Ameliorating β-thalassaemia by manipulating expression of the α-globin gene

Doluweera Sachith Gopeetha Mettananda

Brasenose College

A thesis submitted for the degree of
Doctor of Philosophy
of
University of Oxford

MRC Molecular Haematology Unit
Weatherall Institute of Molecular Medicine
University of Oxford

Trinity Term 2015

38,000 words

This thesis is dedicated to my wife Chamila

Abstract

Doluweera Sachith Gopeetha Mettananda
Brasenose College

Doctor of Philosophy
Trinity Term 2015

β-Thalassaemia is a disorder of haemoglobin production characterised by severe anaemia requiring life-long blood transfusions. The common genetic defects are predominantly based in and around the β-globin gene resulting in reduced or absent β-globin chain synthesis. The resultant excess of free α-globin chains, which precipitates in red blood cells and their precursors and causes ineffective erythropoiesis, is the main pathophysiological mechanism of anaemia in these patients. Clinical and genetic data accumulated over the last 30 years have indicated that reduction of α-globin expression is clinically beneficial to patients with β-thalassaemia, and in the subset of patients with HbE β-thalassaemia it could be transformational. Attention has been centred on pathways that increase γ-globin expression and hence the production of foetal haemoglobin, but I have explored avenues that down regulate the α-globin gene expression without affecting the β-like globin expression, to equalise globin chain imbalance and could have potential clinical applications.

Initially, an *in vitro* erythroid differentiation system in serum-free medium was optimised and characterised. This generates a large number of erythroid cells from human $CD34^+$ cells with minimal non-erythroid contamination. This system appears to be a faithful recapitulation of normal erythropoiesis and was validated with newly developed assays. A library of epigenetically active small molecules and licenced drugs was then screened and this identified a number of interesting compounds including IOX1, a histone demethylase inhibitor. Secondary assays confirmed IOX1 as a selective down regulator of α-globin expression, with no significant adverse effects on β-like globin expression, erythroid cell viability, erythroid differentiation or global erythroid transcriptome.

Next, CRISPR/Cas9 genome editing technique was used to introduce targeted deletions of multispecies conserved sequences (MCS) R2 region, the most critical distant *cis*-acting regulatory element enhancing α-globin expression in humans. High transfection and mutation efficiencies were achieved in human $CD34^+$ cells and single cell assays confirmed high frequency of homozygous and heterozygous deletion events. More importantly, and analogous to patients with an identical mutation, the deletion of MCS-R2 element resulted in selective knockdown of α-globin expression without altering β-globin expression or erythroid differentiation. In conclusion, I have demonstrated that selective down regulation of α-globin expression is plausible through pharmacological and genome engineering approaches and these findings unveil new pathways of therapy for β-thalassaemia.

Acknowledgements

I would like to express thanks to my supervisors Doug Higgs and Richard Gibbons for giving me the opportunity to work in the lab and expertly guiding the work presented in this thesis over the last three years. I am also indebted to Chris Fisher who patiently guided my transition from *bedside* to *bench* and was available for help and advice throughout my time in the lab. I also acknowledge Sir David Weatherall and Hal Drakesmith of my thesis committee for invaluable advice and support. This work would not have been possible without funding from the Commonwealth Scholarship Commission in the UK and Medical Research Council UK.

This thesis has been informed by numerous suggestions from members of the Higgs and Gibbons laboratories both past and present: David Clynes, Caroline Scott, Ricarda Gaentzsch, Barbara Xella, Diu Nguyen, Rosa Mata, Julia Truch, Jim Hughes, Helena Ayyub, Deborah Hay, Noémi Roy, Hsiao Voon, Bryony Graham, Andrew Campbell, Christian Babbs, James Davies, Lars Hanssen, Anthony Cheong to name but a few.

I would also like to thank Jackie Sloane-Stanley, Jackie Sharpe and Sue Butler for invaluable assistance with tissue culture and animal work; Kevin Clark and Paul Sopp for support in flow cytometry and sorting; Philip Hublitz for helping with plasmid cloning; Stephen Taylor for assistance in bioinformatics; Nicki Gray for helping with computational graphics; Jennifer Eglinton for the help with haemoglobin analysis and Udo Oppermann for providing valuable reagents. My gratitude also goes to Sue Harper and Allan Hayle whose efforts made working in the department an enjoyable experience.

None of the work presented here would have been possible without the support of my parents and parents-in-law who provided essential support in child care. Finally my heartfelt gratitude goes to Kalya and Kaushala for being wonderful kids understanding my absence from home and to Chamila for being a supportive, caring and loving wife not only during the last three years but for past eleven years.

List of abbreviations

A	Alanine
ActR	Activin receptor
AHSP	α-haemoglobin stabilizing protein
APC	Allophycocyanin
ATP	Adenosine triphosphate
BCL11A	B-cell lymphoma/leukaemia 11A protein
BM	Bone marrow
C	Cytosine
C9	Cas9-only control
CA2	Carbonic anhydrase 2
Cas	CRISPR-associated
CCR5	C-C chemokine receptor type 5
CD	Cluster of differentiation
cDNA	Complementary DNA
CE-HPLC	Cation exchange high performance liquid chromatography
ChIP	Chromatin Immunoprecipitation
CpG	Cytosine guanine dinucleotide
CRISPR	Clustered, regularly interspaced, short palindromic repeat
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
eGFP	Enhanced green fluorescent protein
FAS	Fas cell surface death receptor
FDA	Food and Drug Administration
FITC	Fluorescein isothiocyanate
FLAG	FLAG Expression System
FS	Forward Scatter
FTH1	Ferritin, Heavy Polypeptide 1
G	Guanine
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GATA1	GATA-binding factor1
GC	Guanine-cytosine

GDF	Growth differentiation factor
GFP	Green fluorescent protein
gRNA	Guide RNA
GVHD	Graft versus host disease
H	Histone
HAT	Histone acetyl transferases
Hb	Haemoglobin
HBA	α-globin
HbA	Adult haemoglobin
HBB	β-globin
HBD	δ-globin
HBE	ε-globin
HbE	Haemoglobin E
HbF	Foetal haemoglobin
HBG	γ-globin
HbH	Haemoglobin H
HBM	μ-globin
HBQ	θ-globin
HBZ	ζ-globin
HDAC	Histone deacetylase
HDACi	HDAC inhibitor
HDR	Homology directed repair
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
Het	Mice heterozygous for the human α-globin cluster
HIF	Hypoxia inducible factor
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HMT	Histone methyltransferase
HPLC	High performance liquid chromatography
HSC	Haematopoietic stem cell
HSCT	Haematopoietic stem cell transplantation
HSP70	Heat shock protein 70
HSPC	Haematopoietic stem and progenitor cell
HTS	High-throughput screening
IC_{50}	Half-maximum inhibitory concentration

IEF	isoelectric focussing
IgG	Immunoglobulin G
JAK2	Janus kinase 2
JmjC	JumonjiC
JMJD3	Jumonji domain-containing protein 3
K	Lysine
kb	Kilo bases
KDM	Histone demethylase
KLF1	Kruppel-like factor 1
LAPTM4A	Lysosomal Protein Transmembrane 4 Alpha
LCI	Labile cellular iron
LPI	Labile plasma iron
LSD1	Lysine specific demethylase 1
MACS	Magnetic activated cell sorting
MCS	Multispecies conserved sequences
MNC	Mononuclear cell
mRNA	Messenger ribonucleic acid
MTC	Mock transfection control
NF-E2	Nuclear factor-erythroid 2
NHEJ	Non-homologous end joining
NHS	National Health Service
NLS	Nuclear localization signal
NTBI	Non-transferrin-bound iron
NTC	Non-transfection control
PABPC1	Polyadenylate-binding protein 1
PAIP	Poly(A) Binding Protein Interacting Protein 2
PB	Peripheral blood
PBMC	Peripheral blood mononuclear cells
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PE	Phycoerythrin
PolII	RNA polymerase II
PRC	Polycomb repressive complex
qRTPCR	Quantitative real-time polymerase chain reaction
QTL	Quantitative trait loci

RBC	Red blood cell
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
ROS	Reactive oxygen species
RP-HPLC	Reversed-phase HPLC
RPL13A	Ribosomal Protein L13a
RPL18	Ribosomal protein L18
RT	Reverse transcription
SCD	Sickle cell disease
SCF	Stem cell factor
SCL	Stem cell leukaemia
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SGC	Structural Genome Consortium
shRNA	Short-hairpin RNA
siRNA	Short-interfering RNA
STAT5	Signal transducer and activator of transcription 5
TALEN	Transcription Activator-Like Effector Nucleases
TBE	Tris/Borate/EDTA
TBS	Tris-buffered saline
TBST	Tris-buffered saline and Tween 20
TF	Transcription factor
UC	Umbilical cord
UCB	Umbilical cord blood
US	United States
WT	Wild type
ZFN	Zinc finger nucleases
β-KO	β-globin knockout
β-LCR	β-locus control region
ΔMCS-R2	MCS-R2 knockout
2OG	2-oxoglutarate

Table of contents

Chapter 1: Introduction 1
1.1 Thalassaemia 1
1.2 Haemoglobin and globin genes 2
1.3 β-Thalassaemia 5
1.3.1 Epidemiology and global burden 5
1.3.2 Molecular basis 6
1.3.3 Cellular pathology 7
1.3.4 Pathophysiology 11
1.3.5 Clinical features and complications 14
1.3.6 Genetic modifiers of disease severity 16
1.4 Treatment of β-thalassaemia 18
1.4.1 Supportive therapy 18
1.4.2 Curative therapy 19
1.4.3 Experimental therapies 21
1.4.4 Limitations of the current and experimental therapies 26
1.5 Reduction of α-globin as a therapeutic option for β-thalassaemia 27
1.5.1 Co-inheritance of α-and β-thalassaemia 27
1.5.2 Effect of excess α-globin chains on β-thalassaemia 30
1.5.3 Evidence from mouse models 30
1.5.4 Desired level of reduction of α-globin 31
1.6 Regulation of α-globin 32
1.6.1 Distal enhancers of the α-globin 32
1.6.2 Epigenetic regulation and histone modifications of α-globin 34
1.6.3 Transcriptional factor control of α-globin 37
1.6.4 Contrasting regulation α- and β-globin expression 37
1.7 Selective silencing of α-globin expression: candidate therapeutic pathways 39
1.7.1 RNA interference strategy 40
1.7.2 Pharmacological strategy through epigenetic drug targeting 42
1.7.3 Genome editing strategy 43
1.8 Summary, hypothesis and aim of the thesis 45

Chapter 2: Materials and Methods 47
2.1 Cell culture methods and associated techniques 47
2.1.1 Purification of human $CD34^+$ HSPCs 47
2.1.2 Erythroid differentiation culture from human $CD34^+$ HSPCs 47
2.1.3 Erythroid differentiation culture from sorted single human HSPCs 48
2.1.4 Erythroblast culture from mouse foetal liver 48
2.1.5 Trypan blue test for viability 49
2.1.6 Benzidine stain for haemoglobin 49
2.1.7 Morphology analysis by cytospin 50
2.1.8 Flow cytometry 50
2.2 Small molecule library and treatment of cells 50
2.3 DNA methods 52
2.3.1 DNA extraction 52
2.3.2 Polymerase chain reaction (PCR) 53
2.3.3 Whole genome DNA amplification from small number of cells 53
2.3.4 Sanger sequencing 54
2.4 RNA methods 54

2.4.1 RNA extraction 54
2.4.2 Reverse transcription (RT) 55
2.4.3 cDNA preparation from small number of cells 55
2.4.4 Quantitative real-time polymerase chain reaction (qRTPCR) 56
2.4.5 High-throughput qRTPCR by Fluidigm Biomark HD system 57
2.4.6 Nanostring gene expression assay 58
2.4.7 Microarray 58
2.5 Protein methods 59
2.5.1 Analysis of haemoglobin 59
2.5.2 Histone protein extraction 59
2.5.3 Western blot 60
2.6 Chromatin Immunoprecipitation (ChIP) 60
2.6.1 Histone protein ChIP 60
2.6.2 ChIP-PCR 61
2.7 Transient gene knock-down methods 62
2.7.1 Transfection of short-interfering RNA (siRNA) 62
2.7.2 Short-hairpin RNA (shRNA) 63
2.8 Genome editing methods 63
2.8.1 Preparation of single-guide RNA (sgRNA) expression construct 63
2.8.2 Transfection 65
2.8.3 Cell sorting 66

Chapter 3: Miniature erythroid differentiation system for high-throughput and single-cell experiments 67
3.1 Introduction 67
3.2 Results 71
3.2.1 Purification of $CD34^+$ HSPCs 71
3.2.2 Optimisation of the small-scale erythroid differentiation system 72
3.2.3 Cell expansion and viability during erythroid differentiation 73
3.2.4 Morphological characterisation of cells 75
3.2.5 Immunophenotypical characterisation of cells 76
3.2.6 Gene expression during erythroid differentiation 79
3.2.7 Characterisation of haemoglobinisation and haemoglobin subtypes 83
3.2.8 System validation 85
3.2.9 Erythroid differentiation from sorted single $CD34^+$ cells 87
3.3 Discussion 89

Chapter 4: Targeted small molecule screen to identify human α-globin silencers 94
4.1 Introduction 94
4.2 Results 96
4.2.1 Selection of the epigenetically active small molecule library 96
4.2.2 Medium-throughput screen of the epigenetic inhibitor small molecule library 96
4.2.3 Identification of compounds which down-regulate α-globin expression 100
4.2.4 Tranylcypromine down-regulates α-globin expression by inhibiting erythroid differentiation 102
4.2.5 HDAC inhibitors induce foetal and embryonic globin gene expression in the β-globin gene cluster 104
4.2.6 DNMT inhibitor induces embryonic globin gene expression 106
4.3 Discussion 107

Chapter 5: Selective silencing of human α-globin by histone demethylase inhibitor, IOX1 111
5.1 Introduction 111
5.2 Results 113
5.2.1 Down-regulation of α-globin by IOX1 is dose-dependent 113
5.2.2 Effect of IOX1 on globin mRNA is predominantly seen at the α-globin locus 114
5.2.3 Effects of IOX1 on erythroid cell proliferation and viability 116
5.2.4 IOX1 does not alter erythroid differentiation 117
5.2.5 IOX1 exerts relatively small effects on the erythroid transcriptome 119
5.2.6 IOX1 increases the H3K27me3 repressive chromatin modification globally and locally at the α-globin locus 124
5.2.7 shRNA mediated knockdown of KDM6B 125
5.2.8 Effect of IOX1 on human and mouse α-globin expression in humanised mouse erythroid cells 127
5.3 Discussion 131

Chapter 6: Histone deacetylase inhibitor, vorinostat down-regulates α-globin expression whilst up-regulating γ-globin 135
6.1 Introduction 135
6.2 Results 138
6.2.1 Effects of vorinostat on globin gene expression 138
6.2.2 Effects of vorinostat on erythroid cell proliferation and viability 141
6.2.3 Effects of vorinostat on erythroid differentiation 142
6.2.4 Effects of vorinostat on erythroid transcriptome 144
6.2.5 Combination of IOX1 and vorinostat 145
6.3 Discussion 146

Chapter 7: CRISPR/Cas9 mediated deletion of the α-globin enhancer in human CD34$^+$ cells 150
7.1 Introduction 150
7.1.1 *Cis*-acting regulatory elements of the α-globin gene 150
7.1.2 MCS-R2 is the most critical enhancer of α-globin 150
7.1.3 Naturally occurring deletion limited to MCS-R2 153
7.1.4 Implications of deletion of MCS-R2 in β-thalassaemia 155
7.1.5 Genome editing in HSCs 156
7.2 Results 159
7.2.1 Design of guide RNAs 159
7.2.2 Preparation of the gRNA expression construct 160
7.2.3 Transfection of human CD34$^+$ cells 161
7.2.4 Erythroid differentiation of CD34$^+$ cells after nucleofection 163
7.2.5 Selection of the most efficient CRISPR/Cas9 plasmid pairs 165
7.2.6 Deletion of MCS-R2 regulatory element selectively down-regulates α-globin expression 166
7.2.7 Deletion of MCS-R2 regulatory element does not have significant effects on erythroid differentiation 169
7.2.8 Analysis of effects of deletion of the MCS-R2 regulatory element at the single cell level 171
7.3 Discussion and future work 176

Chapter 8: General discussion and conclusions 181
8.1 Selective silencing of human α-globin is therapeutically feasible 181
8.2 Limitations of the study 183
8.3 Epigenetic inhibitor therapy to down-regulate α-globin: therapeutic prospects and challenges 186
8.4 Genome editing to down regulate α-globin: therapeutic prospects and challenges 189
8.5 Conclusions 191

References 193

Appendix 211
Appendix 1 – Supplementary Tables 211
Appendix 2 – Publications, presentations and awards arising from this thesis 223

List of figures

Figure 1.1 – Structure of haemoglobin. 2

Figure 1.2 – Developmental expression of the human globin genes..... 3

Figure 1.3 – Schematic diagram of α- and β-globin gene clusters and the types of haemoglobin produced at each developmental stage. 4

Figure 1.4 – Electron microscopy images showing electron-dense inclusions consist of precipitated α-globin chains in erythroid cells of patients with β-thalassaemia 8

Figure 1.5 – Cellular pathology of β-thalassaemia..... 9

Figure 1.6 – Pathophysiology of β-thalassaemia..... 12

Figure 1.7 – Summary of experimental therapies for β-thalassaemia..... 211

Figure 1.8 – *Cis*-acting enhancers of the α-globin gene cluster..... 333

Figure 1.9 – Contrasting epigenetic landscape of α- and β- globin loci..... 366

Figure 2.1 – Map of the plasmid construct with pSpCas9(BB)-2A-GFP backbone and gRNA..... 64

Figure 3.1 – Purity of $CD34^+$ cells after MACS separation..... 71

Figure 3.2 – Growth and viability of erythroid cells differentiated from human UCB and adult $CD34^+$ cells 74

Figure 3.3 – Morphology of cells at different time points representing different stages of erythroid differentiation 76

Figure 3.4 – Immunophenotypical characterisation of cells at different time points in culture..... 78

Figure 3.5 – Expression levels of α-, β- and γ-globin genes in cells at different time points in culture..... 81

Figure 3.6 – Gene expression quantified by Nanostring at different time points in culture..... 82

Figure 3.7 – Haemoglobinisation and characterisation of haemoglobin sub-types..... 84

Figure 3.8 – Validation of the erythroid differentiation system and the globin gene expression detection protocol 86

Figure 3.9 – Erythroid differentiation from single $CD34^+$ cells..... 88

Figure 4.1 – Medium-throughput screen of the epigenetic inhibitor small molecule library..... 98

Figure 4.2 – Validation of Fluidigm BioMark data by standard qRTPCR..... 100

Figure 4.3 – α/β globin mRNA ratios in cells treated with small molecule inhibitors. 101

Figure 4.4 – Effects of tranylcypromine in erythroid cells..... 103

Figure 4.5 – Effects of HDAC inhibition on the β-globin cluster 105

Figure 4.6 – Effect of DNMT inhibition on embryonic globin gene expression 106

Figure 5.1 – Effects of IOX1 on α- and β-globin gene expression in human erythroid cells 114

Figure 5.2 – Effects of IOX1 on globin gene expression in erythroid cell 115

Figure 5.3 – Expansion and viability of erythroid cells treated with IOX1 116
Figure 5.4 – Effect of IOX1 on erythroid differentiation 118
Figure 5.5 – Effects of IOX1 on erythroid transcriptome. 120
Figure 5.6 – Effect of IOX1 on histone methylation 125
Figure 5.7 – KDM6B knockdown by shRNA in erythroid cells..... 126
Figure 5.8 – Schematic diagram of human and mouse α-globin loci 127
Figure 5.9 – Effect of IOX1 on human and mouse α-globin expression in humanised mouse primary erythroid cells 129
Figure 6.1 – Effects of vorinostat on globin gene expression in human erythroid cells 139
Figure 6.2 – Nanostring quantification of globin gene expression in erythroid cell after treatment with vorinostat..... 140
Figure 6.3 – Expansion and viability of erythroid cells treated with vorinostat..... 141
Figure 6.4 – Effects of vorinostat on erythroid differentiation..... 143
Figure 6.5 – Effect of vorinostat on the erythroid transcriptome..... 144
Figure 6.6 – Effects of a combination of IOX1 and vorinostat on globin gene expression in human erythroid cells..... 145
Figure 7.1 – Schematic representation of the α-globin gene cluster and the TF binding motifs in the MCS-R2 core element. 152
Figure 7.2 – Clinical data from individuals with the $(\alpha\alpha)^{ALT}$ (ΔMCS-R2) mutation..... 155
Figure 7.3 – Schematic of workflow of genome editing treatment for β-thalassaemia.... 158
Figure 7.4 – Schematic of gRNA target sites..... 159
Figure 7.5 – Summary of experimental workflow to generate deletions of the MCS-R2 region in human $CD34^+$ cells using the CRISPR/Cas9 system..... 160
Figure 7.6 – Optimisation and efficiency of the transfection protocol..... 162
Figure 7.7 – Cell number optimisation for transfection reactions 163
Figure 7.8 – Erythroid differentiation potential of $CD34^+$ cells after nucleofection.... 164
Figure 7.9 – Gel electrophoresis image of populations of $CD34^+$ cells targeted by different CRISPR/Cas9 plasmid pairs analysed by PCR. 166
Figure 7.10 – Deletion efficiencies and the effects of MCS-R2 deletion on human globin gene expression..... 168
Figure 7.11 – Effects of deletion of MCS-R2 region in human erythroid differentiation 170
Figure 7.12 –Deletion of MCS-R2 region using CRISPR/Cas9 plasmid pair cr1+8; single cell clone analysis..... 173
Figure 7.13 – Deletion of MCS-R2 region using CRISPR/Cas9 plasmid pair cr2+12; single cell clone analysis..... 174
Figure 8.1 – A summary of pharmacological and genome editing approaches, which aid in amelioration of β-thalassaemia uncovered in this thesis. 183

List of tables

Table 1.1 – Differences in the structure and function of the α- and β-globin gene clusters 38

Table 2.1 – List of small molecules included in the epigenetic inhibitor compound library 51

Table 2.2 – Primer sequences used in PCR 53

Table 2.3 – List of inventoried human TaqMan assays used in qRTPCR 56

Table 2.4 – Primer sequences and TaqMan probes used in mouse experiments 57

Table 2.5 – Primer sequences used for ChIP-PCR 62

Table 2.6 – gRNA sequences of CRISPR target sites within and around MCS-R2 enhancer region 65

Table 3.1 – Yield of $CD34^+$ cells from umbilical cord blood units 72

Table 4.1 – Fold changes in the expression of globin genes in erythroid cells treated with small molecule inhibitors 99

Table 5.1 – Number of differentially expressed transcripts in IOX1 treated cells compared to DMSO control cells (n=4) 121

Table 5.2 – Expression levels of genes essential for erythrocyte physiology in IOX1 treated cells compared to DMSO control cells (n=4) 121

Supplementary table 1 – List of genes significantly up-regulated in IOX1 treated cells (in alphabetical order) 2111

Supplementary table 2 – List of genes significantly down-regulated in IOX1 treated cells (in alphabetical order) 2144

Supplementary table 3 – List of genes significantly up-regulated in vorinostat treated cells (in alphabetical order) 2177

Supplementary table 4 – List of genes significantly down-regulated in vorinostat treated cells (in alphabetical order) 2211

Chapter 1: Introduction

1.1 Thalassaemia

Thalassaemia is the most common form of all inherited disorders of the red blood cell (RBC) (Weatherall and Clegg, 1996). It is one of the first human genetic diseases that was characterised at a molecular level and its study is considered to be the forerunner of a completely new era of medical science, which is now widely known as molecular medicine. Thalassaemia has been a subject of interest to many clinician-scientists and molecular biologists since the second half of the 20th century. This has led to an understanding of the remarkable diversity of human molecular pathology and in particular has provided some extremely valuable insights into how human genes are regulated and how these functions may break down in different inherited diseases (Weatherall and Clegg, 2001).

The first clinical description of thalassaemia was reported by Thomas B. Cooley in 1925, in a paper which characterised five children with anaemia, splenomegaly and peculiar bone changes (Cooley and Lee, 1925; Cooley et al., 1927). The term ‘thalassaemia’ was first used by Whipple and Bradford in 1932, to delineate this clinical condition. This term was derived from a combination of two Greek words: ‘thalassa’ meaning ‘the sea’ that is the Mediterranean and anaemia meaning ‘weak blood’ (Weatherall and Clegg, 2001). In 1952, Rich demonstrated the persistence of foetal haemoglobin in patients with thalassaemia and suggested that thalassaemia is due to a defect in the

synthesis of adult haemoglobin (Rich, 1952), which was further verified over the next few years (Weatherall and Clegg, 2001).

1.2 Haemoglobin and globin genes

Haemoglobin (Hb) is a highly specialised protein found within the RBC and is responsible for the delivery of oxygen from lungs to tissues. In humans, haemoglobin is a tetramer of two α-like and two β-like globin chains (Figure 1.1).

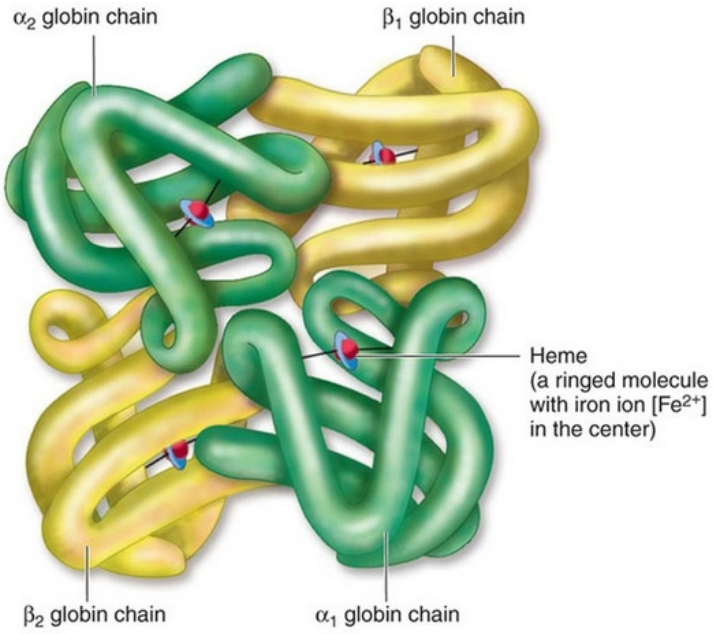


Figure 1.1 – Structure of haemoglobin. Schematic of an adult haemoglobin molecule demonstrating interactions between two α- and two β-globin chains and heme groups. Adapted from (LifestyleProTheme, 2015).

