Thesis Submitted to the University of London in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy

A MOLECULAR GENETIC ANALYSIS OF THE DESMOSOMAL CADHERINS.

By

Mark D. Marsden

National Institute for Medical Research,
Mill Hill, London NW7 1AA, UK.

April 1995
For Eileen and Eva
The desmocollins and the desmogleins belong to the cadherin superfamily of adhesion molecules and are present in the desmosome, a specialised intercellular junction characteristic of epithelium. Three types of desmocollin and three types of desmoglein have been identified and in desmosome-bearing tissues such as epidermis these are differentially expressed. Both DSC1 and DSC2 encode desmocollin isoforms. Whilst the expression of DSC1 is predominantly suprabasal in epidermis, DSC2 is expressed in all desmosome-bearing tissues. To initiate an investigation into the mechanisms that regulate the differential expression of these genes, human genomic libraries were screened to isolate 5' DNA regulatory elements. λ clones containing the DSC2 promoter were identified by screening a human genomic library by hybridisation to the most 5' 197bp of the human DSC2 cDNA. Analysis of such a 5' genomic clone by RNase protection experiments revealed a major transcription initiation site 201bp upstream of the translation start site. Analysis of the 1.9kb DNA sequence upstream of the DSC2 translation start site revealed consensus binding sites for transcription factors associated with the epidermis and motifs common to the promoters of other epidermally expressed genes. Southern analysis confirmed the presence of a CpG island by the presence of clustered restriction sites for the methylation-sensitive endonucleases BssHII and EagI. The activity of the human DSC2 promoter, was examined by coupling it to the luciferase reporter gene. Transient transfection of this reporter construct and also deletions derived from it, into epithelial and non-epithelial cell lines enabled activating and inhibitory DNA elements to be mapped, and identified cell-specific elements between -332 and -525 (relative to translation start site at 0). Other work undertaken as preliminary to gene knockout experiments has been the isolation of a mouse genomic clone for DSC2, and the sequence analysis of a previously isolated mouse desmoglein cDNA.
CONTENTS

TITLE PAGE............................................................................................................. 1
ABSTRACT............................................................................................................... 3
CONTENTS............................................................................................................... 4
LIST OF FIGURES AND TABLES........................................................................ 9
ABBREVIATIONS................................................................................................... 11

CHAPTER ONE - Introduction.
1.1 Introduction............................................................................................ 14
1.2 Intercellular Junctions............................................................................ 15
1.2.1i Desmosomes (Macula Adhaerens, Spot Desmosome)....................... 15
1.2.1ii Molecules of the Desmosome............................................................. 18
Desmosomal Cadherins................................................................................... 18
Plakoglobin....................................................................................................... 18
Desmoplakin, a Putative Link to the Intermediate Filaments.............. 19
Other Components of Desmosomes......................................................... 20
1.2.2 The Adhaerens Junction (Zonula Adhaerens, Belt Desmosome)..... 21
1.2.3 Gap Junctions, Tight Junctions and Hemidesmosomes................... 24
1.3 Structural Organisation of the Desmosomal Cadherins...................... 25
Primary Protein Structure of the Cadherins........................................... 25
Cadherins of the Desmosome; Desmocollins and Desmogleins...... 28
1.4 Molecular Dissection of Desmosomal Cadherin Function............... 29
1.4.1 Intercellular Adhesion............................................................................ 29
Classical Cadherin-Mediated Adhesion......................................................... 29
Desmosomal Cadherin-Mediated Adhesion................................................. 30
1.4.2 Cytoplasmic Interactions........................................................................ 33
Desmosomal Cadherins and the Plaque....................................................... 33
Associations of Plakoglobin................................................................. 34
1.5 Expression Patterns of the Desmosomal Cadherins......................... 34
A Brief Introduction to Keratinisation....................................................... 34
Desmosomal Cadherin Expression in Epidermis......................................... 35

4
CHAPTER TWO - Molecular Cloning and Characterisation of the Human DSC2 Promoter.

2.1 Introduction ............................................................................................ 47
2.2 Results .................................................................................................... 48
2.2.1 Isolation of Genomic Clones ................................................................. 48
2.2.2 Characterisation of the Genomic Clones, λ2.DSC2 and λ4.DSC2.. 51
2.2.3 Southern Analysis of the DSC2 Locus ................................................ 56
2.2.4 Mapping the Transcription Initiation Site ............................................ 59
2.2.5 Primary Structure of the Promoter......................................................... 59
2.2.6 Identification of a CpG Island .............................................................. 66
2.2.7 Human DSC2 Promoter Activity in Cultured Cells ........................... 67
2.2.7.1 Construction of Reporter Constructs ................................................... 67
2.2.7.2 Experimental Controls ...................................................................... 70
2.2.8 Transfection of DSC2 Reporter Constructs into MDCK Cells ....... 75
2.2.9 DSC2 Promoter Activity is Unidirectional............................................. 75
2.2.10 To Test for Enhancer Elements within the Human DSC2 Promoter 78
2.2.11 Transfection of DSC2 Reporter Constructs into Other Cell Lines... 78
2.2.12 DSC2 Reporter Constructs to Study Regulatory Elements In Vivo. 79
2.3 Discussion .............................................................................................. 82
2.3.1 Site of Transcription Initiation.............................................................. 82
2.3.2 Potential Regulatory Transcription Factors .......................................... 83
2.3.3 Activity of the DSC2 Promoter in Transient Transfection Experiments ....................................................................................... 86
2.4 Major Findings ...................................................................................... 89
2.5 Future Work - Identify Proteins that Interact with the Promoter..... 90
6.2 Generation of DNA Recombinant Molecules ...................................... 123
6.2.1 Restriction of DNA with Endonucleases ................................................123
6.2.2 Agarose Gel Electrophoresis ...................................................................123
6.2.3 Extraction of DNA from Agarose Gels ................................................ 124
6.2.4 Alkaline Phosphatase Treatment of Vectors ...........................................124
6.2.5 Ethanol Precipitation of Nucleic Acids.................................................. 125
6.2.6 Filling-in 3’ Recessed Ends ................................................................... 125
6.2.7 Ligation of DNA Fragments into Plasmid Vectors ............................. 125
6.3 Transformation of Competent *Escherichia coli.* .................................... 125
6.3.1 Preparation of Competent Cells...............................................................125
6.3.2 Use of Competent Cells ........................................................................ 126
6.4 Isolation of Plasmid DNA .......................................................................126
6.4.1 Small Scale - Alkaline Lysis ...................................................................126
6.4.2 Large Scale - Qiagen Preparation ...........................................................127
6.4.3 Large Scale - CsCl Banding ................................................................. 127
6.5 Isolation of Human Genomic DNA from Whole Blood ....................... 128
6.6 Estimation of DNA and RNA Concentrations ..................................... 128
6.7 Screening a λFix II Library .......................................................................129
6.7.1 Preparation of Plating Cultures ...............................................................129
6.7.2 Plating out a λFix II Library ................................................................ 129
6.7.3 Replica Plating a λFix II Library ........................................................... 130
6.8 Making High Titre Stocks of λFix II Recombinant Phage................ 130
6.9 Isolation of λ DNA ....................................................................................131
6.10 Southern Blotting .....................................................................................131
6.11 Probing Nylon Membranes .....................................................................132
6.11.1 Hybridisation of Labelled Probes to Nylon Membranes .................... 132
6.11.2 Washing Nylon Membranes .................................................................132
6.11.3 Random Primed Labelling of DNA ..................................................... 132
6.12.1 DNA Sequencing .....................................................................................133
6.12.2 Preparation of Plasmid DNA for Sequencing ...................................... 133
6.13 Polymerase Chain Reaction ...................................................................134
6.14 Vectorette Polymerase Chain Reaction ............................................... 134
FIGURES

CHAPTER ONE - Introduction.

Figure 1.1 Electronmicrograph and Schematic Diagram of a Desmosome.... 17
Figure 1.2 Epithelial Intercellular Junctions....................................................... 23
Figure 1.3 Structural Organisation of the Desmosomal Cadherins........... 27
Figure 1.4 Differential Expression of the Desmosomal Cadherins in Human
Epidermis........................................................................................... 38

CHAPTER TWO - Molecular Cloning and Characterisation of the
Human DSC2 Promoter.

Figure 2.1 Schematic Representation of the Human DSC2a cDNA............. 50
Figure 2.2 Restriction Map of pMDM4, Southern Blot and Restriction
Map of λ2.DSC2.......................................................... 53
Figure 2.3 Southern Blot of Human Genomic DNA and λ2.DSC2 DNA.... 55
Figure 2.4 Southern Analysis of the DSC2 Locus........................................ 58
Figure 2.5 Mapping the Transcription Initiation Site of DSC2...................... 61
Figure 2.6 Nucleotide Sequence of the Human DSC2 Promoter................. 63
Figure 2.7 Southern Blot Analysis of the Human DSC2 Promoter
Revealing a CpG Island............................................................... 69
Figure 2.8 Standard Curve for Luciferase Activity...................................... 74
Figure 2.9 Construction of Human DSC2 Reporter Constructs............... 72
Figure 2.10 Activity of Reporter Constructs in Cell Lines Relative to
pGL2-Control................................................................. 77
Figure 2.11 Activity of Reporter Constructs in Cell Lines Relative to
pGL2-Basic................................................................. 81

CHAPTER THREE - Screening for the DSC1 Promoter.

Figure 3.1 Schematic Representation of the Human DSC1a cDNA............. 95
Figure 3.2 Primary Screen of the ICI YAC Library by PCR...................... 99
CHAPTER FOUR - Isolation and Characterisation of Murine Desmocollin and Desmoglein Clones.

Figure 4.1 Southern Analysis of the Murine Genomic DSC2 Clone............ 105
Figure 4.2 Nucleotide Sequence of the Murine DSG1 Clone Compare with other Desmoglein cDNAs.............................. 107

CHAPTER FIVE - General Discussion.

Figure 5.1 The Human Desmosomal Cadherin Locus............................. 115

CHAPTER SIX - Materials and Methods.

Figure 6.1 RNase Protection Assay on the DSC2 Promoter.................... 142

TABLES

Table 2.I DSC2 Promoter Elements Common to Other Epithelial Promoters................................................................. 65
Table 2.II β-Galactosidase Activity of MDCK Cells................................ 70
Table 2.III Activity of Enhancer Constructs in MDCK cells.................... 78
Table 4.I Comparison of the Murine DSG cDNA with other DSG cDNAs. 108
# LIST OF ABBREVIATIONS AND SYMBOLS

(in alphabetical order)

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>adenosine 5'-triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>cAMP</td>
<td>cyclic adenosine 5'-monophosphate</td>
</tr>
<tr>
<td>CAT</td>
<td>chloramphenicol acetyltransferase</td>
</tr>
<tr>
<td>CBP</td>
<td>cAMP binding protein</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CoA</td>
<td>coenzyme A</td>
</tr>
<tr>
<td>cpm</td>
<td>counts per minute</td>
</tr>
<tr>
<td>CPRG</td>
<td>chlorophenol red β-D-galactopyranoside</td>
</tr>
<tr>
<td>CRBP</td>
<td>cAMP responsive element binding protein</td>
</tr>
<tr>
<td>cRNA</td>
<td>complementary RNA</td>
</tr>
<tr>
<td>CTP</td>
<td>cytosine 5'-triphosphate</td>
</tr>
<tr>
<td>CMV</td>
<td>cytomegalovirus</td>
</tr>
<tr>
<td>dATP</td>
<td>2'-deoxyadenosine 5'-triphosphate</td>
</tr>
<tr>
<td>dCTP</td>
<td>2'-deoxycytosine 5'-triphosphate</td>
</tr>
<tr>
<td>DEPC</td>
<td>diethyl pyrocarbonate</td>
</tr>
<tr>
<td>dGTP</td>
<td>2'-deoxyguanosine 5'-triphosphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribose nucleic acid</td>
</tr>
<tr>
<td>DSC</td>
<td>desmocollin</td>
</tr>
<tr>
<td>DSG</td>
<td>desmoglein</td>
</tr>
<tr>
<td>DSP</td>
<td>desmoplakin</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>dTTP</td>
<td>2'-deoxythymidine 5'-triphosphate</td>
</tr>
<tr>
<td>EB</td>
<td>erythroid body</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EGTA</td>
<td>ethyleneglycol-bis(oxy-ethylenenitrilo)tetraacetic acid</td>
</tr>
<tr>
<td>ES</td>
<td>embryonic stem [cell]</td>
</tr>
</tbody>
</table>
EtBr ethidium bromide
GTP guanosine 5'-triphosphate
HPV human papilloma virus
HSV-tk herpes simplex-thymidine kinase
kb kilobases
kDa kilo Dalton
L-broth Luria broth
Mb megabases
PBS phosphate buffered saline
PBS-A phosphate buffered saline type A
PCR polymerase chain reaction
PEG polyethylene glycol
RACE rapid amplification of cDNA ends
RNA ribonucleic acid
SDS sodium dodecyl sulphate
SSC sodium chloride / sodium citrate solution
TE Tris buffered -EDTA solution
tRNA transfer ribonucleic acid
UTP uridine 5'-triphosphate
X-gal 5-bromo-4-chloro-3-indoyl β-D-galactoside
YAC yeast artificial chromosome
Chapter One - Introduction

1.1 Introduction
1.2 Intercellular Junctions
1.2.1i Desmosomes (Macula Adhaerens, Spot Desmosome)
1.2.1ii Molecules of the Desmosome
   Desmosomal Cadherins
   Plakoglobin
   Desmoplakin, a Putative Link to the Intermediate Filaments
   Other Components of Desmosomes
1.2.2 The Adhaerens Junction (Zonula Adhaerens, Belt Desmosome)
1.2.3 Gap Junctions, Tight Junctions and Hemidesmosomes
1.3 Structural Organisation of the Desmosomal Cadherins
   Primary Protein Structure of the Cadherins
   Cadherins of the Desmosome; Desmocollins and Desmogleins
1.4 Molecular Dissection of Desmosomal Cadherin Function
1.4.1 Intercellular Adhesion
   Classical Cadherin-Mediated Adhesion
   Desmosomal Cadherin-Mediated Adhesion
1.4.2 Cytoplasmic Interactions
   Desmosomal Cadherins and the Plaque
   Associations of Plakoglobin
1.5 Expression Patterns of the Desmosomal Cadherins
   A Brief Introduction to Keratinisation
   Desmosomal Cadherin Expression in Epidermis
   Tissue Distribution of Desmosomal Cadherin Isoforms
   DSC2 Expression in Developing Epithelium; Trophoderm
1.6 Structure of Cadherin Genes
   Location of Cadherin Genes
1.7 Desmosomal Cadherins and Disease
1.8 Focus of this Work
1: INTRODUCTION

1.1 Introduction.
The desmocollins and desmogleins belong to the cadherin superfamily of adhesion proteins and are present in the desmosome, a specialised intercellular junction characteristic of epithelium. Three types of desmocollin and three types of desmoglein have been identified and in desmosome-bearing tissues such as epidermis, these are differentially expressed. The regulated expression of cell adhesion molecules is important in a diverse array of biological phenomena such as cell condensation at compaction, keratinisation, neurite outgrowth, cell sorting and metastasis. Thus it is of fundamental interest to elucidate the mechanisms that control the specific expression of these molecules.

Subsequent to the recent cloning of the desmosomal cadherin cDNAs, their primary structures have been determined, biological function tested, expression patterns probed and their chromosomal loci mapped. Two pieces of evidence suggest that the differential expression of the desmosomal cadherins is under transcriptional controls; in the developing epithelium of the murine preimplantation embryo, the trophectoderm, expression of the desmocollin DSC2 mRNA is detected immediately prior to the protein (Collins et al., 1995). Secondly, in human epidermis expression of the desmosomal cadherin mRNAs is detected within cell layers where the proteins are initially detected (I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data). The mechanisms and factors which generate the differential expression of the desmosomal cadherins through transcriptional controls are currently unknown.

In this introduction a description will be given of the primary structure of the desmocollins and desmogleins which identifies them with the cadherin family of calcium dependent cell adhesion molecules and also recent experiments which have sought to probe their function as adhesive molecules and their interactions with other desmosomal proteins. In addition, experiments will be described, primarily on human epidermis, which demonstrate the differential expression of desmosomal cadherin isoforms. Initially,
however, a description will be given of the unique cellular context of the desmosomal cadherins, the desmosome.

1.2 Intercellular Junctions.

1.2.1i Desmosomes (Macula Adhaerens, Spot Desmosome).

The desmosome is a punctate junction which mediates strong intercellular adhesion and links to the intermediate filament networks of cells. Its distinctive disc-like morphology consists of two electron-dense submembraneous plaques in close apposition between adjacent cells, and an intercellular space also known as the desmoglea (figure 1.1). The desmoglea is highly organised with an almost crystalline structure and characterised by a dense midline which runs parallel to the plasma membrane. Intermediate filaments do not appear to attach directly to the plaque but to terminate or loop back into the cytoplasm at 40-70nm from its inner face. The intervening region is referred to as the satellite zone and appears to be bridged by plaque peripheral filaments. By binding cells together and linking their intermediate filament networks, desmosomes create an extended three dimensional lattice within a cell sheet which confers a tensile strength to tissues and maintains tissue integrity.

The first desmosomes appear in early development between trophectoderm cells (Fleming et al., 1991) and are found in virtually all vertebrate epithelia, the arachnoid and pia meninges (Parrish et al., 1986), cardiac muscle (Cowin et al., 1984a), and the follicular dendritic cells of the lymphoid system (Garrod et al., 1990). Such a tissue distribution reflects a role of the desmosome-intermediate filament association in maintaining integrity where tissues are subjected to mechanical stress. Indeed, as strong sites of intercellular attachment desmosomes are unusually resistant to chemical and mechanical stress; even 8M urea fails to dissociate the adhesive structures during the isolation of desmosomes from bovine muzzle. The identification of desmosome-associated molecules has enabled the structure of this unique intercellular junction to be resolved at the molecular level.
Figure 1.1: Electronmicrograph and Schematic Diagram of a Desmosome.

Top: Electronmicrograph of a transverse section through a desmosome. A dark extracellular midline can be seen within the desmoglea between the two electron-dense submembraneous plaques of the desmosome. A dense material extending into each of the cells from the plaques forms the satelite zones. Intermediate filments are not visible on this micrograph (After Fawcett, 1981). Bottom: Schematic diagram of a desmosome showing the location of the major constituents. (After Garrod, 1993)
1.2.1ii Molecules of the Desmosome.

Desmosomal Cadherins.

The desmosomal cadherins are the molecules which are believed to mediate the strong intercellular adhesion of the desmosome and are the subject of this work. Three genotypically distinct desmogleins are known, Dsg1, Dsg2 and Dsg3 which are encoded by the genes DSG1, DSG2 and DSG3, respectively (Koch et al., 1990; Goodwin et al., 1990; Wheeler et al., 1991a; Nilles et al., 1991; Koch et al., 1991a). cDNAs of all three isoforms have been cloned from human as well as bovine sources; a murine DSG1 cDNA has also been isolated (Buxton et al., 1994). Likewise, three types of human and bovine desmocollin homologues have been cloned; the proteins are referred to as Dsc1, Dsc2 and Dsc3, and the respective genes as DSC1, DSC2 and DSC3 (Collins et al., 1991; Koch et al., 1992; Koch et al., 1991b; Mechanic et al., 1991; Parker et al., 1991; Legan et al., 1994; Kawamura et al., 1994; King, 1994; I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data). A murine DSC2 cDNA has also been cloned (Buxton et al., 1994). The nomenclature used here differs from that of Buxton et al. (1993), since the human desmocollin formerly designated DSC3 (Parker et al., 1991) has a greater sequence identity to bovine DSC2 (Koch et al., 1992) than to human DSC3 (I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data) and bovine DSC3 (Legan et al., 1994). Thus the human desmocollin cloned by Parker et al. (1991) is most probably the species homologue of the bovine desmocollin cloned by Koch et al. (1992), so it is now renamed DSC2.

Plakoglobin.

Although most of the components of the desmosome have probably been characterised and cloned, only one interaction has been resolved between the proteins of the plaque and the transmembrane desmosomal cadherins, that is the binding of plakoglobin with the desmogleins Dsg1 and Dsg3 demonstrated by coimmunoprecipitation and in vitro binding studies (Korman et al., 1989; Mathur et al., 1994). Plakoglobin does not localise just to the desmosome but also to another intercellular junction where cadherins mediate adhesion, the adhaerens junction and the cytosol (Piepenhagen and Nelson, 1993; Gumbiner and McCrea, 1993; Cowin et al., 1986). Plakoglobin may act as an adaptor
molecule between the desmosomal cadherins and the intermediate filament network via proteins such as desmoplakins.

Desmoplakin, a Putative Link to the Intermediate Filaments.

Desmoplakin I and II (DspI and DspII) are unique to the desmosome and the novel syndesmos junction (Complexus adhaerentes of certain lymphatic endothelia, Schmelz and Franke, 1993). Along with plakoglobin these are the major constituents of the desmosomal plaque. Evidence suggests that DspI and DspII may be splice variants of a single gene (DSP) located upon human chromosome 6 (Green et al., 1990; Virata et al., 1992; Arnemann et al., 1991). DspI and DspII share N- and C-terminal domains, but DspI contains a 1797bp insertion which encodes a 130nm coiled coil rod domain. Both splice variants are found in desmosome-bearing tissues, although at widely variable ratios (Virata et al., 1992). As visualised from electron micrographs of rotary shadowed DspI, the molecule appears as a dimer with a rod-like structure with N and C-terminal globular domains (O'Keefe et al., 1989). Transfection of a cDNA encoding the C-terminal domain tagged with the c-myc epitope into desmosome-bearing HeLa cells results in the fusion protein aligning along the intermediate filament network, while a fusion protein containing the N-terminal domain localises to the desmosomal plaque (Stappenbeck et al., 1993). In vitro assays also demonstrate a tight interaction between the DspI C-terminal domain and epidermal type II keratins (Kouklis et al., 1994). Thus the desmoplakins appear to be molecules which may function as adaptors between the intermediate filament network and the desmosomal plaque. Such a role is consistent with antibody studies which localise desmoplakin to the desmosome’s satellite zone (Miller et al., 1987; Jones and Grelling, 1989).

Desmosomes interact with several different types of intermediate filament networks. In epithelia these are predominantly composed of keratin, in cardiac muscle, desmin, while in certain arachnoidal and meningioma cells the desmosome-anchored intermediate filaments contain vimentin. The nature of the desmoplakin-intermediate filament interaction to these different networks is not well understood. In vivo the DspI C-terminal domain associates with keratin and vimentin networks (Stappenbeck et al., 1993), but Kouklis et al. (1994) could only demonstrate desmoplakin binding to type II
epidermal keratins in an *in vitro* binding assay. Moreover, a DspI mutant with 68 residues deleted from the C-terminus transfected into fibroblasts could still coalign with vimentin, but not with keratin in simple epithelial cells (Stappenbeck *et al.*, 1993). Thus desmoplakin-intermediate filament interactions probably differ between different types of network and may require additional proteins.

**Other Components of Desmosomes.**

While the desmosomal cadherins, plakoglobin, desmoplakin and a 680kDa protein which immunolocalises only to the circumference of the plaque known as desmoyokin (Heida *et al.*, 1989), appear to be obligatory components of the desmosome, a number of plaque proteins restricted to desmosomes of stratified epithelia have been reported. Proteins which bind to intermediate filament networks include keratocalmin, also known as desmocalmin, a 240-250kDa protein which binds calmodulin and keratin filaments (Fairley *et al.*, 1991; Cowin *et al.*, 1986), a 140kDa lamin B-like protein which binds vimentin (Cartaud *et al.*, 1990), and a protein of 35kDa which binds desmoplakin and keratins (Chiu and O'Keefe, 1991). Band 6 protein (B6P) or plakophilin 1 can interact directly with intermediate filaments and most interestingly has a degree of sequence identity with plakoglobin (Hatzfeld *et al.*, 1994; Heid *et al.*, 1994). Other proteins include desmonectin which is expressed within the granular layer of epidermis (Zhou and Chaplin, 1993) and a 140kDa protein identified in ultrastructural studies by a monoclonal antibody (Ouyang and Sugrue, 1992). IFAP300 and plectin have both been immunolocalised to the cytoplasmic plaques of the desmosome and the distinct hemidesmosome and may link intermediate filaments to these two very different junctions (Skalli *et al.*, 1994; Wiche *et al.*, 1993).

Abnormalities of the desmosome-keratin complex may have a role in disease. Patients with the autosomal dominant Darier's disease have epidermal lesions where the focal region of cell separation is within the suprabasal layers of epidermis. Histological studies suggest that this disease may be caused by abnormalities in the desmosome-keratin complex (Caulfield and Wilgram, 1963). Furthermore, Darier's disease is linked to human chromosome 12q, significantly at a site distinct and distal to the keratin genes which also map at 12q (Craddock *et al.*, 1993; Bashir *et al.*, 1993; Parfitt *et al.*, 1994).
Thus the abnormalities which give rise to this disease might be a component of the desmosome. Patients with Hailey-Hailey disease, which maps to human chromosome 3q (Ikeda et al., 1994), have similar symptoms to those with Darier’s disease (Burge, 1992) which makes this another candidate where the desmosome is implicated in the disease process.

