

Cranfield University
Institute of BioScience and Technology

PhD Thesis
Academic Years 1998-2001

Sarah Louise Morgan

Towards the Molecular Diagnosis of Bladder and Colorectal
Cancer : Analysis of CD44 Exon Splicing

Supervisor : Dr Anthony C Woodman

August 2002

This thesis is submitted in partial fulfilment of the requirements for the
degree of Doctor of Philosophy

©Cranfield University 2002. All rights reserved. No part of this publication may be
reproduced without the written permission of the copyright owner.

ProQuest Number: 10821001

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10821001

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by Cranfield University.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

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

Abstract

The literature suggests that CD44 has the potential to be used as a non-invasive tumour specific marker for early detection and monitoring. The gene transcript can undergo alternative splicing, leading to the production of a number of isoforms and it has been noted that in neoplastic states this splicing becomes aberrant, leading to the production of a number of variant isoforms. The aim of this study was to analyse CD44 variant expression with the goal of reaching a platform from which a non-invasive assay for routine clinical use could be produced.

In this study, the analysis of exon junction splicing in bladder cancer has also led to the finding of a tumour specific junction expressed in 64% of tumours studied. This exon junction (5/11) was found to be the same as that expressed in colorectal cancer. Though mRNA based assays would provide a satisfactory method for exon junction identification, a protein based approach would provide the platform for a more robust assay.

Translation of gene expression data led to the design of a short synthetic peptide which overlapped the 5/11 junction, and this was used to produce a novel polyclonal antibody for this transitional epitope. The resulting polyclonal antibody was affinity purified and used in a pilot study of 5/11 expression in bladder and colorectal tumours.

In bladder tumours the antibody demonstrated an overall specificity of 76.5% and an overall sensitivity of 73.1 %, comparable to current commercially available early detection assays. In a Dukes staged colorectal cancer study, no link between stage and grade was noted, but the antibody gave an overall specificity of 76%.

Though further analysis is required, it is thought the 5/11 antibody may prove to be a useful tool in the development of a sensitive and specific assay for the non-invasive detection of bladder and colorectal cancer.

Acknowledgements

I would firstly like to thank my supervisor Dr Anthony Woodman for all his help and support through my PhD, for trying to keep me sane and nearly managing it !

Thank you to my research collaborators and sponsors in Northampton General Hospital, Mr Alan Berry and Mr Marek Miller, along with Mr Roger Kunckler and numerous registrars who have aided in my sample collection, especially Jon McFarlane. I would also like to thank the NGH theatre staff and Dr Angus Molyneux for their help.

Thank you to Heidi for also keeping me sane and also to everyone else in the biomedical centre.

Special mention to Dr Jane Rixon, who read parts of my thesis when she should have had her feet up enjoying her pregnancy.

Thanks also to my Mum and Dad, who have always been there for me, and have helped in so many ways.

To Lee thank you for all your love, help and support through the past three years – it would have been difficult without you.

Finally, thanks to all others at Cranfield who have helped me along the way – you're too numerous to name, but, you know who you are !

Contents Page

Abstract	i
Acknowledgements	ii
Contents Page	iii
List of Figures	vi
List of Tables	vii
List of Plates	vii
List of Abbreviations	viii
1 Introduction	1
1.1 Introduction to cancer biology	1
1.1.1 The molecular basis of cancer	2
1.1.2 Pathology of Cancer	8
1.1.3 Diagnosis, detection and screening	9
1.1.4 Tumour markers	11
1.2 CD44	12
1.2.1 Background	12
1.2.2 The CD44 gene	13
1.2.3 Alternative splicing	14
1.2.4 CD44 protein structure	16
1.2.5 CD44 ligands	18
1.2.6 CD44 functions	18
1.2.7 CD44 and cancer	19
1.3 Bladder Cancer	23
1.3.1 Aetiology of bladder cancer	23
1.3.2 The molecular basis of bladder cancer	26
1.3.3 Pathology and staging of bladder cancer	29
1.3.4 Diagnosis of bladder cancer	31
1.3.5 Urinary markers in research	34
1.3.6 CD44 and bladder cancer	37
1.4 Colorectal Cancer	40
1.4.1 Aetiology of colorectal cancer	40
1.4.2 The molecular basis of colorectal cancer	42
1.4.3 Pathology and staging of colorectal cancer	44
1.4.4 Detection and diagnosis of colorectal cancer	45
1.4.5 CD44 and colorectal cancer	49
1.5 CD44 as a marker for the monitoring of recurrence	51
1.6 Aims and Objectives	53

2 Evaluation of CD44 expression in human <i>in vitro</i> models	56
2.1 Introduction	56
2.2 Materials and Methods	58
2.2.1 Cell culture	58
2.2.2 mRNA extraction	59
2.2.3 cDNA production by Reverse Transcription (RT)	60
2.2.4 Standard CD44 PCR	60
2.2.5 Production of standard probe for Southern Blotting	61
2.2.6 Purification of probes from LMP agarose	62
2.2.7 Southern Blotting	63
2.2.8 Protein extraction	64
2.2.9 Immunohistochemistry	65
2.2.10 Western Blotting	66
2.3 Results	68
2.3.1 Transcriptional studies	68
2.3.2 Translational studies	70
2.4 Discussion	78
3 Analysis of CD44 variant exon splicing in bladder cancer	84
3.1 Introduction	84
3.2 Materials and Methods	86
3.2.1 Sample procurement	86
3.2.2 Sample preparation for mRNA extraction	87
3.2.3 mRNA extraction	87
3.2.4 cDNA production	87
3.2.5 Exon link assay	87
3.2.6 PCR-ELISA	88
3.3 Results	90
3.3.1 Exon link assay on RT112 cell line	90
3.3.2 Exon link assay on T24 cell line	92
3.3.3 Exon link on tissues samples	93
3.3.4 Choice of PCR-ELISA probe	93
3.3.5 PCR-ELISA results	96
3.4 Discussion	96
4 Production and characterisation of a polyclonal antibody against a CD44 transitional epitope	101
4.1 Introduction	101
4.2 Materials and methods	102
4.2.1 Peptide design	102
4.2.2 Search for peptide homology	103

4.2.3	Peptide synthesis	103
4.2.4	Polyclonal antibody production	104
4.2.5	Antibody-antigen titration assay	105
4.3	Results	106
4.4	Discussion	108
5	Evaluation of 5/11 junction protein in cell lines and clinical samples	111
5.1	Introduction	111
5.2	Materials and methods	112
5.2.1	Sample collection	112
5.2.2	Immunohistochemistry	112
5.2.3	Western blotting	114
5.3	Results	
5.3.1	5/11 expression in <i>in vitro</i> model	115
5.3.2	5/11 expression in bladder cancer	115
5.3.3	5/11 expression in colorectal cancer	117
5.4	Discussion	126
6	Final Discussion	133
7	Conclusions and further work	139
	References	140
	Appendices	157
Appendix A	Ethical approval and consent forms	157
Appendix B	Solution formulations	161
Appendix C	Miscellaneous protocols	164
C1	Transformation of <i>E. Coli</i> cells	164
C2	Production of competent cells	164
C3	Synopsis of maxiprep plasmid extraction kit	165
C4	Plasmid Linearisation	165
C5	Tissue Sectioning	165
C6	De-waxing, re-hydration and de-hydration of paraffin embedded tissue sections	166
Appendix D	PCR primers	167

List of Figures

	Page
1.1 Pathways of signal transduction, cell cycle control and apoptosis	4
1.2 Structure of the CD44 gene, demonstrating the presence of both standard and variant exons	14
1.3 The pathway from gene to protein via mRNA	15
1.4 Diagrammatic representation of the CD44 protein structure	17
1.5 A model for the progression of transitional cell carcinoma (TCC)	27
1.6 Diagram demonstrating the TNM staging system in bladder cancer	30
1.7 Genetic model for colorectal tumourigenesis	43
2.1 Southern blot by capillary action with 0.4M NaOH	63
2.2 Western blot setup within the semi-dry blotter	67
2.3 Ethidium bromide stained gel of RT112 and T24 cell line RT-PCR	68
2.4 Ethidium bromide stained gel of HT29 and CACO2 cell line RT-PCR	68
2.5 Southern blot of ethidium gel in figure 2.3	70
2.6 Southern blot of ethidium gel in figure 2.4	70
2.7 Western blot with Hermes antibody	76
2.8 Western blot with V3 antibody	76
2.9 Western blot with V5 antibody	77
2.10 Western blot with V6 antibody	77
3.1 Diagram showing primers design for exon link assay	84
3.2 Result of the RT112 exon link assay on an ethidium bromide stained gel	90
3.3 Southern blot of the ethidium gel in figure 3.2	91
3.4 Result of the T24 exon link assay on an ethidium bromide stained gel	92
3.5 Southern blot of the ethidium gel in figure 3.4	92
4.1 Flow diagram showing design from primer sequence to peptide sequence.	103
4.2 Graph showing titration of 5/11 antibody with KLH-peptide	107
4.3 Graph showing titration of 5/11 antibody with BSA-peptide	107
4.4 Graph showing titration of 5/11 antibody with free peptide	107
5.1 Western blot of cell lines with 5/11 antibody	117

List of Tables

	Page
1.1 A comparison of TNM classification ad Dukes stage in colorectal cancer	45
2.1 Analysis of CD44 immunohistochemical results within the RT112, T24, CACO2 and HT29 cell lines	71
3.1 PCR-ELISA results in tumour and non-involved tissue for 5/8, 5/11 and 5/12 probes	94
3.2 Summary of results from table 3.1, along with the results from the normal group	95
5.1 Summary of staining density results for cell lines with 5/11 antibody	115
5.2 Summary of the results of the IHC with 5/11 antibody in bladder sections of varying stage and grade	118
5.3 Summary of western blot results with 5/11 antibody	121
5.4 Summary of results of IHC with 5/11 antibody in Dukes staged colorectal sections	122
5.5 Comparison of sensitivity and specificity values of the 5/11 antibody with commercial FDA approved urinary based assays used in clinical practice	129

List of Plates

	Page
Plate 1 Results of immunohistochemistry with anti-CD44 antibodies on RT112 cells	72
Plate 2 Results of immunohistochemistry with anti-CD44 antibodies on T24 cells	73
Plate 3 Results of immunohistochemistry with anti-CD44 antibodies on HT29 cells	74
Plate 4 Results of immunohistochemistry with anti-CD44 antibodies on CACO2 cells	75
Plate 5 Immunohistochemical staining of all four cell lines with 5/11 antibody	116
Plate 6 Bladder tumour sections stained with 5/11 antibody (1/25 dilution)	119
Plate 7 Colorectal tumour section stained with 5/11 antibody (1/25 dilution)	124

List of Abbreviations

BSA	– Bovine Serum Albumin
CD	– Cluster of Differentiation
cDNA	– Complementary DNA
CEA	– Carcinoembryonic Antigen
CHAPS	– 3-[3-(Cholamidopropyl)dimethylammonio]-1-proanesulfonate
Cis	– Carinoma <i>in situ</i>
DAB	– Diaminobenzidine
DCC	– Deleted in Colon Cancer
Dig	- Digoxigenin
DMBA	– Dimethylbenzoic Acid
DNA	– deoxyribonucleic acid
ECM	– Extracellulat Matrix
EDTA	– Ethylenediaminetetraacetic acid
ELISA	– Enzyme Linked Immunosorbent Assay
FAP	– Familial Adenomatous Polyposis
FCA	– Freunds Complete adjuvant
FDP	– Fibrin Degradation Product
FOB	– Faecal Occult Blood
H and E	– Haematoxylin and Eosin stain
H ₂ O	– Water
HA	– Hyaluronan
Hb	- Haemoglobin
HNPCC	– Hereditary Non-Polyposis Colorectal Cancer
hnRNA	– Heteronuclear RNA
HRP	– Horseradish Peroxidase
IBD	– Irritable bowel disease
ICAM	– Immunoglobulin-like Cell Adhesion Molecule
ICC	– Immunocytochemistry
IgG	– Immunoglobulin G
IHC	– Immunohistochemistry
KLH	– Keyhole Limpet Haemacyanin

LMP – Low Melting Point (agarose)
mAb – Monoclonal antibody
MAPK – Map kinase
MCM – Mini Chromosomal Maintenance Protein
MgCl₂ – Magnesium Chloride
mRNA – Messenger RNA
NaCl – Sodium Chloride
NaOH – Sodium Hydroxide
NCAM – Neural cell adhesion molecule
NCBI – National Centre for Bioinformatics Information
NMP22 – Nuclear Matrix Protein 22
NUFA – Non-Ulcerative Freund's Adjuvant
NUMA – Nuclear Matrix Apparatus
pAb – Polyclonal antibody
PAH – Polycyclic Aromatic Hydrocarbon
PBS – Phosphate Buffered Saline
PCR – Polymerase Chain Reaction
PEG – Polyethylene Glycol
PSA – Prostate Specific Antigen
RB – Retinoblastoma
RFLP – Restriction Fragment Length Polymorphism
RNA – Ribonucleic Acid
RT-PCR – Reverse Transcription – Polymerase Chain Reaction
SDS-PAGE – Sodium Dodecyl Sulphate – Polyacrylamide Gel Electrophoresis
TBS – Tris Buffered Saline
TCC – Transitional Cell Carcinoma
Tis – Tumour *in situ*
TNF – Tumour Necrosis Factor
TNM – Tumour, Node, Metastases (Tumour classification)
TRAP – Telomeric Repeat Amplification Protocol
UICC – International Union against Cancer
VUC – Voided Urine Cytology
WHO – World Health Organisation

1 Introduction

In 2001 the National Statistics Office published a report examining the trends in cancer rates in England and Wales between 1950 and 1999. A comparison of the mortality rates from cancer and those from three other major causes of death (heart disease, stroke and infectious disease) demonstrated that mortality from cancer exceeds that found with the other conditions. This is not due to a significant increase within mortality rates, but due to a decrease within mortality rates from the three other named conditions. This suggests that though changes in the detection and management of these conditions have been effective, much work is still needed in the field of cancer (National Statistics Office, 2001).

Within the UK, cancer is predominantly a disease of the elderly, and with an increasing population age, mortality rates could be set to increase. As our scientific knowledge about cancer develops, so our ability to treat or slow down its progress becomes greater. We cannot yet eliminate this condition, therefore we need to have a method of detecting the disease at an early stage, before it becomes too advanced to treat.

1.1 Introduction to cancer biology

Cancer has been described as “a disease of uncontrolled proliferation” (Rubin and Farber, 1995). Though this may appear true from a pathological respect, this description gives no indication as to the complexity of the underlying factors involved in its development.

To convey the intricate nature and to present the wide ranging research required to elucidate all mechanisms involved in the development of cancer, an overview of tumour formation, including cellular control in both normal and tumour states is required and described herein. It is the key regulators within these mechanisms that can be used as markers for the presence of cancer, and which can be exploited in assays for early detection and management through the monitoring of recurrence.

1.1.1 The molecular basis of cancer

The etiologic basis of cancer is broadly divided between those that arise due to an inherited susceptibility and those that do not. It is the latter sporadic form which predominates in all populations, and is caused due to the interaction of a number of both internal and environmental factors, through a process described as carcinogenesis.

Carcinogenesis

Carcinogenesis is a multistep process which constitutes the first step towards neoplastic tumour development. The first step within the carcinogenic pathway is initiation. It is at this point that an initiating factor (or carcinogen) whether chemical, form of radiation or viral, generates a genetic mutation within a cell. This mutation does not take effect until a secondary factor, a promoting agent, is present. These factors stimulate increased growth and proliferation, resulting in the presence of an increased number of mutated cells (King, 2000).

This mechanism was elucidated through the use of skin painting experiments in mice, and these exemplify the simplest carcinogenic process. Classically DMBA (a polycyclic hydrocarbon) when painted onto the skin generates an inheritable mutation. No phenotypic effect is seen until croton oil (the promoting agent) is painted on, weeks after the initial DMBA treatment. The croton oil induces signalling cascades within the cell which increase its proliferative ability.

Initiation is therefore the point at which the primary genetic mutation occurs. Cells may differ genotypically, but phenotypically are no different to those in the surrounding population. Cells may remain in this initiated yet latent stage for a number of years until acted upon by a promoting agent. These are numerous, and vary from cell specific factors to hormones and normal growth promoters. The aim of the promoting agent is to stimulate and then increase a cells proliferative ability. This may involve the effects of a number of agents over a period of time, rather than one single event. Once a cell begins to proliferate it enters the stage of progression, where an increasing number of mutations and loss of growth regulation lead to eventual tumour formation if left unchecked.

Cell growth and proliferation

Cellular growth and proliferation in normal tissues is highly regulated. The cell cycle (Figure 1.1b) is central to this and is composed of four main phases G1, S, G2 and M, though cells may also undergo a G0 or quiescent phase where cells are neither growing or proliferating. Of these stages S and M could be considered the most important, as it is during S phase that DNA synthesis occurs, in preparation for the cell to divide during M (Mitosis) phase. G1 and G2 are growth phases during which the cell grows and accumulates the necessary proteins for S and M phases to proceed. The cell cycle is controlled through the interaction of three protein families – cyclins (phosphatases), cyclin-dependent serine / threonine kinases and cyclin-dependent-kinase inhibitors. The cyclin-dependent kinases are activated via the cyclins, or inhibited via the kinase inhibitors. These proteins are present throughout the cell cycle, with different family members present at each phase (Pardee, 1989).

A cell enters G1 when stimulated to grow by cell-specific growth factors, hormones or cell contacts. Stimulatory signals induce signalling cascades within the cell, resulting in transcription of regulatory proteins for cell cycle progression to begin. During G1 a cell may undergo a number of growth and arrest periods before passing the restriction point (R) at which a cell commits to division (Pardee, 1989; Hartwell and Weinert, 1989). Before entering S phase the cell must traverse a “checkpoint” where the inhibitory force of Rb (the protein product of the retinoblastoma gene) is removed and to ensure DNA integrity before synthesis at S phase. The protein p53 mediates this process and delays cell cycle progression if DNA damage is detected. This occurs via stimulation of p21 protein to inhibit DNA synthesis, and subsequently to stimulate DNA repair mechanisms. If the damage is irreparable then p53 mediates the entry of cells into apoptosis (cell death). Progression through S phase is rapid, to minimise the possibility of DNA damage due to external factors during replication. A second checkpoint is present between G2 and M, to assess DNA integrity after S phase. Any aberrant cells will not enter M phase, but will enter the apoptotic pathway, again mediated by p53 (Hartwell and Weinert, 1989).

A number of signal transduction pathways may be responsible for the signalling cascades which trigger cell growth (figure 1.1a). The most well characterised of all pathways is the Ras / Map kinase pathway, which involves binding of a signalling

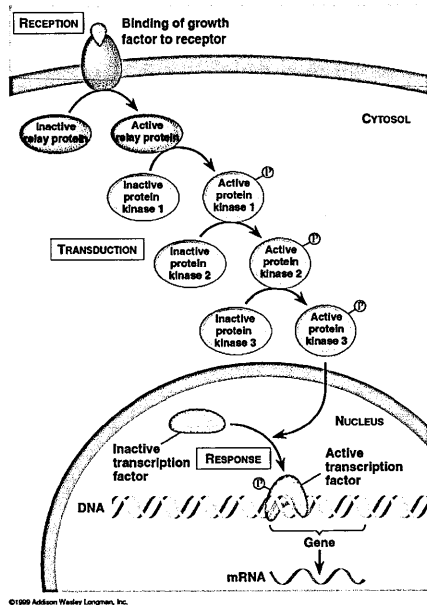


Figure 1.1a – Generalised signal transduction pathway, demonstrating the multiple phosphorylation and de-phosphorylation events which transduce the growth factor signal, causing an increase in gene expression

Figure adapted from :
http://departments.oxy.edu/.../11-27-00_files/image018.jpg

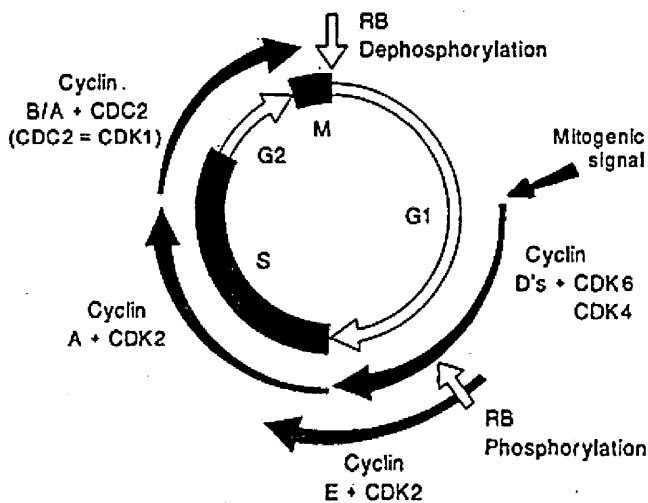


Figure 1.1b – The cell cycle. This is regulated through the interaction of two protein families, the cyclins and the cyclin dependent kinases. Different members of these families are present through each stage of the cycle.

Figure adapted from :
www.entu.cas.cz/fyziol/students/mammals.jpg

Figure 1.1c – Apoptotic pathway. If the cellular DNA is found to be irreparably damaged at either the G1/S or the S/G2 checkpoint within the cell cycle, the cell enters the apoptotic pathway via the influence of p53.

Figure adapted from :
www.bioscience.org/2000/v5/d/somasound/fulltext.htm

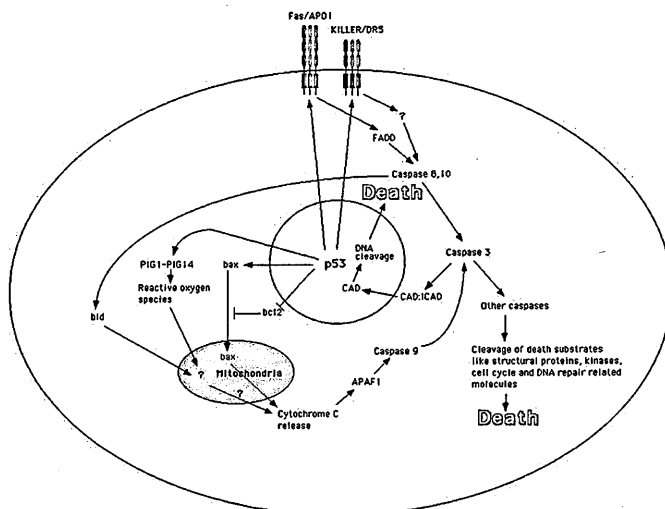


Figure 1.1 – Pathways of signal transduction, cell cycle control and apoptosis

molecule to a specific transmembrane receptor. This activates a Ras protein by phosphorylation, either through activation of a G-linked protein or via intrinsic tyrosine kinase activity. Ras is a monomeric GTPase which is membrane-bound, and activation of this protein starts a serine / threonine phosphorylation cascade within the cytoplasm via the Raf and Mek protein kinases, and a tyrosine / threonine phosphorylation, to activate Map kinase. Activation of Map kinase (MAPK) allows it to enter the nucleus where it can phosphorylate specific transcription factors such as Myc, Fos and Jun (Lodish *et al*, 1995).

In normal tissues, cellular proliferation does not occur on its own, but occurs in balance with cellular death. Apoptosis is the process by which cells which have sustained lethal or irreversible damage may be removed from tissue without causing damage to surrounding cells. It is a highly regulated process which can be divided into three phases – induction, commitment and degradation.

Apoptosis

A cell may enter apoptosis due to a number of inducing factors including a decrease or complete withdrawal of growth signals, the influence of “death signals” from TNF (Tumour Necrosis Factor) and Fas ligand, or via p53 due to DNA damage detected during the cell cycle (figure 1.1.c). The movement from the induction to commitment phase occurs due to the activation of one of a group of proteins known as the caspases (cysteine aspartate proteins) within the cytoplasm. This caspase cleaves a protein, Bid, which then enters the mitochondrial membrane where it interacts with a group of regulatory proteins – the BCl family. This family consists of BCl2, an anti-apoptotic protein, and the pro-apoptotic proteins, BAX and Bad. If the concentration of BAX is high enough, then the cell will enter the degradation phase from which it cannot return; whilst if BCl2 is present in a high enough concentration then a cell will not enter apoptosis. Cells that enter apoptosis via p53 enter the effector phase at this BCl2 / BAX level, as p53 induces an increase in the presence of BAX at the mitochondrial membrane. If a cell enters the degradation phase then a change in the permeability of the mitochondrial membrane results in the release of cytochrome C and calcium ions from within the mitochondria. Cytochrome C activates procaspase 9 which leads to the activation of a number caspases within the cytoplasm, resulting in the breakdown of membranes and the cytoskeleton. The release of calcium ions leads to the

degradation of DNA by calcium ion sensitive nucleases. The end result of apoptosis is cellular fragmentation resulting in the production of apoptotic bodies which are engulfed by neighbouring cells (Bossy-Wetzel and Green, 1999; King, 2000).

For the balance between cellular proliferation and cellular death to be maintained, all mechanisms discussed must be in place. In cancer, loss of controls in favour of increased proliferation and decreased apoptosis lead to tumour formation. The loss of control may occur at any level such as signal transduction or G1 phase, and is due to either mutation or the complete loss of any number of proteins required for these mechanisms. At a DNA level, the genes encoding these proteins have been divided into two groups – oncogenes and tumour suppressor genes.

Oncogenes and tumour suppressor genes

Oncogenes were discovered after investigations into the effect of transforming retroviruses on animal cells found that viral genes responsible for certain changes have cellular homologues. To distinguish between oncogene types, those of viral origin are denoted as v-onc, and those of cellular origin as c-onc. Mutations within the oncogene result in a gain of function, therefore the “normal” non-mutated forms are described as proto-oncogenes. This gain of function is termed transforming potential and is due to mutations within the coding regions and / or regulatory sequences of the oncogene. The resultant phenotype changes may be either qualitative (altered product), quantitative (change in levels of product) or a combination of both. For example, within the Ras family of proteins a number of possible mutations can occur, but all result in a permanently active GTPase. There are five classes of oncogenes, grouped according to function of the proto-oncogene protein product :

Class I – Growth factors eg sis (Platelet derived growth factor)

Class II – Receptors eg erbB (Epidermal growth factor receptor)

Class III – Intracellular transducers. There are four classes of transducer; tyrosine kinases, serine/threonine kinases, Ras proteins and adaptor proteins.

Class IV – Nuclear transcription factors eg myc, fos.

Tumour suppressor genes were found to be responsible for the inhibition of certain processes, and their regulatory functions must be reversibly inactivated for cell growth to occur. Tumour suppressor genes are therefore directly involved in the

control of cellular proliferation. Instead of a gain of function due to mutation, tumour suppressors suffer loss of function. p53 and Rb are the two most widely researched tumour suppressor genes, though others are known. Of all the genes mutated in cancer, p53 is the most frequently changed and over 2000 inactivating mutations have been noted this far. It is one of a group of proteins, but the role of others within this family is not yet fully understood. Due to its involvement in both cell proliferation and apoptosis it is considered to be of importance within the development of most cancers (Lodish *et al* 1995; King 2000).

Another mechanism which is linked to proliferation and apoptosis is that of differentiation. This can be defined as “changes in gene expression by which a precursor cell becomes a distinct specialised cell type”. Proliferation and differentiation are processes which do not occur at the same time and not all differentiated cells retain the ability to proliferate. Cell cycle regulators such as Rb, p53 and some cyclins are thought to be involved in the control of differentiation, the importance being that tumour cells are often found to be poorly differentiated though they may have originated from a well differentiated cell type.

Cell adhesion molecules

As proliferation increases and tumour formation begins, a number of changes in the interaction with of cells with their surroundings occur. Cells are surrounded either by other cells or extracellular matrix (ECM), and interact with these through the use of cell adhesion molecules. These are specialised transmembrane proteins which cause cells to adhere to the ECM and to maintain tissue architecture by cell to cell adhesion. There are four main classes; integrins, cadherins, immunoglobulin superfamily and selectins.

Integrins are expressed in all cell types and are involved in a number of normal cellular functions including cell proliferation and apoptosis, through interaction of the cell with the ECM. They are involved in signalling and have been shown to initiate the MAPK cascade and hence influence gene expression and cell cycle control. As cells become malignant it has been shown that integrin expression is either changed or lost completely. Therefore they are thought to have the potential to mediate tumour

formation and metastasis and have also been shown to be positive mediators of angiogenesis.

Cadherins are intracellular adhesion molecules and maintain tissue architecture and differentiation status of epithelial tissues. The major epithelial form is E-Cadherin and it has been shown that tumours which show a radical decrease or complete loss of this form are poorly differentiated and highly invasive. Similarly, loss of members of the immunoglobulin superfamily increases the metastatic potential of the cell. Members of the immunoglobulin superfamily and selectins are also involved in migration of tumour cells from the site of origin resulting in the formation of metastases (Lodish *et al*, 1995; Miyasaka, 1995)

1.1.2 Pathology of cancer

Pathological stages in the development of cancer

Pathologically, a number of distinct stages can be noted once cells enter progression towards the formation of a multicellular tumour. It is here that an important distinction needs to be made. The presence of a tumour mass does not equate to the presence of a cancer. A benign tumour is a mass formed due to overproliferation of cells, but which remains within its tissue boundaries. A cancerous (or malignant) tumour is an overproliferation of cells, which has then gained the ability to invade surrounding tissue areas and metastasise (Weinberg, 1995).

The initial pathological change seen in the development of a tumour is hyperplasia, where a cell begins to proliferate with no apparent control. As cellular defects increase in number, less control exists on which cells enter the cycle and pass through checkpoints. Cells with a genetic mutation begin to increase in number and show phenotypic change, a stage known as dysplasia. From the dysplastic stage cells then enter a stage of metaplasia or anaplasia (depending on tissue type), where an increasing phenotypic change is noted. The next stage is described as a cancer *in situ*, as a tumour mass is formed of cells phenotypically different to those in the surrounding cell population, but still confined to the boundaries of that population. Further mutations may then enable it to invade the surrounding population, forming

an invasive cancer. Invasive tumours can then enter the blood vessels or lymphatic system allowing tumour cells to circulate around the body and form tumours within other tissue sites. If these tumours are still attached to the primary tumour then these are described as secondary tumours. If the new tumours are found not to be attached, and at distant sites within the body, they are described as metastases (Weinberg, 1995).

Clinical classification of tumour stage

To aid in the diagnosis of cancer, there is a need for classification of tumour type. Classification must identify tumour type according to cellular origin, pathological grade and whether it has entered blood or lymph nodes.

The need for such a system was met by the TNM classification system developed by Pierre Denoix and subsequently published by the International Union Against Cancer (UICC). The first booklet published in 1958 gave recommendations for classification of tumours of the breast and larynx, with the first complete book, classifying tumours in 23 body sites, published in 1968 (Sobin, 2001). The classification book has now reached its 6th edition, and TNM is the system promoted by the World Health Organisation (WHO). It is an anatomic system, based on the extent of the tumour at the primary site (T), presence or absence of tumour in regional lymph nodes (N) and presence or absence of tumour beyond regional lymph nodes (M). The objectives of the system are to provide information useful in the planning of treatment, indication of prognosis and also to evaluate results of treatment. The system covers all known tumour types, but does not classify each tumour in the same way (Sobin, 2001). Other similar methods of classification are available for some tumours. For example, colorectal cancer can also be classified under a system known as Dukes staging, this being the method of choice within the UK.

1.1.3 Diagnosis, detection and screening

Research into the underlying causes of cancer leads directly to an important factor in its clinical management, namely early detection and diagnosis of the disease. To discuss these topics, a clear definition of these terms is required. Diagnosis is defined as the “identification of a disease from its symptoms” (Oxford Dictionary), therefore

the disease has progressed to the stage where clinical symptoms are present. Early detection determines whether or not a disease is present before its symptoms appear, usually in the form of a screening test. Screening involves the testing of clinically asymptomatic individuals to determine whether the disease, or markers of that disease are present.

In the case of cancer, reaching the stage of diagnosis due to clinical symptoms may be considered to be too late, as the tumour may be well established and difficult to treat. It may also have become invasive or metastatic and therefore spread to other body tissues distal to its site of origin. Early detection strategies aim to detect the tumour at an early stage in its progression (dysplasia) at a point where it can easily be treated and has caused no threat to other body tissues.

Current diagnostic tests vary greatly, and strategies for diagnosis often involve the interaction of a number of methods. For many cancers imaging of the affected organ may be the first step eg in cancers of the oesophagus, colon and bladder, endoscopic methods are used to investigate the presence of a tumour mass and find its exact location within that organ. Similarly ultrasound and x-ray techniques are invaluable in the detection of gynaecological malignancies and tumours of the lung. Once the tumour has been located the next step is the taking of a small section of tumour, a biopsy, and the subsequent evaluation of this tissue by trained pathology staff to assess whether the tumour is malignant or benign, and if found to be malignant, assess its stage and grade. These methods are often described as the diagnostic “gold standards” and any new diagnostic methods described must be comparable to these in terms of specificity, sensitivity and reliability.

Tests for monitoring of malignancy, through early detection, are not intended to replace diagnostic methods, but to complement them. A successful test must be able to detect the presence of disease markers reliably, such that the positive presence of a marker then requires further examination to determine at what stage the disease is; A negative result can definitively state, that at the point of testing, the patient is considered disease free. This ensures that the more invasive, involved, expensive procedures of full investigation are only undertaken on patients who require this in-depth study, and confirms that patients who do not appear to require these tests really

do not require them. The reliability of the assay is of importance, as a high number of false positives will result in a high number of unnecessary investigations and conversely a high number of false negatives will result in a large number of diseased patients being missed, therefore leading to increases in late stage treatment and possibly mortality. Though the same requirements can be applied to assays for population screening, these require a higher sensitivity and specificity than is expected within monitoring assays. The requirement for cheap, quick non-invasive assays had led to the search for markers of disease which can be used in such assays. A number of routine diagnostic tests now employ such markers, but many are not sufficiently sensitive for monitoring purposes due to their inability to detect tumours at sufficiently early stage.

1.1.4 Tumour markers

The molecules investigated for their potential within detection or diagnostic uses have often been described as tumour markers. These markers indicate the presence of cancer and may give some insight into prognosis. For a marker to be of use, it must be easily detectable and must show a difference in expression when compared to normal tissue, either a loss or gain of expression or a significant change in the level of expression. These markers may be either causative within tumour development (eg p53 tumour suppressor gene mutation) or be useful as a consequence of the changes occurring (eg Carcinoembryonic antigen production).

The most widely investigated group of markers are those involved in the regulation of normal mechanisms such as proliferation and apoptosis, where mutational change causes a change in function which aids tumourigenesis. These are the oncogenes, tumour suppressor genes and cell adhesion molecules which were discussed in section 1.1.2. As they are thought to be involved in the development of most cancers, investigations into their use as general markers of cancer have been undertaken. p53 is one of the most widely investigated marker at both gene and protein level and has been shown to provide information about the stage of tumour as well as prognosis (King, 2000).

As more information is gained about the development of cancer, the number of markers which can be exploited for diagnostic purposes will also increase. Our knowledge to this point has provided a variety of markers which can, and have been investigated in terms of diagnostic and prognostic potential. Some (eg Prostate Specific Antigen (PSA)) are specific for one tumour type and are being used clinically, others (eg telomerase) have been found to be present in a wide range of tumours but are currently research tools only.

The work presented in this thesis investigates the potential of one marker, CD44 as a detection tool for both bladder and colorectal cancer. A review of current knowledge regarding CD44 will now be presented, followed by reviews of bladder and colorectal cancer, including aetiology, pathology and current detection techniques.

1.2 CD44

CD44 is a transmembrane glycoprotein which is expressed on virtually all human cell types. It is a cell adhesion molecule with Hyaluronan, a component of the ECM, being its primary ligand. The gene can undergo alternative splicing, a process which can result in a number of variable isoforms, leading to the production of variable protein isoforms, a number of which have been shown to differ in their adhesive qualities. The loss of regulation and overexpression of these variant isoforms within tumour cells has implicated CD44 in the formation of metastases, and has led to its investigation as a possible tumour marker (Lesley and Hyman, 1998; Goodison and Tarin, 1998(b)).

1.2.1 Background

The presence of CD44 protein was first noted in 1980 by Dalchau *et al*, through the use of a monoclonal antibody directed against it. It was first described as brain-granulocyte-T lymphocyte antigen due to its expression on these cell types, though later work demonstrated its presence on all haematopoietic cells and most other cell types (Dalchau *et al*, 1980). Many other groups detected similar proteins, again through the use of monoclonal antibodies leading to a variety of labels for the same

protein. Confirmation that all antibodies detected the same protein led to CD44 gaining the CD designation in 1987 (Naor *et al*, 1997; Haynes *et al*, 2001).

Initial investigation into its role demonstrated its ability to bind to hyaluronan, and later studies also noted binding of CD44 to fibronectin and collagen (Haynes *et al*, 2001). The finding by Jalkanen *et al* in 1986 that CD44 protein (here defined as hermes, due to the monoclonal antibody used in its detection) was involved in lymphocyte interaction with endothelial venules led to its description as a lymphocyte homing receptor (Jalkanen *et al*, 1986).

Early studies on the expression of CD44 at a gene level suggested the presence of an alternative splicing mechanism, highlighted by the difference in size of the protein form on haematopoietic cells and that on epithelial cells.

1.2.2 The CD44 gene

In humans, the gene for CD44 has been mapped to the short arm of chromosome 11, in close proximity to a region involved in suppression of metastasis in some tumour types (Günthert, 2001). Analysis of cDNA transcripts from a variety of cell types suggests the involvement of an alternative splicing mechanism in the arrangement of the gene due to differences in transcript size. Sequencing of the gene by Screaton *et al* in 1992 demonstrated a gene containing at least 19 exons, of which 12 were thought to be alternatively spliced (Screaton *et al*, 1992). It has since been noted that there are at least 20 exons in total, spanning approximately 60kb (Günthert, 2001). Of these 20 exons, 10 are constitutively expressed and 10 are expressed through alternative splicing. The presence of 10 variant exons was confirmed by Tolg *et al* in 1993, and it was this group that suggested a standard and variant nomenclature for the exons. The 10 constitutively expressed exons (1-5 and 16-19) were described as standard exons and formed the standard form of the gene – CD44S. The 10 alternative exons (6-15) were described as variant, and inclusion of these into the standard backbone produces the variant form of the gene – CD44V (Tolg *et al*, 1993). The gene structure as described here is illustrated in figure 1.2 below.

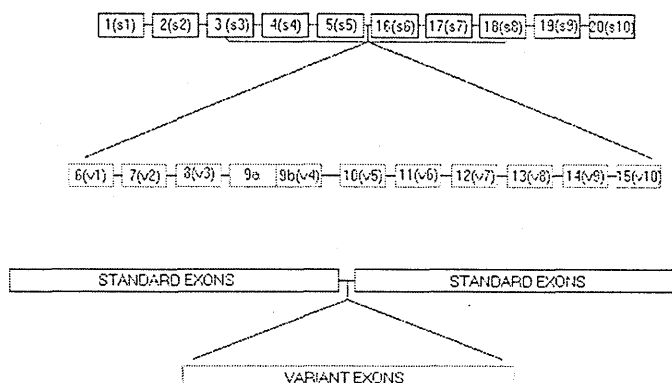


Figure 1.2 – Structure of the CD44 gene, demonstrating the presence of both standard and variant exons. Exons 1-5 and 16-20 produce the standard backbone, whilst the presence of variant exons (numbered exons 6-15) between exons 5 and 16 produces the variant form of the gene. Figure adapted from Woodman *et al*, 1996.

It is with the exon nomenclature that some confusion can arise. Variant exons can either be described by their exon number within the gene structure, ie 6-15, or they may also be described by their variant number ie 1-10. In this thesis, the standard numbering system of exons 1-20 will be employed in the main.

Alternative splicing of the gene occurs at exon 5, with alternative splicing and donor sites present at exon 5 and exon 7 (variant 2). It was also noted by Screaton *et al* (1992) that some alternative splicing took place at exons 18 and 19, though this was later shown to also include exon 20 (Screaton *et al*, 1992). This splicing results in a change within the length of the cytoplasmic tail of the CD44 protein, with the short form encoded by exon 19, and the longer predominant form encoded by exon 20 (Günthert, 2001).