According to the changing needs of growth, development and survival, various haemoglobin isoforms and different α-like and β-like globin chains are synthesised in a developmental-stage-specific manner. Hb Gower-I ($\zeta_2\varepsilon_2$), Hb Gower-II ($\alpha_2\varepsilon_2$) and Hb Portland ($\zeta_2\gamma_2$) are produced during the embryonic stage and are then switched to HbF ($\alpha_2\gamma_2$) in the foetal stage and subsequently to HbA ($\alpha_2\beta_2$) and HbA_2 ($\alpha_2\delta_2$) in the adult stage (Figure 1.2 and Figure 1.3)(Higgs and Wood, 2011).

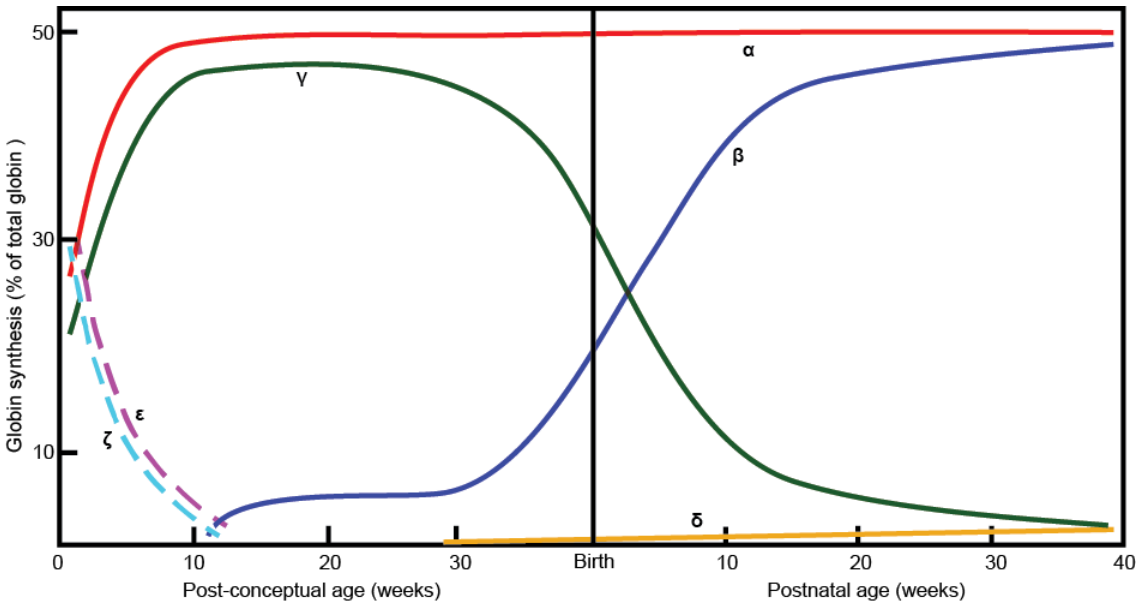


Figure 1.2 – Developmental expression of the human globin genes. ζ and ε are the first globin genes to be expressed during embryonic stage and are then switched to α and γ in the foetal stage. The second globin switch from γ to β occurs during the first few months after birth (Higgs et al., 2012).

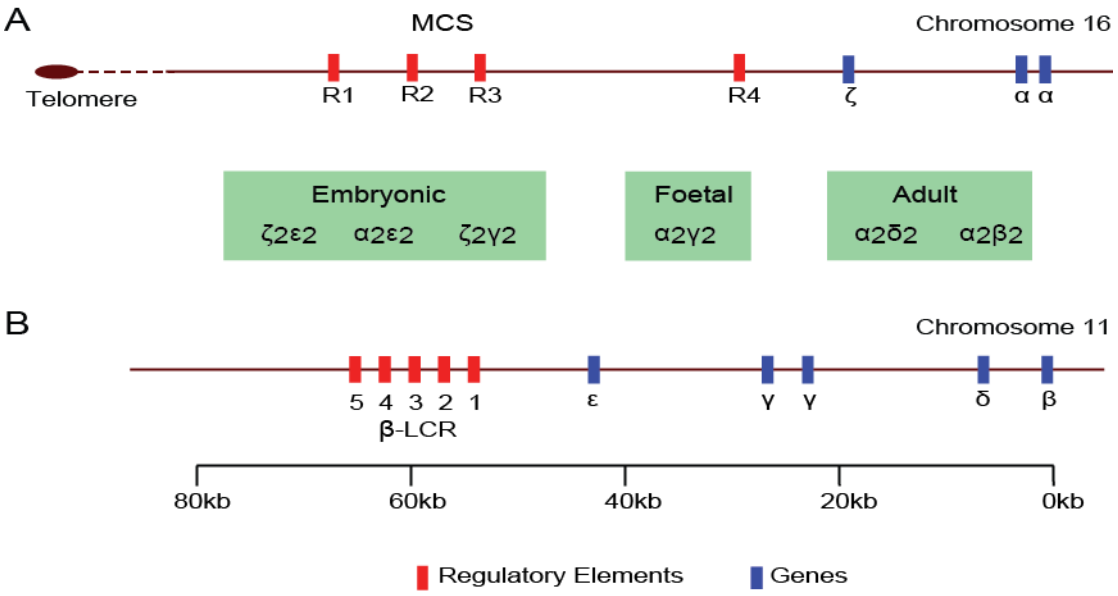


Figure 1.3 – Schematic diagram of α- and β-globin gene clusters and the types of haemoglobin produced at each developmental stage. Genes are arranged along the chromosome in the order in which they are expressed during development; (A) in the α-cluster ζ (embryonic) and α (foetal and adult); (B) in the β-cluster ε (embryonic), γ (foetal), δ and β (adult). The four upstream regulatory elements of the α-locus are known as multispecies conserved sequences (MCS) R1 to R4, whereas the five regulatory elements of the β-locus are collectively referred to as β-locus control region (β-LCR). Adapted from (Mettananda et al., 2015).

At the molecular level, the synthesis of globin chains is controlled by two multigene clusters located on chromosome 16 (encoding the α-like globin genes) and chromosome 11 (encoding the β-like globin genes). In each cluster these genes are arranged along the chromosome in the order in which they are expressed during development (Figure 1.3). Expression of these genes is controlled by several distant upstream *cis*-acting regulatory elements. In the human α-globin locus four such enhancers have been characterised and they are known as multispecies conserved sequences (MCS) R1, R2, R3 and R4 (discussed further in section 1.6.1). The five corresponding regulatory elements of the β-globin locus are collectively referred to as β-locus control region (β-LCR) (Higgs et al., 2012).

Thalassaemias are caused by genetic defects involving these globin genes which result in either absent or reduced production of globin proteins and haemoglobin. More precisely and commonly, defective synthesis of α-globin causes α-thalassaemia and that of β-globin results in β-thalassaemia (Piel and Weatherall, 2014).

1.3 β-Thalassaemia

1.3.1 Epidemiology and global burden

Thalassaemia is a global health problem and nearly seventy thousand children are born with severe forms of the disease each year (Weatherall, 2010). One third of these births are due to homozygous or compound heterozygous states of β-thalassaemia and a similar proportion is due to the most common subgroup of β-thalassaemia known as HbE β-thalassaemia. Although, the incidence rates over the past years have not shown a rising trend, thalassaemia presents an increasing health, financial and social burden to many countries in the world. This is a result of global improvements in hygiene, nutrition and public health and reductions in child mortality rates, which have led to the survival of infants with β-thalassaemia, that would previously not have been recognized (Weatherall, 2011).

β-Thalassaemia was originally confined to the tropical and subtropical regions of the world, distributed throughout the Mediterranean, sub-Saharan Africa, the Middle East and the southern regions of Asia. This is believed to be due to the

protection against *Plasmodium falciparum* malaria in heterozygous states of thalassaemia and the distribution of thalassaemia closely matches that of malarial infection (Haldane, 1949; Weatherall, 2008). However, due to relatively recent migration patterns, many north European and north American countries are now home to growing numbers of patients with thalassaemia (Michlitsch et al., 2009).

1.3.2 Molecular basis

Worldwide, nearly 300 mutations in and around the β-globin gene are reported to cause β-thalassaemia (Thein, 2013). Mutations that completely abolish the β-globin production are known as β^0-thalassaemia whereas mutations that result in reduced production of β-globin are termed β^+-thalassaemia. Most of the mutations causing β-thalassaemia are point mutations, either in the gene or its immediate flanking regions which impair transcription, processing or translation of β-globin messenger ribonucleic acid (mRNA) resulting in β^0 or β^+-thalassaemia (Thein, 2013). Several deletions of variable lengths removing all or part of the β-globin gene, commonly resulting in β^0-thalassaemia are also reported (Varawalla et al., 1991). Some rare β-thalassaemias are due to the deletions removing the upstream regulatory element β-LCR, leaving the globin genes intact (Driscoll et al., 1989).

A mutation associated with the structural haemoglobin variant HbE also results in a phenotype of β-thalassaemia (Fucharoen and Weatherall, 2012). This mutation is due to the substitution of guanine (G) by adenine (A) at codon 26 (GAG to AAG) in exon one of the β-globin gene which results in structurally

abnormal HbE consisting of $\alpha_2\beta^E_2$ globin chains. However, the abnormal sequence also activates a cryptic splice site that causes aberrant splicing which leads to a 16-nucleotide deletion of the 3′ end of the exon one creating a stop codon to reduce the correctly spliced β^E-globin mRNA giving rise to the phenotype of β-thalassaemia (Orkin et al., 1982). Inheritance of a β-thalassaemia allele from one parent and the structural variant haemoglobin E from the other parent results in HbE β-thalassaemia which accounts for nearly half of the patients with severe β-thalassaemia (Nienhuis and Nathan, 2012).

1.3.3 Cellular pathology

Severe anaemia due to the premature destruction of erythroid cells is the cardinal pathophysiological feature of β-thalassaemia. The erythroid cell destruction occurs at two different stages by distinct pathological processes (Rund and Rachmilewitz, 2005). Firstly, the circulating mature RBCs are destroyed predominantly in the spleen through extravascular haemolysis. Secondly, a large-scale destruction of RBC precursors occurs in the bone marrow by ineffective erythropoiesis, which is characterized by an expansion of early erythroid precursors, a maturation arrest and death of erythroid precursors at polychromatic stages limiting the number of RBCs released to the circulation (Rivella, 2009). The relative contributions of these two pathologic processes differ in various forms of thalassaemia (Rund and Rachmilewitz, 2005).

Unbalanced production of α- and β-globin chains within RBCs is the main causative factor which leads to haemolysis and ineffective erythropoiesis in thalassaemia (discussed below). During every stage of normal development,

productions of α-like and β-like globin chains are closely balanced to prevent an accumulation of free globin chains which are damaging to the RBCs. However, in patients with β-thalassaemia, impaired production of β-globin chains results in an excess of free α-globin chains which undergo denaturation and degradation in mature erythroid cells and their precursors to play a critical role in the causation of anaemia (Ribeil et al., 2013). The central role of free α-globin chains in the pathogenesis of β-thalassaemia was first described by Fesses and colleagues in 1960s. They demonstrated the presence of inclusion bodies in erythroid cells of patients with β-thalassaemia which were later confirmed to consist of precipitated α-globin chains (Figure 1.4) (Fessas, 1963; Fessas et al., 1966; Fessas et al., 1965).

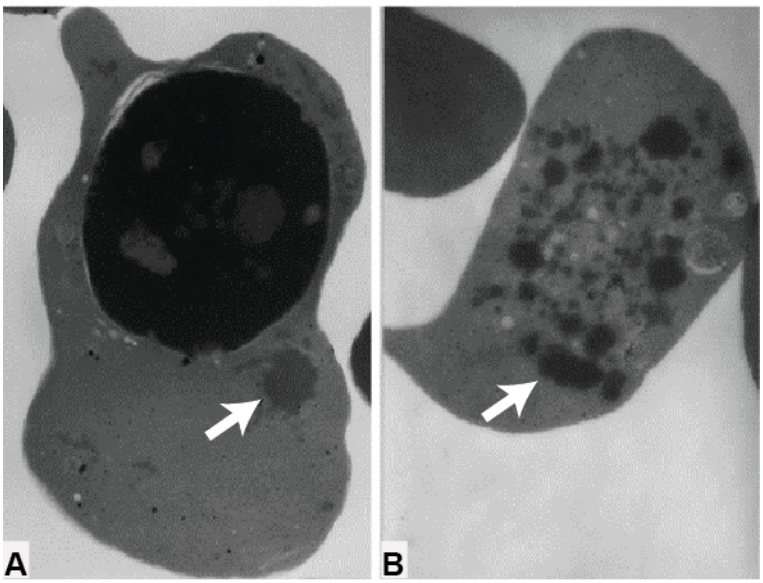


Figure 1.4 – Electron microscopy images showing electron-dense inclusions consist of precipitated α-globin chains in erythroid cells of patients with β-thalassaemia. (A) Erythroblast. (B) Erythrocyte. White arrows point to electron dense inclusions. Adapted from (Beris et al., 1999).

In normal human RBCs, the deleterious effects of small amounts of free α-globin chains are negated by the erythroid specific molecular chaperone, α-haemoglobin stabilizing protein (AHSP). AHSP specifically binds to α-globin and stabilizes free α-chains by promoting protein folding and resistance to protease digestion (Yu et al., 2007). However, in patients with β-thalassaemia, the levels of free α-globin chains are far in excess of the capacity of AHSP. Once this capacity is exceeded, highly unstable free α-globin molecules undergo auto-oxidation forming α-hemichromes (α-globin monomers which contain oxidized ferric iron) and reactive oxygen species (ROS) to trigger a cascade of events leading to haemolysis and ineffective erythropoiesis (Figure 1.5) (Shinar and Rachmilewitz, 1993; Voon and Vadolas, 2008).

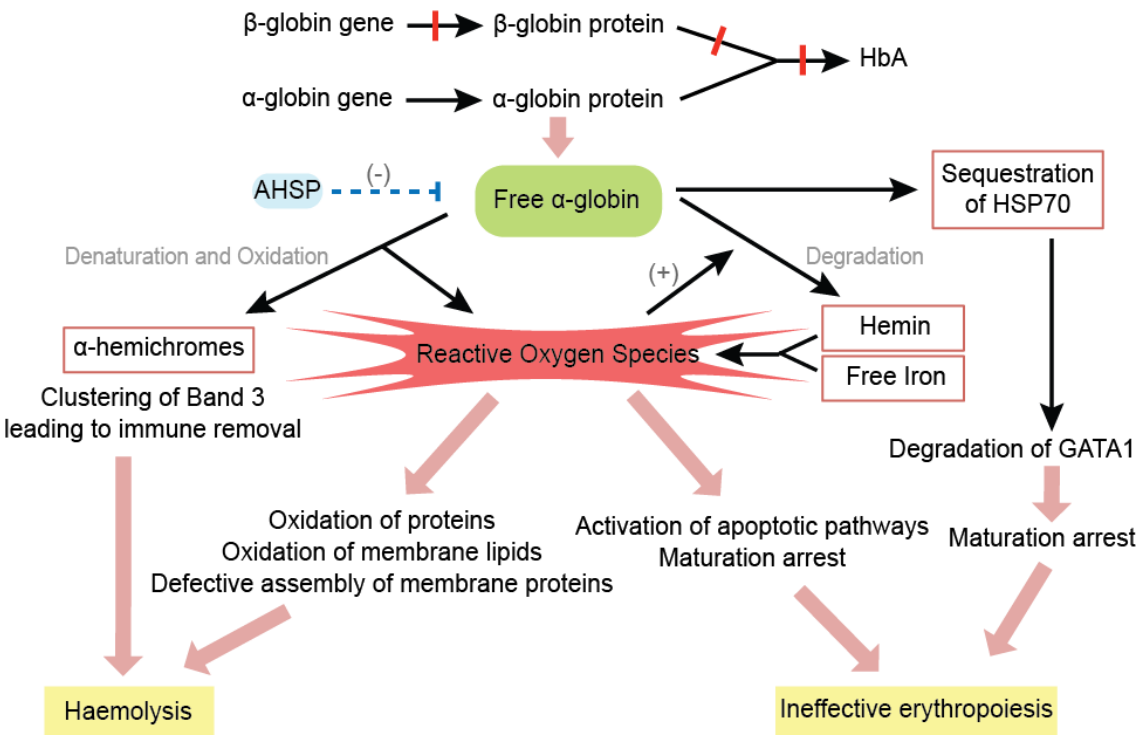


Figure 1.5 – Cellular pathology of β-thalassaemia. Absent or reduced β-globin production leads to an unbalanced excess of α-globin chains, which then triggers a cascade of events through the generation of reactive oxygen species resulting in haemolysis of mature red blood cells and destruction of immature erythroid precursors in the bone marrow (ineffective erythropoiesis). Adapted from (Mettananda et al., 2015).

The α-globin monomers are degraded via several pathways including an adenosine triphosphate (ATP) and ubiquitin dependent proteolytic pathway, an autophagy pathway and a non-enzymatic pathway triggered by ROS that results in release of hemin (heme containing oxidized ferric iron) and free iron which lodges in the cell membrane (Khandros et al., 2012; Nagababu et al., 2008; Nagababu and Rifkind, 2004; Shaeffer, 1988). In addition, α-hemichromes are also found bound to the cytoskeleton (Shinar and Rachmilewitz, 1993). In this unstable conformation, both the heme group and the iron are able to participate in redox reactions leading to further generation of ROS which damage cellular proteins, lipids and nucleic acids (Schrier, 1997). ROS induced alterations in membrane deformability and stability through partial oxidation of protein band 4.1 and defective assembly of spectrin-actin-band 4.1 membrane skeleton complex are believed to be the primary mechanism of haemolysis in mature RBCs in β-thalassaemia (Schrier, 2002). In addition, oxidant injury leads to clustering of band 3 which in turn produces a neoantigen that binds to immunoglobulin G and complement, thereby signalling macrophages to remove them from the circulation (Yuan et al., 1992).

Ineffective erythropoiesis in β-thalassaemia however, is not explained by membrane damage and appears mainly to be due to an enhanced activation of the apoptosis pathways (Ribeil et al., 2013). In non-erythroid cell models, ROS have been reported to induce apoptosis by activating apoptosis signal regulating kinase 1 and Jun-kinase and this mechanism may contribute to erythroid cell apoptosis in thalassaemia (Fujisawa et al., 2007). Furthermore,

there is evidence that apoptosis in thalassaemic RBCs is mediated by the Fas cell surface death receptor (FAS) and FAS-ligand pathway which is triggered by high levels of ROS. In addition, overexpression of growth differentiation factor (GDF) 11 in response to high ROS, results in terminal erythroid maturation arrest, contributing to ineffective erythropoiesis. The action of GDF11 is believed to be mediated via Activin receptor (ActR) IIA, which is a surface receptor found on the erythroid cell membrane (Dussiot et al., 2014). Additionally, a recent study has demonstrated that heat shock protein 70 (HSP70) interacts directly with free α-globin chains and thereby becomes sequestered in the cytoplasm. This, in turn prevents HSP70 from performing its normal physiological role of protecting GATA-binding factor 1 (GATA1) from proteolytic cleavage. The resulting premature degradation of GATA1 results in maturation arrest and apoptosis of polychromatic erythroblasts (Arlet et al., 2014). In summary, the existing evidence clearly suggests that the excess free α-globin chains are directly responsible for haemolysis and ineffective erythropoiesis, which are the two primary pathophysiological mechanisms that cause anaemia in patients with β-thalassaemia.

1.3.4 Pathophysiology

Anaemia reduces the oxygen carrying capacity of blood which in turn leads to tissue hypoxia. Hypoxia increases the secretion of erythropoietin from renal erythropoietin-producing cells through transcriptional induction mediated via hypoxia inducible factors (HIF), a group of transcription factors which orchestrates the response to hypoxia (Souma et al., 2015). Erythropoietin is a glycoprotein hormone that induces erythropoiesis. Although the aim of

stimulation of erythropoiesis is to increase the RBC mass in the circulation, in β-thalassaemia, it becomes counterproductive, worsens ineffective erythropoiesis and triggers a cascade of events which leads to body iron overload (Figure 1.6).

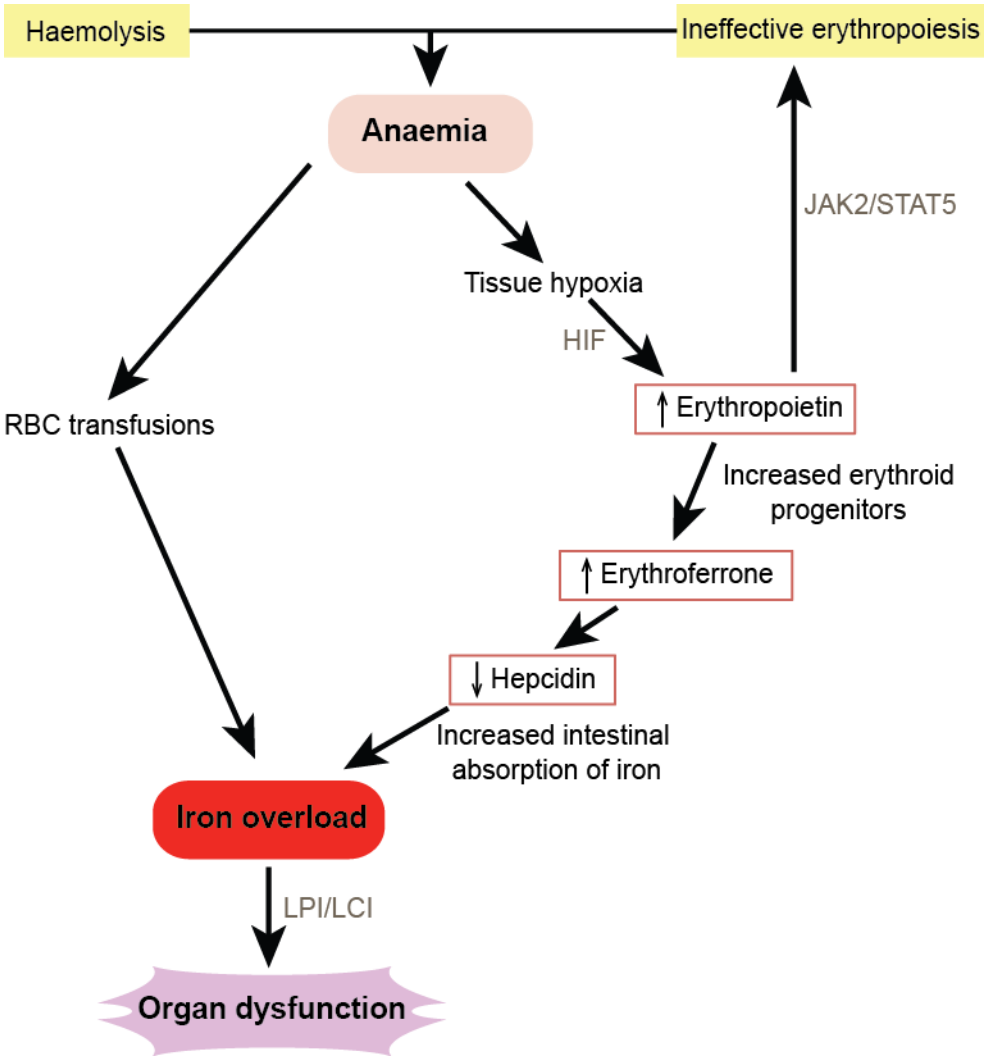


Figure 1.6 – Pathophysiology of β-thalassaemia. Increased secretion of erythropoietin in response to anaemia leads to further worsening of ineffective erythropoiesis and triggers a cascade of events which leads to body iron overload. Regular RBC transfusions also contribute to iron overload. Iron overload through accumulation of toxic iron metabolites (labile plasma iron [LPI] and labile cellular iron [LCI]) consequently leads to tissue damage and organ dysfunction.

Ineffective erythropoiesis is directly linked to high erythropoietin levels. Binding of erythropoietin to its receptor in erythroid progenitor cells activates Janus kinase 2 (JAK2) which in turn activates Signal transducer and activator of

transcription 5 (STAT5), which turns on cell proliferation and inhibits erythroid differentiation, contributing to ineffective erythropoiesis (Rivella, 2015; Socolovsky et al., 2001). Ineffective erythropoiesis then leads to hyperplasia and expansion of the bone marrow as well as proliferation of extramedullary haematopoietic sites giving rise to clinical features of β-thalassaemia.

Stimulation of erythropoiesis also suppresses hepcidin, the master regulator of iron homeostasis. This inhibition is believed to be mediated via erythroferrone, which is a member of the tumour necrosis factor-related protein family. In murine models, erythroferrone is secreted from erythroblasts in the bone marrow and the spleen and is responsible for hepcidin suppression during erythropoietic activity stimulated by endogenous or exogenous erythropoietin (Kautz et al., 2014). Furthermore, erythroferrone has shown to be responsible for the iron overload in a mouse model of β-thalassaemia through suppression of hepcidin (Kautz et al., 2015). GDF15 which is released by erythroid precursors has also been implicated in the down-regulation of hepcidin (Tanno et al., 2007). Hepcidin is a peptide hormone synthesised in the liver which inhibits the flow of iron from enterocytes and macrophages into the plasma by binding to ferroportin (the only known cellular iron exporter) thereby causing its internalization and degradation by the ubiquitin pathway (Coates, 2014). Suppression of hepcidin in β-thalassaemia thereby, leads to an increased iron absorption and body iron overload.

Most patients with severe β-thalassaemia receive regular RBC transfusions (discussed further in section 1.4.1) throughout life. One unit of packed RBC

contains ~200mg of iron (Rachmilewitz and Giardina, 2011) and because humans have no mechanisms for the excretion of iron this leads to a progressive transfusional iron accumulation (Coates, 2014). The resultant iron overload leads to tissue damage and organ dysfunction and is the main cause of morbidity and mortality related to thalassaemia (Weatherall and Clegg, 2001).

Transferrin which binds to two molecules of ferric iron is the main iron transporter protein in the body. About 20–30% of transferrin is normally bound to iron and this transferrin-bound iron is the primary source of iron available to cells. Non-transferrin-bound iron (NTBI) refers to a heterogeneous group of potentially toxic iron complexes which can be detected in plasma only when the transferrin saturation increases beyond 35%. A fraction of NTBI which is known as labile plasma iron (LPI) is loosely bound to proteins, highly redox active, and thought to be the main species that causes iron-mediated oxidative damage. In the presence of iron overload, once transferrin becomes saturated, NTBI and LPI levels rise significantly and can easily enter many cell types, resulting in increased levels of labile cellular iron (LCI) which are highly reactive and causes organ damage and failure (Coates, 2014).