1.2.2 The Adhaerens Junction (Zonula Adhaerens, Belt Desmosome).

The desmosome and adhaerens junction are the two major types of adhesive intercellular junction in epithelial cells. In columnar epithelia, the adhaerens junction encircles the apical region of the cells and forms part of a greater junctional complex flanked by the more apical tight junction and more basal desmosomes and gap junctions (figure 1.2). Like desmosomes, adhaerens junctions contain cadherins as the transmembrane adhesive component such as E- or P-cadherin. Plakoglobin localises to the adhaerens submembraneous plaque and binds to E-cadherin in vitro (Cowin et al., 1986; Hülsken et al., 1994). Whereas the desmosome is linked to the intermediate filament network, adhaerens junctions link to the contractile actin filament network (Tsukita et al., 1992), perhaps to fulfil a function in the coordination of contractile activity among groups of cells. Such a role is illustrated by compaction in the murine preimplantation embryo, where adherent cells flatten extensively against one another (Fleming and Johnson, 1988). The catenins are proteins which coimmunoprecipitate with cadherins of the adhaerens junction; three types of catenin are known, α, β and γ. α-catenin may interact with actin microfilaments and shows some homology to vinculin, a protein known to associate with actin filaments at their sites of attachment to the plasma membrane in focal contacts (Nagafuchi et al., 1991; Hirano et al., 1992). β catenin is the vertebrate homologue of the Drosophila segment polarity gene armadillo which is also related to plakoglobin (Gumbiner and McCrea, 1993). γ-catenins have been identified with plakoglobin, since two dimensional gel electrophoresis resolves the γ-catenin protein into two protein bands, one of which is plakoglobin based upon immunoreactivity, apparent molecular mass and isoelectric point (Piepenhagen and Nelson, 1993).
Figure 1.2: Epithelial Intercellular Junctions. Schematic diagram showing the major types of intercellular junctions in a typical cell of a simple polarised epithelium. (After Garrod, 1986).
Apical Surface Microvilli

Tight Junction

Zonula Adhaerens

Desmosome

Gap Junction

Non-junctional Membrane Adhesion

Basal Lamina Hemidesmosome
1.2.3 Gap Junctions, Tight Junctions and Hemidesmosomes.

Gap junctions provide physical channels of communication that link the cytoplasm of adjacent cells (Bennett et al., 1991). This is functionally significant in the coordinated contraction of tissues such as cardiac muscle and the uterine wall. Connexin proteins in these junctions make up the intercellular transmembrane channels that link two adjacent cells together. Tight junctions (zonula occludens) are characteristic of polarised epithelial cells and have the appearance of two adjacent plasma membranes being pinched firmly together. In contrast to gap junctions they restrict free-movement of substances creating a permeability barrier between adjacent cells (Citi, 1993). Tight junctions also create a boundary in the plasma membrane bilayer that separates the cell surface into biochemically and functionally distinct apical and basolateral membrane domains, enabling the cells to carry out polarised transport. Thus it is only after their formation that the developing epithelium of the murine preimplantation embryo becomes polarised, with the Na/K-ATPase segregating to the basolateral plasma membrane (Watson and Kidder, 1988). The polarity that the tight junction defines can also be seen by the fact that other types of intercellular junctions are basally located with respect to the tight junction (figure 1.2).

Hemidesmosomes belie their name; although their punctate appearance is similar to half a true desmosome they neither partake in intercellular adhesion nor share any common major components with the desmosome, except perhaps for the recently reported IFAP300 protein and plectin (Skalli et al., 1994). Hemidesmosomes are found in stratified and transitional epithelia where they connect the basal face of cells to the basement membrane and link to the intermediate filament network (Legan et al., 1992). A protein which localises to the hemidesmosomal plaque, known as the bullous pemphigoid antigen I, has some homology with desmoplakin and its gene is also located upon human chromosome 6 (Sawamura et al., 1990; Arnemann et al., 1991). The bullous pemphigoid antigen I may mediate binding of the intermediate filament network to the hemidesmosomal plaque (Tanaka et al., 1991). Another protein plectin, which is associated with intermediate filaments is another member of this family of proteins (Green et al., 1992). The principal molecule of adhesion in the hemidesmosome is the integrin α6β4 (Stepp et al., 1990). Thus although desmosomes and hemidesmosomes
appear similar under the electron microscope, molecular analysis has revealed that they contain different proteins fulfilling analogous functions.

1.3 Structural Organisation of the Desmosomal Cadherins.

Primary Protein Structure of the Cadherins.

The desmocollins and the desmogleins form two distinct subfamilies of the cadherin family of calcium dependent cell adhesion molecules. Cadherins are a rapidly expanding group that are typified by a core of classical cadherins represented by N-, E- and P-cadherin; other members include such diverse species as the c-ret tyrosine kinase and the Drosophila fat gene product (Schneider, 1992; Mahoney et al., 1991). Classical cadherins are type I transmembrane glycoproteins with an extracellular domain composed of four unique cadherin repeats each of approximately 110 residues and a domain proximal to the plasma membrane (figure 1.3). Fat contains 34 cadherin repeats, the function of which are unknown. Intercellular adhesion mediated by these molecules is predominantly homophilic and may be important in the development and differentiation of tissues through the formation of selective interactions (Takeichi, 1991). Structural data obtained by nuclear magnetic resonance spectroscopy (NMR) and X-ray crystallography for the N-terminal ectodomain of E- and N-cadherin, respectively, show the calcium binding site in the region linking adjacent ectodomains (Overduin et al., 1995; Shapiro et al., 1995). These data demonstrate that although calcium is necessary for adhesive function it may not take part in adhesive interactions; instead calcium ions are likely to be important in stabilising the structure of the ectodomains which permits such interactions.

The cytoplasmic domains of the classical cadherins are highly conserved with up to 90% identity between them. It is this region where the classical cadherins interact with α, β and γ catenins. Since an E-cadherin deletion mutant lacking the catenin binding region was not adhesive when it was expressed in fibroblasts (Ozawa et al., 1990b), the catenins appear to be important for cadherin-mediated adhesion. Furthermore the adhesive properties of PC9 lung cancer cells, which have a null mutation in the α-catenin gene, could be restored by transfection with α-catenin cDNA (Hirano et al., 1992).
Figure 1.3: Structural Organisation of the Desmosomal Cadherins. Structural organisation of the desmogleins Dsg1, Dsg2 and Dsg3, and the desmocollin Dsc2 in comparison with the classical cadherin, N-cadherin. The mature cleavage site of the precursor peptide (P) is shown by (▼). Sites of putative N-glycosylation and cysteine residues are marked. The cell adhesion recognition tripeptide is labelled in the most N-terminal ectodomain of the mature protein (E1). N-cadherin is organised into five ectodomains made up of four repeated ectodomains (E1-E4) and a less well conserved extracellular anchor region (EA), a transmembrane portion (TM), an intracellular anchor (IA), an intracellular cadherin-type domain and a terminal domain (TER). The desmocollin splice variants Dsc2a and Dsc2b have a similar organisation and size to N-cadherin. Dsc2b contains an unique C-terminus. The extracellular serine/threonine rich domain of Dsg1 (EST) is a potential site for O-glycosylation. Desmogleins contain an intracellular repeat region (IR) and a proline-rich intracellular linker (IPL). The C-termini of Dsg1 and Dsg2 are glycine and serine rich (IG). (After Wheeler et al., 1991b)
Structural Organisation of the Desmosomal Cadherins

**Desmogleins**

Desmoglein 1 (DSG 1)

Desmoglein 2 (DSG 2)

Desmoglein 3 (DSG 3)

**Desmocollins : DSC2**

Desmocollin 1 (YAT)

Desmocollin 2 (FAT)

Desmocollin 3 (YAS)

**Classical Cadherins**

N-Cadherin

- **Cleavage-site**
- **N-glycans**
- **O-glycans**?
- **Cys residues**
Cadherins of the Desmosome; Desmocollins and Desmogleins.

Unlike the classical cadherins, the desmocollins are alternatively spliced (Collins et al., 1991; Parker et al., 1991). Each desmocollin gene produces two splice variants, named $a$ and $b$ which differ only in the encoded C-termini. The longer variant $a$ contains a potential serine phosphorylation site and in MDCK cells Dsc2a has been shown to be phosphorylated (Parrish et al., 1990). Interestingly phosphorylation of E-cadherin may be important in mouse compaction. This event does not coincide with an increase in E-cadherin expression but with an increase in E-cadherin phosphorylation (Sefton et al., 1992), suggesting that phosphorylation may be important in regulating adhesive function.

Desmocollin splice variant $b$ is shorter because of an extra exon with in-frame stop codon and has a unique C-terminus. In DSC1 and DSC2 this exon is 46bp, whilst in DSC3 it is a slightly shorter 43bp. Each type of desmocollin differs in the potential cell adhesion recognition (CAR) site present in the most N-terminal cadherin repeat. Peptides containing this sequence inhibit cadherin-mediated processes such as embryo compaction and neurite outgrowth (Blaschuk et al., 1990). In the classical cadherins this is H-A-V, while in Dsc1 it is Y-A-T, Dsc2 F-A-T and Dsc3 Y-A-S; this difference may reflect alternative adhesive functions.

The most outstanding novel feature of the desmogleins is their large cytoplasmic domain. This extends beyond the region of similarity with the desmocollins or classical cadherins and features a unique repeated element of approximately 29 residues with a core consensus of N-V-V/I-V-T-E-R/S-V-I/V. The repeated elements are predicted to fold into antiparallel $\beta$-sheets followed by a glycine-rich sequence. Dsg1 has five such repeats and by electronmicroscopy this portion of the molecule looks like a globular head attached to a thin tail (Rutman et al., 1994). Dsg2 has six repeats, while Dsg3 contains only two. At the N-terminal end of the desmoglein repeat region lies a proline-rich segment and at the C-terminus of this region in Dsg1 and Dsg2 lies a glycine-rich repetitive sequence. Desmoglein CARs differ from other cadherins, Dsg1 and Dsg3 both have R-A-L, whilst DSG2 has Y-A-L. The fifth cadherin ectodomain adjacent to the plasma membrane of DSG1 and DSG3 is shorter in the desmogleins by approximately 50 amino acids, and in human Dsg1 the region comprises a putative site for O-linked glycosylation (Penn et al., 1987). This may have implications in intercellular adhesion.
The primary structure of the desmosomal cadherins therefore reveals them to have related extracellular domains whereas parts of their cytoplasmic domains are unique. The significance of these differences is unknown, but suggests that the two types of desmosomal cadherin evolved to perform related yet distinct functions in the desmosome.

1.4 Molecular Dissection of Desmosomal Cadherin Function.

1.4.1 Intercellular Adhesion.

Classical Cadherin-Mediated Adhesion.

Experiments involving chimaeric molecules between E- and P-cadherin demonstrate that the 113 most N-terminal residues are essential for binding (Nose et al., 1990). Adhesive function would therefore appear to reside in the N-terminal cadherin ectodomain. Two pieces of evidence, however, suggest that other portions of the extracellular domain of cadherins may also be involved in intercellular adhesion. Firstly intercellular adhesion can be disrupted by both a monoclonal antibody, DECMA-1, which recognises an E-cadherin epitope within the ectodomain adjacent to the plasma membrane, and secondly by reducing agents which may act upon disulphide bonds (Ozawa et al., 1990a). Cadherin-mediated adhesion may also require the molecule to be anchored by the cytoplasmic catenins (Nagafuchi and Takeichi, 1988; Ozawa et al., 1990b), although this has been challenged by Roh and Stanley (1995); murine fibroblasts transfected with a transgene encoding the extracellular portion of E-cadherin and the intracellular domain of Dsg3, which does not bind catenins appear to be adhesive. Furthermore these chimaeric molecules are soluble in nonionic detergent in transfected cells, suggesting that they are not associated with the cytoskeleton. Thus interactions between cadherins and the cytoskeleton are not absolutely necessary for cadherin-mediated adhesion.

The structure of the N-terminal ectodomain has recently been solved for E- and N-cadherin and has ramifications as to how desmosomal adhesion might occur. By NMR, the structure of the 145 N-terminal amino acids of E-cadherin in solution is a seven-stranded \( \beta \)-barrel with two short \( \alpha \) helices (Overduin et al., 1995). Using X-ray diffraction of crystallised recombinant N-cadherin, the structure of the analogous ectodomain is composed of seven \( \beta \)-strands with a single \( \beta \)-helix of 1.5 turns (Shapiro et al., 1995). What is particularly surprising is that these models of cadherin repeat structure are very
similar to the Ig fold of the constant and variable domains within the immunoglobulin family of cell adhesion molecules, even though both families share little sequence homology. Thus this structural motif seems to have arisen through convergent evolutionary processes in these two types of adhesion molecule.

So how do cadherins bind one another? Overduin et al. (1995) suggest that the monomeric state of the recombinant N-terminal ectodomain of E-cadherin in solution is consistent with other parts of the molecule also being involved in adhesive interactions, since dimers might be expected if molecules were able to self-associate. Interestingly, the isolated single adhesive domains of other cell adhesion molecules such as CD2 and CD4 are also monomeric in solution (Driscoll et al., 1991; Ryu et al., 1990). In contrast Shapiro et al. (1995) suggests that packing of the N-cadherin monomers within the crystals is consistent with adhesive interactions mediated solely through the N-terminal ectodomain. The predicted interaction would be relatively loose, suggesting that cadherin adhesion is co-operative, which would be consistent with the observation that cadherins are often found concentrated at junctions to perform their function of strong adhesion. Furthermore, the predicted length of the N-cadherin extracellular domain from the structural data is estimated as 145Å (14.5nm) which is just under half the width of the desmosomal intercellular space (30-40nm). Since the sizes of the desmosomal cadherin extracellular domains are similar to the extracellular domain of N-cadherin, the predicted length of the cadherin extracellular domain is consistent with desmosomal cadherin adhesion mediated by interactions in the most N-terminal domain of the molecules. A model where adhesive interactions are within the N-terminal domain would also be consistent with the primary structure of Dsg1 and Dsg3 which do not contain the conserved ectodomain proximal to the plasma membrane.

Desmosomal Cadherin-Mediated Adhesion.

Early evidence that the desmosomal cadherins are cell adhesion molecules was indirect. Fab' fragments of polyclonal anti-desmocollin antibodies inhibit reformation of antibody-stainable desmosomes in MDBK cells (Cowin et al., 1984b). Antibodies against Dsg1 and Dsg3 are present in the plasma of patients with the autoimmune blistering skin diseases pemphigus foliaceous and pemphigus vulgaris. IgG from patients with
pemphigus vulgaris causes loss of cell adhesion in skin organ culture (Hashimoto et al., 1983). Likewise, Fab' fragments from patients with endemic pemphigus foliaceous cause blistering in neonatal mice (Rock et al., 1990).

Adhesive function of the classical cadherins is demonstrated by transfection of their cDNA into non-adhesive fibroblasts, such as L-cells (Nagafuchi et al., 1987; Miyatani et al., 1989; Nose et al., 1988). Ectopic expression of the cadherin results in aggregation of the fibroblasts when they are gently released from the substratum. However, the cytoplasmic domain of the cadherin must be intact in order for them to mediate cell adhesion (Nagafuchi and Takeichi, 1988; Ozawa et al., 1990b), presumably since this binds the catenins which links the cadherin to the actin microfilament network. L-cells express the catenins, but the desmosomal cadherins do not bind them (Ozawa et al., 1989). Therefore to utilise this system Amagai et al. (1994) constructed a chimaeric gene encoding the extracellular domain of DSG3 and the intracellular portion of E-cadherin. Transfection of this transgene into L-cells results in homophilic binding, but it is much weaker than that observed with L-cells transfected with an E-cadherin expression vector. This observation is surprising as desmosomal adhesion is strong and suggests that desmogleins might require coexpression of other desmosomal proteins, such as desmocollins and/or desmoplakin, for proper adhesive function. Alternatively desmoglein-mediated adhesion may be co-operative and the level of expression of transfected chimaeric genes may not be great enough to elicit a strong adhesive response.

Another possibility stems from the observations that antibodies which inhibit E-cadherin adhesion or anti-E-cadherin plus anti-P-cadherin antibodies also delay desmosome formation in cultured keratinocytes, which suggests that classical cadherin adhesion may be a prerequisite of desmosome-mediated adhesion (Wheelock and Jensen, 1992; Lewis et al., 1994). Furthermore in keratinocytes where classical cadherin adhesion has been disrupted by the introduction of a dominant negative mutation, desmosome formation is delayed upon increasing calcium levels that normally permit desmosome formation (Amagai et al., 1995). This apparent inhibition of desmosome assembly is likely to be post-transcriptional, since the desmocollin gene DSC2 is transcribed in preimplantation
embryos where anti-E-cadherin antibodies inhibit E-cadherin mediated adhesion (Collins et al., 1995)

C. Marcozzi (pers. comm.) has established L-cell cell lines stably transfected with full length cDNAs of either DSG1 or DSC2a, the latter is under the control of inducible promoters. Upon induction cells express Dsc2a ectopically and upon release from the substratum they gradually form small aggregates, demonstrating a weak homophilic adhesion mediated by a desmocollin. The Dsg1 transfectants provide similar results. However, cells stably transfected with both cDNAs express both proteins ectopically and appear to aggregate much more efficiently when released from the substratum. Although the relative levels at which these clones express each desmosomal cadherin will be important, a tentative conclusion would be that heterotypic together with homotypic interactions are more efficient than just homotypic interactions mediated by the desmosomal cadherins.

It is known that the requirement for an exclusive homophilic interaction between classical cadherins is not absolute. If cells expressing N-cadherin are mixed with cells expressing E-cadherin, there is a preference for the formation of homotypic contact, yet heterotypic and therefore heterophilic junctions are also found (Volk et al., 1987). It has yet to be formally addressed whether desmogleins and desmocollins are present within the same desmosome, yet since both types of desmosomal cadherin are found in the same desmosome-bearing tissues and cell lines, this seems highly likely. Desmosomal adhesion may therefore be mediated either via heterodimers present within the same cell or heterophilic interactions. Insofar as adhesion in both the adhaerens junction and desmosome is presumably mediated by cadherins, it is particularly intriguing that the desmosomes may contain more than one type of cadherin in contrast to the sole type within the adhaerens junction. The roles of the desmogleins and desmocollins in desmosomal adhesion await further elucidation of their respective functions.
1.4.2 Cytoplasmic Interactions.

Desmosomal Cadherins and the Plaque.

Chimaeric genes encoding the cytoplasmic domains of either Dsclα, Dsclβ or Dsg1 linked to the transmembrane and extracellular domains of connexin-32 have been used to examine the associations of the desmosomal cadherins with the desmosomal plaque (Troyanovsky et al., 1993). Transfection of chimaeric constructs containing the desmocollin Dsclα domain into desmosome-bearing A431 cells results in the formation of plaque-like structures that tether bundles of intermediate filaments and contain both desmoplakin and plakoglobin. Transfection of the shorter Dsclβ domain, however, results in no plaque formation. Thus the residues present within Dsclα but not within Dsclβ are functionally significant in the assembly of the desmosomal plaque. The function of the shorter desmocollin β variant is all the more intriguing since both variants of Dsc2 are expressed in the same tissues (Nuber et al., 1995). Meanwhile transfection of the Dsg1 chimaera into A431 cells is much more dramatic and results not only in no junction formation but also the disassembly of all endogenous desmosomes. It is unlikely that this effect is due to the depletion of some cytoplasmic factor necessary for desmosome assembly, as a synaptophysin (a transmembrane protein of presynaptic vesicles) chimaera which localises the protein to small cytoplasmic vesicles is still able to bind plakoglobin but not inhibit desmosome assembly. Plaque assembly therefore appears to be a carefully regulated process sensitive to the position of the cytoplasmic tails of the desmosomal cadherins.

The site where plakoglobin interacts with Dsg1 has been mapped by examining the ability of plakoglobin to bind to a series of recombinant Dsg1 cytoplasmic domains with nested deletions from the C-terminus (Mathur et al., 1994). The site is located within a sequence of 19 residues, between 207 and 188 residues from the transmembrane region, within the 469 residues of the cytoplasmic domain. These residues are necessary for the plakoglobin-Dsg1 interaction but other residues towards the transmembrane region could also be involved which cannot be resolved by nested deletions alone. Plakoglobin also binds E-cadherin although the binding site has yet to be resolved. However, 10 of these 19 residues within Dsg1 have some similarity to the catenin binding domain of E-cadherin. Both plakoglobin and β-catenin are highly related with 63% identity between
them, and since E-cadherin can only bind one or the other at a given time (Hinck et al., 1994; Hülsken et al., 1994), the two sites may either be the same or be very close so that the two proteins are sterically inhibited from both interacting with E-cadherin simultaneously.

Transfection of a transgene encoding the extracellular domain of E-cadherin and only the intracellular anchor of Dsg3 (40 amino acid residues) into murine fibroblasts results in intercellular adhesion (Roh and Stanley, 1995). Since the intracellular domain is highly conserved in the desmogleins, this may suggest an important function of this subdomain such as enabling these chimaeric molecules to aggregate for efficient adhesion. This may also be important in desmosome formation.

Associations of Plakoglobin.
Plakoglobin interacts with the desmosomal cadherins Dsg1 and Dsg3, E-cadherin and the product of the APC tumour suppressor gene (Hülsken et al., 1994). Interactions of plakoglobin between E-cadherin and APC are competitive; furthermore plakoglobin can also bind the cytoskeleton-associated protein of the adhaerens junction, α-catenin. Plakoglobin is therefore at an important link between cadherin-mediated cell adhesion, interactions with the cytoskeleton and tumour progression.

Plakoglobin may also have a role in signal transduction since its homologue armadillo (65% identity) is part of a signal cascade in Drosophila, namely from wingless (wnt-1 being the vertebrate homologue) to engrailed, involved in determining segment polarity. Interestingly, expression of wnt-1 in mammalian cells results in an accumulation of plakoglobin and β-catenin at the cell membrane and a strengthening of cadherin mediated adhesion (Hinck et al., 1994; Bradley et al., 1993), suggesting that similar signalling events also occur in vertebrate development.

1.5 Expression Patterns of the Desmosomal Cadherins.
A Brief Introduction to Keratinisation.
The epidermis is a complex multilayered epithelium where the unit cell is the keratinocyte. Only keratinocytes attached to the basement membrane within the basal
layer undergo mitosis. At some point they exit the cell cycle and are pushed towards the skin's surface undergoing a program of differentiation known as keratinisation. There are marked changes in the keratinocytes' morphology and biochemical composition as they migrate from the basal layer through the spinous and granular layers until they reach the stratum corneum where proteases are activated and the cell structure degraded. The dead squames are then sloughed from the surface and replaced by differentiating cells from below in a continuous and steady state process. Desmosomes are present in the basal layer and as keratinisation continues their number and size increases and their electron-density changes (Skerrow et al., 1989; Chapman and Walsh, 1990).

The differential expression of the keratins, which form the intermediate filament network within the keratinocytes and are their major constituent, typifies keratinisation. Basal cells express keratin types 5 and 14 and as keratinisation continues these are down-regulated and differentiation-specific keratins are expressed (Fuchs and Green, 1980). The type of differentiation-specific keratin may depend upon the location of the skin; most body regions express keratins 1 and 10, while in palmar and plantar skin keratin 9 is found. During wound-healing keratins 6 and 16 are induced suprabasally (Mansbridge and Knapp, 1987). Other epidermal genes also have a stratification-dependent expression. The proteins filaggrin, loricrin and involucrin are expressed in the granular layer (Rothnagel and Steinert, 1990; Hohl et al., 1991; Rice and Green, 1979). Filaggrin bundles the keratin filaments into macrofibrillar cables, while both loricrin and involucrin are secreted and become substrates for calcium-dependent transglutaminases which cross-link them to form the cornified envelopes (Hohl et al., 1991).