1.2.3 Alternative splicing

It has been suggested that the number of possible variant forms due to alternative splicing of the CD44 gene could be over 1000. This makes it the most variable adhesion molecule, when compared to other such molecules (including Fibronectin NCAM and CD45) that are known to be alternatively spliced.

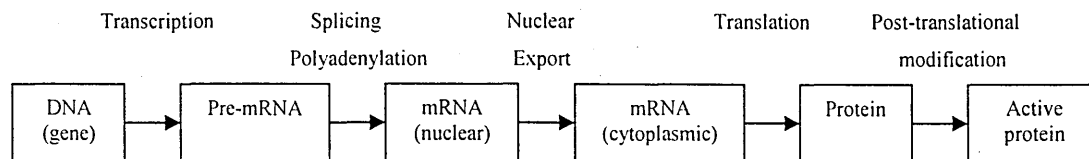


Figure 1.3 – The pathway from gene to protein, via mRNA. The genetic information is transcribed into RNA within the nucleus, and this is then translated to the protein within the endoplasmic reticulum (figure adapted from Philips and Cooper, 2000).

Figure 1.3 above illustrates the pathway from DNA to active protein. The first RNA transcript, which is produced in the nucleus, is described as heterogenous RNA (hnRNA) and consists of both introns and exons bound to proteins. hnRNA will be both pre-mRNA transcripts for protein production, and other RNA which will not progress to translation eg misprocessed transcripts. After post-transcriptional modification of the RNA, by capping at the 5' end and polyadenylation of the 3' end, splicing out of the introns begins. For the splicing reaction to occur, a spliceosome must be formed, and proteins and ribonucleoproteins responsible for this process assembled. Splicing starts with the cutting of the 5' end of the intron and formation of a lariat. This is where the 5' end of the intron loops back on itself and binds to a base within a sequence described as the branch site. The 3' end of the intron is then cut, released (still in lariat formation), and the exons joined. Regulation of splicing is through the presence of short consensus sequences at the 5' and 3' splice sites, and at the branch site (Lewin *et al*, 2000; Phillips and Cooper, 2000).

It is during this process that alternative splicing occurs. For the standard form of CD44 to be produced splicing out of all variant exons will occur at the same time as splicing out of introns between standard exons. During alternative splicing, not all variant exons are spliced out. The exact mechanisms which govern alternative splicing within CD44 are unknown, though it is known that alternative splice sites are generally weaker in affinity than constitutive splice sites, leading to the hypothesis that a change must occur in the splicing mechanism to favour the weaker site.

Some insight has been given into splicing in CD44 by Stickeler *et al* (1999) through investigation of splicing in mammary tumourigenesis. Analysis of the expression of

the SR splicing factors demonstrated an increase of these factors in line with an increase in alternative splicing of CD44. SR proteins are a group of serine-arginine proteins that are involved in the splicing of pre-mRNA and are known to influence alternative splicing in other genes (Stickeler *et al*, 1999).

A further question in the alternative splicing of CD44 is whether the exon assembly patterns noted are random, or have influencing factors.

Bell *et al* (1998) studied CD44 isoform expression and found that though there is a possible splicing pattern of over 1000, some isoforms were seen very commonly. They found that there was a distinct hierarchy of expression, with 3' variant exons being included the most and 5' variant exons the least. Exon 15 (V10) was observed approximately 100 times more frequently than exon 6 (V1). They also noted that if an alternative exon was present it was most likely to be included together with other variant exons found downstream in sequence. Finally it was found that intron length played a major role in which exons were included together, with those having shorter introns being more likely to be included together than those with large introns (Bell *et al*, 1998).

Matsumara *et al* (1994) investigated the possibility that intron retention was also a factor in the production of large variant forms of the mRNA in tumours. Investigation of CD44 transcripts in bladder cancer samples demonstrated that this was the case. Of tumour samples investigated, 60% demonstrated retention of intron 9, and preliminary studies (results of which were not given in detail within this paper) had also demonstrated retention of introns 14 and 15. This could have implications for the final protein product, with the production of a radically different form of CD44 protein. It was unknown however, whether these transcripts were translated to form mature CD44 protein (Matsumara *et al*, 1994).

1.2.4 CD44 protein structure

The standard form of CD44 is a type I transmembrane protein which is synthesised as a 37 – 38 kDa protein which undergoes extensive post-translational modification.

These changes occur either due to glycosylation resulting in a protein of 85 – 95 kDa, or through the addition of chondroitin sulphate producing a protein of 180-200 kDa. The cytoplasmic tail of the protein may also undergo modification and be phosphorylated.

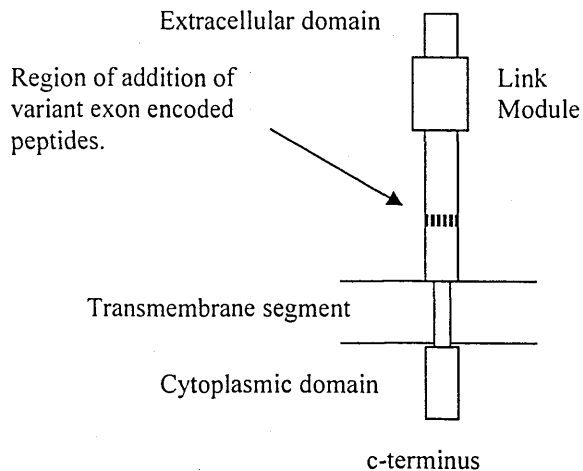


Figure 1.4 – Diagrammatic representation of CD44 protein structure. This illustrates the presence of three domains and indicates the area of variability due to variant exon translation (adapted from Bajorath, 2000).

The protein can be divided into three domains; cytoplasmic, transmembrane and extracellular (see figure 1.4 above). The cytoplasmic domain is composed of 72 amino acids and is encoded by exons 19 and 20. It contains six potential phosphorylation sites on serine residues (Naor *et al*, 1997). The transmembrane domain is single-pass in nature and composed of 21 amino acids, encoded by exon 18. The extracellular domain is the largest, consisting of approximately 207 amino acids in CD44S. It is encoded by exons 1-17, though in CD44V amino acids encoded by variant exons may be inserted. The protein has a “link module” for hyaluronan (HA) binding (named due to its resemblance to the HA binding domain of cartilage link protein) and a number of sites for potential *N*- and *O*- glycosylation. Variant amino acids are included in the middle of the membrane proximal region of the protein (Naor *et al*, 1997; Lesley and Hyman, 1998).

1.2.5 CD44 ligands

HA is the principle ligand for CD44, but although all CD44 proteins express the link module, not all will bind HA. Three activation states for HA binding have been noted:-

Active	– can bind HA
Inducible	– can only bind HA in presence of inducing monoclonal antibodies
Inactive	– unable to bind HA, even in presence of inducing factors.

These activation states appear to be determined according to the cell type in which CD44 is expressed. Changes within all regions of the CD44 protein (other than the link module) have been shown to affect HA binding and it has been suggested that the main regulating factor in this binding is the glycosylation state of the protein (Lesley and Hyman, 1998; Bajorath, 2000).

Several other ligands have also been noted for CD44, including chondroitin sulphate, collagen and fibronectin (via its heparin binding domain). It has also been shown that CD44 proteins which contain exon 8 (V3) have heparan sulphate side-chains which are able to bind heparin-binding growth factors and chemokines. The addition of other variants has also been shown to increase the level of glycosylation due to the presence of extra potential O-linked glycosylation sites (Lesley and Hyman, 1998; Bajorath, 2000).

It has been suggested that the cytoplasmic tail is involved in interaction with cytoskeletal components. *In vitro* studies have demonstrated CD44 binding to ankyrin and have correlated this to HA binding. There is also some evidence for actin microfilament binding, but this is still inconclusive (Lesley and Hyman, 1998).

1.2.6 CD44 functions

The main function associated with CD44 is adhesion. Its primary ligand in standard form is HA and therefore CD44 has been associated with binding to the ECM in a

number of cell types. Initial studies observed the role of CD44 in lymphocyte adhesion to vascular endothelial walls and also in lymphocyte homing. This is thought to occur in an organ specific manner when lymphocyte recruitment is required. Due to its involvement in lymphocyte homing CD44 has also been shown to be present during wound healing. Other suggested functions include involvement in cell motility, degradation of HA and as a signal transducer for growth signals. Binding of growth factors to CD44 were demonstrated to cause an increase in intracellular calcium and protein phosphorylation, resulting in cell activity, including HA binding and proliferation (Naor *et al*, 1997; Haynes *et al*, 2001).

Further examination of CD44 and proliferation showed a positive correlation between CD44 expression and proliferation in colorectal epithelial cells, both normal and neoplastic in nature. CD44 was detected (by a mAb) in only the proliferative zones of normal colorectal crypts and proliferative status was assessed through the use of Ki-67 mAb (Abassi *et al*, 1993).

1.2.7 CD44 and Cancer

The first definitive link between CD44 and cancer arose from a study by Günthert *et al* (1991). Their study investigated CD44 expression in a metastasising rat carcinoma cell line, from which a library of cDNA clones was created and their protein expression assessed. One clone coded for a protein with an extra 152 amino acids in the extracellular domain. Upon insertion of this clone into a non-metastasising cell line the cells became metastatic, with expression of the extra domain at both RNA and protein level. It was also noted that co-injection of the metastatic cell line with a mAb specific for CD44 stopped the formation of tumours *in vivo*. In their opinion this demonstrated a direct relationship between the presence of CD44 variants and the metastatic process (Günthert *et al*, 1991).

Further work by this group investigated the presence of this metastasing variant in a number of human cell lines. Most cell lines were tumour-derived including; colon, breast, lung and cervix. For normal comparison, two immortalised cell lines (fibroblast and keratinocyte) and peripheral blood leucocytes were studied. Most cell lines expressed CD44, but not all tumour lines demonstrated variant expression. Of

the cell lines that did express CD44V it was noted that they did not all express the same variants. The keratinocytes were shown to express CD44V, a finding which may link the keratinocyte differentiation pathway to that of metastatic cells as they share the properties of decreased adhesion to the ECM and the production of matrix degrading enzymes. At the end of this study the group suggested that CD44 variants may be a normal part of the differentiation pathway, but when expressed incorrectly can cause metastasis (Hofmann *et al*, 1996).

Guo *et al* (1994) were able to demonstrate the metastatic potential of CD44 within melanoma cell lines, again one metastasising and one non-metastasising. In *in vitro* studies only the metastasising cell line expressed CD44, and its growth *in vitro* could be halted by the addition of an mAb directed against CD44. They were also able to correlate expression of CD44 with increased tumour growth and metastasis. Again it was suggested that CD44 was linked with metastasis, but also that interactions between CD44 and its ligands were important for tumour growth and metastasis (Guo *et al*, 1994).

The first investigation of CD44 variant expression within real human tissue samples was undertaken by Matsumara and Tarin (1992). They investigated expression of CD44 in both neoplastic and non-neoplastic tissue using the technique of reverse transcription polymerase chain reaction (RT-PCR). This technique is frequently used in the investigation of gene expression and evaluates the level of mRNA transcription. Unlike normal PCR, where genomic DNA is the template, mRNA is the starting point, and this is reverse transcribed to complementary DNA (cDNA) to generate the template for PCR. Using PCR primers anchored at standard exons (one either side of the variant insertion site) Matsumara and Tarin (1992) could amplify the whole gene, thereby including any variant exons which may be inserted. All tumour samples expressed CD44, but the presence of a number of variant forms was noted. General expression levels were increased when compared to levels within normal tissues and most of the CD44 forms noted were above the expected size for that of CD44S or the normal epithelial form (which contains exons 12-14 (V7-9)). These differences began to suggest that CD44 could have potential as a diagnostic tool (Matsumara and Tarin, 1992).

In a review of CD44 expression within a number of tissues, Tarin *et al* (1995) demonstrated that the two general features of expression within tumours that were CD44 positive, were general overexpression and expression of a number of variant forms, described as “aberrant”. The tissues reviewed were breast, gastrointestinal (GI), lymph, cervix, urological (mainly bladder) and lung. Expression in breast and GI tissues showed not just an increase in expression, but a change that could be correlated with the severity of disease. They concluded by stating that CD44, in their opinion, could be useful diagnostically with the presence of deranged expression being a possible marker for early detection. If the link between metastasis and variant expression could be proven, they also suggested CD44 could be used for prognosis (Tarin *et al*, 1995).

The correlation between CD44 expression and tumour burden has been investigated in a number of tissues, but the relationship appears to differ among tissue types. In colorectal cancer overexpression of CD44V indicates the presence of a high grade aggressive cancer, yet in bladder cancer this was found not to be the case. It has been observed that as cancers of the bladder become more invasive they lose CD44 expression, though an initial overexpression was present (Sugino *et al*, 1996).

The overexpression of a number of CD44 variants at an mRNA level as detected by RT-PCR could be expected to result in the overexpression of protein forms. However, in a study of comparative expression in breast, bladder and colon has shown this not to be the case. An increase of both standard and variant forms was seen, though only in tumour cells but not in the normal cells surrounding the tumour. When numbers of mRNA forms were compared with the number of protein forms present, less protein forms were noted. This suggests that not all regulatory mechanisms for translation were lost within the cancer cells (Woodman *et al*, 1996).

The role of CD44 in metastasis and its associated regulation is still unclear. Okamoto *et al* (1999) demonstrated that cancer cell migration in metastasis was not simply due to degradation of the ECM by metalloproteases. They showed that there was also cleavage of CD44 at the membrane proximal region, thereby removing its link module and rendering it incapable of HA binding (Okamoto *et al*, 1999). Bruno *et al* (2000) have noted that CD44 plays a role in proliferation of squamous cells in

squamous cell carcinoma, though this is independent of form or density on the cell surface. Binding of a specific mAb to the protein triggered a signal transduction pathway, resulting in proliferation (Bruno *et al*, 2000).

Though the exact role of CD44 within tumour development and the regulatory mechanisms which govern its splicing are not yet fully understood, it is under investigation as a potential tumour marker. The splicing patterns within variant forms have not been noted for each tumour type, yet the presence of deranged patterns is considered to be enough to investigate its diagnostic potential. When figures for sensitivity and specificity of CD44 detection have been compared to those for p53, H-ras and K-ras, the results appear to be promising (Goodison and Tarin, 1998(a)).

A number of methods have been used to study CD44 expression and as such, all can be considered as possible candidates for diagnostic purposes. Due to the variety of antibodies produced for initial studies of CD44, a large array of both monoclonal and polyclonal antibodies exist for the standard form of the protein, though later work has also produced a number of antibodies directed against variant portions of the protein.

At an mRNA level, RT-PCR and southern blotting have been employed with excellent results. Sensitivity and specificity in many of these studies has been 90% or above, but these techniques though useful in a research setting are not compatible with clinical use.

Goodison and Tarin (1998 (a and b)) have suggested that the most useful formats for diagnostic tests would be either enzyme linked immunosorbent assay (ELISA) for protein detection, or RT-PCR-ELISA for RNA detection. They have also suggested that single marker assays do not provide enough information and that for the best diagnostic use addition of other new markers would be required (Goodison and Tarin, 1998 a and b).

1.3 Bladder Cancer

Bladder cancer is one of the most common malignancies in the UK. It predominates within the male population, with an incidence ratio of 3 to 1 in comparison with the female population. As such in England and Wales, it is the fourth most common malignancy in males and ninth most common in females. Incidence is rare below the age of 45, but rises sharply up to the age of 60 until it peaks at 85 years. Though incident rates differ between sexes, incident age does not (National Statistics Office, 2001).

1.3.1 Aetiology of bladder cancer

A number of factors believed to be causative in the development of bladder cancer have been identified, and can be broadly divided into occupational and non-occupational factors.

Occupational factors

Occupational causes of bladder cancer have been linked to chemical exposure within the workplace, the main chemicals implicated being the aromatic amines. The link between exposure to such chemicals and bladder carcinogenesis was established in 1954, in a study investigating the effects of aromatic amine exposure on workers in the dye industries, and also in those involved in the production of rubber, textiles and chemicals. It was noted that there was a 10-50 fold increase in the risk of bladder cancer after exposure to these chemicals (including 2-naphthylamine, 4-amino-biphenyl and benzidine). Further studies at later dates found similar results, but with the implementation of working conditions which promote less exposure to such chemicals, the risk level has decreased (Johansson and Cohen, 1997; Silverman *et al*, 1999). A study by Sorahan *et al* (2000) in a company producing chemicals for the rubber industry demonstrated that incidence rates for bladder cancer among a cohort of male staff was not too dissimilar to the general population. However, exposure to one chemical (phenyl- β -naphthylamine) was suspected to have caused a small number of cases noted (Sorahan *et al*, 2000).

Though a change in working conditions has decreased cancer rates in some industries, it has been noted that drivers of trucks, buses and taxis, and workers in the aluminium industry appear to be showing increased rates. Polycyclic Aromatic Hydrocarbons (PAHs) present in exhaust fumes are thought to be the influencing factor in the development of cancer within drivers, with an increased rate of cancer positively correlating with length of employment. Within the aluminium industry the effects are thought to be caused by volatiles released from coal-tar pitch during the production process. The constituents of these volatiles are not fully known, but aromatic amines are thought to be present (Silverman *et al*, 1999).

In summary, occupational exposure to a number of chemical carcinogens is a factor in the development of bladder cancer in a sizeable section of the population. Though risk levels have decreased in some occupations, regular screening of the workforce for bladder cancer is undertaken. It could be expected that the majority of persons within these occupations are male, therefore this could explain the higher incidence rates noted in this gender.

Non-occupational factors

Non-occupational factors are those that can affect the whole population, and the factors that have been investigated for bladder cancer include cigarette smoking, diet and alcohol intake.

The link between cigarette smoking and a number of cancers (especially those of the lung) has been known for a number of years, and has also been implicated as a factor in the development of bladder cancer. The risk increases as the number of pack years increases, but as with other cancers, this risk decreases upon cessation of smoking. Some variance in the level of risk has been noted eg a higher risk has been associated with people who smoke stronger tobacco, along with those who smoke unfiltered cigarettes and inhale the smoke deeply. The mechanisms involved in the development of bladder cancer are not fully understood, but appear to be linked to constituents of the smoke such as the aromatic amines and PAHs (Johanssen and Cohen, 1997; Silverman *et al*, 1999).

Caffeine and saccharin are two dietary factors that have been implicated within bladder cancer development, but with no substantial evidence to suggest these are carcinogenic within humans. A number of epidemiological studies have investigated the effects of caffeine intake, but study cohorts contained large numbers of smokers, a factor which could have affected any trends noted. Saccharin was found to be carcinogenic in studies on rats, though no such evidence has arisen from human studies (Johanssen and Cohen, 1997; Silverman *et al*, 1999). Studies of urothelial cancers within obese subjects has shown a general increase in risk for such tumours, but not specifically for bladder cancer (Mydlo *et al*, 1999). Other dietary related studies have investigated the effects of fruit and vegetable consumption on risk levels, to determine whether these provide a protective effect. It was noted in one such study that some vegetables did appear to cause a slight decrease in risk, but that most appeared to have no effect (Zeegers *et al*, 2001).

Though alcohol has been demonstrated to be of importance within the development of other cancers, it appears to have no major effect on the development of bladder cancer (Silverman *et al*, 1999).

The final risk factor known to be of importance is infection with the parasite *Schistosoma haematobium*, though the exact mechanisms by which this occurs are not known. Though not a problem in the UK, it is of great importance in areas of the world where this parasite is endemic eg Africa and the Middle East (Johanssen and Cohen, 1997).

Though there are no inherited disorders which are directly involved in the development of bladder cancer, there are some genetic mutations which make their carriers more susceptible. One such mutation occurs within the gene for *N*-acetyltransferase, an enzyme responsible for a process described as *N*-acetylation. This process is important within the de-toxification pathway of aromatic amines. These chemicals must be metabolised to produce their carcinogenic effect, but acetylation of the metabolites allows them to be excreted before they can exert their effect. Two acetylation phenotypes are noted in humans; normal acetylators (who carry the normal gene) and slow acetylators (who are homozygous for the mutated

gene). As slow acetylators are unable to detoxify chemicals as efficiently as those without the mutated gene, they are at a higher risk of carcinogenic damage. Other mutations investigated are absence of a functional human glutathione *S*-transferase and changes to the cytochrome P-450 enzymes. Again these are enzymes known to be involved in de-toxification pathways and have been linked to aromatic amine breakdown. However, studies on the links between the mutated forms of these genes and the development of bladder cancer are inconclusive (Silverman *et al*, 1999).

1.3.2 The molecular basis of bladder cancer

The most noted feature in the underlying basis of this cancer is the number of alterations detected within chromosome 9. These changes range from monosomy (complete loss of one chromosome) to specific allelic losses, occurring on both arms of the chromosome.

Czerniak *et al* (1999) studied the losses noted in this chromosome in detail, with a view to utilising the information for early detection. No specific patterns of alteration were noted, though loss of certain regions did appear frequently. Linking the losses to tumour stage demonstrated they were more likely to be found in early tumours suggesting that alterations within chromosome 9 are an early tumour event. This hypothesis that was proved through the study of normal urothelial cells adjacent to tumour cells. Microscopically these cells appeared to be normal, but further investigation demonstrated similar chromosomal alterations to those in the tumour cells. The reasons for the importance of these alterations are not fully understood, but a number of putative tumour suppressor genes have been mapped to some of the deleted loci (Czerniak *et al*, 1999).

Other chromosome alterations have also been noted in bladder cancer. The initial move to investigate these chromosomal changes was to produce a model genetic pathway for the progression of transitional cell carcinoma (TCC) of the bladder, similar to that produced by Fearon and Vogelstein (1990) for colorectal cancer (see section 1.4 “Colorectal cancer”). Knowles (1999) has now produced such a model, which summarises all genetic knowledge at that point in time. The model is

demonstrated in figure 1.5 below and includes all known chromosomal alterations at the various pathological stages.

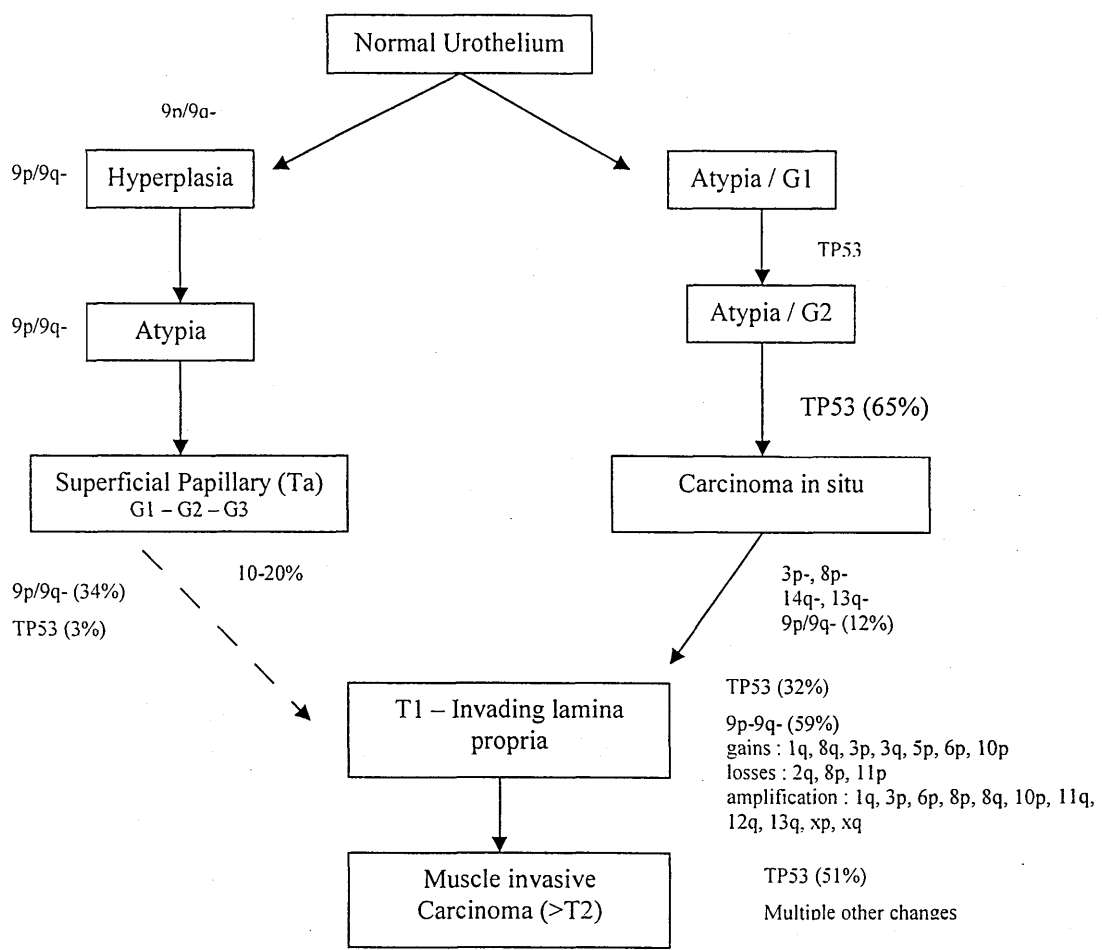


Figure 1.5 – A model for the progression of transitional cell carcinoma (TCC). Evidence suggests that there are two possible pathways for bladder tumour development. This model summarises many of the chromosomal changes noted thus far, with loss of a chromosome described using the minus sign (-). Figure taken from Knowles, 1999.

The model consists of two pathways as evidence suggests most muscle invasive tumours do not arise from superficial tumours, but instead from carcinoma *in situ*. This evidence is mainly histopathological in origin and more study is required to explain this at a genetic level. Since the publication of this model, alterations in chromosomes 3,7, 16 and 17 have been noted (Watters *et al*, 2000; Yu *et al*, 2001; Wada *et al*, 2001). Most recently Hoglund *et al* (2001) have studied chromosomal changes at specific tumour stages. They were able to correlate specific changes with

tumour stage and grade and also demonstrated the presence of two pathways as illustrated by Knowles (1999, above). They noted that the two pathways are well separated at the early stages, but during the later stages, the changes noted become more similar (Hoglund *et al*, 2001).

Oncogene expression within this cancer has been related mainly to late stage invasive tumours. Overexpression of H-Ras, erbB-2 and c-myc has been demonstrated though their roles in the development of TCC specifically are still unknown. The finding that these oncogenes are expressed in late stage tumours has suggested their use as prognostic markers. It is possible that alterations of chromosomes may provide information on other putative oncogenes specific for bladder cancer, but further work is required.

A number of tumour suppressor genes have also been implicated in bladder cancer, including a group of putative genes found through chromosomal analysis in areas of change. Of the known tumour suppressor genes expressed, p53 and RB are the best characterised, though as for the oncogenes, these also appear to be more highly expressed in late stage tumours. Again, the roles specifically undertaken in TCC by these are not yet fully understood.

Of the adhesion molecules expressed in normal urothelium, studies have demonstrated that it is changes within E-Cadherin expression that predominate in the progression of TCC. E-Cadherin is involved in cell-cell adhesion via interactions with members of the catenin family of adhesion molecules. It has been described as a tumour suppressor gene in relation to bladder cancer, as its loss appears to correlate with increased stage and grade of tumour, leading to an invasive phenotype. Similarly, loss of catenins and the catenin-cadherin complex has demonstrated these findings. The loss of function can be attributed to a number of factors, including chromosomal loss, mutation of the gene and degradation of the protein (Cohen *et al*, 1997; Davies *et al*, 1999; Bindels *et al*, 2000).

Expression of the other major adhesion families, the selectins, integrins and the immunoglobulin superfamily has been investigated, but conflicting results regarding their exact role in the progression of bladder cancer have arisen from different studies.

It has however been suggested that a number of these proteins may be of use in a diagnostic or prognostic setting (Davies *et al*, 1999).

One further adhesion molecule, CD44 does appear to be of importance in bladder cancer as both overexpression and deranged expression are seen at early stages within the progression of TCC. The expression of this protein in bladder cancer is discussed in section 1.3.6.

1.3.3 Pathology and Staging of bladder cancer

The majority of bladder cancers are epithelial in origin, with the most common being transitional cell carcinoma (the epithelial layer of the bladder, or urothelium, being transitional in nature). This epithelial layer consists of a layer of basal epithelial cells, approximately four layers of epithelial cells and covered by larger flattened epithelial cells described as umbrella or superficial cells.

TCC is divided into two distinct pathological forms; papillary and non-papillary, both of which may be either invasive or non-invasive. Papillary tumours tend to be less aggressive in nature and are the least likely to metastasise, but are the type most often associated with recurrent disease.

Early pre-neoplastic changes are noted within the bladder, comprising of hyperplasia, metaplasia and dysplasia, though development of tumours may not occur through this ordered pathway. Hyperplastic urothelium is mainly noted as an increase in the thickness of the urothelium, due to an increase in cell layers. Dysplastic tissue may show this thickening, or may be normal in cell layer number. It is the cells themselves that begin to differ, with a number of possible nuclear abnormalities.

The genetic progression model described in the previous section illustrated two pathways for tumour progression. The pathological features of bladder cancer progression are the main factor in the determination of these two pathways. Hyperplastic tissue may lead directly to the formation of a papillary TCC, or it may progress to dysplasia. Dysplastic tissue may also arise directly from normal tissue, with no hyperplastic stage, and can be graded into three types. Urothelial dysplasia is

thought to be the precursor for a highly invasive form of TCC, progressing to a carcinoma *in situ*, or tumour *in situ* (Tis), the non-papillary form of TCC.

There may also be a number of pathological changes appearing within one bladder, eg the presence of Tis, a papillary tumour and areas of hyperplastic tissue. These are described as field changes and are thought to account for the high level of recurrence noted for this cancer. Staging of bladder cancer is carried out using the TNM staging method, as described in section 1.1.2. Figure 1.6 demonstrates the extent of tumour growth for each stage.

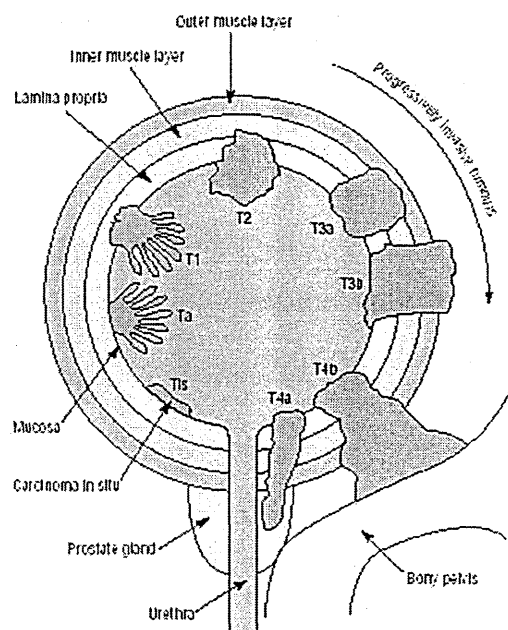


Figure 1.6 – Diagram demonstrating the TNM staging system in bladder cancer. From the Tis towards the base of the bladder, the tumour stage increases as the diagram moves clockwise around the bladder. The majority of tumours presented by patients are Ta stage (Figure from van der Meijden, 1998).

Staging of bladder cancer is carried out under the TNM classification system, as described in section 1.1.2. Non-invasive papillary tumours are described as Ta, and carcinoma *in situ* as Tis. Stages then increase from T1 – T4, depending on invasion into surrounding tissue, beginning with subepithelial connective tissue (T1), muscle invasion (T2), invasion of perivesical tissue (T3) and invasion of other organs such as the prostate and uterus (T4). If multiple tumours exist, then “m” is also included in the description. Tumour grading depends on the differentiation status of the tumour, and

is graded I-III, with I being well-differentiated, and III being the least differentiated (Sobin and Wittekind, 1997).

1.3.5 Diagnosis of bladder cancer

Clinical signs and symptoms

The most common symptom seen in bladder cancer is the presence of blood within the urine, to an extent where it can be noted by the patient. This is described as macroscopic or frank haematuria and is usually a painless blood loss, occurring intermittently, which may also be accompanied by frequency and urgency of urination. If Tis or invasive bladder cancer is present the patient may suffer from some thigh pain. The presence of haematuria does not equal the presence of cancer, but all presenting patients should be treated as such.

Gold standards for diagnosis

The gold standards for diagnosis of bladder cancer are cystoscopy and voided urine cytology. Cystoscopy is the ultimate diagnostic method, and allows full visualisation of the bladder surface and the ability to biopsy for pathological evaluation. This method is carried out under anaesthetic, and in the case of some small bladder tumours, resection of the whole tumour can be carried out concurrently with visualisation.

Voided urine cytology is carried out as an adjunct to cystoscopy, and examines urothelial cells shed from the lining of the bladder (and other areas of the urinary tract) and excreted in the urine. Cytology has been demonstrated to be very sensitive and specific for well differentiated high grade tumours, but not for the lower grade tumours. The reliability of this method has been questioned on a number of occasions, with sensitivity and specificity results varying greatly between some studies. Farrow (1990) examined the technique, and states areas where problems arise. Extrinsic factors he noted which influenced the result included sample collection and sample preparation, areas where cells could become damaged and degraded, thereby causing changes to their morphology. Intrinsic factors within the samples themselves were difficulties in distinguishing low grade tumour cells from

normal cells and the limited number of cells which some tumours shed (Farrow, 1990). Paez *et al* (1999) went further, and assessed the reliability of three cytologists in assessing sample stage and grade. They noted clear differences in the cytologists abilities with the heterogeneity of results gained (Paez *et al*, 1999). Worries about cytology have led to variations within the method through the use of antibody mediated tests, but it may also be replaced by the new generation of tumour marker assays currently available or still at the research stage.

Urinary based marker assays currently in clinical use

BTA stat and TRAK assays

These assays detect human complement factor H-related protein (hCFHrp), a protein first detected in the culture supernatant of a number of human bladder cancer cell lines. It is thought that production of this protein provides a growth advantage to the cells by facilitating evasion of immune surveillance, due to the role of complement factor H within the immune system.

The BTA *stat* was the first to be produced and is a disposable immunochromatographic assay which does not require laboratory use and provides a qualitative result. Urine is placed into the well of the plastic test device and reacts with a colloidal gold-conjugated antibody. This antibody-antigen complex is captured by a secondary antibody and a visible coloured line is formed within five minutes if the antigen is present (Pode *et al*, 1999). Studies of the assay have demonstrated a specificity of 64 – 68.9% and sensitivity of 65-82.8%. In comparison with cytology it has been shown to be more sensitive overall, but it is more specific and sensitive than cytology for lower grade tumours, and both sensitivity and specificity increase with higher stage and grade tumours (Leyh *et al*, 1999; Pode *et al*, 1999).

BTA TRAK is a laboratory based assay. It is a solid-phase sandwich ELISA which provides a quantitative result. This assay can be correlated to grade and stage but is more sensitive and specific for high stage, high grade larger tumours. Studies evaluating this assay have demonstrated a specificity of 50 – 69%, and a sensitivity of 66 – 72%. Again, when compared to cytology, this is the more sensitive assay (Heicappell *et al*, 1999; Thomas *et al*, 1999).

In a comparative study, Irani *et al* (1999) found BTA TRAK to be the most sensitive of the two assays, especially in lower stage and grade tumours. Both showed limited specificity when investigating patients with other genitourinary disorders but no bladder tumour (Irani *et al*, 1999). All studies thus far have suggested more work is required to improve sensitivity and specificity, but all believed either assay to be a useful adjunct to cystoscopy in place of cytology.

NMP22

This is a component protein of the nuclear mitotic apparatus protein (NUMA). The proteins which make up this complex are members of the nuclear matrix protein family, a group of proteins which constitute part of the non-chromatin scaffolding within the cell nucleus. In TCC, NUMA is overexpressed and NMP22 (a component of this apparatus) is found to be present in levels 10 times greater in the urine of patients with tumours than those without (Hughes and Cohen, 1999). Detection of NMP22 is carried out through a lab based ELISA assay and is quantitative in nature. Studies on this assay have demonstrated no significant relationship between stage and grade of tumour, with an overall specificity of 80-89% and sensitivity of 64.3 – 90.8%. These values vary depending on the cut off value designated for units of NMP22 / ml. This assay was found to be sensitive for high stage and grade tumours, along with those of a larger size (Miyanga *et al*, 1997; Sanchez-Carbayo *et al*, 1999). Due to its sensitivity in detecting recurrent disease, it has been evaluated as a screening test for use in routine patient assessments following tumour resection. Results appear promising, though it is suggested that it is still used as an adjunct to cytology (Soloway *et al*, 1996).

Fibrin-degradation products

Bladder cancers are known to have well developed microvasculature within developing tumours, and as such produce a number of factors such as fibrinogen and other clotting factors. Breakdown of fibrin clots produces fibrin degradation products (FDP) which have been shown to be present in increased concentrations within the urine of tumour patients (Foresman and Messing, 1997; McCabe *et al*, 2001). In a study by Ramakumar *et al* (1999) an assay for FDP was shown to have sensitivity of 52% and a specificity of 95%, with a positive correlation to grade of tumour, but not

to stage. Again this assay is affected by the presence of some urinary tract disease (inflammatory disease), though frank haematuria did not appear to have a significant impact. Comparing their results with previous studies they stated it was no greater in sensitivity than cytology, but could be used as an adjunct to cystoscopy due to its rapid test time (seven minutes).

Immunocyt

This is the newest assay to be approved for use and is an immunocytochemical assay which uses three fluorescent-labelled monoclonal antibodies. The antibodies are directed against a high molecular weight carcinoembryonic antigen (19A211, labelled with Texas red) and two mucins (M0.344 and LDQ10, labelled with fluorescein). Mucins are high molecular weight glycoproteins that are found on epithelial cell membranes, whilst CEA is commonly produced by a number of tumours. The mucins detected in this assay are not found on normal bladder urothelium and are only expressed on tumours. The assay is an adjunct to cytology, and utilises voided urine samples which are prepared and placed on slides as for cytology. Cells are first stained with a modified papanicolaou stain before incubation with the immunocyt antibodies. Results are gained through inspection of the slides under a fluorescent microscope, where CEA presence is indicated by red fluorescence, and mucin presence by green fluorescence. Samples are considered positive when at least one green or one red cell is present (Mian *et al*, 1999).

Urinary markers in research

There is a continuing body of research to discover and investigate new markers which are more specific and sensitive than the last generation. A number of markers are currently being studied, including telomerase, cytokeratins and CD44. These markers are currently proving to be the most promising when compared to the markers in clinical use at this time.

Telomerase

To avoid degradation of chromosomes during successive cell replications, the ends of chromosomes are capped with telomeres. These are tandem repeats of a TTAGGG sequence, that are slowly lost with each replication cycle. Once the telomeres have

become too shortened the cells will enter senescence. In cancer cells due to their rapid proliferation telomere shortening would be expected to occur quickly, thereby causing cell senescence and slowing tumour progression. To avoid this event, cells start producing telomerase, a ribonucleoprotein enzyme which synthesises telomeric repeats, causing the cell to become “immortal”. As telomerase is absent from normal cells (except during embryonic development) it has been suggested as a useful marker for a number of cancers.

The most widely used protocol is the TRAP (telomeric repeat amplification protocol) assay, a PCR based assay. Yoshida *et al* (1997), demonstrated that this assay gave a sensitivity of 62% and a specificity of 96.4%, with the only false positives being in patients with benign prostatic hyperplasia. Ohyashiki *et al* (1998) investigated expression using a semi-quantitative method in a variety of urothelial cancers. They also found no expression within normal samples and though they gave no indication as to levels of telomerase activity noted, did suggest its usefulness as both a diagnostic and prognostic indicator.

A comparative study by Saad *et al* (2002) investigated the possible use of telomerase (in the form of a TRAP assay) as a non-invasive detection assay along with NMP22, BTA stat and a haemoglobin (hb) dipstick test. The aim was to analyse their abilities as replacements for voided urine cytology (VUC). Of these assays, Telomerase, NMP22 and BTA stat demonstrated higher levels of both sensitivity and specificity than voided urine cytology. Telomerase alone demonstrated a sensitivity of 84% (the highest noted) and a specificity of 93% (again the highest) and it was found that combining this assay with NMP22 increased the specificity to 96%. Saad concluded that the detection of superficial tumours by these methods was greatly increased and in comparison to VUC, and that values were comparable to those gained with cystoscopy. Telomerase therefore appears to be a possible contender for the role of a marker for early detection (Saad *et al*, 2002).