1.3.5 Clinical features and complications

β-Thalassaemia has a wide spectrum of clinical severity. Individuals with the least severe phenotype, β-thalassaemia trait (commonly due to heterozygous β-thalassaemia mutations) are asymptomatic and have minimal haematological abnormalities with mild anaemia. The most severe phenotype, β-thalassaemia major (commonly due to homozygous or compound heterozygous states of β-

thalassaemia mutations) is characterised by transfusion dependent hypochromic-microcytic anaemia beginning in the first two years of life (discussed below). The onset of anaemia correlates with the γ- to β-globin and foetal to adult haemoglobin switch that occurs in the first few months of life. The intermediate phenotype is termed β-thalassaemia intermedia and is characterised by moderate hypochromic-microcytic anaemia requiring infrequent or occasional RBC transfusions (Galanello and Origa, 2010).

Early clinical features of β-thalassaemia major are due to severe anaemia, haemolysis, bone marrow hyperplasia and extramedullary haematopoiesis. Clinical manifestations of anaemia include pallor, poor feeding, irritability and growth retardation. Jaundice is the cardinal sign of haemolysis. Bone marrow hyperplasia is clinically manifested as skeletal abnormalities including bossing of the skull, hypertrophy of the maxilla and prominent malar eminences giving rise to characteristic ‘thalassaemic facies’. Extramedullary haematopoiesis mainly occur in the liver and the spleen which results in variable degrees of hepatomegaly and splenomegaly. Since the spleen is the predominant site of haemolysis, enlargement of the spleen is associated with an increased destruction of RBCs leading to worsening of anaemia (Rachmilewitz and Giardina, 2011).

From late childhood and the second decade of life, patients with β-thalassaemia are at risk of developing severe complications related to iron overload. Complications of iron overload include growth retardation, failure of sexual maturation and dysfunction of vital organs including: heart (dilated

cardiomyopathy and pericarditis); liver (chronic hepatitis, fibrosis, and cirrhosis); pancreas (insulin-dependent diabetes mellitus) and other endocrine glands (insufficiency of the parathyroid, thyroid, pituitary, and, less commonly, adrenal glands). Other complications of β-thalassaemia include venous thrombosis, osteoporosis and pulmonary hypertension (Cao and Galanello, 2010).

Organ dysfunction leads to organ failure which is refractory to medical management and results in an extremely poor quality of life and premature death. Myocardial disease is the most serious life-limiting complication and approximately 50-60% of patients die from cardiac failure (Borgna-Pignatti et al., 2004). Consequently, many patients with thalassaemia in underdeveloped nations die in childhood or adolescence (Rund and Rachmilewitz, 2005). Even in developed countries where the management is optimal, patients with β-thalassaemia consistently report lower quality of life (Amid et al., 2015).

1.3.6 Genetic modifiers of disease severity

β-Thalassaemia is the archetypal monogenic disorder and, as such, has established the principle that such disorders display remarkable clinical heterogeneity. This can be partly explained by the nature of the mutations. β^0-thalassaemia mutations which cause complete abolition of β-globin gene expression give rise to more severe forms of the disease, whereas β^+-thalassaemia with variable reduction of β-globin chain synthesis produces a less severe phenotype (Danjou et al., 2011). However, there is also remarkable clinical variability even in patients who inherit identical β-globin genotypes.

As discussed before, the basic pathology in β-thalassaemia is due to the accumulation of un-paired α-globin chains in erythroid cells hence, clinical severity is expected to be directly related to the degree of imbalance between α-like and β-like globin chain synthesis. Any condition which decreases the α-globin production or increases β-like globin synthesis therefore, should ameliorate the severity of anaemia. In fact, two naturally occurring phenomena which improve the α-like/β-like globin chain balance have been recognised as disease modifiers of β-thalassaemia.

Firstly, the co-inheritance of genetic determinants which increase the production of γ-globin chains and HbF increases the β-like globin chain pool in erythroid cells, improves the α-like/β-like globin chain balance and reduces the severity of β-thalassaemia. Genetic polymorphisms identified in quantitative trait loci (QTL) including the Xmn1-HBG2 variant in the β-globin cluster, the HBS1L-MYB intergenic region on chromosome 6 and the gene encoding B-cell lymphoma/leukaemia 11A (BCL11A) protein on chromosome 2 (Menzel et al., 2007; Sankaran et al., 2008; Uda et al., 2008) as well as the Kruppel-like factor 1 (KLF1) mutations (Borg et al., 2010) are all known to be associated with high HbF levels and milder forms of β-thalassaemia (Galarneau et al., 2010; Liu et al., 2014; Thein, 2008).

Notably, co-inheritance of factors which lower α-globin production, reduces the free intracellular α-globin pool and improves α-like/β-like globin chain balance and has been postulated to ameliorate the severity of β-thalassaemia. Indeed, a plethora of high-quality clinical studies performed over several decades

(discussed below) have highlighted how a natural reduction of α-globin output in the form of α-thalassaemia improves the disease severity in patients with β-thalassaemia. This remarkable natural phenomenon forms the corner stone of the underlying hypothesis related to the work described in this thesis.

1.4 Treatment of β-thalasaemia

1.4.1 Supportive therapy

Despite being one of the first molecular diseases to be identified and the pathophysiology understood, the management of β-thalassaemia still largely depends on supportive care with RBC transfusions and iron chelation. Patients with severe forms of β-thalassaemia require regular RBC transfusions (usually monthly) to maintain their haemoglobin at safe levels to enable normal growth and development, inhibit bone marrow expansion and suppress clinically significant extramedullary haematopoiesis (Trompeter and Cohen, 2014). However, regular transfusions carry significant risks and lead to a number of complications. Although, the risk of transfusion-transmitted infections has been greatly reduced by implementation of routine screening of transfusion donors, transmission of HIV, hepatitis and rare pathogens is still possible due to flaws and limitations in screening strategies (Rachmilewitz and Giardina, 2011; Shyamala, 2014). Transfusions also lead to various types of reaction including febrile (frequency 1/100), allergic (frequency 1/100), haemolytic (frequency 1/25,000) and anaphylactic (frequency 1/50,000) which can sometimes be life-threatening (Trompeter and Cohen, 2014). Moreover, in many regions of the

world, RBC donor availability limits regular transfusion therapy (Sankaran and Weiss, 2015).

Transfusional iron overload, which leads to multiple organ damage (discussed in section 1.3.4), remains the most challenging complication associated with RBC transfusions. Iron overload is currently managed by iron chelator medications – deferoxamine, deferiprone and deferasirox – which are used in monotherapy or in combination (Porter and Viprakasit, 2014). However, none of these are effective and safe enough to restore normal iron homeostasis. Deferoxamine is still the most efficacious iron chelator but is only available in a parenteral form and requires administration for prolonged periods. Deferiprone is less effective and has serious side effects including agranulocytosis and arthropathy. A new oral iron chelator, deferasirox, is believed to be as efficacious as deferoxamine but its high cost limits its use in resource poor countries (Aydinok et al., 2014). Therefore, in summary, the currently available supportive therapies do not have the potential to transform the natural course of the disease and this leads to multiple complications and significant morbidity and mortality in the majority of patients with β-thalassaemia.

1.4.2 Curative therapy

Allogeneic haematopoietic stem cell transplantation (HSCT) remains the only curative treatment option available for β-thalassaemia (Angelucci et al., 2014b). The best results have been obtained with Human leukocyte antigen (HLA) - identical sibling HSCTs and outcomes today have substantially improved compared to past three decades. More than 90% of patients now survive HSCT

and disease free survival is around 80% (Angelucci et al., 2014a). However, a lack of compatible sibling donors limits the usefulness HSCT in most of the patients and so far only about 3000 transplants have been performed over a 30-year period, worldwide (Angelucci, 2010). Although there is an interest and limited experience in using haematopoietic stem cells (HSC) from alternative donors such as matched unrelated donors, matched unrelated cord blood and mismatched related donors, poor donor selection can lead to a number of complications and increased risk of mortality and morbidity (Angelucci et al., 2014a; Rachmilewitz and Giardina, 2011).

Transplant related morbidity and mortality is not uncommon even with HLA-identical sibling HSCTs. Graft versus host disease (GVHD) is the most common and clinically significant complication following transplantation. A recent large-scale, international study reported incidences of acute and chronic GVHD to be 20% and 12% respectively, following HLA-identical sibling HSCTs (Locatelli et al., 2013). Graft failure (frequency 7%) and severe life threatening infections are the other serious complications of HSCT (Lucarelli et al., 2012).

In addition, substantial financial and health care resources are necessary for successful HSCT. Transplant recipients require intensive care support and careful follow-up especially during the immediate post-transplant period, which are not available in most developing countries. Similarly, the cost associated with a HSCT is extremely high and is reported to be around US$ 112,000 – 150,000 in developed countries (Angelucci et al., 2014b). Although, the cost of transplantation can vary significantly around the world, and is reported to be

around US$ 20,000 in India (Chandy, 2008; Sharma et al., 2014), HSCT is not a feasible option in most of the patients in resource-poor countries where the disease burden is greatest.

1.4.3 Experimental therapies

Over the past 30 years many new and experimental approaches to the treatment of β-thalassaemia have been considered and developed; however none has progressed to the level of significant clinical use. Most of the emphasis has been on induction of foetal haemoglobin by derepressing the γ-globin, a β-like globin gene which is expressed during the foetal life. In addition, several new strategies and pathways are being considered which target various stages in the pathogenesis of β-thalassaemia (Figure 1.7)

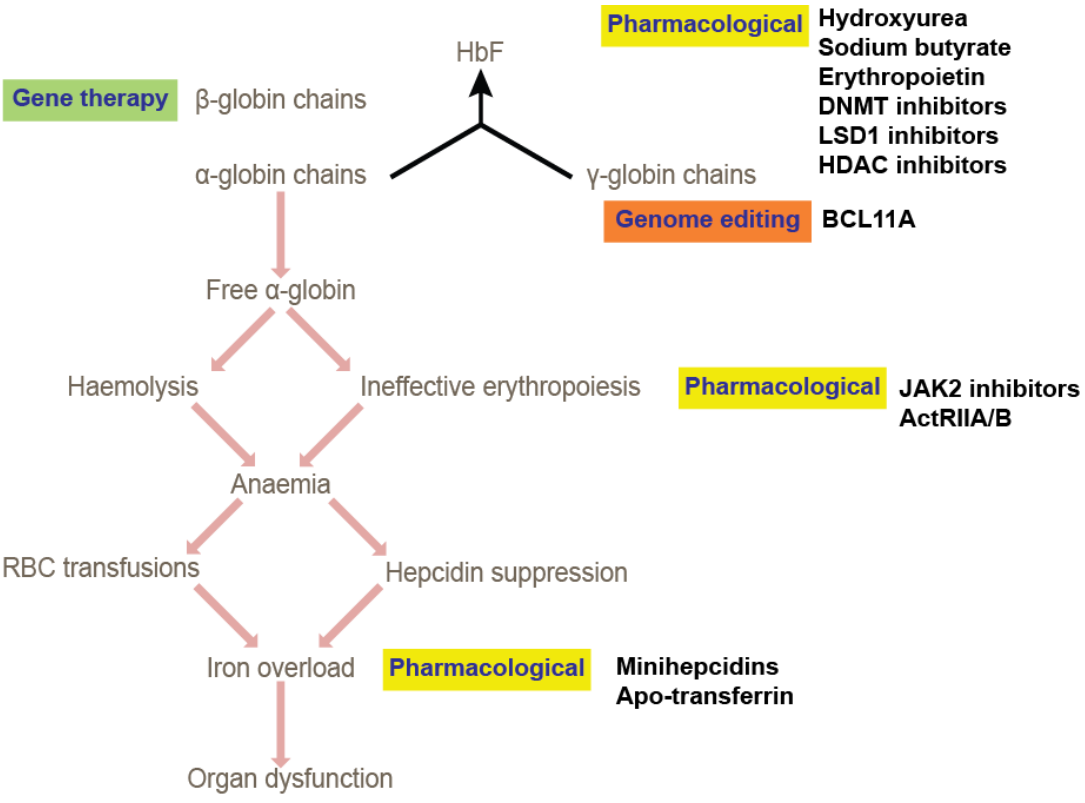


Figure 1.7 – Summary of experimental therapies for β-thalassaemia. The therapies which are currently being developed are shown with the corresponding target site in relation to the pathophysiology of the disease.

1.4.3.1 Pharmacological *strategies which aim to induce foetal haemoglobin*

Pharmacological augmentation of foetal haemoglobin (HbF) production has been a long-standing therapeutic objective and is already in clinical use for patients with severe sickle cell disease (Thein, 2012). The rationale for this strategy in β-thalassaemia is to increase the production of γ-globin chains, which would then mop-up the excess of α-globin protein in order to reduce the pool of free α-globin chains, which are detrimental to the erythroid cells. A few different classes of foetal haemoglobin inducing agents have already been tested in patients with β-thalassaemia. The chemotherapeutic agent hydroxyurea is one of the very first and most widely used compounds shown to induce γ-globin (Bradai et al., 2003; Fibach et al., 1993; Platt et al., 1984). In addition, DNA methyltransferase inhibitors, 5-azacytidine and decitabine (Ley et al., 1982; Olivieri et al., 2011), a short chain fatty acid derivative, sodium phenylbutyrate (Perrine et al., 1993) and human recombinant erythropoietin (Bourantas et al., 1997; Rachmilewitz and Aker, 1998) have been tested in humans. However, none of these has been able to produce a sustainable clinical response in patients with β-thalassaemia (Musallam et al., 2013b).

Several novel approaches to promote HbF production have emerged since the identification of QTLs and the epigenetic mechanisms responsible for haemoglobin switching and γ-globin suppression (Sankaran and Weiss, 2015). The target pathways amenable to pharmacological manipulation currently being considered include, inhibition of epigenetic enzymes: histone demethylase (KDM) 1A (also known as lysine specific demethylase 1[LSD1]), histone

deacetylase (HDAC) 1 and HDAC2. The already licenced, anti-depressant drug tranylcypromine has been shown to induce HbF *in vitro* in primary human erythroid cells by inhibition of LSD1 but concerns have been raised as this pathway retards normal erythropoiesis (Shi et al., 2013). HDACs are ubiquitous enzymes which play a wide variety of roles in a number of tissues and hence, the effects of its inhibition may not be trivial. Therefore, in summary, approaches for HbF induction have not been straight forward and despite a lot of effort, none have made significant progress towards clinical use in patients with β-thalassaemia.

1.4.3.2 Pharmacological strategies which aim to decrease ineffective erythropoiesis

As ineffective erythropoiesis plays a major role in the pathogenesis of anaemia in β-thalassaemia, a number of strategies targeting mediators of ineffective erythropoiesis (discussed in section 1.3.3 and 1.3.4) have been reported. One such strategy is to inhibit JAK2. Although, chemical inhibition of JAK2 has been shown to reduce splenomegaly in β-thalassaemia mice (Libani et al., 2008), the clinical trials with JAK2 inhibitors in myeloproliferative disorders have reported several side effects including severe thrombocytopenia and anaemia (Mesa and Cortes, 2013). This suggests that in patients with β-thalassaemia, treatment of JAK2 might be counter-productive and worsen the severity of anaemia.

Another potential strategy is to block the action of GDF11 by using a ligand to its receptor, ActRIIA. This pathway was uncovered following an incidental

finding of a phase I clinical trial which tested sotatercept- a ligand trap for the ActRIIA- for osteoporosis. Interestingly and unexpectedly, in this trial sotatercept dramatically increased the haemoglobin and RBC counts in healthy human subjects (Sherman et al., 2013). In subsequent studies, sotatercept and a ligand trap of ActRIIB (ACE-536) showed significant improvements in anaemia in β-thalassaemia mouse models (Dussiot et al., 2014; Suragani et al., 2014). However, these findings were predominantly from animal models and further human studies are necessary for a better understanding of the roles of JAK2 and GDF11 in the pathophysiology of β-thalassaemia.

1.4.3.3 Pharmacological strategies which aim to decrease iron overload

As stated before, suppression of hepcidin is one important mechanism leading to iron overload in β-thalassaemia. Hepcidin analogues minihepcidins have been shown to reduce the iron overload in animal models (Ramos et al., 2012; Schmidt and Fleming, 2014). In another study in β-thalassaemia mice, injections of apo-transferrin resulted in improvements in the thalassaemia phenotype and specifically demonstrated normalization of LPI concentrations, increase in hepcidin expression, normalization of RBC survival and increase in the haemoglobin production (Li et al., 2010). However both these strategies aim to improve iron overload and might not be helpful in reducing the transfusion requirements or improving anaemia in patients with β-thalassaemia.

1.4.3.4 Gene therapy

Gene therapy offers an alternative approach for cure in patients with β-thalassaemia. The aim of gene therapy is to transfer a normal β-globin gene to the HSCs of patients. This approach involves harvest of autologous HSCs from patients, *ex vivo* gene transduction and infusion of gene-transduced cells back to the patient. Although the initial progress of the field was much slower than expected, the first successful gene therapy for β-thalassaemia was carried out in 2007 (Cavazzana-Calvo et al., 2010). Since then, a limited number of patients have undergone gene therapy and clinical trials are in progress (Cavazzana-Calvo et al., 2010; Dong et al., 2013; Finotti et al., 2015; Thompson et al., 2014).

However, a number of challenges are still outstanding. The main concerns include improvement of the efficiency of gene delivery and variable expression and non-oncogenic insertion of the β-globin gene. Furthermore, the reported risks associated with the procedure which include development of leukaemia by activation of proto-oncogenes means that much further work is required before gene therapy can be used as a common treatment for β-thalassaemia (Persons, 2010).

1.4.3.5 Genome editing strategies

Recapitulation of genetic mutations that result in foetal haemoglobin induction (discussed in section 1.3.6) using genome editing technologies (discussed in detail in section 1.7.3) has been an attractive area for therapeutic development.

The target genes currently being considered are the erythroid transcription factors (TFs) *BCL11A, MYB and KLF1*. However, in general, suppression of TFs is considered challenging because they are difficult to target by small molecules (Sankaran, 2011). In addition, *BCL11A* has essential functions in non-erythroid cells such as lymphocytes and neurons. Although erythroid specific, *MYB* has pleiotropic roles in haematopoiesis and manipulation of this may also be toxic. Similarly, manipulating *KLF1* could give rise to variable phenotypes as some of the *KLF1* mutations also cause anaemia (Borg et al., 2011). Therefore, at present, manipulating these genes to obtain an erythroid specific phenotype seems a distant reality.

In addition to the above attempts there has been a report of activation of δ-globin gene expression in an animal model to enhance the production of HbA_2 (Manchinu et al., 2014). However, the clinical usefulness this approach needs further consideration.

1.4.4 Limitations of the current and experimental therapies

The information presented in the preceding sections, clearly demonstrates that the current treatment options for β-thalassaemia are sub-optimal and do not alter the natural course of the disease in a majority of patients. Except for gene therapy, all other experimental therapies are aimed at managing complications or are indirect approaches for altering the α/β globin chain balance. None of the approaches explore the possibility of directly reducing the excess α-globin chains, which is central in the pathophysiology of the disease. Hence, it is clear that there is still a great need for new, alternative and effective therapeutic

strategies for the treatment of this life limiting disease, to render patients transfusion independent and live a normal life.

1.5 Reduction of α-globin as a therapeutic option for β-thalassaemia

As stated before, in patients with β-thalassaemia, the primary damage to the RBCs and their precursors is mediated via excess α-globin chains which accumulate when the β-globin expression is reduced (Nathan and Gunn, 1966; Weatherall and Clegg, 2001). Therefore, we hypothesise that if a reduction of α-globin chains can be achieved therapeutically, it should reduce the α-globin excess in erythroid cells, improve the α/β globin chain imbalance and lessen the clinical severity of β-thalassaemia. In fact, a number of case-control and cohort studies (discussed below) have demonstrated that a natural reduction in α-globin chain output, resulting from co-inherited α-thalassaemia, is beneficial in patients with β-thalassaemia. Therefore it is likely that a therapeutic reduction of α-globin gene expression is a potential and promising pathway to develop new therapies for β-thalassaemia.

1.5.1 Co-inheritance of α-and β-thalassaemia

Since the initial descriptions by Fessas (Fessas et al., 1961) and Kan and Nathan (Kan and Nathan, 1970), co-inheritance of α-thalassaemia has been predicted to ameliorate the severity of β-thalassaemia. Co-inheritance of α- and β-thalassaemia is common in areas where the gene frequencies of both types

of thalassaemia are high, especially in regions of the world where heterozygous thalassaemia provides a natural selective advantage in the presence of endemic *Plasmodium falciparum* malaria (Weatherall, 2012). Results from a number of family, cohort and case-control studies in various ethnic groups which explored the outcomes of the co-inheritance of α- and β-thalassaemia can be summarized as follows: deletion of two α-globin genes (--/αα or –α/-α) is associated with a milder clinical phenotype in most patients with β-thalassaemia while deletion of a single α-globin gene (-α/αα) is beneficial in patients other than those with the most severe reduction in β-globin synthesis (β^0/β^0 genotype) (Camaschella et al., 1995; Galanello et al., 1989; Gringras et al., 1994; Ho et al., 1998; Thein et al., 1988; Wainscoat et al., 1983; Winichagoon et al., 2000). However, a recent study performed as a follow up of a genome-wide association study among a relatively uniform group of patients with the β^0/β^0 genotype revealed that even a single α-globin gene deletion significantly increases the chance of having a thalassaemia intermedia phenotype as opposed to thalassaemia major (Galanello et al., 2009). Furthermore, this study also showed that the presence of a mutated α-globin allele provided an additive effect to other modifier alleles responsible for modulating high HbF levels. Similar results were found in another study which used multivariate analysis and concluded that α-thalassaemia is a significant independent predictor for a milder phenotype (Badens et al., 2011). Co-inheritance of three α-globin gene deletion with β-thalassaemia is extremely rare. However in a case report which described this rare co-inheritance, the patient had a completely balanced α/non-α globin chain biosynthesis ratio, was

clinically normal and was not diagnosed until 50 years of age (Kanavakis et al., 2004).

Of all β-thalassaemia subtypes identified throughout the world, the most common genotypes are those associated with HbE β-thalassaemia. This disease, which is seen commonly throughout parts of South and Southeast Asia constitutes about 50% of births of severe β-thalassaemia. Importantly, co-inheritance of α-thalassaemia appears to have its most pronounced and unequivocal effect on patients with HbE β-thalassaemia with substantial clinical benefit even when only a single α-globin gene deletion (-α/αα) is co-inherited. In large scale prospective studies, all patients with co-inherited α-thalassaemia displayed a milder phenotype, older age at presentation, smaller splenic and hepatic sizes, normal physical and sexual maturation and significantly reduced transfusion requirements (Bandyopadhyay et al., 2001; Premawardhena et al., 2005a; Sharma and Saxena, 2009; Sripichai et al., 2008; Suresh et al., 2013; Winichagoon et al., 2000). Furthermore, in HbE β-thalassaemia patient cohorts, the frequency of α-thalassaemia is significantly lower than in the normal matched population and it is extremely rare to find patients with two α-globin gene deletions (--/αα or –α/-α) (Sripichai et al., 2008; Suresh et al., 2013). These observations suggest that individuals with HbE β-thalassaemia, who co-inherit only two, rather than the normal four, α-globin genes have a very mild phenotype and rarely come to medical attention.

1.5.2 Effect of excess α-globin chains on β-thalassaemia

Contrary to the beneficial effect demonstrated by co-inheritance of α and β-thalassaemia, inheritance of excess α-globin genes in the presence of β-thalassaemia worsens the disease severity (Origa et al., 2013; Premawardhena et al., 2005b; Sollaino et al., 2009). Patients with heterozygous β-thalassaemia mutations, who would otherwise have a β-thalassaemia minor phenotype, develop a thalassaemia intermedia phenotype when they inherit more than four functional copies of the α-globin gene. This is because the inheritance of excess α-globin genes in these patients increases the amount of free α-globin within RBCs and worsens the α-like/β-like globin chain imbalance. Increase in free α-globin chains exacerbates the premature erythroid cell destruction and worsens the severity of anaemia. This observation further emphasizes that the number of functional α-globin genes has clear and direct effects on the clinical severity of patients with β-thalassaemia.

1.5.3 Evidence from mouse models

Amelioration of β-thalassaemia by reducing α-globin has also been recapitulated in the β-globin knockout (β-KO) mouse model. Voon and colleagues generated a double heterozygous mouse knocked-out for both α- and β-globin genes (α++/−−; β+/−) and compared the haematological parameters to those of the β-KO mice (α++/++; β+/−) (Voon et al., 2007). Interestingly, all the observed parameters including haemoglobin level, haematocrit, red cell distribution width, reticulocyte number and the percentage of erythroid progenitors in bone marrow and spleen were corrected to near

normal values in the double heterozygous mouse. This observation further emphasises that the co-inheritance of α-thalassaemia might have profound improvements to the β-thalassaemia phenotype and provides an experimental proof that α-globin is a suitable target to develop therapies.

1.5.4 Desired level of reduction of α-globin

The clinical and genetic evidence discussed in the preceding sections strongly supports the idea that reducing α-globin expression in patients with β-thalassaemia is an important area for developing new therapies. Reducing α-globin output to levels comparable to a single or two α-globin gene deletion should be clinically beneficial to patients with β-thalassaemia, and in patients with HbE β-thalassaemia it could be transformational. Among patients with α-thalassaemia, single and two α-globin gene deletions correspond to 75% and 50% α-globin levels respectively compared to normal controls at both the level of mRNA and protein (Chaisue et al., 2007; Harteveld and Higgs, 2010). Therefore, these levels could be considered as rational therapeutic aims in β-thalassaemia.

Clearly, several obstacles need to be considered if this natural phenomenon is to be translated effectively into therapy. Firstly, the α-globin loss should not be complete and any decrease beyond 25% of the normal level (levels seen in patients with three α-globin gene deletions) would be counterproductive and worsen anaemia with loss of sufficient α-globin chains to produce haemoglobin. Secondly, the production of β-like globins should not be affected. Advances in our understanding of the genetic and epigenetic regulation of the α-globin gene

cluster suggest that it might be possible to selectively control α-globin expression to the appropriate degree to be useful for patients with β-thalassaemia.

1.6 Regulation of α-globin

1.6.1 Distal enhancers of the α-globin gene

Transcriptional enhancers play a central role in establishing the cell type-specific expression patterns of most human genes. Enhancers are functionally defined as DNA sequences that have the potential to enhance basal transcription levels from gene promoters and transcriptional start sites (Heinz et al., 2015). Enhancers are often localized large distances from the genes they regulate ranging from hundreds of bases to megabases upstream or downstream. They can be intragenic or intergenic, or even in nonrelated genes, and they do not necessarily regulate transcription of their nearest gene (van den Heuvel et al., 2015). The identification of DNA sequences which act as potential enhancers is commonly done using specific histone methylation signatures marks. In general, enhancers display enrichment of histone H3 lysine 4 monomethylation (H3K4me1) and depletion of H3K4me3 compared with promoters (Heintzman et al., 2007).

