteome. In addition, they could be a pointer towards potential neoplastic changes before they have fully developed in the epithelium, thus potentially being useful as a biomarker to identify high-risk individuals.

1.3.2.4 Marker of disease activity and response to treatment

Kanmura *et al* investigated patients with IBD to identify any serological biomarkers of either disease activity or response to treatment [105]. The study included 48 patients with UC, and 22 with CD and compared them to healthy controls and CRC patients. Using a SELDI-TOF/MS Protein Chip system, 3 proteins were identified as human neutrophil peptides 1–3 (HNP1-3). Enzyme-linked immunosorbent assays demonstrated that the mean plasma concentration of HNP 1–3 was significantly higher in patients with active UC than in patients whose UC was in remission or patients with CD, infectious colitis, or healthy subjects, and tended to be higher than in patients with colon cancer. There was a significant correlation between the HNP 1–3 levels and the UCDAI score or the white blood cell count (WBC) of UC patients ($r = 0.54$, $P < 0.01$; $r = 0.55$, $P < 0.01$, respectively). In addition, the levels of HNP1-3 were also assayed in patients to identify any change with treatment with corticosteroids. In patients who failed to respond to therapy, baseline HNP 1–3 levels were lower than those in the responder group and were not changed

after treatment whereas the elevated HNP 1–3 levels in the responder group were reduced after corticosteroid therapy. Using a cutoff value for HNP 1-3 of 100 ng/mL, a sensitivity of 89% and a specificity of 80% when HNP 1-3 was used as a tool to discriminate responder UC patients from nonresponders.

IHC with anti-HNP1-3 of colonic tissues showed strong staining of the colonic mucosa, lamina propria, muscle layer, and crypt abscesses in patients with active UC; the likely source of the HNP1-3 being the neutrophilic infiltrate. These results suggest a possible role of HNP 1–3 as a novel biomarker in diagnosing patients with active UC and for predicting response to treatment.

In a pilot study using hydrophobic H4 protein chip arrays combined with SELDI-TOF mass spectrometer, Meuwis *et al* investigated response to anti TNF therapy in Crohn's disease [106]. Serum samples of ten patients with active disease showing clear clinical response to Infliximab were compared before and after treatment (week 4 for non-fistulising and week 10 for fistulising patients) to non-responders. Although SELDI spectra identified Platelet aggregation factor 4 (PF4) as a potential biomarker on multivariate analysis as a predictor of response to treatment, subsequent validation tests failed to reveal any significant difference with ELISA for PF4, sCD40L (another marker of platelet activation) or IL-6. This could be due to small numbers investigated or lack of homogeneity of cases.

Gazouli *et al* aimed at measuring proteomic changes in serum in a small cohort of CD patients aiming at identifying a panel of candidate protein biomarkers of CD that might predict response to IFX therapy [107]. Serum samples were subjected to 2DGE followed by MALDI- TOF-MS. Several proteins were noted to be up regulated in primary non-responders and responder groups - apolipoprotein A-I (APOA1), apolipoprotein E (APOE), complement C4-B (CO4B), plasminogen (PLMN), serotransferrin (TRFE), beta-2-glycoprotein 1 (APOH), and clusterin (CLUS). These protein levels were however unchanged in the remission group compared to baseline. In this group up regulated proteins included leucine-rich alpha-2-glycoprotein (A2GL), vitamin D-binding protein (VTDB), alpha-1B-glycoprotein (A1BG) and complement C1r subcomponent (C1R) whereas proteins downregulated in

this group included complement C3 (CO3), transthyretin (TTHY) and fibrinogen alpha chain (FIBA).

Using label free quantitative MS, enriched low mass plasma proteome (<25 kDa) was assessed to identify markers of inflammation [108]. Several peptides were noted differentiating controls from IBD (secreted phosphoprotein 24, SPP 24); those in remission and healthy can be differentiated in UC by SPP24, α-1-microglobulin (AMBP) and CD by SPP24. UC and CD can also be differentiated by Guanylin, and Secretogranin-1. Active and quiescent disease can also be differentiated in UC and CD by CHGB, SPP24and AMBP.

A recent study in the interleukin-10 knockout [IL-10(-/-)] mouse investigating circulating protein biomarkers, identified 15 different proteins differentially expressed in serum samples from mid- to late-stage IL-10(-/-) mice compared to early non-inflamed IL-10(-/-) mice. This suggests a role for protein profiling in assessing severity of the disease [96].

1.3.2.5 Conclusion

Several cell line, murine and human proteomic studies have been undertaken in the recent years investigating various aspects of IBD including pathogenesis of IBD (as well as colitis associated cancers) and identifying biomarkers for diagnosing IBD (and differentiating UC from CD). These studies have identified proteins, the potential roles of which have been discussed previously. Despite significant development of the technology in proteomics and potential significant role, the practical clinical application yet remains in its infancy. IBD is a disease characterised by marked phenotypic variability. This is possibly the prime factor as it affects selection of cases without controlling for variables like disease severity, duration of disease, concomitant drugs and co-morbidities. In addition, the recent studies mentioned previously have used diverse methodologies, often with small patient numbers and lack of clear-cut patient characterisation to address phenotypic variability. Hence despite the promise of being a useful technology of identifying various biomarkers in IBD, it remains a difficult and expensive

technology, whose role is yet not fully defined in understanding the pathogenesis of IBD or its complications. Although a variety of proteins have been demonstrated as up or downregulated in various conditions, their significance is unclear. This could be potentially overcome by investigating functional relationship of differentially expressed proteins using protein network programs like STRING (http://string-db.org). Nevertheless, some of the studies do identify certain proteins the role of which would need further investigation. These include mitochondrial proteins, proteins involved in tissue repair and energy generation as well as a type 3 intermediate filament proteins, vimentin which has been noted to be down-regulated in inflamed colonic tissue.

1.4 Keratins and intermediate filaments

The cellular cytoskeleton, a scaffolding within the cytoplasm of animal cells is composed of various proteins viz. intermediate filaments (IF), microfilaments and microtubules [109, 110]. Intermediate filaments themselves are subdivided into 5 major types based on sequence of amino acid and protein structure. Types I–IV principally constitute cytoplasmic IFs, whereas type V IF proteins, the lamins are part of the nucleoskeleton [111, 112]. Keratins form the largest subgroup of IF proteins. They are composed of polymerised dimers of a Type I (acidic) and Type II (basic) keratin [113]. Keratins are expressed in different epithelial cells [114]; the principal keratins in the intestinal epithelia being K8, K18 and K19 [115].

1.4.1 Isolation and solubilisation of keratins

Keratins differ from other cytoskeletal filament proteins (such as actin, myosin, and tubulin) in that they are resistant to non-denaturing detergent buffers and high concentrations of ions, including chaotropic salts such as potassium chloride or iodide [116-118]. This is extremely important while attempting to study keratin expression since they exist in a soluble form as well as an insoluble filamentous state. Usually 5% of the keratin pool exists in a soluble state [119], although the state of solubility is determined by various factors including cellular stress [120] and the presence of post-translational modifications (described below). The insoluble fraction can be solubilised with high concentrations of urea (9-10 M), or guanidine hydrochloride (4-6 M) to produce soluble tetrameric or oligomeric subunits to result in solubilisation of the denatured monomers [121, 122].

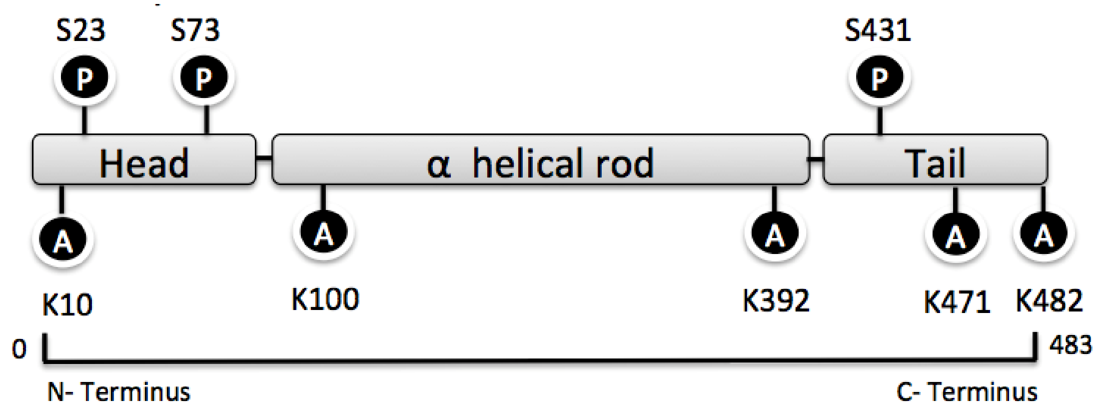
1.4.2 Post-translational modifications of keratin

Following translation, most proteins undergo further changes through post-translational modifications [123]. Such modifications modulate the structure and function of proteins in a manner which is not directly coded for by genes [123]. Keratins are highly dynamic and undergo alterations due to a variety of stimuli like cellular stresses as well as during apoptosis and mitosis [124] [125]. Important PTMs of keratins include serine phosphorylation [126-128] threonine phosphorylation [129], tyrosine phosphorylation [130, 131], lysine

acetylation [132], serine/threonine glycosylation [133], cleavage [134, 135] and sumoylation [136].

Phosphorylation is the most studied of the PTMs. Common PTM sites of K8 molecule are summarised in Figure 1.4.

Figure 1.4. *Keratin 8 molecule with phosphorylation(P) and acetylation sites(A) [137]. Particularly significant are the three serine phosphorylation sites at Ser23, Ser 73 and Ser 431.*



Phosphorylation of K8 occurs primarily at serine residues in the following sites - Ser 431, Ser 73, Ser 23 on K8 and Ser 33, Ser 52 on K18 [138]. Serine phosphorylation is believed to the critical modulator of filament formation and depolymerisation in response to cellular stresses [139-143]. Different protein kinases phosphorylate keratins. The K8 residue, Ser 73 is a substrate for phosphorylation by p38 kinase and c-Jun N-terminal kinase (both of which are involved in cellular responses to stresses, cytokines and apoptosis) [144-146] and this has likewise been shown to regulate filament organisation [141, 147]. Serine residue (Ser 431) on the other hand is phosphorylated by the kinase ERK1 (MAP kinase) mediated by stimulation by epidermal growth factor [127]. These changes not only affect the solubility of keratins, but hyper phosphorylation may protect K18 from degradation by caspases [148]. Hyperphosphorylation of K8 has been noted to act as a “phosphate sponge” for stress kinases resulting in protection from apoptosis [149].

Dephosphorylation of K8 too in CRC tumour cells has been associated with tumour progression [139]. A recent study has also identified phosphorylation of highly conserved tyrosine residues Tyr-267 in the K8 “rod” domain playing too having a role on IF assembly and solubility [131].

Acetylation is another important PTM [150]. Acetylation involves balance of lysine acetyl transferase and histone deacetylases (HDAC) [151]. Acetylation regulates protein-protein interaction [152] and protein stabilisation [153]. It has been hypothesised that the acetylation noted may have a role in stabilisation of intermediate filaments, thus controlling the loss of epithelial cells from the mucosa [132]. Keratin 8 acetylation(at highly conserved lysine-207), by modulating keratin phosphorylation has also been noted to modulate filament organisation and solubility [154].

1.4.3 Function of keratin in normal epithelium

The function of keratins varies according to the organ. In liver K8/K18 confers protection from mechanical and non-mechanical stresses [155]. This also include protection from apoptosis [156]. IFs may have a role in cell-death signalling pathways, in particular apoptosis mediated by tumour necrosis factor-alpha (TNF-α) and Fas [157]. Epithelial cells lacking K8 and K18 are significantly more sensitive to TNF-mediated apoptosis [158].

In the intestine, the role of keratins is poorly understood, although the K8-null mouse phenotype suggests they may have a potential role in cell growth, differentiation, or targeting of proteins to the apical compartment of the colonic epithelial cells [159, 160]. K8-null ($K8^{-/-}$ and $K8^{+/-}$) mice in comparison to $K8^{+/+}$ mice have abnormal colonic active ion transport (Na^+ and Cl^-). These findings suggest a role of colonic keratins in regulating electrolyte transport [161].In another study, mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase 2 (HMGCS2), an enzyme which catalyses the rate-limiting step of ketogenesis, has been noted to be downregulated in $K8^{-/-}$ mice colonocytes suggesting a role in metabolic pathways in the colonocytes [162].

1.4.4 Keratins and inflammatory bowel disease

Studies suggest an association of keratins with IBD. Tumour necrosis factor alpha (TNF-α) a pro-inflammatory cytokine, is involved in the pathogenesis of IBD [163-165]. K8 and K18 co-localises with cytoplasmic domain of TNF receptor 2 (TNFR2) and thus modulates intracellular signalling and NFκB activation [166]. Normal and malignant epithelial cells lacking in K8 and K18 are sensitive to TNF mediated cell death [167]. Negative regulation of toll like receptor (TLR)/NF-κB signaling to limit inflammatory response by by K8 has been shown [168]. Down-regulation of K8 in mice enhanced the TLR-mediated responses, associated with increased levels of inflammatory cytokines and tissue damage. K8 suppresses TLR-induced nuclear factor (NF)-κB activation and interacts with the TNF receptor-associated factor 6 (TRAF6) to prevent its polyubiquitination.

Effects of interleukin 6 (IL6) an immunoregulatory cytokine on keratin expression and function and their influence on intestinal mucosal barrier function have been investigated [169]. Intestinal disorders including IBD are often linked to alteration in barrier function of the epithelial cells. Such loss of barrier function results in sensitisation of the immune system to antigens previously localised to the intestinal lumen [170, 171]. Wang *et al* demonstrated that IL-6 significantly induces expression of insoluble cytoskeletal fraction of K8 and K18, and also serine phosphorylation of K8 at phosphoserine 431 and 73, suggesting PTM of K8 in response to IL-6 [172]. IL6 null mice also demonstrated significantly increased intestinal permeability with dextran sodium sulfate administration. These findings demonstrate regulation of colonic expression of K8 and K18 by IL-6 as well as the role of K8 in mediating effects of IL-6 on the barrier function.

In an animal study [173] homozygous mK8$^{-/-}$ FVB/N mice developed colonic hyperplasia, colitis and rectal prolapse. Histology showed hyperplasia of the crypt cells, suggesting the inflammatory component of mK8^{-} bowel disease could be secondary to colorectal hyperplasia. The colonic inflammation in K8

null mice suggests a model of IBD which could be a primary epithelial disorder rather than an immune cell defect [174].

Keratin 8 deficient mice ($K8^{-/-}$) also develop chronic T- helper type 2 (Th2 colitis) [175]. A significant increase in TCRβ (T cell receptor beta) positive $CD4^{+}$ T cells in colon lamina propria have been noted in $K8^{-/-}$ mice. Colonic epithelial cells in $K8^{-/-}$ mice expressed MHC class II antigens; such inflammation was prevented with antibiotics suggesting a role of luminal bacteria in triggering colitis in mice with a primary epithelial defect [176].

Absence of keratin filaments in the colon is associated with abnormalities in ion transport, protein mis-targeting and diarrhoea prior to occurrence of hyperproliferation and inflammation [161]. $K8^{-/-}$ mice develop diarrhoea with significantly higher stool water content compared with $K8^{+/+}$ mice. Colonic tissue conductance, a measure tight junction activity, is not significantly different suggesting the diarrhoea is unrelated to an alteration of the paracellular transport or increased tight junction permeability. Analysis of protein and mRNA levels of candidate ion transporters in the colonic mucosa of K8 knockout mice (K8-/-) shows decreased/absent levels in the colonic mucosa compared to wild type (K8+/+) mice. This suggests a role of K8 on chloride transport mediated via action on DRA and contributing to the diarrhoeal phenotype noted in K8 knockout mice [177].

Mistargeting in $K8^{-/-}$ colons which normalises with antibiotic treatment suggests that luminal bacteria and/or their consequent inflammatory response may promote the observed protein mis-targeting [176]. The mis-targeted ion transporters may be responsible in creating an environment for pathogenic bacteria to thrive, which in turn stimulate the colitis and maintenance of the mis-localised transport proteins. These studies support the role of luminal bacteria in pathogenesis of colonic inflammation and colitis. Although the aetiology of IBD is multifactorial, luminal microflora particularly an altered host tolerance to them is believed to play an important role in its pathogenesis [178].

In another study, heterozygote keratin knockout mice ($K8^{+/-}$), show 50% less K8 expression compared to wild type mice($K8^{+/+}$) [179]; K7 and K18 expression too are reduced significantly. Although there is crypt epithelial hyper proliferation in the $K8^{+/-}$ phenotype, the mucosa does not show any spontaneous inflammation, in contrast to $K^{-/-}$ mice. These mice however are more sensitive to DSS induced colitis, suggesting a role of K8 in protection against stress.

A study investigated genomic DNA from patients with UC and Crohn's disease, 50 with sporadic and 47 with familial IBD [180]. Heterozygous mis-sense mutations in K8 were identified in five (5.2%) unrelated IBD patients in this study. Three separate KRT8 mutations (G62C) were identified; two were in unrelated UC patients with a family history of IBD and one was in a Crohn's disease patient with no family history. A sporadic Crohn's disease patient was found to have mutation in K8 (I63V); whereas another mutation resulting in amino acid substitution in the tail domain of K8 (K464N) was found in an UC patient with family history of IBD. Four patients were noted to have heterozygous sequence variation in KRT18 resulting to K18 (S230T) in the L12 linker region. All these changes were noted to lie within the non-helical domains, which are usually associated with milder forms of disease. The K8 (K464N) mutation was not found in 194 controls.

Another study investigated 217 patients with Crohn's disease, 131 patients with UC and 560 healthy controls [181]. Heterozygous G62C mutation in K8 was detected in five (2.3%) and three (2.3%) patients with Crohn's disease and UC respectively, whereas 9 (1.6%) control subjects were noted to carry such mutation. None of the patients or controls was homozygous for this mutation nor did they show the Y54H mutation. The IBD patients demonstrating the heterozygous mutation did not show any phenotypic difference to those without such mutation.

A further study investigated the association of mutations in KRT8 and KRT19 in patients with IBD [182]. 16 of the 184 patients with IBD in the study

demonstrated heterozygous variants in KRT8 (G62C, R341H); 4 of 70 unaffected volunteers too carried KRT8 (R341H) mutation. KRT8 variants (G62C, R341H and R341C) when tested in 682 independent nuclear families (with both parents and at least one IBD affected offspring) and 273 controls failed to show any significant departure from random transmission. This study thus failed to provide a conclusive evidence of keratin mutations in the pathogenesis of IBD.

Differential and dynamic regulation of keratins and post translational changes in K8 in the colonic mucosa have been noted in DSS model of colitis [183]. In this study, in response to inflammation, down regulation of total K8 was noted, while an increase in stress responsive K8 serine-74 phosphorylation (K8 pS74) was noted. Following treatment with antibiotics, to eliminate colonic microflora, up regulation of K8/K18 levels along with K8 pS74 is seen.

1.4.5 Keratins in colitis associated colorectal cancer

K8 null (-/-), compared to wild type (K8+/+) or heterozygote(K8+/-) mice do demonstrate increased CRC risk. However, a dramatic increase in distal colonic tumours is noted in azoxymethane (AOM) mice suggesting a role of keratins in inflammation induced CRC. Interleukin 22 (IL22) pathway which plays a key role inflammation and proliferation is activated.in is noted to be activated. Pro-caspase 1 a component of the inflammasome in the colon is co-immuneprecipitated with K8/K18 suggesting a role of keratins in modulating inflammasome activity and protecting against inflammation and inflammation induced carcinogenesis [184].

Expression of K7 and K20 in dysplastic and neoplastic mucosa in UC has been investigated [185, 186]. Immunohistochemical examination of tissue sections in have demonstrated, in contrast to sporadic adenocarcinomas, vast majority of UC associated neoplasms express K7, including patients with low and high grade dysplasia [185]. In another study from northern Sweden in UC patients undergoing colonoscopic surveillance, K7 and K20 expression by IHC were noted in colonic biopsies. K20 expression was increased in lower aspect of colonic crypts in neoplasia-associated lesions; while 2 out of 5 patients with colorectal cancer, 3 out of 6 patients with HGD and 7 out of 10

patients with LGD were positive for K7 [186]. These changes may represent keratins as possible markers of neoplastic changes in patients with UC.

1.5 Hypothesis of the study

The hypothesis of the study is that in patients with ulcerative colitis, changes in keratin expression and post translational modifications in the colonic mucosa are modulated by the severity of the inflammation and longer duration of disease, and thus predispose to the pathogenesis of cancer and dysplasia.

CHAPTER 2

PATIENTS AND METHODS

2.1 Patient recruitment

Patients with UC (undergoing elective or urgent colonoscopic / flexible sigmoidoscopic evaluation) and healthy controls were recruited prospectively from the dedicated inflammatory bowel disease clinic, general gastroenterology outpatient clinics and from in-patient wards at the Royal Hallamshire Hospital, Sheffield, UK. Suitable patients were identified using the inclusion and exclusion criteria specified below. The identified patients and their primary care physicians were invited to participate in the study, and informed consent was obtained.

2.1.1 Inclusion criteria and exclusion criteria.

2.1.1.1 Inclusion criteria

1. Age 18-75 years
2. Patients with suspected or histologically proven UC at clinical presentation or
3. Control patients: patients without pre-existing colonic disease, undergoing colonoscopic examination for non-specific gastrointestinal symptoms. This group of patients had normal colonoscopy including normal colonic biopsies.

2.1.1.2 Exclusion criteria

1. Patients unable to give signed informed consent
2. Non-correctable coagulopathy (prothrombin time > 12 seconds / platelet count < 90 x 10^9 / L)
3. Pregnant patient
4. Inadequate bowel preparation
5. Severe colitis or toxic dilatation of the colon or clinical condition where colectomy is highly likely
6. Patients with a diagnosis of indeterminate colitis, Crohns disease

7. Frail individuals unable to tolerate a lower GI endoscopy
8. Previous surgery for colitis

2.1.2 Patient groups

The recruited patients and biopsy samples were categorised into the following study groups:

1. Longstanding pancolitis (LSPC): patients in remission (clinically and endoscopically) undergoing surveillance. These patients were categorised according to the duration of disease since diagnosis.
2. Quiescent recent onset colitis (ROUC) – patients with UC with disease duration less than 5 years since diagnosis.
3. Patients with UC and co-existent PSC (PSC)
4. Patients with active distal (proctitis / proctosigmoiditis) or left sided colitis (ACT)
5. Proximal colonic biopsies from the endoscopically (and histologically) normal proximal colonic mucosa in patients with distal active distal disease (INACT)
6. Patients with long standing disease who were noted to have dysplastic / neoplastic changes in the colon (DT and DR)
7. Healthy controls (CON)

The patients with long standing pancolitis had pan-colonic disease, with disease duration between 20-40 years. They were in remission, clinically, endoscopically and on histology. The ROUC group included patients with recently diagnosed colitis (within preceding 5 years of the endoscopic assessment), who were also in remission.

We also included a group of patients with histologically proven dysplastic lesions at colonoscopy. Biopsies were obtained from the dysplastic areas and categorised as DT (dysplasia tumour group); additional biopsies were taken from the uninvolved rectal mucosa in the same group of patients (to identify

any proteomic field changes) and categorised as DR (dysplasia rectum group).

2.1.3 Initial assessment

Recruited patients were assessed to obtain data about the following parameters

1. Duration of disease since diagnosis
2. Activity of disease (Mayo scoring system for assessment of ulcerative colitis activity (Disease activity index (DAI)/Mayo score)) [187]
3. Endoscopic scoring of severity of the disease at time of colonoscopy (Baron's score) [188]
4. Endoscopic extent of the disease in the colon
5. Drug history particularly use of 5 ASA (dosage & duration), or UDCA in patients with PSC
6. Family history of colorectal cancers

2.1.4. **Colonoscopy technique**

Routine colonoscopy was undertaken as per the British Society of Gastroenterology (BSG) guidelines [2]. After ileo-caecal intubation detailed inspection of the colonic mucosa was performed. Any visible abnormal lesions were resected using standard techniques, otherwise they were biopsied if not considered resectable. Field biopsies were taken from the immediate vicinity of dysplastic lesions to exclude any dysplastic changes in the adjacent mucosa. Details of protocol for obtaining colonic biopsies is mentioned below (section 3.6)

2.1.5. **Flexible Sigmoidoscopy technique**

For patients with active colitis, it is not always clinically appropriate to undertake total ileo-colonoscopy. These patients underwent limited flexible

sigmoidoscopy, with study biopsies, as outlined below obtained from the inflamed distal colon (sigmoid and rectum), and where appropriate, from the proximal non-inflamed segment (in patients with distal / left sided colitis).

2.1.6 Biopsy / histopathology protocol

2.1.6.1. Patients undergoing colonoscopic surveillance

At extubation, as per standard practice, 2 biopsies per 10 cm colorectal segment were taken according to British Society of Gastroenterology UC surveillance guidelines [2]– corresponding to:

- caecum
- proximal ascending colon
- distal ascending colon
- proximal transverse colon
- distal transverse colon
- proximal descending colon
- distal descending colon
- sigmoid colon
- rectum

Each set of biopsies was processed in individual formalin pots with the distance from anus and anatomical colorectal segment clearly documented. Study biopsies (for immunohistochemistry and proteomic studies) were taken in addition from rectum and mid-sigmoid colon with additional biopsies from the dysplastic regions.

2.1.6.2. Patients with active disease

Biopsies were taken according to clinical indications from inflamed and non-inflamed regions (with at least one set of study biopsies from the rectum and mid-sigmoid colon).

2.1.6.3. Healthy controls

A group of healthy controls were included in the study. These patients without pre-existing colonic disease, underwent colonoscopic examination for non-specific gastrointestinal symptoms. Colonic biopsies in these patients were normal. Apart from routine biopsies (as clinically indicated) additional matched study biopsies were obtained from the rectum and mid-sigmoid.

Five biopsies were obtained from each site in the colon designated for the study. Two biopsies from each site were fixed in formalin (for immunohistochemical studies), the rest were snap frozen in liquid nitrogen and transferred to –80°C freezer (for proteomic assays).

2.1.7 Categorising patients into groups for proteomic studies

The selected patients were categorised into the following groups.

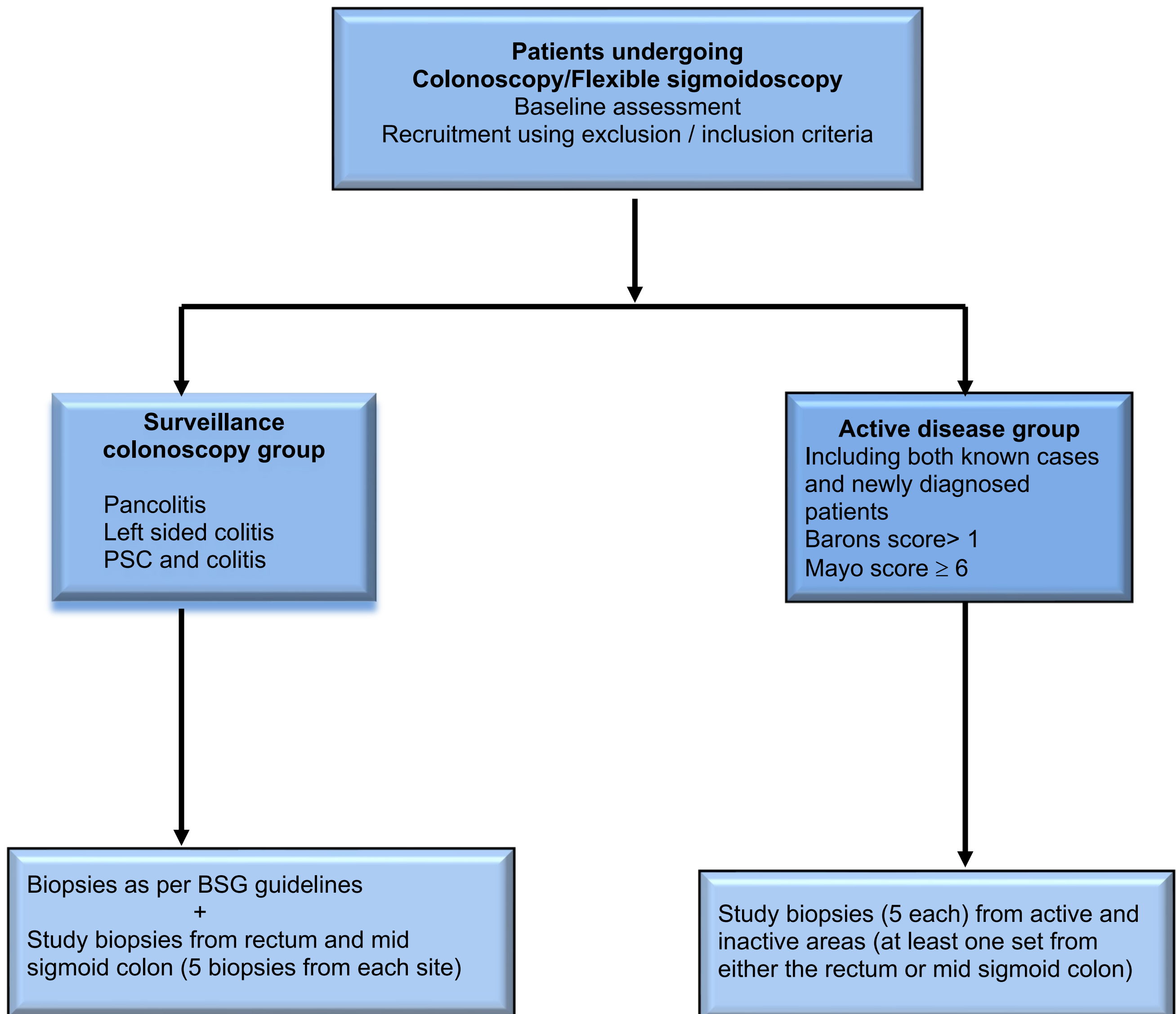
1. Long-standing (20-40years) pancolitis (LSPC) in remission.
2. Recent onset (<5 years) UC (ROUC) with quiescent disease.
3. UC with primary sclerosing cholangitis (PSC).
4. Pancolitis with dysplasia with biopsies taken from the rectum (DR).
5. Biopsies obtained from dysplastic/neoplastic lesions (DT) in patients with UC.
6. Distal active disease (ACT).
7. Inactive proximal colonic segment in those with active distal colitis (INACT).
8. Healthy controls (CON).

2.1.8 Pathologic evaluation

After the biopsies were routinely reported within the standard NHS environment, a designated specialist gastrointestinal pathologist blinded to the endoscopic diagnosis examined the biopsy specimens and undertook a histological assessment of inflammatory activity. Dysplasia was defined as per Riddell *et al* [20]. Histological assessment of inflammation was done as per standard criteria (Mayo scoring system for assessment of ulcerative colitis activity (Disease activity index (DAI) / Mayo score)) [48].