Cytokeratin 20

Cytokeratin 20 is an intermediate filament involved in cell adhesion. In normal urothelial cells, its expression is related to cellular differentiation and as such it is

only found in superficial and occasionally intermediate cells of the urothelium (Harnden *et al*, 1999). The presence of cytokeratin 20 in a number of epithelial tumours was first noted by Moll *et al* (1992) through immunohistochemical analysis of paraffin-embedded tissue sections. Their demonstration of its presence within bladder cancer was mirrored in a number of studies by Harnden *et al* (1995, 1996, 1999). In bladder tumours, cytokeratin 20 is present throughout the urothelium, a change which can be noted immunohistochemically. Two PCR-based methods for the detection of cytokeratin 20 have been developed (Buchumnesky, 1998; Inoue, 2001), though there has been worry over the reliability of the assay as normal superficial cells express it, and these are the cells to be removed first. These methods could be used in a quantitative method, but cut-off points between normal and tumour samples would have to be decided upon, and these may involve an increase in false results.

Mini chromosomal maintenance proteins (MCMs)

This protein family is part of the pre-replicative complex essential for initiating eukaryotic DNA replication. They have been shown to be novel markers for proliferation, providing better information than more commonly used immunohistochemical markers such as Ki-67. Antibodies to MCMs 2, 3, 4, 5, 6 and 7 have been produced and of these MCM 2, 5 and 6 have been demonstrated to be present in bladder cancer. In normal urothelium expression of these proteins mainly occurs within the basal cells, though occasional superficial cell staining has been demonstrated. In non-invasive TCC, staining is dependent on grade, with well-differentiated tumours showing staining in 40% of cells, with an increase to 85% in cells of poorly differentiated tumours. In invasive cancers a similar staining pattern is seen therefore suggesting a link between MCM expression and grade. Though much of this work had been undertaken on paraffin sections, the authors alluded to preliminary work on urinary cytology samples which suggested the usefulness of antibodies to these proteins in a urine-based assay (Freeman *et al*, 2001).

Urinary bladder cancer antigen

This assay detects fragments of cytokeratin present in urine, specifically cytokeratins 8 and 18. The assay is in an ELISA format, marketed as the Ideal monoclonal UBC ELISA. It is a two step sandwich assay, where exfoliated cell pellets are added to an

antibody-coated plate, and then detected using a labelled secondary antibody and substrate system. Sanchez-Carbayo *et al* (1999) demonstrated a sensitivity of 73.1% when a 95% specificity cut off point for absorbance was used. There was some relation to stage and grade, with Grade I and II being similar, though an increase was noted at grade III. The expression increased with stage, but decreased sharply at stage 3 upwards. Some false negatives were noted from other urological malignancies and those from other areas of the body, thought to be due to kidney filtration (Sanchez-Carbayo *et al*, 1999). Sumi *et al* (2000) also investigated this assay and demonstrated a sensitivity of 86.9% and a specificity of 69.9%. Both groups suggested its use as an adjunct to cystoscopy in place of cytology due to its higher sensitivity. Sumi *et al* (2000) compared the ideal UBC ELISA with other available urinary tests, declaring it to be better than BTA *stat*, but comparable to the NMP22 and TRAP assays.

This is simply a selection of some of the markers currently being researched as possible urinary markers for the detection of bladder cancer. A number of other markers are under investigation, and these have been extensively reviewed by Koenig *et al* (2000), Grossman and Dinney (2000) and Saad *et al* (2001).

The most important requirement for any urine-based assay is the ability to detect the early stages of cancer development, an ability which many of the current urinary based assays do not have. It has been noted for a number of years that CD44 expression may prove to be a useful marker for early detection of this and a number of other cancers. The next section summarises current knowledge regarding CD44 expression within bladder tumours.

1.3.6 CD44 and Bladder Cancer

The detection of CD44 within exfoliated cells as a possible early detection assay was undertaken by Matsumara *et al* (1994) using an RT-PCR and southern blotting system. Through the use of these techniques they demonstrated the presence of large variant forms of the gene within tumour cells and focused on a 735bp product (the standard form gives a PCR product of 482bp) as a positive marker for tumour presence. An initial study with four normal and four tumour samples demonstrated a definitive difference in expression, and through serial dilution of samples, the assay

was able to detect at a very low copy number. In a larger scale study, using the 735bp product as the positive marker, a sensitivity of 91% and a specificity of 83% was noted. Tests were carried out on a number of patients with very early stage disease. Some expression was detected in patients with no clinical signs of the disease, suggesting CD44 changes occur early in tumour development. Though this assay was of high specificity and sensitivity, it was suggested that this assay format may be difficult to undertake in a clinical environment due to the care required in sample procurement and preparation with regard to mRNA protection (Matsumara *et al*, 1994).

Kan *et al* (1995) undertook a similar study, though their positive marker of choice was a 1500bp product, never found in normal tissues. Though they demonstrated both qualitative and quantitative changes within CD44 expression, they could find no direct link between specific variants to correlate with progression or metastasis. With their assay a sensitivity of 81% and specificity of 63.6% was achieved. Though this is lower than that found by Matsumara *et al* (1994), this study utilised non-involved tissue (taken from areas of the bladder normal to the eye) rather than normal tissue, therefore some neoplastic change may have already begun in these areas (Kan *et al*, 1995).

To investigate non-invasive detection in a more stable system, Matsumara *et al* (1994) moved onto protein based detection of CD44 expression (Sugiyama *et al*, 1995) using antibodies directed against the standard form of CD44 and a variant form, 11 (V6). Their study was based on western blot analysis of urinary cell pellets, immunocytochemistry of these cells and immunohistochemistry of snap frozen tissues. With western blotting, detection using the hermes (CD44S) antibody demonstrated the presence of a number of high intensity bands of high molecular weight (ranging from 70-220kDa) in samples from tumour patients, compared with bands of 70-90kDa in normal samples. When the presence of bands above 160kDa were used as a positive marker, the assay demonstrated a sensitivity of 75 % (100% specificity). IHC of normal tissue samples with both hermes and V6 antibody demonstrated staining within the basal layer of the urothelium, with some additional staining in higher layers of the urothelium with hermes. Use of the antibodies within tumour tissues illustrated a different pattern of staining, with strong

membrane staining through the full urothelial layer. Upon ICC using V6, only tumour samples were found to be positive, thereby demonstrating that the protein is retained once the cell has been exfoliated. Of the samples tested, 66% of early stage tumours and severe dysplasia expressed this abnormal CD44 expression, again emphasising that CD44 changes occur at an early stage (Sugiyama *et al*, 1995).

The expression of CD44 as shown by both a standard antibody and V6 antibody has been investigated by two other groups and they have also demonstrated similar patterns of staining in normal and tumour tissues (Ross *et al*, 1996 and Lipponen *et al*, 1998). Ross *et al* (1996) showed that as stage and grade increase there is a loss of reactivity with both antibodies, compared to the dense staining seen in both membranes and cytoplasm at a low stage and grade. Lipponen *et al* (1998) compared the expression of CD44 to a number of prognostic factors such as mitotic index, grade and DNA ploidy and, from these comparisons, concluded that though CD44 expression appears to be related to differentiation state and hence prognosis indication, the relationship is not clear cut.

Another variant under investigation is V2 (exon 7), to which an antibody was produced by Borgya *et al* (1995) and in an investigation by Muller *et al* (1997) was found to be present in tumours as homogenous staining in all layers of tumour tissue though no stage or grade dependent differences were noted.

Woodman *et al* (2001) have produced the first clinically relevant protein based assay detecting CD44 expression within urine-voided exfoliated cells. The assay is an ELISA assay which uses a double capture sandwich technique. Antibodies directed against exon 1 and V6 are attached to the plate, followed by addition of the cell protein, then a labelled Hermes antibody and detection system. To assess the diagnostic potential, the highest absorbance level from a patient known to be tumour free was used as the cut off point for the assay, where all values above this point were taken to be caused by tumour presence. This gave the assay a sensitivity of 81.1% (and a specificity of 100%). Therefore it is one of the most sensitive urinary based systems being researched (Woodman *et al*, 2001).

In conclusion, the current literature suggests that a protein based assay for urinary detection of CD44 should produce a highly sensitive assay for the early detection of bladder cancer.

1.4 Colorectal cancer

In England and Wales, colorectal cancer accounts for 13% of all cancers, being the second most common cancer in women and the third most common in men. Unlike bladder cancer, incidence is about equal in both sexes. Its incidence steadily increases with age from about 45 years, but rises sharply from 55 years and above. The incidence of this cancer is three to four times greater in developed countries such as the UK, USA and Australia, when compared to less well developed countries. Survival rates within England and Wales are approximately 40% after 5 years, decreasing to 33% in 10 years (National Statistics Office, 2001).

1.4.1 Aetiology of colorectal cancer

Two main forms of colorectal cancer have been noted in humans; hereditary and non-hereditary. The majority of colorectal cancers are non-hereditary (sporadic) and a number of risk factors involved in their development have been identified. These can be summarised as diet, physical activity levels, alcohol intake and cigarette smoking (Midgley and Kerr, 1999). This work focuses on the sporadic form, though the inherited forms will be briefly discussed in the next section (1.4.2).

The risk factors named above for colorectal development have been researched extensively in number of different populations, yet no firm biological basis for many of these factors has been determined. Dietary factors believed to be involved include decreased intake in fruit, vegetables, fibre and folate and increased consumption of meat (especially red meat) and fat. A Finnish study of dietary factors in a cohort of men investigated the effects of fibre, meat, vegetable and fat consumption. Compared to many developed countries, Finland has a lower colorectal cancer rate in both males and females, thought to be due to increased fibre intake. Previous studies from this group demonstrated that alcohol, especially at high levels, low folate levels and low

protein levels within a diet increased the risk of colorectal cancer. This study demonstrated no links between intake of fat etc., however calcium and vitamin D levels appeared to have a preventative effect. In comparison to other study cohorts, the Finnish population have a high dairy intake, especially milk. It is known that vitamin D and calcium act together to bind bile acids and fatty acids and this is thought to be the source of the protective effect. One strange result noted in this study was that consumption of brassica vegetables (cabbage and broccoli) appeared to increase cancer risk, whereas other countries have stated these decrease cancer risk. It was noted however that in the Finnish population the white brassica (cauliflower, white cabbage) are consumed more than the green brassica (Pietinen *et al*, 1995). Conversely a study within a Swiss population demonstrated a direct association between intake of meats, especially red meats and colorectal cancer, along with a protective effect with intake of fruit, vegetables and wholegrain (Levi *et al*, 1995). The protective effect gained from eating fruit and vegetables is thought to be due to both the presence of anti-oxidants and the intake of fibre. Fibre is thought to increase faecal transit time, thereby reducing the time that harmful carcinogens may remain in the colon and rectum.

Folate is known to be essential for normal synthesis and metabolism of amino acids, purines, pyrimidines and lipids. It has also been demonstrated to be a co-factor in the production of s-adenosylmethionine, which regulates methylation of specific bases to block gene expression. In rats, when folate levels are decreased breaks and hypomethylation of the p53 tumour suppressor gene occur. In a study of a group of US women it was found that decreased incidence of colorectal cancer occurred with an increase in serum folate levels (Kato *et al*, 1995). In comparison another American study demonstrated inconsistencies in the positive effect of folate, though investigation of smoking and alcohol within the same study did show an antagonistic effect upon folate levels by these two factors (Baron *et al*, 1995)

Le Marchand *et al* (1995) investigated the effects of physical activity and obesity on colorectal cancer development, demonstrating a slight increase in risk from obesity, but a strong link with physical activity. Though this activity had some link with occupation, with those in sedentary occupations eg office workers, having an increased risk of cancer development, the strongest protective effect came from

recreational physical activity in adulthood. This study also gave insight into how cancer risk changes when people move from one population to another. Immigrants to the US (an area of high cancer risk) who moved from areas of low colorectal cancer risk (eg Japan) demonstrated an increase in cancer risk, highlighted more when first generation people were studied. This effect has also been noted in countries where westernisation is occurring, thought to be due to dietary change including an increase in the intake of fast food (Le Marchand *et al*, 1995).

It has also been suggested that an increase in cancer risk can be attributed to socio-economic status and social class. This is thought to be due to a decrease in physical activity due to level of occupation and an increase in consumption of foods such as meat and eggs. Though these factors proved to be true in a study by Tavani *et al* (1999) in Italy, this theory has not been proven in any other population.

Though a number of attributing factors have been identified and investigated in a number of populations, results are conflicting for the effects of all factors. The main problem appears to be the variation in diet between both countries and populations of people within those countries. Aetiologic factors can therefore be considered as population specific, though dramatic changes within a populations diet appear to greatly affect incidence rates of this cancer type.

1.4.2 Molecular basis of colorectal cancer

Hereditary colorectal cancer

Two inherited conditions are known to predispose to colorectal cancer, namely familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC). FAP is caused by a germ line mutation in the APC gene, causing the growth of a large number of adenomatous polyps in the colon during the second and third decades of life. If these polyps are not removed then there is a 100% risk of colorectal cancer by the age of 40. These polyps are histologically identical to those that precede sporadic cancer (Lynch and Smyrk, 1995; Midgley and Kerr, 1999). The FAP phenotype presented varies in age of onset (from childhood to middle age) and the number of adenomas present, but a number of non-malignant extra-colonic features are also noted including osteomas, cysts and desmoid tumours (Lynch and

Smyrk, 1995). The exact function of the APC gene is unknown, but it is known to interact with beta-catenin and may therefore be involved in adhesion (Thomas, 1995). The type of mutation affects the disease features, though some environmental interaction also occurs, as noted in families where members all demonstrate the same mutation, but different phenotypes.

HNPCC is caused by germline mutations in genes that code for mismatch repair proteins. The best characterised are hMLH1 and hMSH2, though mutations in hPMS2 and hMSH6 have also been implicated. Mismatch repair protein inactivation leads to an accumulation of mutations which, if found to occur in critical tumour suppressor genes, can increase the possibility of malignant tumour development. HNPCC is characterised by early onset of cancer (typically around 40 years), with a number of tumours present throughout the colon and rectum. Other tumours may arise including endometrial, ovarian, stomach and pancreatic cancers. HNPCC accounts for approximately 5-10% of all colon cancers within the western population and often displays distinctive pathology when compared to sporadic cancers (Lynch and Smyrk, 1995; Wang *et al*, 1995).

Sporadic cancer

Colorectal cancer can be considered to be the most well characterised human cancer, as a genetic model for its development was published in 1990 (Fearon and Vogelstein, 1990) and it has since been adopted (with variations) as a general model for cancer development.

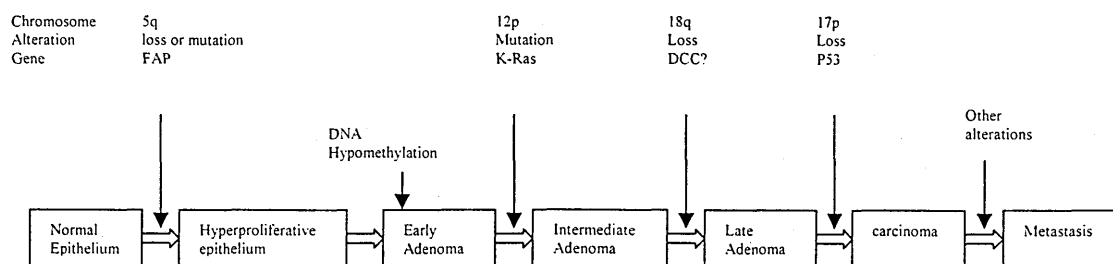


Figure 1.7 – Genetic model for colorectal tumourigenesis (from Fearon and Vogelstein, 1990). This model describes the genetic changes known to occur at each pathological stage within the development of colorectal cancer.

Figure 1.7 illustrates the model presented by Fearon and Vogelstein (1990). It demonstrates the pathological stages from normal colorectal epithelium to metastatic tumour, and demonstrates the key genetic mutations thought to influence the development of each stage. Though these mutations have been integrated into the model at specific points, some mutations may occur at slightly different timepoints. It was therefore suggested that the main factor is accumulation of mutations rather than order.

The first change is either mutation or loss of the APC gene. This induces hyperproliferation within the epithelium, and combined with DNA hypomethylation leads to the formation of early adenomas. It has been suggested that this inhibits chromosomal condensation and results in changes in chromosome number. It has since been noted that some hypermethylation may also occur, therefore clarification of the exact role of methylation in colorectal carcinogenesis is required (Navaratnam *et al*, 1999). K-Ras is a member of the Ras family of oncogenes and mutations within this gene lead to a permanent activation state, hence affecting the signalling pathways of which it is an integral part. Ras mutations are commonly found in approximately 50-80% of all colorectal tumours and it has been noted that adenomas with Ras mutations are more likely to progress than those without (Fearon and Vogelstein, 1990). DCC (deleted in colon cancer) mutations have been noted in 80% of colorectal carcinomas, though the exact role of the protein product is unknown. It does however show structural homology with the adhesion molecule ICAM1 and is therefore believed to be involved in cell adhesion, though it has also been suggested it expresses a tumour suppressor gene function. p53 mutations are noted in many tumours, with 70-80% of colorectal tumours demonstrating these.

1.4.3 Pathology and staging of colorectal cancer

The most common form of colorectal carcinoma is adenocarcinoma, that is epithelial in origin, though with gland-like tissue involvement. As can be seen from the genetic model, colorectal cancer develops through a number of distinct stages. The first change within the epithelium is hyperproliferation of a group of cells, though phenotypically they appear no different. Formation of an early adenoma can be compared with dysplastic features in other tumour types. From this stage it then

progresses through late adenoma to carcinoma with increasing phenotypic change until reaching the point of metastatic tumour. Tumours of the colon and rectum readily metastasise due to their close proximity to both blood and lymphatic systems, based on their functions as absorptive organs.

Staging of colorectal cancer is carried out under two schemes – TNM and the Dukes system. TNM tumour stages range from T0 and Tis to T4, depending on depth of invasion. Grading within this form of cancer is also highly variable, from GX, where differentiation cannot be assessed, to G1 which is well differentiated to G4, undifferentiated. Due to the proximity of the lymphatic system to these organs, lymph node involvement is classed at higher levels than for other tumours, with the highest level (N2) illustrating metastasis in 4 or more regional lymph nodes. Dukes staging is classified as either A, B, C and D (Raraty and Winstanley, 1998). A comparison between this method of staging and TNM is illustrated in table 1.1 below.

Table 1.1 – A comparison of TNM classification and Dukes stage in the staging of colorectal cancer (Adapted from Raraty and Winstanley, 1998).

TNM stage	Tumour	Dukes stage
0	Tis	-
I	T1 T2	A
II	T3 T4	B
III	Any T, with nodal involvement	C
IV	Any T, with nodal involvement and metastasis	D

1.4.4 Detection and diagnosis of colorectal cancer

Clinical signs and symptoms

These include changes in bowel habit which lasts for more than a few days, rectal bleeding or presence of blood in stool and stomach pain, either cramping or steady. Unfortunately, these signs alone could lead to diagnosis of a number of different conditions, including infection, inflammatory bowel disease (eg crohns, ulcerative

colitis) and haemorrhoids. Further thorough investigation is therefore required for any patient demonstrating one or more of these symptoms (Dashwood, 1999).

Gold standards for diagnosis

The first step in diagnosis is digital rectal examination, followed by sigmoidoscopy and colonoscopy (Dashwood, 1999).

Sigmoidoscopy is an endoscopic method, rigid or flexible, which allows visualisation of the whole rectum to the descending (sigmoid) colon. It demonstrates high specificity and sensitivity for tumour detection and is considered positive if the presence of a cancer or polyp over 1 cm in diameter is noted. When using rigid sigmoidoscopy, biopsy of affected areas is possible so that histological examination can be undertaken. Adenoma presence requires full colon examination as it has been noted that finding on sigmoidoscopy give an indication as to the pathology expected within the colon. For example, the presence of adenomas less than 1cm in size has been correlated with the absence of advanced pathology within the colon in most cases. This technique however does have drawbacks. Firstly due to its invasive nature, patient compliance for undergoing this form of examination may be low, though patients with a known history of colorectal disease are more willing to participate. Secondly, complications can arise with perforation of the colon occurring at a rate of 2 in every 10,000 examinations (Anwar *et al*, 1998).

Colonoscopy is the next stage of examination and involves the use of a flexible endoscope which examines the rectum, entire colon and sometimes part of the terminal ileum. Again this is a sensitive and specific technique and also allows for therapeutic intervention as polypectomy (removal of polyps) can be undertaken at the same time as the examination. This technique however also has similar drawbacks with patient compliance and perforation risk. An added problem is complications due to haemorrhaging after polypectomy, though this is rare (Abulafi and Marks, 1998; Anwar *et al*, 1998).

Non-invasive methods of detection

The only current clinical method for non-invasive detection is determination of the presence or absence of faecal occult blood (FOB). Colorectal cancers are known to

bleed intermittently, with 2/3 of cancers bleeding at least once every week. The most common of all FOB tests is the Haemoccult test, which is a qualitative assessment of faecal haem through pseudoperoxidase activity testing. The haematin component of haemoglobin catalyses the oxidation of Guaiac in the presence of H_2O_2 to produce a blue colour. The test fails to detect 20-25% of cancers and only provides an indication of disease presence; it is therefore only useful in a screening capacity as an adjunct to sigmoidoscopy and cystoscopy. Though the test is more acceptable for patients due to its non-invasive nature, the dietary restrictions which are imposed on patients for the benefit of the assay decreases compliance somewhat. It has been noted that a number of foods can produce false positives including red meat, broccoli, potatoes and bananas. It is also susceptible to false negatives from vitamin C and from the fact that not all cancers bleed, or bleed to a sufficient level for detection (Anwar *et al*, 1998).

Markers for non-invasive detection in research

There are two approaches to the study of tumour markers for colorectal cancer detection: assessment of markers in stool samples (through analysis of exfoliated cells) and serum markers. A large proportion of published work is based on the use of markers as prognostic indicators for disease progression, but most markers investigated appear to be “dual use” i.e. both diagnostic and prognostic.

Aside from the FOB test, stool marker assays are based on extraction of either protein or DNA from exfoliated colorectal cells found within stool samples. Most protocols are DNA based, detecting mutational changes within key genes involved in the development and progression of colorectal cancer. A number of methods for both extraction and subsequent analysis have been investigated, though results for some markers appears to be quite variable.

K-ras has been highly investigated due to its early involvement in colorectal cancer progression and its low variability in mutation. 40-50% of all colorectal tumours demonstrate a mutation within K-ras and most of these occur within codon 12. Through the use of an RFLP-PCR based method Puig *et al* (2000) were able to detect mutations in stool samples from 55% of adenocarcinomas and 27% of adenomas, whilst detecting no mutations within samples from patients with non-neoplastic

disease and no disease. When corresponding tissue samples were analysed, there was a direct match between stool and tissue result. Within this study K-ras detection demonstrated 100% specificity, with sensitivity appearing to be linked to stage of tumour. In a study by Dong *et al* (2001) a sensitivity of 47% was noted, again with a specificity of 100%, through the use of another PCR-based method.

Dong *et al* (2001) within the same study also investigated expression of p53 and a microsatellite region (BAT26) for detection purposes. Much of the work on p53 within colorectal cancer is based on its prognostic value, as its presence and mutation status is thought to be linked to tumour aggressiveness (Seregni *et al*, 2001). As a detection marker, mutations within the p53 gene could be considered to be occurring at a considerably late stage according to the model presented by Fearon and Vogelstein (1990 – see section 1.4.2). Investigation of eight mutations however, did detect 59% of colorectal samples, with mutations in stool again matching those in the tissue. These tumours were of Dukes stage B and above (Dong *et al*, 2001). Microsatellite instability was first noted within HNPCC patients, but the presence of these mutations was subsequently seen within sporadic cancer cases. Within this study, only 6% of samples showed a mutation within stool samples again mirrored by tissue results. When the results of all three assays (K-ras, p53 and BAT26) were combined 71% of all colorectal patients demonstrated some form of mutation and it was suggested that a multiple marker form of assay such as this would be the most useful for colorectal cancer detection.

The main serum marker linked with colorectal cancer is CEA, though it now appears to be generally accepted that whilst this is a useful marker for prognostic purposes it is of little or no value within early detection or screening. This is due to the finding that sensitivity of CEA detection depends on tumour stage and ranges from about 30% in Dukes A to approximately 80% in Dukes D. It has also been noted that a number of non-neoplastic conditions may also demonstrate high serum levels of CEA (Duffy *et al*, 2001).

From the literature there does not appear to be a large body of research on protein analysis for marker detection from exfoliated cells in stool samples. DNA assays have been demonstrated to be quite robust and as protein is less labile than DNA, protein based assays should present no problem.

1.4.5 CD44 and colorectal cancer

Yoshida *et al* (1996) demonstrated that the expression of CD44 on colorectal cancer cells exfoliated during luminal washings was abnormal. Though these luminal washings were carried out on resected sections, it was thought that the technique could be adapted for use with exfoliated cells in stool. The expression of CD44 was assessed using RT-PCR and southern blotting, with probes directed against exons V6 (exon 11), V7 (exon 12) and intron 9. Expression of V6 and V7 was noted in 73% of luminal washing samples, which included tumour and inflammatory bowel disease (IBD) samples. No expression of CD44 variants was noted in the IBD luminal wash samples, but when corresponding tissue samples were assayed, some low level expression was detected (Yoshida *et al*, 1996).

To investigate the localisation of CD44 expression within both normal colorectal tissue and tumour tissue, Gorham *et al* (1996a) employed the use of *in situ* hybridisation, using a standard and variant probe, and immunohistochemistry using a Hermes antibody (detects standard form of CD44) and an antibody directed against exon 11. In normal cells, standard CD44 at both mRNA and protein level was located within the epithelial cells lining the bases of crypts. No variant mRNA expression was noted, though some weak staining with exon 11 antibody was observed. In comparison, levels of standard CD44 mRNA and protein were higher in tumour tissue, and was present in all epithelial tissue within the tumour, though no association between Dukes stage was noted. Variant mRNA was also increased, though throughout the carcinoma as a whole, but no signal was detected within connective tissue and inflammatory cells, findings mirrored by the antibody results. High expression was especially noted in areas of invasion. This study again suggested the clinical use of this marker for early detection (Gorham *et al*, 1996a).

Goodison *et al* (1997) investigated variant exon splicing within CD44 isoforms present in colorectal tumours, with the aim of producing a specific assay for early detection based on the splicing patterns noted. Using a PCR based technique with a primer anchored at exon 3 and a second primer anchored to the end of junction 5 and the start of a variant exon, they investigated which variant exons were spliced next to standard exon 5. Through examination of both tumour and normal tissue they were able to detect a distinct pattern of CD44 expression. Within the tumour tissue, the presence of most variant exon junctions (5/7, 5/8 etc) was noted, though 5/7, 5/9 and 5/11 were found in elevated levels compared to the other junctions present. In normal tissue the 5/8 junction was most highly expressed. This demonstrated that there was a specific pattern of expression which could be exploited in a diagnostic assay (Goodison *et al*, 1997).

Ioachim *et al* (1999) noted that increased CD44 expression appears earlier than K-ras and p53 mutations within a significant number of tumours. An immunohistochemical study of CD44 expression compared with markers such as cathepsin D, ECM components, p53, Rb, and bcl-2 demonstrated a positive correlation between CD44 and size of tumour, cathepsin D and Rb protein expression, with an inverse correlation with bcl-2 (Ioachim *et al*, 1999). Though the exact nature of these interactions and their influence on tumour development is unknown, it was suggested that there may be some interaction between CD44 and Rb. Early expression of CD44 was also noted in colorectal polyps, in low and high grade polyps as well as those demonstrating adenomatous features. CD44 expression was highest in the earlier grade tumours using both standard and V5 (exon 10) antibodies. V6 (exon 11) was also investigated, but did not appear to be present in comparable levels (Neumayer *et al*, 1999).

CD44 has also been investigated as a prognostic indicator, through use of a serum based assay for V6 (exon 11) or immunohistochemical analysis. Serum V6 levels demonstrated a worse prognosis for those with elevated levels of this marker, and also correlated with V6 expressing tissue (Yamane *et al*, 1999).

It can therefore be seen that CD44 variant expression is a potential marker for early detection of colorectal cancer. The demonstration of specific splice patterns by Goodison *et al* (1997) could provide the basis for development of an assay either at mRNA or protein level. Translation of these patterns into peptide sequences should facilitate the production of a more specific antibody for detection than those currently in use.

1.6 CD44 as a marker for the monitoring of recurrence

As can be seen from the literature reviewed, a number of groups have demonstrated that CD44 is a potential marker of the presence of both bladder and colorectal tumours. Furthermore, its presence has been demonstrated in early stage tumours, rendering it a useful marker in the early detection of tumours for monitoring of recurrent disease.

With both cancers discussed, recurrence of disease is common-place, and sufficient monitoring of this presents a challenge to health professionals. Non-invasive methods of monitoring are required, and the use of CD44 presents one answer to this problem.

Though bladder and colorectal tumours are very separate in disease terms, current work on CD44 has provided the potential for “crossover” in the studies carried out thus far. Investigation of exon splicing within colorectal tumours has provided specific information which could be exploited for detection purposes. In bladder cancer, a protein based detection system based on standard and variant expression has been shown to produce high levels of sensitivity and specificity.

By employing the exon-splicing analysis procedure used in colorectal cancer, tumour specific markers could be analysed within bladder cancer. Comparison of these may demonstrate the presence of common splice variants which may have wider implications in other cancers. Identification of such markers could lead to assay development, whether DNA or protein based (as in the current bladder cancer ELISA) through translation of information. The production of antibodies specific for the

variant proteins formed as a result of this splicing could potentially result in a highly specific and sensitive assay.

Due to this crossover in knowledge and its potential benefits, it was decided to bring the study of these two tumour types together. The analysis of CD44 splicing in both these tumours is therefore the focus of this thesis.

1.5 Aims and Objectives

Project Aim

- *To analyse CD44 expression in neoplastic bladder and colorectal tissue for the presence of novel tumour specific patterns and to assess their potential as tumour markers.*
- *To utilise this information in reaching a platform from which an assay for the monitoring of bladder and colorectal cancer could be developed.*

Analysis of junction expression has demonstrated the presence of clinically relevant junctions within colorectal tumours, yet no protein based detection system has been designed using this information. A protein based detection system for bladder cancer has been designed and tested, yet not specific information on the splice junction expression within bladder tumours has been demonstrated. This project therefore aims to complete this information with the intention of providing a basis for further work on non-invasive assays for detection of these tumours suitable for use within district general hospitals.

Project Objectives

The initial objective is the establishment of an *in vitro* model for CD44 expression, using human tumour derived cell lines of both bladder and colorectal origin. Once initial expression studies at both mRNA and protein level have been assessed, the study will move on to the second objective, an analysis of exon junction expression in bladder cancer. This will be met by analysing expression at the mRNA level through the use of the *in vitro* model and patient samples (tumour, non-involved and normal). Once a tumour specific junction (or junctions) is discovered the third objective of translating mRNA expression data into a protein based system will be undertaken. This will involve the design of a synthetic peptide and the subsequent production of a polyclonal antibody. The final objective is to evaluate the novel translational antibody using the *in vitro* model and patient samples of bladder and colorectal tissue.

These objectives will each be achieved through a defined chapter of work, as described below.

Chapter 1 – Introduction

This chapter aims to give an overview on basic cancer biology at a molecular and cellular level, including an insight into diagnostic methods. It then investigates CD44 and its use as a tumour marker. The final sections review the molecular and cellular basis of bladder and colorectal cancer, including current diagnostic methods and current knowledge on CD44 expression within these tumours.

Chapter 2 – Evaluation of CD44 expression in relevant *in vitro* models

To fully investigate CD44 expression, an *in vitro* model based on human tumour derived cell lines was established. This gives insight into expression patterns within bladder and colorectal tissue, with variation in the study arising from the use of two cell lines of each tumour type. Expression studies at both mRNA (as an indicator of gene expression) and protein level give the fullest account of expression possible.

Chapter 3 – Analysis of CD44 variant exon splicing in bladder cancer

This chapter describes the analysis of exon splicing based on the method described by Goodison *et al* (1997). Use of the *in vitro* model established in chapter 2 provides the initial study material, followed by investigation of patient derived samples of neoplastic and non-involved tissue. The study moves from the exon link PCR method (as described by Goodison) to a PCR-ELISA method. For normal comparison exfoliated cells from voided-urine samples of a group of healthy volunteers was utilised.

Chapter 4 – Polyclonal antibody production and characterisation

The finding of a potential clinically useful marker in chapter 3 led to the design of a synthetic peptide to be used for antibody production. A polyclonal rabbit antibody was produced which was affinity purified to increase its specificity. Initial evaluation

of the antibody was undertaken using an antigen-antibody titration assay, the results of which are presented within this chapter.

Chapter 5 – Evaluation of 5/11 junction protein in cell lines and clinical samples

This chapter is the final study chapter and presents the results of the initial evaluation of the 5/11 antibody within a number of clinical samples. Bladder cancer evaluation was undertaken through use of two methods – western blotting and IHC, using exfoliated urothelial cells in urine and paraffin embedded tissue sections. Colorectal cancer evaluation involved IHC alone, and was carried out using paraffin embedded sections of dukes staged tumour types.

Chapter 6 – Final discussion

This chapter evaluates the work presented in this thesis and relates its potential diagnostic application. This includes a brief overview of the latest methods for marker detection and analysis of detection requirements. A discussion on which form of assay would be most viable for use within a district general hospital setting brings the chapter to an end.

Chapter 7 – Conclusions and further work

This chapter draws conclusions on the work as a whole, derived from the final discussion. The potential and possibilities for further work within the project will also be examined.

2 Evaluation of CD44 expression in human *in vitro* models

2.1 Introduction

The aim of this section of work was to investigate the expression of CD44, in both standard and variant forms, within cells of bladder and colorectal origin, and to establish an *in vitro* model within which subsequent work on detailed exon expression could be carried out.

This work focused upon the use of four cell lines, two of each tumour type and all human in origin. This *in vitro* model system may provide insight into CD44 expression within human tissues. Due to their ease of growth, large numbers of cells can be cultivated in a relatively short period of time, allowing extraction of sizeable quantities of mRNA and protein for study purposes. Although some variation within cell line growth is thought to occur over successive generations, this variation should not be too great as to lessen reproducibility.

The use of two cell lines of each cancer type may also provide insight into the variability of expression within tumours of differing stage and grade, though an assessment of this may prove to be difficult. For example, it is known that of the two bladder cell lines used, RT112 cells will form tumours in nude mice, whereas T24 cell lines will not, suggesting that the RT112 cell line is more metastatic when compared to T24.

Two of the cell lines utilised within this study (namely RT112 and HT29) had been used in previous studies of CD44 expression, and were therefore known to express CD44. The T24 and CACO2 cell lines had not, and therefore their expression needed to be investigated to ensure suitability for this project.

To fully investigate CD44 expression, assessment at two cellular levels was carried out: studies focusing upon gene expression (transcriptional studies) and those focusing upon protein expression (translational studies).

The transcription of RNA and subsequently mRNA from DNA within the cell nucleus is the first major step towards the cellular production of a protein. mRNA provides the template from which ribosomes can then translate the genetic code into a peptide sequence. Once this sequence has been formed it then undergoes post-translational modification within the endoplasmic reticulum eg glycosylation, before entering the golgi complex and being transported to its place within the cell as a fully formed protein. This is especially important in the case of CD44 as this protein undergoes a significant level of post-translational modification. Both glycosylation and chondroitin sulphate addition occur, resulting in a protein of high molecular weight.

Transcriptional and translational studies are both required to give the fullest account of expression possible. The presence of mRNA within a cell suggests that it intends to produce that protein, yet without investigating the protein presence, we cannot be sure that the protein is being produced. Also, as aberrant protein expression may be noted, in the case of CD44, seen as the presence of a large number of variant isoforms, it may prove to be important to find the basis of this aberrance ie whether caused at post-transcriptional or post-translational stages.

Transcriptional studies

These studies focused upon the presence of mRNA using the technique of reverse transcription – polymerase chain reaction (RT-PCR). mRNA is extracted from whole cells and is then reverse transcribed into complementary DNA (cDNA) using a reverse transcriptase (RT) enzyme. This cDNA is then used as the template within a standard PCR reaction. CD44 amplification was carried out using a primer pair which amplifies from exon 3 to exon 17, thereby encompassing the whole variant region. To ensure the integrity of CD44 amplification, southern blotting was carried out on all RT-PCRs using a “standard” probe aimed at exons 4 and 5 (see figure 1.2 for diagram of gene assembly).

Translational studies

Protein expression was investigated through two antibody based methods using a number of antibodies aimed at both standard and variant epitopes shown to be of importance in previous studies. To investigate the presence of standard and variant forms, total protein whole cell extracts were electrophoresed and then blotted (western

blotting) providing information which can be directly compared to that gained from RT-PCR. To investigate the distribution of CD44 protein within cells, immunohistochemistry was undertaken.

2.2 Materials and Methods

All buffer and solution formulations can be found in appendix B.

2.2.1 Cell Culture

All cell culture media and additives were supplied by Sigma (Poole, Dorset, UK) and all cell culture plastics were produced by Nalge Nunc (Fisher Scientific, Loughborough, UK).

Cell lines

Four cell lines were used during this project as model systems for CD44 expression, all human in origin. The bladder cell lines were RT112/84 (derived from a TCC of the bladder with epithelial morphology from a female patient, age unknown – ECACC : 85061106) and T24 (derived from a TCC of the bladder with epithelial morphology from an 81 year old caucasian female – ECACC : 85061108, also known by the names EJ138 and MGH-U1). The colorectal cell lines were HT29 (derived from a grade II adenocarcinoma, epithelial in origin from a 44 year old female caucasian – ECACC : 91072201) and CACO2 (derived from an adenocarcinoma with epithelial-like morphology from a 72 year old caucasian male – ECACC : 86010202). All cells were grown from seed culture stored under liquid nitrogen and were cultured in Dulbeccos modified eagles medium (DMEM) containing HAMS F12 nutrient mix (Sigma ref : D6421) , with the exception of T24 cells which were cultured in RPMI 1640 containing sodium bicarbonate (Sigma ref : R8758). All media was supplemented with 10% heat inactivated foetal calf serum, 1 ml of 200mM glutamine and 100µl of penicillin/streptomycin solution per 500ml bottle . Cultures were grown at 37° C in a 5% CO₂ atmosphere and were maintained by passage when they had grown to 70-80% confluence (typically every three to four days).

Passage

As all four lines were adherent, passaging required the use of Trypsin/EDTA to detach the cells from the surface of the culture flasks. Cultures were first washed in sterile PBS followed by the addition of 5 ml of 1x Trypsin/EDTA solution to each flask. This was incubated at 37° C for approximately one minute, or until detachment could be detected, upon which the cells were removed from the flask and placed into a centrifuge tube containing 5 ml of media (containing 10% FCS, as above), to inactivate the trypsin. This was then centrifuged at 218 xg for 5 minutes at 4° C. The supernatant was discarded and the cells resuspended in an appropriate amount of media and subcultured into fresh flasks.

Cell growth for Immunohistochemistry

Cells required for immunohistochemistry (IHC) were grown in slide flasks, small culture flasks with a detachable polystyrene slide as their base. Cells were seeded into these flasks and were grown for three to four days, by which point the base of the flask was covered by a monolayer of cells. The media was removed and the cells were washed three times in cold sterile PBS, before detaching the flasks from the slides. These slides were then placed in cold methanol for 10 minutes to fix the cells and left to dry at 4° C. They were then wrapped in foil and stored at -20° C until required.

2.2.2 mRNA extraction

mRNA was extracted using the mRNA micro-fastrack kit (Invitrogen, Paisley, Scotland, UK) according to manufacturers guidelines.

Cell Preparation

Cells for mRNA extraction were grown in T75 flasks (area of 75cm²) until approximately 80% confluent. Culture media was removed and cells were washed with sterile PBS. Fresh sterile PBS (10ml) was added to the flask and cells were detached from the surface using a cell scraper. The PBS (containing the detached cells) was then removed and placed in a centrifuge tube. This was centrifuged at 218 xg for 5 minutes at 4° C. The supernatant was discarded and the cells resuspended in sterile PBS (1ml), then transferred to a 2ml γ -irradiated tube. This was then

centrifuged for 5 minutes at 4° C. The supernatant was discarded and the cell pellet then used for extraction.