The human α-globin gene cluster is located on the short arm of chromosome 16 (16p13.3) very close (~150kb) to the telomere (Figure 1.3). The *cis*-acting regulatory network of this gene cluster has been well characterised and it is

now known that the expression of the α-globin genes is controlled by distant enhancers located 30-70kb upstream of the genes. Distal enhancer regulators of α-globin first emerged in the early 1990s, following a report of a family with α-thalassaemia due to a large (62kb) deletion upstream of the α-globin locus, but leaving the α-globin genes intact (Hatton et al., 1990; Higgs et al., 1990). A more recent comparative sequence analysis of 22 species identified a conserved syntenic region of ~135kb containing the α-globin cluster and all of the *cis*-acting regulatory elements including MCSs required for full tissue- and developmental-stage specific expression of α-globin (Hughes et al., 2005). Mapping the positions of DNAse1 hypersensitive sites to the MCSs confirmed that there are four erythroid specific distant enhancers situated 10 (MCS-R4), 33 (MCS-R3), 40 (MCS-R2) and 48 (MCS-R1) kilo bases upstream of the gene locus enhancing human α-globin expression (Figure 1.8) (Higgs and Wood, 2008; Hughes et al., 2005).

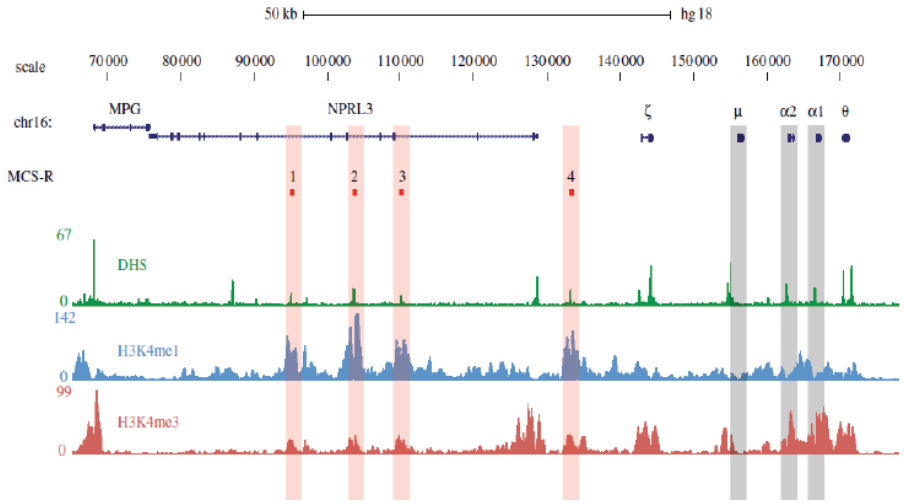


Figure 1.8 – *Cis*-acting enhancers of the α-globin gene cluster. A 100kb region of the human chromosome 16 from 70–170 kb of genome build hg18. RefSeq genes are shown as blue lines for gene extent, and blue ticks for exons. The MCS-R track shows the MCSs associated with the α-globin regulatory elements (MCS-R1–4), and their alignments to DNase-seq and ChIP-seq features are indicated with vertical red bars. Below this are plots of DNase-seq (green) and chromatin ChIP-seq (H3K4me1 in light blue, H3K4me3 in red) data for the K562 cell line from the ENCODE project on genome build hg18. The position of HBM gene (μ) and HBA genes (α1 and α2) are indicated by vertical grey bars. Adapted from (Hughes et al., 2013).

The four upstream enhancers vary in their capacity to enhance α-globin expression in human erythroid cells. A variety of experiments including transient transfections and transgenics combined with naturally occurring human mutations (discussed further in chapter 7) have suggested the MCS-R2 enhancer which lies 40kb upstream to the genes is the most critical regulatory element and is capable of enhancing α-globin expression on its own (Higgs and Wood, 2008).

1.6.2 Epigenetic regulation and histone modifications of α-globin gene

The term epigenetics refers to heritable changes in gene expression that do not involve changes to the underlying DNA sequence. Epigenetics explains how cells in a multicellular organism have identical DNA sequences (and therefore the same genetic instruction sets), yet maintain different terminal phenotypes (Riddihough and Zahn, 2010). Epigenetic signals which generally include DNA methylation and posttranslational modifications of amino acids on the amino-terminal tail of histones are responsible for the establishment, maintenance, and reversal of transcriptional states (Bonasio et al., 2010).

Chromatin environment in and around the human α-globin gene cluster is well characterised. The α-globin cluster is located on chromosome 16 in a gene dense, early replicating and open chromatin area of the genome surrounded by a number of housekeeping genes which are expressed in all cell types (Higgs et al., 2005). DNA at this locus is unmethylated and the promoters of all α-like globin genes are associated with unmethylated cytosine-guanine dinucleotide (CpG) rich islands. Because the α-globin is in an open chromatin environment,

the gene is actively silenced through a transcriptional repressive complex known as polycomb repressive complex (PRC) 2 in non-erythroid cells in which α-globin is not expressed (Garrick et al., 2008).

Human PRC2 is a protein complex which comprises a catalytic enhancer of zeste homolog 2 (EZH2) protein and the following core subunits: embryonic ectoderm development (EED), suppressor of zeste 12 protein homolog (SUZ12) and retinoblastoma binding protein 4 (RBBP4) (Klose et al., 2013). The transcriptional repressor action of PRC2 is mediated through trimethylation of histone H3 lysine 27 (H3K27me3) (Simon and Kingston, 2013). The mechanisms involved in the binding of PRC2 complex to the target genomic sites is not fully understood but is believed to be associated with the unmethylated CpG island promoters (Lynch et al., 2012).

In non-erythroid cells, PRC2 is recruited to the α-globin promoter and thereby through increasing the H3K27me3 chromatin modification signals transcriptional silencing (Figure 1.9) (Garrick et al., 2008). In erythroid cells, when the α-globin gene is activated, PRC2 is detached and the H3K27me3 chromatin mark is erased through both passive and active mechanisms. The active mechanism of demethylation of H3K27me3 is supposed to be mediated through the specific demethylase enzyme histone lysine demethylase (KDM) 6B (also known as Jumonji domain-containing protein 3 [JMJD3]) (Vernimmen et al., 2011). In addition, as cells differentiate into erythroid cells a dramatic increase in the levels of the H3K4me3 chromatin modification which signals transcriptional activation is seen at the promoter of the α-globin gene (De

Gobbi et al., 2011). This is believed to be mediated through the action of Set/MLL histone methyltransferases. The net effect of these chromatin modifications and the recruitment of TFs (discussed below) is the activation of transcription of α-globin to produce massive amounts of globin proteins required for the erythroid cells.

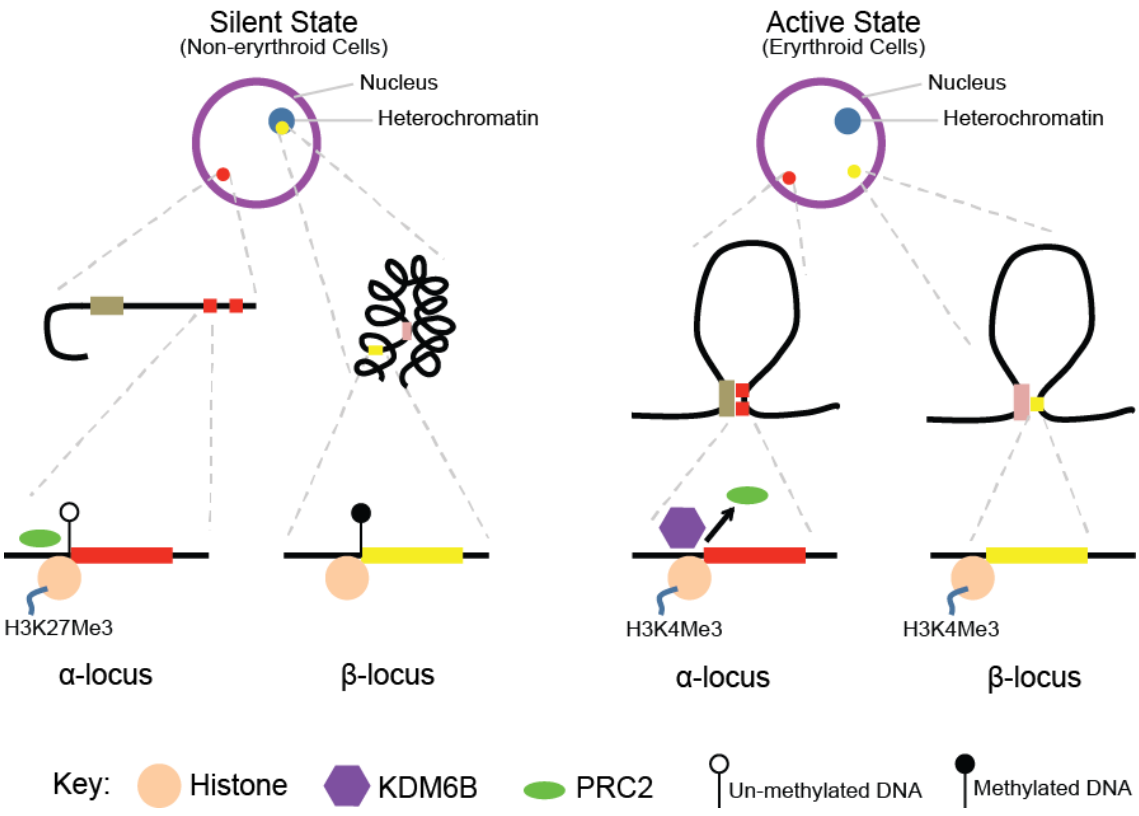


Figure 1.9 – Contrasting epigenetic landscape of α- and β- globin loci. In the silent, non-expressed state (non-erythroid cells) the α-globin locus (red) is in an open chromatin environment whereas the β-globin locus (yellow) is in a closed chromatin conformation incorporated into heterochromatin (dark blue circle within the nucleus). The promoter of α-globin is un-methylated, bound by PRC2 and the associated histone has the H3K27me3 modification. By contrast, the β-globin locus is heavily methylated and does not show binding of PRC2 or the H3K27me3 chromatin modification. In the active state (erythroid cells) both loci are located away from heterochromatin, form loop structures to facilitate respective enhancer-promoter interactions and both promoters show H3K4me3 active chromatin modification. However, only in the α-globin locus, PRC2 is detached and KDM6B is recruited to facilitate removal of the H3K27me3 repressive chromatin modification. Adapted from (Mettananda et al., 2015)

1.6.3 Transcriptional factor control of α-globin gene

Transcription factors (TF) are a group of proteins that bind to specific DNA sequences, thereby controlling tissue-specific expression of genes. During human erythroid differentiation the key TFs, GATA-binding factor 1 (GATA1), KLF1, stem cell leukaemia (SCL), transcription factor E2-alpha (E2A), Lim-only 2 (LMO2) and LIM domain binding protein 1 (Ldb1) are recruited to the four upstream enhancers of the α-globin. In addition, another erythroid TF, nuclear factor-erythroid 2 (NF-E2) is bound to two of the enhancers; MCS-R2 and to a lesser extent MCS-R1 (De Gobbi et al., 2007). Finally, general transcription factors together with RNA polymerase II (PolII) (collectively referred to as the preinitiation complex) are recruited to the MCS-R1 and MCS-R2 enhancers and then subsequently to the α-globin promoters thereby initiating transcription. During this process a physical interaction between enhancers and promoters are believed to occur through the mechanism of chromatin looping (Vernimmen, 2014; Vernimmen et al., 2009).

1.6.4 Contrasting regulation α- and β-globin expression

Despite being expressed in the same cell type in a similar developmental-stage-specific manner, the regulation of β-globin demonstrates significant differences to the regulation of α-globin. The human β-globin gene is located on chromosome 11 in a relatively gene sparse, late replicating region of the genome which, in non-erythroid cells, is in a closed heterochromatic environment (Higgs et al., 2005; Higgs et al., 1998). The promoter of β-globin is methylated and not associated with CpG islands. Therefore, in contrast to α-

globin, in non-erythroid cells the promoter of β-globin is not bound by PRC2 and does not have the repressive chromatin mark H3K27me3 (Chang et al., 2013; Garrick et al., 2008). Instead, silencing of the β-globin gene is mediated through DNA methylation and co-localisation with nuclear heterochromatin (Higgs et al., 2005) (Figure 1.9 & Table 1.1).

Table 1.1 – Differences in the structure and function of the α- and β-globin gene clusters

	α- globin cluster	*β-globin cluster*
Location	16p13.3 telomeric	11p15.5 interstitial
Guanine-cytosine (GC) content	54%	39%
CpG islands	Common	None
Gene density	High	Low
Chromatin	Open	Closed
Replication timing	Early	Late
Predominant mutations	Deletions	Point mutations
Evolution of intergenic regions	Rapid	Slow
Expression in hybrids	Early	Late
Binding of polycomb repressive complex	Present	Absent
H3K27me3 chromatin modification	Present	Absent
KDM6B enzyme	Recruited	Not recruited

In erythroid cells when the β-globin is activated, in a process similar to α-globin, the distant enhancers (β-LCR) and the gene promoters are bound by erythroid specific TFs which include (but not limited to) GATA1, NF-E2, KLF1 and SCL (Cantor and Orkin, 2002; Kim and Dean, 2012). Similar to α-globin the interaction of the enhancers and promoters are believed to be maintained by chromosomal looping (Palstra et al., 2003). The promoter is marked by the H3K4me3 active chromatin modification. However, in contrast to α-globin, the

repressive chromatin modification H3K27me3 is not present at the β-locus and therefore, histone demethylases do not play a role in the activation of β-globin (Garrick et al., 2008).

1.7 Selective silencing of α-globin expression: candidate therapeutic pathways

The current understanding of the regulation of the α and β-globin genes (discussed in section 1.6) in erythroid cells make it possible to consider new strategies to selectively control α-globin expression to the appropriate degree to be useful for patients with β-thalassaemia. Several pathways warrant careful consideration. Post-transcriptional gene silencing through RNA interference (RNAi) is a potential avenue which has been explored to some extent in the past (Voon et al., 2007). The other plausible pathways including pharmacological targeting of epigenetic regulation and selective disruption of enhancer elements by genome editing have not been examined before. The development of new drugs targeting specific epigenetic pathways and the advent of genome engineering using programmable, sequence-specific endonucleases have enable scientists to specifically regulate the expression of a single gene (discussed below). These novel approaches in the context of selective silencing of α-globin expression are examined in great detail in this thesis.

1.7.1 RNA interference strategy

RNA interference mediated through small double stranded RNAs including small interfering RNA (siRNA) and short hairpin RNA (shRNA) has recently emerged as a powerful tool for post-transcriptional gene silencing (Gavrilov and Saltzman, 2012; Hannon, 2002). These double-stranded RNA molecules incorporate into a protein complex known as the RNA-induced silencing complex (RISC) within the cytoplasm, where the strands are separated, and one strand guides RISC to the complementary region of the target mRNA, suppressing gene expression either by degrading mRNA or blocking mRNA translation. Thus, RNAi produces efficient and specific gene silencing. Efforts to translate this new discovery into clinical applications for disease treatment has been reasonably successful and clinical trials of RNAi therapies in treating cancer, cardiovascular diseases, infections and eye diseases are in their early stages (Deng et al., 2014).

Attempts to reduce α-globin expression using this technique have been reported. In an experimental model developed by Voon and colleagues, *in vitro* differentiated β-thalassaemia murine primary erythroid cells were transfected with siRNA targeting the α-globin mRNA. This resulted in a 50% knockdown of α-globin gene expression (Voon et al., 2008). As expected, this was associated with reduction of intracellular ROS and significant phenotypic improvement. Xie and others performed an *in vivo* experiment to knock down α-globin expression by shRNA in a β-thalassaemia mouse model (Xie et al., 2007). Lentiviral vectors with shRNA targeting α-globin were microinjected into β-thalassaemia heterozygous murine single-cell embryos and a transgenic mouse was

generated which produced 20%-35% less α-globin. These mice demonstrated sustained phenotypic improvement in the RBCs with less poikilocytosis, fewer target cells, increased haemoglobin values and red cell counts and reduced reticulocyte counts. The same group then directly injected siRNA plasmids targeting α-globin mRNA into the tail veins of the β-thalassaemia mice and analysed the effects (Xie et al., 2011). Reductions in poikilocytosis, target cells and reticulocyte counts were observed after treatment. Red cell counts and haemoglobin levels were unchanged however.

Significant challenges and a number of barriers still need to be overcome before RNAi based therapies will be useful in the clinic. The key challenge is to develop safe and effective delivery methods (Kanasty et al., 2013). Double stranded RNA molecules are negatively charged and do not readily cross the cell membrane. Traditional delivery methods relied on viral vectors such as retroviruses, lentiviruses, adenoviruses and adeno-associated viruses but several limitations including carcinogenesis, immunogenicity, broad tropism, limited packaging capacity and difficulty of vector production have limited its clinical use (Yin et al., 2014). Similarly, poor stability of siRNA in the extracellular environment, difficulties in targeting them to the cells of interest, limitations in dose-titration, off-target silencing, inherent toxicity and the immunogenic potential of siRNA will also need to be addressed before development of effective therapeutics using this tool (Davidson and McCray, 2011; Gavrilov and Saltzman, 2012).

1.7.2 Pharmacological strategy through epigenetic drug targeting

Epigenetic drug targeting is a promising and rapidly developing area in therapeutics which aims to alter gene expression by affecting the chromosomal environment via specific epigenetic pathways. Many proteins have been identified that write, read or erase epigenetic modifications which include well characterized processes like DNA methylation, histone methylation and histone acetylation (Ivanov et al., 2014). Nucleoside analogue irreversible inhibitors of the DNA methyltransferases (e.g. azacitidine and decitabine), have been approved for clinical use for many years. Shortly after their approval HDAC inhibitors vorinostat and romidepsin were approved and a number of specific inhibitors of other epigenetic pathways are currently under development with some entering clinical trials (Helin and Dhanak, 2013). Although the current main indications for epigenetic drugs are various forms of cancers, some drugs are being tested for non-malignant diseases. In fact, multiple studies including phase 1/2 clinical trials are underway to evaluate the usefulness of vorinostat (ClinicalTrials.gov Identifier NCT01000155) and decitabine (ClinicalTrials.gov Identifier NCT01685515) as HbF inducing agents for the treatment of sickle cell disease.

As outlined earlier, the chromatin environment and silencing mechanisms of the α-globin and β-globin clusters show clear and striking differences. This distinction may be very useful in the development of therapeutic strategies and it is likely that drugs or small molecules which act by altering the chromatin environment will have a differential effect on α- and β-globin expression. More precisely, drugs acting through the PRC2 pathway, histone methylation and

histone demethylation are candidates likely to be useful in this regard. If pharmacologically active small molecules could reduce expression of the α-globin genes without affecting the expression of the β-like globin genes, most barriers faced by RNAi could be overcome. Delivery would not be a major problem and dose titrations may not be too challenging once the pharmacokinetics are determined. Furthermore, transformation into clinical practice would be easier, development would be less costly and the drug would become available for patients in developing countries who are at the greatest need of new approaches to the treatment of thalassaemia.

1.7.3 Genome editing strategy

Genome editing using sequence-specific programmable artificially engineered nucleases has not only revolutionized biomedical research as powerful tools for investigating gene regulation but also provides a realistic approach to the treatment of human genetic diseases. These nucleases create double strand breaks at specific chosen locations in the genome and, when repaired, create mutations at the targeted sites which can be designed to result in either loss of gene function or disruption of non-coding regulatory sequences. There are three main families of engineered nucleases which are currently being used: Zinc finger nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs) and RNA-guided nucleases clustered, regularly interspaced, short palindromic repeat (CRISPR)/CRISPR-associated (Cas) system. Precise targeting to a specific genomic location, high efficiency of mutagenesis and fewer disruptions of the genome are the main advantages of programmable nucleases over traditional gene therapy approaches (Kim and Kim, 2014).

Genome editing using ZFN is already being used successfully in humans to disrupt the C-C chemokine receptor type 5 (CCR5) encoding gene in $CD4^+$ T-cells to promote human immunodeficiency virus (HIV)-1 resistance (Urnov et al., 2010). In a recently published phase 1 clinical trial, a group of HIV-1 infected patients were successfully transfused with autologous ZFN-CCR5-modified T-cells and a follow up phase 2 trial to evaluate safety and tolerance of treatment is currently ongoing (Tebas et al., 2014).

An approach similar to this could be used to modify HSPCs to down-regulate α-globin expression in patients with β-thalassaemia. There is very clear evidence that the upstream enhancer element MCS-R2 plays a critical role in α-globin gene expression. In a rare naturally occurring deletion removing only the MCS-R2 region, individuals with a heterozygous mutation have a reduced α-globin chain output similar to that seen in mild α-thalassaemia, which is desirable to produce a beneficial effect in β-thalassaemia (Coelho et al., 2010). A rare homozygote for this deletion has haemoglobin H (HbH) disease but importantly no other abnormality showing that this element can be removed with no other untoward effects (discussed further in chapter 7; section 7.1.3). Therefore disruption of this single element using programmable nucleases has the potential to be a very useful and simple therapeutic strategy in patients with β-thalassaemia and in particular for those with the common HbE β-thalassaemia genotype.

1.8 Summary, hypothesis and aim of the thesis

Despite being one of the first genetic diseases to be identified, β-thalassaemia still remains an essentially fatal disease without a cure in a majority of patients. Attention to the development of new therapeutic strategies has been centred on pathways that increase γ-globin expression and hence the production of foetal haemoglobin. However, clinical and genetic data accumulated over the last 30 years as well as a precise understanding of the pathophysiological mechanisms have indicated that reduction of α-globin expression is clinically beneficial to patients with β-thalassaemia. Also, the current understanding of the regulation of α- and β-globin genes has identified contrasting regulatory mechanisms, especially the mechanisms that silence the two genes. Therefore, I hypothesise that the selective silencing of the α-globin expression, without affecting the expression of β-like globin genes is therapeutically feasible.

The general aim of this thesis is to identify plausible pathways that selectively down-regulate α-globin expression without affecting the expression of β-like globin genes in human primary erythroid cells, which could lead to greater balanced globin chains and potential clinical applications for β-thalassaemia. In order to achieve this aim, I had three specific objectives. The first objective was to develop, optimise and characterise a small scale human erythroid differentiation system which is suitable to identify effects of external factors which alter globin gene expression. The second was to screen a collection of drugs to identify pharmacological compounds that would down regulate α-globin expression. The third and final objective was to disrupt the critical enhancer

element of the human α-globin using genome editing technology to achieve reductions in α-globin gene expression.

In chapter 3 of this thesis, I have described the development and characterisation of a suitable model to test the effects of various tools altering α-globin expression. This includes the human primary erythroid differentiation system and a globin gene expression detection strategy which were then validated using RNAi and pharmacological methods. In the next three chapters, I have described the pharmacological approach to down regulate α-globin expression. Chapter 4 describes the results of a medium throughput screen of epigenetically active small molecule compound library and the subsequent two chapters describe the further validation and characterisation of two potential compounds which I identified. Next, in chapter 7, I have used a genome editing approach to disrupt the most critical regulatory element of human α-globin gene to selectively silence its expression. Finally I propose that selective down regulation of α-globin expression is feasible through pharmacological and genome engineering approaches and could be used as new pathway of therapy for β-thalassaemia.

Chapter 2: Materials and Methods

2.1 Cell culture methods and associated techniques

2.1.1 Purification of human CD34$^+$ HSPCs

Human umbilical cord (UC) blood and adult peripheral blood buffy coat residues (in the form of component donation leucocyte cones) were purchased from the National Health Service (NHS) Blood and Transplant, Oxford, UK. Ethical approval for the study was granted by North West Research Ethics Committee of NHS National Research Ethics Services. Fresh blood samples were diluted with PBS (1:1 for UC blood; 1:10 for buffy coat residues) and the mononuclear cells (MNCs) were collected from the interface after fractionation on Histopaque-1077 Hybri-Max (Sigma). MNCs were washed twice with PBS and the CD34$^+$ cells were magnetically labelled with CD34 MicroBead Kit (Miltenyl Biotech) and were then purified using LS columns and MidiMACS separator (Miltenyl Biotech) according to the manufacturer's instructions. To increase the purity, two LS columns were used sequentially. Purified CD34$^+$ HSPCs were stored in a freezing solution (90% foetal calf serum and 10% dimethyl sulfoxide) in liquid nitrogen until required.

2.1.2 Erythroid differentiation culture from human CD34$^+$ HSPCs

The CD34$^+$ HSPCs were differentiated into erythroid cells over 21 days using a two-phase liquid culture system. During phase-1, HSPCs were thawed and seeded at a concentration of $5x10^4$/ml in serum-free StemSpan SFEM II (Stemcell technologies) medium supplemented with 100ng/ml stem cell factor

(SCF) (Peprotech), 10ng/ml interleukin-3 (Peprotech), 0.5IU/ml human recombinant erythropoietin, 10µg/ml cholesterol rich lipids (Sigma), 100U/ml penicillin/streptomycin (Gibco) and 2mM glutamine (Gibco). After 7 days, cells were transferred into phase-2 medium approximately at a concentration of $2x10^5$/ml. Phase-2 medium was similar to phase-1 medium except for addition of 0.5mg/ml iron saturated holotransferrin (Sigma) and higher erythropoietin concentration (3U/ml). In both phases, cells were incubated at 37°C in 5% CO_2 environment and the cell concentrations were maintained below $2x10^6$/ml by adding fresh medium every 2-3 days.

2.1.3 Erythroid differentiation culture from sorted single human HSPCs

During genome editing experiments, single cells were sorted into the wells in Terasaki multiwell plates and were cultured in 20µl of phase-1 medium. The cells were incubated at 37°C in a 5% CO_2 environment in a tissue culture box kept in the incubator for up to 14 days without addition or change of medium.

2.1.4 Erythroblast culture from mouse foetal liver

At the E12.5 stage, the mice were euthanized with CO_2 followed by cervical dislocation, and foetuses were extracted under sterile conditions. Foetal livers were dissected out with sterile forceps and the liver cells were seeded in medium containing StemPro (Invitrogen) supplemented with 1U/ml erythropoietin, 50ng/ml SCF, 1µM dexamethasone and glutamine. Cells were expanded for 6 days with addition of fresh medium to maintain cell concentration around $1x10^6$/ml and then the $Ter119^+$ mature erythroid cells

were depleted using mouse anti-Ter-119 MicroBeads (Miltenyl Biotech). Purified erythroblasts were reseeded at a concentration of $1x10^6$/ml in medium containing StemPro (Invitrogen) supplemented with 5U/ml erythropoietin, 0.5mg/ml transferrin and glutamine. Mouse experiments including dissection of mouse foetal livers were done by Jacqline Sharpe and Jackie Sloane-Stanley of Weatherall Institute of Molecular Medicine, University of Oxford.

2.1.5 Trypan blue test for viability

Equal volumes of cell suspension and 0.4% trypan blue (Sigma) were incubated for 3-5 minutes at room temperature. Blue stained (non-viable) and unstained (viable) cells were counted under a light microscope using a haemocytometer.