2.1.9 Summary of investigation in the study

Table 2.1 *Flowchart summarising pathways for investigating in ulcerative colitis patients in the study*



2.2 Reagents and chemicals

2.2.1 Antibodies for western immunoblotting

2.2.1.1 Primary antibodies

Antibody	Supplier	Dilution	Size of protein (kDa)	Species of origin/ antibody type
Keratin 8 [M20] (ab9023)	Abcam, Cambridge, UK	1:1000	50	Mouse monoclonal
Keratin 18 [C-04] (ab668),	Abcam, Cambridge, UK	1:1000	45	Mouse monoclonal
Keratin 19 [A53-B/A2] (ab7754);	Abcam, Cambridge, UK	1:1000	40	Mouse monoclonal
Keratin 8 -phospho S73 [E431-2] (ab32579)	Abcam, Cambridge, UK	1:10000	55	Rabbit monoclonal
Keratin 8 -phospho S431 (ab59434)	Abcam, Cambridge, UK	1:1000	54	Rabbit polyclonal
Keratin 8 -phospho S23 [EP1629Y] (ab76584)	Abcam, Cambridge, UK	1:20000	54	Rabbit monoclonal
Anti-Phosphoserine/threonine antibody (ab17464)	Abcam, Cambridge, UK	1:1000	NA	Rabbit polyclonal
Vimentin V9 clone (MAB3400)	Millipore, UK	1:1000	57-60	Mouse monoclonal
Keratin 8, Acetylated lysine10 (K10)	In-house antibodies[189]	1:1000	54	Rabbit monoclonal
Keratin 8, Acetylated lysine 482 (K482)	In-house antibodies[189]	1:1000	54	Rabbit monoclonal

2.2.1.2 Secondary antibodies

Name	Supplier	Reference/ Catalogue No	Dilution	Species of origin
Polyclonal conjugated rabbit anti mouse / HRP	Dako, Glostrop, Denmark	P026002-2	1:2000	Rabbit
Polyclonal conjugated goat anti rabbit / HRP	Dako, Glostrop, Denmark	P044801-2	1:2000	Goat

2.2.2 Reagents for protein assay

Name	Reference/Catalogue No	Supplier
Protein standard 2mg BSA/ml	P0834	Sigma Aldrich, Gillingham, UK
Bio-Rad protein assay	500-0006	Bio-Rad Laboratoties GMBH, Munchen, Germany

2.2.3 Reagents for cell culture

Name	Strength	Supplier
RPMI 1640 media	85%(v/v)	GIBCO, Invitrogen Paisley, UK
Heat-deactivated foetal calf serum	10%(v/v)	Biosera, Sussex, UK
Penicillin	5% (v/v) (10000 U/mL)	Lonza, Verviers, Belgium
Streptomycin	10000 µg/mL	Lonza, Verviers, Belgium

2.2.4 Materials and reagents for western immunoblotting and gel analysis

Name	Reference/ Catalogue No	Supplier
Albumin from Bovine serum	A7906	Sigma Aldrich, Gillingham, UK
InstantBlue™	ISB1L	Expedeon, Harston, UK
Immobilon™ Western Chemiluminescent HRP substrate	WBKLS0500	Millipore Corporation, Billerica, USA
Immobilon® -P Transfer membrane	IPVH00010	Millipore Corporation, Billerica, USA
ProtoGel 30%(w/v) Acrylamide , 0.8%(w/v) Bis-Acrylamide	EC-890	National Diagnostics
Sodium Dodecyl Sulfate 20% solution	BP1311	Fisher Scientific, NJ, USA
Pierce® SDS-PAGE sample prep kit	89888	Thermo Scientific, Rockford, USA
Ammonium persulfate (APS)	A3678	Sigma Aldrich, Gillingham, UK
Tetramethylethylenediamine (TEMED)	T9281	Sigma Aldrich, Gillingham, UK
All Blue, Precision Plus Protein™ Standards	161-0373	Bio-Rad Laboratories Ltd, Hemel Hempstead, UK
Restore Western Blot Stripping buffer	21059	Thermo Scientific, Rockford, USA

2.2.5 Reagents for mass spectrometry

Name	Unit details	Reference/Catalogue No	Supplier
Water: HPLC Gradient grade	2.5 L	W/0106/17	Fisher Scientific, NJ, USA
Sequencing grade modified Trypsin	20µG/ vial	V5111	Promega Corporation, Madison, USA
Methyl methanethiosulfonate (MMTS) in isopropanol	200mM	23011	Thermo Scientific, Rockford, USA
Tris-(2carboxyethyl) phosphine (TCEP)	50mM	20490	Thermo Scientific, Rockford, USA
Isopropanol	99.9% v/v	53079	VWR, Lutterworth, UK
Trifluroacetic acid (TFA)	NA	T6508	Sigma Aldrich, Gillingham, UK
Acetonitrile	NA	271004	Sigma Aldrich, Gillingham, UK
iTRAQ reagents- 8 plex 113-119, 121	NA	4390811	AB Sciex Pte. Ltd., Foster City, USA

2.2.6 Buffer solutions

2.2.6.1 Phosphate buffer saline (pH 7.4)

Name	Concentration	Supplier
Phosphate buffer saline (pH 7.4)		Sigma Aldrich, Gillingham, UK
NaCl	137mmol/L	
KCl	2.7mmol/L	
Na_2HPO_4	10mmol/L	
KH_2PO_4	1.8mmol/L	

2.2.6.2 Running buffer

Name	Concentration	Supplier
Tris	25mM	Sigma Aldrich, Gillingham, UK
Glycine	192mM	Sigma Aldrich, Gillingham, UK
SDS	0.1%(w/v)	Sigma Aldrich, Gillingham, UK

2.2.6.3 Wash buffer

Name	Reference/ Catalogue No	Final concentration	Supplier
Phosphate-buffered saline	P4417	1 x	Sigma- Aldrich, Gillingham, UK
NaCl		137mmol/L	
KCl		2.7mmol/L	
Na_2HPO_4		10mmol/L	
KH_2PO_4		1.8mmol/L	
Magnesium Chloride	M8266	2 mM	Sigma- Aldrich, Gillingham, UK
Pefabloc®	11 429 868 001	0.5 mM	Roche Diagnostics GmbH, Mannheim, Germany

2.2.6.4 Low Detergent Buffer (LDB)

Name	Reference/ Catalogue No	Final concentration	Supplier
Phosphate-buffered saline	P4417	0.5 x	Sigma- Aldrich, Gillingham, UK
NaCl		137mmol/L	
KCl		2.7mmol/L	
Na_2HPO_4		10mmol/L	
KH_2PO_4		1.8mmol/L	
4-Morpholinepropanesulfonic acid (MOPS) (pH 7)	M1254	50 mM	Sigma- Aldrich, Gillingham, UK
Magnesium Chloride	M8266	10 mM	Sigma- Aldrich, Gillingham, UK
Pefabloc®	11 429 868 001	2 mM	Roche Diagnostics GmbH, Mannheim, Germany
Ethylene glycol tetra acetic acid (EGTA)	03779	1 mM	Sigma- Aldrich, Gillingham, UK
Triton X-100	X100	0.15% v/v	Sigma- Aldrich, Gillingham, UK
Phosphatase inhibitor cocktail 2	P5726	1% (v/v)	Sigma- Aldrich, Gillingham, UK
Sodium β glycerophosphate	G5422	0.02 mM	Sigma- Aldrich, Gillingham, UK
Sodium pyrophosphate decahydrate	221368	0.02 mM	Sigma- Aldrich, Gillingham, UK
Sodium fluoride	S7920	0.1 mM	Sigma- Aldrich, Gillingham, UK

2.2.6.5 High detergent buffer (HDB)

Name	Reference/ Catalogue No	Concentration	Supplier
Phosphate-buffered saline	P4417	0.5 x	Sigma- Aldrich, Gillingham, UK
NaCl		137mmol/L	
KCl		2.7mmol/L	
Na_2HPO_4		10mmol/L	
KH_2PO_4		1.8mmol/L	
4-Morpholinepropanesulfonic acid (MOPS) (pH 7)	M1254	50 mM	Sigma- Aldrich, Gillingham, UK
Magnesium Chloride	M8266	10 mM	Sigma- Aldrich, Gillingham, UK
Pefabloc®	11 429 868 001	2 mM	Roche Diagnostics GmbH, Mannheim, Germany
Sodium chloride (NaCl)	S9888	1 mM	Sigma- Aldrich, Gillingham, UK
Triton X-100	X 100	1% v/v	Sigma- Aldrich, Gillingham, UK
Phosphatase inhibitor cocktail 2 (P5726)	P5726	1% (v/v)	Sigma- Aldrich, Gillingham, UK
Sodium β glycerophosphate	G5422	0.02 mM	Sigma- Aldrich, Gillingham, UK
Sodium pyrophosphate decahydrate	221368	0.02 mM	Sigma- Aldrich, Gillingham, UK
Sodium fluoride	S7920	0.1 mM	Sigma- Aldrich, Gillingham, UK

2.2.6.6 Intermediate filament protein solubilising buffer

Name	Concentration	Supplier
Urea	10M	Sigma-Aldrich, Gillingham, UK
Guanidine Hydrochloride	4M	Sigma- Aldrich, Gillingham, UK
Triethyl ammonium bicarbonate buffer (TEAB)	1M	Sigma- Aldrich, Gillingham, UK

2.2.6.7 Buffers for strong cation exchange

2.2.6.7.1 Buffer A

Name	Concentration	Supplier
Acetonitrile	20% v/v	VWR, Lutterworth, UK
Formic acid	0.1% v/v	Sigma-Aldrich, Gillingham, UK
Water: HPLC Gradient grade W/0106/17	2.5 litre	Fisher Scientific, NJ, USA

2.2.6.7.2 Buffer B

Name	Concentration	Supplier
Acetonitrile	20% v/v	VWR, Lutterworth, UK
Formic acid	0.1% v/v	Sigma-Aldrich, Gillingham, UK
KCl	500 mM	Sigma-Aldrich, Gillingham, UK
Water: HPLC Gradient grade W/0106/17	2.5 litre	Fisher Scientific, NJ, USA

2.2.6.8 Buffers for HPLC

2.2.6.8.1 Buffer A

Name	Concentration	Supplier
Acetonitrile	3% v/v	VWR, Lutterworth, UK
Formic acid	0.1% v/v	Sigma-Aldrich, Gillingham, UK
Water: HPLC Gradient grade W/0106/17	2.5 litre	Fisher Scientific, NJ, USA

2.2.6.8.2 Buffer B

Name	Concentration	Supplier
Acetonitrile	97% v/v	VWR, Lutterworth, UK
Formic acid	0.1% v/v	Sigma-Aldrich, Gillingham, UK
Water: HPLC Gradient grade W/0106/17	2.5 litre	Fisher Scientific, NJ, USA

2.2.6.9 Other buffers and solutions

2.2.6.9.1 Tris-Buffered Saline and Tween 20 (TBS-T)

Name	Concentration	Supplier
Tris	20mM	Sigma- Aldrich, Gillingham, UK
NaCl	137 mM	Sigma- Aldrich, Gillingham, UK
Tween 20	1%(v/v)	Sigma- Aldrich, Gillingham, UK

2.2.6.9.2 Phosphate-buffered saline (1x), pH 7.4

PBS		Sigma- Aldrich, Gillingham, UK
NaCl	137mmol/L	
KCl	2.7mmol/L	
Na_2HPO_4	10mmol/L	
KH_2PO_4	1.8mmol/L	

2.2.7 SDS PAGE solution for preparing gels (12%)

Name	Concentration	Reference/ Catalogue No	Supplier
Tris, pH 8.8	1M	93392	Sigma- Aldrich, Gillingham, UK
Tris, pH 6.8	1M	93362	Sigma- Aldrich, Gillingham, UK
SDS	20%	L3771	Sigma- Aldrich, Gillingham, UK
Acrylamide + Bis acrylamide(37.5:1)	40%	A9099	Sigma- Aldrich, Gillingham, UK
Ammonium persulfate(APS)	10%	A3678	Sigma- Aldrich, Gillingham, UK
(N,N,N′,N′-Tetramethylethylenediamine)TEMED	NA	T9281	Sigma- Aldrich, Gillingham, UK

2.3 Scientific equipment

Name	Supplier
pH meter	Metrohm 780
Cold centrifuge	MIKRO 22R Hettich Zentrifugen GMBH (Germany)
Speed vacuum centrifuge Eppendorf Concentrator 5301	Eppendorf AG, Hamburg, Germany
T75 tissue culture flask	Sigma- Aldrich, Gillingham, UK
T25 tissue culture flask	Sigma- Aldrich, Gillingham, UK
Precellys™ 24	Bertin Technologies, Villeurbanne, France
Precellys™ tube	Bertin Technologies, Villeurbanne, France
Bioruptor® Sonicator	Diagenode, Cambridge, UK
BioTek FLx800	BioTek Instruments, Winooski, VT, USA
Costar 96 well vinyl assay plates	Costar, Cambridge, MA, USA
Lo-Bind Eppendorf	Eppendorf AG, Hamburg, Germany
Mini Trans-Blot® Cell	Bio-Rad Laboratories Limited, Hemel Hempstead, UK
PolySULFOETHYL™ A Column	PolyLC, Columbia, MD
BioLC HPLC unit	Dionex, Surrey, UK
UV Detector	Dionex/LC Packings, Amsterdam, The Netherlands
Ultimate 3000 capillary nano-LC	Dionex, LC Packings, The Netherlands
0.3 x 5 mm trap column (3 µm C18)	Dionex-LC Packings, Amsterdam, The Netherlands
0.075 x 150 mm analytical column (3 µm C18)	Dionex-LC Packings, Amsterdam, The Netherlands
Applied Biosystems QSTAR® XL Hybrid LC/MS/MS System	Life Technologies Limited, Paisley, UK
Chemigenius Bio-Imaging System	Syngene, Cambridge, UK

2.4 Endoscopic equipment / materials required

Equipment/bowel prep	Supplier
Kleanprep™	Norgine Limited, Middlesex UK
Picolax	Ferring, West Drayton, Middlesex, UK
Olympus EndoJaw biopsy forceps	KeyMed, Essex, UK

2.5 Developing a technique for isolation of intermediate filament proteins from colorectal biopsies followed by parallel analyses of the isolated proteins using western immunoblotting and mass spectrometry.

The essential steps of isolating and analysing IF proteins are summarised in table 2.2

Table 2.2 *Steps for isolating and analysing IF proteins*

1	Isolating IF proteins from cell lines (MCF-7) and colonic biopsies
2	Identifying an optimal solubilisation technique for western blotting and mass spectrometry
3	Demonstrating maintenance of post translational modifications by western blot

2.5.1 Cell culture technique

2.5.1.1 Obtaining cell cultures stored in liquid nitrogen.

MCF 7 cells stored in liquid nitrogen were thawed in a water bath at 37°C. The thawed cells were transferred into a conical tube containing 1640 RPMI [which was supplemented with 10%(v/v) heat-deactivated foetal calf serum and antibiotics (5% (v/v) Penicillin (10000 units/mL) and Streptomycin (10000 µg/mL))] warmed media. This was centrifuged at 1000 RPM at 4°C for 5 minutes, the supernatant media was decanted off the pellet. 5 ml of growth media was added and the pellet was re-suspended. This was then transferred onto a T25 flask and incubated at 37°C for 2-3 days.

2.5.1.2 Splitting cell cultures

T75 flasks with MCF 7 cells cultured in the growth media were removed from the incubator and checked for 40-50% confluence under a microscope. The old media was aspirated, and following a wash with PBS (10 mls) and incubation with trypsin at 37°C, the cells were dislodged from the flask wall following administering trypsin. The suspension was centrifuged (1000 RPM, 4°C 5 minutes); the resulting pellet re-suspended in the growth medium and transferred to further T75 flasks.

2.5.2 Intermediate filament protein extraction and solubilisation

In this study we aimed to investigate proteomic changes in IF proteins in mucosa of patients with UC and validate the changes noted using western immunoblotting. Sample preparation is critically important in most proteomic experiments. Keratin and other IF proteins differ significantly from cytoskeletal proteins like actin, myosin, and tubulin in their solubility. They are resistant to dissolution in buffers consisting of non-denaturing detergents and high concentrations of ions [116-118]. However they can be dissolved in high concentrations of urea (9-10M) or guanidine hydrochloride (GuHCl)(4-6M) [115, 117]. Urea is a chaotropic agent which can be removed easily at the stage of reversed-phase liquid chromatography (RP-LC). However, use of urea is fraught with certain inherent disadvantages. The high concentration necessary to solubilise keratins interferes with tryptic digestion of proteins prior to mass spectrometry [190]. In addition, heating of the mixture can result in decomposition of urea to isocyanic acid and carbamylate free amines. The later blocks trypsin digestion sites and also result in peptide modifications indistinguishable from the naturally occurring PTMs [191]. This can affect subsequent analyses for PTMs in IF proteins. Guanidine hydrochloride is compatible with mass spectrometry based analyses [192-194]. However use of GuHCl in WB is limited by the fact that it results in formation of precipitates with SDS [192].

Isolation of intermediate filament proteins from mammalian cell culture monolayers has been previously described using a sequential fractionation technique [122, 195]. We undertook further modification of the technique in order to isolate IF proteins from colorectal biopsies and ensure their effective solubilisation in 4M guanidine hydrochloride. This was done to facilitate parallel analysis of IF proteins by gel-based and gel-free liquid chromatography MS/MS (LC-MS/MS) approaches.

2.5.2.1 Intermediate filament protein isolation from MCF-7 cancer cell lines

IF isolation from cell lines was undertaken using a 'high salt extraction technique' which has been previously described by Achtstaetter *et al*. and subsequently modified by Hermann *et al* [117, 195]. The culture media overlying the MCF-7 cell monolayer in T75 flasks was decanted and gently washed with warm wash buffer. This was then treated with LDB, agitated gently and left to incubate at room temperature for 60-90 seconds. The buffer containing the MCF-7 cells was then centrifuged for 30s at 13000 rpm. The supernatant was retained as fraction 1; the fraction containing the membrane bound soluble protein. The remaining cellular/cytoskeletal components in the flasks were then incubated on ice with HDB (containing benzonase nuclease) (to break down the viscous nucleic acid) for 10 mins. Ice-cold NaCl (250 µL, 5M) was added to the flask and the cytoskeletal components were detached from the flask by gentle pipetting. This fluid constituted the entire cytoskeletal fraction. This was centrifuged at 10000G, for 10 minutes at 4°C; the supernatant containing the high salt soluble fraction (fraction 2) was removed, while the insoluble pellet containing the insoluble cytoskeletal components (fraction 3) was used in our study.

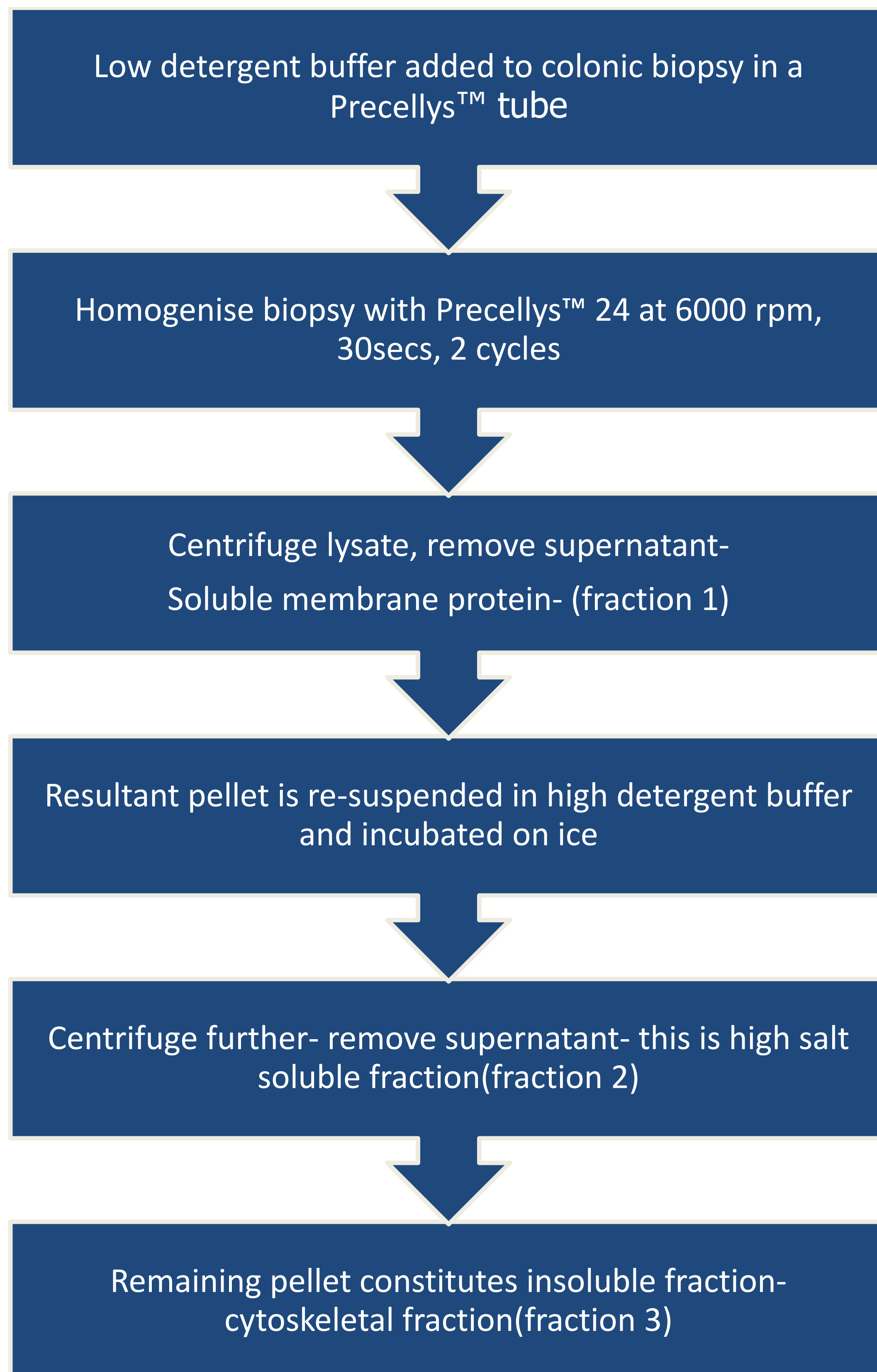
2.5.2.2 Extraction of intermediate filament proteins from human colorectal biopsies (Fig 2.1)

The 'high-salt extraction' technique described above to extract the insoluble cytoskeletal components in cell lines was modified for use on colorectal

biopsies. Precellys™ tubes containing ceramic beads were used to homogenise the biopsies in LDB; a single colonic biopsy was placed in each Precellys™ tube containing 300uL LDB. The tube was then placed in Precellys™ 24 and run at 6000 RPM, 30secs, 2 cycles (with 20 sec rest in between cycles).

The resultant lysate was removed into LoBind eppendorfs®, centrifuged and the supernatant was stored as Fraction 1- the soluble and extractable membrane protein fraction. The insoluble pellet obtained following centrifugation (at 2200 RPM, 5 minutes, 4°C) was washed with further LDB and re-suspended in 200uL HDB (with Benzonase® Nuclease, added at this stage to breakdown residual viscous nucleic acids) followed by incubation on ice for a further 10 minutes. Another centrifugation (at 9020 RPM, 10 minutes, 4°C) yielded a supernatant- the high salt soluble protein (fraction 2); the residual pellet was stored at -80°C as the insoluble cytoskeleton fraction (fraction 3).

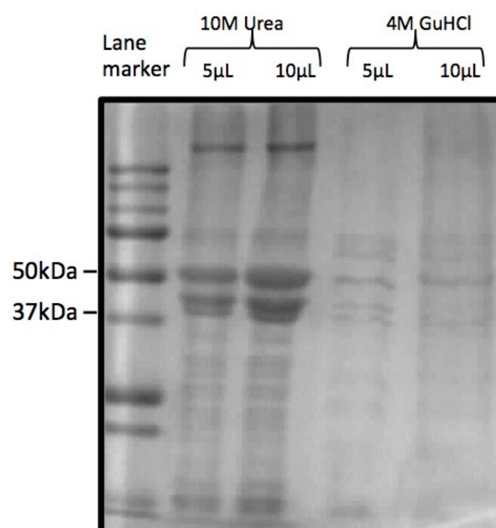
Figure 2.1 *Steps for isolation of insoluble intermediate filament proteins from colorectal biopsies*



2.5.2.3 Solubilising 'detergent insoluble' cytoskeletal fraction (fraction 3) in urea and guanidine hydrochloride (GuHCl)

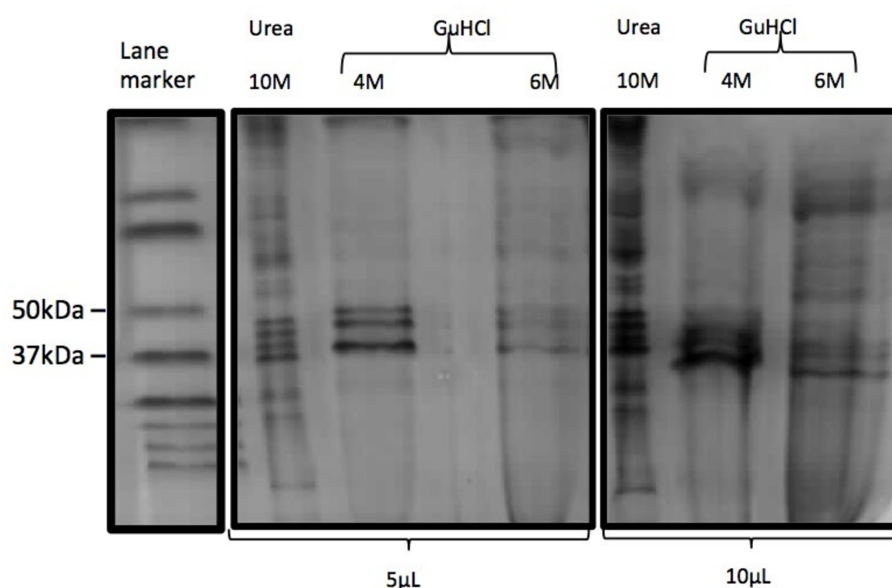
In order to compare solubility of cytoskeletal protein fraction, the extracted fraction 3 obtained from MCF-7 cell lines were dissolved in 20µL of 10M urea and 20µL of 4M GuHCl, the later dissolved in Triethyl ammonium bicarbonate buffer (TEAB) (pH 8.5) respectively. Coomassie stained gel is shown in figure 2.2. Prominent bands between 37-50kDa (which we have previously shown to be keratin-rich) were noted in lanes that had fraction 3 dissolved in urea. Bands obtained with GuHCl dissolved samples were less prominent. Inadequate IF protein solubilisation by the 4M GuHCl was initially hypothesised as the reason for such discrepancy.

***Figure 2.2*. Comparison of band intensity of dissolved cytosketal fraction obtained from MCF-7 cells.** *Following extraction of the cytoskeletal fraction (using a high salt extraction technique) it was dissolved separately in urea (10M) and GuHCl (4M) respectively. Proteins were loaded onto SDS-PAGE gels and separated by gel electrophoresis followed by Coomassie staining of the gel). Increased protein band intensity is noted with samples dissolved in 10M urea.*
GuHCl-Guanidine hydrochloride, SDS-PAGE- sodium dodecyl sulfate polyacrylamide gel electrophoresis



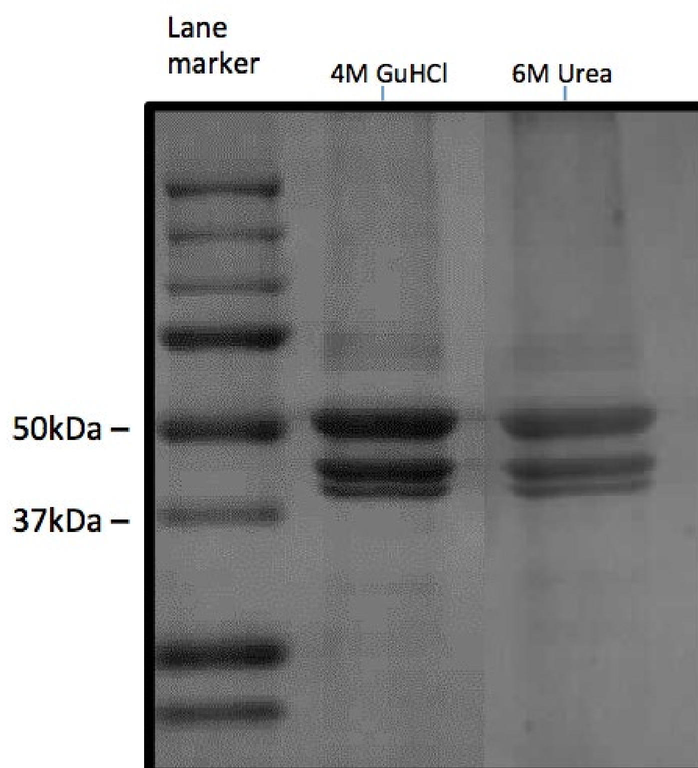
In order to optimise solubilisation in GuHCl, the experiment was then repeated by comparing the solubility of fraction 3 proteins in two strengths of GuHCl (4M or 6M respectively). The results on SDS-PAGE were also compared to equivalent samples dissolved in urea (10M). Band intensity of IF proteins was greater when dissolved in 4M GuHCl than 6M. The intensity (of the bands with the former) was however less than with samples dissolved in urea (Figure 2.3). On heating the dissolved samples with Laemmli buffer, visible precipitates were noted with the GuHCl-dissolved samples.

Figure 2.3 Comparison of band intensity of dissolved cytosketal fraction with varying concentrations of GuHCl. *The extracted cytoskeletal fraction (Fraction 3) was dissolved separately in urea (10M) and increasing concentrations of GuHCl (4M and 6M respectively). Proteins were loaded onto SDS-PAGE gels and separated by gel electrophoresis followed by Coomassie staining of the gel. . Band intensities were compared. With protein dissolved in 6M GuHCl band intensity was less than with 4M GuHCl, due to formation of precipitates with SDS. The band intensity was higher using 10M urea than 4M GuHCl.*
SDS-PAGE- sodium dodecyl sulfate polyacrylamide gel electrophoresis; GuHCl - Guanidine hydrochloride.



In order to overcome issues of sample solubility, the commercially available kit (Pierce® SDS-PAGE Sample Prep Kit) was used to remove interfering buffer component. Band intensities on removal of solvent (GuHCL and Urea) were compared. Results were similar. These results suggest a satisfactory cytoskeletal solubilisation with GuHCl when compared with Urea (Figure 2.4).

Figure 2.4. Band intensity of dissolved cytosketal fraction after removal of solvent. *Extracted cytoskeletal fraction was dissolved separately in urea (6M) and GuHCl (4M) respectively. A commercially available kit (Pierce® SDS-PAGE Sample Prep Kit) was used to remove interfering buffer component. Proteins were loaded onto SDS-PAGE gels and separated by gel electrophoresis followed by Coomassie staining of the gel. Band intensity of separated fraction 3 proteins dissolved in GuHCl (4M) and urea (6M) were noted to be similar. GuHCl- Guanidine hydrochloride, SDS-PAGE- sodium dodecyl sulfate polyacrylamide gel electrophoresis*



2.5.2.4 Assessing suitability of guanidine hydrochloride as a solvent for cytoskeletal proteins obtained from colorectal biopsies

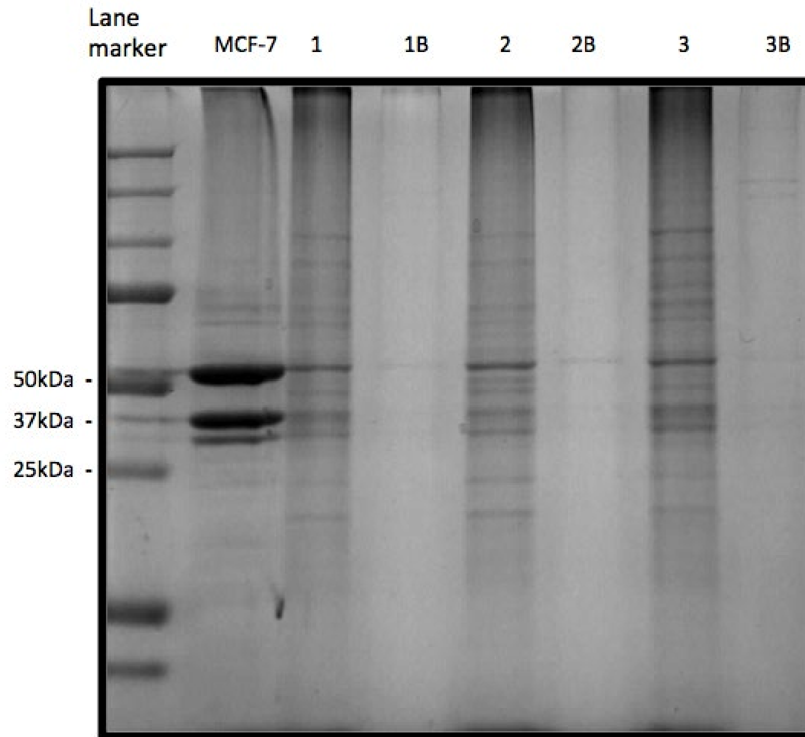
Suitability of GuHCl in solubilising fraction 3 proteins obtained from sigmoid colon biopsies was tested. 50µL of 4M GuHCl (in TEAB) was used to solubilise the cytoskeletal fraction. Most of the pellet obtained was solubilised, although there remained persistent un-dissolved sediment.

2.5.6 Sonication to maximise protein solubilisation

To avoid any significant loss of protein in the sediment and maximise the solubilisation of the insoluble cytoskeletal fraction in either urea or GuHCl, a Bioruptor® Sonicator was used. 3 other colonic biopsies were analysed. Fraction 3 proteins were dissolved in a larger volume GuHCl (200µL of 4M GuHCl). A further 200µL of 4M GuHCl in TEAB was added to the un-dissolved sediment. The supernatant was removed following centrifugation (at 600xG, 1 min at 4°C). The samples were sonicated in ice/water mix for 5 cycles, with each cycle lasting 30 seconds. The samples were then centrifuged (at 2200 RPM, 1 minute, 4°C), the supernatant was analysed to identify the effectiveness of each solubilisation strategy.

Figure 2.5 Results of Coomassie stained gels of extracted cytoskeletal fraction from colorectal biopsies dissolved in GuHCl. *Modification of the 'high salt extraction' technique was undertaken to extract cytoskeletal proteins from colorectal biopsies of 3 patients (1, 2 and 3 respectively). They were dissolved in 200µL of 4M GuHCl. The supernatant with the dissolved IFs was removed by centrifugation followed by removal of solvent using Pierce SDS-PAGE sample prep kit. It was loaded on to the gel (Lane labels 1,2 and 3 respectively). Further 200µL of 4M GuHCl was added to the un-dissolved precipitate. This was sonicated for 5 mins, and following removal of solvent loaded onto the gel (1B, 2B, 3B). Dissolved cytoskeletal fraction obtained from MCF-7 was used as a control. Adequate protein separation was noted in lanes 1, 2 and 3 with negligible protein in 1B, 2B and 3B respectively.*

GuHCl- Guanidine hydrochloride,SDS-PAGE- sodium dodecyl sulfate polyacrylamide gel electrophoresis



2.5.7 Removal of substances (GuHCl and urea) from dissolved cytoskeletal proteins, which could potentially interfere with effective protein separation on SDS-PAGE gels

To eliminate the effects of precipitation of SDS and GuHCl, and compare the solubilisation of fraction 3 in GuHCl, with urea dissolved samples, a commercially available kit Pierce® SDS-PAGE Sample Prep Kit was used. The manufacturers instructions were followed.

2.5.8 Demonstration of adequate band intensity on SDS-PAGE with both urea and GuHCl

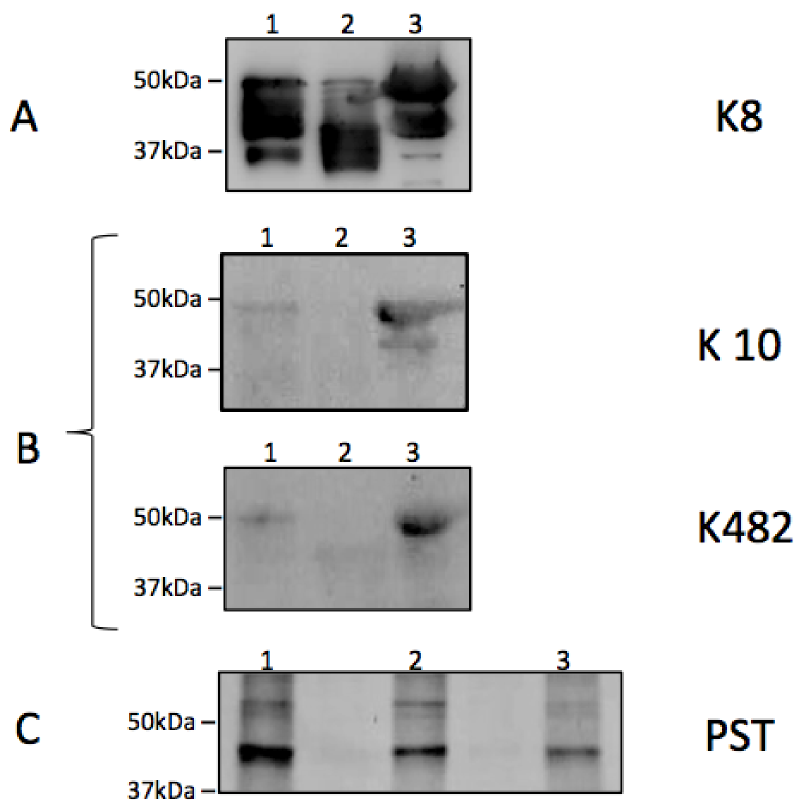
Protein elution was followed by boiling the dissolved protein samples at 95°C for 5 minutes in Laemmli lysis buffer (provided with the kit). They were analysed on 12% SDS-PAGE gels, as previously described with F3 of MCF-7 as control [196]. 5 and 10 µL of protein samples were loaded onto the lanes in the gel. Coomassie staining of the gels was performed using InstantBlue™ in order to visualise the separated the protein bands. Results of the Coomassie stained gels are shown in Figure 2.5. These results showed that the dissolved protein content in each band was satisfactory (Lane 1, 2 and 3 respectively). A negligible amount of protein was noted in the sonicated samples (Lane 1B, 2B and 3B respectively) (figure 2.5)

2.5.9 Demonstration of post- translational modifications (PTMs) in the resolved samples by western immunoblotting

In order to evaluate the retention of PTMs in keratin, in samples dissolved in GuHCl, SDS-PAGE gels were run, western transferred and immunoprobed using the following antibodies- mouse monoclonal antibody to keratin 8, anti-phosphoserine/threonine antibody and antibody to acetylation sites of keratin 8 (lys10 and lys482). The technique for this has been described previously in the literature and is detailed in section 2.6.2 (page 88) [197, 198]. Proteins were separated by SDS–PAGE and transferred onto polyvinylidene difluoride (PVDF) membranes. Primary antibodies used were 1:1000 dilutions each of mouse monoclonal antibody to keratin 8 [M20], polyclonal anti-Phosphoserine/threonine antibody and in-house antibodies raised against acetylated lysine residues of K8 (lys10 and lys482). Cross-reaction was

visualised using polyclonal secondary horseradish peroxidase (HRP) antibodies - conjugated rabbit anti mouse and goat anti-rabbit IgG and enhanced chemiluminescence reagent. Intensity of the bands was analysed with a Chemigenius Bio-Imaging System. Cross-reactions to all four antibodies were seen between 37 and 50kDa. This demonstrates preservation of the modifications of K8 (particularly acetylation) after solubilisation of IF in GuHCl (Figure 2.65).