**Desmosomal Cadherin Expression in Epidermis.**

Immunofluorescence experiments with monoclonal and polyclonal antibodies that recognise human Dsc1 and Dsc2 first suggested that these proteins may have a stratification-related expression in epidermis (King et al., 1991). This was confirmed with further immunofluorescence experiments (King et al., 1993b). However the close identity of these proteins means that they may share epitopes which make it difficult to distinguish differences in their expression by immunological means alone.
Patterns of desmosomal cadherin gene expression in human epidermis has been determined by in situ hybridisation experiments using isoform specific cRNA probes. These also indicated that DSC1 expression was suprabasal whilst DSC2 was expressed within the basal layers of epidermis (Arнемann et al., 1993). However, more recent in situ experiments, whilst confirming the suprabasal expression of DSC1, reveal that DSC2 is expressed throughout the living layers of epidermis (I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data). DSC3 is also expressed throughout the living layers of human epidermis (figure 1.4). A similar pattern is seen in epidermis from human plantar, palmar and forehead (I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data; Theis et al., 1993). Many probes seem to nonspecifically bind to the basal layer of the epidermis which may explain the previous data (Arнемann et al., 1993) which suggested that DSC2 expression might be predominantly within the basal layer of human foreskin. In the case of the desmogleins, in human foreskin, DSG1 mRNA expression is suprabasal, whilst both DSG3 and DSG2, although more weakly expressed, appear to be expressed within the lower spinous layer and basal layer, respectively (Arнемann et al., 1993). Using antibodies against peptides corresponding to unique regions of the desmocollin proteins, it has recently been shown that protein expression of Dsc1, Dsc2 and Dsc3 is similar to the pattern of the mRNAs (I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data). This is consistent with desmocollin expression being controlled at the level of gene transcription.

Subsequent in situ hybridisation experiments on bovine muzzle epidermis describes a similar pattern to human epidermis for bovine DSC1 expression in suprabasal layers (Legan et al., 1994). However, in contrast to human epidermis, bovine DSC2, although present in suprabasal layers, appears to be absent from the basal layer, and DSC3 is restricted to basal layers. Thus in bovine muzzle DSC3 expression is proposed to be basal; in the 5-10 immediately suprabasal layers DSC2 expression is predominant, until this is superceded in more suprabasal layers by DSC1. The discrepancies in the location of DSC2 and DSC3 expression between human and bovine tissues may be due to interspecies differences, the different location of the skin samples or the relatively low expression of DSC2 and DSC3 making the genes' expression difficult to delineate. It is unlikely that these differences would be because the human and bovine genes are not
Figure 1.4: Differential Expression of the Desmosomal Cadherins in Human Epidermis.
Stratum corneum
Stratum granulosum
Stratum spinosum
Stratum basale

Epidermis

Desmogleins
DSG1, DSG2, DSG3

Desmocollins
DSC1, DSC2, DSC3

Filaggrin, Loricrin
Involucrin, Cornifin

K1, K10
K5, K14

Dermis
species orthologues, due to the high levels of identity between the human and bovine DSC2 and DSC3 genes.

The function of this differential expression of the desmosomal cadherins in epidermis remains speculative. Unlike the desmosomal cadherins, neither the desmoplakins nor plakoglobin show stratification-related patterns of expression which suggests that the changes to the desmosome may be predominantly in intercellular adhesion. Thus the desmocollins and desmogleins may modulate the adhesive nature of desmosomes in the different cell layers. Cells of the epidermis form a dynamic barrier which resists both shearing forces and penetration, but also is constantly being sloughed from the skin's surface and regenerated from the lower basal layer. Thus desmosomes of the lower layers could contain less adhesive desmosomal cadherins which allow keratinocytes to move through the cell layers, whilst those of the uppermost layers may contain more adhesive desmosomal cadherins that help to perform the skin's barrier function. In conditions such as psoriasis where the stratum corneum thickens and desquamation may be more inefficient, it could be that the desmosomal cadherins are no longer appropriately expressed.

**Tissue Distribution of Desmosomal Cadherin Isoforms.**

DSC2 and DSG2 transcripts are found in all the desmosome-bearing tissues and cell lines tested and may therefore be ubiquitous to desmosome-bearing cells (Nuber *et al.*, 1995; Schafer *et al.*, 1994). Furthermore in certain tissues such as heart and kidney, DSC2 and DSG2 are the only known desmosomal cadherins to be expressed (Angst *et al.*, 1995; Nuber *et al.*, 1995; Schafer *et al.*, 1994). In contrast DSC1 expression is tightly restricted and transcripts are only detected in epidermis, tongue and lymph node. DSG1 expression likewise is restricted to epidermis, tongue and lymph node, but it is also expressed in tonsil and oesophagus. Both DSC3 and DSG3 have a similar catalogue of tissues to DSG1 where transcripts are detected. Immunofluorescence experiments also show DSC3 to be expressed in all the living layers of stratified epithelia such as hard palate, buccal mucosa, and cervix (I.A. King, K.H. Sullivan, R. Bennett and R.S. Buxton, unpublished data). The overall picture at present is that DSC1 is restricted
mainly to epidermis, while DSG1, DSC3 and DSG3 are associated with certain stratified epithelia.

The similar expression patterns for particular desmocollin and desmoglein isoforms, such as DSC2 and DSG2 in all desmosome-bearing tissues, and DSC1 and DSG1 in the suprabasal layers of epidermis, would be consistent with both types of desmosomal cadherin being necessary for desmosomal function. However, it is not yet known whether desmocollins and desmogleins are expressed within the same desmosome or cell. A more detailed description of desmosomal cadherin protein expression awaits the availability of isoform-specific antibodies. These would also be useful in resolving whether different desmocollin or desmoglein isoforms are present within the same desmosome, since the overlapping expression patterns of isoforms in epidermis and cell lines would make this a possibility.

**DSC2 Expression in Developing Epithelium; Trophectoderm.**

Trophectoderm develops from non-polar blastomeres during preimplantation development. While the first epithelial features are observed at the 8-cell stage with compaction mediated by E-cadherin, desmosomes form relatively late at the 32-cell stage. Desmosome formation is rapid with junctions of full apparent molecular complexity observed at this stage. This is in contrast to tight junctions which begin to form at compaction and are only mature at the 32-cell stage. At this time vectorial transport commences and blastocoel accumulation begins. Desmosomes may therefore have an important function in stabilising the trophectoderm as the blastocoel cavity enlarges (Fleming et al., 1991).

Before desmosomes form, the plaque proteins plakoglobin, DspI and DspII are synthesised and accumulate (Fleming et al., 1991). Plakoglobin is detected as early as the 8-cell stage; this is not too surprising since it is associated with E-cadherin in adherens junctions and may therefore be associated with E-cadherin in early development. DspI and DspII are both detected at the 16-cell stage. However, desmosomal cadherin proteins are not detected before the 32-cell stage which implies that their synthesis may be regulatory for desmosome formation at cavitation. By performing
RT-PCR upon single murine preimplantation embryos, Collins et al. (1995) measured the onset and location of DSC2 transcription from the embryonic genome. DSC2 is transcribed from the 16-cell morula or very early blastocyst (32-cell) stages onwards, immediately prior to the detection of Dsc2 protein. Thus it is most likely that desmosome formation is limited and regulated by transcriptional controls upon DSC2. A similar situation probably exists for a desmoglein, possibly DSG2. Most significantly, expression of DSC2 is likely to be under transcriptional control in early development.

In early blastocysts DSC2 transcription is restricted to the trophectoderm. However when the totipotent cells of the inner cell mass are dissected from the trophectoderm and cultured, DSC2 transcription is activated and upregulated in these cells and Dsc2 protein expressed (Collins et al., 1995). Contact-free cell surfaces may therefore be important in the initiation of DSC2 transcription in trophectoderm at the mid 16-cell stage.

1.6 Structure of Cadherin Genes.

The genes for the classical cadherins E-, N-cadherin and L-CAM (chicken homologue of E-cadherin), although differing in size, share the same genomic organisation, containing 16 exons of similar sizes, and have exon-intron boundaries conserved between the different genes (Ringwald et al., 1991; Miyatani et al., 1992; Sorkin et al., 1988). The organisation of P-cadherin is also very similar to these genes (Hatta et al., 1991), except for the absence of the first intron in the P-cadherin gene and the larger size of its first exon of 1kb, in contrast to the sizes of the first two exons of E- and N-cadherin which together account for 296bp and 504bp, respectively. The second intron (or the first in P-cadherin) splitting exons that encode the precursor peptide, is relatively large in all these genes; in N-cadherin and P-cadherin this is >100kb and >23kb, respectively. In P-cadherin this intron contains significant enhancer activity (Hatta and Takeichi, 1994). The large size of these introns may therefore have some importance in holding DNA regulatory elements that play a role in transcriptional regulation of these genes. The sizes of the genes also vary so that the overall pattern of the introns is similar in all the genes, irrespective of size; the N-cadherin gene is >200kb with a second intron half the size of this at >100kb, the P-cadherin gene is 45kb with a
corresponding intron of >23kb and the L-CAM gene is 9kb with a second intron of 3.5kb.

The genomic organisation of the bovine DSG1 gene has also been reported (Puttagunta et al., 1994). The gene extends over more than 37.5kb and contains fifteen exons. It has a similar organisation of intron-exon boundaries to the classical cadherins, particularly in the region encoding the extracellular domain where the desmogleins are most similar to the classical cadherins. However, the relative sizes of corresponding introns vary, most notably the second intron within the precursor peptide is no more than 400bp in the bovine DSG1 gene. This feature may point to significant differences in the transcriptional control among these genes.

Location of the Cadherin Genes.
Mapping studies have revealed cadherin linkage groups within the human genome. Human E-, P- and M-cadherin all map to human chromosome 16, E-cadherin specifically at 16q22.1 (Mansouri et al., 1988) and M-cadherin at 16q24.1-qter (Kaupmann et al., 1992). Bussemakers et al. (1994a) isolated cosmids which linked the most 3' exon of P-cadherin a mere 32kb upstream of the most 5' E-cadherin exon. The L-CAM and K-CAM (Chicken homologue of P-cadherin) genes are also arranged in tandem in the chicken genome (Sorkin et al., 1991), with less than 700bp separating the presumed poly-A site in the K-CAM gene and the translation initiation site of the L-CAM gene. All the known desmosomal cadherin genes map to human chromosome 18 by PCR analysis on somatic cell hybrids (Arnemann et al., 1991; Arnemann et al., 1992b; Arnemann et al.; 1992a; King et al., 1993a). N-cadherin also maps to human chromosome 18 at q11.2 (Walsh et al., 1990; Wallis et al., 1994). The evolutionary conservation of these gene clusters suggests that the close proximity of the genes may be important for their regulation. Furthermore it probably reflects a role for gene duplications in desmosomal cadherin evolution.

1.7 Desmosomal Cadherins and Disease.
In contrast to other epithelial gene families, such as keratins where a number of skin diseases have been attributed to genetic mutation (McLean and Lane, 1995), no genetic
disease has yet been firmly identified with the desmosomal cadherins. This may simply be that the search has not been extensive enough. Alternatively, mutations of these genes could be lethal, or the function of inactivated isoforms might be compensated by other desmosomal cadherins. There is, however, a report (A. Reis, pers. comm. to J. Arnemann) that some examples of the inherited skin disease epidermolytic palmoplantar keratoderma (EPPK) do not map in the keratin genes, specifically in keratin 9, but are present on chromosome 18 and are closely linked to the DSC and DSG cluster, suggesting that a desmosomal cadherin gene could be a candidate for this disease. Another potential candidate is the mouse bal mutation (Davisson et al., 1994). bal mice display a balding phenotype where hair is lost in patches and as the desmosomal cadherins are expressed in epidermis they are candidates for this mutation. bal also maps very close to the murine desmosomal cadherins upon mouse chromosome 18; indeed it is so close that a recombination event between the mutation and desmosomal cadherins could not be detected.

Desmosomal cadherins are targets of autoimmune diseases. Dsg1 and Dsg3 are antigens for the pathological skin diseases pemphigus foliaceus and the more severe and potentially lethal variant, pemphigus vulgaris, respectively (Rubinstein and Stanley, 1987; Amagai et al., 1991). These diseases result in blistering due to loss of intercellular adhesion, with the position of cell separation correlating to the location of DSG1 and DSG3 transcripts in human epidermis. Interestingly, no specific autoantibodies have yet been described for Dsg2 which would probably be lethal due to the widespread expression of this desmoglein (Schafer et al., 1994). Desmocollins may also be reactive towards certain sera from patients with pemphigus-like diseases, particularly from those with the variant Brazilian pemphigus foliaceus, fogo selvagem (Dmochowski et al., 1993).

Changes in the expression or function of the desmosomal cadherins may be important in the progression of a tumour from a benign to an invasive malignant state, since this involves changes in intercellular adhesion. Tumour metastasis involves two independent events; detachment of a cell from the primary tumour and then reattachment to metastasizing sites. Thus in primary tumours, desmosomal cadherins may be down-
regulated or their adhesive function reduced. Since retrospective studies do not conclusively show a reduction in the number of desmosomes in malignant primary tumours and metastases (Collins et al., 1990; Garrod et al., 1987), desmosomal adhesion mediated by the desmocollins and desmogleins may be altered in such states. By staining primary and metastatic lymph node oesophageal carcinoma with an anti-desmoglein monoclonal antibody, Natsugoe et al. (1995) suggest that desmoglein expression is reduced in metastatic tumours.

E-cadherin is frequently down-regulated in many highly invasive, poorly differentiated carcinomas (Umbas et al., 1992; Mareel et al., 1991). Furthermore, transfection of E-cadherin into invasive cells results in restoration of a non-invasive phenotype, suggesting that E-cadherin may act as a tumour-suppressor gene (Frixen et al., 1991). This is demonstrated dramatically in the finding that E-cadherin is mutated within 50% of diffuse-type gastric carcinomas (Becker et al., 1994). Alterations in E-cadherin expression or function in the progression to malignancy might compromise desmosomal adhesive function, since anti-E-cadherin antibodies affect desmosome reformation in cell culture (Wheelock and Jensen, 1992; Lewis et al., 1994). Other cadherins have also been implicated in cancer. In Drosophila, fat is a tumour suppressor gene which displays a total of 34 cadherin-like tandem repeats (Mahoney et al., 1991). Similar genes with fat-like cadherin repeats exist within vertebrates although their function is unknown (Sano et al., 1993). The human ret gene encodes an extracellular cadherin-like repeat, and as a receptor tyrosine kinase it is also a proto-oncogene (Schneider, 1992).

1.8 Focus of this Work.

As putative adhesion molecules, the desmosomal cadherins are important in maintaining the integrity of developing trophectoderm, differentiating epidermis and fully-formed tissues, such as cardiac muscle. Isoforms are expressed in a cell-specific manner which switches between and within tissues. An investigation into the mechanisms which regulate transcription of the desmosomal cadherins would therefore further understanding of the factors which contribute to events of morphological change. To initiate an investigation into desmocollin transcriptional regulation human genomic libraries were screened for the DSC2 and DSC1 promoters.
In addition, the functional significance of the different desmosomal cadherins and their differential expression is also unknown. To address this issue, experiments preliminary to the targeted inactivation of murine DSC2 and DSG1 were performed. A murine DSG1 cDNA was characterised and a murine genomic DSC2 clone isolated.
Chapter Two - Molecular Cloning and Characterisation of the Human DSC2 Promoter

2.1 Introduction
2.2 Results
2.2.1 Isolation of Genomic Clones
2.2.2 Characterisation of the Genomic Clones, λ2.DSC2 and λ4.DSC2
2.2.3 Southern Analysis of the DSC2 Locus
2.2.4 Mapping the Transcription Initiation Site
2.2.5 Primary Structure of the Promoter
2.2.6 Identification of a CpG Island
2.2.7 Human DSC2 Promoter Activity in Cultured Cells
2.2.7.1 Construction of Reporter Constructs
2.2.7.2 Experimental Controls
2.2.8 Transfection of DSC2 Reporter Constructs into MDCK Cells
2.2.9 DSC2 Promoter Activity is Unidirectional
2.2.10 To Test for Enhancer Elements within the Human DSC2 Promoter
2.2.11 Transfection of DSC2 Reporter Constructs into Other Cell Lines
2.2.12 DSC2 Reporter Constructs to Study Regulatory Elements In Vivo
2.3 Discussion
2.3.1 Site of Transcription Initiation
2.3.2 Potential Regulatory Transcription Factors
2.3.3 Activity of the DSC2 Promoter in Transient Transfection Experiments
2.4 Major Findings
2.5 Future Work - Identify Proteins that Interact with the DSC2 Promoter
2: MOLECULAR CLONING AND CHARACTERISATION OF THE HUMAN DSC2 PROMOTER

2.1 INTRODUCTION

The most important regulatory elements of a gene are found immediately 5' of the transcription start site within the promoter, as it is there that proteins which initiate transcription bind to the DNA. The promoters of eukaryotic protein-coding genes are multipartite and interact with a number of protein factors besides RNA polymerase II. General and promoter-specific transcription factors are required for accurate and appropriate transcription of a gene and their interactions with DNA demarcate a point where gene expression is controlled.

General transcription factors such as the TATA box-binding protein bind core sequences close to the transcription start site and direct RNA polymerase II to the appropriate transcription start site on the DNA (Zawel and Reinberg, 1992). They may bind to motifs such as the TATA box and/or a less well defined initiator element present within some promoters. However, these transcription factors are only sufficient to allow a low level of transcriptional activity. Promoter-specific activators interact with proximal promoter elements within the DNA close to the transcription start site and greatly stimulate the general transcription factors. Typically, activators are composed of a DNA-binding domain that recognises a specific sequence of DNA, and an activation domain required for activating transcription (Ptashne, 1988; Mitchell and Tjian, 1989). These factors are proposed to stabilise the initiation complex as it assembles thus increasing the likelihood of transcription. Several activator binding sites may be present within a promoter.

The protein binding elements within a promoter can be sufficient to allow the appropriate expression of a transgene transfected into cultured cells. For instance, a fragment of the murine E-cadherin promoter extending 178bp upstream of the transcription start site and encompassing the translation start site drives CAT transgene expression within transfected epithelial cells but not fibroblasts (Behrens et al., 1991). Likewise, a 300bp fragment of DNA immediately upstream of the rabbit keratin K3 gene is capable of driving the
expression of the CAT reporter gene in cultured corneal cells but not in kidney epithelial cells (Wu et al., 1993).

Enhancers bind activating proteins which may also be necessary for appropriate gene expression. Enhancer elements are sequences that can be located upstream, downstream or within a transcriptional unit and function in either orientation relative to the transcription initiation site. They stimulate the transcription of a transgene, although they are unable to initiate transcription themselves (Atchison, 1988). Unlike regulatory elements immediately upstream of a gene, the location of enhancers is difficult to predict.

DSC2 is expressed within all the living layers of epidermis (I.A. King, K. H. Sullivan, R. Bennett and R.S. Buxton, unpublished data) and is ubiquitous to all desmosome-bearing tissues and cells (Nuber et al., 1995). Furthermore activation of the DSC2 gene in the murine blastocyst denotes a developmental step; the formation of the first desmosomes as the blastocoeI cavity begins to form and swell with fluid (Collins et al., 1995). The mechanisms which regulate the spatial and temporal expression of the DSC2 gene will involve the interactions between its promoter and transcription factors. To investigate the transcriptional regulation of the human DSC2 gene, sequences immediately 5' of the translation start site were cloned and characterised. This work reports that the cloned human DSC2 promoter was able to drive the expression of the luciferase transgene in transfected cells and identified a minimal DSC2 promoter.

2.2 RESULTS

2.2.1 Isolation of Genomic Clones.

In order to isolate DNA containing 5' promoter elements of the DSC2 gene a 32P-labelled 197bp BamHI-HindII fragment from the most 5' region of the human DSC2 cDNA (figure 2.1, Parker et al., 1991) was used to probe a λFix II male human genomic placenta library. Any genomic clone hybridising to the most 5' portion of the DSC2 cDNA may also contain DNA immediately upstream and adjacent to it. Out of approximately 10^6 plaques screened, the equivalent of representing the diploid human...
Figure 2.1: Schematic Representation of the Human DSC2a cDNA (Parker et al., 1991). The shaded box represents the DSC2a open reading frame with the arrowhead indicating the alternative splice site. The locations of the 5' 197bp BamHI-HindIII and the adjacent 3' 257bp HindII-HindIII fragments that were used as probes are marked. The primer RSB39 was used to sequence DNA upstream of the published DSC2 cDNA sequence within the DSC2 genomic clones.
Human DSC2a cDNA
genome with a three-fold redundancy, two phage, λ2.DSC2 and λ4.DSC2, were identified and purified.

2.2.2 Characterisation of the Genomic Clones, λ2.DSC2 and λ4.DSC2.
Both clones had inserts of more than 15kb and a common 3.2kb EcoRI fragment which, since it hybridised to the most 5' 197bp BamHI-HindII fragment of the DSC2 cDNA but not to an adjacent 3' 257bp HindII fragment, most probably contains DNA immediately upstream of the cDNA probe. This 3.2kb EcoRI fragment was subcloned from λ2.DSC2 into the EcoRI site of pBluescript to create the plasmid, pMDM4 and mapped with restriction endonucleases (figure 2.2).

λ2.DSC2 DNA was digested with restriction endonucleases, Southern blotted and probed with a 5' 0.8kb HindIII-NsiI fragment and a 3' 0.8kb SalI-EcoRI fragment of pMDM4 (figure 2.2). The use of genomic probes enabled a linear map of λ2.DSC2 to be constructed. Southern blotting of λ4.DSC2 digested with SalI revealed the 5' HindII-NsiI pMDM4 probe to hybridise to a 5.4kb SalI fragment whilst with λ2.DSC2 it hybridised to a 7kb SalI fragment. This indicated a SalI site only present within λ4.DSC2 which was probably within the vector. Likewise, the 3' SalI-EcoRI probe from pMDM4 hybridised to a 10.5kb SalI fragment of λ4.DSC2 and with λ2.DSC2 it hybridised to a 8kb SalI fragment, indicating a SalI site only present in λ2.DSC2. Thus both clones overlap but λ2.DSC2 extends further 5' than λ4.DSC2 which contains more 3' DNA (figure 2.2).

During the cloning and amplification of the library it is possible that deletions or rearrangements of DNA sequences could accumulate. To ensure the integrity of the cloned 3.2kb EcoRI fragment, human genomic DNA and λ2.DSC2 DNA were digested with EcoRI and resolved upon the same agarose gel. This was then Southern blotted and probed with the 197bp BamHI-HindII cDNA fragment which had been used to screen the genomic library. The probe hybridised to a 3.2kb band with both digested DNAs (figure 2.3), which confirmed that the cloned EcoRI fragment did not contain any deletions or insertions of DNA that could be resolved from the 3.2kb genomic DNA band on an agarose gel.
Figure 2.2: Restriction Map of pMDM4, Southern Blot and Restriction Map of \( \lambda 2.\text{DSC2} \). Top: a 3.2kb \( \text{EcoRI} \) fragment of \( \lambda 2.\text{DSC2} \) inserted into the \( \text{EcoRI} \) site of pBluescript to create pMDM4, mapped with various endonucleases. Centre: 5' \( \text{HindIII-NsiI} \) and 3' \( \text{SalI-EcoRI} \) fragments of pMDM4 were used to probe a Southern blot of \( \lambda 2.\text{DSC2} \) DNA digested with \( \text{NsiI, Apal, HindIII, PstI, SalI and BamHI} \). Bottom: Map of \( \lambda 2.\text{DSC2} \) derived from Southern blots. A map of \( \lambda 4.\text{DSC2} \) was derived from other Southern blots and is shown aligned below \( \lambda 2.\text{DSC2} \).

**Endonucleases:** \( \text{ApaI, A; AvaI, Av; AvaiII, AvII; BamHI, B; BssHII, Bs; EagI, Eg; EcoNI, En; EcoRI, E; HindIII, H; NsiI, N; PstI, P; SaeII, Sc; SalI, S; SmaI, Sm; SspI, Sp.} \)
Figure 2.3: Southern Blot of Human Genomic DNA and λ2.DSC2 DNA. Human genomic DNA and λ2.DSC2 DNA (labelled Phage DNA) were digested with EcoRI, run within the same agarose gel, Southern blotted and probed with the $^{32}$P-labelled 197bp BamHI-HindII fragment of the DSC2 cDNA.
Sequencing pMDM4 with the primer RSB39 (figure 2.1) verified that this subclone contained 5' sequences continuous with the DSC2 cDNA. It also localised an *AvaI* site already mapped within pMDM4 to be 9bp upstream of the DSC2 translation initiation codon. Thus the 5' extent of the DSC2 translated cDNA was identified within the genomic DNA. As regulatory elements will be present within the DNA upstream of the DSC2 coding region, the 1.9kb *EcoRI-Aval* fragment 5' of the *AvaI* site was blunt-ended and subcloned from pMDM4 into the *EcoRV* site of pBluescript to create the plasmid pMDM7 for further characterisation.

2.2.3 Southern Analysis of the DSC2 locus.

Before the genomic clones λ2.DSC2 and λ4.DSC2 were isolated, the structure of the DSC2 locus was investigated. Human genomic DNA digested with *EcoRI*, *HindIII* and *BamHI* was Southern blotted then probed with the most 5' 197bp *BamHI-HindII* and adjacent 3' 257bp *HindII* fragments of the DSC2 cDNA (figure 2.4A). Restriction fragments of each enzyme which hybridised to the probes could then be orientated relative to one another within the DSC2 locus (figure 2.4B). Since subsequent work identified *EcoRI*, *BamHI* and *HindIII* sites within λ2.DSC2, each map derived from this Southern analysis of human genomic DNA could be aligned with λ2.DSC2 (figure 2.4C). This was consistent with λ2.DSC2 representing sequences 5' of the human DSC2 locus and enabled a *HindII* site bordering both cDNA probes to be localised within the genomic DNA.