2.1.6 Benzidine stain for haemoglobin

Staining solution was prepared by adding 1 part of 30% hydrogen peroxide (Sigma) to 100 parts of 0.2% 3,3′,5,5′-tetramethyl benzidine (Sigma) in 5% acetic acid (Sigma). Equal volumes of this solution and cell suspension were incubated at room temperature for 2-3 minutes. The percentage of blue (haemoglobinised) cells was scored under a light microscope using a haemocytometer.

2.1.7 Morphology analysis by cytospin

Cytospins were prepared by centrifuging 200µl of cell culture at 400rpm for 4 minutes onto a glass slide. The slides were air-dried and stained with modified Wright stain using a Hemateck slide strainer.

2.1.8 Flow cytometry

Cells (100,000) were washed with Phosphate-buffered saline (PBS) (Gibco), resuspended in 200µl of 2% bovine serum albumin (Sigma) in PBS and labelled for 20 minutes on ice with the following anti-human antibodies; allophycocyanin (APC) conjugated anti-CD34 (Miltenyl Biotech; catalogue number 130-090-654), fluorescein isothiocyanate (FITC) conjugated anti-CD71 (BD Biosciences; catalogue number 555536), phycoerythrin (PE) conjugated anti-CD235a (BD Biosciences; catalogue number 340947) and FITC conjugated anti-CD45 (BD Biosciences; catalogue number 560976). Dead cells were identified by Hoechst 33258 pentahydrate nucleic acid stain (Invitrogen) and were excluded. Analysis was done on Cyan ADP (Beckman Coulter) analyser using Summit v4.3 and FlowJo V10 software after gating on viable cells.

2.2 Small molecule library and treatment of cells

The epigenetic inhibitor compound library consisted of 37 epigenetically active compounds and was a kind donation by Udo Oppermann, Structural Genome Consortium, Oxford, UK. During the small molecule screen, primary human erythroid cells at day 7 of culture (corresponding to proerythroblast stage) were

incubated with the compounds for 72 hours at concentrations specified in table 2.1. In the subsequent experiments, incubation times with IOX1 were 72 hours in primary human erythroid cells and 36 hours in mouse primary erythroid cells.

Table 2.1 – List of small molecules included in the epigenetic inhibitor compound library

Small molecule inhibitor	*Inhibitor target/class*	*Working concentration (µM)*
(+)- JQ1	Bromodomains	1
PFI-1	Bromodomains	5
CBP/ BRD4 (0383)	Bromodomains	5
SMARCA	Bromodomains - SMARCA	2.5
I-BET	Bromodomains - BRD2/3/4	1
BAZ2B	Bromodomains - BAZ2B	1
Bromosporin	Bromodomains - broad spectrum	1
CBP probe (0113)	Bromodomains - CBP Enantiomer	5
RVX-208	Bromodomains - BET clinic	5
5-Aza-deoxy-cytidine	DNA methyltransferase - DNMT1/3	5
5-Azacitidine	DNA methyltransferase	10
CXD101	HDAC	1
Valproic acid	HDAC -aliphatic acid compounds	1000
Entinostat (MS-275)	HDAC -ortho-amino anilides	0.5
SRT1720	HDAC -SIRT1 (indirect? Activator)	1
EX527	HDAC -SIRT1	1
Trichostatin A	HDAC -hydroxamic acid Class I & 2	0.5
Vorinostat	HDAC -hydroxamic acid	2.5
C646	Histone acetyltransferase p300/CBP	1
Methylstat	Histone demethylase	2.5
UNC0638	Histone methyltransferase - G9a/GLP	1
GSK343	Histone methyltransferase - EZH2	3
SGC0946 (DOT1L probe)	Histone methyltransferase - DOT1L	7.5
SET7/9-1	Histone methyltransferase - SETD7	2.5
SET7/9-2	Histone methyltransferase - SETD7	2.5
SET7/9-3	Histone methyltransferase - SETD7	2.5

Small molecule inhibitor	*Inhibitor target/class*	*Working concentration (µM)*
Chaetocin	Histone methyltransferase - SUV39H1	0.05
K00135	Kinase inhibitor - ATP competitive -PIM	1
5-Iodotubercidin "HASPIN"	Kinase inhibitor - ATP mimetic - Haspin	1
IOX1	Lysine demethylase - broad range	40
Tranylcypromine	Lysine demethylase - LSD1	20
GSK-J4	Lysine demethylase - KDM6A/B (JMJD3/UTX)	10
UNC1215	Malignant Brain Tumour Domains - L3MBTL3	5
(-)- JQ1	Bromodomains - Negative control	1
Rucaparib	Poly ADP ribose polymerase (PARP)	10
Olaparib	Poly ADP ribose polymerase (PARP)	1
IOX2	Prolyl hydroxylases EGLN1 (PHD2)	10

2.3 DNA methods

2.3.1 DNA extraction

Cells in culture were spun down at 1000rpm for 5 minutes and cell pellets were washed once with PBS and resuspended in 200µl PBS. Genomic DNA from cell pellets were extracted using DNeasy blood and tissue kit (Qiagen) according to manufacturer's protocol and eluted in 50µl of elution buffer. Polymerase chain reaction (PCR) product DNA fragments from agarose gel bands were purified using QIAquick gel extraction kit (Qiagen) according to the manufacturer's protocol. DNA concentration was determined by measuring absorbance at 260nm using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific) and DNA was stored in -20°C for long term storage.

2.3.2 Polymerase chain reaction

Routine PCR reactions were performed with AmpliTaq Gold 360 PCR master mix (Invitrogen). A typical reaction mix contained 12.5µl of AmpliTaq Gold 360 PCR master mix, 0.5µl each of forward and reverse primers (10µM) (Table 2.2), 50ng template DNA with the final volume made up to 25µl with water. PCR conditions were: 95°C for 10 minutes to activate the enzyme, followed by 35 cycles of 95°C 1 minute denaturation, 1 minute annealing, 72°C 1 minute extension with a final extension at 72°C for 10 minutes. PCR amplification products were tested by gel electrophoresis of 5µl product on a 2% agarose gel containing ethidium bromide in TBE buffer at 120V for 1 hour and visualised using video gel documentation system Gel-Doc2000 (BioRad Laboratories). PCR products were purified using QIAquick PCR purification Kit (Qiagen) as per manufacturer's instructions.

Table 2.2 – Primer sequences used in PCR

Target region	*Primer/probe*	*Sequence*
Human MCS-R2 region	Forward primer	TGGTCCTGAAGGATGAGAAG
	Reverse primer	AGCAACAGTCCTTTCTCTGG

2.3.3 Whole genome DNA amplification from small number of cells

In single cell clone experiments, genomic DNA from small cell numbers was amplified directly from cell lysate using illustra single cell GenomiPhi DNA amplification kit (GE Healthcare). Briefly, 10-1000 cells from each clone were washed with PBS, spun down at 1000g for 5 minutes and the cell pellet was

then stored at -70°C. Cells were thawed, resuspended in 1µl PBS and lysed by adding 1µl lysis buffer (provided as part of the kit) and incubating at 65°C for 10 minutes. Whole genome amplification was done by adding amplification reaction mix (prepared by adding 1µl GenomiPhi enzyme mix, 11µl GenomiPhi reaction buffer, 1µl GenomiPhi amplification mix and 4µl water) and incubating at 30°C for 2 hours, followed by heat inactivation at 65°C for 10 minutes. Amplified DNA was purified using QIAquick PCR purification Kit (Qiagen) as per manufacturer's instructions.

2.3.4 Sanger sequencing

Sequencing of DNA fragments was performed in the WIMM DNA sequencing facility using BigDye Terminator v3.1 chemistry in ABI-3730 DNA analyser (Applied Biosystems). Sequences were inspected using the Sequencher application (Gene Codes Corp) and analysed by Chris Fisher of Weatherall Institute of Molecular Medicine, University of Oxford.

2.4 RNA methods

2.4.1 RNA extraction

Cells in culture were spun down at 1000rpm for 5 minutes and cell pellets were stored in -70°C until RNA extraction. Total RNA was purified using RNeasy mini kit (Qiagen) according to the manufacturer's instructions. RNA concentration was determined using Nanodrop ND-1000 spectrophotometer (Thermo Scientific) and RNA quality was assessed using Agilent RNA 6000 Nano Kit

and Agilent 2100 bioanalyser instrument (Agilent Technologies). RNA was stored at -70°C for long term storage. For some of the experiments, contaminating genomic DNA was removed using Turbo DNA-free Kit (Life Technologies) as per manufacturer's instructions.

2.4.2 Reverse transcription (RT)

Complementary DNA (cDNA) was prepared from 500ng of total RNA using high capacity RNA to cDNA kit (Applied Biosystems). In brief, for each reverse transcription reaction, the reaction mix contained 5µl RT buffer, 0.5µl RT enzyme mix and 500ng of RNA in a final volume made up to 10µl with nuclease-free water. The tube was transferred to a thermocycler and incubated with the following temperature profile: 37°C for 60 minutes and then 95°C for 5 minutes. cDNA was stored in -20°C for long term storage.

2.4.3 cDNA preparation from small number of cells

In single cell clone experiments, from small number of cells, cDNA was prepared directly from cell lysate using TaqMan gene expression cells-to-C_T kit (Life technologies). Briefly, 10-1000 cells from each clone were washed with PBS, spun down at 400g for 5 minutes and the cell pellets were stored in -70°C until lysis. Cells were thawed, resuspended in 5µl PBS and were immediately lysed by adding 50µl lysis buffer with DNase I (provided as part of the kit) and incubating at room temperature for 5 minutes. The reaction was stopped by adding 5µl stop solution and incubating at room temperature for 2 minutes. Reverse transcription was carried out in a reaction mixture which contained

25µl RT buffer, 2.5µl RT enzyme mix, 2.5µl nuclease-free water and 20µl cell lysate. The tubes were transferred to a thermocycler and incubated with the following temperature profile: 37°C for 60 minutes and then 95°C for 5 minutes. cDNA was stored in -20°C for long term storage.

2.4.4 Quantitative real-time polymerase chain reaction (qRTPCR)

Validated and inventoried TaqMan assays (Applied Biosystems) (Table 2.3) were used in qRTPCR experiments. In all the assays, primers and probes span exon junctions. qRTPCR was carried out in 10µL reaction volume in technical triplicate using 96-well plates in 7500 fast real time PCR system (Applied Biosystems) according to the manufacturer's protocol. Data was analysed by 7500 software v2.0.6 using delta delta CT method. Ribosomal Protein L13a (RPL13A) was used as the house keeping gene for normalization as a previous study showed it to be constant throughout human erythropoiesis (Merryweather-Clarke et al., 2011). For mouse experiments, previously reported (Wallace et al., 2007) custom made TaqMan assays specific to human α-globin and mouse α- and β-globin were used (Table 2.4).

Table 2.3 – List of inventoried human TaqMan assays used in qRTPCR

Target gene ID	*Target gene name*	*TaqMan assay ID*
HBA	α-globin	Hs00361191_g1
HBB	β-globin	Hs00747223_g1
HBD	δ-globin	Hs00426283_m1
HBE1	ε-globin	Hs00362216_m1
HBG	γ-globin	Hs00361131_g1
HBM	µ-globin	Hs01392876_g1
HBQ	θ-globin	Hs00362218_g1

HBZ	ζ-globin	Hs00923579_m1
RPL13A	ribosomal protein L13a	Hs03043885_g1
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	Hs02758991_g1
KDM6A	lysine (K)-specific demethylase 6A	Hs00958902_m1
KDM6B	lysine (K)-specific demethylase 6B	Hs00996325_g1
KDM4A	lysine (K)-specific demethylase 4A	Hs00206360_m1
KDM4C	lysine (K)-specific demethylase 4C	Hs00909579_m1
FTH1	ferritin, heavy polypeptide 1	Hs01000476_g1
PABPC1	polyadenylate-binding protein 1	Hs01598422_m1
RPL18	ribosomal protein L18	Hs00965812_g1
AHSP	α-hemoglobin stabilizing protein	Hs00372339_g1

Table 2.4 – Primer sequences and TaqMan probes used in mouse experiments (Wallace et al., 2007)

Target gene	*Primer/probe*	*Sequence*
Human α-globin	Forward primer	GCCCTGGAGAGGATGTTCCT
	TaqMan probe	CCTTCCCCACCACCAAGACCTACTTCC
	Reverse primer	CGTGGCTCAGGTCGAAGTG
Mouse α-globin	Forward primer	GCTGAAGCCCTGGAAAGGAT
	TaqMan probe	CTTCCCCACCACCAAG
	Reverse primer	GGCTTACATCAAAGTGAGGGAAGTA
Mouse β-globin	Forward primer	AGTGAGCTCCACTGTGACAAGCT
	TaqMan probe	CATGTGGATCCTGAGAACTTCAGGCTCCT
	Reverse primer	CAGCACAATCACGATCATATTGC

2.4.5 High-throughput qRTPCR by Fluidigm Biomark HD system

During the small molecule screen, high throughput qRTPCR was performed using a Fluidigm 48.48 Gene Expression Chip and inventoried exon-spanning TaqMan assays (Applied Biosystems) for globin and housekeeping genes

(Table 2.3). Data was analysed using Fluidigm Real-Time PCR analysis 3.1.3 software and were normalised to the mean of four housekeeping genes (RPL13A, RPL18, GAPDH and FTH1).

2.4.6 Nanostring gene expression assay

For Nanostring experiments, we purchased custom made capture probe set panel and consumables from Nanostring technologies. In each reaction, 100ng of total RNA was hybridised with capture probe set and reporter code set at 65°C in a thermocycler for 16 hours. Hybridized samples were then processed using an nCounter prep station and nCounter digital analyser (Nanostring Technologies) according to manufacturer's instructions. Raw data was normalized to an internal positive spike-in control to normalize to all platform associated sources of variation and then to the geometric mean of eight housekeeping genes (RPL13A, RPL18, GAPDH, PABPC1, CA2, FTH1, PAIP2 and LAPTM4A).

2.4.7 Microarray

Microarray whole genome gene expression analysis were performed by the high-throughput genomics group at the Wellcome Trust centre for human genetics, University of Oxford using Illumina human HT12v4.0 expression BeadChip. Biotin labelled cRNA synthesized from 80ng of RNA was used for hybridization and the washed chips were then scanned using Illumina iScan Scanner. Data were analysed by Stephen Taylor of Computational Biology Research Group, Radcliffe Department of Medicine, University of Oxford using

Linear models for microarray data (LIMMA) and filtered by adjusted P value of <0.05 to identify differentially expressed genes. Low expressing genes were filtered out when comparing the expression levels of genes in IOX1 treated and untreated cells.

2.5 Protein methods

2.5.1 Analysis of haemoglobin

Weak cation exchange high performance liquid chromatography (CE-HPLC) was performed using Bio-Rad VARIANTnbs system (Bio-Rad Laboratories) according to manufacturer's instructions by Jennifer Eglinton of the Haematology Department, Oxford University Hospitals NHS Trust.

2.5.2 Histone protein extraction

Roughly, $1x10^7$ cells in culture were spun down at 1000rpm for 5 minutes and the cell pellet was washed twice with cold PBS and lysed in ice for 10 minutes in lysis buffer containing 0.2% Triton X-100, 10mM HEPES pH 7.6, 1.5mM $MgCl_2$, 10mM KCl, 10mM sodium butyrate and protease inhibitor. Cell lysate was then spun at 3000rpm for 3 minutes at 4°C and the supernatant discarded. Pelleted nuclei were resuspended in 0.4M HCl to solubilise the histone proteins, which were precipitated overnight in acetone at −20°C. Precipitated histone proteins were dissolved in 8M urea and stored at −20°C for long term storage.

2.5.3 Western blot

Sample mixture was prepared by adding 50µl protein in 8M urea, 10µl 1M dithiothreitol (DTT) (sigma) and 40µl protein loading buffer and was denatured at 95°C for 5 minutes. Samples were loaded into the wells of precast NuPAGE 4-12% gradient bis-tris gel (Invitrogen). Gel was run in MES buffer at a constant voltage of 120V with Page ruler plus protein ladder run adjacent to the protein samples. Size separated protein extracts were transferred to a polyvinylidene difluoride membrane (ImmobilonP; Millipore) by electrophoresis. Subsequently membrane was blocked with 5% milk in TBST for 1 hour, washed and then incubated for 1 hour at room temperature with a primary antibody for H3K27me3 (Millipore; catalogue number 07-449), H3K9me3 (Abcam; catalogue number ab8898) or histone H3 (Abcam; catalogue number ab1791). Then the membrane was incubated with an appropriate horseradish peroxidase-conjugated secondary antibody (Dako) and bands visualised on radiographic film (Kodak) using the EZ-ECL chemiluminescence detection kit (Biological Industries).

2.6 Chromatin Immunoprecipitation (ChIP)

2.6.1 Histone protein ChIP

Quantitative ChIP assays were performed using the Millipore ChIP assay kit (Merck Millipore) according to the manufacturer's instructions. Briefly, 1×10^7 cells per immunoprecipitation were fixed with 0.4% formaldehyde for 10 min at room temperature and the reaction was quenched by adding 0.125M glycine.

Cells were washed twice in cold PBS with protease inhibitors (Roche) and lysed in sodium dodecyl sulfate (SDS) lysis buffer. The lysate is sonicated to obtain DNA fragments of an average size of 500-1000bp using the Bioruptor instrument (Diagenode). Then, 200µl of sonicated chromatin was diluted 1:10 in ChIP dilution buffer and 50µl of the diluted material was taken as input. Diluted chromatin fragments were pre-cleared with protein A: agarose beads followed by immunoprecipitation overnight with one of the following antibodies: H3K27me3 (Millipore; catalogue number 07-449) and H3K9me3 (Abcam; catalogue number ab8898). Protein A: agarose beads were added and successive washes performed with high salt buffer, low salt buffer, LiCl buffer and TE buffer. This was followed by elution of bound chromatin fragments, reversal of crosslinks, proteinase K digestion and purification of DNA by phenol: chloroform extraction.

2.6.2 ChIP-PCR

Precipitated DNA was quantified by qRTPCR assay with previously described (Lynch et al., 2012; Shi et al., 2013) primer pairs for human α-, β- and γ-globin promoter sequences and a primer pair for an intergenic region between the ε- and γ-globin genes as a negative control (Shi et al., 2013)(Table 2.5). Quantitative PCR was performed using SYBR green master mix (Applied Biosystem) in 7500 fast real time PCR system (Applied Biosystems) using in 25µL reaction volume in technical triplicate according to manufacturer's protocol.

Table 2.5 – Primer sequences used for ChIP-PCR

Target gene	*Primer/probe*	*Sequence*
α-globin promoter	Forward primer	GGGCCGGCACTCTTCTG
	Reverse primer	GGCCTTGACGTTGGTCTTGT
β-globin promoter	Forward primer	AACTGTGTTCACTAGCAACCTCAAA
	Reverse primer	ACAGGGCAGTAACGGCAGACT
γ-globin promoter	Forward primer	GCAAATATCTGTCTGAAACGG
	Reverse primer	GTGGAACTGCTGAAGGGTGCTT
ζ-globin promoter	Forward primer	CTGGGTGGACCTAACCCTTG
	Reverse primer	ACATGCCAGTACCCTTCCAC
Negative control β-locus	Forward primer	TCCCACTCTGTGGGTTGTCTGTTT
	Reverse primer	CCCTTCTACACATTGGCTTAGGAAAGG
Negative control 1 α-locus	Forward primer	AACTGCCTTTGCACCCTCAT
	Reverse primer	ACCTTTTCTGGGCCTACAGAC
Negative control 2 α-locus	Forward primer	ACAATCAGAACAACTGTGAGGT
	Reverse primer	TTGCGACACTGACTGTGTGA

2.7 Transient gene knock-down methods

2.7.1 Transfection of short-interfering RNA (siRNA)

Commercially available, validated siRNAs were purchased from Qiagen. Transfections were performed using Amaxa human $CD34^+$ cell nucleofector kit (Lonza) in Amaxa nucleofector I device (Lonza). Briefly, culture medium containing $1.5x10^6$ cells was spun at 1000rpm for 5 minutes. The cell pellet was resuspended in 100µl nucleofection solution and was nucleofected with 100pmol of siRNA in the provided cuvette using U-08 program. Cell suspension was soon transferred back in to the culture medium.

2.7.2 Short-hairpin RNA (shRNA)

Several sequence verified shRNAs cloned into the pLKO.1-puro vector in the form of lentiviral transduction particles (Sigma-MISSION) were purchased. Roughly $4x10^4$ cells were incubated for 24 hours at 37°C with lentiviral transduction particles at multiplicities of infection ranging between 3 to 10. Viral particle containing medium was removed after 24 hours and fresh medium added. After a further 24 hours cells were transferred in to a medium containing 0.5µg/ml puromycin (concentration determined by a kill-curve experiment) to select for puromycin resistant (transfected) cells.

2.8 Genome editing methods

2.8.1 Preparation of guide RNA (gRNA) expression construct

Cloning of gRNA into the Addgene plasmid 48138 (pSpCas9[BB]-2A-GFP [pX458]) backbone vector which contains 2A-GFP fused to Cas9 was done by Philip Hublitz in the Genome Engineering Facility, Weatherall Institute of Molecular Medicine, University of Oxford (Figure 2.1). BbsI restriction enzyme was used in cloning. Addgene plasmid 48138 [pSpCas9(BB)-2A-GFP (PX458)] was a gift from Feng Zhang (Ran et al., 2013b). Top and bottom DNA oligos (10µl each of 100µM) (Table 2.6) were incubated at 99°C for 5minutes and annealed by slowly cooling down. Double stranded DNA was then precipitated by addition of 30µl 3M sodium acetate pH 5.6 and 300µl 100% ethanol. To clone gRNA into the pX458 backbone vector, ligation reaction was set up in 10µl final volume (NEB T4 ligase and associated 10x buffer) with 20ng

backbone vector and 25ng gRNA, incubated at room temperature for 1 hour and transformed in to *E.coli* bacteria. Bacteria were plated on to agar plates containing ampicillin and resistant colonies were picked and grown in liquid cultures. Plasmid DNA was isolated using QIAprep spin miniprep kit (Qiagen) according to manufacturer's instructions.

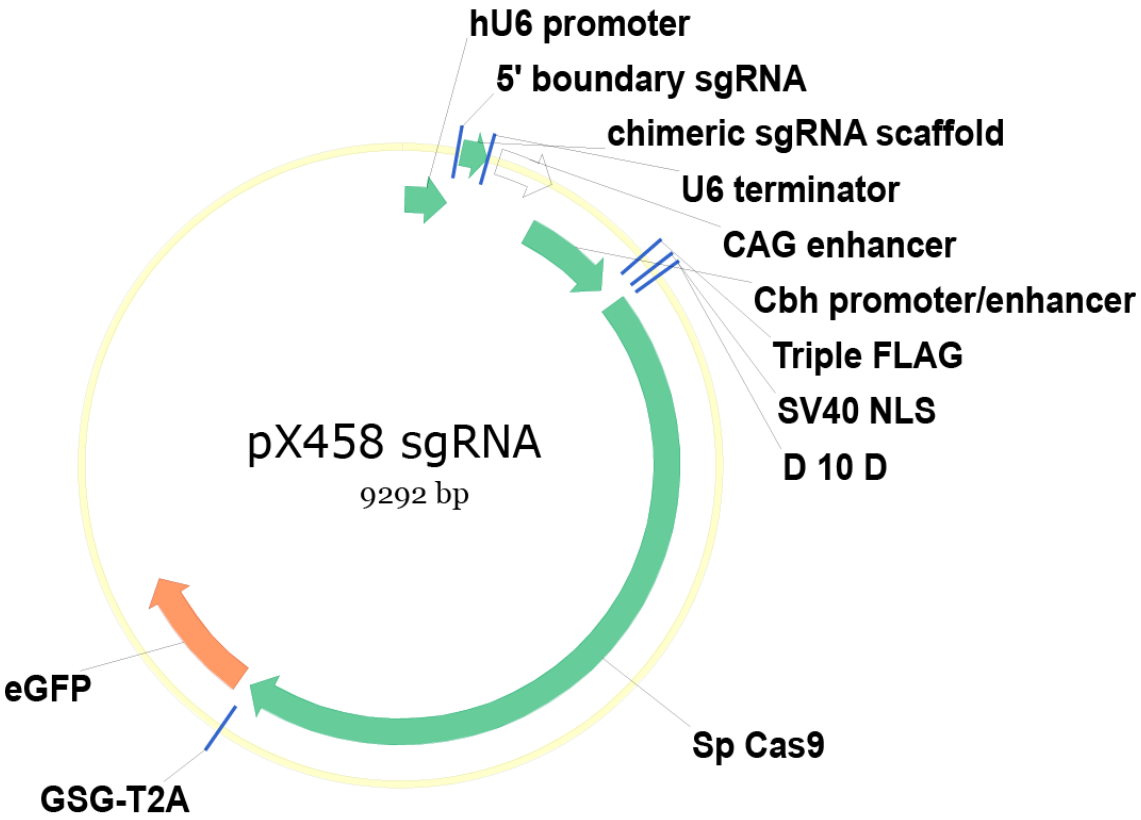


Figure 2.1 – Map of the plasmid construct with pSpCas9(BB)-2A-GFP backbone and gRNA. Map generated by Philip Hublitz, in the Genome Engineering Facility, Weatherall Institute of Molecular Medicine, University of Oxford. Abbreviations: sgRNA, single guide RNA; FLAG, FLAG expression system; NLS, Nuclear localisation signal; eGFP, enhanced green fluorescent protein.

Table 2.6 – gRNA sequences of CRISPR target sites within and around MCS-R2 enhancer region

CRISPR ID	*Target area*	*Target sequence*
1	5′ end	TCGACCCTCTGGAACCTAT
2	1st GATA1 binding site	CGACCCTCTGGAACCTATC
3	2nd GATA1 binding site	ACTTGAGGGAGCAGATAAC
4	1st NF-E2 binding site	AACCATGACTCAGTGCTTC
5	1st NF-E2 binding site	TGACTCAGTGCTTCTGGAG
6	2nd NF-E2 binding site	GACTGCTGAGTCATCCTGT
7	3′ end	CTCCTGTTTATCTGAGAGG
8	3′ end	GACCCAGACAGTAAATACG
9	5′ end	CTTCTGCAACCATGATGAC
10	5′ end	AGAGGGGCCCTCGACCCTC
11	2nd GATA1 binding site	CTTGAGGGAGCAGATAACT
12	3′ end	CCCTCCTGTTTATCTGAGA

2.8.2 Transfection

Transfection of CRISPR plasmid DNA was done using Amaxa human $CD34^+$ cell nucleofector kit (Lonza) and Amaxa nucleofector I device (Lonza). Culture medium containing $3x10^5$ to $1x10^6$ cells was spun at 1200rpm for 5 minutes. The cell pellet was resuspended in 100µl nucleofection solution and was nucleofected with 10µg CRISPR plasmid DNA in the cuvette provided using the U-08 program. Cell suspension was incubated at room temperature for 10 min in nucleofection solution before transferring it to antibiotic-free medium. Transfection efficiency was evaluated after 24 hours.

