



Terms and Conditions of Use of Digitised Theses from Trinity College Library Dublin

Copyright statement

All material supplied by Trinity College Library is protected by copyright (under the Copyright and Related Rights Act, 2000 as amended) and other relevant Intellectual Property Rights. By accessing and using a Digitised Thesis from Trinity College Library you acknowledge that all Intellectual Property Rights in any Works supplied are the sole and exclusive property of the copyright and/or other IPR holder. Specific copyright holders may not be explicitly identified. Use of materials from other sources within a thesis should not be construed as a claim over them.

A non-exclusive, non-transferable licence is hereby granted to those using or reproducing, in whole or in part, the material for valid purposes, providing the copyright owners are acknowledged using the normal conventions. Where specific permission to use material is required, this is identified and such permission must be sought from the copyright holder or agency cited.

Liability statement

By using a Digitised Thesis, I accept that Trinity College Dublin bears no legal responsibility for the accuracy, legality or comprehensiveness of materials contained within the thesis, and that Trinity College Dublin accepts no liability for indirect, consequential, or incidental, damages or losses arising from use of the thesis for whatever reason. Information located in a thesis may be subject to specific use constraints, details of which may not be explicitly described. It is the responsibility of potential and actual users to be aware of such constraints and to abide by them. By making use of material from a digitised thesis, you accept these copyright and disclaimer provisions. Where it is brought to the attention of Trinity College Library that there may be a breach of copyright or other restraint, it is the policy to withdraw or take down access to a thesis while the issue is being resolved.

Access Agreement

By using a Digitised Thesis from Trinity College Library you are bound by the following Terms & Conditions. Please read them carefully.

I have read and I understand the following statement: All material supplied via a Digitised Thesis from Trinity College Library is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of a thesis is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form providing the copyright owners are acknowledged using the normal conventions. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone. This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Exploration of RNAi-based Therapeutic Strategies for Dominant Dystrophic Epidermolysis Bullosa



A thesis submitted to the University of Dublin for the Degree of
Doctor of Philosophy

By

Clare Morgan
October 2011

Under the supervision of Professor Jane Farrar

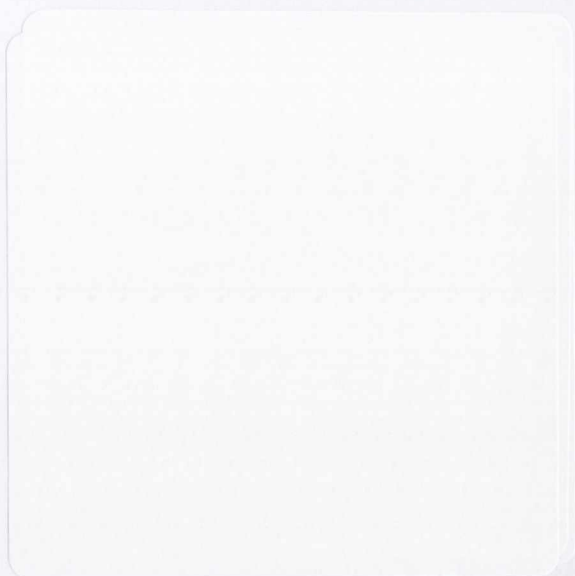
Department of Genetics
Trinity College Dublin



Thesis 9536

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work. I agree to deposit this thesis in the University's open access institutional repository or allow the library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.



Acknowledgements

There are a number of people who have offered me their support and encouragement over the course of my Ph.D. studies and who I would like to take this opportunity to acknowledge. I would like to thank my supervisor Professor Jane Farrar, firstly for allowing me this opportunity to work on a project which was of immense interest to me from the beginning, as well as for her kindness and support, particularly during the final weeks of writing. Special thanks to Danny for his guidance, patience and encouragement and especially for making the lab a more enjoyable and entertaining place to work.

Thank you to everyone in the Farrar and Humphries research groups both past and present, who have extended both their knowledge and friendship throughout the course of my studies, and especially to all of the other students in the lab. In particular I would like to thank Claire and Leanne, both of whom were there from the start, who listened and advised on many occasions and who I would not have survived many a day in the lab without. Thanks also to Carol for the constant friendly advice, lunchtime entertainment and Myzone chats, and to Mary for her help in the proofreading of some of this thesis. I would also like to take this opportunity to acknowledge Dr Alfonso Blanco, for his help with FACS analysis in UCD, and to Debra Ireland for providing me with the funding which allowed me to undertake this project.

Special thanks to all of my other friends, particularly Ciara for listening to my never-ending science and life-related complaints, and to Sara, Denise and Karen for often taking the time out to meet me on my frequent trips to UCD. Karen, your help with the flow cytometry machine was also much appreciated. I would also like to take this opportunity to thank my entire family, particularly my parents Veronica and Paddy, without whose endless loving support and encouragement I would not have achieved as much in life. Lastly I would like to dedicate this thesis to the memory of my grandmother Frances who was an inspiration throughout my life.

Summary

The main focus of this Ph.D. thesis centres on exploration of a potential gene therapy for treatment of the dystrophic form of the rare blistering skin disorder Epidermolysis Bullosa (EB). There are three main subtypes of the disease, characterised by mutations of different genes important for maintaining the integrity of epithelial lined or surfaced tissue layers of the human body, but all of which result in a phenotype of fragility of the skin and mucous membranes (Solovan *et al.*, 2005; Fine, 2010a). All three forms of the disease are genetically heterogeneous and symptoms range in severity according to the mutation in question. The dystrophic form of EB may be inherited in either a dominant (DDEB) or recessive (RDEB) manner and it is the exploration of a potential gene therapy for the dominantly inherited form of dystrophic EB which was under investigation during the course of this study. DEB arises due to mutations of the Collagen Type VII protein-encoding COL7A1 gene (Parente *et al.*, 1991; Van den Akker *et al.*, 2011). Anchoring fibrils, which are important structural components of the skin and function to anchor the layers together, are composed of Collagen VII protein (McGrath *et al.*, 1993; Chung and Uitto, 2010). Dominant COL7A1 mutations give rise to mutant protein which has a dominant negative effect, leading to a weakening of the integrity of these anchoring fibril structures, thus causing the DDEB disease phenotype (Dang and Murrell, 2008). One potential method of treatment of DDEB may possibly be provided by exploitation of the molecular tool RNA interference (RNAi), which is based on the sequence specific suppression of mRNA transcripts and thus which may be useful in the prevention of production of the dominant mutant Collagen VII protein (Kiang *et al.*, 2005; Kim and Rossi, 2007; Angaji *et al.*, 2010). It is an exploration of such an RNAi-based gene therapy approach which is investigated in the course of the studies carried out as described in this Ph.D. thesis.

Chapter 1 of this thesis provides an introduction to the topic and chapter 2 provides details of the materials and methods utilised. Chapter 3 is based on the attempted RNAi-based suppression of COL7A1 mRNA utilising artificial microRNA constructs. Cultured human keratinocyte cells were electroporated with 13 of these RNAi molecules and following FAC sorting of treated cells, real-time rtPCR analysis was used to measure levels of COL7A1 RNA and thus the suppression ability of the artificial microRNAs. The most potent artificial microRNA was found to reduce endogenous human COL7A1 mRNA levels by approximately 50%. While COL7A1 suppression was achieved as described in chapter 3, chapter 4 involved an attempt to identify more potent inhibitors of COL7A1 expression. 4 human COL7A1-targeting siRNAs were delivered to both human keratinocyte and human fibroblast cells *in vitro* and potent suppression of COL7A1 mRNA of up to 80% was achieved, as measured by

real-time rtPCR. This suppression was validated at the protein level following western blot analyses. The attempted suppression of COL7A1 as detailed in chapters 3 and 4 involved a mutation-independent-based approach for DDEB. Thus the suppressors identified would in principle result in a reduction in the levels of both wild-type and mutant transcripts, as RNAi molecules were not designed to distinguish between the two. Thus in theory such a therapeutic would be applicable for the treatment of all patients suffering from DDEB regardless of the COL7A1 mutation in question. However, such a dramatic reduction in the overall levels of Collagen VII protein would likely be harmful and so a strategy for replacement of the wild-type COL7A1 protein would likely be required. The designing of replacement COL7A1 sequences such that they display resistance to the effects of the most potent RNAi suppressors was addressed in chapter 5 of this Ph.D. thesis. The approach exploits the redundancy of the genetic code to generate RNAi-resistant COL7A1 replacement sequences which code for wild-type amino acids. Following confirmation of the resistance of all sequences by flow cytometry analyses of HEK293 cells co-transfected with the relevant siRNAs and small COL7A1 gene fragments fused to GFP, a full length COL7A1 replacement gene was generated which was found to be resistant to the effects of one of the most potent siRNAs at both the mRNA and protein levels.

A further RNAi-based approach for DDEB was investigated in chapter 6 of this thesis, although in contrast to the above, the approach in question was a mutation-dependent one. Two dominantly inherited DEB-causing COL7A1 mutations were chosen and sequence walks of the mutation regions were carried out such that all possible siRNAs that could target the mutant alleles were designed but which in theory may not effectively suppress the wild-type equivalents due to the sequence specificity of RNAi. HEK293 cells were co-transfected with small fragment COL7A1-GFP fusion vectors and siRNA effectiveness was measured by GFP intensity during flow cytometry analysis. Whilst unfortunately no efficient RNAi suppressor of one of the dominant mutations was found, the second mutation which is the most commonly occurring DDEB-causing COL7A1 mutation was reduced by approximately 45% with one of the siRNAs. The same siRNA however also reduced wild-type levels, albeit less potently with at a lower level of 20% suppression. Overall the results presented in this Ph.D. thesis represent significant steps towards development of RNAi-based therapeutics for DDEB, a debilitating disorder with an unmet clinical need.

Abbreviations

AAV: Adeno-associated virus
acGFP: *Aequorea coerulescens* Green Fluorescent Protein
AGO2: Argonaute 2
ASOs: antisense oligonucleotides
ATPase: Adenosine triphosphatase
B2M: Beta-2-microglobulin
BCA: Bicinchoninic acid
BMZ: Basement membrane zone
BPAG: bullous pemphigoid antigen
BSA: Bovine serum albumin
cDNA: complementary Deoxyribonucleic acid
C. elegans: *Caenorhabditis elegans*
CHS: Chalcone synthase
CMV: Cytomegalovirus
C_T: threshold cycle
dH₂O: distilled H₂O
ddH₂O: double distilled H₂O
DDEB: Dominant Dystrophic Epidermolysis Bullosa
DEB: Dystrophic Epidermolysis Bullosa
DGCR8: DiGeorge critical region gene eight protein
DMEM: Dulbecco's modified Eagle medium
DNA: Deoxyribonucleic acid
dNTPs: Deoxynucleotide Triphosphates
dsRBD: double stranded RNA binding domain
dsRNA: double stranded RNA
DTT: Dithiothreitol
EB: Epidermolysis Bullosa
EB simplex D-W: Epidermolysis Bullosa simplex Dowling-Meara
ECM: Extracellular matrix
E.coli: *Escherichia coli*
EDTA: Ethylenediamine tetra acetic acid
EF1: Elongation factor 1
eGFP: enhanced Green Fluorescent Protein
EMA: European Medicines Agency
FACS: Fluorescent Activated Cell Sorting
FCS: Fetal calf serum
FDA: Food and Drug Administration
FGM-2: Fibroblast growth medium-2
g: gram
HaCaT: Human adult low Calcium Temperature keratinocytes
HD: Hemidesmosomes
HEK: Human embryonic kidney
IF: Intermediate filament
iPCSs: induced pluripotent stem cells
JEB: Junctional Epidermolysis Bullosa
L: litre
LB: Luria-Bertani
LTR: Long terminal repeat

mRNA: messenger RNA
 miRNA: micro RNA
 MJD: Machado–Joseph disease
 ml: millilitre
 MLV: Moloney murine leukaemia virus
 NHDF: Normal human dermal fibroblast
 nt: nucleotide
 ORF: Open reading frame
 PAZ: Piwi-argonaute-zwille
 PBS: Phosphate buffered saline
 PC: Pachyonychia congenita
 PCR: Polymerase Chain Reaction
 PolII: Polymerase II
 PolIII: Polymerase III
 pre-miRNA: precursor miRNA
 pri-miRNA: primary miRNA precursor
 PTC: Premature termination codon
 RDEB: Recessive Dystrophic Epidermolysis Bullosa
 RISC: RNA-induced silencing complex
 RNA: Ribonucleic acid
 RNAi: RNA interference
 RNase: Ribonuclease
 rtPCR: reverse transcription Polymerase Chain Reaction
 SAP: Shrimp Alkaline Phosphatase
 SB: Stratum basale
 SC: Stratum corneum
 SCA3: Spinocerebellar ataxia type 3
 SCC: Squamous cell carcinoma
 SCID: Severe combined immunodeficient
 SDS: Sodium dodecyl sulphate
 SDS-PAGE: Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
 SG: Stratum granulosum
 SIN: Self-inactivating
 siRNA: small interfering RNA
 shRNA: short hairpin RNA
 SOC: Super Optimal broth with Catabolite repression
 SS: Stratum spinosum
 ssRNA: single-stranded RNA
 TBP: TATA box binding protein
 TCA: Trichloroacetic acid
 UBC: Ubiquitin C
 µl: microlitre
 UTR: untranslated region
 YWHAZ: tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation
 protein, zeta polypeptide

Table of Contents

	Page
Declaration	ii
Acknowledgements	iii
Summary	iv
Abbreviations	vi
Table of Contents	viii
List of Figures	xiv
Chapter 1: General Introduction	1
1.1 Introduction	2
1.2 The Skin	3
1.2.1 Skin structure	3
1.2.2 The Basement Membrane Zone (BMZ)	5
1.3 Epidermolysis Bullosa	6
1.3.1 EB Simplex	7
1.3.2 Junctional EB (JEB)	8
1.3.3 Dystrophic EB (DEB)	8
1.4 Collagen VII protein and gene	10
1.4.1 Collagen VII protein structure	11
1.4.2 The COL7A1 locus	12
1.4.3 Mutations of the COL7A1 gene	12
1.5 Therapeutic Approaches for Epidermolysis Bullosa	14
1.5.1 Therapeutic approaches for recessive dystrophic EB (RDEB)	16
1.5.2 Therapeutic approaches for dominant dystrophic EB (DDEB)	20
1.6 RNA interference (RNAi)	21
1.6.1 Discovery of RNAi	22
1.6.2 MicroRNAs	23
1.6.3 The RNAi pathway: siRNA- and shRNA-based therapies	23
1.6.4 RNAi delivery	25
1.7 Replacement	27
1.7.1 Gene replacement	27
1.7.2 Protein replacement	28

1.7.3 Mutation-dependent RNAi	28
1.8 Aims	29
Chapter 2: Materials and Methods	39
2.1 General Materials	40
2.1.1 Materials for molecular biology methods	40
2.1.2 Tissue culture materials	42
2.1.3 RNA materials	42
2.1.4 Oligonucleotide synthesis	43
2.1.5 Protein materials	43
2.2 General Methods	46
2.2.1 Cloning	46
2.2.1.1 Restriction enzyme digests	46
2.2.1.2 DNA purification	46
2.2.1.3 Ethanol precipitation	46
2.2.1.4 Gel electrophoresis	46
2.2.1.5 Gel extraction	47
2.2.1.6 Annealing of oligonucleotides	47
2.2.1.7 Ligations	47
2.2.1.8 Making <i>E. coli</i> cells competent for transformation	47
2.2.1.9 Plasmid DNA Transformations	48
2.2.1.10 In-Fusion TM Advantage PCR Cloning	48
2.2.1.11 Plasmid DNA Mini-preparation	49
2.2.1.12 Plasmid DNA maxi-preparation	50
2.2.1.13 Capillary Sequencing	50
2.2.1.14 Commercial sequencing	51
2.2.1.15 Nucleic acid concentration determination	51
2.2.2 Tissue culture techniques	51
2.2.2.1 Sub-culturing of cells	51
2.2.2.2 Haemocytometer cell counting	52
2.2.2.3 Freezing of cell stocks	53
2.2.2.4 Nucleofection of cells	53
2.2.2.5 Generation of stable cell lines	53

2.2.2.6 siRNA transfection	54
2.2.2.7 Co-transfections	54
2.2.2.8 Fluorescent Activated Cell Sorting (FACS)	55
2.2.2.9 Flow cytometry	55
2.2.3 RNA analysis	56
2.2.3.1 RNA isolations	56
2.2.3.2 Real time rtPCR	56
2.2.3.3 Real time rtPCR data analysis	57
2.2.3.4 Statistical analysis	58
2.2.4 Protein analysis	58
2.2.4.1 Isolation of protein from cell media	58
2.2.4.2 Isolation of total cell protein	59
2.2.4.3 Isolation of cytoplasmic protein	59
2.2.4.4 Protein concentration measurement	59
2.2.4.5 TCA precipitation	60
2.2.4.6 Western blotting	60
 Chapter 3: Artificial microRNA-mediated COL7A1 suppression	 64
3.1 Introduction	65
3.2 Materials and methods	69
3.2.1 Materials and methods for artificial microRNAs	69
3.2.1.1 Artificial microRNA DNA oligo design	69
3.2.1.2 Artificial miRNA vector construction	69
3.2.1.3 Nucleofection of HaCaT cells with artificial miRNA vectors	71
3.2.1.4 Analysis of RNA from transfected cells	71
3.2.1.5 Statistical Analysis	72
3.2.1.6 Generation of stable cell lines	72
3.2.1.7 Protein analysis	73
3.2.2 Materials and methods for shRNAs	73
3.2.2.1 shRNA design	73
3.2.2.2 shRNA vector construction	74

3.2.2.3 Nucleofection of HaCaT cells with shRNA vectors	75
3.2.2.4 Statistical Analysis	75
3.3 Results	76
3.3.1 Artificial microRNA vector construction	76
3.3.2 Analysis of COL7A1 mRNA suppression by miRNAs 1-13	76
3.3.3 Analysis of collagen type VII protein suppression by miRNA 4	79
3.3.4 shRNA vector construction	81
3.3.5 Analysis of COL7A1 mRNA suppression by shRNAs 4 and 6	82
3.4 Discussion	83
3.5 Conclusion	87
 Chapter 4: siRNA-mediated COL7A1 suppression	 105
4.1 Introduction	106
4.2 Materials and Methods	110
4.2.1 siRNA design	110
4.2.2 HaCaT siRNA transfection	110
4.2.3 Fibroblast siRNA transfection	111
4.2.4 Analysis of RNA from siRNA-transfected cells	111
4.2.5 Statistical Analysis	111
4.2.6 Protein analysis	112
4.2.7 Analysis of human extracellular matrix and adhesion molecule profiles in siRNA-treated cells	112
4.2.8 shRNA vector construction and RNA analysis	115
4.2.9 Artificial microRNA vector construction and RNA analysis	115
4.2.10 Statistical analysis	116
4.3 Results	117
4.3.1 Analysis of COL7A1 mRNA suppression by siRNAs in keratinocytes	117
4.3.2 Analysis of COL7A1 mRNA suppression by siRNAs in fibroblasts	119
4.3.3 Analysis of collagen type VII protein suppression by siRNA 3	121
4.3.4 Analysis of human extracellular matrix and adhesion molecule profiles in siRNA 3-treated cells	122
4.3.5 shRNA vector construction and RNA analysis	124

4.3.6 Artificial microRNA vector construction and RNA analysis	125
4.4 Discussion	127
4.5 Conclusion	134
 Chapter 5: COL7A1 Gene Replacement	 155
5.1 Introduction	156
5.2 Materials and Methods	161
5.2.1 Replacement sequence design	161
5.2.2 COL7A1 replacement fragment design	161
5.2.3 Mini COL7A1-GFP fusion vector construction	162
5.2.4 Co-transfection of COL7A1-GFP fusion vectors and siRNA	163
5.2.5 Flow cytometry data analysis of co-transfected cells	163
5.2.6 Statistical analysis	163
5.2.7 CMV-full length COL7A1 cloning	164
5.2.8 Full length COL7A1 replacement gene generation	165
5.2.9 Transfection and co-transfection of pTOPO-XLCOL7 and pTOPO-XLCOL7Rep3	166
5.2.10 Analysis of RNA from transfected cells	167
5.2.11 Statistical Analysis	167
5.2.12 Protein analysis	167
5.3 Results	169
5.3.1 Replacement sequence design	169
5.3.2 Mini COL7A1-GFP fusion vector construction	170
5.3.3 Analysis of mini COL7A1-GFP vector resistance to siRNAs	171
5.3.4 Generation of full-length COL7A1 and replacement genes	172
5.3.5 Analysis of siRNA 3 effects on full-length COL7A1 genes: RNA analysis	173
5.3.6 Analysis of siRNA 3 effects on full-length COL7A1 genes: protein analysis	174
5.4 Discussion	175
5.5 Conclusion	179

Chapter 6: Mutation-specific suppression	196
6.1 Introduction	197
6.2 Materials and Methods	201
6.2.1 siRNA design	201
6.2.2 Allele-specific COL7A1 fragment design	201
6.2.3 Mini COL7A1-GFP fusion vector construction	202
6.2.4 Co-transfection of COL7A1-GFP fusion vectors and siRNA	203
6.2.5 Flow cytometry data analysis of co-transfected cells	203
6.2.6 Statistical analysis	203
6.3 Results	204
6.3.1 siRNA design	204
6.3.2 Mini COL7A1-GFP fusion vector construction	204
6.3.3 Analysis of siRNA effect on mutant DNA expression	205
6.3.4 Analysis of siRNA 10 effect on WT DNA expression	207
6.4 Discussion	208
6.5 Conclusion	215
 Chapter 7: General Discussion	229
7.1 General Discussion	230
7.2 Conclusion	239
 References	242
 Appendices	269
Appendix I	270
Appendix II	271
Appendix III	275
Appendix IV	277

List of Figures

	Page
Chapter 1	
Figure 1.1 Structure of the skin	30
Figure 1.2 Illustration of the complexity of the cutaneous basement membrane zone	31
Figure 1.3 Clinical features of various epidermolysis bullosa subtypes	32
Figure 1.4 Some clinical features of DDEB	33
Figure 1.5 Type VII collagen domain organization	34
Figure 1.6 Schematic representation of the polymerization of anchoring fibrils	35
Figure 1.7 Anchoring fibril assembly and the consequences of the major types of mutations on type VII collagen protein synthesis in dystrophic epidermolysis bullosa (DEB)	36
Figure 1.8 RNA interference	38
Chapter 2	
Figure 2.1: Vector map of pSuper.gfp/neo	62
Figure 2.2: Restriction map and Multiple Cloning Site (MCS) of pAcGFP1-N In-Fusion Ready (Linear) Vector	63
Chapter 3	
Figure 3.1 Map of pcDNA TM 6.2-GW/EmGFP-miR vector	89
Figure 3.2 Engineered pre-miRNA structure	90
Figure 3.3 Artificial microRNA sequences	91
Figure 3.4 Overview of Artificial miRNA vector construction	92
Figure 3.5 Primer Sequences	93
Figure 3.6 Distribution of HaCaT cells during FACS analysis	94
Figure 3.7 Mean changes in COL7A1 expression: effect of artificial microRNAs 1-4 on COL7A1 expression	95
Figure 3.8 Mean changes in COL7A1 expression: effect of artificial microRNA 4 on COL7A1 expression following incubation for a range of time periods	96

Figure 3.9 Mean changes in COL7A1 expression: effect of artificial microRNA 4 on COL7A1 expression after 72 hours	97
Figure 3.10 Mean changes in COL7A1 expression: effect of artificial microRNAs 5-13 on COL7A1 expression	98
Figure 3.11 Mean changes in COL7A1 expression: analysis of RNA from stable cell lines	99
Figure 3.12 Mean changes in collagen type VII protein levels: western blot analysis of collagen type VII protein from conditioned media of stable HaCaT cell lines	100
Figure 3.13 Mean changes in collagen type VII protein levels: western blot analysis of collagen type VII protein from conditioned media of mass-transfected HaCaT cells	101
Figure 3.14 mi-shRNA sequences	102
Figure 3.15 Processing of shRNAs	103
Figure 3.16 Mean changes in COL7A1 expression: effect of mi-shRNA molecules on COL7A1 expression	104

Chapter 4

Figure 4.1 siRNA target sequences	135
Figure 4.2 Layout of the 96-well ECM and adhesion molecule PCR array plate	136
Figure 4.3 shRNA sequences	137
Figure 4.4 Artificial microRNA sequences	138
Figure 4.5 Mean changes in COL7A1 expression: effect of siRNAs 1-4 on COL7A1 expression in HaCaT cells	139
Figure 4.6 Mean changes in COL7A1 expression: effect of a pool of siRNAs 1-4 on COL7A1 expression in HaCaT cells	140
Figure 4.7 Mean changes in COL7A1 expression: effect of a range of siRNA 3 concentrations on COL7A1 expression in HaCaT cells	141
Figure 4.8 Mean changes in COL7A1 expression: effect of siRNA 3 on COL7A1 expression over a range of time-points in HaCaT cells	142
Figure 4.9 Mean changes in COL7A1 expression: effect of siRNAs 1-4 on COL7A1 expression in NHDF cells	143
Figure 4.10 Mean changes in COL7A1 expression: effect of a pool of	

siRNAs 1-4 on COL7A1 expression in NHDF cells	144
Figure 4.11 Mean changes in COL7A1 expression: effect of a range of siRNA 3 concentrations on COL7A1 expression in NHDFs	145
Figure 4.12 Mean changes in COL7A1 expression: effect of siRNA 3 on COL7A1 expression over a range of time-points in NHDF cells	146
Figure 4.13 Mean changes in Collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of HaCaT cells	147
Figure 4.14 Mean changes in Collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of NHDF cells	148
Figure 4.15 3D Profile of expression levels of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated HaCaTs	149
Figure 4.16 Volcano Plot of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated HaCaTs	150
Figure 4.17 3D Profile of expression levels of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated NHDFs	151
Figure 4.18 Volcano Plot of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated NHDFs	152
Figure 4.19 Mean changes in COL7A1 expression: effect of shRNA molecules on COL7A1 expression	153
Figure 4.20 Mean changes in COL7A1 expression: effect of artificial microRNA molecules on COL7A1 expression	154

Chapter 5

Figure 5.1 Replacement sequence design for siRNA 1	181
Figure 5.2 Replacement sequence design for siRNA 3	182
Figure 5.3 Replacement sequence design for siRNA 4	183
Figure 5.4 In-Fusion cloning method	184
Figure 5.5 Chapter 5 primer sequences	185
Figure 5.6 Vector map of pTOPO-XLCOL7	186

Figure 5.7 Distribution of HEK293 cells during flow cytometry analysis	187
Figure 5.8 Mean changes in COL7A1-GFP expression: effect of siRNA 1 on WT1-AcGFP, Rep1Max-AcGFP and Rep1-AcGFP vector expression	188
Figure 5.9 Mean changes in COL7A1-GFP expression: effect of siRNA 3 on WT1-AcGFP and Rep3-AcGFP expression	189
Figure 5.10 Mean changes in COL7A1-GFP expression: effect of siRNA 4 on WT4-AcGFP, Rep4Max-AcGFP and Rep4-AcGFP vector expression	190
Figure 5.11 Mean differences in COL7A1 expression	191
Figure 5.12 Mean changes in COL7A1 expression: effect of siRNA 3 on full-length COL7A1 gene expression in HEK293 cells	192
Figure 5.13 Mean changes in COL7A1 expression: effect of siRNA 3 on COL7A1 replacement gene expression in HEK293 cells	193
Figure 5.14 Mean changes in collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of HEK293 cells co-transfected with wild-type full-length COL7A1 and siRNA	194
Figure 5.15 Mean changes in collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of HEK293 cells co-transfected with the COL7A1 replacement gene and siRNA	195

Chapter 6

Figure 6.1 6863del16 siRNA sense sequences	217
Figure 6.2 G2043R siRNA sense sequences	218
Figure 6.3 Schematic diagram of the COL7A1-AcGFP fusion expression constructs	219
Figure 6.4 Chapter 6 primer and control siRNA sequences	220
Figure 6.5 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 6863del16 mutation site (24 hours)	221
Figure 6.6 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 6863del16 mutation site (48 hours)	222

Figure 6.7 Mean changes in COL7A1-GFP expression: siRNA 5 and siRNA 6	223
Figure 6.8 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 G2043R mutation site (24 hours)	224
Figure 6.9 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 G2043R mutation site (48 hours)	225
Figure 6.10 Mean changes in COL7A1-GFP expression: siRNA 10 and G2043R-GFP	226
Figure 6.11 Mean changes in COL7A1-GFP expression: siRNA 10 and 73-75-GFP	227
Figure 6.12 6863del16-GFP mRNA secondary structure	228

Chapter 7

Figure 7.1 COL7A1 target regions of all tested RNAi suppressors	241
---	-----

Chapter 1

1.1 Introduction

Epidermolysis Bullosa (EB) comprises a group of rare and heritable human skin diseases, affecting approximately 1 in 17,000 of live births with approximately 500,000 known cases of the disease worldwide (Featherstone, 2007). It is a prototypical mechanobullous disease and was first described by the German physician Kobner in 1886. It is characterised by the presence of recurring blisters on the skin surface and mucous membranes after exposure to minor friction or trauma, due to an increase in fragility of the skin and mucosa (Solovan *et al.*, 2005; Fine, 2010a). EB is classified into a number of different subtypes, of which there exists three main forms. These comprise Dystrophic EB, Junctional EB and EB Simplex. The different forms of EB arise due to mutations in a variety of genes encoding proteins which have important structural functions for the human body (Eady and Dunnill, 1994; Nagy and McGrath, 2010). Such genetic mutations result in either complete absence of a particular protein or the presence of a mutant protein, thus weakening the skin structure and causing the disease phenotype for which there is currently no known cure.

While no therapies currently exist for Epidermolysis Bullosa, great advances have been made in the past few decades including identification of the causative mutant genes. Therapeutic strategies have been developed in recent years, including clinical trials in which gene therapy approaches directed at correcting the initial genetic defect have been undertaken (Mavilio *et al.*, 2006; Siprashvili *et al.*, 2010). The dominant dystrophic form of EB (DDEB) arises due to mutations of the COL7A1 gene which encodes Type VII Collagen (Parente *et al.*, 1991; Van den Akker *et al.*, 2011). The resulting mutant protein displays a dominant negative effect, causing the disease phenotype. A potential gene therapy approach for DDEB would likely require suppression of expression of the mutant COL7A1 gene: no such gene-therapy based trial has yet been reported. The method of RNA interference (RNAi) represents one such molecular tool which may be suitable for achieving such suppression. RNAi is based on the sequence specific binding of RNA molecules to endogenous mRNA transcripts, resulting in their subsequent degradation (Tiemann and Rossi, 2009; Pecot *et al.*, 2011). Thus RNA interference technology, which has been shown to hold great promise for the treatment of a wide range of dominant diseases (Kiang *et al.*, 2005;

Kim and Rossi, 2007), potentially offers a therapeutic strategy for treatment of this particular form of the group of disorders that fall under the umbrella term EB.

1.2 The skin

The integumentary system is the bodily system comprising the skin and its associated structures including the hair, nails, sweat glands and sebaceous glands (Figure 1.1). It is both the largest and perhaps the most complex organ of the human body, with a variety of cell types contributing to its structure (Menon, 2002). The skin is a complex organ, covering the entire surface area of the body and thus accounting for approximately 15% of total adult body weight. It is continuous with the mucous membranes lining the orifices of the human body. The skin has a wide range of functions, the main one being that of protection. This includes the bodies' physical, chemical, immune, UV radiation and free radical defences. Furthermore, the skin functions as a sensory organ, plays a role in thermoregulation and also has endocrine functions due to its ability to synthesise Vitamin D. The skin has two main layers, the epidermis and dermis, below which lies a layer of subcutaneous fat (the hypodermis). It presents both a challenge and an opportunity in terms of the delivery of drugs to the human body, due to its protective barrier function and large surface area (Chuong *et al.*, 2002).

1.2.1 Skin structure

The skin is composed of two primary layers. The dermis, which forms the bulk of skin, is composed of connective tissue which acts to cushion the body by providing it with flexibility and strength. It is highly vascular and contains the sweat and oil glands as well as the blood vessels, and contains nerve endings, providing the skin with the senses of heat, cold, pain, pressure and touch. The dermis is composed of collagen, elastin and glycosaminoglycans which collectively are termed the extracellular matrix (ECM), as well as fibroblast cells which elaborate the ECM. The dermis also includes dermal adipose cells, the pilosebaceous units, mast cells and infiltrating leucocytes (Menon, 2002). The upper, papillary layer of the dermis contains a thin arrangement of collagen fibers while the lower, reticular layer, is

thicker and made of thick collagen fibers that are arranged parallel to the surface of the skin (Prost-Squarcioni, 2006).

The outermost skin layer is known as the epidermis and it acts as the body's protective barrier to the surrounding environment, covering the body surface and thus preventing entry of water, allergens and potentially pathogenic microbes. The major cell type of the epidermis is known as the keratinocyte, which makes up roughly 90-95% of all epidermal cells. These layers of densely packed keratinocytes produce the protein keratin. The rest of the epidermal layer constitutes melanocytes which produce the pigment melanin, protecting the skin from damage by UV rays, dendritic Langerhans cells which play a role in the body's immunity and Merkel cells (mechanoreceptors). The stratified epidermis is approximately 100 to 150µm thick, and is composed of four layers: the stratum basale (SB; basal or germinative cell layer) (single layer), stratum spinosum (SS; spinous cell layer) (5-15 layers), stratum granulosum (SG; granular cell layer) (1-3 layers), and stratum corneum (SC; horny layer) (5-10 layers) (Kanitakis, 2002). Epidermal keratinocytes are formed by mitotic division of stem cells in the stratum basale. As the daughter keratinocytes produced there migrate up through the stratum spinosum and stratum granulosum towards the skin surface, they undergo morphological and biochemical differentiation to form a rigid internal structure of keratin, microfilaments and microtubules (keratinisation) (Lawton, 2006). The outer layer of the epidermis, the stratum corneum, comprises layers of flattened dead cells known as corneocytes, which are lacking a cell nucleus. They are arranged in overlapping layers, imparting a waterproof nature to the skin's surface. These cells are shed from the skin, in a process known as desquamation, with the entire process from formation to desquamation taking approximately 28-30 days (Brysk and Rajaraman, 1992; Kanitakis, 2002; Houben *et al.*, 2007).

Located beneath the dermis, the deepest part of the skin is known as the hypodermis. It is a layer of subcutaneous fat and its thickness varies, depending on the region of the body in question. It functions to supply the upper skin layers with a source of blood vessels and nerves, while also connecting the skin to the underlying muscle and bone. It also plays important roles in thermoregulation, insulation, provision of energy (nutritional store) and protection of the body from mechanical injuries. The main cells of the hypodermis are large, rounded adipocytes (Kanitakis,

2002; Prost-Squarcioni, 2006). Located at the dermal-epidermal adhesion zone is the cutaneous basement membrane zone (BMZ), which stabilizes the association of the epidermal keratinocytes to the underlying dermis. It is a weakness of the skin within the basement membrane zone which leads to the clinical features of EB (Uitto and Richard, 2005). With EB simplex, the resulting lesions are found at the intraepidermal layer of skin, within the basal keratinocytes of the skin's outer layer. Junctional EB differs in that the lesions occur at the epidermal-dermal junction, within the cutaneous basement membrane zone. Finally, dystrophic EB lesions generally occur at the intradermal layer of skin, below the dermo-epidermal basement membrane within the upper papillary dermis (Mitsushashi and Hashimoto, 2003).

1.2.2 The Basement Membrane Zone (BMZ)

The basement membrane zone is an important and intricate region of the skin, comprised of multiple protein complexes which function to stabilise the association between the epidermal and dermal skin layers (Figure 1.2). It represents a narrow interface between the basal keratinocytes and the dermis. The BMZ is composed of two main layers, the upper lamina lucida and the lower lamina densa (Pulkkinen and Uitto, 1999). Several protein molecules are associated with these two layers and extend upwards to the epidermis and downwards to the dermis, providing the integrity and support required for the skin to maintain its protective function. The hemidesmosomes (HDs) are rivet-like structures which extend from the intracellular portion of the upper epidermal basal keratinocytes to the extracellular milieu of the lamina lucida. They comprise two plaques, the inner plaque which is located directly beneath keratin intermediate filaments and the outer plaque, which overlies the cell plasma membrane. The anchoring filaments are thread-like structures which interact with the hemidesmosomes. They traverse the lamina lucida and insert into the lamina densa. Anchoring fibrils are loop structured attachment molecules which extend from the lower section of the lamina densa to the underlying dermis skin layer (Uitto, 2009). The inner and outer plaques of the hemidesmosomes, along with the anchoring filaments and anchoring fibrils are collectively known as the HD-stable adhesion complex or the HD-anchoring filament complex. This complex forms a continuous link between the intermediate keratin filaments of the basal keratinocytes and the

underlying basement membrane zone and dermal components (McMillan *et al.*, 2003).

Each of these important structural molecules is composed of multiple protein macromolecules, each of which plays an important role in the maintenance of normal skin function. At least four distinct proteins are known to interact with each other to form the hemidesmosome. These consist of the 230-kDa BPAG1 (bullous pemphigoid antigen 1) which binds the hemidesmosomes to the intermediate keratin filaments; the 180-kDa BPAG2, a collagenous protein also known as type XVII collagen (COL17A1) which anchors the keratinocytes to the BMZ; an integrin protein known as $\alpha_6\beta_4$, which also serves to attach the epidermis to the BMZ; and plectin, a large multifunctional adhesion molecule. The anchoring filaments are composed of laminin 332 (formerly known as laminin 5 (Aumailly *et al.*, 2005)), from the laminin family of proteins. Laminin 332 consists of three polypeptide chains: α_3 , β_3 and γ_2 . These are encoded by the LAMA3, LAMB3 and LAMC2 genes respectively. The anchoring fibrils consist of type VII collagen, which is encoded by the COL7A1 gene. The importance of each of these protein molecules is displayed when mutations in the genes involved occur. Such genetic alterations may lead to the synthesis of a truncated protein due to a premature termination codon mutation, or missense mutations may alter the proteins conformation and affect important protein-protein interactions. Such genetic mutations generally lead to a weakness in the skin at the epidermal-dermal junction, resulting in clinical manifestation of the symptoms of skin fragility and blistering which is characteristic of the different forms of EB (Uitto and Richard, 2005).

1.3 Epidermolysis Bullosa

Epidermolysis Bullosa constitutes a group of rare disorders which manifest with a mechanical fragility of epithelial lined or surfaced tissues, most notably the skin. It is a genetically heterogeneous disease, allowing for its classification into a number of different types and subtypes according to the genetic mutation in question, as well as the location of lesions and the resulting clinical features which are observed (Fine, 2010b). Mutations in at least ten distinct genes have so far been implicated. Three

major forms of the disease have been identified and these comprise EB simplex, junctional EB and dystrophic EB. EB may be further subdivided, according to the specific gene mutations involved, all of which encode proteins which are intimately involved in the structural stability of keratinocyte cells or adhesion of keratinocytes to the underlying dermis (Mitsuhashi and Hashimoto, 2003).

Clinical symptoms of EB vary widely and range in severity, from localised blistering of the hands and feet to generalised blistering of the skin and oral cavity, as well as injury to many internal organs (Figure 1.3). Diagnosis of the disease is initially based on examination of the clinical symptoms and elicitation of the personal and family medical history of the patient, while final diagnosis mainly depends on DNA-based mutation analysis (Sawamura *et al.*, 2010). While no cure currently exists, treatment of affected patients is based around the protection of susceptible tissues against trauma, treatment of lesions with sophisticated wound care dressings and nutritional support. Prognosis varies significantly and is dependent on both the subtype of the disease and the overall health of the patient in question (Fine, 2010b). Epidermolysis Bullosa is a life-long debilitating illness, which has profound effects on many aspects of the lives of both the patient and their surrounding family members.

1.3.1 EB Simplex

Epidermolysis Bullosa Simplex is the most common form of the disease (approximately 1 case per 25,000 live births), as well as being the least severe. It is the basal layer of epidermal keratinocytes which are affected and an inherited defect causing EB simplex results in their rupture when the epidermis is subjected to mechanical stress (Coulombe, 2009). The major subtypes of EB Simplex are inherited in an autosomal dominant manner, due to dominant negative mutations in the keratin 5 (KRT5) and keratin 14 (KRT14) genes. These genes encode type I and type II intermediate filament (IF) proteins which are expressed in the epidermal basal keratinocytes and related complex epithelia and these keratin filaments are associated with the basement membrane zone complex (Fuchs, 1995). The major subtypes are EB simplex-generalised (in which the distribution of blistering is generalised over the body), EB simplex-localised (in which the distribution of blistering is localised) and EB simplex Dowling-Meara (in which the distribution of blistering is generalised but

shows a distinct clustered pattern). Other rarer subtypes of EB Simplex arise due to mutations of the plectin gene. These are typically recessively inherited and are often associated with late-onset muscular dystrophy (Coulombe, 2009; Reznicek *et al.*, 2010). Clinically, EB simplex D-W is the most severe subtype, with symptoms at birth including widespread blistering, erosions and areas of denuded skin (Uitto and Richard, 2004).

1.3.2 Junctional EB (JEB)

In the case of the junctional form of Epidermolysis Bullosa, tissue separation occurs within the cutaneous basement membrane zone. There are two main subtypes of Junctional EB: the Herlitz (lethal) subtype and the non-Herlitz (non-lethal) subtype. The more severe form is JEB-Herlitz, which has the highest risk of infant mortality among EB neonates, infants and young children, often due to overwhelming infection. However the most common Junctional EB subtype is that of the non-Herlitz variety (Fine, 2010b). It is most commonly mutations of the LAMA3, LAMB3 and LAMC2 genes encoding the constituent polypeptides $\alpha 3$, $\beta 3$, and $\gamma 2$ of laminin 332 protein, which are implicated in Junctional EB and the disease usually follows an autosomal recessive pattern of inheritance. As previously described, the anchoring filaments which are composed of laminin 332 protein are important constituents of the basement membrane zone and thus are vital for the skin to maintain its structural support role. JEB-Herlitz usually arises due to premature termination codon mutations in these genes. The non-Herlitz form generally harbours mutations in the same genes but these are usually missense or splice site mutations, which have milder consequences at the mRNA and protein levels (Varki *et al.*, 2006). Additionally, the COL17A1 gene which encodes bullous pemphigoid antigen, a transmembranous collagen, has been implicated in some cases of JEB. Both recessively (McGrath *et al.*, 1995) and dominantly (Almaani *et al.*, 2009) inherited COL17A1 mutations have been identified as causing JEB.

1.3.3 Dystrophic EB (DEB)

The dystrophic form of epidermolysis bullosa (DEB), in which tissue separation occurs in the dermis, is characterised by blistering, scarring and milia (small benign

keratinous cysts; Berk and Bayliss, 2008) formation. DEB is separated into two major subtypes based on the mode of transmission: DEB can be inherited in either an autosomal dominant or autosomal recessive pattern. Both forms are caused by mutations in a single gene, COL7A1. This gene encodes type VII collagen, which is the major component of the anchoring fibrils. Absence of the normal protein product of the COL7A1 gene or the presence of an aberrant protein product leads to a weakness of the basement membrane zone and thus fragility of the skin of such an affected individual, resulting in the displayed symptoms of dystrophic EB (McGrath *et al.*, 1993). Dystrophic EB has been further divided into subtypes, in accordance with variations in the particular mutations of the COL7A1 gene and the clinical features observed (Mitsunishi and Hashimoto, 2003).

A number of subtypes of recessive dystrophic EB (RDEB) have been identified, the main ones being RDEB-severe generalised and RDEB-generalised other (Fine *et al.*, 2008). The more severe subtype is RDEB-severe generalised, which represents one of the most devastating and debilitating multi-organ genetic diseases of mankind. This form of the disease may result in such symptoms as generalised skin blistering at birth, progressive and mutilating scarring of the skin, blistering of other epithelial tissues leading to severe injury of gastrointestinal and genitourinary tracts, blistering and scarring of the cornea, mouth blistering, growth retardation and failure to thrive, esophageal strictures and pseudosyndactyly of the hands and feet (fusion of the hands and feet resulting in mitten-like deformities and thus severe loss of function) (Fine, 2010b). RDEB-severe generalised patients are also at a high risk of development of squamous cell carcinoma (SCC) as chronic inflammation of affected skin cells may cause mutations in the cells DNA. The lifetime risk of aggressive SCC is approximately 90%, with most patients dying of metastatic SCC within five years of the initial diagnosis (Fine *et al.*, 2009). The generalised other-recessive subtype has a similar but less severe phenotype, with localisation of blistering to hands, feet, knees and elbows and usually without the severe scarring seen in RDEB. RDEB-severe generalised patients usually have a complete lack of COL7A1-encoded anchoring fibril structures while RDEB-generalised other patients usually have some COL7A1 expression, but at reduced levels compared to non-affected individuals. Other rarer RDEB subtypes include RDEB inversa, RDEB pretibial, RDEB pruriginosa and RDEB centripetalis (Fine *et al.*, 2008).

The main subtype of dominantly inherited dystrophic EB (DDEB) is generalised DDEB, which encompasses the two previously separate subtypes Pasini and Cockayne-Touraine DDEB (Fine *et al.*, 2008). The phenotype of DDEB is generally milder in comparison to that of severe generalised-RDEB. DDEB patients usually show generalised blistering at birth, atrophic or hypertrophic scarring and nail dystrophy (Figure 1.4). Blistering and erosion of the esophagus is also common, leading to constriction of the esophagus and difficulty in swallowing (dysphagia). Other rarer subtypes include DDEB accral, DDEB pretibial and DDEB pruriginosa (Fine *et al.*, 2008). Approximately 70% of patients diagnosed with DDEB are reported to have an affected parent (Pfundner and Lucky, 2006). DDEB mutations are strongly associated with glycine substitutions in one allele of the COL7A1 gene encoding type VII collagen protein (see section 1.4).

1.4 Collagen VII protein and gene

An abundance of collagen proteins, known as the collagen superfamily, are found in the extracellular matrix of the human body, where their main function is to provide strength and support. They are an important element of vertebrate tissues such as cartilage, tendon, bone and skin. Collagens also play a significant role in cell migration, angiogenesis, cancer, tissue morphogenesis and tissue repair (Kadler *et al.*, 2007). These proteins are subdivided, according principally to their structure, although all members are similar in that they are trimeric molecules composed of three α polypeptide chains containing the tripeptide repeat (Gly-X-Y)_n, where X is frequently proline and Y is frequently hydroxyproline. The repeats allow formation of a triple helical structure, characteristic of the proteins of the collagen superfamily (Ricard-Blum and Ruggiero, 2005). Flanking the triple helical regions (also known as Col domains) are non-(Gly-X-Y) regions, termed non-collagenous domains (Gordon and Hahn, 2010). Each polypeptide chain of the triple helical domain forms a left handed helix, with 18 amino acid residues per turn, and the three chains are wound around a central axis to form the triple helical structure. Being the smallest of the amino acids, the non-bulky glycine residues are generally located in the centre of the triple helix, allowing the collagen molecule to form a tight compact structure, consistent with its role in providing support to organs such as the skin (Gelse *et al.*,

2003). The importance of the collagen proteins is illustrated by the resulting phenotypes in a wide range of diseases caused by collagen-encoding gene mutations. These include Osteogenesis Imperfecta (Collagen I; Marini *et al.*, 2007), Alport syndrome (Collagen IV; Kashtan, 1999) and as previously mentioned, dystrophic Epidermolysis Bullosa (Collagen VII).

1.4.1 Collagen VII protein structure

The collagen superfamily is divided into two major groups based mainly on the types of supramolecular organisations which they form. This includes the fiber-forming fibrillar collagens, which are composed of uninterrupted triple helical domains, and the non-fibrillar collagens, which are composed of α -chains with both triple helical and non-triple helical interrupting regions of varying lengths. Type VII Collagen is a non-fibrillar collagen molecule and is a homotrimer of three $\alpha 1(\text{VII})$ chains (Christiano *et al.*, 1994a). Each of these chains is composed of 2944 amino acids, with a molecular mass of approximately 300 kDa. The glycine-rich collagenous segment of the polypeptide spans 1530 residues, although it contains some interruptions, including a 39 residue non-collagenous “hinge” region (Pulkkinen and Uitto, 1999). At the amino-terminal region of the polypeptide is the non-collagenous NC-1 region of approximately 145 kDa, while at the carboxy-terminal end the non-collagenous NC-2 domain, of roughly 18 kDa, resides. The NC-1 domain consists of sub-elements which have homology to known adhesive proteins, including cartilage matrix protein, fibronectin type III-like domains and the A domains of von Willebrand factor, while the NC-2 domain has homology to the Kunitz protease inhibitor (Christiano *et al.*, 1992; Greenspan, 1993).

Initially, each $\alpha 1(\text{VII})$ chain is translated to a 350 kDa polypeptide, termed pro $\alpha 1(\text{VII})$ (Figure 1.5). Each of these three chains subsequently associates with one another to form homotrimeric type VII collagen monomers which are secreted from the keratinocyte cells where they are primarily synthesised. These monomers may then interact to form antiparallel dimers which overlap at the carboxy-terminal ends and are stabilised by intermolecular disulfide bonds. Cleavage of a segment of the NC-2 domain also occurs at this stage, resulting in the approximate 300 kDa $\alpha 1(\text{VII})$ chains (Figure 1.6). It is then possible for these collagen dimer protein molecules to

assemble into anchoring fibrils at the basement membrane zone (Bruckner-Tuderman *et al.*, 1999; Varki *et al.*, 2007). Intertwining of the anchoring fibrils between interstitial collagen fibers, consisting mainly of types I, III and V collagens, allows for the stable association of the dermal-epidermal basement membrane zone to the underlying dermis (Pulkkinen and Uitto, 1999).

1.4.2 The COL7A1 locus

The human type VII collagen gene (COL7A1) has been mapped to the short-arm of chromosome 3, region 3p21.1 (Parente *et al.*, 1991). Strong genetic linkage between the COL7A1 locus and dystrophic EB was identified after the discovery and exploitation of COL7A1 intragenic and flanking polymorphic markers, thus implicating mutations of COL7A1 in the dominant and recessive DEB disease phenotypes (Gruis *et al.*, 1992; Hovnanian *et al.*, 1992). The gene is 31,132 nucleotides in length from the transcription start site to the site of polyadenylation. After COL7A1 is transcribed and post-transcriptionally modified, the resulting mRNA sequence is approximately three times smaller in length than its DNA precursor, measuring approximately 9kb. The extensive length of the transcript suggests that the introns of the gene are relatively small in comparison with other genes, and a large part of the gene is composed of coding exons. There are in fact 118 exons within the type VII collagen gene and the average length of introns within COL7A1 is just 188 nucleotides (Christiano *et al.*, 1994a). By comparison, the much larger 2.4mb dystrophin gene is composed of just 79 exons (Roberts *et al.*, 1993). The NC-1 domain of COL7A1 is encoded by 27 exons. The triple helix coding region of the gene is composed of 84 exons with the last of these, exon 112, along with 6 additional exons, encoding the beginning of the NC-2 domain. Exons 70, 71 and 72 encode non-triple helical interruptions of the triple helix coding domain. All other exons of the triple helix coding domain begin with a codon for glycine (Christiano *et al.*, 1994a).

1.4.3 Mutations of the COL7A1 gene

A number of mutations of the human collagen type VII gene have been identified, resulting in either autosomal dominant or recessive dystrophic EB, the symptoms of

which may range in severity according to the particular mutation in question (Figure 1.7). Currently over 300 pathogenic mutations have been identified, although novel mutation sites are still continuously being detected (Dang and Murrell, 2008, Jerábková *et al.*, 2010). With regard to recessive DEB, the majority of cases are caused by either nonsense mutations, or small deletions or insertions, resulting in either a premature stop codon (PTC) or a frameshift mutation in the DNA of affected individuals. Such mutations have been identified at numerous nucleotide positions throughout the COL7A1 gene (Varki *et al.*, 2007). The RDEB-severe generalised form of the disease usually arises due to PTC mutations, resulting in either nonsense-mediated decay of the mRNA or truncated polypeptides which are unable to assemble into functional anchoring fibrils (Pulkkinen and Uitto, 1999; Dang and Murrell, 2008). When an individual has one mutated allele and one wild type allele, such mutations are generally silent. However, when an individual is homozygous for the mutant allele, the resulting significant alteration in levels of collagen type VII protein leads to the skin fragility associated with the most severe form of recessive DEB (Christiano *et al.*, 1997). The milder forms of RDEB generally arise due to compound heterozygous COL7A1 mutations. A PTC mutation of one allele along with a missense mutation in the other, results in production of full-length type VII collagen polypeptides by the affected individual. These polypeptides however have different conformations, which affects the stabilisation of the anchoring fibrils due to structural changes (Dang and Murrell, 2008).

Dominant dystrophic EB generally arises as a result of missense mutations at the COL7A1 locus, although deletions and splice-site mutations have also been implicated in some cases. The missense mutations usually occur within the triple-helix coding domain of the gene and typically result in substitutions of glycine residues. Approximately 60 glycine substitutions in COL7A1 have been shown to result in a dominantly inherited DEB phenotype (Nakamura *et al.*, 2004). Although such missense mutations have been identified at a range of positions inside the triple-helix coding region, they occur most frequently within exons 73-75 (Varki *et al.*, 2007). As such changes in the DNA sequence do not generally result in nonsense mutations, the entire collagen type VII gene may be transcribed and translated, but the resulting protein molecule is altered. Such mutated polypeptides are capable of interacting with other wild type polypeptides to form collagen homotrimeric

structures, as well as having the ability to form antiparallel dimers with other wild type homotrimers. However, the dominant negative interference which arises as a result of these glycine substitutions leads to a weakening of the anchoring fibrils and thus increased fragility of the skin. Assuming equal expression of both the mutant and wild-type alleles, seven out of every eight collagen VII trimeric molecule includes at least one mutant pro- α 1 chain, while just one-eighth consists solely of normal polypeptides (Dang and Murrell, 2008). The fact that it is still possible for these anchoring fibrils to be synthesized despite the existence of mutations in the COL7A1 gene explains the reason why the severity of the symptoms of some cases of dominant DEB is relatively mild in comparison with other forms of the disease such as certain severe cases of recessive DEB, where the anchoring structures may be completely absent. The severity of the symptoms displayed by an affected DDEB is also affected by the position within the triple helix of the mutated amino acid and the structure of the particular residue which replaces the wild type glycine residue (Varki *et al.*, 2007).

1.5 Therapeutic approaches for Epidermolysis Bullosa

Advances in the field of molecular biology in recent decades, including the establishment of technology which allows for the successful cloning and manipulation of genes *in vitro*, have led to studies into the possible treatment of a number of genetic diseases, in a process known as gene therapy (Weissman, 1992). Dominant dystrophic epidermolysis bullosa represents just one of such diseases which may benefit as a result of continuing advances in the field of gene therapy. Gene therapies directed towards recessive disorders often involve attempts to introduce the expression of a wild-type gene which in essence has been suppressed in the mutated form, causing a disease phenotype. In contrast, gene therapy-based strategies for dominant disorders may require prevention of expression of the mutant gene along with the possible need for concurrent delivery of the wild-type gene. Alternatively, attempts to directly repair the defective gene in question may be possible in both dominant and recessive cases of a particular disease. For gene therapy-based strategies, an optimal delivery strategy needs to be considered (viral vs. physical/chemical) depending on the tissue in question, as well as the method of correction (*in vivo/ex vivo*). Other therapeutic

approaches which are currently being investigated for the treatment of genetic disorders such as Epidermolysis Bullosa include cell-based and protein replacement therapies.

One initial success, in the treatment of recessive junctional EB, has been reported in recent years. This study involved an *ex vivo* viral delivery gene therapy-based approach (Mavilio *et al.*, 2006). As previously described, this form of EB is often caused by mutations in one of the three chains of laminin 332, a heterotrimeric protein made of $\alpha 3$, $\beta 3$, and $\gamma 3$ chains, and which is a key constituent of the dermal-epidermal junction. Scientists successfully introduced expression of the LAMB3 gene into cultured epidermal stem cells taken from an affected patient with a LAMB3 mutation, using a Moloney murine leukaemia virus (MLV)-derived retroviral vector carrying the full-length LAMB3 cDNA under the control of the viral long terminal repeat (LTR) promoter. Subsequent transplanting of corrected skin grafts onto the upper-leg regions of the patients body resulted in stabilisation of the epidermis and relief from the symptoms of junctional EB in the treated areas. More than 5 years of follow up has seen persistence of stable, non-blistered skin in the transplanted area, either spontaneously or upon being rubbed, compared to the chronic blistering lesions visible in the surrounding non-treated regions (Tamai *et al.*, 2009). Continued expression of laminin 332 protein has been detected and additionally, no inflammation or immune response against the transgene product has been observed (De Luca *et al.*, 2009).

The success of this initial clinical trial seems promising for future gene therapy methods which may be directed at treatment of other forms of EB and it is hoped that the general applicability of such an approach will soon be realised. Regarding dystrophic EB however, a number of issues have arisen which need to be considered for any gene therapy approach for this form of the disease. Firstly, the human COL7A1 mRNA transcript is large, at approximately 9kb in length, posing technical challenges for effective gene delivery. Furthermore, duration of expression of any delivered COL7A1 gene and its long-term effect on human epidermal homeostasis and immune surveillance needs to be further investigated (Siprashvili *et al.*, 2010). It is also of note that the precise level of COL7A1 expression required for a wild-type phenotype is currently not fully understood. Individuals who are

heterozygous carriers of a nonsense COL7A1 mutation exhibit a normal non-blistering phenotype, suggesting that the levels of type VII collagen required to prevent blistering lies somewhere in the region of 0 to 50% (Tidman and Eady, 1985). Furthermore, whilst COL7A1-null mice usually die within the first two weeks of life (Heinonen *et al.*, 1999), transgenic mice which express collagen type VII at 10% of normal levels exhibit a milder phenotype and survive to adulthood (Fritsch *et al.*, 2008). Therefore, it is possible that the presence of a relatively low amount of normal collagen VII protein may suffice for the maintenance of a normal dermal-epidermal adhesion zone.

1.5.1 Therapeutic approaches for recessive dystrophic EB (RDEB)

Gene therapy approaches for recessive dystrophic EB have involved research into the use of lentiviral vectors, retroviral vectors and non-viral based vectors for restoration of collagen type VII protein levels. Transduction of lentiviral vectors carrying the COL7A1 gene into collagen VII-deficient recessive DEB keratinocyte and fibroblast cell lines has resulted in the persistent production of type VII collagen. Transplantation of these gene-corrected cells onto immune-deficient mice resulted in successful formation of anchoring fibrils *in vivo* at the basement membrane zone (Chen *et al.*, 2002). Subsequent direct intradermal injection of this lentiviral vector into RDEB skin grafted onto muscle fascia of severe combined immunodeficient (SCID) mice (Wang *et al.*, 2000), resulted in restoration of collagen type VII protein production and anchoring fibril formation. A single lentiviral vector injection resulted in stable type VII collagen protein at the BMZ for at least 3 months. Thus these data demonstrated efficient and long-term type VII collagen gene transfer *in vivo* using direct intradermal injection of an engineered lentiviral vector (Woodley *et al.*, 2004). A non-viral vector approach has involved use of the phi C31 bacteriophage integrase, which resulted in stable integration of the COL7A1 cDNA into primary epidermal progenitor cells from four unrelated RDEB patients. Skin regenerated using these cells resulted in a correction of the RDEB disease phenotype including expression of type VII collagen, formation of anchoring fibril structures and cohesion of the dermal-epidermal skin layers (Ortiz-Urda *et al.*, 2002).

Retroviral vectors have also been utilised in the transfer of the 9kb COL7A1 cDNA into primary human RDEB keratinocytes, resulting in successful secretion of functional collagen type VII protein. Fibrin-based skin equivalents made with the transduced RDEB keratinocytes and grafted onto SCID mice generated cohesive and orderly stratified epithelia with all the characteristics of normal human epidermis, including formation of anchoring fibrils (Gache *et al.*, 2004). More recently, further significant advances have been made with the use of retroviral vectors for delivery of the COL7A1 gene. Self-inactivating (SIN) retroviral vectors have recently been developed in which the COL7A1 cDNA is under the control of the human elongation factor 1 (EF1) or COL7A1 promoter. It is thought that these new minimal self-inactivating vectors would minimise the risk of oncogenic events associated with the use of a retrovirus. Using these vectors, long-term expression of type VII collagen along with anchoring fibril formation and dermal-epidermal adherence was achieved, upon *ex vivo* correction of RDEB keratinocytes and fibroblasts and subsequent grafting of the corrected cells onto immunodeficient mice (Titeux *et al.*, 2010). In a separate study, retroviral delivery of full-length COL7A1 cDNA to RDEB human keratinocytes followed by regeneration of corrected skin grafts onto immunodeficient mice resulted in long-term type VII collagen expression *in vivo*. Type VII collagen expression was maintained for 1 year, leading to restoration of anchoring fibrils at the BMZ and thus correction of the RDEB phenotype. No noticeable tissue toxicity was observed and so this study has shown corrective, durable gene delivery to long-lived epidermal progenitor cells without causing any adverse reactions. Therefore these preclinical studies have provided the foundation for a human phase I clinical trial of *ex vivo* gene delivery in RDEB, which has been approved by the United States Food and Drug Administration (Siprashvili *et al.*, 2010), the results of which have yet to be reported.

Recent investigations into a method of correction of type VII collagen expression in dystrophic EB have also shown some promising results (Murauer *et al.*, 2011). Such an approach may be applicable not only for the treatment of the recessive form of the disease, but may also represent a method of correction of those mutations which are inherited in a dominant manner. In this *in vitro* study using RDEB keratinocytes, retroviral transduction of the cells with a 3' pre-*trans*-splicing molecule resulted in correction of full-length collagen type VII expression. The mechanism in

question of *trans*-splicing utilises the cells own spliceosome machinery to recombine an endogenous target pre-mRNA and an exogenous RNA molecule called a pre-*trans*-splicing molecule (PTM). A new reprogrammed mRNA containing the desired nucleotide sequence is produced by *trans*-splicing of the PTM into the 3' or 5' sequence of the target pre-mRNA (Puttaraju *et al.*, 1999; Wally *et al.*, 2008). In this COL7A1 correction study, normal localisation of type VII collagen protein was found at the basement membrane zone of skin equivalents from transduced cells and this resulted in assembly of anchoring fibril-like structures (Murauer *et al.*, 2011). The delivered PTM used encompassed just 3.3kb and thus eliminates the need for delivery of the large 9kb full-length COL7A1 cDNA. The single PTM construct in question also in theory could allow for correction of one third of all COL7A1 mutations, of both recessive and dominant types.

The strategy of protein replacement therapy for EB is based on the concept of administration of recombinant protein to the skin by topical application, local injection or systemic administration to the circulation. Such an approach may be amenable to both dominant and recessive forms of the disease as it would potentially allow for either replenishment of an absent protein or replacement of a defective protein (Uitto, 2009). Recent studies have involved investigation of the direct delivery of recombinant human type VII collagen protein to COL7A1-null mice, an approach which would bypass requirement of delivery of the large COL7A1 gene. Intradermal injection of purified human collagen VII into RDEB mice resulted in stable incorporation of the protein into the basement membrane zone, decreased skin fragility and markedly prolonged survival compared to non-treated mice. Untreated COL7A1-null mice usually die within their first week of life due to extreme fragility of the skin and mucous membranes due to absence of anchoring fibrils at the BMZ. However intradermal injection of the recombinant type VII protein resulted in a significantly prolonged survival of these mice, with some surviving as long as 20-25 weeks (Remington *et al.*, 2009). Additionally, despite development of circulating anti-human type VII collagen antibodies in treated mice, these antibodies did not appear to be pathogenic as they failed to bind to the mouse BMZ nor did they prevent incorporation of newly injected human collagen VII protein into the BMZ. Furthermore, formation of these antibodies was prevented by treatment of the injected mice with an anti-CD40L monoclonal antibody MR1. Thus it is conceivable that

protein replacement therapy may represent an option in the future treatment of DEB (Remington *et al.*, 2009).

Further recent attempts to develop molecular strategies for the treatment of DEB have involved research into cell-based therapies. Human fibroblast cells have been shown to migrate to wounded areas of skin upon their intravenous injection into RDEB mice, where they deliver collagen type VII and promote wound healing (Woodley *et al.*, 2007). Following these findings, pilot studies in which five RDEB patients were given single intradermal injections of autologous or allogeneic fibroblasts have seen some promising results (Wong *et al.*, 2008). The skin of treated patients showed a reduced tendency to blister following injection, with an increase in local type VII collagen expression. The benefits of these cell injections were found to last for at least three months. However, it was later determined that these benefits had arisen due primarily to cytokine-mediated up-regulation of mutant type VII collagen expression in those patients who had residual levels of activity from their mutant alleles. Synthesis of normal type VII collagen protein from the newly introduced cells was not apparent (Wong *et al.*, 2008). These studies suggest that those RDEB patients who display some baseline synthesis of partially functional collagen VII protein may benefit from this cell-based therapy, whereas those patients completely lacking type VII collagen expression might not (Uitto, 2008; Wong *et al.*, 2008).

Another cell-based therapy approach for RDEB which has recently been investigated involves bone marrow stem cell transfer. Bone marrow-derived cells have been found to have the ability to differentiate into cell types which are important for both skin maintenance and for aiding in repair of damaged skin (Badiavas, 2004). Studies in which GFP-expressing mice were used as a source of bone marrow cells for transplantation of RDEB mice have recently provided evidence that a population of bone marrow cells may signify a potentially valuable method of treatment of RDEB. COL7A1-null mice were found to produce collagen type VII protein after injection at birth or within a few days of birth of isolated subpopulations of these bone marrow cells. Such injections resulted in formation of anchoring fibrils, reduced skin fragility and reduced lethality (Tolar *et al.*, 2009). In a similar study, embryonic bone marrow cell transfer to RDEB mice resulted in amelioration of the RDEB phenotype at birth, with an extension of survival of up to several weeks compared to non-treated control

RDEB mice (Chino, 2008). Based on the results of these animal studies, human clinical trials using allogeneic stem cell transplantation have recently been undertaken for the treatment of RDEB patients. Seven children suffering from RDEB were treated with immunomyeloablative chemotherapy and allogeneic stem cell transplantation. Although further studies are required in order to assess the long-term risks and benefits of such a therapeutic approach, initial results do seem promising. While one of the patients died of cardiomyopathy before transplantation, all six recipients were found to have substantial proportions of donor cells in the skin and none had detectable anti-collagen VII antibodies. In five of the six recipients, an increase in collagen VII protein was also observed at the dermal-epidermal junction (Wagner *et al.*, 2010). These early observations suggest that in the future, bone marrow cell populations may be capable of correction of the basement membrane defect in human RDEB patients. Recent investigations into generation of induced pluripotent stem cells (iPSCs) from both normal human fibroblast cells as well as fibroblasts from RDEB patients have also yielded promising results (Itoh *et al.*, 2011). These iPSCs were successfully differentiated into keratinocyte cells at high efficiency and were used to generate 3D skin equivalents, highlighting the potential for iPSCs to provide a novel approach for applying regenerative medicine to cutaneous diseases such as DEB (Itoh *et al.*, 2011; Uitto, 2011a).

1.5.2 Therapeutic approaches for dominant dystrophic EB (DDEB)

Gene therapy strategies aimed at the treatment of dominant disorders generally differ considerably from those used to treat recessive ones, as it is likely that a gene replacement approach will not suffice regardless of the efficiency of the replacement system. Dominant disorders may be caused by a decline in the amount of wild-type protein which is being synthesised; the disease phenotype may arise due to a gain of function mutation; or a combination of the two. Thus gene therapies directed at the treatment of dominant disorders may require suppression of the mutant allele while also maintaining expression of the wild type allele (Millington-Ward *et al.*, 1997). This is likely the case for DDEB due to the dominant negative nature of the mutations in question, causing the disease phenotype. No gene therapy-based clinical trials for treatment of DDEB patients have yet been reported. Gene correction by either *trans*-splicing or homologous recombination represents possible methods of investigation

for silencing the mutant allele. As discussed in section 1.5.1, recent investigations into the use of pre-*trans*-splicing molecules (PTMs) for correction of mutant COL7A1 have demonstrated some success *in vitro* and hence may hold promise for the provision of a tool for future treatment of both the dominant and recessive forms of dystrophic EB (Murauer *et al.*, 2011). Immense progress has also been made in the recent past with regard to homologous recombination due to the introduction of zinc-finger nuclease technology (Durai *et al.*, 2005; Davis and Stokoe, 2010; Le Provost *et al.*, 2010). This class of DNA-binding proteins facilitates editing of the genome by creating double-stranded breaks in DNA, thus stimulating homology-directed gene repair. Thus in principle, it is possible to replace a region of mutated DNA by delivery of a copy of a corrected version, which is incorporated into the chromosomal site by homologous recombination (Kim *et al.*, 1996; Kandavelou *et al.*, 2009; Van Nierop *et al.*, 2009). Another possible method of silencing the dominant mutated COL7A1 allele may involve the use of RNA interference (RNAi) technology, which is based on sequence specific suppression of target mRNA molecules and which has been shown to hold great promise for the treatment of a wide range of dominant diseases (Kiang *et al.*, 2005; Kim and Rossi, 2007; Boudreau and Davidson, 2010). Such an approach may require concurrent introduction of the wild-type genes protein product for some dominantly inherited disorders (see section 1.7).

1.6 RNA interference (RNAi)

RNA interference (RNAi) is a relatively simple and rapid process which occurs in a wide range of organisms, whereby the expression of targeted genes may be silenced. RNAi involves inhibition of the production of proteins due to suppression of mRNA molecules, mediated by double stranded RNA. The field of RNAi has grown rapidly in the past decade, with the elucidation of the RNAi pathway and the discovery of its involvement in crucial biological events. The discovery of RNAi has provided a potent tool with which to study gene function and which has the potential to revolutionise the treatment of many diseases. Despite it being just slightly more than a decade since its discovery, RNA interference is already being exploited therapeutically in human clinical trials. Currently, several RNAi-based clinical trials are underway for treating human diseases such as age-related macular degeneration,

acute renal failure, skin disease and even cancer (Davis, 2009; Kaiser *et al.*, 2010; Leachman *et al.*, 2010).

1.6.1 Discovery of RNAi

The process of RNAi was first reported in 1990 when attempts to overexpress chalcone synthase (CHS), a protein involved in the production of the purple colour of petunias, by introduction of a CHS transgene, resulted in the production of some white flowers instead of the expected darker coloured purple ones. It was theorised that the endogenous CHS gene was being cosuppressed by the newly introduced transgene (Napoli *et al.*, 1990). Similar cosuppression events were observed in *Neurospora Crassa* (Romana and Macino, 1992) and *Caenorhabditis elegans* (Guo and Kemphues, 1995) but it was not until 1998 that an explanation was finally provided. While investigating methods of manipulating gene expression in *C. elegans*, Andrew Fire and Craig Mello introduced both single stranded (ss) and double stranded (ds) RNA molecules into the *C. elegans* genome in an effort to interfere with the *unc-22* gene. They found that silencing of *unc-22* by dsRNA was significantly more effective than the silencing achieved by ssRNA molecules and thus they deduced that the observed gene silencing was occurring as a result of the presence of double stranded RNA molecules, in the process which is now referred to as RNA interference (Fire *et al.*, 1998). While Fire and Mello were unsure of the precise mechanism of RNA-mediated interference in *C. elegans*, studies were soon undertaken on identifying the components involved in the process. Multiple genes were soon determined to be involved in the *C. elegans* system (Tabara *et al.*, 1999) and it is now known that the mechanism of RNA mediated silencing involves a pathway in which several proteins act on the double stranded RNA, resulting in the eventual degradation of sequence specific target mRNA or inhibition of its translation. It is not just the nematode in which such a system occurs; similar pathways are now known to exist in a range of organisms from plants to humans. The cells endogenous RNAi pathway is thought to have arisen as a mechanism of inhibition of transposon mobilization (Tabara *et al.*, 1999) and as an antiviral mechanism in plants (Kasschau and Carrington, 1998).

1.6.2 MicroRNAs

MicroRNAs (miRNAs) represent one class of silencing molecule which is endogenous to the human genome. They function in the regulation of gene expression at the post-transcriptional level by inhibiting the expression of mRNAs comprising fully or partly homologous target sequences. Thus microRNAs play important roles in numerous cellular processes including development, stem cell division, apoptosis, cancer and disease (Zimmerman and Wu, 2011). It is estimated that approximately 3% of human genes code for microRNAs, while these miRNAs may regulate a further 30% of the protein coding genes (Wang and Wu, 2009). Advancement in research in the area of silencing molecules resulted in a merging of the RNAi pathway with the miRNA pathway by showing that core components are closely shared.

miRNAs are initially transcribed by RNA polymerase II as primary miRNA precursors (pri-miRNA) of several kilobases in length containing a stem-loop structure. After cleavage of the hairpin structure by a complex consisting of Drosha, an RNase III enzyme and DiGeorge critical region gene eight protein (DGCR8), small hairpin precursor miRNAs (pre-miRNAs) of approximately 70bp are generated (Zeng *et al.*, 2005a). These pre-miRNAs are subsequently exported from the nucleus by exportin-5, a nuclear transport receptor (Yi *et al.*, 2003), and further processed by numerous protein complexes until they are incorporated into a microRNA RNA-induced silencing complex (RISC), where they may function to down regulate the activity of homologous mRNAs (Cullen, 2005). It is components of this RISC which the miRNA pathway shares with the RNAi pathway (section 1.6.3). MicroRNAs regulate gene expression by either inhibiting translation or promoting degradation of specific mRNA transcripts, depending on the degree of complementarity between the microRNA and its target mRNA transcript. Higher complementarity generally results in mRNA degradation while lower complementarity generally results in translational attenuation (Zimmerman and Wu, 2011).

1.6.3 The RNAi pathway: siRNA- and shRNA-based therapies

Artificial RNAi molecules offer a powerful tool for the down regulation of various mRNA molecules. Such suppression would be extremely useful in the treatment of various diseases, by preventing translation of mutated protein molecules. It is now

possible to synthesise small double stranded RNA molecules to target almost any gene within the human genome (Davidson and Paulson, 2004). Gene silencing methods using artificial double stranded RNA molecules may involve the direct introduction of small interfering RNAs (siRNAs) into the cell, or the introduction of short hairpin RNAs (shRNAs) which first must be processed in the cell's nucleus into siRNAs (Kim and Rossi, 2007). Each of these molecules is then processed by cytoplasmic machinery which is endogenous to the cell (Figure 1.8). shRNAs are generally introduced into the cell by means of vectors. Either plasmids or viral vectors which have been engineered to express the shRNA are incorporated into the nucleus. shRNAs generally consist of a sense sequence composed of about 21 bases, followed by a 6 to 8 base non-complementary loop and an antisense sequence which is complementary to the sense sequence (Davidson and Paulson, 2004). The shRNA is transcribed from a specific promoter and is then processed and exported from the nucleus where it is further processed by Dicer, an RNAseIII-like enzyme, resulting in a short double stranded siRNA molecule. The Dicer enzyme is composed of a double stranded RNA binding domain (dsRBD), two catalytic RNase III-like domains, a ssRNA-binding piwi-argonaute-zwille (PAZ) domain and an RNA helicase/ATPase domain (Zhang *et al.*, 2002).

In the cytoplasm of the cell, siRNAs, which are usually composed of approximately 21-23 base pairs and have 2-nucleotide 3' overhangs, interact with a multisubunit ribonucleoprotein complex known as the RISC, as defined previously. The siRNA antisense strand (known as the guide strand) is then incorporated into the RISC and binds to its complementary mRNA target. Selectivity of strand incorporation is dependent on differential thermodynamic stabilities of the ends of the siRNA; the end which is least thermodynamically stable unwinds the 5' end of the guide strand and is cleaved (Geusens *et al.*, 2009). If the two mRNA and siRNA guide strand are perfect complements around the site of cleavage, which is usually roughly 10 to 11 bases upstream of the 5' end of the guide strand, the mRNA will subsequently be completely degraded by endonuclease digestion (Elbashir *et al.*, 2001a). If on the other hand the two strands are not perfect matches, as is often the case with the cell's endogenous microRNAs, the mRNA target is not cleaved, but its translation is suppressed (Olsen *et al.*, 1999; Lares *et al.*, 2010). The RNase H enzyme Argonaute 2 (AGO2) is the RISC subunit which is responsible for cleavage

of target mRNA in humans (Rana, 2007). The cleavage site of the mRNA target sequence is aligned with the AGO2 PIWI endonuclease domain and AGO2 cleaves the phosphodiester bond on the mRNA in the middle of the siRNA-mRNA recognition site. The cleaved mRNA fragments are subsequently released and recognised as aberrant transcripts, resulting in their degradation. The siRNA-loaded RISC can generally be used for multiple rounds of silencing, making RNAi molecules particularly suitable for therapeutic purpose (Manjunath and Dykxhoorn, 2010).