Both cDNA probes hybridised to a 6.5kb *BamHI* and a 2kb *HindIII* fragment (figure 2.4A), which meant that the *HindII* site that bordered both cDNA probes must have been within these fragments. An alignment of the restriction maps with that of λ2.DSC2 implied there was a short region common to the 6.5kb *BamHI* and 2kb *HindIII* fragments where the *HindII* site would probably map (figure 2.4C). As this region of overlap was approximately 6kb from the *AvaI* site 9bp from the translation initiation codon mapped within λ2.DSC2 (figure 2.2), there may be approximately 6kb of intron sequence between the translation start codon and this *HindII* site in the genomic DNA, whereas in the cDNA the distance between these two points was 155bp.
Figure 2.4: Southern Analysis of the DSC2 Locus. (A) Southern blot of human genomic DNA digested with HindIII, EcoRI and BamHI and probed with the 5' 197bp BamHI-HindIII and adjacent 3' 257bp HindIII-HindIII fragment of the human DSC2 cDNA. (B) Restriction maps of the DSC2 locus derived from (A). (C) An alignment of λ2.DSC2 and genomic maps. The predicted region of overlap between BamHI and HindIII fragments that hybridise to both the 5' and 3' cDNA probes is marked.

Endonucleases: BamHI, B; EcoRI, E; HindIII, H; NsiI, N; SalI, S.
Blot probed with 5' cDNA

Blot probed with 3' cDNA

Overlapping genomic HindIII and BamHI fragments
2.2.4 Mapping the Transcription Initiation Site.
The major transcription initiation site of the human DSC2 gene was mapped using RNase protection of a mRNA fragment by a riboprobe. This was performed using total RNA from cultured primary keratinocytes; these cells are known to express the DSC2 gene since a human cDNA library prepared using RNA purified from cultured primary keratinocytes had been used to clone the original DSC2 cDNA clone (Parker et al., 1991). pMDM7 was linearised using *AvaiI* and a riboprobe of 561nt synthesised from the phage T7 promoter. After gel purification, a total of $10^5$cpm were hybridised with 20µg of total RNA for fifteen hours at 49°C. After RNase digestion, a protected fragment equivalent to a 192nt ssDNA molecule on a sequencing ladder was identified (figure 2.5). This predicts that the major transcription initiation site lies 201bp upstream of the translation initiation site.

Although the RNase protection technique is highly sensitive and specific it does have limitations, since it does not distinguish between the 5' extent of a mRNA and the end of an intron. To complement RNase protection experiments, primer extension and 5' RACE experiments were performed, but were unsuccessful. Both these techniques generate cDNA using a reverse transcriptase to extend a specific primer to the 5' extent of a mRNA to which it is annealed. Regions of DNA with a high G+C content, such as that upstream of the DSC2 translation start site, may form stable secondary structures which inhibit the progression of reverse transcriptase along the mRNA. Within the p53 promoter, for example, a large stem and loop structure forms and can be distinguished by S1 nuclease sensitive sites 3' and 5' and at the loop. However when primer extension experiments were performed, extension terminated 3' of the stem (Bienz-Tadmoor et al., 1985).

2.2.5 Primary Structure of the Promoter.
Both DNA strands of the 1.9kb upstream of the DSC2 translation initiation site were sequenced (figure 2.6). Transcription factor consensus binding sites and putative regulatory motifs were identified using the GCG FINDPATTERNS program with the TFD database release 6.5 (5/93). The human DSC2 promoter was devoid of canonical TATA or CAAT boxes. Three potential initiator sequences corresponding to the
Figure 2.5: Mapping the Transcription Initiation Site of DSC2. A 561nt antisense riboprobe extending as far 5' as the AvaII site (-497) was generated from the T7 promoter of pMDM7 and used in RNase protection experiments. Total keratinocyte RNA or yeast tRNA was hybridised to the riboprobe before treatment with RNase and subsequent sizing upon a 6% polyacrylamide gel alongside a DNA sequencing ladder. The arrowhead indicates a major protected band of keratinocyte RNA equivalent to a ssDNA molecule of 192nt.
DNA Ladder

Keratinocyte RNA

tRNA

(-497)
AvalI

(-9)
Aval

T7 RNA polymerase

Riboprobe

561 nt

Protected fragment

192 nt
Figure 2.6: Nucleotide Sequence of the Human DSC2 promoter. Nucleotide position -1 is assigned to the first nucleotide upstream of the ATG translation start codon. The amino acid translation of the first exon is shown beneath the nucleotide sequence using the one letter code. (¶) indicates the predicted major transcription start site at -201. The primers MDM24 used to map the 5' end of the first intron, and RSB39 used to sequence part of the promoter are overlined. (⇒) denotes the 5' extent of the first intron.

Consensus binding sites for the following transcription factors and motifs are highlighted and labelled; AP-1, AP-2, SP-1, NFκB, Pit-1, CK-8-mer and GC-Box. Keratin 14-, DSG1- and DSG3-like sequences are underlined and labelled. Consensus sequences for retinoic acid response elements (RAREs) are also highlighted and labelled. Two types of repeated sequence, 1 and 2, are underlined and labelled. Endonuclease sites for EcoRI, HindIII, NsiI, AvaII, EagI, EcoNI, BssHII, and AvaI are shown.
EcoRI

HindIII

AP-1

Pit-1

DGS3-like

AP-2

SP-1

AP-2

AvaI

RSB39
consensus C/TC/TCAC/TC/TC/TC/TC/T (Carcamo et al., 1991) were located at -121, -852 and -1055. A large number of putative cis-regulatory elements were identified within the DSC2 promoter, these included potential binding sites for ubiquitous transcription factors AP-1 at -1045, -1178 and -1565 (TGANTA/CA; Faisst and Meyer, 1991), an exact SP-1 consensus site at -489 with other less well conserved SP-1 sites present at -59, -567 and -1858 (GGGCGGAG; Gidoni et al., 1985). A GC-box which may bind SP-1 was present at -383 (G/TA/GGGCGG/TA/GA/GC/T; Briggs et al., 1986). Common to the promoters of other epithelially expressed genes were consensus binding sites for AP-2 at +151, -56, -403 and -491 (C/GC/GCNG/TGGC/GGA/G; Imagawa et al., 1987) and a CK-8-mer motif at -1267 (AANCCAAA; Blessing et al., 1987). There were three consensus binding sites for Pit-1, a factor expressed in the anterior pituitary that belongs to the POU family of transcription factors at -790, -1044 and -1522 (ATGA/GATAA/T; Elsholzt et al., 1990). Sites at -184 and -301 may bind NFkB, a rel-related protein expressed in B cells and activated T cells (GGGA/GNTC/TC/TC; Lenardo and Baltimore, 1989). The significance of these motifs and consensus transcription binding sites remains to be tested.

Unusual sequences within the human DSC2 promoter included an octamer (GGGCAGGG) repeated three times between -533 and -427 and a hexamer (TCTTAA) repeated directly at -250 and -257. The likelihood of such repeats existing so close together is quite low. A number of DNA-binding proteins interact with DNA as dimers making contact with repeated recognition sequences. Examples include a family of nuclear receptors responsive to retinoic acid, thyroid hormone and vitamin D3. This family binds similar DNA sequences which differ by the distance and orientation between the repeated sequences (Näär et al., 1991). Three putative retinoic acid response elements (RAREs) which bind retinoic acid receptors were located at -1475, -529 and -504 (AGGGCA; Umesono et al., 1991).

The site of the first intron was mapped to +69 by sequencing pMDM4 with the primer MDM24, which is within the DNA encoding the putative signal peptide of the translated protein (figure 2.6). The sequence at the 5' end of the first intron conforms to the consensus 5' splice site of 5'-AG GTGAGT-3', where GT are invariant (Padgett et al.,
### Table 2.1: DSC2 Promoter Elements Common to Other Epithelial Promoters.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Common Sequence Element</th>
<th>Position Relative to Transcription Start Site</th>
<th>% Identity of the Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSC2</td>
<td>CCGGCTGTTAATCTCGCCTGCGG</td>
<td>-202</td>
<td>2</td>
</tr>
<tr>
<td>Keratin 14</td>
<td>CCCCCTGTGAAATCAGCCTGCGG</td>
<td>-192</td>
<td>83%</td>
</tr>
<tr>
<td>DSC2</td>
<td>GGATGAGAGG</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td>Keratin 14</td>
<td>GGATGAGAGG</td>
<td>66</td>
<td>100%</td>
</tr>
<tr>
<td>DSC2</td>
<td>ATTAATTTTTAATTTA</td>
<td>-1214</td>
<td>65</td>
</tr>
<tr>
<td>DSG3</td>
<td>ATTAATTTATTATTTA</td>
<td>Not Determined</td>
<td>85%</td>
</tr>
<tr>
<td>DSC2</td>
<td>ACACCGGGGGGCTGGGTTGGG</td>
<td>-179</td>
<td>65</td>
</tr>
<tr>
<td>DSG3</td>
<td>ACGCCAGGGAGGGGCTGGG</td>
<td>Not Determined</td>
<td>78%</td>
</tr>
<tr>
<td>DSC2</td>
<td>AGGAGTGAGTCTT</td>
<td>-255</td>
<td>65</td>
</tr>
<tr>
<td>DSG1</td>
<td>AGGAGTGAAATCTT</td>
<td>Not Determined</td>
<td>92%</td>
</tr>
</tbody>
</table>
1986). The first intron within bovine DSG1, murine E-cadherin, murine N-cadherin and chicken L-CAM is also located within the putative signal peptide (Puttagunta et al., 1994; Ringwald et al., 1991; Miyatani et al., 1992; Sorkin et al., 1988).

A comparison of the human DSC2 promoter was made with the promoters of several epithelially expressed genes using the GCG BETTERFIT program (Table 2.1). Two elements within the human keratin 14 promoter (Leask et al., 1990) of 24bp and 10bp had 83% and 100% identity, respectively, to sequences at -403 and -172 within the DSC2 promoter. Both elements, particularly the 24bp element were present at similar positions relative to the transcription start sites of both genes. The DSC2 promoter was also compared with the sequences available for regions immediately 5' of the human DSG1 and DSG3 genes (M.A. Adams, pers. comm.). The putative human DSG3 promoter contained two sequences of 18bp and 16bp with 77% and 81% identity, respectively, to regions found in the human DSC2 promoter, whilst a comparison with the DSG1 promoter uncovered a 13bp element with 92% identity to a sequence within the DSC2 promoter.

2.2.6 Identification of a CpG Island.

CpG islands are short sequences of about 1.5kb present towards the 5' ends of at least 50% of mammalian genes (Bird, 1986). They are distinct from the bulk of the genome in that they show no depletion of the CpG dinucleotide, are G+C-rich with a content over 60%, and are generally unmethylated. The low frequency of CpGs within the genome is thought to have arisen from deamination and subsequent mutation of methylated CpG residues by repair mechanisms. No example of a CpG island not associated with a gene has yet been published.

The 500bp 5' of the DSC2 translation initiation site was 66% G+C-rich and the CpG dinucleotide occurred at a frequency of 1:11 dinucleotides, in contrast to a frequency of 1:100 for the most 5' 500bp of pMDM7 and 1:25 for the whole of pMDM7. The statistically predicted frequency of a dinucleotide is 1:16. The sequence of the DSC2 promoter also revealed the recognition sites for a number of rare-cutting endonucleases including BssHII, EagI and SacII. These are good indicators of CpG islands because
they have a number of CpG dinucleotides in their recognition sites and are sensitive to methylation.

To examine if BssHII and Eagl could cut the DSC2 promoter, human genomic DNA isolated from whole blood was digested with EcoRI and either BssHII or Eagl and analysed by Southern blot hybridisation to a \(^{32}\)P-labelled HindIII-NsiI fragment of pMDM7 (figure 2.7). Human genomic DNA was used as this would be in its native methylation state unlike cloned DNA prepared from bacterial hosts. EcoRI was used to generate small DNA fragments that could be resolved by conventional agarose gels. Both Eagl+EcoRI and BssHII+EcoRI digests restricted human genomic DNA to yield single bands of 1.4kb and 1.7kb, respectively, which were much shorter than the 3.2kb band of genomic DNA digested with only EcoRI. This indicated that Eagl and BssHII did cleave human genomic DNA at their predicted sites within the DSC2 promoter and therefore must be unmethylated on both alleles. Genomic DNA digested with BssHII or Eagl alone was too large to be resolved and remained in the wells of the agarose gel. The primary structure and methylation-state of the BssHII and Eagl restriction sites of the human DSC2 promoter demonstrates the region to be a CpG island within the genome.

2.2.7 Human DSC2 Promoter Activity in Cultured Cells.

2.2.7.1 Construction of Reporter Constructs.

The DSC2 promoter was initially linked to the lacZ reporter gene to determine its ability to drive expression in cultured cells. lacZ encodes bacterial \(\beta\)-galactosidase whose activity can be measured quantitatively, and is able to covert the X-Gal substrate to an insoluble indigo blue product making it ideal for localising reporter gene expression in vivo. However, the level of endogenous \(\beta\)-galactosidase varies greatly between cell lines. This endogenous activity is predominantly lysosomal and active at low pH, thus partial discrimination can be achieved by performing quantitative assays at >pH 8 (Bronstein et al., 1994; Alam and Cook, 1990). It was apparent that MDCK cells had a high endogenous \(\beta\)-galactosidase-like activity which would make measuring low bacterial \(\beta\)-galactosidase activity inaccurate (Table 2.II).
Figure 2.7: Southern Blot Analysis of the Human DSC2 Promoter Revealing a CpG island. **Bottom:** Restriction map of pMDM7 that contains a 1.9kb *EcoRI-Aval* insert immediately 5' of the DSC2 translation start site, including recognition sites for the methylation-sensitive endonucleases *EagI* and *BssHII* deduced from the nucleotide sequence. **Top:** Southern blot of human genomic DNA digested with *EcoRI*, *EcoRI+EagI*, *EcoRI+BssHII*, *BssHII*, *EagI*, probed with the 1.0kb *EcoRI-NsiI* fragment of pMDM7.
Table 2.II: \(\beta\)-Galactosidase Activity of MDCK Cells.

<table>
<thead>
<tr>
<th>Reporter Construct Transfected into MDCK Cells</th>
<th>(\beta)-Galactosidase Activity(^1) of MDCK Cells (% of pMDM14(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMDM14(^2) (CMV Promoter + lacZ)</td>
<td>100</td>
</tr>
<tr>
<td>pPD46.19 (Promoterless lacZ)</td>
<td>34 (±5)</td>
</tr>
<tr>
<td>Mock - no reporter</td>
<td>35 (±3)</td>
</tr>
</tbody>
</table>

\(^1\)Relative specific activity of \(\beta\)-galactosidase. Transfections were performed in duplicate.

\(^2\)pMDM14 was created by inserting the CMV promoter immediately upstream of the Nuclear Localisation Signal 5’ of the lacZ reporter gene within pPD46.19

The DSC2 promoter was therefore linked to the luciferase reporter gene to measure its activity in transfected cultured cells. The luciferase reporter gene should be ideal for detecting and measuring low levels of transcription from reporter constructs, since there is virtually no endogenous luciferase activity in vertebrate cells. Moreover, the assay is linear for at least 2.5 orders of magnitude until the signal saturates the luminometer (figure 2.8). Reportor constructs were derived from pMDM7 which did not contain the DSC2 translation start codon to ensure that translation of the reporter gene would always initiate from its own start codon and be in frame. To localise regions important in the regulation of DSC2 transcription a series of reporter constructs were created with deletions from the 5’ end of the DSC2 promoter (figure 2.9).

2.2.7.2 Experimental Controls.

Since the efficiency of transfection can vary, pCAT-Control was cotransfected with all reporter constructs. Subsequent to the extraction of the transfected cells, luciferase activity was normalised against CAT activity which then served as a control for the relative efficiency of transfection. To control for the effects of titrating transcription factors, reporter constructs were transfected into cell lines in equimolar ratios. The activity of different constructs was measured relative to pGL2-Control which contained
Figure 2.8: Standard Curve for Luciferase Activity. A series of dilutions was prepared from a 1mg.ml\(^{-1}\) luciferase solution and 1X Reporter Lysis Buffer in duplicate. The activity of 20\(\mu\)l of diluted luciferase solution was then measured under standard assay conditions described in section 6.17.2 and the activity plotted as Relative Light Units vs. Amount of Luciferase Protein.
Figure 2.9: Construction of Human DSC2 Reporter Constructs. All DSC2 reporter constructs were derived from pMDM7. pMDM7: 1.9kb EcoRI-AvaI fragment of pMDM4 blunt-ended and inserted into the EcoRV site of pBluescript SK. Note - the AvaI site is 9bp 5' of the DSC2 translation start site. pMDM15: 1.9kb SacI-SalI inserted into the SacI and XhoI sites of pGL2-Basic. pMDM16: 1.7kb HindIII fragment inserted into the HindIII site of pGL2-Basic. pMDM17: 0.9kb NsiI-SalI fragment inserted into the PstI-SalI sites of pBluescript; 0.9kb SacI-SalI fragment of this intermediate plasmid inserted into the SacI-XhoI sites of pGL2-Basic. pMDM18: 0.5kb AvaII-XhoI fragment blunt-ended and inserted into the SmaI site of pGL2-Basic. pMDM19: 0.3kb EcoNI-XhoI fragment blunt-ended and inserted into the SmaI site of pGL2-Basic. pMDM20: 0.1kb BssHII-XhoI fragment blunt-ended and inserted into the SmaI site of pGL2-Basic. pMDM21: 1.9kb SacI-KpnI fragment inserted into the same sites of pGL2-Basic so that the 5' end of the luciferase gene is adjacent to the 5' end of the DSC2 promoter. pMDM23: 1.9kb SacI-XhoI fragment inserted into the same sites of pGL2-Promoter. pMDM24: 1.9kb SacI-KpnI fragment inserted into the same sites of pGL2-Promoter so that the 5' end of the minimal promoter is adjacent to the 5' end of the DSC2 promoter. pMDM8: 1.9kb PstI-SalI fragment inserted into the same sites of pPD46.21. NLS - encodes the nuclear localisation propeptide. PROM - Minimal promoter of pGL2-Promoter.
the luciferase gene driven by the SV40 early promoter and enhancer, and pGL2-Basic which contained a promoterless luciferase gene and from which reporter constructs were also derived.

2.2.8 Transfection of DSC2 Reporter Constructs into MDCK Cells.

MDCK cells were transfected with DSC2 reporter constructs to enable the activity of the putative promoter to be measured in cultured epithelial cells. MDCK cells are a well characterised canine kidney simple epithelial cell line that express both DSC2 and DSC3, although DSC2 expression is predominant (G.A. Roberts, pers. comm.). The high degree of inter-species conservation between desmocollin and desmoglein orthologues (Buxton et al., 1993) makes it most likely that the genes are controlled by similar mechanisms in different species. The results presented in figure 2.10 are from a single experiment carried out in triplicate. MDCK cells transfected with DSC2 reporter constructs upon six other occasions demonstrated similar trends.

The human DSC2 promoter was active in MDCK cells (figure 2.10). Cells transfected with the longest construct pMDM15 (-1883) had a luciferase activity 27% of pGL2-Control, whilst those transfected with pMDM18 (-525) produced the highest activity of 68%. There was more than a three-fold increase in luciferase expression between pMDM19 (-332) and pMDM18 (-525) from 17% to 68%, respectively. Both pMDM17 (-874) and pMDM16 (-1697) had similar activities of 51% and 53%, respectively. Luciferase activity of cells transfected with pMDM15 (-1883) was half that of those transfected with pMDM16 (-1697). An activity of 5% for pMDM20 (-134) transfected cells was consistent with a predicted transcription initiation site at -201.

2.2.9 DSC2 Promoter Activity is Unidirectional.

Promoters initiate transcription in one direction only. To test this the DSC2 promoter was inverted and cloned upstream of the luciferase reporter gene, creating pMDM21 (figure 2.9). All cell lines transfected with pMDM21 had negligible luciferase activity (figure 2.10). This indicated that the putative promoter did have an unidirectional activity and that the activity observed in 5'-3' constructs was not a consequence of cloning random DNA.
Figure 2.10: Activity of Reporter Constructs Transfected into Cell Lines Relative to pGL2-Control. MDCK and NIH-3T3 cells were transfected with reporter constructs in triplicate. L929 and SVK-14 cells were transfected once.
2.2.10 To Test for Enhancer Elements within the Human DSC2 Promoter.
The DSC2 promoter was inserted in each orientation 5' of a minimal promoter in the pGL2-Promoter vector, creating pMDM23 and pMDM24 (figure 2.9). These were then transfected into MDCK cells and the resulting luciferase activity measured (Table 2.III).

<table>
<thead>
<tr>
<th></th>
<th>pMDM23</th>
<th>pMDM24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity as % of pGL2-Control</td>
<td>55.2% (±2.1)</td>
<td>52.0% (±1.3)</td>
</tr>
<tr>
<td>Activity as % of pGL2-Promoter</td>
<td>220.0% (±6.0)</td>
<td>210.0% (±5.0)</td>
</tr>
</tbody>
</table>

Both pMDM23 and pMDM24 had similar activities within transfected MDCK cells that could be measured as either two-fold greater than the luciferase expression of pGL2-Promoter, or half that produced by pGL2-Control. This relatively small increase in luciferase transcription did not reveal any strong enhancer-like activity and may have been a result of the inserted sequences abolishing the repressive effect of the vector's poly-A sequence upon the minimal promoter.

2.2.11 Transfection of DSC2 Reporter Constructs into Other Cell Lines.
To test for cell-specific regulatory elements within the cloned DSC2 promoter, reporter constructs were transfected into murine fibroblastic NIH 3T3 cells in triplicate (figure 2.10). The trends observed were consistent with those from two other transfections. Reporter constructs were also transfected into Jurkat-J6 cells derived from human T-cells. However the efficiency of transfection by electroporation into this cell line was too low to quantitate luciferase and CAT activities. L929 murine fibroblasts and the human epithelial SVK-14 cell line were transfected with reporter constructs once (figure 2.10).

All the DSC2 reporter constructs, except for the inverted construct pMDM21, were active in NIH 3T3 cells (figure 2.10). The longest construct pMDM15 generated a luciferase activity 23% of the SV40 early promoter and enhancer within pGL2-Control and was the most active DSC2 promoter reporter construct in this cell line. pMDM20 and pMDM19 had similar activities within NIH 3T3 cells as MDCK cells of 6% and
12%, respectively (compared with 5% and 17% in MDCK cells). The most active reporter construct within MDCK cells, pMDM18, had only 21% the activity of pGL2-Control in NIH 3T3 cells. Both pMDM17 and pMDM16 has similar activities to each other of 13% and 10%, respectively.

Expressing the activity of cells transfected with reporter constructs as a percentage of pGL2-Control activity allowed trends between the reporter constructs to be observed and their strength relative to the SV40 early promoter and enhancer to be measured. However to compare reporter construct activity as fractions of pGL2-Control activity between cell lines presupposes that pGL2-Control activity is constant within different cell lines. To take this into account luciferase activity was also expressed as a multiple of pGL2-Basic activity in transfected cells (figure 2.11). This was possible because pGL2-Basic did generate some luciferase activity in transfected cells. Luciferase activity generated by pGL2-Control did appear to vary between cell lines and ranged from 44 times greater than pGL2-Basic activity in L929 cells to 110 times greater in NIH 3T3 cells and 164 times greater in MDCK cells.