2.8.3 Cell sorting

Cell sorting was done in BD FACSAria Fusion cell sorter (BD Biosciences) by Kevin Clark of the flow cytometry facility, Weatherall Institute of Molecular Medicine, University of Oxford. Cells were spun at 1200rpm for 5 minutes 24 hours after nucleofection and resuspended in 300µl of medium. Green Fluorescent Protein (GFP) positive live cells were sorted as single cells in to wells of Terasaki multiwall plates which contained 20µl of culture medium. Remaining cells were sorted in bulk into 24-well plates.

Chapter 3: Miniature erythroid differentiation system for high-throughput and single-cell experiments

3.1 Introduction

Erythropoiesis is a complex biological process. At every stage of development in life, there is a continued need to replace senescent RBCs that are lost in the peripheral circulation. During human erythropoiesis, multi-potential HSCs differentiate through various progenitor stages to the earliest morphologically identifiable erythroid precursors, proerythroblasts, within the bone marrow (Higgs and Wood, 2011). Thereafter, during terminal differentiation, proerythroblasts become progressively smaller, undergo nuclear condensation and simultaneous haemoglobinisation to produce basophilic, polychromatic and orthochromatic erythroblasts in sequential order. Then, these nucleated RBCs enucleate to form reticulocytes and mature RBCs which are released into the circulation to fulfil their long journey through the blood vessels, which lasts approximately 120 days.

The development of *in vitro* erythroid differentiation models to recapitulate erythropoiesis is essential for understanding normal and abnormal erythropoiesis and to test external factors affecting the RBCs. A number of protocols have been developed for the production of erythroblasts in culture. The two main culture systems in use are those requiring semi solid or stromal cell support and liquid cultures that do not require supporting stroma (Dorn et al., 2008).

Since the initial description by Fibach and colleagues (Fibach et al., 1989), several liquid culture systems have been developed that do not require supporting stroma . Most of these systems utilize animal serum or conditioned medium hindering the exact characterization of the composition and thus their reproducibility (Fibach and Prus, 2005; Ronzoni et al., 2008). Furthermore, the initial systems used isolated MNCs as the starting cell source which is a mixture of monocytes, lymphocytes and only a small fraction of HSPCs (Fibach and Prus, 2005; Migliaccio et al., 2002). The resultant cultures have rarely been pure erythroid because of the contaminating cells such as macrophages. To overcome these drawbacks, new erythroid differentiation systems have been developed which use pure populations of $CD34^+$ HSPCs as starting material and fully defined serum-free growth media. These systems produce cultures with increased cell numbers, superior purity, enhanced terminal differentiation and greater reproducibility (Dorn et al., 2008; Giarratana et al., 2005; Neildez-Nguyen et al., 2002).

Current approaches for screening of drugs and chemical compounds heavily rely on high-throughput platforms, which are automated, convenient and efficient. In order to minimise the cost and the wastage of material, these platforms essentially require validated culture systems that can be used in small scale. More demandingly, newer molecular biology methods increasingly utilize single cell experiments, which require robust and reproducible systems to generate erythroid cells from single HSPCs in liquid cultures. In my project, I intend to screen libraries of small molecules in a high-throughput fashion, which requires a well characterised and validated miniature erythroid differentiation

system. An ideal erythroid differentiation model should generate a viable, synchronous, pure and plentiful population of erythroid cells in small tissue culture plates, with minimal non-erythroid contamination.

To date, only a handful of erythroid differentiation systems have been tested in small-scale experiments. Cheung and others described a small-scale serum-free liquid culture model of erythropoiesis to assess the effects of exogenous toxins (Cheung et al., 2007). However, this 14-day long, single-phase system produced limited number of cells that did not progress through terminal differentiation. The model system used by Bradner and colleagues, to screen for compounds that upregulate γ-globin expression, has only being characterised for the initial 10 days, at which point the cells were midway through differentiation (Bradner et al., 2010). The globin gene detection strategy used by Bradner and colleagues was a fluorescent microsphere detection system, which lacks sensitivity and involved several cumbersome steps. Even though this system was suitable to identify the marked up-regulation of the normally suppressed γ-globin expression, it is not sensitive enough to identify subtle changes in α-globin expression.

The objective of the first part of my project was to optimise, characterise and validate a miniature erythroid differentiation system and a globin gene expression detection protocol suitable for high-throughput and single-cell experiments. In this chapter, I have presented the characterisation of a serum-free erythroid differentiation system starting from $CD34^+$ HSPCs, suitable for small-scale and high-throughput experiments. I then validated this system's

ability with known reagents to detect changes in globin gene expression brought about by pharmacological and RNAi manipulation. Finally, with further modifications, I used this system to generate erythroid cells from sorted single $CD34^+$ HSPCs as micro cultures.

3.2 Results

3.2.1 Purification of CD34$^+$ HSPCs

The three main primary tissue sources of human CD34$^+$ HSPCs available for research are umbilical cord blood (UCB), adult bone marrow (BM) and adult peripheral blood (PB). This culture protocol was initially developed and optimised using CD34$^+$ cells from UCB which was then adapted for CD34$^+$ cells derived from adult PB. MNCs were first separated from UCB using density gradient centrifugation and the CD34$^+$ cell fraction was then sorted using magnetic-activated cell sorting (MACS). Purification using a single MACS column was suboptimal and the sorted cells were contaminated with large numbers of CD34$^-$ cells. However, the use of two sequential MACS columns improved the CD34$^+$ cell purity to 90% (Figure 3.1). The yield of CD34$^+$ cells from each UCB unit demonstrated considerable variation, probably due to the difference in input blood volumes and biological variation. The numbers of CD34$^+$ cells obtained by a representative sample of UCB units are shown in table 3.1.

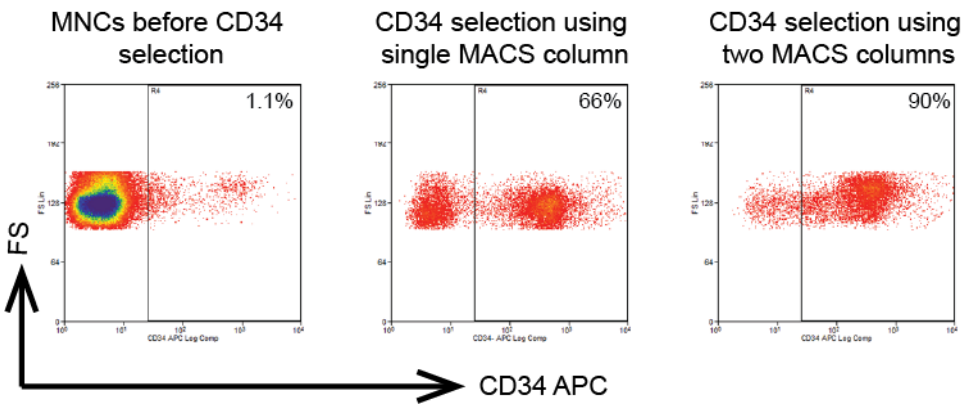


Figure 3.1 – Purity of CD34$^+$ cells after MACS separation. Representative flow cytometry analysis of cell before and after the purification using one or two MACS columns. Purity was 66% with a single MACS column and improved to 90% by using two sequential columns.

Table 3.1 – Yield of CD34$^+$ cells from umbilical cord blood units

Sample ID	Blood volume (ml)	Number of MNCs	Number of CD34$^+$ cells	CD34$^+$ cells / MNC (%)
1	105	216,000,000	1,800,000	0.83%
2	102	400,000,000	1,620,000	0.41%
3	84	400,000,000	2,880,000	0.72%
4	74	120,000,000	1,080,000	0.90%
5	80	113,400,000	1,036,800	0.91%
6	128	500,000,000	1,200,000	0.24%

3.2.2 Optimisation of the small-scale erythroid differentiation system

In our hands, the CD34$^+$ cells differentiated using the Bradner protocol (Bradner et al., 2010), lacked terminal differentiation and died prematurely after about 12 days. Therefore, I optimised this culture system by testing the effect of a number of additional growth factors and cytokines which have been used in other erythroid differentiation models. The addition of insulin like growth factor-1 and human plasma or increasing the erythropoietin concentration up to 10U/ml did not improve the outcome of the cells in culture (data not shown). However, the addition of iron-saturated holotransferrin to phase-2 medium promoted haemoglobinisation and improved terminal erythroid differentiation. Furthermore, the dilution of cultured cells by adding fresh medium every 2-3 days to maintain cell concentration below 2X10^6/ml improved cell survival. Therefore, these modifications were introduced to the erythroid differentiation system (as described in chapter 2; section 2.1.2); this protocol was characterised and validated in a 96-well format using a 200µL culture volume as shown in detail below.

3.2.3 Cell expansion and viability during erythroid differentiation

This optimized protocol resulted in high cell expansion. The cell growth was slow during the first 3 days, increased exponentially thereafter until day 18 to reach a plateau of a 500,000-fold mean amplification in UCB-derived cells (Figure 3.2A). A decrease in cell growth related to cell death was observed after day 18. The pattern of kinetics was comparable to the observations made with most other primary erythroid culture protocols (Giarratana et al., 2005; van den Akker et al., 2010). Adult $CD34^+$ cells demonstrated a lower expansion potential compared to the cells derived from UCB, again consistent with the previous reports (Giarratana et al., 2005).

During the first seven days of culture cell viability was over 95% and remained high (above 90%) for 14 days. Cell death increased thereafter and overall viability dropped to around 75% on day 21 (Figure 3.2B). No significant difference in viability was observed between the cells of UCB and adult origin.

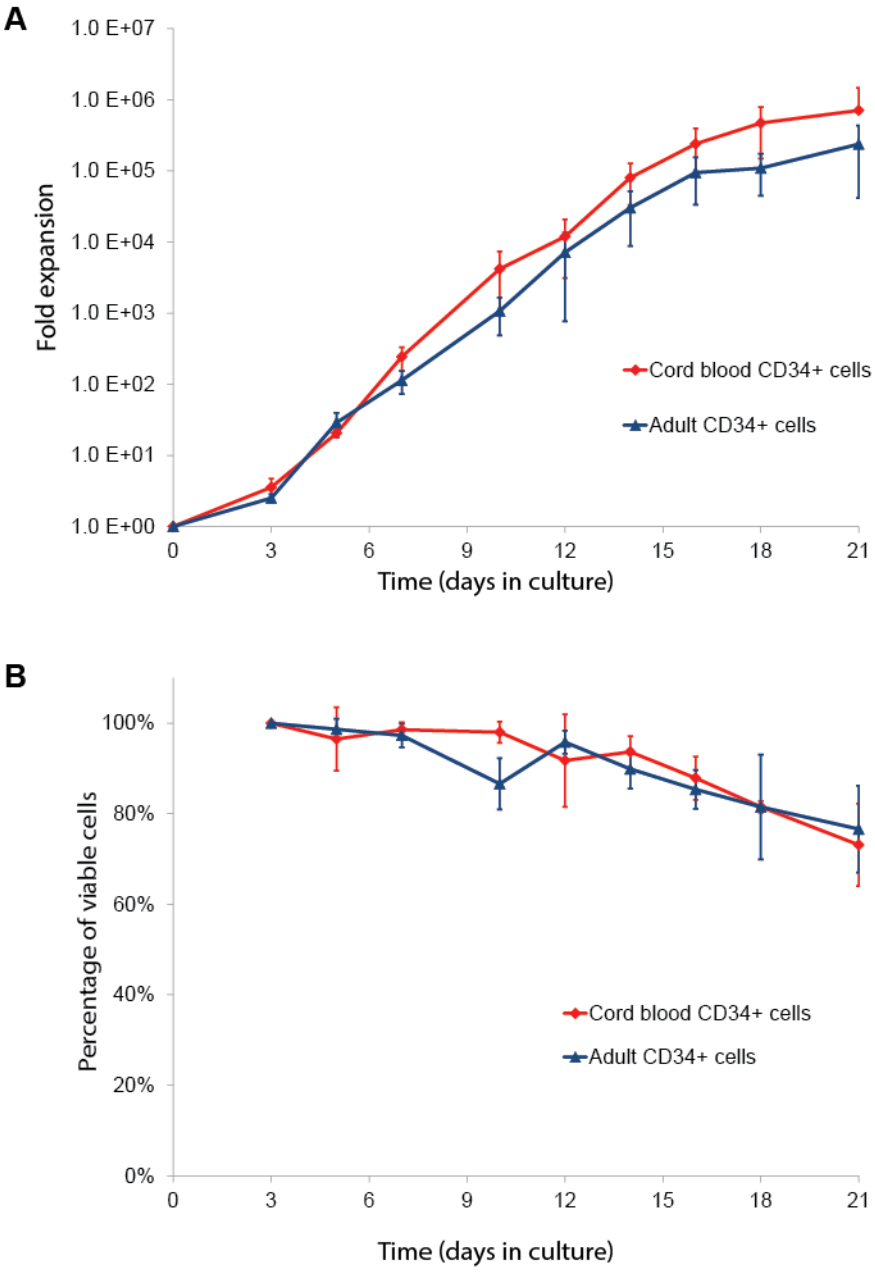


Figure 3.2 – Growth and viability of erythroid cells differentiated from human UCB and adult CD34$^+$ cells. (A) Mean fold expansion from the initial cell input at indicated time points is shown; error bars represent standard deviation (SD) from 3-4 independent biological repeats. (B) Percentage of viable cells assessed by trypan blue test at indicated time points. Mean of 3-4 independent biological repeats is shown; error bars represent SD.

3.2.4 Morphological characterisation of cells

Erythroid differentiation was morphologically evident from around day seven of the culture when a majority (80%) of the cells were proerythroblasts (Figure 3.3). After a few divisions these cells differentiated into smaller basophilic erythroblasts which constituted a majority (59%) of the cells by day 10. As the cells proceeded through differentiation, the cytoplasm gradually turned pink and the nucleus condensed to form polychromatic erythroblasts, which first appeared on day 14 and accounted for 40% of cells by day 18. The percentage of orthochromatic erythroblasts, the final nucleated stage of erythroblasts, increased gradually from day 14 (from 1% on day 14 to 55% on day 21). The non-erythroid cell percentages were very low from day seven until the end of culture.

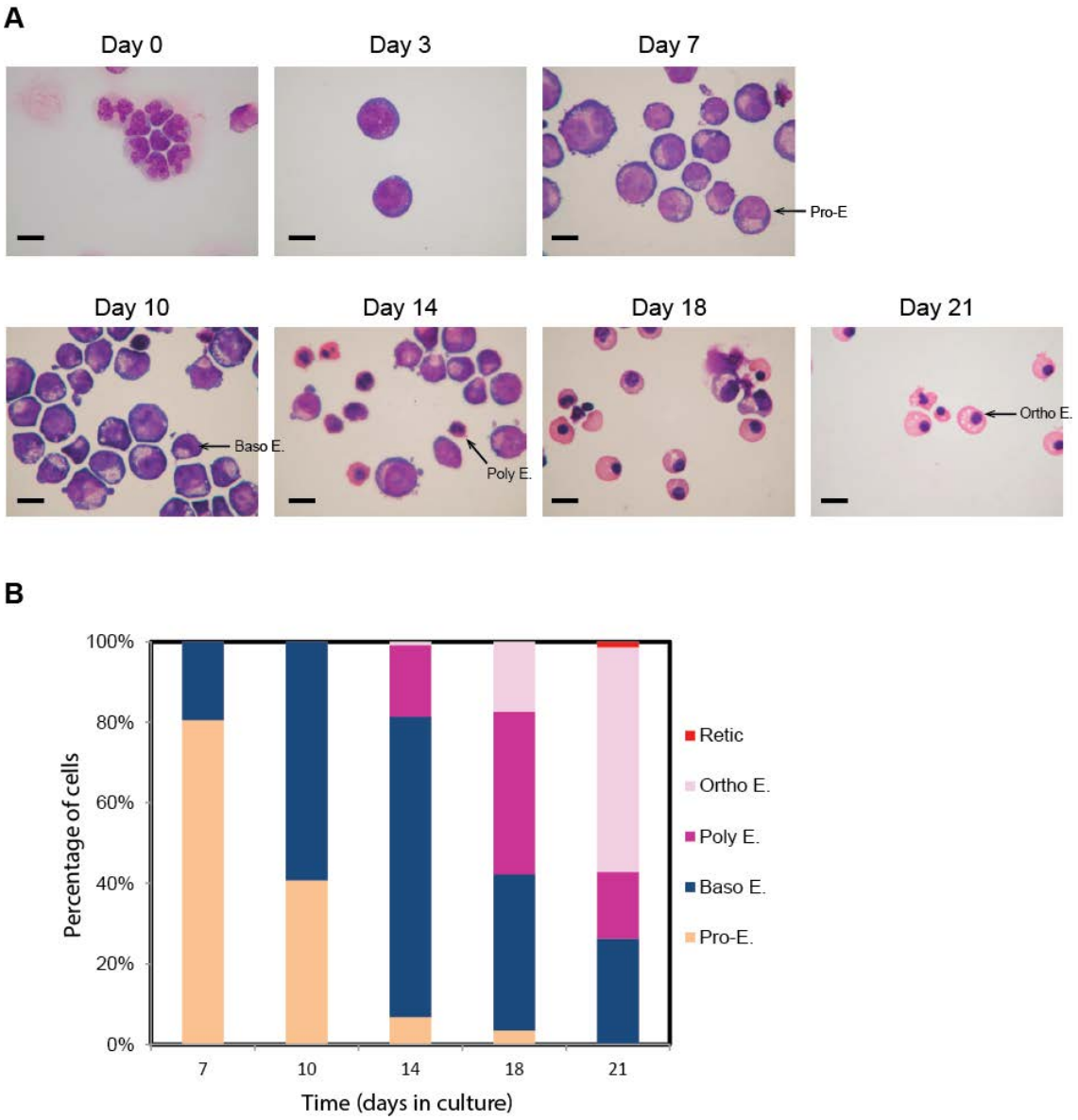


Figure 3.3 – Morphology of cells at different time points representing different stages of erythroid differentiation. (A) Representative cytospins of cells at different time points of culture stained with modified Wright stain demonstrating progression through stages of pro, basophilic and polychromatic to orthochromatic erythroblasts; scale bar – 10μm. (B) Percentages of morphologically distinct erythroid precursors at different time points of culture. At each stage a total of 250-300 cells were counted to analyse the morphological stage. Abbreviations: Pro-E, proerythroblasts; Baso E., basophilic erythroblast; Poly E., polychromatic erythroblast; Ortho E., orthochromatic erythroblast; Retic, reticulocyte.

3.2.5 Immunophenotypical characterisation of cells

Erythroid differentiation was further confirmed by immunophenotypical characterization of cell populations using antibodies to CD71, CD235a (also

known as glycophorin A [GPA]) and CD34 cell surface markers. CD71, the transferrin receptor, is expressed in erythroid progenitor cells from the colony forming unit-erythroid stage and is absent in reticulocytes and mature RBCs (Higgs and Wood, 2011). In contrast, CD235a is a marker of more mature erythroid cells which first appears at the proerythroblast stage and its expression increases throughout the rest of erythroid differentiation (Higgs and Wood, 2011). In this culture system, the expression of CD71 was first observed by flow cytometry on day three and was maximal by day 14 (Figure 3.4A&C). This was accompanied by the later appearance of CD235a, from day seven onwards. After day 14, CD71 expression declined, while the expression of CD235a remained high, resembling normal erythropoiesis. The very high expression of CD235a (98%) at day 21 suggests that at the terminal stages of differentiation the cells are purely erythroid.

The expression of CD34 was high (>90%) during the initial three days, and demonstrated a gradual decline thereafter to very low levels (0.2%) by day 14 (Figure 3.4B&C). This was expected as CD34 is a marker of early HSPCs and in the erythroid lineage is not expressed after the colony forming unit-erythroid stage (Higgs and Wood, 2011; Iskander et al., 2015). Furthermore, at later stages of culture, almost all the cells (over 99% at day 21) were negative for CD45. This, again, confirmed minimal non-erythroid cell contamination as CD45 is a pan-leukocyte antigen expressed in leukocytes and platelets but not in erythroid cells.

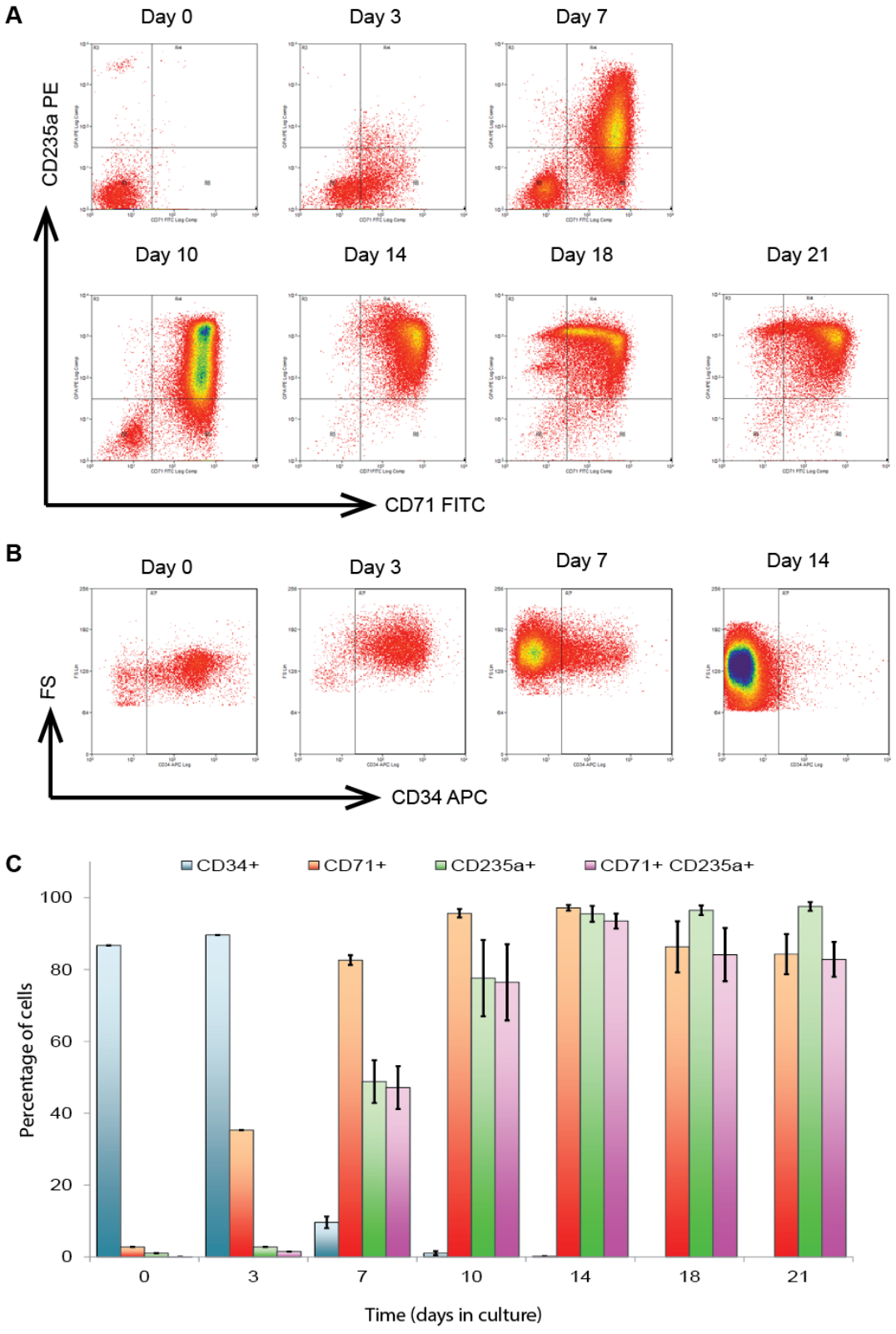


Figure 3.4 – Immunophenotypical characterisation of cells at different time points in culture. (A) Representative flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-CD235a antibodies (B) Representative flow cytometry plots of cells stained with APC-conjugated anti-CD34 antibody plotted against forward scatter. (C) Percentages of cells expressing CD34, CD71 and CD235a; means of 3 independent biological repeats are shown; error bars represent SD. Data on day 0 and 3 are from single experiments.

3.2.6 Gene expression during erythroid differentiation

Globin genes are essential markers of erythroid differentiation and during erythropoiesis their mRNA levels rise gradually from the proerythroblast stage. In this culture system, the relative expression levels of α-, β- and γ-globin were initially quantified using qRTPCR, which showed a steady increase in expression throughout the culture (Figure 3.5). In the erythroid cells differentiated from UCB $CD34^+$ cells, the expression levels of γ-globin were 10-fold higher than the expression levels of β-globin, throughout the culture. In contrast, the cells differentiated from adult $CD34^+$ cells predominantly expressed β-globin and the levels of γ-globin were comparatively low. However, the expression of α-globin was similar in both cell types. These findings clearly demonstrate that the globin gene expression in the erythroid cells in culture closely resembled the developmental background of the cells and adds more weight to the validity of the system.

Next, the gene expression was analysed using the more sensitive Nanostring technique, which directly quantifies the mRNA levels of each gene without reverse transcription or amplification. In cells differentiated from UCB $CD34^+$ cells, the expression of all globin genes increased throughout the culture. Although the expression levels of embryonic globin genes (ζ and ε), μ-globin and θ-globin increased during differentiation, they were expressed 10 to 100-fold lower than that of α, β and γ-globin, suggesting that their expression is relatively negligible (Figure 3.6A). Interestingly, the mRNA levels of γ-globin were roughly 10 folds higher than that of α-globin levels, throughout the culture. This was, however, surprising as it has been long believed that α-like and β-like

globin chains are produced in equal proportions in erythroid cells. If all of this mRNA is translated into protein, the cells would have a great excess of γ-globin, which would precipitate in cells as γ-globin tetramers, also known as haemoglobin Barts. This was not supported by the analysis of protein by isoelectric focussing (IEF) and CE-HPLC that did not reveal high levels of haemoglobin Barts, suggesting that the excess level of γ-globin mRNA is not translated into protein.

In addition to the globin genes, the expression levels of other erythroid specific genes (AHSP, KLF1 and FTH1) demonstrated a gradual rise throughout the culture (Figure 3.6B). In summary, these findings suggest that this differentiation system recapitulates erythroid gene expression at the molecular level and is a reasonable *in vitro* model in which to evaluate the effects of external factors on human globin gene expression.