1.6.4 RNAi delivery

Possibly the biggest challenge in the use of RNAi-based therapies is the difficulty of delivery into the cell of these small molecules. Direct delivery of naked RNA may be possible, but although the simplest method, does not always result in high delivery efficiencies due to the difficulty associated with penetrating cellular lipid membranes (Li *et al.*, 2006). Thus the use of vectors, which facilitate the delivery of nucleic acids into the target cells and their nuclei, offer an extremely useful tool for RNAi based therapy methods. Such vectors are required to be safe for human exposure while also remaining safe themselves from degradation or from an immune attack posed by the infected host. Additionally, it should be possible to purify such vectors at high concentrations in high amounts at a relatively inexpensive cost. At present two major kinds of vector hold great promise for successful treatment of numerous genetic diseases. These include viral vectors such as lentivirus and adeno-associated virus and non-viral vectors, such as plasmids (Gardlík *et al.*, 2005).

Plasmid vectors which are designed to participate in RNA interference-based therapies are generally modified resulting in the cloning of shRNA sequences into the plasmid. Such plasmids have the ability to express the shRNAs as they are placed under the control of promoters dependent on RNA Polymerase III, such as U6 or H1 promoters. Once these shRNA molecules are expressed they may undergo intracellular processing, resulting in generation of cytoplasmic siRNA molecules which may enter the RISC and eventually degrade target mRNA (Wall and Shi, 2003). One such useful mammalian expression vector which can be used to direct the expression of siRNA molecules is pSuper (suppression of endogenous RNA) (Brummelkamp *et al.*, 2002a). Vectors which utilize the RNA Polymerase III H1 gene

promoter generally produce small RNA transcripts which lack a polyA tail. The promoter contains a distinct transcription start site and a termination signal of five thymidine nucleotides. Cleavage of the transcript occurs at the second uridine nucleotide at the termination site, resulting in an shRNA molecule with 2-nucleotide 3' overhangs on each strand. The shRNA transcript may then fold back on itself, with the two complementary strands binding to form the stem loop structure typical of an shRNA (Paddison *et al.*, 2002).

An additional form of delivery of shRNA molecules has been exploited in recent years, and this relatively new technology is based on the cell's own endogenous microRNA system. The shRNA are designed such that upon delivery to the cell they are released to form an intra-molecular stem loop structure similar to the structure of the cell's own endogenous pre-miRNA, which is then processed by the cell's RNAi machinery into a mature artificial miRNA. Complementary mRNA sequences may then be targeted for degradation. Such a system has a number of benefits, namely due to the polymerase II promoters in use. While the previously mentioned shRNAs are expressed from polymerase III promoters and thus are ubiquitously expressed in all tissues, the artificial microRNAs are under the control of polymerase II promoters, which can be replaced by other tissue specific or inducible promoters. This allows for a greater level of control of the potential therapeutic. Furthermore, these miRNA expression cassettes are thought to possibly represent a delivery method with reduced toxicity in comparison to previously widely used vectors, due to the fact that they are modeled on the cell's own endogenous miRNA machinery (McBride *et al.*, 2008; Boudreau *et al.*, 2009). shRNAs have been found to be expressed at significantly higher levels than artificial microRNAs and so concerns have arisen over the safety profiles of the shRNA molecules. High levels of processed and unprocessed shRNA may saturate the RNAi machinery, thus interfering with the cell's endogenous miRNA processing (Boudreau *et al.*, 2008). MicroRNA-based approaches seem to be less prone to interfering with such processes, thus suggesting that artificial miRNAs may be less toxic (McBride *et al.*, 2008; Boudreau *et al.*, 2008; Boudreau *et al.*, 2009).

1.7 Replacement

A major problem associated with the use of RNAi molecules as a therapeutic is the fact that the siRNA may not have the ability to distinguish between normal and mutated forms of the same allele. Whilst some success has been achieved in the isolation of siRNAs which distinguish between transcripts which differ by only a single nucleotide (Miller *et al.*, 2003; Hickerson *et al.*, 2008), often the optimal suppressor molecules for a particular mRNA target are found to target a region of the transcript which is common to both wild-type and mutant copies, and not the area in which the disease-causing mutation lies. Thus RNAi-mediated suppression of a particular gene may result in complete absence of that genes protein product, which in most cases would have detrimental effects on an individual's health if applied in human clinical trials. Furthermore, many disorders, including dominantly inherited dystrophic EB, are genetically heterogeneous, with the existence of hundreds of different mutations of the one gene affecting different families. Indeed, novel DEB-causing COL7A1 mutations are still continuously being detected (Jerábková *et al.*, 2010; Galehdari *et al.*, 2010). The establishment of different therapies for the treatment of each separate mutation may not be viable and so any potential therapy would generally likely require a mutation-independent knockdown approach, which would result in suppression of both the wild type and mutant gene (Kiang *et al.*, 2005; Palfi *et al.*, 2010; Millington-Ward *et al.*, 2011). Thus, central to most mutation-independent RNAi suppression therapies is the requirement for replacement of the wild type protein.

1.7.1 Gene replacement

Any mutation-independent RNAi-based therapeutic strategy, in which both the mutant and wild-type forms of the gene are suppressed, would likely require concurrent delivery of the wild-type gene. Such a gene however, must be resistant to the effects of the RNAi molecules in question. Taking advantage of the degeneracy of the genetic code is one method by which this resistance problem may be overcome. Replacement genes may be introduced which are genetically altered at the wobble base positions of codons. This results in the introduction of a gene which differs from the wild type gene in terms of the nucleotide sequence, but leads to translation of the same

sequence of amino acids. A gene of this nature would in principle be resistant to the effects of RNAi-mediated suppression. Such an approach would ultimately result in the absence of the RNAi-suppressed mutant gene and the presence of the genes wild-type protein product (Millington-Ward *et al.*, 1997; O'Reilly *et al.*, 2007; Millington-Ward *et al.*, 2011).

1.7.2 Protein replacement

It may however be possible to circumvent the need for a replacement gene, by direct introduction of the protein product of the gene being repressed. As previously mentioned (section 1.5.1), direct injection of collagen type VII protein is a possible viable therapeutic for the treatment of recessive dystrophic Epidermolysis Bullosa. Studies into such an approach are currently underway and some therapeutic benefit has been observed in COL7A1-null mice following intradermal injection with recombinant human type VII collagen protein (Remington *et al.*, 2009). With regard to the dominant form of the disease, in conjunction with suppression of both the mutant and wild-type alleles with an allele-unspecific RNAi approach, introduction of recombinant collagen VII protein may represent an alternative therapeutic approach, and one which would be independent of the effects of the RNAi molecules.

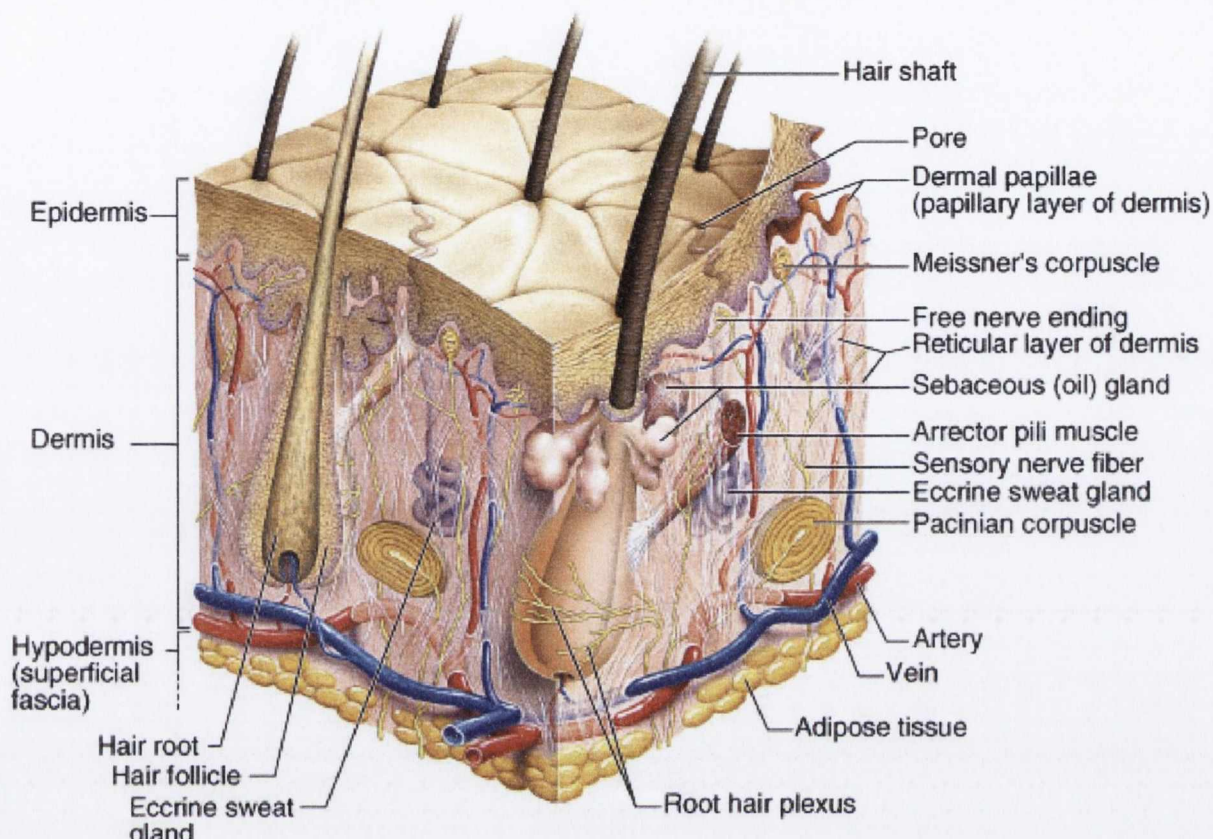
1.7.3 Mutation-dependent RNAi

One possible therapeutic approach in an RNAi-based study which would circumvent the requirement for concurrent delivery of a replacement gene or protein would be the undertaking of a mutation-dependent RNAi-based suppression approach. In such a case, RNAi molecules would be designed such that only the mutant allele is susceptible to suppression, whilst the wild-type allele should be resistant and thus transcribed and translated in the usual manner. A number of such studies have been investigated and reported (Ohnishi *et al.*, 2008; Scholefield *et al.*, 2009; Hohjoh, 2010; Atkinson *et al.*, 2011) including research into allele-specific RNAi for a genetic skin disease termed Pachyonychia Congenita (PC). This method involves the testing of a range of siRNAs covering the region of the mutation site, in an attempt to identify an siRNA molecule which suppresses the mutant allele but not the wild-type copy (Hickerson *et al.*, 2008). Following identification of an siRNA which was found

to inhibit expression of the mutant keratin 6a allele without affecting wild-type gene expression, initial reports of a human clinical trial have recently been published. Injection of the siRNA in question into the callused foot of a patient has resulted in callus regression. This represents the first use of siRNA in human skin and the results appear promising for the treatment of other dominant-negative skin diseases (Leachman *et al.*, 2010). As mentioned above (section 1.7), such an allele-specific RNAi approach may not be viable for heterogenous diseases in which multiple mutations in the one gene have been identified. However for the most commonly occurring mutations, which may affect a significant percentage of total patients, this mutation-specific strategy may represent a viable option for exploration in the quest for RNAi-based therapeutics for EB.

1.8 Aims

At the outset of the studies undertaken for this Ph.D., whilst the molecular basis underlying the dystrophic form of EB was understood, there was no published research with the specific aim of development of a gene therapy strategy for the dominant form of the disease. It was thought that this form of the genetic skin disease may possibly represent a good candidate for an attempted suppression and replacement therapeutic approach using the powerful RNA interference (RNAi) molecular tool. The focus of the research described in this Ph.D. thesis therefore centres on identification and development of suitable potential RNAi suppressors of the COL7A1 gene, the gene implicated in the disease phenotype. An investigation into a range of potential suppressor types was undertaken, including analysis of the effects of siRNAs, shRNAs and artificial microRNAs on COL7A1 expression. Following identification of effective mutation-independent COL7A1 suppressors, it would then be of interest to investigate development of a replacement gene refractory to the effects of the RNAi whilst still encoding wild-type collagen type VII protein. An additional mutation-dependent RNAi study for DDEB was carried out, with the aim of identifying potent suppressors of certain chosen dominant mutant COL7A1 alleles which do not affect expression of the wild-type allele. The data obtained represent significant steps in the progression towards development of a suitable therapeutic for DDEB.



Copyright © 2002 Pearson Education, Inc., publishing as Benjamin Cummings.

Figure 1.1 Structure of the skin: Diagram representing the layers of skin (epidermis, dermis and hypodermis) and their principle components (Cummings, 2002).

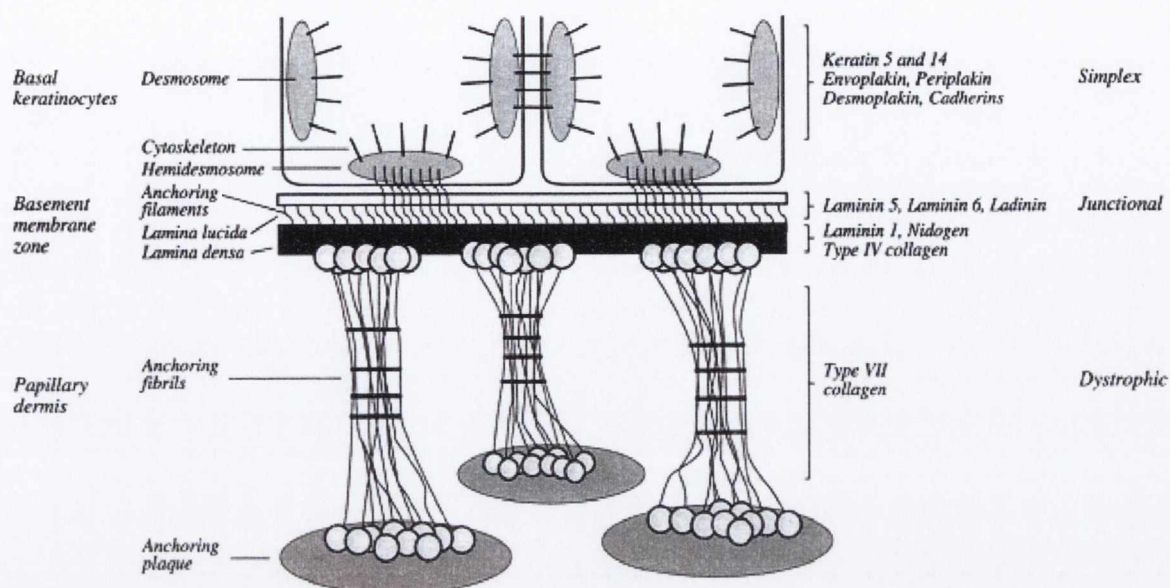


Figure 1.2 Illustration of the complexity of the cutaneous basement membrane zone: The schematic representation depicts the presence of basal keratinocytes overlying the papillary dermis, with the dermal-epidermal basement membrane separating these two compartments. Ultrastructurally recognisable attachment structures within this adhesion zone are indicated on the left, and individual protein components comprising these structures are indicated on the right. Note the level of tissue separation in the Simplex, Junctional and Dystrophic forms of Epidermolysis Bullosa (figure and legend modified from Pulkkinen and Uitto, 1999).



Figure 1.3 Clinical features of various epidermolysis bullosa subtypes: (a) In very mild Epidermolysis bullosa (EB) simplex, blisters occur occasionally after friction, and light scaling is seen after healing of a blister. (b) A 12-year-old patient with Kindler syndrome (a skin fragility syndrome similar to EB) shows significant skin atrophy on the hand. (c) Dominant dystrophic EB (DEB) leads to nail loss, and (d, e) recessive DEB to severe trauma-induced blistering on the foot and the hand. (f) In recessive mutilating DEB, scarring leads to contractures of the fingers and functional deficits (figure and legend modified from Aumailley *et al.*, 2006).



Figure 1.4 Some clinical features of DDEB: (a) Hypertrophic scarring in a patient with generalised DDEB. (b) Atrophic scarring and post-inflammatory hypopigmentation on the extremity of a patient with DDEB. (c) Dystrophy of all twenty nails in a patient with DDEB (figures from Fine, 2010b).

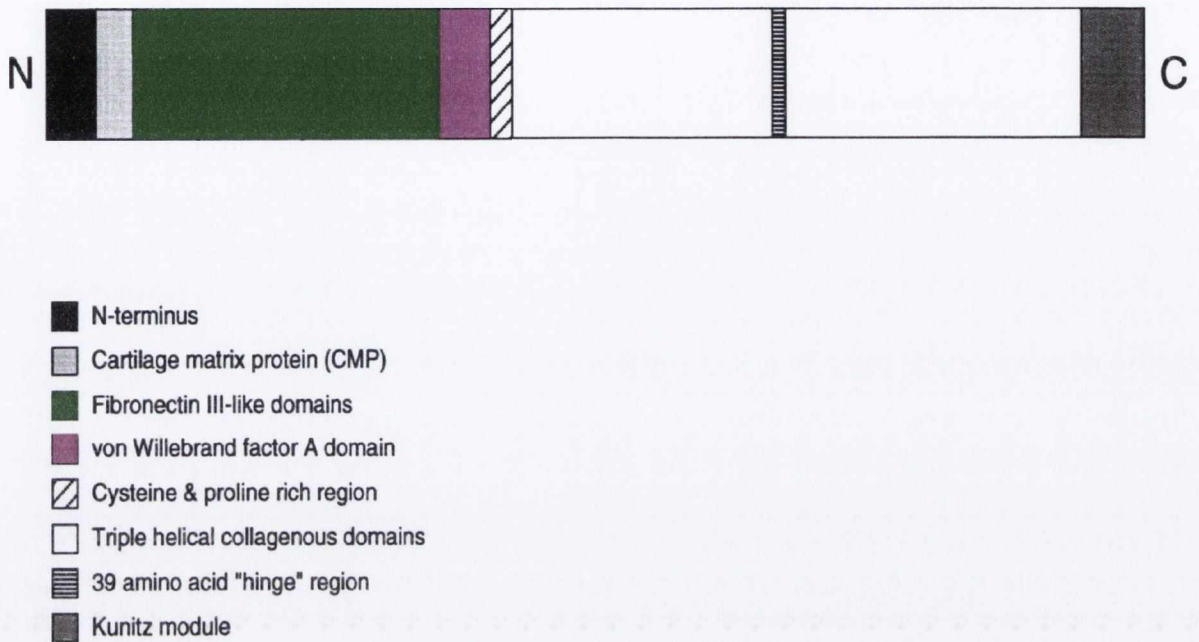


Figure 1.5 Type VII collagen domain organisation: Schematic representation of type VII collagen domain organization. The type VII collagen molecule consists of a central collagenous triple-helical domain which has a 39-amino acid non-collagenous interruption ('hinge' region). The collagenous domain is flanked by amino-terminal and carboxy-terminal globular domains, with submodules with homology to known protein modules, as indicated on the lower left corner (figure and legend modified from Pulkkinen and Uitto, 1999).

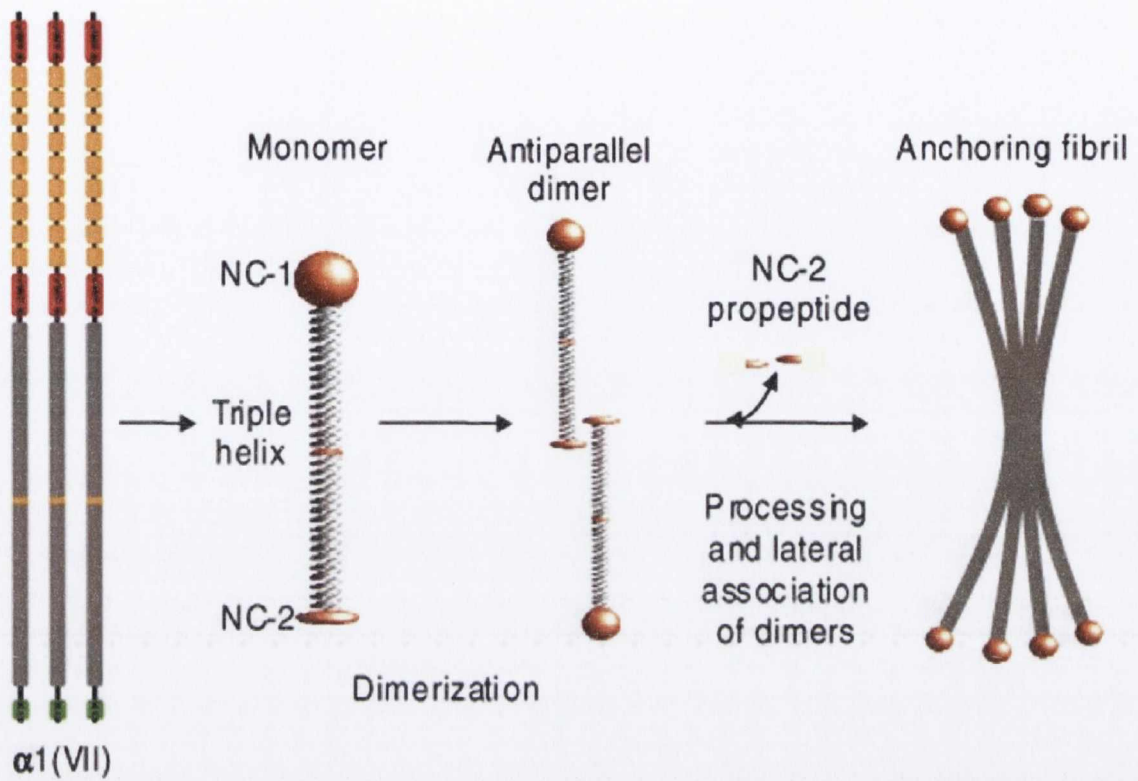


Figure 1.6 Schematic representation of the polymerization of anchoring fibrils:

The type VII procollagen monomer consists of three $\alpha 1(\text{VII})$ polypeptide chains folded into a triple helix. Two monomers form an antiparallel dimer, from which the NC-2 propeptides are removed proteolytically. Finally, the mature dimers laterally aggregate into anchoring fibrils (Bruckner-Tuderman, 2009).

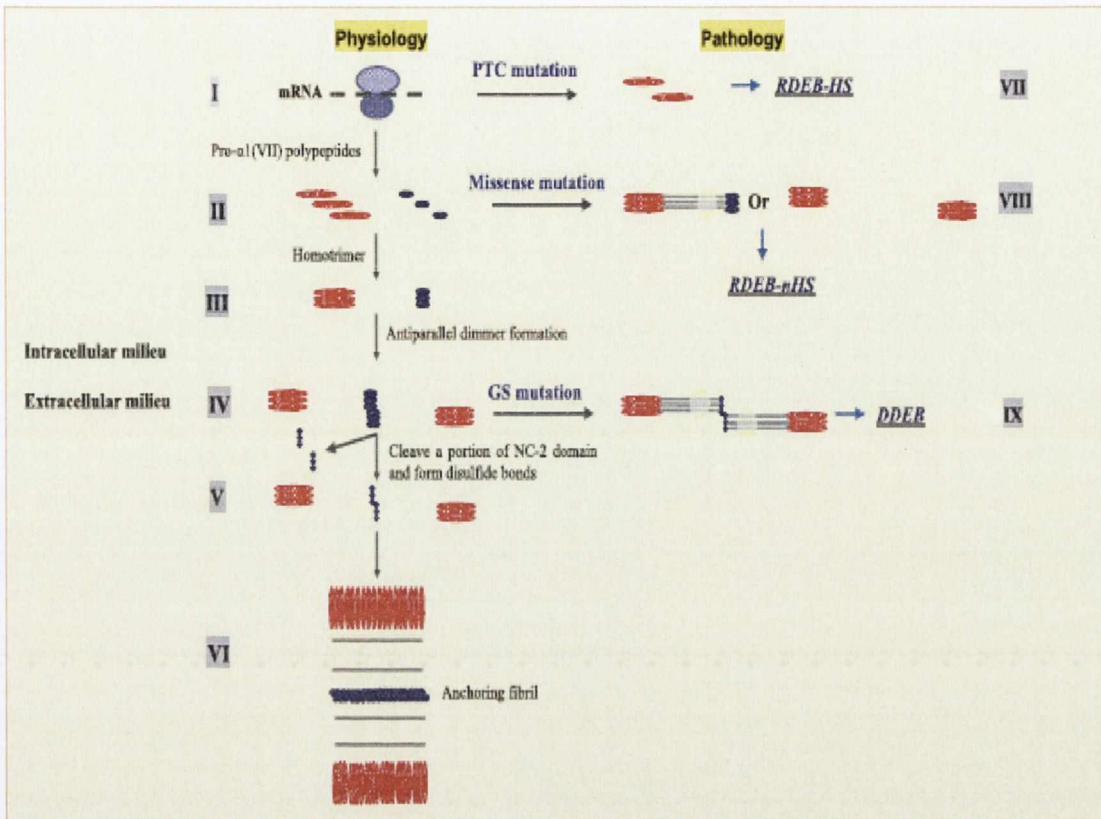


Figure 1.7 Anchoring fibril assembly and the consequences of the major types of mutations on type VII collagen protein synthesis in dystrophic epidermolysis bullosa (DEB): The left side shows the physiology of type VII collagen and the right side shows the pathology. *I:* proα1 (VII) polypeptides are synthesized in ribosomal complex. *II:* Three of these chains assemble into a triple helical type VII collagen molecule – homotrimers. At stages *III* & *IV*, two homotrimers form antiparallel tail-to-tail dimers with a central carboxy-terminal overlap and with the amino-termini outwards, a portion of the NC-2 domain is removed, and the association of the monomers is stabilized by intermolecular disulphide bonds. Stages *V* & *VI*: a large number of dimer molecules assemble into anchoring fibrils and the complete NC-1 domain keeps the adhesive property at both ends. Mature anchoring fibrils are stabilized by transglutaminase cross-links in vivo. Stage *VII*: premature termination codon mutations (PTC) decrease the amount of the mutated transcripts and result in truncated non-functional polypeptides which are unable to assemble into anchoring fibrils, then causing Hallopeau-Siemens recessive dystrophic epidermolysis bullosa (RDEB-HS; also known as RDEB-severe generalised). *VIII*: Missense mutations alter

homotrimer formation and/or subsequent stabilization of the dimer molecules by disulphide bonds result in decreased stability and/or alter function of VII collagen leading to milder types of recessive dystrophic epidermolysis bullosa (RDEB-nHS). *IX*: Glycine substitutions often happen in triple helix region of COL7A1 affecting the correct folding and the secretion of type VII collagen, resulting in dominant dystrophic epidermolysis bullosa (DDEB) (figure and legend modified from Dang and Murrell, 2008).

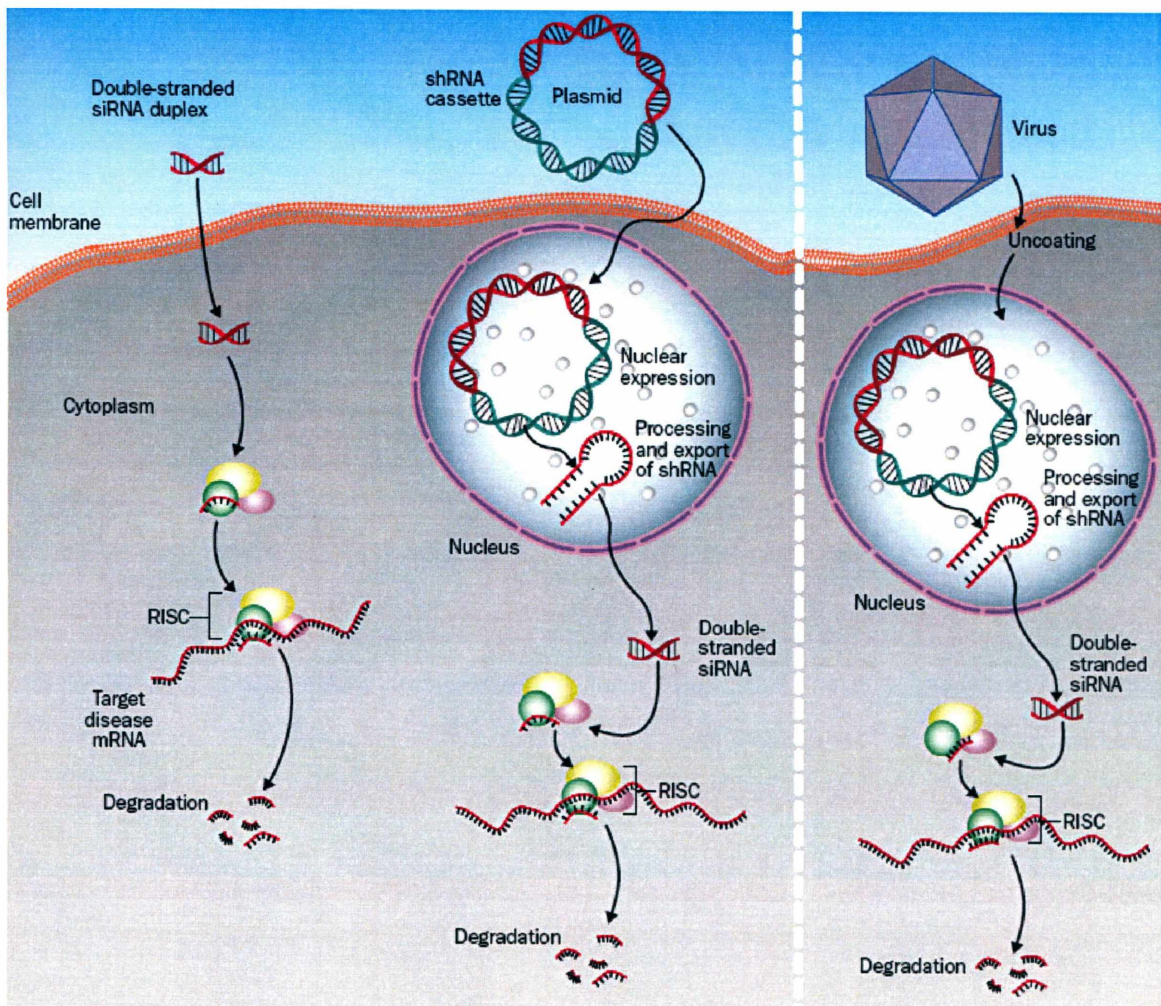


Figure 1.8 RNA interference: Three ways of inhibiting protein expression with RNAi. *Left:* cells are transfected with pre-formed siRNA duplexes, generally in combination with a transfection reagent. The transfected siRNA becomes incorporated into the RISC following cell entry, leading to degradation of target gene expression. *Middle:* transfected plasmid vectors must reach the nucleus for transcription of shRNA. The shRNA is processed and exported to the cytoplasm where it is incorporated into the RISC for directed gene silencing. *Right:* viral vectors encoding siRNA or shRNA bind to the target cell via a receptor, followed by receptor-mediated endocytosis. Uncoating of the viral protein coat (adeno-associated viruses) or fusion of the viral envelope with cellular membranes (lentivirus) releases viral genomes and allows trafficking to the nucleus (figure and legend modified from Davidson and Paulson, 2004).

Chapter 2

Materials and Methods

2.1 General Materials

2.1.1 Materials for molecular biology methods

10X TAE Buffer: 400mM Tris-Acetate and 10mM EDTA (Invitrogen).

Xylene Cyanol: 0.25% Xylene Cyanol (BDH Laboratory Supplies), 30% glycerol (Sigma Aldrich) in dH₂O.

Bromophenol Blue: 0.25% Bromophenol Blue (BDH Laboratory Supplies), 40% (w/v) sucrose (Sigma Aldrich) in dH₂O.

DNA Oligo Annealing Buffer: 100mM KOAc, 30mM HEPES K-OH, 2mM Mg-OAc, pH7.4 (Xeragon).

Crimson Taq 1 polymerase (NEB)

1X Crimson Taq Reaction Buffer: 12.5mM Tricine, 42.5mM KCl, 1.5mM MgCl₂, 6% Dextran, Acid Red, pH8.5 at 25°C (NEB)

dNTPs (Roche Diagnostics)

Luria-Bertani (LB) Medium: 2g Bacto-tryptone (OXOID), 1g Yeast extract (OXOID) and 2g NaCl made up to 200ml with ddH₂O before autoclaving at 121°C for 15 minutes.

Luria-Bertani (LB) Agar: 4g Bacto-tryptone (OXOID), 2g Yeast extract (OXOID), 4g NaCl and 6g Agar made up to 400ml with ddH₂O before autoclaving at 121°C for 15 minutes.

SOC medium: 2% Tryptone (OXOID), 0.5% Yeast extract (OXOID), 10mM NaCl, 2.5mM KCl, 10mM MgCl₂·6H₂O, 20mM glucose.

NZY⁺ medium: 10g NZ amine (casein hydrolysate; Unipath Ltd.), 5g yeast extract (OXOID) and 5g NaCl was dissolved in ddH₂O, the pH was adjusted to 7.5 with NaOH and the medium made up to 1litre with ddH₂O. Following autoclaving at 121°C for 15 minutes, the following filter-sterilised supplements were added: 12.5ml of 1M MgCl₂, 12.5ml of 1M MgSO₄ and 20ml of 20% (w/v) glucose.

Antibiotics

	<u>Stock concentration</u>	<u>Working concentration</u>
Tetracycline (Sigma Aldrich)	10mg/ml in 50% ethanol	25mg/ml
Spectinomycin (Sigma Aldrich)	10mg/ml	50µg/ml
Ampicillin (Sigma Aldrich)	100mg/ml	100µg/ml
Kanamycin (Sigma Aldrich)	10mg/ml	50ug/ml

Spectinomycin, Ampicillin and Kanamycin stock solutions were made in ddH₂O, filter sterilized using a sterile 0.2 µm syringe filter unit (Sartorius), aliquoted and stored frozen at -20°C

0.5M CaCl₂: 36.755g of CaCl₂·2H₂O (Sigma Aldrich) was dissolved in ddH₂O, made to 500ml and autoclaved for 15 minutes at 121°C.

1M Tris-Cl, pH 8.0: 60.55g of Tris base (Sigma Aldrich) was dissolved in ddH₂O. The pH was adjusted to 8.0 by addition of concentrated hydrochloric acid (HCl). The solution was made up to a final volume of 500ml with ddH₂O and autoclaved at 121°C for 15 minutes.

0.5M EDTA, pH 8.0: 93.05g of disodium ethylenediaminetetra-acetate·2H₂O (Sigma Aldrich) was dissolved in ddH₂O. The pH was adjusted to 8.0 with NaOH and the volume was made up to 500ml with ddH₂O, followed by autoclaving at 121°C for 15 minutes.

3M Sodium Acetate, pH 5.2 and pH 7.0: 123.045g of sodium acetate anhydrous (Sigma Aldrich) was dissolved in ddH₂O. The pH was adjusted to either 5.2 with glacial acetic acid or to 7.0 with dilute acetic acid. The volume was made up to 500ml with ddH₂O and the solution was autoclaved at 121°C for 15 minutes.

E. coli cell Lysis Solution I: 50mM Tris pH 8.0, 25% sucrose and 2mM EDTA pH 8.0.

M-STET buffer: 5% triton-x, 50mM EDTA pH 8.0, 5% sucrose and 50mM Tris pH 8.0.

2.1.2 Tissue culture materials

Dulbecco's Modified Eagle Medium (DMEM)+: 500ml Dulbecco's Modified Eagle Medium (Lonza) was supplemented with 50ml foetal calf serum (Invitrogen, Carlsbad, CA, USA), 5ml of 100mM sodium pyruvate (Lonza) and 5ml of 200mM L-glutamine (Lonza).

Fibroblast Growth Medium-2 (FGM-2)+: 500ml FGM-2 medium (Lonza) was supplemented with 0.5ml insulin (Lonza), 0.5ml r-human fibroblast growth factor-B (Lonza), 0.5ml gentamicin-sulfate (Lonza) and 10ml fetal bovine serum (Lonza).

2X Freezing media: 25% DMSO (Sigma Aldrich), 50% DMEM or FGM-2 and 25% fetal calf serum.

4% Paraformaldehyde: 130ml of Phosphate Buffered Saline (PBS) was heated to 65°C on a heated plate. The heat was turned off prior to addition of 8g of powdered paraformaldehyde (Sigma Aldrich) with constant stirring using a magnetic stirrer. NaOH was added dropwise, to aid solubilisation of the paraformaldehyde. Once the solution had reached room temperature, the pH was adjusted to 7.4 using HCl and the final volume was adjusted to 200ml using PBS.

Cell lines: Normal Human Dermal Fibroblast cell line (NHDF) was purchased from Lonza. Human Embryonic Kidney 293 cell line (HEK293) was purchased from ATCC. Human Epidermal Keratinocyte cell line (Human adult low Calcium Temperature keratinocytes; HaCaTs) was a kind gift from Professor Fusenig at the DKFZ (German Cancer Research Center).

2.1.3 RNA materials

Real-time rtPCR DNA primer sequences

COL7A1 Fw: 5'-GAT GAC CCA CGG ACA GAG TT-3'

COL7A1 Rev: 5'-ACT TCC CGT CTG TGA TCA GG-3'

UBC Fw: 5'-ATT TGG GTC GCG GTT CTT G-3'

UBC Rev: 5'-TGC CTT GAC ATT CTC GAT GGT-3'

TBP Fw: 5'-CAC GAA CCA CGG CAC TGA TT-3'

TBP Rev: 5'-TTT TCT TGC TGC CAG TCT GGA-3'

B2M Fw: 5'-GAT GAG TAT GCC TGC CGT GTG-3'

B2M Rev: 5'-CAA TCC AAA TGC GGC ATC T-3'

YWHAZ Fw: 5'-ACT TTT GGT ACA TTG TGG CTT CAA-3'

YWHAZ Rev: 5'-CCG CCA GGA CAA ACC AGT AT-3'

2.1.4 Oligonucleotide synthesis

Oligonucleotides were supplied by Sigma-Aldrich. Primers were resuspended in ddH₂O at a final stock concentration of 100 µM and diluted appropriately for use.

2.1.5 Protein materials

1X TBS: 137mM NaCl, 25mM Tris-HCl pH7.5, and 3mM KCl were dissolved in 350ml ddH₂O and the pH was adjusted to 7.5 with HCl. The final volume was then adjusted to 400ml.

Total protein lysis buffer: 62mM Tris, 2% SDS, 10mM DTT. One protease-inhibitor-cocktail tablet (Complete; Roche Diagnostics, UK) added to 10ml of the solution.

Cytoplasmic Protein Lysis Buffer: 1% triton, 1X TBS, 1mM EDTA. 300µl of a protease inhibitor stock solution (1 Roche Diagnostics protease inhibitor cocktail tablet in 1.5ml ddH₂O) was added to 1.7ml cytoplasmic protein lysis buffer.

Bicinchoninic acid (BCA) assay reagents

Reagent A: 500ml containing sodium carbonate, sodium bicarbonate, bicinchoninic acid and sodium tartrate in 0.1M sodium hydroxide.

Reagent B: 25ml containing 4% cupric sulfate.

Albumin standard ampoules: 2.0mg/ml bovine serum albumin (BSA) in 0.9% saline and 0.05% sodium azide.

1.5M Tris HCl, pH 8.8: 36.33g of tris was dissolved in 150ml ddH₂O. The pH was adjusted to 8.8 with HCl and the volume was made up to 200ml with ddH₂O.

0.5M Tris HCl, pH 6.8: 12.11g of tris was dissolved in 15ml of ddH₂O. The pH was adjusted to 6.8 with HCl and the volume was made up to 200ml with ddH₂O.

10% SDS, pH 7.2: 20g of sodium dodecyl sulphate (SDS, Sigma Aldrich) was dissolved in 150ml ddH₂O by heating. The pH was adjusted to 7.2 with NaOH and/or HCl and the final volume was made up to 200ml with ddH₂O.

6% polyacrylamide gel:

1.5M Tris HCl, pH 8.8	5ml
Acrylamide/Bis 30% Solution (Sigma Aldrich)	4ml
10% SDS (Sigma Aldrich)	200µl
10% Ammonium persulfate (Sigma Aldrich)	200µl
TEMED (N,N,N',N'-tetramethylethylenediamine, Sigma Aldrich)	20µl

4.5% stacking gel:

0.5M Tris HCl, pH 6.8	1.67ml
Acrylamide/Bis 30% Solution (Sigma Aldrich)	1.33ml
10% SDS (Sigma Aldrich)	100µl
10% Ammonium persulfate (Sigma Aldrich)	100µl
TEMED (N,N,N',N'-tetramethylethylenediamine, Sigma Aldrich)	10µl

2X Sample Buffer: 125mM Tris pH 6.8, 2% SDS, 10% glycerol, 0.2% bromophenol blue, 50mM DTT.

1X TBS-tween: 137mM NaCl, 25mM Tris-HCl pH 7.5, 3mM KCl, 0.05% (v/v) Tween-20 (Sigma Aldrich).

1X Electrophoresis running Buffer: 25mM Tris pH 8.9, 192mM Glycine, 0.2% SDS.

1X Western Blot Transfer Buffer: 25mM Tris pH 8.9, 192mM Glycine, 0.05% SDS.

Ponceau S Solution: 0.1% Ponceau S (w/v) in 5% (v/v) acetic acid.

Restore Western Blot Stripping Buffer (Thermo Scientific)

PageBlue[™] Protein Staining Solution (Fermentas Life Sciences)

2.2 General Methods

2.2.1 Cloning

2.2.1.1 Restriction enzyme digests

Restriction enzyme digest reactions were typically carried out in a final volume of 50µl. Between 1-2 Units of the required restriction enzyme was used per µg DNA, with the appropriate buffer at 1X concentration. Reactions were typically incubated in a 37°C water bath for 2 hours unless otherwise stated. Restriction enzymes and buffers were obtained from New England Biolabs (Frankfurt, Germany).

2.2.1.2 DNA purification

In order to purify DNA fragments from enzymatic reactions, purification was carried out using the Qiagen QIAquick Enzyme reaction Purification Kit Protocol. 5 volumes of Buffer PB was added to 1 volume of the enzymatic reaction and placed in a QIAquick spin column. Samples were centrifuged for 1 minute at 13,000rpm (IEC Micromax, Thermo-Electron), washed with 0.75ml Buffer PE and centrifuged for an additional 1 minute. DNA was eluted in 50µl Buffer EB.

2.2.1.3 Ethanol precipitation

Ethanol precipitation of DNA was typically carried out to concentrate DNA fragments that were too small for use of the Qiagen QIAquick PRC Purification Kit (<100bp). In order to precipitate DNA, 3X volume of 100% ethanol and 0.1X volume of sodium acetate (pH 7.0) were added to the DNA sample. Samples were left to stand at room temperature for 30 minutes, followed by centrifugation at 13,000rpm (IEC Micromax, Thermo-Electron) for 20 minutes. Supernatant was removed and samples were left to airdry, followed by resuspension in ddH₂O.

2.2.1.4 Gel electrophoresis

1 or 2% agarose gels were prepared by heating the appropriate amount of agarose in 1X TAE Buffer (Invitrogen): 1% for separation of large and 2% for separation of smaller fragments. 2µl of 10,000X Sybr[®] Safe (Invitrogen) was added to 25µl agarose to allow visualization of DNA under UV light. Prior to loading, 6X loading dye (either xylene cyanol or bromophenol blue) was added to DNA samples.

Bromophenol blue and xylene cyanol migrate in a 1% gel at 400-500bp and 4-5Kb respectively. Samples together with a DNA size marker (100bp or 1Kb; New England Biolabs) were electrophoresed at a constant voltage of 100V. DNA was then visualized under a UV light source and images were captured using a UV lamp (MiniBIS Pro, DNR Bio-Imaging systems Ltd.) with GelCapture software (DNR Bio-Imaging systems Ltd.).

2.2.1.5 Gel extraction

To isolate specific fragments from restriction enzyme fragmented DNA, samples were electrophoresed on a 1% agarose gel, the required fragment was excised using a scalpel blade and the DNA purified from the gel using the Qiagen QIAquick Gel Extraction Kit Protocol according to the manufacturer's instructions.

2.2.1.6 Annealing of oligonucleotides

Oligonucleotides were re-suspended in ddH₂O to a stock concentration of 100pmol/ μ l (100 μ M). 2 μ l of both sense and its corresponding antisense sequence were annealed in 46 μ l annealing buffer (Xeragon; see section 2.1.1). The annealing reaction involved heating the samples to 90°C for 4 minutes and step cooling to 70°C for 10 minutes, 37°C for 20 minutes followed by incubation at 10°C until samples were ready for use.

2.2.1.7 Ligations

Ligation reactions were typically set up in total volumes of 10 μ l unless otherwise stated. 1 μ l of T4 DNA ligase enzyme (Roche Diagnostics) was used along with ligation buffer (Roche Diagnostics) at a concentration of 1X. Typically 100ng of DNA was used in ligations and various ratios of vector DNA: insert DNA (typically 1:10 to 1:20) were utilised. Appropriate control reactions eg restriction enzyme digested-vector with and without ligase were included. Reactions were incubated overnight at 18°C.

2.2.1.8 Making *E. coli* cells competent for transformation

10ml LB broth, containing 25 μ g/ml tetracycline was inoculated with a single colony of *E. coli* XL1-blue cells and incubated overnight at 37°C in an orbital shaker at 220rpm (Innova 4300; New Brunswick Scientific). 0.25ml of this overnight culture

was subcultured into 25ml fresh LB containing 25µg/ml and grown at 37°C in the orbital shaker until the OD₆₀₀ reached 0.3-0.7. The cells were centrifuged for 5 minutes at 3,000rpm (Centrifuge 5810; Eppendorf) and the supernatant was discarded. The pellet was resuspended in 25ml cold 50mM CaCl₂, left on ice for approximately 15 minutes, centrifuged for 5 minutes at 3,000rpm, and the supernatant was discarded. The pellet was resuspended in 5ml cold CaCl₂. The competent cells were used for plasmid DNA transformation experiments within 24 hours.

2.2.1.9 Plasmid DNA Transformations

Approximately 100ng of DNA in 5µl of ligation solution was added to 200µl of competent cells in 1.5ml eppendorf tubes. Samples were placed on ice for approximately 30 minutes, heat shocked in a 42°C water bath for 5 minutes and returned to ice. 1ml of LB broth was added to each tube, followed by gentle mixing. Samples were incubated at 37°C for 1hr. 200µl of each sample was plated onto LB agar plates containing the required selective antibiotic. Plates were incubated overnight at 37°C.

2.2.1.10 In-Fusion™ Advantage PCR Cloning

The manufacturer's instructions were followed for cloning using the In-Fusion™ Advantage PCR Cloning kit (Clontech). DNA inserts were synthesised and cloned in plasmid vectors by GeneArt AG (Germany). Polymerase chain reactions were carried out on a T3000 Thermocycler (Bio-Sciences) in a total volume of 20µl in Crimson Taq Reaction Buffer (New England BioLabs), containing the GeneArt vectors as template, 10pmol each of forward and reverse primers, 0.2mM each of dATP, dTTP, dCTP and dGTP (Roche Diagnostics), and 1 unit of Crimson Taq DNA polymerase (New England BioLabs). The standard PCR cycle run was as follows; an initial denaturation step at 95°C for 5 minutes followed by 35 cycles of 95°C for 30 seconds for denaturation, 55°C for 30 seconds for annealing and 72°C for 1 minute for DNA extension, ending with a final incubation at 72°C for 10minutes to ensure maximum yield of product. Following verification of the fragment size which was amplified during the PCR by loading 5µl of the PCR product on a 2% agarose gel, 5µl of the PCR product was added to 2µl of kit-supplied Cloning Enhancer. This mixture was then incubated at 37°C for 15 minutes, followed by 80°C for 15 minutes on a T3000 Thermocycler. Ligations were then carried out in a final volume of 10µl in In-Fusion

Reaction Buffer containing 1µl In-Fusion Enzyme, 100ng vector plasmid DNA, 1µl Cloning Enhancer-treated PCR insert. Samples were incubated for 15 minutes at 37°C, followed by 15 minutes at 50°C, then placed on ice. Reaction volumes were then brought to 50µl with TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0), and mixed well by pipetting.

Transformations were subsequently carried out, using the kit-supplied Fusion-Blue™ competent cells and SOC medium. Cells were thawed on ice and 50µl was transferred to falcon tubes (BD Falcon). 2.5µl of ligation reactions was added to cells and tubes were placed on ice for 30 minutes. Cells were heat-shocked for 45 seconds at 42°C and then placed on ice for 1-2 minutes. 447.5µl pre-warmed SOC medium was added to tubes and cells were incubated with shaking at 220rpm for 1 hour at 37°C. Transformed cells were plated neat (100µl per plate and 50µl per plate) and 10X concentrated (100 µl per plate). Plates were incubated overnight at 37°C.

2.2.1.11 Plasmid DNA Mini-preparation

Plasmid mini-preps were carried out on selected colonies. Firstly, each colony was inoculated in 10ml of selective LB broth and incubated overnight at 37°C in an orbital shaker at 220rpm (Innova 4300; New Brunswick Scientific). The following day, cultures were centrifuged at 3,000rpm (Centrifuge 5810; Eppendorf) for 10mins and the supernatant discarded. Pellets were resuspended in 60µl *E. coli* cell lysis solution 1, and transferred to 1.5ml eppendorf tubes. 20µl lysozyme solution (40mg/ml in 250mM Tris, pH 8, freshly prepared) was added to each sample, which were then mixed by pipetting and left at room temperature for 10 minutes. 550µl M-STET buffer was added and samples were gently mixed and left at room temperature for a further 10 minutes. Samples were then boiled for 90 seconds and centrifuged at 13,000rpm (IEC Micromax, Thermo-Electron) for 15 minutes. The bacterial cell debris pellet was then removed with a toothpick and discarded. 1µl RNase-A (10mg/ml; Sigma Aldrich) was added to each tube in order to degrade any RNA and samples were incubated at 37°C for 15 minutes. An equal volume of phenol was added to samples, which were then mixed by vortexing. Samples were then centrifuged at 13,000rpm (IEC Micromax, Thermo-Electron) for 5 minutes. 400µl of the top aqueous layer was transferred to 1.5ml eppendorf tubes and 0.6X volume of

isopropanol was added, followed by vortexing of each tube for a few seconds. Samples were left at room temperature for 15mins and then spun at 13,000rpm for a further 15 minutes. Isopropanol was poured off, pellets were washed with 500µl 96% ethanol and left to stand for 10 minutes. Ethanol was discarded, samples were left to air dry, and plasmid DNA was resuspended in 100µl dH₂O.

Alternatively, after overnight cultures were pelleted, minipreps were carried out using the Qiagen QiaPrep® Spin Miniprep kit according to the manufacturer's instructions.

2.2.1.12 Plasmid DNA maxi-preparation

5ml of selective LB was inoculated with a bacterial colony containing a plasmid of interest and incubated for 8 hours in a 37°C orbital shaker at 220rpm (Innova 4300; New Brunswick Scientific). 200µl of the starter culture was transferred to 200ml of fresh LB with the appropriate antibiotics and grown overnight in a 37°C orbital shaker. Cells were centrifuged at 3,000rpm (Centrifuge 5810; Eppendorf) for 15 minutes and plasmid DNA was prepared using the Qiagen HiSpeed® Plasmid Maxi Kit according to the manufacturer's instructions.

2.2.1.13 Capillary Sequencing

Sequencing was performed using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Each reaction mixture contained 250ng plasmid DNA, 2pmol primer, 0.4µl 5X BigDye® Sequencing Buffer, 2µl BigDye® Terminator mix in a final aqueous volume of 10µl. Sequencing reactions were performed on a PTC-225 Peltier Thermal Cycler (MJ Research, Inc.) for 2 hrs. The steps were as follows: an initial denaturation step at 96°C for 2 minutes, followed by 25 cycles of 96°C for 10 seconds for denaturation, 50°C for 5 seconds for annealing and 60°C for 4 minutes for DNA extension, followed by a 15 minute 4°C incubation period. The DNA was subsequently precipitated by addition of one-tenth volume 3M NaAc (pH5.2) and 3X volumes 96% ethanol. Samples were incubated for 30 minutes at room temperature and subsequently centrifuged at 13,200rpm (IEC Micromax, Thermo-Electron) for 20 minutes. The pelleted DNA was allowed to airdry. Samples were resuspended in EDTA + formamide (95% deionised formamide, 10 mM EDTA, pH 8.0), heated to 95°C for 5 minutes and immediately cooled to 4°C. Sequencing was performed using

an ABI 373A DNA sequencing machine according to the manufacturer's instructions and data was processed using ABI data collection and analysis software. Sequencing was undertaken with the assistance of Alex McKee of the Smurfit Institute of Genetics, Trinity College Dublin.

2.2.1.14 Commercial sequencing

Plasmid DNA and primers were sent to either GATC Biotech or Source BioScience LifeScience for sequencing, using an ABI 3730 xl sequencing machine. Data was processed using GATCViewer™ software.

2.2.1.15 Nucleic acid concentration determination

DNA and RNA concentrations were measured using a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific), by measuring the absorbance at a wavelength of 260nm (A260). All samples were blanked with 2µl of ddH₂O or the relevant buffer in which samples were resuspended. 2µl of each sample was then loaded on the surface of the plate reader and the absorbance readout in ng/ml was recorded. Nucleic acid concentration was calculated based on the Beer Lambert Law, which relates the amount of light absorbed at 260nm to the concentration of dsDNA (50µg/OD260) and RNA (40µg/OD260), based on the different extinction coefficients of the nucleic acids. The ratio between the readings A260/A280 was used to determine the purity of the DNA and RNA samples. A ratio of 1.8 and 1.9-2.1 was achieved for DNA and RNA samples respectively, an indication of the purity of the samples.

2.2.2 Tissue culture techniques

2.2.2.1 Sub-culturing of cells

Spontaneously immortalised human keratinocytes (HaCaT cells), normal primary human skin fibroblasts (NHDFs) and HEK293 cells were maintained in a sterile humidified environment at 37°C and 5% CO₂. All cell culture experiments were performed in a laminar flow hood (ESI Securigarde 1200; Thermo Electron) under sterile conditions unless otherwise stated. Cells were grown in flasks of either 75cm² or 175cm² unless otherwise stated.

HaCaT cells were grown in supplemented Dulbecco's Modified Eagle media (DMEM+). Adherent HaCaT cells were passaged by removal of DMEM and incubation in 0.02% EDTA (Sigma) at 37°C for 20 minutes. EDTA was removed and cells were incubated in Trypsin/EDTA (0.25mg/ml, Lonza) for 5 minutes in order to dissociate cells from the plate. Trypsin was neutralised by addition of DMEM+ and cells were seeded onto fresh plates.

Normal Human Fibroblasts were incubated in FGM-2 medium. For passaging of fibroblast cells, medium was removed and cells were washed in HEPES Buffered Saline Solution (30mM, Lonza). Cells were incubated in trypsin/EDTA (0.25mg/ml, Lonza) at 37°C for approximately 2 minutes or until cells were dissociated from the plate, followed by neutralisation of trypsin by addition of an equal volume of Trypsin Neutralizing Solution (Lonza). FGM-2 was added and cells were seeded onto fresh plates.

HEK293 cells were incubated in DMEM+. Adherent cells were passaged by removal of media and washing of cells in 1X phosphate buffered saline (PBS, Lonza). PBS was removed and cells were incubated in trypsin/EDTA at 37°C for approximately 2 minutes or until cells were dissociated from the plate, followed by neutralisation by addition of DMEM+. Cells were then seeded onto fresh plates.

2.2.2.2 Haemocytometer cell counting

Cells were dissociated from plates as previously described. Following neutralisation, cells were centrifuged at 1,000rpm (IECCentra-3C centrifuge, DJB Labcare) for 5 minutes and resuspended in the appropriate medium. 100µl of cells were taken and diluted 1 in 2 in Trypan blue solution (Sigma Aldrich). A haemocytometer slide (Hawksley) was prepared by moistening the slide and affixing a coverslip over the cell counting area, creating a firm seal. Cells were gently pipetted under the edge of the coverslip, resulting in their movement throughout the area beneath the coverslip by capillary action. The slide was then viewed under a microscope using 10X magnification (Olympus CK30 microscope). The number of cells in all 8 quadrants was counted and the average calculated. The haemocytometer is designed so that the number of cells in one quadrant is equivalent to the number of cells $\times 10^4$ /ml.

Therefore the number of cells per ml was calculated as the average cell count of all quadrants x dilution factor x 10^4 .

2.2.2.3 Freezing of cell stocks

Cells were dissociated and pelleted as previously described, followed by resuspension in DMEM+ or FGM-2. Cells were then counted using a haemocytometer and 1×10^6 cells were resuspended in medium and an equal amount of 2X freezing medium in labelled CryoTubes (Nunc). Tubes were placed in a Cryo 1°C Freezing Container (Nalgene), which allows for the gradual cooling of cells at a rate of -1°C per minute, and subsequently transferred to a -80°C freezer.

2.2.2.4 Nucleofection of cells

HaCaT and fibroblast cells were transfected using the Lonza Cell Line Nucleofector™ Kit V for HaCaT cells and the Lonza Human Dermal Fibroblast Nucleofector Kit respectively. The required number of cells was prepared using a haemocytometer slide and centrifuged at 1,000rpm for 5mins in DMEM or FGM-2. The cells were resuspended in the appropriate Nucleofector™ Solution to a concentration of $1-5 \times 10^6$ cells/100 μl for HaCaTs or $4-5 \times 10^5$ cells/100 μl for fibroblasts. For each sample, 100 μl cell suspension was mixed with the required amount of plasmid DNA (1-5 μg) and transferred to an Amaxa certified cuvette which was then placed in the nucleofection machine (Nucleofector® II; Amaxa) and the U-20 or U-023 programs were selected for HaCaTs and fibroblasts respectively. Following transfection 500 μl medium was added and the cells transferred to a well of a 6 well plate containing 2ml medium, using a plastic pipette. The plate was rocked gently to mix cells and incubated at 37°C .

2.2.2.5 Generation of stable cell lines

Prior to generation of stable cell lines, a kill curve was set up, in order to determine the minimum concentration of antibiotic required to kill untreated cells within 1-3 weeks. A range of antibiotic concentrations was tested, selective medium was replenished every 3-4 days and the percentage of surviving cells analysed daily. When the appropriate antibiotic concentration was determined, cells were transfected by nucleofection as previously described (section 2.2.2.4). 8 wells of 6 well plates (each well being approximately 2cm^2) were transfected per construct. 24 hours post

transfection, cells were washed with PBS and DMEM+ was replaced with fresh medium. 48 hours post transfection cells were passaged from eight wells to a tissue culture flask of 175cm². 3 hours later, media was removed and replaced with DMEM+ containing the appropriate pre-determined concentration of antibiotic. Cells were fed with selective medium every 3-4 days until all non-transfected cells were killed and an antibiotic-resistant stable cell line population was formed.

2.2.2.6 siRNA transfection

HaCaT cells and normal human fibroblasts were transfected with siRNAs, using HiPerFect Transfection Reagent (Qiagen). For HaCaTs, 1×10^5 cells per well of a 24-well plate were seeded in 0.5ml of DMEM+ and incubated under normal conditions. For each well, the appropriate amount of siRNA was diluted in 100µl serum-free DMEM+ and HiPerFect Transfection Reagent was added. Samples were then mixed by vortexing and incubated for 10 minutes at room temperature to allow formation of transfection complexes. The complexes were added drop-wise onto the cells and the plate was swirled gently to ensure uniform distribution of the transfection complexes. Cells were then incubated at 37°C.

For normal human fibroblast cells, 6×10^4 cells per well of a 24-well plate were seeded in 0.5ml FGM-2 and incubated under normal conditions. For each well the appropriate amount of siRNA was diluted in 100µl culture medium and mixed by vortexing. HiPerFect Transfection Reagent was then added to the diluted siRNA and samples were mixed by vortexing for 10 seconds. Samples were incubated at room temperature for 10 minutes and added drop wise onto the cells. Cells were then incubated under their normal growth conditions.

2.2.2.7 Co-transfections

HEK293 cells were co-transfected with both a DNA plasmid and siRNA, using Attractene Transfection Reagent (Qiagen). 5×10^5 cells per well of a 6-well plate were seeded in 2ml DMEM+ and incubated under normal conditions. For each well, the appropriate amount of plasmid DNA was diluted in serum-free DMEM+ in 1.5ml eppendorf tubes, to a total volume of 100µl and the appropriate volume of siRNA was added. Samples were mixed by vortexing. Attractene Transfection Reagent was then added to each tube and following vortexing, samples were incubated at room

temperature for 10-15 minutes. Following incubation samples were added drop wise onto cells, which were then incubated under their normal growth conditions.

2.2.2.8 Fluorescent Activated Cell Sorting (FACS)

Following transfection of cells and incubation for the appropriate amount of time, HaCaT cells in each well of a 6 well plate were treated with EDTA (0.02%) and trypsin, pelleted, and resuspended in 800µl sterile PBS in a 1.5ml eppendorf. Samples were then placed on ice and transported to the Flow Cytometry Core Facility of University College Dublin, where FACS was carried out with the assistance of Dr. Alfonso Blanco. Immediately prior to sorting, the cell suspension was filtered using a 50µm Filcon Filter (DakoCytomation) and transferred to 5ml polystyrene round-bottom falcon tubes (BD Falcon). Cells were sorted based on green fluorescent protein (GFP) fluorescence, using a BD FACSAria Cell Sorter (BD Biosciences). The sorter is designed with 3 lasers: 488 nm (blue), 635 nm (red) and 405 nm (violet), enabling detection of two scattered signals. As cells flow past a focused laser beam the beam is deflected enabling the detection of two scatter signals, side scatter and forward scatter. Forward scatter is a measure of cell size and side scatter is a measure of cell granularity or general shape. An FITC optical filter was used to detect GFP fluorescence and a sorting flow rate was chosen to maximise the purity of the final samples. A non-fluorescent sample was used as a control to enable identification of scatter profiles and negative populations for the fluorochrome. Approximately 100,000-200,000 GFP positive cells were collected per sample. Sorted cells were collected in a 1.5ml eppendorf tube and 1ml of RNAlprotect Cell Reagent (Qiagen) was added immediately. This stabilizes RNA in sorted or cultured cells for up to 1 day at 30°C, 7 days at 15-25°C and 4 weeks at 2-8°C.

2.2.2.9 Flow cytometry

Following co-transfections, HEK293 cells were incubated for the required amount of time (24 or 48 hours) and subsequently collected and pelleted as previously described. Cells were resuspended in 400µl PBS/4% paraformaldehyde (3:1) and transferred to 5ml polystyrene round-bottom falcon tubes, placed on ice and transported to the Flow Cytometry Core Facility of University College Dublin. Flow cytometry was carried out using a CyAn ADP machine, which is designed with 3 lasers: 488nm (blue), 635nm (red) and 405nm (violet). The CyAn ADP measures forward light scatter and

side light scatter and a FITC optical filter was used to detect fluorescence. A non-fluorescent sample was used as a control, to enable the determination of the auto-fluorescence threshold. For each sample, mean GFP fluorescence of 15,000 cells was measured. Analysis of samples was carried out using Summit 4.3 software.

2.2.3 RNA analysis

2.2.3.1 RNA isolations

Mammalian cells suspended in RNAprotect cell reagent were centrifuged at 8,000rpm (IEC Micromax, Thermo-Electron) for 5 minutes, supernatant was discarded and cells were resuspended in Buffer RLT from the Qiagen RNeasy[®] Plus mini kit. The RNA isolation was then performed according to the Qiagen RNeasy[®] Plus mini kit protocol. In order to homogenise the lysate QIAshredder[™] spin columns (Qiagen) were utilised.

Alternatively, RNA from cells growing in a monolayer was isolated by adding Buffer RLT directly to the cells. Cells were then scraped with a pipette tip in order to dislodge them from the surface of the plate, and transferred to a 1.5ml eppendorf. The RNA isolation protocol was then followed using the Qiagen RNeasy[®] Plus mini kit according to the manufacturer's instructions.

2.2.3.2 Real time rtPCR

Quantitative real time rtPCR was carried out on RNA isolated from mammalian cells using the QuantiTect[®] SYBR[®] Green RT-PCR kit (Qiagen). This allows for measurement of changes in expression of a target gene relative to an untreated control. SYBR Green binds double-stranded DNA, and upon excitation emits light. Thus as a PCR product accumulates, fluorescence increases, forming the basis of quantification. The reaction was carried out using the Applied Biosystems 7300 Real Time PCR System, in ABI PRISM[™] 96-Well Optical Reaction Plates (Applied Biosystems). Primers for the target gene of interest were resuspended to 100pmole/ μ l in nuclease-free H₂O. A range of internal control genes were included, against which RNA levels of the transcript of interest could be standardised. Each well contained a total volume of 20 μ l, including 4.5 μ l RNA and 15.5 μ l Master Mix. The Master Mix contained 10picomoles of forward and reverse primers, 10 μ l SYBR Green (Qiagen),

0.147 μ l RT-enzyme (Qiagen), and 4.353 μ l nuclease-free H₂O per reaction. The master mix was added to each well first using reverse pipetting technique, followed by addition of the RNA. The 96-well plate was covered with MicroAmp™ Optical Adhesive Film (Applied Biosystems) and centrifuged (Centrifuge 5810; Eppendorf) for a few seconds to ensure thorough mixing in each well. Analysis was performed using the Applied Biosystems 7300 Real Time PCR System employing the following reaction protocol:

Reverse transcription: 50°C for 20 minutes;

Activation of DNA polymerase: 95°C for 15 minutes;

cDNA amplification: 37 cycles of 95°C for 15 seconds and 60°C for 1 minute;

Dissociation: 95°C for 15 seconds and 60°C for 15 seconds.

The data collected during the real-time rtPCR was presented as a graph of fluorescence versus cycle number, on the Applied Biosystems 7300 Sequence Detection Software, Version 1.4. Melting curves were analysed, which represent the rate of change of fluorescence with time versus temperature, peaking at the melting temperature (T_m). Following a run, a melt curve is automatically calculated whereby the temperature is raised by a fraction of a degree, causing the two strands of DNA to separate resulting in a subsequent decrease in fluorescence which is measured. As all PCR products for a particular primer pair should have the same melting temperature and since SYBR green does not distinguish between different DNA samples, it is important that all samples have a similar T_m . If peaks are not similar, contamination of samples, mispriming or presence of primer-dimer artifacts may have occurred. For analysis of quantification data, only fluorescence values from the log-linear phase of the PCR amplification were considered. The cycle number at which the increase in fluorescence, and thus cDNA accumulation, is exponential is known as the threshold and the point at which the fluorescence crosses this fixed threshold is known as the C_T . The more abundant the target transcript is in a given sample, the lower the C_T number.

2.2.3.3 Real time rtPCR data analysis

Real time rtPCR data was analysed using the comparative C_T method, also known as the $2^{-\Delta\Delta C_T}$ method, as described by Livak and Schmittgen (Livak and Schmittgen,

2001). The relative quantification method of data analysis which was employed relates the PCR signal of the target transcript in a treatment group to that of a control sample. The threshold cycle (C_T) indicates the fractional cycle number at which the amount of amplified target reaches a fixed threshold. Therefore, the C_T value of any single well is a measure of the stage at which the number of amplicons in that well has accumulated to be at a statistically significant point above the baseline. To obtain a measure of the fold change, the data was transformed to a linear scale, as C_T data is expressed on a $\log_2$ scale. The amount of target, normalised to an endogenous reference and relative to a calibrator, is given by $2^{-\Delta\Delta C_T}$. The $\Delta\Delta C_T$ value is obtained from the following formula:

$$(C_{T\text{gene of interest}} - C_{T\text{internal control gene}})_{\text{Treated sample}} - (C_{T\text{gene of interest}} - C_{T\text{internal control gene}})_{\text{Untreated control sample}}$$

The comparative C_T method relies on the assumption that the amplification efficiencies of the target and control transcripts are approximately equal, and that the efficiency of the PCR is close to one.

2.2.3.4 Statistical analysis

Statistical analysis was performed using Minitab 15 Statistical Software. The mean, standard deviation and standard error were calculated for each set of samples. A two-sample t-test was used in order to determine significance. Values of $p < 0.05$ were taken as statistically significant.

2.2.4 Protein analysis

2.2.4.1 Isolation of protein from cell media

Cells at 80-100% confluency were incubated in serum-free DMEM for 24 hours. Medium containing secreted collagen type VII protein was collected and a protease-inhibitor-cocktail (Complete; Roche, UK), 2mM N-ethylmaleimide (Sigma Aldrich) and 10mM EDTA (Sigma Aldrich) were added. Protein was concentrated using Amicon® Ultra centrifugal filter devices (Millipore) with a molecular weight cut-off of 100,000. Filtered protein was collected and stored at -80°C .

2.2.4.2 Isolation of total cell protein

Medium was removed from cells in wells of a 6 well-plate and cells were washed with ice-cold PBS twice. 300µl of Total Protein Lysis Buffer was added to each well and pipetted up and down until detachment of the cells from the plate surface had occurred. Cell lysates were transferred to eppendorf tubes and left on ice for 20 minutes. Samples were then centrifuged at 14,000rpm (IEC Micromax, Thermo-Electron) for 20 minutes and the supernatant collected into a fresh eppendorf.

2.2.4.3 Isolation of cytoplasmic protein

Medium was removed from cells in wells of a 6-well plate and cells were washed twice with ice-cold PBS. 300µl of Cytoplasmic Protein Lysis Buffer was added to each well and pipetted up and down until detachment of the cells from the plate surface had occurred. Cell lysates were then transferred to eppendorf tubes containing small magnetic stirring bars (5mm x 2mm; Fisher Scientific). Samples were left stirring on ice for 1 hour. Samples were then transferred to a fresh eppendorf and spun at 12,000rpm at 4°C (Microfuge 22R Centrifuge, Beckman Coulter) for 30 minutes. The supernatant was collected into a fresh eppendorf.

2.2.4.4 Protein concentration measurement

The concentration of protein isolated from cells was measured using a BCA Protein Assay Kit (Pierce). Concentrations were determined with reference to a series of dilutions of known concentration of bovine serum albumin (BSA) protein. The BSA standard curve was set up on a 96-well plate with the following concentrations: 0mg/ml, 0.25mg/ml, 0.5mg/ml, 0.75mg/ml, 1mg/ml, 1.5mg/ml and 2mg/ml. Isolated protein samples of unknown concentration were diluted one in five and 10µl of each sample as well as each protein standard and a blank sample were loaded. 200µl of BCA working reagent (50:1 dilution of reagent A: reagent B) was then added to each well. All samples were loaded in triplicate. Plates were incubated at 37°C for 1 hour and subsequently measured at 595nm on a Rosys 2010 Anthos Plate reader and data was retrieved using Anthos 2010 platereader MS-DOS according to the manufacturer's instructions. A standard curve was prepared by plotting the average blank-corrected measurement for each BSA standard versus its concentration in mg/ml. This standard curve was used to establish the concentration of each unknown protein sample using the equation $y=mx+c$, where y is the OD, x is the concentration

of protein and m and c are the standard curve determined slope and axis intercept respectively.

2.2.4.5 TCA precipitation

An equal volume of 20% Trichloroacetic acid (TCA) was added to collected, filtered media protein. Samples were incubated on ice for 30 minutes, and subsequently centrifuged at 13,000rpm for 15 minutes (IEC Micromax, Thermo-Electron). 200 μ l ice-cold acetone was added and samples were then centrifuged at 13,000rpm for 5 minutes at 4°C (Microfuge 22R Centrifuge, Beckman Coulter). Supernatant was then removed and the pellet was left to airdry, followed by freezing. For western blot analysis, pellets were resuspended in one volume of Sample buffer.

2.2.4.6 Western blotting

Two glass plates and 1.5mm spacers were assembled in a dedicated gel electrophoresis apparatus according to the manufacturer's instructions (BioRad). A 6% separating polyacrylamide gel solution was prepared and carefully poured between the glass plates to a point approximately 4cm from the top of the inner plate. The gel was overlaid with ddH₂O and left to polymerise at room temperature for approximately 20 minutes. Upon polymerisation of the gel, the top layer of water was removed and a 4.5% stacking gel was cast with loading combs inserted in the resolving gel. The gel was left to polymerise at room temperature for approximately 20 minutes.

Prior to loading onto the polyacrylamide gel, protein samples were diluted in 2X Sample buffer and heated at 90°C for 5 minutes. Up to 25 μ l of protein samples were loaded onto the gel per well, as well as 20 μ l of a molecular weight marker (Novex® Sharp Pre-Stained Protein Standards; Invitrogen). The gel was run at 120V in 1X Running buffer until the desired separation was obtained.

Protein samples were electrotransferred from the polyacrylamide gel onto a nitrocellulose membrane (Whatman) in 1X transfer buffer in a dedicated apparatus. Samples were transferred at 400mA for 2 hours at 4°C. Membranes were stained in Ponceau S solution (Sigma) for approximately 30 seconds in order to assess the

quality of transfer. Membranes were then washed 2x5 minutes followed by 2x10 minutes in 1X TBS and destained for 5 minutes in ddH₂O.

Non-specific antibody binding to the membrane was prevented by incubation in 5% dried milk powder (marvel) diluted in TBS-tween, at room temperature with agitation for one hour. Blocked membranes were transferred to 5% marvel containing the appropriate primary antibody at a 1 in 500 dilution, and incubated overnight at 4°C with gentle agitation. Membranes were washed 2x5 minutes followed by 2x10 minutes in TBS-tween, and incubated in a solution of 5% marvel containing the appropriate peroxidase-conjugated anti-rabbit secondary antibody (Sigma) at a dilution of 1 in 2,000 for 1 hour at room temperature with gentle agitation. Membranes were subsequently washed again for 2x5 minutes in TBS-tween followed by 2x10 minutes in TBS-tween. Membranes were then incubated in ECL Plus Mix solution (Amersham Biosciences) for 5 minutes, sealed in a plastic sheet and exposed to chemiluminescence sensitive film (Fuji Medical X-Ray Film, Fujifilm) in a dedicated cassette from a few seconds to 30 minutes at room temperature.