Synopsis of micro-fastrack protocol (Invitrogen)

A lysis buffer containing a protein and RNase degrader was preheated to 45° C then added to the cell pellet. After incubation at 45 °C (where cells were lysed and their proteins degraded), DNA was precipitated out of solution by addition of NaCl and then sheared by passing through an 18g needle. The solution was then incubated with oligo-dT cellulose (to which the mRNA binds), and subsequently transferred to a column. The oligo-dT was washed a number of times, before eluting the mRNA; mRNA was then precipitated using ethanol and sodium acetate and snap frozen in liquid nitrogen. After centrifugation the ethanol mixture was removed and the mRNA allowed to dry slightly before resuspending in 10µl of elution buffer.

Due to the small volumes gained from this procedure it was not possible to quantify the mRNA.

2.2.3 cDNA production by Reverse Transcription (RT)

cDNA production was carried out using a cDNA cycle kit (Invitrogen), according to the kit protocol.

mRNA (2µl) from cell extraction was mixed with 8.5 µl of water and 1 µl of oligo dT primer. This was incubated at 65° C for 10 minutes, then for 2 minutes at room temperature. The following reagents were then added :

RNase Inhibitor	1ml
5xRT Buffer	4µl
DNTPs	1µl
Sodium Pyrophosphate	1µl
AMV RT	0.5µl

This mix was then incubated at 42° C for 1 hour, then at 95°C for 2 minutes.

2.2.4 Standard CD44 PCR

cDNA from the above cycle was then used in a standard CD44 PCR reaction.

The PCR mastermix used was as follows :

10x PCR buffer (with MgCl ₂)	10µl
DNTPs	4µl
P1 primer	1µl
P4 primer	1µl
Nuclease free water	80µl
CDNA	3µl

The mix was spun briefly before being placed into the PCR machine. The PCR cycles were as follows :

Hot Start	94 °C for 1 minute
1 cycle of	85 °C for 1 minute
	[<i>Taq</i> polymerase (1µl) then added]
29 cycles of	94 °C for 1 minute
	55 °C for 1 minute
	72 °C for 2 minutes
1 cycle of	94 °C for 1 minute
	55 °C for 1 minute
	72 °C for 5 minutes

Taq polymerase was purchased from Qiagen (Crawley, West Sussex, UK). Samples were then placed on ice before analysis by agarose gel electrophoresis, in a 1.2% gel, including a molecular weight marker (100 bp ladder, Promega, Southampton, UK) with bands from 100bp to 1500bp, to ensure the PCR product was of the correct size. A positive result was indicated by the presence of a number of bands, the smallest of which is 482 bp.

Positive and negative controls were included in each reaction, a negative control included everything within the PCR mix except the cDNA to ensure that no contamination was present. The positive control used was a Beta-Actin amplification. Beta actin is an abundant protein within eukaryotic cells and is therefore readily detectable by PCR via its mRNA.

2.2.5 Production of standard probe for Southern Blotting

To produce a probe for the purpose of Southern blotting, a CD44 plasmid (based on pbluescript II SK, Stratagene, Gorham *et al*, 1996a), containing all standard exons,

was grown using the method of bacterial transformation. The plasmid was linearised using restriction digestion, and this was then the template DNA for a PCR to produce the final probes. The protocols for plasmid growth and linearisation can be found in appendix C.

Probe PCR

The standard probe hybridises to sections of CD44 exons 4 and 5.

The PCR reaction mix used was :

10 x buffer (with MgCl ₂)	10µl
DNTPs	4µl
P2 primer	1µl
Plasmid DNA	10µl
Nuclease free H ₂ O	74µl

The PCR was carried out using the same basic cycle as for a standard CD44 PCR (see section 2.4), but with 34 cycles for step 2 (instead of 29). Taq polymerase (1µl) was added after the hot start. PCR products were then analysed by agarose gel electrophoresis. Probes were purified using low melting point (LMP) agarose and phenol : chloroform extraction.

2.2.6 Purification of Probes from LMP agarose

A 1.2% gel was run at 4°C until the first band of the loading buffer reached the bottom of the gel. After staining with ethidium bromide, bands were excised from the gel using a clean scalpel and were placed into a 15 ml γ-irradiated centrifuge tube and their volume weighed. Five volumes of a 20mM Tris, 1mM EDTA solution was added to the agarose and this was incubated at 65 °C for five minutes to melt the gel. The solution was cooled to room temperature, and an equal volume of phenol (equilibrated to pH 7.9, Sigma) was added and the solution vortexed for 20 seconds. The aqueous phase was then recovered by centrifugation at 2060 xg for 10 minutes at room temperature. The aqueous phase was then removed and placed into a clean centrifuge tube. This step was repeated once with a mixture of phenol : chloroform (1:1) and once with chloroform alone. The final aqueous phase was placed in a clean tube and 0.2 volumes of sodium acetate and 2 volumes of 100% ethanol were added. This mixture was snap frozen and after thawing was centrifuged at 8870 xg for 10 minutes at 4 °C. The ethanol mix was removed and after allowing the DNA to dry

briefly the pellet was resuspended in 40 µl of nuclease free water. The probe concentration was assessed by reading the absorbance of the solution at 260 nm (in a quartz cuvette) and using the equation :

$$\text{Concentration (in } \mu\text{g per ml)} = A_{260} \times \text{dilution factor} \times 50$$

2.2.7 Southern Blotting

A standard CD44 RT-PCR was run as described in section 2.2.4. The products were run on a 1.2% agarose gel and the DNA transferred to a nylon membrane (Hybond N+, Amersham Biosciences, Little Chalfont, Bucks, UK) using capillary action with a 0.4M NaOH solution overnight, then hybridised with a labelled standard probe. Figure 2.1 below shows the apparatus set up for southern blotting.

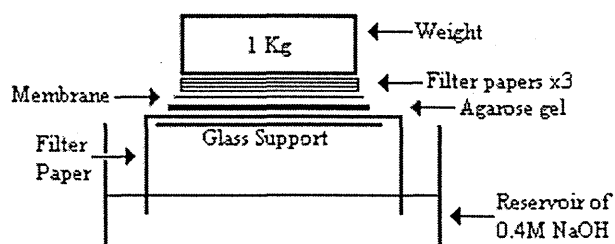


Figure 2.1 : Southern blot by capillary action with 0.4M NaOH. The agarose gel is placed on a filter paper wetted with NaOH. The membrane followed by three filter papers again wetted with NaOH are then placed over the gel. Finally a stack of paper followed by a 1kg weight are applied and the apparatus left overnight.

Probe labelling

Probes were labelled using an ECL direct labelling kit (Amersham Biosciences) according to the kit protocol. Probes were diluted to a concentration of 10ng/µl and 100ng (10µl) was denatured by heating for 5 minutes in a boiling water bath. The DNA was cooled immediately and kept on ice for 5 minutes, before being spun briefly. 10µl of DNA labelling reagent was added to the cool DNA and mixed gently after which 10µl of Gluteraldehyde were added and the mixture spun briefly. Probes were then incubated at 37°C for 10 minutes, and if not used immediately, were kept on ice for 10-15 minutes.

Hybridisation

Hybridisation was carried out using Gold Hybridisation Buffer (Amersham Biosciences) containing NaCl (1.461g/50ml) and blocking reagent (5g/50ml – supplied with the buffer). Membranes were first washed briefly in 2 times SSC before being placed in pre-warmed hybridisation buffer (42°C) and pre-hybridised for 1 hour at 42°C with gentle agitation. Labelled probe was added directly to the buffer (ensuring that the probe does not come in contact with the membrane directly) and incubation continued with the same conditions overnight. Membranes were then washed in two different wash solutions. Membranes were placed into wash solution 1 for two washes of twenty minutes with gentle agitation at 42°C and were then placed in wash solution 2 for two five minute washes at room temperature, again with gentle agitation.

Detection of DNA for Southern blotting was carried out using ECL detection reagents (Amersham Biosciences). Equal volumes of reagents one and two were mixed, then placed onto the membrane for 1 minute. The solution was then blotted off and the membranes placed into plastic bags before being placed into film cassettes. Film (ECL Hyperfilm, Amersham Biosciences) was preflashed before placing in the cassettes for one hour, then developed using Kodak developer and fixer.

2.2.8 Protein extraction

Protein was extracted from all four cell lines, once they had reached 70-80% confluence (typically, cultures for protein extraction were grown in T75 flasks, an area of 75cm²). All culture media was removed from the flasks and cells were washed with PBS before addition of PBS/protease inhibitor (1 ml of protease inhibitor in 50 ml of PBS, inhibitor cocktail produced by ICN Biomedicals, Basingstoke, UK) and scraping of cells off the surface of the flask using a flexible cell scraper (Nalge Nunc). Once all the cells were removed from the surface, the solution was transferred into a 15 ml centrifuge tube and the cells pelleted by centrifugation at 554 xg for 10 minutes at room temperature. The supernatant was removed and discarded and the pellet resuspended in 1.5ml of PBS/inhibitor. This was then transferred to an eppendorf tube and spun at 12073 xg for 15 minutes. The supernatant was removed and the pellet resuspended in 1.5 ml of CHAPS lysis buffer and snap frozen. This

was then thawed and incubated on ice for 1 hour before being centrifuged at 12073 xg for 30 minutes. The resulting supernatant is the cellular protein extract and after quantification was stored at -80°C.

Protein Quantification

Protein quantification was achieved through the use of the Bio-Rad protein assay dye (a modified bradford assay - Bio-Rad Laboratories, Hemel Hempstead, UK). The dye was diluted 1/5 with H₂O and 1.5 ml was added to samples and the mixture vortexed gently. After incubating at room temperature for a minimum of five minutes (maximum of 1 hour), the absorbance of the sample was read at 595 nm using plastic cuvettes. To quantify the protein a standard curve (consisting of 6 standards from 0 – 1000 µg of protein/ml) was constructed using an IgG protein standard (1 mg/ml, Bio-Rad). Aliquots of the sample were typically diluted 1/5 with H₂O unless the absorbance was above that of the highest standard. The final calculation of protein concentration was gained from reading the concentration of samples from the standard curve and multiplying by the dilution factor.

2.2.9 Immunohistochemistry

Prepared slides were removed from storage at -20°C and allowed to stand at room temperature (whilst still wrapped in foil) for 10 minutes, before being divided into a grid using an ImmEdge pen (Vector Laboratories, Peterborough, UK). To block non-specific antibody reactions, a 1:5 dilution of appropriate serum in TBS was applied to the slides (50µl/grid) which were then incubated in a humid atmosphere at room temperature for 1 hour (appropriate serum being the serum of the species that was used to raise the secondary antibody). Blocking serum was removed and the primary antibody then added to the slides. For CD44, 6 primary antibodies were used, all of which are aimed at portions of the protein that correspond with specific exons. All antibodies were monoclonal mouse IgG.

Antibody	Protein raised against	Source
Exon 1	Exon 1 (Clone VFF-10-442)	Novocastra / Boehringer Mannheim
Hermes	Exon 5	Gift from Dr E Butcher *
V2	Exon 7	Borgaya et al, 1995

V3	Exon 8 (Clone 3G5)	R + D Systems
V5	Exon 11 (Clone VFF-8)	Bender Med Systems
V6	Exon 12 (Clone VFF-18)	Bender Med Systems

*Stanford University

A 1/50 dilution of each antibody was made using tris-buffered saline (TBS) containing 1% serum as the diluent. 50µl of antibody was placed on its assigned grid and slides were placed in a humid atmosphere at 4°C to incubate overnight.

The next day slides were incubated at room temperature for 30 minutes followed by three five minute washes in TBS. As the detection system is linked to horseradish peroxidase, slides were incubated for 10 minutes in a 3% hydrogen peroxide / 100% methanol solution to exhaust endogenous peroxidases, followed by another set of three TBS washes. Secondary antibody (Biotinylated rabbit-anti-mouse IgG, Dako, Ely, Cambridge, UK) was applied in a 1/200 dilution (again in serum/TBS) and slides were incubated at room temperature for 1 hour, again in a humid atmosphere. Slides were subsequently washed in TBS (as above) before addition of the first stage of the detection system, the ABC (avidin-biotin conjugate) incubation. The ABC conjugate (Dako) was made up 30 minutes in advance, consisting of equal amounts of solution A (4µl) and solution B (4µl) with 500µl of TBS, and once placed on the slides was left for 30 minutes at room temperature (again in a humid atmosphere). Slides were then washed as before and incubated in Diaminobenzidine (DAB) solution for approximately 5-10 minutes (depending on speed of colouration). Slides were then washed in water before undergoing counterstaining for 30 seconds in haematoxylin, followed by washing in running water to produce the true colour of the counterstain. Excess water was removed from the slides before being mounted using a glycerol based mountant (Dako). Staining was evaluated using a minimum of a x20 objective.

2.2.10 Western Blotting

Electrophoresis

SDS-PAGE was carried out using a non-denaturing gel, consisting of a 4% stacking gel and a 6% resolving gel. Total cell protein samples (undiluted) were diluted 3 to 1 with a 4x loading buffer and boiled for five minutes. They were then spun for five seconds and placed on ice until required. 10µl of sample was placed in each well, and a set of high molecular weight markers was included (Cruz Markers, Santa Cruz

Biotechnology inc., Autogen Bioclear). The gel was run firstly at 75V for approximately 15 minutes, to allow samples to move through the stacking gel, then at 100V for about 1 hour or until the loading dye front reached the bottom of the gel.

Electroblotting

For the transfer of proteins to a membrane a semi-dry blotter was used (Hoefer TE70 series, Amersham). The gel was placed in a stack within the blotter consisting of four filter papers, membrane, gel and four more filter papers (see figure 2.2 below).

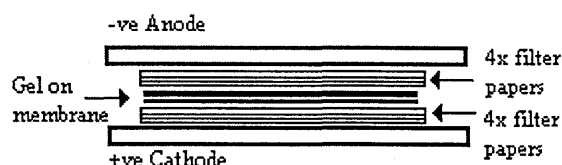


Figure 2.2 –Western blot setup within the semi-dry blotter. The SDS-PAGE gel is placed on top of the membrane and sandwiched between eight filter papers which have previously been soaked in transfer buffer.

Before placing in the stack, filter papers and gel were soaked in transfer buffer for approximately five minutes. The membrane (Immobilon-P, Millipore, Watford, UK) was prepared according to manufacturers instructions, by firstly wetting in methanol until opaque, soaking in water for five minutes, then into transfer buffer for five minutes. The stack was rolled with a serological pipette to ensure removal of air bubbles before placing the lid on the blotter. The blotter was run at 44mAmps for 55 minutes.

Detection and Identification

Detection and identification were carried out using specific antibodies raised against portions of the protein product of specific exons, Hermes, V3, V5 and V6 (see section 2.7 “IHC” for details). To block non-specific reactions membranes were firstly placed into a 5% non-fat milk/TBS solution for 1 hour, shaking gently. Primary antibodies were diluted 1/1000 in 5% non-fat milk/TBS and membranes placed into the solution to incubate overnight at 4°C. The next day membranes were incubated at

room temperature for 30 minutes, with agitation. Membranes were then washed five times for five minutes in a 0.1% tween/TBS solution), before being placed in the secondary antibody solution (Anti-mouse IgG peroxidase conjugate, Sigma) at a dilution of 1/1000 in 5% non-fat milk/TBS. A separate secondary antibody was used for detection of the molecular weight markers (goat anti-rabbit peroxidase conjugate, Santa Cruz Biotechnology Inc., Autogen Bioclear,) which required the membrane to be split. The antibody was used in a 1/500 dilution, again in 5% non-fat milk/TBS. Membranes were washed as before. Detection was carried out using an ECL Western Blotting Detection Reagent kit (Amersham Biosciences), the mix consisting of equal amounts of solution 1 and solution 2. Excess liquid was blotted off the membrane before applying the detection mix, which was left for 1 minute and then removed by blotting. The membrane was then placed into a plastic bag before placing in a film cassette and the film (Hyperfilm ECL, Amersham) exposed for 1-2 hours before being developed (developer and fixer by Kodak, obtained from Sigma).

2.3 Results

2.3.1 Transcriptional Studies

RT-PCR

Figures 2.3 and 2.4 show a Beta actin and CD44 RT-PCR for the RT112/T24 and HT29/CACO2 cell lines respectively. Beta actin was used a positive control to ensure integrity of mRNA.

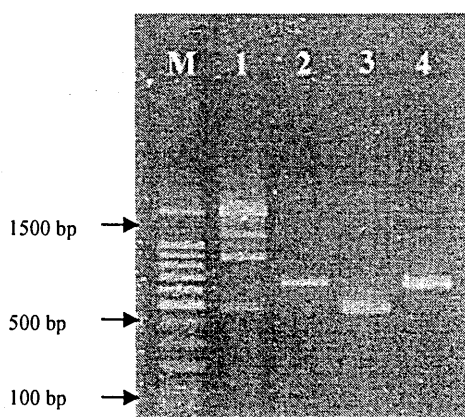


Figure 2.3 – Ethidium bromide stained gel of RT112 and T24 cell line RT-PCR. Lane M is the marker, lane 1 and 2 is the RT112 cell line and lanes 3 and 4 is the T24 cell line

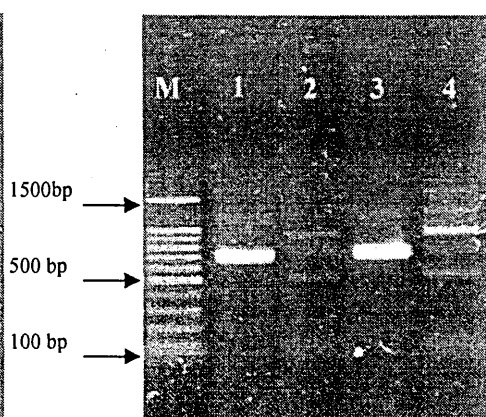


Figure 2.4 – Ethidium bromide stained gel of HT29 and CACO2 cell line RT-PCR. Lane M is the marker, lane 1 and 2 is the HT29 cell line and lanes 3 and 4 is the CACO2 cell line

Each gel shows two amplifications for each cell line. Lane M in both figures is the marker (Promega). Beta actin is shown as a band of approximately 600bp in lanes 2 and 4 (figure 2.3) and 1 and 3 (figure 2.4). The other lanes show a standard CD44 amplification.

The standard RT-PCR for CD44 was carried out using the primer pair P1-P4, which amplifies along the whole gene, whether standard or variant in form. If the standard form alone is present, then one would expect to see one single band, but if variant forms are present then multiple banding should be noted.

From the above figure it can be seen that all four cell lines express CD44 at an mRNA level, though expression patterns differ between each.

The RT112 cell line (figure 2.3, lane 1) expresses a variety of isoforms, as illustrated by the presence of a number of bands of varying molecular weight. The lowest band has a size of approximately 480bp and represents the standard form of the gene. Above this band approximately five bands can be seen, with a particularly bright and dense band at approximately 1,600 bp, suggesting that this is the predominant isoform. The majority of isoforms must therefore be variant in nature.

The T24 cell line (figure 2.3, lane 3) demonstrates a very different pattern of expression than that of RT112, with only one band noted on agarose gel analysis. This band is of approximately 480bp in size and relates to the standard form of the gene. The band is much brighter and more dense than that seen in the RT112 cell line.

The expression pattern for HT29 (figure 2.4, lane 2) is similar to that of RT112, with a number of bands of varying size again being noted. The lowest band seen is of approximately 480bp (again corresponding to the standard form) and is the least distinct band present. The largest band is approximately 1,600bp, with a distinct bright band seen at 900bp, suggesting that this is the predominant isoform. All bands vary in both size and intensity.

Figure 2.4, lane 4 illustrates CD44 expression within the CACO2 cell line, demonstrating the presence of a similar banding pattern to those for RT112 and HT29, but with less distinct presence. Though the band intensity is decreased in comparison, the lower band is again of approximately 480bp, and a distinct band of 900bp is noted, as seen for HT29.

Southern Blotting

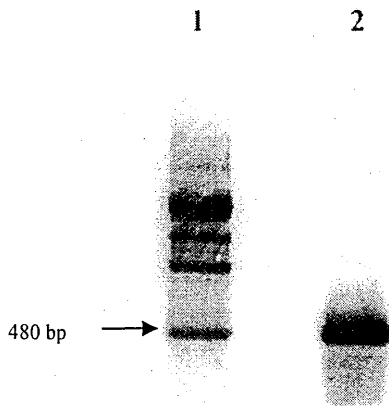


Figure 2.5 – Southern blot of ethidium gel in figure 2.3. Lane 1 is the CD44 banding in RT112 cells, whilst lane 3 demonstrates the pattern in T24 cells

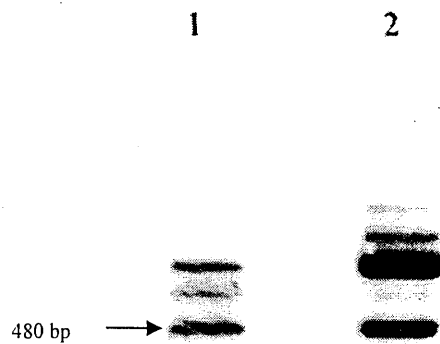


Figure 2.6 – Southern blot of ethidium gel in figure 2.4. Lane 1 is the banding seen with HT29 cells, whilst lane 3 demonstrates the pattern in CACO2 cells.

Figure 2.5 and 2.6 shows a Southern blot of the ethodium bromide stained gels seen in figures 2.3 and 2.4. The Southern blot demonstrates two things. Firstly all banding seen at an mRNA level with agarose gel analysis are true positive results, being due to CD44 amplification alone. Secondly, PCR products present in low concentrations can be visualised. All banding seen in all four cell lines with agarose gel analysis can be seen here, but extra clarity is provided for banding at the 1000-1600 bp region, and the presence of an extra band within the T24 cell line can be noted.

2.3.2 Translational studies

Immunohistochemistry

Plates 1-4 on pages ?? illustrate the immunohistochemical (IHC) results with each cell line, with six CD44 antibodies, E1, Hermes, V2, V3, V5 and V6.

The results with each antibody are summarised in table 2.1. Staining intensity is denoted through the use of “+”, with “+” denoting weak staining, “++” moderate staining and “+++” strong staining. The staining description is divided between membrane (M) and cytoplasm (C).

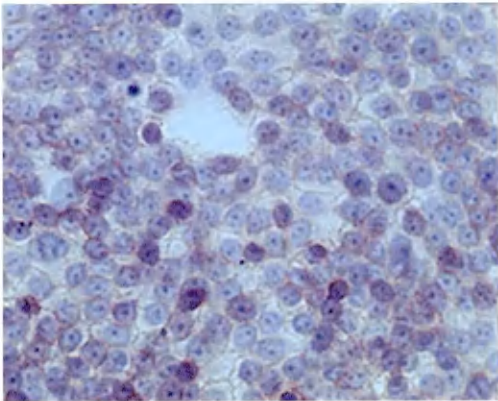
Table 2.1– Analysis of CD44 immunohistochemical results within the RT112, T24, CACO2 and HT29 cell lines.

		Antibody					
		E1	Hermes	V2	V3	V5	V6
Cell Line	RT112						
	M	++	+++	+	+++	++	++
	C	-	++	+	++	+	+
	T24						
	M	-	++	-	-	+	++
	C	-	+	-	-	+	+
	HT29						
	M	++	+++	+++	+	++	++
	C	++	+	+	+	+	+
	CACO2						
	M	+	++	+	++	+	+
	C	-	-	-	+	-	-

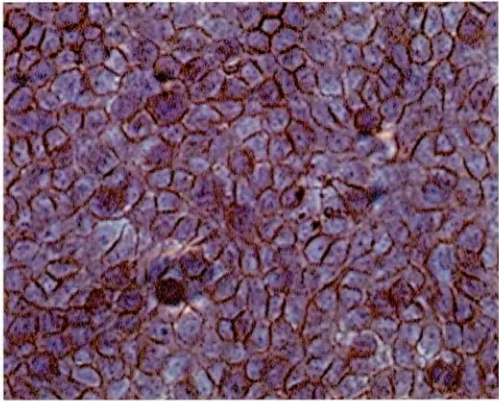
All cell lines expressed CD44 protein, though as for the mRNA expression, patterns differed between each. As CD44 is a cell adhesion molecule, it would be expected that all staining would be concentrated within the cell membrane, but as can be seen in the table above, cytoplasmic staining was also noted with some antibodies. E1 and Hermes antibodies detect the standard form of CD44, with V2, V3, V5 and V6 detecting variant forms.

The E1 antibody detected protein in all cell lines, except T24. Membrane staining intensity differed between the three cell lines, with RT112 showing the strongest and CACO2 the weakest. Some cytoplasmic staining was noted in HT29 cells, giving cells a grainy appearance. Though both this antibody and Hermes detect standard forms of the protein, there was a great difference in staining patterns. Using the Hermes antibody, protein expression was noted in all cell lines in both membrane and cytoplasm, except for CACO2 cells, where staining was noted in the membrane only. RT112 and HT29 showed strong membrane staining, with moderate staining in

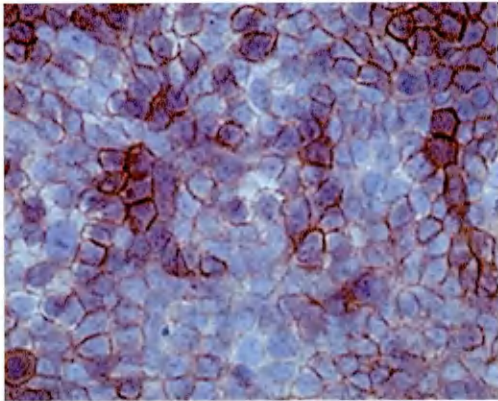
Plate 1 : Results of immunohistochemistry with anti-CD44 antibodies on RT112 cells



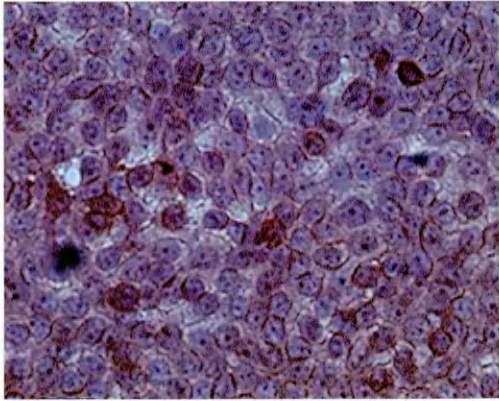
1A – RT112 cells stained with E1 antibody



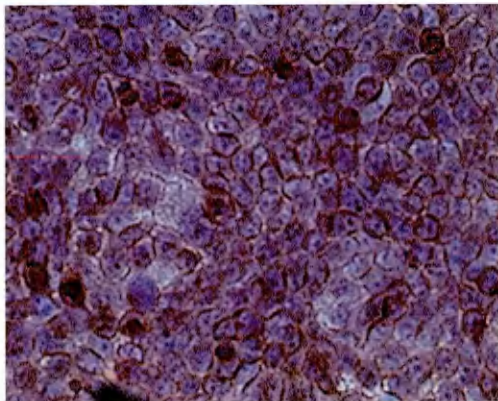
1B – RT112 cells stained with Hermes antibody



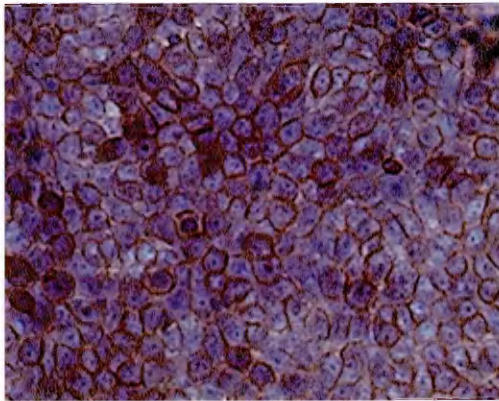
1C – RT112 cells stained with V2 antibody



1D – Cells stained with V3 antibody

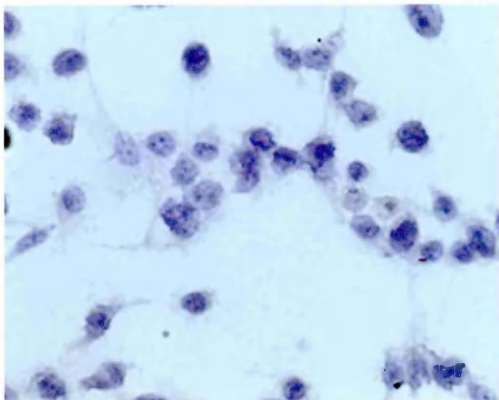


1E – RT112 cells stained with V5 antibody

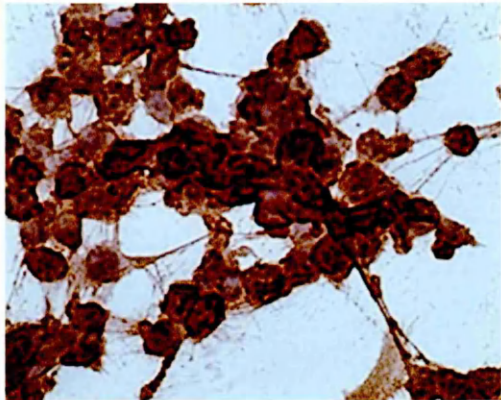


1F – RT112 cells stained with V6 antibody

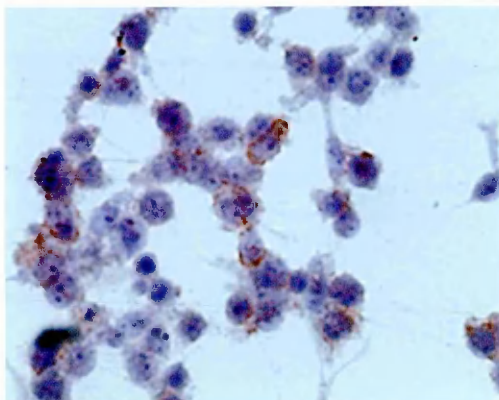
Plate 2 : Results of Immunohistochemistry with anti-CD44 antibodies on T24 cells



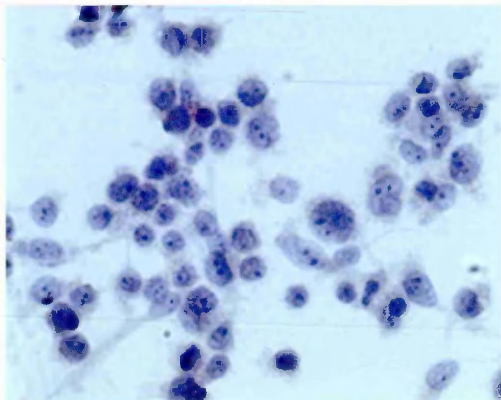
2A – T24 cells stained with E1 antibody



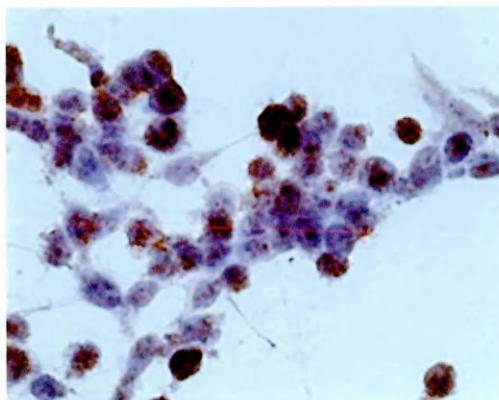
2B – T24 cells stained with hermes antibody



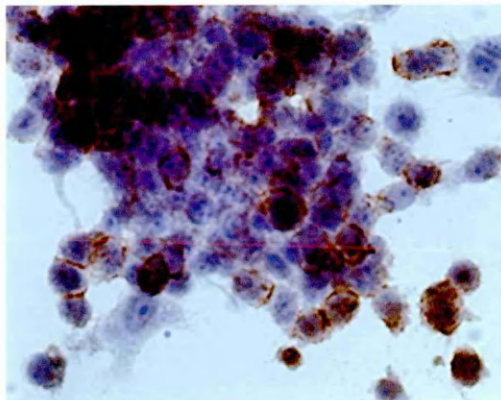
2C – T24 cells stained with V2 antibody



2D – T24 cells stained with V3 antibody

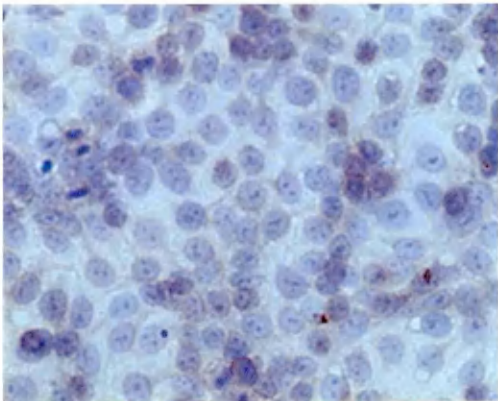


2E – T24 cells stained with V5 anitbody

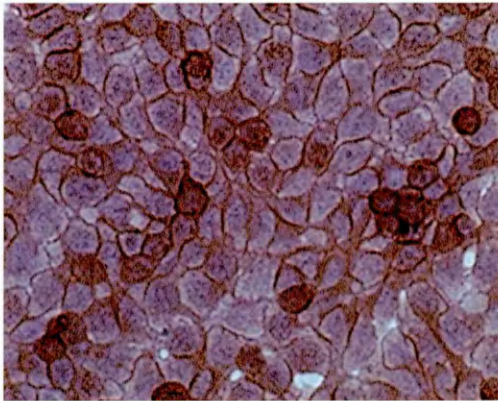


2F – T24 cells stained with V6 antibody

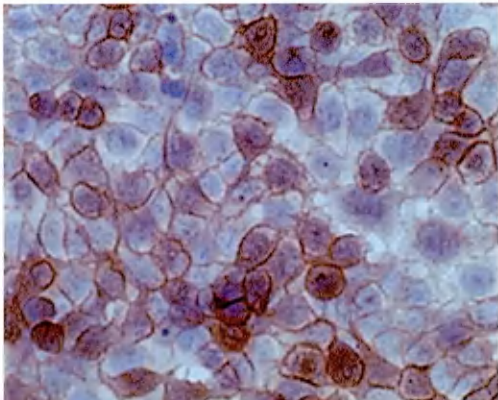
Plate 3 : Results of Immunohistochemistry with anti-CD44 antibodies on HT29 cells



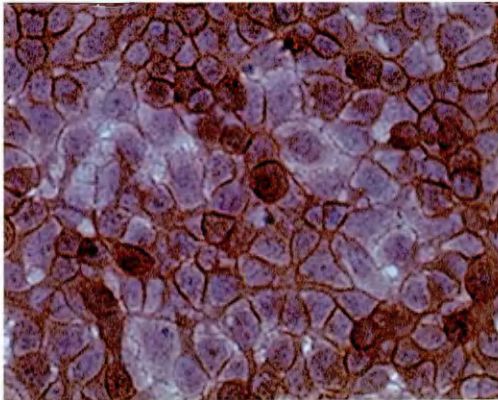
3A – HT29 cells stained with E1 antibody



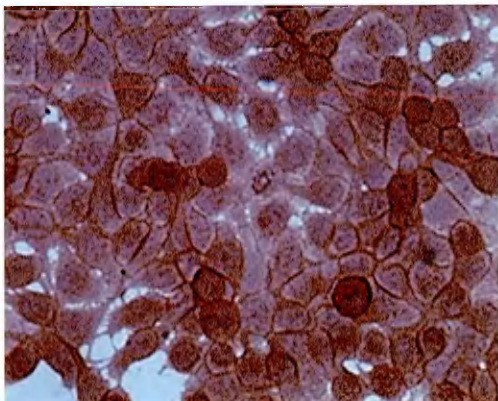
3B – HT29 cells stained with Hermes antibody



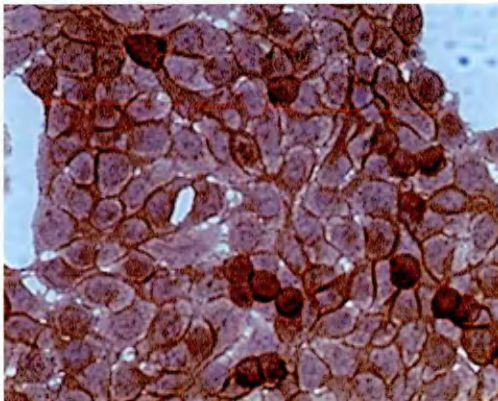
3C – HT29 cells stained with V2 antibody



3D – HT29 cells stained with V3 antibody

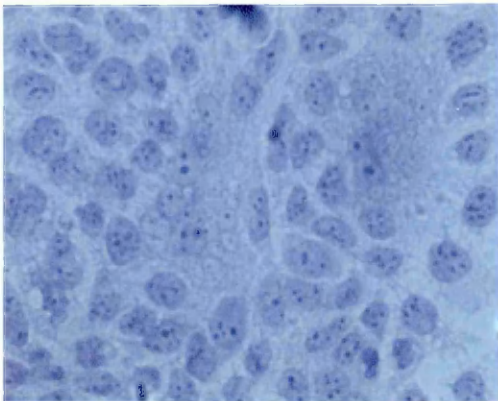


3E – HT29 cells stained with V5 antibody

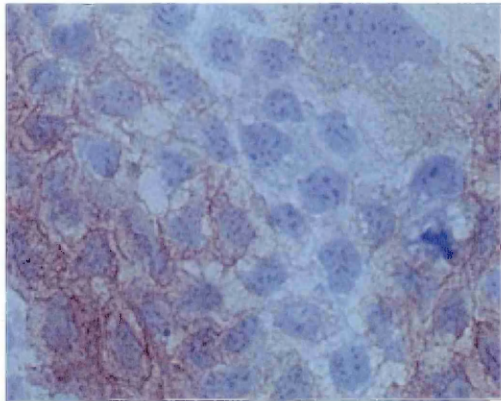


3F – HT29 cells stained with V6 antibody

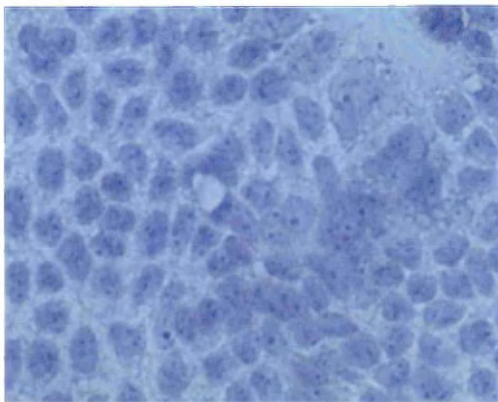
Plate 4 : Results of Immunohistochemistry with anti-CD44 antibodies on CACO2 cells



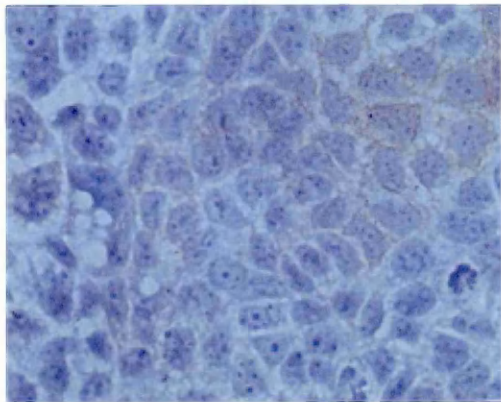
4A – CACO2 cells stained with E1 antibody



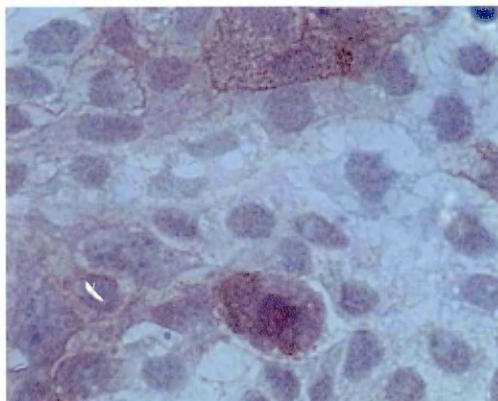
4B – CACO2 cells stained with Hermes antibody



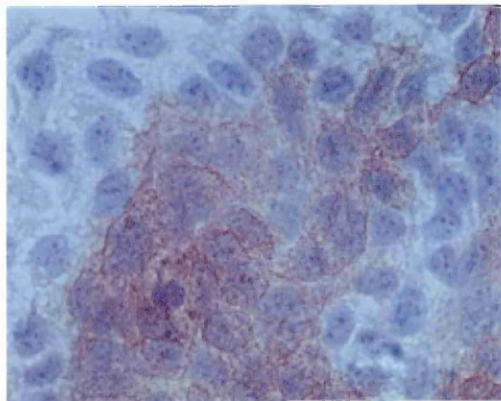
4C – CACO2 cells stained with V2 antibody



4D – CACO2 cells stained with V3 antibody



4E – CACO2 cells stained with V5 antibody



4F – CACO2 cells stained with V6 antibody

CACO2 and T24 cells. Cytoplasmic staining also varied, with RT112 showing the strongest.