Figure 2.6 Demonstration of K8 expression and maintenance of PTMs in K8 in extracted cytoskeletal fraction from colonic biopsies. *The extracted cytoskeletal fraction from 3 patients was dissolved in GuHCl followed by removal of solvent, gel electrophoresis on SDS-PAGE gels, western transfer and immunoprobing using the following antibodies- mouse monoclonal antibody to K8 (Panel A), antibody to acetylation sites of keratin 8 (lys10 and lys482)(Panel B) and anti-phosphoserine/threonine antibody(Panel C). Results show adequate K8 expression and maintenance of PTMs.*
K- keratin, GuHCl- Guanidine hydrochloride, SDS-PAGE- sodium dodecyl sulfate polyacrylamide gel electrophoresis



A summary of the essential steps performed in intermediate filament extraction and dissolution. The steps were optimised to be compatible with SDS-PAGE and for MS analysis, including analysis of phosphorylation and acetylation, which are preserved (figure 2.7 and 2.8).

Figure 2.7 *Essential steps in developing a technique for extracting and dissolving IF proteins using MCF 7 cell lines and colorectal biopsies to allow analysis using SDS-PAGE and mass spectrometry.*

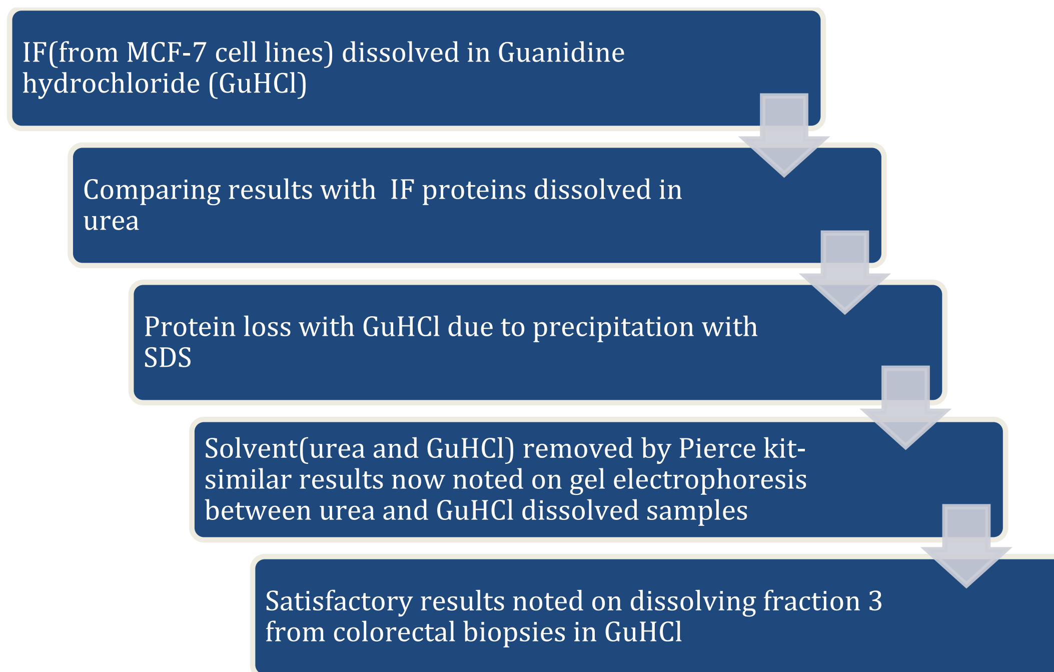
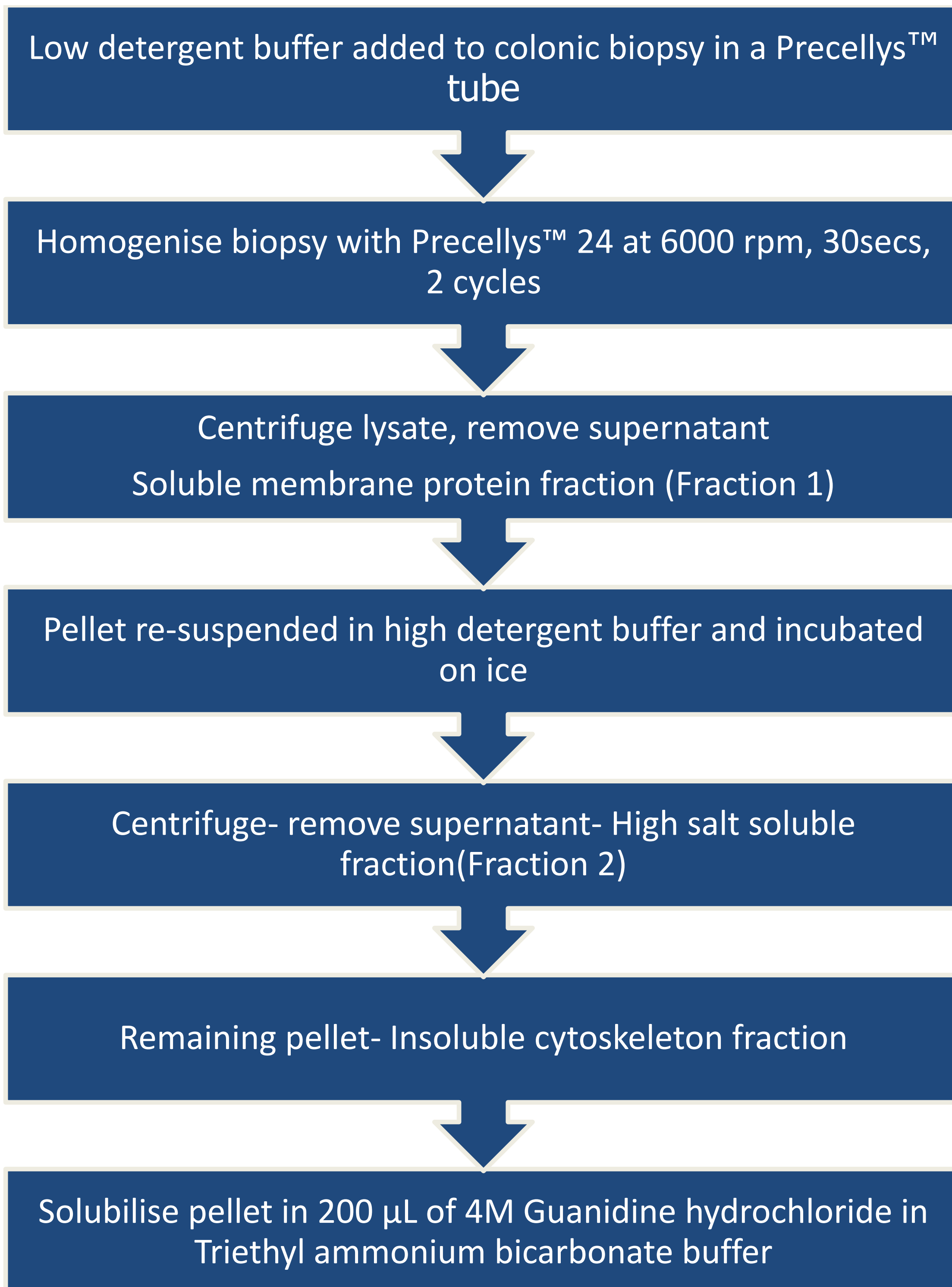


Figure 2.8 *Steps in processing and solubilising colonic biopsies*



2.6 Analysing colorectal biopsies using mass spectrometry and orthogonal validation using western blotting.

The workflow for the proteomics experiment is described in figure 2.9.

2.6.1 Protein concentration assay

Protein concentration in solubilised cytoskeletal proteins was assayed in triplicate using Costar 96 well vinyl assay plates and Bio-Rad protein assay reagent according to manufacturer's instructions. Absorbance was measured at 595 nm (reference at 450 nm) against standard BSA curves using a BioTek FLx800 multi-detection reader.

2.6.2 Preparing samples for analysis by mass spectroscopy

The dissolved cytoskeletal fraction obtained from individual colon biopsy samples were pooled into eight groups being investigated as described in section 3.7. Hundred(100) μg of protein pooled in each group was diluted with TEAB to 2M GuHCl to ensure equal volumes for trypsin digestion.

2.6.2.1 Reduction and alkylation of proteins

The cysteine residues in the samples were blocked by adding 50mM Tris (2-carboxyethyl) phosphine hydrochloride (TCEP-HCl), followed by alkylation using 10mM Methyl methanethiosulfonate (MMTS) (Thermo Scientific, Rockford, IL, USA).

2.6.2.2 Trypsin digestion step

Sequencing grade modified trypsin was prepared at 10 μg/ml by adding 50 μl trypsin to suspension buffer according to the manufacturers' instructions. Trypsin digestion was carried out at a 1:20 ratio for 16 hours at 37°C. The solution from the overnight digest containing peptide was used for iTRAQ labelling.

2.6.2.3 iTRAQ (8-plex) labelling

Isopropanol (50 µl) was added separately to the 8-iTRAQ reagents (113-119, 121). The resulting digested peptides were iTRAQ labelled with 8-plex iTRAQ reagents according to manufacturers' protocols as previously published [199].

The reagents were then added to the trypsin digested samples as follows:

113: Controls (CON)

114: Active disease (ACT)

115: Inactive disease proximally (INACT)

116: Long standing Pancolitis (LSPC)

117: Recent onset UC (ROUC)

118: PSC and colitis (PSC)

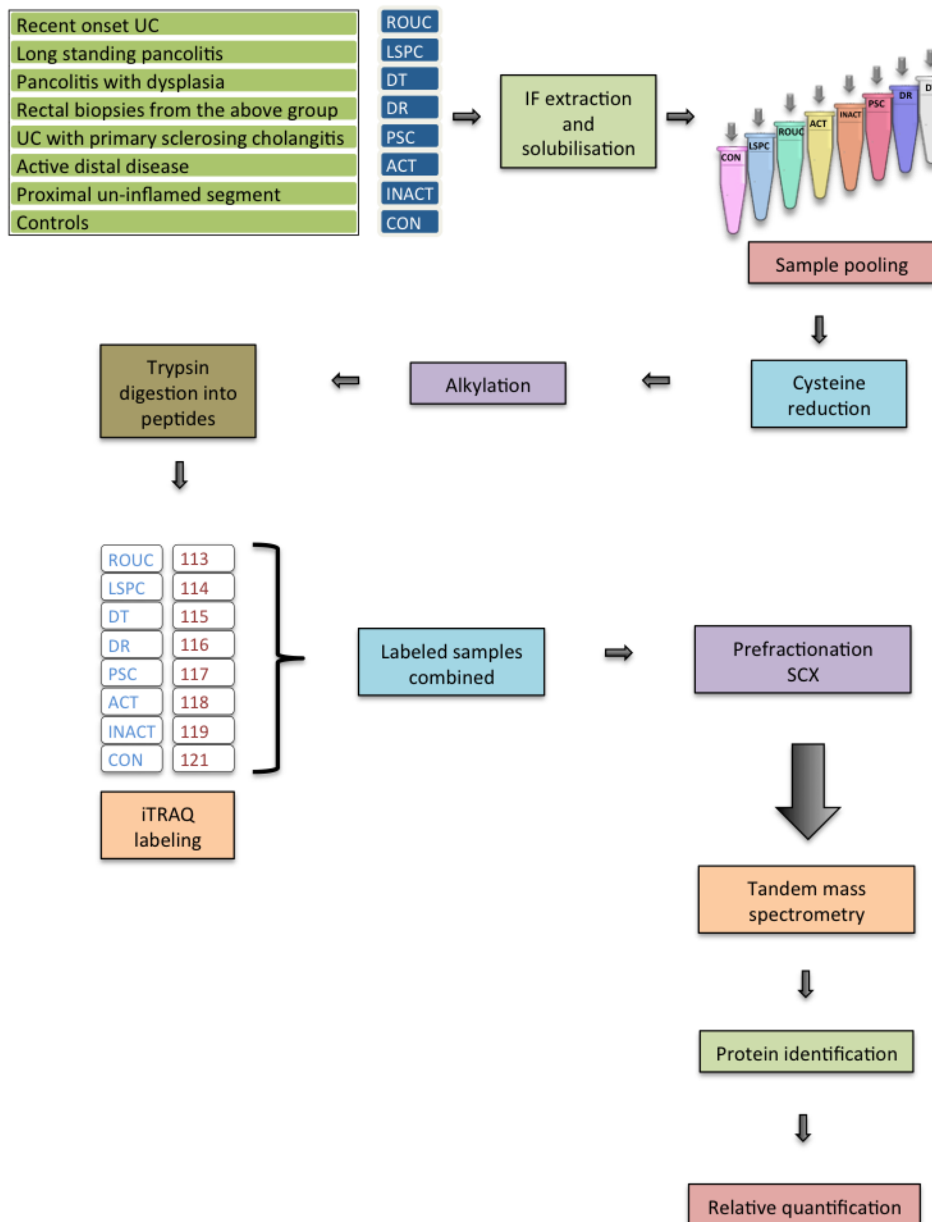
119: Patients with dysplasia- biopsies obtained from the rectum (DR)

121: Biopsies from the dysplastic lesions (DT)

iTRAQ labelled samples were incubated at room temperature for 2 hours followed by pooling in a 1:1 ratio. The pooled mixture was dried in a vacuum centrifuge and was reconstituted in 90 µl of 20% acetonitrile (ACN), 0.1% formic acid and the pH of the solution was adjusted to 2 (using formic acid). This reconstituted peptide mixture underwent sonication followed by centrifugation to remove insoluble material.

Figure 2.9 Proteomic workflow for analysing colorectal biopsies*. 8 pooled samples from the patient groups were processed. Labelled peptides from pooled patients were analysed by SCX-LC-MS/MS and data reconstituted in GeneBio Phenyx. All runs were searched against Uniprot database.*[200]

(CON- control, LSPC- long standing pancolitis, ROUC- recent onset UC, PSC-UC associated with primary sclerosing cholangitis, DR- rectal biopsy in patients with dysplasia, DT- biopsies obtained from the dysplastic tissue, ACT- active disease, INACT- proximal inactive segment in the ACT patients)



2.6.2.4 Strong cation exchange (SCX) fractionation of peptides

Strong cation exchange was carried out on a BioLC HPLC unit prior to MS/MS analyses. This technique has been described previously in the literature by Pham *et al* [200].

2.6.2.4.1 Peptide fractionation

Strong cation exchange (SCX) was achieved using a PolySULFOETHYL A Pre-Packed Column (PolyLC, Columbia, MD) with a 5 µm particle size and a column dimension of 100 mm.4.6 mm i.d., 200 Å pore size, on a BioLC HPLC (Dionex, Surrey, U.K.). Sample was loaded onto the column and washed for at least 60 minutes at a flow rate of 400 µL /min with 100 % SCX Buffer A (20 % acetonitrile, 0.1 % Formic Acid) to remove salts, TCEP and unincorporated iTRAQ reagent. Peptides were then separated using a gradient of SCX Buffer B (20% acetonitrile, 0.1% formic acid, 0.5 M KCl) at the same flow rate of 400 µL /min. Buffer B levels increased from 0% to 25% from 5 minutes to 30 minutes then from 25% to 100% over 5 minutes, followed by an increase from 26% to 100% over the next 15 min. Buffer B was held for another 5 min for isocratic washing prior to column re-equilibration with buffer A. The sample injection volume was 100 µL, and the liquid flow rate was 400 µL/min. The SCX chromatogram was monitored using UVD170U ultraviolet detector and Chromeleon software v. 6.50 (Dionex, LC Packings, The Netherlands). Fractions were collected using a Foxy Jr. (Dionex) fraction collector in 1 min intervals.

2.6.2.4.2 Sample desalting

Fractions were desalted using buffers: A (97% H_2O, 0.1% formic acid); B (97% acetonitrile, 0.1% formic acid) and C (3% acetonitrile, 0.1% trifluoroacetic acid) and MiniSpin™ columns according to the protocol outlined by The Nest Group (The Nest Group, MA, USA). Fractions were vacuum-concentration prior to LC-MS/MS analysis.

2.6.2.4.3 LC-MS/MS analysis

Fractions collected from offline separation techniques were eluted through the Famos-Ultimate 3000 nano-LC system (Dionex, LC Packings, The Netherlands) interfaced with a QSTAR® XL (Applied Biosystems; MDS-Sciex) tandem ESI-QUAD-TOF MS. Vacuum dried fractions were resuspended in loading buffer (3% acetonitrile, 0.1% trifluoroacetic acid), injected and captured into a 0.3.5 mm trap column (3 μm C18 Dionex-LC Packings). Trapped samples were then eluted onto a 0.075.150 mm analytical column (3 μm C18 Dionex-LC Packings) using an automated binary gradient with a flow of 300 nL/min from 95% buffer A (3% acetonitrile, 0.1% formic acid), to 35% buffer B (97% acetonitrile, 0.1% formic acid) over 90 min, followed by a 5 min ramp to 95% buffer II (with isocratic washing for 10 min). Predefined 1 s 350−1600 m/z MS survey scans were acquired with up to two dynamically excluded precursors selected for a 3 s MS/MS (m/z 65−2000) scan. The collision energy range was increased by 20% as compared to the unlabeled peptides in order to overcome the stabilizing effect of the basic N-terminal derivatives, and to achieve equivalent fragmentation as recommended by Applied Biosystems.

2.6.2.4.4 Protein identification and relative quantification

MS/MS data generated from the QSTAR® XL was converted to generic MGF peaklists using the mascot.dll embedded script (version 1.6 release no. 25) in Analyst QS v. 1.1 (Applied Biosystems, Sciex; Matrix Science). Further processing of the data was undertaken using an in-house Phenyx algorithm cluster (binary version 2.6; Geneva Bioinformatics SA) at the ChELSI Institute, University of Sheffield, against the *Homo sapiens* UniProt protein knowledgebase (SwissProt and Trembl (41070 and 71449 entries respectively, downloaded July 2012) to derive peptide sequence and hence protein identification. These data were then searched within the reversed *Homo sapiens* database to estimate the false-positive rate. Peptides identifications at 1% false discovery rate were accepted. The iTRAQ reporter ion intensities were exported. Protein quantifications were obtained by computing the geometric means of the reporters' intensities. Median correction was subsequently applied to every reporter in order to compensate for systematic errors *e.g.* if a sample happened to have been loaded at a largely different total concentration. The reporter intensities, in each individual MS/MS scan, were also median corrected using the same factors, with the rationale that if the total concentration of a sample A was half that of another sample B, the intensities of sample A's reporter have to be doubled to allow for a fair comparison. *t*-tests applied to determine alterations in protein level between samples use these corrected intensities since these were was carried out for every protein and because of the multiple times each test was performed, the threshold (α=5%) used for significance was corrected for data

mining. A standard Bonferroni correction was used to minimise false positive results. The workflow was developed in house [200].

2.7 Orthogonal validation using western immunoblotting.

2.7.1 Transferring of resolved proteins onto PVDF membranes and immunoblotting

Following elution of protein using the Pierce SDS-PAGE Sample Prep Kit, the fraction 3 in individual eppendorfs were boiled with non-reducing sample buffer (provided with the kit) at 95C for 5 min, and then centrifuged at 600g for 1 min. The supernatant obtained was pooled into individual groups as described in 3.7; 5µG of pooled protein was loaded onto 12% SDS PAGE gels. In addition, 5µG of fraction 3 of MCF 7 cells was loaded onto a lane as an internal control. The SDS PAGE gel was prepared as per standard techniques. The gel was transferred into the tank and filled with running buffer and electrophoresed at 100V for 1.5 -2 hours.

The proteins separated by SDS–PAGE was transferred to polyvinylidene difluoride (PVDF) membranes using a wet transfer technique. The transfer was achieved in a tank containing transfer buffer and run at 100 V for 60-90 minutes. The membrane was blocked overnight at 4°C with 5% w/v bovine serum albumin in Tris-Buffered Saline with Tween® 20 (TBST), pH 8.0.

2.7.2 Probing of membranes with primary and secondary antibodies and analysing band intensities

The primary antibodies antibodies described before were used. Cross-reaction was visualised using polyclonal secondary horseradish peroxidase (HRP) antibodies - conjugated rabbit anti-mouse or goat anti-rabbit IgG as appropriate along with enhanced chemiluminescence reagent.

2.7.3 Re-probing PVDF membrane

The PVDF membranes were washed in TBST to remove the chemiluminescence agent. They were then incubated in 5 mls of Restore Western Blot Stripping buffer at room temperature with gentle rocking. Following a further 3 washes in TBST for 10 minutes each, the membranes were re-probed.

2.7.4 Densitometric analysis of immunoblots and calculating acetyl and phosphorylation ratios

Using Chemigenius Bio-Imaging System and Gene Tools software, the membranes were imaged. Individual bands of interest were selected manually and their pixel densities obtained from the imaging software. The software automatically corrected for any background activity. Densitometry results to calculate the proportion of phosphorylated/acetylated K8 were normalised to the band obtained with MCF-7 and calculated as follows:

X=. Phosphorylated/acetylated K8 band intensity/MCF-7 band intensity

Y= Total K8 band intensity/MCF-7 band intensity

X/Y= phosphorylated/acetylated K8 ratio

2.8 Setting up the study and obtaining ethical approval

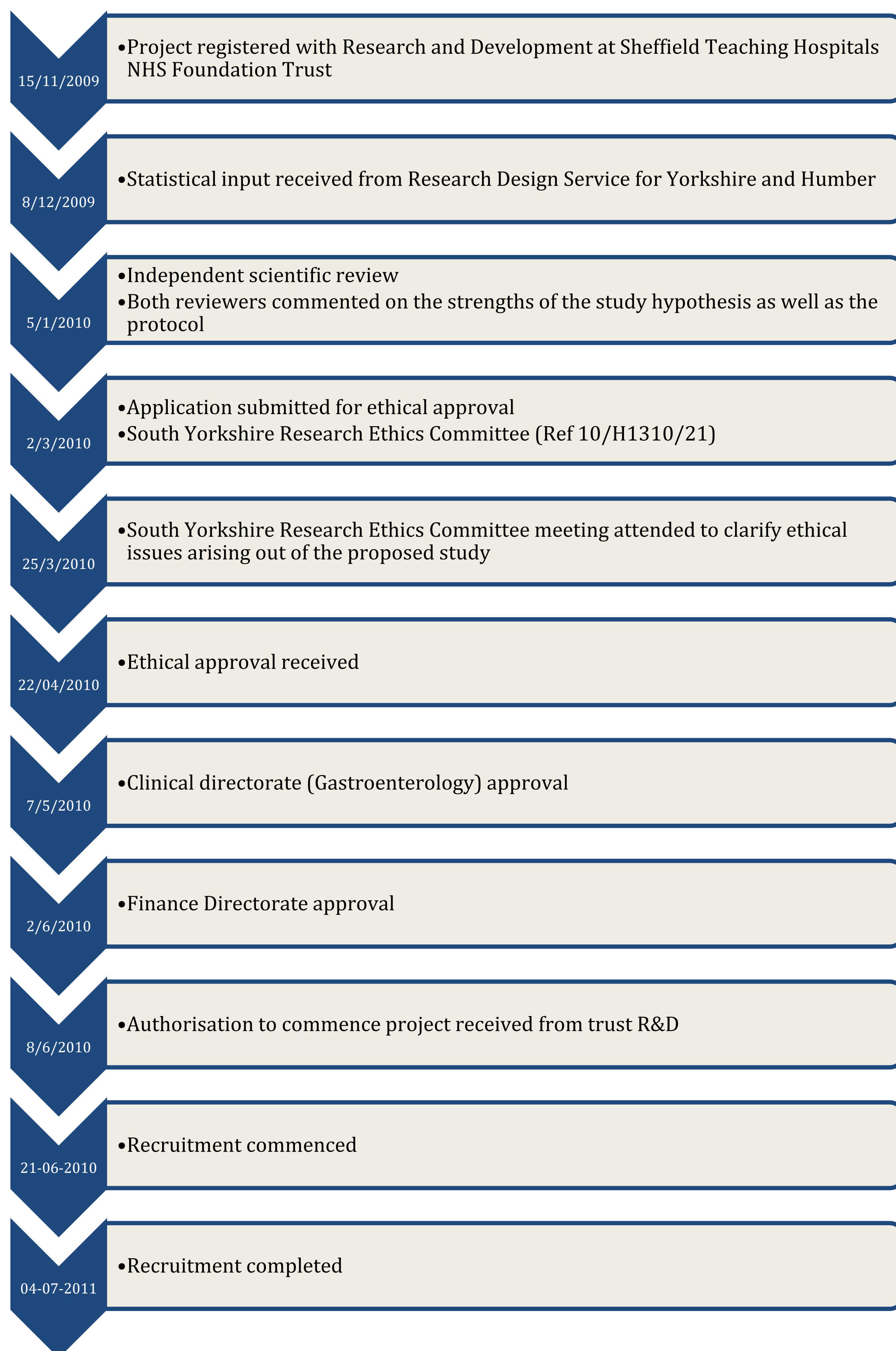
The clinical study was undertaken at Department of Gastroenterology, Royal Hallamshire Hospital, Sheffield Teaching Hospitals NHS Foundation Trust and Department of Oncology, University of Sheffield. Prior to recruitment of patients in the study, we had to complete several steps:

1. Registering the project and developing a protocol for investigation
2. Ensuring successful independent peer review
3. Obtaining favourable ethical approval
4. Obtaining research funding grant for consumables and experimental work.

2.8.1 Time line for registering project, obtaining ethics approval prior to commencement of recruitment of study patients (figure 2.10)

The study was registered with the Sheffield Teaching Hospital NHS Foundation Trust, Research and Development (R&D) department on 15-11-2009. Successful ethical approval from the South Yorkshire Research Ethics Committee was obtained on 22-04-2010 and, recruitment of patients was commenced from 21-6-2010. Further details of the timeline of activities undertaken prior to commencing recruitment have been summarised in figure 2.10.

Figure 2.10. *Timelines with activities undertaken since the start of the project have been summarised.*



2.8.2 Ethical assessment

The research ethics committee received the study favourably. Apart from a few minor clarifications, no further major changes were needed and ethical approval was obtained on 22/04/2011(appendix 7 page 187).

2.8.3 Grant application

A grant application was made to Bardhan Research and Education Trust of Rotherham Ltd on 14/1/2010. This was successful and we received notification from the grant administrator on 13/05/2010.

2.9 Recruitment of patients

Recruitment of patients for the study was commenced on 21/6/2010. Informed consent was obtained from all patients meeting the inclusion and exclusion criteria (Section 2.1.1, page 55). All patients recruited had the the study explained to them in details by the investigators. They were all recruited either by Dr Majumdar (student investigator) or Dr Lobo (Principal investigator). They were also encouraged to read the 'Patient/Control information leaflet' (Appendix 3 & 4 page 172 and page 177 respectively) prior to consenting to the endoscopic procedures.

Colonoscopic /flexible sigmoidoscopic examination and biopsies were performed as per standard clinical practice and indication. Most of the colonoscopic / flexible sigmoidoscopic examinations were undertaken by Dr Majumdar.

2.9.1 Categorisation of patients

The recruited patients were assessed for the following parameters

1. Duration of colitis since diagnosis
2. Clinical disease activity (Mayo score) [187]
3. Mucosal disease activity at endoscopic (Baron's score) [188]
4. Endoscopic extent of the disease in the colon
5. Family history of colorectal cancers

Colonoscopy or flexible sigmoidoscopy was undertaken. Colonic biopsies were obtained according to British Society of Gastroenterology guidelines [2]. Study biopsies (for immunohistochemistry and proteomic studies) were taken, in addition from rectum and mid-sigmoid colon with additional biopsies from the dysplastic regions. In patients with active distal disease, biopsies were taken from inflamed and non-inflamed regions. Any visible abnormal lesions were resected or biopsied with field biopsies taken from the immediate vicinity of dysplastic lesions. Details of the protocol for obtaining colonic biopsies are shown in section 2.1.6 (page 58) and section 2.1.9 (page 60). Study biopsies from each site were either fixed in formalin (for subsequent immunohistochemical studies), or were frozen in liquid nitrogen (for proteomic assays).

In order to analyse samples by quantitative proteomics, patients were categorised into the following phenotypic groups.

1. **ROUC** – Recent onset (duration <5 years) UC
2. **LSPC** – Long standing quiescent pancolitis (duration of disease >20 years) to evaluate effect of disease duration on insoluble intermediate filament protein profile in comparison with ROUC
3. **PSC** – PSC associated with UC: to identify changes in this high-risk group
4. **ACT** – Active distal colitis: to identify effects of acute inflammation in the same group of patients when compared to the inactive proximal segment
5. **INACT** – Inactive proximal colon in patients with distal active disease
6. **DT** – Biopsies from the dysplastic lesions
7. **DR** – Rectal biopsies in patients with dysplasia to investigate for field changes in intermediate filaments
8. Healthy controls **(CON)**

2.9.2 Clinical, histological and endoscopic data

2.9.2.1 Ulcerative colitis patients and controls recruited in the study

A total of 100 patients with UC were recruited along with 18 healthy controls. Details of UC patients are provided in table 5.1. The median age of patients with UC was 54 years (range 20-75). Patients were categorised according to disease extent and other phenotypic characteristics. Number of patients in individual groups was as follows pancolitis (n=47), subtotal colitis (colitis involving most but not all of the colon)(n=6), left sided colitis (n=19), proctosigmoiditis (n=12), proctitis (n=9) and PSC and colitis (n=7). The median duration of disease (in years) in each group is described in table 2.3. There were 4 patients with evidence of dysplasia histologically. Additional biopsies were obtained in these patients from the dysplastic lesions (n=3).

The median age of recruited control patients was 60 years (range 41-75) with a male to female ratio of 3:7.

Patients in each phenotype were also categorised according to the duration of disease. Details are shown in table 2.4.

Table 2.3 *Summary of UC patients recruited in the study*

	No	Median age(years)	Sex M:F	Median duration of disease
Pancolitis	47	57(20-74)	33:14	21(1-42)
Subtotal colitis	6	62.5(51-75)	1:1	31(19-45)
Left sided colitis	19	41(25-71)	11:8	9(0-38)
Proctosigmoiditis	12	49.5(24-70)	1:1	15.5(0-30)
Proctitis	9	42(30-69)	1:8	7(2-25)
PSC and colitis	7	56(38-71)	5:2	17(2-35)

Table 2.4 *Patient groups categorised according to duration of disease*

Duration of disease	No	Pancolitis	Subtotal colitis	Left sided colitis	Proctitis	Procto-sigmoiditis	PSC colitis
<5	19	6	0	7	3	2	1
5-10	14	5	0	3	2	3	1
11-19	23	11	1	4	3	1	3
20-29	23	13	1	2	1	5	1
30-39	18	10	3	3	0	1	1
>40	3	2	1	0	0	0	0

2.9.2.2 Categorisation of UC phenotype according to disease activity

Endoscopic (clinical) and histological activity was assessed in all patients with UC. Details are provided in table 2.5.

Table 2.5 *Median disease activity as assessed clinically (endoscopically) and histologically*

	Barons score		Histological score	
	Rectum	Sigmoid	Rectum	Sigmoid
Pancolitis	0(0-3)	0(0-3)	0(0-3)	0(0-2)
Subtotal colitis	1(0-2)	1(0-2)	1(0-3)	0(0)
Left sided colitis	2(0-3)	2(0-3)	1(0-3)	0(0-1)
Proctitis	2(1-3)	0(0-1)	1(0-3)	0(0)
Proctosigmoiditis	0(0-2)	0(0-2)	0(0-2)	0(0-1)
PSC and colitis	0(0-2)	0(0-2)	0(0)	0(0)

2.9.2.3 Patients with distal active disease only

A group of patients with distal active disease (proctitis/proctosigmoiditis) were included in the study. This group was chosen in order to compare proteomic changes in inflamed mucosa to uninflamed proximal mucosa in the same patient. These patients had biopsies taken from the rectum, mid-sigmoid colon and in addition from proximal un-involved colon. A total of 10 such patients were recruited. Details of disease extent and endoscopic activity are shown in table 2.6

Table 2.6 *Patients with distal active distal disease*

Patient No	Sex	Age	Extent of disease	Disease duration (years)	Baron score	
					Rectum	**Proximal colon**
UC0004	F	71	Left sided	30	2	0
UC0008	F	39	Proctosigmoiditis	5	2	0
UC0016	F	30	Proctitis	13	3	0
UC0018	M	33	Proctitis	3	2	0
UC0021	F	69	Proctitis	2	3	0
UC0026	F	56	Left sided	45	2	0
UC0045	M	33	Left sided	9	2	0
UC0076	F	23	Proctitis	14	2	0
UC0082	F	45	Proctitis	5	2	0
UC0097	F	26	Proctosigmoiditis	1	3	0

2.9.3 Categorisation of patients for proteomic study

In order to analyse samples by quantitative proteomics, recruited patients were categorised into the phenotypic groups described in section 2.9.1(page 99). Eight patients with recent onset (duration of disease <5 years) quiescent disease were included, with a median age of 31 years (range 22-52) and a median duration of disease of 2 years (range 0-4) (Table 2.7). There were 10 patients in the LSPC group with a median age of 59.5 years (range 48-73 years) and median duration of disease of 29 years (range 21-37) (table 2.8). The other two groups included patients with PSC [median age 56 years (range 38-71), duration of disease 17 years (range 5-35)](table 2.9) and dysplasia [median age 60 years (range 20-74), duration of disease 21 years (range 9-41)](table 2.10).