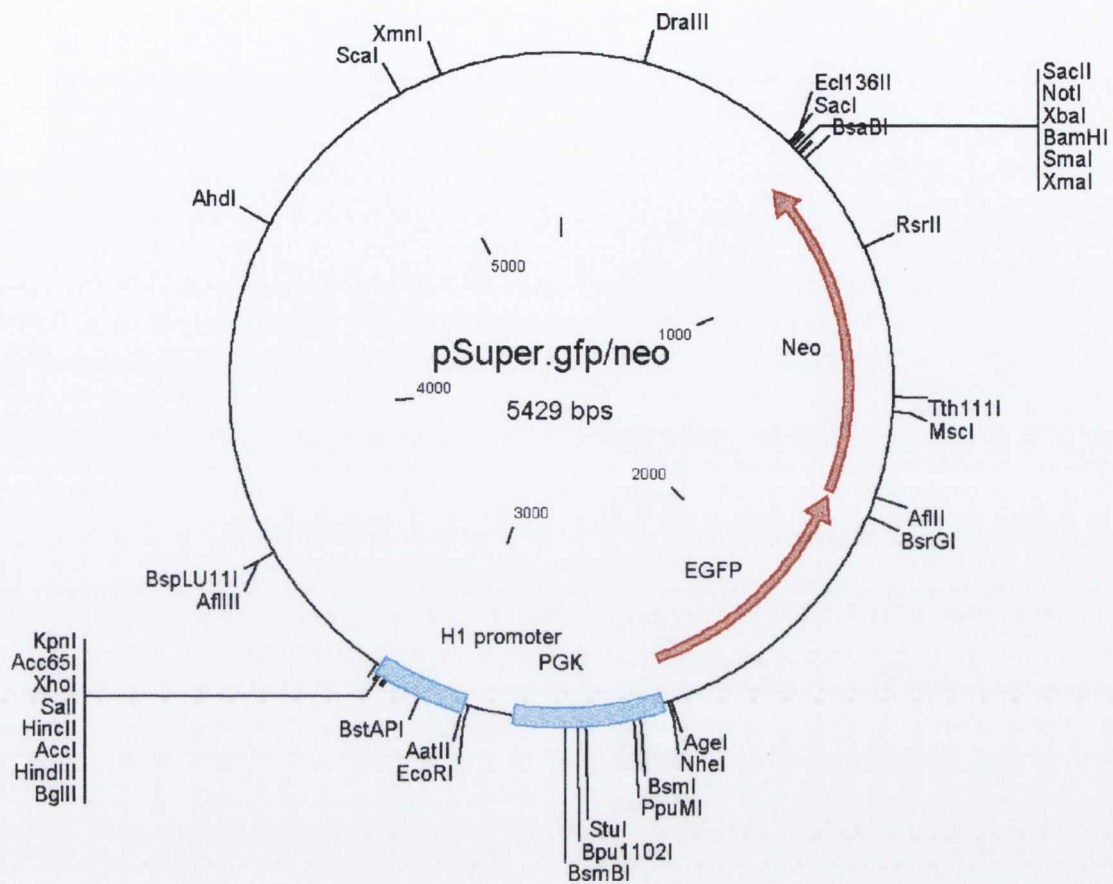


Figure 2.1: Vector map of pSuper.gfp/neo (OligoEngine). The key restriction sites are illustrated.

Artificial microRNA-mediated COL7A1 suppression

3.1 Introduction

Dystrophic Epidermolysis Bullosa (DEB) is caused by mutations of the COL7A1 gene which encodes type VII collagen protein (section 1.3). A number of these COL7A1 mutations, which may be genetically inherited or arise *de novo* in any affected individual, act in a dominant negative manner, giving rise to the blistering phenotype which is typical of all forms of EB. More than 100 missense mutations that result in a glycine substitution have been described in the collagenous domain of the collagen type VII protein, approximately half of which have a dominant-negative effect causing DDEB (Christiano *et al.*, 1995b; Varki *et al.*, 2007). Collagen type VII protein is a homotrimer of three $\alpha 1(\text{VII})$ chains, which assemble as dimers into anchoring fibrils at the basement membrane zone. These anchoring fibrils provide structural integrity to the epithelial tissue of the human body (Chung and Uitto, 2010; section 1.4). In the case of COL7A1-linked DDEB, in cells which are expressing both a wild-type and mutant allele and thus are producing equal amounts of both wild-type and mutant collagen VII protein, seven out of every eight collagen VII trimeric molecule produced will include at least one mutant pro- $\alpha 1$ chain, while just one-eighth will consist solely of normal polypeptides. This leads to disruption of type VII collagen anchoring fibril organisation, impairing the structural integrity of the skin and thus causing the disease phenotype (Dang and Murrell, 2008). Due to the difficulties associated with treating DEB, one of the most effective approaches to relieving the pathology associated with the disorder may be to prevent incorporation of the mutant $\alpha 1(\text{VII})$ chains into the triple helix.

In the past number of decades, a variety of molecular tools have become applicable in the sequence-specific targeting of mRNA for degradation. These include antisense oligonucleotides (Hnik *et al.*, 2009; Visser *et al.*, 2010), ribozymes (Mastroiannopoulos *et al.*, 2010; Mulhbachter *et al.*, 2010) and RNAi technology (section 1.6). Whilst some success has been found with antisense oligonucleotides (ASOs) they are limited by their short half-life and toxicity associated with chemical modifications which have been employed to extend the ASO half life. Similarly, advances in the use of ribozymes as therapeutics have been limited by the fact that *in vitro* suppression efficiencies have failed to be replicated *in vivo* (Millington-Ward *et al.*, 2005). Advances in the field of RNAi in recent years have proved extremely

promising and so in this study, it was decided to explore the use of RNAi technology for down-regulation of COL7A1 expression: no such therapeutic approach directed towards the treatment of DDEB has yet been reported. RNAi is an endogenous sequence-specific mRNA silencing pathway, induced by the presence of double-stranded RNA (dsRNA). Whilst long dsRNA was found to induce an RNAi response in a range of non-mammalian cell types including plants, nematodes, protozoan and invertebrate species, it was originally thought that this pathway could not be induced in mammalian cell lines (Caplen *et al.*, 2000). This was owing to the triggering of an interferon response and ultimate death of the cells following introduction of long dsRNA (Stark *et al.*, 1998). However it was subsequently discovered that synthesised small interfering RNAs (siRNAs) of approximately 21 nucleotides in length could induce RNAi-mediated gene silencing in mammalian cells without an immunological response (Elbashir *et al.*, 2001b). Thus this observation led to the possibility of the application of RNAi technology to many areas of the field of human genetics.

Apart from the need for the identification of potent suppressors for the therapeutic approach under evaluation in this study, DDEB presents additional challenges when considering the design of effective therapeutics. DDEB is a genetically heterogeneous disease; of the approximate 300 DEB-causing mutations, thus far more than 60 of these mutations of the COL7A1 gene have been implicated in the DDEB disease phenotype (Varki *et al.*, 2007; Van den Akker *et al.*, 2011). Therefore, the design of RNAi suppressor molecules to target each individual dominant COL7A1 mutation may not be a viable option, due in particular to the financial costs and technical difficulties involved. One method of overcoming this which is addressed in the current chapter of this study, involves a mutation-independent approach. As discussed in section 1.6, the suppression studies described in the current chapter of this Ph.D. thesis are focused upon targeting the coding sequence of the COL7A1 gene such that expression of both wild-type and mutant alleles should be reduced. It is of note however, that the level of suppression required of the dominant negative COL7A1 allele in order to reverse the DDEB phenotype is currently not yet fully known.

At the outset of this study, an increasing amount of research in the RNAi field was being published which focused on the use of artificial microRNAs (miRNAs) as

suppressor agents (Zeng *et al.*, 2005b; Du *et al.*, 2006; Scherr and Eder, 2007). These artificial miRNAs comprise miRNA-based expression cassettes expressing RNAs which resemble the stem-loop structure of endogenous precursor miRNAs (pre-miRNAs). This allows these artificial miRNAs to enter the miRNA pathway and be processed into mature siRNAs. The use of artificial miRNAs in suppression studies is thought to have a number of benefits over the use of previously widely used shRNA molecules, as described in section 1.6.4, namely due to the ability to more tightly regulate their expression as well as the fact that they are thought to represent a less toxic RNAi delivery methodology (McBride *et al.*, 2008; Boudreau *et al.*, 2009). For these reasons, it was decided to commence this COL7A1 RNAi suppression study with the testing of COL7A1-targeting artificial miRNA constructs. Extensive studies expressing endogenous microRNAs 30 and 155 from vectors had previously demonstrated that artificial miRNAs could be expressed from synthetic stem-loop precursors based on the miRNA-30 and miRNA-155 pre-miR sequences (Zeng *et al.*, 2002; Stegmeier *et al.*, 2005; Chung *et al.*, 2006). One commercially available vector which has emerged from these studies is the pcDNA 6.2-GW/EmGFP-miR construct, which has been utilised in this study for the expression of the designed COL7A1-targeting artificial miRNA.

pcDNA 6.2-GW/ EmGFP-miR is a vector system based on the murine pri-miR-155 sequence (Lagos-Quintana *et al.*, 2002), that mediates the expression of synthetic miRNAs. Previous characterisation of the miR-155 gene had identified the promoter regions and the minimal sequence length of the pre-miR-155 sequence required for efficient production of mature miR-155, including the stem-loop pre-miR and adjacent flanking regions (Tam, 2001). The pcDNA 6.2-GW/ EmGFP-miR vector system retains the sequence requirements of miR-155, whilst incorporating slight modifications. Optimisation of the stem-loop structure from the 5 nucleotide/3 nucleotide internal loop found in the native miR-155 loop to a 2 nucleotide internal loop was carried out in order to improve knockdown rates. Thus this vector allows for insertion of engineered pre-miRNA duplexes within the pre-miR-155 sequence, replacing both miR-155 and its complementary strand, whilst maintaining the miR-155 loop and its flanking regions (figure 3.1). The expression of the artificial miRNA vector is under the control of the cytomegalovirus (CMV) Pol II promoter, which allows for a high level of expression in various mammalian cells (Boshart *et al.*,

1985). Additionally, the EmGFP (Emerald Green Fluorescent Protein) is incorporated within the pre-miRNA expression cassette in order to allow efficient tracking of miRNA expression.

The aim of the study described in the current chapter of this thesis is therefore the generation and testing of artificial microRNA constructs designed to target the human COL7A1 gene and analysis of their ability to achieve this objective. Experiments were undertaken to analyse COL7A1 gene expression in human keratinocyte (HaCaT) cells, which endogenously express COL7A1, at both the RNA and protein levels following treatment with artificial miRNA constructs. It was also of interest to compare the potency of the COL7A1-targeting artificial miRNAs which were found to be the most effective, with their equivalent shRNA molecules.

3.2 Materials and methods

3.2.1 Materials and methods for artificial microRNAs

3.2.1.1 Artificial microRNA DNA oligo design

Artificial microRNA sequences were designed using the Invitrogen online BLOCK-iT™ Pol II miR RNAi Designer program, which identifies optimal target sites within a gene for an artificial microRNA to induce gene knockdown. Following input of the RefSeq ID of the gene of interest (COL7A1 in this case; accession number NM_000094), the Designer automatically generated high probability 60bp DNA duplexes based on the endogenous mouse miRNA-155 sequence, that once processed, would have 100% homology to the human COL7A1 gene. These double-stranded oligos encoding the engineered pre-miRNA and based on miRNA-155 comprise a 4 nucleotide 5' overhang (TGCT) complementary to the vector; a 5'G + 21 nucleotide antisense sequence (mature miRNA) derived from the target gene; a short spacer of 19 nucleotides to form the terminal loop; a sense target sequence with 2 nucleotides removed to create an internal loop and a 4 nucleotide 5' overhang (CAGG) complementary to the vector (figure 3.2). In total, 13 COL7A1-targeting artificial microRNA sequences were chosen (see figure 3.3 for sequences).

3.2.1.2 Artificial miRNA vector construction

The pcDNA 6.2-GW/EmGFP-miR vector (Invitrogen) mediates the expression of shRNAs which have characteristics of miRNAs and thus are processed by the miRNA processing pathway. The vector allows for the insertion of an engineered pre-miRNA ds-oligonucleotide (section 3.2.1.1) between the 5' and 3' flanking regions of the miRNA-155 sequence (figure 3.4). The incorporation of the miR-155 flanking sequences facilitates the processing of the inserted pre-miRNA sequence by the Drosha enzyme (see section 1.6.2). Circularised pcDNA 6.2-GW/EmGFP-miR was linearised using the *BsaI* restriction enzyme and 1X NEB Buffer 3. Digestion was carried out at 50°C for 2 hours followed by 65°C for 20 minutes in order to heat inactivate the restriction enzyme. *BsaI* cleaved the vector at two sites (5'..GGTCTC(N)₁'..3'), resulting in removal of a 26bp stuffer fragment (Appendix I) and linearisation of the vector. This creates an insertion site for the pre-miRNA cassette which has a 5'TGCTG and 3'GTCC overhang. The digested vector was

separated from the 26bp stuffer insert by gel electrophoresis (1%), the linearised vector was excised from the gel and the DNA extracted using a QIAquick column (see section 2.2.1.5).

The top and bottom oligo DNA sequences for each artificial miRNA expression construct were resuspended in nuclease-free H₂O to a concentration of 100µM and annealed according to section 2.2.1.6, resulting in a 50µM stock solution. This stock solution was then further diluted to 10nM for ligation into the linearised pcDNA 6.2-GW/EmGFP-miR vector. Ligations were carried out in a final volume of 20µl, containing 2µl linearised vector, 4µl ds oligo (10nM), 2µl 10X ligation buffer, 1µl T4 DNA ligase and 11µl ddH₂O; and incubated overnight at 18°C.

Competent XL1-blue *E.coli* bacteria were prepared (see section 2.2.1.8) and transformed using 5µl of ligated insert and vector (see section 2.2.1.9). 200µl of each transformation was plated onto LB agar plates containing spectinomycin (50µg/ml) and incubated at 37°C overnight. Transformants were examined and a number of colonies were chosen for mini plasmid DNA preparation (see section 2.2.1.11). The sequence of each plasmid was verified using the recommended forward and reverse primers (see figure 3.5 for primer sequences) to determine the sequence and orientation of the engineered pre-miRNA cassette, as per section 2.2.1.13.

Validated positive clones were identified and transformations were again carried out using 5µl of a 1 in 50 dilution of the appropriate mini preparation (section 2.2.1.9). The Qiagen High Speed Plasmid Maxi Kit was used for large-scale plasmid purification, as outlined in section 2.2.1.12. The concentration and purity of the DNA samples was determined using the NanoDropTM 1000 Spectrophotometer (section 2.2.1.15) and the constructs were again sequenced as outlined above, to guarantee that the expected DNA construct was isolated. In total, 13 of these COL7A1-targeting artificial microRNA constructs were generated (mi1-mi13), as well as a non-targeting negative control vector which contains a pre-miRNA structure which is predicted to not target any known vertebrate gene (figure 3.3).

3.2.1.3 Nucleofection of HaCaT cells with artificial miRNA vectors

Frozen stocks of HaCaT cells were thawed and grown in supplemented DMEM (see section 2.2.2.1). Cells were transfected with artificial miRNA constructs by the process of nucleofection, as described in section 2.2.2.4. The pcDNA 6.2-GW/EmGFP-miR vector containing a pre-miRNA sequence which is predicted not to target any known vertebrate gene was employed as a negative control and transfected in parallel with the COL7A1-targeting constructs. The required number of cells was collected and the pellet was resuspended in Nucleofector™ Solution to a concentration of 1×10^6 cells/100µl. For each sample, 100µl cell suspension was mixed with 5µg of the DNA of interest and transferred to an Amaxa certified cuvette whereby the process of nucleofection was subsequently carried out (section 2.2.2.4). After transfer of each nucleofected cell sample to a well of a 6-well plate, cells were incubated under normal growth conditions for the appropriate amount of time.

Cells were then collected and FAC sorting was carried out, whereby only fluorescent cells and thus cells expressing the RNAi constructs were isolated (section 2.2.2.8). Originally, four of these artificial micro-RNA vectors (mi1-mi4) were tested along with the negative control and the transfection of each construct was undertaken in duplicate in 6-well plates. FAC sorting was carried out 24 hours post-nucleofection, with the experiment being undertaken three times. Additionally, a time-point assay was carried out, in which what was found to be the optimal suppressor and the non-targeting construct were transfected in sextuplicate and RNA was directly isolated from duplicate transfected sample wells after 24, 48 and 72 hours. The optimal suppressor was also transfected in duplicate along with the negative control construct and FAC sorting was carried out following incubation of cells for 72 hours. Following this set of experiments, another nine COL7A1-targeting constructs were cloned and transfected, along with the negative control. Transfection of each construct was undertaken in duplicate in 6-well plates and this was carried out either two or three times for each construct. FAC sorting was carried out 24 hours post transfection.

3.2.1.4 Analysis of RNA from transfected cells

RNA from transfected cells was extracted either by direct treatment of cells in 6-well plates with Buffer RLT using the Qiagen RNeasy® Plus mini kit (see section 2.2.3.1) or in the case of FAC sorted cells, by pelleting of cells suspended in RNAlater Cell

Reagent and resuspension in Buffer RLT using the Qiagen RNeasy[®] Plus mini kit (see section 2.2.3.1). The concentration and purity of the RNA samples was determined using the NanoDrop[™] 1000 Spectrophotometer (section 2.2.1.15). Real-time rtPCR was used to analyse expression levels of COL7A1 mRNA according to section 2.2.3.2 using COL7A1 primers previously available in the laboratory (section 2.1.3). Expression was normalised using 4 endogenous internal control genes (section 2.1.3): Ubiquitin C (UBC), TATA box binding protein (TBP), Beta-2-microglobulin (B2M) and tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ). The real-time rtPCR steps which were carried out are outlined in section 2.2.3.2 and the Cycle Threshold (CT) values were used for RNA quantification as outlined in section 2.2.3.3.

3.2.1.5 Statistical Analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

3.2.1.6 Generation of stable cell lines

Blasticidin-resistant HaCaT stable cell lines expressing the most potent artificial miRNA plasmid and the negative control miRNA plasmid were generated as described in section 2.2.2.5, made possible by the presence of a blasticidin-resistance gene on the pcDNA 6.2-GW/EmGFP-miR construct. A kill-curve, in which a range of antibiotic concentrations (from 1 µg/ml to 10 µg/ml) were tested, demonstrated that the minimum concentration of blasticidin antibiotic required to kill untreated HaCaT cells within 1-3 weeks was 5 µg/ml. Following nucleofection of cells with the two separate constructs, the selection process was carried out as detailed in section 2.2.2.5, until all cells conferred blasticidin resistance.

Cells were treated for media protein isolation as described in section 2.2.4.1 and following removal and collection of media, cells were pelleted and counted as previously described (section 2.2.2.2) and resuspended in RNeasy Protect Cell Reagent (2×10^6 cells/ml) for RNA extraction and quantification by real-time rtPCR (sections 2.2.3.1 and 2.2.3.2). Expression was normalised using the endogenous internal control genes (section 2.1.3) Ubiquitin C (UBC), TATA box binding protein (TBP) and

tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide (YWHAZ).

3.2.1.7 Protein analysis

Protein analysis of artificial microRNA-treated cells was carried out by western blotting; two experimentation models were employed. Firstly, as described above, stable cell lines were established and media was isolated from these cells expressing the most potent artificial COL7A1-targeting microRNA and cells expressing the negative control artificial microRNA, for western blot analysis. Additionally, a large-scale transfection experiment was undertaken, in which 25 nucleofections of the most potent artificial microRNA and 25 nucleofections of the non-targeting control were carried out (see section 2.2.2.4 and section 3.2.1.3). Instead of transferral of nucleofected samples to 6-well plates, each set of 25 nucleofections were transferred to two T175 tissue culture plates and incubated under normal growth conditions.

Following collection of media from both sets of cells, measurement of protein concentration using a BCA Protein Assay Kit (section 2.2.4.4) was carried out. Western blotting was then undertaken in order to measure any differences in collagen type VII protein levels between cells treated with COL7A1-targeting and those treated with non-targeting artificial microRNA expression constructs as described in section 2.2.4.6. A 1 in 500 dilution of anti-collagen type VII rabbit polyclonal antibody (Calbiochem) was used to detect collagen type VII protein. An exposure time of approximately 5 seconds at room temperature was generally adequate for collagen type VII protein to be detected on chemiluminescence sensitive film.

3.2.2 Materials and methods for shRNAs

3.2.2.1 shRNA design

shRNAs were designed using the pSuper RNAi SystemTM (OligoEngine) based on the shRNA design algorithms as described by Brummelkamp *et al.*, 2002a. These double-stranded oligos comprise a 15 nucleotide 5' sequence including a *Bam*HI restriction site, such that upon *Bam*HI digestion of the annealed oligo, a 5' *Bgl*II end will result; the 21 nucleotide target sequence (sense); a 9 nucleotide (TTCAAGAGA) hairpin sequence, the 21 nucleotide target sequence (antisense) and an 18 nucleotide 3'

sequence including a *HindIII* restriction site such that upon *HindIII* digestion of the annealed oligo, a 3' *HindIII* end will result (see figure 3.14 for sequences).

3.2.2.2 shRNA vector construction

Circular pSuper.gfp/neo vector (OligoEngine) was linearised by sequential digests with *HindIII* and *BglIII* restriction enzymes. Following digestion with *HindIII* using NEB Buffer 2 (section 2.2.1.1), the vector was then purified (section 2.2.1.2) and subsequently digested with the *BglIII* enzyme, using 1X NEB Buffer 3 as per section 2.2.1.1. The digested vector was excised from an agarose gel (1%) and the DNA extracted using a QIAquick column (section 2.2.1.5). In order to remove the phosphate groups from the ends of the linearised vector, DNA was then treated with Shrimp Alkaline Phosphatase (SAP). The 30µl of linearised vector, along with 5µl SAP enzyme (Affymetrix, Inc.), 5µl 10X SAP buffer (Affymetrix, Inc.) and 10µl ddH₂O, were incubated at 37°C for 2 hours, followed by heating to 65°C for 20 minutes in order to inactivate the enzyme. The linearised and SAP-treated vector was then purified (section 2.2.1.2).

The top and bottom oligo DNA sequences for each shRNA were resuspended in nuclease-free H₂O to a concentration of 100µM and annealed according to section 2.2.1.6. Annealed oligos were then digested with *HindIII* and NEB Buffer 2, as per section 2.2.1.1 and subsequently concentrated by carrying out an ethanol precipitation and resuspending in 42.5µl ddH₂O (section 2.2.1.3). *BamHI* digestion of the oligos was then carried out using NEB Buffer 3 (section 2.2.1.1). 1X Bovine Serum Albumin (BSA; NEB) was also included in the reaction, which prevents adhesion of the restriction enzyme to reaction tubes and pipette surfaces. Following digestion, DNA was concentrated by ethanol precipitation (section 2.2.1.3) and resuspended in 50µl ddH₂O.

Ligations were carried out in a final volume of 10µl containing either 1 or 3µl digested ds oligo, 1µl linearised pSuper vector, 1µl 10X ligation buffer, 1µl T4 DNA ligase and either 6 or 4µl ddH₂O; and incubated overnight at 18°C. Competent XL1-blue *E.coli* bacteria were prepared (see section 2.2.1.8) and transformed using 5µl of ligated insert and vector (see section 2.2.1.9). 200µl of each transformation was plated

onto LB agar plates containing ampicillin (100µg/ml) and incubated at 37°C overnight. Transformants were examined and a number of colonies were chosen for mini plasmid DNA preparation (see section 2.2.1.11). The sequence of each plasmid was verified using the recommended forward and reverse primers (see figure 3.5) to determine the sequence and orientation of the shRNA inserts, as per section 2.2.1.14.

Validated positive clones were identified and transformations were again carried out using 5µl of a 1 in 50 dilution of the appropriate mini preparation (section 2.2.1.9). The Qiagen High Speed Plasmid Maxi Kit was used for large-scale plasmid purification, as outlined in section 2.2.1.12. The concentration and purity of the DNA samples was determined using the NanoDrop™ 1000 Spectrophotometer (section 2.2.1.15) and the constructs were again sequenced as outlined above, to guarantee that the expected DNA construct was isolated. Two of these shRNA constructs were generated, based on the same COL7A1 target sequences as the two optimal artificial microRNA vectors, as well as a non-targeting negative control shRNA vector which contains a sequence which is predicted to not target any known vertebrate gene.

3.2.2.3 Nucleofection of HaCaT cells with shRNA vectors

Nucleofection of HaCaT cells with shRNAs was carried out as for artificial microRNA constructs (section 3.2.1.3). The pSuper.gfp/neo vector containing an shRNA sequence which is predicted not to target any known vertebrate gene was employed as a negative control and transfected in parallel with the COL7A1-targeting constructs. Transfection of each construct was carried out in duplicate. RNA was extracted directly from 6-well plates 24 hours after transfection and RNA was analysed as per section 3.2.1.4. Expression was normalised using 2 endogenous internal control genes: Ubiquitin C (UBC) and TATA box binding protein (TBP).

3.2.2.4 Statistical Analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

3.3 Results

3.3.1 Artificial microRNA vector construction

The pcDNA 6.2-GW/EmGFP-miR vector was used to generate the artificial microRNA constructs. The terminal regions of the arms of endogenous miRNA-155 are contained in the linearised vector, which provide overhangs for insertion of artificial miRNA sequences. Web designed insert structures designed to target the COL7A1 gene are outlined in figure 3.2 and their sequences are displayed in figure 3.3. The Invitrogen online BLOCK-iT™ Pol II miR RNAi Designer which was used incorporates their own design rules into a proprietary algorithm which claims to design pre-microRNA sequences which have optimal silencing potential for the target gene of interest and which are compatible for use upon cloning into the pcDNA 6.2-GW/EmGFP-miR vector. The programme claims that more than 70% of designed miRNAs reduce target gene expression by at least 70%. Originally, 4 sequences predicted by the Invitrogen RNAi Designer programme to suppress expression of the COL7A1 transcript were ordered and cloned into the miRNA expression vector (miRNA1-4) as well as a negative control insert in which the pre-miRNA sequence is not predicted to target any known mammalian gene (NT miRNA). Constructs were designed and cloned as outlined in figure 3.4. Double-stranded oligos were annealed and inserted into the linearised vector using T4 DNA ligase. Positive clones were sequenced (see figure 3.5 for primer sequences) in order to ensure the correct sequence and orientation and DNA preparations of each construct were made and used in further analysis. Following testing of these 4 constructs, another 9 COL7A1-targeting sequences were chosen and constructs were generated and tested (miRNA5-13). These 13 constructs are designed to target a broad range of areas across the open reading frame (ORF) of the COL7A1 transcript.

3.3.2 Analysis of COL7A1 mRNA suppression by miRNAs 1-13

Originally, miRNAs 1-4 were cloned and tested *in vitro* in human keratinocyte (HaCaT) cells along with the negative control plasmid NT miRNA. Cells were transfected with 5µg DNA by the process of nucleofection and the transfection of each construct was undertaken in duplicate in 6-well plates. HaCaTs were collected 24 hours post-nucleofection and FAC sorting for emGFP expression was carried out, with the experiment being repeated three times. FAC sorting of treated cells showed

that the process of nucleofection of HaCaT cells with artificial miRNA constructs results in a relatively high transfection efficiency, with approximately 70-80% of cells from each well emitting fluorescence (see figure 3.6). Total RNA was extracted from cells which were found to express GFP during the cell sorting process using the RNeasy® Plus mini kit and real-time rtPCR was used to analyse expression levels of COL7A1 mRNA. Expression was normalised using expression levels of 4 endogenous control genes: UBC, TBP, B2M and YWHAZ. Data was analysed by the $2^{-\Delta\Delta C_T}$ method, in order to compare differences in levels of COL7A1 mRNA expression between the non-targeting negative control-treated cells and the COL7A1-targeting artificial microRNA-treated cells. Of the 4 original constructs tested, the most potent suppressor of COL7A1 mRNA was miRNA 4, treated cells with which were found to result in an approximate mean 47% decrease of COL7A1 transcript levels compared to cells which were treated with NT miRNA (figure 3.7). A two-sample t-test, with which p values of less than 0.05 were taken to be statistically significant, found this COL7A1 mRNA suppression to be statistically significant ($p < 0.0005$). Statistically significant COL7A1 suppression was also achieved with miRNAs 2 and 3 ($p = 0.039$ and $p = 0.034$ respectively), although the levels of COL7A1 mRNA suppression were much lower than those found with miRNA 4, with mean COL7A1 suppression of approximately 16% and 25% respectively. miRNA 1 on the other hand, failed to significantly down regulate the COL7A1 transcript ($p = 0.140$) (figure 3.7).

In order to examine whether or not extended periods of cell incubation following nucleofection and before isolation of RNA results in a higher knockdown rate of COL7A1 mRNA, a time-point assay was carried out. Six wells in total of the most potent suppressor miRNA 4, along with NT miRNA, were nucleofected into HaCaT cells. RNA from duplicate samples of each construct was collected directly from 6-well plates without FAC sorting, at 24, 48 and 72 hours and analysed by real-time rtPCR. While mean suppression of COL7A1 mRNA was found to increase slightly over time (approximate mean suppression of 34%, 42% and 44% at 24, 48 and 72 hours respectively), the suppression seen at 72 hours was not significantly different from that seen at 24 hours ($p = 0.401$) (figure 3.8). As these samples represented a mixed cell population of both cells expressing and those not expressing the artificial microRNA constructs due to the fact that FAC sorting was not carried

out, it was of interest to further investigate the effects of the 72 hour incubation timepoint on COL7A1 expression. Both miRNA 4 and NT miRNA were transfected by nucleofection in duplicate as previously described and incubated for 72 hours. FAC sorting was then carried out and RNA from cells expressing the miRNA constructs was isolated and analysed by real time rtPCR. Statistically significant ($p < 0.0005$) mean suppression values of approximately 42% were seen in cells treated with miRNA 4 after 72 hours, compared to those treated with miRNA NT (figure 3.9). Thus overall, incubation of cells treated with the COL7A1-targeting miRNA for increased time periods did not seem to have a significant impact on the expression levels of the cell's COL7A1 mRNA.

Following analysis of the effects of four artificial miRNA vectors which were designed to target and downregulate COL7A1 gene expression, it was notable that levels of COL7A1 suppression of greater than 50% failed to be achieved. This was despite Invitrogen's claims that more than 70% of their web-designed artificial miRNAs reduce target gene expression by at least 70%. For this reason, it was subsequently decided to clone and test nine more of these vectors, in an attempt to identify a vector with increased potency compared to those 4 already tested, and which would match the suppression levels claimed to be achievable by Invitrogen. Following generation of these 9 constructs (mi5-13), their effects on COL7A1 expression were analysed in the same manner as for the previous 4 vectors. Transfection of each construct was undertaken in duplicate in 6-well plates and this was repeated either two or three times for each construct. FAC sorting was carried out 24 hours post transfection and following RNA isolation, COL7A1 mRNA levels of treated cells were evaluated by real-time rtPCR compared to those of NT miRNA-treated cells. Significant COL7A1 mRNA suppression was achieved by all but one of these constructs (miRNA 13; $p = 0.702$). However, of the 8 remaining constructs, the highest suppression achieved was with miRNA 6, which reduced COL7A1 mRNA by approximately 37% ($p < 0.0005$) (Figure 3.10). Thus overall, the greatest mean COL7A1 suppression yielded by any of the artificial miRNA constructs was in the order of 50% with miRNA 4.

3.3.3 Analysis of collagen type VII protein suppression by miRNA 4

In order to measure the effects of the most potent tested suppressor, miRNA 4, on collagen type VII protein levels in treated HaCaT cells, it was firstly of interest to generate a reliable and effective method of measurement of collagen type VII protein. Following generation of a western blot protocol, a range of various parameters were altered and optimised in an effort to maximise its efficiency. This included testing and comparing total protein from HaCaT cells (see section 2.2.4.2), cytoplasmic protein (see section 2.2.4.3) and media secreted protein (see section 2.2.4.1), the testing of multiple anti-collagen VII-specific antibodies, and investigations into the amount of HaCaT cells required for the ability to adequately detect collagen VII protein. Following optimisation of the protocol, as well as personal communication with Professor Edel O'Toole (Professor of Molecular Dermatology at Barts and The London School of Medicine and Dentistry, London), such an assay was successfully established from media-secreted protein using an anti-collagen type VII rabbit polyclonal antibody (Calbiochem). An adequate signal was only detected when a large number of cells were utilized (approximately 1×10^7 cells). Detection of collagen VII from this cell line was not possible from a smaller quantity of cells, even with precipitation and thus concentration of the media protein using Trichloroacetic acid (see section 2.2.4.5). Hence a method of measurement of collagen type VII protein was ascertained, allowing for quantification of data for all future studies, a valuable capability for the validation of any suppression study (see section 2.2.4.6).

In an attempt to further characterise the levels of suppression achieved with the miRNA 4 construct, two methods of analysis of the effects of the suppressor on protein levels were carried out. Firstly, stable cell lines of HaCaT cells expressing the miRNA 4 construct and the non-targeting control were established, made possible by the presence of a Blasticidin-resistance gene on the miRNA expression plasmid (see section 3.2.1.6). Selection was complete after approximately 3 weeks of treatment of nucleofected cells with $5 \mu\text{g/ml}$ or $7.5 \mu\text{g/ml}$ Blasticidin. As the protein is expressed from this cell line at quite low levels, a large number of HaCaT cells (approximately 1×10^7) are required for adequate analysis of protein levels. This is unlike the previously discussed RNA studies, of which analysis of COL7A1 mRNA levels is possible from just a small number of cells (a sufficient amount of RNA was isolated from samples comprising approximately 1×10^5 cells after FAC sorting). Additionally,

similar to RNA analysis from FAC-sorted cells, in principle this stable cell line model would allow for testing of a population consisting exclusively of cells expressing the transfected construct. Collagen type VII protein is a large protein of 290kDa, which is secreted from HaCaT cells after its synthesis. For this reason, media from cells of the stable cell line populations was collected and treated as described in section 2.2.4.1, in order to perform western blot analysis of collagen VII protein levels. After collection of the protein-containing media, cells were lysed and RNA was isolated using the Qiagen RNeasy® Plus mini kit as per section 3.2.1.6. However, upon analysis of RNA from the stable cell line populations, the expected approximate 50% deduction of COL7A1 mRNA, as previously observed in transfected and FAC-sorted samples, was not replicated. No significant difference in COL7A1 mRNA levels between miRNA 4-treated and NT miRNA-treated cell populations was observed despite the fact that these cell populations conferred blasticidin resistance ($p=0.0667$) (figure 3.11). Thus as expected, following probing for collagen type VII protein by western blotting using secreted media protein from stable cell line populations, there appeared to be no visible difference in collagen VII protein levels between samples (figure 3.12). Thus these stable cell lines were found not to be a suitable model for analysis of suppression of collagen type VII protein.

Further to this preliminary work with stable cell lines, another method of testing for suppression of collagen type VII protein was undertaken. A large scale transfection of the miRNA 4 and non-targeting constructs was carried out, with 25 nucleofection samples of each construct added to single T175 tissue culture plates, in principle sufficient for adequate detection of the COL7A1 protein product (see section 3.2.1.7). Western blot analysis was then carried out on media protein secreted from these transfected cells as described in section 3.2.1.7. However again no significant suppression of collagen type VII protein was seen in miRNA 4-treated cells compared to the NT miRNA-treated control cells (figure 3.13). It is of note that due to the nature of the collagen type VII protein no sufficient control protein had been identified at this point which would ensure proof of equal loading of samples in all wells. As mentioned, collagen type VII is a large protein and is secreted from the cells into the media and so a protein with these mutual properties would be required. Thus control proteins such as β -actin which are routinely used to control for loading during western

blot analyses for many other proteins would not be relevant in this case (see section 3.4).

3.3.4 shRNA vector construction

It is of note that levels of suppression of COL7A1 mRNA of greater than 50% were not achieved despite the testing of 13 different artificial microRNA constructs and despite Invitrogens claims that more than 70% of designed miRNAs reduce target gene expression by at least 70% (Invitrogen BLOCK-iT™ Pol II miR RNAi Expression Vector Kits user manual, Version E, June 2007). An increase in RNAi-mediated suppression of COL7A1 would be ideal and so two COL7A1-targeting shRNA molecules were then designed (mi-sh4 and mi-sh6) using the OligoEngine pSuper RNAi system which is based on the optimal shRNA-design algorithms as described by Brummelkamp *et al.*, 2002a. These constructs were designed to target the same regions of the COL7A1 transcript as two of the most potent artificial miRNAs, miRNA 4 and 6. Thus, once cloned and tested *in vitro*, it would be possible to directly compare the efficiencies of both shRNA constructs with those of their equivalent artificial miRNA-based vectors in an effort to identify any suppression-limiting effects elicited by the microRNA characteristics of the artificial miRNA vector. A non-targeting control vector was also designed, which similar to mi-sh4 and 6, would be processed to the same siRNA transcript as its equivalent non-targeting artificial miRNA.

shRNAs were designed according to the OligoEngine pSuper RNAi system, which provides a mammalian expression vector that directs intracellular synthesis of siRNA-like transcripts (pSuper manual: A Vector System for Expression of Short Interfering RNA, OligoEngine). The pSuper.gfp/neo vector was utilised, which uses the polymerase-III H1 gene promoter. It produces a small RNA transcript lacking a polyadenosine tail and has a well-defined start of transcription and a termination signal consisting of five thymidines. Cleavage of the transcript occurs after the second uridine of the termination site, resulting in a sequence resembling the ends of synthetic siRNAs. The top oligos of the two shRNA molecules were designed with their 21 nucleotide COL7A1 target sequences in both sense and antisense orientation, separated by a 9 nucleotide spacer sequence and *Bgl*III and *Hind*III 5' and 3' ends respectively (figure 3.14). The resulting transcript is predicted to fold back on itself to

form a stem-loop structure which is cleaved in the cell to produce a functional siRNA which is stably expressed (figure 3.15). For the construction of the shRNA vectors, double-stranded oligos were annealed, digested and inserted into the SAP-treated linearised pSuper.gfp/neo vector using T4 DNA ligase. Positive clones were sequenced (see figure 3.5 for primer sequences) in order to ensure the correct sequence and orientation and DNA preparations of each construct were made and used in further analysis.

3.3.5 Analysis of COL7A1 mRNA suppression by shRNAs 4 and 6

Upon cloning of the two COL7A1-targeting and one non-targeting shRNA molecules into the pSuper.gfp/neo vector, HaCaT cells were transfected by the method of nucleofection in the same manner as the testing of artificial microRNAs (see section 3.2.2.3). Transfection of each construct was carried out in duplicate and RNA was extracted directly from 6-well plates 24 hours after transfection. COL7A1 mRNA levels were analysed as per section 3.2.1.4. Interestingly, no significant difference in COL7A1 expression between mi-shRNA 4 and NT mi-shRNA was observed ($p=0.675$), despite the equivalent artificial miRNA 4 having yielded approximately 50% suppression (section 3.3.2). In contrast, mi-shRNA 6 significantly suppressed COL7A1 transcript levels, with a mean suppression of approximately 37% ($p<0.0005$) (figure 3.16). This level of COL7A1 suppression is comparable to that obtained with the equivalent artificial miRNA 6 which similarly reduced COL7A1 mRNA expression by approximately 37% compared to the non-targeting control vector.

3.4 Discussion

The mechanism of RNA interference represents a powerful tool with which to prevent the natural cellular process of translation and thus achieve suppression of production of a particular protein, as ably demonstrated by the myriad of publications using this suppression technology over the past decade (Angaji *et al.*, 2010; Stieger and Lorenz, 2010; Lochmatter and Mullis, 2011). Dominantly inherited dystrophic EB, caused by mutations of the COL7A1 gene, is characterised by the production of a mutant type VII collagen protein. This leads to a weakening of the Basement Membrane Zone, causing a phenotype of blistering and erosions of the skin and mucous membranes. Thus far, no significant advances have been made in the suppression of COL7A1 in any gene therapy directed approach for DDEB. The mechanism of RNAi holds great potential for the provision of a possible means of treatment of this form of EB, by acting to suppress the mutant type VII collagen protein. It is such a strategy that has been employed in this study and the focus of this particular chapter of this Ph.D. thesis has centred on the use of artificial microRNAs as suppressor agents. The artificial microRNAs in question while based on the murine pri-miR-155 sequence, were designed such that upon their processing by the cells endogenous RNAi machinery, the resulting double-stranded siRNA molecules are complementary to regions of the COL7A1 transcript and thus in principle, should target COL7A1 mRNA for degradation.

Initially, it was decided to test the efficiency of artificial microRNA molecules for their ability to suppress expression of the collagen VII gene. The effectiveness of 13 such constructs was assessed in human epidermal keratinocyte cells (HaCaTs), a spontaneously immortalised human epithelial cell line from adult skin, which exhibits full epidermal differentiation capacity (Boukamp *et al.*, 1988). HaCaTs represent an ideal cell line for use in these studies due to the fact that they endogenously produce collagen type VII protein as well as their ease of availability, genetic homogeneity and comparable behaviour to primary keratinocytes (Jacobsen *et al.*, 2006). The artificial miRNA constructs were delivered to cells by the method of nucleofection, a non-viral electroporation-based gene transfer method which delivers directly to the cell's nucleus (Trompeter *et al.*, 2003; Martinet *et al.*, 2003). As keratinocyte cells are notoriously difficult to transfect (Nead and McCance, 1995; Compton *et al.*, 2000) the

highly efficient method of transfection by nucleofection was chosen for delivery of the RNAi molecules in this study. Such a delivery method is also useful for the constructs in question in this study, as in principle direct delivery of miRNA plasmids to the nucleus should allow for the rapid initiation of their processing from pri-miRNAs to pre-miRNAs by the nuclear enzyme Drosha (Lee *et al.*, 2003). The presence of an emGFP marker gene on the artificial miRNA plasmid allowed for segregation of positive cells which were expressing the constructs, following employment of the method of fluorescent activated cell sorting (FACS). After isolation of RNA, COL7A1 mRNA levels were analysed by real-time rtPCR in targeting and non-targeting-treated cells. Expression was normalised using expression levels of a range of endogenous control genes which had previously been found in the Trinity College laboratory to be stably expressed in HaCaT cells and thus appropriate for normalisation of expression data (Allen *et al.*, 2008). Of the thirteen artificial miRNA constructs generated and tested, the most potent suppressor reduced mean COL7A1 mRNA expression levels by approximately 50% ($p < 0.0005$; figure 3.7).

As decreased levels of COL7A1 mRNA should coincide with decreased levels of collagen type VII, the transcript's protein product, validation of the achieved RNA findings with protein data was attempted. In this case however, neither a blasticidin-resistant stable cell line model nor a large scale transfection resulted in such a validation at the protein level. With regard to the stable cell line, the previously seen significant suppression of approximately 50% in transient transfections by miRNA 4 at the RNA level following FAC sorting, failed to be replicated in this stable cell line (figure 3.11). In this case it is possible that the artificial miRNA vector may have integrated into an area of the genome from which only low levels of expression were achievable, thus preventing adequate levels of expression of the targeting sequence to allow COL7A1 suppression. While it would have been possible to attempt to explain if this may be the case using, for example, RNase protection assays, or by measuring whether or not the cells were expressing the emGFP reporter gene contained in the artificial microRNA vector by fluorescence microscopy, it was deemed that the focus of the work should be the characterisation of one or more potent suppressor of the COL7A1 gene rather than exploring reasons for absence of potency with some suppressors. Large scale transfection of cells with miRNA 4 also failed to adequately detect the construct's suppression abilities at the protein level. However, due to the

nature of this experimental model using transient transfections, it is likely that any suppression which is occurring at the protein level may be somewhat difficult to validate. As the transfection efficiency of the artificial microRNA plasmid is not at 100% efficiency (it is in the order of approximately 70%, as found during the original undertaking of FAC sorting of nucleofected cells) and thus not all cells are expressing the miRNA 4 construct, combined with the fact that maximum suppression levels of just approximately 50% were obtained at the RNA level from FACs purified cells, any reduction in collagen type VII protein may possibly be somewhat masked. It is of note that at this point, no control protein was probed for, in order to ensure proof of equal loading of samples in all wells. Collagen type VII is a large 290kDa protein and is secreted from HaCaT cells into the media in which the cells are incubated. Following collection of this media, the protein is concentrated by filtering, and any protein below 100kDa is excluded. Thus during these preliminary suppression studies, a large protein which is secreted from HaCaT cells and is detectable by western blotting was not identified and so evidence of equal loading of all wells was difficult to attest. However, following these studies the Laminin 332 protein was later identified as a possible control protein (see chapter 4). Its large 165kDa $\alpha 3$ chain and 145kDa $\beta 3$ chain as well as the fact that it is secreted from HaCaT cells make it an ideal choice. Thus all future studies involving measurement of collagen VII protein levels from HaCaT cells by western blotting have included the additional step of probing for Laminin 332 protein.

It was primarily decided to initiate this study with the testing of artificial microRNA vectors due to their increasing recognition in recent times as having a number of advantages over previously widely used polymerase III-based shRNA suppressor molecules. As well as in principle allowing for a greater level of control of the therapeutic due to the tissue-specific or inducible expression that polymerase II promoters permit, these miRNA expression cassettes are also thought to represent an RNAi methodology with reduced toxicity in comparison with previously widely used shRNA vectors, due to the fact that the artificial microRNA vectors are modelled on the cells own endogenous miRNA machinery (McBride *et al.*, 2008). However despite the testing of thirteen such molecules, suppression levels of the COL7A1 transcript of greater than 50% failed to be realised with any of the constructs. This was despite the claims by Invitrogen that at least 70% suppression with more than

70% of their designed artificial miRNAs may be achieved using their online BLOCK-iT™ Pol II miR RNAi Designer program. Subsequent to these findings, target sequences of the two most potent artificial microRNAs were used in the conversion of miRNA 4 and 6 to equivalent polymerase III-based shRNA molecules, cloned into the pSuper.gfp/neo vector and driven by the H1 promoter. No significant improvement in suppression ability was seen with the shRNAs when compared to their equivalent artificial microRNAs (figure 3.16). In fact, while equivalent suppression levels were obtained with both mi-shRNA 6 and miRNA 6 (approximately 37% COL7A1 mRNA suppression in both cases), no significant suppression was obtained with mi-shRNA 4 compared to the non-targeting control shRNA. It is notable that analysis of the efficiency of suppression of shRNAs was undertaken using RNA samples isolated directly from cells in 6-well plates, rather than RNA from those which had been FAC sorted. It is possible that an improvement in COL7A1 suppression levels would be detected in RNA isolated from FACs-purified cells which had been treated with the shRNA constructs. A substantial increase in measured suppression levels in such samples would be unlikely however, considering the high transfection efficiency in the order of 70% (see figure 3.6) which was generally found to arise using this experimental model (and considering no notable differences in size exist between the pcDNA 6.2-GW/EmGFP-miR and pSuper.gfp/neo vectors which are approximately 5.7kb and 5.4kb in size respectively).

The combined artificial microRNA and shRNA COL7A1 suppression data obtained during these studies suggests that the levels of suppression achieved thus far may not be due to limitations associated with the microRNA vector specifically. It is possible that the particular 21bp regions of the COL7A1 gene being targeted may not have been particularly amenable to RNAi-mediated suppression, possibly due to the mRNA secondary structure. The endogenous RNAi machinery of HaCaT cells may also be a limiting factor whereby processing of both the artificial miRNA and shRNA vectors to siRNAs in the nucleus and in the cytoplasm may be restricted. The affects of these constructs have only been tested in the HaCaT cell line, and due to differences in expression patterns between cell types, there is the possibility that analysis of their affect in other cell lines may result in an improvement in COL7A1 suppression levels. Moreover, the affects of these constructs in an *in vivo* model have yet to be examined, and other published findings have found significantly greater

efficiencies associated with artificial microRNA-based shRNA constructs *in vivo*, in a mouse model, in comparison with *in vitro* data (McBride *et al.*, 2008). Thus these artificial microRNA molecules still hold some potential for use in further studies, despite their current seeming suppression limitations, particularly due to the previously mentioned benefits of their use over other potential suppressor molecules.

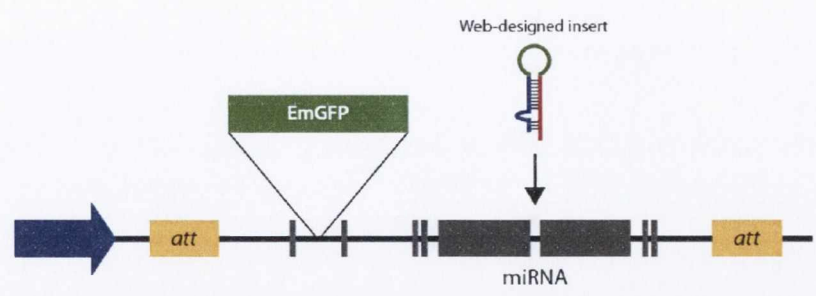
It is also of note that due to the nature of the disease pathology, there is a possibility that controlled shifting of wild type to mutant collagen VII ratios, rather than complete obliteration of the mutant protein, may have therapeutic potential for DDEB. In the case of DDEB, the mutant protein has a dominant negative affect as it accounts for 50% of collagen VII polypeptides and incorporates into the triple-helical homotrimers, affecting folding of normal polypeptides. Thus if more wild-type than mutant protein is expressed, it is predicted that more wild-type collagen VII polypeptides would be incorporated into collagen VII homotrimers, and so there would be an increase in the number of fully functional and stable anchoring fibril structures. As mentioned in chapter 1, the level of collagen type VII expression required for a wild-type phenotype is not yet fully known. However, the fact that individuals who are heterozygous carriers of a nonsense COL7A1 mutation do not display symptoms of DEB despite only producing 50% of normal levels of the protein (Tidman and Eady, 1985), suggests that levels of collagen type VII of between 0 and 50% of normal levels should in principle be sufficient to prevent the dystrophic EB phenotype. Initial published data from recent studies have shown that counteracting the dominant negative interference of dominant COL7A1 mutations may be possible by shifting the balance between wild-type and mutant chains forming the collagen triple helix (Fritsch *et al.*, 2009). One such way of shifting the wild-type to mutant ratio would be to suppress the collagen VII gene. In such as case, the 50% suppression capability of the most potent tested artificial microRNA construct may represent an ideal tool for creating such a shift in the balance.

3.5 Conclusion

Artificial microRNA expression constructs were generated, designed to reduce COL7A1 mRNA expression levels, through the application of the pcDNA6.2 vector system. The results presented in the current chapter of this Ph.D. thesis represent a

preliminary work-up of potential RNAi-based suppressors for COL7A1 and demonstrate that suppression of COL7A1 mRNA of approximately 50% was achieved with the optimal artificial microRNA suppressor. These constructs may potentially be suitable for use in a future gene therapy-based therapeutic for the dominant form of Dystrophic EB. Given the levels of COL7A1 suppression achieved with the artificial miRNA construct however, it was of interest to ascertain whether or not a greater reduction of COL7A1 mRNA expression could be attained. The work associated with addressing this question forms the basis of the studies described in Chapter 4 of this Ph.D. thesis.

a)



b)

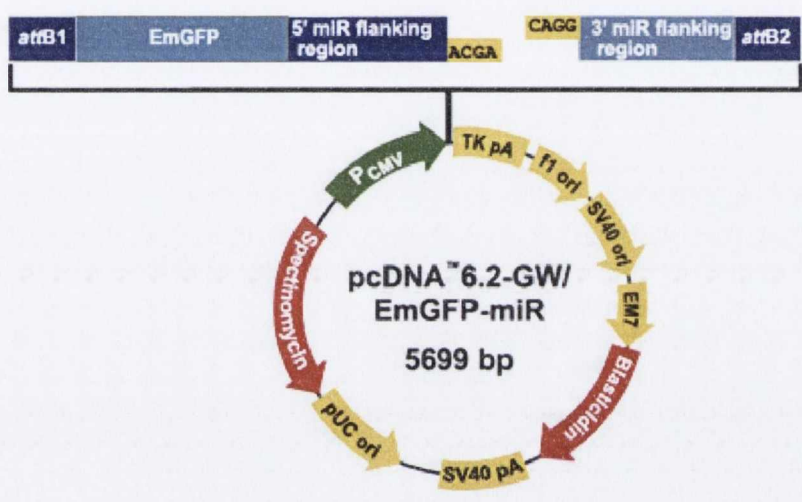


Figure 3.1 Map of pcDNA™6.2-GW/EmGFP-miR vector: a) 60bp web-designed inserts are inserted into the pcDNA6.2 vector between mir-155 flanking regions. EmGFP is located upstream of the insert. b) The linearised 5.7kp vector has ACGA and CAGG overhangs, allowing for insertion of inserts with complementary overhangs. Inserts are under the control of a CMV promoter. The plasmid contains gene sequences conferring spectinomycin resistance for selection in *E. coli* and blasticidin resistance for selection in *E. coli* and mammalian cells (Figure from BLOCK-iT™ Pol II miR RNAi Expression Vector Kits user manual, Invitrogen, Version E, June 2007).



Figure 3.2 Engineered pre-miRNA structure: Structural features of the ds oligonucleotide encoding the engineered pre-miRNA. The 5' and 3' overhangs are complementary to the vector and required for directional cloning. The 21 nucleotide antisense sequence (mature miRNA) is derived from the target gene, and a short spacer of 19 nucleotides forms the terminal loop. Finally, the short sense sequence has 2 nucleotides removed to create an internal loop. (Figure and Legend modified from Invitrogen BLOCK-iT™ Pol II miR RNAi Expression Vector Kits user manual, Version E, June 2007).

StrandName	Sequence 5'- 3'
mi1 Top	TGCTGTCCAGGTGATAGCTGCTGAGCGTTTTGGCCACTGACTGACGCTCAGCATATCACCTGGA
mi1 Bottom	CCTGTCCAGGTGATATGCTGAGCGTCAGTCAGTGGCCAAAACGCTCAGCAGCTATCACCTGGAC
mi2 Top	TGCTGTAAGACAGCAGGCCAACCTGAGTTTTGGCCACTGACTGACTCAGGTTGCTGCTGTCTTA
mi2 Bottom	CCTGTAAGACAGCAGCAACCTGAGTCAGTCAGTGGCCAAAACCTCAGGTTGGCCTGCTGTCTTAC
mi3 Top	TGCTGAATACTCGGAGTATTCAGAGTGTTTTGGCCACTGACTGACACTCTGAACTCCGAGTATT
mi3 Bottom	CCTGAATACTCGGAGTTCAGAGTGTTCAGTCAGTGGCCAAAACCTCTGAATACTCCGAGTATTC
mi4 Top	TGCTGATCCCAAGGAGCTTCAGGGTCGTTTTGGCCACTGACTGACGACCCTGACTCCTTGGGAT
mi4 Bottom	CCTGATCCCAAGGAGTCAGGGTCGTCAGTCAGTGGCCAAAACGACCCTGAAGCTCCTTGGGATC
mi5 Top	TGCTGTCCAGTAAGAACACAATGTGAGTTTTGGCCACTGACTGACTGACATTGTTCTTACTGGA
mi5 Bottom	CCTGTCCAGTAAGAACAATGTGAGTCAGTCAGTGGCCAAAACCTGACATTGTGTTCTTACTGGAC
mi6 Top	TGCTGAGAAGAAGAAGTCACGTGGTGGTTTTGGCCACTGACTGACCCACCAGTCTTCTTCTTCT
mi6 Bottom	CCTGAGAAGAAGAAGACTGGTGGTTCAGTCAGTGGCCAAAACCCACCAGTACTTCTTCTTCTC
mi7 Top	TGCTGTATTCTGGATGGTCAGTTCGGTTTTGGCCACTGACTGACCGGAAGTCATCCAGAATA
mi7 Bottom	CCTGTATTCTGGATGCAGTTCGGTCAGTCAGTGGCCAAAACCGGAAGTCATCCAGAATAC
mi8 Top	TGCTGTAATGGTGACAGATGAGCTGCGTTTTGGCCACTGACTGACGCAGCTCATGTACCATTAC
mi8 Bottom	CCTGTAATGGTGACATGAGCTGCGTCAGTCAGTGGCCAAAACGCAGCTCATCTGTACCATTAC
mi9 Top	TGCTGTGTAAGAGACGCCAGGCTCTAGTTTTGGCCACTGACTGACTAGAGCCTCGTCTCTTACA
mi9 Bottom	CCTGTGTAAGAGACGAGGCTCTAGTCAGTCAGTGGCCAAAACCTAGAGCCTGGCGTCTCTTACAC
mi10 Top	TGCTGATGACTGTAAGACAGCAGGCCGTTTTGGCCACTGACTGACGGCCTGCTCTTACAGTCAT
mi10 Bottom	CCTGATGACTGTAAGAGCAGGCCGTCAGTCAGTGGCCAAAACGGCCTGCTGTCTTACAGTCATC
mi11 Top	TGCTGAGCATCTGGTGCCAACATGTAGTTTTGGCCACTGACTGACTACATGTTTACCAGATGCT
mi11 Bottom	CCTGAGCATCTGGTGAACATGTAGTCAGTCAGTGGCCAAAACCTACATGTTGGCACCAGATGCTC
mi12 Top	TGCTGTGAATATGTCACCTCTCAAGGGTTTTGGCCACTGACTGACCCCTTGAGATGACATATTCA
mi12 Bottom	CCTGTGAATATGTCATCTCAAGGGTCAGTCAGTGGCCAAAACCCCTTGAGAGGTGACATATTAC
mi13 top	TGCTGACCACATTAAGCCCAGAAGCCGTTTTGGCCACTGACTGACGGCTTCTGCTTAATGTGGT
mi13 Bottom	CCTGACCACATTAAGCAGAAGCCGTCAGTCAGTGGCCAAAACGGCTTCTGGGCTTAATGTGGTC
NT top	TGCTGAAATGTACTGCGGTGGAGACGTTTTGGCCACTGACTGACGCTCTCCACGCAGTACATTT
NT bottom	CCTGAAATGTACTGCGTGGAGACGTCAGTCAGTGGCCAAAACGCTCTCCACGCAGTACATTTTC

Figure 3.3 Artificial microRNA sequences: DNA sequences of top and bottom strands of artificial microRNA oligonucleotides. Red nucleotides indicate overhangs following annealing of top and bottom strands. Green nucleotides indicate the terminal loop region based on the endogenous murine miRNA miR-155.

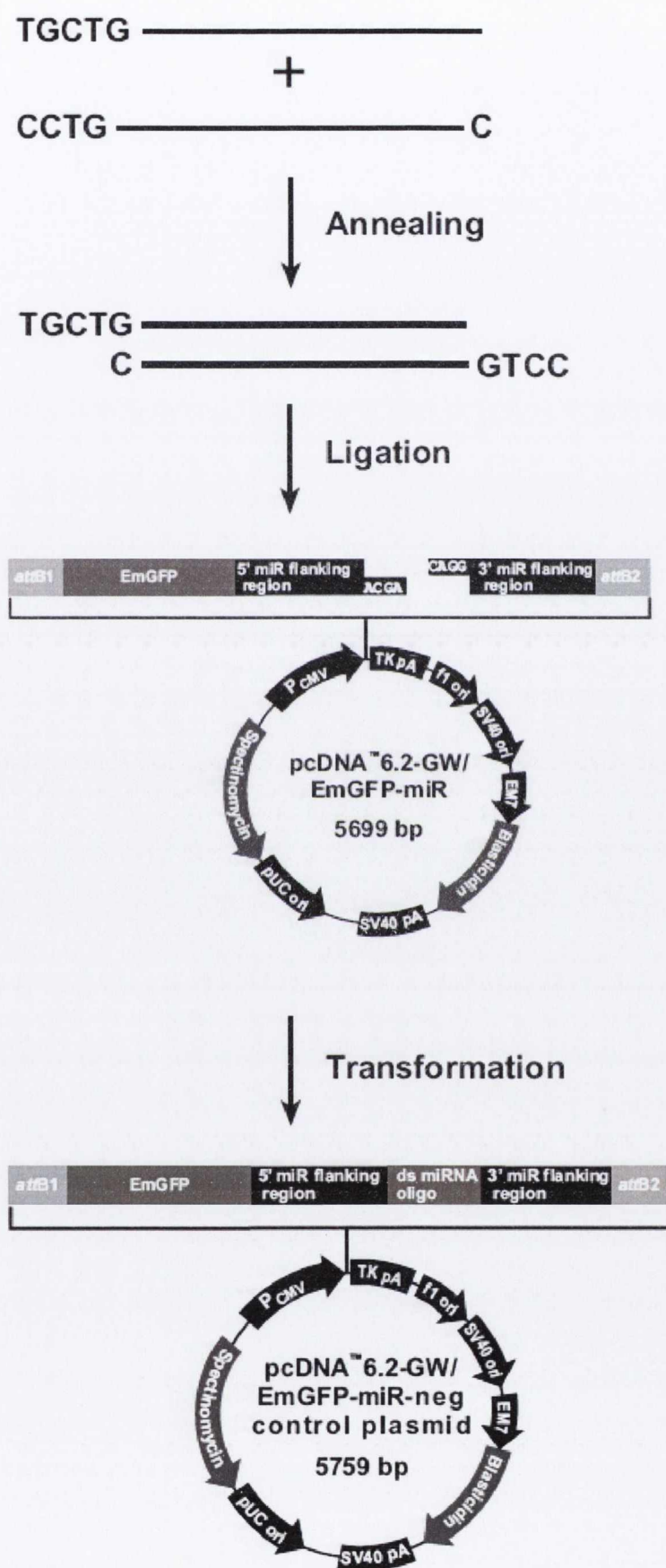


Figure 3.4 Overview of Artificial miRNA vector construction

	Sequence 5'-3'
EmGFP forward sequencing primer	GGCATGGACGAGCTGTACAA
miRNA reverse sequencing primer	CTCTAGATCAACCACTTTGT
pSuper forward sequencing primer	CTTGCGACTGCAGTAGTTGG
pSuper reverse sequencing primer	AACAGCTATGACCATG

Figure 3.5 Primer Sequences: DNA sequences of primers used in Chapter 3. For primers used for real-time rtPCR, see section 2.1.3.

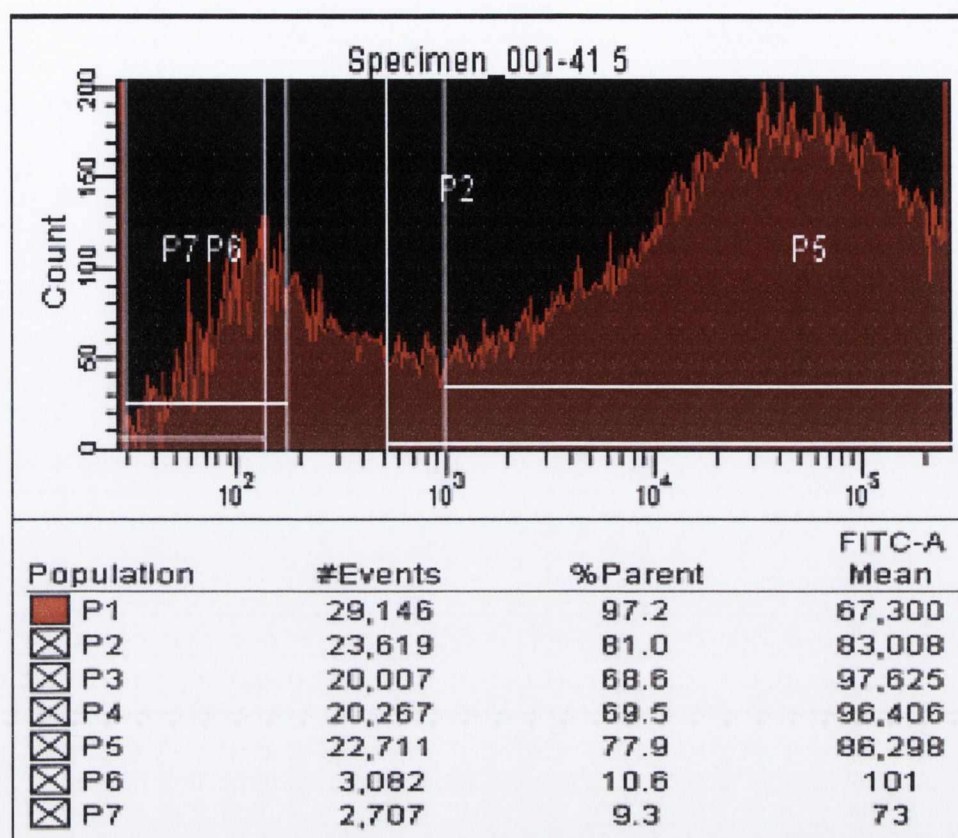


Figure 3.6 Distribution of HaCaT cells during FACS analysis: HaCaT cells were transfected with an artificial microRNA construct expressing emGFP. The P5 population of cells indicates those cells expressing emGFP and thus the artificial microRNA construct (77.9% of cells in this case). The P7 population of cells indicates the cells which do not express emGFP and thus are not fluorescent, representing those cells which did not receive the artificial microRNA construct during the nucleofection process. Cells from the P5 population were collected and analysed for RNAi-mediated COL7A1 suppression.

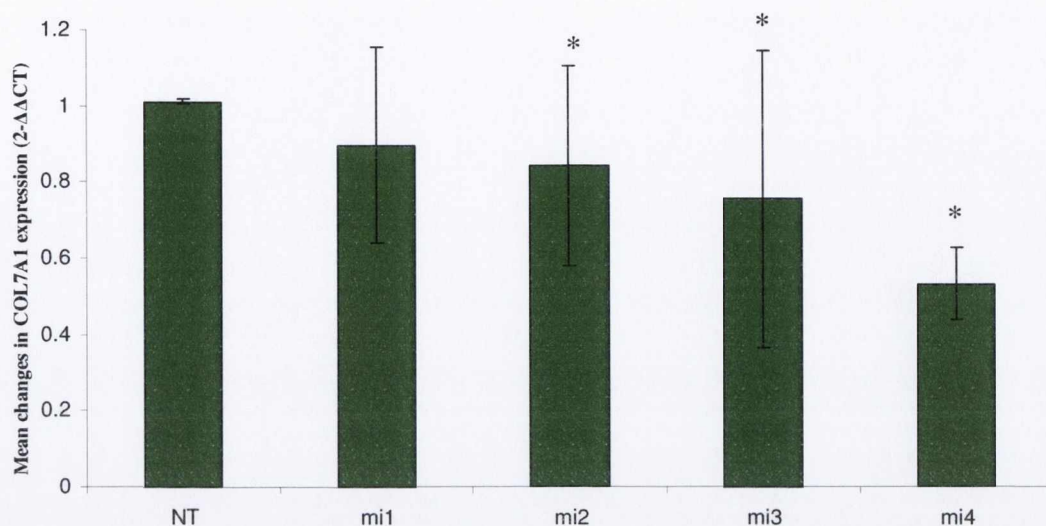


Figure 3.7 Mean changes in COL7A1 expression: effect of artificial microRNAs 1-4 on COL7A1 expression. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5μg DNA (artificial miRNA 1, 2, 3, 4 or NT). 24 hours later, cells expressing the artificial miRNA constructs were isolated by FAC sorting and their COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate and the experiment was repeated three times. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. While miRNA 1 failed to significantly suppress COL7A1 mRNA when compared to a non-targeting negative control artificial miRNA ($p=0.140$), significant COL7A1 suppression was achieved with miRNA 2 ($p=0.039$), miRNA 3 ($p=0.034$) and miRNA 4 ($p<0.0005$). These constructs reduced COL7A1 suppression by approximately 16%, 25% and 47% respectively.

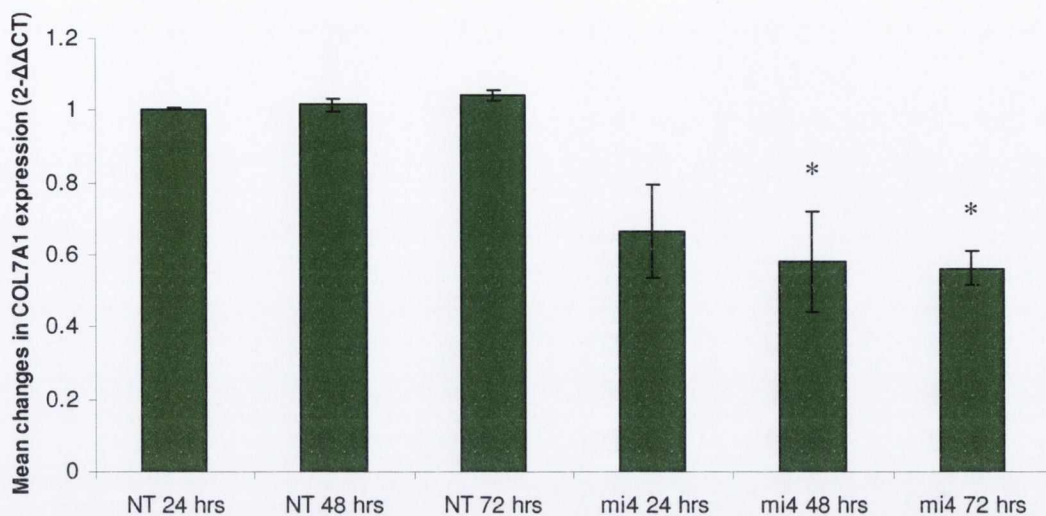


Figure 3.8 Mean changes in COL7A1 expression: effect of artificial microRNA 4 on COL7A1 expression following incubation for a range of time periods. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5μg DNA (artificial miRNA 4 or NT). RNA was directly isolated from duplicate transfected sample wells after 24, 48 and 72 hours and their COL7A1 mRNA levels were measured by real-time rtPCR. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Mean suppression of COL7A1 mRNA was found to increase slightly over time (approximate mean suppression of 34%, 42% and 44% respectively; $p = 0.065$, 0.049 and 0.005 respectively). However, COL7A1 mRNA suppression visible at 72 hours was not significantly different from that seen at 24 hours ($p = 0.401$).

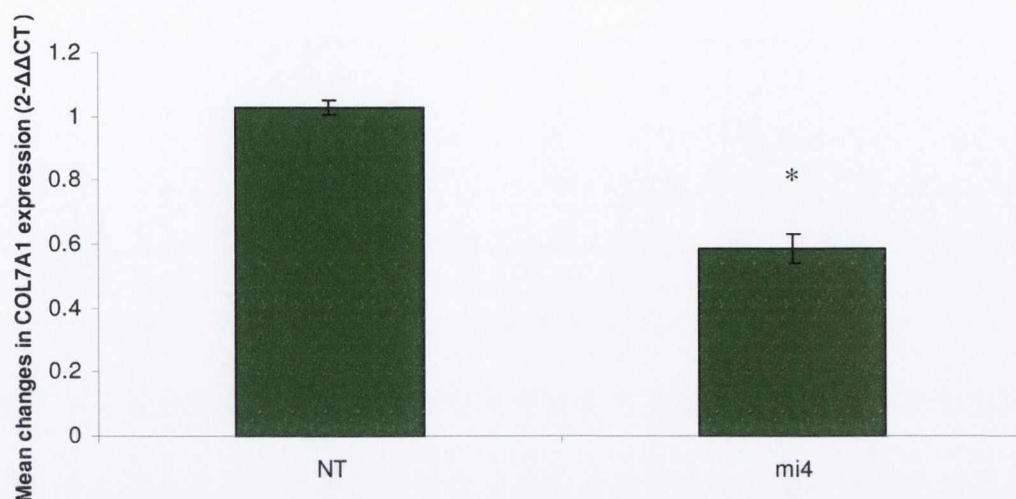


Figure 3.9 Mean changes in COL7A1 expression: effect of artificial microRNA 4 on COL7A1 expression after 72 hours. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5μg DNA (artificial miRNA 4 or NT). 72 hours later, cells expressing the artificial miRNA constructs were isolated by FAC sorting and their COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant mean suppression values of approximately 42% were seen in cells treated with miRNA 4 after 72 hours compared to those treated with miRNA NT ($p < 0.0005$).

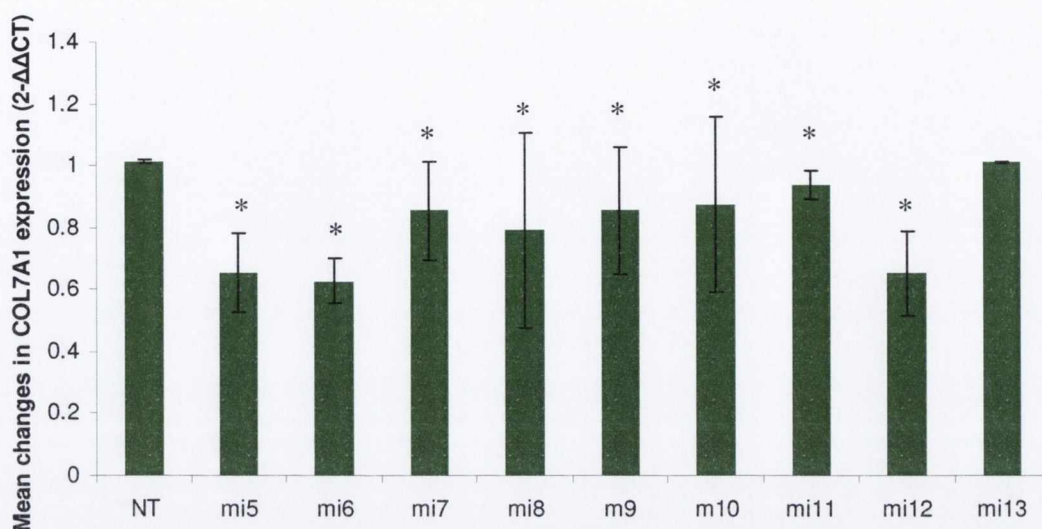


Figure 3.10 Mean changes in COL7A1 expression: effect of artificial microRNAs 5-13 on COL7A1 expression. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5 μ g DNA (artificial miRNA 5-13 or NT). 24 hours later, cells expressing the artificial miRNA constructs were isolated by FAC sorting and their COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate and this was repeated either 2 or 3 times for each construct. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Significant COL7A1 mRNA suppression was achieved with all but one of these constructs (miRNA 13; $p = 0.702$). Of the 8 remaining constructs, miRNA 6 was the optimal suppressor. It reduced COL7A1 mRNA expression by approximately 37% ($p < 0.0005$).

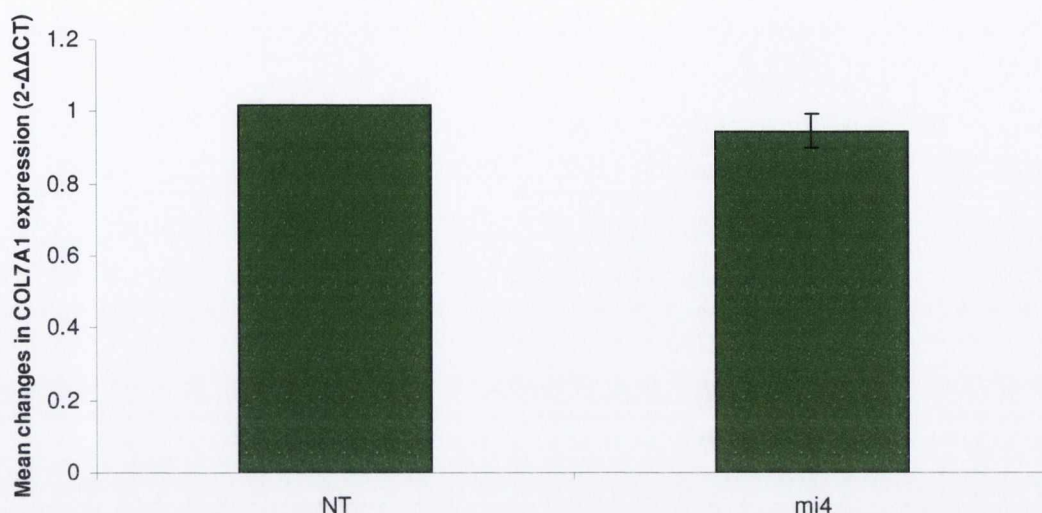


Figure 3.11 Mean changes in COL7A1 expression: analysis of RNA from stable cell lines. Stable cell lines were established expressing NT miRNA and miRNA 4. Establishment of such cell lines was possible due to the presence of a blasticidin-resistance gene on the pcDNA 6.2 vector. RNA was isolated from stable cell line cells and COL7A1 mRNA levels measured by real-time rtPCR. Error bars represent standard deviation. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. No significant difference in COL7A1 mRNA levels between miRNA 4-treated and NT miRNA-treated cells was observed despite the fact that these cell populations were blasticidin-resistant ($p = 0.0667$).

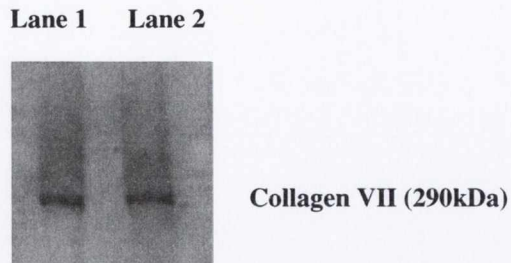
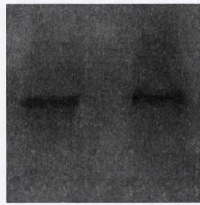


Figure 3.12 Mean changes in collagen type VII protein levels: western blot analysis of collagen type VII protein from conditioned media of stable HaCaT cell lines. Cells were transfected by nucleofection with NT miRNA (Lane 1) and miRNA 4 (Lane 2) and cells expressing the constructs were stably selected for, made possible by the presence of the blasticidin-resistance gene on the plasmids in question. Following establishment of stable cell lines, cells in tissue culture plates of 175cm² were incubated in 13ml serum-free media and 24 hours later, media was collected, conditioned and filtered. Media protein samples were subjected to 6% SDS-PAGE. The positions of the bands corresponding to full-length 290kDa collagen type VII are indicated. Western blot analysis was carried out on three separate occasions.