2.2.12 DSC2 Reporter Constructs to Study Regulatory Elements In Vivo.

DSC2 expression is more complex than just being cell-specific; it is also sensitive to temporal and spatial controls in tissues such as stratifying epithelium and developing trophoderm (Nuber et al., 1995; Collins et al., 1995). At present these systems are difficult or not feasible to model using cultured cells. To localise DNA elements sensitive to factors which generate the specific expression of the DSC2 gene in vivo, the cloned promoter and nested deletions of it were linked to the lacZ reporter gene. The lacZ reporter construct with the longest portion of the DSC2 promoter, pMDM8 (figure 2.8) was subsequently used to create transgenic mice by injection into the pronuclei of fertilized oocytes. A number of mice have been identified by Southern analysis to have the pMDM8 construct integrated into their genome (M.A. Adams, pers. comm.) and are currently under investigation (R. Ali, pers. comm.).
Figure 2.11: Activity of Reporter Constructs in Cell Lines Relative to pGL2-Basic.
2.3 DISCUSSION

The desmosomal cadherins probably evolved through duplication and subsequent mutation of a common ancestral gene. Similar evolutionary processes where genetic changes result in the inactivation of a gene may produce pseudogenes (Rosenberg et al., 1988). Pseudogenes may also be generated by the reverse transcription and integration of a mRNA into the genome, creating what is termed as a processed pseudogene (Kulesh and Oshima, 1988). A number of examples have been identified within the human genome, including keratins 14, 16, 17 and 18 pseudogenes (Rosenberg et al., 1988; Savtchenko et al., 1990; Troyanovsky et al., 1992; Kulesh and Oshima, 1988). Although there is no evidence to suggest that desmocollin pseudogenes exist, this remains a possibility and the DSC2 cDNA probe may have hybridised to a DSC2-like gene during the screening of the λFix II library. It is unlikely that there is more than one DSC2 gene because the 3' DSC2 cDNA probe hybridised to a single band upon a Southern blot of human genomic DNA digested with EcoBl or BamHl (figure 2.4A). If there were DSC2-like genes that hybridised to the DSC2 cDNA it would be unlikely that single bands would be observed, since the likelihood of different loci containing the same restriction sites at similar distances apart, particularly within introns, should be quite small. The ability of the cloned DSC2 promoter to recruit transcription factors that drive the expression of a reporter gene in transfected cells (figure 2.10), implied that it was part of an actively transcribed gene. Furthermore the cloned DSC2 promoter represented sequences that constitute a CpG island within the genome (figure 2.7) which are usually associated with transcribed genes. For instance, the 5' region of the active human keratin 17 gene comprises a CpG island, whilst those of the two human keratin 17 pseudogenes are fully methylated (Troyanovsky et al., 1992). Thus there is no evidence from my results that the cloned DSC2 promoter is part of a DSC2 pseudogene.

2.3.1 Site of Transcription Initiation.

RNase protection experiments predict the transcription initiation site of the DSC2 gene to be 201bp upstream of the translation initiation codon. However this prediction does not take into account the small difference in the mobility of the protected single stranded RNA product and that of the single stranded DNA molecules used to size the RNA within
a sequencing gel. Zacksenhaus et al. (1993) measured the mobility of riboprobes of known sizes within a sequencing gel and calculated that single stranded RNA migrates at 95.75% the rate of single stranded DNA. This means that the predicted site could be between -192 and -201.

TATA-less promoters may contain an initiator sequence that encompasses the transcription start site and directs the assembly of the transcription factors which form the initiation complex. This DNA element is poorly conserved within eukaryotic promoters and has a consensus sequence of C/TC/TCAC/TC/TC/TC/TC/T (Carcamo et al., 1991), where transcription starts at A. The murine E-cadherin promoter contains such an element at its predicted transcription start site (Behrens et al., 1991). Three consensus initiator elements are present within the DSC2 promoter, with the element at -131 closest to the site predicted by RNase protection experiments.

A major transcription initiation site within the region between -201 and -131 would be consistent with the low luciferase activity of cells transfected with pMDM20, since this reporter construct contains sequences extending no more than to -134 (figure 2.8). However it has yet to be shown that the cloned DSC2 promoter within the transfected reporter constructs initiates transcription at the same site as the endogenous gene within human keratinocytes. A transcription initiation site between -131 and -201 for human DSC2 compares with transcription start sites of -124 and -128 for human and murine E-cadherin, respectively (Bussemakers et al., 1994b; Faraldo and Cano, 1993) and -68 for murine P-cadherin (Faraldo and Cano, 1993). All these three latter sites are within a conserved initiator sequence.

2.3.2 Potential Regulatory Transcription Factors.
Proteins which interact with the DSC2 promoter form part of the regulatory cascade that culminates in DSC2 transcription. Since transcription factors bind to specific sequences of DNA, defined consensus binding sites may easily be recognised and amongst the many potential transcription factor binding sites and motifs found within the DSC2 promoter, there were several that were particularly interesting. These included potential sites for the proteins AP-2, SP-1, AP-1, NFκB, Pit-1, and the CK-8-mer motif.
There were a total of four AP-2 sites clustered within 350bp of the predicted transcription start site, one of which was located within the first intron. AP-2 binding sites are found in the promoters of a large number of epidermally expressed genes, including keratins K1, K5, K6, K14, K16 and K17 (Leask et al., 1991; Snape et al., 1991). AP-2 itself is an activator of transcription with a tissue-specific pattern of expression. In Xenopus an AP-2-like factor appears during the embryonic development of epidermis and is predominant in the adult skin (Snape et al., 1991). In the murine embryo AP-2 is expressed within cells of epidermal and neural crest lineages (Mitchell et al., 1990) and closely precedes and mimics expression of epidermal genes such as keratins 5 and 14 (Byrne et al., 1994). A function for AP-2 in epidermal gene regulation is further implied by in vitro studies where an AP-2 expression vector is cotransfected with CAT reporter genes driven by either a K5 or K14 promoter into hepatocytes. These cells do not normally express the K5 or K14 genes, but do express the CAT reporter gene when AP-2 is expressed (Byrne et al., 1994). However AP-2 is not sufficient to induce epidermal specificity since HeLa cells which express AP-2 do not express epidermal markers, such as the keratin 14 gene (Leask et al., 1991).

The CK-8-mer motif at -1267 was initially defined by Blessing et al. (1987) as a sequence conserved in the promoters of keratins predominantly expressed in epidermis and involucrin. Although a 90bp fragment of the bovine keratin 4 promoter containing this element provided a fourfold increase in the expression of a CAT reporter gene (Blessing et al., 1989), it has yet to be shown that it binds cell-specific factors (Cripe et al., 1990). Furthermore Ishiji et al. (1992) were unable to demonstrate function of a CK-8-mer within the E6/E7 promoter of HPV 16. Thus this motif may either indicate other less well conserved regulatory elements or require other DNA elements present for it to function.

Pit-1 is predominantly expressed within the anterior pituitary and has been proposed to be important in the terminal differentiation of pituitary cells. It belongs to the POU family of transcription factors which bind to very similar DNA sequences, characteristically the canonical octamer ATGCAAAAT (Rosenfeld, 1991; Scholer, 1991). Thus it is possible that POU proteins other than Pit-1 might interact with the DSC2
promoter. POU proteins are distinguished by a 147 to 156 amino acid POU domain, consisting of a 75 to 82 residue POU-specific segment and a 60 to 62 residue POU homeodomain separated by a short linker segment of more variable size and sequence. The POU domain may be involved in site-specific DNA binding and protein-protein interactions. Outside the POU domain the proteins are hypervariable which probably contributes to differences in their regulatory effects. Several members have been cloned using degenerate POU domain-specific probes and found to be enriched within epidermis, namely Oct-6, Skn1a and Skn1i (Fans et al., 1994; Anderson et al., 1993). Oct-6 is expressed in epidermis and other stratified epithelia. Its function as a regulator of epidermal gene expression was demonstrated by cotransfecting an Oct-6 expression vector with CAT reporter genes linked to the promoters of either human keratin 5 or 14 into cultured keratinocytes. This resulted in a down-regulation of CAT expression. Skn1i and Skn1a may be alternative transcripts of the same gene. Skn1i expression is predominantly suprabasal in murine epidermis. Cotransfection of a Skn1a expression vector with a luciferase reporter gene linked to the keratin 10 promoter into HeLa cells results in a 15-fold increase in luciferase expression (Anderson et al., 1993).

The DSC2 promoter contains consensus binding sites for the ubiquitous factors AP-1, SP-1 and NFκB which are all activators of transcription. SP-1 is able to self-associate, bringing together distant DNA segments, and may therefore stimulate transcription by bringing other activators together (Mastrangelo et al., 1991). SP-1 binds to the keratin 16 promoter (Magnaldo et al., 1993), but may not be essential for epithelial expression since its consensus binding site is absent from other epithelial promoters such as the keratin 5 promoter (Eckert and Rorke, 1988). NFκB is a ubiquitous transcription factor found in keratinocytes. It is involved with tissue-specific and differentiation-dependent expression of several defence molecules including β-interferon, TNFα, TNFβ, κ light chain of immunoglobulin and T-cell receptor. The rabbit keratin 3 promoter also contains a NFκB consensus binding site that may actually bind this factor (Wu et al., 1993).

Apart from the CK-8-mer and AP-2 binding site, no other motifs appear to be common to epithelial promoters; even a comparison between the promoter sequences of the
coordinately expressed keratins 5 and 14 revealed no unique common element. Epithelial genes may therefore be specifically activated by a number of ubiquitous factors that bind epithelial promoters in a specific pattern. The involvement of multiple factors would offer a mechanism to finely tune gene expression in the many types of epithelia. It could also be envisaged that such ubiquitous factors may be modulated by interactions with other proteins within epithelial cells. Indeed, Mack and Laimins, (1991) identified a keratinocyte-specific factor that interacts with AP-1 to direct HPV 18 gene expression.

The E1a gene product of the adenovirus may also interact with transcription factors that activate epithelial genes. This protein does not appear to bind directly to DNA but can induce epithelial genes, including DSC2 and DSG2, to be activated when it is transfected into certain non-epithelial cell lines (Frisch, 1994). Recent work has begun to uncover mechanisms of how E1a may function. E1a binds to the proteins p300 and CBP and one effect of this interaction is to inhibit their coactivator function upon the transcription factor CREB (Arany et al., 1995; Lundblad et al., 1995). One effect of E1a may therefore be to inactivate cAMP-responsive gene transcription. Since some cell differentiation pathways are regulated by transcription factors responsive to cAMP (Boshart et al., 1991; Jones et al., 1991), E1a may deregulate tissue-specific expression which would also imply that the epithelial phenotype observed by Frisch, (1994) is a default pathway. An E1a-like activity exists in the murine blastocyst (Suemori et al., 1988) which suggests that there is a cellular analogue.

2.3.3 Activity of the DSC2 Promoter in Transient Transfection Experiments.
The human DSC2 promoter is active in MDCK cells and this activity is relatively strong compared to the SV40 early promoter and enhancer (figure 2.10). A threefold increase in luciferase activity between MDCK cells transfected with pMDM19 and pMDM18 suggests that there are activating elements within the 193bp between -332 and -525 (figure 2.10). This region of the DSC2 promoter contains potential binding sites for the activators AP-2 and SP-1 (figure 2.6), as well as elements with identity to the human K14, DSG1 and DSG3 promoters which suggests that they are conserved and important in epithelial transcription. In the keratin 14 promoter the 24bp element common to DSC2 and keratin 14 is located within a 110bp activating region of the K14 promoter, adjacent
to a functional AP-2 binding site (Leask et al., 1990). The 193bp region may also harbour binding sites for as yet uncharacterised transcription factors. MDCK cells transfected with pMDM17 (-874) and pMDM16 (-1697) have similar luciferase activities (figure 2.10). It is possible that this belies that the 823bp region of non-overlap between the two reporter constructs may contain binding sites for both activating and inhibitory factors. This could result in the small change in net activity observed. To test for clustered regulatory elements a series of reporter constructs with shorter nested deletions could be made, thus enabling the promoter to be mapped with a greater resolution. The decrease in activity observed between MDCK cells transfected with pMDM16 and pMDM15 may reflect inhibitory elements within the 186bp between -1697 and -1883.

The supposition that the pGL2-Control had a constant level of activity within different mammalian cell lines did not appear to be true. Although all the cell lines tested were transfected with the same preparation of pGL2-Basic and pGL2-Control, the activity of different cell lines transfected with pGL2-Control varied when measured relative to pGL2-Basic (figure 2.11). The trends observed in figure 2.10 and 2.11 however are similar, specifically in that reporter constructs larger than pMDM18 are more active in MDCK cells than NIH 3T3 cells.

DSC2 reporter constructs were active in fibroblastic NIH 3T3 cells which do not express endogenous DSC2. NIH 3T3 cells and MDCK cells transfected with pMDM20 (-134) and pMDM19 (-332) have similar low luciferase activities (figure 2.11). However, NIH 3T3 cells transfected with pMDM18 (-525) have less than a third of the activity of MDCK cells transfected with this reporter construct (figure 2.11) which implies that the activating elements identified between -332 and -525 are also cell-specific. These results are very similar to those of Sawamura et al. (1994) who observed a three-fold greater CAT activity in PAM 212 cells (a murine transformed keratinocyte cell line) transfected with a reporter construct containing the CAT reporter gene driven by the bullous pemphigoid antigen I promoter, than in NIH 3T3 cells transfected with the same construct.
The seemingly high luciferase activities of L929 and SVK-14 cells transfected with DSC2 reporter constructs when activity was measured relative to pGL2-Control (figure 2.10) were probably due to a low pGL2-Control activity within these cell lines. In the case of the L929 cells, when activity was expressed as a multiple of pGL2-Basic activity (figure 2.11) the activity induced by pMDM18 and other longer DSC2 reporter constructs was at the most half of the activity of MDCK cells transfected with the same constructs. This was consistent with the DSC2 promoter containing cell-specific elements.

It is not known which desmosomal cadherins SVK-14 cells express and at what levels, although it is likely that they do express DSC2 since this appears to be ubiquitous to desmosome-bearing tissues and cell lines (Nuber et al., 1995). A comparison between MDCK cells and SVK-14 cells of reporter construct activity measured as a multiple of pGL2-Basic activity showed that the DSC2 promoter is more active in MDCK cells, especially in cells transfected with reporter constructs larger than pMDM18 (figure 2.11). This may mean either SVK-14 cells do not express DSC2 at all or only at low levels, otherwise the luciferase activities of transfected MDCK and SVK-14 cells would be similar. Unlike in the MDCK and L929 cell lines, luciferase activity increased sharply between SVK-14 cells transfected with reporter constructs pMDM18 (-525) and pMDM17 (-874) (figure 2.10). The suggestion in these very preliminary findings would be that the activating element between -332 and -525 that function within MDCK cells, do not work in SVK-14 cells where elements between -525 and -874 enhance reporter expression. This may be attributed to peculiarities of each cell line. Alternatively, if SVK-14 cells do indeed express the human DSC2 gene, differences observed between these cells and MDCK cells transfected with human DSC2 reporter constructs may arise from different mechanisms involved in the expression of the DSC2 gene within each species. These transfection experiments into SVK-14 cells need to be repeated to be substantiated.

The cloned human DSC2 promoter does impart some tissue-specificity, since reporter constructs containing promoter sequences that extended beyond -332 were expressed at higher levels within MDCK cells. However reporter constructs were also active in NIH 3T3 cells. Several reasons may account for the activity of the DSC2 promoter in cell lines which do not express the endogenous DSC2 gene. Regulatory elements, particularly
those that inhibit inappropriate DSC2 promoter activity may be missing from the reporter constructs; these may be further upstream, within the first intron or even 3' of the gene. Also, regulatory elements within the constructs may require a chromosomal context for their function. Both methylation and chromosomal structure may affect genomic DNA and be the reason why some genomic sequences, such as the human keratin 18 promoter, can sometimes exhibit tissue-specificity in transgenic mice, but lack such specificity in \textit{in vitro} transfection assays (Neznanov \textit{et al.}, 1993; Abe and Oshima, 1990; Kulesh and Oshima, 1988). Another explanation may be that cell lines may express activating factors not normally associated with the cell-type from which they were derived or fail to express the appropriate repressor. Since the DSC2 promoter contains many consensus transcription binding sites, some of these may become functional in a particular cellular environment and give rise to an apparent promiscuous expression.

It has yet to be shown that transcription initiates from the same site in MDCK cells as the endogenous gene within human keratinocytes, which would confirm that transcription initiation is the same in the transgene model and endogenous gene.

In the short term work should ratify the presence of cell-specific activators between -332 and -525 by further transient transfection experiments of reporter constructs into other cell lines. It may also be useful to use a human cell line known to express DSC2 such as HaCaT cells (personal observations, and Nuber \textit{et al.}, 1995). Larger reporter constructs may also be made using \textlambda 2.DSC2 or \textlambda 4.DSC2 to include more 5' sequences and sequences within the first intron to investigate other regulatory elements close to the 5' end of the DSC2 gene.

2.4 Major Findings.

The human DSC2 promoter has been cloned and characterised. A transcription start site was predicted to be 201bp 5' of the translation start site. The promoter constituted a CpG island within the genome and contained consensus binding sites for transcription factors proposed to be important in epithelial development and differentiation, namely AP-2 as well as POU proteins. Sequences between -332 and -525 had identity with the human keratin 14, DSG1 and DSG3 promoters. This 193bp region also held cell-specific
elements that increase the DSC2 promoter’s activity in MDCK cells, but not in NIH 3T3 cells.

2.5 Future Work - Identify Proteins that Interact with the DSC2 Promoter.

This work has uncovered potential binding sites for transcription factors within the DSC2 promoter that could contribute to an epithelial-specific expression. To test if these specific factors are able to affect DSC2 expression, expression vectors encoding transcription factors such as AP-2 or Oct-6 could be co-transfected into non-epithelial cultured cells. An enhancement of DSC2 promoter activity when it was transfected with a particular transcription factor expression vector would infer a role for the factor in DSC2 expression. However such an experiment could not reveal at which part of the regulatory cascade culminating in DSC2 transcription the factor may act. Direct DNA:protein interactions would have to be investigated using band shift assays.

Band shift assays demonstrate protein binding to DNA, since the migration of DNA through an agarose gel is retarded when it interacts specifically with protein. Thus these assays are useful in localising regions of protein interaction and allow any binding of purified transcription factors such as AP-2, Oct-6 to fragments of the DSC2 promoter to be demonstrated directly. Moreover, uncharacterised factors which bind to the DSC2 promoter could be fractionated from the nuclear extracts of cells expressing DSC2, using band shift experiments as an assay for DNA:protein interactions. Since the fragment between -332 and -525 may contain cell-specific activating elements it is likely to bind transcription factors and could be used in initial experiments. It is possible that alternative tissue-specific transcription factors may bind to the DSC2 promoter in tissues such as heart and epithelium. Thus the source of nuclear extract could be important in investigating what factors bind to the promoter in particular tissues. HaCaT cells are a potential source of human epithelial transcription factors, since they express DSC2 (unpublished data, and Nuber et al., 1995) and are more easily cultured than alternatives such as human primary keratinocytes. Once the region of DNA:protein interaction has been localised and the protein identified and purified, the nucleotides of the binding site could be defined using footprinting experiments and mutagenesis experiments. Consequently a consensus would be established and recognised in other DNA sequences.
An alternative approach to isolating proteins that interact with the DSC2 promoter would be to screen an expression library with a fragment of the promoter. Which fragment of the DSC2 promoter to use as a probe would have to be considered, since larger probes would increase the levels of non-specific hybridisation. Regions likely to bind interesting factors could be localised by mapping of the promoter for important regulatory elements with a higher resolution by transient transfection of reporter constructs with shorter nested deletions into cultured cells. The ability of these elements to bind protein from crude nuclear extracts could then be demonstrated using band shift assays and these regions of DNA subsequently used to probe an expression library.
Chapter Three - Screening for the Human DSC1 Promoter

3.1 Introduction
3.2 Results
3.2.1 Screening a λFix II Human Genomic Library
3.2.2 Construction and Screening of Vectorette Genomic Libraries
3.2.3 Screening the ICI YAC Library by PCR
3.3 Discussion
3: SCREENING FOR THE HUMAN DSC1 PROMOTER

3.1 INTRODUCTION
The expression of the DSC1 gene is much more restricted and tissue- and differentiation-specific than that of the apparently ubiquitous DSC2 gene; DSC1 is expressed in lymph nodes and stratified epithelia, whilst DSC2 is expressed within all tested desmosome-bearing tissues (Nuber et al., 1995). A comparison of the DSC1 and DSC2 5' regulatory sequences would initiate an investigation into the mechanism that causes the differential expression of these genes. Furthermore the cloned DSC1 promoter could also provide a useful reagent for targeted expression of genes to the suprabasal layers of epidermis.

3.2 RESULTS
When initial screening for the human DSC1 promoter began only part of the human DSC1 cDNA had been isolated. It had been cloned from an oligo-dT primed cDNA library and did not extend as far 5' as to the translation initiation start codon (King et al., 1991). From comparisons with the bovine DSC1 cDNA which did include the translation start codon, it was estimated that approximately 1kb of cDNA was needed to reach the human DSC1 translation start codon (figure 3.1). Any genomic clones isolated using what was the most 5' human DSC1 cDNA as a probe may not contain 5' regulatory elements because of intron sequences between the probe and sequences upstream of the translation initiation site. The initial DSC1 cDNA probe was therefore not ideal.

3.2.1 Screening a λFix II Human Genomic Library.
To isolate DNA containing the human DSC1 promoter a human genomic library in λFix II was screened by hybridisation to the most 5' cloned sequence of the human DSC1 cDNA, a 2.7kb EcoRI-BamHI fragment of pJA87T3 (figure 3.1). More than $10^6$ plaques were screened but none hybridised to the probe. In order to screen the human genomic library with a more 5' DSC1 cDNA probe, a 1.7kb EcoRI-SacI fragment from the bovine DSC1 cDNA (Collins et al., 1991) was used. A comparison of the bovine and human DSC1 cDNAs showed them to have 80% identity within their sequenced open reading frames. Hybridisation was carried out at 55°C and washing conditions of reduced
Figure 3.1: Schematic Representation of the Human DSC1a cDNA. DSC1a cDNA and subclones pJA87T3 and pMDM6. The 5' extent of the cDNA represents the extent of the clone K24 (King et al., 1993a). The shaded box represents the open reading frame and does not contain the putative start of translation. The arrowhead indicates the alternative splicing site. pJA87T3 was used in initial screening for 5' human genomic DSC1 clones and extended as far 5' as the cloned DSC1 cDNA at that time. pMDM6 was created by PCR using the DSC1 specific primer RSB86 and the λ primer RSB87, the product was restricted at internal EcoRI and PstI sites and subcloned into pBluescript. Both IK29 and MDM6 were used individually as the specific primer for Vectorette PCR. JA141 and JA142 were used to screen the ICI YAC library by the PCR.
Human DSC1a cDNA

Diagram showing the location of restriction enzymes and cloning sites within the cDNA sequence, with annotations for RSB87, IK29, RSB86, JA141, JA142, EcoRI, PstI, SalI, and BamHI, along with pMDM6 and pJA87T3 vectors.
stringency (0.6X SSC) were used to accommodate differences in the human and bovine cDNAs and interspecies hybridisation. Above the increased level of background no plaques hybridised to the 5' bovine DSC1 cDNA probe. A phage clone, K24, was isolated which extended the human DSC1 cDNA 791bp more 5', although it still did not reach the translation start codon (King et al., 1993a). To enable direct cloning of the region from the phage into a plasmid vector, a PCR using the K24 clone as a template and the internal λ primer RSB87 and the DSC1 cDNA primer RSB86 (figure 3.1) was performed. The product was restricted at internal endonuclease sites with EcoRI and PstI and the resulting 450bp fragment subcloned into the same restriction sites of pBluescript to generate pMDM6. When the λFix II library was rescreened a number of plaques hybridised to pMDM6 in the primary screen, although these turned out to be false positives in subsequent rounds of plaque purification. The λFix II library did contain clones representing at least the 3' portion of the DSC1 gene, since a PCR upon the library using the primers JA141 and JA142 (figure 3.1) produced a single product of 911bp. The primers JA141 and JA142 were used as these were known to yield a product with genomic DNA.

3.2.2 Construction and Screening of Vectorette Genomic Libraries.
Vectorette PCR was used to screen a human genomic library and clone DNA 5' of the DSC1 cDNA without the need of a 5' cDNA probe. This technique allows the amplification of regions of DNA adjacent to those of known sequence. Four vectorette libraries were constructed from human genomic DNA isolated from whole blood using the restriction enzymes EcoRI, HindIII, ClaI and BamHI. Vectorette PCR was carried out upon each library using either MDM6 or IK29 (figure 3.1) as the specific primer. Regardless of attempts to optimise the PCR conditions, by using various annealing temperatures, lengths of time for annealing and product extension, hot-start PCR to aid specific annealing, no specific products which would hybridise to the pMDM6 probe upon Southern analysis were generated from any of the vectorette libraries.

3.2.3 Screening the ICI YAC Library by PCR.
As an alternative to the human genomic λFix II library which may not have contained the DSC1 promoter region, the ICI YAC library was screened using PCR. This library is
maintained as part of the Human Genome Mapping Programme by the MRC. The YACs carry human genomic DNA and have an average insert size of 300kb with an estimated 10% chimaeric for different human chromosomes. Primary screening was carried out upon forty YAC pools using the primers JA141 and JA142 which amplify the 3' untranslated region of the human DSC1 cDNA to yield a 911bp product. It is likely that a YAC clone will carry both the promoter and 3' untranslated region of the DSC1 gene. Out of the forty pools screened three primary pools - 8, 14 and 24, yielded a product of 911bp, with the band of pool 8 being the brightest on a stained agarose gel (figure 3.2). Surprisingly, none of the nine secondary pools which had been generated from primary pool 8 gave a product with the PCR under the same conditions. A blot of DNA from all the individual YAC clones of primary pool 8 was obtained from the HGMP and screened by hybridisation to the ^32P-labelled insert of pMDM6. Putative clones that appeared to hybridise to the probe were identified and the individual YACs obtained. Further analysis of these clones by the PCR with JA141 and JA142 proved to be negative.