Table 2.7 *Patients with recent onset UC (ROUC)*

Patient No	Sex	Age	Extent of disease	Duration of disease	Baron score Rectum	Histological score Rectum
UC0005	F	25	Left sided	<1	0	0
UC0006	M	52	Pancolitis	3	0	0
UC0009	F	42	Proctitis	2	1	1
UC0025	F	31	Left sided	<1	1	0
UC0032	F	22	Pancolitis	3	1	0
UC0053	M	44	Pancolitis	1	1	0
UC0062	F	31	Pancolitis	4	1	0
UC0089	M	24	Procto-sigmoiditis	1	1	0

Table 2.8 *Patients with long standing pancolitis (LSPC)*

Patient No	**Sex**	**Age**	**Duration of disease**	**Baron score Rectum**	**Histological score Rectum**
UC0002	F	56	21	0	0
UC0007	F	70	42	0	0
UC0013	M	62	21	1	0
UC0017	F	51	31	1	1
UC0020	M	46	37	0	0
UC0031	M	57	27	0	0
UC0033	M	73	23	0	0
UC0034	M	62	33	0	0
UC0040	M	73	23	0	0
UC0041	M	52	31	0	0

Table 2.9 *Patients with primary sclerosing cholangitis and UC*

Patient No	Sex	Age	Duration of disease	Baron score Rectum	Histological score Rectum
UC0003	M	61	5	0	0
UC0011	M	71	17	0	0
UC0024	F	52	32	0	0
UC0058	M	56	20	0	0
UC0068	M	44	35	1	0
UC0079	F	38	16	2	0
UC0094	M	58	13	0	0

Table 2.10 *Patients with pancolitis and dysplasia*

Patient No	Sex	Age	Duration of disease	Baron score rectum	Histological score rectum
UC0036	F	52	10	0	0
UC0078	M	68	32	0	0
UC0083	F	74	41	0	0
UC0101	M	20	9	2	0

CHAPTER 3
MASS SPECTROMETRIC ANALYSIS OF INTERMEDIATE FILAMENT PROTEINS

3.1 Introduction

'Proteomics' has emerged as an important technique for large-scale analysis of proteins in biological tissues [79]. It has enabled a better understanding of expression of proteins, their structure and alterations and interactions with other proteins as well as their physiological and pathophysiological functions in biological tissues and samples (Figure 1.2 page 33) [79, 83]. Several cell line, murine and human proteomic studies have been undertaken investigating various aspects of IBD including pathogenesis of IBD (as well as colitis associated cancers) and identifying biomarkers for diagnosing IBD (and differentiating UC from CD) (Section 1.3.2, page 36).

The central technique used in all proteomic strategies is tandem mass spectrometry (MS/MS) [78]. This enables measurement of the ratio of mass to charge for ionised constituents like peptides. A variety of mass spectrometric techniques have been developed including quantitative proteomic [(isotope coded affinity tagging (ICAT), stable isotope labeling with amino acids in cell cultures (SILAC) and isobaric tag for relative and absolute quantitation (iTRAQ)], protein/antibody arrays, and multi-epitope-ligand cartograph for identifying and quantifying proteins[78]. 'Discovery proteomics' based on iTRAQ (isobaric tagging for relative and absolute quantification) employs chemical labeling of peptides and relative quantification at the level of tandem mass spectrometry (MS/MS) thus enabling protein profiling, based on the simultaneous identification and relative quantification, in parallel from multiple biological samples [78]. Sample preparation and identifying the clinical context are also critically important in most proteomic experiments.

Intermediate filament proteins are key cellular cytoskeletal components [109] with keratins and vimentin predominating in epithelial tissue [115] and of mesenchymal tissues [201, 202] respectively. Studies have shown association of keratin intermediate filament proteins in the pathogenesis of IBD [175, 180]. Altered keratin expression is also seen in CAC [185, 186] as well as sporadic CRC [203].Increased expression of vimentin in the stroma in sporadic CRC has been noted to be associated with increased malignant potential [204]. In view of the demonstrated association of keratins with IBD pathogenesis as well as a potential role of keratins and vimentin as markers

of CRC, we investigated proteomic changes in IF proteins in a well characterised group of UC patients with differing phenotype.

Following development and validation of an iTRAQ-compatible extraction and solubilisation protocol for insoluble IF proteins, labelled peptides from pooled patients were analysed by SCX-LC-MS/MS and data obtained was reconstituted using GeneBio Phenyx. All data was searched against Uniprot database (http://www.uniprot.org). Inter-group comparisons were made using in-house algorithms based on *t*-testing with multiple test correction developed previously [200]. iTRAQ analyses was run once as the results were subsequently validated orthogonally by western blotting.

3.2 Experimental design

The dissolved cytoskeletal fraction obtained from individual colon biopsy samples, were pooled into eight groups being investigated. Hundred (100) μg of protein pooled in each group was diluted with TEAB to obtain a strength of 2M GuHCl. Reduction of proteins was done by adding 50mM TCEP-HCl followed by alkylation using 10mM MMTS. After overnight trypsin digestion, iTRAQ (8-plex) labelling was done followed by SCX fractionation of peptides on a BioLC HPLC unit prior to MS/MS analyses. This technique has been described previously in the literature by Pham *et al* [182]. Reverse phase high performance liquid chromatography (RP-HPLC) on line to a QSTAR XL tandem mass spectrometer was undertaken. Steps were undertaken as described by Pham *et al* [182]. MS/MS data generated from the QSTAR® XL was first converted to generic MGF peaklists using the mascot.dll embedded script (version 1.6 release no. 25) in Analyst QS v. 1.1. Further processing of the data was undertaken using an in-house Phenyx algorithm cluster (binary version 2.6; Geneva Bioinformatics SA) at the ChELSI Institute, University of Sheffield, against the Homo sapiens UniProt protein knowledgebase (SwissProt and Trembl (41070 and 71449 entries respectively, downloaded July 2012) to derive peptide sequence and hence protein identification. These data were then searched within the reversed Homo sapiens database to estimate the false-positive rate. Peptides identifications at 1% false discovery rate were accepted. The iTRAQ reporter ion intensities were exported. Protein quantifications were obtained by computing the geometric means of the

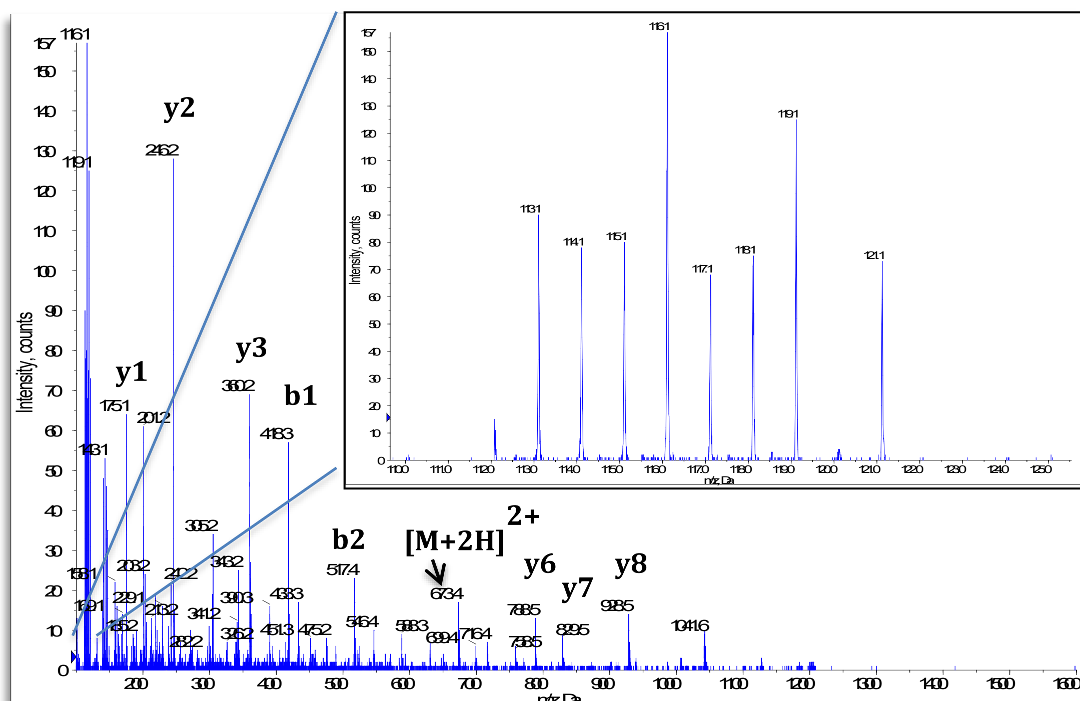
reporters' intensities. The essential steps are shown in figure 2.9(page 90).

3.3 Results

3.3.1 Protein expression profile in patient categories compared to controls on iTRAQ analysis

Forty aliquots of SCX fractions were analysed as described in the methods section (Section 2.6.2.4 page 91) [200]. Eight hundred and sixty-two (862) peptides were identified and quantified using an applied false discovery rate of 1%. This resulted in identifying 52 proteins, of which 32 proteins (with ≥ 2 quantifiable peptides) were quantified and were taken forward for further analysis. Alterations in protein level were determined by *t*-test, utilising peptide-level intensities of the iTRAQ reporter ions to determine which proteins are significantly altered. A comprehensive list of all the proteins identified is described below (table 3.1) and an example of a representative 8-plex iTRAQ labelled peptide is depicted in figure 3.1. Proteins identified included intermediate filaments K8, K18 and K19, intermediate filament associated proteins Spectrin β chain brain 4, Xin actin-binding protein 1 and KRT1B as well as histone proteins and collagen.

***Figure 3.1**. Representative 8-plex iTRAQ labeled peptide. Tandem mass spectrum generated by fragmentation on doubly charged ion of m/z 673.4. The sequence ions and iTRAQ reporter ion (inset) are shown and the y and b ion series annotated for the K19 peptide sequence IVLQIDNAR*



IVLQIDNAR

***Table 3.1:** List of proteins identified and relatively quantified in the iTRAQ proteomic datasets.*

Accession	# peptides	# valid peptide sequence	% Coverage	Score	Mass (kDa)	Description
Q96KK5	78	3	17	25.1	13775	Histone H2A type 1-H
O60814	68	8	55	60.2	13759	Histone H2B type 1-K
B4E380	62	8	41	62.0	12918	Histone H3
Q53SW3	40	1	2	7.5	52094	Putative uncharacterized protein DPYSL5
P62805	38	6	47	48.1	11236	Histone H4 [CHAIN 0]
P05787	35	11	20	95.3	53573	Keratin, type II cytoskeletal 8
P08727	31	13	25	122.0	44092	Keratin, type I cytoskeletal 19
P30989	24	1	2	5.4	46259	Neurotensin receptor type 1
Q147W7	23	1	3	11.0	37522	KRT1B protein

P05783	17	4	7	38.1	47927	Keratin, type I cytoskeletal 18
D3DP13	15	2	4	16.0	39737	Fibrinogen beta chain
P12111	15	19	6	138.3	321350	Collagen alpha-3(VI) chain
P08123	13	4	3	29.3	91755	Collagen alpha-2(I) chain
B4E241	6	1	7	9.4	14203	Splicing factor, arginine/serine-rich 3
O75084	5	1	1	6.6	60730	Frizzled-7
P11678	5	1	2	7.4	53422	Eosinophil peroxidase heavy chain
Q702N8	5	1	1	7.2	122134	Xin actin-binding repeat-containing protein 1
Q9NRC6	5	2	0	12.7	416835	Spectrin beta chain, brain 4
B3KRK8	5	4	6	37.2	46976	cDNA FLJ34494 fis, highly similar to vimentin
B4DJC3	4	1	5	9.3	21548	Histone H2A
P07355	4	2	7	16.2	38473	Annexin A2
P59665	4	2	35	16.0	3377	Neutrophil defensin 2
Q9HAM5	4	1	3	8.8	21598	cDNA FLJ11359 fis, moderately similar to HYPOXIA-INDUCIBLE FACTOR 1 ALPHA;
B4DUR6	3	1	1	7.0	76094	cDNA FLJ58493, highly similar to Centaurin-gamma 2
P12644	3	1	2	6.0	46555	Bone morphogenetic protein 4
Q68CJ9	3	1	2	7.0	44153	Processed cyclic AMP-responsive element-binding protein 3-like protein 3
Q9H9P2	3	1	3	6.4	25979	Chondrolectin
D3DWL0	3	2	1	12.2	234119	Plectin 1, intermediate filament binding protein 500kDa
Q5T1W7	2	1	3	7.5	21473	Beta-transducin repeat containing
Q6ZNC2	2	1	3	6.1	30533	cDNA FLJ16223 fis, clone CTONG3003598
B3KR37	2	1	1	5.7	55951	cDNA FLJ33617 fis, clone BRAMY2019055,

						highly similar to Epidermal growth factor receptor substrate 15
B4DVQ0	1	5	13	42.0	37349	cDNA FLJ58286, highly similar to Actin, cytoplasmic 2
C9JLJ5	1	1	5	6.1	16083	Putative uncharacterized protein TRIP12
D3DX73	1	1	1	6.2	101590	HCG1994636, isoform CRA_b
P04053	1	1	2	5.7	58309	DNA nucleotidylexotransferase
P52961	1	1	3	6.8	30538	GPI-linked NAD(P)(+)--arginine ADP-ribosyltransferase 1
Q03167	1	1	1	7.3	91316	Transforming growth factor beta receptor type 3
Q13032	1	1	2	6.7	80137	Transporter
Q14118	1	1	4	5.9	26557	Beta-dystroglycan
Q15161	1	1	5	6.1	12747	DNA-directed RNA polymerase
Q59GU6	1	1	2	6.4	50099	Sorting nexin 1 isoform a variant
Q5CZ99	1	1	1	6.8	85623	Putative uncharacterized protein DKFZp686I1370
Q5T911	1	1	4	5.3	27077	Mediator complex subunit 4
Q5TEC6	1	5	29	38.1	15430	Histone H3
Q5VZM2	1	1	2	5.7	40172	Ras-related GTP-binding protein B
Q6ZSP3	1	1	1	8.5	133571	cDNA FLJ45330 fis, highly similar to Potential phospholipid-transporting ATPase IB
Q8TE58	1	1	1	5.9	80515	A disintegrin and metalloproteinase with thrombospondin motifs 15
Q8TES4	1	2	1	12.8	153599	FLJ00119 protein
Q96Q15	1	1	0	6.8	269425	Serine/threonine-protein kinase SMG1
Q9HDB4	1	1	6	6.0	13200	Putative uncharacterized protein CATX-1

3.3.2 Longstanding colitis is associated with increased mucosal levels of cytoskeletal proteins, keratins and vimentin.

In order to control for altered total protein levels in groups, presence of protein in the dissolved samples was confirmed by protein assay (in triplicate) as described in the methods section (Section 2.6.1, page 88). Table 3.2 and figure 3.2 summarises linear fold change in IF protein levels in individual groups (eg 2.2 = 2.2 linear fold increase) compared to control patients. There is a marked increase in mucosal level of K8 (1.7x), K19 (1.7x) and vimentin (2.2x) in the rectal mucosa of patients with LSPC with quiescent disease. In contrast patients with ROUC, despite having macroscopically and microscopically quiescent disease, demonstrated significantly reduced keratin levels compared to controls (K8, K18, K19 and VIM 0.3, 0.5, 0.4 and 0.5 fold respectively, all $p < 0.05$).

In PSC there is a significant reduction in keratin and vimentin levels (0.8, 0.7, 0.8 and 0.9 fold for K8, K18 and K19 and VIM respectively; $p < 0.05$. Analysis of dysplastic tissue (DT) and field rectal mucosa (DR) (in patients with dysplasia) also showed a similar pattern of changes in DR (0.8, 0.7, 0.8 fold for K8, K18 and K19) and DT (0.5, 0.7 and 0.4 fold for K8, K18 and K19) relative to controls.

In order to investigate effects of disease duration on IF proteins, fold changes in proteins in LSPC patients were compared with ROUC patients (figure 3.3). In comparison to ROUC, a significant increase in levels of IFs (4.3, 1.2, 4 and 3.5 fold for K8, K18, K19 and VIM respectively; $p < 0.05$) and IF associated proteins Xin and Spectrin was noted in LSPC. These results indicated that

longstanding quiescent colitis is associated with an increased mucosal level of keratins, vimentin and IF associated proteins.

Figure 3.2 Results of mass spectrometry analysis showing significant fold changes in protein levels in the patient groups relative to controls. *IF extracts from colorectal biopsies were subject to proteomic analysis. Mass spectrometry results showing significant fold changes (p<0.05) of IF proteins between different patient groups (LSPC, ROUC, and PSC and those with dysplasia-DR and DT respectively) in comparison to control. . LSPC shows a significant increase in levels of K8, K19 and VIM; in ROUC, PSC, DR and DT the levels of K8, K18 and K19 are significantly decreased.*

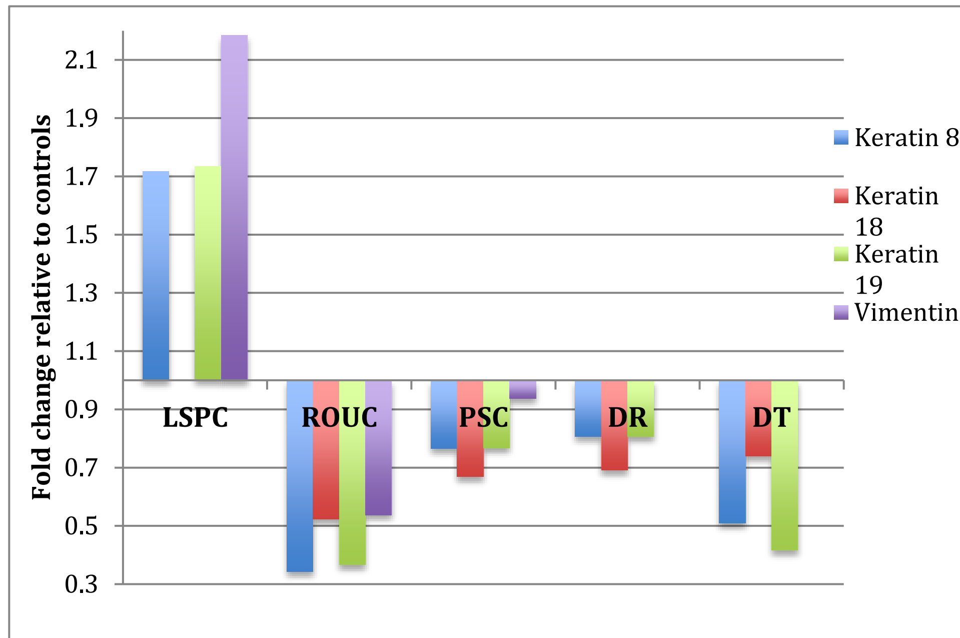


Figure 3.3 Effect of disease duration on mucosal levels of IFs, comparing patients with LSPC to ROUC. *Cytoskeletal protein extracts from colorectal biopsies were subject to proteomic analysis. Mass spectrometry results are shown describing significant fold changes (p<0.05) of IF proteins and associated proteins The graph shows an samples exhibited a marked increase in all IF proteins and IF associated proteins (Spectrin, Xin) in LSPC compared to ROUC.*

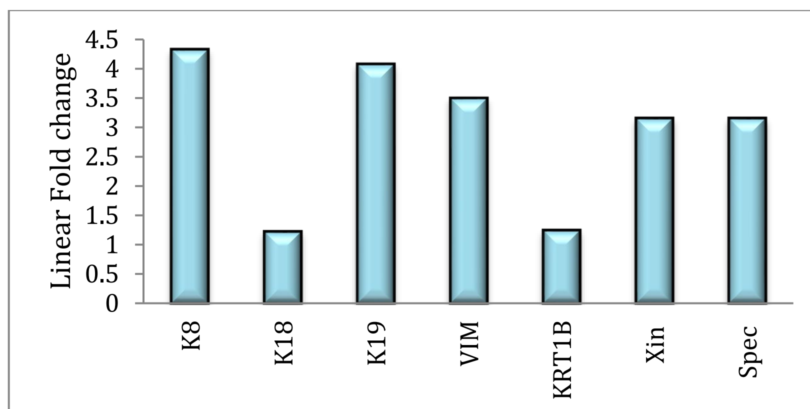


Table 3.2 *Mass spectrometry results summarising fold changes(linear) in levels of key IF proteins and associated proteins between different patient groups. Fold change in all cases are compared to control samples. Multiple test correction (Bonferroni) is applied to the data. Fold changes shown have reached at least P=0.05, and higher significance thresholds where indicated.*
(*¶ p<0.001 NS- not significant* K8- Keratin8, K18- Keratin 18, K19- Keratin 19, VIM- Vimentin, Spec- Spectrin)

Protein	Accession no	Gene symbol	MW(Da)	Peptide no	Fold change in levels (p<0.05 unless stated)						
					Control vs					LSPC vs	DT vs
					LSPC	ROUC	PSC	DR	DT	ROUC	DR
K8	P05787	KRT8	53704	35	1.7	0.3	0.8	0.8	0.5	4.3	0.7
K18	P05783	KRT18	48058	17	NS	0.5	0.7	0.7	0.7	1.2	NS
K19	P08727	KRT19	44106	31	1.7	0.4	0.8	0.8	0.4	4.1	0.6
VIM	B3KRK8	VIM	46976	4	2.2¶	0.5	0.9¶	NS	NS	3.5	NS
KRT1B protein	Q147W7	KRT1B	37522	23	NS	0.6	NS	0.8¶	0.7	1.2	NS
Xin	Q702N8	XIRP1	198561	5	0.7¶	0.2	NS	NS	0.4¶	1.2	0.5
Spectrin	Q9NRC6	SPTBN5	416835	5	0.7¶	0.2	NS	NS	0.4¶	3.2	0.5

3.3.3 Keratins 8, 18 and 19 are reduced in primary sclerosing cholangitis and in dysplasia

In this study we evaluated proteomic changes not only in the dysplastic tissue (DT) but also in the rectal mucosa to identify any field changes in the IF proteome. Dysplastic tissue (DT) revealed a significant reduction in levels of K8, K18, K19 as well as KRT1B, Xin and Spectrin compared to controls (figure 3.2). Interestingly rectal mucosa (DR) in patients with dysplasia too showed reduced levels, though less than in the dysplastic tissues themselves. Such changes suggest a possible pan colonic change in IF proteome in patients with dysplasia.

When compared to LSPC, patients with dysplasia showed a reduction in levels of K8, K19 and VIM ($p<0.05$) (figure 3.4). This was more pronounced in the dysplastic tissue, which showed a significant reduction in K8, K19, Xin and Spectrin compared to the field mucosa (figure 3.5). In PSC group too there was a significant reduction in keratin levels compared to controls (figure 3.2). In comparison to LSPC patients, PSC showed a significant reduction($p<0.05$) in levels of K18, K18 and K19 (figure 3.4). Interestingly VIM levels were increased in this high-risk group compared to LSPC.

Figure 3.4 Proteomic changes in patients with PSC and dysplasia compared to LSPC group. *Cytoskeletal protein extracts from colorectal biopsies were subject to proteomic analysis. Mass spectrometry results are shown describing significant fold changes ($p<0.05$) of IF proteins and associated proteins There is a significant reduction ($p<0.05$) in most IF protein levels compared to LSPC in all 3 groups. Directionality of VIM alteration distinguishes PSC group from DR and DT.*

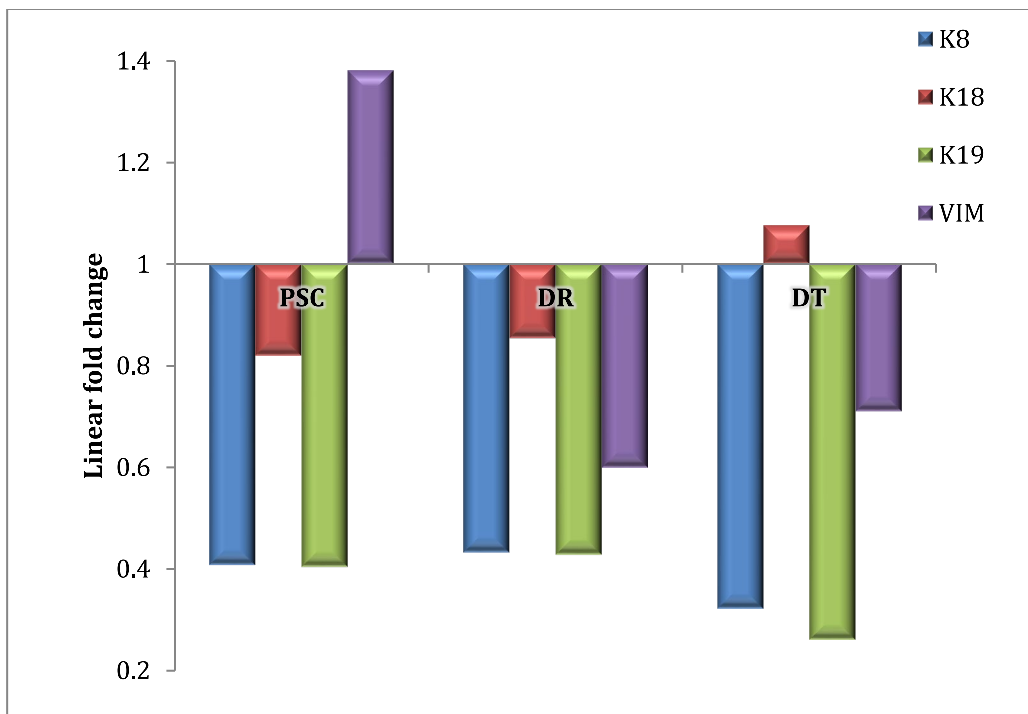
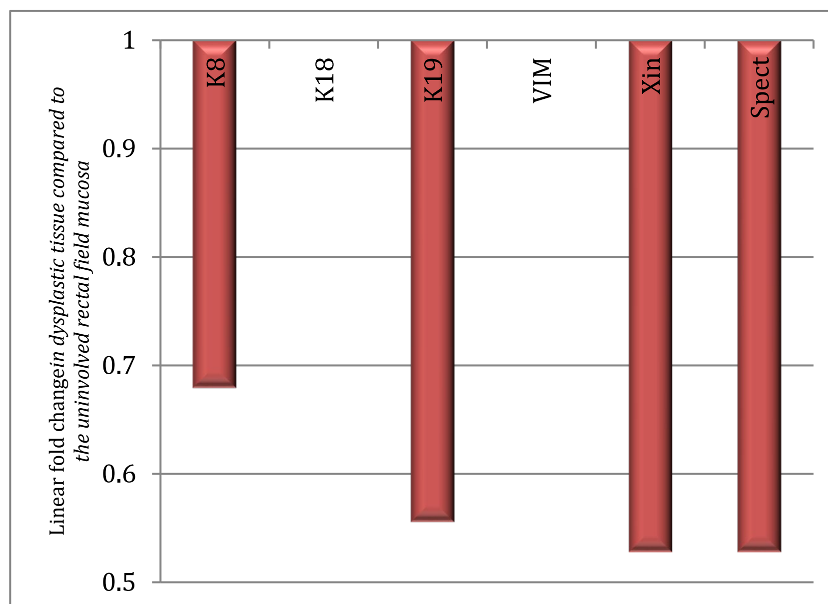


Figure 3.5 Proteomic changes in dysplastic tissue (DT) compared to non-dysplastic rectal field mucosa (DR). *Cytoskeletal protein extracts from colorectal biopsies were subject to proteomic analysis. Mass spectrometry results are shown describing significant fold changes (p<0.05) of IF proteins and associated proteins. There is significant reduction (p<0.05) in K8, K19, Xin and Spectrin levels in dysplastic tissue compared to the uninvolved rectal field mucosa.*



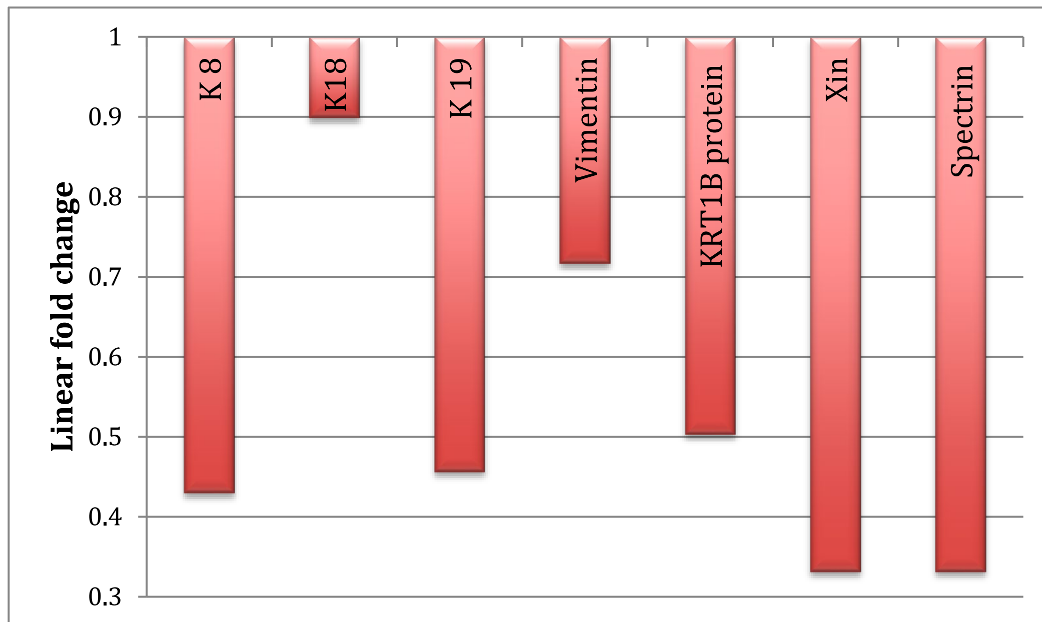
3.3.4 Acute inflammation in the colorectal mucosa is associated with reduced levels of mucosal cytoskeletal proteins

In order to identify the effect of inflammation on the IF proteins in the colonic mucosa, we investigated a group of patients with active distal colitis (ACT) who had normal proximal colonic mucosa, both histologically and endoscopically (INACT). In this group, mass spectrometry revealed significantly decreased fold change in levels of K8 (0.4 fold), K18 (0.7 fold), K19 (0.4 fold) and VIM (0.6 fold) (all p<0.05) in inflamed mucosa in comparison to controls as well as to the un-involved proximal colon (Table 3.3) (figure 3.6 a & b). Thirteen proteins in the insoluble protein fraction were unchanged between the patient samples compared; levels of two cytokine proteins, neutrophil defensin 1, and bone morphogenetic protein 4 were increased only in actively inflamed segment. As an experimental control, other proteins were also noted raised in the inflamed mucosa (H2A, ND1 and EP). The endoscopically and histologically un-inflamed mucosa did not show any significant difference in protein levels from the mucosa of controls (Table 3.3).

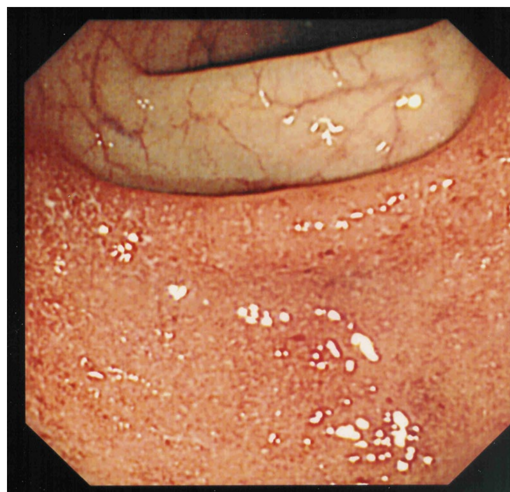
Table 3.3 *Mass spectrometry results showing significant fold changes (applying quantification stringency of MTC on, p<0.05) in levels of intermediate filament proteins and associated proteins in acutely inflamed mucosa compared to controls. ¶ p<0.001, NS- not significant*

Protein	Accession no.	Gene symbol	MW (Da)	Peptide no	Fold change in levels-(p<0.05 unless stated)		
					Control versus		**ACT vs**
					ACT	**INACT**	**INACT**
K8	P05787	KRT8	53,704	35	0.4	NS	0.4
K18	P05783	KRT18	48,058	17	0.7	NS	0.9
K19	P08727	KRT19	44,106	31	0.4	NS	0.5
VIM	B3KRK8	VIM	46,976	4	0.6¶	NS	0.7
KRT1B protein	Q147W7	KRT1B	37,522	23	0.4	NS	0.5
Xin actin-binding protein 1	Q702N8	XIRP1	198,561	5	0.2¶	NS	0.3
Spectrin β chain brain 4	Q9NRC6	SPTBN5	416,835	5	0.2¶	NS	0.3

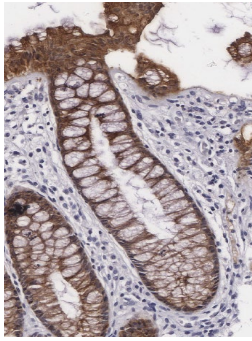
Figure 3.6 (a) iTRAQ data showing alteration in mucosal intermediate filament protein levels in acute inflammation compared to proximal un-inflamed mucosa. *Cytoskeletal protein extracts from colorectal biopsies were subject to proteomic analysis. Mass spectrometry results are shown describing significant fold changes (p<0.05) of IF proteins and associated proteins comparing colonic mucosa in acute inflammation to proximal un-inflamed mucosa (p<0.05). There is a significant decrease in (p<0.05) in K8, K18, K19, vimentin, KRT1B, Xin and Spectrin levels in inflamed mucosa.*



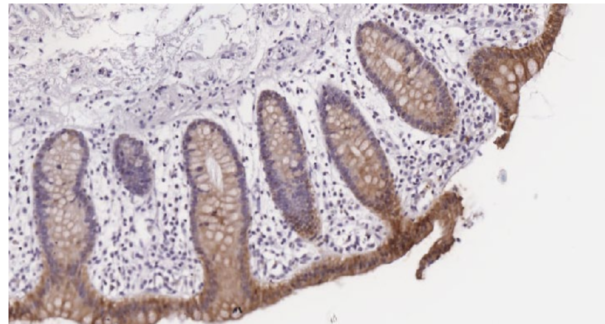
(b) *Figure shows a representative colonic mucosa showing inflamed and un-inflamed segments in the colonic mucosa of the same patient.*



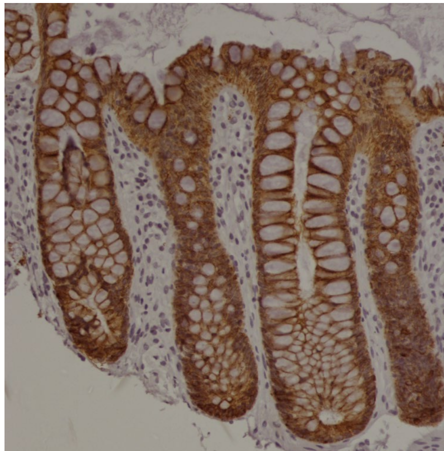
(c) Immunohistochemical staining showing normal distribution of K8, K18, K9 in the colonic crypts and VIM in the mesenchymal tissue in normal colonic mucosa of control patients. Following deparafinisation, rehydration and antigen retrieval of the tissue sections, they were treated with the appropriate primary (K8, K18, K19 and vimentin) and secondary antibodies (Biotinylated Goat Anti-Mouse IgG antibody).