Lane 1 Lane 2



Collagen VII (290kDa)

Figure 3.13 Mean changes in collagen type VII protein levels: western blot analysis of collagen type VII protein from conditioned media of mass-transfected HaCaT cells. A large scale nucleofection experiment was undertaken, whereby 25 transfections with NT miRNA (Lane 1) and 25 transfections with miRNA 4 (Lane 2) were carried out and both sets of transfections were transferred to tissue culture plates of 175cm². 24 hours post transfection, cells were incubated in 13ml serum-free media and a further 24 hours later, media was collected, conditioned and filtered. Media protein samples were subjected to 6% SDS-PAGE. The positions of the bands corresponding to full-length 290kDa collagen type VII are indicated. Western blot analysis was carried out on three separate occasions.

Strand name	Sequence
mi-sh4 Top 5'-3'	GAGTACAGGATCCCGACCTGAAGCTCCTGGGATTTCAAGAGAATCCCAAGGAGCTTCAGGGTCTTTTTAAGCTTTGAGTACA
mi-sh4 Bottom 3'-5'	CTCATGTCTAGGGCTGGGACTTCGAGGAACOCCTAAGTTCTCTTAGGGTTCCTCGAAGTCCACAGAAAAATTCGAACTCATGT
mi-sh6 Top 5'-3'	GAGTACAGGATCCCCCACAGTGACTTCTTCTTCAAGAGAAAGAAAGTCACTGGTGGTTTTAAGCTTTGAGTACA
mi-sh6 Bottom 3'-5'	CTCATGTCTAGGGGGGTGGTCACTGAAAGAGAAAGAAGTTCTCTTCTTCTCAGTGACCAACAAAAATTCGAACTCATGT

Figure 3.14 mi-shRNA sequences: DNA sequences of top and bottom strands of shRNA oligonucleotides. Red letters indicate *Bam*HI and *Hind*III restriction sites. Following annealing of oligos, digesting with *Bam*HI and *Hind*III results in 5' *Bgl*II and 3' *Hind*III overhangs (5'-GATC-3' and 3'-TCGA-5' respectively), which allows for cloning into the pSuper vector. Green letters indicate the hairpin. The first seven and last seven nucleotides of each oligo represent random sequences.

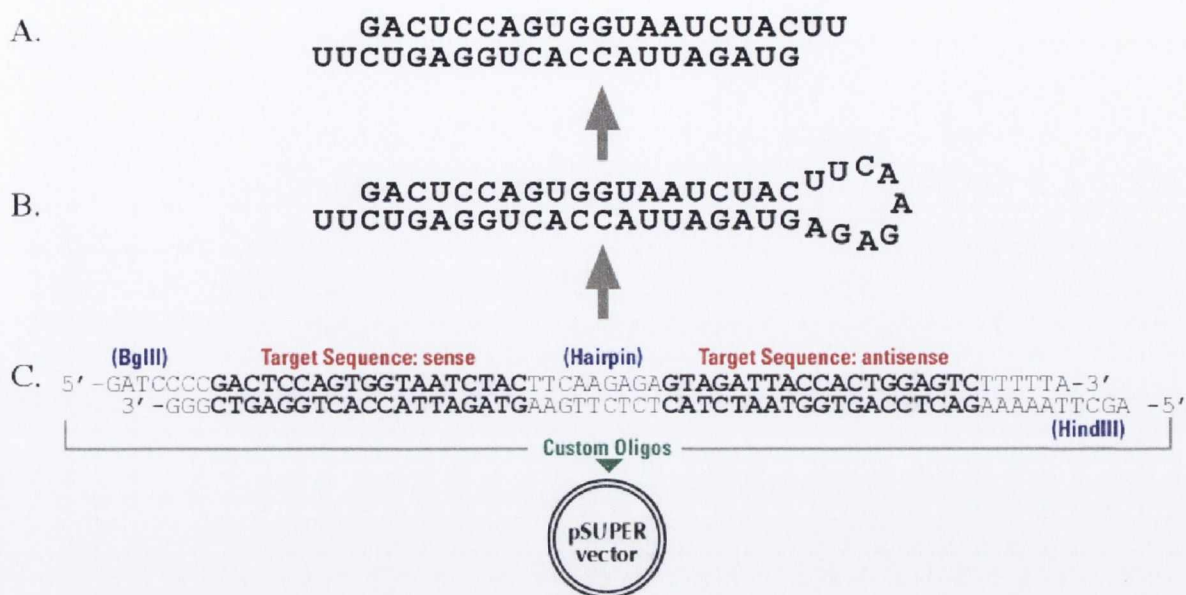


Figure 3.15 Processing of shRNAs: Diagram shows the design structural features of shRNA sequences. shRNAs are transcribed and fold back on themselves to form hairpin structures. Processing of shRNAs leads to formation of functional siRNAs. The 19nt target sequence in this diagram is derived from the pSuper RNAi System™ manual. In the case of mi-shRNA 4 and 6, the target sequences based on the target sequences of miRNA 4 and 6, were 21 nucleotides in length (figure derived from pSuper RNAi System™ manual).

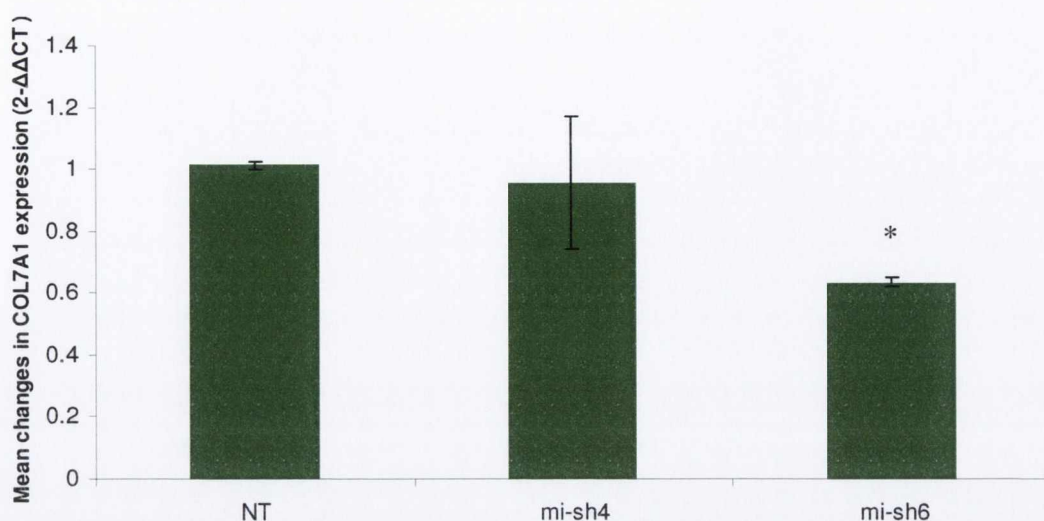


Figure 3.16 Mean changes in COL7A1 expression: effect of mi-shRNA molecules on COL7A1 expression. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5μg DNA (mi-shRNA 4, 6 and NT). RNA was directly isolated from duplicate transfected sample wells after 24 hours and their COL7A1 mRNA levels were measured by real-time rtPCR. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. No significant difference in COL7A1 mRNA levels was observed between mi-shRNA 4 and NT mi-shRNA ($p = 0.675$). mi-shRNA 6 significantly suppressed COL7A1 mRNA by approximately 37% ($p < 0.0005$). In comparison, the equivalent artificial microRNAs (miRNA 4 and 6) suppressed COL7A1 mRNA by approximately 47% ($p < 0.0005$) and 37% ($p < 0.0005$) respectively (see figures 3.7 and 3.10).

Chapter 4

siRNA-mediated COL7A1 suppression

4.1 Introduction

Since the 1990s, insights into the pathway involved in the process of RNAi and efforts to exploit this natural cellular process have become major areas of interest in the field of biology and genetics (section 1.6). RNAi not only offers itself as a valuable biological research tool for deciphering gene function or uncovering genes implicated in disease, it also has potential for use as a therapeutic for gene-related diseases. A number of both preclinical and clinical studies have arisen in recent years, which take advantage of the sequence-specific suppression of mutant genes afforded by RNAi technology (Davis, 2009; Kaiser *et al.*, 2010; Millington-Ward *et al.*, 2011). As outlined in chapter one of this thesis, patients suffering with the dominant form of the genetic disease dystrophic EB may be one such set of people who may potentially benefit from advances in the field of RNAi. It is dominant negative mutations of the COL7A1 gene which are implicated in this disease, the wild-type function of which is the production of anchoring fibrils composed of collagen type VII protein (Hovnanian *et al.*, 1992). DDEB-causing mutations lead to a phenotype of fragility of the skin and mucous membranes, the severity of which ranges between individuals and is dependent on the COL7A1 mutation in question, of which approximately 60 have so far been identified (Nakamura *et al.*, 2004).

RNAi is a natural regulatory mechanism which occurs within eukaryotic cells, mediated by RNAi molecules known as microRNAs (section 1.6.2). The process involves suppression of mRNA translation and this homology-dependent control of gene expression is carried out by small double-stranded RNA molecules. The RNAi pathway has been found to be artificially triggered in two different manners; the first by introduction into a cell of synthetic 21-23 basepair duplexes (siRNAs) and the second by introduction of DNA expression vectors which transcribe short hairpin RNAs (shRNAs). shRNA expression systems are typically driven by Polymerase III promoters, usually U6 or H1. The sense and antisense sequences of the single transcript self-anneal to form a hairpin structure (Brummelkamp *et al.*, 2002a; Paddison *et al.*, 2002). These hairpin molecules are then exported from the cell's nucleus by means of the cell's endogenous miRNA machinery, and upon entry to the cytoplasm they are processed into siRNAs. It is the siRNA antisense strand which is then incorporated into the RISC for RNAi-mediated suppression of its target mRNA,

whilst the siRNA sense strand is degraded (Geusens et al, 2009). The focus of chapter 3 of this thesis involved an investigation into suppression of the COL7A1 gene utilising artificial microRNA vectors. These artificial microRNAs which were described and tested in chapter 3 are based on the same design as shRNAs, although they differ slightly in structure such that they should have characteristics of the cell's own endogenous miRNAs as well as being driven by polymerase II promoters. The potential benefits of using such constructs over regular shRNA molecules are outlined in sections 1.6.4 and chapter 3.

Artificial miRNA and shRNA systems offer an advantage over the direct introduction of siRNAs into a cell, due to the fact that their continuous expression from their delivery DNA plasmids will increase the longevity of their effects. RNAi-mediated suppression using synthesised siRNAs has been found to be transient, with suppression only lasting approximately 3-5 days (Holen *et al.*, 2002; Amarzguioui *et al.*, 2005). While chemical modifications can be used to significantly augment longevity (Chernolovskaya and Zenkova, 2010; Bramsen and Kjems, 2011), a continuous cell-produced supply of RNAi molecules may be preferable for many applications of RNAi. Plasmid delivered RNAi molecules have an inherent higher stability and large amounts of shRNAs may be transcribed from each individual template plasmid. Furthermore, it is possible to achieve stable expression of shRNA-expressing vectors by either their integration into the genome or in episomal form. In principle, this should allow for constitutive expression of the RNAi suppressor molecules (Uprichard, 2005). The ability to include selectable markers or reporter genes within shRNA expression plasmids may also assist in the identification of plasmid-expressing cells and thus generation of cell lines stably expressing shRNAs.

However, an advantage of directly using chemically synthesised siRNAs for suppression studies is the fact that they may be chemically modified in order to improve their effectiveness. Such modifications may increase the RNAi molecule's stability, prevent specific off-target effects by blocking binding to unintended target transcripts, and reduce general off-target effects by preventing potential immune stimulatory effects (Castanotto and Rossi, 2009). Increased *in vivo* resistance of siRNAs to ribonucleases has been achieved by replacement of all 2'-OH residues on the RNA with 2'-fluoro, 2'-O-Methyl or 2'-deoxy residues (Morrissey *et al.*, 2005a;

Morrissey *et al.*, 2005b). Incorporation of 2-O-Methyl residues into siRNAs has been found to inhibit Toll-like receptor (TLR)-mediated immune stimulation which is often seen following delivery of unmodified siRNAs, without a reduction in gene silencing efficacy (Robbins *et al.*, 2007). Specific off-target effects mediated by siRNAs have also been reduced by chemical modifications which have been found to decrease or block the activity of the siRNAs sense strand. RISC-loading of the sense strand may be reduced by blocking 5'-phosphorylation by methylation (Chen *et al.*, 2008), while incorporation of DNA nucleotides at the 3' end of the sense strand had been found to significantly reduce sense strand-mediated off-target effects (Ui-Tei *et al.*, 2008). Other siRNA modifications have included addition of lauric acid, lithocholic acid and cholesterol derivatives, which has resulted in increased cellular uptake of siRNAs (Lorenz *et al.*, 2004). Other chemical modifications of siRNAs which have resulted in increased silencing efficiency of the target mRNA compared to unmodified equivalents without increasing toxicity have included incorporation of two 4'-C-hydroxymethyl-DNA (HM) residues at the 3' end of the sense strand together with two locked nucleic acid (LNA) residues (nucleosides in which the ribose ring is "locked" by a methylene bridge connecting the 2'-O atom and the 4'-C atom) at the 3' end and one HM residue towards the 5' end of the antisense strand, all of which are thought to result in favoured antisense incorporation into the RISC (Bramsen *et al.*, 2009). Increased gene silencing activity relative to unmodified siRNAs has also been seen upon the introduction of boron atoms in place of one of the non-bridging oxygen atoms of the siRNA residues (Hall *et al.*, 2004; Corey, 2007).

The aim of the study described in this chapter is the testing of chemically synthesised siRNA molecules designed to target the COL7A1 gene and analysis of their effects *in vitro*. Whilst approximately 50% suppression of COL7A1 mRNA expression was achieved using artificial microRNAs as outlined in chapter 3, it was of interest to examine whether or not higher suppression of the COL7A1 transcript could be achievable using COL7A1-targeting siRNAs. Both COL7A1-expressing human keratinocyte (HaCaT) and primary fibroblast (Normal Human Dermal Fibroblasts) cell lines were utilised in this investigation. An analysis of potential off-target effects elicited by COL7A1-targeting siRNA was also carried out, in which the effects of the most potent siRNA on the expression of a range of extracellular matrix and adhesion molecules were measured. In addition, it was of interest to generate and analyse the

effects of shRNA and artificial microRNA expression vectors which were designed based on the target sequences of those siRNAs which were found from this study to be the most potent suppressors of the human COL7A1 gene and to ascertain their relative potencies of suppression in comparison to the siRNAs and to the constructs tested in chapter 3.

4.2 Materials and Methods

4.2.1 siRNA design

4 siRNA molecules (Qiagen) designed to target the human COL7A1 gene were tested. Following input of the COL7A1 RefSeq ID (NM_000094) to the online program GeneGlobe Web portal (Qiagen), 4 FlexiTube GeneSolution siRNAs were recommended. All 4 COL7A1-targeting siRNAs (siRNAs 1-4), as well as a tested and validated AllStars negative control siRNA (NT siRNA; Qiagen) with no homology to any known mammalian gene, were provided annealed (see figure 4.1 for 21 nucleotide target sequences). siRNAs were resuspended in nuclease-free H₂O to a concentration of 10 μ M.

4.2.2 HaCaT siRNA transfection

Human keratinocyte HaCaT cells were transfected with siRNAs using the HiPerFect Transfection Reagent (Qiagen) as per section 2.2.2.6. HiPerFect is a blend of neutral and cationic lipid molecules which allow for effective uptake of siRNAs and their efficient release once inside cells (HiPerFect Transfection Reagent Handbook, 2008). 1x10⁵ cells were seeded in each well of a 24-well plate and 5nM of siRNA (siRNA 1-4 and NT siRNA) was added along with 3 μ l transfection reagent (section 2.2.2.6). Each siRNA was transfected in duplicate, with the experiment being carried out four times. Transfected cells were incubated under normal growth conditions for 24 hours before isolation of RNA from cells. A pool of all four siRNAs was also tested whereby 5nM of all 4 COL7A1-targeting siRNAs (1.25nM of each siRNA) were transfected into a single well of a 24-well plate, in quadruplicate. RNA was isolated from all 4 wells, as well as 4 wells of 5nM NT siRNA after 24 hours. An experiment was also undertaken in which a range of concentrations of siRNA were tested. siRNA 3 and NT siRNA were transfected in duplicate at concentrations of 1nM, 5nM, 10nM and 20nM and RNA was isolated from cells 24 hours later. Additionally, a time-point assay was carried out in which siRNA 3 and NT siRNA were transfected in sextuplicate and RNA was isolated from duplicate sample wells after 24, 48 and 72 hours.

4.2.3 Fibroblast siRNA transfection

Normal Human Dermal Fibroblast (NHDF) cells were transfected with siRNAs using the HiPerFect Transfection Reagent (Qiagen) as per section 2.2.2.6. 6×10^4 cells were seeded in each well of a 24-well plate and 5nM siRNA (siRNA 1-4 and NT siRNA) was added along with 3 μ l transfection reagent (section 2.2.2.6). Each siRNA was transfected in duplicate, with the experiment being carried out four times. Transfected cells were incubated under normal growth conditions for 24 hours before isolation of RNA from cells. A pool of all four siRNAs was also tested whereby 5nM of all 4 COL7A1-targeting siRNAs (1.25nM of each siRNA) were transfected into a single well of a 24-well plate, in quadruplicate. RNA was isolated from all 4 wells, as well as 4 wells of 5nM NT siRNA after 24 hours. An experiment was also undertaken in which a range of concentrations of siRNA were tested. siRNA 3 and NT siRNA were transfected in duplicate at concentrations of 1nM, 5nM, 10nM and 20nM and RNA was isolated from cells 24 hours later. Additionally, a time-point assay was carried out in which siRNA 3 and NT siRNA were transfected in sextuplicate and RNA was isolated from duplicate sample wells after 24, 48 and 72 hours.

4.2.4 Analysis of RNA from siRNA-transfected cells

RNA from HaCaT and NHDF cells transfected with siRNAs using HiPerFect Transfection Reagent was extracted using the Qiagen RNeasy[®] Plus mini kit. 350 μ l Buffer RLT was added directly to each well of a 24-well plate and the RNeasy[®] Plus mini kit protocol was followed (section 2.2.3.1). The concentration and purity of RNA samples was determined using the NanoDrop[™] 1000 Spectrophotometer (section 2.2.1.15). Real-time rtPCR was used to analyse expression levels of COL7A1 mRNA according to section 2.2.3.2 using COL7A1 primers previously available in the laboratory (section 2.1.3). Expression was normalised using the endogenous internal control gene Ubiquitin C (UBC) (section 2.1.3). The real-time rtPCR steps which were carried out are outlined in section 2.2.3.2 and the Cycle Threshold (CT) values were used for RNA quantification as outlined in section 2.2.3.3.

4.2.5 Statistical Analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

4.2.6 Protein analysis

Protein analysis of siRNA-treated cells was carried out by western blotting. For HaCaT cells, 1×10^7 cells were seeded in a T175 plate and cells were transfected with siRNA 3 at a concentration of 5nM, along with 300 μ l HiPerFect Transfection Reagent, according to section 2.2.2.6. A separate T175 plate was transfected with NT siRNA at a concentration of 5nM. For NHDF cells, 12×10^6 cells were seeded in a T175 plate and cells were transfected with siRNA 3 at a concentration of 5nM. A separate T175 plate was transfected with NT siRNA at a concentration of 5nM. For both cell lines, 24 hours later, cells at approximately 90% confluency were washed with PBS and incubated in serum-free DMEM for a further 24 hours. Media protein was then collected and treated, as described in section 2.2.4.1, and protein concentration was measured using a BCA Protein Assay Kit (section 2.2.4.4). Western blotting was then undertaken in order to measure any differences in collagen type VII protein levels between cells treated with COL7A1-targeting siRNA and those treated with non-targeting siRNA as described in section 2.2.4.6.

A 1 in 500 dilution of anti-collagen type VII rabbit polyclonal antibody (Calbiochem) was used to detect collagen type VII protein. An exposure time of approximately 5 seconds at room temperature was generally adequate for collagen type VII protein to be detected on chemiluminescence sensitive film. Following probing for collagen type VII protein, nitrocellulose membranes were incubated in Restore Western Blot Stripping Buffer (Thermo Scientific) with gentle agitation for 20 minutes, in order to remove primary and secondary antibodies from the membrane. Membranes were then washed 3x10 minutes in 1X TBS and incubated overnight in 5% marvel containing a 1 in 500 dilution of an anti-laminin 332 rabbit polyclonal antibody (Abcam). Membranes were then processed according to section 2.2.4.6 and exposure times of approximately 30 minutes were generally required for adequate laminin 332 detection. Transfection and western blot analysis was repeated on three separate occasions for each cell line.

4.2.7 Analysis of human extracellular matrix and adhesion molecule profiles in siRNA-treated cells

HaCaT cells were transfected with 5nM of either siRNA 3 or NT siRNA, as described in section 4.2.2. NHDF were transfected in 6-well plates with 5nM siRNA3 or NT

siRNA, at a density of 24×10^6 cells per well. Following extraction of RNA from cells 24 hours later using the Qiagen RNeasy[®] Plus mini kit as described in section 2.2.3.1, RNA was converted to cDNA using the RT² First Strand Kit (SABiosciences). Firstly, approximately 1µg RNA was treated with 5X gDNA Elimination Buffer in a total volume of 10µl, in order to remove any possible genomic DNA contamination. Samples were mixed gently by pipetting and incubated at 42°C for 5 minutes, followed by chilling on ice for at least one minute. An RT Cocktail was prepared, in a final volume of 10µl for each reaction, containing RT Buffer 3 (4µl), Primer & External Control Mix (1µl), RT Enzyme Mix 3 (2µl) and H₂O (3µl). The 10µl of RT Cocktail was subsequently added to the 10µl of genomic DNA Elimination mixture and mixed gently by pipetting. Samples were then incubated at 42°C for 15 minutes and the reaction was immediately stopped by heating at 95°C for 5 minutes. 91µl H₂O was then added to each 20µl of cDNA synthesis reaction and mixed well by pipetting. First Strand cDNA Synthesis reaction samples were stored at -20°C until use.

Real-time PCR was carried out on cDNA samples using the SABiosciences's RT² qPCR Master Mix. 102µl of diluted First Strand cDNA Synthesis reaction samples were mixed with 1350µl of 2X SABiosciences RT² qPCR Master Mix and made to a final volume of 2.7ml with H₂O. The master mix contains Sybr Green dye which, as previously described (section 2.2.3.2), binds to double-stranded DNA and fluoresces, allowing for a monitoring of the quantity of the gene of interest. The Experimental cocktail was then dispensed to an RT² PCR Array Loading Reservoir (SABiosciences) and 25µl of the cocktail was added to each well of the PCR Array plate (SABiosciences) using an eight-channel pipettor. The PCR array in question is the Human Extracellular Matrix & Adhesion Molecules RT² Profiler PCR Array, which is a set of optimized real-time PCR primer assays on a 96-well. It profiles the expression of 84 genes important for cell-cell and cell-matrix interactions, including the COL7A1 gene (see figure 4.2 and Appendix II). The remaining wells of the 96-well plate include a profile of the expression of 5 control housekeeper genes, a genomic DNA control, a reverse transcription control (in triplicate) and a positive PCR control (in triplicate). Following loading of the plate, the PCR array was tightly sealed with optical thin-wall 8-cap strips (SABiosciences) and centrifuged for 1 minute at room temperature at 1000rpm (Centrifuge 5810; Eppendorf). The plate was

then loaded onto the machine (Applied Biosystems 7300 Real Time PCR System) and the reaction steps were as follows:

Activation of DNA polymerase: 95°C for 10 minutes;

cDNA amplification: 37 cycles of 95°C for 15 seconds and 60°C for 1 minute;

Dissociation: 95°C for 15 seconds and 60°C for 15 seconds.

The data collected during the real-time PCR was presented as described in section 2.2.3.2. The SABiosciences's PCR Array Data Analysis Web Portal, which automatically performs calculations and provides an interpretation of the data upon inclusion of threshold cycle data from the real-time instrument, was used to analyse the data (<http://www.SABiosciences.com/pcrarraydataanalysis.php>). This web-based software package for the PCR Array System automatically performs all $\Delta\Delta C_t$ based fold-change calculations from the uploaded raw threshold cycle data. The average raw threshold cycle data of five housekeeping genes were used to normalise the PCR Array data. The genomic DNA control primer set incorporated into the array was used to specifically ensure the absence of any non-transcribed genomic DNA contamination. The reverse transcription control (RTC) wells were used to test the efficiency of the RT² First Strand Kit reaction with a primer set designed to detect template synthesised from the kit's built-in external RNA control: any impurities which affect the reverse transcription of the external RNA control would also affect the reverse transcription of experimental RNA. A positive PCR control (PPC) was also included to test the efficiency of the polymerase chain reaction itself using a pre-dispensed artificial DNA sequence and the primer set that detects it, as any impurities which affect the positive control PCR amplification will also affect amplification of the gene-specific products of interest. Furthermore, the two sets of replicate control wells (RTC and PPC) on the PCR Array also test for inter-well, intra-plate consistency. In total, 6 PCR arrays were analysed for each cell line. 3 of these arrays involved analysis of cDNA converted from the RNA isolated from cells treated with siRNA 3, and 3 of which involved analysis of cDNA converted from the RNA isolated from cells treated with NT siRNA.

4.2.8 shRNA vector construction and RNA analysis

shRNA constructs were generated using the pSuper.gfp/neo vector, based on the target sequences of siRNAs 1, 3 and 4. shRNA oligonucleotides were designed in the same manner as those described in chapter 3 (section 3.2.2.1; see figure 4.3 for sequences). shRNA vector construction was carried out in the same manner as described in section 3.2.2.2 and HaCaT and NHDF cells were nucleofected with all three shRNA molecules, as well as the non-targeting shRNA, as described in section 2.2.2.4. HaCaT cells were resuspended in Nucleofector™ Solution V to a concentration of 1×10^6 cells/100 μ l and along with 5 μ g DNA, cells were transfected using the U-20 program on the nucleofection machine. NHDF cells were resuspended in the Human Dermal Fibroblast Nucleofector® Solution to a concentration of 5×10^5 cells/100 μ l and cells were transfected with 5 μ g DNA using the U-023 program on the nucleofection machine. 24 hours after nucleofection, cells were collected and FAC sorting was carried out, as described in section 2.2.2.8. Transfection of each construct was undertaken in duplicate in 6-well plates and for HaCaT cells the experiment was repeated 3 times. RNA was extracted from FAC sorted cells and analysed as per sections 2.2.3.1, 2.2.3.2 and 2.2.3.3. Expression was normalised using the endogenous internal control genes Ubiquitin C (UBC) and TATA box binding protein (TBP).

4.2.9 Artificial microRNA vector construction and RNA analysis

Artificial miRNA constructs were generated using the pcDNA 6.2-GW/EmGFP-miR vector, based on the COL7A1 target sequences of siRNAs 1, 3 and 4. Artificial miRNA oligonucleotides were based on the same design as described in section 3.2.1.1 (see figure 4.4 for sequences) and vector construction was carried out in the same manner as described in section 3.2.1.2. HaCaT cells were resuspended in Nucleofector™ Solution V to a concentration of 1×10^6 cells/100 μ l and cells were transfected with 5 μ g DNA using the U-20 program on the nucleofection machine. Cells were then incubated under normal growth conditions and 24 hours later, cells were collected and FAC sorting was carried out, as described in section 2.2.2.8. Transfection of each artificial miRNA construct, as well as the non-targeting artificial miRNA (see figure 3.3 for sequence) was undertaken in duplicate in 6-well plates and the experiment was repeated twice. RNA was extracted from FAC sorted cells and analysed as per sections 2.2.3.1, 2.2.3.2 and 2.2.3.3. Expression was normalised using

the endogenous internal control genes Ubiquitin C (UBC) and TATA box binding protein (TBP).

4.2.10 Statistical analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

4.3 Results

4.3.1 Analysis of COL7A1 mRNA suppression by siRNAs in keratinocytes

Given that despite the guarantee provided by Invitrogen regarding potent suppression using the online BLOCK-iT™ Pol II miR RNAi Designer program, COL7A1 suppression studies using artificial microRNA vectors as described in chapter 3 of this Ph.D. thesis failed to result in a greater than 50% suppression level of COL7A1 mRNA. Thus it was decided to test synthesised siRNA molecules targeting COL7A1 and the ability of these molecules to suppress collagen type VII production *in vitro*. Predesigned and validated siRNAs for the target gene were chosen and ordered from Qiagen, a company which claims to provide at least 70% gene knockdown of target sequences. HaCaT cells were transfected using the HiPerFect Transfection Reagent (Qiagen), using 1×10^5 cells per well of a 24 well plate (each well being approximately 2cm^2) as described in section 4.2.2. Four siRNA molecules were tested (termed siRNAs 1, 2, 3 and 4), as well as a non-targeting Allstars Negative Control siRNA (Qiagen), which has no homology to any known mammalian gene and is incorporated into the RISC. Each siRNA transfection involved addition of siRNA at a concentration of 5nM into a single well, with duplicate wells for each siRNA. Each transfection experiment was carried out on four separate occasions. Following transfection, cells were incubated with the transfection complexes under their normal growth conditions and after 24 hours, RNA was isolated directly from the wells. siRNA gene silencing effects were monitored by real-time rtPCR and COL7A1 expression was normalised using expression levels of the Ubiquitin C (UBC) endogenous control gene as described in section 4.2.4.

COL7A1 expression levels were found to be statistically significantly different in cells transfected with each of the siRNA molecules in comparison with those treated with the negative control ($p < 0.0005$ for siRNA 1, siRNA 2, siRNA 3 and siRNA 4; figure 4.5). siRNA 1, siRNA 3 and siRNA 4 appeared to be the optimal suppressors, with mean COL7A1 suppression levels of approximately 65% and 68% being achieved with siRNA 1 and siRNA 4 respectively. Up to 83% suppression was achieved with siRNA 3 and the estimated mean COL7A1 mRNA suppression obtained was 71% (figure 4.5). A pool of all 4 siRNAs (total siRNA concentration of 5nM) was transfected in quadruplicate in a 24-well plate in order to ascertain whether

or not the simultaneous transfection of all four synthesised siRNAs might possibly result in increased levels of COL7A1 suppression in comparison to those transfections using individual siRNAs. 24 hours after transfection, RNA was isolated and compared to RNA from cells treated with 5nM NT siRNA. Statistically significant mean COL7A1 mRNA suppression of 69% was achieved using the siRNA pool compared to the non-targeting control ($p < 0.0005$), which does not appear to be significantly different to the levels of COL7A1 suppression achieved using individual suppressor molecules (figure 4.6). A range of concentrations of one of the optimal suppressors siRNA 3 were also tested in an effort to measure whether or not treatment of cells with reduced or increasing amounts of siRNA would lead to any changes in potency of suppression. siRNA 3 and NT siRNA were transfected in duplicate at concentrations of 1nM, 5nM, 10nM and 20nM. RNA was isolated from cells 24 hours later and following analysis of COL7A1 mRNA levels, it was found that variation of siRNA concentration did not seem to significantly alter siRNA gene silencing efficiency, with a tight range of COL7A1 suppression of between 46% and 49% being achieved in this single-transfection experiment (figure 4.7). It is of note that a reduction in suppression efficacy seems to have occurred during this particular experimental procedure and as greater suppression was consistently seen with this siRNA on other occasions, it is possible that lower transfection efficiency may have occurred during the undertaking of this particular experiment resulting in reduced levels of COL7A1 suppression compared to previously observed levels. No statistically significant differences in levels of suppression achieved were observed when comparing across concentrations, with the p-value when comparing the lowest concentration to the highest being 0.935. Additionally, a time-point assay was carried out in which siRNA 3 and NT siRNA were transfected in sextuplicate and RNA was isolated and analysed from duplicate sample wells after 24, 48 and 72 hours in order to analyse the time point at which optimal suppression is achieved (figure 4.8). Mean reductions in COL7A1 expression of 66% ($p = 0.006$), 59% ($p = 0.002$) and 61% ($p = 0.01$) were seen at 24, 48 and 72 hours respectively. Comparison among time points did not show statistically significant differences in levels of suppression ($p = 0.311$ when comparing the 24 hour and 48 hour time points; $p = 0.610$ when comparing the 24 hour and 72 hour time points and $p = 0.727$ when comparing the 48 hour and 72 hour time points).

The siRNA suppression data described above represents levels of suppression of COL7A1 RNA from a heterogeneous population of cells, due to the absence of a fluorescent marker which enabled performance of FAC sorting in the microRNA and shRNA experiments described in chapter 3 and thus isolation of a population of positively transfected cells. As the transfection efficiency of synthesised siRNA may not be 100%, the suppression data obtained represents a marked improvement in suppression of expression of COL7A1 with the siRNA molecules, in comparison with the previously tested artificial microRNAs and their equivalent shRNA constructs.

4.3.2 Analysis of COL7A1 mRNA suppression by siRNAs in fibroblasts

Following evaluation of siRNA activity in the human keratinocyte HaCaT cell line, it was subsequently decided to test these same constructs and ensure their potency in another collagen type VII-expressing cell line: Normal Human Dermal Fibroblasts (NHDFs). NHDF cells were transfected with siRNAs 1-4 as well as the negative control NT siRNA in the same manner as for HaCaT cells. Each siRNA transfection involved addition of 5nM of siRNA into a single well containing 6×10^4 cells, with duplicate wells for each siRNA. The transfection experiment was carried out on four separate occasions. Following transfection, cells were incubated with the transfection complexes under their normal growth conditions and after 24 hours, RNA was isolated directly from the wells. siRNA gene silencing effects were monitored by real-time rtPCR and COL7A1 expression was normalised using expression levels of the UBC endogenous control gene as described in section 4.2.4. COL7A1 expression levels were found to be statistically significantly different in NHDF cells transfected with each of the siRNA molecules in comparison with those treated with the negative control ($p < 0.0005$ for siRNA 1, siRNA 2, siRNA 3 and siRNA 4; figure 4.9). As was seen with siRNA-transfection of HaCaT cells (section 4.3.1), siRNA 1, siRNA 3 and siRNA 4 appeared to be the optimal COL7A1 suppressors. Mean approximate COL7A1 suppression levels of 74% and 76% were achieved with siRNA 1 and siRNA 5 respectively. Meanwhile, the highest reduction in COL7A1 expression was achieved with siRNA 3, with which mean COL7A1 mRNA suppression of 79% was found (figure 4.9).

NHDF cells were also transfected with a pool of all 4 siRNAs (total siRNA concentration of 5nM) in quadruplicate in a 24-well plate and 24 hours after

transfection, RNA was isolated and compared to RNA from cells treated with 5nM NT siRNA. Statistically significant mean COL7A1 mRNA suppression of 70% was achieved using the siRNA pool compared to the non-targeting control ($p<0.0005$), which does not appear to be significantly different to the levels of COL7A1 suppression achieved using individual suppressor molecules (figure 4.10). As for the analysis of siRNAs in HaCaT cells, a range of concentrations of one of the optimal suppressors siRNA 3 were also tested in NHDFs in an effort to measure whether or not treatment of cells with reduced or increasing amounts of siRNA would lead to any changes in potency of suppression. siRNA 3 and NT siRNA were transfected in duplicate at concentrations of 1nM, 5nM, 10nM and 20nM. RNA was isolated from cells 24 hours later and COL7A1 mRNA levels were analysed. An increase in siRNA concentration did not seem to correlate with an increase in COL7A1 mRNA suppression, as COL7A1 levels were reduced by approximately 63% ($p=0.008$), 77% ($p<0.0005$), 65% ($p=0.002$) and 80% ($p=0.002$) with siRNA 3 transfected at concentrations of 1nM, 5nM, 10nM and 20nM respectively (figure 4.11). Analysis of the data when comparing across concentrations found no statistically significant differences in levels of suppression achieved, with the p-value when comparing the lowest concentration to the highest concentration being 0.130. It is entirely possible however that an increase in the sample number in this case may result in statistical significance. Finally, a time-point assay was carried out in NHDF cells in which siRNA 3 and NT siRNA were transfected in sextuplicate and RNA was isolated and analysed from duplicate sample wells after 24, 48 and 72 hours in order to analyse the time point at which optimal suppression is achieved. Mean reductions in COL7A1 expression of 71% ($p<0.0005$) and 70% ($p=0.005$) were seen at 24 and 48 hours respectively while mean suppression increased to 82% ($p=0.003$) following 72 hours of post-transfection incubation (figure 4.12). Whilst it appears that there is an increase of suppression following incubation of transfected cells for 72 hours, a comparison of the data between time points found no statistical significance between time points ($p=0.839$ when comparing the 24 hour and 48 hour time points; $p=0.135$ when comparing the 24 hour and 72 hour time points and $p=0.206$ when comparing the 48 hour and 72 hour time points).

4.3.3 Analysis of collagen type VII protein suppression by siRNA 3

Following significant suppression of COL7A1 mRNA in both human keratinocyte and human fibroblast cell lines using synthesised siRNAs, validation of this RNA knockdown at the protein level was attempted by western blot analysis. As a large number of cells is required for adequate detection of collagen type VII protein levels, 10^7 HaCaT cells were seeded in tissue culture plates of 175cm². For NHDF cells, 6×10^6 cells were seeded in T175 plates. Cells were transfected with the COL7A1 mRNA suppressor siRNA 3 or the negative control NT siRNA at a concentration of 5nM and 48 hours later, serum-free media was collected and conditioned as described in section 4.2.6. Collagen VII protein was detected using an anti-collagen type VII rabbit polyclonal antibody (Calbiochem). Laminin 332 protein was used as a loading control and following probing of the membranes for collagen type VII, membranes were incubated overnight in a rabbit polyclonal antibody to Laminin 332 (Abcam) and probed the following day (section 4.2.6). Transfection and western blot analysis was repeated on three separate occasions for both HaCaT and NHDF cell lines.

Western blot analysis on conditioned media from siRNA 3- and NT siRNA-treated HaCaT cells showed significant suppression of collagen type VII protein. While a distinct band of protein was visible in the lane in which NT-treated cell media was loaded which corresponds to a size of approximately 290kDa according to the pre-stained protein standards, only a very faint band of this size was visible in the lane in which siRNA 3-treated cell media was loaded. Bands corresponding to the 165kDa Laminin 332 $\alpha 3$ chain and the 145kDa Laminin 332 $\beta 3$ chain appear to be of equal density in both lanes, suggesting loading of equal amounts of protein in each well (figure 4.13). Thus siRNA 3 appears to effectively suppress COL7A1 at both the mRNA and protein levels. For NHDF cells, while a distinct protein band of approximately 290kDa was detected in the NT siRNA-treated cell media lane, this band was relatively faint in comparison to that detected from HaCaT cells media protein. No band was detected in lanes in which media from siRNA 3-treated NHDF cells was loaded. Furthermore, an adequate signal of Laminin 332 protein was failed to be detected in any of the NHDF media protein samples and so equal loading of all protein samples during the SDS-PAGE experiment could not be demonstrated (figure 4.14). This western blot analysis was attempted on three separate occasions and the results obtained were despite the testing of the exposure of membranes to the

chemiluminescence sensitive film overnight in an attempt to detect the presence of even a low concentration of protein. Thus it is possible that the levels of Laminin 332 expressed by this cell line are too low for detection during western blot analysis.

4.3.4 Analysis of human extracellular matrix and adhesion molecule profiles in siRNA 3-treated cells

Analysis of the expression profiles of a range of extracellular matrix and adhesion molecules was undertaken in both the HaCaT and NHDF cell lines using the SABiosciences Human Extracellular Matrix (ECM) & Adhesion Molecule RT² Profiler PCR Array according to section 4.2.7. Using real-time PCR, the array in question allows for analysis of expression of a selection of 84 genes related to cell adhesion and the ECM. The array analyses expression of a range of ECM proteins including basement membrane constituents, collagens, and other genes defining the structure of the ECM; matrix and other metalloproteinases that remodel the ECM as well as their inhibitors; and integrins, selectins, cell-adhesion molecule family members, and other genes important to cell adhesion and cytoskeleton bridging such as the catenins. Among the genes in question which are profiled in this array is COL7A1. For both HaCaT and NHDF cells, following siRNA 3- or NT siRNA-transfection as described in section 4.2.7, RNA was isolated and 1µg RNA from each sample was converted to cDNA as described in section 4.2.7. Following loading of samples onto the 96-well PCR array plates, real-time PCR was carried out and data was analysed as described in section 4.2.7. For both HaCaT and NHDF cells, 3 plates of cDNA from siRNA 3-treated cells and 3 plates of cDNA from NT-treated cells were analysed and compared in order to measure the effects of siRNA 3 on a range of ECM and related proteins.

In HaCaT cells, analysis of the expression profiles of the range of 84 ECM and adhesion molecule-encoding genes in siRNA 3-treated cells compared to NT siRNA-treated cells resulted in some variations (figures 4.15 and 4.16). As expected, COL7A1 was significantly reduced in cDNA from siRNA 3-treated cells by a mean level of 54% ($p=0.018$). While the vast majority of the 84 genes examined were unaffected, expression levels of a small number of genes were found to be altered in siRNA 3-treated cells. The COL1A1 gene, which encodes type I collagen and functions in provision of the strength characteristic of healthy bone (Solomon *et al.*,

1984; Di Lullo *et al.*, 2002), was also found to be decreased in siRNA 3-treated cells by approximately 41% ($p=0.046$). Statistically significant differences in expression were also found with the LAMA2 and LAMB1 genes ($p=0.033$ in both cases), with LAMA2 being up-regulated by approximately 32% and LAMB1 being down-regulated by approximately 18% compared to the non-targeting control siRNA. LAMA2 and LAMB1 encode laminin alpha 2 and laminin beta 1 proteins respectively, which encode subunits of the large multidomain laminin proteins which function in the development and maintenance of cellular organization and supramolecular structure (Engel, 1992).

In NHDF cells, variations were seen in the expression levels of a number of genes in siRNA 3-treated compared to NT siRNA-treated cells, some of which differed from the expression of similarly treated HaCaTs (figures 4.17 and 4.18). As was observed with cDNA analysed from treated HaCaT cells, both the COL7A1 and LAMA2 genes were statistically significantly altered in cDNA from siRNA 3-treated NHDFs compared to those treated with the negative control. COL7A1 expression was reduced by approximately 82% ($p=0.0006$), while LAMA2 was over expressed by approximately 18% ($p=0.0405$). The differences in expression of the COL1A1 and LAMB1 genes between treated groups which were observed in HaCaTs were not observed in NHDF cells; expression profiles of both genes were unaltered in siRNA 3-treated and NT siRNA-treated samples from NHDF cells. However, variations in a number of different genes were observed in the fibroblast cell line. The COL11A1 gene, which encodes collagen type XI protein and is required for normal skeletal morphogenesis in humans (Li *et al.*, 1995), was down-regulated by approximately 56% in siRNA 3-treated NHDFs ($p=0.001$). ITGA8, which encodes a member of the integrin family of adhesion receptors (Schnapp *et al.*, 1995), was found to be down-regulated by 20% ($p=0.04989$) and the SPP1 gene, which acts as a ligand for an integrin adhesion receptor (Bayless *et al.*, 1998), was down-regulated by 12% ($p=0.0397$) in cDNA from siRNA 3-treated fibroblasts. Furthermore, the MMP12 gene was statistically significantly up-regulated by approximately 27 fold in siRNA 3-treated samples compared to NT-treated samples ($p=0.033096$). MMP12 encodes a matrix metalloproteinase, which although the function of which has not yet been fully characterised, plays a role in the binding and degradation of collagens (Taddese *et al.*, 2010).

4.3.5 shRNA vector construction and RNA analysis

Following potent suppression of the COL7A1 gene at both the mRNA and protein levels with synthesised siRNAs, it was decided in parallel to the work being undertaken analysing ECM gene expression profiles to generate vectors expressing these siRNAs as short hairpin RNAs (shRNAs). Such shRNA constructs offer possible advantages over the direct delivery of siRNAs in terms of delivery options and suppression longevity, as described in section 4.1. In addition, shRNAs may hold some advantages in terms of minimising potential stimulation of innate immunity pathways such as the Toll-like receptor 3 (TLR3) pathway (Kleinman *et al.*, 2008; Cho *et al.*, 2009). shRNAs were designed according to the OligoEngine pSuper RNAi system, which provides a mammalian expression vector that directs intracellular synthesis of siRNA-like transcripts. Three shRNAs were designed and generated, based on the COL7A1 target sequences of the three optimal siRNAs 1, 3 and 4 (section 4.2.8). HaCaTs and NHDFs were nucleofected with shRNAs 1, 3 and 4 as well as a negative control non-targeting shRNA construct and 24 hours post-transfection, cells were FAC-sorted based on eGFP fluorescence given the presence of the eGFP gene in the shRNA plasmid constructs. RNA from shRNA-expressing FACS-purified cells was isolated and real time rtPCR was used to analyse expression levels of COL7A1 mRNA as described in section 4.2.8.

Statistically significant suppression of COL7A1 mRNA was found to have been achieved following analysis by real-time rtPCR of RNA from shRNA 1, 3 and 4-treated HaCaT cells compared to NT shRNA-treated cells ($p < 0.0005$ in all cases). However, suppression levels differed from those achieved following direct delivery of synthesised siRNAs to HaCaTs. shRNA 1 appeared to be the optimal shRNA suppressor, with a mean reduction of COL7A1 levels of approximately 73%. Mean COL7A1 suppression of approximately 57% was achieved with shRNA 3 and mean suppression of approximately 49% was achieved with shRNA 4 (figure 4.19). Validation of this RNAi-mediated suppression data by analysis of protein levels was not carried out due to the nature of the experimental model required, as detailed in section 3.3.3. Difficulties were encountered when attempting to validate the mRNA suppression data obtained with artificial microRNA 4 tested in chapter 3 of this thesis by western blot analysis, and both stable cell lines and mass transfection experiments were undertaken, in a vain attempt to achieve this aim. Thus, as siRNA-mediated

COL7A1 suppression potency at both the mRNA and protein levels had already been shown in this cell line (figures 4.5 and 4.13), attempted western blot analysis of COL7A1-targeting shRNA-transfected cells was deemed not to be worthwhile. Furthermore, analysis of the effects of the shRNAs on COL7A1 mRNA levels in NHDFs was not achievable, due to the low level of transfection efficiency following shRNA nucleofection. Only a small percentage (<1%) of cells were found to be expressing the suppressor constructs as seen during the FACS process, and so a sufficient number of cells from which RNA could be extracted and analysed was not obtainable despite three attempts. This was somewhat surprising as although primary cell lines such as NHDFs have been reported as being somewhat difficult to transfect (Ovcharenko *et al.*, 2005; Jordan *et al.*, 2008), potent COL7A1 knockdown was achieved following siRNA delivery to NHDF cells (figure 4.9) suggesting high transfection efficiency, as well as claims by Lonza that up to 70% transfection efficiency following plasmid electroporation of NHDF cells may be achieved.

4.3.6 Artificial microRNA vector construction and RNA analysis

In recent years, an artificial microRNA approach has been employed in some RNAi suppression studies, whereby shRNAs have been designed based upon a microRNA scaffold. This is thought to perhaps represent a system of increased safety due to the fact that it is modeled on the cells own endogenous microRNA machinery, as described in chapter 3 of this thesis. Following testing of thirteen artificial microRNA vectors designed to target the COL7A1 gene with which suppression of no greater than 50% was achieved, it was decided to generate artificial microRNAs based on the target sequences of siRNAs 1, 3 and 4 and to determine whether or not the higher levels of suppression obtained with siRNAs was achievable with this artificial microRNA vector format. Artificial microRNA constructs were designed and generated using the pcDNA 6.2-GW/EmGFP-miR system, as described in section 4.2.9. HaCaT cells were transfected with artificial microRNAs 1, 3 and 4 as well as a non-targeting control artificial microRNA by nucleofection and following FAC sorting 24 hours later, RNA was isolated and analysed as described in section 4.2.9.

Statistically significant COL7A1 mRNA suppression failed to be achieved with artificial microRNAs 1 and 3 compared to the negative control construct ($p=0.291$ and $p=0.121$ respectively). Of the three constructs tested, artificial

microRNA 4 was found to be the optimal suppressor, with which a mean reduction of COL7A1 mRNA of 51% was achieved ($p < 0.005$) (figure 4.20). Thus while suppression levels of COL7A1 obtained with shRNA 4 and artificial microRNA 4 were of approximate equal levels (49% and 51% respectively), suppression of the COL7A1 transcript in the target regions of siRNAs 1 and 3 were significantly reduced when the suppressors were in artificial microRNA format.

4.4 Discussion

Following advances in the field of RNAi in recent decades, the paradigm of which offers the potential to suppress translation of any of the thousands of genes in the human body, the focus of the work described in this thesis has centred on attempted repression of expression of the COL7A1 gene. The ability to achieve such suppression would possibly result in the overcoming of one obstacle in the pursuit of a potential therapeutic for the treatment of those suffering from the dominant form of Dystrophic Epidermolysis Bullosa, due to the dominant negative effects of the disease-causing mutations in question. Following results of previous suppression studies which failed to result in a greater than 2-fold decrease in COL7A1 mRNA using artificial microRNA-based suppressor constructs (chapter 3 of this Ph.D. thesis), it was hypothesised that due to the nature of the COL7A1 transcript the structure may possibly not be optimally amenable to RNAi-mediated suppression: the COL7A1 transcript has quite a high GC content and is large, with a length of approximately 9kb (Christiano *et al.*, 1994b). Consequently it was subsequently decided to test the activity of synthetic siRNAs *in vitro* and the ability of these molecules to suppress collagen type VII production. These studies found that alternatively, simply additional studies were required to characterise a potent RNAi-based COL7A1 suppressor.

Four siRNA molecules were initially tested in HaCaT cells and delivered using the lipid-based HiPerFect Transfection Reagent. Significant suppression of COL7A1 mRNA was attained with all four siRNAs tested, with mean levels of approximately 71% suppression being achieved with the optimal suppressor, siRNA 3 (figure 4.5). Western blot analysis was carried out on cells transfected with what was found to be one of the most effective siRNAs, siRNA 3, and cells transfected with a non-targeting siRNA. After previous attempts to adequately show suppression at the protein level were unsuccessful during artificial microRNA-based COL7A1 suppression analysis, siRNA-mediated collagen type VII protein suppression was successfully demonstrated, likely due mainly to the increase in suppression levels obtainable with the siRNAs compared to the artificial microRNAs. This was supported by successful probing of laminin 332 protein for use as a loading control protein (figure 4.13). The result of this protein study appears to correlate with the

previous RNA data, and thus strengthens the potential of these prospective therapeutic molecules as a means to elicit COL7A1 suppression. It is also of note that whilst the half-life of collagen VII protein, although not fully determined, is thought to be greater than 2 months *in vivo* (Kern *et al.*, 2009; Remington *et al.*, 2009), a long *in vitro* half-life is not likely to affect the results obtained during western blot analysis of collagen VII levels in the studies described in this Ph.D thesis due to the nature of the experimental procedure carried out. As stated, collagen VII protein is secreted into the media in which the cells are incubated and this secretion process seems to be quite rapid, seen by the low signals obtained from analysis of intracellular collagen VII levels compared to secreted levels in reported studies such as those of Saito *et al.*, 2009. In the studies described in this thesis, 24 hours post siRNA transfection, this media was removed and replaced with fresh media, thus indicating that any collagen VII protein already produced and secreted or in the process of being secreted prior to transfection would be removed and thus should not interfere with the suppression results obtained.

Hitherto, testing of all RNAi molecules was carried out *in vitro*, in the human epidermal keratinocyte (HaCaT) cell line. As discussed in section 3.4, the fact that these cells undergo differentiation upon growth to confluency and express genes which are specific to the upper layers of healthy skin is just one property which makes them an ideal candidate for use in human epidermal keratinocyte studies (Boukamp *et al.*, 1988). However, the skin is composed of multiple cell types and so analysis of the effects of these RNAi molecules in other cells types would be ideal. Another potential target skin cell for gene therapy is fibroblasts, the principal cells of connective tissue in the human body. Their main function is to provide collagen to areas of the skin which have been wounded. Upon injury of the skin or another organ, fibroblasts multiply and migrate to the area of the wound, where they deposit collagen to promote healing (Schäffer *et al.*, 1997). Therefore, Normal Human Dermal Fibroblasts (NHDFs) were transfected with all four siRNA molecules which were previously tested in HaCaT cells, using the HiPerFect Transfection Reagent. Any drastically reduced potency of these suppressors in this skin cell line would likely rule out the use of the siRNAs in any potential therapeutic for dominant dystrophic EB.

Notably, siRNA-mediated suppression of COL7A1 mRNA was also achieved in the fibroblast cell line. Moreover, a slight increase in levels of suppression of COL7A1 was seen for each siRNA compared with previously observed suppression data in HaCaT cells, although a comparison of the effects of each siRNA between the two cell lines found this data not to be statistically significant ($p>0.05$ for each siRNA). An average of 79% suppression was achieved with the most effective siRNA molecule, siRNA 3 (figure 4.9). These slight differences in suppression of the COL7A1 transcript between the HaCaT and NHDF cell lines may perhaps be due to differences in expression patterns between the two cell types or differences in the RNAi machinery between the two cell lines. Furthermore, transfection efficiency may have been higher in the fibroblast cell line compared with the HaCaT cell line, and so if a larger proportion of cells were undergoing RNAi-mediated suppression of COL7A1 expression in fibroblasts, the overall levels of suppression achieved from a transfected population of cells would likely be noticeably higher. These data represent further evidence of the potential of these RNAi molecules for further exploration in a gene therapy-directed dominant dystrophic EB therapeutic approach. The data also refutes the possibility that the COL7A1 transcript may not be readily amenable to high levels of suppression by the process of RNAi, as was originally speculated following testing of thirteen artificial microRNAs, none of which achieved higher than 50% suppression of COL7A1. Results from western blot analysis of siRNA-mediated COL7A1 suppression in NHDF cells were not as clear as those obtained in the HaCaT cell line (figure 4.14). The collagen VII protein band obtained during western blot analysis of NHDFs was significantly fainter than that obtained from HaCaTs, suggesting less collagen VII protein is present in these samples. The likely reason for this is due to the amount of cells which are utilized in the experiment. As the fibroblast cells are larger in size than HaCaTs and thus a smaller number of cells fill a 175cm² tissue culture plate (1×10^7 HaCaTs compared to 6×10^6 NHDFs), it is more difficult to pick up a collagen VII protein signal from the fibroblast cells. As the 175cm²-sized tissue culture plate was also the maximum sized plate available and 15ml was the maximum amount of media which could be filtered using the Amicon® Ultra centrifugal filter devices, it was not possible to attempt this experiment using a larger number of NHDF cells. This also may possibly explain the reason for lack of detection of the laminin 332 loading control protein in the NHDF cell line during western blot analysis. It is likely that a larger number of cells is required for adequate

detection of this protein in this cell line and thus for its use as a loading control protein when measuring changes in collagen type VII protein levels. It is worth noting that it was observed during analysis of the ECM gene profiling array data, among the primer sets included in the array were those specific for both the laminin $\alpha 3$ and $\beta 3$ subunits of laminin 332, and the threshold cycle values for both genes were considerably higher in the NHDF data compared to the HaCaT data (Laminin $\alpha 3$: average C_T of 20.88 in HaCaTs and 25.16 in NHDFs; Laminin $\beta 3$: average C_T of 18.69 in HaCaTs and 25.67 in NHDFs). This suggests that both genes are expressed at lower levels in the NHDF cell line, which may thus facilitate our understanding of why the protein could not be adequately detected in media samples from this particular cell line.

While each siRNA was transfected *in vitro* at a concentration of 5nM and their effects on COL7A1 mRNA analysed 24 hours later, a number of different parameters were subsequently tested in an effort to identify the optimal conditions for the siRNAs to reach their maximum suppression potential. In both HaCaT and NHDF cells, all four siRNA molecules were pooled together and their potency tested, in order to examine whether or not the combination of suppressors used together would allow for an increase in COL7A1 suppression to those already seen with the individual molecules. However in both cell lines, the suppression levels obtained with the siRNA pool did not exceed those achieved with the most potent suppressors, with statistically significant 69% and 70% mean suppression of COL7A1 mRNA being achieved in HaCaTs and NHDFs respectively (figures 4.6 and 4.10 respectively). In this COL7A1 suppression study, one of the optimal suppressors, siRNA 3, was also tested at 24, 48 and 72 hours time points in an effort to identify the time point at which optimal mRNA knockdown is achieved. In tested HaCaT cells, mean COL7A1 suppression seemed to decrease slightly from 66% after 24 hours to 59% and 61% after 48 and 72 hours respectively (figure 4.8), although statistical analysis of the data was not significant ($p > 0.05$ in each case of time point comparison). In NHDFs however, an increase in suppression was seen at the 72 hour time point compared to 24 and 48 hours (82% suppression at 72 hours, compared to 71% and 70% suppression at 24 and 48 hours respectively) (figure 4.12). However, similar to the HaCaT data, these differences were not found to be statistically significant ($p > 0.05$ in all cases). Following testing of a range of siRNA concentrations from 1nM to 20nM, a

significant increase in suppression was not seen with treatment of cells with increased siRNA concentration in HaCaT or NHDF cell lines (figures 4.7 and 4.11). Furthermore, the data during comparison of concentrations was not statistically significant ($p > 0.05$ in all cases for both cell lines).

Some of the research described in chapter 4 of this thesis also involved an examination of the effects of one of the most potent siRNA suppressors compared to the negative control siRNA on expression of a range of structural genes which encode proteins involved in the extracellular matrix and adhesion of the skin layers, in both keratinocyte and fibroblast cells. Among those genes profiled is the collagen type VII-encoding COL7A1 gene, an important basement membrane zone constituent as described in chapter one. As expected, in both cell lines COL7A1 gene expression was statistically significantly reduced. While expression of the majority of the 84 genes profiled remained unaffected in both cell lines, there were a small number of genes whose expression were altered in cDNA from siRNA 3-treated cells compared to cDNA from those treated with NT siRNA. A number of genes were found to be down-regulated in siRNA 3-treated cells, including two collagen-encoding genes COL1A1 and COL11A1, although not to the same extent as down-regulation of COL7A1. It was hypothesised that these genes which are down-regulated may possibly bear some homology to the COL7A1 gene and thus the siRNA may be binding to their gene's transcript product, albeit at reduced efficiency compared to its binding to the COL7A1 transcript. Therefore, a homology analysis of the COL7A1 mRNA sequence with both the COL1A1 and COL11A1 transcripts was carried out using an online tool (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html; ClustalW2-Multiple Sequence Alignment). These sequence alignments found the COL7A1 and COL1A1 sequences to be just 38.2% similar, with 8 mismatches over the siRNA 3 target region. Similarly, the COL7A1 and COL11A1 sequences were just 36.7% similar, with 10 mismatches over the siRNA 3 target region. It is unlikely therefore that siRNA 3 would bind to these regions due to the low degree of complementarity between the siRNA and transcript.

Alternatively, these siRNA off-target effects may arise due to insufficient degradation of the siRNA sense strand, allowing it to have somewhat of an RNAi effect on the cells transcript population. This may also explain differences between

the expression profiles of certain genes in keratinocyte and fibroblast cell lines, as differences in the cells RNAi machinery may result in differences in processing of the siRNA upon delivery to cells. It is also possible that the significant down-regulation of COL7A1 as a result of RNAi-mediated suppression may cause a stress response in the cell due to the altered environment, causing alterations in gene expression levels. This may explain the dramatic increase (approximately 27 fold) in levels of the MMP12 gene in siRNA 3-treated NHDF cells. Furthermore, although the function of the MMP12 gene has not been fully elucidated, several MMP proteins have been found to function in the degradation of collagen proteins (Yeh *et al.*, 2010; Barascuk *et al.*, 2011). Therefore, it is possible that the decrease in COL7A1 mRNA levels directly causes a response in expression levels of the MMP12 gene, although the exact mechanism by which this may be occurring is not fully understood. As COL7A1 encodes a protein with a structural function, possible interactions between COL7A1 and other genes at the mRNA level have not been elucidated and may require further investigation. Overall, the relevance of these observations in terms of changes in gene expression profiles after treatment of HaCaT and NHDF cells with COL7A1-targeting siRNAs is somewhat difficult to interpret without substantial further experimentation. However it is beyond doubt that the majority of gene-based therapies will have effects on the expression profiles of multiple genes. Given the nature of RNAi-based drugs in terms of direct modulation of gene expression profiles, it may be that regulatory bodies such as the FDA (Food and Drug Administration) or EMA (European Medicines Agency) may require additional toxicology studies above those required for non-nucleotide based drugs prior to moving such molecules into human clinical trial.

In parallel with the research undertaken on evaluation of gene expression profiles post-treatment with COL7A1-targeting siRNAs, it was decided to generate shRNA vector versions of the COL7A1 siRNAs. Conversion of synthetic siRNAs to a system such that they may be expressed from vectors offers advantages over direct delivery of siRNAs, as described in section 4.1. While it is possible that direct delivery of siRNAs to the skin in a gene-based therapy for dominant Dystrophic EB may represent a more favourable approach, it was of interest to also generate constructs expressing the most potent suppressors in both shRNA and artificial microRNA formats. As sufficient delivery of constructs to the fibroblast cell line was

not achieved by the method of nucleofection as discussed in section 4.3.5, the two plasmid types were tested in HaCaT cells exclusively. Interestingly, in some cases the effectiveness of the suppressors seemed to be altered when converted to shRNA format compared to the effect of their equivalent siRNAs. Whilst shRNA 1 retained the potency of its equivalent siRNA by effectively reducing mean COL7A1 levels by approximately 73% (which was not statistically significantly different from the suppression achieved with the equivalent siRNA 1; $p=0.170$), the shRNAs designed to target the same transcript regions as siRNAs 3 and 4 were not as effective and the data comparing the two to their siRNA counterparts was statistically significant ($p=0.002$ for siRNA/shRNA 3 and $p=0.001$ for siRNA/shRNA 4). Suppression by siRNA 3 resulted in an approximate 71% reduction in COL7A1 levels, whilst the equivalent shRNA 3 reduced COL7A1 levels by approximately 57%. Similarly, siRNA 4 yielded suppression levels of approximately 68% compared to mean reductions of 49% achieved by shRNA 4. Thus it may be possible that the process of transcribing these two shRNA molecules in the nucleus of the cell followed by their export to the cytoplasm & subsequent processing to siRNAs may be a limiting factor in the effort to achieve high levels of suppression of the COL7A1 gene in the HaCaT cell line. Meanwhile, conversion of siRNAs to artificial microRNA format, similar to those tested in chapter 3 of this thesis, resulted in maximum suppression of just 51%, achieved by the construct based on siRNA 4. This is similar to the optimal suppression achieved by those constructs tested in chapter 3. However significant suppression of the COL7A1 gene was not achieved with either artificial microRNA 1 or 3 and so overall it appears that there is a trend towards reduced potency upon conversion of siRNAs to artificial microRNAs, based on the tested miR-155 format (all data comparing potency of siRNAs to their equivalent artificial microRNAs was statistically significant; $p=0.001$ for siRNA/artificial microRNA 1, $p<0.0005$ for siRNA/artificial microRNA 3 and $p=0.005$ for siRNA/artificial microRNA 4). Whilst there have been some reports of potent silencing achieved with artificial microRNAs based on mir-155 (Shan *et al.*, 2009a) including potent gene suppression using artificial microRNAs generated using the same BLOCK-iT™ RNAi Designer program which was used in the artificial microRNA studies described in this thesis (Shan *et al.*, 2009b; Wei *et al.*, 2010), the results generated during the course of this thesis are conflicting to these publications. The testing of a large number of COL7A1-targeting artificial microRNAs without achieving greater than 50% suppression

suggests that this particular vector is not ideal for achieving potent and high knockdown of the COL7A1 gene.

4.5 Conclusion

Synthetic siRNAs designed to reduce COL7A1 mRNA expression levels were tested in both keratinocyte and fibroblast cell lines. The results presented in this chapter of this PhD thesis demonstrate that statistically significant suppression of COL7A1 mRNA of up to 80% was achieved with the optimal siRNA. These molecules were found to be statistically significantly more effective at reducing COL7A1 expression than equivalent shRNA and artificial microRNAs and may potentially be suitable for use in a future gene therapy-based therapeutic for the dominant form of Dystrophic EB. As the siRNAs in question are designed to achieve mutation-independent knockdown of the COL7A1 gene, both wild-type and mutant genes would be targeted for degradation, the effects of which would likely be damaging due to the requirement of collagen VII protein for maintenance of the structural integrity of the skin. Thus it would be of interest to generate constructs which are resistant to the effects of the siRNA, whilst retaining the ability to express normal type VII collagen protein. The work associated with addressing this matter forms the basis of the study described in Chapter 5 of this PhD thesis.

Name	Target sequence 5'-3'
siRNA 1	CAGCCAATCCTTGAGAGTACA
siRNA 2	CCGGGCAGCAATGGTGACCAA
siRNA 3	CCGAATGGTGCTGCAGGCAAA
siRNA 4	CTCGATCGACTCGGTGACTTT

Figure 4.1 siRNA target sequences: 5'-3' sequences of the COL7A1 gene which are targeted by siRNAs 1-4.

ADAMTS1 A01	ADAMTS13 A02	ADAMTS8 A03	CD44 A04	CDH1 A05	CNTN1 A06	COL11A1 A07	COL12A1 A08	COL14A1 A09	COL15A1 A10	COL16A1 A11	COL1A1 A12
COL4A2 B01	COL5A1 B02	COL6A1 B03	COL6A2 B04	COL7A1 B05	COL8A1 B06	VCAN B07	CTGF B08	CTNNA1 B09	CTNNB1 B10	CTNND1 B11	CTNND2 B12
ECM1 C01	FN1 C02	HAS1 C03	ICAM1 C04	ITGA1 C05	ITGA2 C06	ITGA3 C07	ITGA4 C08	ITGA5 C09	ITGA6 C10	ITGA7 C11	ITGA8 C12
ITGAL D01	ITGAM D02	ITGAV D03	ITGB1 D04	ITGB2 D05	ITGB3 D06	ITGB4 D07	ITGB5 D08	KAL1 D09	LAMA1 D10	LAMA2 D11	LAMA3 D12
LAMB1 E01	LAMB3 E02	LAMC1 E03	MMP1 E04	MMP10 E05	MMP11 E06	MMP12 E07	MMP13 E08	MMP14 E09	MMP15 E10	MMP16 E11	MMP2 E12
MMP3 F01	MMP7 F02	MMP8 F03	MMP9 F04	NCAM1 F05	PECAM1 F06	SELE F07	SELL F08	SELP F09	SGCE F10	SPARC F11	SPG7 F12
SPP1 G01	TGFB1 G02	THBS1 G03	THBS2 G04	THBS3 G05	TIMP1 G06	TIMP2 G07	TIMP3 G08	CLEC3B G09	TNC G10	VCAM1 G11	VTN G12
B2M H01	HPRT1 H02	RPL13A H03	GAPDH H04	ACTB H05	HGDC H06	RTC H07	RTC H08	RTC H09	PPC H10	PPC H11	PPC H12

Figure 4.2 Layout of the 96-well ECM and adhesion molecule PCR array plate: Wells A01 to G12 contain primers specific to 84 different genes which encode Extracellular matrix (ECM) proteins (including constituents of the basement membrane, collagens, and other genes involved in defining the ECM structure) and other genes involved in cell adhesion and cytoskeleton bridging. Row H represents control wells, which contain primers specific to 5 housekeeping genes (wells H01 to H05), as well as human genomic contamination (well H06), reverse transcription (wells H07-H09) and positive PCR (wells H10-H12) controls. See appendix II for the description of all genes.

Name	Sequence
shRNA 1 Top 5'-3'	GAGTACAGGATCCCCAGCCAAATCCTTGAGAGTACATTCAAGAGATGTACTCTCAAGGATTGGCTGTTTTTAAAGCTTGAGTACA
shRNA 1 Bottom 3'-5'	CTCATGTCTAGGGGGTGGTTAGGAATCTCTCATGTAAGTTCTCTACATGAGAGTTCTTAACCGACAAAAATTCGAACTCATGT
shRNA 3 Top 5'-3'	GAGTACAGGATCCCCCGAATGGTGTGCAGGCAAAATCAAGAGATTTGCCTGCAGCACCATTCCGGTTTTTAAAGCTTGAGTACA
shRNA 3 Bottom 3'-5'	CTCATGTCTAGGGGGGCTTACCACGACGTCCGTTTAAAGTTCTCTAAACGGACGTCGTGTAAGCCAAAAATTCGAACTCATGT
shRNA 5 Top 5'-3'	GAGTACAGGATCCCCCTCGATCGACTCGGTGACTTTTTCAGAGAAAAAGTCAACCGAGTCGATCGAGTTTTTAAAGCTTGAGTACA
shRNA 5 Bottom 3'-5'	CTCATGTCTAGGGGGAGCTAGCTGAGCCACTGAAAAAGTTCTCTTTTTCAGTGGCTCAGCTAGCTCAAAAAATTCGAACTCATGT

Figure 4.3 shRNA sequences: DNA sequences of top and bottom strands of shRNA oligonucleotides. Red letters indicate *Bam*HI and *Hind*III restriction sites. Following annealing of oligos, digesting with *Bam*HI and *Hind*III results in 5' *Bgl*III and 3' *Hind*III overhangs (5'-GATC-3' and 3'-TCGA-5' respectively), which allows for cloning into the pSuper vector. Green letters indicate the hairpin. The first seven and last seven nucleotides of each oligo represent random sequences.

StrandName	Sequence 5'- 3'
mi1 Top	TGCTGTGTA
mi1 Bottom	CCTGTGTA
mi3 Top	TGCTGTGTA
mi3 Bottom	CCTGTGTA
mi4 Top	TGCTGTGTA
mi4 Bottom	CCTGTGTA

Figure 4.4 Artificial microRNA sequences: DNA sequences of top and bottom strands of artificial microRNA oligonucleotides. Red nucleotides indicate overhangs following annealing of top and bottom strands. Green nucleotides indicate the terminal loop region based on the endogenous murine microRNA miR-155.

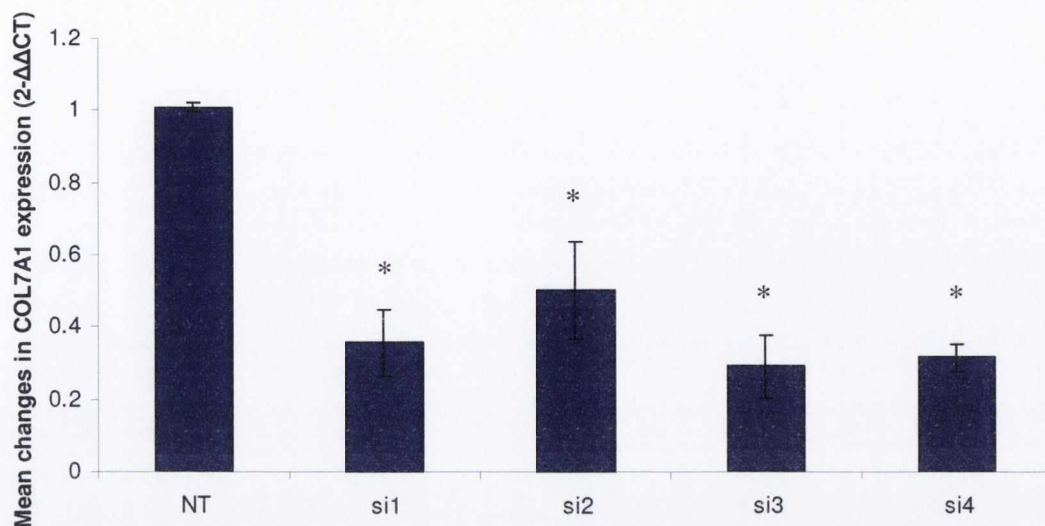


Figure 4.5 Mean changes in COL7A1 expression: effect of siRNAs 1-4 on COL7A1 expression in HaCaT cells. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with siRNA at a concentration of 5nM. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate and the experiment was repeated three more times. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved with all four siRNAs when compared to a non-targeting negative control siRNA ($p < 0.0005$ in all cases). siRNAs 1-4 reduced COL7A1 expression by an approximate average of 65%, 50%, 71% and 68% respectively.

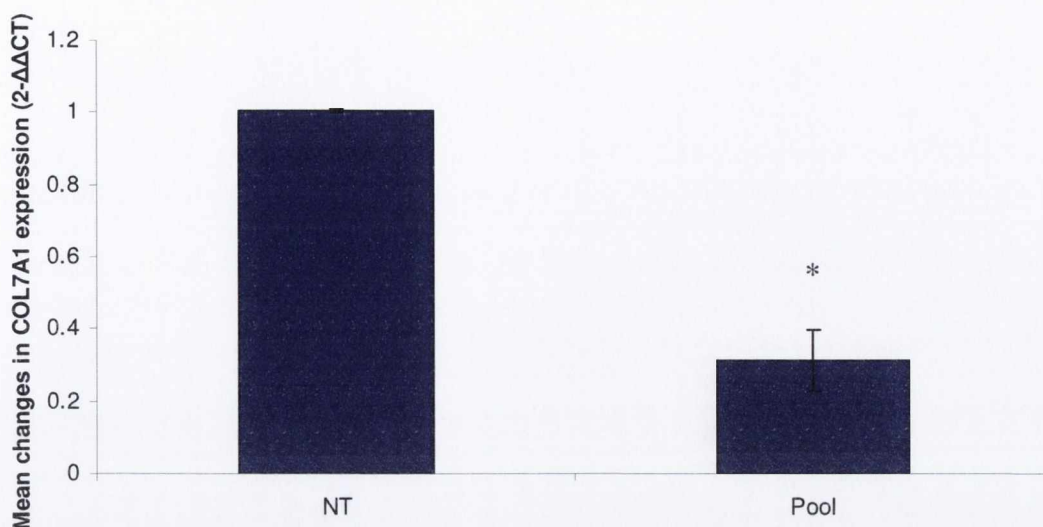


Figure 4.6 Mean changes in COL7A1 expression: effect of a pool of siRNAs 1-4 on COL7A1 expression in HaCaT cells. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5nM siRNA pool and 5nM NT siRNA. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection was undertaken in quadruplicate. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved using the pool of all four siRNAs when compared to a non-targeting negative control siRNA ($p < 0.0005$). The siRNA pool reduced COL7A1 expression by an average of 69%.

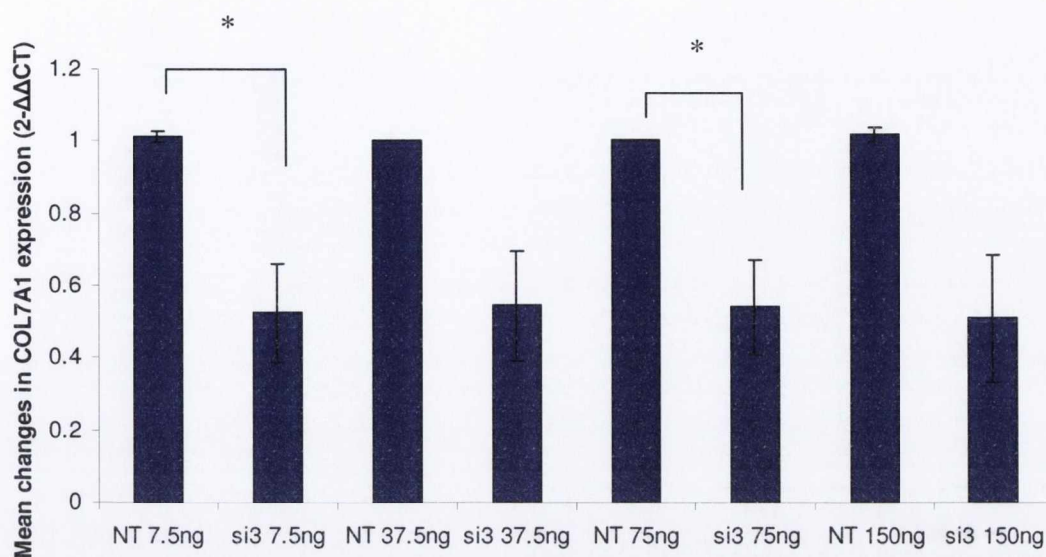


Figure 4.7 Mean changes in COL7A1 expression: effect of a range of siRNA 3 concentrations on COL7A1 expression in HaCaT cells. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 7.5ng, 37.5ng, 75ng or 150ng of siRNA 3 or NT siRNA, equating to final siRNA concentrations of 1nM, 5nM, 10nM and 20nM respectively. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate. COL7A1 mRNA levels were reduced by approximately 48% ($p=0.037$), 46% ($p=0.051$), 46% ($p=0.039$) and 49% ($p=0.055$) with siRNA 3 transfected at concentrations of 7.5ng, 37.5ng, 75ng and 150ng respectively. Whilst suppression with 37.5ng and 150ng siRNA was not found to be statistically significant, it is likely that this is due to the low sample number ($n=2$) and significance would likely arise with an increase of the n number. No statistically significant differences were observed in levels of suppression achieved when comparing across concentrations, with the p -value when comparing the lowest siRNA concentration (7.5ng) to the highest (150ng) being 0.935.