Staining with the variant antibodies V2, V3, V5 and V6 also demonstrated varied protein expression within all cell lines. RT112 was shown to express all protein variants, though strongest staining was seen with V3 antibody. V5 and V6 gave moderate staining, with V2 showing the weakest, cytoplasmic staining was noted with all antibodies. In comparison, the T24 cell line showed no expression with V2 and V3, and weak and moderate staining with V5 and V6. Weak cytoplasmic staining was also noted. HT29 cells showed the same expression patterns as RT112 with V5 and V6 antibodies, but showed its strongest staining with V2 antibody and its weakest with V3. Weak cytoplasmic staining was noted for all antibodies. CACO2 only demonstrated cytoplasmic staining with V3 antibody, which also gave the strongest staining of all the variants, moderate in nature. All other antibodies resulted in weak staining.

Western Blotting

Figures 2.7 – 2.10 demonstrate the results of the western blotting with the Hermes, V3, V5 and V6 antibodies.



Figure 2.7 Western Blot with Hermes Antibody. Lane 1 = RT112, followed By T24, HT29 and CACO2 (lane 4).

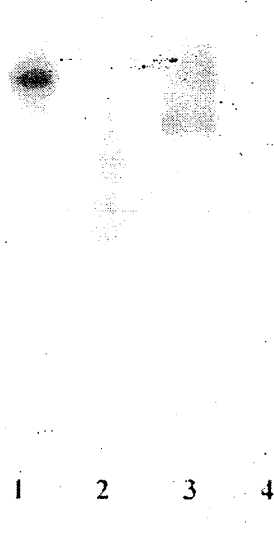


Figure 2.8 Western blot with V3 antibody. Lane 1 = RT112, followed by T24, HT29 and CACO2 (lane 4).

From the western blot in figure 2.7, it can be seen that all cell lines express CD44, as would be expected from the IHC results. Each cell lines demonstrates a different banding pattern, suggesting the presence of a variety of isoforms. T24 (lane 2) demonstrates the largest number of isoforms, with a banding pattern of seven bands. Of these bands, four are below the size of the standard 90kDa glycosylated form of CD44. In the other three cells lines, the lowest band seen is of approximately 90 kDa. But unlike the T24 cell line, it is not the strongest band seen. In the HT29 and RT112 cell lines the number of bands present are difficult to dissect out and count. Figure 2.8 demonstrates a western blot with V3 antibody, and a differen pattern of expression. The T24 cell lines still appears to express the most bands, with approximately five present. The HT29 and RT112 cell lines also express less bands, whilst the CACO2 cell line expression is very low, with only one very light band present.

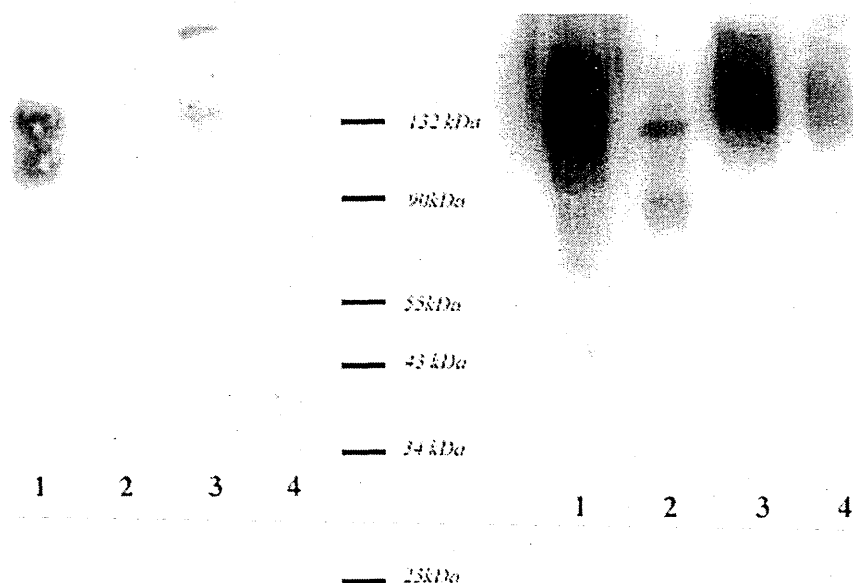


Figure 2.9 Western Blot with V5 antibody.
Lane 1 = RT112, followed by T24, HT29
and CACO2 (lane 4).

Figure 2.10 Western blot with V6 antibody.
Lane 1 = RT112, followed by T24, HT29 and
CACO2 (lane 4).

The expression of V5 protein (figure 2.9) only appears to occur within RT112 and HT29 cells, though some expression was noted upon IHC of the T24 and CACO2 cells. The RT112 presence is as a dark smear, whilst HT29 shows two very light bands. Figure 2.10 however demonstrates a more similar pattern to that seen in

Hermes for the RT112 and HT29 cell lines. The T24 cell lines only demonstrates three bands, but these can be directly compared to those present within the Hermes blot. Caco2 cells also demonstrate less banding, though these can again be compared to the hermes blot.

2.4 Discussion

From the results presented, it can be seen that all cell lines used in the *in vitro* model express CD44, at both an mRNA and protein level. However, this expression does vary, between both tissue type and cell lines of each tissue type.

Transcriptional studies exhibited the standard form of CD44 via the presence of a band of approximately 480bp in size (as seen in figures 2.3 to 2.6) and which is concurrent with the findings of Screaton *et al*, (1992). A number of bands above this size are also present and these must therefore be considered to be variant isoforms. These are larger than the standard form as they contain all ten standard exons plus any number of variant exons. These variant isoforms may be numerous in number and vary greatly in size. The presence of both standard and variant isoforms is also seen in translational studies where western blot results demonstrate similar multiple banding patterns (figures 2.7 – 2.10). In RT112, HT29 and CACO2 cell lines, the lowest band seen is that of the glycosylated standard form, approximately 90 kDa in size (Sugiyama *et al*, 1995), detected with the Hermes antibody. The T24 result is different in that a number of bands are seen below this 90 kDa size, though the reasons for this are unclear. The presence of variant isoforms in itself differs from the patterns of expression found within normal tissues, except for certain tissues where specific variant forms are found eg CD44v7-9 (addition of exons 12-14) within epithelial tissues.

It was noted by Woodman *et al* (1996) that the number of bands seen at an mRNA level, relating to transcript number, do not correspond directly with the number of bands seen at a protein level. They showed that though a number of transcripts can be demonstrated, the protein forms seen are less in number, suggesting that not all

transcripts are fully translated. This was the case with the RT112, HT29 and CACO2 cell lines, but not with the T24 cells. At an mRNA level, only one band was seen, but with Hermes antibody detection at western blot, approximately 7 bands can be seen, with the majority being below the expected standard form molecular weight. Detection with V6 antibody also demonstrated the presence of more than one band. The presence of bands of a lower molecular weight than expected could be indicative of mis-translation of mRNA transcripts, or incomplete processing at the level of post-translational modification. It is most likely that the latter is true however, as the bands range from the approximate size of the un-glycosylated standard form of CD44 up to and above the glycosylated standard form.

The lack of post-translational modification could be due to decreased function of the golgi, resulting in incorrect or poor modification of proteins. It would be expected that this would cause number of inherent problems in the cells in general through the mis-modification of other proteins. It may be that as CD44 is such a heavily glycosylated protein (the standard form more than doubling in molecular weight when fully modified) it is more severely affected than other proteins which require modification. Experimentally, some information regarding the cause could be gained through the use of a transfection assay where protein functionality can be assessed. Transfection of native protein from T24 cells into cells known to contain working golgi bodies should result in fully glycosylated molecules if the problem is due to the golgi alone. If full glycosylation is not noted, then this could point towards the presence of mutations within the transcript which when translated result in a protein which cannot undergo full modification as required.

The expression of larger isoforms of protein within these cell lines can be attributed to both the inclusion of extra variant exons within the transcript and extra post-translational modification which is known to occur as a result of their exons presence. When variant exons are present, extra sites of both O- and N-linked glycosylation are included into the protein, along with extra sites for both chondroitin and heparan sulphate side-chain addition. This extra modification will increase the molecular weight of the protein substantially. The finding by Woodman *et al* (1996) of decreased protein banding when compared to mRNA transcript number could also be related to this finding. If certain variant exons are more likely to provide extra sites

for modification, then if a protein contains the product of a large exon with no modification, but a protein with a smaller exon product which includes a modification point are both present, this may result in proteins of the same molecular weight, yet different in peptide structure.

The mechanisms responsible for the change in alternative splicing towards more aberrant expression are not yet fully understood. These could in their simplest form be due to mutational change within the regulatory mechanisms, resulting in decreased or complete loss of control. Due to the continued presence of the standard form of CD44, even when a large number of variant isoforms are present, a decrease in control is probably the more likely mechanism. It has been demonstrated however, that a number of growth factors and cytokines can modulate CD44 expression within tumour cells, causing both increased and decreased levels of standard and variant forms at an mRNA level. Studies upon a number of cultured human tumour cell lines have shown these effects. For example, the human neuroblastoma cell line SK-N-SH has been shown to increase its expression of CD44V3,5,6,7 and 8 (exons 8,9,10,11 and 12) in the presence of insulin-like growth factors (IGF) and / or platelet derived growth factor (PDGF); whilst the presence of nerve growth factor (NGF) within this line resulted in decreased expression of CD44V3,6 and 9 (Hamada *et al*, 1998).

During tumour progression, a number of growth factors and cytokines are present, either produced by surrounding cells, or by the tumour cells themselves. This interaction could allow modulation of expression throughout the neoplastic process, enabling increased production of CD44 when required. Whether caused by mutation, growth factor interaction or both, major changes are known to be present.

The main consequence of alternative splicing within the neoplastic process appears to be the change or modification of CD44 function. Due to its involvement in both cell-to-cell and cell-to-ECM adhesion, it is suspected of involvement within the process of invasion and the formation of metastases.

Changes in a number of cell adhesion molecules occur throughout tumour progression. The first stage of progression is over-proliferation, and to accommodate the growing mass of cells within the tissue boundary, a certain number of changes to

cell adhesion must occur. Once cells begin to invade surrounding tissue, cells must be able to decrease their adhesive ability and become motile to allow invasion of, and movement through, the basement membrane and underlying ECM tissue. Once the lymph or blood vessels are reached then the ability to bind to and penetrate these vessels is required. Once inside these vessels the cells must aggregate to reduce damage and detection by immune system surveillance until the secondary growth site is reached.. The ability to modulate function and expression of adhesion molecules is crucial throughout this whole process, and this is possible with the CD44 protein.

A number of possible functional changes may occur to CD44 due to the presence of variant exons. Firstly, the extra peptide may in itself change function by affecting the tertiary structure. Changes in structure may alter the link module, by which CD44 adheres to its main ligand, hyaluronan. This immediately affects binding to the ECM. Increased glycosylation and side-chain addition may further increase tertiary structure change. Secondly, the addition of chondroitin sulphate side chains are known to allow CD44 to bind to collagen, laminin and fibronectin within the ECM. Thirdly, binding of heparan sulphate side chains allows binding of heparin binding growth factors, such as hepatocyte growth factor. These heparin binding growth factors are involved in growth, motility and invasion of tumour cells, along with the possible stimulation of angiogenesis. CD44 is known to act as a signal transducer and as a co-factor for growth factor receptors. It may therefore play a role within proliferation etc of the tumour cells, and may act in an autocrine manner. It has been demonstrated that the oncogene product c-erb-B2, a constitutively active growth factor receptor has the ability to modulate CD44 expression on ovarian tumour cells. It was further noted that in some cells there appears to be a bond between the two proteins, which allows CD44 to be used as a co-factor by the receptor. Binding of hyaluronic acid to CD44 stimulates signalling cascades via the c-erb-B2 receptor which in turn increase cell growth (Hamada *et al*, 1998).

As mentioned earlier, it is not just aberrant expression of protein that is noted, but also over-expression, as can be seen from the IHC results. This suggests that controls upon overall levels of expression are also affected, and it has been noted by some groups that as tumour stage and grade increases, so CD44 levels rise to a peak point, where at very high stage and grade expression suddenly drops (Sugino *et al*, 1996).

The IHC results demonstrated the extent of this protein over-expression, as both membrane and cytoplasmic staining were noted. As CD44 is a transmembrane protein, with the major section of protein being extracellular it would be expected to find the majority of CD44 presence on the membrane alone. In all cell lines when incubated with the Hermes antibody, all showed expression in both membrane and cytoplasm of moderate to high density. In some cells this cytoplasmic staining was seen to be "grainy" in nature, suggesting aggregation of protein was occurring. The presence of strong membrane staining suggests that high levels of CD44 protein are being incorporated within the membrane. The cytoplasmic staining could be due to a build-up of functional protein which is unable to integrate in the membrane due to saturation. Alternatively, staining could be due to the presence of numbers of defective or immature proteins, and though mechanisms are in place to remove damage proteins, due to the high concentrations thought to be present, these may be ineffective. If large volumes of immature proteins were present, it would be expected that more unusual western blot banding patterns (similar to that of T24) would be seen in all cell lines, but as this is not the case, then the former explanation of high protein volumes may be the more likely. If the unusual banding seen in the T24 cells is due to differential glycosylation, then this may affect the proteins ability to integrate into the membrane. This could be demonstrated by immunohistochemistry as a lack of membrane staining, with the whole cell demonstrating the same density of staining. This is not the case with the majority of antibodies tested. This pattern of more distinct membrane staining was also seen with the majority of all antibodies in the other three cell lines. Where cytoplasmic staining appears grainy, this could be due to removal mechanisms within the cells, where proteins are encapsulated within proteosomes to be degraded.

As stated in the introduction to this chapter, the main aim of this section of work was the establishment of an *in vitro* model for use in subsequent CD44 analysis. These cell lines can be considered as such due to the fact that firstly CD44 is expressed by all, and that this expression is aberrant in nature and therefore mirrors that found in malignant tissues. The use of two lines of each tumour type does provide variation as was expected, but the extent of the heterogeneity, especially within the bladder cell lines could be thought of as greater than expected. This heterogeneity would also be

expected within patient tissue samples, due to differences within tissue micro-environments between different persons. These cell lines can therefore be considered as a good working *in vitro* model in which to undertake further CD44 studies.

3 Analysis of CD44 variant exon splicing in bladder cancer

3.1 Introduction

To investigate the potential of CD44 as a marker for the early detection of bladder cancer a more detailed analysis of variant expression was deemed necessary. The presence of variant forms, and over expression of these, was noted in the *in vitro* model evaluated in chapter 2. Though this overexpression itself could be useful diagnostically, it was felt that analysis of expression of specific variant exons could provide evidence of a characteristic pattern which could be exploited within a screening test. It was necessary to ensure that variant forms present within tumour-derived tissue differed to that found in normal bladder tissue, as the lining of the bladder is a specialised epithelium and epithelial tissues are known to display characteristic CD44 variants. It was then necessary to determine which exon or exons would be of most use in distinguishing between normal and tumour-derived tissues.

A number of studies have investigated the presence of CD44 variant exons, using both nucleic acid and protein based techniques. The majority of studies have focused upon the presence of only one or two exons, yet results suggested the expression of a number of variant exons at any one time (see chapter 2, RT-PCR). To investigate expression fully a technique examining all variant exons present was required. Goodison et al (1997) carried out an investigation of CD44 exon splicing within colorectal tumours, focusing upon splicing at the standard / variant exon border. This exon linkage was determined through the use of a PCR-based assay (termed an “Exon link assay”) which uses one primer anchored at standard exon 3, and a second primer which overlaps the junction between standard exon 5 and a variant exon (exons 7-14). This primer design is shown in figure 3.1 below.

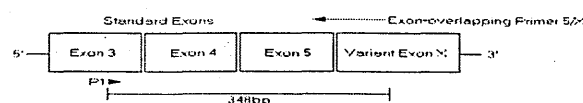


Figure 3.1 Diagram showing primer design for exon link assay (Goodison et al, 1997). One primer (P1) hybridises to a standard exon (3) whilst the second primer hybridises to the junction between standard exon five and one of variant exons 7-14.

Their findings established a characteristic pattern of splicing within colorectal tumours when compared to normal tissue, therefore giving rise to potentially useful tumour markers.

It was decided to utilise the exon link assay to analyse splicing patterns within bladder tumours, to determine the presence of tumour specific markers which could be exploited within an early detection and monitoring assay. This section of work therefore fulfills the second objective as stated in chapter 1.

In addition to the exon link assay a second method for analysis was employed, namely the PCR-ELISA. It was suggested by Goodison *et al* (1997) that newer more rapid methods for detection could supersede the exon link assay in its present format. The PCR-ELISA is a simple and rapid assay (in comparison with the exon link) thought to show comparable sensitivity to the exon-link method.

The PCR-ELISA was first described by Hølmstrom *et al* (1993), and involved the labelling of a PCR product with Digoxigenin (Dig) by incorporation of labelled dUTP into the new strands. At the same time the PCR product was also labelled with biotin through the use of a biotinylated primer. Biotinylated PCR products were attached to a streptavidin coated microtitre plate, followed by the addition of an alkaline phosphatase labelled anti-Dig-antibody. The final result could be either colourimetric, fluorescent or luminescent in nature. A number of variations of this basic method have been recorded eg labelling of PCR product with biotin, and detection through the use of a fluorescein labelled probe; but the principle remains the same.

The assay used for this study was developed by Boehringer Mannheim (now Roche Biochemicals) and marketed in early 1996. The kit is based on a Dig labelling and detection system, similar to that described above. A PCR reaction is carried out during which a Dig label is incorporated into all PCR product via a labelled nucleotide. A sequence-specific probe, end labelled with biotin (customised to the study) is hybridised to this product, and the resulting mixture added to a microtitre plate coated with streptavidin, to which the biotinylated probe binds. An anti-dig-

HRP-linked antibody is then added to each well, and the final step is a colorimetric reaction using an HRP substrate (ABTS). The result is gained through spectrophotometry.

This method is rapid compared to Southern blotting, and is thought to be 100 times more sensitive than agarose gel analysis (Roche Biochemicals technical literature), and comparable to that of Southern blotting (unpublished finding, Churchman and Woodman)

The study was undertaken in a series of stages, these being :

- Investigation of exon linkage within the *in vitro* model system
- Investigation of exon linkage within patient samples (tumour tissue only)
- Design of probe for PCR-ELISA
- Use of the PCR-ELISA to determine specific exon junction expression within tumour derived and non-involved tissues
- Analysis of exon junction expression within normal exfoliated urothelial cells by PCR-ELISA to determine tumour specificity.

3.2 Materials and Methods

3.2.1 Sample procurement

Samples of both tumour-derived and non-involved tissue were collected with the help of the Urology department at Northampton General Hospital (NGH). Ethical approval had been gained for the project from the local ethics committee at NGH and all patients were asked to sign consent forms before tissues were taken.

The determination of tumour specificity was undertaken using exfoliated urothelial cells. These were obtained from voided urine samples donated by of a group of normal healthy volunteers from within the Institute of Bioscience and Technology, Cranfield University. This group consisted of 22 people over the age of 30, with an equal mix of males and females. All volunteers signed consent forms in accordance with the ethical approval from NGH.

Copies of ethical approval and both patient and volunteer consent forms can be found in appendix A.

3.2.2 Sample preparation for mRNA extraction

Extraction from tissue samples

For tissue sample extraction, two methods were used. Where samples were large enough, the tissue was embedded in OCT embedding fluid (Bright Instruments, Cambridge, UK) and sectioned using a cryostat to aid homogenisation of the sample for extraction. Typically 80 – 100 70µm sections were taken from each tissue (full details of this procedure can be found in appendix C) Samples which were too small to section (eg biopsies) were immersed in lysis buffer and homogenised using a scalpel and 18g needle and syringe before the preliminary incubation step.

Recovery of exfoliated urothelial cells from Urine

Urine was centrifuged at 554 xg for 15 minutes to pellet urothelial cells within the sample. The pellet was resuspended in cold sterile PBS containing protease inhibitor (1ml) and then transferred to a 2ml γ-irradiated tube and centrifuged at 12073 xg for 30 minutes at 4°C to re-pellet cells.

3.2.3 mRNA extraction

mRNA was extracted using the mRNA micro-fastrack kit (Invitrogen) as detailed in chapter 2 (section 2.2.2).

3.2.4 cDNA production

cDNA was produced using the same method detailed in chapter 2 (section 2.2.3). For RT112 and T24 cell lines 2 µl of mRNA was used, but this was increased to 3µl in the case of tissue and exfoliated cells.

3.2.5 Exon link assay

This was divided into two sections of work – exon link PCR and agarose gel analysis, and southern blotting.

Exon link PCR

The exon link PCR was carried out using the PCR mastermix shown below. mRNA extraction and cDNA production were carried out as previously described.

The PCR mastermix was as follows :

10x buffer (with MgCl ₂)	10 µl
DNTPs	4 µl
P1 primer	1 µl
2 nd primer	1 µl
Nuclease free water	79µl
cDNA	4µl

The 2nd primers used are as follows : P4, E2, 5/7, 5/8, 5/9, 5/10, 5/11, 5/12, 5/13 and 5/14. Primer sequences can be found in appendix D.

The reaction cycle used was the same as the standard CD44 PCR cycle detailed in chapter 2. 1 µl *Taq* polymerase was added after the hot start. Samples were analysed by electrophoresis using a 1.2 % agarose gel and subsequent staining with ethidium bromide.

Southern Blotting

All exon link assays were blotted to ensure amplification of CD44. Southern blotting was carried out using the same method as described in chapter 2, using a standard CD44 probe.

3.2.6 PCR ELISA

The PCR ELISA was carried out according to manufacturers guidelines. There are two steps to this protocol ; DIG label PCR and DIG detection PCR ELISA.

DIG labelling PCR

mRNA extraction and cDNA production was carried out as previously described.

The PCR mastermix was as follows:

10 x buffer	10 µl
MgCl ₂	5 µl
DIG label mix	10 µl
P1 primer	1 µl
P4 primer	1 µl
cDNA	2-5 µl
Nuclease free water	70.5 – 67.5 µl

The amount of cDNA varied according to whether the origin of the mRNA was tissue, exfoliated cells or cell culture, with cell culture samples requiring the least amount.

The PCR reaction was carried out using the standard CD44 PCR cycle as detailed in chapter 2. 0.5 ul of *Taq* Polymerase was added after the hot start.

PCR products were stored at 4°C until PCR ELISA analysis.

DIG detection ELISA

All kit reagents including microtitre strips were brought to room temperature before use. Solutions were then prepared as follows :

Hybridisation buffer	1 µl of 100pmol probe per ml of buffer
Anti-Dig antibody	1 volume of antibody in 99 volumes of conjugate buffer
ABTS solution	1 tablet in 5 ml of substrate buffer
Wash solution	1 tablet in 2 litres of RO H ₂ O

Samples were first denatured by adding between 5 and 10 µl of sample to 20 µl of denaturation solution. This was then spun briefly and incubated for 10 mins at room temperature. Hybridisation buffer up to a volume of 250 µl was added to the denatured samples and 200µl of this hybridisation mix was added to the microtitre

strip. All samples were carried out in duplicate, and each strip contained a negative control (containing all solutions except PCR product) and two positive controls (RT112 and T24 PCR products). Microtitre strips were covered in film to prevent evaporation then incubated at 37°C in a shaking incubator for 3 hours.

After three hours all solutions were removed from the strips by aspiration and the wells washed using 250µl of wash buffer, again removing the wash solution by aspiration. Strips were washed five times and after the final wash, they were inverted and tapped dry. 200ul of Anti-DIG antibody was then added to each well and the strips again incubated at 37°C, in a shaking incubator, for 30 minutes. Wells were washed five times again as described, and finally 200 µl of ABTS solution was added to each well. This was once again incubated at 37°C, in a shaking incubator, for 30 minutes, but the plate was wrapped in foil to allow the colorimetric reaction to develop in the dark.

The results of the assay were gained by reading the absorbance of the wells using a plate reader set at 405nm.

3.3 Results

3.3.1 Exon Link assay on RT112 cell line

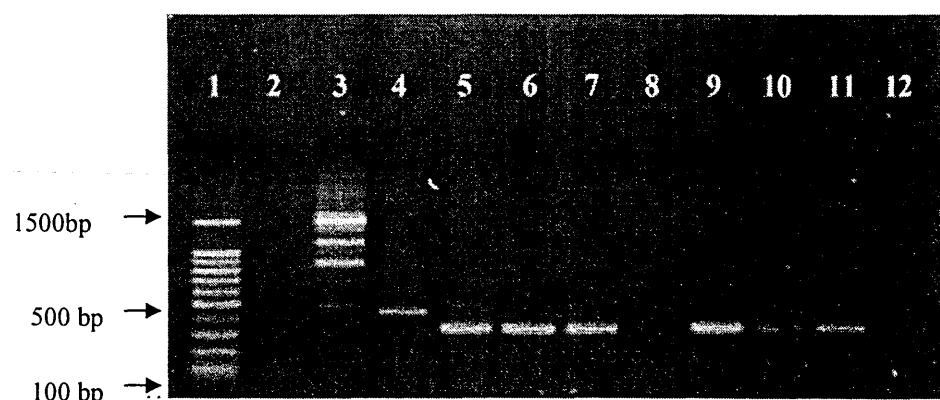


Figure 3.2 – Result of the RT112 exon link assay on an ethidium bromide stained gel. Lanes numbered 1-12 from left to right.

Figure 3.2 shows the result of an RT112 exon link assay. Lane 1 contains a DNA ladder (Promega 100bp ladder, 100 – 1500 bp in size). Lane 2 is a negative control,

whilst lane 3 shows a standard CD44 amplification, using primers P1 and P4, used here as a positive control. This is illustrated as the presence of a number of bands, the lowest of which is approximately 480 bp (CD44s), though this band is quite faint. The band in lane 4 is an internal positive control for this assay, and should only be present if the 5/7 junction, (lane 5) is also present. Lanes 5 – 12 indicate the presence or absence of exon junctions between 5/7, 5/8, 5/9, 5/10, 5/11, 5/12, 5/13 and 5/14.

It can be seen from the figure above that the RT112 cell line expresses exon junctions between standard exon 5 and variant exons 7,8,9,11,12 and 13. This is demonstrated by the presence of a band of approximately 348 bp in each relevant lane.

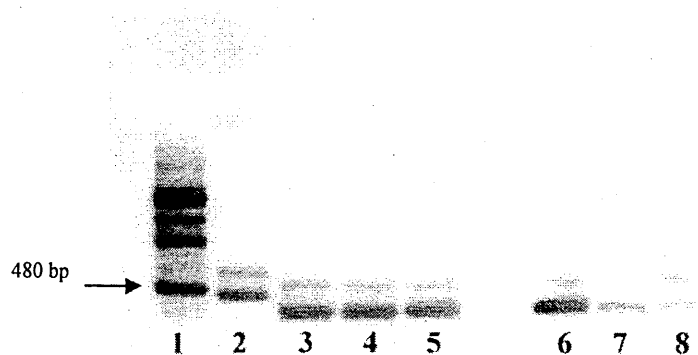


Figure 3.3– Southern Blot of the ethidium gel in figure 3.2. Lanes where bands are present have been highlighted, by numbering 1-8 from left to right.

A Southern blot of figure 3.2 is shown in figure 3.3 All exon junction bands that were considered to be positive results (those that were approximately 348 bp in size), are present on this blot. There are however double bands in some lanes, with the top band being above the expected 348 bp size. These are not expected, and the reason for their presence is unclear.

3.3.2 Exon link assay on T24 cell line

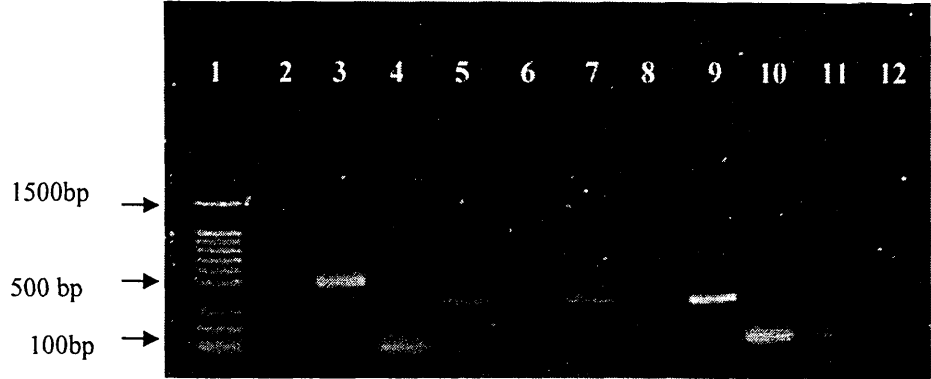


Figure 3.4 – Result of the T24 exon link assay on an ethidium bromide stained gel. Lanes numbered 1-12 from left to right

Figure 3.4 shows the results of the exon link assay from the T24 cell line. Again Lane 1 is a DNA ladder (Promega, as above), with lanes 2 and 3 being negative and positive controls. The positive control is a P1-P4 amplification, and in this cell line is demonstrated by the presence of one band of approximately 480 bp (CD44s). Lane 4 is the internal positive control, whilst lanes 5 – 12 are indicative of the presence or absence of exon junctions 5/7 – 5/14.

Through the use of agarose gel analysis, three distinct bands are seen in lanes 5,7, and 9 which relate to exon junctions 5/7, 5/9 and 5/11.

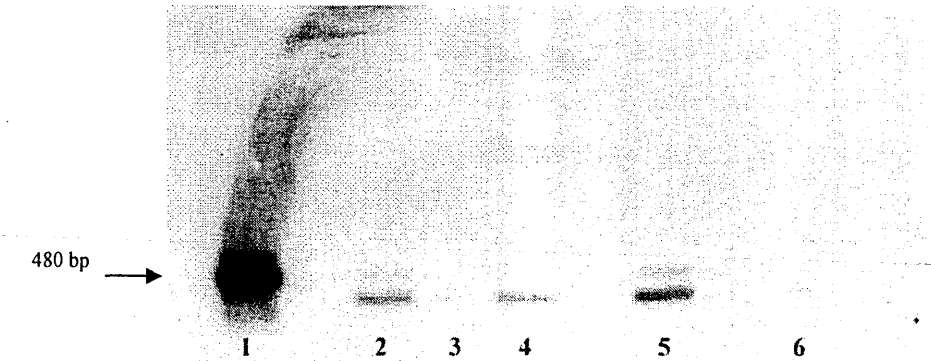


Figure 3.5 – Southern blot of the ethidium gel in figure 3.4. Lanes where bands are present have been highlighted, by numbering 1-6 from left to right.

When the above gel is blotted, a number of differences are seen. Firstly, another band is noted in the P1-P4 (positive control) amplification (lane 1), above that of the band seen on the agarose gel above, suggesting the presence of variant forms, though lesser in number than in RT112. The three junctions noted in figure 3.4 are still present in

lanes 2, 4 and 5, but two extra bands, though not very distinct, are noted in lanes 3 and 6. These correlate to exon junctions 5/8 (lane 3) and 5/13 (lane 6). Of all the junctions seen, 5/11 appears to be the most distinct.

3.3.3 Exon link assay on tissue samples

Exon link analysis of tissue samples was only undertaken in a small group of samples at the early stages of the study. The need for Southern Blotting was highlighted by these samples, as most showed no result with agarose gel analysis, but when blotted, some bands were seen (data not presented).

3.3.4 Choice of PCR-ELISA probe

From the cell line and early tissue exon link results, four exon junctions were initially chosen as possible markers for CD44. These were 5/8, 5/11, 5/12 and 5/13. The presence of exon junctions 5/8, 5/11 and 5/13 were demonstrated within this study. Junction 5/12 was chosen as previous work (unpublished, personal communication, Churchman and Woodman) had suggested this was a promising marker for use with PCR-ELISA.

Initial results with 5/8 and 5/13 did not appear promising, therefore for the remainder of the study 5/11 and 5/12 were employed. A standard probe was also included in the study, based on the P1 primer, detecting the presence of exon 3.

3.3.5 PCR-ELISA results

All tumour, non-involved tissues and normal exfoliated urothelial cells underwent PCR-ELISA investigation. The sample sizes were small, with a total of 14 tumour tissues, 6 non-involved tissues and 22 normal samples.

For the purpose of this study the PCR-ELISA was used qualitatively. For use in this manner, a cut-off point was required to establish whether results could be considered positive or negative. Investigation of expression of the 5/11 marker within the normal sample group demonstrated the presence of a low level of background false positives. It was decided therefore to take the highest absorbance result gained in a 5/11 assay within the normal group as the defining cut-off point for all assays. This cut off point

was defined at 0.06 absorbance units (the highest result in the normal group being 0.057).

Once this cut off point had been defined, all data was checked and tissues designated either positive or negative for all markers. Before reaching this point some data treatment had been carried out. As all sample assays were carried out in duplicate, the mean result for each sample (including positive and neagive results) was calculated. To remove the effect of background levels of absorbance the mean of the negative control was then subtracted from each sample. As each assay undertaken contained its own negative control, each assay had its own background levels removed.

Table 3.1 shows all results gained using standard, 5/11 and 5/12 assays on both tumour and non-involved tissues. Normal results have not been included, as for the 5/11 assay, all results were taken as negative after definition of the cut-off point. The RT112 and T24 cell lines were used as positive controls for each assay, and absorbance units in excess of 1 were routinely recorded for both cell lines.

Table 3.1 PCR-ELISA results in tumour and non-involved tissue for 5/8, 5/11 and 5/12 probes. Results for each assay are indicated as absorbance units at 405 nm

Tumour	standard	- / +	5/11	- / +	5/12	- / +
T1	0	-	0.182	+	0.066	+
T4	0	-	0.144	+	0.051	-
T10	0	-	0	-	0	-
T13	0.251	+	0.258	+	0.112	+
T15	0.019	-	0	-	0	-
T16	0.232	+	0.609	+	0.38	+
T18	0.255	+	0.019	-	0.03	-
T19	0.074	+	0.255	+	0.166	+
T21	0.034	-	2.129	+	1.42	+
B0	0.009	-	0.158	+	0.134	+
X0	0.236	+	2.697	+	2.116	+
X23	0.01	-	0.079	+	0.052	-
X27	0	-	0.001	-	0	-
X3	0.601	+	0.005	-	0.018	-
	Total Positive	6/14	Total Positive	9/14	Total Positive	7/14

(Table 3.1 continued)

Non-involved	standard	- / +	5/11	- / +	5/12	- / +
T11	0.055	-	0	-	0	-
T12	0.17	+	0	-	0	-
T14	0.123	+	0.447	+	0.206	+
T17	0.027	-	0	-	0	-
T20	0.531	+	0.087	+	0.099	+
T22	0.007	-	0.001	-	0.004	-
	Total Positive	3/6	Total Positive	2/6	Total Positive	2/6

Table 3.2 Summary of results from table 3.1, along with the results from the normal group. Expression has been shown as both a percentage and number of samples in each group.

Marker	Non-involved	Tumour	Normal
Std	50% [3/6]	42% [6/14]	9% [2/22]
5/11	33% [2/6]	64% [9/14]	0% [0/22]
5/12	33% [2/6]	50% [7/14]	---

Statistical analysis of the data was deemed irrelevant due to the small sample sizes involved.

The first observation that can be made from table 3.2 is that a percentage of both tumour derived and normal tissues gave a positive result with all three probes, though exact numbers differed between groups. All tissues displayed standard form expression, though the non-involved tissues showed the highest level (50% of tissues). It could be expected that all tissues that expressed 5/11 and 5/12 would give a positive result for the standard assay, but as can be seen from the table, this is not the case. Of the normals tested only 9% (2/22) gave a positive result for the standard assay.

Of the two exon junctions investigated, 5/11 showed the highest expression in tumour tissues, with 64% of tissues expressing this junction, compared to 50% of tissues expressing the 5/12 junction. Though both these markers were found to be present in

the non-involved tissues, the levels in tumour tissues were found to be higher, with nearly double the amount of tumour tissues expressing 5/11.

Looking more closely at the absorbance measurements for each tissue type, it can be seen that there is variability within the levels of junction expression in each tissue. For example, in the tumour tissues tested for 5/11, the absorbance ranges from 0.019 – 2.637 units above the cut off point, though the majority of tissues in the range up to 0.3 absorbance units.

From these results, the exon junction chosen as a marker is 5/11, as this was expressed in the most tumour-derived tissues.

3.4 Discussion

From the initial results with the *in vitro* model (RT112 and T24 cell lines) it can be seen that a number of variant exon junctions are expressed, suggesting that a characteristic splice pattern can be demonstrated within bladder cancer. In the RT112 cell line seven out of eight possible exon 5/variant exon junctions were found to be present, namely 5/7, 5/8, 5/9, 5/11, 5/12, 5/13 and 5/14. The T24 cell line appeared to express three variant junctions when analysing the result of the exon link assay using agarose gel analysis. Upon Southern blotting however, the presence of one further junctions was noted; the full expression profile therefore being 5/7, 5/9, 5/11 and 5/13. The number of exon junctions present within each cell line can be compared to the characteristic banding noted when carrying out a standard P1-P4 amplification across the whole gene. RT112 shows multiple banding with varied product size, therefore the presence of a number of exon junctions could be expected. Conversely, the T24 cell line demonstrates only one major band, until southern blotting is undertaken where a further band is seen. This band is not much greater in size than the standard band, therefore the presence of a lower number of variant exons could again be expected. The banding patterns seen in this standard amplification can therefore be attributed directly to the presence of variant exons.

The exon link assay has a distinct advantage over previous studies on exon expression as the design of the assay allows for a characteristic pattern of exon expression to be

established. The pattern produced is simple to identify and can be easily exploited for diagnostic purposes. Previous studies focusing upon exon expression have analysed the presence of the exon alone, with no thought of the splicing, whether characteristic or otherwise which may be occurring. The initial exon link assay by Goodison *et al* (1997) on colorectal tissue demonstrated that both normal and tumour tissues exhibit low levels of all exons linked with exon 5. It is known from work by Bell *et al* (1999) that certain variant exons are known to be spliced in “blocks”, and exons rarely appear to be spliced alone in the mRNA transcript. If it is the case that exons splice in full groups, ie the presence of exon 8 may also result in the presence of 9, 10, 11; then a simple analysis of exon presence may not give rise to any form of distinguishable pattern. Results may simply demonstrate a similar level of expression for each exon due to the block splicing present. Analysis of exon junctions allows for detection of a predominant pattern which may be tumour specific and / or tissue specific. As can be seen in both colorectal and bladder tissue a predominant pattern was noted and when this was compared to normal tissue a tumour specificity was found.

In bladder tissue the predominant junctions were found to be 5/8, 5/11 and 5/13 therefore sequences analogous to these were used as probes within the PCR-ELISA assay. The junction 5/12 was included as exon 12 had previously been demonstrated to be a potentially useful marker for bladder cancer. Exon 12 is known to be expressed within epithelial tissues, though it is unknown how expression of this exon changes during tumourigenesis. Initial results with 5/8 and 5/13 were not promising (data not presented) and as a result these were omitted from the study. The final chosen junction is 5/11 as this was the junction most commonly found in tumour samples. It is interesting that this junction predominates in both bladder and colorectal cancer especially as exon 11 (V6) has been implicated in the formation of metastases (Günthert, 2001).