3.3 DISCUSSION

A number of different strategies were used to isolate human genomic clones of the DSC1 promoter region. A PCR upon the λFix II library with the primers JA141 and JA142 confirmed the presence of the 3' portion of the DSC1 gene within this library. However both the bovine and human cDNA clones used to probe the λFix II library contained sequences 5' of this region, and none hybridised to any plaques when the library was plated out. This could have been because the 5' portion of the DSC1 gene was not represented within the library or when the library was plated out.

Vectorette PCR may have been unsuccessful for a number of reasons. The distance between restriction sites for the endonucleases used to make each library may have been too great for a product to be amplified efficiently. Alternatively, the Vectorette primer may have hybridised non-specifically during the first and subsequent cycles of the PCR. Since specific primer sequences were taken from the cDNA sequence, they may have crossed exon/intron boundaries and therefore may not have hybridised to the genomic template. It may be more efficient to use an endonuclease that cuts more frequently than
Figure 3.2: Primary Screen of the ICI YAC Library by PCR. Each of the forty pools was screened using the primers JA141 and JA142. Lanes M, 1kb λ ladder markers; lane +, positive control using human genomic DNA as the template for the PCR; lane -, negative control without DNA template; lanes 1 to 40, PCRs using each of the forty fractions of the YAC library as templates.
the hexamer endonucleases used, to reduce the potential distance between the vectorette and target sequences. An enriched source of the template such as a YAC for the single specific primer may also be required for this technique to work. Specific DNA sequences can be amplified directly from genomic DNA, as can be seen with PCRs using JA141 and JA142 (figure 3.2). However since the vectorette strategy uses only one specific primer this may halve the probability of the reaction priming from the target sequence.

No YAC clones with the 3' portion of the DSC1 gene could be identified within the ICI YAC library. PCR using the secondary subpools of the primary pool 8 did not yield any product that could be visualised upon a stained agarose gel. This may have been because the target region was present in very low amounts. Subsequent work where two consecutive PCRs were carried out, using the product of the first as the template for the second reaction generated a product of 911bp. Individual YAC clones containing the DSC1 gene have now been isolated (C.M.E. Cowley and R.S. Buxton, pers. comm.).

YACs are difficult to subclone directly because they are maintained in low copy numbers within yeast cells. They are also difficult to manipulate and need to be supported within agarose blocks to prevent the DNA from shearing. Individual YAC clones should be a more suitable source of genomic DNA for vectorette PCR than DNA purified from whole blood because they are an enriched source of target sequences. This approach may allow the cloning of genomic DNA 5' of the DSC1 gene.
Chapter Four - Isolation and Characterisation of Murine Desmocollin and Desmoglein Clones

4.1 Introduction
4.2 Results
4.2.1 Isolation of Murine Genomic DSC2 Clones
4.2.2 Sequence Analysis of a Murine Desmoglein cDNA
4.3 Discussion
4.3.1 Murine Genomic DSC2 Clone
4.3.2 Future Work - Targeted Disruption of DSC2 Alleles
4: ISOLATION AND CHARACTERISATION OF MURINE DESMOCOLLIN AND DESMOGLEIN CLONES

4.1 INTRODUCTION

The desmosomal cadherins are expressed in a complex and specific pattern that is likely to be functionally significant. An important approach to studying gene function in vivo is the analysis of cells or organisms homozygous for loss-of-function mutations. Embryonic stem (ES) cells have been used to introduce targeted mutations into the germ line of mice (Larue et al., 1994). In addition to their use for functional studies in vivo, ES cells provide a powerful in vitro model of embryonic development (Baribault and Kemler, 1989). In suspension culture, ES cells differentiate to organised structures known as embryoid bodies (EBs) containing an outer layer of endoderm and an inner ectoderm layer separated by a basal lamina. These EBs expand into large cystic structures reminiscent of the visceral yolk sac both morphologically and biochemically. Ultrastructural analysis shows an apparently normal epithelium with desmosomes and tight junctions. Since DSC2 expression is induced in trophectoderm in vivo (Collins et al., 1995) and appears to be ubiquitous in desmosome-bearing cells (Nuber et al., 1995), it is likely to be induced in vitro as ES cells differentiate to EBs.

As yet the specific functions of the desmocollin and desmoglein isoforms have not been resolved. The targeted disruption of a desmosomal cadherin gene may yield some answers as to the importance of particular isoforms throughout the development and differentiation of tissues. Moreover, the diversity of desmosomal cadherin function may be investigated to see if the isoforms' functions are sufficiently similar to compensate for the loss of one of their number, and if, like the keratins, the desmocollins and desmogleins form obligate heterodimers. A mouse that displays a phenotype may also provide a useful model of human disease. Preliminary to the study of desmosomal cadherin function in early embryonic development by the inactivation of desmosomal cadherin alleles, a murine DSC2 genomic clone was isolated and characterised and a murine desmoglein cDNA sequenced.
4.2 RESULTS

4.2.1 Isolation of Murine Genomic DSC2 Clones.

A murine genomic library in the λFix II vector was screened by hybridisation to a 900bp EcoRI fragment of the murine DSC2 cDNA (Buxton et al., 1994). This cDNA clone encoded part of the extracellular domain of Dsc2. Since the ES cells commonly used are derived from 129 strain of mice, it was important to use a genomic library that had been prepared using 129 mouse DNA. This would optimise homologous recombination between disrupted cloned genomic DNA and chromosomal DNA during targeted inactivation of DSC2 alleles within ES cells. Out of approximately $10^6$ plaques screened one clone, λM.DSC2, with an insert of approximately 18kb was identified and purified. A gene may be disrupted by insertion of heterologous sequences, usually a selectable gene, into an exon (Zimmer, 1992). The murine DSC2 cDNA contained a single NheI site. To investigate the presence of this NheI site within the genomic clone, λM.DSC2 DNA was cut with NheI, blotted and probed with the $^{32}$P-labelled murine DSC2 cDNA. A band of 1.2kb and one greater than 26kb hybridised to the cDNA probe (figure 4.1).

4.2.2 Sequence Analysis of a Murine Desmoglein cDNA.

A murine desmoglein cDNA had been isolated by PCR upon cDNA prepared from newborn mice using degenerate primers of known desmoglein sequences in the region encoding a cytoplasmic portion of the protein (G.N. Wheeler, pers comm., and Buxton et al., 1994). The 681bp product was sequenced to identify which type of desmoglein it represented and to provide data for future targeted disruption of this desmoglein. The murine desmoglein cDNA sequence was compared with human DSG1, DSG2, DSG3 and bovine DSG1 cDNAs using the GCG programs, GAP and PILEUP. The GAP program creates an alignment of two sequences with the largest number of matched bases and calculates the percentage identity (Table 4.1), whilst the PILEUP program displays the best alignment of several sequences (figure 4.2). Therefore from a comparison of desmoglein nucleotide sequences, the isolated murine desmoglein represents a DSG1 isoform. This cDNA could be used to screen a genomic library for reagents to inactivate the murine DSG1 gene.
Figure 4.1: Southern Analysis of the Murine Genomic DSC2 Clone. (A) Southern blot of \( \lambda M.DSC2 \) digested with \( NheI \) or \( SalI \) probed with a 450bp \( EcoRI \) fragment of murine DSC2 cDNA. (B) Tentative map of \( \lambda M.DSC2 \) with the possible position of two \( NheI \) sites marked by arrowheads, the site towards the left of the page is likely to be present within an exon. \( \lambda \) FixII vector sequence is represented by the shaded boxes.
Vector --------------- Insert ---------------- Vector

- 26kb
- 13kb
- 1.2kb

Sal I
Nhe I

Nhe I
Nhe I

Vector ─────────── Insert ─────────── Vector

4kb
Figure 4.2: Nucleotide Sequence of the Murine DSG1 Clone Compared with other Desmoglein cDNAs. All the sequence is from the cytoplasmic part of the molecules and compared using GAP of the GCG programs. Nucleotides with identity to murine DSG1 are shaded. Murine DSG clone (mdsg), human DSG1 (hdsg1), bovine DSG1 (bdsg1), human DSG3 (hdsg3) and human DSG2 (hdsg2).
**Table 4.1: Comparison of DSG cDNAs with the Murine cDNA.**

<table>
<thead>
<tr>
<th>Percentage Identity of cDNAs</th>
<th>Human DSG1</th>
<th>Human DSG2</th>
<th>Human DSG3</th>
<th>Bovine DSG1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murine DSG cDNA</td>
<td>82%</td>
<td>52%</td>
<td>49%</td>
<td>80%</td>
</tr>
</tbody>
</table>

**4.3 DISCUSSION**

**4.3.1 Murine Genomic DSC2 Clone.**

The clone λM.DSC2 was isolated from a murine genomic library using part of the mouse DSC2 cDNA. Southern analysis shows that: (1) the NheI site present within the cDNA is probably also present within λM.DSC2, since the DSC2 cDNA probe hybridised to two bands of NheI-digested λM.DSC2 DNA; (2) the NheI site is 1.2kb from another NheI site within λM.DSC2. The most likely positions of the NheI sites identified within λM.DSC2 are shown in figure 4.1, where the NheI site present within the cDNA is 1.2kb from a site within the shorter arm of the λFixII vector. Subsequent work has confirmed this interpretation (V.K. Sahota, pers. comm.), as the 1.2kb NheI band is cut by both SalI and NorI which have recognition sites between the NheI of the shorter λ arm and the vector’s insert. Furthermore sequence analysis confirmed that the NheI site was the same site as in the cDNA. Thus λM.DSC2 contains at least one exon of the murine DSC2 gene, with a NheI site within this exon which is less than 1kb from the shorter arm of the vector.

**4.3.2 Future Work - Targeted Disruption of DSC2 Alleles.**

To mutate a chromosomal gene, the cloned gene is first mutated *in vitro* and then used to replace part of the chromosomal allele. The mutation is often accomplished by the insertion of heterologous sequences into an exon. Depending upon the site of the insertion within the gene, the mutation will either result in premature termination of translation or complete inactivation of the gene’s transcription.
The *NheI* site identified within $\lambda$M.DSC2 could be used to introduce the disrupting selectable *neo* gene. First, a fragment which encompasses this site could be subcloned into a targeting vector. Whilst the right hand flanking sequence (with respect to figure 4.1) is restricted by the position of the *NheI* site 1kb from the vector DNA, the length of the left hand flanking sequence depends upon which restriction sites are present within $\lambda$M.DSC2. The sizes of these sequences are important since these correlate to the efficiency with which they replace part of the chromosomal gene by homologous recombination. Total lengths of homology between 1.3kb and 6.8kb have been reported to be optimum for efficient homologous recombination (Hasty *et al.*, 1991). As homologous recombination between exogenous DNA and chromosomal DNA is a relatively rare event, a gene targeting vector that would select for cells where homologous recombination has occurred would be used (Manosour *et al.*, 1988). Recombination is selected for by the integration of the disrupting *neo* gene into the genome, whilst non-specific recombination events are selected against by integration of the HSV-tk gene from the targeting vector into the genome. The HSV-tk gene product converts gancyclovir into a cytotoxic product. Integration of the disruptive sequences may be tested by Southern analysis. Once the mutated sequences have integrated into a DSC2 allele, any phenotype could be observed either *in vitro* in ES cells and EBs, or *in vivo* after the ES cells have been injected into murine blastocysts to create chimaeric mice. To investigate homozygous phenotypes *in vitro* both DSC2 alleles could be disrupted by increasing the selective pressure for integration of the *neo* gene (Gu *et al.*, 1994). Alternatively homozygous mice could be created by crossing heterozygous mice for disrupted DSC2 alleles.

As the DSC2 cDNA used to isolate $\lambda$M.DSC2 encodes sequences only within the Dsc2 extracellular domain, it is likely that the exon to be disrupted also encodes a portion of the Dsc2 extracellular domain. The targeted disruption of sequences in this portion of the gene might result in the partial translation of the DSC2 mRNA before termination. Since no anchoring transmembrane sequences would be translated the product would be secreted from the cell and might result in a dominant negative phenotype. Complete inactivation of the DSC2 gene could be accomplished by disrupting sequences within the
most 5' portion of the mRNA which would require the isolation of more DSC2 cDNA and genomic clones.

Other cadherin gene have been disrupted by targeted mutation. In the case of E-cadherin, mice heterozygous for an E-cadherin null allele are apparently normal, while homozygous embryos do not form a trophectoderm leading to prenatal lethality (Larue et al., 1994). One could predict that a homozygous desmosomal cadherin null mutation may also be lethal, especially of genes like DSC2 whose expression is both widespread and activated during early development. Alternatively, desmocollins may be able to compensate for each other resulting in a subtle or no apparent phenotype.
Chapter Five - General Discussion

5.1 Searching for an Epithelial-Specifier
5.2 Long Range Regulatory Elements
5.3 Investigating Long Range Regulatory Elements
5.4 Applications of this Work
5: GENERAL DISCUSSION

The major focus of this work has been to identify regulatory elements within the region of DNA immediately upstream of the human DSC2 gene. A direct extension of this would be to characterise DNA:protein interactions within the promoter and identify transcription factors important in both the temporal and spatial control of DSC2 expression.

Searching for an Epithelial-Specifier.

Although all epithelia in higher organisms have a distinct origin in development, exhibit polarity and express specific protein markers, the transcription factors which define an epithelial phenotype are unknown. In muscle, a family of transcription factors related to MyoD is responsible for the distinct pattern of muscle gene expression. In the pituitary a single transcription factor, Pit-1 has been described to drive the tissue-specific phenotype (Kuo et al., 1992). Epithelia are particularly rich in a number of transcription factors such as AP-2, Basonuclin and the Oct-related proteins, Oct-6, Skn-1a and Skn-1i (Mitchell et al., 1990; Tseng and Green, 1994; Faus et al., 1994; Anderson et al., 1993). However, none of these been associated with a function in epithelia analogous to MyoD in muscle. This may be because (1) factors remain to be detected, (2) ubiquitous factors direct epithelial specific gene expression, (3) the epithelial phenotype is a default and all other tissue-specific phenotypes require repression of this and activation by tissue-specific factors.

An epithelial-specific transcription factor might bind to the DSC2 promoter since, unlike the expression of other epithelial genes such as keratins which are relatively exclusive to cell-type within epithelia, DSC2 expression is ubiquitous in desmosome-bearing tissues and in epithelia (Nuber et al., 1995). Preliminary sequence analysis of desmosomal cadherin promoters, namely DSC2, DSG1 and DSG3 (M.Adams, pers. comm.) shows regions of conservation. In addition, since DSC2 and DSG2 have very similar patterns of expression they too might have conserved elements which would help to predict the location of DNA elements important in their transcriptional regulation and sites which might bind an epithelial-specifier.
DSC2 expression in heart is also of particular interest since unlike epithelia, heart is derived from mesoderm and DSC2 transcription may be activated by alternative mechanisms.

**Long Range Regulatory Elements.**

Recent mapping experiments of the human desmosomal cadherin genes have uncovered a highly organised structure to the chromosomal locus with implications for their regulation. By isolating several YAC clones spanning the desmosomal cadherin locus, it has been shown (Simrak *et al.*, 1995; C.M.E. Cowley, pers. comm.) that the desmosomal cadherins are tightly linked within a cluster no larger than 580kb upon human chromosome 18 at q12.1 (figure 5.1). Both the desmocollin and desmoglein genes have subsequently been ordered and the direction in which the genes are transcribed identified. These genes form two clusters with the direction of gene transcription in opposite directions. The clustering of this subfamily of cadherin genes is reminiscent of type II keratins upon human chromosomes 12q.13, where more closely related genes, with respect to their open reading frames, are also physically closer together (Yoon *et al.*, 1994). The close proximity of the desmocollin and desmoglein genes to one another is probably a consequence of the recent duplication and mutation of an ancestral desmosomal cadherin.

The desmocollins and desmogleins share approximately 30% identity in their open reading frames, the same amount as both types of desmosomal cadherins share with the classical cadherins; nevertheless the two types have remained tightly linked. This suggests a selective pressure keeping the two gene clusters together, which could be generated by common regulatory elements. If the desmosomal cadherins were to form obligate partners like the keratins, common regulatory elements would offer a point where pairwise expression could be regulated.

If regulatory elements that act upon the whole desmosomal cadherin locus have been instrumental in keeping the genes so tightly linked these could be similar to the locus control region (LCR) of the \( \beta \)-globin gene cluster. The five globin genes are arranged upon human chromosome 11 in the order in which they are expressed during
Figure 5.1: The Human Desmosomal Cadherin Locus. Chromosomal location of the desmosomal cadherins at q12.1 and the direction of their transcription. The chromosomal location of N-cadherin (N-CAD) at q11.2 is also marked. Tr is a sequence tag site (STS), a mapped marker at 18q12.1.
Human Chromosome 18

18q12.1

N-CAD
Tr
DSG2
DSG3
DSG1
DSC1
DSC2
DSC3

Direction of Transcription
development; expression switches from embryonic (e) in the yolk sac, to foetal (γ) during intra uterine life, and to adult (δ and β) after birth. The LCR is located upstream of the e-globin gene and has been proposed to play a key role in the transcriptional regulation, replication timing and chromatin structure of the entire β-globin domain (Orkin, 1990). LCR elements are distinct from classical enhancer elements in that they confer high levels of expression of reporter genes only after integration into chromosomal DNA and this high level of expression is independent of the site of integration (Grosfeld et al., 1987).

The β-globin gene cluster upon human chromosome 11 and the four HOX clusters upon human chromosomes 2, 7, 12 and 17 demonstrate a temporal and/or spatial colinearity (Hanscombe et al., 1991; Scott, 1992; Manak and Scott, 1994). Thus the order of the genes reflects the spatial and/or temporal patterns of expression. A similar pattern may be demonstrated by the desmoglein genes in their expression in epidermis, with DSG2 expressed in the basal layer, DSG3 more suprabasally and then DSG1 much more suprabasally.

**Investigating Long Range Regulatory Elements.**

Long range regulatory elements in the intervening region between the desmocollin and desmoglein clusters could be investigated by a physical or functional approach. DNase hypersensitive sites often mark regions within chromatin where DNA has a more open conformation, presumably to allow for interactions with transcription factors. Such sites could be mapped between the two desmosomal cadherin clusters in a search for long range regulatory elements. Alternatively fragments of the intervening region could be subcloned adjacent to reporter genes and their effects observed in cultured cells or in transgenic mice. However there are limitations to this strategy to study regulatory elements that may be at some distance from the affected gene. Transient transfection experiments may be redundant if regulatory elements require integration into chromosomal DNA to function. Also, since the spatial distance between the putative regulatory element and reporter gene may differ from that of the element and the chromosomal gene, it is difficult to demonstrate that the element affects the target chromosomal gene *in vivo*. Reporter constructs that integrate into chromosomal DNA,
on the other hand, can also be affected by surrounding chromosomal sequences depending upon the site of integration.

To overcome some of these difficulties vectors capable of carrying putative long range regulatory elements and their target genes could be used, such as YACs or P1 clones which can carry up to 1Mb or 95kb of insert DNA, respectively (Pierce and Sternberg, 1992). Long range elements within the intervening region could then be mapped by their effect upon human desmosomal cadherin expression in transgenic mice, since the technology to create such models directly from YACs and P1 clones already exists (Scheldl et al., 1992; Peterson et al., 1993; McCormick et al., 1994). However this approach could still be limited by the possibility of positional effects when clones integrate into the murine genome. An alternative would be to study putative long range regulatory elements in situ by targeted mutation. Targeted gene mutation has been limited by the low frequency of homologous recombination of introduced DNA and chromosomal alleles, however recent advances in recombination technology may make this a more viable option.

Applications of this Work.

Finally, although the significance of this work is primarily to initiate an investigation into the regulation of the desmosomal cadherins, it may also be advantageous to consider the proteins that bind to the promoters of the desmosomal cadherins as targets for therapy. Most of the changes induced in hyperprolific skin disorders may be a result of changes in gene transcription. In psoriasis for example, involucrin and transglutaminase expression appears precociously (Bernard et al., 1986; Parent et al., 1990), whilst the onset of early differentiation markers, such as keratins 1 and 10, can be abolished in favour of keratins 6 and 16 (Bernerd et al., 1992). It could be that the inappropriate expression of particular desmosomal cadherins contributes to certain skin diseases. The treatment of such skin diseases may involve the use of potent drugs such as steroids, retinoids or vitamin D derivatives which have pleiotropic effects. Transcription factors that modulate the expression of specific genes would make an ideal focus for targeting aberrant gene expression. The initial steps would be to define the specific interactions of target genes and which transcription factors regulate them.
Chapter Six - Materials and Methods

6.1 Reagents
6.1.1 Bacterial Hosts, Vectors and Mammalian Cell Lines
6.1.2 Primers
6.1.3 Bacteriological Media
6.1.4 Agarose Gel Electrophoresis Reagents
6.1.5 Standard Molecular Biology Reagents
6.1.6 Mammalian Cell Culture Media

6.2 Generation of DNA Recombinant Molecules
6.2.1 Restriction of DNA with Endonucleases
6.2.2 Agarose Gel Electrophoresis
6.2.3 Extraction of DNA from Agarose Gels
6.2.4 Alkaline Phosphatase Treatment of Vectors
6.2.5 Ethanol Precipitation of Nucleic Acids
6.2.6 Filling-in 3' Recessed Ends
6.2.7 Ligation of DNA Fragments into Plasmid Vectors

6.3 Transformation of Competent E.coli
6.3.1 Preparation of Competent Cells
6.3.2 Use of Competent Cells

6.4 Isolation of Plasmid DNA
6.4.1 Small Scale - Alkaline Lysis
6.4.2 Large Scale - Qiagen Preparation
6.4.3 Large Scale - CsCl Banding

6.5 Isolation of Human Genomic DNA from Whole Blood

6.6 Estimation of DNA and RNA Concentrations
6.7 Screening a λFix II library
6.7.1 Preparation of Plating Cultures
6.7.2 Plating out a λFix II Library
6.7.3 Replica Plating a λFix II Library

6.8 Making High Titre Stocks of λFix II Recombinant Phage

6.9 Isolation of λ DNA

6.10 Southern Blotting

6.11 Probing Nylon Membranes
6.11.1 Hybridisation of Labelled Probes to Nylon Membranes
6.11.2 Washing Nylon Membranes
6.11.3 Random Primed Labelling of DNA
6.12.1 DNA Sequencing
6.12.2 Preparation of Plasmid DNA for Sequencing
6.13 Polymerase Chain Reaction
6.14 Vectorette Polymerase Chain Reaction
6.15 Cell Biology Methods
6.15.1 Maintenance of Adherent Cell Cultures
6.15.2 Maintenance of Jurkat-J6 Cells
6.16 Transfection of Mammalian Cells
6.16.1 Lipofection
6.16.2 Electroporation
6.17 Enzyme Assays
6.17.1 Cell Extraction
6.17.2 Assay for Luciferase Activity
6.17.3 Assay for Chloramphenicol Acetyl Transferase Activity
6.17.4 Assay for β-Galactosidase Activity
6.17.5 Quantitation of Total Protein
6.18 Isolation of RNA from Cultured Cells
6.19 RNase Protection Assay
6.19.1 Riboprobe Synthesis
6.19.2 Hybridisation of Riboprobe and RNA in Solution
6.19.3 Digestion of Single-Stranded RNA
6: MATERIALS AND METHODS

6.1 Reagents.

6.1.1 Bacterial Hosts, Vectors and Mammalian Cell Lines.

*Escherichia coli* K-12 Strains.

TG1  
\[F'\text{traD36 proAB lacF' lacZA}\Delta M15 \Delta (lac-pro) \text{ supE thi-1 hsdDS} \]

SURE  
\[e14 (mcrA) \Delta (mcrCB-hsdSMR-mrr)171 \text{ endA1 supE44 thi-1 gyrA96 relA1 lac recB recJ sbcC umuC::Tn5(kan') uvrC } [\text{F'proAB}, \text{lacF'ZA}\Delta M15, Tn10(tet')] \]

SRB  
\[sbcC recJ uvrC umuC::Tn5(kan') \text{ supE44 lac gyrA96 relA1 thi-1 endA1 e14 (mcrA) } \Delta (mcrBC, hsdRMS, mrr)171 \text{ [F'proAB lacF'ZA}\Delta M15] \]

SRB(P2)  
SRB (P2) lysogen

Phage Strains.