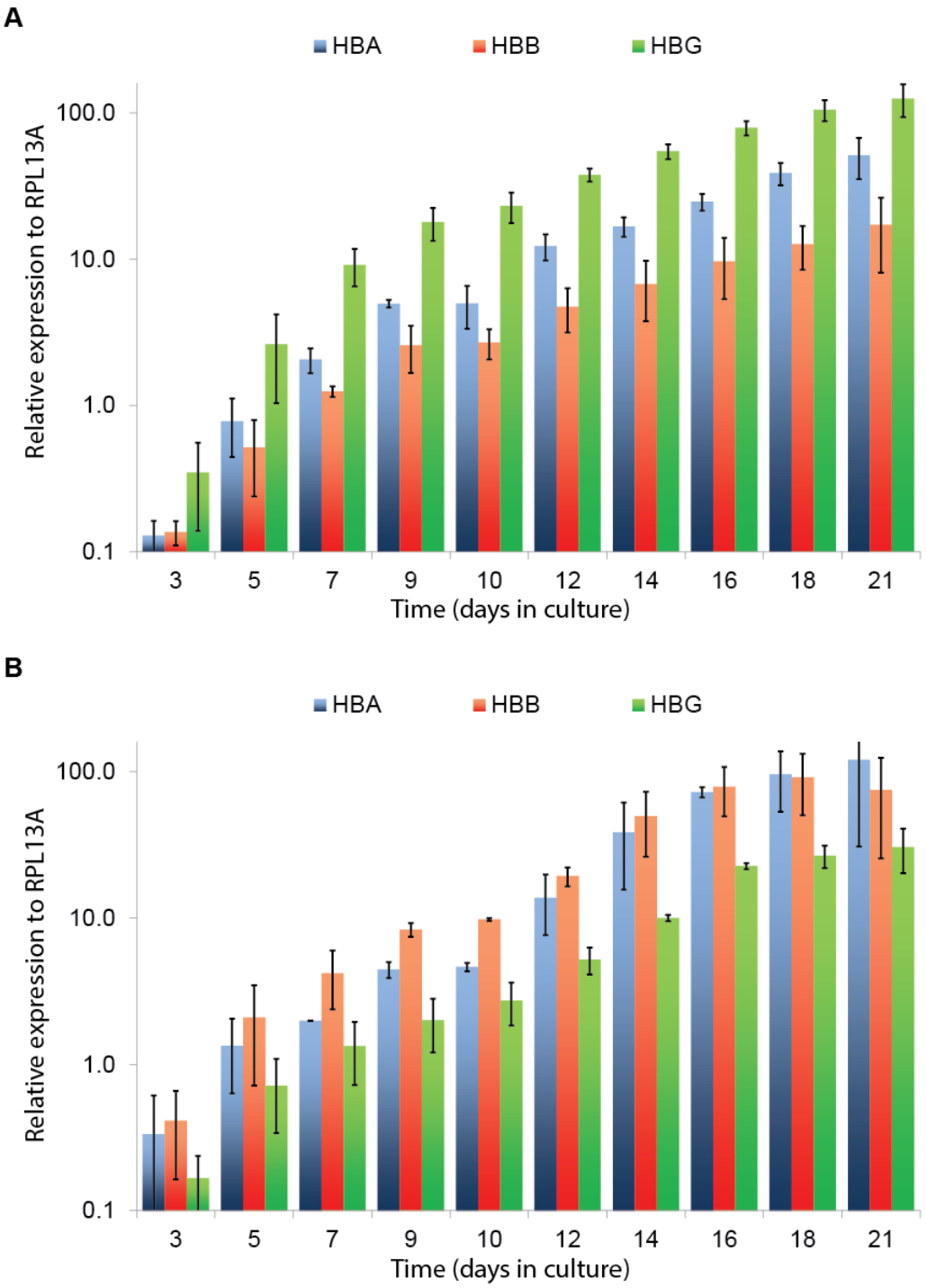


Figure 3.5 – Expression levels of α-, β- and γ-globin genes in cells at different time points in culture. Mean expression levels relative to RPL13A from 3 independent biological repeats quantified using qRTPCR is shown; error bars represent SD. (A) Cells differentiated from UCB derived $CD34^+$ cells. (B) Cells differentiated from adult $CD34^+$ cells. Abbreviations: HBA, α-globin; HBB, β-globin; HBG, γ-globin; RPL13A, ribosomal protein L13a.

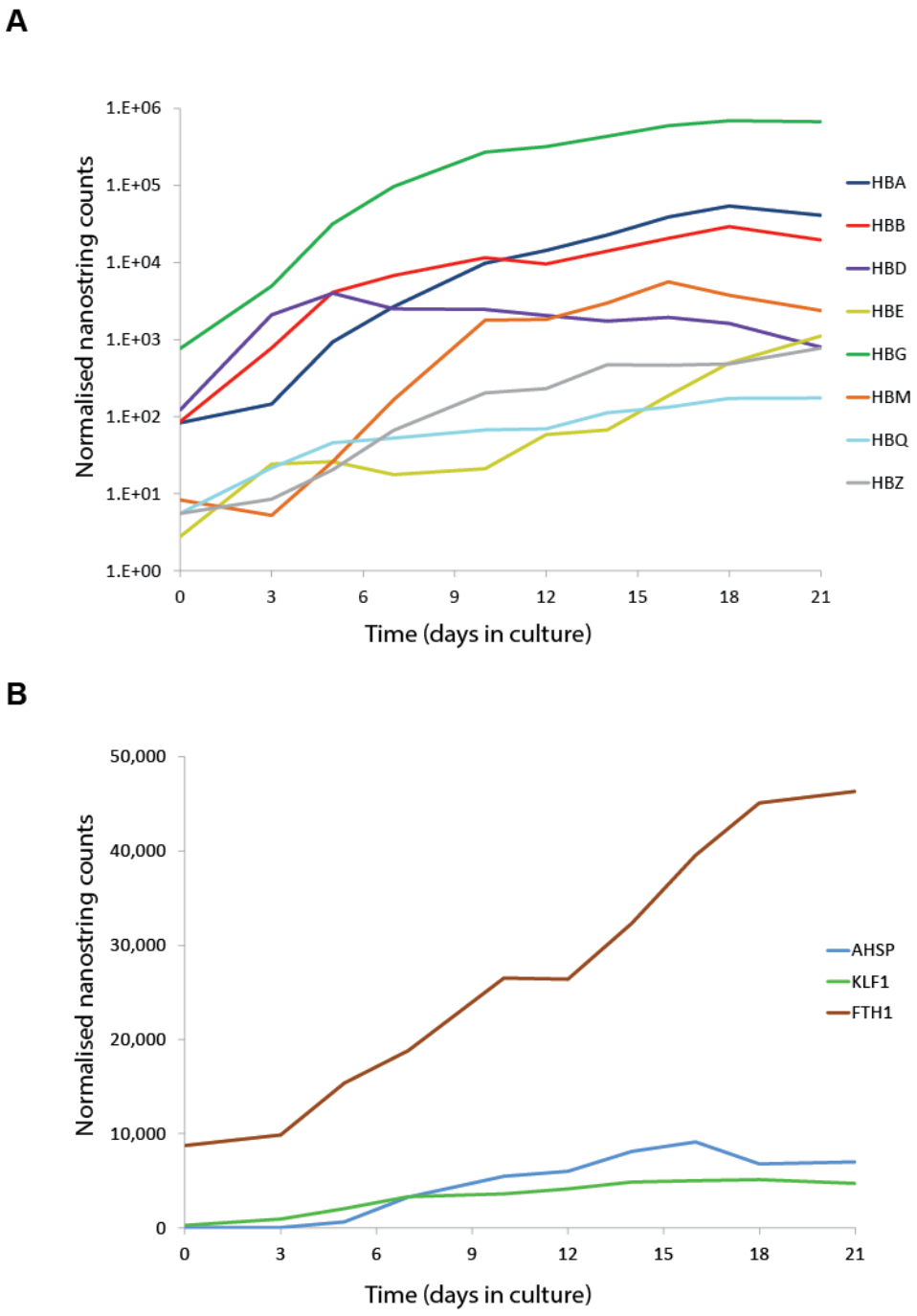


Figure 3.6 – Gene expression quantified by Nanostring at different time points in culture. Data is from erythroid cells differentiated from UCB CD34^{+} cells. (A) Expression of globin genes. (B) Expression of other erythroid specific genes. Abbreviations: HBA, α-globin; HBB, β-globin; HBD, δ-globin, HBE, ε-globin; HBG, γ-globin; HBM, μ-globin; HBQ, θ-globin; HBZ, ζ-globin; AHSP, α-haemoglobin stabilising protein; KLF1, Kruppel-like factor 1 and FTH1, Ferritin, Heavy Polypeptide 1.

3.2.7 Characterisation of haemoglobinisation and haemoglobin subtypes

Next, I characterised the cells at the protein level. Differentiating erythroid cells in culture produced haemoglobin from around day seven and the colour of the cell pellet changed gradually from white through pink to red at later stages. The haemoglobinisation was also assessed quantitatively using a benzidine wet stain which demonstrated a rise in the percentage of haemoglobinised cells from 7% on day 7 to over 90% on days 16 and 18 (Figure 3.7 A&B).

To characterize the relative proportions of haemoglobin subtypes that are produced, cultured erythroid cells were lysed and the extracted haemoglobin was analysed by IEF and CE-HPLC using VARIANTnbs system (Figure 3.7C&D) (This was done by Jennifer Eglinton of the Haematology Department, Oxford University Hospitals NHS Trust). Comparable results were observed with IEF and CE-HPLC which showed the cells derived from UCB predominantly produce foetal haemoglobin (HbF) and the cells differentiated from adult $CD34^+$ cells mostly produce HbA. The production of haemoglobin subtypes appropriate to the developmental stage is a further validation that the system recapitulates erythropoiesis at the molecular level in a developmental stage-specific manner. Also, no unaccounted bands in IEF or unusual elutes in CE-HPLC were seen confirming the absence of unusual haemoglobin subtypes.

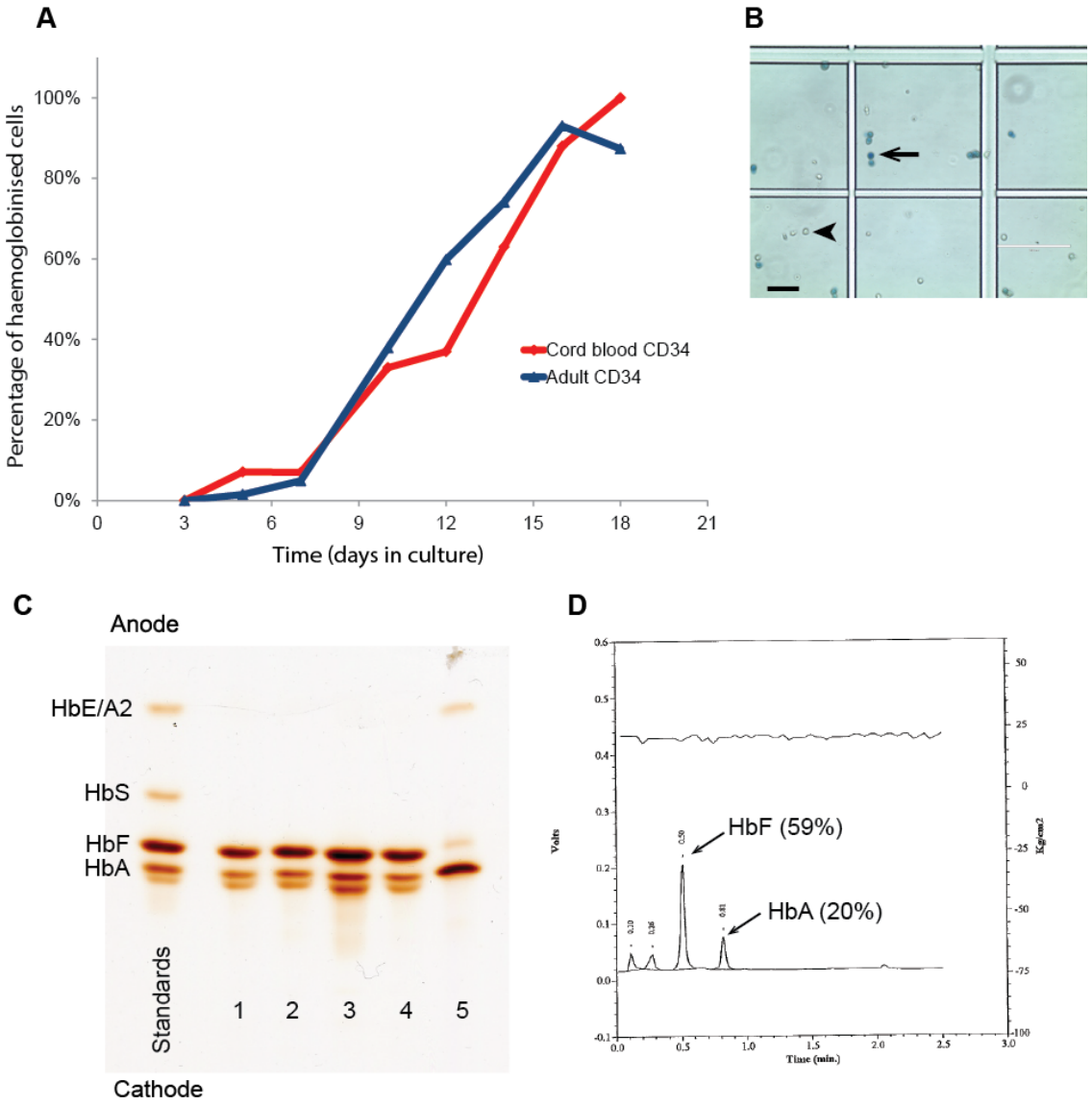


Figure 3.7 – Haemoglobinisation and characterisation of haemoglobin sub-types. (A) Percentage of haemoglobinised cells at different stages of culture analysed by benzidine staining. (B) Representative microscopy image of a haemocytometer containing cells incubated with benzidine. Haemoglobinised cells are stained blue (arrow) while non-haemoglobinsed cells are not stained (arrow head); black scale bar – 100 µm. (C) Haemoglobin subtypes of the erythroid cells differentiated from UCB and adult $CD34^+$ cells analysed by IEF. The samples were run against a commercial set of standards. Lanes 1 to 4 represent cells differentiated from UCB $CD34^+$ cells and the lane 5 represents cells generated from adult $CD34^+$ cells. (D) Representative CE-HPLC tracing of haemoglobin of cells differentiated from UCB CD34+ cells.

3.2.8 System validation

My aim in developing this erythroid differentiation system was to use it to screen libraries of small molecules to identify those down-regulating human α-globin expression. Therefore, the next step was to validate the culture system and the globin gene expression detection protocol to demonstrate that they are suitable and sensitive enough to detect changes in α-globin expression. However, there were no small molecules known to down-regulate α-globin expression to use as positive controls. Therefore, two independent approaches were used for validation, including a method that utilised the regulation of other globins.

Firstly, the effects of two small molecules already known to alter γ-globin expression were examined. Hydroxyurea and sodium butyrate are two potent compounds that upregulate γ-globin expression and has been used in patients to induce foetal haemoglobin (Musallam et al., 2013b). Incubation of cells at day seven of culture with either hydroxyurea or sodium butyrate demonstrated a dose-dependent up-regulation of γ-globin expression and an increase in the γ/β mRNA ratio, consistent with previous reports in *in vitro* culture models and in patients (Figure 3.8A&B) (Bradner et al., 2010; Italia et al., 2009).

Secondly, using RNAi, the cells were transfected with a pair of previously characterized siRNAs targeting α-globin mRNA (Figure 3.8C-E). This, as expected, resulted in selective knockdown of α-globin expression by up to 80% compared to a negative control siRNA, suggesting that the erythroid cells

generated in this culture system and the qRTPCR gene expression detection protocol behaves in a manner consistent with previous findings.

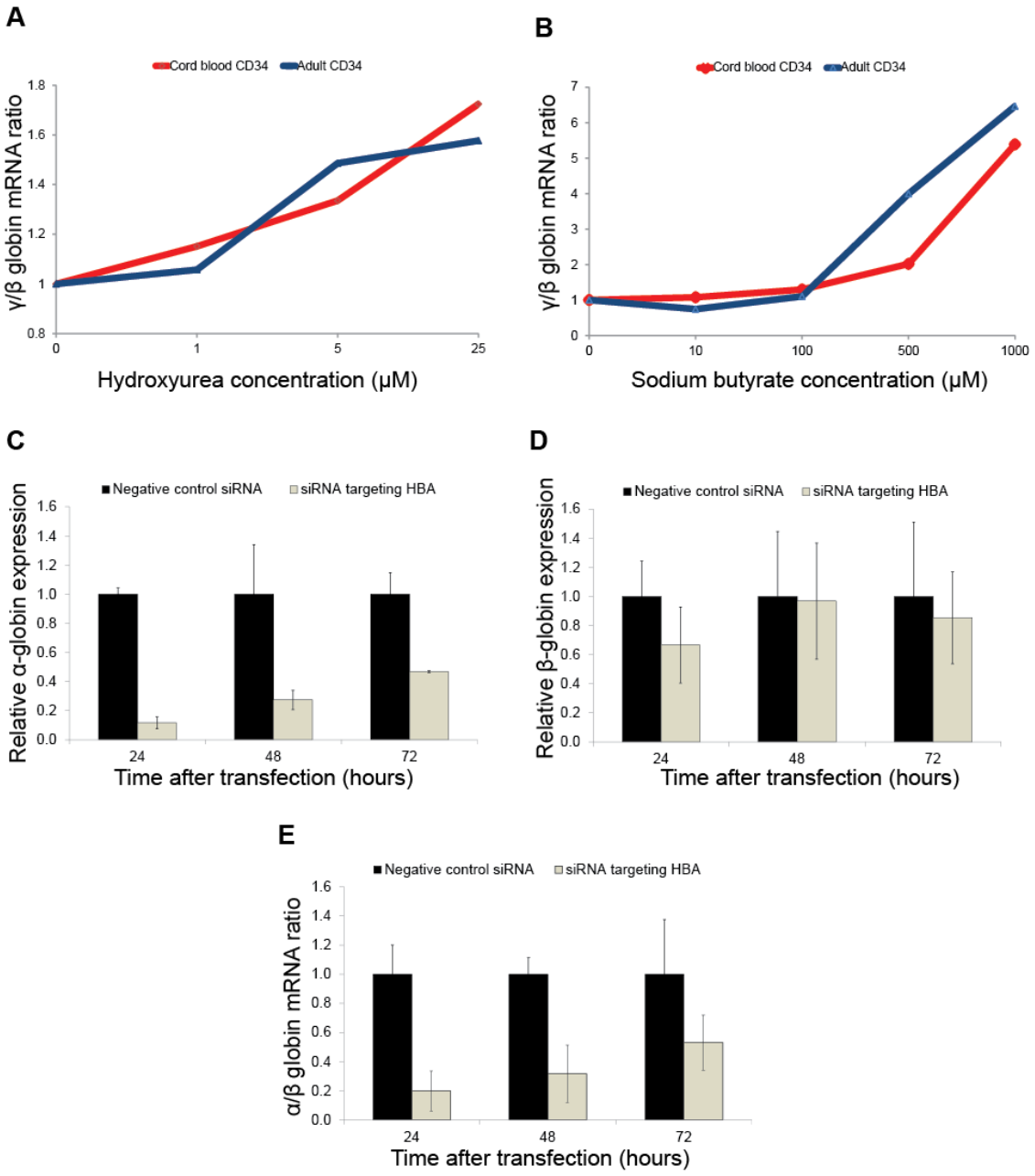


Figure 3.8 – Validation of the erythroid differentiation system and the globin gene expression detection protocol. (A) and (B) γ/β mRNA ratios of cultured erythroid cells after incubating with a dose range of hydroxyurea (A) and sodium butyrate (B). (C-D) α- and β-globin expression in erythroid cells transfected with a pair of siRNAs targeting human α-globin (HBA) and a negative control siRNA. Mean level of globin gene expression relative to the expression of RPL13A is shown; error bars represent SD from 2 independent biological repeats. (E) α/β globin mRNA ratio of erythroid cells transfected with a pair of siRNAs targeting human α-globin and a negative control siRNA; error bars represent SD from 2 independent biological repeats.

3.2.9 Erythroid differentiation from sorted single CD34$^+$ cells

Finally, attempts were made to further miniaturise the erythroid differentiation system to produce erythroid cells from a single CD34$^+$ cell. This was achieved using a slightly modified culture protocol (chapter 2; section 2.1.3). Here, single CD34$^+$ cells were sorted by flow cytometry into the wells of Terasaki multiwell tissue culture plates and cultured in 20µL medium for up to 14 days without addition or change of medium (Figure 3.9A&B). Clones appeared in about half of the wells in each plate (mean-50.7%; SD±7.0%; n=5) by day 7. All of them differentiated into erythroid colonies by day 14 and were visibly red at that point (Figure 3.9C). These clones demonstrated large cell expansions and grew in size to cover between half and the full surface area of the well (Figure 3.9D). Although slightly advanced, the erythroid differentiation progressed in the same sequence as the usual culture protocol and morphologically, a majority of the cells were orthochromatic erythroblasts by day 14 (Figure 3.9E).

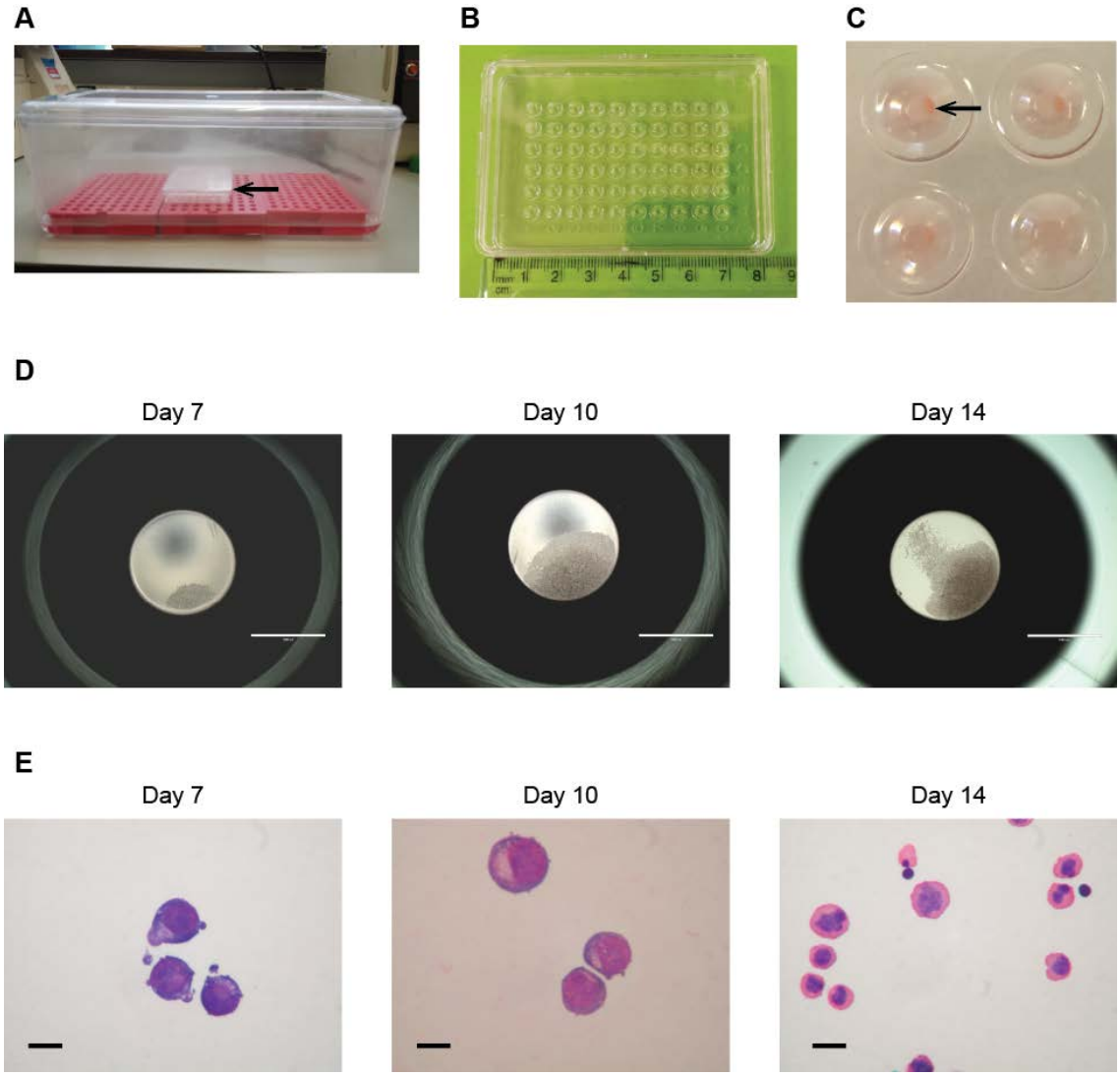


Figure 3.9 – Erythroid differentiation from single CD34$^+$ cells. (A) Photograph of a Terasaki multiwell tissue culture plate placed inside a humidified tissue culture box. (B) Close-up photograph of a Terasaki multiwell tissue culture plate (C) Clones grown and differentiated from single CD34$^+$ cells were visibly red (arrow) after 14 days suggesting erythroid differentiation and successful haemoglobinisation. (D) Representative microscopic images of a single well in the Terasaki plate demonstrating the size of the clone at different time points; white scale bar – 1mm. (E) Representative cytospins of the cells obtained from the clones and stained with modified Wright stain at different time points; black scale bar – 10µm.

3.3 Discussion

In order to perform high-throughput and single-cell experiments it has been a challenge to develop an *in vitro* system that recapitulates the events occurring in the bone marrow and that can generate large numbers of pure erythroid cells in small tissue culture wells. In this chapter, I have presented the optimisation, characterisation and validation of such a system.

This erythroid differentiation system has a number of advantages. It utilises human $CD34^+$ as the starting material, instead of MNCs which has been used in many previous liquid erythroid cell culture systems (Fibach and Prus, 2005; Migliaccio et al., 2002). MNCs are an admixture of monocytes, and lymphocytes with just a small fraction of HSPCs, and this results in impure populations of erythroid cells which are commonly contaminated with large number of macrophages. In contrast, CD34 is a cell surface marker expressed in HSCs and progenitor cells and has been used as a marker of transplantable HSCs in human bone marrow transplantations for many years (Dzierzak, 2011). Since $CD34^+$ cells are enriched for HSPC the purity of erythroid cells generated from my culture system was high throughout. Furthermore, erythroid differentiation occurred with reasonable synchrony, when characterised using CD71 and CD235a, the two most commonly used cell surface marker proteins in human erythroid differentiation staging. Similarly, this system does not use animal or human serum or other conditioned medium, but instead uses a fully characterised serum-free culture medium to induce erythroid differentiation, which is a huge benefit as it greatly improves reproducibility (Cheung et al., 2007). Animal serum is a collection of unknown growth factors and cytokines

which may vary in its composition and its absence facilitates analysing the effects of external factors and small molecules on erythroid cells by minimising the unknown variables which might perturb the system.