The increase and overexpression of these splice variants is often described as aberrant or disorderly, suggesting diminished function or even complete loss of the mechanisms responsible for regulation of alternative splicing. These ideas would complement the changes seen in other regulatory mechanisms normally operational within cells when tumourigenesis begins. What is often not considered is that the factors responsible for the regulatory mechanism itself could change, ie variations

within the protein population responsible for the regulation. An investigation of the SR family of splicing factors in mammary tumourigenesis demonstrated that changes in these factors could be detected when this aberrant splicing is seen.

The SR factors are a family of proteins known to influence alternative splicing when changes in their relative concentrations occur. Stickeler *et al* (1999) specifically investigated the changes that occur when considering the alternative splicing of CD44. They noted that though general changes occurred in SR expression there was a predominant protein present in the tumour tissue, which was not seen in the pre-tumour tissue. Looking more closely at the CD44 splicing they found that the changes in splicing were due to a change in the sub-set of splicing factors required for CD44 exon recognition. No change in the splicing machinery itself was noted.

The change in regulation could be a consequence of metastatic change, or could be an influencing factor in this. The predominance of certain variant exon junctions over others suggests that there may be a specific reason for their presence. This is highlighted by the changes which occur to the protein as a result of variant exon peptide addition. As discussed in chapter 2, variant exon addition can result in increased glycosylation and heparan sulphate addition, leading to possible conformational change and hence a change in binding abilities. The ability to modulate CD44 binding would be of advantage to the tumour cells as their metastatic potential increases, or indeed to increase this potential.

The exact factors or mechanisms which influence these changes are not known, and much further work is required to fully understand alternative splicing and its role in tumourigenesis, both in CD44 and other proteins.

The use of the PCR-ELISA within this study did provide a simpler more rapid method for the analysis of exon junction expression, though initially the exon link assay was required. Though the initial aim was to undertake both assays in parallel, a move towards the use of the PCR-ELISA as the main assay was made, prompted by technical difficulties in extraction of sufficient amounts of mRNA from tissue samples. These difficulties were due to the decreasing size of samples procured, leading to an inability to section samples for extraction. As an alternative to

sectioning, samples were immersed in lysis buffer and homogenised through the use of a scalpel and needle. This homogenisation was difficult to perform and in most cases was incomplete. Unlysed tissue was therefore still present after incubation in the lysis buffer, meaning not all mRNA may have been extracted. Another factor thought to have influenced these extraction problems is the difference in procurement practices between staff at NGH. It was found that a number of samples were not being snap frozen immediately after resection, but were kept at room temperature for periods of up to 2 hours. As temperatures within surgical theatres are generally higher than normal room temperature this may have led to a degree of mRNA degradation. Carrying out both the exon link assay and the PCR-ELISA requires a number of PCR reactions to be carried out. Therefore, due to the increased amounts of RNA required for cDNA production, not enough sample was present for carrying out both assays. It was decided to undertake PCR-ELISA alone, but testing only for those junctions which appeared to be promising in the early part of the study. Use of this assay alone also decreased the turnaround time it can typically be done in one day, if the PCR reaction is carried out on the previous day, and a number of samples can be analysed at once.

Interestingly a confounding result has arisen from the PCR-ELISA study. This is the lack of "standard" in many of the tissue cases investigated, both tumour and non-involved. This was further highlighted when considering the results gained with the RT112 and T24 cell lines for both standard and variant junction. Though these results are not presented, it is interesting to point out that the RT112 cell lines gave a lower absorbance for the standard probe than that noted for a variant probe. In T24 this was reversed, with the standard giving the highest result, and the variant the lowest. It was also seen that the RT112 result was often half the absorbance of the T24. Analysis of CD44 expression in RT112 cell lines demonstrates the presence of a variety of CD44 transcripts. Not all would be expected to express the particular variant junctions under analysis, but due to the nature of the standard probe (detection of portions of exons 3 and 4) it would be expected that all isoforms present would be detected. This would result in an absorbance level at least equal, but presumably higher than that found with a variant probe. An explanation of this standard probe result is not simple, and clarification would require much further work. One possibility is that there is a link to washing stringency due to folding differences within the DNA whilst in

solution. It is suggested that DNA does not remain linear when in solution, and may fold and form complex structures. Due to the difference in areas to which the probes hybridise, the structure near the standard probe hybridisation site may differ to that surrounding the variant. This difference in folding could lead to a weaker hybridisation of standard probe, which when washing steps are introduced, results in a loss of captured DNA.

A lack of standard exon presence was also noted in the normal exfoliated urothelial samples, but this is expected. In normal urothelium, CD44 is predominantly found in the layers closest to the basement membrane, with CD44 expression being lost as cells move outwards towards the top layer of urothelium. As the majority of cells present within the urine are exfoliated from the outer layer then it would not be expected that any standard CD44 would be in these samples.

4 Production and characterisation of a polyclonal antibody against a CD44 transitional epitope

4.1 Introduction

The aim of the project overall, was to produce a protein-based assay to assess the presence of CD44 variants within clinical samples. The information gained thus far regarding specific variant junction expression was at an mRNA level, therefore to produce a protein-based assay, there was a need to translate this information into a format which could be used for antibody production.

Antibody production is either polyclonal or monoclonal in nature depending on the production technique. Polyclonal antibodies are produced *in vivo* in a variety of different animals including rabbit, sheep and horse. The chosen antigen is injected intramuscularly into the animal and its immune system produces the required antibody. To increase the titre of the antibody, the animal producing it is given a number of booster immunisations before terminal bleeding and purification. Monoclonal antibodies are produced *in vitro* through the fusion of immunised mouse spleen cells and myeloma cells to produce hybridomas. Immunisation of the spleen cells stimulates antibody formation whilst fusion with the myeloma cells allows antibody production to be constant during cell culture. Polyclonal antibody production results in a heterogenous mix of antibodies directed against a number of epitopes, though affinity purification can be applied to the final quantity of animal serum to reduce the variety of antibodies present. Monoclonal antibodies are a homogenous selection, as hybridoma clones can be selectively grown to produce a single high affinity antibody against the chosen antigen. Both techniques differ in complexity and cost, and each form of antibody has their advantages and disadvantages, which will be discussed in section 4.4.

For this study a polyclonal antibody was chosen as it was decided that this would give a good indication as to the potential of the protein marker, with the intention of undertaking monoclonal antibody production in the future.

To produce the specific antibody required it was decided to design and produce a synthetic peptide antigen. Purification of a protein containing the 5/11 epitope would not be easy and likely result in the production of a too diverse variety of antibodies. The peptide was based on the translated sequence of the 5/11 primer and its design is shown in section 4.2. The peptide was produced through the use of a solid phase synthesis method described as fmoc chemistry, where peptide chains are built by adding one amino acid at a time onto a solid resin phase. As the peptide is the specific epitope required for antibody production, the degree of variety within the serum should be lower than if the whole protein was used.

After affinity purification of the antibody an initial assessment of its sensitivity and specificity was undertaken through the use of an antibody : antigen ELISA titration assay.

4.2 Materials and Methods

The synthetic peptide and polyclonal antibody were produced by Immune Systems Ltd (Paignton, Devon). Due to confidentiality reasons it was not possible to gain access to the exact protocols used, therefore the protocols that follow are based on current standard protocols.

4.2.1 Peptide Design

The starting point for the design of this peptide was the 5/11 primer used within the exon link assay. The primer is 24 bases long, which would translate to a peptide of 8 amino acids. It was decided that this may not produce a peptide specific enough for antibody production, therefore the primer sequence was located within the genomic sequence (Screaton *et al*, 1992) and the primer extended outwards by 12 bases at the 5' end and 6 bases at the 3' end. The resulting sequence was transcribed to an mRNA sequence, then finally to an amino acid sequence, based on the reading frame noted within the genomic sequence. The final product was a peptide of 15 amino acids. Figure 4.1 overleaf demonstrates the design from primer to peptide.

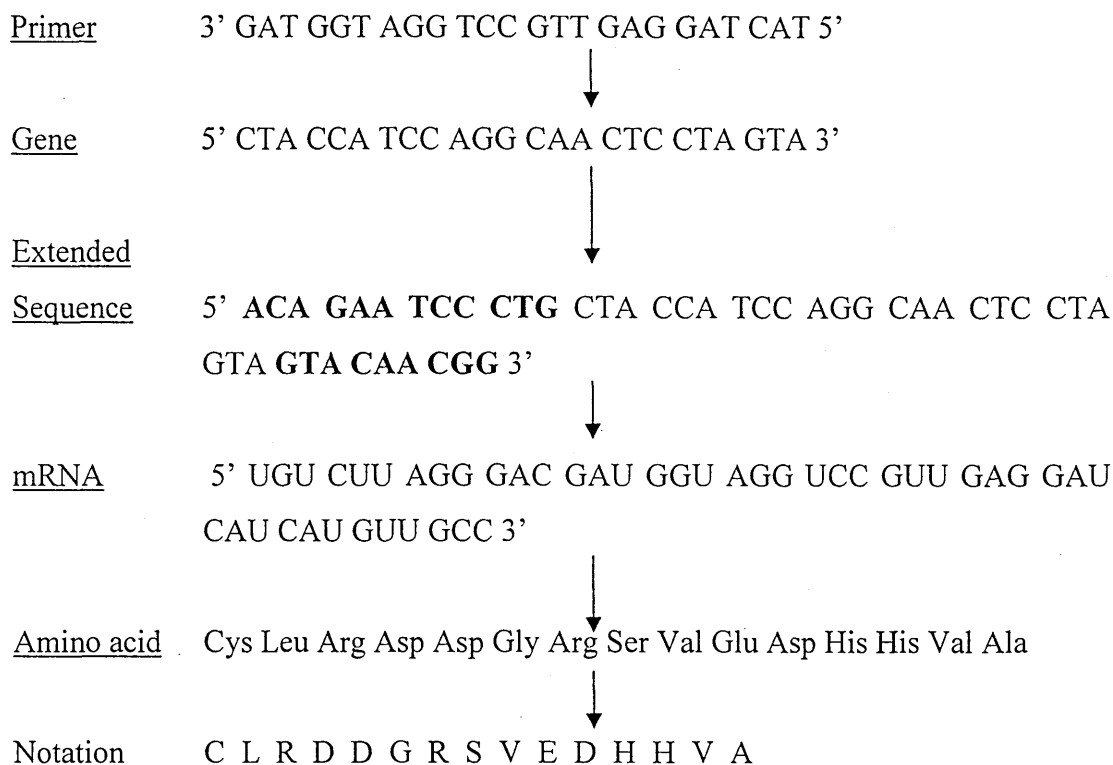


Figure 4.1 – Flow diagram showing design from primer sequence to peptide sequence.

4.2.2 Search for peptide homology

The peptide was checked for homology to other human proteins using a protein BLAST search. The BLAST search facility is a web based search engine hosted by the US National Center for Biotechnology Information (NCBI) which can be found at www.ncbi.nlm.nih.gov/BLAST. Upon entering the protein search engine (BLASTp), the protein sequence is typed into the search box and the relevant database selected, in this case swissprot. The “BLAST” button is pressed, and the results are presented in approximately 1-2 minutes. The data is presented as percentage homology to sequences within the database, starting with the highest percentage homology. The data can be used to predict the possible cross-reactivity of the antibody produced using the designed peptide to other cellular proteins.

4.2.3 Peptide Synthesis

This was carried out using a solid phase synthesis technique, described as fmoc chemistry. Solid phase synthesis involves the production of the peptide attached to a solid phase support, in this case PEG-polystyrene. The first amino acid within the peptide sequence attaches to this support via a linker attached to its carboxyl terminal, with subsequent amino acids joining to the amino terminal of the preceding one.

Fmoc chemistry is so called, as the amino terminal of each amino acid is protected by a 9-fluorenylmethoxycarbonyl (fmoc) group, which is cleaved to allow the addition of another amino acid, forming a peptide. Fmoc groups are removed under different conditions from those required for cleavage of the solid phase linker, so that no loss of peptide through loss of solid phase support occurs during amino terminal de-protection (Fields and Noble, 1990).

The carboxyl terminal of the amino acid is attached to the PEG polystyrene support via its carboxyl group and its amino fmoc protecting group is removed. An excess of the second amino acid is then added and its carboxy group activated so it can bond to the amino terminal of the first one. Excess reagents are removed by washing, then the next amino acid is added. This continues until the full peptide sequence has been produced. Due to the presence of side chains on some amino acids, which could cross react during peptide production, many are masked with protecting groups. The final step in sequence production is therefore removal of these side-chain protecting groups and cleavage of the peptide from its solid support. To ensure stability of the peptide and to decrease the risk of cross reactivity of side chains at this point, side chain protecting groups and resin linkage groups are chosen so that they are all removed under the same conditions (Fields and Noble, 1990).

The peptide is then isolated by precipitation using cold diethyl ether and purified using HPLC. The purity of the resulting peptide is then assessed by mass spectrometry, where the expected molecular weight of the peptide (in Daltons) is compared to the actual molecular weight. For antibody production, whether polyclonal or monoclonal, a purity of 70% or above is required.

4.2.4 Polyclonal antibody production

Antibody production was carried out using a female New Zealand white rabbit.

Due to the very small size of the peptide, it was bound to a large carrier protein, keyhole limpet haemocyanin (KLH), to ensure the immune response within the animal would be fully elicited. As injection of the antigen alone may not produce a

large enough immune response, therefore antigen is mixed with an adjuvant, a liquid mixture which heightens the response. For rabbits, the most common adjuvant in use is Freund's complete adjuvant (FCA), which is known to specifically elicit a large IgG response. FCA consists of mineral oil (Bayol F), a suspension of heat inactivated mycobacterium (either *M. butyricum* or *M. tuberculosis*) and an emulsifier (Arlacel A). This adjuvant is now available in a non-ulcerative format (non-ulcerative Freund's adjuvant or NUFA) which contains an attenuated strain of mycobacterium and reduces the painful ulceration seen in many immunised animals with traditional FCA. It is this NUFA form that was utilised in the production of the 5/11 polyclonal antibody. The antigen (peptide-KLH) is added to the NUFA and emulsified, producing antigen-water microdroplets within the oil layer. The NUFA-antigen mix is then injected intramuscularly and a total of four immunisations with this mix are carried out at two week intervals. After the 3rd and 4th immunisations, test bleeds are carried out after 7 days to assess the strength of the immune response in terms of the ability of the animal sera to detect the presence of free and bound peptide. If after the 4th immunisation a satisfactory level has been reached, a final booster immunisation is given and after 7 days a terminal bleed of the animal is undertaken.

Due to the heterogeneity of the anti-sera produced through polyclonal techniques, the serum was affinity purified, using a column containing the immobilised peptide. The final antibody solution was eluted in PBS with 0.05% sodium azide.

4.2.5 Antibody-antigen titration assay

To assess the specificity and sensitivity of the purified antibody, an ELISA titration assay was carried out using serial dilutions of free peptide, peptide bound to KLH (used to elicit the immune response) and peptide bound to BSA, against three concentrations of antibody.

Antigen concentrations ranged from 100µg to 1 ng (100µg, 10µg, 1µg, 100ng, 10ng, 1ng) and were produced by serial dilution in a carbonate-bicarbonate buffer (pH 9.6). Antibody dilutions ranged from 1/500 to 1/2000 (1/500, 1/1000, 1/2000) and were diluted in blocking agent (Pierce superblock in PBS). A negative control (no antigen, no antibody) was carried out for each antigen assay.

The ELISA was designed as an antigen:antibody assay, where the antigen was bound to the plate, followed by primary antibody incubation and the addition of the detection system using an HRP-linked secondary antibody and ABTS solution.

The peptide was bound to the ELISA plate (Maxisorp – Nalge Nunc, Rochester NY, USA) by aliquoting 100µl of the relevant antigen concentration per well and incubating overnight at 4°C. The following day the peptide solution was removed and the plate blocked by the addition of 200µl of blocking solution (Pierce Superblock in PBS, Pierce Biotechnology inc, Cheshire, UK) per well. This was immediately removed by inversion of the plate and the process repeated twice. Primary antibody, again of the correct dilution was then added (200µl / well) and the plate incubated for 1 hour at room temperature. After removal of the antibody solution, the plate was washed three times using a PBS / Tween wash solution (PBS pH 7.4, 0.05% Tween 20) before addition of the secondary antibody (Anti-rabbit IgG-HRP linked, Sigma – 200µl / well) at a concentration of 1/2000 in blocking buffer. Again this was incubated for 1 hour at room temperature. Following removal of the secondary antibody solution the plate was washed three times as before and ABTS solution added (200µl / well). The plate was then incubated for 30 minutes at room temperature. The final result was gained through reading the absorbance of the plate spectrophotometrically at 405nm. Results were placed into an Excel spreadsheet and the plotted as a scatter graph.

4.3 Results

Upon a protein BLAST search of the final peptide sequence, using the swissprot database, no significant homology was noted.

Peptide production resulted in a 70% pure peptide which was then used in the production of the polyclonal antibody. The final antibody yielded 2.98g of protein after affinity purification and was presented as 21ml of a 142µg/ml concentrate solution.

The results of the antigen : antibody titration undertaken with the final affinity

purified antibody are demonstrated in figures 4.2 – 4.4.

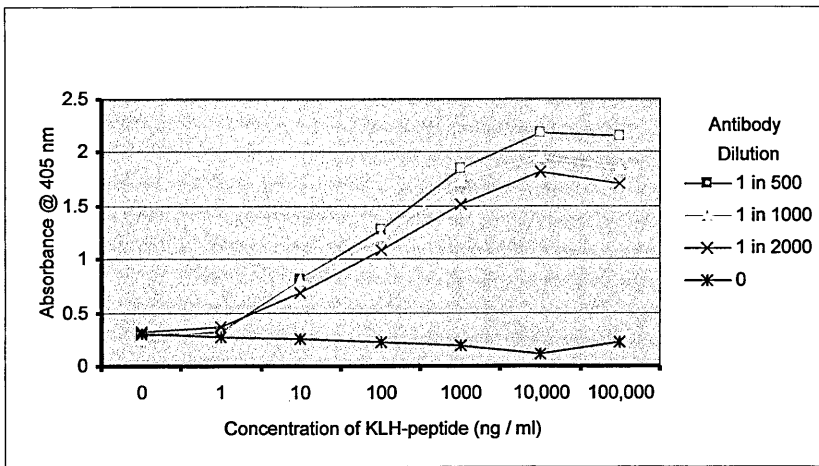


Figure 4.2 – Graph showing titration of 5/11 antibody with KLH-peptide

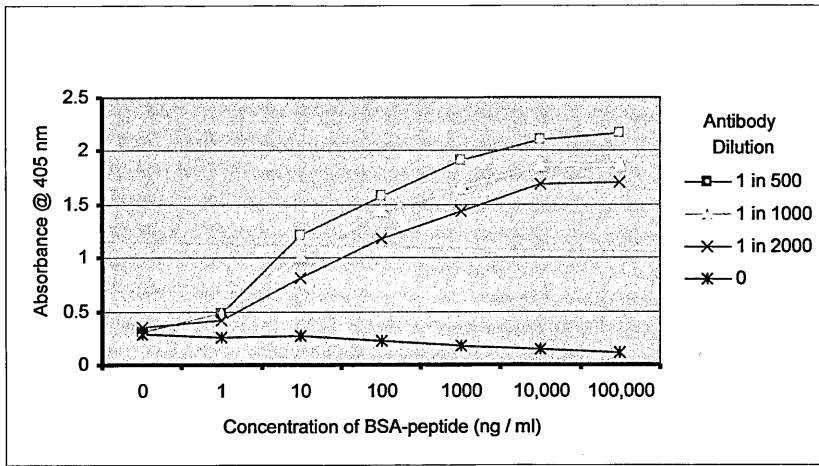


Figure 4.3 – Graph showing titration of 5/11 antibody with BSA-peptide

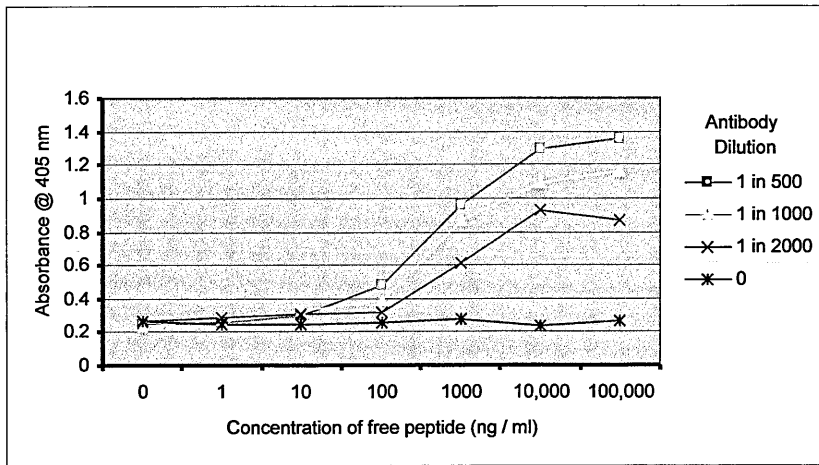


Figure 4.4 – Graph showing titration of 5/11 antibody with free peptide

Figures 4.1.2 – 4.1.4 : Results of the antigen : antibody titration assays carried out with the 5/11 antibody and KLH-peptide conjugate, BSA-peptide conjugate and free peptide. Results are demonstrated as absorbance units at 405 nm against peptide concentration in ng / ml. Each figure demonstrates four plots, which correspond to the antibody dilutions used.

Titration of the antibody with the KLH-peptide demonstrates the classic curve which would be expected for this assay. The curve starts reasonably flat, as there is more antibody than antigen present; but as the antigen concentration increases, so the curve rises sharply. When the concentration of antigen becomes too high the saturation point is reached and the curve flattens out once more and begins to drop slightly. All three antibody concentrations demonstrate a similarly shaped plot, with the curve rising sharply from a 1ng concentration of antigen until it plateaus and decreases slightly at 10 ug.

The BSA-peptide plot is similar to that KLH, though the distinction between the antibody concentrations is clearer with a more obvious difference between the three curves. The curve again rises sharply at 1ng concentration of antigen and begins to plateau at 10 ug, though no decrease of absorbance is noted.

The titration of antibody with free peptide produces the same characteristic plot, though the curve does not start to rise sharply until 10ng of antigen in the 1/5000 and 1/1000 antibody dilution, and until 100ng of antigen in the 1/2000 dilution. At the 10 ug antigen point the curve is still rising for the 1/500 and 1/1000 antibody dilution, though a lesser gradient than before 10 ug. In contrast, the 1/2000 antibody dilution demonstrates a decrease between 10ug and 100ug of antigen, of a similar gradient to the increase seen within the other antibody dilutions. Here the plots of each antibody dilution showing the greatest difference from the other two dilutions.

4.4 Discussion

The finding of no significant homology of the 5/11 peptide with the other proteins suggests that this antibody should be of high specificity. If homology had been noted, then it would have increased the possibility of false positives when used with samples. It was not expected to note any homology with the CD44 protein, as the protein form noted within the databases may be of standard form only therefore exon 11 would not be present. Alternatively, if variant forms are included, they may not contain the 5/11 junction.

The antibody itself demonstrates a good response when titrated against both bound and free antigen. All three assays give similar results in terms of plot shape, but some differences were noted. The differences between the titration curves for KLH-peptide and free peptide may be due to the fact that not all antibody present is directed against the 5/11 peptide alone. Due to the antibody being polyclonal in nature, a number of antibodies would have been produced, against both the peptide and the KLH carrier. Though the antibody has been affinity purified, there will still be a certain percentage of antibody which detects the carrier. This means that there is more antigen present for the antibody to detect, therefore the curve starts to rise more sharply than that with just free peptide alone. Differences in the end of the curve will be due to the same reasons and also as the curve rise begins later in the 5/11 free peptide curve, then it may naturally plateau later.

From this assay, it can be seen that the antibody can detect the 5/11 antigen down to a level of between 10 and 100ng with the optimal antibody concentration being between a 1/1000 and 1/5000 dilution.

Though polyclonal in nature, this does not appear to affect the ability of the antibody to detect antigen too greatly. A monoclonal antibody could be considered to be the more useful antibody in most cases, and from the literature does appear to be the antibody of choice currently. Polyclonal antibodies do have an advantage over monoclonal antibodies in a number of respects. In terms of cost and time required, polyclonal antibodies are the best, but a degree of variation between batches may always occur due to differences between individual animals. If therefore, large amounts of antibody are required then this may be the optimal method. Monoclonal antibodies are time consuming to produce in terms of finding a hybridoma clone which produces an antibody with the required characteristics, but once found, these can be frozen and thawed and subcultured when required.

In terms of antibody affinity, for whole proteins polyclonal antibodies may produce the highest affinity types as they rely on the body's immune system to pick out the best epitopes for antibody production. Though this will happen with monoclonal

antibodies as each separate clone must be screened, the clone which produces the best result may be missed.

Monoclonal antibodies are the most specific as only one antibody type is produced, yet polyclonal antibodies such as 5/11 can be thought of as being “almost” monoclonal. As the immunogen is a short peptide which is also used to affinity purify the antibody solution. This was demonstrated in the titration assay, in comparison of KLH-peptide and free peptide titrations. For the assays which this antibody has been produced (IHC and western blot) polyclonal antibodies have been demonstrated to be suitable, though some variation between protein type exists. No true judgement can be made about the antibody until it has been tested on real samples.

5 Evaluation of 5/11 junction protein in cell lines and clinical samples

5.1 Introduction

The initial project aim was to reach a platform from which a protein based diagnostic assay, for the monitoring of bladder and colorectal cancer, could be produced. The requirement was for a simple assay that could be easily carried out within a district general hospital. Through the work presented in the early chapters, analysis of CD44 expression within bladder tumours by the investigation of variant exon linkage, has provided a potentially useful marker. The variant exon junction found in bladder tumours was the same as that found in colorectal tumours, therefore one polyclonal antibody was produced – 5/11.

The aim of this final chapter was to assess expression patterns of the 5/11 junction protein and its tumour specificity as suggested by the earlier mRNA study. This was carried out using the 5/11 antibody as described in chapter 4.

The initial evaluation was undertaken using the *in vitro* model system (established in Chapter 2) which consists of the RT112, T24, HT29 and CACO2 cell lines. Once the expression patterns within these had been established, the antibody was tested on clinical samples. Expression in bladder tissue was determined through the use of both bladder tissue sections (paraffin wax embedded sections) and urothelial cells in voided urine samples. Tissue samples were tumour origin only, but urine samples were taken from patients with confirmed tumours, patients with no known tumours and a normal healthy control group from within the Institute of Bioscience and Technology (IBST), Cranfield University. For colorectal cancer evaluation a staged study was carried out with paraffin embedded sections on tissue ranging from polyps and Dukes stage A tumours to Dukes stage D. It was not possible to gain any normal or non-involved colorectal tissue for this section of work.

To investigate 5/11 expression immunohistochemistry and western blotting were employed. Immunohistochemistry allowed visualisation of the 5/11 protein within

the cell, giving an indication of expression levels and as to whether the protein was fully formed and able to integrate into the cell membrane, or it was a mis- or poorly-processed protein. These are assessed by its presence both in membrane and cytoplasm, as demonstrated in Chapter 2 with other anti-CD44 antibodies. Western blotting gave an indication of relative levels of the protein expression, how many variant forms expressed this protein junction, and their relative size.

Investigation of expression within cell lines provided an opportunity to optimise the assays in terms of relative antibody concentration before application to the clinical samples.

5.2 Materials and Methods

5.2.1 Sample Collection

All clinical samples were gained from consenting patients at Northampton General Hospital, or consenting adults at IBST, Cranfield University. Copies of relevant consent forms and ethical information can be found in appendix A. An attempt at age and sex matching was undertaken in the bladder study, though this proved to be difficult. To achieve the most comprehensive investigation of normal expression a number of males over the age of 35 and up to 65 were included. It is at the age of 40 that the risk of bladder cancer begins to rise sharply, therefore this allowed for a cross section of samples around the risk age.

5.2.2 Immunohistochemistry

Immunohistochemistry was carried out on all cell lines and on paraffin sections of both colorectal and bladder tissue.

Slide Preparation

Cultured cells

Fixed slides were removed from storage at -20°C and incubated on the bench, still wrapped in foil, for 10 minutes to bring them to room temperature. A grid was then

drawn on each slide using an Immedge pen (Vector Laboratories) to allow a number of reactions to be carried out on one slide at the same time. Of these reactions, a negative grid was always included to check for the extent of background staining.

Paraffin sections

Only one reaction could be carried out per section, though every batch of reactions included a positive and negative control in the form of a slide of the relevant cell line – RT112 and T24 for bladder samples, HT29 and CACO2 for colorectal samples. Due to the tissue sections being fixed in paraffin wax, sections had to undergo pre-treatment before immunohistochemistry could be undertaken. Whilst still coated in wax the sections are hydrophobic, therefore no antibody interaction can occur as this requires the use of solutions. De-waxing through the use of xylene and subsequent re-hydration in a decreasing concentration of ethanol samples produces less hydrophobic sections (see Appendix B for protocol). To ensure all possible antigen-antibody interactions occurred, all antigenic proteins were un-masked through the use of an antigen retrieval method.

Antigen retrieval

Antigen retrieval is often required on paraffin embedded sections due to artefacts caused by the formalin fixation which such tissue undergoes. Formalin fixation results in the formation of methylene bridges which crosslink proteins and could lead to the masking of antigenic sites. Antigen retrieval breaks these methylene bridges, unmasking the antigenic site and allows antibody binding to occur.

Slides were boiled in a sodium citrate buffer (pH 6.0 – see appendix B) for 15 minutes (carried out in a microwave). After boiling (by incubating at power 9 for 9 minutes and power 8 for 5 minutes) slides were left in the hot buffer for 20 minutes. They were then removed and rinsed three times in TBS, at which point they were ready to begin IHC.

Immunohistochemistry

The basic protocol was that detailed in chapter 2.

As the 5/11 antibody is a rabbit polyclonal, the blocking serum for this was normal goat serum diluted 1 in 5 with TBS.

The primary antibody (5/11) was diluted 1/25 with neat normal goat serum and the secondary antibody (biotinylated goat anti-rabbit - Dako) was diluted 1/100 again with neat normal goat serum.

DAB incubation was increased, to a time of approximately 25 minutes for tissue sections.

All other parts of the protocol were unchanged.

Mounting of slides

Cultured cell slides were mounted using a glycerol based mountant (Dako) and stored at 4°C. Paraffin sections required de-hydration (see Appendix C for protocol) before mounting with DPX (Fisher Life Sciences Ltd, Loughborough, UK) and allowed to dry overnight.

5.2.3 Western Blotting

Western Blotting was carried out on all four cell lines and on proteins extracted from exfoliated urothelial cells in naturally voided urine samples.

Protein Extraction

Proteins from cell lines were extracted as detailed in chapter 2.

Urothelial cells were pelleted by centrifugation of urine samples at 450xg for 15 minutes. The supernatant was removed and the cell pellet resuspended in PBS / Protease inhibitor (1ml – ICN Biomedical) solution. This suspension was then placed into a 2ml eppendorf tube and centrifuged at 12073xg for 30 minutes.

The resulting pellet then underwent extraction using CHAPS lysis buffer (see appendix B) as for cultured cells.

Final protein samples were concentrated using a microconcentrator.

Western Blotting

The technique was carried out using the same protocol as detailed in chapter 2.

The 5/11 antibody was used in a dilution of 1/100 with 5% non-fat milk/TBS. The secondary antibody was an HRP-linked goat-anti rabbit (Autogen Bioclear).

5.3 Results

5.3.1 5/11 expression in *in vitro* model

The expression of 5/11 protein within the cell lines of the *in vitro* model was assessed using both immunohistochemistry and western blotting. All cell lines were found to express the protein, though differences were noted. Immunohistochemical analysis demonstrated the presence of the protein in both membrane and cytoplasm, as was noted with other anti-CD44 antibodies in Chapter 2. Staining density differed between all four cell lines, and the results are summarised in table 5.1 below. Plate 5 (page 116) shows the immunohistochemical staining seen with this antibody.

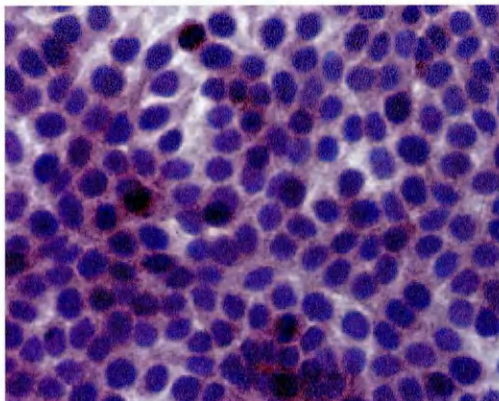
Table 5.1 – Summary of staining density results for cell lines with 5/11 antibody

	Cell Line	Membrane	Cytoplasm
Bladder	RT112	+++	+++
	T24	++	++
Colorectal	HT29	+++	+++
	CACO2	+	+

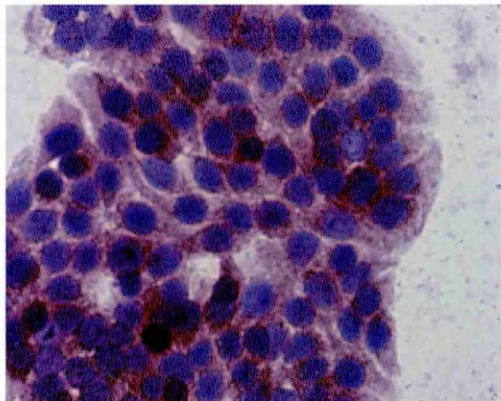
Key : +++ = strong staining, ++ = moderate staining, + = weak staining

RT112 and HT29 showed the strongest staining for the 5/11 protein, with T24 showing moderate staining and CACO2 being the weakest stained. These patterns of expression are very similar to those seen with the Hermes antibody as demonstrated in Chapter 2.

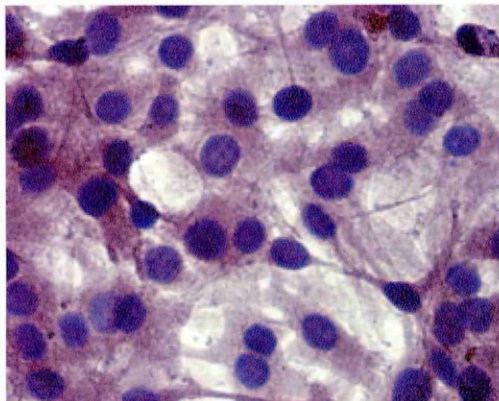
Plate 5 : Immunohistochemical staining of all four cell lines with 5/11 antibody



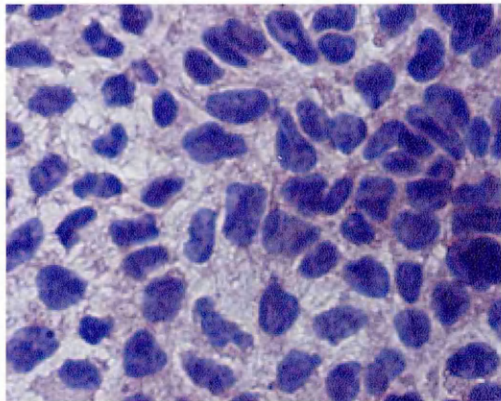
5A – RT112 cells with 1/10 dilution of antibody



5B – HT29 cells with 1/10 dilution of anitbody



5C –T24 cells with 1/10 dilution of antibody



5D – CACO2 cells with 1/10 dilution of antibody

Western blotting demonstrated the presence of a large protein of approximately 130 kDa in size, as demonstrated in figure 5.3 below. The bands produced by this protein are not very dense, though those in the RT112 and T24 cell lines are stronger.

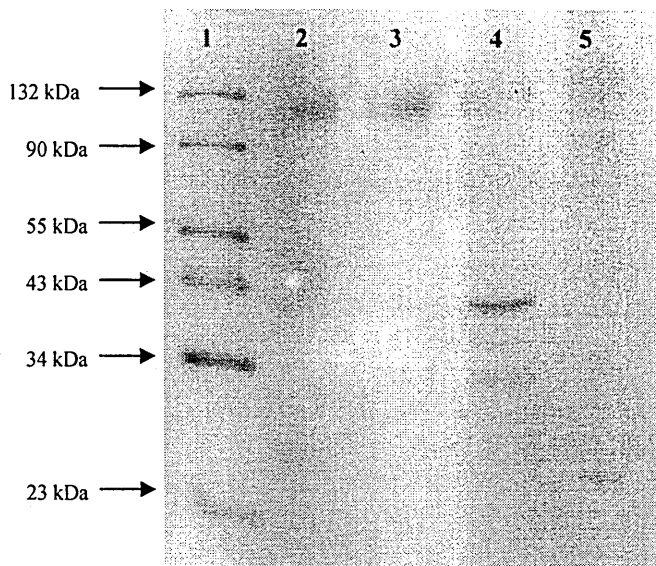


Figure 5.1 – Western blot of cell lines with 5/11 antibody. Lane 1 is a molecular weight marker (Cruz marker, Santa Cruz Biotechnology), whilst Lanes 2,3,4 and 5 are the RT112, T24, HT29 and CACO2 lines respectively.

The presence of further faint bands can be noted in all cell lines, but some, more definite bands can also be seen. In particular there is a band of approximately 34kDa within the T24 cells, three bands between approximately 34 and 43 kDa in HT29 cells, and two bands between 23 and 34kDa within the CACO2 cell line.

5.3.2 5/11 expression in bladder cancer

The expression of 5/11 protein in bladder cancer and its comparison with non-tumour samples was assessed using immunohistochemistry (IHC) and western blotting. Tumour expression was investigated using paraffin embedded sections which underwent immunohistochemical staining with the 5/11 antibody. From earlier immunohistochemical staining with other CD44 antibodies, protein was noted in both membrane and cytoplasm, and a similar pattern of staining was expected here.

Table 5.2 summarises the immunohistochemical results in all bladder sections, whilst plate 6 on page 119 demonstrates some of the staining seen visually.

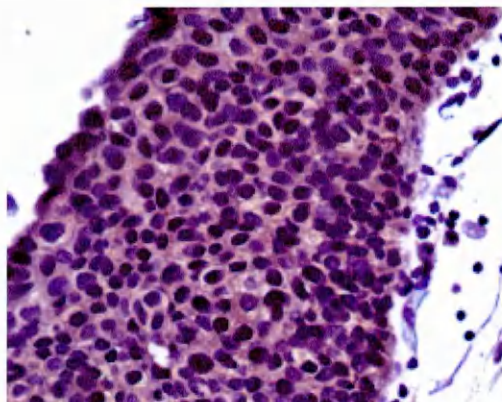
Table 5.2 Summary of the results of the IHC with 5/11 antibody in bladder sections of varying stage and grade. Results are shown as density of staining in both membrane and cytoplasm, percentage distribution across the section and where the staining was noted.

	Sex	Stage/Grade	Stain (M)	Stain (C)	Distribution	Area
A	M	1G in situ	+++ / +	+++ / +	C	mid to top of urothelium + substrata
B	F	1A	+	++	C	FTU
C	F	1A	++ / +++	++	D	FTU
D	F	1A	++ / +++	++	B	FTU
E	M	1A	+++	++	D	FTU
F	M	1A	+++	++	D	FTU
G	M	1A	+	+	D	FTU + substrata
H	F	1A	++	+ / ++	D	FTU
I	M	1A	+ / ++	+ / ++	D	FTU
J	M	1A	+++ / ++	++ / +	B	Strong in basal layer, weakening upwards
K	M	1A	+	+	A	Background
L	M	1A	+++	+++	A	FTU (v thin urothelium)
M	F	1A	+++	+++	B	FTU some substrata
N	F	1F	+++	++	D	FTU + substrata
O	F	2A	++ / +++	++ / +	D	FTU + substrata
P	F	2A	+++	+++	C	Strong in folds, getting weaker towards top

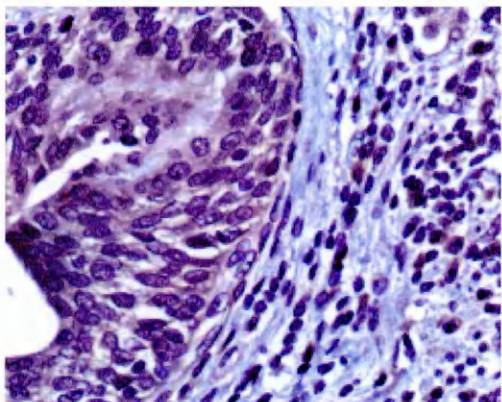
Key to table 5.2
 + = weak staining, ++ = moderate staining, +++ = strong staining
 A = 0 – 25%, B = 26 – 50%, C = 51 – 75%, D = 76 – 100%
 FTU = full thickness of urothelium.