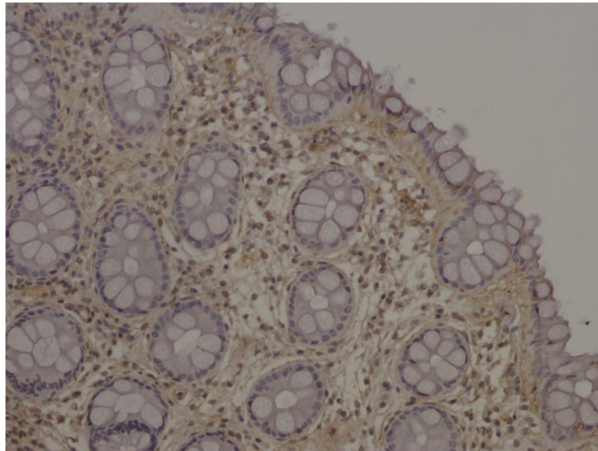
K8



K18

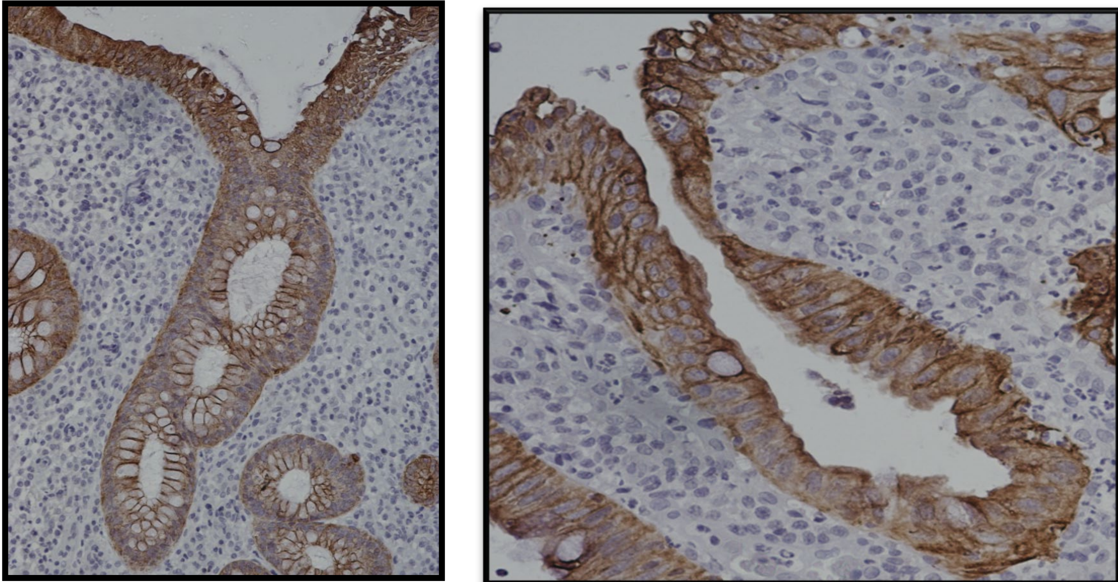


K19



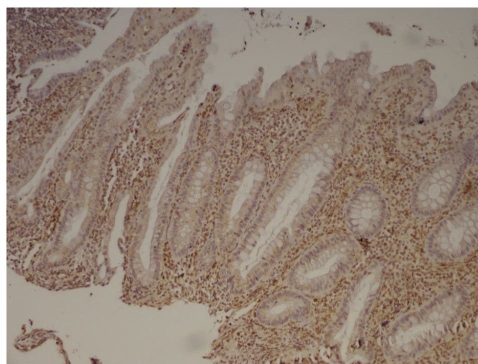
VIM

d) Figure showing immunohistochemical staining of the colonic crypts with K8 in patients with active inflammation and quiescent disease. Following deparafinisation, rehydration and antigen retrieval of the tissue sections, they were treated with the appropriate primary (K8, K18, K19 and vimentin) and secondary antibodies (Biotinylated Goat Anti-Mouse IgG antibody).



Inactive Inflammation K8 **Active Inflammation K8**

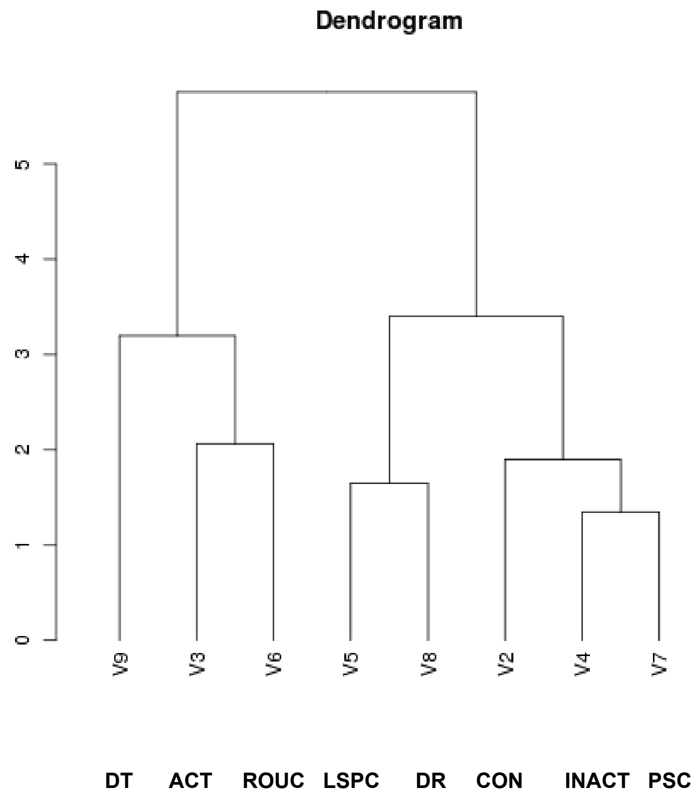
e) Figure showing mesenchymal vimentin distribution on immunohistochemical staining in a patient with dysplasia. Following deparafinisation, rehydration and antigen retrieval of the tissue sections, they were treated with the appropriate primary (vimentin) and secondary antibodies (Biotinylated Goat Anti-Mouse IgG antibody).



3.3.5 Hierarchical cluster analysis of patient groups based on identified peptides

Hierarchical clustering was performed in parallel to the analyses of the *iTRAQ* data by other researchers in the group. This is based on similarity of identified peptides to group the patients into similar categories based on their proteomic characteristics. Agglomerative clustering used a ‘bottom up’ approach with squared Euclidean distance between $\log_{10}$ value of iTRAQ ratios and smallest intercluster dissimilarity linkage procedure was performed (Mathematica 7.0.0 for Mac), to generate the dendrogram shown in figure 3.7. The dendrogram demonstrates the following clustering into two broad groups in terms of proteomic similarity– DT, ACT and ROUC; CON, INACT and PSC, LSPC and DR (figure 3.7). A clear cut pattern is not seen in the proteomic characteristics in the different phenotypic sub clusters within each group. This could be explained by the fact that the number of proteins in the IF fraction is much smaller than would be obtained from whole cell or soluble proteome analysis, but there is potential for trend data to emerge. Additionally, proteins other than insoluble IF proteins were also identified in the analyses which could have influenced the results. Nevertheless, a similarity in proteomic characteristics was noted between ACT and ROUC, changes which reflect the subsequent western blotting experiment and results from a possible failure of restoration of keratins in the ROUC group despite apparent mucosal healing.

Figure 3.7 *Dendogram demonstrating clustering of patient groups (V2 CON, V3 ACT, V4 INACT, V5 LSPC, V6 ROUC, V7 PSC, V8 DR, V9 DT) based on proteomic characteristics. Similarities in proteomic results, with clustering are noted between DT, ACT and ROUC.*



3.4 Conclusion

In this study, using a quantitative proteomic approach with an *iTRAQ* based proteomic workflow we have identified characteristic pattern of alterations in insoluble IF proteins in different phenotypic groups of UC patients with differing colon cancer risk. Rectal mucosa of patients with quiescent long-standing colitis shows a significant increase in levels of IFs compared to controls. The data thus suggests an influence of disease duration on mucosal level of IF proteins. Although studies have demonstrated duration of disease as a significant risk factor for CAC [22], in our study, other high risk group (PSC) shows a reduced level of IF proteins; similar reduction is also noted in dysplastic tissue.

Acute inflammation is associated with significant reduction in levels of IF proteins and IF associated proteins (Xin and Spectrin). Additionally, patients with ROUC, despite having macroscopically and microscopically quiescent disease, demonstrated significantly reduced keratin levels in rectal mucosa compared to controls. The later changes suggest a lag in recovery of inflammation related changes in IF proteins despite apparent endoscopic and microscopic healing. Taken together the changes noted in acutely inflamed mucosa, along with those in `high-risk phenotype point towards a central role of inflammation in the pathogenesis of CAC. Role of mucosal inflammation (both endoscopic and histological) in pathogenesis of CAC is now well recognised [47, 48] and our findings supports those previous studies. The raised keratin levels in LSPC, on the other hand could suggest a protective role of raised keratins in the mucosa.

Dysplastic mucosa, and (un-involved) field tissue around dysplasia showed a general reduction in IF protein levels. These changes in the rectal biopsies in patients with dysplasia point towards a possible pan-colonic field change in IF proteome. This is in keeping with previous studies demonstrating a molecular field defect throughout the colonic mucosa [70, 71, 205]. Colonic field effects have also been noted in the context of sporadic CRC and adenomatous polyps [206]. A separate cross-sectional study from our group, investigating proteomic changes in colonic mucosa of patients with adenomatous polyps

have shown reduced K8, K18 and K19 expression. These levels reduce with progressive proximity to the lesion [207]. Taken together with our study results it supports a potential use of keratin alteration in the colonic mucosa as biomarkers of CACs.

Overall these changes identified by quantitative proteomics need further characterisation and validation including investigating for PTMs. This may help obtain further insight into the pathogenesis of colitis associated cancer

.

CHAPTER 4

VALIDATION OF ITRAQ RESULTS BY WESTERN BLOTTING AND DEMONSTRATION OF POST-TRANSLATIONAL MODIFICATIONS IN KERATIN

4.1 Introduction

Ulcerative colitis is a chronic inflammatory disease process of the colon associated with an increased risk of of colitis-associated cancer [2, 19]. A variety of factors increase the risk, including duration of disease since diagnosis. Meta-analysis has demonstrated a cumulative risk of CAC of 2% at 10 years, 8% at 20 years and 18% at 30 years [22]. Other important factors include extent of disease in the colon [29, 30], family history of CRC [34], associated PSC [53, 54] and active inflammation in the colon, both endoscopic and on histological examination [47]. Current surveillance strategies for identifying patients with CAC involving colonoscopy and biopsies are limited in their efficacy due to the non targeted nature of such investigations [65]. As genomic abnormalities are noted elsewhere in the colon in patients with dysplastic and neoplastic changes a pan-colonic biomarker could potentially enable diagnose such patients with limited rectal/distal colonic biopsies.

‘Proteomics’ the study of expression, alterations and interactions of proteins has emerged as a technique for analysing proteins in biological tissues [79, 83]. This is because proteins not only constitute the structural elements in tissues but also act as messengers of various signalling pathways[81].

Intermediate filament proteins are key components of the cellular cytoskeleton [109] with keratins and vimentin predominating in tissues of epithelial [115] and mesenchymal origin [201, 202] respectively. Keratin predominantly is in an insoluble filamentous form in cells, although solubility is affected by a variety of factors including cellular stress [120] and post-translational modifications. Recent studies have shown an association of keratin IF proteins in the pathogenesis of IBD. In an animal study, homozygous mK8-/-FVB/N mice developed colonic hyperplasia with a histology showing inflammation [173]. In another study, heterozygous missense mutation in K8 gene has also been noted in 5.2% of IBD patients [180]. Altered keratin expression in colonic mucosa on immunohistochemistry has been noted in patients CAC[185, 186]. Additionally differential expression of K8 has been shown in polyps and in the surrounding field mucosa in patients with sporadic CRC [206]. Vimentin is a type III IF protein, present in cells of mesenchymal

origin [208]. It is now regarded as a marker of epithelial- mesenchymal transformation. High VIM expression in CRC tumour stroma may reflect a higher malignant potential of the tumour[204].

Following translation of proteins, they undergo further changes in tissues. Such changes are termed post-translational modifications (PTMs) and help to modulate protein functions not encoded for by genes [123]. Keratins too demonstrate wide variety of PTMs including phosphorylation[126, 128, 138], acetylation [132], glycosylation [133], sumoylation [136] and cleavage [134].

In this study we have demonstrated using a quantitative mass spectrometric strategy, several key changes in IF proteins in the mucosa. analysis of insoluble IF proteins in mucosa of patients with UC.

Using an *iTRAQ* based quantitative proteomic approach using colonic biopsies from patients with UC, we have shown alteration in IF proteins. LSPC patients show elevated mucosal levels of keratins, whereas inflammation is associated with reduced levels in the inflamed mucosa only. Similar changes are also seen in dysplasia and in rectal mucosa of those with dysplasia. Hence in this study we investigated those changes using western blot to validate the results noted on *iTRAQ* analysis and to investigate for PTMs.

4.2 Aims and experimental design

In this study we aimed to validate the results of the *iTRAQ* analysis using WB. In addition, we aimed to investigate for PTMs primarily phosphorylation and acetylation using specific antibodies. Details of the steps taken are mentioned in in section 2.7.2 (page 95). Briefly, following elution of the proteins on SDS-PAGE gel, they were transferred onto a membrane followed by probing with primary antibodies to K8, K18, K19, VIM, phosphorylated K8 (K8-pS23, K8-pS73, K8-pS431) and acetylated K8. Densitometric analysis of the band intensity was undertaken as described previously (section 2.7.4, page 95). All WB experiments were run twice.

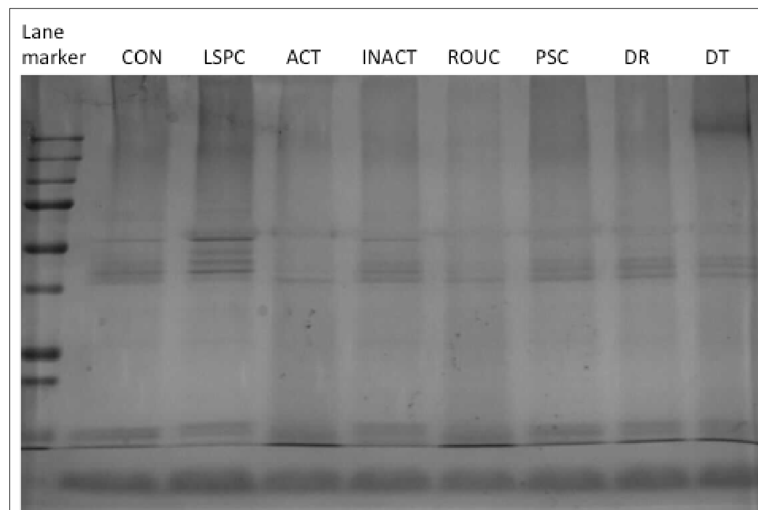
4.3 Results

4.3.1 Validation of *iTRAQ* results by western blotting and demonstration of post-translational modifications in keratins

Orthogonal validation of the *iTRAQ* results was undertaken by western immunoblotting using antibodies described previously (section 2.2.1.1 page

61). An initial Coomassie staining of a gel was undertaken to visualise the loaded gels. Results of the Coomassie staining are shown in figure 4.1.

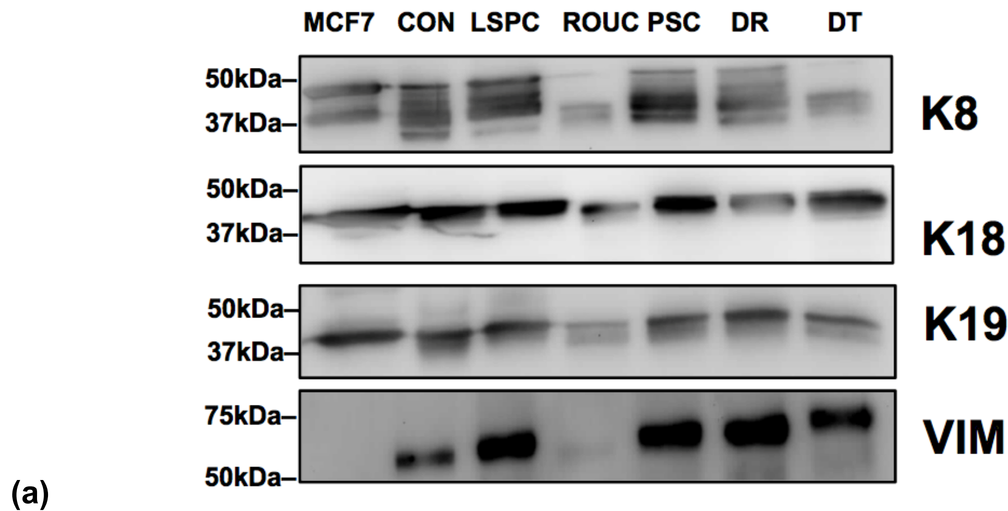
Figure 4.1 Coomassie stained gel to visualise the protein bands and ensure adequate loading of the gel with samples*. Extracted and solubilised cytoskeletal protein samples from individual patients were pooled into 8 groups. 5µG protein loaded into each lane from each group was loaded onto a SDS-PAGE gel followed by staining with Coomassie stain. This shows adequate visualisation of protein bands in each group.*



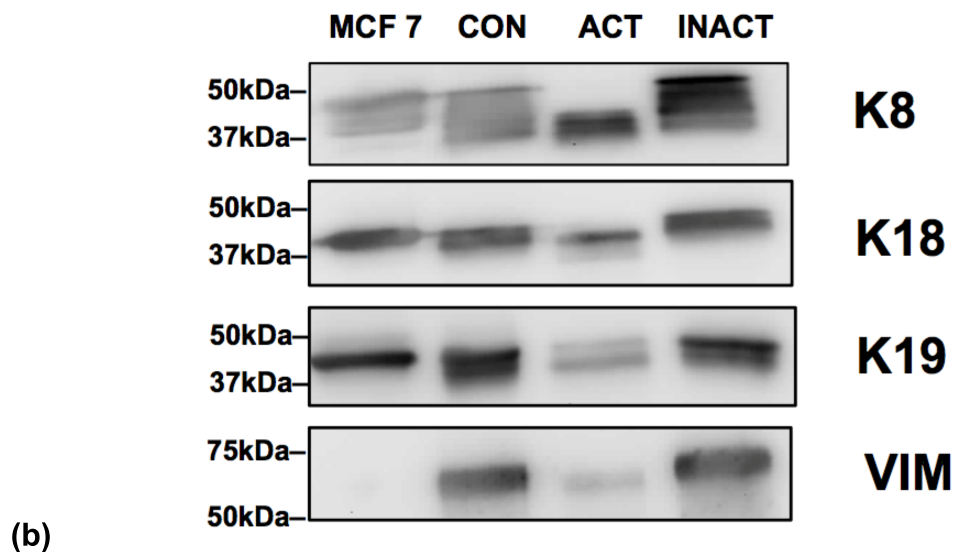
4.3.2 Intermediate filament proteins show different levels of expression in health and in disease states (figure 4.2 a and b)

Keratin 8 levels are reduced in patients with quiescent recent onset UC (ROUC) compared to controls. Similar changes are also noted in the DT group. In ROUC, K19 and more markedly, VIM levels are reduced. In comparison to controls, VIM levels were however increased in LSPC, PSC, DR and DT. In the acutely inflamed mucosa (ACT), there was similar to iTRAQ results, a reduction in K8, K18 and K19 levels in the inflamed mucosa compared to controls and the proximal healthy mucosa (figure 4.2 b).

Figure 4.2. Results of western immunoblotting with antibody to K8, K18, K19 and vimentin. *The extracted cytoskeletal fraction from recruited patients was dissolved in GuHCl followed by removal of solvent, gel electrophoresis on SDS-PAGE gels, western transfer and immunoprobing using the following antibodies- mouse monoclonal antibody to K8, K18, K19 and VIM. Panel (a) shows results of WB in LSPC, ROUC, PSC and in patients with dysplasia (DT and DR) and Panel (b) shows WB results in patients with active distal disease compared to controls and proximal inactive mucosa. 5μG protein was loaded onto each lane. Multiple isoforms of K8 are noted in all the patient groups including healthy controls with lower levels and low molecular weight forms predominating in ROUC and DT. K8 levels are reduced in ROUC and in DT. In acute inflammation there is a reduction in total insoluble IF protein levels compared to CON and INACT, with K8 being predominantly in the lower molecular weight forms (37kDa) in inflamed tissue. K18, K19 and VIM are also reduced in acutely inflamed mucosa*



(a)

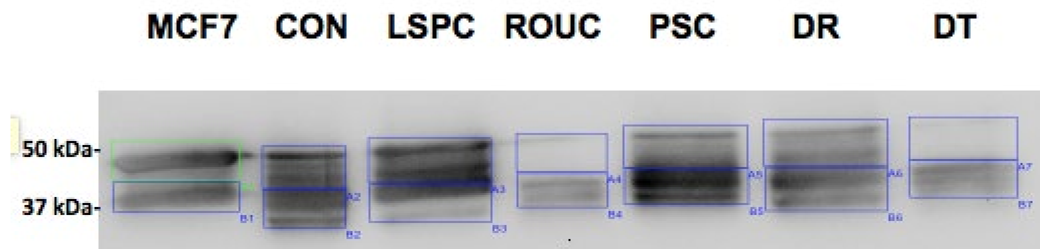


(b)

4.3.3 Keratin 8 exists in multiple isoforms in controls as well as patients with UC

In all groups K8 was noted to be present in multiple isoforms corresponding to several discrete bands on WB (figure 4.3). These bands were noted between the molecular weights 37-50 kDa. This could indicate post-translational modifications in K8, including in normal healthy colonic mucosa. In contrast to controls, the isoforms in LSPC were evident at higher molecular weights whereas in PSC and DR they tended to predominate at lower molecular weights. In DT and ROUC, there was overall reduction in K8 level, which too was predominant in lower MW forms. In ACT the expressed K8 again was present in lower MWs.

***Figure 4.3 Western immunoblotting using K8 antibody demonstrating multiple isoforms of K8, including in control patients**. The extracted cytoskeletal fraction from recruited patients was dissolved in GuHCl followed by removal of solvent, gel electrophoresis on SDS-PAGE gels, western transfer and immunoprobing using the following antibodies- mouse monoclonal antibody to K8. 5µG protein was loaded onto each lane. Isoforms in LSPC are predominant in higher molecular weight.*



4.3.4 Post-translational modification in Keratin 8 (phosphorylation and acetylation) occurs in colonic mucosa

In order to investigate origins of the K8 isoforms, we undertook further WB to identify common PTMs in K8 in the form of phosphorylation at pS23, pS431 and pS73 as well as acetylation at lys10 and lys482 respectively.

4.3.4.1 Phosphorylation in Keratin 8 occurs in healthy colonic mucosa as well in mucosa of patients with ulcerative colitis

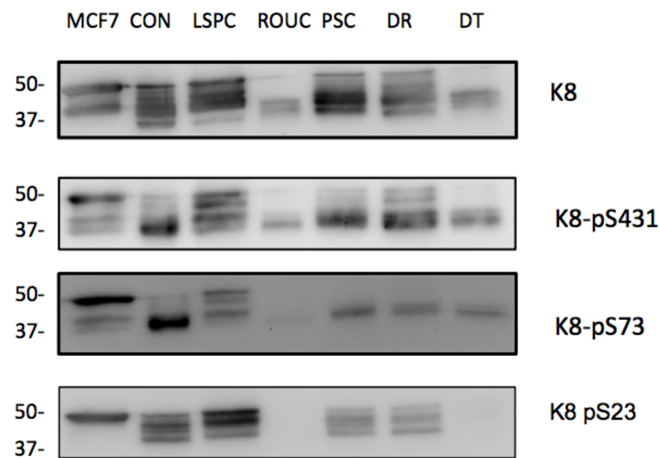
We investigated changes in K8 phosphorylation at three common phosphorylation sites, Ser 23, Ser 73 and Ser 431. The results are shown in figure 4.4 (a & b). Intense phosphorylation at Ser73 was noted at 37kDa in healthy controls, similar phosphorylation in controls at lower molecular weight (37kDa) was noted at Ser 431, while diffuse phosphorylation was noted at K8 Ser 23. These changes demonstrate physiological phosphorylation of K8 at multiple sites in the keratin molecule in healthy mucosa.

In contrast to intense phosphorylation in K8 Ser 73 in controls, other patient groups also showed altered phosphorylation. LSPC patients showed phosphorylated K8 at higher molecular weights (predominantly with pS431 and pS23). With pS431 there were multiple phosphorylated isoforms of K8 in LSPC. Scant phosphorylation in keeping with reduced total K8 levels was noted in ROUC. PSC and DR showed lower levels of total phosphorylation and lower molecular weight forms with pS431. In DT there was absence of phosphorylation at S23, similar to S73.

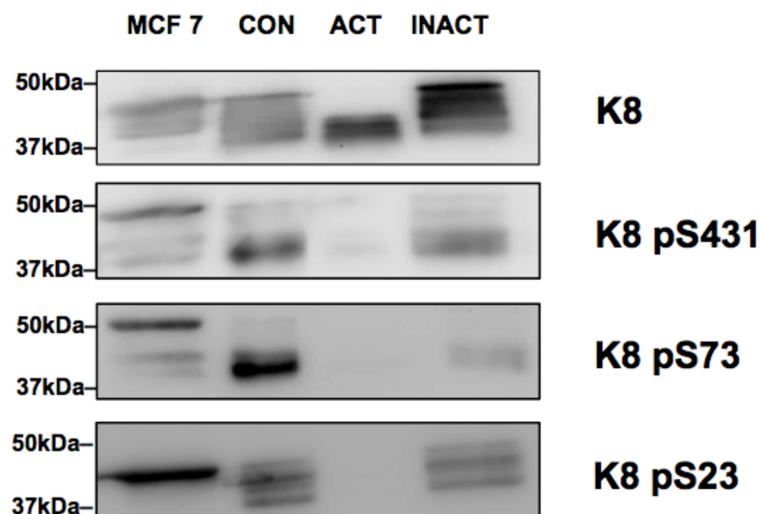
Active inflammation results in scant or absent phosphorylation at all three phosphorylation sites (figure 4.4 b) suggesting potential dephosphorylation of keratin 8 molecule in presence of inflammation.

Figure 4.4 Western immunoblotting using antibodies to phosphorylated K8. *The extracted cytoskeletal fraction from recruited patients was resolved by gel electrophoresis on SDS-PAGE gels, followed by western transfer and immunoprobing using the following antibodies to phosphorylated K8- K8-pS23, K8-pS73 and K8-pS431. Panel (a) shows results of WB in LSPC, ROUC, PSC and in patients with dysplasia (DT and DR) and Panel (b) shows WB results in patients with active distal disease compared to controls and proximal inactive mucosa. 5μG protein was loaded onto each lane. Panel a shows intense phosphorylation in healthy control mucosa at lower molecular weight (predominantly at a single band around 37kDa) with K8-pS73 and K8-pS431. Phospho Ser 23 antibody cross-reacts with a more diverse set of molecular weights. Phosphorylated K8 is reduced in ROUC and DT (Panel A) and in acute inflammation (Panel B).*

(a)



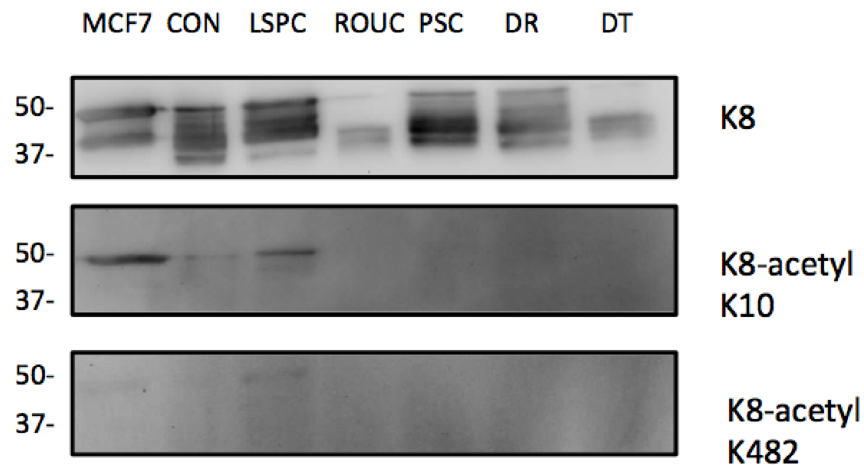
(b)



4.3.4.2 Acetylation of K8 is noted only in LSPC

Acetylation has previously been demonstrated as an important post-translational modification of K8 (section 1.4.2, page 47). We undertook WB using anti acetyl lysine antibodies to lys10 and lys 482 of K10. Although there was no evidence of acetylation at K482, LSPC patients only demonstrated the PTM when probed with K10 antibody (figure 4.5).

Figure 4.5. Western immunoblotting demonstrating acetylation of K8. *The extracted cytoskeletal fraction from recruited patients was resolved by gel electrophoresis on SDS-PAGE gels, followed by western transfer and immunoprobing using the following antibodies to acetylated K8- K8-acetyl K10 and K8-acetyl K482. Acetylated K8 cross reacting to - K8-acetyl K10 is noted in LSPC group only. 5μG protein was loaded onto each lane.*



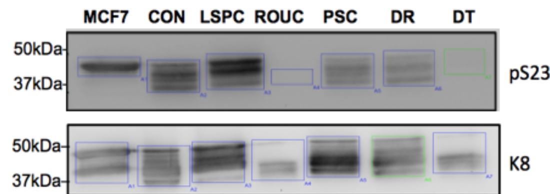
4.3.5. Densitometry analysis of keratin 8 phosphorylation in comparison to total K8 levels identifies distinct phenotypic patterns.

Densitometry was undertaken for imaging the membranes in order to obtain a quantitative assessment of changes in relative phosphorylation (Figures 4.6-4.9). Individual bands of interest were selected manually with Chemigenius Bio-Imaging System and their pixel densities obtained from the imaging software (Gene Tools).

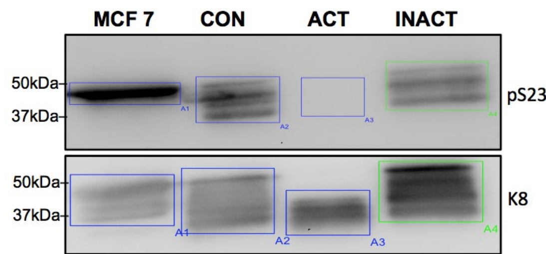
K8 phosphorylation is reduced or absent in ROUC. Diminished phosphorylation levels are also noted in PSC, DR and DT compared to LSPC. In general, phosphorylated K8 is reduced or absent (K8pS23) in presence of

acute inflammation. The relative phosphorylation at sites S431, S23 and S73 in INACT, ACT and control tissues was also evaluated by western immunoblot and relative densitometry was done to a control loading of MCF-7 IF extract. The data suggests a progressive dephosphorylation of K8 in association with inflammatory status. K8pS23 levels relative to total K8 are comparable in CON and LSPC while a reduced ratio is seen in ROUC & DT.