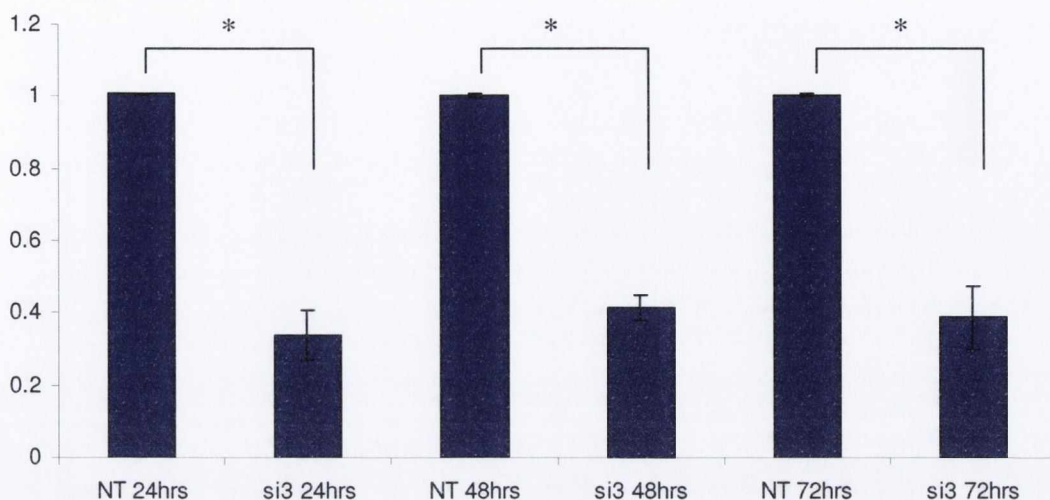


Figure 4.8 Mean changes in COL7A1 expression: effect of siRNA 3 on COL7A1 expression over a range of time-points in HaCaT cells. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5nM siRNA. RNA was isolated from duplicate wells of transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR 24, 48 and 72 hours after transfection. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved with siRNA 3 at all three timepoints when compared to the non-targeting negative control siRNA ($p = 0.006$ for 24hrs, $p = 0.002$ for 48hrs and $p = 0.010$ for 72hrs). Mean reductions in COL7A1 expression of 66%, 59% and 61% were seen at 24, 48 and 72 hours respectively. Comparison among time points did not show statistically significant different levels of suppression ($p = 0.311$ when comparing the 24 hour and 48 hour time points; $p = 0.610$ when comparing the 24 hour and 72 hour time points and $p = 0.727$ when comparing the 48 hour and 72 hour time points).

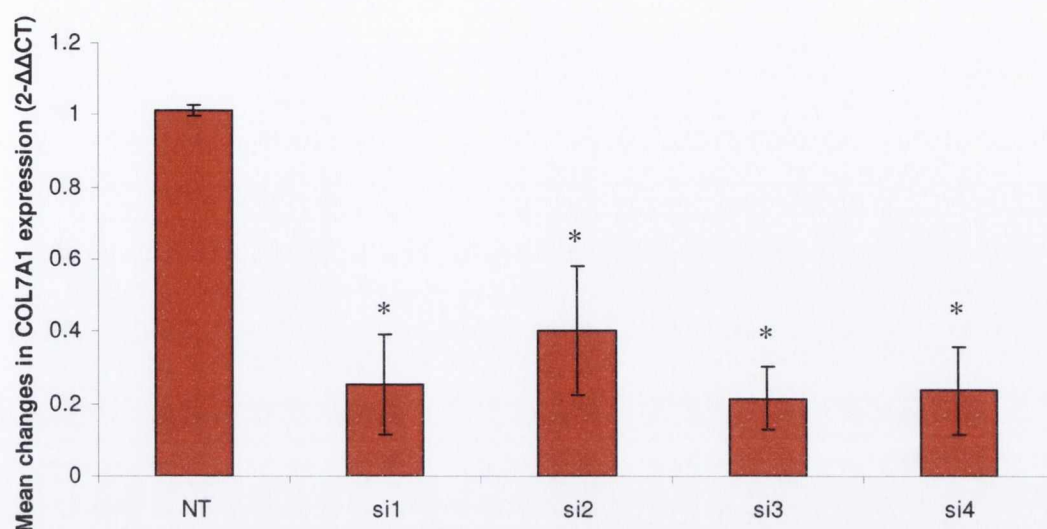


Figure 4.9 Mean changes in COL7A1 expression: effect of siRNAs 1-4 on COL7A1 expression in NHDF cells. Error bars represent standard deviation and stars represent statistical significance. NHDF cells were transfected with 5nM siRNA. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate and the experiment was repeated three more times. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved with all four siRNAs when compared to a non-targeting negative control siRNA ($p < 0.0005$ in all cases). siRNAs 1-4 reduced COL7A1 expression by an approximate average of 74%, 60%, 79% and 76% respectively.

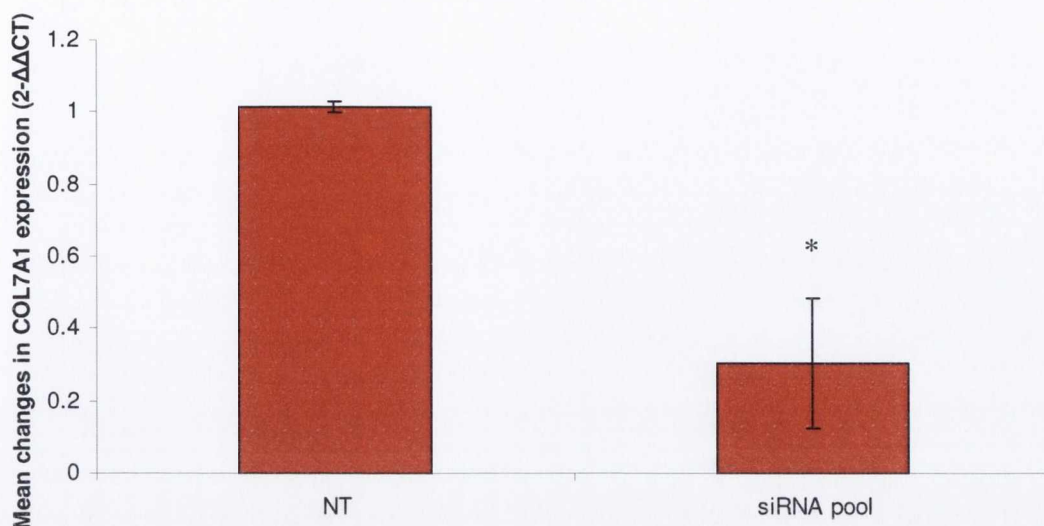


Figure 4.10 Mean changes in COL7A1 expression: effect of a pool of siRNAs 1-4 on COL7A1 expression in NHDF cells. Error bars represent standard deviation and stars represent statistical significance. NHDF cells were transfected with 5nM siRNA pool and 5nM NT siRNA. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection was undertaken in quadruplicate. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved using the pool of all four siRNAs when compared to a non-targeting negative control siRNA ($p < 0.0005$). The siRNA pool reduced COL7A1 expression by an average of 70%.

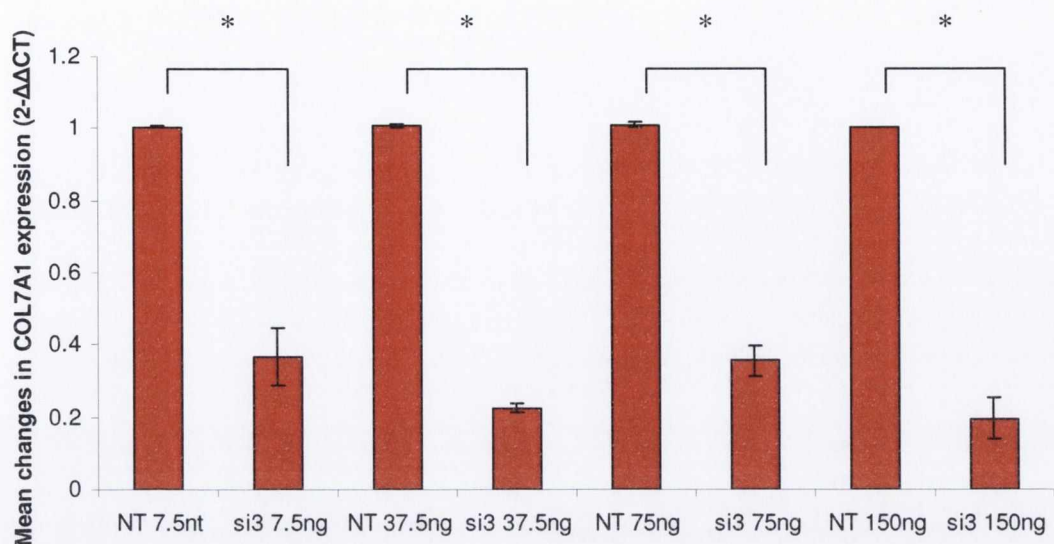


Figure 4.11 Mean changes in COL7A1 expression: effect of a range of siRNA 3 concentrations on COL7A1 expression in NHDFs. Error bars represent standard deviation and stars represent statistical significance. NHDF cells were transfected with 7.5ng, 37.5ng, 75ng or 150ng of siRNA 3 or NT siRNA, equating to final siRNA concentrations of 1nM, 5nM, 10nM and 20nM respectively. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate. COL7A1 mRNA levels were reduced by approximately 63% ($p=0.008$), 77% ($p<0.0005$), 65% ($p=0.002$) and 80% ($p=0.002$) with siRNA 3 transfected at concentrations of 7.5ng, 37.5ng, 75ng and 150ng respectively. No statistically significant differences were observed in levels of suppression achieved when comparing across concentrations, with the p-value when comparing the lowest siRNA concentration (7.5ng) to the highest (150ng) being 0.130. It is entirely possible however that an increase in the sample number ($n=2$ in this case) may result in statistical significance.

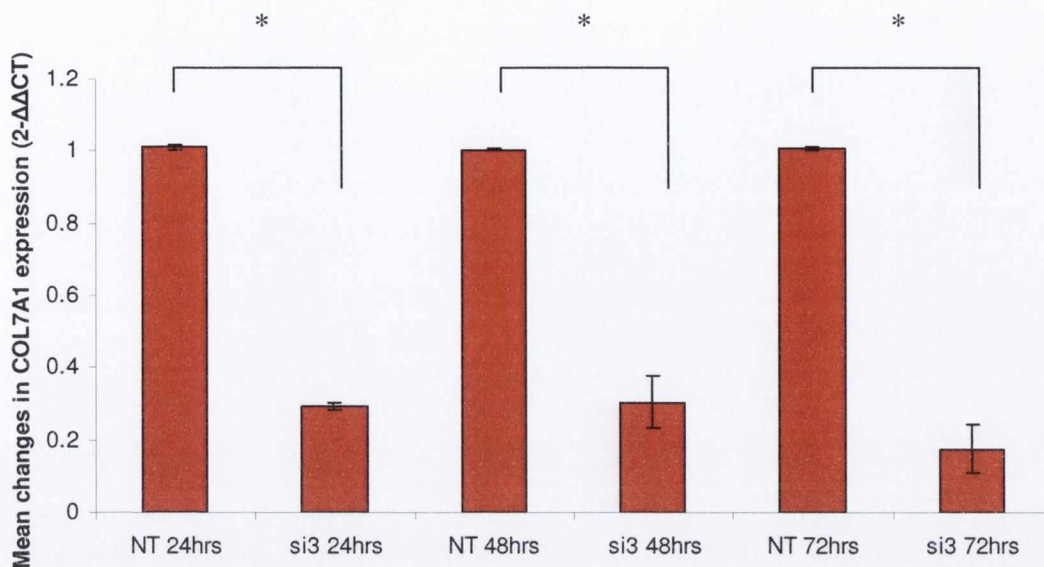


Figure 4.12 Mean changes in COL7A1 expression: effect of siRNA 3 on COL7A1 expression over a range of time-points in NHDF cells. Error bars represent standard deviation and stars represent statistical significance. NHDF cells were transfected with 5nM siRNA. RNA was isolated from duplicate wells of transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR 24, 48 and 72 hours after transfection. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved with siRNA 3 at all three time points when compared to the non-targeting negative control siRNA ($p < 0.0005$ for 24hrs, $p = 0.005$ for 48hrs and $p = 0.003$ for 72hrs). Mean reductions in COL7A1 expression of 71%, 70% and 82% were seen at 24, 48 and 72 hours respectively. Whilst it appears that there is an increase of suppression following incubation of transfected cells for 72 hours, a comparison of the data between time points found no statistical significance between time points ($p = 0.839$ when comparing the 24 hour and 48 hour time points; $p = 0.135$ when comparing the 24 hour and 72 hour time points and $p = 0.206$ when comparing the 48 hour and 72 hour time points).

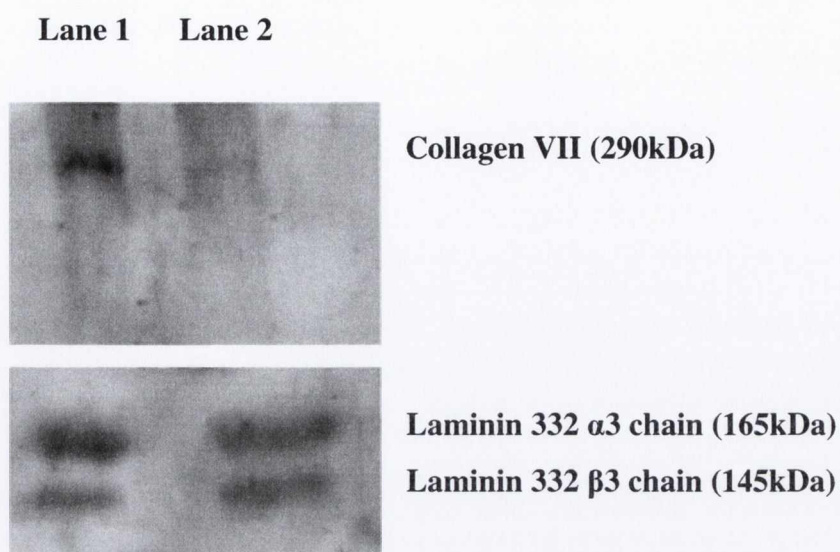
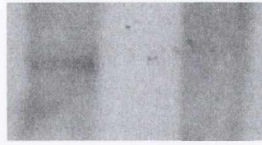


Figure 4.13 Mean changes in Collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of HaCaT cells. Cells were transfected with NT siRNA (Lane 1) and siRNA 3 (Lane 2). Laminin 332 was used as a loading control protein. 48 hours post transfection serum-free media was collected from cells, conditioned and filtered. Media protein samples were subjected to 6% SDS-PAGE. The position of the bands corresponding to full-length 290kDa Collagen type VII, the 165kDa Laminin 332 α 3 chain and the 145kDa Laminin 332 β 3 chain are indicated. Transfection and western blot analysis was repeated on three separate occasions.

Lane 1 Lane 2



Collagen VII (290kDa)

Figure 4.14 Mean changes in Collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of NHDF cells. Cells were transfected with NT siRNA (Lane 1) and siRNA 3 (Lane 2). 48 hours post transfection serum-free media was collected from cells, conditioned and filtered. Media was subjected to 6% SDS-PAGE. The position of the band corresponding to full-length 290kDa Collagen type VII is indicated. Transfection and western blot analysis was repeated on three separate occasions.

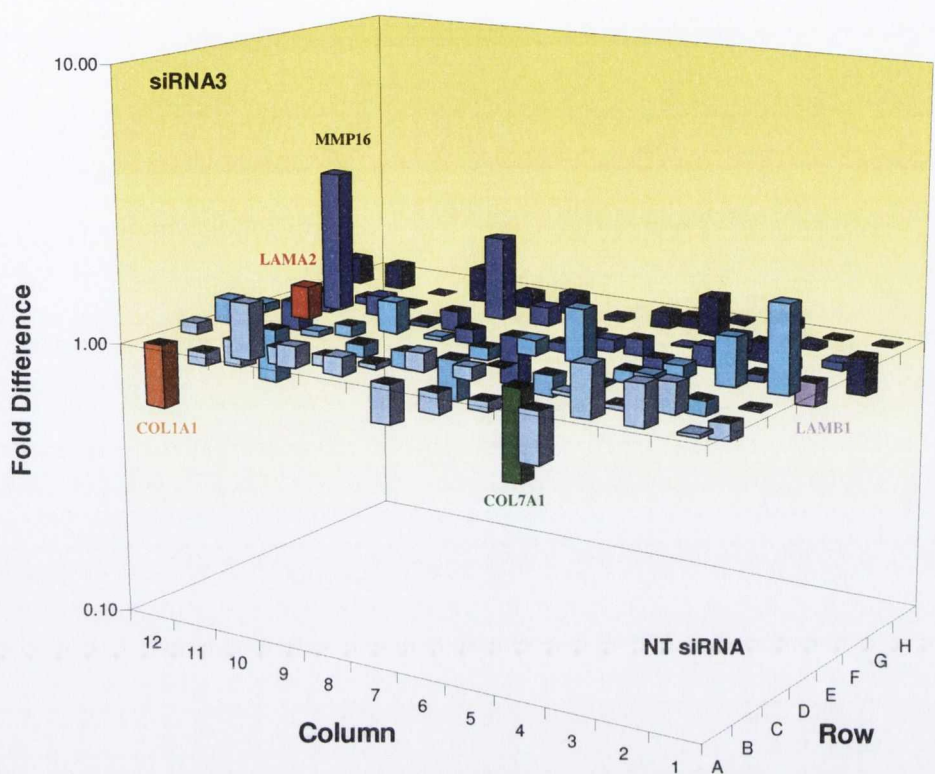


Figure 4.15 3D Profile of expression levels of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated HaCaTs: The 3D Profile graphs the fold difference in expression of each gene between the two samples in the 96-well format of the PCR Array. Columns pointing up (with z-axis values > 1) indicate an up-regulation of gene expression, and columns pointing down (with z-axis values < 1) indicate a down-regulation of gene expression in the siRNA 3-treated sample relative to the NT-treated control sample. Columns in which statistically significant changes in gene expression were found are indicated; COL7A1 was down-regulated by 54% ($p=0.018$), COL1A1 was down-regulated by 41% ($p=0.046$), LAMB1 was down-regulated by 18% ($p=0.033$) and LAMA2 was up-regulated by 32% ($p=0.033$). While MMP16 appears to be up-regulated by approximately 3-fold, this gene's average threshold cycle is relatively high, meaning that its relative expression level is low in both control and test samples, and the p-value for the fold-change is relatively high ($p=0.297$). All other genes were not statistically significantly altered. The layout of the plate and its full list of genes are indicated in figure 4.2.

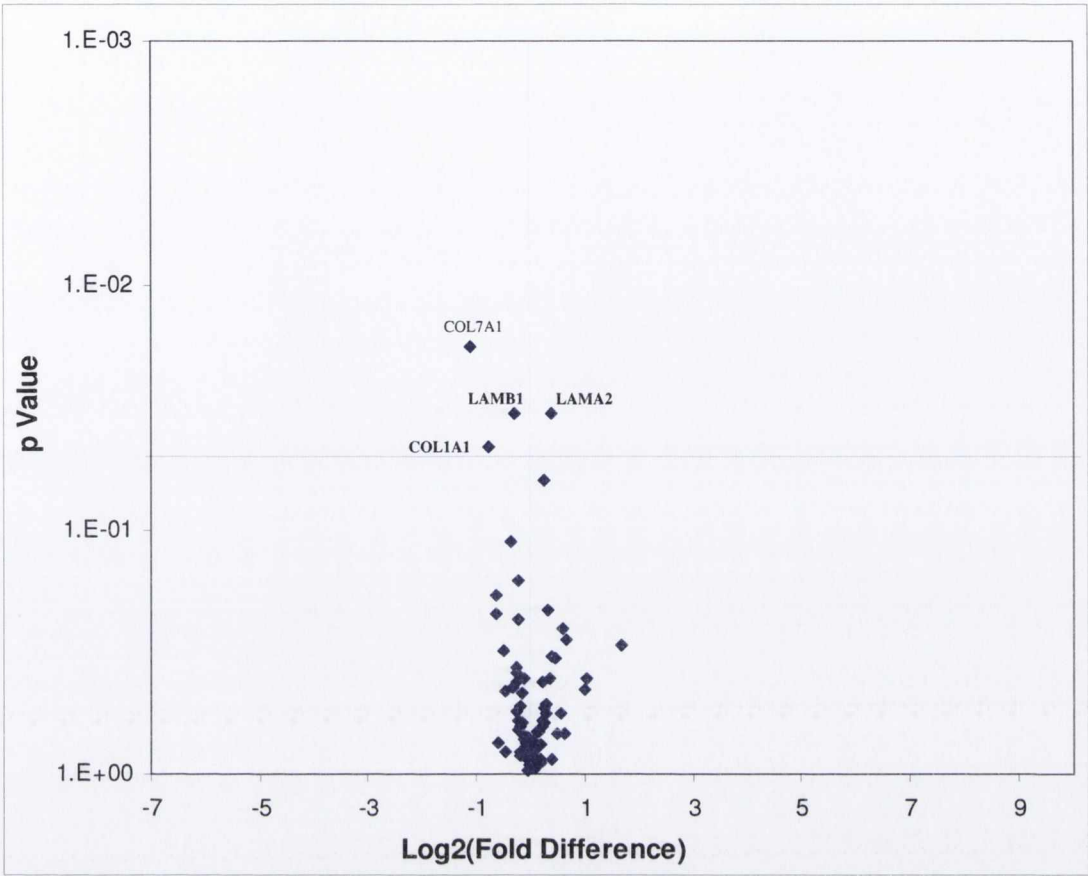


Figure 4.16 Volcano Plot of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated HaCaTs: The Volcano Plot graphs the log2 of the fold change in each gene's expression between the samples versus its p value from the t-test. The pink line indicates fold changes of 1. The blue line indicates the desired p value of the t-test threshold, which is set at $p=0.05$. As indicated, statistically significant changes in expression was found for 4 of the genes tested: COL7A1 ($p=0.018$), COL1A1 ($p=0.046$) and LAMB1 ($p=0.033$), which are all statistically significantly down-regulated in siRNA 3-treated samples and LAMA2 which is statistically significantly up-regulated ($p=0.033$).

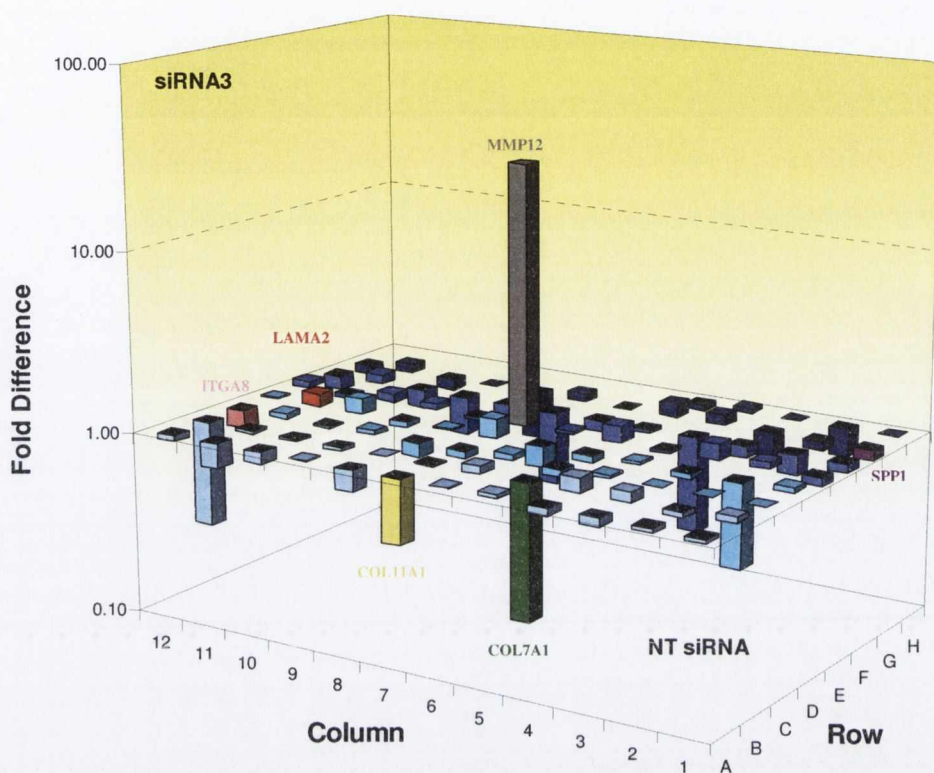


Figure 4.17 3D Profile of expression levels of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated NHDFs: The 3D Profile graphs the fold difference in expression of each gene between the two samples in the 96-well format of the PCR Array. Columns pointing up (with z-axis values > 1) indicate an up-regulation of gene expression, and columns pointing down (with z-axis values < 1) indicate a down-regulation of gene expression in the siRNA 3-treated sample relative to the NT-treated control sample. Columns in which statistically significant changes in gene expression were found are indicated; COL7A1 was down-regulated by 82% ($p=0.0006$), COL11A1 was down-regulated by 56% ($p=0.001$), ITGA8 was down-regulated by 20% ($p=0.04989$) and SPP1 was down-regulated by 12% ($p=0.0397$). LAMA2 was up-regulated by 18% ($p=0.0405$) and MMP12 was up-regulated by approximately 27 fold ($p=0.033096$). All other genes were not statistically significantly altered. The layout of the plate and its full list of genes are indicated in figure 4.2.

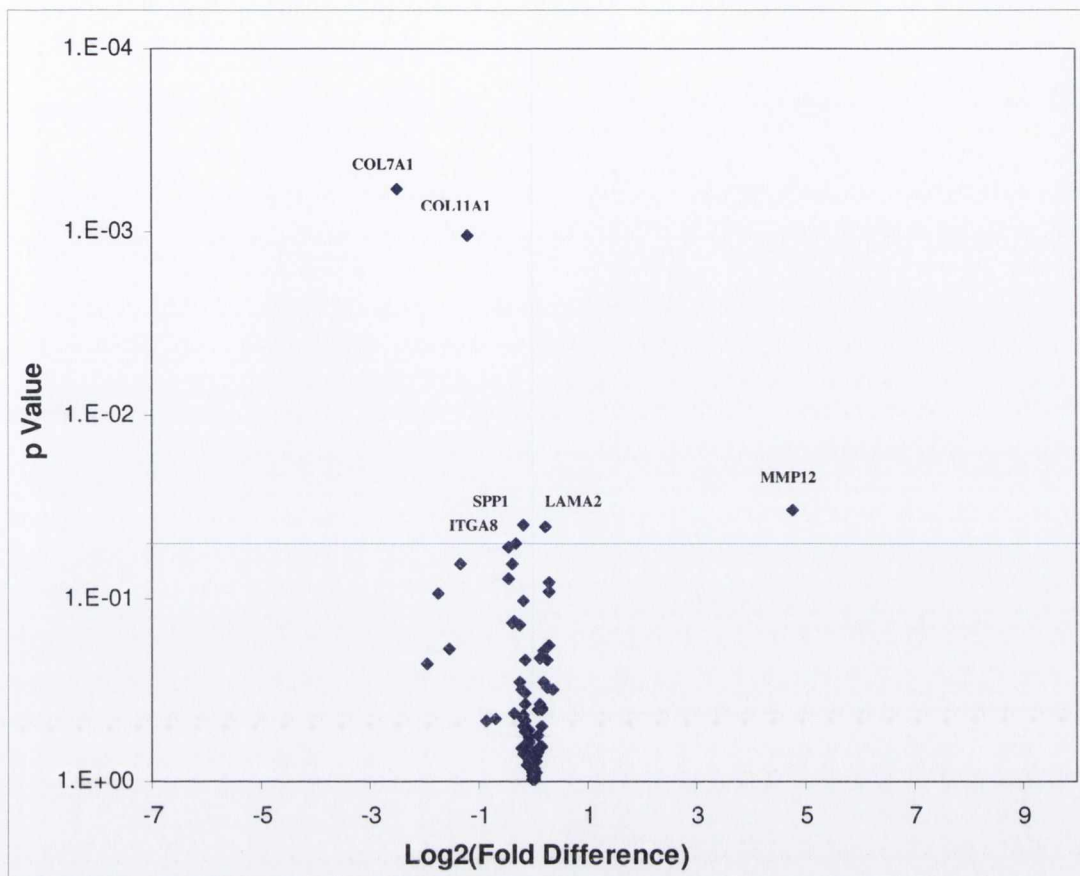


Figure 4.18 Volcano Plot of the 84 ECM and adhesion molecule genes in siRNA 3-treated and NT siRNA-treated NHDFs: The Volcano Plot graphs the log2 of the fold change in each gene's expression between the samples versus its p value from the t-test. The pink line indicates fold changes of 1. The blue line indicates the desired p value of the t-test threshold, which is set at $p=0.05$. As indicated, statistically significant changes in expression were found for 6 of the genes tested: COL7A1 ($p=0.0006$), COL11A1 ($p=0.001$), SPP1 ($p=0.0397$) and ITGA8 ($p=0.04989$), which are all statistically significantly down-regulated in siRNA 3-treated samples, and LAMA2 ($p=0.0405$) and MMP12 ($p=0.033096$) which are statistically significantly up-regulated.

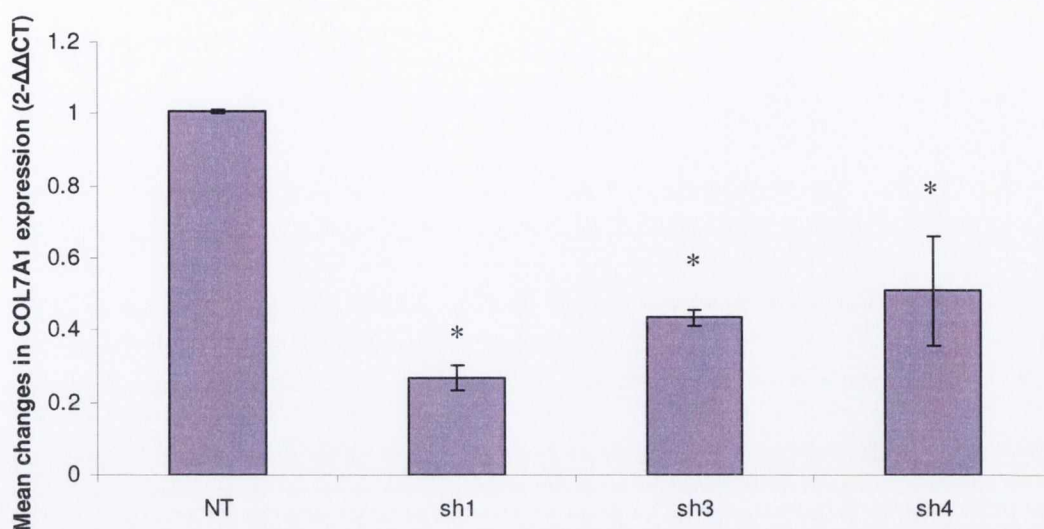


Figure 4.19 Mean changes in COL7A1 expression: effect of shRNA molecules on COL7A1 expression. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5μg DNA (shRNA 1, 3, 4 or NT). RNA was directly isolated from duplicate transfected sample wells after 24 hours and their COL7A1 mRNA levels were measured by real-time rtPCR. The experiment was repeated 3 times. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Significant differences in COL7A1 mRNA levels were observed between all three COL7A1-targeting shRNAs and NT shRNA ($p < 0.005$ in all cases). shRNA 1 significantly suppressed COL7A1 mRNA by approximately 73%, shRNA 3 significantly suppressed COL7A1 mRNA by approximately 57% and shRNA 4 significantly suppressed COL7A1 mRNA by approximately 49%.

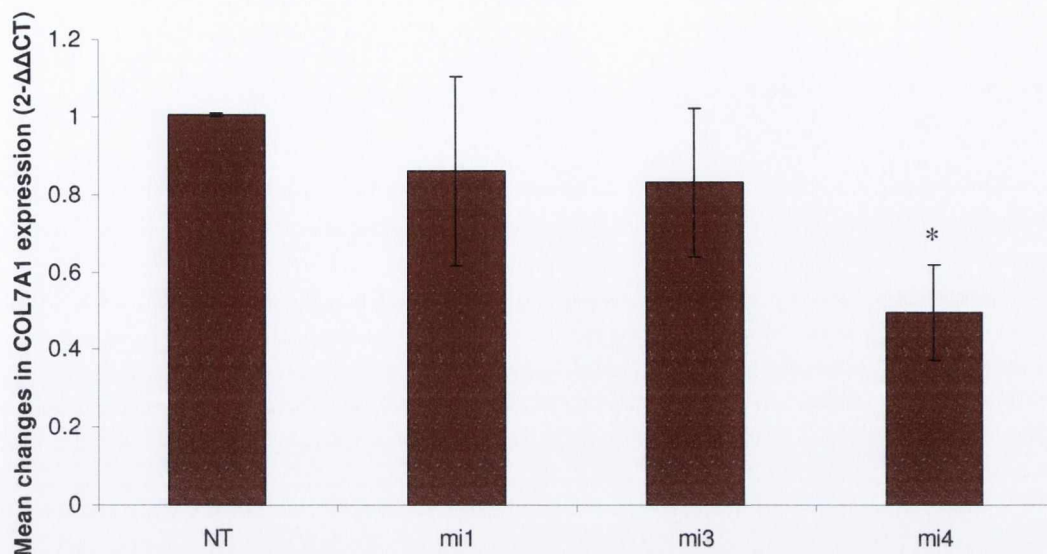


Figure 4.20 Mean changes in COL7A1 expression: effect of artifical microRNA molecules on COL7A1 expression. Error bars represent standard deviation and stars represent statistical significance. HaCaT cells were transfected with 5μg DNA (artificial microRNA 1, 3, 4 or NT). RNA was directly isolated from duplicate transfected sample wells after 24 hours and their COL7A1 mRNA levels were measured by real-time rtPCR. The experiment was repeated twice. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Significant differences in COL7A1 mRNA levels were observed between artificial microRNA 4 and NT artificial microRNA ($p < 0.005$). Artificial microRNA 4 significantly suppressed COL7A1 mRNA by approximately 51%. Artificial microRNAs 1 and 3 reduced COL7A1 mRNA by average levels of 14% and 17% respectively, although this data was not statistically significant ($p = 0.291$ and $p = 0.121$ respectively).

Chapter 5

COL7A1 Gene Replacement

5.1 Introduction

A major complexity related to the development of gene therapies for dominant diseases is the fact that the disease-causing alleles associated with such conditions often involve vast intragenic heterogeneity (Millington-Ward *et al.*, 1997). Over 100 mutations of the rhodopsin gene and over 50 mutations of the peripherin gene have been found to cause autosomal dominant Retinitis Pigmentosa (Kiang *et al.*, 2005; Palfi *et al.*, 2006) and over 150 mutations of the COL1A1 and COL1A2 genes result in dominantly inherited Osteogenesis Imperfecta (Forlino and Marini, 2000). As previously discussed, the dominant dystrophic form of Epidermolysis Bullosa (DDEB) represents another such heterogeneous disease. Over 60 dominant mutations of the COL7A1 gene have so far been implicated, the majority of which involve substitution of one of the many glycine residues within the collagenous domain of the protein, which are important for maintaining the structural integrity of collagen type VII protein to allow for its role as an essential structural protein (Persikov *et al.*, 2004). Moreover, further DDEB-causing COL7A1 mutations are continuously being described (Riedl *et al.*, 2009; Koga *et al.*, 2010). Whilst it may be possible to target individual mutations in some cases, particularly commonly occurring mutations which may account for a significant proportion of the total number of affected patients, development of a therapeutic strategy which would be suitable regardless of the mutation in question would be ideal. Such a general gene therapy-based approach should in principle be less time-consuming and less technically and financially demanding in comparison to the development of individual therapies for each disease-causing mutation.

The focus of the research carried out as described in this thesis so far has centred upon an exploration of the potential of RNAi technologies to mediate down-regulation of expression of the COL7A1 gene. Chapters 3 and 4 involve discussion of a mutation independent-based therapeutic approach for dominant DEB, which has centred on suppression of the COL7A1 transcript regardless of mutant or wild-type status. Use of siRNAs, shRNAs and artificial microRNAs to achieve such suppression has been investigated. In order to treat a dominant negative disorder such as DDEB, it would most likely be necessary to eliminate the negatively interfering protein. In the case of the mutation-independent approach which has been employed in these studies,

an additional step of restoration of expression of the wild-type protein would likely be required in conjunction with COL7A1 suppression. As discussed in chapter 1, there are a number of potential methods by which it may be possible to achieve such replacement. Direct delivery of human collagen VII protein may represent one possible viable approach, and one which would be independent of and resistant to the effects of the RNAi suppressors. Investigations into such a delivery mechanism have been reported (Bruckner-Tuderman, 2009; Remington *et al.*, 2009), whereby COL7A1-null mice were injected with recombinant human collagen type VII protein, and initial findings have proved promising. The phenotype of the treated mice was found to markedly improve in comparison to non-treated control mice, as shown by decreased blistering and prolonged survival of the animals. A number of other gene-based replacement methods have also been suggested and investigated *in vitro* for the treatment of dominant negative genetic disorders (Millington-Ward *et al.*, 1997; Millington-Ward *et al.*, 2004; Kubodera *et al.*, 2011), all of which would potentially avoid the requirement to target individual mutations for genetic suppression.

One such possible strategy for gene replacement would involve targeting suppressors to regions of the transcript of interest that are transcribed but which are not translated (UTRs), resulting in suppression of both mutant and wild-type alleles. Thus concurrent introduction of a replacement gene with an altered UTR sequence while maintaining the sequence coding for amino acids would allow for wild-type gene expression which is resistant to the effects of the suppressor therapeutic. Another proposed possible method of treatment of dominant genetic human diseases involves using intragenic polymorphisms to exclusively target the mutant allele, the idea being that a significant proportion of patients despite having different disease-causing mutations may have the same polymorphism. In this case expression of the wild-type allele would be maintained as suppression is targeted to a single allele and so delivery of a replacement gene would only be required in cases whereby some disease pathology was due to reduced levels of wild-type protein. Finally, a third method of treating a dominant disease in a mutation-independent manner would involve suppression in the coding sequence of wild-type and mutant alleles while simultaneously introducing a replacement gene which is resistant to the effects of the suppressor, by exploiting the degeneracy of the genetic code. It is the latter approach which has been employed in this research project, the results of which are described

in the current chapter of this thesis, in an effort to generate replacement genes encoding wild-type type VII collagen protein but which are resistant to the effects of the most potent COL7A1 suppressors discussed in chapter 4 of this thesis. It is of note that the approach of suppression and replacement to treat dominantly inherited disorders has been explored previously by a number of groups developing gene-therapies for other inherited disorders (Palfi *et al.*, 2006; Chadderton *et al.*, 2009; Kubodera *et al.*, 2011; Millington-Ward *et al.*, 2011).

Exploitation of the degeneracy of the genetic code in the creation of a replacement gene results in a gene which differs in DNA sequence to the wild-type version but which still encodes normal wild-type protein. Whilst most synonymous changes which may be made are single changes at wobble base positions, it may be possible to alter two or even all three nucleotides within a codon while maintaining the encoded amino acid. The sequence specificity of RNAi technology means that such a gene which is altered at the nucleotide level should in theory be refractory to suppression. When designing such changes in a gene sequence it may be necessary to take into account codon usage for that particular gene, as codon usage has been shown previously to affect gene expression. Codon usage patterns have been analysed for multiple genes from multiple species and significant codon usage biases have been detected in humans (Eyre-Walker, 1991; Shabalina *et al.*, 2006; Ren *et al.*, 2007; Schmid and Flegel, 2011). While alterations made to the nucleotide sequence during the design of a replacement gene should not result in any amino acid sequence alterations, it is possible that such nucleotide changes may affect transcript features such as its half life or three dimensional structure (O'Reilly *et al.*, 2008). Thus in the design of a replacement gene, utilisation of codons which are commonly used at other sites within that gene may limit possible undesirable effects of such modifications.

With regard to a suppression and replacement therapeutic strategy for DDEB, a major obstacle associated with the replacement aspect of the therapy may be the need to deliver full-length COL7A1 to the target cells. In the quest for development of a therapeutic for the recessive form of DEB, whereby the complete absence of collagen type VII protein may lead to the requirement for delivery of the full length gene, a wide range of delivery strategies have been investigated including viral and non-viral based delivery mechanisms (Ortiz-Urda *et al.*, 2002; Siprashvili *et al.*, 2010;

Titeux *et al.*, 2010), as discussed in section 1.5.1. As previously described, the size of the COL7A1 cDNA is large, at a length of approximately 9kb (Christiano *et al.*, 1994a). This large size, as well as the repetitive nature of the cDNA sequence has hampered development of a gene therapy for the disease, largely due to difficulties in accommodation of the full-length COL7A1 cDNA by therapeutically-suitable viral vectors as well as the risk of occurrence of genetic rearrangements (Lanuti *et al.*, 2011). In recent studies however, progress has been made in the *ex vivo* delivery of full-length COL7A1, which has laid the basis for the commencement of a phase I human clinical trial for RDEB (Siprashvili *et al.*, 2010), as discussed in section 1.5.1, the results of which have yet to be reported. Delivery of full-length COL7A1 cDNA to RDEB human keratinocytes using the LZRSE-COL7A1 retroviral vector has resulted in type VII collagen expression *in vivo*, following grafting of corrected skin grafts onto immunodeficient mice. In fact, type VII collagen expression was found to be maintained for up to one year following grafting, and correction of the disease phenotype was detected at the levels of cell and tissue morphology as well as anchoring fibril formation. Efficient gene transfer to progenitor cells present in the basal layer of the skin allowed for the long-term expression of the COL7A1 gene, as it is these keratinocyte progenitors which are involved in the continuous self-renewal of normal human epidermis.

The work described in the current chapter of this Ph.D. thesis involves an attempt to design, generate and test COL7A1 replacement genes, following the results obtained in chapter 4. Significant suppression of the COL7A1 transcript has been obtained with three siRNAs tested in chapter 4 via the process of RNA interference. However if used as a therapeutic for dominant dystrophic EB this approach would result in down-regulation of not only the mutant transcript, but also of the wild-type gene copy. Collagen type VII protein is essential for the maintenance of healthy skin, as seen by the severe blistering phenotype of recessive dystrophic EB patients with two COL7A1 nonsense mutations and thus a complete absence of collagen VII protein (Hovnanian *et al.*, 1994; Christiano *et al.*, 1997). Therefore, concurrent replenishment of collagen type VII levels following knockdown with the RNAi suppressor molecules is likely to be necessary. One possible method of such replenishment which has been investigated during the course of the studies outlined in this chapter involves delivery of a COL7A1 replacement gene which has been

designed to be resistant to the effects of the RNAi suppressors by exploitation of the degeneracy of the genetic code. Due to difficulties associated with the testing of the full-length COL7A1 cDNA, predominantly due to its large size and repetitive nature, this replacement study originally involved testing of small fragments of the COL7A1 gene corresponding to regions targeted by three of the most potent siRNAs tested and characterised in chapter 4. These small COL7A1 DNA fragments were altered in nucleotide sequence such that the associated COL7A1 transcripts would display resistance to siRNA-mediated suppression. Following this proof of principle of the general strategy for COL7A1, it would then be worthwhile to progress to generation of and testing of a full-length replacement COL7A1 gene.

5.2 Materials and Methods

5.2.1 Replacement sequence design

DNA sequences for replacement genes were designed by altering nucleotides within the siRNA target sites of the COL7A1 gene. By taking advantage of the degeneracy of the genetic code, in principle this allows for translation of wild-type collagen VII protein despite differences in gene sequence. These altered nucleotides were designed taking codon usage of COL7A1 into account (Sequence Manipulation Suite: Codon Usage; http://www.ualberta.ca/~stothard/javascript/codon_usage.html). Replacement sequences were designed to be resistant to siRNAs 1, 3 and 4. For siRNAs 1 and 4, two separate sequences were designed due to the fact that it was possible to make a large number of nucleotide changes without changing the translational product. Thus sequences with the maximum possible number of changes, (13 nucleotide changes in the case of replacement 1 and 12 nucleotide changes in the case of replacement 4) were designed and denoted Rep1Max and Rep4Max, as well as sequences with a smaller amount of changes (7 nucleotide changes in the case of replacement 1 and 6 nucleotide changes in the case of replacement 4), denoted Rep1 and Rep4. One replacement sequence was designed for siRNA 3, denoted Rep3, which incorporated 6 nucleotide changes into the design. See figures 5.1, 5.2 and 5.3 for replacement sequences.

5.2.2 COL7A1 replacement fragment design

Fragments of the COL7A1 gene, each encompassing approximately 400-500bp, were ordered pre-synthesised from GeneArt AG. For siRNA 1, 3 fragments encompassing exons 5-7 of the COL7A1 gene were ordered: one fragment containing the wild-type sequence (WT1), one in which 13 nucleotide changes had been made (Rep1Max), and one in which 7 nucleotide changes had been made (Rep1). For siRNA 3, 2 fragments were ordered encompassing exons 57-66 of the COL7A1 gene, one containing the wild-type sequence (WT3) and the other containing 6 nucleotide changes (Rep3). Finally, for siRNA 4, 3 fragments were ordered which encompassed COL7A1 exons 21-23. These fragments contained the wild-type sequence (WT4), the sequence with 12 changes (Rep4Max) and the sequence in which 6 changes were made (Rep4). Each GeneArt DNA sequence was designed with *NotI* and *XbaI* restriction sites towards the 5' and 3' ends respectively. Full ordered sequences are listed in Appendix III. Pre-

synthesised DNA were supplied in vector format, in either standard GeneArt pMA or pMK-RQ plasmids and resuspended to a concentration of 0.1µg/µl.

5.2.3 Mini COL7A1-GFP fusion vector construction

COL7A1-GFP fusion vectors were constructed using the In-Fusion™ Advantage PCR Cloning Kit (Clontech), as outlined in figure 5.4. All GeneArt-ordered COL7A1 DNA fragments were cloned into the pAcGFP1-N In-Fusion Ready vector, which allows for the expression of the COL7A1 sequence as a fusion protein to the N-terminus of AcGFP1, a Green Fluorescent Protein from *Aequorea coerulescens*. The In-Fusion cloning method was utilised, in which PCR primers were designed which share 15 bases of homology with the sequence at the ends of the linearised pAcGFP1-N vector. These primers were used to amplify the COL7A1 DNA fragments in the GeneArt vectors, as detailed in section 2.2.1.10. Primer sequences are listed in figure 5.5a. Following loading of a small sample of the PCR products on a 2% agarose gel to ensure that the correct sized fragments had been amplified, the amplified products were treated with Cloning Enhancer which removes background template DNA and PCR residue, thus eliminating the need to gel purify the PCR products prior to cloning (section 2.2.1.10).

Ligations were carried out using the kit-supplied In-Fusion enzyme and buffer, and subsequent transformations, as detailed in section 2.2.1.10, were plated on kanamycin-agar plates. A number of kanamycin-resistant colonies were chosen for mini plasmid DNA preparation (section 2.2.1.11) and plasmid sequences and orientation were verified using the pAcGFP1-N sequencing primers (figure 5.5a). Validated positive clones were identified and transformations were again carried out using 5µl of a 1 in 50 dilution of the appropriate mini-preparation (section 2.2.1.9). The Qiagen High Speed Plasmid Maxi Kit was used for large-scale plasmid purification, as outlined in section 2.2.1.12. The concentration and purity of the DNA samples was determined using the NanoDrop™ 1000 Spectrophotometer (section 2.2.1.15) and the constructs were again sequenced as outlined above to guarantee that the expected DNA plasmid had been isolated. In total, 8 of these COL7A1-GFP fusion vectors were generated (WT1, Rep1Max, Rep1, WT3, Rep3, WT5, Rep5Max and Rep5).

5.2.4 Co-transfection of COL7A1-GFP fusion vectors and siRNA

HEK293 cells were co-transfected with COL7A1-GFP fusion DNA and siRNA using Attractene transfection reagent (Qiagen). Attractene is a non-liposomal lipid that allows for transfection of adherent cells including those that are usually difficult to transfect, and has been found to be suitable for co-transfection of DNA with siRNA. It enables highly efficient DNA transfection while ensuring low cytotoxicity (Attractene Transfection Reagent Handbook, 2008). Co-transfections were carried out as described in section 2.2.2.7. Following seeding of 5×10^5 cells per well of a 6-well plate, 0.6 μ g DNA was diluted in serum-free DMEM to a total volume of 100 μ l and the appropriate siRNA at a final concentration of 1.25 nM was added. Samples were mixed, treated with 4.5 μ l Attractene reagent and incubated for 10-15 minutes at room temperature. Cells were then treated with prepared samples and incubated under normal growth conditions. Co-transfections were carried out with WT1, Rep1Max and Rep1 along with siRNA 1 or NT siRNA; WT3 and Rep3 along with siRNA 3 or NT siRNA; and WT4, Rep4Max and Rep4 along with siRNA 4 or NT siRNA. Transfection of each construct with the relevant siRNA was carried out on at least three separate occasions.

5.2.5 Flow cytometry data analysis of co-transfected cells

24 hours following co-transfections, cells were collected and pelleted as previously described. Cells were resuspended in 300 μ l PBS and 100 μ l 4% paraformaldehyde and placed on ice until flow cytometry analysis was carried out, as detailed in section 2.2.2.9. Summit 4.3 software was used for the analysis of samples by flow cytometry. The mean fluorescent intensities (MFI) of GFP from cell samples were taken and used for data analysis. Data obtained from 15,000 cells treated with siRNA were normalised to the values obtained from 15,000 cells treated with NT siRNA.

5.2.6 Statistical analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

5.2.7 CMV-full length COL7A1 cloning

The full-length COL7A1 gene was kindly provided by Dr Alain Hovnanian of Institut National de la Santé et de la Recherche Médicale, Toulouse, France. The COL7A1 cDNA provided was contained in the pTOPO vector, termed pTOPO-XLCOL7 (figure 5.6). As the COL7A1 gene was not under the control of any promoter in this format, it was firstly required to clone a promoter into pTOPO-XLCOL7 to allow for COL7A1 expression. The CMV promoter was chosen and isolated from the pcDNA3.1+ plasmid which was previously available in the laboratory. 10µg of both pcDNA3.1+ and pTOPO-XLCOL7 were firstly digested with 1µl of *HindIII* restriction enzyme and NEB Buffer 2 in total volumes of 50µl at 37°C for 2 hours. Following purification of the digested vectors as described in section 2.2.1.2, plasmids were further digested with 2µl *MluI* restriction enzyme, along with NEB Buffer 3 in final volumes of 50µl, at 37°C for 2 hours. Whilst digested pTOPO-XLCOL7 was then purified as described in section 2.2.1.2, the 683bp CMV promoter-containing fragment was separated from the pcDNA3.1+ vector by gel electrophoresis (1%). The promoter fragment was excised from the gel and the DNA extracted as described in section 2.2.1.5.

Ligations were carried out in a final volume of 10µl, containing 5µl CMV insert DNA, 1µl linearised pTOPO-XLCOL7 plasmid, 1X Ligation buffer, 1µl T4 DNA ligase and 2µl ddH₂O; and incubated overnight at 18°C. Competent XL1-blue *E.coli* bacteria were prepared (section 2.2.1.8) and transformed using 5µl of ligated insert and vector (section 2.2.1.9). 200µl of each transformation was plated onto LB agar plates containing kanamycin (50µg/ml) and incubated overnight at 37°C. Transformants were examined and a number of colonies were chosen for mini plasmid DNA preparation (section 2.2.1.11). The sequence of each plasmid was verified using forward and reverse sequencing primers (see figure 5.5b), as per section 2.2.1.14. Validated positive clones were identified and transformations were again carried out using 5µl of a 1 in 50 dilution of the appropriate mini preparation (section 2.2.1.9). Large-scale plasmid purification was then carried out using the Qiagen High Speed Plasmid Maxi kit, as outlined in section 2.2.1.12. The concentration and purity of the DNA was determined (section 2.2.1.15) and sequencing was carried out again to ensure that the expected DNA construct was isolated.

5.2.8 Full length COL7A1 replacement gene generation

A replacement full-length COL7A1 gene which is resistant to the effects of siRNA 3, termed pTOPO-XLCOL7Rep3, was generated using the QuikChange II XL Site-Directed Mutagenesis kit (Agilent Technologies). The full-length replacement gene design incorporated the same changes in COL7A1 sequence as were made in Rep3 (section 5.2.1) and as outlined in figure 5.2. The QuikChange kit in question utilises primers which are designed to incorporate the desired mutations into the construct of interest, thus allowing for site-specific mutation of the double-stranded plasmid. The primer sequences used for creation of the siRNA 3 full-length replacement gene, as outlined in figure 5.5b, were designed using the web-based QuikChange Primer Design Program (<http://www.agilent.com/genomics/qcpd>). Following uploading of the sequence and the desired changes to the design programme, forward and reverse mutagenic primers were designed containing the desired mutations and which anneal to the same sequence on opposite strands of the pTOPO-XLCOL7 plasmid.

The PCR mutagenesis reaction was set up using 25ng plasmid DNA template, 125ng of forward mutagenesis primer, 125ng of reverse mutagenesis primer, 1µl dNTP mix, 1X reaction buffer and 3µl QuikSolution, to a final volume of 50µl. 1µl (2.5 Units) *PfuUltra* HF DNA polymerase was then added and the PCR reaction was carried out as follows: one cycle of 95°C for 1 minute; 18 cycles of 95°C for 50 seconds, 60°C for 50 seconds and 68°C for 19.5 minutes (1.5 minutes per kb of plasmid length); and one cycle of 68°C for 7 minutes. Following temperature cycling, the reaction tubes were placed on ice for 2 minutes in order to cool the reaction to below 37°C. PCR reactions were then treated with 1µl of *DpnI* restriction enzyme (10 Units), in order to digest the parental DNA template and to select for mutation-containing synthesised DNA. Samples were mixed thoroughly by pipetting, centrifuged at 13,200rpm for 1 minute (IEC Micromax, Thermo-Electron) and transferred to a 37°C water bath for 1 hour.

Following *DpnI* digestion, transformations were carried out, using the kit-supplied XL10-Gold ultracompetent cells. Cells were thawed on ice and for each reaction, 45µl cells were transferred to prechilled 14ml BD Falcon polypropylene round-bottom tubes. 2µl of the kit-supplied β-Mercaptoethanol mix was added to each 45µl aliquot of cells and the cells were incubated on ice for 10 minutes, with gentle

swirling every 2 minutes. 2µl of the *DpnI*-treated DNA from each sample was transferred to separate aliquots of the ultracompetent cells. Samples were swirled gently to mix and incubated on ice for 30 minutes. Tubes were then heat-pulsed in a 42°C water bath for 30 seconds, followed by incubation of samples on ice for 2 minutes. 0.5ml of NZY⁺ broth (section 2.1.1) preheated to 42°C was added to each tube and samples were incubated at 37°C for 1 hour with shaking at 225rpm. 250µl of transformation reactions were then spread on agar plates containing kanamycin (50µg/ml) and plates were incubated overnight at 37°C.

Plates were examined and a number of colonies were chosen for mini plasmid DNA preparation (section 2.2.1.11). The sequence of each plasmid was verified using forward and reverse sequencing primers in order to identify plasmids which had incorporated the desired mutant sequences (see figure 5.5b), as per section 2.2.1.14. Validated positive clones were identified and transformations were again carried out using 5µl of a 1 in 50 dilution of the appropriate mini preparation (section 2.2.1.9). Large-scale plasmid purification was then carried out using the Qiagen High Speed Plasmid Maxi kit, as outlined in section 2.2.1.12. The concentration and purity of the DNA was determined (section 2.2.1.15) and sequencing was carried out again to ensure that the expected DNA construct was isolated.

5.2.9 Transfection and co-transfection of pTOPO-XLCOL7 and pTOPO-XLCOL7Rep3

pTOPO-XLCOL7 and pTOPO-XLCOL7Rep3 were transfected into HEK293 cells using Attractene transfection reagent, either alone or along with a targeting (siRNA 3) or non-targeting siRNA (NT siRNA). Co-transfections were carried out as described in section 2.2.2.7. Following seeding of 5×10^5 cells per well of a 6-well plate, the appropriate amount of DNA was diluted in serum-free DMEM to a total volume of 100µl and in the case of co-transfections, the appropriate siRNA at a concentration of 1.25nM was added. Samples were mixed, treated with 4.5µl Attractene reagent and incubated for 10-15 minutes at room temperature. Cells were then treated with prepared samples and incubated under normal growth conditions. Transfection of each construct either alone or with the relevant siRNA was carried out in duplicate wells on three separate occasions.

5.2.10 Analysis of RNA from transfected cells

RNA from HEK293 cells transfected with pTOPO-XLCOL7 and pTOPO-XLCOL7Rep3 either alone or co-transfected with siRNAs using Attractene Transfection Reagent was extracted using the Qiagen RNeasy[®] Plus mini kit. 350µl Buffer RLT was added directly to each well of a 6-well plate and the RNeasy[®] Plus mini kit protocol was followed (section 2.2.3.1). The concentration and purity of RNA samples was determined using the NanoDrop[™] 1000 Spectrophotometer (section 2.2.1.15). Real-time rtPCR was used to analyse expression levels of COL7A1 mRNA according to section 2.2.3.2 using COL7A1 primers previously available in the laboratory (section 2.1.3). Expression was normalised using the endogenous internal control gene Ubiquitin C (UBC) (section 2.1.3). The real-time rtPCR steps which were carried out are outlined in section 2.2.3.2 and the Cycle Threshold (C_T) values were used for RNA quantification as outlined in section 2.2.3.3.

5.2.11 Statistical Analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

5.2.12 Protein analysis

Protein analysis of cells treated with pTOPO-XLCOL7 and pTOPO-XLCOL7Rep3 along with siRNAs was carried out by western blotting. HEK293 cells growing in a T175 tissue culture plate at approximately 80% confluency were transfected by addition of 8µg DNA and 90µl Attractene Transfection Reagent, prepared as described in section 5.2.9. Co-transfections were carried out with 8µg DNA along with siRNA 3 or NT siRNA at a concentration of 1.25nM and 90µl Attractene Transfection Reagent. 24 hours later, cells were washed with PBS and incubated in serum-free DMEM for a further 24 hours. Media protein was then collected and treated, as described in section 2.2.4.1, and protein concentration was measured using a BCA Protein Assay Kit (section 2.2.4.4). Western blotting was then undertaken as described in section 2.2.4.6. A 1 in 500 dilution of anti-collagen type VII rabbit polyclonal antibody (Calbiochem), which binds to the NC1 domain of the protein, was used to detect collagen type VII protein. An exposure time of approximately 5 seconds at room temperature was generally adequate for collagen type VII protein to

be detected on chemiluminescence sensitive film. For each western blot experiment, identical samples were loaded on duplicate gels and while one gel was used for transfer of samples onto nitrocellulose membrane and probed for collagen VII protein as described, the second identical gel was stained with PageBlue[™] Protein Staining Solution (Fermentas Life Sciences) for use as a loading control. PageBlue[™] Protein Staining Solution allows for staining of all proteins separated in polyacrylamide gels. The gel was firstly washed in ddH₂O with gentle agitation for 3x10 minutes for removal of SDS, incubated in PageBlue solution overnight with gentle agitation and then de-stained in ddH₂O. Transfections and western blot analyses were carried out on three separate occasions.

5.3 Results

5.3.1 Replacement sequence design

Following achievement of significant suppression of the COL7A1 gene in both human keratinocyte and fibroblast cell lines by introduction of synthesised siRNA suppressor molecules as discussed in chapter 4 of this Ph.D. thesis, it was of interest to generate replacement genes resistant to the effects of the siRNAs in question. Such a suppression and replacement-based gene therapy approach may in principle be suitable for the treatment of the dominant form of dystrophic Epidermolysis Bullosa. Replacement gene sequences were designed by exploitation of the degeneracy of the genetic code, whilst taking into account codon usage data relating to the COL7A1 gene. Replacement sequences were designed in an effort to create constructs resistant to the effects of siRNAs 1, 3 and 4, whilst still encoding wild-type type VII collagen protein. Thus despite incorporating a number of changes into the COL7A1 nucleotide sequence, all changes made were synonymous, allowing for maintenance of the wild-type amino acid residue sequence. Upon commencement of the designing of the replacement sequences, both siRNAs 1 and 4 were found to target regions of the COL7A1 transcript in which it would be possible to make quite a high number of synonymous changes. Thus it was decided to design two possible COL7A1 replacement sequences for siRNAs 1 and 4, one with the maximum number of nucleotide changes possible and one in which a smaller number of changes were made. In this case it would be possible to compare the resistance to the desired siRNA of the two separate genes with differing numbers of nucleotide changes. For siRNA 1, sequences with both 13 and 7 nucleotide changes were designed (Rep1Max and Rep 1 respectively; figure 5.1) whilst for siRNA 4, sequences with both 12 and 6 nucleotide changes were designed (Rep4Max and Rep4 respectively; figure 5.3). For siRNA 3, the target sequence in question just allows for a smaller number of synonymous changes to be made and so in this case, just one replacement sequence was designed, which incorporated 6 changes into the nucleotide sequence (Rep3; figure 5.2). During the design process of all COL7A1 replacement genes, the nucleotide changes were made whilst taking into account codon usage of wild-type COL7A1.

5.3.2 Mini COL7A1-GFP fusion vector construction

In order to test the efficiency of the replacement design sequences in their resistance to the effects of the siRNAs in question, it was decided to firstly test vectors containing small fragments of the COL7A1 gene encompassing the siRNA target regions. This initial proof-of-principle experiment would circumvent the requirement to work with the extremely large full length COL7A1 gene in the testing of all 5 of the designed replacement sequences (Rep1, Rep1Max, Rep3, Rep4, and Rep4Max). The vector in question was pAcGFP1-N In-Fusion Ready, a linearised mammalian expression vector which encodes a Green Fluorescent Protein from *Aequorea coerulescens*. The vector is designed such that if a nucleotide sequence is added in-frame into the In-Fusion cloning site, this sequence will automatically be in-frame with the AcGFP1 sequence upstream. Therefore, the incorporated sequence of interest will be expressed as a fusion protein to the N-terminus of AcGFP1 (pAcGFP1-N In-Fusion™ Ready Vector Information, 2006).

Fragments of the COL7A1 gene, each encompassing approximately 500bp, were pre-synthesised by GeneArt and supplied in vector format. Fragments were then cloned into the pAcGFP1-N In-Fusion Ready vector, thus allowing for creation of mini COL7A1-GFP fusion vectors. 8 such vectors were generated, based on wild-type sequences containing the siRNA1, 3 and 4 target regions (exons 5-7, 57-66 and 21-23 respectively; denoted WT1, WT3 and WT4 respectively) and equivalent sequences with the incorporated nucleotide changes designed for Rep1, Rep1Max, Rep3, Rep4 and Rep4Max. The cloning process of the mini COL7A1-AcGFP fusion vectors was based on In-Fusion™ Advantage PCR, which allows for the joining of multiple pieces of DNA that have 15 bases of homology at their linear ends (In-Fusion™ Advantage PCR Cloning Kit User Manual, 2010). Therefore in this case, primers were designed to amplify the GeneArt COL7A1 DNA fragments, which shared 15 bases of homology with the sequence at the ends of the linearised pAcGFP1-N In-Fusion Ready vector. PCR products were then combined with the linearised vector in the In-Fusion cloning reaction and transformed into *E.coli*, as outlined in figure 5.4. Positive clones were sequenced (figure 5.5a) in order to ensure the correct sequence and orientation and DNA preparations of all 8 constructs were made and used in further analysis.

5.3.3 Analysis of mini COL7A1-GFP vector resistance to siRNAs

Following cloning of the relevant fragments of the COL7A1 gene into the pAcGFP1-N vector, testing of these fusion genes and their resistance to the relevant siRNAs was examined. The fusion of the COL7A1 fragments to a fluorescent marker allows for a direct measurement of the levels of expression of the COL7A1 mini genes by measuring GFP intensity by flow cytometry. Co-transfection experiments were carried out in 6-well plates, with 0.6 μ g of DNA along with 1.25nM siRNA being delivered to HEK293 cells using the lipid-based Attractene transfection reagent. Cells were collected 24 hours post transfection and flow cytometry analysis was carried out. Each COL7A1 construct was transfected with both the relevant COL7A1-targeting and non-targeting siRNAs on at least three separate occasions. A non-fluorescent sample of non-transfected cells was used as a control in order to determine the auto-fluorescence threshold. Mean fluorescent intensities (MFI) of AcGFP for each treated sample were used for data analysis (figure 5.7). The data obtained from each construct treated with the relevant siRNA was normalised to the data obtained from the same construct treated with the non-targeting siRNA.

Co-transfection experiments of the WT1-AcGFP fusion vector and siRNA 1 indicated significant suppression of the WT sequence, as measured by GFP intensity during flow cytometry analysis. Mean GFP intensity was reduced by approximately 57.3% compared to cells treated with WT1-AcGFP and NT siRNA ($p=0.006$). Meanwhile, neither mean GFP intensities of Rep1Max-AcGFP nor Rep1-AcGFP vectors treated with siRNA 1 were statistically significantly different from mean GFP intensities of the same vectors treated with NT siRNA ($p=0.721$ and $p=0.759$ respectively) (figure 5.8). Co-transfection of the WT3-AcGFP fusion vector and siRNA 3 also indicated significant suppression of the WT sequence, as measured by GFP intensity during flow cytometry analysis. Mean GFP intensity was reduced by approximately 76.45% compared to HEK293 cells treated with WT3-AcGFP and NT siRNA ($p=0.001$). However the mean GFP intensity of the Rep3-AcGFP vector treated with siRNA 3 was not statistically significantly different from the mean GFP intensity of the same vector treated with NT siRNA ($p=0.757$) (figure 5.9). Significant suppression of the WT4-AcGFP vector sequence by siRNA 4 was also indicated during flow cytometry analysis, following transfection of the WT4-AcGFP fusion vector along with siRNA 4. Mean GFP intensity was statistically significantly reduced

by mean levels of 77%, compared to cells treated with the same vector along with NT siRNA ($p=0.001$). Conversely, mean GFP intensities of Rep4Max-AcGFP and Rep4-AcGFP did not significantly differ regardless of siRNA 4 or NT siRNA treatment ($p=0.736$ and $p=0.448$ respectively) (figure 5.10). These data suggest that the introduction of nucleotide changes in the COL7A1 sequence over the target site for RNAi-based suppression had eliminated suppression.

5.3.4 Generation of full-length COL7A1 and replacement genes

Following evidence of resistance of all of the designed replacement gene sequences to siRNAs targeting the wild-type regions of the sequences by testing of small fragments of the COL7A1 gene fused to GFP, it was subsequently decided to proceed to evaluation of full length COL7A1 genes. The full-length wild-type COL7A1 cDNA was kindly obtained from Dr. Alain Hovnanian (Institut National de la Santé et de la Recherche Médicale, Toulouse, France), contained in the cloning vector pTOPO (figure 5.6). In order to allow for expression of the COL7A1 gene, it was firstly required to insert a promoter sequence into the pTOPO vector. The CMV promoter was chosen, the sequence was isolated from the pcDNA3.1+ plasmid which was available in the laboratory, and the CMV sequence was subsequently cloned into the pTOPO-XLCOL7 vector to allow for efficient expression of COL7A1 cDNA. It was then decided to proceed with creating and testing a full-length replacement gene, based on the design of Rep3 (figure 5.2), designed to be resistant to the effects of what was found to be one of the most potent COL7A1 suppressors from the studies described in chapter 4 of this thesis, siRNA 3. As mentioned, previous studies using small COL7A1 gene fragments suggest that this altered nucleotide sequence in the target region of siRNA 3 renders transcripts from the replacement gene indeed resistant to the effects of the siRNA suppressor (figure 5.8). A full-length replacement COL7A1 gene which is resistant to the effects of the siRNA would be required for this proposed suppression and replacement gene-based therapeutic for DDEB, and so the full-length replacement gene, termed pTOPO-XLCOL7Rep3, was generated. Mutagenic primers were used to incorporate the desired sequence changes into the pTOPO-XLCOL7 gene by the process of PCR mutagenesis (section 5.2.8), to create the pTOPO-XLCOL7Rep3 construct. Following transformation of PCR products using XL10-Gold ultracompetent bacterial cells, DNA preparations of colonies were sequenced (see figure 5.5b for primer sequences) in order to isolate a construct of the

correct sequence. More specifically, 6 nucleotide changes were incorporated into the COL7A1 gene pTOPO-XLCOL7Rep3 over the siRNA 3 target site (figure 5.2).

5.3.5 Analysis of siRNA 3 effects on full-length COL7A1 genes: RNA analysis

Following generation of full-length wild-type and replacement COL7A1 genes under the control of the CMV promoter, testing of these COL7A1 constructs and their expression abilities as well as differences in their expression patterns following treatment with siRNA 3 were examined. In order to firstly ensure efficient expression of the COL7A1 gene from the pTOPO vector following insertion of the CMV promoter, HEK293 cells were transfected with either 0.6µg or 1.2µg pTOPO-XLCOL7 DNA in duplicate wells using the lipid-based Attractene transfection reagent. 24 hours later, RNA was isolated directly from transfected wells as well as from duplicate wells of non-transfected HEK293 cells. COL7A1 levels were measured by real-time rtPCR and normalised using expression levels of the endogenous control gene UBC. In both 0.6µg and 1.2µg DNA-treated HEK293 cells, levels of COL7A1 expression were increased by between approximately 1500-2000 fold, compared to COL7A1 levels of non-treated HEK293 cells (figure 5.11).

As the ability of the pTOPO-XLCOL7 plasmid to express COL7A1 cDNA was confirmed by real-time rtPCR, analysis of the effects of siRNA 3 on both pTOPO-XLCOL7 and pTOPO-XLCOL7Rep3 was carried out. HEK293 cells were transfected with 0.4µg DNA, along with 1.25nM siRNA using Attractene transfection reagent. 24 hours later, RNA was isolated and real-time rtPCR was used to measure COL7A1 mRNA levels. Both wild-type and replacement gene plasmids were transfected with siRNA 3 and NT siRNA in duplicate wells, and each transfection experiment was carried out on 3 separate occasions. As expected, statistically significant suppression of the wild-type COL7A1 gene was obtained with RNA samples from cells co-transfected with pTOPO-XLCOL7 and siRNA 3, compared to those co-transfected with pTOPO-XLCOL7 and NT siRNA ($p < 0.0005$). A mean reduction in COL7A1 mRNA levels of approximately 77% was achieved with siRNA 3 (figure 5.12). Conversely, statistically significant suppression of the replacement COL7A1 gene was not obtained with siRNA 3. Following analysis of RNA samples from cells co-transfected with pTOPO-XLCOL7Rep3 and siRNA 3, there was no significant difference in COL7A1 mRNA levels compared to levels in RNA from

cells treated with pTOPO-XLCOL7Rep3 and NT siRNA ($p=0.796$). Thus, the full-length replacement gene appears to be resistant to the effects of siRNA 3 (figure 5.13).

5.3.6 Analysis of siRNA 3 effects on full-length COL7A1 genes: protein analysis

Following significant suppression of the wild-type full-length COL7A1 gene and resistance of the replacement gene to siRNA 3 as analysed by real-time rtPCR, validation of this RNA data at the protein level was attempted by western blot analysis. Approximately 1×10^7 HEK293 cells were co-transfected with $8 \mu\text{g}$ DNA (either pTOPO-XLCOL7 or pTOPO-XLCOL7Rep3) along with siRNA 3 or NT siRNA at a concentration of 1.25 nM using Attractene Transfection Reagent as described in section 5.2.12. Media protein was collected and collagen VII protein was detected using an anti-collagen VII rabbit polyclonal antibody (Calbiochem). Co-transfections and western blot analyses were carried out three times. In order to ensure that equal amounts of protein had been loaded into each well, two gels were loaded during each western blot experiment and one of these gels was stained with PageBlue™ Staining Solution, which stains all proteins which have been separated in a polyacrylamide gel. Western blot analysis on conditioned media from cells treated with the wild-type gene pTOPO-XLCOL7 showed significant suppression of collagen type VII protein in those cells co-transfected with siRNA 3, compared to those co-transfected with the negative control NT siRNA. Only a very faint band of approximately 290 kDa was visible in the lane in which siRNA3-treated cell media was loaded (figure 5.14). In contrast, there did not appear to be any significant difference in densities of collagen type VII bands following western blot analysis of media samples from cells treated with pTOPO-XLCOL7Rep3 and either siRNA 3 or NT siRNA (figure 5.15). The data suggest that siRNA 3 can elicit suppression of wild-type COL7A1 at both the RNA and protein levels while resistance to RNAi-mediated suppression can be engineered into replacement COL7A1 genes using nucleotide changes at degenerate positions over the target site for RNAi-based suppression.

5.4 Discussion

Following investigations into the effects of a range of synthesised COL7A1-targeting siRNAs, significant suppression of the COL7A1 transcript at both the mRNA and protein levels was achieved and validated by real-time rtPCR and western blot analyses respectively as discussed in chapter 4. An average of 71% COL7A1 mRNA suppression was achieved with one of the most potent suppressors siRNA 3 in the human keratinocyte HaCaT cell line, while an average of 79% suppression was achieved with the same molecule in a primary fibroblast cell line (NHDFs). Thus such a suppressor may represent a valuable tool for use in a gene therapy-based approach for the dominant form of Dystrophic Epidermolysis Bullosa. However, use of the suppressor alone in this mutation-independent RNAi-based approach would result in suppression of both mutant and wild-type alleles, the consequence of which would likely be catastrophic for any individual and may actually worsen the DDEB disease phenotype. Thus, concurrent introduction of a collagen type VII protein-encoding gene is most likely necessary for such an approach, and the gene in question would be required to display a property of resistance to the effects of the RNAi suppressor. It is the designing and generation of such a replacement gene which is central to the research described in the current chapter of this Ph.D. thesis. Generation of a COL7A1 gene displaying such an attribute is made possible by utilising the degeneracy associated with the genetic code.

As potent suppression of the COL7A1 transcript was achieved with three separate siRNAs as detailed in chapter 4, initially it was decided to design and test replacement sequences which would be resistant to the effects of all three of these suppressor agents. As described in section 5.3.1, such a feat was achieved by exploiting the degeneracy of the genetic code which allows for the making of synonymous changes to the nucleotide sequence of the siRNA target regions. Thus such alterations should theoretically result in translation of the same amino acid sequence as the wild-type protein, despite differences in the nucleotide sequence. It would be ideal to incorporate a sufficient number of changes into the nucleotide sequence of a replacement gene such that the RNAi molecule is completely ineffective in its role as a suppressor, whilst the number of changes made should not be too great such that this would cause any negative effects, such as changes in the

structure or efficiency of the transcript and translation of the sequence. For both siRNAs 1 and 4, as the sequences in question allowed for quite a large number of synonymous changes to be made (13 and 12 nucleotide changes respectively), two different replacement sequences were designed. One set of sequences incorporated the maximum number of changes possible; whilst the second set of sequences had a reduced number of nucleotide changes (7 and 6 nucleotide changes for siRNAs 1 and 4 respectively). If the sequence with a smaller number of changes was as effective at maintaining resistance as the sequence with a greater number of changes, it is the former which would likely be more suitable for use in a therapeutic approach, due to its closer similarity to the wild-type gene sequence. Meanwhile for siRNA 3, the maximum number of synonymous changes which could be made was 7 and so the replacement sequence design for this siRNA incorporated just 6 changes into the nucleotide sequence. During the designing of these COL7A1 replacement gene sequences, particularly the sequences with the minimum number of incorporated changes, decisions on the nucleotide changes to be made included consideration of codon usage of the wild-type COL7A1 gene. In cases where more than one possible codon could be used whilst maintaining the amino acid sequence, the codon which is more commonly used throughout the COL7A1 gene was generally chosen. As discussed in section 5.1, significant codon usage biases have been detected in humans and so as it is not fully known whether or not changes in codons could have negative effects on expression of the gene in question. Therefore significant alterations in codon usage patterns were kept to a minimum during the replacement gene design process.