λFix II  
λ replacement vector capable of carrying 9 - 23kb of cloned DNA (Stratagene)

Plasmid Strains.

pBluescript SK II  
cloning vector (Stratagene)

pBluescript KS II  
cloning vector (Stratagene)

pGL2-Basic  
luciferase reporter gene vector (Promega)

pGL2-Control  
luciferase reporter gene under the control of the SV40 early promoter and enhancer (Promega)

pGL2-Promoter  
luciferase reporter gene under the control of the SV40 early promoter (Promega)

pCAT-Control  
CAT reporter gene under the control of the SV40 early promoter and enhancer (Promega)

pPD46.19  
lacZ reporter gene vector (Fire et al., 1990)

pCMV-βGal  
lacZ reporter gene under the control of the cytomegalovirus constitutive promoter

Cell Lines.

L929  
Murine fibroblasts (ATCC catalogue)
6.1.2 Primers.
All oligonucleotides were synthesised by the Institute’s central facility using an Applied Biosystems 380B DNA synthesiser. Oligonucleotides were supplied in ammonia which was removed under vacuum in a Hetovac concentrator and the oligonucleotide resuspended in 50μl of dH₂O. Their concentration was estimated by measuring the absorbance of an oligonucleotide solution at 260nm using a Perkin-Elmer Lambda Bio UV/VIS spectrophotometer, calculating the molar extinction coefficient of the oligonucleotide (absorbance of a 1M solution at 260nm), then dividing the measured absorbance by the calculated molar coefficient to give the molarity of the oligonucleotide solution.

DSC2 cDNA - RSB 39 (5’-GGGCTCCGTTCCAGGAG-3’),
DSC1 cDNA - RSB 86 (5’-TGCTAATTATGAAGCTTCC-3’),
DSC1 cDNA - RSB 87 (5’-ATTCTAAGCTTCTCCTGGAGCCCGTC-3’),
DSC1 cDNA - MDM 6 (5’-GTGTAAATTGAGCCATCT-3’),
DSC1 cDNA - IK 29 (5’-GAAAACAAGTCTCCTA-3’),
DSC1 cDNA - JA 141 (5’-CCACAGATGCATAAGTAGG-3’),
DSC1 cDNA - JA 142 (5’-CATCAGGGCTAGTTGTCC-3’).

All enzymes were purchased from Northumbria Biologicals Limited or Boehringer Mannheim. Chemicals were purchased from Sigma or BDH. Plasticware for mammalian cell-culture was purchased from Nunc, Corning and Cel-Cult.

6.1.3 Bacteriological Media.
L-Agar: tryptone 15g, yeast extract 5g, sodium chloride 10g, Difco agar 15g, pH adjusted to 7.2 with NaOH, made up to 1 litre with dH₂O.
NZY-Agar: NaCl 5g, MgCl₂.6H₂O 1.837g, casamino acids 10g, yeast extract 5g, Difco agar 15g, made up to 1 litre with dH₂O.

L-Broth: tryptone 15g, yeast extract 5g, NaCl 5g, pH adjusted to 7.2 with NaOH, made up to 1 litre with dH₂O.

NZY-Broth: NaCl 5g, MgCl₂.6H₂O 1.837g, casamino acids 10g, yeast extract 5g, made up to 1 litre with dH₂O.

NZY top agarose: NZY-broth with 0.8% low melting point agarose

Terrific-Broth: tryptone 12g, yeast extract 24g, glycerol 4ml, KH₂PO₄ 2.31g, K₂HPO₄ 12.54g.

λ Buffer: 10mM Tris pH7.5, 10mM MgSO₄, 50mM NaCl, 0.01% Difco gelatin.

6.1.4 Agarose Gel Electrophoresis Reagents.

Electrophoresis Buffer: 40mM Tris acetate pH8.3, 20mM sodium acetate, 1mM EDTA.

Loading Buffer: Ficoll-400 2.500g, xylene cyanol 0.042g, bromophenol blue 1.000g, made up to 10ml with dH₂O to make a 10X solution.

6.1.5 Standard Molecular Biology Reagents.

Denaturing solution: 1.5M NaCl, 0.5M NaOH.

Neutralising solution: 0.5M Tris pH 7.2, 1.5M NaCl, 0.001M EDTA.

20X SSC: 6M NaCl, 0.6M trisodium citrate pH 7.0.

20X SSC+Denhardt’s solution: Ficoll-400 4g, BSA 4g, polyvinylpyrrolidone 4g, made up to 1 litre with 20X SSC.

Southern prehybridisation solution: 1% v/v SDS, 4X SSC+Denhardt’s, made up to required volume with dH₂O.

Southern hybridisation solution: 1% v/v SDS, 4X SSC+Denhardt’s, 4% v/v dextran sulphate, made up to required volume with dH₂O.

TE: 10mM Tris pH 8.0, 0.1mM EDTA.

Solutions for plasmid preparations:

Solution 1: 500mM glucose, 25mM Tris pH 8.0, 10mM EDTA

Solution 2: 0.2M NaOH, 1% SDS.

Solution 3: 3M KCH₂COOH, 5M acetic acid.
6.1.6 Mammalian Cell Culture Media.
Antibiotics + glutamine: penicillin 6.00g, streptomycin 10.00g, L-glutamine 29.20g, made up to 1L with dH₂O.
Cell culture medium: Dulbecco’s Modified Eagle’s Medium (Gibco BRL), 10% v/v foetal calf serum, 1% antibiotics + glutamine.
Jurkat culture medium: RPMI (Gibco BRL), 10% v/v foetal calf serum.
PBS-A: NaCl 10g, KCl 0.25g, Na₂HPO₄ 1.437g, KH₂PO₄ 0.25g, made up to 1L with dH₂O.
Trypsin versene: NaCl 8g, KCl 0.2g, Na₂HPO₄ 1.15g, KH₂PO₄ 0.2g, EDTA 0.1g, trypsin 1.25g, phenol red 0.01g, made up to 1L with dH₂O

6.2 Generation of DNA Recombinant Molecules.
6.2.1 Restriction of DNA with Endonucleases.
Restriction endonucleases cleave DNA molecules at sequence-specific sites. Enzymes were diluted at least 10X in the reaction volume to dilute out glycerol and were used with the recommended buffers under the suggested optimum incubation temperature for 2-3 hours. Samples could then be used in agarose gel electrophoresis.

6.2.2 Agarose Gel Electrophoresis.
This technique is used to resolve, visualise and size linear DNA molecules. A 0.8-2.0% w/v agarose solution was made by boiling powdered agarose in electrophoresis buffer and allowed to cool to 60°C. The solution was then poured into a preassembled gel-cradle with a comb inserted at one end to create a series of wells. Once set, the gel was transferred to the electrophoresis apparatus and just covered with electrophoresis buffer. Three systems were used (Pharmacia GNA-200, BRL Horizon 11.14 and 58) depending on the size of gel required. Loading buffer was added to the sample before loading it on to the gel. As DNA has a negative charge it will migrate through the gel towards the anode. The gel acts as a molecular sieve, retarding the movement of larger molecules, the rate of migration being inversely proportional to the log₁₀ of the molecules’ size. DNA of known sizes was loaded alongside samples to size molecules of unknown lengths. The gel was stained by soaking in a 0.1mg/ml EtBr solution for 30 minutes and then the DNA was visualised using an ultra-violet transilluminator with a wavelength of
302nm. Gels were photographed using a Polaroid Land camera on Kodak technical pan film 4445 and processed in an AGFA Gevaert developer.

6.2.3 Extraction of DNA from Agarose Gels.

DNA was purified from cut agarose bands using the Geneclean kit (Bio 101 Inc.). The technique uses the principle that double-stranded DNA will stick to glass powder in the presence of high salt concentrations. The DNA band was excised from the stained gel using a sterile scalpel then placed into a microfuge tube. A volume of 4M NaI approximately three times that of the gel slice was added to the tube, then the tube placed in a 55°C water bath. The contents were vortexed occasionally until the agarose had dissolved completely. 1µl of glass powder solution was added for every 2µg of DNA in suspension, vortexed, then left on ice for at least 5 minutes. The glass powder was pelleted by spinning at high speed for 5s and the supernatant was then drawn off. The pellet was washed three times by resuspending in fifty volumes of ethanol wash solution (50% ethanol, 0.1M NaCl, 10mM Tris pH 7.5, 1mM EDTA) which was stored at -20°C, microfuging at high speed for 5s and discarding the supernatant. After the third wash all the supernatant was carefully drawn off and the DNA eluted by adding the same volume of dH₂O as the volume of powdered glass solution and incubating at 55°C for 3 minutes. The technique was unsuitable for molecules below 500bp, in which instance agarose gel electrophoresis was performed using a low melting point agarose (FMC Bioproducts). Subsequent reactions were then carried out directly on the excised band of DNA after heating the agarose for five minutes in a boiling water bath and diluting with three volumes of dH₂O.

6.2.4 Alkaline Phosphatase Treatment of Vectors.

To prevent compatible ends of cut plasmids rejoining during ligation reactions they were dephosphorylated using calf intestinal phosphatase (Boehringer Mannheim). One unit of enzyme was added for the last hour of endonuclease restriction of plasmid DNA and subsequently removed during gel electrophoresis.
6.2.5 Ethanol Precipitation of Nucleic Acids.
In order to purify either DNA or RNA they were precipitated by adding 2.5 volumes of 100% ethanol and 0.1 volume of 3M sodium acetate, mixing, then placing at -70°C for 30 minutes. After centrifuging at 14000 rpm for 15 minutes the DNA or RNA pellet was washed with ice-cold 70% ethanol, allowed to air-dry for 5 minutes, then resuspended in an appropriate volume of dH₂O.

6.2.6 Filling-in 3' Recessed Ends.
To create blunt-ended molecules from DNA fragments with 5' overhangs, approximately 1µg of the fragment was incubated with 1 unit of Klenow in a 20µl volume containing 20mM of each dATP, dCTP, dGTP and dTTP and Klenow buffer (10x buffer: 500mM Tris-HCl at pH 7.5, 100mM MgCl₂, 10mM DTT, 500µg/ml BSA), for 15 minutes at 25°C. The product was ethanol precipitated, washed, then resuspended in 10µl dH₂O. Blunt-ended molecules could then be ligated into compatible vectors.

6.2.7 Ligation of DNA Fragments into Plasmid Vectors.
DNA fragments and cut plasmid DNA, both with either compatible sticky ends or blunt ends, were mixed and joined to produce a closed circular covalent molecule. One unit of T4 DNA ligase was used in a 20µl reaction volume with the recommended buffer. The DNA fragment and vector were added in either a 3:1 ratio if ends were sticky or 1:3 ratio if the molecules were blunt-ended. The reaction was incubated overnight at 17°C before diluting to 100µl with dH₂O. 20µl were used to transform competent cells.

6.3 Transformation of Competent E.coli.
6.3.1 Preparation of Competent Cells (Chung and Miller, 1988).
A number of techniques exist that increase the ability of E.coli to take up DNA. The method used allowed a readily available frozen stock of highly competent cells to be made. Bacterial cells were grown to early log phase (OD₆₀₀ = 0.4) in L-broth then centrifuged and resuspended in 1/10 of the volume in pre-chilled transformation and storage buffer (10% PEG-3350, 5% DMSO, 10mM MgCl₂, 10mM MgSO₄, in L-broth). Subsequent to incubation on ice for 20 minutes the cells were aliquotted and rapidly
frozen in a dry ice/ethanol bath then stored until required at -70°C. PEG induces competence, whilst DMSO acts as a cryoprotectant.

6.3.2 Use of Competent Cells.
The cells were thawed slowly on ice as their level of competence is temperature sensitive. 20μl of diluted ligation reaction or 0.5μg of uncut plasmid were added to 100μl of cells in a pre-cooled microfuge tube and gently mixed by pipetteing. The tube was then returned to ice for 30 minutes before 700μl of L-broth were added and the tube transferred to a 37°C water bath for at least 1 hour. 200μl of the cells were plated on L-agar containing a selective antibiotic to select for transformants.

6.4 Isolation of Plasmid DNA.
6.4.1 Small Scale - Alkaline Lysis (Birnboim and Doly, 1979).
This technique allows the purification of small amounts of plasmid DNA suitable for both endonuclease restriction and DNA sequencing. It is based upon the rapid denaturation of both chromosomal and plasmid DNA followed by the selective renaturation of plasmid DNA following neutralisation. 5ml of L-broth containing a selective antibiotic were inoculated and then incubated overnight at 37°C with vigourous shaking. 1.5ml were taken and microfuged at high speed for 1 minute, the supernatant was then aspirated off and the pellet resuspended in 100μl of solution 1. The tube was incubated at room temperature for 5 minutes before adding 200μl of freshly prepared solution 2, gently mixed by inversion and left on ice for a further 5 minutes. 150μl of ice-cold solution 3 were then added and the tube was vortexed for 10 seconds and left on ice for 5 minutes. The white precipitate of genomic DNA and protein was spun at high speed for 10 minutes in a microfuge and the supernatant transferred to a fresh microfuge tube taking care to avoid the pellet. 5μg of DNase-free RNase A were then added and incubated for 30 minutes at 37°C. After the addition of an equal volume of phenol:chloroform (1:1 v/v) the tube was vortexed then microfuged at high speed for five minutes. The aqueous phase was carefully transferred to a fresh microfuge tube where the plasmid DNA was precipitated by 2.5 volumes of ice-cold ethanol. The pellet was then washed with 1ml of ice-cold 70% ethanol, dried in a Hetovac concentrator for 5 minutes, and then dissolved in 50μl of dH₂O.
6.4.2 Large Scale - Qiagen Preparation.
A Qiagen plasmid maxikit was employed to purify up to 750μg of plasmid DNA from either 500ml cultures in L-broth or 100 ml cultures in Terrific broth as directed by the manufacturer's instructions. This kit was based upon the alkaline lysis method with a column packed with an anion-exchange resin purifying the plasmid DNA further through binding, washing and selective elution of the plasmid DNA with 1.6M NaCl.

6.4.3 Large Scale - CsCl Banding.
Plasmid DNA used for transfection experiments was prepared by alkaline lysis followed by centrifugation to equilibrium in a CsCl-EtBr density gradient to purify supercoiled plasmid DNA. 500ml of Terrific-broth containing the selective antibiotic were inoculated with the bacterial host, grown overnight and the bacteria were pelleted in a JA10 rotor at 5000 rpm at 4°C for 10 minutes, resuspended in 9ml of Solution 1 and transferred to a 50ml Oakridge centrifuge tube. 1ml of a freshly prepared lysozyme solution (10mg/ml in Solution 1) was added and the tube allowed to stand at room temperature for 5 minutes. 20ml of freshly prepared Solution 2 were then added and mixed by gentle inversion several times, before placing the tube on ice for 10 minutes. 15ml of Solution 3 were added and the tube mixed by shaking several times and replaced on to the ice for a further 10 minutes. The tube was then spun at 15000 rpm, 4°C for 10 minutes, the supernatant removed and its volume measured. 0.6 volumes of isopropanol were added, mixed and incubated at room temperature for 15 minutes. The tube was then spun in a JA20 rotor at 15000 rpm for 15 minutes, the supernatant carefully removed and the pellet gently washed with 70% ice-cold ethanol and allowed to air-dry for 10 minutes before being resuspended in 2ml of dH₂O. The DNA was then phenol:chloroform extracted once and transferred to a Universal bottle. 2g of CsCl and 0.8ml of EtBr (10mg/ml) were added to the DNA and mixed thoroughly before being transferred to a Beckman polyallomer Quick-Seal centrifuge tube. The tube was topped up with 1g/ml CsCl solution, heat-sealed and spun in a Ti70.1 rotor at 60000 rpm, 18°C for at least 14 hours. Closed circular plasmid DNA is resolved from the linear bacterial DNA and nicked circular plasmid DNA because the supercoiled nature of closed circular plasmid DNA restricts the amount of EtBr that can intercalate between adjacent bases, giving it a greater density. When the DNA is visualised under UV illumination, two bands can be
seen; a weaker upper band of genomic DNA and nicked plasmid DNA and a broader more intense lower band of closed circular plasmid DNA. This lower band was recovered using a 18S needle and 5ml syringe by piercing the tube just below the band in the upright tube and drawing the band of closed circular DNA off. The ethidium bromide was removed by successive extractions with water-saturated butanol and the DNA precipitated using three volumes of ice-cold 70% ethanol (CsCl precipitates if 100% ethanol is used), washed twice using 70% ice-cold ethanol to remove salt, and resuspended in 500µl dH₂O.

6.5 Isolation of Human Genomic DNA from Whole Blood.
This method is based upon that of Miller, et al. (1988) which uses a saturated salt solution to precipitate the cellular proteins of nucleated blood cells. 25ml of fresh blood were mixed with 200µl 0.5M EDTA to inhibit clotting and frozen until used. When the blood was thawed it was split into two 50ml centrifuge tubes (Falcon) and 28ml 0.05M KCl were added, mixed well then incubated at 37°C for ten minutes. The white blood cells were purified from the mixture by centrifuging the tube in a MSE Mistral 3000i at 2500rpm, 4°C for 10 minutes, the supernatant was discarded and the pellet washed carefully twice with 0.05M KCl before resuspending it in 15ml dH₂O. 250µg proteinase K and 1.5ml 10% SDS were added to digest and denature proteins, and carefully mixed with the resuspended cells and incubated overnight at 37°C. 5.5ml of a saturated NaCl solution (approximately 6M) were added to the tube and mixed vigorously before centrifuging at 3000rpm, 4°C for 15 minutes. The supernatant was removed using a pasteur pipette and the genomic DNA carefully precipitated by adding two volumes of propanol to the DNA-rich solution in a large beaker and carefully swirling the mixture as the DNA precipitates and clumps together. This DNA was then transferred to a 2ml microfuge tube and washed with 1ml of ice cold 70% ethanol before being dissolved in 300µl of dH₂O. Genomic DNA dissolved slowly and was stored at 4°C to prevent shearing from successive freeze-thaw cycles.

6.6 Estimation of DNA and RNA Concentrations.
This was done using a Perkin-Elmer Lambda Bio UV/VIS spectrophotometer to measure the absorbance of an aqueous solution at 260nm. An absorbance of 1.000 is equivalent
to concentrations of 47.5μg/μl for DNA or of 20.0μg/μl for RNA. The ratio of the absorbances at 260nm:280nm gives an indication of the purity of the sample; DNA should have a ratio of 1.8, whilst RNA should have a ratio of 2.0, contaminating proteins or salts may give a ratio which deviates from these.

6.7 Screening a λFix II Library.

6.7.1 Preparation of Plating Cultures.
The SRB(P2) E. coli strain restricts the growth of wild type λ phage which are spi⁺ (spi = sensitive to P2 interference) phage. Only phage where the recombination genes that give the spi⁺ phenotype (red, gamma, delta) are deleted will be propagated. λFix II is a replacement vector lacking these genes and accepts inserts of DNA between 9 and 23kb.

10ml of L-broth with 0.2% v/v maltose and 10mM MgSO₄ were inoculated with the SRB(P2) strain and grown overnight with shaking at 32°C. Maltose induces the expression of maltose binding protein to which the phage will adhere during infection, whilst the low temperature ensures the cells do not overgrow as the phage will also adhere to dead cells. The cells were then centrifuged at 4°C for 10 minutes at 2000rpm. The supernatant was discarded and the cells gently resuspended in 10ml 10mM MgSO₄, then respun with the supernatant again being discarded and the cells resuspended in 4ml of 10mM MgSO₄. Plating cultures were stored at 4°C and could be used for a fortnight.

6.7.2 Plating out a λFix II Library.
Aliquots of λ buffer containing 10³ pfu were mixed with 200μl of the plating culture cells in 15ml centrifuge tubes (Falcon) then incubated at 37°C for fifteen minutes. 7ml of melted NZY top-agarose with 10 mM MgSO₄ cooled to 48°C were added to each tube of infected bacteria then quickly mixed by inversion and plated on to two-day old 150mm NZY-agar plates. Once the agarose had set the plates were incubated at 39°C for 8 hours, which was optimal for plaque formation, after which the plates were chilled to 4°C for at least 2 hours prior to replica plating.
6.7.3 Replica Plating a λFix II Library.

Positively charged nylon membranes (Hybond N+, Amersham) were used to blot the library. A membrane was carefully placed on the surface of the agarose using forceps, then asymmetrically marked with a sterile needle to allow future orientation of the filter. After one minute the membrane was removed and placed clean side down upon two sheets of 3MM Whatman filter paper soaked in denaturation solution for seven minutes. The membrane was then transferred to two sheets of 3MM paper soaked in neutralising solution, clean side down, and left for three minutes. This step was repeated on a fresh stack of filter paper soaked in neutralising solution. Membranes were then washed of any excess agarose in 2X SSC and the DNA fixed on to them by placing them on two sheets of 3MM paper soaked in 0.4M NaOH for twenty minutes with subsequent rinsing in 5X SSC for no more than one minute. A duplicate was made of each plate by repeating the replica plating procedure with contact between the agarose surface and the membrane extended to 3 minutes.

Membranes were probed using 32P-labelled cDNA probes (see section 6.10 Southern Blotting). Putative positive plaques were picked by removing a 1cm square of top agarose from the site on the plate using an alcohol-flamed cooled spatula. Agarose squares were placed in 5ml of λ buffer and allowed to stand for at least three hours at room temperature for the phage to diffuse out. Putative positive phage were replated, replica plated and probed until all the plaques on a single plate were positive. This usually took at least three rounds of screening.

6.8 Making High Titre Stocks of λFix II Recombinant Phage.

A single plaque was picked from a plate using a sterile pasteur pipette into 1 ml of λ buffer. The phage were allowed to diffuse out at room temperature for at least three hours. 100μl of this solution were plated out as described in section 6.7.2. After incubation an overnight at 39°C, 5ml of λ buffer were pipetted on to the surface of the agarose and the plate placed upon an orbital shaker at low speed for one hour. A pasteur pipette was used to collect as much of the phage-rich λ buffer as possible, which was then transferred to a 15ml centrifuge tube. 100μl of chloroform were added and the tube vortexed briefly. Cell debris was removed by centrifuging the lysate at 3000rpm, 4°C
for 10 minutes. The supernatant was transferred to another tube and another 100μl of chloroform added. The phage were then stored at 4°C.

6.9 Isolation of λ DNA.

50ml of L-broth with 0.2% v/v maltose and 10mM MgSO₄ were inoculated with 330μl of the phage stock solution and 10μl of SRB plating culture cells then shaken overnight at 37°C. SRB cells were used as these give higher titres of λ FixII than SRB(P2) cells. 0.5ml of chloroform was added and the flask shaken for a further twenty minutes. The lysed culture was transferred to a 50ml Oakridge centrifuge tube and centrifuged in a JA20 rotor at 10000rpm, 4°C for 10 minutes to remove the cell debris. After carefully decanting the supernatant to a fresh tube it was centrifuged for two hours at 16000rpm, 4°C. The pellet was resuspended in 700μl of λ buffer and transferred to a microfuge tube to which 35μl 20% v/v SDS and 7μl 0.5M EDTA were added. After incubating at 68°C for fifteen minutes to break open the phage protein coat 700μl of phenol/chloroform (50:50 v/v) were added and mixed to remove protein. The aqueous layer was transferred to a fresh tube where the λ DNA was ethanol precipitated, washed and dried before resuspension in 50μl of dH₂O.

6.10 Southern Blotting (Southern, 1975).

This technique allows the identification of discrete bands of nucleic acids on an agarose gel through hybridisation with a labelled probe. The stained and photographed agarose gel was first immersed in denaturing solution and gently shaken for 30 minutes, then soaked in neutralising solution for a further 30 minutes. If DNA fragments to be transferred were greater than 15kb, the gel was initially soaked in 0.25M HCl for fifteen minutes only, before immersion in the denaturation solution. HCl hydrolyses larger fragments to smaller, more transferable ones. A piece of 3MM Whatman filter paper was cut just larger than the width of the gel but long enough to form wicks into the buffer of the blot apparatus. This was positioned on a glass plate above the buffer reservoir containing 20X SSC with the wicks dipping into it. The treated gel was positioned on the filter paper taking care to avoid air bubbles and surrounded by Saran Wrap to ensure buffer is only drawn through the gel. A piece of nylon membrane (Hybond N, Amersham) and three sheets of 3MM paper were cut to the exact size of the
gel. The membrane was carefully placed on the gel and air bubbles were removed by rolling a pipette over it. The three sheets of filter paper were soaked in 2X SSC and placed on top of the membrane. A stack of paper towels at least 10 cm high was put on top of the filter papers and a 1 kg weight applied and the assembly left overnight. The membrane was marked for orientation and washed in 2X SSC, dried and wrapped in Saran Wrap. DNA was fixed to the membrane by irradiating with 120mJ/cm² in a Stratalinker UV crosslinker (Stratagene). The membrane was then ready to probe or could be stored at -20°C.

6.11 Probing Nylon Membranes.

6.11.1 Hybridisation of Labelled Probes to Nylon Membranes.

The membranes were soaked in 2X SSC and rolled between nylon nets, taking care to avoid air bubbles and placed in hybridisation bottles (Hybaid). 20ml of prehybridisation solution were added to each bottle which was then incubated in a rotary hybridisation oven (Hybaid) for at least one hour at the hybridisation temperature; 68°C for human or 65°C for mouse. Prehybridisation solution was replaced by prewarmed hybridisation solution and the labelled DNA probe added. The bottles were incubated in the oven at the hybridisation temperature overnight.