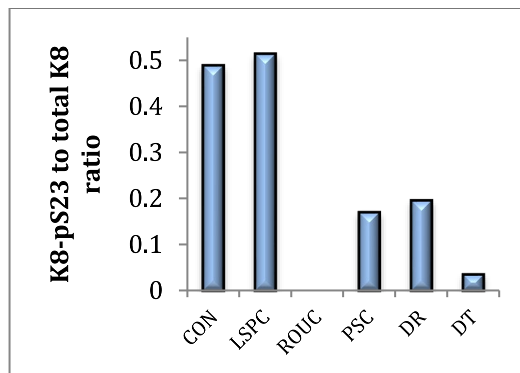
Figure 4.6 Densitometry analysis comparing phosphorylated K8 (pS23) to total K8. *The extracted cytoskeletal fraction from recruited patients (in pooled groups) was resolved by gel electrophoresis on SDS-PAGE gels, followed by western transfer and immunoprobing using the following antibody to phosphorylated K8- K8-pS23. 5μG protein was loaded onto each lane. Densitometric analysis of the band intensity was undertaken. Panel (a and c) shows results of WB and densitometry ratios in LSPC, ROUC, PSC and in patients with dysplasia (DT and DR) and Panel (b and d) shows WB results and densitometry results in patients with active distal disease compared to controls and proximal inactive mucosa. Similar changes in the ratio are noted in LSPC and controls with reduced ratios in ACT, ROUC and DT.*



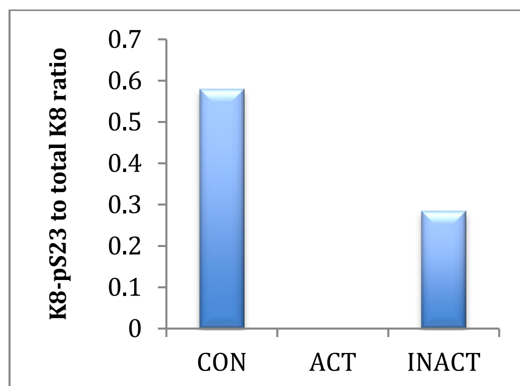
(a)



(b)



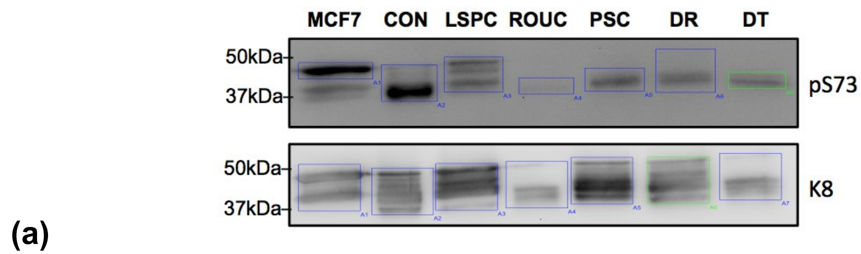
(c)



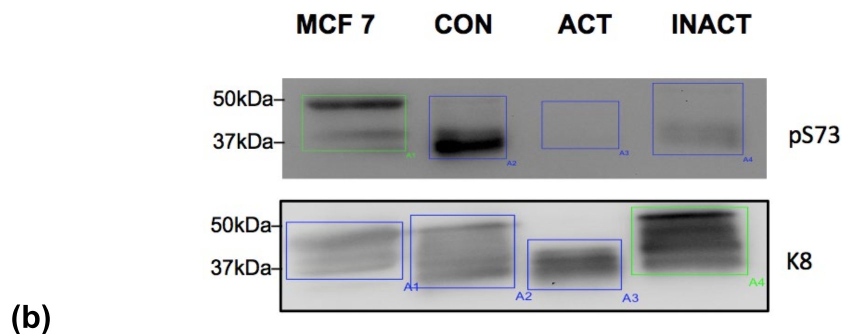
(d)

Figure 4.7 *Densitometry analysis comparing phosphorylated K8(pS73) to total K8. The extracted cytoskeletal fraction from recruited patients (in pooled groups) was resolved by gel electrophoresis on SDS-PAGE gels, followed by western transfer and immunoprobing using the following antibody to phosphorylated K8- K8-pS73. 5μG protein was loaded onto each lane. Densitometric analysis of the band intensity was undertaken. Panel (a and c) shows results of WB and densitometry ratios in LSPC, ROUC, PSC and in patients with dysplasia (DT and DR) and Panel (b and d) shows WB results and densitometry results in patients with active distal disease compared to controls and proximal inactive mucosa.*

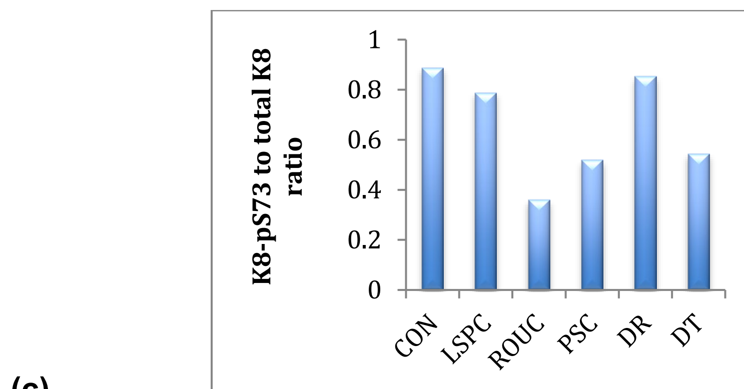
(a)



(b)



(c)



(d)

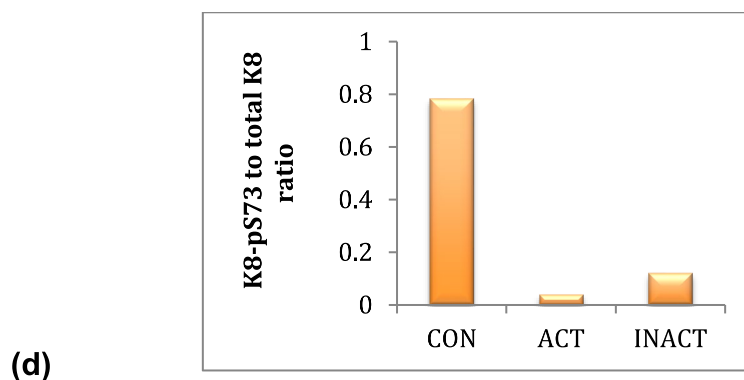
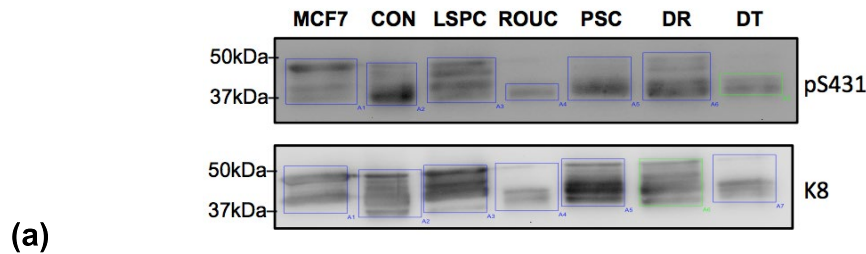
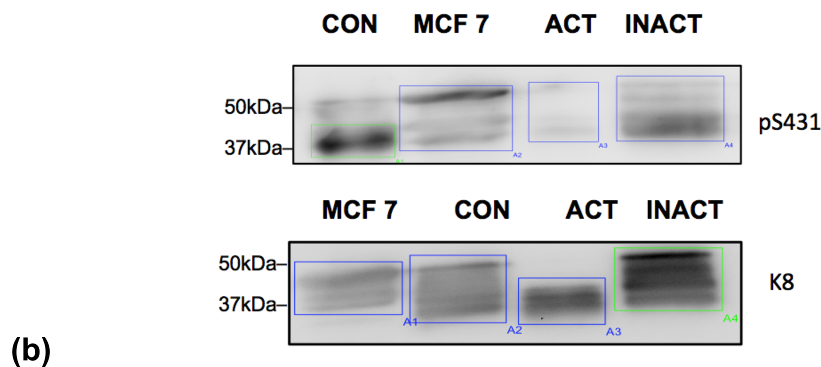


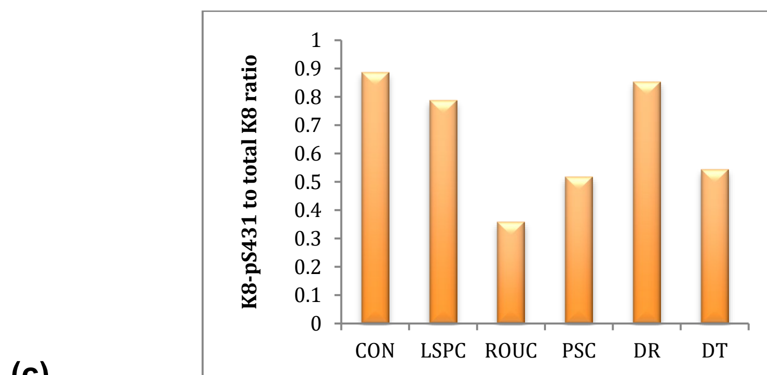
Figure 4.8 *Densitometry analysis comparing phosphorylated K8 (pS431) to total K8. The extracted cytoskeletal fraction from recruited patients (in pooled groups) was resolved by gel electrophoresis on SDS-PAGE gels, followed by western transfer and immunoprobing using the following antibody to phosphorylated K8- K8-pS431. 5µG protein was loaded onto each lane. Densitometric analysis of the band intensity was undertaken. Panel (a and c) shows results of WB and densitometry ratios in LSPC, ROUC, PSC and in patients with dysplasia (DT and DR) and Panel (b and d) shows WB results and densitometry results in patients with active distal disease compared to controls and proximal inactive mucosa.*



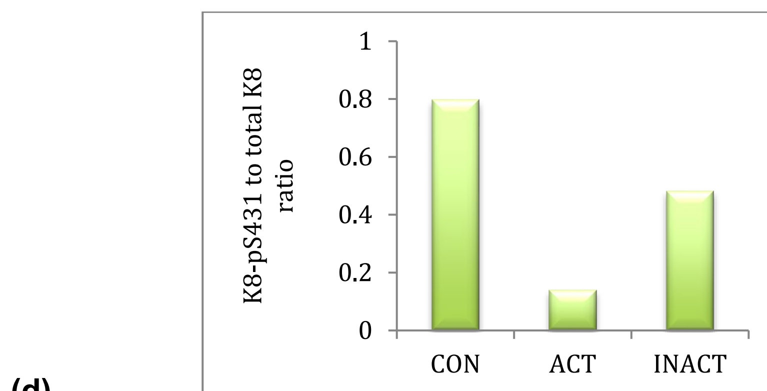
(a)



(b)



(c)



(d)

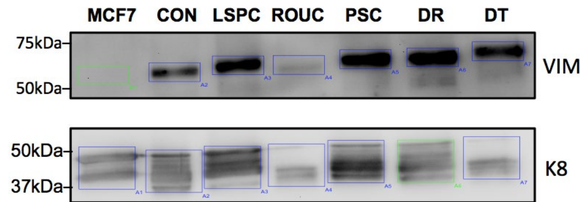
4.3.6. Mucosal vimentin levels are altered in UC

A densitometry analysis for regions quantified (figure 4.9a) of VIM and K8 immunoblots was undertaken. The ratio of VIM to K8 (figure 4.9b) was reduced in ROUC, in comparison to controls. There is however a progressive increase in VIM/K8 ratios from LSPC to PSC with markedly increased ratios noted in DT and DR compared to LSPC or ROUC (figure 4.9b). This increased ratio of K8:VIM in high risk phenotypes suggest a possible role of mucosal VIM:K8 ratio in identifying patients at risk of colitis associated cancer.

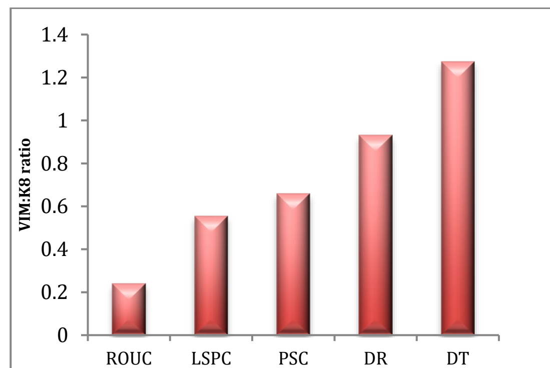
In the acutely inflamed mucosa (figure 4.9 c & d) there is a reduction in both VIM as well as total K8 levels; the VIM:K8 ratio is thus markedly low in acute inflammation. A relative increase (in comparison to controls) in total K8 is seen in the un-inflamed mucosa. This is reflected in an increased VIM:K8 ratio in the healthy controls relative to un-inflamed proximal mucosa;

Figure 4.9 *Densitometry analysis of K8 and VIM in different patient groups with ratio of VIM relative to total K8 in colonic mucosa. The extracted cytoskeletal fraction from recruited patients (in pooled groups) was resolved by gel electrophoresis on SDS-PAGE gels, followed by western transfer and immunoprobing using the antibodies to K8 and vimentin. 5µG protein was loaded onto each lane. Densitometric analysis of the band intensity was undertaken(Panel a&c). There is a progressive increase in VIM:K8 ratio in phenotypes associated with increased risk of CRC and in patients with dysplasia(Panel c).*

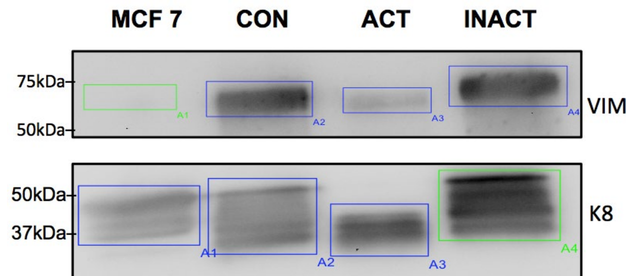
(a)



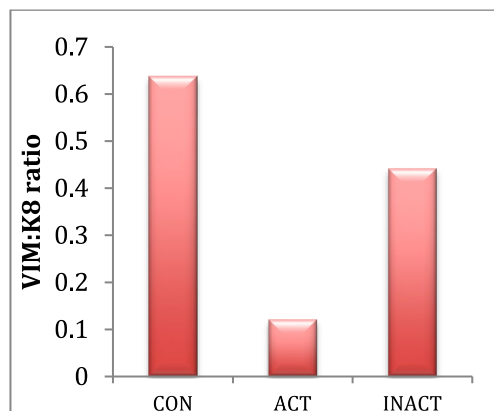
(b)



(c)



(d)



4.4 Conclusion

In this study, using WB we have validated the results of *iTRAQ* analysis. We have demonstrated increased levels of IFs in LSPC, and also a reduction in K8, K18 and K19 levels in the inflamed mucosa compared to controls (and proximal healthy colonic mucosa). In addition, we have identified changes in K8 expression including the presence of K8 in multiple isoforms corresponding to discrete bands on WB. This has been noted in diseased as well as in healthy mucosa, although the pattern of distribution (between 37-50kDa) has varied in diseased states.

In the actively inflamed mucosa, there is reduced expression of total K8 and also absent or reduced phosphorylation of K8. Similar changes have also been noted in noted in ROUC and DT. This is particularly noted with phosphorylated K8 at pS23. There is intense phosphorylation in controls and in LSPC. Although LSPC is believed to be an important factor contributing to the pathogenesis of CAC, it is likely that elevated total K8 and K8-pS23 levels might have a protective role. Dephosphorylation of K8 is now known to be associated with CRC progression [139]. Recurrent inflammation and on-going active inflammation is also an important risk factor for development of CAC [47] [48]. Recurrent bouts of inflammation and failure of recovery of phosphorylated K8 could be a contributing factor in the pathogenesis of CAC.

Vimentin is a type III IF proteins expressed in mesenchymal cells such as fibroblasts, endothelial cells and hematopoietic cells [201, 202]. It is considered as a marker of epithelial-mesenchymal transformation (EMT), a biological process whereby epithelial cells show features of mesenchymal cells [209-211]. One of the features of EMT is the association with more aggressive disease phenotype in cancers including in tumour invasion and metastasis [212]. Investigation of murine (CT26) and rat (IEC-6) intestinal cancer cell line has shown high vimentin protein expression with negligible K8/K18 levels [213]. Higher levels of vimentin have also been noted in longstanding UC associated with up-regulation of other tissue fibrosis markers [214].

In comparison to controls, VIM levels are increased in higher risk groups, LSPC, PSC, DR and DT. VIM levels are reduced in the acutely inflamed mucosa as well as in ROUC. More importantly on densitometry analysis of WB results, there is a progressive increase in VIM:K8 ratio in higher risk phenotypes. A relative overexpression of VIM in the colonic mucosa may thus be a potential factor in pathogenesis of CAC. In breast cancers, aggressive tumours (with high histological and nuclear grades) shows similar increase in vimentin levels with decrease in keratin levels [215]. There is also a potential role of VIM:K8 ratio as a mucosal biomarker for identifying individuals at a higher risk of colitis associated cancer.

CHAPTER 5

DISCUSSION

5.1 Introduction

Ulcerative colitis is a chronic inflammatory disease process characterised by in majority of patients, a chronic quiescent disease process often interspersed with episodes of acute inflammation [1]. It is associated with increased risk of colorectal cancer [18]. A variety of factors are associated with increased risk including prolonged duration of disease [22], extent of colitis [29, 30], presence of associated PSC [53] or family history of CRC [34]. One of the factors now considered to play a significant role in pathogenesis of colitis-associated cancer is on-going active inflammation with both endoscopic and histological inflammation being important [47, 48]. In ulcerative colitis the phenotypic presentation can be varied (often with varying disease distribution) and with differing risk of cancer in individual phenotypes. We investigated proteomic changes in insoluble intermediate filaments in colonic mucosa for the following reasons:

I. Varying phenotypic presentation of patients with differing cancer risk
II. Suitability of categorising patients into phenotypic groups
III. Early dysplastic changes are limited to the mucosa, and blood and other non-invasive tests poorly reflect mucosal changes. Hence investigating molecular changes in UC patient subsets might provide greater insights into the pathogenesis of disease progression including development of CAC.
IV. The mucosal changes of inflammation tend to be continuous reducing the chances of sampling error at endoscopy
V. Inflammation in a subset of patients with distal disease tends to show a distinct demarcation between the inflamed and non-inflamed segment. This allows direct comparison in the same group of patients.
VI. Patients with dysplasia often tend to have pan-colonic genomic changes also known as field changes [21]

Current strategies of surveillance in patients with long standing colitis include regular surveillance colonoscopies along with random (non-targeted) biopsies in order to identify dysplasia and early cancers. This technique is insensitive

due to the random nature of the biopsies and is fraught with the risk of missing malignant or pre-malignant changes [62, 65]. As a result, the focus of management of these patients should be towards development of mucosal biomarkers, which can act as a surrogate marker of dysplasia or cancer in colonic biopsies. As genomic field defects occur throughout the colonic mucosa in patients with dysplasia [70, 71], a pan-colonic biomarker could potentially avoid the pitfalls of errors associated with sampling the mucosa with random biopsies.

The importance of intermediate filament protein K8 in gastrointestinal function was demonstrated in K8 knockout mice has shown high rates of GI epithelial inflammation and hyperplasia [173]. In humans, heritable predispositions to ulcerative colitis (UC) have been mapped to the KRT8/18 loci. *In vitro* reconstitution of these mutations (KRT8: G62C, I63V, K464N; KRT18: S230T) reduced their ability to polymerise correctly [180]. In addition to playing a role in UC, derangements of keratins have been noted in morphologically normal fields around colorectal adenoma [206].

In view of the evidence implicating inflammation and cancer in colitis, and potential association of keratins in pathogenesis of IBD [174, 180], we investigated mucosal proteomic changes in insoluble intermediate filaments in well-characterised group of UC patients with varying cancer risk.

Quantification of protein (quantitative proteomics) is an important goal of proteomic experiments [78]. Currently there are two main techniques of relative quantification used in MS-based proteomics analyses. These include label-based and label-free approaches [78, 216]. In the former, quantification is done by introducing stable isotopes. Such labeling of peptides or proteins with isotopes (isotope incorporation) includes metabolic labeling [stable isotope labeling of amino acids in cell culture, (SILAC)], chemical labeling [isotope-coded affinity tag (ICAT), isobaric tag for relative and absolute quantification (ITRAQ)] or enzymatic labeling [216, 217].

Both, labeled and label-free quantification have been widely applied in proteomic research. Labeling techniques allows simultaneous analyses of

multiple samples in a single MS run (multiplexing) reducing variability in analyses. This is particularly relevant with iTRAQ labeling, where up to eight samples can be analysed simultaneously in a single experiment. Due to the isobaric nature of label, labeled peptides appear as a single peak in the full MS scan. Upon peptide fragmentation at MS/MS, the isotope-containing reporter ions are released and distinguished according to their masses based on the label composition [217]. On the other hand, in the label-free approach quantification is done based on spectral counting and intensity-based measurements [216].

Several studies have compared label-free and iTRAQ approaches [218-221]. The label free approach provides both a higher number and a higher percentage of differentially abundant proteins with a higher protein sequence coverage. Nevertheless, iTRAQ labeling as a quantification strategy for shotgun proteomics has an inherent advantages of multiplexing. This can potentially reduce MS analytical time and the total cost of the experiments compared to the label free techniques. Since all the samples are analysed simultaneously in a single MS run, inter-run variations in protein identification and quantification is avoided [216]. Studies have also shown that with pre-fractionation, an improved identification rate can also be achieved, with results similar to that obtained with label free techniques [216]. However relative quantification by iTRAQ can lead to under-estimation for a number of technical reasons that result in ratio compression [222-224]. Nevertheless the direction of fold change ie increase or decrease is correct and iTRAQ (and other isobaric tags) remains popular as a discovery tool for clinical proteomics [225].

In this study we undertook an iTRAQ-based approach in view of the pooled samples categorised into 8 groups, which enabled direct comparison between groups and hence reducing inter run variations. Additionally, multiplexing enabled the comparisons in a cost effective and timely fashion. As we also used a pre-fractionation step to isolate the insoluble intermediate filament proteins, we were able to offset the disadvantage of protein sequence coverage.

5.2 Developing a technique for intermediate filament isolation from biopsies and parallel analysis of IF proteins by western blotting and mass spectrometry

Intermediate filament proteins mostly exist in epithelial cells in an insoluble filamentous form; only 5% of it is usually soluble [127]. Previous studies have described extraction and solubilisation of keratins from cell line material with sequential use of high and low detergent buffers [94, 97, 99]. As biological samples tend to be complex, and the size of colorectal biopsies are very small, we modified the cell line technique and adapted them for colorectal biopsies.

In this study we aimed to develop a technique for solubilisation of insoluble intermediate filament proteins obtained from colorectal biopsies to enable parallel assessment by western immunoblotting and quantitative proteomics. This involved adapting IF protein extraction techniques used for breast cancer cell lines (MCF-7) to colonic biopsies. IF isolation from cell lines was undertaken using a 'high salt extraction technique' which has been previously described by Achtstaetter *et al*. and subsequently modified by Hermann *et al* [117, 195]. The 'high-salt extraction' technique was then modified for use on colorectal biopsies.

We then had to ensure adequate solubilisation of the IFs to allow further analyses. An important characteristic, which distinguishes IFs including keratins from other cytoskeletal filament proteins such as actin, myosin, and tubulin, is their resistance to buffers having non-denaturing detergents and high concentrations of ions [116-118]. High concentrations of urea (9-10 M) or guanidine hydrochloride (GuHCl)(4-6 M) can enable solubilisation IFs [115, 117]. Although the latter is of particular value due to compatibility with downstream mass spectrometry based analyses [192-194], its use in WB is affected due to formation of precipitates with SDS [192]. On the other hand heating can result in decomposition of urea to isocyanic acid and carbamylate free amines. Carbamylation blocks trypsin digestion sites and also result in peptide modifications indistinguishable from the naturally occurring PTMs [191]. This can affect subsequent analyses for PTMs in IF proteins.

We developed and validated a technique for isolation of IFs and effective solubilisation in GuHCl. This enabled assay of IF proteins by mass spectrometry. We adapted a technique for removal of GuHCl from dissolved samples to enable gel electrophoresis. This modified protocol thus enabled parallel analysis of IFs using both gel electrophoresis and 'shotgun proteomics' approach of mass spectrometry while ensuring preservation of PTMs.

5.3 Implication of altered keratin levels in the mucosa: an apparent failure of restoration of keratins could predispose to pathogenesis of colitis-associated cancer

Acute inflammation in colonic mucosa has been shown to reduce the levels of mucosal K8 in comparison to controls and proximal un-inflamed mucosa. Similar changes are also noted in ROUC, recently diagnosed patients with (endoscopically and histologically) quiescent disease and with recent evidence of active mucosal inflammation. The dysplastic tissue has also shown a similar pattern of changes. On the other hand, quiescent long-standing disease is associated with an elevated level of mucosal K8.

Keratins are dynamic and are involved in regulating various inflammatory pathways. They may have a role in cell-death signalling pathways, in particular apoptosis mediated by tumour necrosis factor-alpha (TNF-α) and Fas [157]. Epithelial cells lacking K8 and K18 are significantly more sensitive to TNF-mediated apoptosis [158]. The reduced levels in acute inflammation and low levels in ROUC and DT could signify an apparent failure of restoration of keratins in the mucosa following acute inflammation, despite macroscopic mucosal healing. We hypothesise that in recurrent bouts of inflammation and consequent failure of restoration of keratin levels in the mucosa could be a predisposing factor to CAC development. On the other hand quiescent LSPC with restored keratin levels may be at a less risk of CAC. Previous studies have also shown persistent cellular and molecular damage with activated kinase and transcription factor involving signaling pathways in the colon, despite apparent microscopic healing [226]. A proposed pathway of pathogenesis of CAC is shown in figure 5.1.

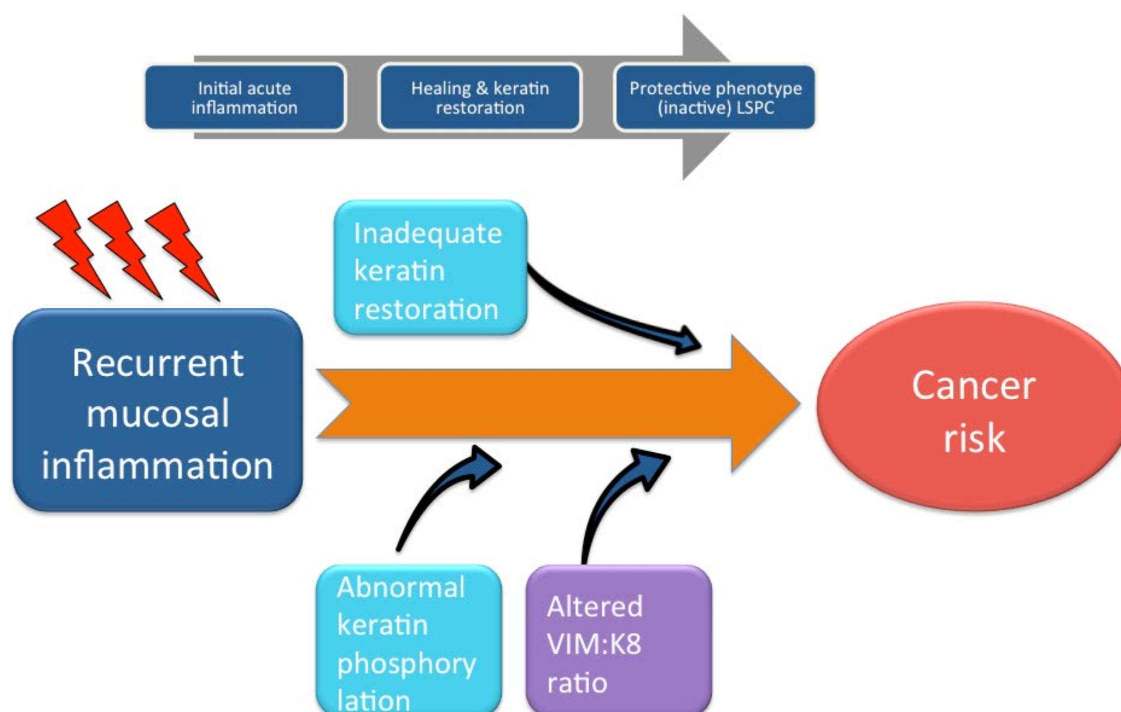
Phosphorylation of K8 is an important PTM, which affects its solubility and function [127, 227]. Dephosphorylation of K8 has been previously been shown to be associated with CRC progression [139]. K8 is a substrate of phosphatase of regenerating liver-3 (PRL-3); the later plays a role in cancer migration. Dephosphorylation of K8 associated is associated with high PRL-3 expression in aggressive tumours and thus contributes to colorectal cancer progression

In our study, there is evidence of absent or reduced phosphorylation in acutely inflamed mucosa. Similar changes are once again noted in ROUC and DT. This is particularly noted with phosphorylated K8 at pS23. There is intense phosphorylation in controls and in LSPC. These changes suggest a possible protective role of physiological phosphorylation of K8 at Ser 23. Although LSPC is believed to be an important factor contributing to the pathogenesis of CAC, it is likely that elevated total K8 and K8-pS23 levels could be protective and primarily patients with recurrent bouts of inflammation and failure of recovery of phosphorylated K8 are more prone to CAC. Acute inflammation, and particularly recurrent mucosal inflammation could be associated with loss of such phosphorylation.

5.4 Altered vimentin to keratin ratio: a marker of aggressive disease?

Vimentin is a type III intermediate filament protein of mesenchymal origin [208]. The role of VIM in the pathogenesis of various diseases has emerged [228]. In particular this is now considered a marker of epithelial-mesenchymal transformation associated with more aggressive disease phenotype in cancers particularly colorectal cancers [204]. Higher VIM expression has been shown in sera of patients with colorectal cancer [229]. In our study VIM levels are reduced in the acutely inflamed mucosa as well as in ROUC. More importantly on densitometry analysis of WB results, there is a progressive increase in VIM:K8 ratio with more aggressive phenotype. These results suggest that relative overexpression of VIM in the colonic mucosa may be an important factor in pathogenesis of CAC (figure 5.1).

Figure 5.1 Proposed pathway of pathogenesis of colorectal cancer in ulcerative colitis. *This demonstrates two pathways, the less aggressive quiescent disease being associated with restored keratin levels, which confer mucosal protection. On the other hand the aggressive phenotype is associated with recurrent bouts of inflammation, associated inadequate restoration of keratin levels, keratin dephosphorylation and relative excess of vimentin (relative to keratin).*



5.5 Conclusion

This is a prospective cross sectional study in which we recruited a large number of UC patients who were then categorised based on their phenotypic characteristics. We then carefully preselected patients into further groups (based on phenotypic features) to enable comparison using proteomic studies. This is the only quantitative proteomic study described in literature which has investigated changes in insoluble IF levels in the mucosa in patients with UC. Although previous studies have focussed on the soluble fraction of proteins, we undertook subcellular fractionation of the relatively insoluble fraction since this constitutes the bulk of the cellular keratin pool. In order to analyse (by quantitative proteomics and western blotting) such insoluble proteins from colorectal biopsies, we modified an extraction technique previously described in cell lines. We then validated a method for

solubilising the proteins to enable parallel analyses by quantitative proteomic techniques and western immunoblotting.

Our results identified characteristics which were unique to each group, in particular changes in total K8, vimentin and phosphorylated K8 which could not only be useful as biomarkers of phenotypic states, but also could provide insight into the the pathogenesis of CAC. We hypothesised that recurrent acute inflammation by reducing keratin 8 (and phosphorylated K8) levels in the mucosa, affects overall mucosal IF protein integrity. These changes in K8 lag behind apparent clinical, microscopic and endoscopic recovery (as noted in ROUC group). Recurrent inflammation with the consequent failure of restoration of K8 to normal healthy state might be an important step in pathogenesis of CAC. In addition, the role of vimentin is also important. Vimentin has been described as a marker of epithelial-mesenchymal transformation; higher levels have also been noted in patients with CRC. Hence the altered VIM:K8 ratio noted in our study could suggest a potential role of its use as a mucosal marker of more aggressive disease.

We would however like to also point out that despite the significant results, there are certain limitations of this project. This is a cross sectional pooled study. Although inflammation related decrease in keratin levels, and a failure to restore mucosal keratin levels in quiescent recent onset disease, has been shown in this study, the long term effects of such changes are unknown. As dysplastic and neoplastic changes in the mucosa often develop in longstanding disease, it will require prospective studies following up mucosal proteomic changes over a longer period of time.

Pooling of samples was undertaken in order to obtain information about lead biomarkers in the 8 groups identified. Although, this suffers from a disadvantage of loss of information due to biological variation, we aimed to identify biomarkers (and subsequently validate the results) in a cost and time effective manner. These results have since been orthogonally validated in pooled groups in this study and subsequently as individual patients by both western/dot blotting and immunohistochemistry (Assadsangabi A *et al*, BSG 2015; Marsh A *et al.* BMedSci project). Additionally, we have tried to minimise

clinical heterogeneity by pre-screening to ensure that homogenous clinical groups were selected.

The numbers of patients in certain groups in this study, particularly in the dysplastic groups and PSC groups are small. This could have impacted the overall results in these groups by providing false positive results due to heterogeneity in clinical samples. Nevertheless, the information obtained from the proteomic analysis of these groups do provide information about trend in the biomarker levels. These undoubtedly need to be investigated further by prospective studies involving larger patient numbers.

CHAPTER 6

BIBLIOGRAPHY

References

1. Silverberg, M.S., et al., *Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology.* Can J Gastroenterol, 2005. **19 Suppl A**: p. 5-36.
2. Carter, M.J., A.J. Lobo, and S.P. Travis, *Guidelines for the management of inflammatory bowel disease in adults.* Gut, 2004. **53 Suppl 5**: p. V1-16.
3. Loftus, E.V., Jr., *Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences.* Gastroenterology, 2004. **126**(6): p. 1504-17.
4. Rubin, G.P., et al., *Inflammatory bowel disease: epidemiology and management in an English general practice population.* Aliment Pharmacol Ther, 2000. **14**(12): p. 1553-9.
5. Fiocchi, C., *Inflammatory bowel disease: etiology and pathogenesis.* Gastroenterology, 1998. **115**(1): p. 182-205.
6. Kucharzik, T., et al., *Recent understanding of IBD pathogenesis: implications for future therapies.* Inflamm Bowel Dis, 2006. **12**(11): p. 1068-83.
7. Xavier, R.J. and D.K. Podolsky, *Unravelling the pathogenesis of inflammatory bowel disease.* Nature, 2007. **448**(7152): p. 427-34.
8. Sartor, R.B., *Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis.* Nat Clin Pract Gastroenterol Hepatol, 2006. **3**(7): p. 390-407.
9. Calkins, B.M., *A meta-analysis of the role of smoking in inflammatory bowel disease.* Dig Dis Sci, 1989. **34**(12): p. 1841-54.
10. Koutroubakis, I.E., I.G. Vlachonikolis, and E.A. Kouroumalis, *Role of appendicitis and appendectomy in the pathogenesis of ulcerative colitis: a critical review.* Inflamm Bowel Dis, 2002. **8**(4): p. 277-86.
11. Duchmann, R., et al., *Tolerance exists towards resident intestinal flora but is broken in active inflammatory bowel disease (IBD).* Clin Exp Immunol, 1995. **102**(3): p. 448-55.
12. Gaya, D.R., et al., *New genes in inflammatory bowel disease: lessons for complex diseases?* Lancet, 2006. **367**(9518): p. 1271-84.
13. Ho, G.T., et al., *Allelic variations of the multidrug resistance gene determine susceptibility and disease behavior in ulcerative colitis.* Gastroenterology, 2005. **128**(2): p. 288-96.
14. Farrell, R.J., et al., *High multidrug resistance (P-glycoprotein 170) expression in inflammatory bowel disease patients who fail medical therapy.* Gastroenterology, 2000. **118**(2): p. 279-88.
15. Panwala, C.M., J.C. Jones, and J.L. Viney, *A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis.* J Immunol, 1998. **161**(10): p. 5733-44.
16. Solberg, I.C., et al., *Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN Study).* Scand J Gastroenterol, 2009. **44**(4): p. 431-40.
17. Farmer, R.G., K.A. Easley, and G.B. Rankin, *Clinical patterns, natural history, and progression of ulcerative colitis. A long-term follow-up of 1116 patients.* Dig Dis Sci, 1993. **38**(6): p. 1137-46.

18. Crohn, U.B. and H. Rosenberg, *The sigmoidoscopic picture of chronic ulcerative colitis (non-specific).* American Journal of the Medical Sciences, 1925. **170**: p. 220-228.
19. Choi, P.M. and M.P. Zelig, *Similarity of colorectal cancer in Crohn's disease and ulcerative colitis: implications for carcinogenesis and prevention.* Gut, 1994. **35**(7): p. 950-4.
20. Riddell, R.H., et al., *Dysplasia in inflammatory bowel disease: standardized classification with provisional clinical applications.* Hum Pathol, 1983. **14**(11): p. 931-68.
21. Bernstein, C.N., F. Shanahan, and W.M. Weinstein, *Are we telling patients the truth about surveillance colonoscopy in ulcerative colitis?* Lancet, 1994. **343**(8889): p. 71-4.
22. Eaden, J.A., K.R. Abrams, and J.F. Mayberry, *The risk of colorectal cancer in ulcerative colitis: a meta-analysis.* Gut, 2001. **48**(4): p. 526-35.
23. Rutter, M.D., et al., *Thirty-year analysis of a colonoscopic surveillance program for neoplasia in ulcerative colitis.* Gastroenterology, 2006. **130**(4): p. 1030-8.
24. Winawer, S., et al., *Colorectal cancer screening and surveillance: clinical guidelines and rationale-Update based on new evidence.* Gastroenterology, 2003. **124**(2): p. 544-60.
25. Leighton, J.A., et al., *ASGE guideline: endoscopy in the diagnosis and treatment of inflammatory bowel disease.* Gastrointest Endosc, 2006. **63**(4): p. 558-65.
26. Cairns, S.R., et al., *Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002).* Gut. **59**(5): p. 666-89.
27. Lutgens, M.W., et al., *High frequency of early colorectal cancer in inflammatory bowel disease.* Gut, 2008. **57**(9): p. 1246-51.
28. Farraye, F.A., et al., *AGA technical review on the diagnosis and management of colorectal neoplasia in inflammatory bowel disease.* Gastroenterology. **138**(2): p. 746-74, 774 e1-4; quiz e12-3.
29. Gyde, S.N., et al., *Colorectal cancer in ulcerative colitis: a cohort study of primary referrals from three centres.* Gut, 1988. **29**(2): p. 206-17.
30. Ekbom, A., et al., *Ulcerative colitis and colorectal cancer. A population-based study.* N Engl J Med, 1990. **323**(18): p. 1228-33.
31. Mathy, C., et al., *Gross versus microscopic pancolitis and the occurrence of neoplasia in ulcerative colitis.* Inflamm Bowel Dis, 2003. **9**(6): p. 351-5.
32. Heuschen, U.A., et al., *Backwash ileitis is strongly associated with colorectal carcinoma in ulcerative colitis.* Gastroenterology, 2001. **120**(4): p. 841-7.
33. Haskell, H., et al., *Pathologic features and clinical significance of "backwash" ileitis in ulcerative colitis.* Am J Surg Pathol, 2005. **29**(11): p. 1472-81.
34. Askling, J., et al., *Family history as a risk factor for colorectal cancer in inflammatory bowel disease.* Gastroenterology, 2001. **120**(6): p. 1356-62.
35. Bergeron, V., et al., *Risk Factors for Neoplasia in Inflammatory Bowel Disease Patients With Pancolitis.* Am J Gastroenterol.
36. Nuako, K.W., et al., *Familial predisposition for colorectal cancer in chronic ulcerative colitis: a case-control study.* Gastroenterology, 1998. **115**(5): p. 1079-83.