As described in sections 5.3.2 and 5.3.3, original testing of the replacement sequences involved fusion of short fragments of the COL7A1 gene encompassing the relevant siRNA target regions to AcGFP1, allowing for their expression as fusion proteins to the fluorescent AcGFP1 protein's N-terminus. Although it was possible that the use of small fragments of the COL7A1 gene rather than the full-length form may not provide a true representation of the effects of the siRNA, possibly due to differences in secondary structure, if reliable such a screening method would be faster and thus more efficient than the cloning and subsequent testing of all 5 full-length COL7A1 replacement genes. Hence, it was thought that such a method may possibly result in the ruling out of the effectiveness of any of the replacement genes, which

would thus prevent the inappropriate investment of time in generating and testing the full-length equivalent. Each siRNA was firstly co-transfected along with the wild-type form of the relevant COL7A1-AcGFP fusion plasmids in order to ensure that the result resembled that seen with the endogenous COL7A1 in HaCaT and fibroblast cells and thus that this experimental model was applicable for testing of the replacement gene sequences. Co-transfection experiments were carried out in HEK293 cells, a widely-used cell line which was originally generated by transformation of cultures of human embryonic kidney cells with sheared fragments of adenovirus type 5 DNA (Graham *et al.*, 1977). HEK293 cells were thought to be ideal for use in such an experiment due to their quick replication and ease of maintenance, their amenability to transfection using a range of methods including the ability to achieve high transfection efficiencies, as well as their faithful translation and processing of proteins (Thomas and Smart, 2005).

Co-transfections of HEK293 cells with wild-type COL7A1-AcGFP DNA and the pertinent siRNAs closely emulated those suppression results seen following transfection of HaCaTs and NHDFs with the same suppressors (figures 5.8, 5.9 and 5.10). Significant suppression was obtained with all 3 siRNAs, as measured by GFP intensity during flow cytometry analysis. Whilst the suppression levels obtained with siRNAs 3 and 4 of WT3-AcGFP and WT4-AcGFP respectively were similar to levels of the suppression achieved of the endogenous COL7A1 gene with the same siRNAs, the siRNA 1-mediated suppression of WT1-AcGFP achieved was not to the same extent as that seen with the endogenous gene (overall 65% COL7A1 suppression in HaCaTs and 74% suppression in NHDFs compared to 57.3% suppression of WT1-AcGFP). This reduction of suppression achieved with the siRNA in these co-transfection experiments, although not statistically significant ($p=0.408$ when comparing siRNA 1-mediated suppression of endogenous COL7A1 in HaCaTs and of WT1-AcGFP; $p=0.156$ when comparing siRNA 1-mediated suppression of endogenous COL7A1 in NHDFs and of WT1-AcGFP) was consistently observed. This is unlikely to be as a result of low transfection efficiency of the siRNA, as due to the nature of the experimental setup it is more likely that the larger DNA molecule rather than the small siRNA would be the limiting factor in terms of transfection efficiency. Thus it is possible that the secondary structure of WT1-AcGFP mRNA is slightly less amenable to siRNA 1-mediated suppression compared to that of the

endogenous full-length COL7A1 mRNA. Nonetheless, following co-transfections of COL7A1-AcGFP DNA and the three siRNAs, the results obtained overall showed the experimental model in question to be viable for testing of the replacement gene sequences and their resistance to the same siRNAs (figures 5.8, 5.9 and 5.10).

As described in section 5.3.3, subsequent testing of the replacement gene sequences and the effects of the relevant siRNAs was carried out in the same manner as for the wild-type COL7A1-AcGFP plasmids. All five designed replacement gene sequences were found be unaffected following co-transfection of the targeting siRNAs and thus all sequences were found to display resistance to their equivalent siRNAs. In the cases of both siRNAs 1 and 4 in which two replacement sequences were designed with differing numbers of nucleotide changes, both sets of sequences displayed equal resistance to the relevant siRNAs. As the sequences with the smaller number of changes were equally resistant to the effects of the RNAi molecules as the sequences with the larger number of changes, it is those sequences with fewer nucleotide changes which would likely be more suitable for use in a potential DDEB therapeutic approach due to their closer similarity to the wild-type counterpart. Following establishment of the resistance of the replacement sequences, it was subsequently decided to progress with just one of these sequences for further development of a DDEB therapeutic. As siRNA 3 represents one of the most potent COL7A1 suppressors characterised as described in chapter 4 of this thesis, the Rep3 sequence was chosen for creation of a full-length COL7A1 replacement gene.

Subsequent to obtaining the promoter-less full-length COL7A1 gene pTOPO-XLCOL7, a cloning procedure which would allow for a method of expression of the gene in mammalian cells was firstly required to be carried out. The cytomegalovirus (CMV) promoter, which is one of the strongest promoters described, was chosen for use. The fact that the CMV promoter directs high-level constitutive expression of genes and is active in a broad range of mammalian cell lines would make it ideal for use in these *in vitro* experiments (Boshart *et al.*, 1985; Nelson *et al.*, 1987; Lai *et al.*, 2010). Whilst HEK293 cells were found to express low levels of endogenous COL7A1, delivery of the CMV-driven full-length COL7A1 gene to this cell line using Attractene transfection reagent resulted in an up to 2,000-fold increase in COL7A1 expression levels, compared to the non-transfected HEK293 cells, as

evaluated by real-time rtPCR (figure 5.11). These results show both the efficiency of the transfection reagent for delivery of this large plasmid to HEK293 cells as well as the efficiency of the CMV promoter for allowing gene expression. Using a PCR mutagenesis strategy, the full-length COL7A1 replacement gene designed to be resistant to siRNA 3 was generated from the wild-type sequence, and denoted pTOPO-XLCOL7Rep3. The replacement gene in question, which had incorporated 6 synonymous nucleotide changes into the siRNA target region, was shown to be resistant to siRNA 3 following real-time rtPCR and western blot analysis of co-transfected cells (figures 5.13 and 5.15). There was no significant difference in expression levels between cells co-transfected with plasmid DNA and siRNA 3 compared to those transfected with plasmid DNA and the non-targeting siRNA. Whilst the Laminin 332 protein was found to be ideal for use as a loading control protein during western blot analysis of transfected HaCaT cells (chapter 4), adequate detection of this protein could not be achieved from media obtained from HEK293 cells. Thus as no other control protein could be identified which would allow for proof of equal loading of protein samples in all wells, a second gel was loaded with the same protein samples during western blot analysis and stained with PageBlue™ Protein Staining Solution. This allows for staining of all proteins separated in polyacrylamide gels and thus allowed for visualisation of adequate loading of equal amounts of protein to each well.

5.5 Conclusion

The focus of the research described in the current chapter of this Ph.D. thesis centred on an investigation into the generation of replacement COL7A1 sequences. These replacement gene sequences were designed to be resistant to the effects of the most potent COL7A1 siRNA suppressors which were identified and discussed in chapter 4 of this thesis. Resistance to RNAi-based suppression was engineered into COL7A1 replacement genes by exploiting the degeneracy of the genetic code. Replacement gene sequences were designed for three of the siRNAs and these COL7A1 sequences were chosen to minimise perturbations in codon usage. In addition, a full length replacement COL7A1 gene was generated and tested *in vitro* for resistance to one of the most potent suppressors, siRNA 3. The results obtained mark an important step

towards development of a mutation-independent suppression and replacement strategy for the dominantly inherited form of Dystrophic EB.

siRNA1												
Target	AGC	AGC										
Amino Acid	S	S	CAA	TCC	TTG	AGA	GTA	CAG				
			Q	S	L	R	V	Q				
	AGC 0.22	AGC 0.22	CAA 0.46	TCC 0.30	TTG 0.03	AGA 0.28	GTA 0.21	CAG 0.54				
	AGT 0.16	AGT 0.16	CAG 0.54	TCT 0.10	TTA 0.04	AGG 0.41	GTT 0.30	CAA 0.46				
	TCT 0.10	TCT 0.10		TCA 0.09	CTT 0.25	CGT 0.06	GTC 0.35					
	TCC 0.30	TCC 0.30		TCC 0.12	CTC 0.32	CGC 0.04	GTG 0.14					
	TCA 0.09	TCA 0.09		AGT 0.16	CTA 0.12	CGA 0.08						
	TCG 0.12	TCG 0.12		AGC 0.22	CTG 0.23	CGG 0.13						
Actual gene	AGC	AGC	CAA	TCC	TTG	AGA	GTA	CAG				
Replacement Max Changes	AGT	TCG	CAG	AGT	CTC	CGG	GTC	CAG			13 changes	
Replacement Min Changes	AGT	TCC	CAG	TCC	CTG	AGG	GTC	CAG			7 changes	

Figure 5.1 Replacement sequence design for siRNA 1: Replacement sequences were designed such that they would, in principle, be resistant to the effects of the siRNA whilst still encoding wild-type collagen VII protein. Synonymous nucleotide changes were made possible due to the degeneracy of the genetic code and taking into account codon usage of collagen type VII. Letters in bold indicate the 21-nucleotide wild-type COL7A1 target sequence of the siRNA. Letters in red indicate nucleotide changes from wild-type. Possible codons encoding the amino acid in question are listed, along with the fraction by which each codon is used throughout the COL7A1 gene. Two replacement sequences were generated for siRNA 1 (Rep1Max and Rep1) due to the high number of synonymous nucleotide changes that were possible.

siRNA4											
Target	ACC	TCG	ATC	GAC	TCG	GTG	ACT	TTG			
Amino Acid	T	S	I	D	S	V	T	L			
	ACC 0.46	TCG 0.12	ATC 0.64	GAC 0.68	TCG 0.12	GTG 0.14	ACT 0.26	TTG 0.03			
	ACG 0.13	AGT 0.16	ATA 0.11	GAT 0.33	AGT 0.16	GTA 0.21	ACG 0.13	TTA 0.04			
	ACA 0.15	AGC 0.22	ATT 0.25		AGC 0.22	GTT 0.30	ACA 0.15	CTG 0.23			
	ACT 0.26	TCA 0.09			TCA 0.09	GTC 0.35	ACC 0.46	CTA 0.12			
		TCT 0.10			TCT 0.10			CTT 0.25			
		TCC 0.30			TCC 0.30			CTC 0.32			
Actual gene	ACC	TCG	ATC	GAC	TCG	GTG	ACT	TTG			
Replacement Max Changes	ACT	AGC	ATT	GAT	AGC	GTC	ACC	CTG			12 changes
Replacement Min Changes	ACC	TCC	ATC	GAT	TCC	GTC	ACC	CTG			6 changes

Figure 5.3 Replacement sequence design for siRNA 4: Replacement sequences were designed such that they would, in principle, be resistant to the effects of the siRNA whilst still encoding wild-type collagen VII protein. Synonymous nucleotide changes were made possible due to the degeneracy of the genetic code and taking into account codon usage of collagen type VII. Letters in bold indicate the 21-nucleotide wild-type COL7A1 target sequence of the siRNA. Letters in red indicate nucleotide changes from wild-type. Possible codons encoding the amino acid in question are listed, along with the fraction by which each codon is used throughout the COL7A1 gene. Two replacement sequences were generated for siRNA 4 (Rep4Max and Rep4) due to the high number of synonymous nucleotide changes that were possible.

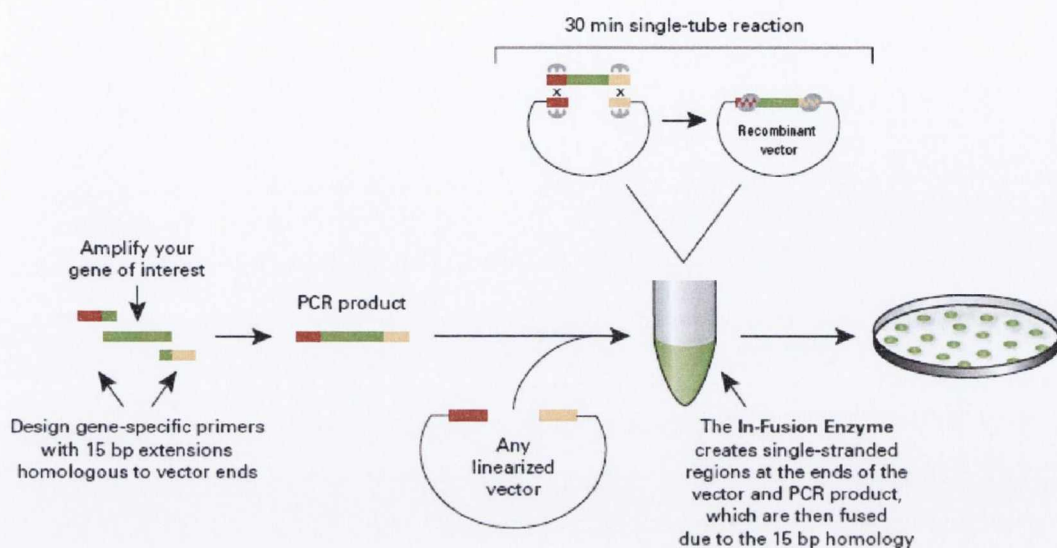
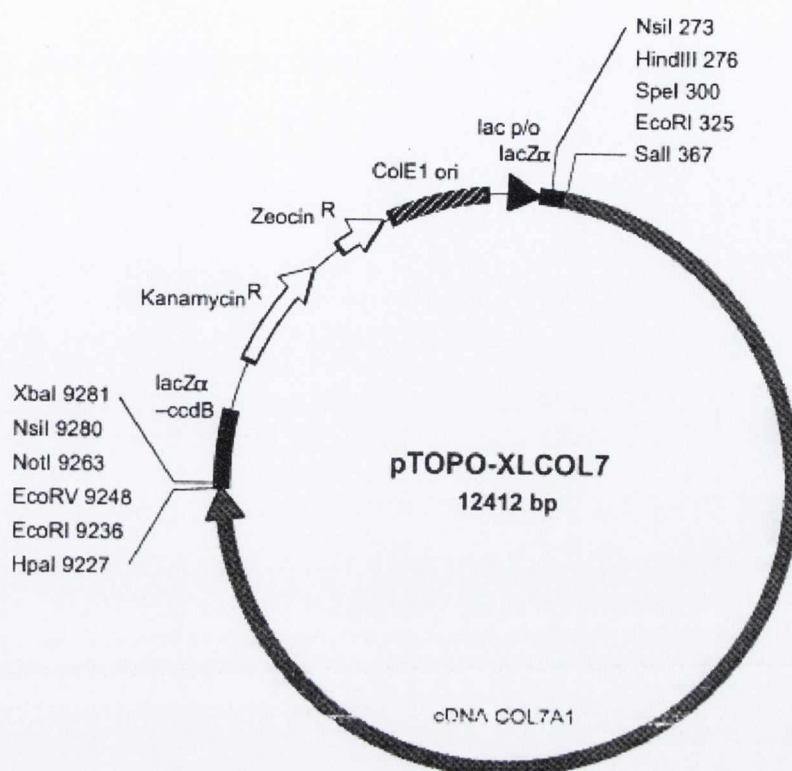


Figure 5.4 In-Fusion cloning method: Fragments of the desired insert were amplified by PCR, using primers with 15bp extensions homologous to the ends of the linearised pAcGFP1-N In-Fusion Ready vector. Both vector and PCR insert were then combined along with the In-Fusion enzyme which creates single-stranded regions at the ends of both the vector and PCR product, allowing for fusion of the homologous ends of the insert and vector. Following transformation, positive colonies were identified by sequencing (Figure derived from In-Fusion™ Advantage PCR Cloning Kit User Manual, 2010).

Primer name	Sequence 5'-3'
(a)	
Rep1 sense	AAGGCCTCTGTGCGACATGATCAAGAATGCTGACCCTGAGG
Rep1 antisense	AGAATTCGCAAGCTTAGTGGTCCGAGCTGTCCCG
Rep3 sense	AAGGCCTCTGTGCGACATGGTAGACACAGGACCTGGAGCCA
Rep3 antisense	AGAATTCGCAAGCTTACCTTCTCTCCCAGAGGCGCCTGAA
Rep4 sense	AAGGCCTCTGTGCGACATGGGCCAGGAACAGTCCCGG
Rep4 antisense	AGAATTCGCAAGCTTACGGGGCGGACACTGGCGT
pAcGFP1-N FW	GAGCTGGTTTAGTGAACCGTC
pAcGFP1-N Rev	ACGATGCCGGTGAACAGCTCG
(b)	
pTOPO-XLCOL7 sequencing FW	CTTTTGTCTATGGTGGCTGTGG
pTOPO-XLCOL7 sequencing Rev	AAATTGCTGCGGCCAATGGATG
PCR mutagenesis FW	GGCCCTCTGGGCCCCAACGGGGCCGCCGGGAAGCTGGGGACC
PCR mutagenesis Rev	GGTCCCCAGCTTTCCCGGCGGCCCCGTTGGGCCAGAGGGCC
pTOPO-XLCOL7Rep3 sequencing FW	GGACCGCGGACAAGAGGGTC
pTOPO-XLCOL7Rep3 sequencing Rev	GGTCCTGCAGGGCCAGCAAG

Figure 5.5 Chapter 5 primer sequences: (a) Primers used for the PCR of GeneArt fragments WT1, Rep1Max, Rep1 (Rep1 sense and Rep1 antisense), WT3 and Rep3 (Rep3 sense and Rep3 antisense) and WT4, Rep4Max and Rep4 (Rep5 sense and Rep4 antisense). Primers had 15bp of homology with the ends of the pAcGFP1-N In-Fusion linearised vector, to allow for the cloning of the PCR products into the vector. pAcGFP1-N FW and Rev primers were used for verification of the correct sequences and orientation following cloning of PCR fragments into the pAcGFP1-N vector. (b) pTOPO-XLCOL7 sequencing FW and Rev primers were used for sequencing of the pTOPO-COL7 construct and verification of the correct sequence and orientation following cloning of the CMV promoter into the plasmid. PCR mutagenesis FW and Rev primers were used for generation of pTOPO-XLCOL7Rep3 by PCR mutagenesis. pTOPO-XLCOL7Rep3 FW and Rev primers were used for verification of generation of the full-length replacement COL7A1 gene (pTOPO-XLCOL7Rep3) following site-directed PCR mutagenesis.



Human collagen VII cDNA (8894 bp)	337–9230
– Translation initiation codon	390–392
– Signal peptide region	390–458
– Stop codon	9222–9224

Figure 5.6 Vector map of pTOPO-XLCOL7: The full-length COL7A1 gene was kindly provided by Dr Alain Hovnanian of Institut National de la Santé et de la Recherche Médicale, Toulouse, France.

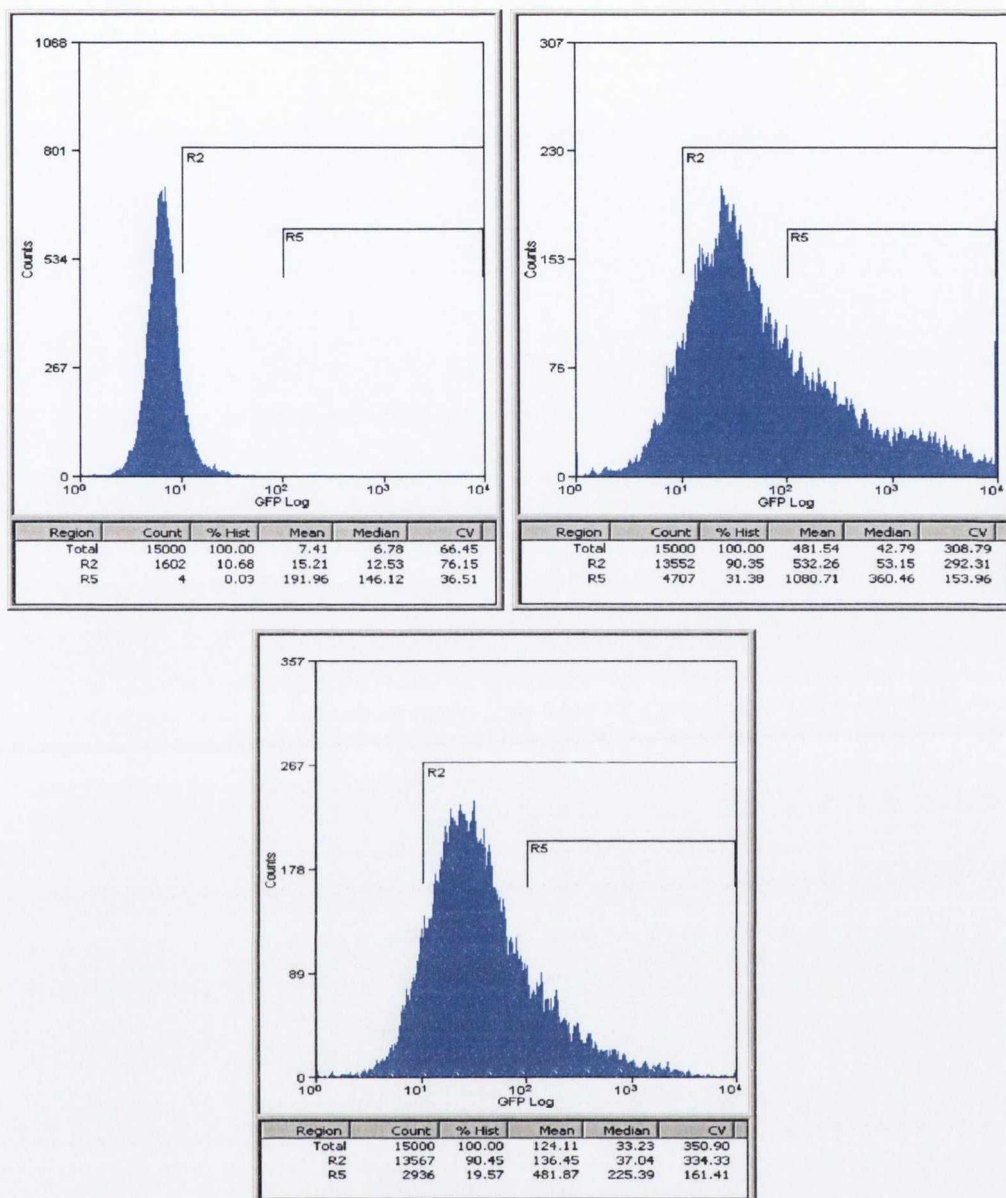


Figure 5.7 Distribution of HEK293 cells during flow cytometry analysis: Example of flow cytometry analysis of a co-transfection experiment with HEK293 cells. Cells were either (a) untreated, (b) co-transfected with a GFP-expressing plasmid and a non-targeting siRNA or (c) co-transfected with a GFP-expressing plasmid and a targeting siRNA. Histograms show the distribution of cells (blue populations) and the number of cells versus GFP intensity. Mean fluorescence intensities (7.41, 481.54 and 124.11 in figures a, b and c respectively) of the total cell populations were used for data analysis.

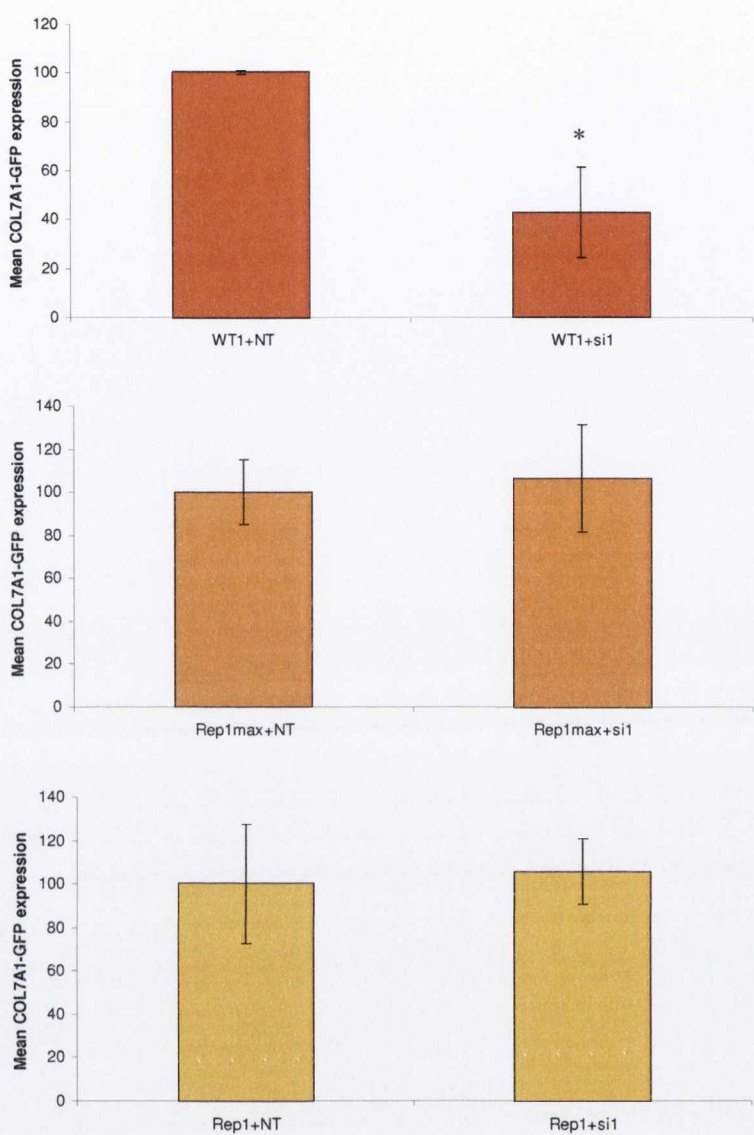


Figure 5.8 Mean changes in COL7A1-GFP expression: effect of siRNA 1 on WT1-AcGFP, Rep1Max-AcGFP and Rep1-AcGFP vector expression. Error bars represent standard deviation and stars represent statistical significance. HEK293 cells were co-transfected with plasmid DNA as well as either siRNA 1 or NT siRNA. GFP intensity of samples was measured by flow cytometry and data obtained from cells treated with siRNA were normalised to the values obtained from cells treated with NT siRNA. Each co-transfection was carried out on at least 3 separate occasions. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. siRNA 1 significantly reduced WT1-AcGFP by mean levels of 57.3% ($p = 0.006$). Neither Rep1Max-AcGFP nor Rep1-AcGFP were statistically significantly suppressed by siRNA 1, suggesting resistance of both constructs to the effects of siRNA 1 ($p = 0.721$ and $p = 0.759$ respectively).

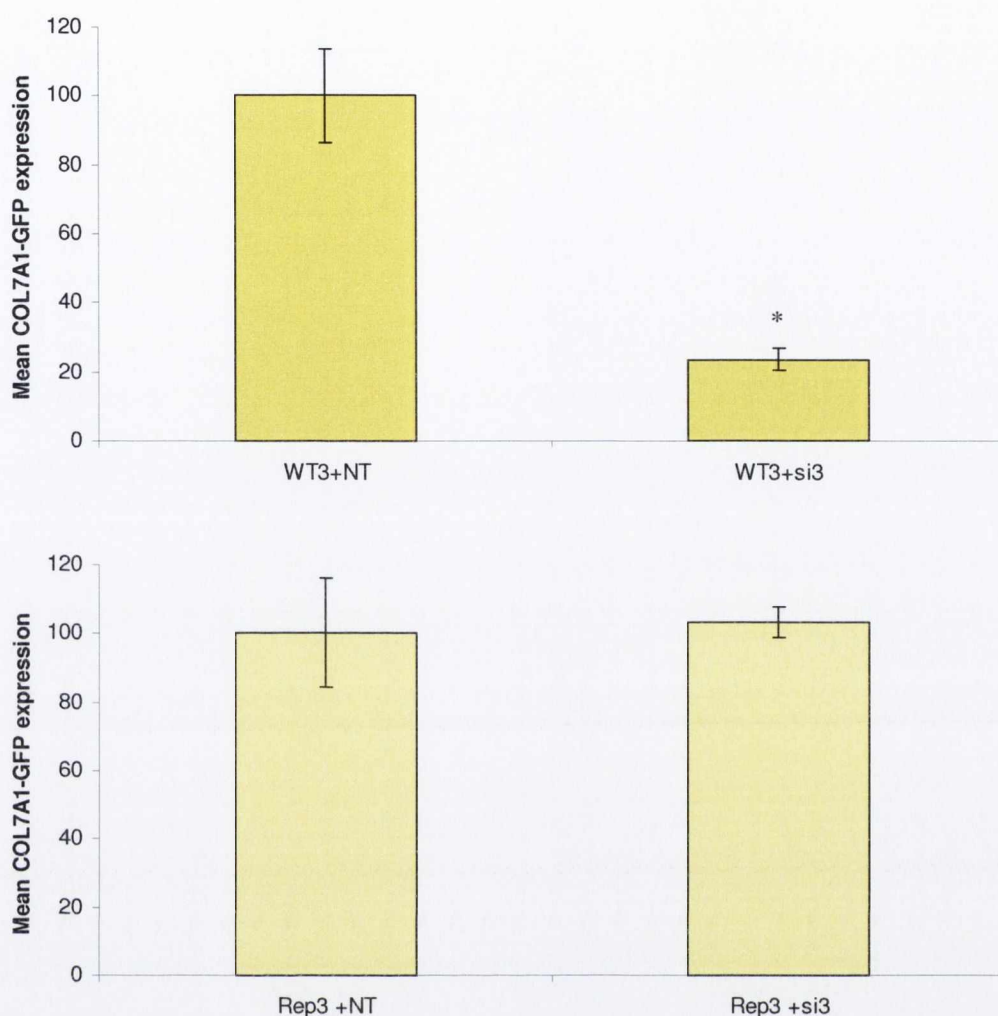


Figure 5.9 Mean changes in COL7A1-GFP expression: effect of siRNA 3 on WT1-AcGFP and Rep3-AcGFP expression. Error bars represent standard deviation and stars represent statistical significance. HEK293 cells were co-transfected with plasmid DNA as well as either siRNA 3 or NT siRNA. GFP intensity of samples was measured by flow cytometry and data obtained from cells treated with siRNA 3 were normalised to the values obtained from cells treated with NT siRNA. Each co-transfection was carried out on at least 3 separate occasions. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. siRNA 3 significantly reduced WT3-AcGFP by mean levels of 76.45% ($p = 0.001$). Rep3-AcGFP was not statistically significantly suppressed by siRNA 3, suggesting its resistance to the effects of siRNA 3 ($p = 0.757$).

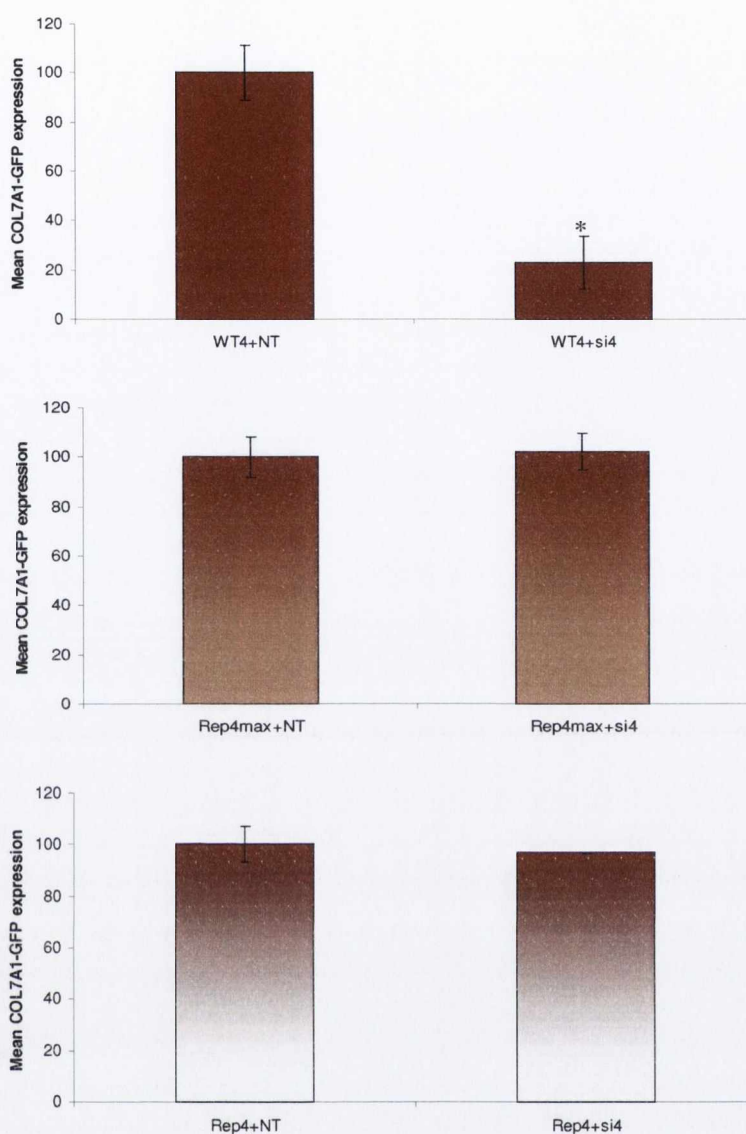
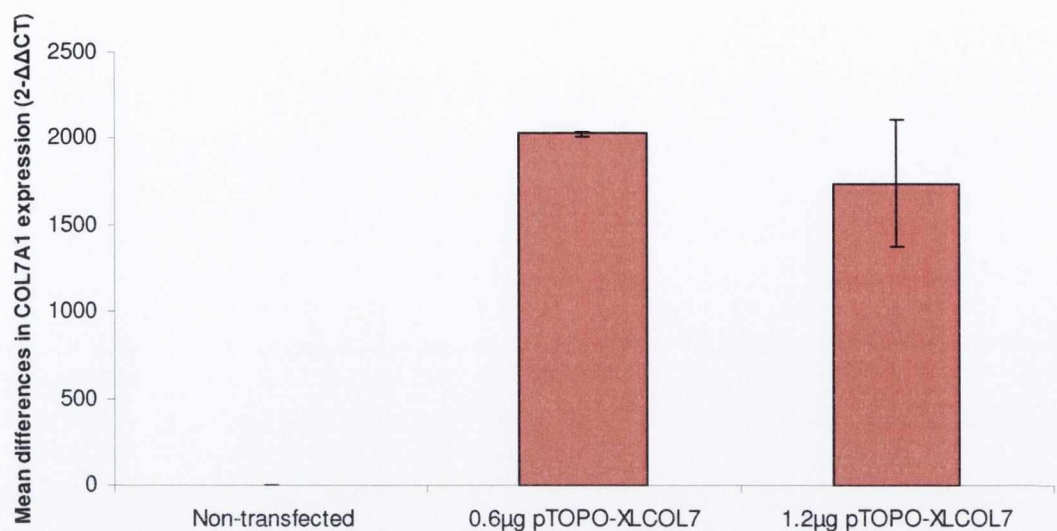


Figure 5.10 Mean changes in COL7A1-GFP expression: effect of siRNA 4 on WT4-AcGFP, Rep4Max-AcGFP and Rep4-AcGFP vector expression. Error bars represent standard deviation and stars represent statistical significance. HEK293 cells were co-transfected with plasmid DNA as well as either siRNA 4 or NT siRNA. GFP intensity of samples was measured by flow cytometry and data obtained from cells treated with siRNA were normalised to the values obtained from cells treated with NT siRNA. Each co-transfection was carried out on at least 3 separate occasions. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. siRNA 4 significantly reduced WT4-AcGFP by mean levels of 77% ($p = 0.001$). Neither Rep4Max-AcGFP nor Rep4-AcGFP were statistically significantly suppressed by siRNA 4, suggesting resistance of both constructs to the effects of siRNA4 ($p = 0.736$ and $p = 0.448$ respectively).



	Non-transfected	0.6μg pTOPO-XLCOL7	1.2μg pTOPO-XLCOL7
2-ΔΔCT	1.017698365	2008.934618	1998.00058
	1.000315438	2030.885571	1482.306396
Mean	1.009006902	2019.910095	1740.153488
St. Dev.	0.012291586	15.52166769	364.6508548

Figure 5.11 Mean differences in COL7A1 expression: HEK293 cells were transfected with either 0.6μg pTOPO-XLCOL7 DNA or 1.2μg pTOPO-XLCOL7 DNA and RNA was isolated from duplicate wells of transfected and non-transfected HEK293 cells 24 hours later. Levels of COL7A1 mRNA were increased by approximately 2000 fold in DNA-transfected cells compared to COL7A1 levels in non-transfected HEK293 cells.

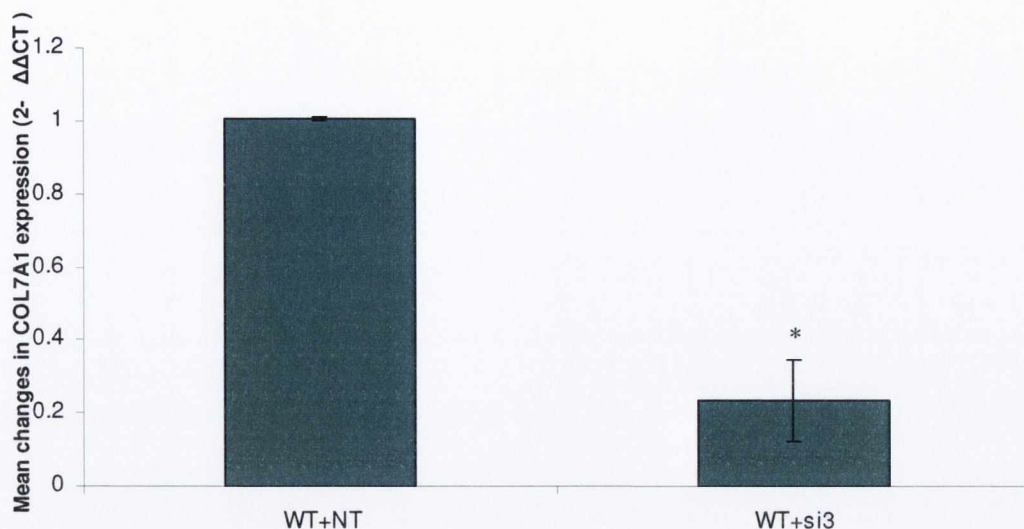


Figure 5.12 Mean changes in COL7A1 expression: effect of siRNA 3 on full-length COL7A1 gene expression in HEK293 cells. Error bars represent standard deviation and stars represent statistical significance. HEK293 cells were co-transfected with 0.4μg pTOPO-XLCOL7 (WT) and 1.25nM siRNA 3. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate and the experiment was carried out on three separate occasions. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant suppression of COL7A1 mRNA was achieved with siRNA 3 when compared to a non-targeting negative control siRNA ($p < 0.0005$). siRNA 3 reduced COL7A1 expression by an approximate average of 77%.

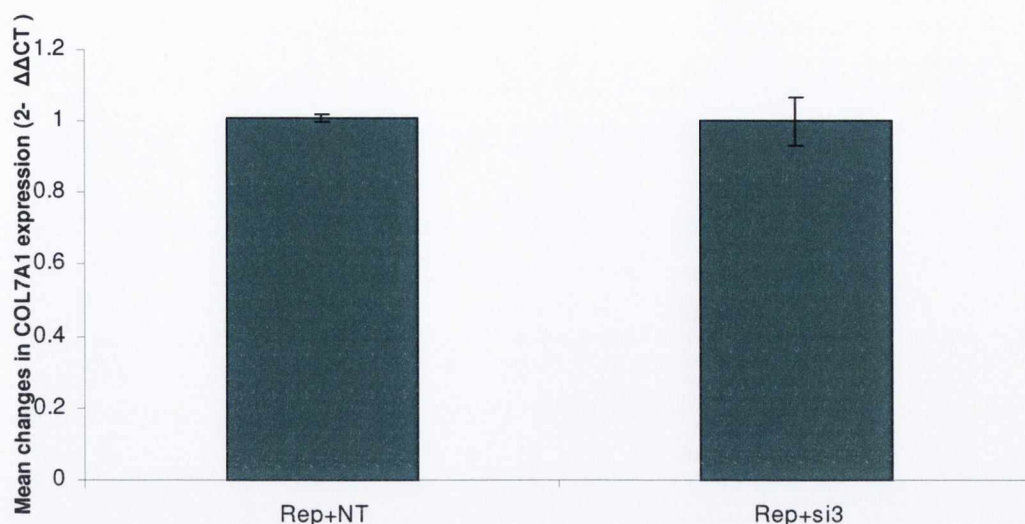


Figure 5.13 Mean changes in COL7A1 expression: effect of siRNA 3 on COL7A1 replacement gene expression in HEK293 cells. Error bars represent standard deviation. HEK293 cells were co-transfected with 0.4μg pTOPO-XLCOL7Rep3 (Rep) and 1.25nM siRNA 3. 24 hours later, RNA was isolated from transfected cells and COL7A1 mRNA levels were measured by real-time rtPCR. Transfection of each construct was undertaken in duplicate and the experiment was carried out on three separate occasions. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. No statistically significant differences in COL7A1 mRNA levels were found in cells treated with pTOPO-XLCOL7Rep3 and siRNA 3 when compared to those treated with pTOPO-XLCOL7Rep3 and NT siRNA ($p = 0.796$).

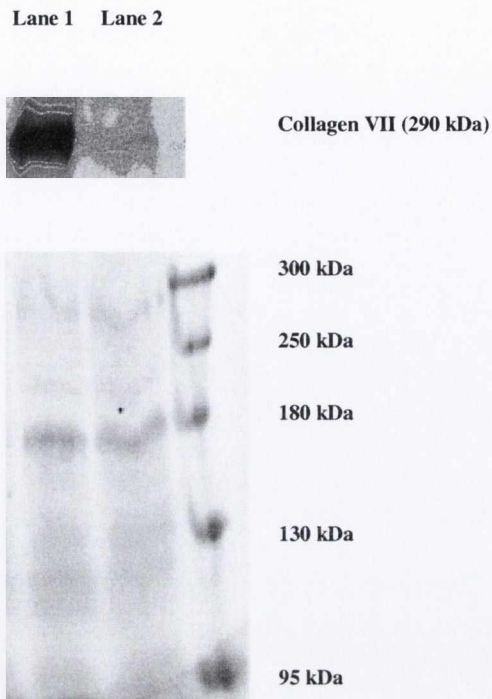


Figure 5.14 Mean changes in collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of HEK293 cells co-transfected with wild-type full-length COL7A1 and siRNA. Cells were co-transfected with pTOPO-XLCOL7 along with either NT siRNA (Lane 1) or siRNA 3 (Lane 2) using Attractene transfection reagent. Media protein samples were subjected to 6% SDS-PAGE. The position of the band corresponding to full-length 290kDa Collagen type VII is indicated. A separate gel was stained with PageBlue™ Protein Staining Solution for use as a loading control and protein ladder standard sizes are indicated. Transfections and western blot analyses were repeated on three separate occasions.

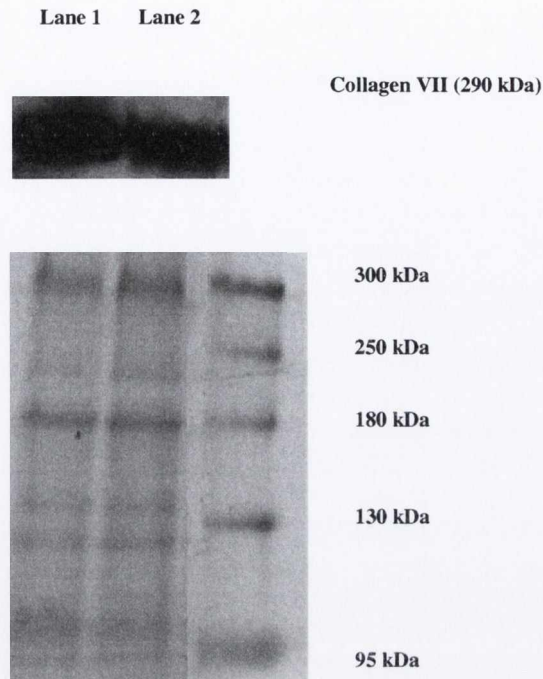


Figure 5.15 Mean changes in collagen type VII protein levels: western blot analysis of Collagen type VII protein from conditioned media of HEK293 cells co-transfected with the COL7A1 replacement gene and siRNA. Cells were co-transfected with pTOPO-XLCOL7Rep3 along with either siRNA 3 (Lane 1) or NT siRNA (Lane 2). Media protein samples were subjected to 6% SDS-PAGE. The position of the band corresponding to full-length 290kDa Collagen type VII is indicated. A separate gel was stained with PageBlue™ Protein Staining Solution for use as a loading control and protein ladder standard sizes are indicated. Transfections and western blot analyses were repeated on three separate occasions.

Chapter 6

Mutation-specific suppression

6.1 Introduction

The focus of the research described in this Ph.D. thesis to date has centred on exploration of a mutation-independent RNA interference-based approach for the dominantly inherited form of the genetic skin disorder Dystrophic Epidermolysis Bullosa. The approach in question involves attempted suppression of human COL7A1, a gene encoding type VII collagen protein. As discussed in chapter 1, collagen type VII assembles as anchoring fibrils in the basement membrane zone, providing structural strength and support for the epithelial layers of the human body (Chung and Uitto, 2010). Dominant COL7A1 mutations generally have a dominant negative effect, leading to the fragility of the epithelial layers associated with the DDEB phenotype (Nakamura *et al.*, 2004; Varki *et al.*, 2007). The allele-unspecific RNAi suppression approach under exploration in this study is likely to require concomitant replacement of the COL7A1 gene, due to the non-discriminatory nature of the RNAi suppressors regarding targeting of wild-type and mutant transcripts. Complete absence of the collagen VII protein would likely have a devastating effect, as seen from the severe blistering phenotype of recessive dystrophic EB patients who completely lack collagen type VII expression usually due to two nonsense COL7A1 mutations (Christiano *et al.*, 1997; Ohashi *et al.*, 2011; Uitto, 2011b). Identification of effective COL7A1 siRNA suppressors and generation of a replacement gene which is resistant to the effects of the most potent suppressor has been achieved *in vitro*, as discussed in chapters 4 and 5. However, an alternate RNAi-based approach is described in the current chapter of this thesis, which centres on exploration of an allele-specific gene therapy-based approach for particular dominant DEB-causing COL7A1 mutations. Such an RNAi approach theoretically should avoid the requirement for simultaneous replacement of the COL7A1 gene, due to the specificity of the suppressors in question. The method relies on identification of suppressors which target the mutant transcript for degradation but which are ineffective at achieving suppression of the equivalent wild-type allele.

In recent years, successful allele-specific silencing has been reported using RNAi technology (Ding *et al.*, 2003; Leachman *et al.*, 2010; Sierant *et al.*, 2011). This approach theoretically offers the potential for treatment of a wide range of dominant disorders, due to the ability of siRNA molecules to distinguish sequences which differ

by even a single nucleotide. Studies which have centred on either targeting of the mutant allele itself or a single nucleotide polymorphism associated with only the mutant allele have been investigated for a range of dominant disorders. Allele-specific silencing of the MJD1 gene, the mutant form of which causes the neurodegenerative Machado–Joseph disease/spinocerebellar ataxia type 3 (MJD/SCA3), has been achieved by taking advantage of a single nucleotide polymorphism in the MJD1 gene which has been found to segregate with the disease allele (Miller *et al.*, 2003; Alves *et al.*, 2008). Targeting of specific mutations themselves has also been achieved for a range of disorders including some cancers, neurodegenerative disorders and even skin diseases including recent initial investigations into allele-specific silencing for the Simplex subtype of Epidermolysis Bullosa (Brummelkamp *et al.*, 2002b; Schwarz *et al.*, 2006; Atkinson *et al.*, 2011). One of perhaps the most relevant cases which has involved the use of allele-specific gene silencing in recent years has involved research into the genetic skin disorder Pachyonychia Congenita (PC; Liao *et al.*, 2007), as discussed in section 1.7.3. There are several recurrent mutations of the KRT6A gene, the most common PC-causing mutation arising at codon N171, often resulting in an amino acid change (Hickerson *et al.*, 2006). Following identification of mutant-specific siRNAs for the dominant N171K mutation of the keratin KRT6A gene, rapid progress in this area has led to the completion of a phase 1b human clinical trial (Leachman *et al.*, 2010). The siRNA in question was found by performing a sequence walk along the mutation site, whereby all possible siRNAs targeting the mutant allele were tested and their effectiveness identified (Hickerson *et al.*, 2008; Leachman *et al.*, 2008). Potent suppression of the mutant allele was achieved with the TD101 siRNA, which had no effect on expression of the wild-type equivalent. In the 17-week clinical trial in which the TD101 siRNA was injected into the callused foot of a single patient, significant callus regression was reported. Thus the observed results of this study appear promising enough to merit further studies into allele-specific silencing using siRNA for both PC and other dominant-negative skin disorders.

The basis of the mutation-independent RNAi-based therapeutic approach which was discussed in chapters 3, 4 and 5 of this Ph.D. thesis may represent a valuable potential tool for treatment of dominant disorders, due to the fact that theoretically this approach would be applicable for any mutation of a given gene. As mentioned previously, this would be useful for heterogeneous disorders such as

dominant dystrophic EB, of which there are over 60 disease-causing dominant mutations (Nakamura *et al.*, 2004; Varki *et al.*, 2007; Van den Akker *et al.*, 2011). Indeed features such as the degeneracy of the genetic code enable the generation of mutation-independent suppression and replacement therapies for mutationally heterogeneous dominant disorders (Palfi *et al.*, 2006; Kubodera *et al.*, 2011; Millington-Ward *et al.*, 2011). However, allele-specific silencing represents an alternative possible and interesting gene-based therapy approach, whereby the therapeutic is designed to target individual mutations. Whilst in principle highly specific, the disadvantage of such an approach is that it limits the number of potential patients who may be treated as each suppressor is targeted to a specific mutation. Exploration of such an approach may indeed be worthwhile however, for targeting of the more commonly occurring mutations, which may represent a significant proportion of the total number of patients. Quick and efficient screening of a range of potential candidate siRNAs may be effectively carried out by linking of wild-type and mutant cDNAs to reporter constructs such as genes encoding fluorescent proteins. Effective siRNAs may then be characterised by fluorescence assays following plasmid co-transfection with candidate siRNAs (Hickerson *et al.*, 2008; Sierant *et al.*, 2011). The ability of siRNAs to distinguish between mutant and wild-type alleles in theory would allow for the treatment of many diseases caused by dominant mutations, assuming sufficient wild-type protein is produced from the untargeted wild-type allele. In the case of dominant DEB, a reduction in the levels of expression of collagen type VII protein to 50% of normal levels should still in principle be sufficient for a wild-type phenotype, considering those individuals who are heterozygous carriers of a nonsense COL7A1 mutation exhibit a normal non-blistering phenotype (Tidman and Eady, 1985).

The results presented in the current chapter of this Ph.D. thesis involve an exploration of an allele-specific therapeutic approach for dominantly inherited COL7A1-linked dystrophic EB. Similar to previous chapters in which a mutation-independent therapeutic approach was investigated for COL7A1-linked DDEB, the methodology employed in this case also involves the use of RNA interference technology. Two mutations in the COL7A1 gene were chosen for investigation of allele-specific suppression, both of which are known to cause a dominant DEB phenotype. One of the chosen mutations is a deletion mutation of 16bp which results

in an in-frame exon skipping (6863del16; Cserhalmi-Friedman *et al.*, 1998; Mellerio *et al.*, 1999; Dang *et al.*, 2007), whilst the second is a common point mutation which results in replacement of a glycine residue with an arginine amino acid (G2043R; Christiano *et al.*, 1995a; Mellerio *et al.*, 1998; Murata *et al.*, 2004). The former was chosen as it involves a significant change in nucleotide sequence from wild-type (16bp deletion), therefore hopefully facilitating the characterisation of a suppressor which discriminates effectively between the mutant and wild-type COL7A1 sequences. The latter was chosen due to its relatively frequent occurrence, representing the most common COL7A1 DEB mutation (Wessagowit *et al.*, 2001). Screening of all possible siRNAs for both mutations was carried out in an effort to identify an siRNA which effectively suppresses the mutant allele, whilst not significantly altering wild-type COL7A1 gene expression.

6.2 Materials and Methods

6.2.1 siRNA design

Two mutations of the COL7A1 gene, known to cause a dominant dystrophic EB phenotype, were chosen for allele-specific gene silencing. Two sets of siRNAs were synthesised by Sigma Aldrich and designed to screen all possible target sequences containing either the 6863del16 or G2043R mutations. siRNAs were provided annealed and resuspended in nuclease-free H₂O to a concentration of 100µM. The 6863del16 COL7A1 mutation results in a 16bp deletion at cDNA nucleotide position 6863, resulting in in-frame exon skipping of the 69 bp exon 87 (Cserhalmi-Friedman *et al.*, 1998). The G>A G2043R point mutation at cDNA nucleotide position 6127 results in conversion of a glycine residue (GGG) to arginine (AGG) (Christiano *et al.*, 1995a). All siRNAs were designed in the laboratory by Dr. Danny Allen. For the 6863del16 deletion mutation, 17 siRNAs were designed in the 19+2 format, comprising 19 nucleotide duplexes with two 3'-thymine DNA overhangs (see figure 6.1 for full 6863del16 siRNA sequences). For the G2043R point mutation, 19 siRNAs were designed in the 21+2 format, comprising 21 nucleotide duplexes with two 3'-thymine DNA overhangs (see figure 6.2 for full G2043R siRNA sequences).

6.2.2 Allele-specific COL7A1 fragment design

Fragments of the COL7A1 gene, each encompassing approximately 300bp, were designed and ordered pre-synthesised from GeneArt AG by Dr. Danny Allen. 2 fragments encompassing exons 83-88 were ordered, one which encoded the wild-type sequence and one which contained the 6863del16 16bp deletion mutation. 2 fragments encompassing exons 73-75 were also ordered, one encoding the wild-type sequence and one containing the G2043R point mutation, resulting in substitution of a glycine residue for an arginine residue. Each GeneArt DNA sequence was designed with *NotI* and *XbaI* restriction sites towards the 5' and 3' ends respectively. Full ordered sequences are listed in Appendix IV. Pre-synthesised DNA were supplied in vector format, in either standard GeneArt pMA, pGA18 or pMA-RQ plasmids and resuspended to a concentration of 0.1µg/µl.

6.2.3 Mini COL7A1-GFP fusion vector construction

COL7A1-GFP fusion vectors for allele-specific silencing were constructed using the In-Fusion™ Advantage PCR Cloning Kit (Clontech), as outlined in figure 5.4 and figure 6.3. All GeneArt-ordered COL7A1 DNA fragments designed by Dr. Danny Allen were cloned into the pAcGFP1-N In-Fusion Ready vector, which allows for the expression of the COL7A1 sequence as a fusion protein to the N-terminus of AcGFP1, a Green Fluorescent Protein from *Aequorea coerulescens* (see section 5.3.2). The In-Fusion cloning method was utilised, in which PCR primers were designed which share 15 bases of homology with the sequence at the ends of the linearised pAcGFP1-N vector. These primers were used to amplify the COL7A1 DNA fragments in the GeneArt vectors, as detailed in section 2.2.1.10. Primer sequences are listed in figure 6.4. Following loading of a small sample of the PCR products on a 2% agarose gel to ensure that the correct sized fragments had been amplified, the amplified products were treated with Cloning Enhancer which removes background template DNA and PCR residue, thus eliminating the need to gel purify the PCR products prior to cloning (section 2.2.1.10).

Ligations were carried out using the kit-supplied In-Fusion enzyme and buffer, and subsequent transformations, as detailed in section 2.2.1.10, were plated on kanamycin-agar plates. A number of kanamycin-resistant colonies were chosen for mini plasmid DNA preparation (section 2.2.1.11) and plasmid sequences and orientation were verified using the pAcGFP1-N sequencing primers (figure 6.4). Validated positive clones were identified and transformations were again carried out using 5µl of a 1 in 50 dilution of the appropriate mini-preparation (section 2.2.1.9). The Qiagen High Speed Plasmid Maxi Kit was used for large-scale plasmid purification, as outlined in section 2.2.1.12. The concentration and purity of the DNA samples was determined using the NanoDrop™ 1000 Spectrophotometer (section 2.2.1.15) and the constructs were again sequenced as outlined above to guarantee that the expected DNA plasmid had been isolated. In total, 4 of these COL7A1-GFP fusion vectors were generated (denoted WT83-88-GFP, 6863del16-GFP, WT73-75-GFP and G2043R-GFP).

6.2.4 Co-transfection of COL7A1-GFP fusion vectors and siRNA

HEK293 cells were co-transfected with COL7A1-GFP fusion plasmid DNA and siRNA using Attractene transfection reagent (Qiagen). Co-transfections were carried out as described in section 2.2.2.7. Following seeding of 5×10^5 cells per well of a 6-well plate, the appropriate amount of plasmid DNA was diluted in serum-free DMEM to a total volume of 100 μ l and the appropriate amount of siRNA was added. Samples were mixed, treated with 4.5 μ l Attractene reagent and incubated for 10-15 minutes at room temperature. Cells were then treated with prepared samples and incubated under normal growth conditions. Co-transfections were first carried out with the mutant DNA constructs in order to isolate any potential suppressors. All 17 possible 6863del16-targeting siRNAs were screened against 6863del16-GFP plasmid DNA as well as the non-targeting negative control siRNA (NT siRNA) in duplicate wells. All 19 possible G2043R-targeting siRNAs were screened against G2043R-GFP as well as NT siRNA in duplicate wells. Any siRNAs which appeared to significantly suppress the mutant construct were then co-transfected with the equivalent wild-type construct. A positive control GFP siRNA (Qiagen) was also used in co-transfection experiments (see figure 6.4 for GFP siRNA target sequence).

6.2.5 Flow cytometry data analysis of co-transfected cells

24 or 48 hours following co-transfections, cells were collected and pelleted as previously described. Cells were resuspended in 300 μ l PBS and 100 μ l 4% paraformaldehyde and placed on ice until flow cytometry analysis was carried out, as detailed in section 2.2.2.9. Summit 4.3 software was used for the analysis of samples by flow cytometry. The mean fluorescent intensities (MFI) of GFP from cell samples were taken and used for data analysis. Data obtained from 15,000 cells treated with siRNA were normalised to the values obtained from 15,000 cells treated with NT siRNA.

6.2.6 Statistical analysis

Analysis of statistical significance was performed using a two-sample t-test and the Statistical Software Minipad 15. Values of $p < 0.05$ were considered statistically significant.

6.3 Results

6.3.1 siRNA design

Unlike the research described in each previous results chapter, the focus of the current chapter of this Ph.D. thesis centres on exploration of a mutation-specific RNAi-based therapeutic approach for DDEB. Two dominant COL7A1 mutations were chosen for investigation: 6863del16 and G2043R. As mentioned previously (section 6.2.1), the 6863del16 mutation results in deletion of 16 nucleotides at COL7A1 cDNA nucleotide position 6863. Whilst such a mutation was originally predicted to cause a frameshift and thus downstream premature termination codon, analysis at the mRNA level actually revealed that the intraexonic deletion leads to skipping of exon 87 and subsequent restoration of the open reading frame (Cserhalmi-Friedman *et al.*, 1998; Mellerio *et al.*, 1999; Dang *et al.*, 2007). This particular DDEB-causing mutation was chosen for investigation of allele-specific silencing as any suppressors of the mutant transcript identified would likely be ineffective against targeting of the wild-type transcript, due to marked differences in sequences between the two alleles. The sense sequences of the 17 possible siRNAs which may potentially differentiate between the mutant and wild-type alleles are listed in figure 6.1. The G2043R COL7A1 mutation, as mentioned previously (section 6.2.1) is caused by substitution of a glycine residue for an arginine in the triple-helical domain of COL7A1, due to a G-to-A transition at cDNA nucleotide position 6127 (Christiano *et al.*, 1995a; Mellerio *et al.*, 1998; Dang *et al.*, 2007; Jerábková *et al.*, 2010). The mutation in question was chosen for this attempted allele-specific silencing as it is reported to be the most common DDEB-causing COL7A1 mutation throughout the world (Wessagowit *et al.*, 2001), accounting for 12.5% of all cases of DDEB reported on the International Dystrophic Epidermolysis Bullosa Patient Registry (<http://www.col7a1.org/>). It also accounts for one of the five definitive cases of DDEB in the Irish population (personal communication with Sharon Foley, Our Lady's Children's Hospital, Crumlin). The sense sequences of the 19 possible siRNAs, which may potentially differentiate between the mutant and wild-type alleles, are listed in figure 6.2.

6.3.2 Mini COL7A1-GFP fusion vector construction

In order to screen both sets of all possible siRNAs so as to isolate any suppressors which block mutant but not wild-type COL7A1 expression in a quick and efficient

manner, vectors containing small fragments of the COL7A1 gene encompassing the target regions and fused to a GFP reporter gene were generated. Four fragments of the COL7A1 gene encompassing exons 83-88 and 73-75 containing both wild-type and mutant 6863del16 and G2043R sequences respectively were pre-synthesised, as described in section 6.2.2. These sequences were cloned into the pAcGFP1-N In-Fusion Ready vector using In-Fusion™ Advantage PCR, which allows for their expression as fusion proteins to the N-terminus of AcGFP, as discussed in section 5.3.2. Thus four vectors were generated, containing wild-type and mutant exon 83-88 and 73-75 COL7A1 fragments, which was confirmed by sequencing (figure 6.4).

6.3.3 Analysis of siRNA effect on mutant DNA expression

Following cloning of the relevant fragments of the COL7A1 gene into the pAcGFP1-N vector, testing of these fusion genes and the effects of the designed siRNAs was examined. As discussed previously (chapter 5), the fusion of the COL7A1 fragments to a fluorescent marker allows for a direct measurement of the expression levels of the COL7A1 fragment genes by measuring GFP intensity by flow cytometry. Co-transfection experiments were initially carried out in 6-well plates, with 0.6µg of 6863del16-GFP or G2043R-GFP DNA along with 1.25nM siRNA being delivered to HEK293 cells using the lipid-based Attractene transfection reagent. Cells were collected 24 hours post transfection and flow cytometry analysis was carried out. All 17 6863del16-targeting siRNAs as well as a non-targeting siRNA were co-transfected along with the 6863del16-GFP DNA in duplicate and all 19 G2043R-targeting siRNAs as well as a non-targeting siRNA were co-transfected along with the G2043R-GFP DNA in duplicate. A non-fluorescent sample of non-transfected cells was used as a control in order to determine the auto-fluorescence threshold and mean fluorescent intensities (MFI) of AcGFP for each treated sample were used for data analysis. The data obtained from DNA treated with the relevant siRNA was normalised to the data obtained from the same construct treated with the non-targeting siRNA.

Following co-transfection of 6863del16-GFP DNA along with all 17 targeting siRNAs and NT siRNA, none of the potential suppressors were found to significantly reduce expression of the COL7A1 sequence, as measured by GFP intensity during flow cytometry analysis 24 hours post transfection (figure 6.5). A single transfection

of each siRNA was subsequently carried out, this time using an increased concentration of siRNA (12.5nM) and with incubation of transfected cells for 48 hours prior to flow cytometry analysis. The non-targeting negative control siRNA as well as a GFP-targeting positive control siRNA were also co-transfected with 6863del16-GFP DNA in duplicate. Whilst DNA expression was reduced by approximately 76% with the GFP siRNA ($p=0.018$), the increase in siRNA concentration as well as the increased incubation period did not seem to significantly affect COL7A1-targeting siRNA suppression efficiency (figure 6.6). Although samples treated with siRNAs 5 and 6 did appear to have lower GFP intensities (approximately 61% and 60% GFP expression respectively compared to NT siRNA-treated cells), this was found to most likely be as a result of lower plasmid DNA transfection efficiencies in these particular samples on this particular occasion resulting in reduced mean GFP intensity, shown by repetition of transfection of siRNAs 5 and 6 in quadruplicate as well as the NT siRNA. This repeat transfection showed mean GFP intensities to be reduced by just 4% and 10% respectively ($p=0.305$ and $p=0.019$ respectively). siGFP on the other hand, reduced mean 6863del16-GFP expression by approximately 64% ($p<0.0005$) (figure 6.7).

Co-transfections of G2043R-GFP DNA were subsequently carried out, along with 1.25nM of each of the 19 designed siRNAs and NT siRNA in order to identify any potential suppressors of the mutant sequence. 24 hours later, GFP intensity was measured by flow cytometry. However none of the potential RNAi-based suppressors were found to significantly reduce expression of the COL7A1 sequence (figure 6.8). A single transfection of each siRNA was subsequently carried out, this time using an increased concentration of siRNA (12.5nM) and with incubation of transfected cells for 48 hours prior to flow cytometry analysis. The non-targeting negative control siRNA as well as a GFP-targeting positive control siRNA were also co-transfected with G2043R-GFP DNA in duplicate (figure 6.9). G2043R-GFP expression was reduced by 58% with the positive control GFP-targeting siRNA. siRNA 10 was found to be the most potent COL7A1 suppressor, with GFP intensity being reduced by 43% in siRNA 10-treated cells compared to those treated with NT siRNA. 12.5nM siRNA 10, siGFP or NT siRNA and 0.3 μ g G2043R-GFP DNA were subsequently co-transfected in triplicate and GFP intensity was measured 48 hours later by flow cytometry. The experiment was carried out on three separate occasions (figure 6.10).

The positive control siGFP reduced G2043R-GFP levels by approximately 58% ($p<0.0005$). Mean GFP intensity of siRNA 10-treated cells was also found to be statistically significantly reduced by approximately 45% compared to NT siRNA treated cells ($p<0.0005$).

6.3.4 Analysis of siRNA 10 effect on WT DNA expression

Following achievement of significant suppression of the G2043R mutant DNA sequence by siRNA 10 of approximately 45%, it was subsequently of interest to determine whether or not this RNAi suppressor was mutant-specific. Thus testing of the resistance of the wild-type sequence to the effects of siRNA 10 was carried out. 12.5nM siRNA 10, siGFP or NT siRNA and 0.3 μ g WT73-75-GFP DNA were co-transfected in triplicate and GFP intensity was measured 48 hours later by flow cytometry. The experiment was carried out on three separate occasions (figure 6.11). The positive control siGFP was found to reduce WT73-75-GFP levels by approximately 59% ($p<0.0005$). Mean GFP intensity of siRNA 10-treated cells was however also found to be statistically significantly reduced and thus WT73-75 DNA was found not to be resistant to the effects of the suppressor, although the suppression achieved was not to the extent of that achieved with mutant G2043R-GFP DNA. GFP intensity of WT73-75-GFP/siRNA 10-transfected cells was found to be reduced by approximately 20% compared to NT siRNA treated cells ($p<0.0005$). Thus overall the data suggest that siRNA 10 can suppress expression of the mutant allele by approximately 45% while the wild-type allele is suppressed by approximately 20%. Furthermore, the difference in the levels of suppression achieved by siRNA 10 of the mutant and wild-type constructs was statistically significant ($p<0.0005$).

6.4 Discussion

Dominantly inherited gene mutations have been implicated in a range of inherited disorders due to the toxic function effects caused by the defective protein. Often such mutations arise in genes encoding essential proteins, thus causing a disease phenotype. The dominantly inherited form of the genetic skin disorder dystrophic EB represents one such disease, which arises due to mutations of the collagen type VII protein-encoding COL7A1 gene. Exploitation of the method of RNA interference may represent one possible method of treatment of such dominant disorders, as discussed in section 1.6, by aiming to reduce expression of the mutant gene. Both mutation-independent and/or mutation-dependent approaches using this powerful RNAi-based molecular tool have been investigated for a range of such disorders, which have highlighted the ability of siRNAs to discriminate between highly homologous sequences and even at times single-nucleotide changes (Miller *et al.*, 2003; O'Reilly *et al.*, 2007; Li *et al.*, 2009; Atkinson *et al.*, 2011; Millington-Ward *et al.*, 2011; de Yñigo-Mojado L *et al.*, 2011). In general, treatment of such dominantly inherited disorders using either of these RNAi therapeutic strategies may require concomitant delivery of the wild-type protein which is resistant to the effects of the RNAi suppressor, or a treatment mechanism which selectively inhibits expression of the mutant allele without affecting wild-type allele expression, respectively. It is the latter approach which forms the basis of the research described in the current chapter of this thesis.

As discussed in section 6.3.1, two COL7A1 mutations which are known to cause a DDEB phenotype were chosen for investigation and for examination of their potential as a target for the development of an allele-specific RNAi therapeutic approach for DDEB. Following scrutiny of the published literature for all reported DDEB-causing COL7A1 mutations and their relative frequencies, the 16bp 6863del16 deletion mutation (Cserhalmi-Friedman *et al.*, 1998; Dang *et al.*, 2007) and G2043R point mutation (Christiano *et al.*, 1995a; Mellerio *et al.*, 1998; Murata *et al.*, 2004) were chosen for attempted allele-specific silencing. The 6863del16 mutation, which causes in-frame skipping of exon 87, was thought to represent one example of a COL7A1 mutation which in theory would be ideal for allele-specific silencing due to the nature of the mutation. The deletion of 16 nucleotides in the DNA sequence of the

mutant allele means that the two nucleotide sequences should differ enough such that any significant siRNA suppressor of the mutant transcript should in theory be ineffective at suppression of the equivalent wild-type allele (see figure 6.1). One particular study of the COL7A1 mutation database was carried out by the Dystrophic Epidermolysis Bullosa Research Association Molecular Diagnostics Laboratory (USA) and involved COL7A1 analysis of DNA from 310 DEB families (Varki *et al.*, 2007). Of the 48 (15.5%) of these families who were considered to be of the autosomal dominant DEB category, one (2.1%) of the patients was found to harbour the deletion mutation 6863del16. The second mutation chosen for attempted allele-specific silencing was the G2043R mutation, which as stated previously arises due to a G-to-A point mutation in exon 73 and causes substitution of a glycine residue with an arginine (Christiano *et al.*, 1995a; Mellerio *et al.*, 1998). This DDEB-causing mutation exemplifies the importance of glycine residues in every third position of the triple helical domain of type VII collagen protein, as discussed in section 1.4. The G2043R substitution mutation has been found to be recurrent among DDEB patients of multiple ethnic backgrounds (Christiano *et al.*, 1995a; Mellerio *et al.*, 1998; Dang *et al.*, 2007) and the COL7A1 mutation analysis study mentioned above (Varki *et al.*, 2007) also found G2043R to be a recurrent mutation, with an allelic frequency of 3.7% of all DEB patients (both dominant and recessive). In recent months, the International Dystrophic EB Patient Registry has been established containing a database of all DEB-associated COL7A1 mutations which have been published in the literature (<http://www.col7a1.org/>), as well as their resulting phenotypes. Of the 579 registered patients, 7 were found to harbour the 6863del16 mutation (1.2% of all registered DEB patients). 5 of these patients exhibited symptoms of DDEB (4.2% of all registered DDEB patients), whilst another 2 of these patients displayed symptoms of recessive DEB due to a second COL7A1 mutation. Additionally, 15 of these patients had the G2043R mutation, representing 2.6% of all registered DEB cases and 12.5% of all registered DDEB cases.

In order to identify effective inhibitors of mutant COL7A1, all possible siRNAs across the two mutation sites were designed and tested. For the 6863del16 mutation site, siRNAs of 19 nucleotides in length were designed. Due to the nature of the wild-type and mutant sequences, it was possible to design just 17 siRNAs which could potentially differentiate between wild-type and mutant transcripts. The first

nucleotide after the 16bp deletion in the mutant sequence is a cytosine, as is the first nucleotide of the non-deleted sequence which exists in the wild-type sequence, thus reducing the number of potential discriminatory siRNAs from 19 to 17 (see figure 6.1). In any case, it is of note the last two nucleotides of an siRNA have been found not to contribute to binding specificity and may actually slow down RNAi-mediated cleavage by slowing the release of products (Elbashir *et al.*, 2001c; Haley and Zamore, 2004). All siRNAs were designed with two 3'-thymine overhangs: it has previously been found that siRNA duplexes with unpaired nucleotides at the 3' ends are more efficient at RNAi-mediated target degradation than their blunt-ended counterparts (Elbashir *et al.*, 2001b). Furthermore, the use of DNA bases at the 3' ends of siRNAs are often chosen because as well as reducing the cost of RNA synthesis, they are also thought to enhance the RNase resistance of siRNAs (Elbashir *et al.*, 2001c). For the G2043R mutation site, siRNAs of 21 nucleotides in length were designed and tested. There are 19 potential siRNAs that could potentially distinguish mutant from wild-type transcripts and so 19 of such siRNAs which potentially would have the ability to distinguish mutant G2043R from wild-type transcripts were designed with two 3'-thymine overhangs and tested *in vitro* (figure 6.2).

Similar to the testing of mini COL7A1 replacement genes as described in chapter 5, two small fragments of the COL7A1 gene encompassing the two chosen mutation regions were cloned into the pAcGFP1-N vector, allowing for their expression as fusion proteins to the N-terminus of AcGFP and thus allowing for a direct measurement of the effects of the potential siRNA suppressors on protein expression. Co-transfection experiments of both DNA and siRNAs were carried out in HEK293 cells, a cell line which was readily available in the laboratory and which has a number of further advantages as discussed in chapter 5, including the fact that it is known to be relatively amenable to transfection (Thomas and Smart, 2005). Following transfections of all 17 6863del16-GFP-targeting siRNAs along with 6863del16-GFP DNA and the undertaking of flow cytometry analysis 24 hours later to measure GFP intensity, unfortunately none of the potential suppressors were found to significantly reduce COL7A1 and thus GFP expression (figure 6.5). Co-transfections were subsequently repeated under different parameters, using an increased concentration of siRNA as well as increasing the incubation period before flow cytometry analysis in an effort to increase the possibility of identifying any

potential siRNA suppressors. Whilst each siRNA was for this particular experiment transfected in just a single well and thus no measure of statistical significance could be carried out, the mean GFP intensity appeared to be significantly lower in siRNA 5 and siRNA 6-treated cells compared to those treated with the NT siRNA (figure 6.6). As it is possible that such a difference in GFP intensity could in theory be due to lower transfection efficiencies of plasmid DNA in these wells compared to the NT siRNA-treated well, siRNA 5, siRNA 6 and NT siRNA transfections were subsequently repeated in quadruplicate, under the same set of conditions. The results of this transfection confirmed the possibility that the original results were most likely due to lower transfection efficiencies (figure 6.7). Whilst no significant differences in GFP intensity between siRNA 5 and NT-treated cells was found, despite statistically significant differences in siRNA 6 and NT-treated cells, GFP intensity was reduced by mean levels of just 10%, a level of suppression which would be unlikely to be sufficient for a mutation-dependent based gene-therapy approach for DDEB patients carrying the 6863del16 mutation. As no potent suppressor of the 6863del16 allele was identified in this study, co-transfections of the wild-type 83-88-GFP and targeting siRNAs were not carried out.