6.11.2 Washing Nylon Membranes.

Membranes were rinsed with 50ml 2X SSC at room temperature in each hybridisation bottle. They were then washed in two 4 litre washes, the first 0.4X SSC, 0.1% SDS, the second 0.4X SSC, both at 5°C below the hybridisation temperature for fifteen minutes. The membranes were wrapped in Saran Wrap and exposed to preflashed film (X-OMAT Kodak) overnight at -70°C between two intensifying screens. The film was then processed in an AGFA Gevaert developer.

6.11.3 Random Primed Labelling of DNA (Feinberg and Vogelstein, 1983).

A megaprime DNA labelling kit (Amersham) was used according to the manufacturer's instructions to label double-stranded DNA fragments with ³²P-labelled dCTP. DNA was isolated from an agarose gel and approximately 25ng boiled with nonamer oligonucleotide primers to denature the DNA and allow annealing of primer to the template upon cooling.
Reaction buffer, dTTP, dGTP, dATP were added as directed at room temperature. 4μl 32P α-dCTP (Amersham AA0005, 10μCi/μl, 3000Ci/mmol) were added prior to 2 units of the Klenow fragment of DNA polymerase. The reaction was carried out in a final volume of 50μl at 37°C for fifteen minutes and stopped by addition of 5μl 0.5M EDTA. The probe was precipitated with addition of 20μl tRNA (1μg/μl), 12μl 3M NaCH3COOH and 300μl ethanol, then microfuged for fifteen minutes. The pellet was redissolved in 200μl dH2O and boiled for 5 minutes with 130μl sheared salmon sperm DNA (10mg/ml) as a non-specific competitor, then snap cooled on ice and added to the hybridisation solution.

6.12.1 DNA Sequencing (Sanger et al., 1977).
All sequencing reactions were carried out using either the Sequenase Version 2.0 kit (United States Biochemical Corp) or the Circumvent Thermal Cycling kit (New England Biolabs Inc.) which both use chain termination with dideoxynucleotides. 35S α-dATP (Amersham, SJ1304, 1000 Ci/mmol) was used to label ssDNA molecules. 6% polyacrylamide gels were produced using Sequagel (National Diagnostics) reagents and run using the BRL Model S2 electrophoresis apparatus. After electrophoresis gels were fixed in 2l of 10% methanol plus 10% acetic acid and dried on a BioRad slab gel drier for one hour at 80°C, exposed to X-OMAT Kodak film overnight without intensifying screens at room temperature, which was then processed in an AGFA Gevaert developer.

Double-stranded plasmid DNA was prepared for sequencing with the Sequenase Version 2.0 kit by incubating approximately 20μg of DNA in a volume of 20μl dH2O with 5μl 1M NaOH for 5 minutes at room temperature. This was then spun to remove the NaOH through a freshly prepared spin-column for 1.5 minutes at 2000 rpm, 4°C, and could then be used immediately or stored on ice. Spin-columns were prepared by piercing the bottom of a 0.5 ml microfuge tube with a 18S needle, packing the lower quarter of the tube with glass wool then filling the rest of the tube with TE-buffered Sepharose CL-6B (Pharmacia). The 0.5ml tube was then placed inside a 1.5ml microfuge tube with a pierced bottom and both tubes placed in a 15ml Falcon tube and spun at 2000rpm, 4°C
for 1.5 minutes in order to pack the column. The spin column was then ready to be used.

6.13 Polymerase Chain Reaction.
The PCR was carried out in a Hybaid thermal cycler. Reactions were carried out in a 50μl volume with 200μM of dATP, dCTP, dGTP and dTTP, 1μM of both primers, and buffer (10X 100mM Tris-HCl pH8.3, 500mM KCl, 15mM MgCl₂, 0.1% w/v gelatin, from Perkin-Elmer Cetus) with 1.67 units of AmpliTaq Taq Polymerase (Perkin-Elmer Cetus). Standard thermal cycling conditions - A total of 32 cycles for each PCR were used. The first cycle: 94°, 3 minutes 50 seconds; annealing temperature, 2 minutes; 72°, 2 minutes. Subsequent 30 cycles: 94°, 1 minute 30 seconds; annealing temperature, 1 minute 30 seconds; 72°, 1 minute 45 seconds. The final cycle: 94°, 1 minute 30 seconds; annealing temperature, 1 minute 30 seconds; 72°, 15 minutes.

Hot Start PCR (D’Aquila et al., 1991) - To reduce the likelihood of products being synthesised from non-specific sequences at lower temperatures, Taq polymerase was added to the PCR only after the reaction mixture had reached 94°C of the first cycle.

Preparation of Phage DNA - So that phage could be used directly in the PCR, they were heated to 70°C for five minutes to remove their protein coats.

Preparation of YAC DNA - YACs were supplied in agarose blocks to support the DNA and prevent shearing. The plugs were washed twice with dH₂O, melted at 65°C for 5 minutes and then diluted at a ratio of 1:5 with dH₂O. 5μl of this diluted YAC DNA were used in the PCR as a template.

6.14 Vectorette Polymerase Chain Reaction.
This technique allows the amplification of DNA of unknown sequence that is adjacent to regions of known sequence. The vectorette is a piece of double-stranded DNA with a region of partial mismatch which is used to make a library that is screened by the PCR. One of the two primers used during the PCR is generated from the region of known sequence whilst the other is the vectorette primer. This vectorette primer spans the mismatched region of the vectorette and will only anneal to DNA synthesised from this region during the first and subsequent rounds of the PCR. Therefore during the first
cycle only the known primer hybridises to the template, in the second and subsequent cycles both the known primer and the vectorette primer may anneal to the template, amplifying the region adjacent to that of the known sequence.

A vectorette kit was purchased from Cambridge Research Biochemicals, Northwich and used as directed by the manufacturer’s instructions. Libraries were constructed by digesting human genomic DNA with either EcoRI, BamHI or HindIII then ligating to the compatible vectorette. As the restriction sites are not recreated, ligation could be optimised by repeating cycles of digestion and ligation.

6.15 Cell Biology Methods.
All procedures were carried out in a laminar flow tissue culture hood. Cells were grown in supplemented cell culture medium at 5% CO₂, 37°C. Medium was replaced every two to three days.

6.15.1 Maintenance of Adherent Cell Cultures.
Passaging Cells - Cells were passaged when they covered approximately 70% of the culture flask. Splitting the cells involved pouring off the culture medium and washing them carefully three times with PBS-A. A thin layer of trypsin-versene was pipetted on to the cells which were then replaced in the incubator until they had detached. A volume of pre-warmed culture medium was then pipetted into the flask and mixed gently before the cells were dispensed into fresh flasks containing medium and then replaced into the incubator.

Frozen Storage of Cells - A 70% confluent 75cm² flask was trypsinised, as above, and 10ml of ice-cold culture medium was added. The cell suspension was pelleted at 500rpm, 4°C for 5 minutes and resuspended in 1.5ml of ice-cold freezing medium (culture medium with 10% DMSO). The suspension was aliquotted into two cryotubes and slowly frozen by placing them inside a Stratacooler (Stratagene) that had been stored at 4°C, then placing the Stratacooler and cryotubes at -70° for 12 hours. The cryotubes were then immersed in liquid nitrogen for long term storage.
6.15.2 Maintenance of Jurkat-J6 Cells.

Jurkat cells were passaged by centrifuging the cells at 500rpm, 4°C for 5 minutes, then washing them by resuspension in PBS-A and recentrifuging them. Cells were then resuspended in pre-warmed Jurkat culture medium and a proportion added to a fresh flask containing medium which was then replaced in the incubator.

6.16 Transfection of Mammalian Cells.

6.16.1 Lipofection.

Adherent cell lines were transfected with plasmid DNA by lipofection. This method of DNA delivery coats the plasmid DNA with cationic lipids which may promote uptake by binding to the cells through electrostatic interactions followed by either direct fusion of the liposome or entry by spontaneous endocytosis.

A 70% confluent 25cm² flask was trypsinised and split 1:30 into the 9.5cm² wells of six-welled multiwell plates. After allowing the cells to reattach overnight, they were washed carefully three times with PBS-A, any excess PBS-A was aspirated off and 1.1ml of unsupplemented F15 medium added to them. Liposomes were prepared by mixing 114μl of serum-free F15 medium to 5.7μl of lipofectamine (Gibco, BRL) and then adding a total of 1.71μg of plasmid and herring sperm DNA (Promega) and mixing by gentle pipetting. The liposomes were left to stand for fifteen minutes before adding them to the serum-free medium of the washed cells and mixing by gently rocking the plate. After eight to twelve hours, during which time some cell-death had occurred, the liposomes were aspirated off the cells, which were then washed three times with PBS-A before 3ml of supplemented culture medium was added. The transfected cells were replaced in the incubator, then extracted 36 hours later.

6.16.2 Electroporation.

Jurkat cells were transfected with plasmid DNA by electroporation. Cells are subjected to a high voltage pulse which causes the transient formation of micropores in their plasma membranes, allowing extraneous DNA to be enter the cells. Electroporation was carried out using a Bio-Rad Gene-Pulser with capacitance extender. Cells were split twice at 24 hour intervals to ensure that they were in their log phase of growth just prior to
transfection. The cells were washed with serum-free RPMI medium and resuspended at a density of $5 \times 10^7$ cell/ml in serum-free RPMI. 200μl of cells were added to 50μl of a RPMI-DNA solution containing 100μg of plasmid DNA and herring sperm DNA in a Biorad 0.4cm electroporation cuvette and mixed by gentle pipetting. This was incubated on ice for ten minutes then pulsed with 250V, 950μF; this should give a time constant within the range of 35-50ms. The cuvette was replaced on ice for ten minutes then the contents added to 10ml of pre-warmed complete Jurkat culture medium in a 25cm² flask. The transfected cells were replaced to the incubator and extracted after 36 hours.

6.17 Enzyme Assays.
6.17.1 Cell Extraction.
Adherent cells were washed three times with PBS-A, any excess PBS-A was aspirated off and 200μl of 1X reporter lysis buffer (Stratagene) added to them. The proprietary reporter lysis buffer allowed assays for CAT, β-galactosidase and luciferase to be performed in the same buffer. The lysis buffer was incubated over the cells for fifteen minutes at room temperature, after which the cells were scraped from the bottom of the well using the wide end of a pipette tip and the whole lysate transferred to a 1.5ml microfuge tube. Cellular debris was pelleted by microfuging at 15000 rpm for 2 minutes and the supernatant used in subsequent assays. Jurkat cells were extracted by centrifuging the cells at 500rpm, 4°C for five minutes, washing them by resuspension in PBS-A and centrifuging them again at 500rpm, 4°C for five minutes. The supernatant was aspirated off and the cellular pellet resuspended in 1X lysis buffer and incubated at room temperature for fifteen minutes. Cellular debris was pelleted by microfuging at 15000 rpm for 2 minutes and the supernatant used in subsequent assays.

6.17.2 Assay for Luciferase Activity (Bronstein et al., 1994).
Luciferase catalyses the oxidation of luciferin in the presence of ATP with the concomitant production of a photon of light.

$$\text{luciferin} + \text{ATP} + O_2 \rightarrow \text{oxyluciferin} + \text{AMP} + PP_i + CO_2 + \text{light}$$

The proprietary assay reagent that was used included coenzyme A (CoA) which produces a more intense and sustained emission of light because of more favourable kinetics with
the luciferyl-CoA as a substrate of oxidation. The reaction was performed at room temperature. 100µl of luciferase assay reagent (Promega) were added to 20µl of cell extract in a transparent polystyrene tube (Starstedt) and mixed by gentle pipetting. Emitted light was measured using a Berthold CliniLumat LB9502 luminometer over a 10s period.

6.17.3 Assay for Chloramphenicol Acetyl Transferase Activity (Neumann et al., 1987).

This assay is based upon the catalysed transfer of labelled acetyl groups to chloramphenicol to yield products which, unlike the substrates, are able to diffuse into a toluene-based scintillation fluid. Cell extracts were incubated at 60°C for 10 minutes to denature endogenous heat-sensitive acetyl CoA hydrolases (CAT is relatively heat-stable). 100µl of treated cell extract was placed in the bottom of a plastic scintillation insert (Sterilin), to which were added 5µl unlabelled acetyl CoA (4.5mM in 0.01M NaCH₃COOH), 1µl tritiated acetyl CoA (NEN CAT assay grade, 200mCi.mmol⁻¹), 50µl 5mM chloramphenicol and 50µl 1M phosphate buffer (pH 7.2). 5ml of Econfluor (NEN) were added down the side of the tube. The tubes were incubated at room temperature and the cpm read at various times in a Wallac 1410 liquid scintillation counter.

CAT catalyses acetylation only at the 3-hydroxyl position of chloramphenicol:

\[ \text{Chloramphenicol} + \text{acetyl-CoA} \rightarrow \text{3-acetylchloramphenicol} + \text{CoA} \] [1]

The 1-acetyl product occurs by nonenzymatic rearrangement:

\[ \text{3-acetylchloramphenicol} + \text{acetyl CoA} \rightarrow \text{1-acetylchloramphenicol} \] [2]

The 3-hydroxyl position is free to be acetylated again:

\[ \text{1-acetylchloramphenicol} + \text{acetyl CoA} \rightarrow \text{1,3-diacetylchloramphenicol} + \text{CoA} \] [3]

6.17.4 Assay for β-Galactosidase Activity (Hollon and Yoshimura, 1989).

180µl of cell extract were added to 20µl of 50mM CPRG (Boehringer Mannheim) and incubated at 37°C for one hour then put on to ice before the absorbance was read at 574nm using a Perkin Elmer Lambda Bio UV/VIS spectrophotometer.
6.17.5 Quantitation of Total Protein.
Total protein of cellular extracts was quantitated using the Pierce Bicinchoninic acid Protein assay reagent as directed in the manufacturer’s instructions. The assay is based upon the Biuret reaction where protein reduces copper ions in alkaline conditions. The reduced copper ions form a complex salt with the bicinchoninic acid reagent that has a strong absorbance at 562nm, which allows the spectrophotometric quantitation of proteins in solution.

6.18 Isolation of RNA from Cultured Cells (Auffray and Rougeon, 1980).
RNA was purified from cell cultures using LiCl to selectively precipitate RNA, 6M urea to denature protein and inhibit endogenous RNases and the anti-coagulant heparin, to prevent clumping of denatured proteins and cellular debris. Disposable plastic-ware was used where possible and solutions were made up using DEPC-treated water to prevent RNase contamination.

A 70% confluent petri-dish of cells was washed three times with PBS-A, then 10ml of ice cold lysis buffer (6M Urea, 3M LiCl, 10µg.ml⁻¹ Heparin) was added and swirled over the cells to cover them completely. The wide end of a plastic pipette tip was used to scrape the cells from the plastic and the viscous mixture then transferred to a 15ml tube and vortexed to ensure the mixture was homogenous. The tube was then left on ice for at least twenty four hours for the RNA to precipitate. The mixture was split into two lots and 2ml of each were spun in a 2ml microfuge tube at 14000 rpm, 4°C for thirty minutes after which the supernatant was drawn-off and another 2ml of the mixture were added to the tube and the spin repeated until all 10ml had been centrifuged. The pellet was resuspended in 500µl of 10mM Tris pH 7.5, 0.5% SDS by gentle pipetting. It was then phenol extracted twice, phenol-chloroform (1:1 v/v) extracted once and then finally chloroform extracted. The RNA was then ethanol precipitated, pelleted, washed and resuspended in 1ml dH₂O. RNA was stored in aliquots at -70°C in ethanol.

6.19 RNase Protection Assay.
This technique can be used to identify the presence of specific RNAs, map both the transcription start site of a gene and 3’ end of a mRNA. Antisense riboprobes (cRNAs)
are generated and allowed to hybridise to the target mRNA under stringent conditions to form stable duplex RNA. Single-stranded RNAs are digested with RNase and the double stranded products visualised and sized on a 6% polyacrylamide sequencing gel alongside a sequencing ladder. Disposable plastic-ware was used where possible and solutions were made up using DEPC-treated dH2O to prevent extraneous RNase contamination. A summary of the protocol used to predict the transcription initiation site of the human DSC2 gene is shown in figure 6.1.

6.19.1 Riboprobe Synthesis.
The DNA template was subcloned adjacent to a RNA polymerase promoter such as T7, T3 or SP6 and the plasmid linearised within the template. It is important that the endonuclease used to linearise the plasmid template does not create either 3' overhangs or blunt ends as these allow the RNA polymerase to turn around at the cut site and synthesise sense riboprobe. The linear plasmid was phenol:chloroform (1:1 v/v) extracted, ethanol precipitated, washed with 70% ethanol then resuspended in DEPC-treated dH2O at 0.5mg/ml.

The components of the synthetic reaction were mixed in a 1.5ml microfuge tube at room temperature to a final volume of 20μl; 10X Transcription buffer (Boehringer Mannheim), 10mM DTT, 0.5mM ATP, CTP and GTP, 20 units of RNasin (Promega), 1μg of DNA template, 5μl 32P α-UTP (Amersham, PB20383, 20μCi/μl, 800 Ci/mM ) 1.5μl of the appropriate RNA polymerase. The mixture was mixed by gentle pipetting and incubated for one hour at 37°C. The DNA template may then be removed by the addition of 1 unit RQ1 DNase (Pharmacia) and incubating at 37°C for fifteen minutes, although this is not always necessary. The riboprobe was then gel-purified. 20μl of RNase-free formamide dyes (deionised formamide, 0.01M EDTA, 0.1% bromophenol blue, 0.1% xylene cyanol) were added, and the riboprobe held at 90°C for two minutes before loading on to a 6% polyacrylamide gel run at 50W for approximately 1.5 hours.

The riboprobe band was visualised by exposing the gel to photographic film and removed by excision using a sterile scalpel blade and placing the gel slice into a 2ml microfuge tube. Elution of the riboprobe from the gel slice was carried out in 400μl elution buffer
Figure 6.1: RNase Protection Assay on the DSC2 Promoter. The human DSC2 promoter was subcloned adjacent to the T7 phage promoter to create the plasmid pMDM7. A template was prepared by linearising pMDM7 with the endonuclease AvaII and a labelled riboprobe subsequently generated. The riboprobe was hybridised to the mRNA containing the DSC2 mRNA creating a RNase-resistant dsRNA duplex and single stranded RNA degraded with the addition of RNases. The product was then sized upon a sequencing gel.
Mapping Transcription Start Site of DSC2: RNase Protection

1. Subclone into plasmid, adjacent to T7 promoter
2. Generate Riboprobes
3. Hybridise Riboprobe to mRNA
4. Digest ssRNA with RNase
5. Product: dsRNA

Results for site should corroborate with either/or primer extension SI Nuclease protection expts.
(0.5M ammonium acetate, 1mM EDTA, 0.1% SDS, 10μg E.coli tRNA), overnight at 37°C. The riboprobe was ethanol precipitated and the pellet resuspended in 20μl of dH₂O. Due to their instability, riboprobes should be used immediately and not more than forty-eight hours after synthesis.

6.19.2 Hybridisation of Riboprobe and RNA in Solution.

RNA and the riboprobe (400000 cpm) were combined in a 0.5ml microfuge tube, ethanol precipitated then resuspended in 20μl of hybridisation buffer (80% formamide, 40mM sodium Pipes pH 6.7, 0.4M NaCl, 1mM EDTA) and overlaid with a drop of mineral oil. The RNA was denatured at 80°C for five minutes then quickly transferred to a waterbath at the hybridisation temperature (empirically determined, usually between 45°C to 50°C) and incubated overnight.

6.19.3 Digestion of Single-Stranded RNA.

400μl of digestion mix (0.5M NaCl, 10mM Tris, 1mM EDTA, 20μg/ml RNase A and 20 units/ml RNase T₁) were added to the hybridised RNA (RNase A cleaves phosphodiester bonds 3'-adjacent to C/U/N, RNase T₁ cleaves phosphodiester bonds 3'-adjacent to G/N), mixed and incubated at 37°C for fifteen minutes. The concentration of RNases could be altered to vary the contrast between the signal and the background. The RNases were removed by extracting with phenol:chloroform (1:1 v/v) and the RNA duplex ethanol precipitated on dry ice using 5μg E.coli tRNA as a carrier. The pellet was carefully washed with ice-cold 70% ethanol, which was then aspirated off completely and the pellet left to air dry for five minutes before being resuspended in 5μl formamide dye. To denature any secondary structure the resulting RNA duplex was boiled for five minutes before loading on to a 6% polyacrylamide sequencing gel alongside a sequencing ladder.
ACKNOWLEDGEMENTS

I thank Roger S. Buxton and Anthony I. Magee for their guidance in my work here at the Institute, as well as members of EKG-C both past and present for their tolerance and support, especially Michael Adams whose work has paralleled my own. I also thank Dai A. Rees and Peter W. J. Rigby for useful comments and David R. Garrod for providing the bovine DSC1 cDNA clone. Allan Watkinson, Clive Harding and Anita Evans ensured my stay at the Colworth Laboratory of Unilever Research was both insightful and enjoyable. I extend my gratitude to my Parents and Garth who have encouraged and supported me throughout my education, and Hilary Thirlwell for her delicious pasta sauces.

This work was funded by a Medical Research Council and Unilever Research collaborative studentship.
BIBLIOGRAPHY


Arnemann, J., Spurr, N. K., Magee, A. I., and Buxton, R. S. (1992b). The human gene (DSG2) coding for HDGC, a second member of the desmoglein subfamily of the desmosomal cadherins, is, like DSG1 coding for desmoglein DGI, assigned to chromosome 18. Genomics 13, 484-486.


Craddock, N., Dawson, E., Burge, S., Parfitt, L., Mant, B., Roberts, Q., Daniels, J.,

Cripe, T. P., Alderborn, A., Anderson, R. D., Parkkinen, S., Bergman, P., Haugen,
papillomavirus promoter by a 88-nucleotide enhancer containing distinct cell-

heating. Nucleic Acids Res. 19, 3749.

new mutation on mouse chromosome 18 causing hair loss and immunological defects.

Desmocollins I and II are recognised by certain sera from patients with various types
of pemphigus, particularly Brazilian pemphigus foliaceus. J. Invest. Dermatol. 100,
380-384.

of domain 1 of rat T lymphocyte CD2 antigen [see comments]. Nature 353, 762-765.

(5) type II keratin reveals an absence of 5' upstream sequence conservation between
coexpressed epidermal keratins. DNA. 7, 337-345.

base change in a POU factor-binding site switches pituitary-specific to lymphoid-
specific gene expression. Genes Dev. 4, 43-51.


154


desmocollin derived from cloned cDNA: a novel subtype of desmosomal cadherins. Differentiation 47, 29-36.


of the protein and glycoprotein components of bovine nasal epithelial desmosomes by
immunoelectron microscopy. EMBO J. 6, 885-889.

for extracting DNA from human nucleated cells. Nucleic Acids Res. 16, 1215.


Transcription factor AP-2 is expressed in neural crest cell lineages during mouse
embryogenesis. Genes Dev. 5, 105-119.

Miyatani, S., Shimamura, K., Hatta, M., Nagafuchi, A., Nose, A., Matsunaga, M.,

Genomic structure and chromosomal mapping of the mouse N-cadherin gene. Proc.

The orientation and spacing of core DNA-binding motifs dictate selective
transcriptional responses to three nuclear receptors. Cell 65, 1267-1279.

by the cytoplasmic domain. EMBO J. 7, 3679-3684.


165


APPENDICES

Appendix 1: Map of Oligonucleotides within the Human DSC2 Promoter
Appendix 2: Map of λ2.DSC2, λ4.DSC2 and subclones
Map of Oligonucleotides within the DSC2 Promoter.
Map of λ2.DSC2, λ4.DSC2 and Subclones. Restriction map of λ2.DSC2 with the CpG island at the 5' end of the DSC2 gene marked, and restriction map of λ4.DSC2 aligned below. pMDM11: a 5kb SalI fragment of λ4.DSC2 subcloned into the SalI site of pBluescript. pMDM5: a 2kb BamHI fragment of λ2.DSC2 subcloned into the BamHI site of pBluescript. pMDM7: a 1.9kb EcoRI-Aval fragment of pMDM4, blunt-ended and inserted into the EcoRV site of pBluescript. pMDM4: a 3.2kb EcoRI fragment of λ2.DSC2 inserted into the EcoRI site of pBluescript. A more detailed restriction map of pMDM4 is shown with the position of the first exon marked underneath.

Endonucleases: Apal, A; Aval, Av; Avall, AvII; BamHI, B; BssHII, Bs; EagI, Eg; EcoNI, En; EcoRI, E; HindIII, H; NsiI, N; SacII, Sc; SalI, S; SmaI, Sm; SspI, Sp.