37. Balkwill, F. and A. Mantovani, *Inflammation and cancer: back to Virchow?* Lancet, 2001. **357**(9255): p. 539-45.
38. Coussens, L.M. and Z. Werb, *Inflammation and cancer.* Nature, 2002. **420**(6917): p. 860-7.
39. Westbrook, A.M., A. Szakmary, and R.H. Schiestl, *Mechanisms of intestinal inflammation and development of associated cancers: lessons learned from mouse models.* Mutat Res. **705**(1): p. 40-59.
40. Cooper, H.S., et al., *Dysplasia and cancer in the dextran sulfate sodium mouse colitis model. Relevance to colitis-associated neoplasia in the human: a study of histopathology, B-catenin and p53 expression and the role of inflammation.* Carcinogenesis, 2000. **21**(4): p. 757-68.
41. Burstein, E. and E.R. Fearon, *Colitis and cancer: a tale of inflammatory cells and their cytokines.* J Clin Invest, 2008. **118**(2): p. 464-7.
42. Fantini, M.C. and F. Pallone, *Cytokines: from gut inflammation to colorectal cancer.* Curr Drug Targets, 2008. **9**(5): p. 375-80.
43. Szlosarek, P., K.A. Charles, and F.R. Balkwill, *Tumour necrosis factor-alpha as a tumour promoter.* Eur J Cancer, 2006. **42**(6): p. 745-50.
44. Allavena, P., et al., *Pathways connecting inflammation and cancer.* Curr Opin Genet Dev, 2008. **18**(1): p. 3-10.
45. Mumm, J.B. and M. Oft, *Cytokine-based transformation of immune surveillance into tumor-promoting inflammation.* Oncogene, 2008. **27**(45): p. 5913-9.
46. Lin, W.W. and M. Karin, *A cytokine-mediated link between innate immunity, inflammation, and cancer.* J Clin Invest, 2007. **117**(5): p. 1175-83.
47. Rutter, M., et al., *Severity of inflammation is a risk factor for colorectal neoplasia in ulcerative colitis.* Gastroenterology, 2004. **126**(2): p. 451-9.
48. Gupta, R.B., et al., *Histologic inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study.* Gastroenterology, 2007. **133**(4): p. 1099-105; quiz 1340-1.
49. Rutter, M.D., et al., *Cancer surveillance in longstanding ulcerative colitis: endoscopic appearances help predict cancer risk.* Gut, 2004. **53**(12): p. 1813-6.
50. Velayos, F.S., et al., *Predictive and protective factors associated with colorectal cancer in ulcerative colitis: A case-control study.* Gastroenterology, 2006. **130**(7): p. 1941-9.
51. Broome, U., et al., *Natural history and prognostic factors in 305 Swedish patients with primary sclerosing cholangitis.* Gut, 1996. **38**(4): p. 610-5.
52. Okolicsanyi, L., et al., *Primary sclerosing cholangitis: clinical presentation, natural history and prognostic variables: an Italian multicentre study. The Italian PSC Study Group.* Eur J Gastroenterol Hepatol, 1996. **8**(7): p. 685-91.
53. Broome, U., et al., *Primary sclerosing cholangitis and ulcerative colitis: evidence for increased neoplastic potential.* Hepatology, 1995. **22**(5): p. 1404-8.
54. Soetikno, R.M., et al., *Increased risk of colorectal neoplasia in patients with primary sclerosing cholangitis and ulcerative colitis: a meta-analysis.* Gastrointest Endosc, 2002. **56**(1): p. 48-54.
55. Claessen, M.M., et al., *High lifetime risk of cancer in primary sclerosing cholangitis.* J Hepatol, 2009. **50**(1): p. 158-64.

56. Vera, A., et al., *Colorectal cancer in patients with inflammatory bowel disease after liver transplantation for primary sclerosing cholangitis.* Transplantation, 2003. **75**(12): p. 1983-8.
57. Loftus, E.V., Jr., et al., *Risk of colorectal neoplasia in patients with primary sclerosing cholangitis and ulcerative colitis following orthotopic liver transplantation.* Hepatology, 1998. **27**(3): p. 685-90.
58. Itzkowitz, S.H., *Molecular biology of dysplasia and cancer in inflammatory bowel disease.* Gastroenterol Clin North Am, 2006. **35**(3): p. 553-71.
59. Akare, S., et al., *Ursodeoxycholic acid modulates histone acetylation and induces differentiation and senescence.* Int J Cancer, 2006. **119**(12): p. 2958-69.
60. Kornbluth, A. and D.B. Sachar, *Ulcerative colitis practice guidelines in adults (update): American College of Gastroenterology, Practice Parameters Committee.* Am J Gastroenterol, 2004. **99**(7): p. 1371-85.
61. Rubin, C.E., et al., *DNA aneuploidy in colonic biopsies predicts future development of dysplasia in ulcerative colitis.* Gastroenterology, 1992. **103**(5): p. 1611-20.
62. Connell, W.R., et al., *Factors affecting the outcome of endoscopic surveillance for cancer in ulcerative colitis.* Gastroenterology, 1994. **107**(4): p. 934-44.
63. Eaden, J.A., B.A. Ward, and J.F. Mayberry, *How gastroenterologists screen for colonic cancer in ulcerative colitis: an analysis of performance.* Gastrointest Endosc, 2000. **51**(2): p. 123-8.
64. Rutter, M.D., et al., *Pancolonic indigo carmine dye spraying for the detection of dysplasia in ulcerative colitis.* Gut, 2004. **53**(2): p. 256-60.
65. Zisman, T.L. and D.T. Rubin, *Colorectal cancer and dysplasia in inflammatory bowel disease.* World J Gastroenterol, 2008. **14**(17): p. 2662-9.
66. Rosenstock, E., et al., *Surveillance for colonic carcinoma in ulcerative colitis.* Gastroenterology, 1985. **89**(6): p. 1342-6.
67. Lim, C.H., et al., *Ten year follow up of ulcerative colitis patients with and without low grade dysplasia.* Gut, 2003. **52**(8): p. 1127-32.
68. Mpofu, C., A.J. Watson, and J.M. Rhodes, *Strategies for detecting colon cancer and/or dysplasia in patients with inflammatory bowel disease.* Cochrane Database Syst Rev, 2004(2): p. CD000279.
69. Hung, K.E. and K.H. Yu, *Proteomic approaches to cancer biomarkers.* Gastroenterology. **138**(1): p. 46-51 e1.
70. Rabinovitch, P.S., et al., *Pancolonic chromosomal instability precedes dysplasia and cancer in ulcerative colitis.* Cancer Res, 1999. **59**(20): p. 5148-53.
71. Chen, R., et al., *DNA fingerprinting abnormalities can distinguish ulcerative colitis patients with dysplasia and cancer from those who are dysplasia/cancer-free.* Am J Pathol, 2003. **162**(2): p. 665-72.
72. O'Sullivan, J.N., et al., *Chromosomal instability in ulcerative colitis is related to telomere shortening.* Nat Genet, 2002. **32**(2): p. 280-4.
73. Issa, J.P., et al., *Accelerated age-related CpG island methylation in ulcerative colitis.* Cancer Res, 2001. **61**(9): p. 3573-7.

74. Din, S., et al., *Technology insight: the application of proteomics in gastrointestinal disease.* Nat Clin Pract Gastroenterol Hepatol, 2007. **4**(7): p. 372-85.
75. Pandey, A. and M. Mann, *Proteomics to study genes and genomes.* Nature, 2000. **405**(6788): p. 837-46.
76. Kok, K., P. Stokkers, and P.H. Reitsma, *Genomics and proteomics: implications for inflammatory bowel diseases.* Inflamm Bowel Dis, 2004. **10 Suppl 1**: p. S1-6.
77. Felley-Bosco, E. and M. Andre, *Proteomics and chronic inflammatory bowel diseases.* Pathol Res Pract, 2004. **200**(2): p. 129-33.
78. Brewis, I.A. and P. Brennan, *Proteomics technologies for the global identification and quantification of proteins.* Adv Protein Chem Struct Biol, 2010. **80**: p. 1-44.
79. Wilkins, M.R., et al., *From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis.* Biotechnology (N Y), 1996. **14**(1): p. 61-5.
80. Wilkins, M.R., et al., *Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it.* Biotechnol Genet Eng Rev, 1996. **13**: p. 19-50.
81. Eisenberg, D., et al., *Protein function in the post-genomic era.* Nature, 2000. **405**(6788): p. 823-6.
82. Wilkins, M.R., et al., *High-throughput mass spectrometric discovery of protein post-translational modifications.* Journal of molecular biology, 1999. **289**(3): p. 645-57.
83. Godovac-Zimmermann, J. and L.R. Brown, *Perspectives for mass spectrometry and functional proteomics.* Mass Spectrom Rev, 2001. **20**(1): p. 1-57.
84. Mallick, P. and B. Kuster, *Proteomics: a pragmatic perspective.* Nat Biotechnol, 2010. **28**(7): p. 695-709.
85. Guo, Y., Z. Fu, and J.E. Van Eyk, *A proteomic primer for the clinician.* Proc Am Thorac Soc, 2007. **4**(1): p. 9-17.
86. Hsieh, S.Y., et al., *Comparative proteomic studies on the pathogenesis of human ulcerative colitis.* Proteomics, 2006. **6**(19): p. 5322-31.
87. Fogt, F., et al., *Proteomic analysis of mucosal preparations from patients with ulcerative colitis.* Mol Med Rep, 2008. **1**(1): p. 51-4.
88. Naito, Y., et al., *Identification of inflammation-related proteins in a murine colitis model by 2D fluorescence difference gel electrophoresis and mass spectrometry.* J Gastroenterol Hepatol, 2010. **25 Suppl 1**: p. S144-8.
89. Nanni, P., et al., *Differential proteomic analysis of HT29 Cl.16E and intestinal epithelial cells by LC ESI/QTOF mass spectrometry.* J Proteomics, 2009. **72**(5): p. 865-73.
90. Barcelo-Batllori, S., et al., *Proteomic analysis of cytokine induced proteins in human intestinal epithelial cells: implications for inflammatory bowel diseases.* Proteomics, 2002. **2**(5): p. 551-60.
91. Wong, W.Y., et al., *Proteomic profiling of dextran sulfate sodium induced acute ulcerative colitis mice serum exosomes and their immunomodulatory impact on macrophages.* Proteomics, 2016. **16**(7): p. 1131-45.
92. Nanni, P., et al., *A label-free nano-liquid chromatography-mass spectrometry approach for quantitative serum peptidomics in Crohn's*

disease patients. J Chromatogr B Analyt Technol Biomed Life Sci, 2009. **877**(27): p. 3127-36.
93. Broedl, U.C., et al., *Apolipoprotein A-IV is an independent predictor of disease activity in patients with inflammatory bowel disease.* Inflamm Bowel Dis, 2007. **13**(4): p. 391-7.
94. Meuwis, M.A., et al., *Biomarker discovery for inflammatory bowel disease, using proteomic serum profiling.* Biochem Pharmacol, 2007. **73**(9): p. 1422-33.
95. M'Koma, A.E., et al., *Proteomic profiling of mucosal and submucosal colonic tissues yields protein signatures that differentiate the inflammatory colitides.* Inflamm Bowel Dis, 2011. **17**(4): p. 875-83.
96. Viennois, E., et al., *Longitudinal study of circulating protein biomarkers in inflammatory bowel disease.* J Proteomics, 2015. **112**: p. 166-79.
97. Townsend, P., et al., *Serum Proteome Profiles in Stricturing Crohn's Disease: A Pilot Study.* Inflamm Bowel Dis, 2015. **21**(8): p. 1935-41.
98. Yeo, M., et al., *Loss of transgelin in repeated bouts of ulcerative colitis-induced colon carcinogenesis.* Proteomics, 2006. **6**(4): p. 1158-65.
99. Yasui, Y. and T. Tanaka, *Protein expression analysis of inflammation-related colon carcinogenesis.* J Carcinog, 2009. **8**: p. 10.
100. Suzuki, R., et al., *Global gene expression analysis of the mouse colonic mucosa treated with azoxymethane and dextran sodium sulfate.* BMC Cancer, 2007. **7**: p. 84.
101. Araki, K., et al., *High expression of HSP47 in ulcerative colitis-associated carcinomas: proteomic approach.* Br J Cancer, 2009. **101**(3): p. 492-7.
102. May, D., et al., *Investigating neoplastic progression of ulcerative colitis with label-free comparative proteomics.* Journal of proteome research, 2011. **10**(1): p. 200-9.
103. Pawlak, G. and D.M. Helfman, *Cytoskeletal changes in cell transformation and tumorigenesis.* Curr Opin Genet Dev, 2001. **11**(1): p. 41-7.
104. Brentnall, T.A., et al., *Proteins That Underlie Neoplastic Progression of Ulcerative Colitis.* Proteomics Clin Appl, 2009. **3**(11): p. 1326.
105. Kanmura, S., et al., *Human neutrophil peptides 1-3 are useful biomarkers in patients with active ulcerative colitis.* Inflammatory bowel diseases, 2009. **15**(6): p. 909-17.
106. Meuwis, M.A., et al., *Proteomics for prediction and characterization of response to infliximab in Crohn's disease: a pilot study.* Clin Biochem, 2008. **41**(12): p. 960-7.
107. Gazouli, M., et al., *Serum protein profile of Crohn's disease treated with infliximab.* J Crohns Colitis, 2013. **7**(10): p. e461-70.
108. Wasinger, V.C., et al., *Low Mass Blood Peptides Discriminative of Inflammatory Bowel Disease (IBD) Severity: A Quantitative Proteomic Perspective.* Mol Cell Proteomics, 2016. **15**(1): p. 256-65.
109. Fuchs, E. and D.W. Cleveland, *A structural scaffolding of intermediate filaments in health and disease.* Science, 1998. **279**(5350): p. 514-9.
110. Ku, N.O., et al., *The cytoskeleton of digestive epithelia in health and disease.* The American journal of physiology, 1999. **277**(6 Pt 1): p. G1108-37.
111. Schweizer, J., et al., *New consensus nomenclature for mammalian keratins.* The Journal of cell biology, 2006. **174**(2): p. 169-74.

112. Kim, S. and P.A. Coulombe, *Intermediate filament scaffolds fulfill mechanical, organizational, and signaling functions in the cytoplasm.* Genes & development, 2007. **21**(13): p. 1581-97.
113. Omary, M.B., et al., *Toward unraveling the complexity of simple epithelial keratins in human disease.* The Journal of clinical investigation, 2009. **119**(7): p. 1794-805.
114. Moll, R., et al., *The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells.* Cell, 1982. **31**(1): p. 11-24.
115. Ku, N.O., et al., *Studying simple epithelial keratins in cells and tissues.* Methods in cell biology, 2004. **78**: p. 489-517.
116. Franke, W.W., et al., *Differentiation-related patterns of expression of proteins of intermediate-size filaments in tissues and cultured cells.* Cold Spring Harbor symposia on quantitative biology, 1982. **46 Pt 1**: p. 431-53.
117. Achtstaetter, T., et al., *Separation of cytokeratin polypeptides by gel electrophoretic and chromatographic techniques and their identification by immunoblotting.* Methods in enzymology, 1986. **134**: p. 355-71.
118. Steinert, P., et al., *Isolation and characterization of intermediate filaments.* Methods in cell biology, 1982. **24**: p. 399-419.
119. Chou, C.F., et al., *A significant soluble keratin fraction in 'simple' epithelial cells. Lack of an apparent phosphorylation and glycosylation role in keratin solubility.* Journal of cell science, 1993. **105 (Pt 2)**: p. 433-44.
120. Liao, J., et al., *The 70-kDa heat shock proteins associate with glandular intermediate filaments in an ATP-dependent manner.* The Journal of biological chemistry, 1995. **270**(2): p. 915-22.
121. Achtstaetter, T., et al., *Separation of cytokeratin polypeptides by gel electrophoretic and chromatographic techniques and their identification by immunoblotting.* Methods Enzymol, 1986. **134**: p. 355-71.
122. Ku, N.O., et al., *Studying simple epithelial keratins in cells and tissues.* Methods Cell Biol, 2004. **78**: p. 489-517.
123. Cho, W.C., *Proteomics technologies and challenges.* Genomics Proteomics Bioinformatics, 2007. **5**(2): p. 77-85.
124. Ku, N.O., et al., *The cytoskeleton of digestive epithelia in health and disease.* Am J Physiol, 1999. **277**(6 Pt 1): p. G1108-37.
125. Toivola, D.M., et al., *Intermediate filaments take the heat as stress proteins.* Trends in cell biology, 2010. **20**(2): p. 79-91.
126. Oshima, R.G., *Developmental expression of murine extra-embryonic endodermal cytoskeletal proteins.* The Journal of biological chemistry, 1982. **257**(7): p. 3414-21.
127. Omary, M.B., et al., *Keratin modifications and solubility properties in epithelial cells and in vitro.* Sub-cellular biochemistry, 1998. **31**: p. 105-40.
128. Omary, M.B., et al., *"Heads and tails" of intermediate filament phosphorylation: multiple sites and functional insights.* Trends in biochemical sciences, 2006. **31**(7): p. 383-94.
129. Steinert, P.M., *The dynamic phosphorylation of the human intermediate filament keratin 1 chain.* The Journal of biological chemistry, 1988. **263**(26): p. 13333-9.
130. Feng, L., et al., *Pervanadate-mediated tyrosine phosphorylation of keratins 8 and 19 via a p38 mitogen-activated protein kinase-dependent pathway.* Journal of cell science, 1999. **112 (Pt 13)**: p. 2081-90.

131. Snider, N.T., H. Park, and M.B. Omary, *A conserved rod domain phosphotyrosine that is targeted by the phosphatase PTP1B promotes keratin 8 protein insolubility and filament organization.* J Biol Chem, 2013. **288**(43): p. 31329-37.
132. Leech, S.H., et al., *Proteomic analyses of intermediate filaments reveals cytokeratin8 is highly acetylated--implications for colorectal epithelial homeostasis.* Proteomics, 2008. **8**(2): p. 279-88.
133. Chou, C.F., A.J. Smith, and M.B. Omary, *Characterization and dynamics of O-linked glycosylation of human cytokeratin 8 and 18.* The Journal of biological chemistry, 1992. **267**(6): p. 3901-6.
134. Ditzel, H.J., et al., *Cancer-associated cleavage of cytokeratin 8/18 heterotypic complexes exposes a neoepitope in human adenocarcinomas.* The Journal of biological chemistry, 2002. **277**(24): p. 21712-22.
135. Makowski, G.S. and M.L. Ramsby, *Degradation of cytokeratin intermediate filaments by calcium-activated proteases (calpains) in vitro: implications for formation of Mallory bodies.* Research communications in molecular pathology and pharmacology, 1998. **101**(3): p. 211-23.
136. Snider, N.T., et al., *Keratin hypersumoylation alters filament dynamics and is a marker for human liver disease and keratin mutation.* The Journal of biological chemistry, 2011. **286**(3): p. 2273-84.
137. Corfe, B.M., et al., *Inflammation decreases keratin level in ulcerative colitis; inadequate restoration associates with increased risk of colitis-associated cancer.* BMJ Open Gastroenterol, 2015. **2**(1): p. e000024.
138. Omary, M.B., et al., *Keratin modifications and solubility properties in epithelial cells and in vitro.* Subcell Biochem, 1998. **31**: p. 105-40.
139. Mizuuchi, E., et al., *Down-modulation of keratin 8 phosphorylation levels by PRL-3 contributes to colorectal carcinoma progression.* Int J Cancer, 2009. **124**(8): p. 1802-10.
140. Ku, N.O. and M.B. Omary, *Identification of the major physiologic phosphorylation site of human keratin 18: potential kinases and a role in filament reorganization.* The Journal of cell biology, 1994. **127**(1): p. 161-71.
141. Ku, N.O., S. Azhar, and M.B. Omary, *Keratin 8 phosphorylation by p38 kinase regulates cellular keratin filament reorganization: modulation by a keratin 1-like disease causing mutation.* The Journal of biological chemistry, 2002. **277**(13): p. 10775-82.
142. Toivola, D.M., et al., *Type II keratins are phosphorylated on a unique motif during stress and mitosis in tissues and cultured cells.* Mol Biol Cell, 2002. **13**(6): p. 1857-70.
143. Ridge, K.M., et al., *Keratin 8 phosphorylation by protein kinase C delta regulates shear stress-mediated disassembly of keratin intermediate filaments in alveolar epithelial cells.* J Biol Chem, 2005. **280**(34): p. 30400-5.
144. Ip, Y.T. and R.J. Davis, *Signal transduction by the c-Jun N-terminal kinase (JNK)--from inflammation to development.* Current opinion in cell biology, 1998. **10**(2): p. 205-19.
145. Kyriakis, J.M. and J. Avruch, *Sounding the alarm: protein kinase cascades activated by stress and inflammation.* The Journal of biological chemistry, 1996. **271**(40): p. 24313-6.

146. Davis, R.J., *Signal transduction by the JNK group of MAP kinases.* Cell, 2000. **103**(2): p. 239-52.
147. He, T., et al., *The intermediate filament protein keratin 8 is a novel cytoplasmic substrate for c-Jun N-terminal kinase.* The Journal of biological chemistry, 2002. **277**(13): p. 10767-74.
148. Ku, N.O. and M.B. Omary, *Effect of mutation and phosphorylation of type I keratins on their caspase-mediated degradation.* The Journal of biological chemistry, 2001. **276**(29): p. 26792-8.
149. Ku, N.O. and M.B. Omary, *A disease- and phosphorylation-related nonmechanical function for keratin 8.* J Cell Biol, 2006. **174**(1): p. 115-25.
150. Kouzarides, T., *Acetylation: a regulatory modification to rival phosphorylation?* EMBO J, 2000. **19**(6): p. 1176-9.
151. Sadoul, K., et al., *The tale of protein lysine acetylation in the cytoplasm.* Journal of biomedicine & biotechnology, 2011. **2011**: p. 970382.
152. Chen, H., et al., *Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase.* Cell, 1999. **98**(5): p. 675-86.
153. Martinez-Balbas, M.A., et al., *Regulation of E2F1 activity by acetylation.* The EMBO journal, 2000. **19**(4): p. 662-71.
154. Snider, N.T., et al., *Glucose and SIRT2 reciprocally mediate the regulation of keratin 8 by lysine acetylation.* J Cell Biol, 2013. **200**(3): p. 241-7.
155. Omary, M.B., N.O. Ku, and D.M. Toivola, *Keratins: guardians of the liver.* Hepatology, 2002. **35**(2): p. 251-7.
156. Eriksson, J.E., et al., *Introducing intermediate filaments: from discovery to disease.* The Journal of clinical investigation, 2009. **119**(7): p. 1763-71.
157. Paramio, J.M. and J.L. Jorcano, *Beyond structure: do intermediate filaments modulate cell signalling?* BioEssays : news and reviews in molecular, cellular and developmental biology, 2002. **24**(9): p. 836-44.
158. Caulin, C., et al., *Keratin-dependent, epithelial resistance to tumor necrosis factor-induced apoptosis.* The Journal of cell biology, 2000. **149**(1): p. 17-22.
159. Toivola, D.M., et al., *Keratins modulate colonocyte electrolyte transport via protein mistargeting.* The Journal of cell biology, 2004. **164**(6): p. 911-21.
160. Ameen, N.A., Y. Figueroa, and P.J. Salas, *Anomalous apical plasma membrane phenotype in CK8-deficient mice indicates a novel role for intermediate filaments in the polarization of simple epithelia.* Journal of cell science, 2001. **114**(Pt 3): p. 563-75.
161. Toivola, D.M., et al., *Keratins modulate colonocyte electrolyte transport via protein mistargeting.* J Cell Biol, 2004. **164**(6): p. 911-21.
162. Helenius, T.O., et al., *Keratin 8 absence down-regulates colonocyte HMGCS2 and modulates colonic ketogenesis and energy metabolism.* Mol Biol Cell, 2015. **26**(12): p. 2298-310.
163. Murch, S.H., et al., *Serum concentrations of tumour necrosis factor alpha in childhood chronic inflammatory bowel disease.* Gut, 1991. **32**(8): p. 913-7.
164. Murch, S.H., et al., *Location of tumour necrosis factor alpha by immunohistochemistry in chronic inflammatory bowel disease.* Gut, 1993. **34**(12): p. 1705-9.
165. Braegger, C.P., et al., *Tumour necrosis factor alpha in stool as a marker of intestinal inflammation.* Lancet, 1992. **339**(8785): p. 89-91.

166. Caulin, C., et al., *Keratin-dependent, epithelial resistance to tumor necrosis factor-induced apoptosis.* J Cell Biol, 2000. **149**(1): p. 17-22.
167. Inada, H., et al., *Keratin attenuates tumor necrosis factor-induced cytotoxicity through association with TRADD.* The Journal of cell biology, 2001. **155**(3): p. 415-26.
168. Dong, X.M., et al., *Keratin 8 limits TLR-triggered inflammatory responses through inhibiting TRAF6 polyubiquitination.* Sci Rep, 2016. **6**: p. 32710.
169. Wang, L., et al., *Interleukin-6 induces keratin expression in intestinal epithelial cells: potential role of keratin-8 in interleukin-6-induced barrier function alterations.* The Journal of biological chemistry, 2007. **282**(11): p. 8219-27.
170. Shen, L. and J.R. Turner, *Role of epithelial cells in initiation and propagation of intestinal inflammation. Eliminating the static: tight junction dynamics exposed.* Am J Physiol Gastrointest Liver Physiol, 2006. **290**(4): p. G577-82.
171. Dignass, A.U., D.C. Baumgart, and A. Sturm, *Review article: the aetiopathogenesis of inflammatory bowel disease--immunology and repair mechanisms.* Aliment Pharmacol Ther, 2004. **20 Suppl 4**: p. 9-17.
172. Wang, L., et al., *Interleukin-6 induces keratin expression in intestinal epithelial cells: potential role of keratin-8 in interleukin-6-induced barrier function alterations.* J Biol Chem, 2007. **282**(11): p. 8219-27.
173. Baribault, H., et al., *Colorectal hyperplasia and inflammation in keratin 8-deficient FVB/N mice.* Genes & development, 1994. **8**(24): p. 2964-73.
174. Baribault, H., et al., *Colorectal hyperplasia and inflammation in keratin 8-deficient FVB/N mice.* Genes Dev, 1994. **8**(24): p. 2964-73.
175. Habtezion, A., et al., *Keratin-8-deficient mice develop chronic spontaneous Th2 colitis amenable to antibiotic treatment.* Journal of cell science, 2005. **118**(Pt 9): p. 1971-80.
176. Habtezion, A., et al., *Keratin-8-deficient mice develop chronic spontaneous Th2 colitis amenable to antibiotic treatment.* J Cell Sci, 2005. **118**(Pt 9): p. 1971-80.
177. Asghar, M.N., et al., *Keratin 8 knockdown leads to loss of the chloride transporter DRA in the colon.* Am J Physiol Gastrointest Liver Physiol, 2016. **310**(11): p. G1147-54.
178. Fava, F. and S. Danese, *Intestinal microbiota in inflammatory bowel disease: friend of foe?* World J Gastroenterol, 2011. **17**(5): p. 557-66.
179. Asghar, M.N., et al., *The amount of keratins matters for stress protection of the colonic epithelium.* PLoS One, 2015. **10**(5): p. e0127436.
180. Owens, D.W., et al., *Human keratin 8 mutations that disturb filament assembly observed in inflammatory bowel disease patients.* J Cell Sci, 2004. **117**(Pt 10): p. 1989-99.
181. Buning, C., et al., *Keratin 8 Y54H and G62C mutations are not associated with inflammatory bowel disease.* Dig Liver Dis, 2004. **36**(6): p. 388-91.
182. Tao, G.Z., et al., *Analysis of keratin polypeptides 8 and 19 variants in inflammatory bowel disease.* Clin Gastroenterol Hepatol, 2007. **5**(7): p. 857-64.
183. Helenius, T.O., et al., *Keratins Are Altered in Intestinal Disease-Related Stress Responses.* Cells, 2016. **5**(3).

184. Misiorek, J.O., et al., *Keratin 8-deletion induced colitis predisposes to murine colorectal cancer enforced by the inflammasome and IL-22 pathway.* Carcinogenesis, 2016. **37**(8): p. 777-86.
185. Tatsumi, N., et al., *Cytokeratin 7/20 and mucin core protein expression in ulcerative colitis-associated colorectal neoplasms.* Virchows Arch, 2006. **448**(6): p. 756-62.
186. Stenling, R., et al., *Altered expression of CK7 and CK20 in preneoplastic and neoplastic lesions in ulcerative colitis.* APMIS, 2007. **115**(11): p. 1219-26.
187. Schroeder, K.W., W.J. Tremaine, and D.M. Ilstrup, *Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study.* N Engl J Med, 1987. **317**(26): p. 1625-9.
188. Baron, J.H., A.M. Connell, and J.E. Lennard-Jones, *Variation between Observers in Describing Mucosal Appearances in Proctocolitis.* Br Med J, 1964. **1**(5375): p. 89-92.
189. Drake, P.J., et al., *Application of high-content analysis to the study of post-translational modifications of the cytoskeleton.* J Proteome Res, 2009. **8**(1): p. 28-34.
190. Chen, E.I., et al., *Optimization of mass spectrometry-compatible surfactants for shotgun proteomics.* J Proteome Res, 2007. **6**(7): p. 2529-38.
191. Chen, E.I., et al., *Optimization of mass spectrometry-compatible surfactants for shotgun proteomics.* Journal of proteome research, 2007. **6**(7): p. 2529-38.
192. Tanford, C., *Protein denaturation.* Advances in protein chemistry, 1968. **23**: p. 121-282.
193. Marston, F.A. and D.L. Hartley, *Solubilization of protein aggregates.* Methods in enzymology, 1990. **182**: p. 264-76.
194. Pepinsky, R.B., *Selective precipitation of proteins from guanidine hydrochloride-containing solutions with ethanol.* Analytical biochemistry, 1991. **195**(1): p. 177-81.
195. Herrmann, H., L. Kreplak, and U. Aebi, *Isolation, characterization, and in vitro assembly of intermediate filaments.* Methods in cell biology, 2004. **78**: p. 3-24.
196. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
197. Griffiths, G.J., et al., *Cell damage-induced conformational changes of the pro-apoptotic protein Bak in vivo precede the onset of apoptosis.* The Journal of cell biology, 1999. **144**(5): p. 903-14.
198. Towbin, H., T. Staehelin, and J. Gordon, *Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications.* Proceedings of the National Academy of Sciences of the United States of America, 1979. **76**(9): p. 4350-4.
199. Stensjo, K., et al., *An iTRAQ-based quantitative analysis to elaborate the proteomic response of Nostoc sp. PCC 7120 under N2 fixing conditions.* J Proteome Res, 2007. **6**(2): p. 621-35.
200. Pham, T.K., et al., *A quantitative proteomic analysis of biofilm adaptation by the periodontal pathogen Tannerella forsythia.* Proteomics, 2010. **10**(17): p. 3130-41.

201. Costa, L.C., et al., *Expression of epithelial-mesenchymal transition markers at the invasive front of oral squamous cell carcinoma.* J Appl Oral Sci, 2015. **23**(2): p. 169-78.
202. Satelli, A. and S. Li, *Vimentin in cancer and its potential as a molecular target for cancer therapy.* Cell Mol Life Sci, 2011. **68**(18): p. 3033-46.
203. Chu, P.G. and L.M. Weiss, *Keratin expression in human tissues and neoplasms.* Histopathology, 2002. **40**(5): p. 403-39.
204. Ngan, C.Y., et al., *Quantitative evaluation of vimentin expression in tumour stroma of colorectal cancer.* Br J Cancer, 2007. **96**(6): p. 986-92.
205. Lai, L.A., et al., *Pan-colonic field defects are detected by CGH in the colons of UC patients with dysplasia/cancer.* Cancer Lett, 2012. **320**(2): p. 180-8.
206. Polley, A.C., et al., *Proteomic analysis reveals field-wide changes in protein expression in the morphologically normal mucosa of patients with colorectal neoplasia.* Cancer Res, 2006. **66**(13): p. 6553-62.
207. Evans, C.A., et al., *Reduced keratin expression in colorectal neoplasia and associated fields is reversible by diet and resection.* BMJ Open Gastroenterol, 2015. **2**(1): p. e000022.
208. Osborn, M., *Intermediate filaments as histologic markers: an overview.* J Invest Dermatol, 1983. **81**(1 Suppl): p. 104s-9s.
209. Chaffer, C.L., E.W. Thompson, and E.D. Williams, *Mesenchymal to epithelial transition in development and disease.* Cells Tissues Organs, 2007. **185**(1-3): p. 7-19.
210. Guarino, M., B. Rubino, and G. Ballabio, *The role of epithelial-mesenchymal transition in cancer pathology.* Pathology, 2007. **39**(3): p. 305-18.
211. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.
212. Zeisberg, M. and E.G. Neilson, *Biomarkers for epithelial-mesenchymal transitions.* J Clin Invest, 2009. **119**(6): p. 1429-37.
213. Kwan, R., K. Looi, and M.B. Omary, *Absence of keratins 8 and 18 in rodent epithelial cell lines associates with keratin gene mutation and DNA methylation: Cell line selective effects on cell invasion.* Exp Cell Res, 2015. **335**(1): p. 12-22.
214. Ippolito, C., et al., *Fibrotic and Vascular Remodelling of Colonic Wall in Patients with Active Ulcerative Colitis.* J Crohns Colitis, 2016. **10**(10): p. 1194-204.
215. Thomas, P.A., et al., *Association between keratin and vimentin expression, malignant phenotype, and survival in postmenopausal breast cancer patients.* Clin Cancer Res, 1999. **5**(10): p. 2698-703.
216. Bantscheff, M., et al., *Quantitative mass spectrometry in proteomics: a critical review.* Anal Bioanal Chem, 2007. **389**(4): p. 1017-31.
217. Latosinska, A., et al., *Comparative Analysis of Label-Free and 8-Plex iTRAQ Approach for Quantitative Tissue Proteomic Analysis.* PLoS One, 2015. **10**(9): p. e0137048.
218. Wang, H., S. Alvarez, and L.M. Hicks, *Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two Chlamydomonas reinhardtii strains of interest for biofuels engineering.* J Proteome Res, 2012. **11**(1): p. 487-501.