The first finding from this table is that there appears to be no relation between staining intensity or distribution and stage and grade. The majority of sections tested were of the same stage and grade, yet staining was not consistent between these. When compared to the few sections of different stage and grade, no major contrast can be highlighted. All section demonstrate the presence of 5/11 protein except one, section K, which appeared to produce background staining within the substratum alone.

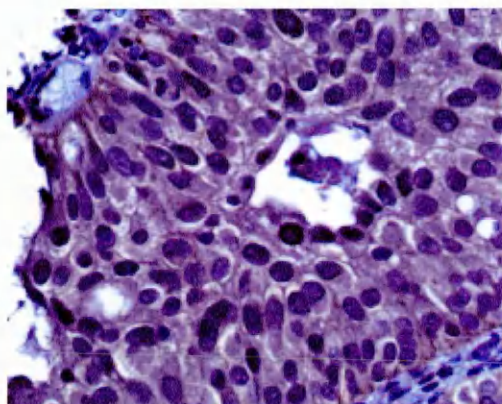
Plate 6 : Bladder tumour sections stained with 5/11 antibody (1/25 dilution)



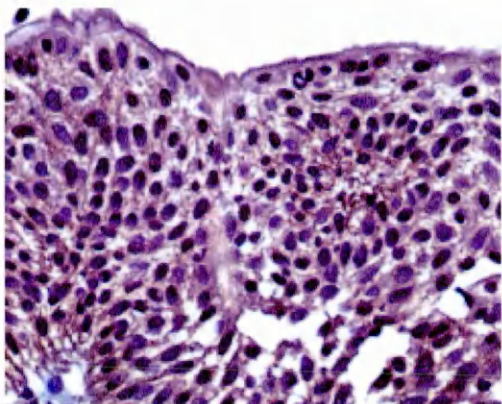
6A – 1G *in situ* section. Staining in full thickness of urothelial layer.



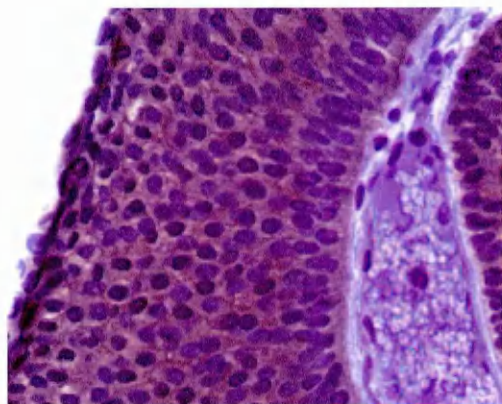
6B – 1F section. Staining in urothelium and Some light staining in substrata



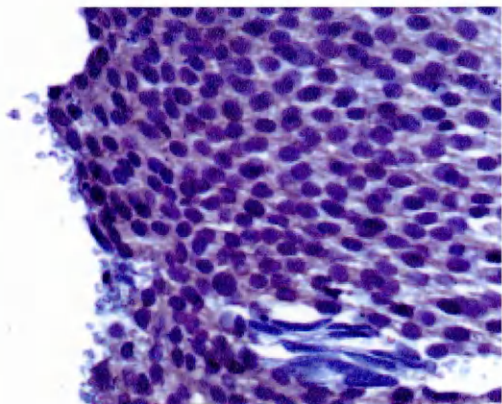
6C – 1A section. Heavy staining of urothelium.



6D – 1A section. Medium staining of urothelium, including umbrella cells.



6E – 2A section. Very heavy staining throughout urothelial layer.



6F – 2A section. Lighter staining, again Throughout urothelium.

Consistent with earlier findings with anti-CD44 antibodies (see chapter 2), both cytoplasmic and membrane staining was noted, though this staining was not consistent through the whole section in many cases. In the majority of these cases, the staining was either between weak and moderate or moderate to strong, but in one case (section A), there was high inconsistency, with staining varying between weak and strong in both cytoplasm and membrane.

The distribution of staining was relatively high for many sections, with most showing a staining distribution of 50% and above. This is not surprising as the majority of samples were trans-urethral resections of bladder tumours (TURBT's), therefore the majority of tissue present is urothelial in nature. The majority of this staining occurred across the full extent of the urothelial layers, from basal layer to umbrella cells, but there were some exceptions. Some sections demonstrated a strong basal layer staining, which became weaker towards the top of the urothelium. The carcinoma *in situ* section demonstrated the strangest pattern of expression, with one (section A) demonstrating staining from the middle to the top of the urothelium alone, and the other (section B) demonstrating staining throughout the urothelium with the exception of the umbrella cells (top layer of urothelium) which demonstrate no staining at all. Where substrata layers were found to be present, a number did demonstrate some staining though this was very weak in nature.

Due to the lack of normal tissues available for immunohistochemical analysis, western blotting was employed to investigate the presence of 5/11 protein in exfoliated urothelial cells. These were extracted from voided urine samples of patients without bladder cancer, a larger group of healthy patients and a small number of tumour patients.

The group consisted of 44 people, 16 females and 28 males. Of these, 10 were patients with known tumours (4 female / 6 male), 11 were from patients with no known tumour (5 female / 6 male) and 23 normal healthy volunteers (7 female / 16 male, with 14 being over the age of 35 years). The tumour group consisted of a group of pre-operative patients with a visually confirmed tumour, but not histologically confirmed at this point. No quantification of protein took place as after concentration,

samples yielded varying amounts of protein, ranging from little or no concentrate to large volumes of concentrate. For this reason it was decided not to undertake quantification.

Table 5.3 summarises the results in each group in terms of both sample group and individual gender.

Table 5.3 – Summary of western blot results with 5/11 antibody. Percentages below male and female results are for that gender, whilst overall percentage figures combine both male and female results.

	Female		Male		Overall	
Tumour	+ve	-ve	+ve	-ve	+ve	-ve
n=10	2	2	2	4	4	6
	50%	50%	33.3%	66.7%	40%	60%
Normal Patient	+ve	-ve	+ve	-ve	+ve	-ve
n=11	1	5	3	2	4	7
	16.7%	83.3%	60%	40%	36.4%	63.6%
Normal	+ve	-ve	+ve	-ve	+ve	-ve
n=23	4	3	0	16	4	19
	57%	43%	0%	100%	17.4%	82.6%

In the tumour group, not all samples gave a positive result, which is unexpected due to the IHC results presented with the same antibody. Only 50% of the female group gave a positive result, though the male result was lower at 33.3%. This gave an overall positive percentage for both males and females of 40%. A similar overall positive rate was noted in the normal patient group, though this was due to a much increased positive group of males, with 60% of males in this group providing a positive result. Only 16.7% of females gave a positive result, and the overall percentage was 36.4%.

Within the healthy control group no males gave a positive result, giving a 100% negative result as would be expected. Interestingly in the females, 57% demonstrated a positive result, increasing the overall positive rate to 17.4%.

From the western blot results, combining the results of the normal patient and normal healthy group gives the antibody an overall specificity of 76.5%. Its sensitivity is 40% from this group, but when these are combined with the IHC results, the overall sensitivity is 73.1% (16 IHC results, of which 1 could be considered to be negative plus 4 out of 6 positive western blot results).

When looking specifically at gender, the overall specificity and sensitivity does change. In females (a group consisting of 12 tumours and 13 normals), the specificity is found to be only 61.5%, though the sensitivity increases to 83.3%. In males (a group consisting of 15 tumours and 21 normals), the figures are reversed, with a high specificity of 85.7%, but a lower sensitivity at 66.7%. These figures have possible implications for the use of this antibody within screening test format.

5.3.3 5/11 expression in colorectal Cancer

Analysis of 5/11 expression within colorectal cancer was again carried out using IHC on paraffin sections, though a stage-specific analysis was undertaken, using 5 sections each of polyp and Dukes stage A to D. Table 5.4 summarises the colorectal immunohistochemistry results, whilst plate 7 (pages 124-125) demonstrate the a selection of the results visually.

Table 5.4 Summary of results of IHC with 5/11 antibody in Dukes staged colorectal sections.

	Section	Stain (M)	Stain (C)	Distribution	Area
Polyp	P1	+	+	C	Epithelium
	P2	-	-	-	-
	P3	++	++	D	Epithelium and substrata
	P4	++	++	D	Epithelia and some sub
	P5	++	++	C	Epithelial and some sub
Dukes A	A1	-	-	-	-
	A2	+	+	A	Top layer of epithelium
	A3	+	+	A	Top layer of epithelium
	A4	-	-	-	-
	A5	+	+	A	Top layer of epithelium
Dukes B	B1	-	+	A	Epithelial
	B2	++	++	C	Epithelium and substratum
	B3	+	+	B	Epithelium + substratum
	B4	-	-	-	-
	B5	-	-	-	-
Dukes C	C1	-	-	-	-
	C2	+	+	B	Epithelium and substrata
	C3	+	++	A	Lower epithelial and substrata
	C4	++	++	A	Epithelium
	C5	++	+	A	Top epithelial layer

Dukes D	D1	++	++	C	Epithelium + substrata
	D2	++	+++	C	Epithelium of tumour, some non-involved
	D3	+	+++	B	Epithelium, some substrata
	D4	+	++	A	Epithelium
	D5	+	+ / ++	C	Epithelium

Key to table 5.4

+ = weak staining, ++ = moderate staining, +++ = strong staining

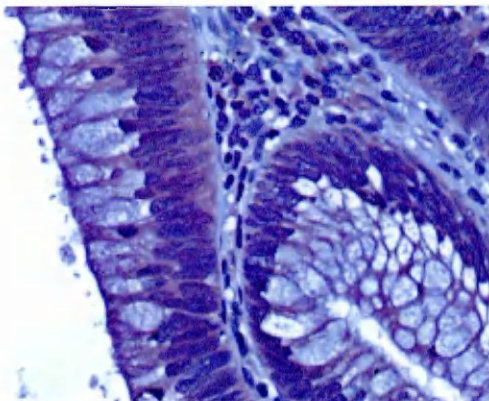
A = 0 – 25%, B = 26 – 50%, C = 51 – 75%, D = 76 – 100%

As was noted in the bladder cancer sections, there appears to be no correlation between 5/11 protein expression and tumour stage. Not all sections express the 5/11 protein, but these negative sections are not all in one staged group. The polyp group contained one negative section, along with the Dukes C group. Dukes A had two negatives, whilst Dukes B showed the highest number with three negative sections. Only the Dukes D demonstrated a positive result on all sections.

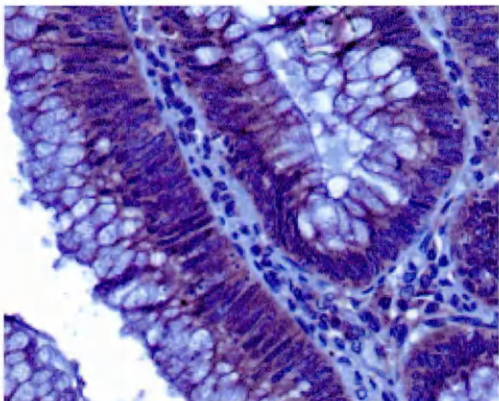
The staining density was generally lower than that found in the bladder tumours and in the HT29 cell line. There was a slight increase in staining in the Dukes D sections, but only one of these showed strong staining, this being cytoplasmic. These densities were not consistent amongst sections of the same stage, with the Dukes D ranging from across the full spectrum from weak to strong. Across separate sections the staining was consistent unlike that seen in the bladder samples. The distribution of staining also differed, with the majority of Dukes staged sections showing staining of between 0 and 50%. This is mainly due to section differences, with the colorectal tissue being a cross section of the organ, containing epithelial, ECM and muscle layers. Polyp sections gave a higher distribution as these consisted of polyp material alone, which like the TURBT sections are epithelium alone. This gave rise to higher distribution of staining in these when compared to the Dukes staged samples.

As no normal or non-involved tissue was available for study a full analysis of the antibody is not possible. From these sections, the specificity for each stage can be calculated as : polyp – 80%, Dukes A – 60%, Dukes B – 40%, Dukes C – 80% and Dukes D – 100%. From these results it can be seen that the protein is expressed in detectable levels in polyp or Dukes C and D ie either very low or very high stage

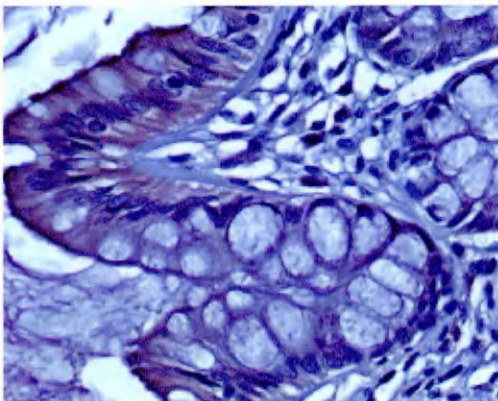
Plate 7 : Colorectal tumour sections stained with 5/11 antibody (1/25 dilution)



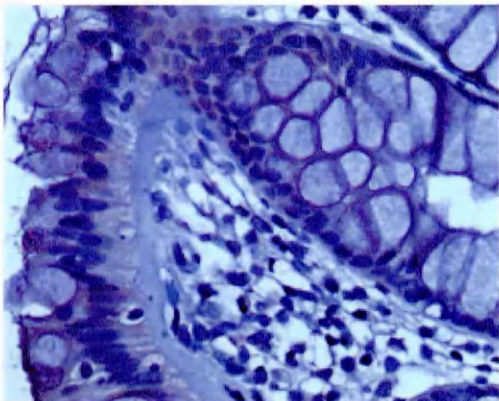
7A – Polyp. Medium staining seen in epithelium, some staining of substrata



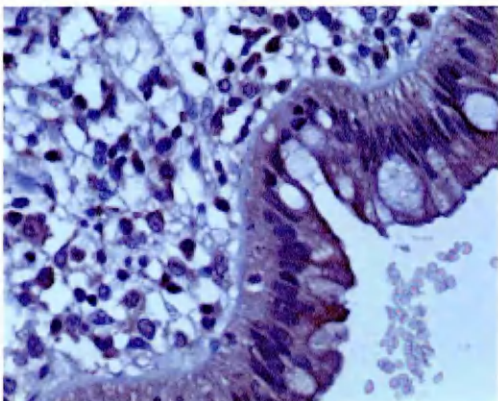
7B – Polyp. Heavy staining of epithelium, at base of cells. Some substratal staining.



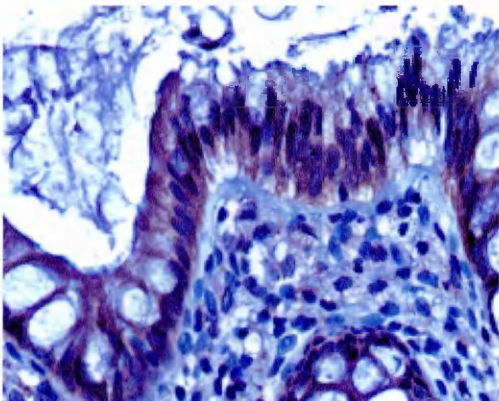
7C – Dukes A. Staining through epithelial layer.



7D – Dukes A. Staining in epithelial, lighter than In 7C. No substratal staining.

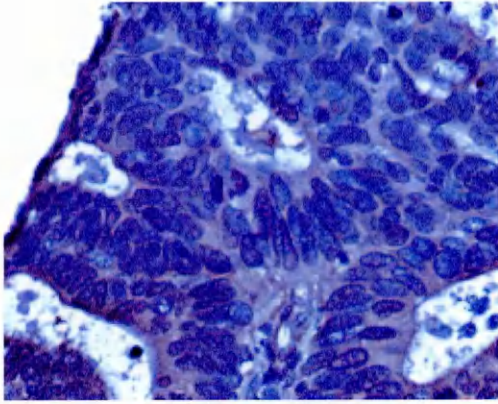


7E – Heavy staining in epithelium.

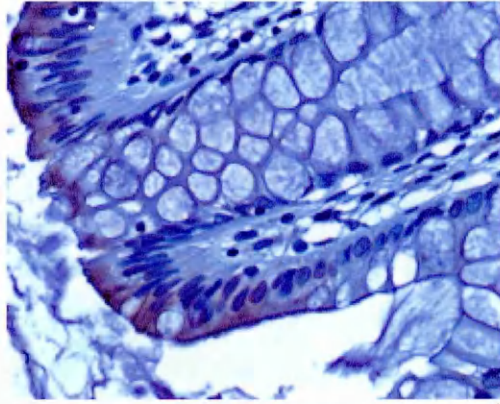


7F – Medium staining in epithelium.

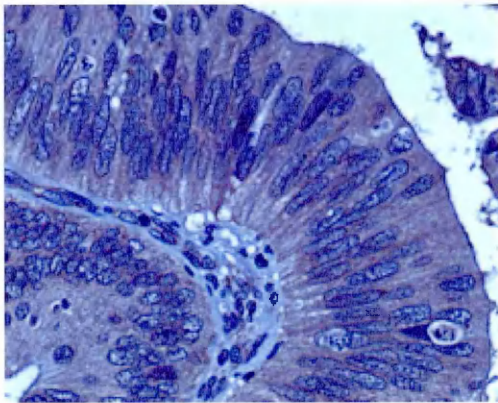
Plate 7 : Colorectal tumour sections stained with 5/11 antibody (1/25 dilution)



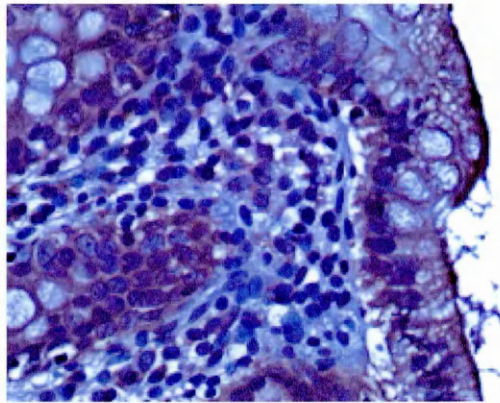
7G – Dukes C. Medium epithelial staining



7H – Dukes D. Epithelial staining, concentrated in some areas.



7I – Dukes D. Epithelial staining throughout.



7J – Dukes D. Strong epithelial staining.

tumours. An assessment of antibody sensitivity is not possible, due to lack of normal tissue samples.

5.4 Discussion

The testing of the 5/11 antibody in both the *in vitro* model and patient samples has demonstrated the presence of a tumour specific protein, the 5/11 junction protein.

All cell lines of the *in vitro* model expressed the protein, though staining density varied between them. When staining densities are compared to those seen for Hermes and V6 antibody (see Chapter 2), a similar trend in expression is noted. Some background staining was noted at higher antibody levels (data not presented) but optimisation of the antibody to a 1/25 dilution lowered this significantly. As is noted with other anti-CD44 antibodies, staining is seen in both membrane and cytoplasm, demonstrating an overexpression of protein within single cells. Unlike the Hermes antibody however, membrane staining is less distinct than the cytoplasmic staining. The possible reasons for cytoplasmic staining were discussed in chapter 2, and could be due to a simple overexpression of complete CD44 protein, or increased levels of incorrectly or poorly processed proteins.

Western blotting with the antibody demonstrated a protein of approximately 130kDa. The formation of a large protein may be expected due to exon 11 being one of the largest variant exons (132bp), and if, as previously hypothesised it is always spliced with exon 12 (102bp) this would add to its size. The main band of 130kDa is not the only one seen, but the reasons for the presence of these other bands is not entirely clear. As was discussed in chapter 2, these could be mis-translated or poorly processed proteins.

Immunohistochemical results for the bladder sections demonstrated staining throughout the thickness of the urothelium in the majority of tumour samples. This finding alone can be thought of as tumour specific, as in normal tissue CD44 staining is evident in the basal layer of the urothelium only (Woodman *et al*, 1996 and Sugino

et al, 1996). This finding compares to the pattern of staining observed with other anti-CD44 antibodies (Woodman *et al*, 1996 and Sugino *et al*, 1996). Again both membrane and cytoplasmic staining are evident, though in many samples there is heterogeneity in the staining densities noted. The majority of this variation in staining is noted within the membranes, but some variation is also noted within the cytoplasm. Though a homogenous density of staining was demonstrated within the cell lines, this pattern would not be expected within tissue samples. Cell lines can be thought of as a homogenous cell population, as cells are all of one type, arising from the same tumour. Differences within the consistency of staining within the section as a whole may be due to the presence of both tumour and non-involved tissue, and inconsistencies in the state of differentiation of the tumour cells present. Where some sections of the substrata or lamina propria below the urothelium cells are found some background staining is also noted. Previous studies have noted this staining with CD44 standard antibodies, such as Hermes (Tarin *et al*, 1995 and Woodman *et al*, 1996), and this was attributed to detection of the haematopoietic form of CD44 on lymphocytes present within the substratum. In the case of the 5/11 antibody, the cause of the staining is unknown.

In the Gis, sample, a strange pattern of urothelial staining was noted. Staining in some areas only occurred at either the very top layer of the urothelium or from top to mid urothelial layer. In most other samples, where staining was not full thickness, it was concentrated at the lower edge. The exact reason for this is not clear.

For analysis of 5/11 expression in normal tissue, exfoliated urothelial cells from voided urine samples underwent western blotting. The normal group consisted of both normal healthy volunteers and patients with a urological problem which was known not to be due to the presence of a tumour. The inclusion of the patient group allows some insight into false-positive potential and the group as a whole aids in determination of tumour specificity.

The majority of normal volunteers gave negative results when tested. However, false positives were seen, and these fell into two distinct groups. Many of the false positives were found in females and these form the first distinct group. In this study, all women who gave a positive result are well below the risk age for development of bladder cancer, and when this information is combined with the decreased risk in

women over males, these cannot be accepted as true positive results. The reasons for the false positives are not fully understood, but it is a finding that has been noted previously (personal communication, Woodman, 2002). The main reason attributed to these results is the contamination of urine samples with desquamated cells from the vagina and cervix. It has been noted that there is a high expression of CD44V isoforms within normal cervical squamous epithelium (Gorham *et al*, 1996(b)). Exfoliation of these cells would lead to them entering vaginal secretions and contaminating the urine sample.

The causative factors in the other group of false positives are well documented. It was noted during sample preparation that a number of urine samples contained microscopic haematuria. It has been shown that haematuria is a cause of false positives and it has been suggested that regardless of the outcome of such an assay, anyone who presents with haematuria should immediately undergo cystoscopy to ensure the cause is not a bladder tumour. Another cause of false positives are patients who have undergone BCG intravesical therapy for treatment of bladder cancer. BCG stimulates an inflammatory response within the bladder tissue, and any residual inflammatory response which may be present could produce a false positive.

A small number of patients with known tumours was also included within the western blot group, but not all gave the expected positive result. It must be remembered that these were pre-operative patients with a visually confirmed tumour only, no histological data was available. It may be that lesions thought to be tumour at first visual examination may in fact be ulcerative and would therefore be expected to give a negative result. In a study on telomerase expression in urine it was suggested that urine samples be taken in duplicate over a set time period, as there are differences in exfoliation of normal and neoplastic cells. This may result in a larger proportion of normal cells present in one sample, but a higher proportion of tumour cells in another (Yoshida *et al*, 1997).

These results may also be due to problems in sampling procedures, with samples being taken from patients who have ingested large volumes of fluid, resulting in a very dilute sample with low cell presence. When these samples are processed using a protein microconcentrator, there is often little or no concentrated protein present within the sample. Similarly if samples were taken and left for too long before being

prepared, a certain amount of protein degradation will occur. It was also noted that some samples formed a heavy precipitate when left, which did cause some problems with the subsequent protein extraction. The speed by which such samples are processed have been shown to be of importance in another study on CD44 proteins in urine, therefore a simple yet robust protocol for collection and processing is required (Sugiyama *et al*, 1995). In this study, time did differ between patient and healthy volunteer groups sample processing due in part to logistical reasons.

Specificity and sensitivity for the antibody varied between the sexes, largely due to the high number of false positives noted within the female normal group. When compared to the gold standard for urinary bladder cancer diagnosis, voided urine cytology (VUC), the overall sensitivity is higher. The 5/11 antibody also compares favourably to other urinary based assays such as BTA STAT and Trak, NMP22 and Immunocyt, as demonstrated in table 5.4 below.

Table 5.5 – Comparison of sensitivity and specificity values of the 5/11 antibody with commercial FDA approved urinary based assays used in clinical practice.

Assay	Sensitivity	Specificity
Voided urine cytology	38.4%	94.2%
BTA-stat	65-82.8%	64-68.9%
BTA trak	66-72%	50-69%
NMP22	64.3-90.8%	80-89%
5/11 antibody	73.1%	76.5%

From this table it can be seen that at this early stage the antibody does compare favourably to other urinary based assays, and with further optimisation within a complete assay format may well provide a much more sensitive and specific assay.

The colorectal study within this chapter was undertaken using paraffin embedded sections. It was not possible to procure normal samples or fresh frozen tissue to carry out western blotting. The analysis of expression was carried out using a staged study on sections of Dukes A – D tumours, and polyps.

The finding of 5/11 protein expression within colorectal cancer in one way confirms the exon linkage results gained by Goodison *et al* (1997). It also furthers their argument that specific CD44 splice variants would be of use for cancer detection, and that a protein based assay using exon linkage data is possible. The pre-dominant junctions within the colorectal exon linkage assay were 5/7 and 5/11 which were both found to be present in 85% of tumours. The presence of 5/11 protein was however found in a smaller percentage of patients, the total being 60% (tumours only – does not include polyp group).

The reasons for this decreased percentage may be due to differences in sample cohorts studied. This study was based on Dukes stage, with a defined number of samples for each stage assessed. No comment was made by Goodison *et al* (1997) on the stage and grade of tumours examined within the exon link study, therefore this may be a reason for the variance. If certain tumour stages express the 5/11 protein at a higher percentage than other stages, a sample cohort where these tumours are in the majority would skew the results in this manner.

It has however been noted that in this pilot study no correlation with stage of tumour could be established. Expression was found at all stages, including polyp, yet no trend could be distinguished in terms of either staining presence or density. Results for each stage were heterogenous, with the majority of groups demonstrating staining of between two and four sections (40 – 80%). Only one stage (Dukes D) demonstrated staining in all sections examined. Where sections were shown to express 5/11 protein there was again variation between the stage in terms of staining density. Though there was a slight upward trend in expression as stage increased, this can not be taken as a positive correlation. The cohort is small, and therefore any trends present may in fact be exaggerated due to the group size. Also, no analysis of section grade was included with the staging information, and this would probably affect the 5/11 expression.

This lack of stage correlation was noted by Gorham *et al* (1996a) in a study of mRNA transcript expression within colorectal cancer. They investigated a number of staged and graded tumours and found no positive correlation between CD44 variant expression and tumour stage / grade. It was however noted by Wielenga *et al* (1993)

that CD44 variant protein expression did in fact correlate with tumour progression in colorectal cancer. In studying CD44 variant expression using variant specific antibodies, they summarised that different variants are expressed at different stages within tumour development. Their findings do conflict with the results gained thus far from this study.

Wielenga *et al* (1993) stated that variants of CD44 without V6 (exon 11) do not appear at early stages, including adenomas and polyps. As tumour stage increased, so the variant expression changed, and variants containing V6 were found, especially at the latest stages (Dukes C and D). This study demonstrates the presence of 5/11, which contains the exon 11 junction at very early stages, and throughout all intervening stages. Though the reasons for the differences noted are unclear, this strengthens the reasons for investigating a tumour specific exon junction rather than the presence of the exon alone, as discussed in chapter 3.

The protein itself was noted in the colorectal epithelial layer in the majority as was noted by previous studies of variant antibody expression (Woodman *et al*, 1996; Gorham *et al*, 1996a). Some substratal staining was noted, which when found with standard antibodies can be attributed to the presence of haematopoietic cells within the lamina propria. With the 5/11 antibody this would not be expected, therefore the exact cause of this staining has yet to be elucidated.

Expression of the 5/11 protein with over half the polyp samples (3/5) suggests that this antibody may be useful within an early detection system. It is known that many polyps are pre-cancerous in nature, therefore the detection of 5/11 protein within these samples suggests it can be used for the detection of very early changes. Tumour sensitivity of this antibody was found to be at its highest when examining very low and very high grade tumours (polyps, Dukes C and Dukes D). The apparent reduction in detection ability at the middle stages should not necessarily prove to be a problem considering the required use of this marker. As it is to be used in the monitoring of patients who have already received treatment for one or more tumours, it is early tumours alone which should be under investigation.

Further examination of the 5/11 protein within colorectal cancer is required, including the investigation of normal and non-involved tissue. Without the testing of these tissues, no assessment of tumour specificity can be made. To assess the potential of the 5/11 antibody for non-invasive detection of colorectal cancer, a study of expression within exfoliated cells extracted from faeces is also required. It has been demonstrated that analysis of such samples is possible, and should present no more problems in terms of sample preparation than those noted with urine samples.

6 Final Discussion

The initial aim of the project was to analyse CD44 expression for the presence of a tumour specific marker. Once found, this information would then be used to reach a platform from which a working diagnostic could be developed. This aim has been achieved, through meeting all of the objectives established within chapter 1. The scientific impact and importance of each work objective has been discussed in each associated chapter. It is the aim of this final discussion chapter to employ a more holistic approach to the examination of this work, drawing it together in the context of its potential diagnostic application. The natural progression at the end of this project as it now stands is the production of a working early detection tool, yet a number of points need to be considered before reaching this final goal.

The search for tumour markers is an area of research that looks set to continue until we can at least identify the causative and consequential factors in all tumours. The ways in which tumour markers are being investigated have changed rapidly as new technologies and methods have arisen. From blood and serum markers, to mutation based markers. The latest trends in research are the use of genomics and proteomics to identify possible markers from a wide range of target genes or proteins (Srinivas *et al*, 2001 (a)).

Genomics, as its name suggests, focuses upon the whole cellular genome, the main technique in use being the expression microarray. A microarray in its simplest form is a solid surface upon which either single stranded cDNA or oligonucleotide probes have been dotted. A DNA or RNA extraction from a tissue is flooded over the array, and areas where hybridisation occurs are detected, usually through the use of fluorescent labelling. Comparative studies of normal and tumour tissue can highlight genes which show differential expression and may be potential markers. Thousands of genes can be tested, resulting in large volumes of data being produced. At this stage though, no gene functions are known (Onyango, 2002).

Proteomics is the study of cellular protein and the main techniques in use are either 2D-electrophoresis (followed by mass spectrometry) or protein arrays. Again

comparative studies are employed and the proteins unknown. With both methods profiling of certain tumour types has been undertaken, resulting in a large number of possible markers, but with unknown function. Other methods which have developed alongside these include Laser Capture Microdissection, where single cells or small groups of cells of interest can be removed from a tissue sample and the nucleic acid or protein extracted (Srinivas *et al*, 2001 (b) and Pennington and Dunn, 2001).

The format of assays have also changed, with a move away from single marker assays to those involving multiple markers. Due to the complexity of changes seen within tumourigenesis this is perhaps not so surprising. Multiple marker assays would provide a more comprehensive picture of the events which unfold within a patient, and may be able to provide information of both diagnostic and prognostic value in one step.

The ultimate hope for a multiple marker assay is the development of a completely non-invasive test which is 100% specific and sensitive. It must also be able to detect at the earliest stages of cancer development, and if possible provide some form of prognostic information, as well as indicators of stage and grade. Lokeshwar and Solway (2001) describe the ideal non-invasive test (for bladder cancer) as “sensitive, specific, rapid, technically simple and have low intra and inter-assay variability”. Invasive methods would still be required, but could be used for imaging and biopsy of tumours.

Within the NHS, patient testing is changing, with “near patient testing” becoming increasingly important. This sees many tests moving out of the laboratories and onto the wards, allowing the most rapid turnaround time for assays possible.

Assays for early detection are being discussed more in terms of screening assays for large scale screening programmes for the general population. Screening programmes are only used for certain sets of disease, the main factors in their choice being ability to screen reliably and the ability to treat the disease satisfactorily if found. In the UK, only two screening programmes are currently being run – screening for cervical cancer and for breast cancer. In both cases these are aimed at a particular subset of the population, females of certain ages – above the age of 21 for cervical, and over 45

for breast. The newest screening programme which has just finished its pilot trial is screening for colorectal cancer, which gave some promising results, and another pilot currently underway is for prostate cancer screening (National Screening Committee, 2002). These new screening programmes and the general search for early detection assays within the UK links in with the NHS plan set out by the government in 2000. In this document they pledged to reduce the levels of cancer in the under 75 age groups “..by at least one fifth by the year 2010” (NHS Plan, 2000)

Linked in with both early detection assays and screening tests are those used for monitoring of cancer recurrence. Typically patients who have had a tumour removed are monitored regularly so recurrences can be detected and treated with speed. In the case of bladder cancer, patients are tested every three to six months, whilst colorectal cancer patients are monitored every six to twelve months.

As stated earlier, the aim of this project was to provide a platform from which a non-invasive assay for use within a district general hospital (DGH) setting, specifically aimed at the monitoring of recurrence, could be developed. The specification of a DGH as the setting for the assay provides some constraints on the type of assay required. Initially there appears to be two distinct choices of assay : DNA or protein based, yet not all are suitable.

The exon link assay in its original form does not lend itself easily to diagnostic use. Though the PCR assay is simple to carry out, the requirement for Southern blotting within the procedure for confirmatory reasons discounts the use of this assay due to lack of rapidity. The PCR-ELISA is therefore a much better method as is it just as simple to perform, but takes less time. If the PCR is set up as staff are leaving the laboratory one night, then it will be ready to place into the ELISA the following day, resulting in a fast turnaround time. Unfortunately this assay does have many disadvantages if it were to be used within a clinical lab. To undertake the ELISA, mRNA extraction from exfoliated urothelial cells is required. This extraction could take a large amount of time if a number of samples are being treated concurrently. Time between sample collection and sample preparation must be fairly rapid due to mRNA degradation. A cDNA step is also required before the PCR assay can be undertaken. PCR based assays require optimisation to ensure the assay is robust and

will always give the expected result. However, although PCR is fairly reliable, results can be easily affected if any form of contamination is present. Finally, many DGHs do not have the facilities for undertaking PCR based assays, therefore a test based on DNA would not be viable.

Consequently, this reduces the choice to a protein based assay, yet there are a number of formats in which this could be carried out. Sample preparation would be much simpler as extracted protein would be less prone to degradation than mRNA. Addition of a protease inhibitor either to the sample pot before urine collection or directly after would safeguard against a high protein loss within the sample. Protein extraction itself is simple. Concentration of the protein would be required, but similarly this is a simple matter. One choice of assay could be western blotting, but as with Southern blots, this is overly time consuming. The next choice of format would be an ELISA, similar to that designed by Woodman *et al* (2000). Many laboratory-based assays are already in this format, and therefore it should easily integrate into the testing repertoire of the lab in question. For a monitoring assay, an ELISA would only be required to be qualitative if a tumour specific marker has been adopted. From this method, further work could produce an even simpler assay based on a lateral flow assay, commonly known as a “dipstick” format, allowing near patient testing and more rapid results.

As a result of this project, a marker which appears to be tumour specific has been detected, and initial results with an antibody directed against the protein product of the variant exon junction have suggested its potential use as a marker for monitoring purposes.

The use of initial exon link results and PCR-ELISA demonstrated the 5/11 exon junction was potentially the more “useful” marker, as compared to 5/12. The study cohort was small, and results of the PCR-ELISA were not as distinguishing as had been hoped. The 5/11 junction was the most predominant, but the result margin between that of 5/11 and 5/12 was not very great. However, as the aim was to move towards a protein based assay fairly quickly, it was decided that a more in-depth study was not appropriate.

One major point to have come out of both this and later sections of the study was a need for simple and well organised sample procurement protocols. The staff of the urology department at NGH were unused to organising tissue sample collections, therefore protocols and guidelines for such a scheme had to be arranged. These protocols had to be simple to undertake as the majority of time it was the clinicians themselves (mainly registrars) who were responsible for this collection, whilst treating patients. The ideal person for collection of samples would be a nursing member of theatre staff, who would be responsible for collection in the correct manner and for recording the required sample details. Towards the latter part of the study, this was starting to happen. Though cohorts in this study were affected by these problems, future studies within this setting should prove easier as staff are now aware of the needs of research sample collection and the procedures involved.

The 5/11 junction was chosen as the marker and subsequently an antibody against the protein product of the junction was raised, to be used in both bladder and colorectal assays. Polyclonal antibody production could have presented a number of problems due to the heterogeneous nature of such an antibody, but affinity purification appears to have removed many of the concerns (Catty and Raykundalia, 1995). The presence of high levels of background is often associated with the use of a polyclonal antibody, but within this study it has not proved to be a major issue. Optimisation of antibody dilutions coupled with a polyclonal antibody-specific method for IHC appears to have lessened the problem considerably. Some background staining within tissue sections was noted, but this did not affect the detection as it was relatively low level in most cases.

From initial results in bladder samples, the antibody appears to be tumour specific for this tissue. Though false positives were seen they could all be explained and were not thought to be cause for concern. The finding that urothelial cell protein extraction could be used for antibody detection of the junction protein through the use of western blot suggests that the antibody could be applied to an ELISA. Further work in a larger cohort of patients would be required before development of an ELISA could take place. It would be interesting to carry out further investigation into stage and grade expression of the protein. This may prove be difficult as the majority of

tumours with which people present are early superficial tumours, these are therefore the main tumour types resected and hence collected.

The ELISA developed by Woodman *et al* (2000) was a multiple capture assay using an antibody directed against a protein epitope of exon 1, and more interestingly a protein epitope of exon 11. The assay gave a specificity of 100% and a sensitivity of 80.2%. Adapting this assay for use with the 5/11 antibody would not prove difficult, and it may increase the sensitivity of the current assay (Woodman *et al*, 2000).

Within colorectal tissue, results again looked promising though no link to stage could be made. The collection of further samples, including frozen tissues for protein extraction for western blot would provide a better insight into expression, though a large group of normal samples would also be required, to ensure tumour specificity. Though an assessment of sensitivity could be made from this work, due to the lack of normal samples no assessment of specificity could be undertaken.

This project has fulfilled its objectives and hence reached its aims as set out in chapter 1. The main emphasis throughout was on bladder cancer, and in terms of the above aim, not much work is required to produce a crude form of diagnostic assay. An ELISA based on the design of Woodman *et al* (2000) but using the 5/11 antibody would be simple to set up. It is the optimisation of the assay which would take time. Single marker assays are not considered to present enough information, therefore other markers should also be included in the final assay. The MCM family of proteins are currently of interest, as these have been shown to be diagnostically useful in the detection of a number of cancers, including bladder and colorectal cancer. Inclusion of such markers into an ELISA format could possibly provide a much more sensitive and specific assay, whilst retaining rapidity and simplicity.

7 Final Conclusions and Further Work

In conclusion, the work presented here has met the initial aim as set out for this project – a platform has been reached from which a protein based assay could now be developed. The finding of a tumour specific exon-junction and the subsequent development of a novel antibody, as discussed in this thesis, provides the basis for such future work. Some preliminary studies have been undertaken on the use of this antibody to detect protein from urothelial cells, and its possible usefulness in a non-invasive setting demonstrated. As a prototype ELISA is already available, it can be considered to be a fairly short step from this platform to a working assay. However, a large body of further work arises from this project, including optimisation and clinical evaluation of such an assay, and this is summarised below :

1. Investigation of 5/11 protein expression in a larger cohort for the bladder study; with the linked aim to undertake a stage and grade study.
2. Investigation of tumour specificity in colorectal samples through inclusion of normal samples into the study.
3. Analysis of non-invasive use of 5/11 in colorectal cancer by investigating protein extracted from exfoliated cells present in stool samples.
4. Development of the ELISA assay based on the format adopted by Woodman *et al* (2000).
5. Inclusion of further markers into the CD44 5/11 assay eg MCM proteins
6. Clinical evaluation of ELISA in both colorectal and bladder cancers.
7. If ELISA is clinically useful, possibility of “dipstick” assay development for near patient testing.