The use of a positive control GFP-targeting siRNA in this experiment, with which up to 76% suppression of 6863del16-GFP DNA was achieved, suggests that it is not the experimental model in question which is somehow masking or preventing achievement of suppression of the transcript by any of the 17 tested 6863del16-targeting siRNAs. The fact that a small fragment of the COL7A1 gene encompassing just six of the 118 exons was used in this study instead of the full length gene does however allow for the possibility that a true representation of the suppression potential of these siRNAs has not been obtained, as differences such as those involving secondary structure could exist between the full length COL7A1 and small fragment COL7A1-GFP transcripts. As the efficiency of the mechanism of RNAi is thought to be influenced by the secondary structure of the mRNA in question and thus the accessibility of the siRNA to its target transcript (Luo and Chang, 2004; Ameres *et al.*, 2007; Shao *et al.*, 2007), the Mfold programme was used to predict the secondary structure of the 6863del16-GFP transcript (Zuker, 2003; <http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>). Analysis of the predicted structure plot did not however appear to highlight any major structural obstacles that

may prevent access of the siRNAs to the transcript's target region (figure 6.12). This suggests that it is likely that it is the actual siRNAs themselves which are limiting factors and are inefficient in the sequence-specific suppression of 6863del16-GFP. It is worth noting that the mutation-specific methodology employed in this study is similar to those which have been undertaken for other dominant genetic diseases including other skin disorders, resulting in successful identification of discriminatory mutant allele RNAi suppressors (Hickerson *et al.*, 2008; Leachman *et al.*, 2008; Atkinson *et al.*, 2011). Thus in principle this allele-specific approach in which a sequence walk of the mutation site is performed and a fluorescence-based reporter gene system is used, represents an appropriate therapeutic strategy worthy of investigation for many dominant disease-causing mutations.

Co-transfections were subsequently carried out of G2043R-GFP DNA and the 19 designed siRNAs in an attempt to identify potent inhibitors of the mutant G2043R allele. Similar to siRNA-6863del16-GFP co-transfections, flow cytometry analysis was carried out on duplicate samples of co-transfections of G2043R-GFP DNA with 1.25nM of each siRNA 24 hours post transfection (figure 6.8), as well as on single transfections of DNA along with 12.5nM of siRNA 48 hours after transfection (figure 6.9). Whilst none of the siRNAs were found to statistically significantly suppress the transcript when cells were analysed 24 hours post transfection, incubation of transfected cells for 48 hours appeared to be more valuable in the detection of the effects of the siRNAs. The results of this siRNA screening identified siRNA 10 as a significant G2043R-GFP inhibitor. By increasing the incubation time period by 24 hours, the intensity of GFP expression was found to increase, due to an accumulation of GFP expression. This in turn was found to allow for easier identification of siRNA-mediated suppression as differences in GFP expression were more marked between targeting and non-targeting siRNA-treated samples. Whilst siRNA 14 appeared to be the most potent suppressor following flow cytometry analysis 24 hours after transfection (approximately 39% G2043R-GFP suppression compared to NT/G2043R-GFP treated cells; figure 6.8), this was subsequently found to likely be due to low plasmid DNA transfection efficiencies in those particular wells on the occasion of that particular experiment resulting in lower GFP intensity, rather than RNAi-mediated suppression. Following subsequent co-transfections and incubation of the siRNA 14-treated co-transfection samples for 48 hours before flow cytometry

analysis, no significant difference in GFP intensity between siRNA 14/G2043R-GFP-treated cells in comparison to the NT siRNA/G2043R-GFP-co-transfected cells was observed (figure 6.9).

Following repetition of transfection of cells with G2043R-GFP along with siRNA 10 on three further occasions, mean statistically significant suppression of the mutant transcript of 45% was achieved ($p < 0.0005$). The suppression results achieved with siRNA 10 against the mutant transcript deemed it worthwhile to then examine the effects of siRNA 10 against the equivalent wild-type transcript in order to ensure its ineffectiveness at suppression of transcript levels of the wild-type allele. Co-transfections of 73-75-GFP DNA and siRNA 10 were carried out in triplicate on three occasions and despite the siRNA and transcript sequences not being fully complementary, a statistically significant reduction of mRNA levels of approximately 20% was still found ($p < 0.0005$). This maintenance of some suppression ability of the siRNA indicates that single base pair mismatches between the siRNA guide strand and target transcript may be tolerated. Studies have suggested that the ability of siRNAs to tolerate mismatches may be influenced by the particular nucleotide mispairing and the position along the guide strand of the mismatch in question (Schwarz *et al.*, 2006; Du *et al.*, 2005; Dykxhoorn *et al.*, 2006; De Yñigo-Mojado *et al.*, 2011). Purine:purine mismatches between the siRNA and target mRNA have been found to greatly reduce the efficiency of the suppressor, and it is thought that this is largely due to disruption of the central A-form helix between the mRNA and siRNA which is required for cleavage of the target mRNA (Chiu and Rana, 2002; Haley and Zamore, 2004). Indeed in the work described in chapter 5 of this thesis, the replacement sequence designed to be resistant to siRNA 3, Rep3, had incorporated 6 nucleotide changes from wild-type, two of which represented purine:purine mismatches. This altered sequence was found to display complete resistance to the effects of siRNA 3. In the case of this G2043R-directed allele specific silencing, the U:G mismatch of the siRNA antisense strand:mRNA represents a pyrimidine:purine binding. Such a change is less structurally drastic than a purine:purine or pyrimidine:pyrimidine mismatch and so this may explain the ability of siRNA 10 to maintain a significant level of suppression of the wild-type mRNA, albeit at a reduced level of efficiency compared to its suppression of the mutant transcript.

Furthermore, it has been reported that for a U nucleotide in an siRNA's guide strand, the most tolerated mismatch is a U:G mismatch (Huang *et al.*, 2009). As it is such a mismatch in question in the case of the binding of siRNA 10 to the 73-75-GFP transcript, this may further explain the suppression ability of the wild-type transcript by siRNA 10 despite incomplete complementarity. Mismatch position sensitivity also seems to apply to the degree of discrimination which an siRNA can impart, with mismatches 3' to the seed region of the siRNA (the seed region being nucleotides 2-8 of the guide strand, which are responsible for binding and recognition (Jackson *et al.*, 2003; Lewis *et al.*, 2005)), having been found to provide the most robust single nucleotide discrimination (Schwarz *et al.*, 2006; Huang *et al.*, 2009). In particular, a number of studies have found the nucleotide located at position 10 from the 5' end of the guide strand to confer high position-dependent single-nucleotide sensitivity (Dykxhoorn *et al.*, 2006; Schwarz *et al.*, 2006; Hickerson *et al.*, 2008; Huang *et al.*, 2009). In the case of 73-75-GFP-targeting siRNA 10, the mismatch is located at position 11 from the 5' end of the guide strand (see figure 6.2), a nucleotide position which has been reported to be less sensitive to mismatches than that of position 10 (Schwarz *et al.*, 2006; Huang *et al.*, 2009) and thus may allow for some level of suppression to occur, as was observed with the 20% reduction in 73-75-GFP mRNA levels by siRNA 10.

It is worth noting that the concentrations of siRNA used in the course of the work described in this chapter may not have been optimal for the desired allele-specific silencing. Other similar studies investigating allele-specific silencing for dominant skin disorders have generally tested siRNAs at concentrations ranging from 0.1nM to 6nM, with the optimal concentration for allele-specific silencing generally being approximately 1nM (Hickerson *et al.*, 2008; Atkinson *et al.*, 2011). The work described in the current chapter of this Ph.D thesis however involved testing of siRNAs at just two concentrations-1.25nM and 12.5nM. In the case of the 6863del16 studies, no significant suppression of the mutant allele was achieved with any siRNA at either of the concentrations tested. In the case of the G2043R allele studies, no significant suppression of the mutant allele was achieved with any of the siRNAs tested at the 1.25nM concentration. Increasing this concentration to 12.5nM however resulted in identification of a potent suppressor of the mutant transcript, siRNA 10, which suppressed the G2043R allele by 45%. However, at this concentration, a

reduction in wild-type allele levels of 20% was also visible with this siRNA. It is possible in this case that reducing the concentration of siRNA 10 to somewhere between 1.25nM and 12.5nM may result in maintenance of mutant allele suppression while reducing the suppression levels of the non-complementary wild-type transcript, and so may be worthy of future investigation.

Overall, if lowering the siRNA concentration were found not to alter the results obtained thus far, although siRNA 10 does not suppress the wild-type mRNA to the same extent as the mutant transcript, the suppression effect is significant and may potentially limit the use of this siRNA in any mutation-dependent therapeutic approach for G2043R-causing DDEB. Such a molecule, which has the ability to reduce wild-type expression by 20% and mutant expression by 45%, would in theory cause a shift in the ratio of wild-type: mutant transcript levels from 50:50 in the untreated disease form to approximately 60:40 in the treated form. Whilst just one-eighth of collagen VII trimeric molecules would consist solely of normal polypeptides in the DDEB state assuming equal expression of mutant and wild-type alleles (Dang and Murrell, 2008), this shift in ratios would theoretically increase the frequency of normal polypeptides to approximately one in every five. As discussed in section 3.4, it is not currently known what ratio would be necessary for amelioration or complete recovery of the disease phenotype and so it is still possible that siRNA 10 could provide some therapeutic benefit for DDEB patients harbouring the G2043R mutation.

6.5 Conclusion

A mutation-dependent approach for DDEB was explored during the course of the research described in the current chapter of this Ph.D. thesis, with two DDEB-causing COL7A1 mutations being chosen for investigation. Following generation of fusion vectors containing the relevant COL7A1 gene fragments fused to GFP, all possible siRNAs designed to target the mutant but not wild-type alleles were tested. Flow cytometry analyses following co-transfection of DNA and siRNAs found no sufficient suppressor of the 6863del16 deletion mutation transcript. For the G2043R mutation, the most common DDEB-causing COL7A1 mutation identified so far, one siRNA

was identified which suppressed the mutant transcript by approximately 45%. However, this suppressor also significantly reduced wild-type transcript levels by approximately 20%. Thus, whilst such a therapeutic molecule may potentially be useful in the amelioration of symptoms of DDEB, it is unlikely that it would lead to complete reversal of the DDEB phenotype in those patients harbouring the G2043R mutation.

WT COL7A1: 5'-GCCTGGCCCTGTCGGACCTAAAGGAGAACCTGGCCCCACGGGGGCCCTGG		
6868DEL16: 5'-GCCTGGCCCTGTCGGAC		CCCCACGGGGGCCCTGG
6863del16.1:	GCCUGGCCCUGUCGGAC	CCd[T]d[T]
6863del16.2:	CCUGGCCCUGUCGGAC	CCCd[T]d[T]
6863del16.3:	CUGGCCCUGUCGGAC	CCCCd[T]d[T]
6863del16.4:	UGGCCCUGUCGGAC	CCCCAd[T]d[T]
6863del16.5:	GGCCCUGUCGGAC	CCCCACd[T]d[T]
6863del16.6:	GCCCUGUCGGAC	CCCCACGd[T]d[T]
6863del16.7:	CCCUGUCGGAC	CCCCACGGd[T]d[T]
6863del16.8:	CCUGUCGGAC	CCCCACGGGd[T]d[T]
6863del16.9:	CUGUCGGAC	CCCCACGGGGd[T]d[T]
6863del16.10:	UGUCGGAC	CCCCACGGGGGd[T]d[T]
6863del16.11:	GUCGGAC	CCCCACGGGGGcd[T]d[T]
6863del16.12:	UCGGAC	CCCCACGGGGGGcd[T]d[T]
6863del16.13:	CGGAC	CCCCACGGGGGGCCcd[T]d[T]
6863del16.14:	GGAC	CCCCACGGGGGGCCcd[T]d[T]
6863del16.15:	GAC	CCCCACGGGGGGCCCCud[T]d[T]
6863del16.16:	AC	CCCCACGGGGGGCCCCUGd[T]d[T]
6863del16.17:	C	CCCCACGGGGGGCCCCUGGd[T]d[T]

Figure 6.1 6863del16 siRNA sense sequences: 17 siRNAs (19+2 format, comprising 19 nucleotide duplexes with two 3'-thymine DNA overhangs) were designed and synthesised to screen all possible target sequences containing the COL7A1 deletion mutation 6863del16 but not the wild-type sequence. Antisense sequences were designed to be complementary to the 19 nucleotide long sense target sequences, with two 3'-thymine DNA ends.

WT COL7A1: 5'-CTGGAAAGCCTGGTATTCCCGGGCTCCCAGGCAGGGCTGGG
G2043R: 5'-CTGGAAAGCCTGGTATTCCCAGGCTCCCAGGCAGGGCTGGG

G2043R.1: UGGAAAGCCUGGUAUUCCCaGd[T]d[T]
G2043R.2: GGAAAGCCUGGUAUUCCCaGGd[T]d[T]
G2043R.3: GAAAGCCUGGUAUUCCCaGGCd[T]d[T]
G2043R.4: AAAGCCUGGUAUUCCCaGGCUd[T]d[T]
G2043R.5: AAGCCUGGUAUUCCCaGGCUCd[T]d[T]
G2043R.6: AGCCUGGUAUUCCCaGGCUCCd[T]d[T]
G2043R.7: GCCUGGUAUUCCCaGGCUCCAd[T]d[T]
G2043R.8: CCUGGUAUUCCCaGGCUCCAd[T]d[T]
G2043R.9: CUGGUAUUCCCaGGCUCCAGd[T]d[T]
G2043R.10: UGGUAUUCCCaGGCUCCAGGd[T]d[T]
G2043R.11: GGUAUUCCCaGGCUCCAGGCd[T]d[T]
G2043R.12: GUAUUCCCaGGCUCCAGGCAd[T]d[T]
G2043R.13: UAUUCCCaGGCUCCAGGCAGd[T]d[T]
G2043R.14: AUUCCCaGGCUCCAGGCAGGd[T]d[T]
G2043R.15: UUCCCaGGCUCCAGGCAGGGd[T]d[T]
G2043R.16: UCCCaGGCUCCAGGCAGGGCd[T]d[T]
G2043R.17: CCCaGGCUCCAGGCAGGGCUd[T]d[T]
G2043R.18: CCaGGCUCCAGGCAGGGCUGd[T]d[T]
G2043R.19: CaGGCUCCAGGCAGGGCUGGd[T]d[T]

Figure 6.2 G2043R siRNA sense sequences: 19 siRNAs (21+2 format, comprising 21 nucleotide duplexes with two 3'-thymine DNA overhangs) were designed and synthesised to screen all possible target sequences containing the COL7A1 point mutation G2043R (G>A) but not the wild-type sequence. Antisense sequences were designed to be complementary to the 21 nucleotide long sense target sequences, with two 3'-thymine DNA ends.

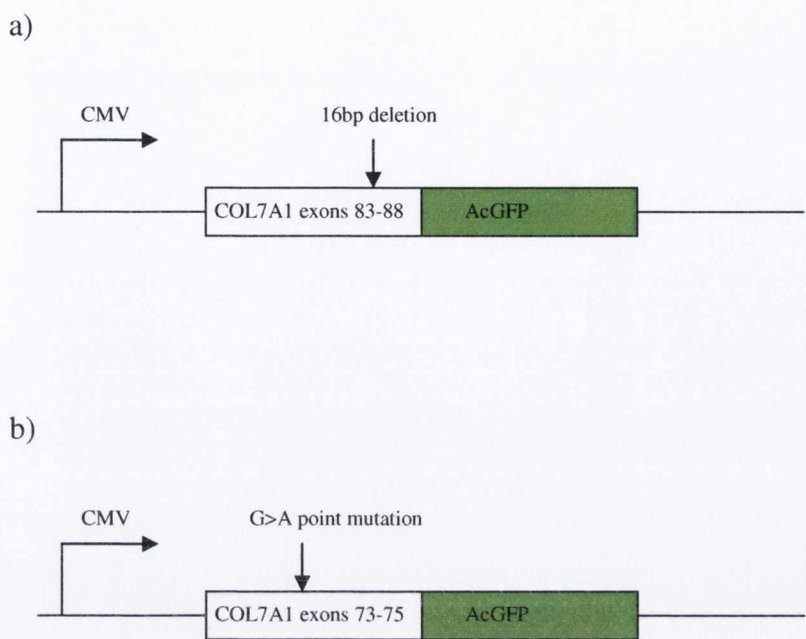


Figure 6.3 Schematic diagram of the COL7A1-AcGFP fusion expression constructs. Fragments of the COL7A1 gene comprising (a) exons 83-88 and (b) 73-75 were fused to the AcGFP gene using the pAcGFP1-N In-Fusion vector. Wild-type and mutant fusion constructs were created. Mutations, comprising a 16bp deletion (exons 83-88) and a G>A point mutation (exons 73-75) are indicated on the diagrams. Expression constructs were placed under the control of the CMV promoter.

Primer name	Sequence 5'-3'
Exons 83-88 sense	AAGGCCTCTGTGCGACATGGGGGAGCCTGGAGAG
WT Exons 83-88 antisense	AGAATTCGCAAGCTTCTTCTCTCCCTTTGCTCC
6863del16 antisense	AGAATTCGCAAGCTTACTTCTCTCCCTTTGCTC
Exons 73-75 sense	AAGGCCTCTGTGCGACATGGGCCCCATCGGCTTT
Exons 73-75 antisense	AGAATTCGCAAGCTTCTTGGGGCCAGGGGGTCC
pAcGFP1-N FW	GAGCTGGTTTAGTGAACCGTC
pAcGFP1-N Rev	ACGATGCCGGTGAACAGCTCG
siRNA	Target sequence 5'-3'
siGFP	AAGCAGCACGACTTCTTCAAG

Figure 6.4 Chapter 6 primer and control siRNA sequences: Primers used for the PCR of GeneArt fragments 83-88 (exons 83-88 sense and WT exons 83-88 antisense), 6863del16 (Exons 83-88 sense and 6863del16 antisense) and 73-75 and G2043R (exons 73-75 sense and exons 73-75 antisense). Primers had 15bp of homology with the ends of the pAcGFP1-N In-Fusion linearised vector, to allow for the cloning of the PCR products into the vector. pAcGFP1-N FW and Rev primers were used for verification of the correct sequences and orientation following cloning of PCR fragments into the pAcGFP1-N vector. siGFP was used as a positive control siRNA in co-transfection experiments.

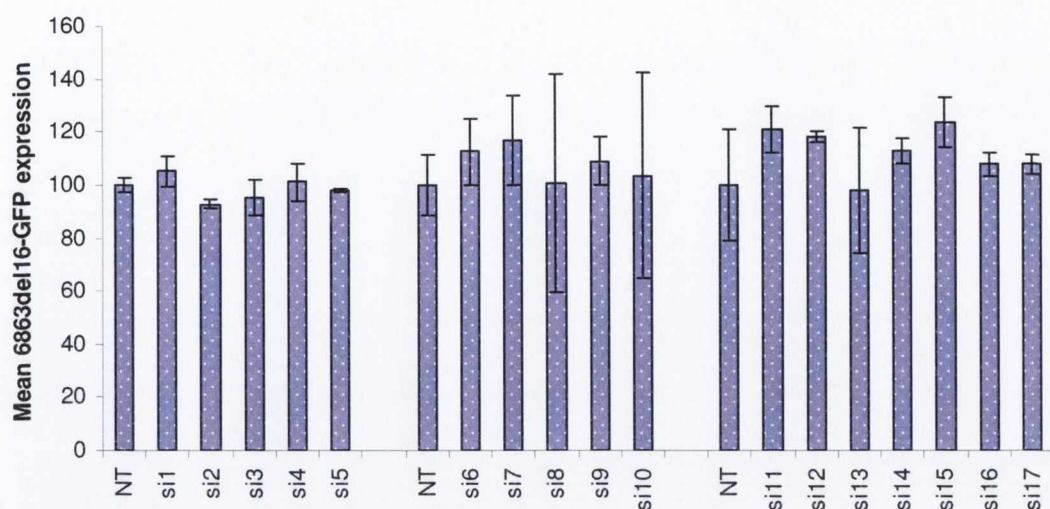


Figure 6.5 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 6863del16 mutation site (24 hours). Error bars represent standard deviation. HEK293 cells were co-transfected with plasmid DNA (6863del16-GFP) and 1.25nM of all 17 possible targeting siRNAs and the negative control NT siRNA in duplicate. GFP intensity of samples was measured by flow cytometry 24 hours post-transfection and data obtained from cells treated with siRNA were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant COL7A1 suppression was not achieved with any of the 17 tested siRNAs.

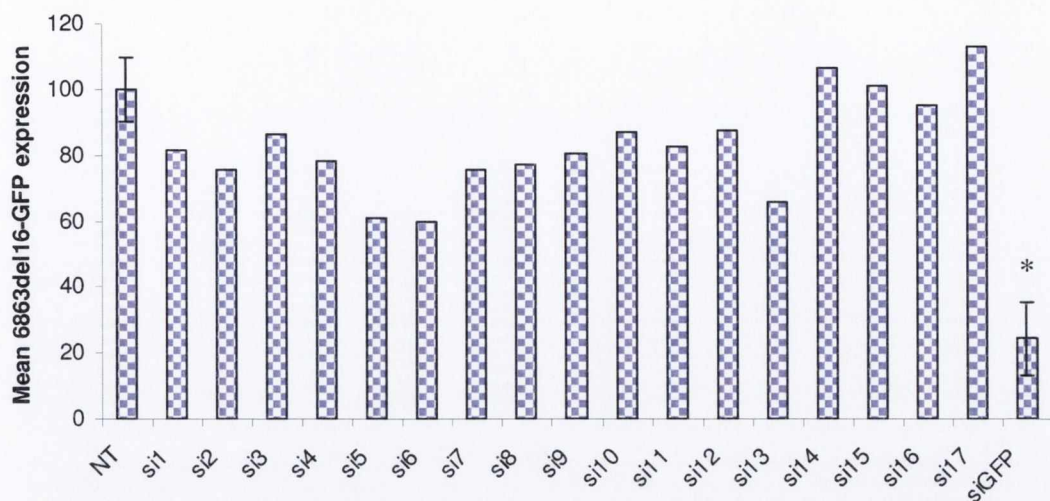


Figure 6.6 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 6863del16 mutation site (48 hours). Error bars represent standard deviation and stars indicate statistical significance. HEK293 cells were co-transfected with plasmid DNA (6863del16-GFP) as well as 12.5nM of all 17 possible targeting siRNAs. The negative control NT siRNA and a positive control GFP-targeting siRNA were also transfected in duplicate. GFP intensity of samples was measured by flow cytometry 48 hours post-transfection and data obtained from cells treated with siRNA were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. The positive control siGFP reduced 6863del16-GFP expression by approximately 76% ($p = 0.018$). Although samples treated with siRNAs 5 and 6 did appear to have lower GFP intensities (approximately 61% and 60% GFP expression respectively compared to NT siRNA-treated cells), this was found to most likely be as a result of lower plasmid DNA transfection efficiencies in these particular samples on this particular occasion resulting in reduced mean GFP intensity (see figure 6.7).

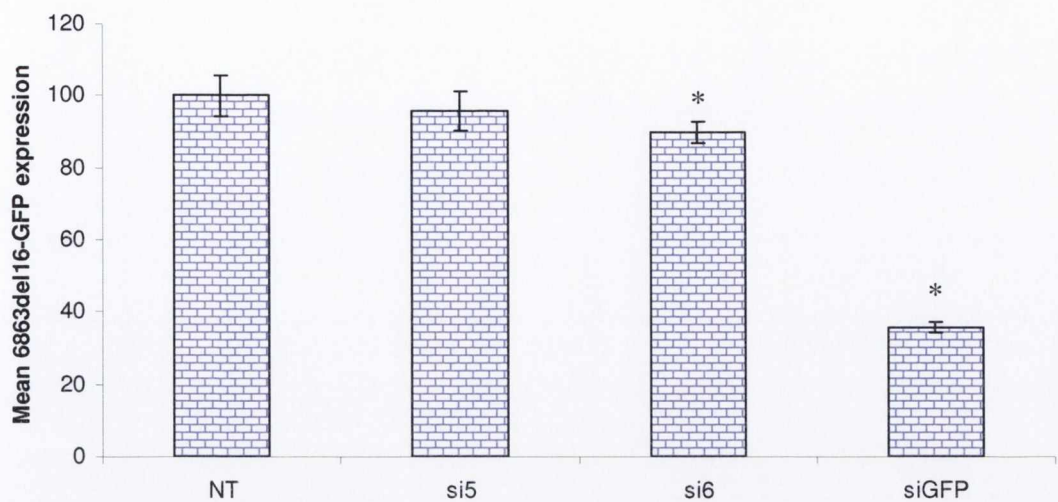


Figure 6.7 Mean changes in COL7A1-GFP expression: siRNA 5 and siRNA 6.

Error bars represent standard deviation and stars indicate statistical significance. HEK293 cells were co-transfected in quadruplicate with plasmid DNA (6863del16-GFP) as well as 12.5nM of siRNA 5, siRNA 6, siGFP or NT siRNA. GFP intensity of samples was measured by flow cytometry 48 hours post-transfection and data obtained from cells treated with siRNA 5, siRNA 6 and siGFP were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. The positive control siGFP reduced 6863del16-GFP expression by approximately 64% ($p < 0.0005$). siRNA 5 did not statistically significantly reduce 6863del16-GFP expression ($p = 0.305$) and while siRNA 6 did statistically significantly reduce 6863del16-GFP expression ($p = 0.019$), expression was reduced by mean levels of only 10%.

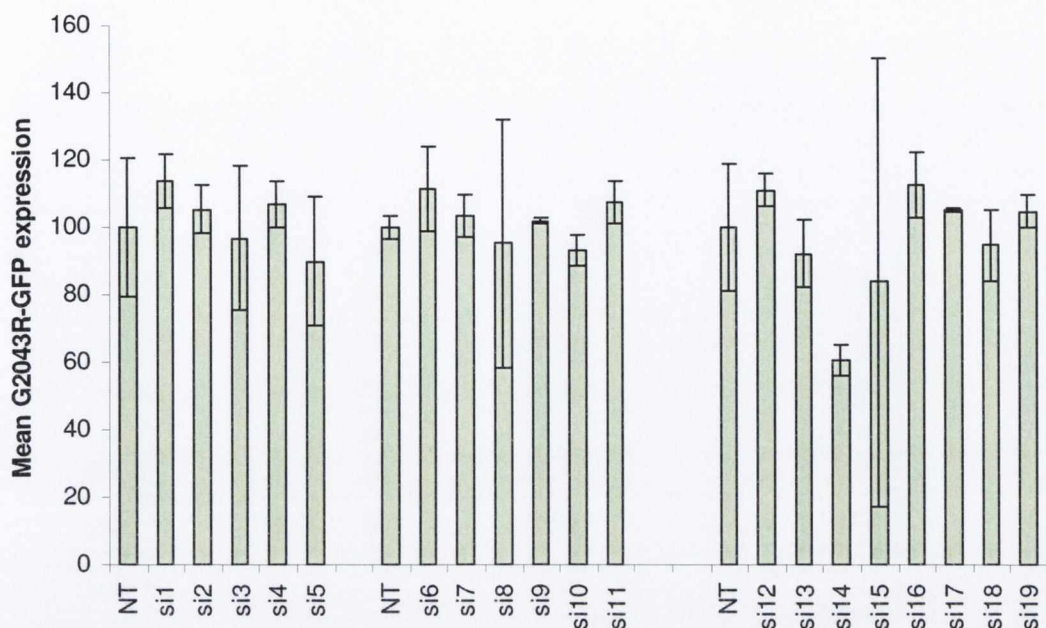


Figure 6.8 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 G2043R mutation site (24 hours). Error bars represent standard deviation. HEK293 cells were co-transfected with plasmid DNA (G2043R-GFP) and 1.25nM of all 19 possible targeting siRNAs and the negative control NT siRNA in duplicate. GFP intensity of samples was measured by flow cytometry 24 hours post-transfection and data obtained from cells treated with siRNA were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. Statistically significant COL7A1 suppression was not achieved with any of the 19 tested siRNAs.

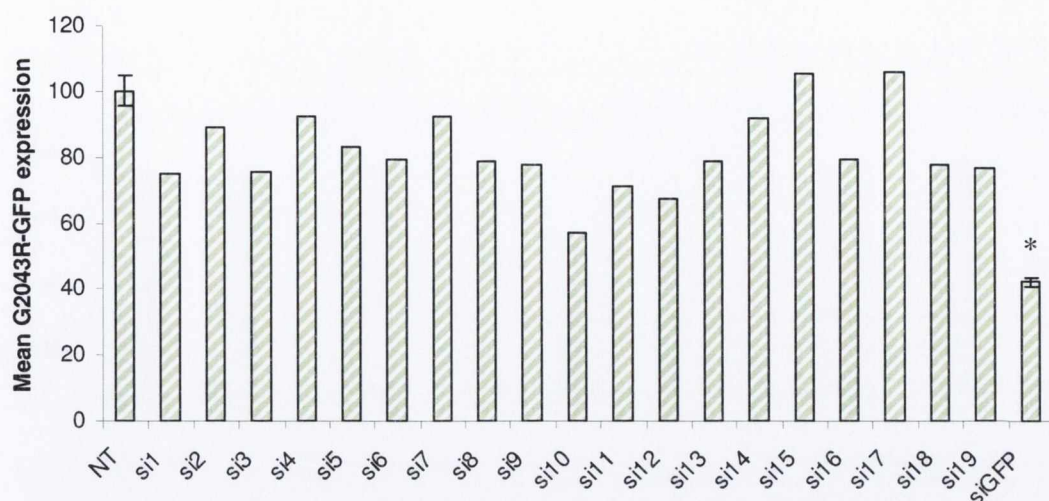


Figure 6.9 Mean changes in COL7A1-GFP expression: siRNA sequence walk of the COL7A1 G2043R mutation site (48 hours). Error bars represent standard deviation and stars indicate statistical significance. HEK293 cells were co-transfected with plasmid DNA (G2043R-GFP) as well as 12.5nM of all 19 possible targeting siRNAs. The negative control NT siRNA and a positive control GFP-targeting siRNA were also transfected in duplicate. GFP intensity of samples was measured by flow cytometry 48 hours post-transfection and data obtained from cells treated with siRNA were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. The positive control siGFP reduced G2043R-GFP expression by approximately 58% ($p = 0.003$). GFP intensity was reduced by 43% by the most potent COL7A1-targeting siRNA tested, siRNA 10.

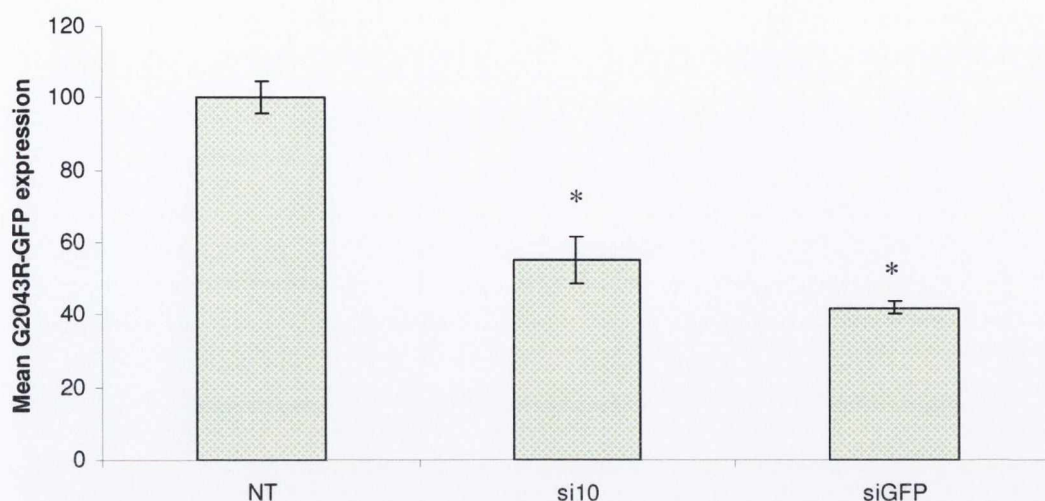


Figure 6.10 Mean changes in COL7A1-GFP expression: siRNA 10 and G2043R-GFP. Error bars represent standard deviation and stars indicate statistical significance. HEK293 cells were co-transfected in triplicate with plasmid DNA (G2043R-GFP) as well as 12.5nM of siRNA 10, siGFP or NT siRNA. GFP intensity of samples was measured by flow cytometry 48 hours post-transfection and data obtained from cells treated with siRNA 10 or siGFP were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. The positive control siGFP reduced G2043R-GFP expression by approximately 58% ($p < 0.0005$). siRNA 10 reduced G2043R-GFP expression by approximately 45% ($p < 0.0005$).

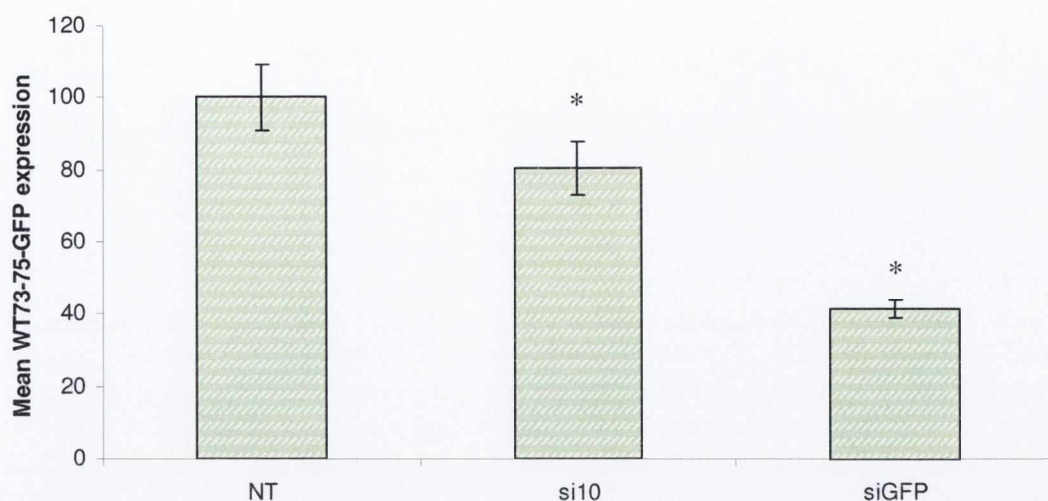


Figure 6.11 Mean changes in COL7A1-GFP expression: siRNA 10 and 73-75-GFP. Error bars represent standard deviation and stars indicate statistical significance. HEK293 cells were co-transfected in triplicate with plasmid DNA (73-75-GFP) as well as 12.5nM of siRNA 10, siGFP or NT siRNA. GFP intensity of samples was measured by flow cytometry 48 hours post-transfection and data obtained from cells treated with siRNA 10 or siGFP were normalised to the values obtained from cells treated with NT siRNA. Analysis of statistical significance was performed using a two-sample t-test and values of $p < 0.05$ were considered statistically significant. The positive control siGFP reduced WT73-75-GFP expression by approximately 59% ($p < 0.0005$). siRNA 10 reduced WT73-75-GFP expression by approximately 20% ($p < 0.0005$).

Chapter 7

General Discussion

7.1 General Discussion

The realisation in recent decades that small RNA molecules play an important role in the control of gene expression through Watson-Crick base-pairing with complementary messenger RNAs has led to marked advances in our understanding of eukaryotic gene function and regulation. The discovery of this phenomenon, known as RNA interference, has allowed for its exploitation in the uncovering of gene function as well as in the treatment of disease. Indeed, despite the fact that the RNAi process was only relatively recently unearthed with the discovery that double-stranded RNA may trigger down regulation of complementary mRNA in *Caenorhabditis elegans* (Fire *et al.*, 1998), extrapolation of the process to mammalian cells has already resulted in the undertaking of RNAi-based human clinical trials for treatment of human diseases including age-related macular degeneration and the skin disorder Pachyonychia Congenita (Kaiser *et al.*, 2010; Leachman *et al.*, 2010). The enormity of the discovery of RNAi was highlighted by the fact that the 2006 Nobel Prize in Physiology or Medicine was awarded to Andrew Fire and Craig Mello who played a major role in the elucidation of the mechanism of RNAi. The focus of the research carried out as described in this Ph.D. thesis has centred on utilisation of this powerful molecular tool RNAi, in an effort to investigate its potential in a gene-based therapy for the autosomal dominant disorder Dystrophic Epidermolysis Bullosa (DEB).

As described previously, EB is a rare genetic blistering skin disorder for which there is currently no known cure. The dystrophic form of the disease, the phenotype of which arises due to tissue separation at the dermal layer of skin, represents one of three major subtypes and a wide range of both dominant and recessive DEB-causing mutations have been identified (Varki *et al.*, 2007; Van den Akker *et al.*, 2011). Mutations of the COL7A1 gene are implicated, the function of which is to encode Type VII Collagen protein. Dominant DEB mutations result in production of an aberrant protein which has a dominant negative interfering effect, leading to a weakening of the integrity of the anchoring fibril structures which are pertinent for maintenance of normal skin function (McGrath *et al.*, 1993; Chung and Uitto, 2010). Thus the basis of the RNAi-themed research as described in this thesis is based on attempted suppression of expression of this dominant mutant protein in order to prevent its incorporation into the anchoring fibril structures where it exerts its

negative effect. Two separate RNAi-based methods were investigated in the course of this study, involving both a mutation-independent and a mutation-dependent therapeutic approach for DDEB. The existence of a wide range of DDEB-causing COL7A1 mutations highlights the relevance of the mutation-independent approach, a method which should in essence be applicable for the treatment of all DDEB patients, regardless of the mutation in question. Such an approach is based on RNAi targeting of both the mutant and wild-type transcripts and thus would require the simultaneous delivery of a replacement gene which is refractory to suppression but whilst still encoding wild-type protein. Meanwhile, the second mutation-dependent approach investigated involved an attempt to exploit the specificity of RNAi, by testing suppressors which are designed to target the mutant but not the wild-type alleles of two chosen DDEB-causing COL7A1 mutations. This retention of wild-type gene expression may theoretically be sufficient to result in a wild-type phenotype despite levels of wild-type Collagen VII protein being half those of normal levels and thus such an approach would not require concurrent delivery of a replacement COL7A1 gene. Such a mutation-specific approach would however limit the number of potential patients who would benefit.

The mutation-independent RNAi approach firstly involved attempted identification of potent suppressors of the COL7A1 gene. Figure 7.1 indicates the COL7A1 target regions of each potential suppressor tested. Of the 13 artificial microRNA constructs tested, which would be potentially beneficial mainly in terms of their safety profiles due to their similarity to the cells endogenous microRNAs, none of the constructs successfully reduced COL7A1 levels by any more than approximately 50%. A number of COL7A1-targeting siRNAs were subsequently tested *in vitro* in both human keratinocyte and fibroblast cell lines and suppression of up to 80% of the COL7A1 gene was achieved. Following conversion of three of the most potent of these siRNAs to artificial microRNA format, based on the same vector as the 13 originally tested, the suppression levels achieved did not resemble those obtained with the original three siRNAs. In fact, once again suppression of no higher than 50% was obtained. In this case it may be that the pcDNA 6.2-GW/ EmGFP-miR vector in question, which produces transcripts based on the murine pri-miR-155 sequence (Lagos-Quintana *et al.*, 2002), may not be capable of achieving higher levels of suppression, regardless of the sequence of the mature siRNA end product.

This may possibly be due to the intracellular processing of the construct, which may yield low levels of siRNAs and thus would be unable to potently suppress the COL7A1 gene. It would be possible to explore some aspects of this using for example RNase protection assays, however the focus of the work was to characterise a potent RNAi suppressor for COL7A1 and hence further research on underlying reasons for absence of potent suppression was not undertaken. It is possible that the use of another artificial microRNA vector, based on a different endogenous microRNA, may facilitate more potent COL7A1 suppression. The use of an artificial microRNA scaffold based on human microRNA-30 has reportedly resulted in potent reduction of up to 70% of the human Huntington's disease (HD) gene (McBride *et al.*, 2008) and it is possible that the delivery of the COL7A1 siRNAs using this miRNA-30 shuttle may achieve higher COL7A1 knockdown and thus possibly more favourable results.

Conversion of the siRNAs to shRNA format also, in some cases, resulted in a reduction of suppression efficacy in comparison to the suppression achieved with direct delivery of siRNAs. Whilst the potency of siRNA 1 was retained in shRNA format, shRNAs 3 and 4 were not quite as effective as their siRNA counterparts (the data in both cases was statistically significant; $p=0.002$ when comparing differences in observed levels of COL7A1 suppression between siRNA 3 and shRNA 3 and $p=0.001$ when comparing differences in observed levels of COL7A1 suppression between siRNA 4 and shRNA 4), possibly due to limitations in the processing of these two shRNAs, as discussed in section 4.4. Whilst each class of artificial inhibitory RNAs tested was capable of mediating COL7A1 gene silencing, albeit at somewhat different levels, differences also exist between siRNAs and shRNAs or artificial microRNAs, namely due to the delivery method in question. This may affect the duration of both expression and thus gene silencing. The mode of delivery of a potential RNAi-based therapeutic largely depends on the target tissue or cell population and to date both viral and nonviral approaches have been employed to achieve RNAi-mediated gene silencing (Boudreau and Davidson, 2010). Delivery of RNAi may be accomplished by direct injection of siRNAs, either naked or complexed with lipid-based transfection reagents (Wolff and Bukder, 2005; Kim *et al.*, 2009; Krebs and Alsberg, 2011) although such an approach may be limited by the ability of siRNAs to diffuse throughout the tissue and enter its target cells. Additionally, the half-life of the siRNA affects the duration of suppression and it is likely that in a

therapeutic setting, repeated injections may be required in order to maintain gene silencing. Conversely, in some cases it may be possible to achieve long-term gene silencing from a single injection of stem-loop RNAs, although the ability of the RNAi to diffuse within tissue and transduce target cells may also be limited in both viral and non-viral delivery formats (Raoul *et al.*, 2006). As well as representing a possible method of achieving long-term gene silencing, viral delivery of either shRNA or artificial microRNA molecules to target tissues would also need to be safe and efficient in order to exploit this endogenous machinery for therapeutic purposes. Viral vector systems based on adeno-associated vectors (AAV) and lentiviruses have been found to represent suitable mediators of RNAi as they have the ability to safely transduce a wide range of tissues as well as provide sustained levels of gene expression (Couto and High, 2010; Sliva and Schnierle, 2010). However whilst AAV vectors are thought to be apathogenic to humans, they are also non-integrating and so their delivery information may be lost during repetitive cell division (Koerber *et al.*, 2008; Sliva and Schnierle, 2010). Lentiviral vectors in contrast, stably integrate their genomes into the host cell and so may be more applicable for long-term therapeutic gene expression in organs such as the skin (Teo *et al.*, 2009). In recent years, safer versions of such lentiviral vectors have been developed, in which strong viral enhancers present in the long terminal repeat (LTR) regions of the lentiviral genome have been deleted (termed self-inactivating or SIN vectors). This should in theory prevent potentially dangerous altered gene expression in the vicinity of the integration site by these viral enhancers and as transgene expression is driven by cellular promoters, this should allow for increased safety during delivery of the potential therapeutic (Di Nunzio *et al.*, 2008; Almarza *et al.*, 2011).

In the case of delivery of RNAi to the skin, the main target organ for any potential DEB therapeutic, it is notable that some success has been achieved with the direct delivery of siRNAs to the skin by injection, in a Phase 1b clinical trial for the dominant skin disorder Pachyonychia Congenita (PC) (Leachman *et al.*, 2010). The symptoms displayed by affected patients include painful plantar calluses and a dose-escalated series of intralesional injections of siRNA into a callus on the foot of a single patient to the level of the superficial dermis was carried out twice weekly over a 17 week time period. This resulted in significant callus regression, providing evidence for the potential effectiveness of siRNA delivery directly to the skin. One of

the drawbacks of such a method, which would likely be mirrored in a similar delivery approach for treatment of DDEB, was the degree of pain experienced by the patient during the injection process. An alternative delivery mechanism such as non-invasive topical delivery of the therapeutic would be ideal, although for a delivery mechanism which requires intradermal injection of the therapeutic, a reduction in the number of required injections would also likely be beneficial. A range of approaches for *in vivo* delivery of siRNA have been investigated and reported, including use of lipoplexes, liposomes and polymer-based Nanoparticles (NPs) (De Fougères, 2008; Whitehead *et al.*, 2009) and more recently a biodegradable polymer NP has been tested and found to result in sustained release of siRNA following intradermal injection to the footpads of mice (Jacobson *et al.*, 2010). In this particular study the biodegradable polymer poly(L-lactic acid) (L-PLA) together with a copolymer of L-PLA-poly(ethylene glycol) (PEG) were used, with supercritical carbon dioxide (SC-CO₂) being used as an anti-solvent for particle formation and encapsulation. The addition of PEG to L-PLA was intended to improve *in vivo* circulation and reduce agglomeration of the NPs in solution. The initial findings of this study have shown high encapsulation efficiency and slow release of an siRNA mimic for up to 80 days *in vivo*, thus providing initial promising evidence that such NPs may prove to be suitable therapeutic agents for delivery of siRNAs and thus control of gene expression in the skin. Other potential delivery mechanisms of siRNA to skin include the use of microneedles, jet injection and sonophoresis (Geusens *et al.*, 2009).

Following identification of potent suppressors of the COL7A1 gene the next step required in this mutation-independent therapeutic approach involved generation of a replacement collagen type VII-encoding gene, as described in chapter 5 of this thesis. Replacement sequences were designed for siRNAs 1, 3 and 4 such that they would encode wild-type collagen protein, whilst the nucleotide sequence would differ from the wild-type version of the gene such that the siRNAs would be ineffective suppressors of the replacement gene. Originally the testing of these replacement sequences and their resistance to the relevant siRNAs was carried out using small fragments of the COL7A1 gene as opposed to the full-length sequence. These small fragments were cloned into the pAcGFP1-N vector such that they would be expressed as fusion proteins to AcGFP. Thus a measure of siRNA resistance could be directly measured by flow cytometry analysis of GFP intensity of cells co-transfected with

plasmid DNA and siRNA. All tested sequences appeared to display resistance to their equivalent siRNAs and so further to these studies, it was decided to progress with one of the replacement sequences for generation of a full-length COL7A1 replacement gene. The gene in question incorporated the nucleotide changes such that it would be resistant to RNAi-targeting by siRNA 3, one of the most potent of the suppressors described in chapter 4 of this thesis. Real-time rtPCR and western blot analysis confirmed the resistance of the full-length replacement COL7A1 gene, denoted pTOPO-XLCOL7Rep3, to the suppressor effects of siRNA 3.

During analysis of the effects of the full-length replacement gene, expression of the gene *in vitro* was achieved by the placing of the gene sequence under the control of the CMV promoter. This promoter was chosen largely due to its ability to control high level constitutive gene activity in a variety of cell lines (Boshart *et al.*, 1985; Nelson *et al.*, 1987). In a therapeutic setting however, the use of another more appropriate promoter sequence may be required. The use of a promoter which is as short as possible while still maintaining adequate levels of expression of the gene in both basal keratinocytes and dermal fibroblasts would be ideal and the COL7A1 promoter itself represents one such potential sequence (Vindevoghel *et al.*, 1997; Titeux *et al.*, 2010). The 616bp COL7A1 promoter sequence contains all the essential functional elements required and has been shown to ensure both tissue-specific and regulated expression of type VII collagen protein (Verrecchia *et al.*, 2001; Pendaries *et al.*, 2003). More specific expression of the gene may also be required, and the human keratin 14 (K14) or mouse pro- α 2 chain of type I collagen (col1a2) promoters may be useful in allowing for restricted transgene expression to either the epidermal (keratinocyte) or dermal (fibroblast) skin layers respectively (Ito *et al.*, 2009). The mode of delivery of the full-length replacement gene is also something which would need to be considered in any therapeutic approach for DDEB. Extensive research has been carried out in an effort to deliver full-length COL7A1 for treatment of DEB patients suffering from the recessive form of the disease, as discussed in section 1.5.1 and any success in these studies could likely be extrapolated to delivery of the replacement COL7A1 gene for DDEB patients. Due to the large size of the COL7A1 transcript however, difficulties have arisen during attempted delivery of the full-length gene. Recent studies using retroviral vectors including self-inactivating (SIN) retroviral vectors has yielded some success however (Siprashvili *et al.*, 2010; Titeux

et al., 2010) as detailed in chapter 1, with the results of one *in vivo* study leading to commencement of a phase I human clinical trial investigating *ex vivo* gene delivery of COL7A1 (Siprashvili *et al.*, 2010). Thus although the results of this clinical trial have yet to be reported, these findings are extremely promising for the delivery of COL7A1 for the treatment of not only RDEB patients, but also for those suffering from the dominantly inherited form of the disease and who may therefore require a COL7A1 suppression and replacement-based therapeutic strategy.

The research carried out as described in this Ph.D. thesis has involved *in vitro* studies investigating the potential of an RNAi-based therapeutic for DDEB. Undoubtedly further to these studies it would be worthwhile to progress to testing of the RNAi in an *in vivo* model in order to ensure their efficacy, safety and their phenotypic effects. To date such studies have been hampered due to the absence of the relevant DDEB animal model. More specifically, an animal model presenting with an autosomal dominant DEB phenotype due to the presence of a mutant human COL7A1 transgene would ideally be required. Such an animal model would allow for the *in vivo* evaluation of the potency of the RNAi suppressor and replacement constructs relevant to the treatment of human patients. At present there exist a number of transgenic mice developed as models for DEB, but each of these mice recapitulate the features of the recessive form of the disease (Bruckner-Tuderman *et al.*, 2010). Such mice, including the collagen VII knockout mouse (Heinonen *et al.*, 1999) or the more recently developed collagen VII hypomorphic mouse (Fritsch *et al.*, 2008) which has a longer life span than the knockout mouse, may be useful for the testing of the full-length replacement gene generated as described in chapter 5 of this thesis. For analysis of the effects of the COL7A1 suppressor molecules, it may also be worthwhile to attempt their delivery by intradermal injection to a mouse model which expresses the human COL7A1 gene, such as that described by Ito *et al.*, 2009. Thus if a phenotype of blistering were to be observed following delivery of the RNAi, evidence for the *in vivo* activity of the RNAi would potentially be demonstrated. For overall evaluation of this mutation-independent suppression and replacement strategy however, the availability of a DDEB mouse model would be extremely invaluable.

In parallel to the mutation-independent RNAi approach which was carried out as detailed in chapters 3, 4 and 5 of this Ph.D. thesis, a mutation-specific RNAi

approach was also investigated. As discussed above and in detail in chapter 6, central to such a methodology is the requirement to identify a potent and specific inhibitor which reduces expression of the mutant allele whilst retaining wild-type allele expression. Thus the need for concurrent delivery of a replacement gene would likely be extraneous in such an allele-specific approach, due largely to the fact that levels of COL7A1 gene expression of half those of wild-type levels may be tolerated, as demonstrated by the fact that heterozygous carriers of a nonsense COL7A1 mutation exhibit a wild-type non-blistering phenotype (Tidman and Eady, 1985). Of the range of approximately 60 DDEB-causing COL7A1 mutations so far identified (Nakamura *et al.*, 2004) two were chosen for examination of allele-specific silencing in this particular study. Unfortunately the first DDEB-causing 16bp deletion mutation which was examined, denoted 6863del16, resulted in failure to identify any potent inhibitor despite testing 17 such siRNAs which were complementary to the mutant deletion allele, ruling out the possibility of progression towards further development of a potential allele-specific therapeutic for this particular dominantly inherited COL7A1 mutation. The second mutation investigated, a point mutation denoted G2043R, represents the most commonly occurring DDEB-causing mutation (Wessagowit *et al.*, 2001) and some degree of success was achieved in this initial allele-specific COL7A1 RNAi silencing study. A single siRNA was found to suppress the mutant COL7A1 transcript expression by approximately 45%, although suppression of the wild-type allele of approximately 20% was also found to be significant. Thus in the disease state the use of such a suppressor in principle would lead to a shift in the balance of wild-type: mutant transcript ratios from 50:50 to approximately 60:40. As the ratio of wild-type:mutant transcript required for amelioration of the disease phenotype is currently unknown, it is not quite known if such a suppressor would be in any way beneficial for use in a mutation-specific RNAi-based therapeutic approach for the G2043R mutation. However it would of course be preferable to characterise a more potent suppressor. Further to these studies, it may be beneficial to test the effect of this siRNA suppressor in a human skin cell line isolated from a DDEB patient harboring the G2043R mutation and thus expressing the full-length mutant transcript, in an effort to investigate whether the levels of suppression achieved of both mutant and wild-type transcripts resemble those found with the experimental model in question as described in chapter 6. As mentioned above, the existence of a DDEB animal model would be extremely valuable and required for development of any gene-based therapy

for DDEB. As the G2043R mutation has so far been reported to be the most commonly occurring mutation, the generation of a transgenic animal model harboring this mutation may be extremely relevant. As well as allowing for investigations of the mutation-independent RNAi-based therapeutics for DDEB as discussed in chapters 4 and 5 of this thesis, such a model would also be useful for analysis of the phenotypic affects of the suppressor identified in this allele-specific silencing study. It is also of note that only two COL7A1 mutations implicated in DDEB were tested in the current study, despite the existence of approximately 60 such mutations (Nakamura *et al.*, 2004; Van den Akker *et al.*, 2011). Thus it is entirely possible that greater success in allele-specific silencing may be achieved with some of these other mutations and so investigations into allele-specific RNAi-based therapeutics for some of the more common of these mutations may be worthwhile.

During analysis of the literature for reported DDEB-causing COL7A1 mutations and examination of a database of DEB-associated COL7A1 mutations (<http://www.col7a1.org/>) during the course of the research carried out as described in this thesis, a region of the COL7A1 gene in which clustering of dominant dystrophic EB-causing mutations occurs was observed. In the region of the gene encompassing exons 70-75, approximately 50% of all dominant dystrophic EB-causing mutations were found to reside. Furthermore, these represent some of the most frequently occurring dominant mutations, accounting for 82.9% of all cases in one particular study of 48 families classified as having autosomal dominant dystrophic EB (Varki *et al.*, 2007). Furthermore, in this particular study this region also accounted for approximately 13% of all recessive COL7A1 mutations. Thus it was noted that a gene correction-based therapeutic may represent a promising alternative potential therapeutic option for DEB which may be worthy of investigation in the future. One possible strategy which may be employed is based on the relatively new zinc finger nuclease technology. Zinc finger nucleases are a class of DNA-binding proteins which facilitate editing of the genome by creating double-stranded breaks in DNA (Kandavelou *et al.*, 2009; Rémy *et al.*, 2010). In theory, zinc finger nucleases offer a means of delivering a site-specific double-stranded break to any genomic site (Kim *et al.*, 1996; Reyon *et al.*, 2011). One mechanism by which eukaryotic cells heal such breaks is by homology-directed repair: induction of site-specific double-stranded chromosomal breaks has been shown to stimulate homology-directed gene repair, an

otherwise highly accurate but infrequent mechanism (Van Nierop *et al.*, 2009). Thus in theory it is possible to replace a region of mutated DNA by delivery of a copy of a corrected version, which is incorporated into the chromosomal site by homologous recombination (Kandavelou *et al.*, 2009; Li *et al.*, 2011). Sigma-Aldrich is one commercial company who provide a service whereby a single pair of zinc finger nucleases may be designed, assembled and validated to edit a target gene of interest. Therefore, it is noteworthy that in the future it may be worthwhile to investigate utilisation of zinc finger nucleases designed to target the COL7A1 gene, thus potentially allowing for correction of mutant COL7A1 DNA sequences via homologous recombination. In principle, targeted correction of exons 70-75 may possibly represent a viable therapeutic for greater than 50% of all dominant dystrophic EB patients. Hence, it may be possible to undertake a gene therapy approach which utilises zinc finger nucleases to create a double stranded break around this region, thus allowing for targeted correction of exons 70-75 by homologous recombination with a donor wild-type molecule. Such an approach would represent a potential therapeutic model for treatment of a significant number of cases of both dominant and recessive forms of dystrophic EB. On a cautionary note, this represents a relatively new technology and the potency of the approach for any region of the genome has to be fully explored (Rahman *et al.*, 2011).

7.2 Conclusion

Important advances in research into inherited skin disorders such as Pachyonychia Congenita (PC) and Epidermolysis Bullosa (EB) amongst others over the past number of decades have resulted in elucidation of the many genes implicated in the major forms of these inherited diseases. Progress in the field of gene therapy in recent years, with increasing employment of such powerful tools as RNA interference have perhaps laid the basis for the pursuit of a successful treatment strategy for dominant dystrophic EB. It is hoped that initial advances in the treatment of the Junctional form of the disease (JEB), shown by the success achieved in a single patient *ex vivo* human clinical trial for JEB (Mavilio *et al.*, 2006), may be mirrored in other forms of the genetic skin disease including DEB. Unfortunately, a ban on the use of γ -retroviral vectors in clinical trials following development of leukaemia-like lymphoproliferative

disease in patients being treated in a clinical trial for X-SCID, has hampered further development of this strategy for JEB patients. However the use of SIN vectors, which self-inactivate the retroviral long-terminal repeats, represents an attractive alternative with an improved safety profile, and offers the potential for further progress in the treatment of those affected by JEB as well as other forms of EB (Almaraz *et al.*, 2011). Indeed it has recently been reported that phase 1 clinical trials, involving both gene-based and protein-based COL7A1 delivery are currently underway for the recessive form of DEB (Siprashvili *et al.*, 2010), although the details and results of these clinical trials have yet to be reported. The focus of the research carried out as described in this Ph.D thesis was the development of gene therapy strategies for the autosomal dominantly inherited form of the disease and overall a number of essential components required for the evaluation of such an objective were achieved. One such approach as discussed, involved an RNAi-mediated mutation-independent suppression and replacement method and identification of effective inhibitors of the COL7A1 gene as well as generation of a full-length COL7A1 gene which is resistant to the effects of the most potent RNAi inhibitor was achieved. Additionally, a mutation-dependent RNAi-based approach was investigated for two DDEB-causing COL7A1 mutations, although the potency of suppression achieved using this approach was rather limited. Overall, additional components are undoubtedly still needed for further progression of these therapeutic strategies using the potentially powerful molecular tools identified to date. Most notably the requirement for an animal model exhibiting the human DDEB genotype and phenotype still exists: such a transgenic animal would provide an important model to test the potential therapeutic strategies explored during this study together with other possible therapies under consideration. It now remains for the findings thus far as described in this Ph.D. thesis to be further developed, exploited and perfected in an effort to pave the way for the specific curative treatment of individuals affected by this devastating and debilitating skin disorder.

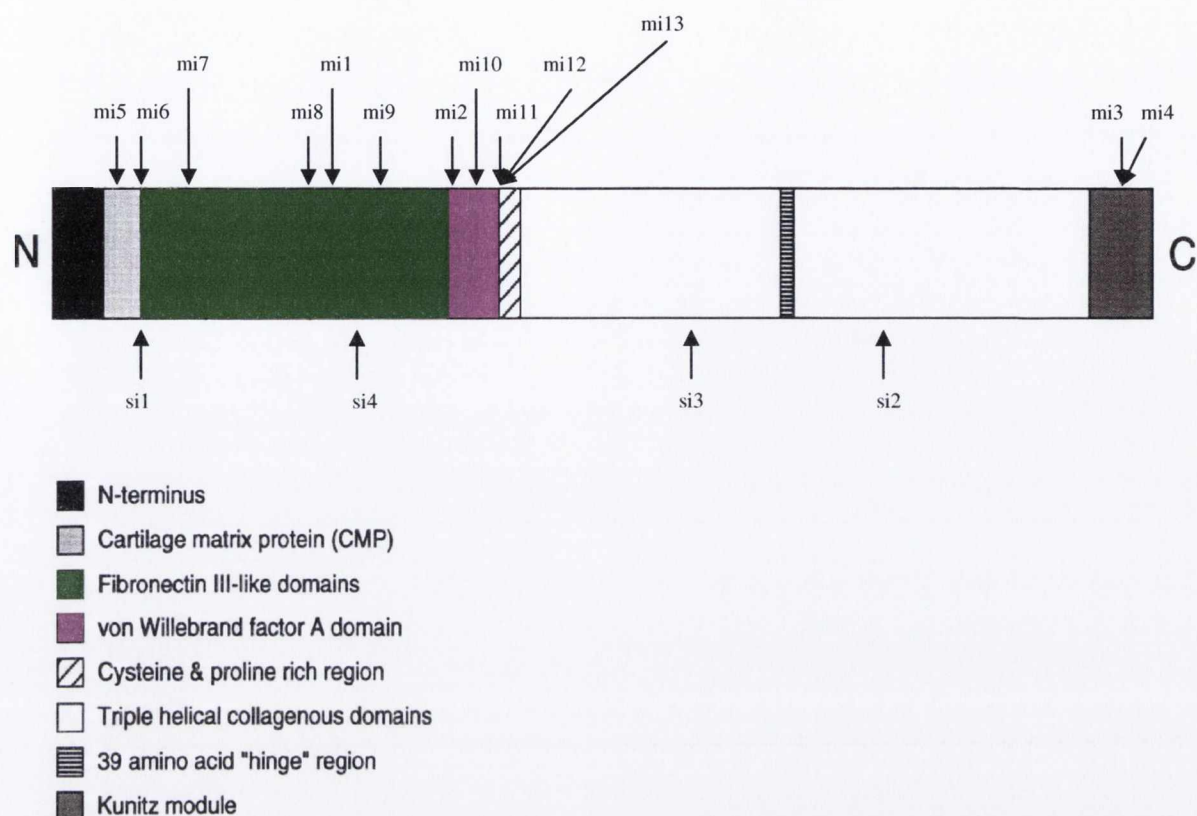


Figure 7.1 COL7A1 target regions of all tested RNAi suppressors: Schematic representation of type VII collagen domain organization. The type VII collagen molecule consists of a central collagenous triple-helical domain which has a 39-amino acid non-collagenous interruption ('hinge' region). The collagenous domain is flanked by amino-terminal and carboxy-terminal globular domains, with submodules with homology to known protein modules, as indicated on the lower left corner. Arrows represent the COL7A1 target regions of all 13 artificial microRNAs tested (mi1-13; chapter 3) and all 4 siRNAs tested (si1-4; chapters 4 and 5) (figure and legend modified from Pulkkinen and Uitto, 1999).

References

References

- Allen D, Winters E, Kenna PF, Humphries P, Farrar GJ (2008) Reference gene selection for real-time rtPCR in human epidermal keratinocytes. *Journal of Dermatological Science* **49**(3): 217-225.
- Almaani N, Liu L, Dopping-Hepenstal PJ, Lovell PA, Lai-Cheong JE, Graham RM, Mellerio JE, McGrath JA (2009) Autosomal dominant junctional epidermolysis bullosa. *The British Journal of Dermatology* **160**(5): 1094-1097.
- Almarza D, Bussadori G, Navarro M, Mavilio F, Larcher F, Murillas R (2011) Risk assessment in skin gene therapy: viral-cellular fusion transcripts generated by proviral transcriptional read-through in keratinocytes transduced with self-inactivating lentiviral vectors. *Gene Therapy* **18**(7): 674-681.
- Alves S, Nascimento-Ferreira I, Auregan G, Hassig R, Dufour N, Brouillet E, Pedroso de Lima MC, Hantraye P, Pereira de Almeida L, Déglon N (2008) Allele-specific RNA silencing of mutant ataxin-3 mediates neuroprotection in a rat model of Machado-Joseph disease. *PloS One* **3**(10): e3341.
- Amarzguioui M, Rossi JJ, Kim D (2005) Approaches for chemically synthesized siRNA and vector-mediated RNAi. *FEBS Letters* **579**(26): 5974-5981.
- Ameres SL, Martinez J, Schroeder R (2007) Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* **130**(1): 101-112.
- Angaji SA, Hedayati SS, Poor RH, Madani S, Poor SS, Panahi S (2010) Application of RNA interference in treating human diseases. *Journal of Genetics* **89**(4): 527-537.
- Atkinson SD, McGilligan VE, Liao H, Szeverenyi I, Smith FJ, Moore CB, McLean WH (2011) Development of Allele-Specific Therapeutic siRNA for Keratin 5 Mutations in Epidermolysis Bullosa Simplex. *The Journal of Investigative Dermatology* **131**(10): 2079-2086.
- Attractene Transfection Reagent Handbook, Qiagen, January 2008.
- Aumailley M, Bruckner-Tuderman L, Carter WG, Deutzmann R, Edgar D, Ekblom P, Engel J, Engvall E, Hohenester E, Jones JC, Kleinman HK, Marinkovich MP, Martin GR, Mayer U, Meneguzzi G, Miner JH, Miyazaki K, Patarroyo M, Paulsson M, Quaranta V, Sanes JR, Sasaki T, Sekiguchi K, Sorokin LM, Talts JF, Tryggvason K, Uitto J, Virtanen I, von der Mark K, Wewer UM, Yamada Y,

- Yurchenco PD (2005) A simplified laminin nomenclature. *Matrix Biology* **24**(5): 326-332.
- Aumailley M, Has C, Tunggal L, Bruckner-Tuderman L (2006) Molecular basis of inherited skin-blistering disorders, and therapeutic implications. *Expert Reviews in Molecular Medicine* **8**(24): 1-21.
- Badiavas EV (2004) The potential of bone marrow cells to orchestrate homeostasis and healing in skin. *Blood Cells, Molecules and Diseases* **32**(1): 21-23.
- Barascuk N, Vassiliadis E, Larsen L, Wang J, Zheng Q, Xing R, Cao Y, Crespo C, Lapret I, Sabatini M, Villeneuve N, Vilaine JP, Rasmussen LM, Register TC, Karsdal MA (2011) Development and validation of an enzyme-linked immunosorbent assay for the quantification of a specific MMP-9 mediated degradation fragment of type III collagen-A novel biomarker of atherosclerotic plaque remodeling. *Clinical Biochemistry* **44**(10-11): 900-906.
- Bayless KJ, Meininger GA, Scholtz JM, Davis GE (1998) Osteopontin is a ligand for the $\alpha 4 \beta 1$ integrin. *Journal of Cell Science* **111**(Pt 9): 1165-1174.
- Berk DR, Bayliss SJ (2008) Milia: a review and classification. *Journal of the American Academy of Dermatology* **59**(6): 1050-1063.
- BLOCK-iT™ Pol II miR RNAi Expression Vector Kits user manual, Invitrogen, Version E, June 2007.
- Boshart M, Weber F, Jahn G, Dorsch-Häsler K, Fleckenstein B, Schaffner W (1985) A very strong enhancer is located upstream of an immediate early gene of human cytomegalovirus. *Cell* **41**(2): 521-530.
- Boudreau RL, Monteys AM, Davidson BL (2008) Minimizing variables among hairpin-based RNAi vectors reveals the potency of shRNAs. *RNA* **14**(9): 1834-1844.
- Boudreau RL, Martins I, Davidson BL (2009) Artificial MicroRNAs as siRNA Shuttles: Improved Safety as Compared to shRNAs *in vitro* and *in vivo*. *Molecular Therapy* **17**(1): 169-175.
- Boudreau RL, Davidson BL (2010) RNAi therapeutics for CNS disorders. *Brain Research* **1338**: 112-121.
- Boukamp P, Petrussevska RT, Breitkreutz D, Hornung J, Markham A, Fusenig NE (1988) Normal keratinization in a spontaneously immortalized aneuploid human keratinocyte cell line. *The Journal of Cell Biology* **106**(3): 761-771.
- Bramsen JB, Kjems (2011) Chemical modification of small interfering RNA. *Methods in Molecular Biology* **721**: 77-103.

- Bramsen JB, Laursen MB, Nielsen AF, Hansen TB, Bus C, Langkjaer N, Babu BR, Højland T, Abramov M, Van Aerschot A, Odadzic D, Smicius R, Haas J, Andree C, Barman J, Wenska M, Srivastava P, Zhou C, Honcharenko D, Hess S, Müller E, Bobkov GV, Mikhailov SN, Fava E, Meyer TF, Chattopadhyaya J, Zerial M, Engels JW, Herdewijn P, Wengel J, Kjems J (2009) *Nucleic Acids Research* **37**(9): 2867-2881.
- Bruckner-Tuderman L, Höpfner B, Hammami-Hauasli N (1999) Biology of anchoring fibrils: lessons from dystrophic epidermolysis bullosa. *Matrix Biology* **18**(1): 43-54.
- Bruckner-Tuderman L (2009) Can type VII collagen injections cure dystrophic epidermolysis bullosa? *Molecular Therapy* **17**(1): 6-7.
- Bruckner-Tuderman L, McGrath JA, Robinson EC, Uitto J (2010) Animal models of epidermolysis bullosa: update 2010. *The Journal of Investigative Dermatology* **130**(6): 1485-1488.
- Brummelkamp TR, Bernards R, Agami R (2002a) A System for Stable Expression of Short Interfering RNAs in Mammalian Cells. *Science* **296**(5567): 550-553.
- Brummelkamp TR, Bernards R, Agami R (2002b) Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* **2**(3): 243-247.
- Brysk MM, Rajaraman S (1992) Cohesion and desquamation of epidermal stratum corneum. *Progress in Histochemistry and Cytochemistry* **25**(1): 1-53.
- Caplen NJ, Fleenor J, Fire A, Morgan RA (2000) dsRNA-mediated gene silencing in cultured *Drosophila* cells: a tissue culture model for the analysis of RNA interference. *Gene* **252**(1-2): 95-105.
- Castanotto D, Rossi JJ (2009) The promises and pitfalls of RNA-interference-based therapeutics. *Nature* **457**(7228): 426-433.
- Chadderton N, Millington-Ward S, Palfi A, O'Reilly M, Tuohy G, Humphries MM, Li T, Humphries P, Kenna PF, Farrar GJ (2009) Improved retinal function in a mouse model of dominant retinitis pigmentosa following AAV-delivered gene therapy. *Molecular Therapy* **17**(4): 593-599.
- Chen M, Kasahara N, Keene DR, Chan L, Hoeffler WK, Finlay D, Barcova M, Cannon PM, Mazurek C, Woodley DT (2002) Restoration of type VII collagen expression and function in dystrophic epidermolysis bullosa. *Nature Genetics* **32**(4): 670-675.

- Chen PY, Weinmann L, Gaidatzis D, Pei Y, Zavolan M, Tuschl T, Meister G (2008) Strand-specific 5'-O-methylation of siRNA duplexes controls guide strand selection and targeting specificity. *RNA* **14**(2): 263-274.
- Chernolovskaya EL, Zenkova MA (2010) Chemical modification of siRNA. *Current Opinion in Molecular Therapeutics* **12**(2): 158-167.
- Chino T, Tamai K, Yamazaki T, Otsuru S, Kikuchi Y, Nimura K, Endo M, Nagai M, Uitto J, Kitajima Y, Kaneda Y (2008) Bone marrow cell transfer into fetal circulation can ameliorate genetic skin diseases by providing fibroblasts to the skin and inducing immune tolerance. *The American Journal of Pathology* **173**(3): 803-814.
- Chiu YL, Rana TM (2002) RNAi in human cells: basic structural and functional features of small interfering RNA. *Molecular Cell* **10**(3): 549-561.
- Cho WG, Albuquerque RJ, Kleinman ME, Tarallo V, Greco A, Nozaki M, Green MG, Baffi JZ, Ambati BK, De Falco M, Alexander JS, Brunetti A, De Falco S, Ambati J (2009) Small interfering RNA-induced TLR3 activation inhibits blood and lymphatic vessel growth. *Proc. Natl. Acad. Sci. USA* **106**(17): 7137-7142.
- Christiano AM, Rosenbaum LM, Chung-Honet LC, Parente MG, Woodley DT, Pan T-C, Zhang RZ, Chu M-L, Burgeson RE, Uitto J (1992) The large non-collagenous domain (NC-1) of type VII collagen is amino-terminal and chimeric. Homology to cartilage matrix protein, the type III domains of fibronectin and the A domains of von Willebrand factor. *Human Molecular Genetics* **1**(7): 475-481.
- Christiano AM, Hoffman GG, Chung-Honet LC, Lee S, Cheng W, Uitto J, Greenspan DS (1994a) Structural Organization of the Human Type VII Collagen Gene (COL7A1), Composed of More Exons Than Any Previously Characterized Gene. *Genomics* **21**(1): 169-179.
- Christiano AM, Greenspan DS, Lee S, Uitto J (1994b) Cloning of human type VII collagen. Complete primary sequence of the alpha 1(VII) chain and identification of intragenic polymorphisms. *The Journal of Biological Chemistry* **269**(32): 20256-20262.
- Christiano AM, Morricone A, Paradisi M, Angelo C, Mazzanti C, Cavalieri R, Uitto J (1995a) A glycine-to-arginine substitution in the triple-helical domain of type VII collagen in a family with dominant dystrophic epidermolysis bullosa. *The Journal of Investigative Dermatology* **104**(3): 438-440.

- Christiano AM, Lee JY, Chen WJ, LaForgia S, Uitto J (1995b) Pretibial epidermolysis bullosa: genetic linkage to COL7A1 and identification of a glycine-to-cysteine substitution in the triple-helical domain of type VII collagen. *Human Molecular Genetics* **4**(9): 1579-1583.
- Christiano AM, Amano S, Eichenfield LF, Burgeson RE, Uitto J (1997) Premature termination codon mutations in the type VII collagen gene in recessive dystrophic epidermolysis bullosa result in nonsense-mediated mRNA decay and absence of functional protein. *The Journal of Investigative Dermatology* **109**(3): 390-394.
- Chung HJ, Uitto J (2010) Type VII collagen: the anchoring fibril protein at fault in dystrophic epidermolysis bullosa. *Dermatologic Clinics* **28**(1): 93-105.
- Chung KH, Hart CC, Al-Bassam S, Avery A, Taylor J, Patel PD, Vojtek AB, Turner DL (2006) Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucleic Acids Research* **34**(7): e53.
- Chuong CM, Nickoloff BJ, Elias PM, Goldsmith LA, Macher E, Maderson PA, Sundberg JP, Tagami H, Plonka PM, Thestrup-Pedersen K, Bernard BA, Schröder JM, Dotto P, Chang CH, Williams ML, Feingold KR, King LE, Kligman AM, Rees JL, Christophers E (2002) What is the 'true' function of skin? *Experimental Dermatology* **11**(2): 159-187.
- ClustalW2- Multiple Sequence Alignment
(http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html).
- Compton SH, Mecklenbeck S, Mejía JE, Hart SL, Rice M, Cervini R, Barrandon Y, Larin Z, Levy ER, Bruckner-Tuderman L, Hovnanian A (2000) Stable integration of large (>100 kb) PAC constructs in HaCaT keratinocytes using an integrin-targeting peptide delivery system. *Gene Therapy* **7**(18): 1600-1605.
- Corey DR (2007) Chemical modification: the key to clinical application of RNA interference? *The Journal of Clinical Investigation* **117**(12): 3615-3622.
- Coulombe PA, Kerns ML, Fuchs E (2009) Epidermolysis bullosa simplex: a paradigm for disorders of tissue fragility. *The Journal of Clinical Investigation* **119**(7): 1784-1793.
- Couto LB, High KA (2010) Viral vector-mediated RNA interference. *Current Opinion in Pharmacology* **10**(5): 534-542.
- Cserhalmi-Friedman PB, McGrath JA, Mellerio JE, Romero R, Salas-Alanis JC, Paller AS, Dietz HC, Christiano AM (1998) Restoration of open reading frame resulting from skipping of an exon with an internal deletion in the COL7A1 gene.