219. Sjodin, M.O., et al., *Comparative study of label and label-free techniques using shotgun proteomics for relative protein quantification.* J Chromatogr B Analyt Technol Biomed Life Sci, 2013. **928**: p. 83-92.
220. Patel, V.J., et al., *A comparison of labeling and label-free mass spectrometry-based proteomics approaches.* J Proteome Res, 2009. **8**(7): p. 3752-9.
221. Li, Z., et al., *Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos.* J Proteome Res, 2012. **11**(3): p. 1582-90.
222. Ow, S.Y., et al., *iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly".* J Proteome Res, 2009. **8**(11): p. 5347-55.
223. Evans, C., et al., *An insight into iTRAQ: where do we stand now?* Anal Bioanal Chem, 2012. **404**(4): p. 1011-27.
224. Christoforou, A.L. and K.S. Lilley, *Isobaric tagging approaches in quantitative proteomics: the ups and downs.* Anal Bioanal Chem, 2012. **404**(4): p. 1029-37.
225. Westbrook, J.A., et al., *Quantitation with chemical tagging reagents in biomarker studies.* Proteomics Clin Appl, 2015. **9**(3-4): p. 295-300.
226. Treton, X., et al., *Altered endoplasmic reticulum stress affects translation in inactive colon tissue from patients with ulcerative colitis.* Gastroenterology, 2011. **141**(3): p. 1024-35.
227. Majumdar, D., et al., *Keratins in colorectal epithelial function and disease.* Int J Exp Pathol, 2012. **93**(5): p. 305-18.
228. Zajchowski, D.A., et al., *Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells.* Cancer Res, 2001. **61**(13): p. 5168-78.
229. Bukhari, S., et al., *Affinity proteomics led identification of vimentin as a potential biomarker in colon cancers: insights from serological screening and computational modelling.* Mol Biosyst, 2015. **11**(1): p. 159-69.
230. Majumdar, D., et al., *An integrated workflow for extraction and solubilization of intermediate filaments from colorectal biopsies for proteomic analysis.* Electrophoresis, 2012. **33**(13): p. 1967-74.

CHAPTER 7

APPENDICES

Appendix 1

Extraction of intermediate filaments from colorectal biopsies and optimisation of lysis strategies for protein recovery from colonic biopsies [230]

Four different lysis strategies were compared to optimise global protein recovery from colonic pinch biopsies: 1. freeze-thaw, 2. mechanical homogenisation, 3. probe sonication (combined with freeze-thaw and homogenisation) and 4. homogenisation using the Precellys® platform (Bertin Technologies, Villeurbanne, France).Variations in cycles for homogenisation were tested. Individual frozen colonic biopsies were weighed; following lysis using each technique the recovered protein was assayed by Bradford assay (using Bio-Rad protein assay, Hemel Hempsted UK) and the percentage protein recovered (per biopsy) to biopsy weight was used as the endpoint. Since it is not possible to estimate the total amount of protein in a tissue prior to solubilisation, an indirect measure has to be used to determine the yield of extracted proteins. This has been undertaken by measuring percentage solubilisation of total wet weight of the tissue.

Appendix 2

Consent form

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

STH Project No	15550	Patient ID No			-				

CONSENT FORM

Title of Project: Expression and acetylation of keratins 8 and 18 in colonic mucosa of patients with Ulcerative Colitis: role in pathogenesis of colitis associated colorectal neoplasia and dysplasia.

Name of Principal Researcher: Dr Alan Lobo

1. I confirm that I have read and understand the information sheet dated 01/10/2009 Version 1.0) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, without my medical care or legal rights being affected
3. I understand that relevant sections of my medical notes and data collected during the study, may be looked at by individuals from regulatory authorities, or from the NHS Trust, where it is relevant to my taking part in this research, as well as Dr Lobo and Dr Majumdar. I give permission for these individuals to have access to my records.
4. I agree to my GP being informed of my participation in the study.
5. I agree to take part in the above study.

Name of Patient	Date	Signature
Name of Person taking consent (if different from researcher)	Date	Signature
Researcher	Date	Signature

When completed, 1 for patient; 1 for researcher site file; 1 (original) to be kept in medical notes

STH 15550
08/04/2010 Version 2.0

Appendix 3 Participant information sheet (Control patient)

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

PARTICIPANT INFORMATION SHEET (CONTROL PATIENT)

Part 1.

1. **Study title**

Expression and acetylation of keratins 8 and 18 in colonic mucosa of patients with Ulcerative Colitis: role in pathogenesis of colitis associated colorectal neoplasia and dysplasia.

2. **Invitation paragraph**

'You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please feel free to talk to others about the study if you wish.

- Part 1 tells you the purpose of this study and what will happen to you if you take part.
- Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

3. **What is the purpose of the study?**

In this study we are investigating specific chemicals (called keratins) and changes in their chemical structures (termed acetylation) in the large bowel of patients with a condition in their bowel termed ulcerative colitis.

This study aims to try and understand some of the changes that might occur in the bowel lining due to their disease (ulcerative colitis) and how they might relate to previous or current inflammation. Understanding these changes might help us in tailoring our treatment strategies for such patents in the future.

In order to interpret such changes, it is important that we compare the results (from patients with colitis in this study) with patients who do not have any obvious abnormality in their bowel. This will help us to ensure that any changes in the above chemicals (if observed) in the patients with colitis, are not just by chance. You have been requested to participate in the study as a 'control' patient. This means that you do are not known to have the clinical condition under investigation (i.e. ulcerative colitis) nor any other significant abnormality in your bowel. Your doctor however has referred you for a colonoscopy, which is a test to investigate your bowel because of your bowel symptoms. You have been chosen because we do not anticipate any significant abnormality in your bowel at colonoscopy either.

4. **Why have I been chosen?**

Because you are due for a colonoscopy or a flexible sigmoidoscopy for bowel symptoms in the absence of any known previous bowel ailment

5. **Do I have to take part?**

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

If you choose not to take part, you will undergo colonoscopy/flexible sigmoidoscopy following our current practice, and, again, it will not affect your future treatment in anyway.

6. **What will happen to me if I take part?**

If you participate in this study you will have your colonoscopy as usual. It is likely that the individual performing the colonoscopy will obtain a few biopsies (samples from the bowel lining) from different areas in your bowel. The number of such biopsies may vary depending on your clinical condition. For the study we would like to obtain 10 additional biopsies (5 each, from 2 different sites in the lower large bowel). Each biopsy is very small and not felt by the patient. We expect an average of 5 minutes to be added to the test in order to obtain the additional study biopsies. The research biopsies will be processed separately. The standard biopsies will be processed as usual. You will be informed about the results of the biopsies once they had been processed by the pathology laboratory, as we would usually do.

This study is a cross-sectional observational study. This means that you would not be required to attend for any further investigations or make any changes to medications or treatment as a result of this study – though clearly your doctors will want to assess the results of the colonoscopy and standard biopsies.

All data from the study will be recorded for the purposes of ensuring the highest quality procedure. They will be stored on NHS computers with passwords known only to the investigators of the study and in the electronic version will be fully anonymised. All paper versions of records will be kept in the NHS trust.

Expenses and payments:

You will not have to make any additional visits to the hospital, beyond that which you would have made for your planned colonoscopy.

7. **What do I have to do?**

You will not have to do anything different to your normal colonoscopy. You will be sent out a diet sheet and bowel preparation to be taken prior to attending for the colonoscopy. You will attend for the colonoscopy on a day case basis (either morning or afternoon) unless, for clinical reasons, it is needed as an in-patient. You will be given instructions after the colonoscopy – but these will not differ from the usual instructions given to patients after a colonoscopy.

8. **What is the drug, device or procedure that is being tested?**

We are investigating a protein called keratin (particularly two forms of it, keratins 8 and 18) in biopsies from the colon at the time of colonoscopy.

9. **What are the alternatives for diagnosis or treatment?**

The alternative to this study would be for you to undergo colonoscopy and biopsy as usual, without the study biopsies.

10. **What are the side effects of any treatment received when taking part?**

The risks of the procedure are the same as for the colonoscopy that you were originally planned to undergo. These are detailed in the accompanying leaflet explaining the colonoscopy.
Biopsies from the colon can very rarely be associated with bleeding from the site of the biopsy. This risk is very small and has been estimated to vary between 0.008% to 0.03%. Because we are taking additional biopsies, there may be a risk of bleeding associated with this, but we feel that the risk of this is tiny and not significantly greater than if only the standard biopsies are taken. This is supported by recent research, which has shown that extra biopsies taken during endoscopy for research is safe, well tolerated, and does not appear to add additional risk to the procedure.

11. What are the other possible disadvantages and risks of taking part?

None

12. **What are the possible benefits of taking part?**

It is unlikely that you will receive any immediate benefit from taking part in this study. It is hoped however that by performing this study, we will obtain a greater insight into changes in the large bowel in patients with colitis that might result from inflammation, This may lead to new preventative treatments and improve the way that we carry out surveillance in long-standing Ulcerative Colitis and as such improve the quality of service we offer to them in the future.

13. What happens when the research study stops?

On completion of the study the data will be analysed with a view to publication in peer-reviewed medical journals. We also intend to disseminate the information through patient support groups, in particular the Yorkshire cancer research group or National Association for Colitis and Crohn's Disease.

14 What if there is a problem?

If you have any complaints or concerns regarding this study they can be pursued through the Royal Hallamshire PALS (Patient Advice and Liaison) service:

Patient Advice and Liaison Service
Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF
Tel: 0114 2712450

15. Will my taking part in the study be kept confidential?

Yes. All the information about your participation in this study will be kept confidential. The details are included in Part 2.

16. Contact Details:

For further information please contact:

Dr Alan Lobo
Consultant Gastroenterologist
Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF

Tel: 0114 2713160

This completes Part 1 of the Information Sheet.
If the information in Part 1 has interested you and you are considering participation, please continue to read the additional information in Part 2 before making any decision.

Part 2

17 What if relevant new information becomes available?

Not applicable.

18. What will happen if I don't want to carry on with the study?

It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.
If you choose not to take part, you will undergo colonoscopy following our current practice, and, again, it will not affect your future treatment in anyway.

19. What if there is a problem?

Complaints:
If you have any complaints or concerns regarding this study they can be pursued through the Royal Hallamshire PALS (Patient Advice and Liaison) service:

Patient Advice and Liaison Service
Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF

Tel: 0114 2712450

Harm:
As we are taking the biopsies for the study during a planned colonoscopy or flexible sigmoidoscopy, we do not anticipate procedure-related complications other than those normally associated with the procedure.

In the event that something does go wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Sheffield Teaching Hospitals NHS Foundation Trust but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

20. **Will my taking part in this study be kept confidential?**

Your taking part in this study will be totally confidential. Once entered into the study you will be assigned a study number that will be used for the storage of all study results. All data pertaining to the study will be stored on a Trust PC that is password protected. When processing the data, confidential and sensitive data will not be identifiable. All paper versions of study data will be kept in appropriate secure rooms in the Trust as per usual practice. As per current guidance all data will be stored for 10 years to enable retrospective review should it be deemed necessary.

Involvement of the General Practitioner/Family doctor (GP)

A formal report of the colonoscopy, its findings and the results of the regular biopsies will be sent to your GP. This is no different to our standard procedures. The report will specify which type of procedure you underwent.

21. What will happen to any samples I give?

The study samples taken will be assessed for the presence of proteins (keratins) and changes in them (termed acetylation). Samples taken as part of your planned colonoscopy/flexible sigmoidoscopy will be assessed as usual.

22. Will any genetic tests be done?

No.

23. What will happen to the results of the research study?

The results of this study will be submitted to peer review journals for consideration of publication. Results will be disseminated locally through patient support groups, in particular the Yorkshire cancer research group or National Association for Colitis and Crohn's Disease.

24. Who is organising and funding the research?

Sheffield Teaching Hospitals NHS Foundation Trust is sponsoring the study. We aim to obtain further funding for the study.

25. **Who has reviewed the study?**

This study has undergone independent scientific peer review and has been reviewed and approved by the South Yorkshire Research Ethics Committee.

If you wish to take part in this study, you will be given a copy of this leaflet to keep and will need to sign a consent form, a copy of which you will also keep.

Thank you for your time.

Appendix 4 Participant information sheet

Sheffield Teaching Hospitals NHS
NHS Foundation Trust

PARTICIPANT INFORMATION SHEET

Part 1.

1. **Study title**

Expression and acetylation of keratins 8 and 18 in colonic mucosa of patients with Ulcerative Colitis: role in pathogenesis of colitis associated colorectal neoplasia and dysplasia.

2. **Invitation paragraph**

'You are being invited to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please feel free to talk to others about the study if you wish.

- Part 1 tells you the purpose of this study and what will happen to you if you take part.
- Part 2 gives you more detailed information about the conduct of the study.

Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

3. **What is the purpose of the study?**

In patients, like yourself, who have had ulcerative colitis affecting the large bowel (colon) it is currently recommended that you undergo regular examinations of the bowel (colonoscopy) commencing 8-10 years of having the disease. This is to ensure that there is no evidence of change within the bowel, that may be associated with the future development of bowel cancer. Currently this is done by performing a colonoscopy and taking multiple samples (biopsies) throughout the colon, to be looked at under the microscope to ensure there are no signs of cancer or pre-cancerous change. This is of utmost clinical importance, as detection of early cancerous changes in the colon allows earlier and more effective treatment.

This study aims to try and understand some of the steps that might occur in the bowel lining prior to such changes and how they might relate to previous or current inflammation. Understanding these steps should help in preventing or treating these changes in the future.

Several proteins called keratins (and changes in their chemical structure, called acetylation) have been shown in several studies to predict or to be associated with cancerous changes in the bowel. It is well known that the effects of the damage of the bowel in colitis are quite widespread. Hence we aim to develop simple tests looking at these specific proteins in biopsies from a few selected areas in the bowel rather than taking samples throughout the colon, using a much fewer number of biopsies.

These biopsies will be analysed in the laboratory with particular scientific expertise in how to analyse for them, and then related to factors such as the amount of time that you have had colitis, and the amount of inflammation in the bowel.

We will also approach a number of patients without colitis to allow a comparison group.

4. **Why have I been chosen?**

Either

A) You have ulcerative colitis and are due to have a routine surveillance colonoscopy as you have had the disease for a considerable period of time OR
B) You have ulcerative colitis and are due to have a colonoscopy or flexible sigmoidoscopy to assess the activity or extent of inflammation OR
C) There is a strong clinical suspicion that you may have ulcerative colitis and either a flexible sigmoidoscopy or colonoscopy is being done as a confirmatory test

5. **Do I have to take part?**

No. It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.

If you choose not to take part, you will undergo colonoscopy/flexible sigmoidoscopy following our current practice, and, again, it will not affect your future treatment in anyway.

6. **What will happen to me if I take part?**

If you participate in this study you will have your colonoscopy as usual. In addition to the usual biopsies, which would be taken normally (on an average varying between 20-40 biopsies taken randomly throughout the colon), 10 additional biopsies will also be taken for the study (5 each, from 2 different sites in the lower large bowel). Each biopsy is very small and not felt by the patient. We anticipate an average of 5 minutes to be added to the test in order to obtain the additional study biopsies. The research biopsies will be processed separately. The standard biopsies will be processed as usual. You will be informed about the results of the biopsies once they had been processed by the pathology laboratory, as we would usually do.

This study is a cross-sectional observational study. This means we are only looking at the effects of your disease on your bowel at the time of your colonoscopy. You would not be required to attend for any further investigations or make any changes to medications or treatment as a result of this study – though clearly your doctors will want to assess the results of the colonoscopy and standard biopsies.

All data from the study will be recorded for the purposes of ensuring the highest quality procedure. They will be stored on NHS computers with passwords known only to the investigators of the study and in the electronic version will be fully anonymised. All paper versions of records will be kept in the NHS trust.

Expenses and payments:

You will not have to make any additional visits to the hospital, beyond that which you would have made for your planned colonoscopy.

7. **What do I have to do?**

You will not have to do anything different to your normal colonoscopy. You will be sent out a diet sheet and bowel preparation to be taken prior to attending for the colonoscopy. You will attend for the colonoscopy on a day case basis (either morning or afternoon) unless, for clinical reasons, it is needed as an in-patient. You will be given instructions after the colonoscopy – but these will not differ from the usual instructions given to patients after a colonoscopy.

8. **What is the drug, device or procedure that is being tested?**

We are investigating a protein called keratin (particularly two forms of it, keratins 8 and 18) in biopsies from the colon at the time of colonoscopy. Studies have shown that these proteins may be associated with cancerous changes elsewhere in the bowel. Also changes in the above proteins (called acetylation) have been shown in laboratory studies to have an effect on the risk of developing bowel cancers.

9. **What are the alternatives for diagnosis or treatment?**

The alternative to this study would be for you to undergo colonoscopy and biopsy as usual, without the study biopsies.

10. **What are the side effects of any treatment received when taking part?**

The risks of the procedure are the same as for the colonoscopy that you were originally planned to undergo. These are detailed in the accompanying leaflet explaining the colonoscopy.

Biopsies from the colon can very rarely be associated with bleeding from the site of the biopsy. This risk is very small and has been estimated to vary between 0.008% to 0.03%. Because we are taking additional biopsies, there may be a risk of bleeding associated with this, but we feel that the risk of this is tiny and not significantly greater than if only the standard biopsies are taken. This is supported by recent research, which has shown that extra biopsies taken during endoscopy for research is safe, well tolerated, and does not appear to add additional risk to the procedure.

11. What are the other possible disadvantages and risks of taking part?

None

12. What are the possible benefits of taking part?

It is unlikely that you will receive any immediate benefit from taking part in this study. It is hoped however that by performing this study, we will obtain a greater insight into changes that might result from inflammation, and then contribute to the risk of developing colon cancer. This may lead to new preventative treatments and improve the way that we carry out surveillance in long-standing Ulcerative Colitis and as such improve the quality of service we offer to you in the future.

13. What happens when the research study stops?

On completion of the study the data will be analysed with a view to publication in peer-reviewed medical journals. We also intend to disseminate the information through patient support groups, in particular the Yorkshire cancer research group or National Association for Colitis and Crohn's Disease.

14 What if there is a problem?

If you have any complaints or concerns regarding this study they can be pursued through the Royal Hallamshire PALS (Patient Advice and Liaison) service:

Patient Advice and Liaison Service
Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF
Tel: 0114 2712450

15. Will my taking part in the study be kept confidential?

Yes. All the information about your participation in this study will be kept confidential. The details are included in Part 2.

16. Contact Details:

For further information please contact:

Dr Alan Lobo
Consultant Gastroenterologist
Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF

Tel: 0114 2713160

This completes Part 1 of the Information Sheet.
If the information in Part 1 has interested you and you are considering participation, please continue to read the additional information in Part 2 before making any decision.

Part 2

17 What if relevant new information becomes available?

Not applicable.

18. What will happen if I don't want to carry on with the study?

It is up to you to decide whether or not to take part. If you do, you will be given this information sheet to keep and be asked to sign a consent form. You are still free to withdraw at any time and without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive.
If you choose not to take part, you will undergo colonoscopy following our current practice, and, again, it will not affect your future treatment in anyway.

19. What if there is a problem?

Complaints:
If you have any complaints or concerns regarding this study they can be pursued through the Royal Hallamshire PALS (Patient Advice and Liaison) service:

Patient Advice and Liaison Service
Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF

Tel: 0114 2712450

Harm:
As we are taking the biopsies for the study during a planned colonoscopy or flexible sigmoidoscopy, we do not anticipate procedure-related complications other than those normally associated with the procedure.

In the event that something does go wrong and you are harmed during the research study there are no special compensation arrangements. If you are harmed and this is due to someone's negligence then you may have grounds for a legal action for compensation against Sheffield Teaching Hospitals NHS Foundation Trust but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

20. Will my taking part in this study be kept confidential?

Your taking part in this study will be totally confidential. Once entered into the study you will be assigned a study number that will be used for the storage of all study results. All data pertaining to the study will be stored on a Trust PC that is password protected. When processing the data, confidential and sensitive data will not be identifiable. All paper versions of study data will be kept in appropriate secure rooms in the Trust as per usual practice. As per current guidance all data will be stored for 10 years to enable retrospective review should it be deemed necessary.

Involvement of the General Practitioner/Family doctor (GP)

A formal report of the colonoscopy, its findings and the results of the regular biopsies will be sent to your GP. This is no different to our standard procedures. The report will specify which type of procedure you underwent.

21. What will happen to any samples I give?

The study samples taken will be assessed for the presence of proteins (keratins) that may indicate an increased risk of cancer. Samples taken as part of your planned colonoscopy/flexible sigmoidoscopy will be assessed as usual.

22. Will any genetic tests be done?

No.

23. What will happen to the results of the research study?

The results of this study will be submitted to peer review journals for consideration of publication. Results will be disseminated locally through patient support groups, in particular the Yorkshire cancer research group or National Association for Colitis and Crohn's Disease.

24. Who is organising and funding the research?

Sheffield Teaching Hospitals NHS Foundation Trust is sponsoring the study. We aim to obtain further funding for the study.

25. **Who has reviewed the study?**

This study has undergone independent scientific peer review and has been reviewed and approved by the South Yorkshire Research Ethics Committee.

If you wish to take part in this study, you will be given a copy of this leaflet to keep and will need to sign a consent form, a copy of which you will also keep.

Thank you for your time.

Appendix 5 Letter to patient's GP

!

To,

Dr
GP Surgery

Dear *Dr*

Re: *Patient name*

Date of birth: *date of birth*

The Department of Gastroenterology, Sheffield Teaching Hospitals NHS Foundation Trust is currently undertaking a study entitled 'Expression and acetylation of keratins 8 and 18 in colonic mucosa of patients with Ulcerative Colitis: role in pathogenesis of colitis associated colorectal neoplasia and dysplasia'.

This is an exciting, novel approach to examining why patients with colitis might develop colon cancer. It combines the clinical research background of our department with the basic science expertise of a University laboratory with a range of novel molecular techniques to address the questions raised in our study. This is a cross sectional study. Patients with proven or suspected ulcerative colitis undergoing colonoscopy or flexible sigmoidoscopy will have additional biopsies taken, apart from the standard biopsies. These biopsies will be then analysed in specialised laboratory at the University of Sheffield.

Your patient, *patient name*, has agreed to take part in the study. Participation however, will not involve any change in their therapy or in their overall treatment strategy. They will be given adequate information regarding the study including printed patient information leaflets.

If you would like any further information about this project, please contact me using the details below.

Yours sincerely

Dr Alan J Lobo
Principal Investigator
Consultant Gastroenterologist and Clinical Lead for Gastroenterology
Royal Hallamshire Hospital
Sheffield Teaching Hospitals NHS Foundation Trust

Email: alan.lobo@sth.nhs.uk
Tel: 0114 2712353
Fax:0114 2713160!

STH 15550
Version 1.0, 1/2/2010

Appendix 6 Letter to study participant

!

To,

Patient name
Address

Dear *patient*

Sub: Participation in a clinical study in ulcerative colitis

I am about to start a research project entitled 'Expression and acetylation of keratins 8 and 18 in colonic mucosa of patients with Ulcerative Colitis: role in pathogenesis of colitis associated colorectal neoplasia and dysplasia' at the Department of Gastroenterology, Sheffield Teaching Hospitals NHS Foundation Trust.

This is an exciting, novel approach to examining why patients with colitis might develop colon cancer. It combines the clinical research background of our department with the basic science expertise of a University laboratory with a range of novel molecular techniques to address the questions raised in our study.

You have been contacted for participation in the study, since you undergo regular examination of your bowel (colonoscopy) to ensure that there is no evidence of change within the bowel that may be associated with the future development of bowel cancer.

I am attaching the following documents along with this letter.

1. Detailed patient information sheet
2. Consent form

Please take time to read all the information provided. If you have any questions at any stage feel free to contact me at the contact details given below. One of the research investigators will contact you shortly before the colonoscopy and will answer any questions that you may have regarding the study. If you are agreeable to participate he will obtain written consent from you. Please remember, that participation (or non-participation) in the study will not affect your treatment or endoscopic examination in any way.

With kind regards

Yours sincerely

Dr Alan J Lobo
Principal Investigator
Consultant Gastroenterologist and Clinical Lead for Gastroenterology

Royal Hallamshire Hospital
Sheffield Teaching Hospitals NHS Foundation Trust

Email: alan.lobo@sth.nhs.uk
Tel: 0114 2712353
Fax:0114 2713160

Appendix 7 Approval from the Ethics committee

National Research Ethics Service

South Yorkshire Research Ethics Committee
1st Floor Vickers Corridor
Northern General Hospital
Herries Road
Sheffield
S5 7AU

Telephone: 0114 226 9153
Facsimile: 0114 256 2469
Email: joan.brown@sth.nhs.uk

22 April 2010

Dr Debabrata Majumdar
Department of Gastroenterology
P Floor, Royal Hallamshire Hospital
Glossop Road
Sheffield
S10 2JF

Dear Dr Majumdar

Study Title:	**Expression and acetylation of keratins 8 and 18 in colonic mucosa of patients with Ulcerative Colitis: role in pathogenesis of colitis associated colorectal neoplasia and dysplasia.**
REC reference number:	**10/H1310/21**
Protocol number:	**1.0**

Thank you for your letter of 12 April 2010responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

<u>Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.</u>

This Research Ethics Committee is an advisory committee to Yorkshire and The Humber Strategic Health Authority
The National Research Ethics Service (NRES) represents the NRES Directorate within the National Patient Safety Agency and Research Ethics Committees in England

For NHS research sites only, management permission for research ("R&D approval") should be obtained from the relevant care organisation(s) in accordance with NHS research governance arrangements. Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at http://www.rdforum.nhs.uk. *Where the only involvement of the NHS organisation is as a Participant Identification Centre, management permission for research is not required but the R&D office should be notified of the study. Guidance should be sought from the R&D office where necessary.*

Sponsors are not required to notify the Committee of approvals from host organisations.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	*Version*	*Date*
Covering Letter		01 March 2010
REC application		02 March 2010
Protocol	1.0	01 September 2009
Investigator CV		15 February 2010
Student CV - Dr Debabrata Majumdar		01 March 2010
Academic Supervisor's CV - Dr Bernard Corfe		01 March 2010
Letter of invitation to participant	1.0	01 March 2010
Referees or other scientific critique report		
GP Letter	1.0	01 February 2010
Flow Chart	1.0	01 March 2010
Participant Information Sheet	2.0	08 April 2010
Participant Information Sheet: Controls	2.0	08 April 2010
Participant Consent Form	2.0	08 April 2010
Covering letter		13 April 2010

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Now that you have completed the application process please visit the National Research Ethics Service website > After Review

You are invited to give your view of the service that you have received from the National Research Ethics Service and the application procedure. If you wish to make your views known please use the feedback form available on the website.

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

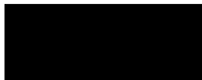
- Notifying substantial amendments
- Adding new sites and investigators
- Progress and safety reports
- Notifying the end of the study

The NRES website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

We would also like to inform you that we consult regularly with stakeholders to improve our service. If you would like to join our Reference Group please email referencegroup@nres.npsa.nhs.uk.

10/H1310/21	Please quote this number on all correspondence

Yours sincerely



pp **Miss Jo Abbott**
Chair

Enclosures: "After ethical review – guidance for researchers" SL-AR2

Copy to: STH R&D Department

Dr Alan J Lobo, Consultant Gastroenterologist and Clinical Lead for Gastroenterology, RHH

This Research Ethics Committee is an advisory committee to Yorkshire and The Humber Strategic Health Authority

The National Research Ethics Service (NRES) represents the NRES directorate within The National Patient Safety Agency and Research Ethics Committees in England

CHAPTER 8

TABLES AND FIGURES

8.1 List of Tables

1.1 Montreal classification of ulcerative colitis based on disease extent in the colon.

1.2: Classification of dysplasia in Ulcerative colitis

2.1 Flowchart of investigation in ulcerative colitis patients

2.2.1.1 Primary antibodies

2.2.1.2 Secondary antibodies

2.2.2 Reagents for protein assay

2.2.3 Reagents for cell culture

2.2.4 Materials and reagents for western immunoblotting and gel analysis

2.2.5 Reagents for mass spectrometry

2.2.6.1 Phosphate buffer saline (pH 7.4)

2.2.6.2 Running buffer

2.2.6.3 Wash buffer

2.2.6.4 Low Detergent Buffer (LDB)

2.2.6.5 High detergent buffer (HDB)

2.2.6.6 Intermediate filament protein solubilising buffer

2.2.6.7 Buffers for strong cation exchange

    2.2.6.7.1 Buffer A

    2.2.6.7.2 Buffer B

2.2.6.8 Buffers for HPLC

    2.2.6.8.1 Buffer A

    2.2.6.8.2 Buffer B

2.2.6.9.1 Tris-Buffered Saline and Tween 20 (TBS-T)

2.2.6.9.2 Phosphate-buffered saline (1x), pH 7.4

2.2.7 SDS PAGE solution for preparing gels (12%)

2.2. Scientific equipment

2.4 Endoscopic equipment / materials required

2.2 Steps for isolating and analysing IF

2.3 Summary of patients recruited in the study

2.4 Patient groups categorised according to duration of disease

2.5 Median disease activity as assessed clinically (endoscopically) and histologically

2.6 Patients with distal active distal disease

2.7 Recent onset UC (ROUC)

2.8 Long standing pancolitis (LSPC)

2.9 Primary sclerosing cholangitis and UC

2.10 Patients with pancolitis and dysplasia

3.1 Differentially regulated proteins from colorectal biopsies identified on mass spectrometry, in patients with ulcerative colitis compared to controls.

3.2 Mass spectrometry results summarising fold changes in levels of key IF proteins and associated proteins between different patient groups.

3.3 Mass spectrometry results showing significant fold changes in levels of intermediate filament proteins and associated proteins in acutely inflamed mucosa compared to controls.

8.2 List of Figures

Figure 1.1 Various factors, which interact and contribute to the pathogenesis of ulcerative colitis in genetically susceptible individuals.

Figure 1.2 The central role of proteomics in identifying expression, alteration and functions of proteins in biological tissues.

Figure 1.3 Proteomic analysis of biological samples commences with sample preparation.

Figure 1.4. Keratin 8 molecule with phosphorylation and acetylation sites.

Figure 2.1 Steps for isolation of insoluble intermediate filament proteins from colorectal biopsies.

Figure 2.2 Coomassie stained gel comparing band intensity of cytoskeletal fraction of MCF-7 cells dissolved in 10M urea and 4M Guanidine hydrochloride(GuHCl).

Figure 2.3 Fraction 3 proteins were dissolved in urea (10M) and GuHCl (4M and 6M respectively). Band intensity with protein dissolved in 6M GuHCl was less than 4M GuHCl, due to formation of precipitates with SDS. The band intensity was greater with 10M urea than 4M GuHCl.

Figure 2.4 The band intensity of separated fraction 3 proteins dissolved in GuHCl (4M) and urea (6M) were noted to be similar (after removal of the solvent).

Figure 2.5 Following trial of GuHCl with in vitro material, the effectiveness of GuHCl as solvent for colonic biopsies was tested. Protein content in each band (from the dissolved sample) was satisfactory (Lane 1, 2 and 3 respectively). Following further attempt at sonication and dissolving the sediment this was analysed for protein content. A negligible amount of protein was noted in the sonicated samples (Lane 1B, 2B and 3B respectively)

Figure 2.6 Demonstration of Keratin 8 expression (from 3 colonic biopsies)(A), and post translational modification (acetylation) in keratin 8 using anti acetyl lysine antibodies (K10 and K482 respectively) to K8(B). Phosphorylation has also been shown using phospho-serine threonine antibody (PST)(C).

Figure 2.7 Essential steps in developing a technique for extracting and dissolving IF proteins using MCF 7 cell lines and colorectal biopsies to allow analysis using SDS-PAGE and mass spectrometry.

Figure 2.8 Steps in processing and solubilising colonic biopsies

Figure 2.9 Proteomic workflow for analysing colorectal biopsies.

Figure 2.10. Timelines with activities undertaken since the start of the project have been summarised.

Figure 3.1 Representative 8-plex iTRAQ labeled peptide.

Figure 3.2 Mass spectrometry results showing significant fold changes of IF proteins and associated proteins between different patient groups (LSPC, ROUC, and PSC and those with dysplasia-DR and DT respectively) in comparison to control.

Figure 3.3 Effect of duration of on mucosal levels of IFs.

Figure 3.4 Proteomic changes in patients with PSC and dysplasia.

Figure 3.5 Proteomic changes in dysplastic tissue (DT) and surrounding non-dysplastic field mucosa (DR).

Figure 3.6 (a) iTRAQ data showing significant changes (decreased fold changes) in intermediate filament protein levels in colonic mucosa in acute inflammation compared to proximal un-inflamed mucosa. **(b)** colonic mucosa showing inflamed and un-inflamed segments

Figure 3.7 Dendogram demonstrating clustering of patient groups (V2 CON, V3 ACT, V4 INACT, V5 LSPC, V6 ROUC, V7 PSC, V8 DR, V9 DT) based on proteomic characteristics.

Figure 4.1 Coomassie stained gel demonstrating equivalent loading of the gel with samples (5μG protein loaded into each lane).

Figure 4.2 Results of western immunoblotting with antibody to K8, K18, K19 and vimentin showing alteration in intermediate filament protein levels in colonic mucosa in (a) patients with longstanding, recent onset disease, PSC and in those with dysplasia and (b) in patients with active distal disease.

Figure 4.3 Western immunoblotting using K8 antibodies demonstrating multiple isoforms of K8, including in control patients. Isoforms in LSPC are predominant in higher molecular weight.

Figure 4.4 Western immunoblotting using antibodies to phosphorylated K8 (K8-pS23, K8-pS73 and K8pS431).

Figure 4.5 Western immunoblotting demonstrating acetylation of K8 (K8-acetyl K10) in LSPC group.

Figure 4.6 Densitometry analysis comparing phosphorylated K8 (pS23) to total K8.

Figure 4.7 Densitometry analysis comparing phosphorylated K8(pS73) to total K8.

Figure 4.8 Densitometry analysis comparing phosphorylated K8 (pS431) to total K8.

Figure 4.9 Densitometry analysis of K8 and VIM in different patient groups.

Figure 5.1 Proposed pathway of pathogenesis of colorectal cancer in ulcerative colitis.