Reference List

- Abassi A.M., Chester K.A., Talbot I.C., Macpherson A.S., Boxer G., Forbes A., Malcolm A.D.B., and Begent R.H.J. (1993). CD44 is associated with proliferation in normal and neoplastic human colorectal epithelial cells. *European Journal of Cancer* 29A, 1995-2002.
- Abulafi A.M. and Marks C.G. (1998). The role of colonoscopy in coloproctological practice. *Hospital Medicine* 59, 612-616.
- Anwar S., Hall C., and Elder J.B. (1998). Screening for colorectal cancer : present, past and future. *European Journal of Surgical Oncology* 24, 477-486.
- Bajorath J. (2000). Molecular organization, structural features, and ligand binding characteristics of CD44, a highly variable cell surface glycoprotein with multiple functions. *Proteins* 39, 103-111.
- Baron J.A., Sandler R.S., Haile R.W., Mandel J.S., Mott L.A., and Greenberg E.R. (1995). Folate intake, alcohol consumption, cigarette smoking and risk of colorectal adenomas. *Journal of the National Cancer Institute* 90, 57-62.
- Bell M.V., Cowper A.E., Lefranc M.-P., Bell J.I., and Screaton G.R. (1998). Influence of intron length on alternative splicing of CD44. *Molecular and Cellular Biology* 18 : 10, 5930-5941.
- Bindels E.M.J., Vermey M., van den Beemd R, Dinjens W.N.M., and van der Kwast T.H. (2000). E-Cadherin promotes intraepithelial expansion of bladder carcinoma cells in an *in vitro* model of carcinoma *in situ*. *Cancer Research* 60, 177-183.
- Borgya A., Woodman A., Sugiyama M., Donie F., Kopetzki E., Matsumara Y., and Tarin D (1995). Isolation and characterisation of antibodies which specifically recognise

the peptide encoded by exon 7 (v2) of the human CD44 gene. *Clinical Molecular Pathology* 48, M241-M250

Bossy-Wetzel E. and Green D.R. (1999). Apoptosis : checkpoint at the mitochondrial frontier. *Mutation Research* 434, 243-251.

Bruno S., Fabbi M., Tiso M, Santamaria B., Ghiotto F., Saverino D., Tenca C., Zarcone D., Ferrini S., Ciccone E., and Grossi C.E. (2000). Cell activation via CD44 occurs in advanced stages of squamous cell carcinogenesis. *Carcinogenesis* 21, 893-900.

Buchumnesky V., Klein A., Zemer R., Kessler O.J., Zimlichman S., and Nissenkorn I. (1998). Cytokeratin 20 : A new marker for early detection of bladder cell carcinoma. *The Journal of Urology* 160, 1971-1974.

Catty D. and Raykundalia C. (1989). Production and quality control of polyclonal antibodies. In *Antibodies : A practical approach*. Catty D., ed. (IRL Press), pp. 19-79.

Cohen M.B., Griebeling T.L, Ahaghotu C.A., Rokhlin O.W., and Ross J.S. (1997). Cellular adhesion molecules in urologic malignancies. *American Journal of Clinical Pathology* 107, 56-63.

Czerniak B., C.V., Li L., H.S., Johnston D., Ro J.Y., Luthra R., Logothetis C., Von Eschenbach A.C., Grossman H.B., Benedict W.F., and Batsakis J.G. (1999). Superimposed histologic and genetic mapping of chromosome 9 in progression of human urinary bladder neoplasia : implications for a genetic model of multistep urothelial carcinogenesis and early detection of urinary bladder cancer. *Oncogene* 18, 1185-1196.

- Dalchau R., Kirkley J, and Fabre J.W. (1980). Monoclonal antibody to a human brain-granulocyte-T lymphocyte antigen probably homologous to the W 3/13 antigen of the rat. *European Journal of Immunology* 10 , 745-749.
- Dashwood R.H. (1999). Early detection and prevention of colorectal cancer. *Oncology Reports* 6, 277-281.
- Davies G., Jiang W.G., and Mason M.D (1999). Cell-cell adhesion molecules and their associated proteins in bladder cancer cells and their role in mitogen induced cell - cell dissociation and invasion. *Anticancer Research* 19, 547-552.
- Dong S.M., Traverso G., Johnson C, Geng L., Favis R., Boynton K., Hibi K., Goodman S.N., D'alessio M., Paty P., Hamilton S.R., Sidransky D., Barany F., Levin B., Shuber A., Kinzler K.W., Vogelstein B., and Jen K. (2001). Detecting colorectal cancer in stool with the use of multiple genetic targets. *Journal of the National Cancer Institute* 93, 858-865.
- Farrow G.M. (1990). Urine cytology in the detection of bladder cancer : A critical approach. *Journal of occupational medicine* 32, 817-821.
- Fearon E.R. and Vogelstein B. (1990). A genetic model for colorectal tumorigenesis. *Cell* 61, 759-767.
- Fields G.B. and Noble R.L. (1990). Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *International journal of peptide and protein research* 35, 16-214.
- Foresman W.H. and Messing M.D. (1997). Bladder cancer : Natural history, tumour markers and early detection strategies. *Seminars in Surgical Oncology* 13, 299-306.

- Freeman A., Morris L.S., Mills A.D., Stoeber K., Laskey R.A., Williams G.H., and Coleman N. (2001). Minichromosome maintenance proteins as biological markers of dysplasia and malignancy. *Clinical Cancer Research* 5, 2121-2132.
- Goodison S., Yoshida K., sugino T., Woodman A., Gorham H., Bolodeoku J., Kaufmann M., and Tarin D. (1997). Rapid analysis of distinctive CD44 RNA splicing preferences that characterize colonic tumours. *Cancer Research* 57, 3140-3144.
- Goodison S. and Tarin D. (1998a). Clinical Implications of anomalous CD44 gene expression in neoplasia. *Frontiers in Bioscience* 3, e89-e109
- Goodison S. and Tarin D. (1998b). Current status of CD44 variant isoforms as cancer diagnostic markers. *Histopathology* 32, 1-6.
- Gorham H., Sugino T, Bolodeoku J., Yoshida K., Goodison S., and Tarin D (1996b). Distribution of CD44 messenger RNA in archival paraffin wax embedded tumours and normal tissues viewed by in situ hybridisation. *Clinical Molecular Pathology* 49, M147-M150
- Gorham H., sugino T., Woodman A.C., and Tarin D. (1996a). Cellular distribution of CD44 gene transcripts in colorectal carcinomas and in normal colonic mucosa. *Journal of Clinical Pathology* 49, 482-488.
- Grossman H.B. and Dinney C.P.N. (2000). Markers of bladder cancer - state of the art. *Urologic Oncology* 3-10.
- Guo Y., Ma J., Wang J., Che X., Narula J., Bigby M., Wu M., and Sy M-S. (1996). Inhibition of human melanoma growth and metastasis *in vivo* by anti-CD44 monoclonal antibody. *Cancer Research* 54, 1561-1565.
- Günthert U (2001). CD44 : A multitude of isoforms with diverse functions. *Current Topics in Microbiology and Immunology* 184, 45-62.

- Günthert U., Hofmann M., Wolfgang R., Reber S, Zöller M., Hausmann I., Matzku S., Wenzel A., Ponta H., and Herrlich P. (1991). A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. *Cell* 65, 13-24.
- Hamada J, Sawamura Y., and Van Meir E.G (1998). CD44 expression and growth factors. *Frontiers in Bioscience* 3, d657-d664
- Harnden P., Allam A., Joyce A.D., Patel A., Selby P., and Southgate J. (1995). Cytokeratin 20 expression by non-invasive transitional cell carcinomas : potential for distinguishing recurrent from non-recurrent disease. *Histopathology* 27, 169-174.
- Harnden P., Eardley I., and Joyce A.D. (1996). Cytokeratin 20 as an objective marker of urothelial dysplasia. *British Journal of Urology* 78, 870-875.
- Harnden P., Mahmood N., and Southgate J. (1999). Expression of cytokeratin 20 redefines urothelial papillomas of the bladder. *The Lancet* 353, 974-977.
- Hartwell L.H. and Weinert T.A. (1989). Checkpoints: Controls that ensure the order of cell cycle events. *Science* 246, 629-634.
- Haynes B.F., Telen M.J., Hale L.P, and Denning S.M. (2001). CD44 - A molecule involved in leukocyte adherence and T-cell activation. *Immunology Today* 10, 423-428.
- Heicappell R., Wettig I.C., Schostak M., Muller M., Steiner U., Sauter T., and Miller K. (1999). Quantitative detection of the human complement factor H-related protein in transitional cell carcinoma of the urinary bladder. *European Urology* 35, 81-87.
- Hofmann M., Rudy W., Zöller M., Tölg C., Ponta H., Herrlich P., and Günthert U. (1996). CD44 splice variants confer metastatic behaviour in rats : homoogous

sequences are expressed in human tumour cell lines. *Cancer Research* 51, 5292-5297.

Hoglund M., Sall T., eim S., itelman F., M, dahl N., F, and l-Elmula I. (2002). Identification of cytogenetic subgroups and karyotypic pathways in transitional cell carcinoma. *Cancer Research* 61, 8241-8246.

Holmstrøm K., Rossen L., and Rasmussen O.F (1993). A highly sensitive and fast non-radioactive method for detection of polymerase chain reaction products. *Analytical Biochemistry* 209, 278-283.

Hughes J.H., Katz R.L., Rodriguez-Villanueva J., Kidd L., Dinney C., Grossman H.B., and Fritsche H.A. (1999). Urinary nuclear matrix protein 22 (NMP22) : A diagnostic adjunct to urine cytologic examination for the detection of recurrent transitional-cell carcinoma of the bladder. *Diagnostic Cytopathology* 20, 285-290.

Hughes J.H. and Cohen M.B. (1999). Nuclear matrix proteins and their potential applications to diagnostic pathology. *American Journal of Clinical Pathology* 111, 267-274.

Inoue T., Nakanishi H., Inada K-I, Hioki T., Tatematsu M., and Sugimura Y. Real time reverse transcription polymerase chain reaction of urinary cytokeratin 20 detects transitional cell carcinoma cells. *The Journal of Urology* 166, 2134-2141. 2002.
(GENERIC)

Ref Type: Generic

Ioachim E., Goussia A., and Agnantis N.J. (1999). Glycoprotein CD44 expression in Colorectal Neoplasms. *Virchows Archiv - An International Journal of Pathology* 434, 45-50.

- Irani J., Desgrandchamps F., Millet C., Toubert M-E., Bon D., Aubert J., and le Duc A. (1999). BTA *stat* and BTA TRAK : A comparative evaluation of urine testing for the diagnosis of transitional cell carcinoma of the bladder. *European Urology* 35, 89-92.
- Jalkanen S., Jalkanen M., Bargatze R, Tammi M., and Butcher E.C. (1987). Biochemical properties of glycoproteins involved in lymphocyte recognition of high endothelial venules in man. *Journal of Immunology* 141, 1615-1623.
- Johansson S.L. and Cohen S.M. (1997). Epidemiology and etiology of bladder cancer. *Seminars in Surgical Oncology* 13, 291-298.
- Kan M., Furukawa A., Aki M., Kanayama H., and Kagawa S. (1995). Expression of CD44 splice variants in bladder cancer. *International Journal of Urology* 2, 295-301.
- Kato I., Dnistrian A.M., Schwartz M., Toniolo P., Koenig K, Shore R.E., Akhmedkhanov A., Zeleniuch-Jacquotte A., and Riboli E. (1995). Serum folate, homocysteine and colorectal cancer risk in women : a nested case-control study. *British Journal of Cancer* 79, 1917-1921.
- King R.J.B. (2000). *Cancer Biology* (Harlow, Prentice Hall).
- Knowles M.A. (1999). Identification of novel bladder tumour supressor genes. *Electrophoresis* 20, 269-279.
- Knowles M.A. (1999). The genetics of transitional cell carcinoma : progress and potential clinical application. *BJU International* 84, 412-427.
- Koenig F., Jung K., Schnorr D., and Loening S.A. (2000). Urinary markers of malignancy. *Clinica Chimica Acta* 297, 191-205.

- Le Marchand L., Wilkens L.R., Kolonel L.N., Hankin J.H., and Lyu L-C. (1995). Associations of sedentary lifestyle, obesity, smoking, alcohol use and diabetes with the risk of colorectal cancer. *Cancer Research* 57, 4787-4794.
- Lesley J. and Hyman R. (1998). CD44 structure and function. *Frontiers in Bioscience* 3, d616-d630
- Levi F., Pasche C., La Vecchia C., Lucchini F., and Franceschi S. (1995). Food groups and colorectal cancer risk. *British Journal of Cancer* 79, 1283-1287.
- Lewin B. (2000). *Genes VII* (Oxford University Press).
- Leyh H., Marberger M., Conort P., Sternberg C., Pansadoro V., Pagano F., Bassi P., Boccon-Gibod L., Ravery V., Treiber U., and Ishak L. (1999). Comparison of the BTA *stat*TM test with voided urine cytology and bladder wash cytology in the diagnosis and monitoring of bladder cancer. *European Urology* 35, 52-56.
- Lipponen P., Aaltoma A., Kosma V.-M., Ala-Opas M., and Eskelinen M. (1998). Expression of CD44 standard and variant-v6 proteins in transitional cell bladder tumours and their relation to prognosis during a long-term follow-up. *Journal of Pathology* 186, 157-164.
- Lodish H., Baltimore D., Berk A., Zipursky S.L., Matsudaira P., and Darnell J. (1995). *Molecular Cell Biology* (Scientific American Books).
- Lokeshwar B.L. and Soloway M.S. (2001). Current bladder tumour tests : Does their projected utility fulfill clinical necessity ? *The Journal of Urology* 165, 1067-1077.
- Lynch H.T. and Smyrk T.C. (1995). Hereditary colorectal cancer. *Seminars in Oncology* 26, 478-484.

- Matsumura Y. and Tarin D (1992). Significance of CD44 gene-products for cancer-diagnosis and disease evaluation. *Lancet* 340 , 1053-1058.
- Matsumura Y., Hanbury D., Smith J., and Tarin D (1994). Non-invasive detection of malignancy by identification of unusual CD44 gene activity in exfoliated cancer cells. *British Medical Journal* 308, 619-624.
- McCabe R.P., Lamm D.L., Haspel M.V., Pomato N, Smith K.O., Thompson E., and Hanna M.G.Jr (2001). A diagnostic-prognostic test for bladder cancer using a monoclonal antibody-based enzyme-linked immunoassay for detection of urinary fibrin(ogen) degradation products. *Cancer Research* 44 , 5886-5893.
- Mian C., Pycha A., Wiener H., Haitel A., Lodde M., and Marberger M. (1999). Immunocyt : A new tool for detecting transitional cell cancer of the urinary tract. *The Journal of Urology* 161, 1486-1489.
- Midgley R. and Kerr D. (1999). Colorectal cancer. *Lancet* 353, 391-399.
- Miyanga N., Akaza H., Ishikawa S., Ohtani M., Noguchi R., Kawai K., Koiso K., Kobayashi M., Koyama A., and Takahashi T. (1997). Clinical evaluation of Nuclear Matrix Protein 22 (NMP22) in urine as a novel marker for urothelial cancer. *European Urology* 31, 163-168.
- Miyasaka M. (1996). Cancer metastasis and adhesion molecules. *Clinical Orthopaedics and Related Research* 312, 10-18.
- Moll R., Lowe A., Laufer J., and Franke W.W. (1992). Cytokeratin 20 in human carcinomas. *American Journal of Pathology* 140, 427-447.
- Muller M., Heicappell R., Habermann F., Kaufmann M., Steiner U., and Miller K. (1997). Expression of CD44V2 in transitional cell carcinoma of the urinary bladder and in urine. *Urology Research* 25, 187-192.

- Mydlo J.H., Kanter J.L., Kral J.G., and Macchia R.J. (1999). The role of obesity and diet in urological carcinogenesis. *British Journal of Urology International* 84, 225-234.
- Naor D., Sionov R.V., and Ish-Shalom D. (1997). CD44 : Structure, function and association with the malignant process. *Advances in Cancer Research* 71, 241-319.
- Navaratnam R.M., Chowaniec J., and Winslet M.C. (1999). The molecular biology of colorectal cancer development and the associated genetic events. *Annals of the Royal College of Surgeons of England* 81, 312-319.
- Neumayer R., Rosen H.R., Reiner A., Sebesta C., Schmid A., Tuchler H., and Schiessel R. (1999). CD44 expression in benign and malignant colorectal polyps. *Diseases of the Colon and Rectum* 42, 50-55.
- Ohyashiki K., Yahata N., Ohyashiki J.H., Iwama H., Hayashi S., Ando K., Aizawa T., Ito T., Miki M., and Ebihara Y. (1998). A combination of semiquantitative telomerase assay and in-cell telomerase activity measurement using exfoliated urothelial cells for the detection of urothelial neoplasia. *Cancer* 83, 2554-60.
- Okamoto I., Kawano Y., Tsuiki H., Sasaki J., Nakao M., Matsumoto M., Suga M., Ando M., Nakajima M., and Saya H. (1999). CD44 cleavage induced by a membrane-associated metalloprotease plays a critical role in tumour cell migration. *Oncogene* 18, 1435-1446.
- Onyango P. (2002). Genomics and cancer. *Current Opinions in Oncology* 14, 85
- Paez A., Coba J.M., Murillo N., Fernandez P., de la Cal M.A., Lujan M., and Berenguer A. (1999). Reliability of the routine cytological diagnosis in bladder cancer. *European Urology* 35, 228-232.

- Pardee A.B. (1989). G1 events and regulation of cell proliferation. *Science* 246, 603-608.
- Pennington S.R. and Dunn M. J. (2001). Proteomics from protein sequence to function (Bios Scientific publishers).
- Phillips A.V. and Cooper T.A. (2000). RNA processing in human disease. *Cellular and Molecular Life Sciences* 57, 235-249.
- Pietinen P., Malila N., Virtanen M., Hartman T.J., Tangrea J.A., Albanes D., and Virtamo J. (1995). Diet and risk of colorectal cancer in a cohort of Finnish men. *Cancer Causes and Control* 10, 387-396.
- Pode D., Shapiro A., Wald M., Nativ O., Laufer M., and Kaver I. (1999). Noninvasive detection of bladder cancer with the BTA Stat test. *The Journal of Urology* 161, 443-446.
- Quinn M., Babb P., Brock A., rby L., and nes J. Cancer trends in England and Wales 1950 - 1999. 2001. National Statistics, HMSO.
Ref Type: Report
- Ramakumar S., Bhuiyan J., Besse J.A., Roberts S.G., Wollan P.C., Blute M.L., and O'Kane D.J. (1999). Comparison of screening methods in the detection of bladder cancer. *The Journal of Urology* 161, 388-394.
- Raraty M.G.T. and Winstanley J.H.R. (1998). Variation in the staging of colorectal carcinomas : a survey of current practice. *Annals of the Royal College of Surgeons of England* 80, 188-191.
- Ross J.S., del Rosario A.D., Bui H.X., Kallakury B.V.S., Okby N.T., and Figge J. (1996). Expression of the CD44 cell adhesion molecule in urinary bladder transitional cell carcinoma. *Modern Pathology* 9, 854-860.

Rubin E. and Farber J.L. (1998). Pathology (

Saad A., Hanbury D., McNicholas T.A., Boustead G.B., and Woodman A.C. (2001). The early detection and diagnosis of bladder cancer : A critical review of the options. *European Urology* 39, 619-633.

Saad A., Hanbury D., McNicholas T.A., Boustead G.B., Morgan S., and Woodman A.C. (2002). A study comparing various noninvasive methods of detecting bladder cancer in urine. *BJU International* 89, 369-373.

Sanchez-Carbayo M., Herrero E., Megias J., Mira A., Espasa A., Chinchilla V., and Soria F. (1999). Initial evaluation of the diagnostic performance of the new urinary bladder cancer antigen test as a tumor marker for transitional cell carcinoma of the bladder. *The Journal of Urology* 161, 1110-1115.

Screaton G.R., Bell M.V., Jackson D.G., Cornelis F.B., Gerth U., and Bell J.I. (1992). Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. *Proceedings of the National Academy of Sciences* 89, 12160-12164.

Seregini E., Ferrari L., Martinetti A, and Bombardieri E. (2001). Diagnostic and prognostic tumour markers in the gastrointestinal tract. *Seminars in Surgical Oncology* 20, 147-166.

Silverman D., Rothman N., and Devesa S.A (1999). Epidemiology of bladder cancer. In *Bladder cancer : biology, diagnosis and management*. Syrigos K.N. and Skinner D.G., eds. (Oxford University Press), pp. 11-56.

Sobin L.H. (2001). TNM: Principles, history and relation to other prognostic factors. *Cancer (supplement)* 91, 1589-1592.

Sobin L. H. and Wittekind Ch (1997). *TNM : Classification of malignant tumours* (John Wiley and Sons, inc.).

Soloway M.S., Briggman J.V., Carpinito G.A., Chodak G.W., Church P.A., Lamm D.L., Lange P., Mesing E., Pasciak R.M., Reservitz G.B., Rukstalis D.B., Sarosdy M.F., Stadler W.M., Thiel R.P., and Hayden C.L. (1996). Use of a new tumour marker, urinary NMP22, in the detection of occult or rapidly recurring transitional cell carcinoma of the urinary tract following surgical treatment. *The Journal of Urology* 156, 363-367.

Sorahan T., Hamilton L., and Jackson J.R (2000). A further cohort study of workers employed at a factory manufacturing chemicals for the rubber industry, with special reference to the chemicals 2-mercaptobenzothiazole (MBT), aniline, phenyl- β -naphthylamine and *o*-toluidine. *Occupational and Environmental Medicine* 57, 106-115.

Srinivas P.R., Kramer B.S., and Srivastava S. (2001a). Trends in biomarker research for cancer detection. *Lancet Oncology* 2, 698-704.

Srinivas P.R., Srivastava S., Hanash S, and Wright G.L (2001b). Proteomics in early detection of cancer. *Clinical Chemistry* 47, 1901-1911.

Stickeler E., Kittrell F., Medina D, and Berget S.M. (1999). Stage specific changes in SR splicing factors and alternative splicing in mammary tumorigenesis. *Oncogene* 18, 3574-3582.

Sugino T, Gorham H., Yoshida K., Bolodeoku J., Nargund V., Cranston D., Goodison S., and Tarin D. (1996). Progressive loss of CD44 gene expression in invasive bladder cancer. *American Journal of Pathology* 149, 873-882.

Sugiyama M., Woodman A., Sugino T, Crowley S., Ho K., Smith J., Matsumara Y., and Tarin D. (1995). Non-invasive detection of bladder cancer by identification of

- Sugiyama M., Woodman A., Sugino T, Crowley S., Ho K., Smith J., Matsumara Y., and Tarin D. (1995). Non-invasive detection of bladder cancer by identification of abnormal CD44 proteins in exfoliated cancer cells in urine. *Journal of Clinical Pathology : Molecular Pathology* 48, M142-M147
- Sumi S., Arai K., Kithara S, and Yoshida K. (2000). Preliminary report of the clinical performance of a new urinary bladder cancer antigen test: comparison to voided urine in detection of transitional cell carcinoma of the bladder. *Clinica Chimica Acta* 296, 111-120.
- Tarin D., Bolodeoku J., Hatfill S.J., Sugino T, Woodman A.C., and Yoshida K. (1995). The clinical significance of malfunction of the CD44 locus in malignancy. *Journal of Neuro-Oncology*
- Tavani A., Fioretti F., Franceschi S., Gallus S., Negri E., Montella M., Conti E., and La Vecchia C. (1995). Education, socioeconomic status and risk of cancer of the colon and rectum. *International Journal of Epidemiology* 28, 380-385.
- Thomas G. (1995). Genetic predispositions to colorectal cancer. *Pathologie-Biologie* 43, 159-164.
- Thomas L., Leyh H., Marberger M., Bombardieri E., Bassi P., Pagano F., Pansadoro V., Sternberg C.N., Boccon-Gibod L., Ravery V., Le Guludec D., Meulemans A., Conort P., and Ishak L. (1999). Multicenter trial of the quantitative BTA TRAK assay in the detection of bladder cancer. *Clinical Chemistry* 45, 472-477.
- Tölg C., Hofmann M., Herrlich P., and Ponta H. (1993). Splicing choice from ten variant exons establishes CD44 variability. *Nucleic Acids Research* 21, 1225-1229.
- van der Meijden A.P.M. (1998). Bladder Cancer. *British Medical Journal* 317, 1366-1369.

- Wada T., Louhelainen J., Hemminki K., Adolfsson J., Wijkstrom H., Norming U., Borgstrom E., Hansson J., and Steineck G. (2001). The prevalence of loss of heterozygosity in chromosome 3, including *FHIT*, in bladder cancer, using the fluorescent multiplex polymerase chain reaction. *British Journal of Urology International* 87, 876-881.
- Wang Q., Desseigne F., Lasset C., Saurin J-C., Navarro C., Yagci T., Keser I., Bacgi H., Luleci G., Gelen T., Chayvialle J-A., Puisieux A., and Ozturk M. (1995). Germline *hMSH2* and *hMLH1* gene mutations in incomplete HNPCC families. *International Journal of Cancer* 73, 831-836.
- Watters A.D., Ballantyne S.A., Going J.J., Grigor K.M., and Bartlett J.M.S. (2000). Aneusomy of chromosomes 7 and 17 predicts the recurrence of transitional cell carcinoma of the urinary bladder. *British Journal of Urology International* 85, 42-47.
- Weinberg R.A. (1996). How cancer arises. *Scientific American* 275, 32-40.
- Wielenga V.J., Heider K.H., Offerhaus G.J, Adolf G.R., van den Berg F.M., Ponta H., Herrlich P., and Pals S.T. (1993). Expression of CD44 variant proteins in human colorectal cancer is related to tumour progression. *Cancer Research* 53, 4754-4756.
- Woodman A.C., Sugiyama M., Yoshida K., Sugino T, Borgya A., Goodison S., Matsumara Y., and Tarin D. (1996). Analysis of anomalous CD44 gene expression in human breast, bladder and colon cancer and correlation of observed mRNA and protein isoforms. *American Journal of Pathology* 149, 1519-1530.
- Woodman A.C., Goodison S., Drake M., Noble J., and Tarin D. (2000). Noninvasive diagnosis of bladder carcinoma by enzyme-linked immunosorbent assay detection of CD44 isoforms in exfoliated urothelia. *Clinical Cancer Research* 6, 2381-2392.

- Yamane N., Tsujitani S., Makino M., Maeta M., and Kaibara N. (1999). Soluble cd44 variant 6 as a prognostic indicator in patients with colorectal cancer. *Oncology* 56, 232-238.
- Yoshida K., Sugino T., Bolodeoku J., Warren B.F., Goodison S., Woodman A., Toge T., Tahara E., and Tarin D. (1996). Detection of exfoliated carcinoma cells in colonic luminal washings by identification of deranged patterns of expression of the CD44 gene. *Journal of Clinical Pathology* 49, 300-305.
- Yoshida K., sugino T., Tahara H., Woodman A., Bolodeoku J., Nargund V., Fellows G., Goodison S., Tahara E., and Tarin D. (1997). Telomerase activity in bladder carcinoma and its implication for noninvasive diagnosis by detection of exfoliated cancer cells in urine. *Cancer* 79, 362-369.
- Yu D.-S., Hsieh D.-S., and Chang S.-Y (2001). Detection of chromosomal alterations in bladder cancer by comparative genomic hybridization. *British Journal of Urology International* 87, 889-893.
- Zeegers M.P.A., Goldbohm R.A., and van den Brandt P.A. (2001). Consumption of vegetables and fruits and urothelial cancer incidence : a prospective study. *Cancer Epidemiology, Biomarkers and Prevention* 10, 1121-1128.

Internet References

National Screening committee (accessed 15/8/2002)

www.nsc.nhs.uk/uk_nsc/uk_nsc_ind.htm

NHS plan (accessed 15/8/2002)

www.nhs.uk/nationalplan/npch14.htm

Figure 1.1a (accessed 15/8/2002)

<http://departments.oxy.edu/.../11-27-00files/images018.jpg>

Figure 1.1b (accessed 15/8/2002)

www.entu.cas.cz/fyziol/students/mammals.jpg

Figure 1.1c (accessed 15/8/2002)

www.bioscience.org/2000/v5/d/somasound/fulltext.htm

Appendix A – Evidence of Ethical Approval and patient consent forms

NORTHAMPTON MEDICAL RESEARCH/ ETHICS COMMITTEE

Chairman: Mr Fred Evans

Secretary: Miss Michelle Skelton

Our Ref: FE/MS/99/19



17 May 1999

Dr Anthony Woodman
Lecturer in Cell Biology
Cranfield Biotechnology Centre
IBST
Cranfield University
MK43 0AL

Dear Dr Woodman

99/19 THE NON-INVASIVE DETECTION AND MONITORING OF MALIGNANCY USING MULTIPLE-MARKER MOLECULAR ASSAYS

The Northampton Medical Research/Ethics Committee reviewed your application in respect of the above study at their meeting on Thursday 13 May 1999, and I am pleased to inform you that Formal Ethical Approval has been granted.

To complete our records regarding the project, please complete and return the form accompanying this letter.

Please let me know if the study has to be terminated or any ethical considerations arise which need to be discussed further by the Committee.

Yours sincerely

FRED EVANS
Chairman, Northampton Medical Research/
Ethics Committee



Northamptonshire Health Authority, Highfield, Cliftonville Road,
Northampton NN1 5DN Tel: 01604 615000 Fax: 01604 615010
Chairman: Simon Schanschieff OBE, FCA, JP Chief Executive: Lynda Hamlyn BA(Hons)



Copy of Northampton General Hospital patient consent form

Patient information

Background

The staff of Northampton General Hospital and scientists at Cranfield University are working together to find new ways of providing advanced clinical support for the local community.

An area where we feel we can make major advances is in improving our ability to detect cancer. In particular we are working towards being able to detect cancer of the bladder and colon from either urine or faeces. Using advanced rapid tests we hope to be able to easily detect these cancers without the need to bring patients into hospital or undergo long, uncomfortable procedures.

To decide whether a new test will work, we have to know what results to expect. The only way we can achieve this is by looking at samples from as many people as possible.

Please help us

Your doctor has explained to you that he/she will be removing some bladder or bowel tissue during the course of your care. Specialist staff within the hospital will examine this tissue, and your care guided by what they find. After all these examinations have been undertaken, there may be a small amount of this tissue remaining. If you are willing to allow us, we can use this tissue to help develop the new assays.

Additional information

Only samples remaining after all clinical examinations have been completed can be donated for research.

Your care, treatment or method of assessment will not be altered in anyway be your decision to donate or otherwise. Furthermore no additional procedures will be undertaken as a result of your consent.

We require as many samples from people who have no cancer as those that do. Your doctor has already told you what he/she thinks is your diagnosis, and therefore this request does not mean that this diagnosis is unsure.

The scientists at Cranfield University will not know your identity, only your age, sex and what your doctor believes is your diagnosis. Your doctor will not know the results of the tests and therefore your treatment will not be affected or altered by them.

What happens to your samples

Your samples are the most valuable part of our research, without them we are unable to develop new ways of detecting cancer.

Any tissue donated will be frozen and transported to the laboratory. Here we look to see if we can find molecules in the tissue which seem to be more common in cancers. The more samples we collect, the more confident we become that if we see a certain molecule, that a cancer is present.

How do we obtain your consent

Your doctor will have spoken to you about our work. If you have any questions please ask him/her and they will be delighted to discuss these with you.

If you would like to help us and allow us to use some of your sample please sign the attached form.

NB You care will in no way be affected by you decision.

Thank you for you help and co-operation

.....
.....

CONFIDENTIAL

DONATION OF SAMPLES FOR RESEARCH

I have read the attached information and discussed it with my doctor.

Having been fully informed as to what is required:

I request that any tissue samples taken during the course of my care at Northampton General Hospital, may be made available for research after all procedures required for my clinical management have been completed.

I understand that my care will be unaffected by this consent

.....
.....

Name of patient

Hospital number

.....
.....

Signature

Date

Your doctor will retain this form.

Copy of consent form used for volunteers within Cranfield University

Information and Consent Form

Background

I am carrying out a project aimed at producing a rapid non-invasive diagnostic test for the detection of bladder cancer, using molecular markers. To produce a reliable test I need to use normal human samples to ensure the marker is specific to tumour tissue.

Through my link with Northampton Hospital I have been able to get samples of urine from patients with confirmed cancer and age matched patients without, but I need to sample urine from a healthy group of volunteers.

Please help

I require as many urine samples from healthy volunteers as possible – and this is where you come in.

What happens to the samples

Samples will be processed within the department and investigated for marker expression. They will be treated as anonymous, with the only information required being your age and sex.

Consent Forms

If you are happy to provide a sample then please fill in the consent form below. No results from these samples will be given to those who have taken part, and no further investigation will be carried out if any sample appears to be abnormal due to the stage of development which the test is at.

Thank you for your help and co-operation.

.....

CONFIDENTIAL

DONATION OF SAMPLES FOR RESEARCH

I have read the attatched information and discussed it with the researcher involved.

I understand that I will not be told any results gained from testing of any samples.

.....

Name of donor

.....
Signature

.....
Date

Appendix B – Solution Formulations

General Solutions

PBS PBS is made using tablets purchased from Sigma dissolved in 500ml of RO H₂O. PBS for cell culture was sterilised by autoclaving.

TBS This is made in a 20X solution
Trizma base
NaCl
Made to 1 litre with RO H₂O, pH 7.6

Protein Extractions

Protease Inhibitor This arrives in kit form with four separate constituents (EDTA, pepstatin, leupeptin and AEBSF). Each bottle is re-suspended in RO H₂O, and the four bottles mixed. The volume is then made up to 100ml with RO H₂O. The EDTA is not fully soluble, therefore will stay in suspension.

CHAPS lysis buffer 20mM Tris at pH 8.0
0.15M NaCl
5mM EDTA
310 mg of CHAPS
1 ml of protease inhibitor
Made to final volume of 50 ml with RO H₂O

IHC

DAB solution 30mg of DAB
50ml TBS
100µl of hydrogen peroxide

Western Blotting

6% Main Gel 2.5ml of 1.5M Tris
2ml of 30% Bis/Acrylamide
100µl of 10% SDS
50µl of 10% Ammonium Persulphate
10µl of TEMED
5.35ml of RO H₂O

4% Stacking gel 1.25ml of 0.5M Tris
650µl of 30% Bis/Acrylamide
50µl of 10% SDS
25µl of 10% Ammonium Persulphate
10µl of TEMED
3.05 ml of RO H₂O

Loading Buffer 5ml of 0.5M Tris
5ml of Glycerol
10ml of 10% SDS
5ml of 1% bromophenol blue
25ml of RO H₂O

Electrophoresis buffer (Made as a 10x solution)

30.3g of Tris
144g of Glycine
100g of SDS
RO H₂O top 1 litre

Transfer Buffer 2.905g of Tris
1.465g of glycine
0.5g of SDS
100ml of Methanol

RO H₂O to 1 litre
Make fresh each time required.

Bacterial Transformations

LB Broth	10g of Tryptone 10g of NaCl 5g of yeast extract RO H ₂ O to 1 litre
LB Agar	1.5% Agar in LB broth

Southern Blotting

SSC	Made up as a 20x solution Na ₃ Citrate 0.3M NaCl 3 M pH 7.0
Primary wash	4ml 10% SDS solution 25ml 20x SSC solution Made to 1 litre with RO H ₂ O
Secondary wash	2x SSC solution

Appendix C – Miscellaneous protocols

C1 Transformation of bacterial cells

Transformation was carried out using a strain of *E. Coli* (JM83) made competent, and were grown in LB broth and on LB containing agar plates. The plasmid contains an ampicillin resistant gene, therefore growth of bacteria was regulated by inclusion of ampicillin into the LB agar once they had undergone the transformation procedure. All agar and LB broth was autoclaved to ensure sterility before use. All growth incubations were carried out at 37°C unless stated otherwise.

C2 Production of competent cells

The *E.coli* were plated out using a wire loop onto LB agar plates and grown overnight at 37°C. One colony was removed from each plate and was placed into an aliquot of LB broth (5ml) before being placed back into the incubator for overnight growth. An aliquot (1 ml) of this culture was then placed into 100ml of LB broth and placed into a shaking incubator at 150 rpm. The culture was allowed to grow until the absorbance at 550nm reached 0.4 to 0.5 (using LB broth as blank). Once this point was reached the culture was placed on ice for 10 minutes, cells were then harvested by centrifuging 20ml of culture at 2000rpm for 5 minutes. The supernatant was discarded and the pellet re-suspended in cold MgCl₂ (10ml, 0.1M, sterile). The suspension was again spun at 2000rpm for 5 minutes and the supernatant discarded, and cells resuspended in cold CaCl₂ (1 ml, 0.1M, sterile) then kept on ice for two hours. To transform the cells, 200µl of competent cells was added to 5 µl of plasmid and these were incubated on ice for 30 minutes. Cells were then heat shocked at 42°C for two minutes and placed on ice immediately before adding 20µl of the mix to 180µl of LB broth, producing a 1/10 dilution. This was further diluted to 1/100 and 1/1000. 100µl of each dilution were then plated out using a glass spreader onto ampicillin containing LB agar plates.

Agar plates were incubated for 12-14 hours at 37°C at which point the number of colonies present on each plate was recorded, and a number of single colonies from both sets of cells were placed into 50 ml aliquots of LB broth containing ampicillin. These were grown for a further 14 hours, again at 37° in a shaking incubator before isolating the plasmid DNA using a maxiprep kit produced by Qiagen.

C3 Synopsis of Maxiprep plasmid extraction kit (Qiagen)

Bacteria in suspension are pelleted by centrifugation in a 50ml volume and the pellets lysed through use of an alkaline lysis buffer. The resulting lysate is cleared by filtration using and is then passed through a column to which the plasmid DNA binds. The column is washed several times to remove RNA, protein and other impurities and the DNA eluted and precipitated using isopropanol. A qiaprecipitator attached to a syringe captures the precipitated DNA and removes all traces of IPA before the purified DNA is eluted using water.

C4 Plasmid linearisation

Linearisation was achieved through cutting the plasmid at a specific restriction site and from the plasmids restriction map, the chosen site was that of *Hind III*. To digest the plasmid, the following reaction mix was used :

Nuclease Free water	6µl
10 x buffer	2µl
Acetylated BSA	2µl
Plasmid	8µl
Hind III enzyme	1µl

The mix was incubated at 37°C for two hours, then placed on ice. The digestion was analysed by agarose gel electrophoresis using a 0.6 % gel.

C5 Tissue sectioning

Cryostat chucks were first placed into the quick freezer bar within the cryostat to bring them down to a low temperature (approximately -20°C, all tubes and forceps required were also cooled down in this way). A small amount of water based cryomountant (OCT embedding fluid, Bright Instruments) was placed on the chuck

and the tissue sample placed on top. Once the cryomountant began to freeze (as demonstrated by the mountant changing from a clear liquid to a white solid), the temperature was brought down further by spraying with cryospray to ensure tissue and mountant was completely frozen. Once the surface of the tissue was ready for sectioning (having taken some 10µm sections to even up the face of the tissue), 80 – 100 7µm sections was taken of each sample and were placed into a 2ml γ-irradiated tube. Sections that were not being used immediately for extraction were snap frozen and stored at –80 °C until use. During the course of sectioning, four sections were placed on poly-l-lysine coated slides (coated using poly-l-lysine solution, Sigma, according to manufacturers instructions). One section was taken at the beginning, one at the end, and two inbetween. These sections were fixed using cold methanol (incubated for 10 minutes at 4°C, then dried at 4°C) and were then stained with haematoxylin and eosin to clarify the histology of sections taken.

C6 De-waxing, re-hydration and de-hydration of paraffin embedded tissue sections

De-waxing

Slides are incubated for 1 minute in each of :

2 x Histoclear

2 x 100% EtOH

1 x 90% EtOH

1 x 70% EtOH

Slides were then rinsed in RO water and stored in fresh RO water until required.

Dehydration of slides for mounting

Slides are incubated for 1 minute in each of :

1 x 70% EtOH

1 x 90% EtOH

2 x 100% EtOH

2 x Histoclear

Slides are then ready for mounting.