- Laboratory Investigation; a journal of technical methods and pathology* **78(12)**: 1483-1492.
- Cullen BR (2005) RNAi the natural way. *Nature Genetics* **37(11)**: 1163-1165.
- Cummings B (2002) Pearson Education Inc.
- Dang N, Klingberg S, Marr P, Murrell DF (2007) Review of collagen VII sequence variants found in Australasian patients with dystrophic epidermolysis bullosa reveals nine novel COL7A1 variants. *Journal of Dermatological Science* **46(3)**: 169-178.
- Dang N, Murrell DF (2008) Mutation analysis and characterization of COL7A1 mutations in dystrophic epidermolysis bullosa. *Experimental Dermatology* **17(7)**: 553-568.
- Davidson BL, Paulson HL (2004) Molecular medicine for the brain: silencing of disease genes with RNA interference. *Lancet Neurology* **3(3)**:145-149.
- Davis D, Stokoe D (2010) Zinc finger nucleases as tools to understand and treat human diseases. *BMC Medicine* **8**: 42.
- Davis ME (2009) The first targeted delivery of siRNA in humans via a self-assembling, cyclodextrin polymer-based nanoparticle: from concept to clinic. *Molecular Pharmaceutics* **6(3)**: 659-668.
- De Fougères AR (2008) Delivery vehicles for small interfering RNA in vivo. *Human Gene Therapy* **19(2)**: 125-32.
- De Luca M, Pellegrini G, Mavilio F (2009) Gene therapy of inherited skin adhesion disorders: a critical overview. *The British Journal of Dermatology* **161(1)**: 19-24.
- De Yñigo-Mojado L, Martín-Ruiz I, Sutherland JD (2011) Efficient Allele-Specific Targeting of LRRK2 R1441 Mutations Mediated by RNAi. *PloS One* **6(6)**: e21352.
- Di Lullo GA, Sweeney SM, Korkko J, Ala-Kokko L, San Antonio JD (2002) Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen. *The Journal of Biological Chemistry* **277(6)**: 4223-4231.
- Ding H, Schwarz DS, Keene A, Affar el B, Fenton L, Xia X, Shi Y, Zamore PD, Xu Z (2003) Selective silencing by RNAi of a dominant allele that causes amyotrophic lateral sclerosis. *Aging Cell* **2(4)**: 209-217.
- Di Nunzio F, Maruggi G, Ferrari S, Di Iorio E, Poletti V, Garcia M, Del Rio M, De Luca M, Larcher F, Pellegrini G, Mavilio F (2008) Correction of laminin-5

- deficiency in human epidermal stem cells by transcriptionally targeted lentiviral vectors. *Molecular Therapy* **16**(12): 1977-1985.
- Du G, Yonekubo J, Zeng Y, Osisami M, Frohman MA (2006) Design of expression vectors for RNA interference based on miRNAs and RNA splicing. *The FEBS Journal* **273**(23): 5421-5427.
- Du Q, Thonberg H, Wang J, Wahlestedt C, Liang Z (2005) A systematic analysis of the silencing effects of an active siRNA at all single-nucleotide mismatched target sites. *Nucleic Acids Research* **33**(5): 1671-1677.
- Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. *Nucleic Acids Research* **33**(18): 5978-5990.
- Dykxhoorn DM, Schlehuber LD, London IM, Lieberman J (2006) Determinants of specific RNA interference-mediated silencing of human beta-globin alleles differing by a single nucleotide polymorphism. *Proc. Natl. Acad. Sci. USA* **103**(15): 5953-5958.
- Eady RA, Dunnill MG (1994) Epidermolysis bullosa: hereditary skin fragility diseases as paradigms in cell biology. *Archives of Dermatological Research* **287**(1): 2-9.
- Elbashir SM, Lendeckel W, Tuschl T (2001a) RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development* **15**(2): 188-200.
- Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T (2001b) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**(6836): 494-498.
- Elbashir SM, Martinez J, Patkaniowska A, Lendeckel W, Tuschl T (2001c) Functional anatomy of siRNAs for mediating efficient RNAi in *Drosophila melanogaster* embryo lysate. *The EMBO Journal* **20**(23): 6877-6888.
- Engel J (1992) Laminins and other strange proteins. *Biochemistry* **31**(44): 10643-10651.
- Eyre-Walker AC (1991) An analysis of codon usage in mammals: selection or mutation bias? *Journal of Molecular Evolution* **33**(5): 442-449.
- Featherstone C (2007) Epidermolysis Bullosa: From fundamental molecular biology to clinical therapies. *Journal of Investigative Dermatology* **127**(2): 256-259.
- Fine JD, Eady RA, Bauer EA, Bauer JW, Bruckner-Tuderman L, Heagerty A, Hintner H, Hovnanian A, Jonkman MF, Leigh I, McGrath JA, Mellerio JE, Murrell DF,

- Shimizu H, Uitto J, Vahlquist A, Woodley D, Zambruno G (2008) The classification of inherited epidermolysis bullosa (EB): Report of the Third International Consensus Meeting on Diagnosis and Classification of EB. *Journal of the American Academy of Dermatology* **58(6)**: 931-950.
- Fine JD, Johnson LB, Weiner M, Li KP, Suchindran C (2009) Epidermolysis bullosa and the risk of life-threatening cancers: the National EB Registry experience, 1986-2006. *Journal of the American Academy of Dermatology* **60(2)**: 203-211.
- Fine JD (2010a) Inherited epidermolysis bullosa: past, present, and future. *Annals of the New York Academy of Sciences* **1194**: 213-222.
- Fine JD (2010b) Inherited epidermolysis bullosa. *Orphanet Journal of Rare Diseases* **5**: 12.
- Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391(6669)**: 806-811.
- Forlino A, Marini JC (2000) Osteogenesis imperfecta: prospects for molecular therapeutics. *Molecular genetics and metabolism* **71(1-2)**: 225-232.
- Fritsch A, Loeckermann S, Kern JS, Braun A, Bösl MR, Bley TA, Schumann H, von Elverfeldt D, Paul D, Erlacher M, Berens von Rautenfeld D, Hausser I, Fässler R, Bruckner-Tuderman L (2008) A hypomorphic mouse model of dystrophic epidermolysis bullosa reveals mechanisms of disease and response to fibroblast therapy. *The Journal of Clinical Investigation* **118(5)**: 1669-1679.
- Fritsch A, Spasov S, Elfort S, Schlosser A, Gache Y, Meneguzzi G and Bruckner-Tuderman L (2009) Dominant Negative Effects Of COL7A1 Mutations Can Be Rescued By Controlled Over-Expression Of Normal Collagen VII. *The Journal of Biological Chemistry* **284(44)**: 30248-30256.
- Fuchs E (1995) Keratins and the skin. *Annual Review of Cell and Developmental Biology* **11**:123-153.
- Gache Y, Baldeschi C, Del Rio M, Gagnoux-Palacios L, Larcher F, Lacour J-P, Meneguzzi G (2004) Construction of Skin Equivalents for Gene Therapy of Recessive Dystrophic Epidermolysis Bullosa. *Human Gene Therapy* **15(10)**: 921-933.
- Galehdari H, Mohammadian G, Azmoon S, Salehi B, Pedram M (2010) A novel COL7A1 gene mutation in an Iranian individual suffering dystrophic epidermolysis bullosa. *The Journal of Molecular Diagnostics* **12(3)**: 377-379.

- Gardlík R, Pálffy R, Hodosy J, Lukács J, Turňa J, Celec P (2005) Vectors and delivery systems in gene therapy. *Medical Science Monitor* **11(4)**: RA110-121.
- Gelse K, Pöschl E, Aigner T (2003) Collagens--structure, function and biosynthesis. *Advanced Drug Delivery Reviews* **55(12)**: 1531-1546.
- Geusens B, Sanders N, Prow T, Van Gele M, Lambert J (2009) Cutaneous short-interfering RNA therapy. *Expert Opinion on Drug Delivery* **6(12)**: 1333-1349.
- Gordon MK, Hahn RA (2010) Collagens. *Cell and Tissue Research* **339(1)**:247-257.
- Graham FL, Smiley J, Russell WC, Nairn R (1977) Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *The Journal of General Virology* **36(1)**: 59-74.
- Greenspan DS (1993) The carboxyl-terminal half of type VII collagen, including the non-collagenous NC-2 domain and intron/exon organization of the corresponding region of the COL7A1 gene. *Human Molecular Genetics* **2(3)**: 273-278.
- Gruis NA, Bavinck JN, Steijlen PM, van der Schroeff JG, van Haeringen A, Happle R, Mariman E, van Beersum SE, Uitto J, Vermeer BJ, *et al.* (1992) Genetic linkage between the collagen VII (COL7A1) gene and the autosomal dominant form of dystrophic epidermolysis bullosa in two Dutch kindreds. *The Journal of Investigative Dermatology* **99(5)**: 528-530.
- Guo S, Kemphues KJ (1995) Par-1, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. *Cell* **81(4)**: 611-620.
- Haley B, Zamore PD (2004) Kinetic analysis of the RNAi enzyme complex. *Nature Structural and Molecular Biology* **11(7)**: 599-606.
- Hall AH, Wan J, Shaughnessy EE, Ramsay Shaw B, Alexander KA (2004) RNA interference using boranophosphate siRNAs: structure-activity relationships. *Nucleic Acids Research* **32(20)**: 5991-6000.
- Heinonen S, Männikkö M, Klement JF, Whitaker-Menezes D, Murphy GF, Uitto J (1999) Targeted inactivation of the type VII collagen gene (Col7a1) in mice results in severe blistering phenotype: a model for recessive dystrophic epidermolysis bullosa. *Journal of Cell Science* **112(Pt 21)**: 3641-3648.
- Hickerson RP, Smith FJ, McLean WH, Landthaler M, Leube RE, Kaspar RL (2006) SiRNA-mediated selective inhibition of mutant keratin mRNAs responsible for the skin disorder pachyonychia congenita. *Annals of the New York Academy of Sciences* **1082**: 56-61.

- Hickerson RP, Smith FJ, Reeves RE, Contag CH, Leake D, Leachman SA, Milstone LM, McLean WH, Kaspar RL (2008) Single-nucleotide-specific siRNA targeting in a dominant-negative skin model. *The Journal of Investigative Dermatology* **128**(3): 594-605.
- HiPerFect Transfection Reagent Handbook, Qiagen, May 2008, Fifth Edition.
- Hnik P, Boyer DS, Grillone LR, Clement JG, Henry SP, Green EA (2009) Antisense oligonucleotide therapy in diabetic retinopathy. *Journal of Diabetes Science and Technology* **3**(4): 924-930.
- Hohjoh H (2010) Allele-specific silencing by RNA interference. *Methods in Molecular Biology* **623**: 67-79.
- Holen T, Amarzguioui M, Wiiger MT, Babaie E, Prydz H (2002) Positional effects of short interfering RNAs targeting the human coagulation trigger Tissue Factor. *Nucleic Acids Research* **30**(8): 1757-1766.
- Houben E, De Paepe K, Rogiers V (2007) A keratinocyte's course of life. *Skin Pharmacology and Physiology* **20**(3): 122-132.
- Hovnanian A, Duquesnoy P, Blanchet-Bardon C, Knowlton RG, Amselem S, Lathrop M, Dubertret L, Uitto J, Goossens M (1992) Genetic linkage of recessive dystrophic epidermolysis bullosa to the type VII collagen gene. *The Journal of Clinical Investigation* **90**(3): 1032-1036.
- Hovnanian A, Hilal L, Blanchet-Bardon C, de Prost Y, Christiano AM, Uitto J, Goossens M (1994) Recurrent nonsense mutations within the type VII collagen gene in patients with severe recessive dystrophic epidermolysis bullosa. *American Journal of Human Genetics* **55**(2): 289-296.
- Huang H, Qiao R, Zhao D, Zhang T, Li Y, Yi F, Lai F, Hong J, Ding X, Yang Z, Zhang L, Du Q, Liang Z (2009) Profiling of mismatch discrimination in RNAi enabled rational design of allele-specific siRNAs. *Nucleic Acids Research*: **37**(22): 7560-7569.
- In-Fusion™ Advantage PCR Cloning Kit User Manual (2010) Clontech.
- International registry of patients with DEB & database of associated COL7A1 mutations: <http://www.col7a1.org/>.
- Ito K, Sawamura D, Goto M, Nakamura H, Nishie W, Sakai K, Natsuga K, Shinkuma S, Shibaki A, Uitto J, Denton CP, Nakajima O, Akiyama M, Shimizu H (2009) Keratinocyte-/fibroblast-targeted rescue of Col7a1-disrupted mice and generation of

- an exact dystrophic epidermolysis bullosa model using a human COL7A1 mutation. *The American Journal of Pathology* **175**(6): 2508-2517.
- Itoh M, Kiuru M, Cairo MS, Christiano AM (2011) Generation of keratinocytes from normal and recessive dystrophic epidermolysis bullosa-induced pluripotent stem cells. *Proc Natl Acad Sci U S A*. **108**(21): 8797-8802.
- Jackson AL, Bartz SR, Schelter J, Kobayashi SV, Burchard J, Mao M, Li B, Cavet G, Linsley PS (2003) Expression profiling reveals off-target gene regulation by RNAi. *Nature Biotechnology* **21**(6): 635-637.
- Jacobsen F, Mertens-Rill J, Beller J, Hirsch T, Daigeler A, Langer S, Lehnhardt M, Steinau HU, Steinstraesser L (2006) Nucleofection: a new method for cutaneous gene transfer? *Journal of Biomedicine & Biotechnology* **2006**(5): 26060.
- Jacobson GB, Gonzalez-Gonzalez E, Spitler R, Shinde R, Leake D, Kaspar RL, Contag CH, Zare RN (2010) Biodegradable nanoparticles with sustained release of functional siRNA in skin. *Journal of Pharmaceutical Sciences* **99**(10): 4261-4266.
- Jerábková B, Kopecková L, Bucková H, Veselý K, Valícková J, Fajkusová L (2010) Analysis of the COL7A1 gene in Czech patients with dystrophic epidermolysis bullosa reveals novel and recurrent mutations. *Journal of Dermatological Science* **59**(2): 136-140.
- Jordan ET, Collins M, Terefe J, Ugozzoli L, Rubio T (2008) Optimizing electroporation conditions in primary and other difficult-to-transfect cells. *Journal of Biomolecular Techniques* **19**(5): 328-334.
- Kadler KE, Baldock C, Bella J, Boot-Handford RP (2007) Collagens at a glance. *Journal of Cell Science* **120**(Pt 12):1955-1958.
- Kaiser PK, Symons RC, Shah SM, Quinlan EJ, Tabandeh H, Do DV, Reisen G, Lockridge JA, Short B, Guercioli R, Nguyen QD (2010) RNAi-based treatment for neovascular age-related macular degeneration by Sirna-027. *American Journal of Ophthalmology* **150**(1): 33-39.
- Kandavelou K, Ramalingam S, London V, Mani M, Wu J, Alexeev V, Civin CI, Chandrasegaran S (2009) Targeted manipulation of mammalian genomes using designed zinc finger nucleases. *Biochemical and Biophysical Research Communications* **388**(1): 56-61.
- Kanitakis J (2002) Anatomy, histology and immunohistochemistry of normal human skin. *European Journal of Dermatology* **12**(4): 390-399.

- Kashtan CE (1999) Alport syndrome. An inherited disorder of renal, ocular, and cochlear basement membranes. *Medicine* **78**(5): 338-360.
- Kasschau KD, Carrington JC (1998) A counterdefensive strategy of plant viruses: suppression of posttranscriptional gene silencing. *Cell* **95**(4): 461-470.
- Kern JS, Loeckermann S, Fritsch A, Hausser I, Roth W, Magin TM, Mack C, Müller ML, Paul O, Ruther P, Bruckner-Tuderman L (2009) Mechanisms of fibroblast cell therapy for dystrophic epidermolysis bullosa: high stability of collagen VII favors long-term skin integrity. *Molecular Therapy* **17**(9): 1605-1615.
- Kiang A-S, Palfi A, Ader M, Kenna PF, Millington-Ward S, Clark G, Kennan A, O'Reilly M, Tam LCT, Aherne A, McNally N, Humphries P, Farrar GJ (2005) Toward a Gene Therapy for Dominant Disease: Validation of an RNA Interference-Based Mutation-Independent Approach. *Molecular Therapy* **12**(3): 555-561.
- Kim DH, Rossi JJ (2007) Strategies for silencing human disease using RNA interference. *Nature Reviews Genetics* **8**(3): 173-184.
- Kim SS, Garg H, Joshi A, Manjunath N (2009) Strategies for targeted nonviral delivery of siRNAs in vivo. *Trends in Molecular Medicine* **15**(11): 491-500.
- Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to FokI cleavage domain. *Proc Natl Acad Sci U S A*. **93**(3): 1156-1160.
- Kleinman ME, Yamada K, Takeda A, Chandrasekaran V, Nozaki M, Baffi JZ, Albuquerque RJ, Yamasaki S, Itaya M, Pan Y, Appukuttan B, Gibbs D, Yang Z, Karikó K, Ambati BK, Wilgus TA, DiPietro LA, Sakurai E, Zhang K, Smith JR, Taylor EW, Ambati J (2008) Sequence- and target-independent angiogenesis suppression by siRNA via TLR3. *Nature* **452**(7187): 591-597.
- Koerber JT, Jang JH, Schaffer DV (2008) DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. *Molecular Therapy* **16**(10): 1703-1709.
- Koga H, Hamada T, Ishii N, Fukuda S, Sakaguchi S, Nakano H, Tamai K, Sawamura D, Hashimoto T (2010) Exon 87 skipping of the COL7A1 gene in dominant dystrophic epidermolysis bullosa. *The Journal of Dermatology* **38**(5): 489-492.
- Krebs MD, Alsberg E (2011) Localized, Targeted, and Sustained siRNA Delivery. *Chemistry* [Epub ahead of print].
- Kubodera T, Yamada H, Anzai M, Ohira S, Yokota S, Hirai Y, Mochizuki H, Shimada T, Mitani T, Mizusawa H, Yokota T (2011) In vivo application of an RNAi strategy for the selective suppression of a mutant allele. *Human Gene Therapy* **22**(1): 27-34.

- Lagos-Quintana M, Rauhut R, Yalcin A, Meyer J, Lendeckel W, Tuschl T (2002) Identification of tissue-specific microRNAs from mouse. *Current Biology* **12**(9): 735-739.
- Lai MD, Chen CS, Yang CR, Yuan SY, Tsai JJ, Tu CF, Wang CC, Yen MC, Lin CC (2010) An HDAC inhibitor enhances the antitumor activity of a CMV promoter-driven DNA vaccine. *Cancer Gene Therapy* **17**(3): 203-211.
- Lanuti EL, Wikramanayake TC, Kirsner RS (2011) Overcoming obstacles for gene therapy for recessive dystrophic epidermolysis bullosa. *The Journal of Investigative Dermatology* **131**(1): 5.
- Lares MR, Rossi JJ, Ouellet DL (2010) RNAi and small interfering RNAs in human disease therapeutic applications. *Trends in Biotechnology* **28**(11): 570-579.
- Lawton S (2006) Anatomy and function of the skin. Part 2--The epidermis. *Nursing Times*. **102**(32):28-29.
- Leachman SA, Hickerson RP, Hull PR, Smith FJ, Milstone LM, Lane EB, Bale SJ, Roop DR, McLean WH, Kaspar RL (2008) Therapeutic siRNAs for dominant genetic skin disorders including pachyonychia congenita. *Journal of Dermatological Science* **51**(3): 151-157.
- Leachman SA, Hickerson RP, Schwartz ME, Bullough EE, Hutcherson SL, Boucher KM, Hansen CD, Eliason MJ, Srivatsa GS, Kornbrust DJ, Smith FJ, McLean WI, Milstone LM, Kaspar RL (2010) First-in-human mutation-targeted siRNA phase Ib trial of an inherited skin disorder. *Molecular Therapy* **18**(2): 442-446.
- Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S, Kim VN (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**(6956): 415-419.
- Le Provost F, Lillico S, Passet B, Young R, Whitelaw B, Vilotte JL (2010) Zinc finger nuclease technology heralds a new era in mammalian transgenesis. *Trends in Biotechnology* **28**(3): 134-141.
- Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**(1): 15-20.
- Li CX, Parker A, Menocal E, Xiang S, Borodyansky L, Fruehauf JH (2006) Delivery of RNA interference. *Cell Cycle* **5**(18): 2103-2109.
- Li H, Haurigot V, Doyon Y, Li T, Wong SY, Bhagwat AS, Malani N, Anguela XM, Sharma R, Ivanciu L, Murphy SL, Finn JD, Khazi FR, Zhou S, Paschon DE, Rebar

- EJ, Bushman FD, Gregory PD, Holmes MC, High KA (2011) In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* **475**(7355): 217-221.
- Li M, Xu J, Chen X, Sun X (2009) RNA interference as a gene silencing therapy for mutant MYOC protein in primary open angle glaucoma. *Diagnostic Pathology* **4**: 46.
- Li Y, Lacerda DA, Warman ML, Beier DR, Yoshioka H, Ninomiya Y, Oxford JT, Morris NP, Andrikopoulos K, Ramirez F, *et al.* (1995) A fibrillar collagen gene, *Col11a1*, is essential for skeletal morphogenesis. *Cell* **80**(3): 423-430.
- Liao H, Sayers JM, Wilson NJ, Irvine AD, Mellerio JE, Baselga E, Bayliss SJ, Uliana V, Fimiani M, Lane EB, McLean WH, Leachman SA, Smith FJ (2007) A spectrum of mutations in keratins K6a, K16 and K17 causing pachyonychia congenita. *Journal of Dermatological Science* **48**(3): 199-205.
- Livak KJ, Schmittgen TD (2001) Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods* **25**(4): 402-408.
- Lochmatter D, Mullis PE (2011) RNA interference in mammalian cell systems. *Hormone Research in Paediatrics* **75**(1): 63-69.
- Lorenz C, Hadwiger P, John M, Vornlocher HP, Unverzagt C (2004) Steroid and lipid conjugates of siRNAs to enhance cellular uptake and gene silencing in liver cells. *Bioorganic & Medicinal Chemistry Letters* **14**(19): 4975-4977.
- Luo KQ, Chang DC (2004) The gene-silencing efficiency of siRNA is strongly dependent on the local structure of mRNA at the targeted region. *Biochemical and biophysical research communications* **318**(1): 303-310.
- Manjunath N, Dykxhoorn DM (2010) Advances in synthetic siRNA delivery. *Discovery Medicine* **9**(48): 418-430.
- Marini JC, Forlino A, Cabral WA, Barnes AM, San Antonio JD, Milgrom S, Hyland JC, Körkkö J, Prockop DJ, De Paepe A, Coucke P, Symoens S, Glorieux FH, Roughley PJ, Lund AM, Kuurila-Svahn K, Hartikka H, Cohn DH, Krakow D, Mottes M, Schwarze U, Chen D, Yang K, Kuslich C, Troendle J, Dalglish R, Byers PH (2007) Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. *Human Mutation* **28**(3): 209-221.

- Martinet W, Schrijvers DM, Kockx MM (2003) Nucleofection as an efficient nonviral transfection method for human monocytic cells. *Biotechnology Letters* **25(13)**: 1025-1029.
- Mastroiannopoulos NP, Uney JB, Phylactou LA (2010) The application of ribozymes and DNazymes in muscle and brain. *Molecules* **15(8)**: 5460-5472.
- Mavilio F, Pellegrini G, Ferrari S, Di Nunzio F, Di Iorio E, Recchia A, Maruggi G, Ferrari G, Provasi E, Bonini C, Capurro S, Conti A, Magnoni C, Giannetti A, De Luca M (2006) Correction of junctional epidermolysis bullosa by transplantation of genetically modified epidermal stem cells. *Nature Medicine* **12(12)**: 1397-1402.
- McBride JL, Boudreau RL, Harper SQ, Staber PD, Monteys AM, Martins I, Gilmore BL, Burstein H, Peluso RW, Polisky B, Carter BJ, Davidson BL (2008) Artificial miRNAs mitigate shRNA-mediated toxicity in the brain: Implications for the therapeutic development of RNAi. *Proc. Natl. Acad. Sci. USA* **105(15)**: 5868-5873.
- McGrath JA, Ishida-Yamamoto A, O'Grady A, Leigh IM, Eady RA (1993) Structural Variations in Anchoring Fibrils in Dystrophic Epidermolysis Bullosa: correlation with type VII collagen expression. *The Journal of Investigative Dermatology* **100(4)**: 366-372.
- McGrath JA, Gatalica B, Christiano AM, Li K, Owaribe K, McMillan JR, Eady RA, Uitto J (1995) Mutations in the 180-kD bullous pemphigoid antigen (BPAG2), a hemidesmosomal transmembrane collagen (COL17A1), in generalized atrophic benign epidermolysis bullosa. *Nature Genetics* **11(1)**: 83-86.
- McMillan J, Akiyama M, Shimizu H (2003) Epidermal basement membrane zone components: ultrastructural distribution and molecular interactions. *Journal of Dermatological Science* **31(3)**: 169-177.
- Mellerio JE, Salas-Alanis JC, Talamantes ML, Horn H, Tidman MJ, Ashton GH, Eady RA, McGrath JA (1998) A recurrent glycine substitution mutation, G2043R, in the type VII collagen gene (COL7A1) in dominant dystrophic epidermolysis bullosa. *The British Journal of Dermatology* **139(4)**: 730-737.
- Mellerio JE, Ashton GH, Mohammedi R, Lyon CC, Kirby B, Harman KE, Salas-Alanis JC, Atherton DJ, Harrison PV, Griffiths WA, Black MM, Eady RA, McGrath JA (1999) Allelic heterogeneity of dominant and recessive COL7A1 mutations underlying epidermolysis bullosa pruriginosa. *The Journal of Investigative Dermatology* **112(6)**: 984-987.

- Menon GK (2002) New insights into skin structure: scratching the surface. *Advanced Drug Delivery Reviews* **54** (Suppl 1): S3-S17.
- Miller VM, Xia H, Marrs GL, Gouvion CM, Lee G, Davidson BL, Paulson HL (2003) Allele-specific silencing of dominant disease genes. *Proc. Natl. Acad. Sci. USA* **100**(12): 7195-7200.
- Millington-Ward S, O'Neill B, Tuohy G, Al-Jandal N, Kiang A-S, Kenna PF, Palfi A, Hayden P, Mansergh F, Kennan A, Humphries P, Farrar GJ (1997) Strategems *in vitro* for gene therapies directed to dominant mutations. *Human Molecular Genetics* **6**(9): 1415-1426.
- Millington-Ward S, McMahon HP, Allen D, Tuohy G, Kiang AS, Palfi A, Kenna PF, Humphries P, Farrar GJ (2004) RNAi of COL1A1 in mesenchymal progenitor cells. *European Journal of Human Genetics* **12**(10): 864-866.
- Millington-Ward S, McMahon HP, Farrar GJ (2005) Emerging therapeutic approaches for osteogenesis imperfecta. *Trends in Molecular Medicine* **11**(6): 299-305.
- Millington-Ward S, Chadderton N, O'Reilly M, Palfi A, Goldmann T, Kilty C, Humphries M, Wolfrum U, Bennett J, Humphries P, Kenna PF, Farrar GJ (2011) Suppression and Replacement Gene Therapy for Autosomal Dominant Disease in a Murine Model of Dominant Retinitis Pigmentosa. *Molecular Therapy* **19**(4): 642-649.
- Mitsuhashi Y, Hashimoto I (2003) Genetic abnormalities and clinical classification of epidermolysis bullosa. *Archives of Dermatological Research* **295** Suppl 1: S29-S33.
- Morrissey DV, Blanchard K, Shaw L, Jensen K, Lockridge JA, Dickinson B, McSwiggen JA, Vargeese C, Bowman K, Shaffer CS, Polisky BA, Zinnen S (2005a) Activity of stabilized short interfering RNA in a mouse model of hepatitis B virus replication. *Hepatology* **41**(6): 1349-1356.
- Morrissey DV, Lockridge JA, Shaw L, Blanchard K, Jensen K, Breen W, Hartsough K, Machemer L, Radka S, Jadhav V, Vaish N, Zinnen S, Vargeese C, Bowman K, Shaffer CS, Jeffs LB, Judge A, MacLachlan I, Polisky B (2005b) Potent and persistent *in vivo* anti-HBV activity of chemically modified siRNAs. *Nature Biotechnology* **23**(8): 1002-1007.
- Mulhbacher J, St-Pierre P, Lafontaine DA (2010) Therapeutic applications of ribozymes and riboswitches. *Current Opinion in Pharmacology* **10**(5): 551-556.

- Murata T, Masunaga T, Ishiko A, Shimizu H, Nishikawa T (2004) Differences in recurrent COL7A1 mutations in dystrophic epidermolysis bullosa: ethnic-specific and worldwide recurrent mutations. *Archives of Dermatological Research* **295**(10): 442-447.
- Murauer EM, Gache Y, Gratz IK, Klausegger A, Muss W, Gruber C, Meneguzzi G, Hintner H, Bauer JW (2011) Functional correction of type VII collagen expression in dystrophic epidermolysis bullosa. *The Journal of Investigative Dermatology* **131**(1): 74-83.
- Nagy N, McGrath JA (2010) Blistering skin diseases: a bridge between dermatopathology and molecular biology. *Histopathology* **56**(1): 91-99.
- Nakamura H, Sawamura D, Goto M, Sato-Matsumura KC, LaDuca J, Lee JY, Masunaga T, Shimizu H (2004) The G2028R glycine substitution mutation in COL7A1 leads to marked inter-familial clinical heterogeneity in dominant dystrophic epidermolysis bullosa. *Journal of Dermatological Science* **34**(3):195-200.
- Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *The Plant Cell* **2**(4): 279-289.
- Nead MA, McCance DJ (1995) Poly-L-ornithine-mediated transfection of human keratinocytes. *The Journal of Investigative Dermatology* **105**(5): 668-671.
- Nelson JA, Reynolds-Kohler C, Smith BA (1987) Negative and positive regulation by a short segment in the 5'-flanking region of the human cytomegalovirus major immediate-early gene. *Molecular and Cellular Biology* **7**(11): 4125-4129.
- Ohashi M, Shu E, Nagai M, Murase K, Nakano H, Tamai K, Sawamura D, Hiroka T, Seishima M, Kitajima Y, Aoyama Y (2011) Two cases of recessive dystrophic epidermolysis bullosa diagnosed as severe generalized. *The Journal of Dermatology* **38**(9): 893-899.
- Ohnishi Y, Tamura Y, Yoshida M, Tokunaga K, Hohjoh H (2008) Enhancement of allele discrimination by introduction of nucleotide mismatches into siRNA in allele-specific gene silencing by RNAi. *PLoS One* **3**(5): e2248.
- Olsen PH, Ambros V (1999) The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Developmental Biology* **216**(2): 671-680.

- O'Reilly M, Palfi A, Chadderton N, Millington-Ward S, Ader M, Cronin T, Tuohy T, Auricchio A, Hildinger M, Tivnan A, McNally N, Humphries MM, Kiang AS, Humphries P, Kenna PF, Farrar GJ (2007) RNA interference-mediated suppression and replacement of human rhodopsin in vivo. *American Journal of Human Genetics* **81**(1): 127-135.
- Ortiz-Urda S, Thyagarajan B, Keene DR, Lin Q, Fang M, Calos MP, Khavari PA (2002) Stable nonviral genetic correction of inherited human skin disease. *Nature Medicine* **8**(10): 1166-1170.
- Ovcharenko D, Jarvis R, Hunicke-Smith S, Kelnar K, Brown D (2005) High-throughput RNAi screening in vitro: from cell lines to primary cells. *RNA* **11**(6): 985-993.
- Paddison PJ, Caudy AA, Bernstein E, Hannon GJ, Conklin DS (2002) Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. *Genes & Development*. **16**(8): 948-958.
- Palfi A, Ader M, Kiang AS, Millington-Ward S, Clark G, O'Reilly M, McMahon HP, Kenna PF, Humphries P, Farrar GJ (2006) RNAi-based suppression and replacement of rds-peripherin in retinal organotypic culture. *Human Mutation* **27**(3): 260-268.
- Palfi A, Millington-Ward S, Chadderton N, O'Reilly M, Goldmann T, Humphries MM, Li T, Wolfrum U, Humphries P, Kenna PF, Farrar GJ (2010) Adeno-associated virus-mediated rhodopsin replacement provides therapeutic benefit in mice with a targeted disruption of the rhodopsin gene. *Human Gene Therapy* **21**(3): 311-323.
- Parente MG, Chung LC, Ryyänen J, Woodley DT, Wynn KC, Bauer EA, Mattei MG, Chu M-L, Uitto J (1991) Human type VII collagen: cDNA cloning and chromosomal mapping of the gene. *Proc. Natl. Acad. Sci. USA* **88**(16): 6931-6935.
- Pecot CV, Calin GA, Coleman RL, Lopez-Berestein G, Sood AK (2011) RNA interference in the clinic: challenges and future directions. *Nature Reviews Cancer* **11**(1): 59-67.
- Pendaries V, Verrecchia F, Michel S, Mauviel A (2003) Retinoic acid receptors interfere with the TGF-beta/Smad signaling pathway in a ligand-specific manner. *Oncogene* **22**(50): 8212-8220.
- Persikov AV, Pillitteri RJ, Amin P, Schwarze U, Byers PH, Brodsky B (2004) Stability related bias in residues replacing glycines within the collagen triple helix

- (Gly-Xaa-Yaa) in inherited connective tissue disorders. *Human Mutation* **24(4)**: 330-337.
- Pfendner EG, Lucky AW (2006) Dystrophic Epidermolysis Bullosa. In: Pagon RA, Bird TC, Dolan CR, Stephens K, editors. GeneReviews [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2006. [updated 2010].
- Prost-Squarcioni C (2006) Histology of skin and hair follicle. *Médecines Sciences* **22(2)**:131-137.
- pSuper RNAi System™ manual: A vector system for expression of short interfering RNA, OligoEngine.
- Pubmed: <http://www.ncbi.nlm.nih.gov/pubmed/>
- Pulkkinen L, Uitto J (1999) Mutation analysis and molecular genetics of epidermolysis bullosa. *Matrix Biology* **18(1)**: 29-42.
- Puttaraju M, Jamison SF, Mansfield SG, Garcia-Blanco MA, Mitchell LG (1999) Spliceosome-mediated RNA trans-splicing as a tool for gene therapy. *Nature Biotechnology* **17(3)**: 246-252.
- QuikChange Primer Design Program (<http://www.agilent.com/genomics/qcpd>).
- Rahman SH, Maeder ML, Joung JK, Cathomen T (2011) Zinc-finger nucleases for somatic gene therapy: the next frontier. *Human Gene Therapy* **22(8)**: 925-933.
- Rana TM (2007) Illuminating the silence: understanding the structure and function of small RNAs. *Nature Reviews Molecular Cell Biology* **8(1)**: 23-36.
- Raoul C, Barker SD, Aebischer P (2006) Viral-based modelling and correction of neurodegenerative diseases by RNA interference. *Gene Therapy* **13(6)**: 487-495.
- Remington J, Wang X, Hou Y, Zhou , Burnett J, Muirhead T, Uitto J, Keene DR, Woodley DT, Chen M (2009) Injection of Recombinant Human Type VII Collagen Corrects the Disease Phenotype in a Murine Model of Dystrophic Epidermolysis Bullosa. *Molecular Therapy* **17(1)**: 26-33.
- Rémy S, Tesson L, Ménoret S, Usal C, Scharenberg AM, Anegón I (2010) Zinc-finger nucleases: a powerful tool for genetic engineering of animals. *Transgenic Research* **19(3)**: 363-371.
- Ren L, Gao G, Zhao D, Ding M, Luo J, Deng H (2007) Developmental stage related patterns of codon usage and genomic GC content: searching for evolutionary fingerprints with models of stem cell differentiation. *Genome Biology* **8(3)**: R35.

- Reyon D, Kirkpatrick JR, Sander JD, Zhang F, Voytas DF, Joung JK, Dobbs D, Coffman CR (2011) ZFNGenome: a comprehensive resource for locating zinc finger nuclease target sites in model organisms. *BMC Genomics* **12**: 83.
- Rezniczek GA, Walko G, Wiche G (2010) Plectin gene defects lead to various forms of epidermolysis bullosa simplex. *Dermatologic Clinics* **28(1)**: 33-41.
- Ricard-Blum S, Ruggiero F (2005) The collagen superfamily: from the extracellular matrix to the cell membrane. *Pathologie-Biologie* **53(7)**: 430-442.
- Riedl E, Klausegger A, Bauer JW, Foedinger D, Kittler H (2009) A novel glycine mutation in the COL7A1 gene leading to dominant dystrophic epidermolysis bullosa with intra-familial phenotypical heterogeneity. *Pediatric Dermatology* **26(1)**: 115-117.
- Robbins M, Judge A, Liang L, McClintock K, Yaworski E, MacLachlan I (2007) 2'-O-methyl-modified RNAs act as TLR7 antagonists. *Molecular Therapy* **15(9)**: 1663-1669.
- Roberts RG, Coffey AJ, Bobrow M, Bentley DR (1993) Exon structure of the human dystrophin gene. *Genomics* **16(2)**: 536-538.
- Romano N, Macino G (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Molecular Microbiology* **6(22)**: 3343-3353.
- SABiosciences PCR data tool:
<http://www.SABiosciences.com/pcrarraydataanalysis.php>.
- Saito K, Chen M, Bard F, Chen S, Zhou H, Woodley D, Polischuk R, Schekman R, Malhotra V (2009) TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. *Cell* **136(5)**: 891-902.
- Sawamura D, Nakano H, Matsuzaki Y (2010) Overview of epidermolysis bullosa. *Journal of Dermatology* **37(3)**: 214-219.
- Schäffer MR, Tantry U, Ahrendt GM, Wasserkrug HL, Barbul A (1997) Stimulation of fibroblast proliferation and matrix contraction by wound fluid. *The International Journal of Biochemistry & Cell Biology* **29(1)**: 231-239.
- Scherr M, Eder M (2007) Gene silencing by small regulatory RNAs in mammalian cells. *Cell Cycle* **6(4)**: 444-449.
- Schmid P, Flegel WA (2011) Codon usage in vertebrates is associated with a low risk of acquiring nonsense mutations. *Journal of Translational Medicine* **9**: 87.

- Schnapp LM, Hatch N, Ramos DM, Klimanskaya IV, Sheppard D, Pytela R (1995) The human integrin alpha 8 beta 1 functions as a receptor for tenascin, fibronectin, and vitronectin. *The Journal of Biological Chemistry* **270(39)**: 23196-23202.
- Scholefield J, Greenberg LJ, Weinberg MS, Arbuthnot PB, Abdelgany A, Wood MJ (2009) Design of RNAi hairpins for mutation-specific silencing of ataxin-7 and correction of a SCA7 phenotype. *PLoS One* **4(9)**: e7232.
- Schwarz DS, Ding H, Kennington L, Moore JT, Schelter J, Burchard J, Linsley PS, Aronin N, Xu Z, Zamore PD (2006) Designing siRNA that distinguish between genes that differ by a single nucleotide. *PLoS Genetics* **2(9)**: e140.
- Sequence Manipulation Suite: Codon Usage;
http://www.ualberta.ca/~stothard/javascript/codon_usage.html
- Shabalina SA, Ogurtsov AY, Spiridonov NA (2006) A periodic pattern of mRNA secondary structure created by the genetic code. *Nucleic Acids Research* **34(8)**: 2428-2437.
- Shan Z, Lin Q, Deng C, Li X, Huang W, Tan H, Fu Y, Yang M, Yu XY (2009a) An efficient method to enhance gene silencing by using precursor microRNA designed small hairpin RNAs. *Molecular Biology Reports* **36(6)**: 1483-1489.
- Shan ZX, Lin QX, Yang M, Deng CY, Kuang SJ, Zhou ZL, Xiao DZ, Liu XY, Lin SG, Yu XY (2009b) A quick and efficient approach for gene silencing by using triple putative microRNA-based short hairpin RNAs. *Molecular and Cellular Biochemistry* **323(1-2)**: 81-89.
- Shao Y, Chan CY, Maliyekkel A, Lawrence CE, Roninson IB, Ding Y (2007) Effect of target secondary structure on RNAi efficiency. *RNA* **13(10)**: 1631-1640.
- Sierant M, Paduszynska A, Kazmierczak-Baranska J, Nacmias B, Sorbi S, Bagnoli S, Sochacka E, Nawrot B (2011) Specific Silencing of L392V PSEN1 Mutant Allele by RNA Interference. *International Journal of Alzheimer's Disease* **2011**: 809218.
- Siprashvili Z, Nguyen NT, Bezchinsky MY, Marinkovich MP, Lane AT, Khavari PA (2010) Long-term type VII collagen restoration to human epidermolysis bullosa skin tissue. *Human Gene Therapy* **21(10)**: 1299-1310.
- Sliva K, Schnierle BS (2010) Selective gene silencing by viral delivery of short hairpin RNA. *Virology Journal* **7**: 248.
- Solomon E, Hiorns L, Sheer D, Rowe D (1984) Confirmation that the type I collagen gene on chromosome 17 is COL1A1 (alpha 1[I]), using a human genomic probe. *Annals of Human Genetics* **48(Pt 1)**: 39-42.

- Solovan C, Ciolan M, Olariu L (2005) The biomolecular and ultrastructural basis of epidermolysis bullosa. *Acta Dermatoven APA* **14(4)**: 127-135.
- Stark GR, Kerr IM, Williams BR, Silverman RH, Schreiber RD (1998) How cells respond to interferons. *Annual Review of Biochemistry* **67**: 227-264.
- Stegmeier F, Hu G, Rickles RJ, Hannon GJ, Elledge SJ (2005) A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells. *Proc. Natl. Acad. Sci. USA* **102(37)**: 13212-13217.
- Stieger K, Lorenz B (2010) Gene therapy for vision loss -- recent developments. *Discovery Medicine* **10(54)**: 425-433.
- Tabara H, Sarkissian M, Kelly WG, Fleenor J, Grishok A, Timmons L, Fire A, Mello CC (1999) The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. *Cell* **99(2)**: 123-132.
- Taddese S, Jung MC, Ihling C, Heinz A, Neubert RH, Schmelzer CE (2010) MMP-12 catalytic domain recognizes and cleaves at multiple sites in human skin collagen type I and type III. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1804(4)**: 731-739.
- Tam W (2001) Identification and characterization of human BIC, a gene on chromosome 21 that encodes a noncoding RNA. *Gene* **274(1-2)**: 157-167.
- Tamai K, Kaneda Y, Uitto J (2009) Molecular therapies for heritable blistering diseases. *Trends in Molecular Medicine* **15(7)**: 285-292.
- Teo EH, Cross KJ, Bomsztyk ED, Lyden DC, Spector JA (2009) Gene therapy in skin: choosing the optimal viral vector. *Annals of Plastic Surgery* **62(5)**: 576-580.
- Thomas P, Smart TG (2005) HEK293 cell line: a vehicle for the expression of recombinant proteins. *Journal of Pharmacological and Toxicological Methods* **51(3)**: 187-200.
- Tidman MJ, Eady RA (1985) Evaluation of anchoring fibrils and other components of the dermal-epidermal junction in dystrophic epidermolysis bullosa by a quantitative ultrastructural technique. *The Journal of Investigative Dermatology* **84(5)**: 374-377.
- Tiemann K, Rossi JJ (2009) RNAi-based therapeutics-current status, challenges and prospects. *EMBO Molecular Medicine* **1(3)**: 142-151.
- Titeux M, Pendaries V, Zanta-Boussif MA, Décha A, Pironon N, Tonasso L, Mejia JE, Brice A, Danos O, Hovnanian A (2010) SIN retroviral vectors expressing COL7A1 under human promoters for ex vivo gene therapy of recessive dystrophic epidermolysis bullosa. *Molecular Therapy* **18(8)**: 1509-1518.

- Tolar J, Ishida-Yamamoto A, Riddle M, McElmurry RT, Osborn M, Xia L, Lund T, Slattery C, Uitto J, Christiano AM, Wagner JE, Blazar BR (2009) Amelioration of epidermolysis bullosa by transfer of wild-type bone marrow cells. *Blood* **113**(5): 1167-1174.
- Trompeter HI, Weinhold S, Thiel C, Wernet P, Uhrberg M (2003) Rapid and highly efficient gene transfer into natural killer cells by nucleofection. *Journal of Immunological Methods* **274** (1-2): 245-256.
- Ui-Tei K, Naito Y, Zenno S, Nishi K, Yamato K, Takahashi F, Juni A, Saigo K (2008) Functional dissection of siRNA sequence by systematic DNA substitution: modified siRNA with a DNA seed arm is a powerful tool for mammalian gene silencing with significantly reduced off-target effect. *Nucleic Acids Research* **36**(7): 2136-2151.
- Uitto J, Richard G (2004) Progress in Epidermolysis Bullosa: Genetic Classification and Clinical Implications. *Am J Med Genet C Semin Med Genet.* **131C**(1): 61-74.
- Uitto J, Richard G (2005) Progress in epidermolysis bullosa: from eponyms to molecular genetic classification. *Clinics in Dermatology* **23**(1): 33-40.
- Uitto J (2008) Epidermolysis bullosa: prospects for cell-based therapies. *The Journal of Investigative Dermatology* **128**(9): 2140-2142.
- Uitto J (2009) Progress in Heritable Skin Diseases: Translational Implications of Mutation Analysis and Prospects of Molecular Therapies. *Acta Dermato-Venereologica* **89**(3): 228-235.
- Uitto J (2011a) Regenerative medicine for skin diseases: iPS cells to the rescue. *The Journal of Investigative Dermatology* **131**(4): 812-814.
- Uitto J (2011b) In this issue: glycine substitution mutations in the COL7A1 gene: implications for inheritance of dystrophic epidermolysis bullosa - dominant vs. recessive. *Acta Dermato-venereologica* **91**(3): 259-261.
- Uprichard SL (2005) The therapeutic potential of RNA interference. *FEBS Letters* **579**(26): 5996-6007.
- Van den Akker PC, Jonkman MF, Rengaw T, Bruckner-Tuderman L, Has C, Bauer JW, Klausegger A, Zambruno G, Castiglia D, Mellerio JE, McGrath JA, van Essen AJ, Hofstra RM, Swertz MA (2011) The International Dystrophic Epidermolysis Bullosa Patient Registry: an online database of dystrophic epidermolysis bullosa patients and their COL7A1 mutations. *Human Mutation* **32**(10): 1100-1107.

- Van Nierop GP, De Vries AAF, Holkers M, Vrijksen KR, Gonçalves MAFV (2009) Stimulation of homology-directed gene targeting at an endogenous human locus by a nicking endonuclease. *Nucleic Acids Research* **37**(17): 5725-5736.
- Varki R, Sadowski S, Pfendner E, Uitto J (2006) Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants. *Journal of Medical Genetics* **43**(8): 641-652.
- Varki R, Sadowski S, Uitto J, Pfendner E (2007) Epidermolysis bullosa. II. type vii collagen mutations and phenotype/genotype correlations in the dystrophic subtypes. *Journal of Medical Genetics* **44**(3): 181-192.
- Verrecchia F, Vindevoghel L, Lechleider RJ, Uitto J, Roberts AB, Mauviel A (2001) Smad3/AP-1 interactions control transcriptional responses to TGF-beta in a promoter-specific manner. *Oncogene* **20**(26): 3332-3340.
- Vindevoghel L, Chung KY, Davis A, Kouba D, Kivirikko S, Alder H, Uitto J, Mauviel A (1997) A GT-rich sequence binding the transcription factor Sp1 is crucial for high expression of the human type VII collagen gene (COL7A1) in fibroblasts and keratinocytes. *The Journal of Biological Chemistry* **272**(15): 10196-10204.
- Visser ME, Kastelein JJ, Stroes ES (2010) Apolipoprotein B synthesis inhibition: results from clinical trials. *Current Opinion in Lipidology* **21**(4): 319-323.
- Wagner JE, Ishida-Yamamoto A, McGrath JA, Hordinsky M, Keene DR, Woodley DT, Chen M, Riddle MJ, Osborn MJ, Lund T, Dolan M, Blazar BR, Tolar J (2010) Bone marrow transplantation for recessive dystrophic epidermolysis bullosa. *The New England Journal of Medicine* **363**(7): 629-639.
- Wall NR, Shi Y (2003) Small RNA: can RNA interference be exploited for therapy? *Lancet* **362**(9393): 1401-1403.
- Wally V, Klausegger A, Koller U, Lochmüller H, Krause S, Wiche G, Mitchell LG, Hintner H, Bauer JW (2008) 5' trans-splicing repair of the PLEC1 gene. *The Journal of Investigative Dermatology* **128**(3): 568-574.
- Wang CK, Nelson CF, Brinkman AM, Miller AC, Hoeffler WK (2000) Spontaneous cell sorting of fibroblasts and keratinocytes creates an organotypic human skin equivalent. *The Journal of Investigative Dermatology* **114**(4): 674-680.
- Wang V, Wu W (2009) MicroRNA-based therapeutics for cancer. *BioDrugs* **23**(1): 15-23.

- Wei JA, Zeng X, Han L, Huang Y (2010) Establishment of a bladder cancer cell line with toll-like receptor 2 gene knockdown and identification of its biological characteristics. *Journal of Southern Medical University* **30(8)**: 1797-1800.
- Weissman SM (1992) Gene therapy. *Proc. Natl. Acad. Sci. USA* **89(23)**:11111-11112.
- Wessagowit V, Ashton GH, Mohammedi R, Salas-Alanis JC, Denyer JE, Mellerio JE, Eady RA, McGrath JA (2001) Three cases of de novo dominant dystrophic epidermolysis bullosa associated with the mutation G2043R in COL7A1. *Clinical and Experimental Dermatology* **26(1)**: 97-99.
- Whitehead KA, Langer R, Anderson DG (2009) Knocking down barriers: advances in siRNA delivery. *Nature Reviews Drug Discovery* **8(2)**: 129-138.
- Wolff JA, Budker V (2005) The mechanism of naked DNA uptake and expression. *Advances in Genetics* **54**: 3-20.
- Wong T, Gammon L, Liu L, Mellerio JE, Dopping-Hepenstal PJC, Pacy J, Elia G, Jeffery R, Leigh IM, Navsaria H, McGrath JA (2008) Potential of Fibroblast Cell Therapy for Recessive Dystrophic Epidermolysis Bullosa. *The Journal of Investigative Dermatology* **128(9)**: 2179-2189.
- Woodley DT, Keene DR, Atha T, Huang Y, Ram R, Kasahara N, Chen M (2004) Intradermal Injection of Lentiviral Vectors Corrects Regenerated Human Dystrophic Epidermolysis Bullosa Skin Tissue *in vivo*. *Molecular Therapy* **10(2)**: 318-326.
- Woodley DT, Remington J, Huang Y, Hou Y, Li W, Keene DR, Chen M (2007) Intravenously injected human fibroblasts home to skin wounds, deliver type VII collagen, and promote wound healing. *Molecular Therapy* **15(3)**: 628-635.
- Yeh SI, Han KY, Sabri A, Rosenblatt MI, Azar DT, Jain S, Chang JH (2010) MMP-7 knock-in corneal fibroblast cell lines secrete MMP-7 with proteolytic activity towards collagen XVIII. *Current Eye Research* **35(9)**: 799-805.
- Yi R, Qin Y, Macara IG, Cullen BR (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes and Development* **17(24)**: 3011-3016.
- Zeng Y, Wagner EJ, Cullen BR (2002) Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. *Molecular Cell* **9(6)**: 1327-1333.

- Zeng Y, Yi R, Cullen BR (2005a) Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. *The EMBO Journal* **24**(1): 138-148.
- Zeng Y, Cai X, Cullen BR (2005b) Use of RNA polymerase II to transcribe artificial microRNAs. *Methods in Enzymology* **392**: 371-380.
- Zhang H, Kolb FA, Brondani V, Billy E, Filipowicz W (2002) Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. *The EMBO Journal* **21**(21): 5875-5885.
- Zimmerman AL, Wu S (2011) MicroRNAs, cancer and cancer stem cells. *Cancer Letters* **300**(1): 10-19.
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* **31**(13): 3406-3415.

Appendices

Appendix I

Stuffer sequence, which was inserted into the pcDNA 6.2-GW/ EmGFP-miR vector (provided linearised by Invitrogen) to circularise the plasmid. In addition, the stuffer sequence incorporates two restriction sites for the restriction enzyme *BsaI* within this region, for linearisation of the plasmid for cloning.

Forward sequence insert: 5'- TGCTTGAGACCAAAAAAAGGTCTCA

Appendix II

ECM and Adhesion Molecule Genes:

Position	Unigene	GeneBank	Symbol	Description	Gene Name
A01	Hs.643357	NM_006988	ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	C3-C5, KIAA1346, METH1
A02	Hs.131433	NM_139025	ADAMTS13	ADAM metallopeptidase with thrombospondin type 1 motif, 13	C9orf8, DKFZp434C2322, FLJ42993, MGC118899, MGC118900, TTP, VWFCP, vWF-CP
A03	Hs.271605	NM_007037	ADAMTS8	ADAM metallopeptidase with thrombospondin type 1 motif, 8	ADAM-TS8, FLJ41712, METH2
A04	Hs.502328	NM_000610	CD44	CD44 molecule (Indian blood group)	CDW44, CSPG8, ECMR-III, HCELL, IN, LHR, MC56, MDU2, MDU3, MGC10468, MIC4, MUTCH-I, Pgp1
A05	Hs.461086	NM_004360	CDH1	Cadherin 1, type 1, E-cadherin (epithelial)	Arc-1, CD324, CDHE, ECAD, LCAM, UVO
A06	Hs.143434	NM_001843	CNTN1	Contactin 1	F3, GP135
A07	Hs.523446	NM_080629	COL11A1	Collagen, type XI, alpha 1	CO11A1, COLL6, STL2
A08	Hs.101302	NM_004370	COL12A1	Collagen, type XII, alpha 1	BA209D8.1, COL12A1L, DJ234P15.1
A09	Hs.409662	NM_021110	COL14A1	Collagen, type XIV, alpha 1	UND
A10	Hs.409034	NM_001855	COL15A1	Collagen, type XV, alpha 1	FLJ38566
A11	Hs.368921	NM_001856	COL16A1	Collagen, type XVI, alpha 1	447AA, FP1572
A12	Hs.172928	NM_000088	COL1A1	Collagen, type I, alpha 1	OI4
B01	Hs.508716	NM_001846	COL4A2	Collagen, type IV, alpha 2	DKFZp686I14213, FLJ22259
B02	Hs.210283	NM_000093	COL5A1	Collagen, type V, alpha 1	-
B03	Hs.474053	NM_001848	COL6A1	Collagen, type VI, alpha 1	OPLL
B04	Hs.420269	NM_001849	COL6A2	Collagen, type VI, alpha 2	DKFZp586E1322, FLJ46862, PP3610
B05	Hs.476218	NM_000094	COL7A1	Collagen, type VII, alpha 1	EBD1, EBDCT, EBR1
B06	Hs.654548	NM_001850	COL8A1	Collagen, type VIII, alpha 1	C3orf7, MGC9568
B07	Hs.643801	NM_004385	VCAN	Versican	CSPG2, DKFZp686K06110, ERVR, PG-M, WGN, WGN1
B08	Hs.591346	NM_001901	CTGF	Connective tissue growth factor	CCN2, HCS24, IGFBP8, MGC102839, NOV2
B09	Hs.534797	NM_001903	CTNNA1	Catenin (cadherin-associated protein), alpha 1, 102kDa	CAP102, FLJ36832, FLJ52416
B10	Hs.476018	NM_001904	CTNNB1	Catenin (cadherin-associated protein), beta 1, 88kDa	CTNNB, DKFZp686D02253, FLJ25606, FLJ37923
B11	Hs.166011	NM_001331	CTNND1	Catenin (cadherin-associated protein), delta 1	CAS, CTNND, KIAA0384, P120CAS, P120CTN, p120
B12	Hs.314543	NM_001332	CTNND2	Catenin (cadherin-associated protein), delta 2 (neural plakophilin-related arm-repeat protein)	GT24, NPRAP

C01	Hs.81071 NM_004425	ECM1	Extracellular matrix protein 1	-
C02	Hs.203717NM_002026	FN1	Fibronectin 1	CIG, DKFZp686F10164, DKFZp686H0342, DKFZp686I1370, DKFZp686O13149, ED-B, FINC, FN, FNZ, GFND, GFND2, LETS, MSF
C03	Hs.57697 NM_001523	HAS1	Hyaluronan synthase 1	HAS
C04	Hs.643447NM_000201	ICAM1	Intercellular adhesion molecule 1	BB2, CD54, P3.58
C05	Hs.644352NM_181501	ITGA1	Integrin, alpha 1	CD49a, VLA1
C06	Hs.482077NM_002203	ITGA2	Integrin, alpha 2 (CD49B, alpha 2 subunit of VLA-2 receptor)	BR, CD49B, GPIa, VLA-2, VLAA2
C07	Hs.265829NM_002204	ITGA3	Integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	CD49C, FLJ34631, FLJ34704, GAP-B3, GAPB3, MSK18, VCA-2, VL3A, VLA3a
C08	Hs.694732NM_000885	ITGA4	Integrin, alpha 4 (antigen CD49D, alpha 4 subunit of VLA-4 receptor)	CD49D, IA4, MGC90518
C09	Hs.505654NM_002205	ITGA5	Integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	CD49e, FNRA, VLA5A
C10	Hs.133397NM_000210	ITGA6	Integrin, alpha 6	CD49f, DKFZp686J01244, FLJ18737, ITGA6B, VLA-6
C11	Hs.524484NM_002206	ITGA7	Integrin, alpha 7	FLJ25220
C12	Hs.171311NM_003638	ITGA8	Integrin, alpha 8	-
D01	Hs.174103NM_002209	ITGAL	Integrin, alpha L (antigen CD11A (p180), lymphocyte function-associated antigen 1; alpha polypeptide)	CD11A, LFA-1, LFA1A
D02	Hs.172631NM_000632	ITGAM	Integrin, alpha M (complement component 3 receptor 3 subunit)	CD11B, CR3A, MAC-1, MAC1A, MGC117044, MO1A, SLEB6
D03	Hs.436873NM_002210	ITGAV	Integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51)	CD51, DKFZp686A08142, MSK8, VNRA
D04	Hs.643813NM_002211	ITGB1	Integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12)	CD29, FNRB, GPIIA, MDF2, MSK12, VLA-BETA, VLAB
D05	Hs.375957NM_000211	ITGB2	Integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)	CD18, LAD, LCAMB, LFA-1, MAC-1, MF17, MFI7
D06	Hs.218040NM_000212	ITGB3	Integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	CD61, GP3A, GPIIIa
D07	Hs.632226NM_000213	ITGB4	Integrin, beta 4	CD104
D08	Hs.536663NM_002213	ITGB5	Integrin, beta 5	FLJ26658
D09	Hs.521869NM_000216	KAL1	Kallmann syndrome 1 sequence	ADMLX, HHA, KAL, KALIG-1, KMS
D10	Hs.270364NM_005559	LAMA1	Laminin, alpha 1	LAMA
D11	Hs.200841NM_000426	LAMA2	Laminin, alpha 2	LAMM
D12	Hs.436367NM_000227	LAMA3	Laminin, alpha 3	BM600, E170, LAMNA, LOCS, lama3a

E01	Hs.650585NM_002291	LAMB1	Laminin, beta 1	CLM, MGC142015
E02	Hs.497636NM_000228	LAMB3	Laminin, beta 3	BM600-125KDA, FLJ99565, LAM5, LAMNB1
E03	Hs.609663NM_002293	LAMC1	Laminin, gamma 1 (formerly LAMB2)	LAMB2, MGC87297
E04	Hs.83169 NM_002421	MMP1	Matrix metalloproteinase 1 (interstitial collagenase)	CLG, CLGN
E05	Hs.2258 NM_002425	MMP10	Matrix metalloproteinase 10 (stromelysin 2)	SL-2, STMY2
E06	Hs.143751NM_005940	MMP11	Matrix metalloproteinase 11 (stromelysin 3)	SL-3, ST3, STMY3
E07	Hs.1695 NM_002426	MMP12	Matrix metalloproteinase 12 (macrophage elastase)	HME, MGC138506, MME
E08	Hs.2936 NM_002427	MMP13	Matrix metalloproteinase 13 (collagenase 3)	CLG3
E09	Hs.2399 NM_004995	MMP14	Matrix metalloproteinase 14 (membrane-inserted)	MMP-X1, MT1-MMP, MTMMP1
E10	Hs.80343 NM_002428	MMP15	Matrix metalloproteinase 15 (membrane-inserted)	MT2-MMP, MTMMP2, SMCP-2
E11	Hs.546267NM_005941	MMP16	Matrix metalloproteinase 16 (membrane-inserted)	C8orf57, DKFZp761D112, MMP- X2, MT-MMP2, MT- MMP3, MT3-MMP
E12	Hs.513617NM_004530	MMP2	Matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	CLG4, CLG4A, MMP-II, MONA, TBE-1
F01	Hs.375129NM_002422	MMP3	Matrix metalloproteinase 3 (stromelysin 1, progelatinase)	CHDS6, MGC126102, MGC126103, MGC126104, MMP-3, SL-1, STMY, STMY1, STR1
F02	Hs.2256 NM_002423	MMP7	Matrix metalloproteinase 7 (matrilysin, uterine)	MMP-7, MPSL1, PUMP- 1
F03	Hs.161839NM_002424	MMP8	Matrix metalloproteinase 8 (neutrophil collagenase)	CLG1, HNC, PMNL-CL
F04	Hs.297413NM_004994	MMP9	Matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)	CLG4B, GELB, MMP-9
F05	Hs.503878NM_000615	NCAM1	Neural cell adhesion molecule 1	CD56, MSK39, NCAM
F06	Hs.514412NM_000442	PECAM1	Platelet/endothelial cell adhesion molecule	CD31, FLJ58394, PECAM-1
F07	Hs.89546 NM_000450	SELE	Selectin E	CD62E, ELAM, ELAM1, ESEL, LECAM2
F08	Hs.82848 NM_000655	SELL	Selectin L	CD62L, LAM-1, LAM1, LECAM1, LNHR, LSEL, LYAM1, Leu-8, Lyam-1, PLNHR, TQ1, hLHRc
F09	Hs.73800 NM_003005	SELP	Selectin P (granule membrane protein 140kDa, antigen CD62)	CD62, CD62P, FLJ45155, GMP140, GRMP, LECAM3, PADGEM, PSEL
F10	Hs.371199NM_003919	SGCE	Sarcoglycan, epsilon	DYT11, ESG
F11	Hs.111779NM_003118	SPARC	Secreted protein, acidic, cysteine-rich (osteonectin)	ON
F12	Hs.185597NM_003119	SPG7	Spastic paraplegia 7 (pure and	CAR, CMAR, FLJ37308,

				complicated autosomal recessive)	MGC126331, MGC126332, PGN, SPG5C
G01	Hs.313	NM_000582	SPP1	Secreted phosphoprotein 1	BNSP, BSPI, ETA-1, MGC110940, OPN
G02	Hs.369397	NM_000358	TGFBI	Transforming growth factor, beta-induced, 68kDa	BIGH3, CDB1, CDG2, CDGG1, CSD, CSD1, CSD2, CSD3, EBMD, LCD1
G03	Hs.164226	NM_003246	THBS1	Thrombospondin 1	THBS, THBS-1, TSP, TSP-1, TSP1
G04	Hs.371147	NM_003247	THBS2	Thrombospondin 2	TSP2
G05	Hs.169875	NM_007112	THBS3	Thrombospondin 3	MGC119564, MGC119565, TSP3
G06	Hs.522632	NM_003254	TIMP1	TIMP metalloproteinase inhibitor 1	CLGI, EPA, EPO, FLJ90373, HCI, TIMP
G07	Hs.633514	NM_003255	TIMP2	TIMP metalloproteinase inhibitor 2	CSC-21K
G08	Hs.644633	NM_000362	TIMP3	TIMP metalloproteinase inhibitor 3	HSMRK222, K222, K222TA2, SFD
G09	Hs.476092	NM_003278	CLEC3B	C-type lectin domain family 3, member B	DKFZp686H17246, TN, TNA
G10	Hs.143250	NM_002160	TNC	Tenascin C	HXB, MGC167029, TN
G11	Hs.109225	NM_001078	VCAM1	Vascular cell adhesion molecule 1	CD106, DKFZp779G2333, INCAM-100, MGC99561
G12	Hs.2257	NM_000638	VTN	Vitronectin	V75, VN, VNT
H01	Hs.534255	NM_004048	B2M	Beta-2-microglobulin	-
H02	Hs.412707	NM_000194	HPRT1	Hypoxanthine phosphoribosyltransferase 1	HGPRT, HPRT
H03	Hs.523185	NM_012423	RPL13A	Ribosomal protein L13a	-
H04	Hs.592355	NM_002046	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	G3PD, GAPD, MGC88685
H05	Hs.520640	NM_001101	ACTB	Actin, beta	PS1TP5BP1
H06	N/A	SA_00105	HGDC	Human Genomic DNA Contamination	HIGX1A
H07	N/A	SA_00104	RTC	Reverse Transcription Control	RTC
H08	N/A	SA_00104	RTC	Reverse Transcription Control	RTC
H09	N/A	SA_00104	RTC	Reverse Transcription Control	RTC
H10	N/A	SA_00103	PPC	Positive PCR Control	PPC
H11	N/A	SA_00103	PPC	Positive PCR Control	PPC
H12	N/A	SA_00103	PPC	Positive PCR Control	PPC

Appendix III

WT1 GeneArt sequence

CGGAAGGCCCATGAGGCCAGTTAATTAAGAGGTACCGCGGCCGCATGATCAAGAATGCTG
ACCCTGAGGAGCTGAAGCGAGTTGCCTCACAGCCCACCAAGTGACTTCTTCTTCTCGTCAA
TGACTTCAGCATCTTGAGGACACTACTGCCCCCTCGTTTCCCGGAGAGTGTGCACGACTGCT
GGTGGCGTGCCTGTGACCCGACCTCCGGATGACTCGACCTCTGCTCCACGAGACCTGGTGC
TGTCTGAGCCAAGCAGCCAATCCTTGAGAGTACAGTGGACAGCGGCCAGTGGCCCTGTG
ACTGGCTACAAGGTCCAGTACACTCCTCTGACGGGGCTGGGACAGCCACTGCCGAGTGAG
CGGCAGGAGGTGAACGTCCCAGCTGGTGAGACCAGTGTGCGGCTGCGGGGTCTCCGGCCA
CTGACCGAGTACCAAGTGACTGTGATTGCCCTCTACGCCAACAGCATCGGGGAGGCTGTG
AGCGGGACAGCTCGGACCACTCTAGAGAGCTCATGGCGCGCCTAGGCCTTGACGGCCTTC
CG

Rep1Max GeneArt sequence

CGGAAGGCCCATGAGGCCAGTTAATTAAGAGGTACCGCGGCCGCATGATCAAGAATGCTG
ACCCTGAGGAGCTGAAGCGAGTTGCCTCACAGCCCACCAAGTGACTTCTTCTTCTCGTCAA
TGACTTCAGCATCTTGAGGACACTACTGCCCCCTCGTTTCCCGGAGAGTGTGCACGACTGCT
GGTGGCGTGCCTGTGACCCGACCTCCGGATGACTCGACCTCTGCTCCACGAGACCTGGTGC
TGTCTGAGCCAAGTTCGCAAGTCTCCGGGTCCAGTGGACAGCGGCCAGTGGCCCTGTG
ACTGGCTACAAGGTCCAGTACACTCCTCTGACGGGGCTGGGACAGCCACTGCCGAGTGAG
CGGCAGGAGGTGAACGTCCCAGCTGGTGAGACCAGTGTGCGGCTGCGGGGTCTCCGGCCA
CTGACCGAGTACCAAGTGACTGTGATTGCCCTCTACGCCAACAGCATCGGGGAGGCTGTG
AGCGGGACAGCTCGGACCACTCTAGAGAGCTCATGGCGCGCCTAGGCCTTGACGGCCTTC
CG

Rep1 GeneArt sequence

CGGAAGGCCCATGAGGCCAGTTAATTAAGAGGTACCGCGGCCGCATGATCAAGAATGCTG
ACCCTGAGGAGCTGAAGCGAGTTGCCTCACAGCCCACCAAGTGACTTCTTCTTCTCGTCAA
TGACTTCAGCATCTTGAGGACACTACTGCCCCCTCGTTTCCCGGAGAGTGTGCACGACTGCT
GGTGGCGTGCCTGTGACCCGACCTCCGGATGACTCGACCTCTGCTCCACGAGACCTGGTGC
TGTCTGAGCCAAGTTCGCAAGTCTCCGGGTCCAGTGGACAGCGGCCAGTGGCCCTGTG
ACTGGCTACAAGGTCCAGTACACTCCTCTGACGGGGCTGGGACAGCCACTGCCGAGTGAG
CGGCAGGAGGTGAACGTCCCAGCTGGTGAGACCAGTGTGCGGCTGCGGGGTCTCCGGCCA
CTGACCGAGTACCAAGTGACTGTGATTGCCCTCTACGCCAACAGCATCGGGGAGGCTGTG
AGCGGGACAGCTCGGACCACTCTAGAGAGCTCATGGCGCGCCTAGGCCTTGACGGCCTTC
CG

WT3 GeneArt Sequence

CGAATTGAAGGAAGGCCGTCAAGGCCGCATGGTACCGCGGCCGCATGGTAGACACAGGA
CCTGGAGCCAGAGAGAAGGGAGAGCCTGGGGACCGCGGACAAGAGGGTCTCGAGGGCC
CAAGGGTGATCCTGGCCTCCCTGGAGCCCCTGGGGAAAGGGGCATTGAAGGGTTTCGGGG
ACCCCCAGGCCACAGGGGGACCCAGGTGTCCGAGGCCAGCAGGAGAAAAGGGTGACC
GGGGTCCCCCTGGGCTGGATGGCCGGAGCGGACTGGATGGGAAACCAGGAGCCGCTGGG
CCCTCTGGGCGCAATGGTGCTGCAGGCAAAGCTGGGGACCCAGGGAGAGACGGGCTTC
CAGGCCTCCGTGGAGAACAGGGCCTCCCTGGCCCCCTCTGGTCCCCCTGGATTACCGGGAA
AGCCAGGCGAGGATGGCAAACCTGGCCTGAATGGAAAAACGGAGAACCTGGGGACCT
GGAGAAGACGGGAGGAAGGGAGAGAAAGGAGATTCAGGCGCCTCTGGGAGAGAAGGTC
TAGAGAGCTCCTGGGCCTCATGGGCCTTCCTTTCACTGCC

Rep3 GeneArt Sequence

CGAATTGAAGGAAGGCCGTCAAGGCCGCATGGTACCGCGGCCGCATGGTAGACACAGGA
CCTGGAGCCAGAGAGAAGGGAGAGCCTGGGGACCGCGGACAAGAGGGTCTCGAGGGCC
CAAGGGTGATCCTGGCCTCCCTGGAGCCCCCTGGGGAAAGGGGCATTGAAGGGTTTCGGGG
ACCCCCAGGCCACAGGGGGACCCAGGTGTCCGAGGCCCAGCAGGAGAAAAGGGTGACC
GGGGTCCCCCTGGGCTGGATGGCCGGAGCGGACTGGATGGGAAACCAGGAGCCGCTGGG
CCCTCTGGGCCCAACCGGGCGCCGGGAAAGCTGGGGACCCAGGGAGAGACGGGCTTC
CAGGCCTCCGTGGAGAACAGGGCCTCCCTGGCCCCCTCTGGTCCCCCTGGATTACCGGGAA
AGCCAGGCGAGGATGGCAAACCTGGCCTGAATGGAAAAAACGGAGAACCTGGGGACCCT
GGAGAAGACGGGAGGAAGGGAGAGAAAGGAGATTACGGCGCCTCTGGGAGAGAAGGTC
TAGAGAGCTCCTGGGCCTCATGGGCCTTCCTTTCACTGCC

WT5 GeneArt Sequence

GGGCAGTGAGCGGAAGGCCCATGAGGCCAGTTAATTAAGCGGCCGCATGGGCCAGGAAC
AGTCCCGGGTCTCTGGGGCCCCGAGCTCAGCAGCTATCACCTGGACGGGCTGGAGCCAGCGA
CACAGTACCGCGTGAGGCTGAGTGTCTTAGGGCCAGCTGGAGAAGGGCCCTCTGCAGAGG
TGACTGCGCGCACTGAGTCACCTCGTGTTC AAGCATTGAACTACGTGTGGTGGACACCT
CGATCGACTCGGTGACTTTGGCCTGGACTCCAGTGTCCAGGGCATCCAGCTACATCCTAT
CCTGGCGGCCACTCAGAGGCCCTGGCCAGGAAGTGCCTGGGTCCCCGCAGACACTTCCAG
GGATCTCAAGCTCCCAGCGGGTGACAGGGCTAGAGCCTGGCGTCTCTTACATCTTCTCCCT
GACGCCTGTCCTGGATGGTGTGCGGGGTCTGAGGCATCTGTACACAGACGCCAGTGTC
GCCCCGTCTAGAGGCGCGCCTAGGCCTTGACGGCCTTCCGCCAATTTCGCC

Rep5Max GeneArt sequence

GGGCAGTGAGCGGAAGGCCCATGAGGCCAGTTAATTAAGCGGCCGCATGGGCCAGGAAC
AGTCCCGGGTCTCTGGGGCCCCGAGCTCAGCAGCTATCACCTGGACGGGCTGGAGCCAGCGA
CACAGTACCGCGTGAGGCTGAGTGTCTTAGGGCCAGCTGGAGAAGGGCCCTCTGCAGAGG
TGACTGCGCGCACTGAGTCACCTCGTGTTC AAGCATTGAACTACGTGTGGTGGACACTA
GCATTGATAGCGTCAACCTGGCCTGGACTCCAGTGTCCAGGGCATCCAGCTACATCCTAT
CCTGGCGGCCACTCAGAGGCCCTGGCCAGGAAGTGCCTGGGTCCCCGCAGACACTTCCAG
GGATCTCAAGCTCCCAGCGGGTGACAGGGCTAGAGCCTGGCGTCTCTTACATCTTCTCCCT
GACGCCTGTCCTGGATGGTGTGCGGGGTCTGAGGCATCTGTACACAGACGCCAGTGTC
GCCCCGTCTAGAGGCGCGCCTAGGCCTTGACGGCCTTCCGCCAATTTCGCC

Rep5 GeneArt sequence

GGGCAGTGAGCGGAAGGCCCATGAGGCCAGTTAATTAAGCGGCCGCATGGGCCAGGAAC
AGTCCCGGGTCTCTGGGGCCCCGAGCTCAGCAGCTATCACCTGGACGGGCTGGAGCCAGCGA
CACAGTACCGCGTGAGGCTGAGTGTCTTAGGGCCAGCTGGAGAAGGGCCCTCTGCAGAGG
TGACTGCGCGCACTGAGTCACCTCGTGTTC AAGCATTGAACTACGTGTGGTGGACACCT
CCATCGATTCCGTCAACCTGGCCTGGACTCCAGTGTCCAGGGCATCCAGCTACATCCTAT
CCTGGCGGCCACTCAGAGGCCCTGGCCAGGAAGTGCCTGGGTCCCCGCAGACACTTCCAG
GGATCTCAAGCTCCCAGCGGGTGACAGGGCTAGAGCCTGGCGTCTCTTACATCTTCTCCCT
GACGCCTGTCCTGGATGGTGTGCGGGGTCTGAGGCATCTGTACACAGACGCCAGTGTC
GCCCCGTCTAGAGGCGCGCCTAGGCCTTGACGGCCTTCCGCCAATTTCGCC

Appendix IV

Exons 83-88 WT GeneArt sequence

ggtaccgcgccgcggggagcctggagagacaggacctccaggacggggcctgactggacactactggagctgtggga
cttcttgacccccggcccttcaggccttgtgggtccacaggggtctccaggtttgcctggacaagtgggggagacagg
gaagccgggagccccaggtcagatggtgccagtggaaaagatggagacagaggagccctggtgtgccagggtcac
caggtctgcctggccctgtcggacctaaaggagaacctggccccacgggggccctggacaggctgtggtcgggctcc
ctggagcaaagggagagaagtctagagctc

Exons 83-88 6863del16 GeneArt sequence

ggtaccgcgccgcggggagcctggagagacaggacctccaggacggggcctgactggacactactggagctgtggga
cttcttgacccccggcccttcaggccttgtgggtccacaggggtctccaggtttgcctggacaagtgggggagacagg
gaagccgggagccccaggtcagatggtgccagtggaaaagatggagacagaggagccctggtgtgccagggtcac
caggtctgcctggccctgtcggacccccacgggggccctggacaggctgtggtcgggctccctggagcaaagggaga
gaagtctagagctc

Exons 73-75 WT GeneArt sequence

ggtaccgcgccgcggccccatcggtttcttgagaacgcgggctgaagggcgaccgtggagaccctggccctcag
gggccacctggtctggcccttggggagagggggccccccgggccttcggccttgccggggagcctggaaagcctggt
attccgggctcccaggcagggtgggggtgtgggagaggcaggaaggccaggagagaggggagaacggggagag
aaaggagaacgtggagaacagggcagagatggccctcctggactccctggaaccctgggccccccggacccctgg
cccaagtctagagctc

Exons 73-75 G2043R GeneArt sequence

ggtaccgcgccgcggccccatcggtttcttgagaacgcgggctgaagggcgaccgtggagaccctggccctcag
gggccacctggtctggcccttggggagagggggccccccgggccttcggccttgccggggagcctggaaagcctggt
attccgggctcccaggcagggtgggggtgtgggagaggcaggaaggccaggagagaggggagaacggggagag
aaaggagaacgtggagaacagggcagagatggccctcctggactccctggaaccctgggccccccggacccctgg
cccaagtctagagctc