Another important feature of this culture system is its potential to generate a large number of cells. The mean fold expansion (500,000-fold) was far in excess of what has been demonstrated by most of the previous liquid culture systems which reported expansion rates less than 1000-fold (Dorn et al., 2008; Panzenbock et al., 1998), but was similar to the three-step stimulation protocol developed by Neildez-Nguyen et al (Neildez-Nguyen et al., 2002). However, in small-scale studies, Cheung et al only achieved 3000-fold expansion, 100-fold less than we report (Cheung et al., 2007). In summary, the erythroid differentiation system presented in this chapter generates a large number of erythroid cells from $CD34^+$ human HSPCs in 96-well plates, functions in a fully characterized serum-free medium, has minimal non-erythroid contamination and recapitulates *in vivo* erythropoiesis faithfully.

One limitation of this system is the low frequency of enucleation and differentiation into reticulocytes seen in the terminal stages. Interaction of erythroid cells with other cells like macrophages has been shown to induce enucleation (Giarratana et al., 2005) and the frequency of enucleation has always been low in most liquid culture systems that do not have supporting stromal cells (Neildez-Nguyen et al., 2002). Nonetheless, the lack of enucleation does not interfere with the assays for gene expression during

erythropoiesis and in fact an intact nucleus is required for changes in gene expression.

Another interesting, but rather unusual, finding was that the erythroid cells generated from UCB $CD34^+$ cells demonstrated very high expression of γ-globin at every stage of differentiation, not only compared to β-globin, but also relative to α-globin. Although, many erythroid differentiation systems are known to produce high levels of γ-globin and foetal haemoglobin due to culture conditions (Li et al., 2012; Pope et al., 2000), to our knowledge this is the first time that the level of γ-globin has been compared to the level of α-globin. This was possible as we used the more sensitive Nanostring technique for gene expression analysis, which provided an absolute quantification of the mRNA molecules in each sample, rather than the traditional less sensitive and relative quantification methods (Geiss et al., 2008). The extremely high amounts of γ-globin mRNA, if translated directly into protein, in the absence of α-like globins, should produce γ-globin tetramers known as haemoglobin Barts. Haemoglobin Barts is detectable by IEF and HPLC as described in a previous study which confirmed the findings also by mass spectrometry (Viprakasit et al., 2014). However, the analysis of haemoglobin proteins from these cells did not reveal haemoglobin Barts or other unaccounted haemoglobin subtypes raising the possibility of an additional regulatory step at the level of translation at least *in vitro*. However this may be true *in vivo* as well. Efremov and others reported a great excess of γ-globin mRNA (measured as γ/[γ+β]) compared to the percentage of HbF in a group of patients with β-thalassaemia major (Efremov et al., 1994) and similar observations were made by Smetanina and colleagues,

suggesting less efficient translation of γ-globin mRNA into protein (Smetanina et al., 1997). We did not explore this any further, as it did not impose any barriers to our use of this assay to detect changes in α-globin expression, however it would be fascinating to explore why and how the γ-globin mRNA is not translated into protein. This also would have an impact on researches that study the effect of γ-globin induction by various mechanisms.

Once optimisation and characterisation of the erythroid differentiation system was done, the next question was to decide on the time point which is most appropriate to test the effect of factors affecting globin gene expression. Day seven was taken as the optimal time point for a number of reasons. Firstly, it appears as the most convenient time point because it corresponds with the change from phase-1 to phase-2 medium. In addition, on day seven, the cellular morphology was definitely erythroid and the system generated adequate number of cells (200-fold expansion) sufficient for many experiments with excellent viability (98%). Furthermore, to explore the changes of globin gene expression, the cells should not have large amounts of residual globin mRNA or protein that were produced before the time of intervention; before day seven the globin gene expression and haemoglobinisation were very low.

Validation of the erythroid differentiation system, along with the detection protocol for globin expression, was performed using pharmacological compounds known to alter globin gene expression and RNAi techniques. Both hydroxyurea and sodium butyrate upregulated γ-globin expression in a dose dependent manner similar to previous reports (Bradner et al., 2010; Letvin et al.,

1984; Liakopoulou et al., 1995) and a pair of siRNAs targeting α-globin demonstrated the expected down-regulation of α-globin expression suggesting the assay is a valid *in vitro* tool to assess the changes in globin gene expression in human erythroid cells.

The novel molecular biology techniques increasingly use single cell assays to minimise inter-cell variations, therefore methods of differentiating erythroid cells from a single HSPC are immensely valuable. Here, I have miniaturised the erythroid differentiation system to the single cell level in a 20μL volume. Large clones were generated from single HSPCs using this model and notably, they complete the differentiation into pure orthochromatic erythroblasts. Erythroid differentiation was more rapid than was seen in the 96-well plates despite using a low erythropoietin medium throughout. This could be due to increased cell-cell interactions in smaller space, presence of cytokines secreted by adjacent cells promoting differentiation in conditioned medium (as fresh medium was not added) or to the both. Nevertheless, this assay would undoubtedly add strength to research in single cell biology related to erythropoiesis.

In conclusion, I have characterised and validated a small-scale serum-free erythroid differentiation system generating large numbers of pure and synchronous erythroid cells from pools of $CD34^+$ HSPCs or single HSPCs. This system appears to recapitulate *in vivo* erythropoiesis in erythroid morphology, immunology and at the molecular level and was validated as a tool to assess the changes in globin gene expression.

Chapter 4: Targeted small molecule screen to identify human α-globin silencers

4.1 Introduction

The principal difference in the regulation of human α- and β- globin genes lies in their epigenetic environments. Although, both genes are expressed in the same cell type, in a similar developmental-stage-specific manner, the chromatin architecture at the two loci exhibit remarkable disparities (discussed in chapter 1; section 1.6) (Brown et al., 2001). This contrasting chromatin environment gives potential for the discovery of drugs and pathways which selectively control one locus without affecting the other (Higgs et al., 2005; Mettananda et al., 2015). In search for selective human α-globin silencers, we intended to utilise these differences to test a collection of small molecules which are epigenetically active.

The epigenetic mechanisms are increasingly of interest as drug targets (Helin and Dhanak, 2013). Many potent, selective and cell-active small molecules that inhibit known epigenetic pathways which include, DNA and histone methylation and histone acetylation are already available or being developed (Muller and Brown, 2012). A significant proportion of these have entered clinical trials and some have already been approved by the United States (US) Food and Drug Administration (FDA) to be used in patients (Ivanov et al., 2014).

The objective of this chapter was to screen an epigenetically active small molecules compound library to identified compounds that down-regulate α-globin expression in human erythroid cells. Using the already validated miniature erythroid differentiation system, I performed a medium-throughput screen of a compound library and examined the effects on human globin gene expression. This screen, not only identified compounds that down-regulate α-globin expression, but also provided useful insights into the regulation of other globin genes.

4.2 Results

4.2.1 Selection of the epigenetically active small molecule library

Many epigenetic inhibitor small molecule libraries compiled by various biotech and pharmaceutical companies are available on the market and a decision had to be made on which compound library to be used in the screen. We decided to use the epigenetic inhibitor library designed by the Structural Genome Consortium (SGC), Oxford, because it contained a group of carefully selected compounds, all of which have been previously demonstrated as cell-permeable. This library contained 37 compounds that inhibit a wide array of epigenetic pathways, and comprised licensed drugs developed specifically as epigenetic drugs, long-standing drugs recently discovered to have epigenetic inhibitor actions and novel inhibitor molecules with medicinal properties and was kindly donated by Udo Oppermann of the SGC, Oxford (Chapter 2; table 2.1).

4.2.2 Medium-throughput screen of the epigenetic inhibitor small molecule library

Using the validated miniature erythroid differentiation system described in chapter 3, human UCB $CD34^+$ cells were differentiated down the erythroid lineage. Small molecules were added to the liquid cultures on day seven (proerythroblast stage), at previously described concentrations as determined by SGC (Figure 4.1A). Cells were incubated with compounds for 72 hours and the gene expression readout was obtained using Fluidigm BioMark HD high throughput qRTPCR system (Figure 4.1B & Table 4.1). Five compounds which resulted in cytotoxicity in the first experiment were re-assessed at lower

concentrations ($1/5^{th}$) in the subsequent experiments. Three biological and technical replicates were undertaken and sodium butyrate (500µM) was used as the positive control to up-regulate γ-globin gene expression. Gene expression data obtained with the Fluidigm BioMark platform was validated by standard qRTPCR assays, which demonstrated a very high correlation between the Fluidigm and standard qRTPCR results ($r=0.92$, $p<0.0001$) (Figure 4.2).

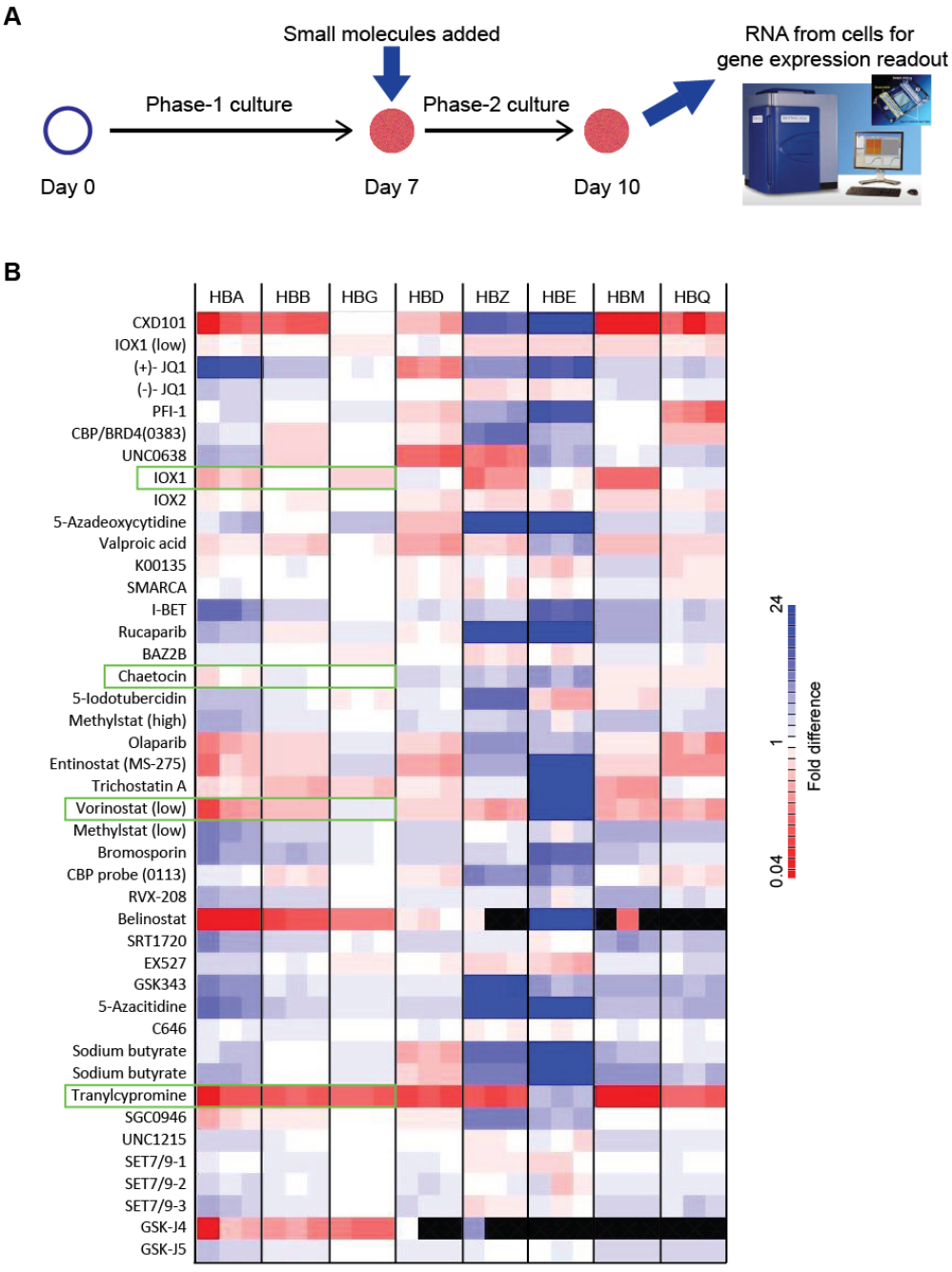


Figure 4.1 – Medium-throughput screen of the epigenetic inhibitor small molecule library. (A) A schematic diagram of the work flow of small molecule screen. (B) Representative heat map (one of 3 biological repeats) demonstrating fold differences of globin mRNA levels in erythroid cells treated with small molecules. The expression levels were normalized to the mean of four housekeeping genes (RPL13A, RPL18, GAPDH and FTH1) and referenced to the vehicle (DMSO) control. Each row represents a single compound and each column represent one of three technical repeats done for each globin gene with colours ranging from dark red (decreased expression) through white (no change) to blue (increased expression). The compounds with lowest α/β ratios are highlighted in green boxes.

Table 4.1 – Fold changes in the expression of globin genes in erythroid cells treated with small molecule inhibitors[1]

Small molecule inhibitor	*HBA*		*HBB*		*HBD*		*HBE*		*HBG*		*HBZ*	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
CXD101	0.6	(0.1)	0.2	(0.0)	0.2	(0.0)	3912	(35.9)	1.6	(0.4)	27.8	(29.5)
(+)- JQ1	11.6	(3.8)	1.8	(0.3)	0.2	(0.0)	9.4	(4.5)	1.3	(0.2)	2.8	(1.2)
(-)- JQ1	1.4	(0.3)	1.1	(0.2)	1.0	(0.1)	1.1	(0.4)	0.9	(0.1)	0.8	(0.2)
PFI-1	3.8	(3.9)	1.4	(0.5)	0.9	(0.4)	8.0	(4.9)	1.4	(0.3)	2.5	(0.6)
CBP/BRD4(0383)	0.7	(0.4)	0.6	(0.1)	0.8	(0.1)	2.5	(0.3)	1.0	(0.3)	4.3	(1.8)
UNC0638	1.7	(0.4)	0.6	(0.2)	0.2	(0.0)	3.3	(1.8)	1.2	(0.3)	0.3	(0.1)
IOX1	0.4	(0.1)	1.2	(0.3)	1.0	(0.1)	0.7	(0.5)	0.6	(0.1)	0.3	(0.1)
IOX2	0.9	(0.2)	0.9	(0.1)	0.8	(0.1)	0.9	(0.3)	1.0	(0.0)	0.7	(0.2)
5-Azadeoxycytidine	2.7	(0.9)	0.9	(0.2)	0.5	(0.1)	1378	(681)	2.7	(0.5)	89.1	(56.0)
Valproic acid	0.9	(0.6)	0.5	(0.1)	0.4	(0.1)	5.5	(2.8)	1.1	(0.3)	0.4	(0.1)
K00135	0.8	(0.3)	1.1	(0.0)	1.0	(0.3)	1.0	(0.7)	1.1	(0.1)	1.6	(0.4)
SMARCA	0.8	(0.1)	0.9	(0.1)	0.9	(0.1)	0.8	(0.1)	1.0	(0.1)	0.9	(0.2)
I-BET	4.3	(1.7)	1.7	(0.3)	1.2	(0.3)	6.7	(2.6)	1.4	(0.3)	1.7	(0.2)
Rucaparib	2.4	(1.1)	0.7	(0.0)	1.0	(0.3)	35.3	(14.5)	1.5	(0.2)	16.8	(7.5)
BAZ2B	1.2	(0.5)	1.0	(0.3)	1.0	(0.2)	1.1	(0.4)	1.0	(0.1)	0.8	(0.2)
Chaetocin	0.6	(0.1)	1.0	(0.3)	1.6	(0.2)	5.0	(2.1)	1.1	(0.0)	1.6	(0.7)
5-Iodotubercidin	1.1	(0.7)	1.0	(0.3)	0.9	(0.4)	0.7	(0.6)	0.9	(0.2)	11.7	(10.4)
Olaparib	0.9	(0.6)	0.8	(0.2)	0.9	(0.3)	2.5	(1.0)	1.2	(0.2)	4.4	(0.7)
Entinostat	0.8	(0.6)	0.7	(0.2)	0.4	(0.2)	105.0	(82.3)	1.8	(0.4)	3.6	(1.8)
Trichostatin A	0.5	(0.1)	0.3	(0.1)	0.5	(0.2)	56.5	(37.8)	0.6	(0.1)	1.2	(0.3)
Vorinostat	0.3	(0.2)	0.5	(0.0)	0.6	(0.1)	34.3	(12.6)	1.2	(0.1)	0.3	(0.0)
Methylstat	2.5	(0.7)	1.5	(0.2)	1.5	(0.2)	1.9	(0.7)	1.3	(0.1)	1.2	(0.4)
Bromosporin	3.1	(0.8)	2.1	(0.3)	1.5	(0.3)	4.8	(2.4)	1.4	(0.3)	1.3	(0.2)
CBP probe (0113)	1.3	(0.6)	0.6	(0.0)	0.8	(0.2)	2.9	(1.4)	0.9	(0.1)	3.4	(0.8)
RVX-208	1.6	(0.4)	1.5	(0.1)	1.4	(0.1)	2.1	(0.7)	1.2	(0.1)	1.0	(0.2)
SRT1720	2.0	(1.3)	1.7	(0.2)	1.9	(0.2)	1.1	(0.2)	1.5	(0.3)	1.6	(0.6)
EX527	1.2	(0.3)	1.0	(0.0)	1.0	(0.1)	1.0	(0.6)	0.9	(0.1)	0.7	(0.2)
GSK343	2.2	(1.6)	1.1	(0.3)	1.2	(0.2)	2.3	(1.0)	1.6	(0.1)	21.9	(3.6)
5-Azacitidine	2.9	(2.1)	1.8	(0.0)	1.9	(0.3)	58.3	(1.5)	1.6	(0.4)	11.9	(10.7)
C646	1.4	(0.7)	1.1	(0.1)	1.1	(0.1)	1.3	(0.9)	1.0	(0.1)	0.9	(0.1)
Sodium butyrate	2.2	(0.5)	1.1	(0.0)	0.5	(0.2)	47.4	(10.9)	2.2	(1.0)	11.3	(10.8)
Tranylcypromine	0.1	(0.0)	0.2	(0.1)	0.1	(0.1)	1.5	(0.8)	0.2	(0.1)	0.1	(0.0)
SGC0946	0.6	(0.2)	0.7	(0.2)	0.7	(0.1)	3.2	(2.0)	0.9	(0.0)	9.0	(8.2)
UNC1215	0.9	(0.7)	0.9	(0.3)	0.9	(0.3)	1.6	(0.6)	1.1	(0.2)	1.2	(0.4)
SET7/9-1	1.0	(0.1)	1.2	(0.1)	1.2	(0.1)	1.4	(0.6)	1.1	(0.2)	1.1	(0.6)
SET7/9-2	1.0	(0.4)	1.1	(0.2)	1.2	(0.1)	0.9	(0.2)	1.1	(0.1)	0.9	(0.2)
SET7/9-3	1.4	(0.5)	1.1	(0.2)	1.1	(0.2)	0.9	(0.1)	0.9	(0.0)	0.7	(0.1)

[1] Mean fold change of the expression levels normalised to four housekeeping genes (RPL13A, RPL18, GAPDH and FTH1) and referenced to the vehicle (DMSO) control is shown. SD is given in parenthesis. Abbreviations: HBA, α-globin; HBB, β-globin; HBD, δ-globin, HBE, ε-globin; HBG, γ-globin; HBZ, ζ-globin.

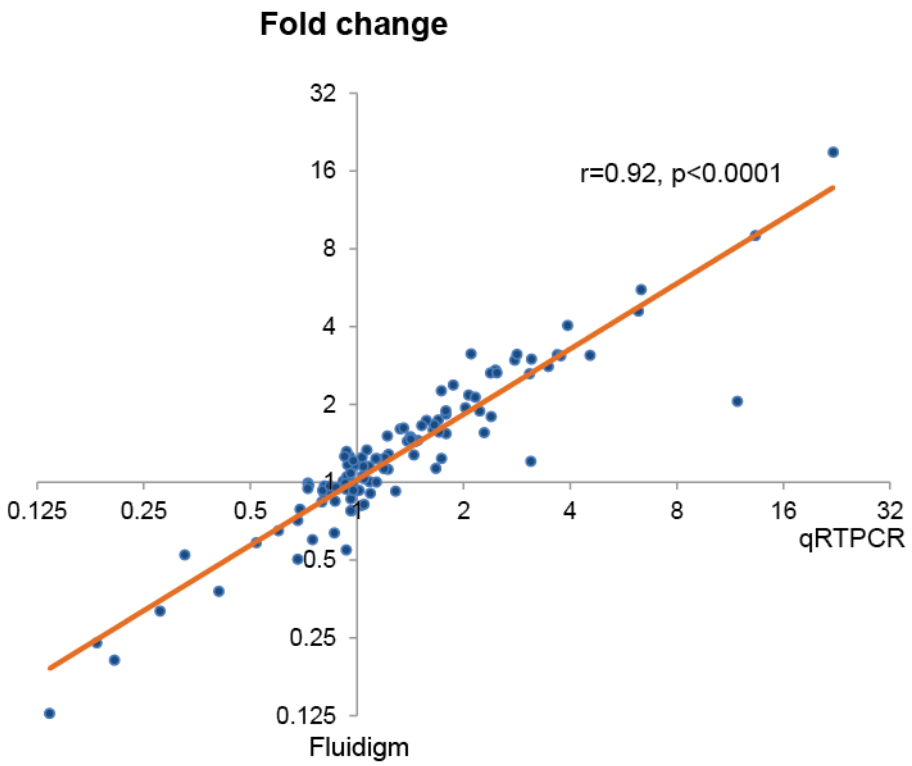


Figure 4.2 – Validation of Fluidigm BioMark data by standard qRTPCR. Fold changes of globin genes of the same cDNA samples used in the Fluidigm analysis were independently quantified by qRTPCR. The data indicates close correlation between the two data sets, and thus validates the Fluidigm analysis. Abbreviations: r, Pearson's correlation coefficient.

4.2.3 Identification of compounds which down-regulate α-globin expression

Down-regulation of α-globin expression without altering the β-globin expression was used as the primary screening measurement and α/β globin mRNA ratio of less than 0.75 was considered as the cut-off level to identify high-scoring compounds. Using these criteria, four compounds were identified (Figure 4.1B & 4.3). IOX1, a histone demethylase (KDM) inhibitor, was the highest scoring compound with the lowest α/β globin mRNA ratio. The second was a histone methyltransferase (HMT) inhibitor, chaetocin, the third was a histone deacetylase (HDAC) inhibitor, vorinostat and lastly was a lysine-specific histone

demethylase 1 (LSD1) inhibitor, tranylcypromine. Multiple confirmatory assays were then performed on these compounds and the detailed evaluation of IOX1 and vorinostat are discussed in chapters 5 and 6. Results of further experiments performed on tranylcypromine are presented in section 4.2.4. The HMT inhibitor, chaetocin was extremely toxic to the erythroid cells, caused decreased viability even at very low concentrations (10nM) and hence was eliminated as a useful candidate for further investigation.

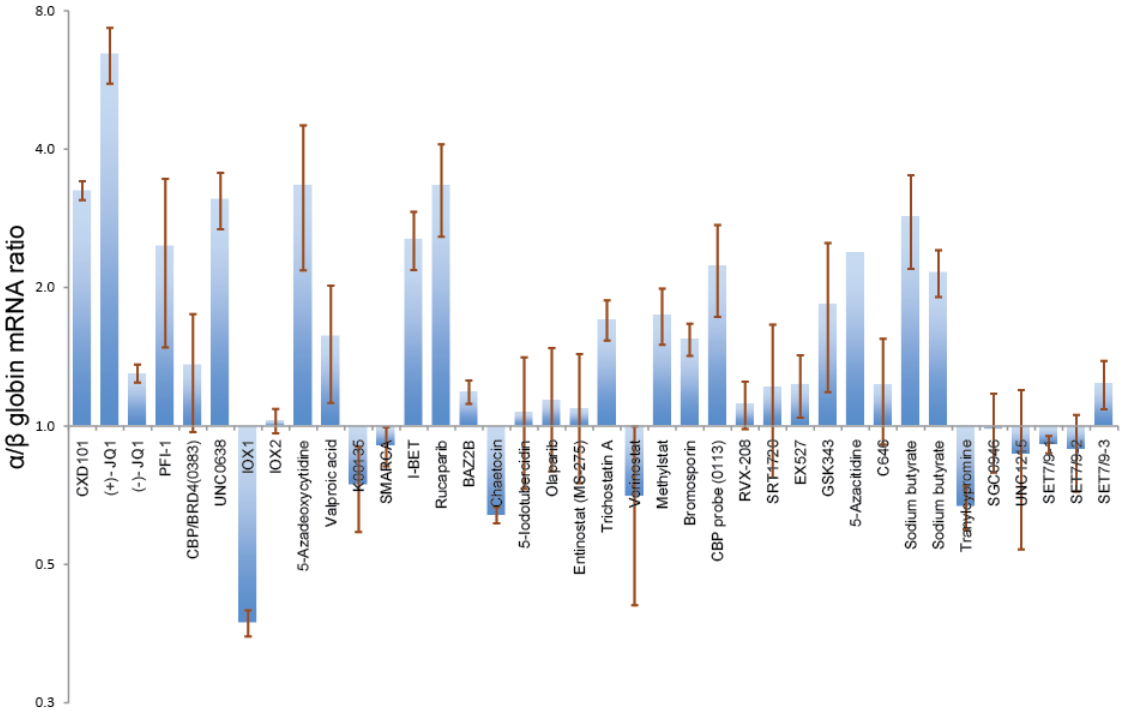


Figure 4.3 – α/β globin mRNA ratios in cells treated with small molecule inhibitors. Mean values of α/β globin mRNA ratio analysed by Fluidigm BioMark platform from 3 independent biological repeats is shown; error bars represent SD.

4.2.4 Tranylcypromine down-regulates α-globin expression by inhibiting erythroid differentiation

Subsequent experiments performed with tranylcypromine demonstrated that it down-regulates both β- and γ-globin expression, in addition to the α-globin (Figure 4.4A). Further analysis demonstrated that the treatment with tranylcypromine retarded erythroid differentiation markedly. On day ten of differentiation, cells treated with tranylcypromine were larger and morphologically immature and resembled proerythroblasts compared to the control cells which were differentiating towards basophilic erythroblasts (Figure 4.4B). This was confirmed by flow cytometry which showed only 8% of the cells treated with tranylcypromine co-expressed CD71 and CD235a compared to 82% of the control cells (Figure 4.4C). Interruption in differentiation was also consistent with a decreased number of haemoglobinised cells in the treatment group (Figure 4.4D).

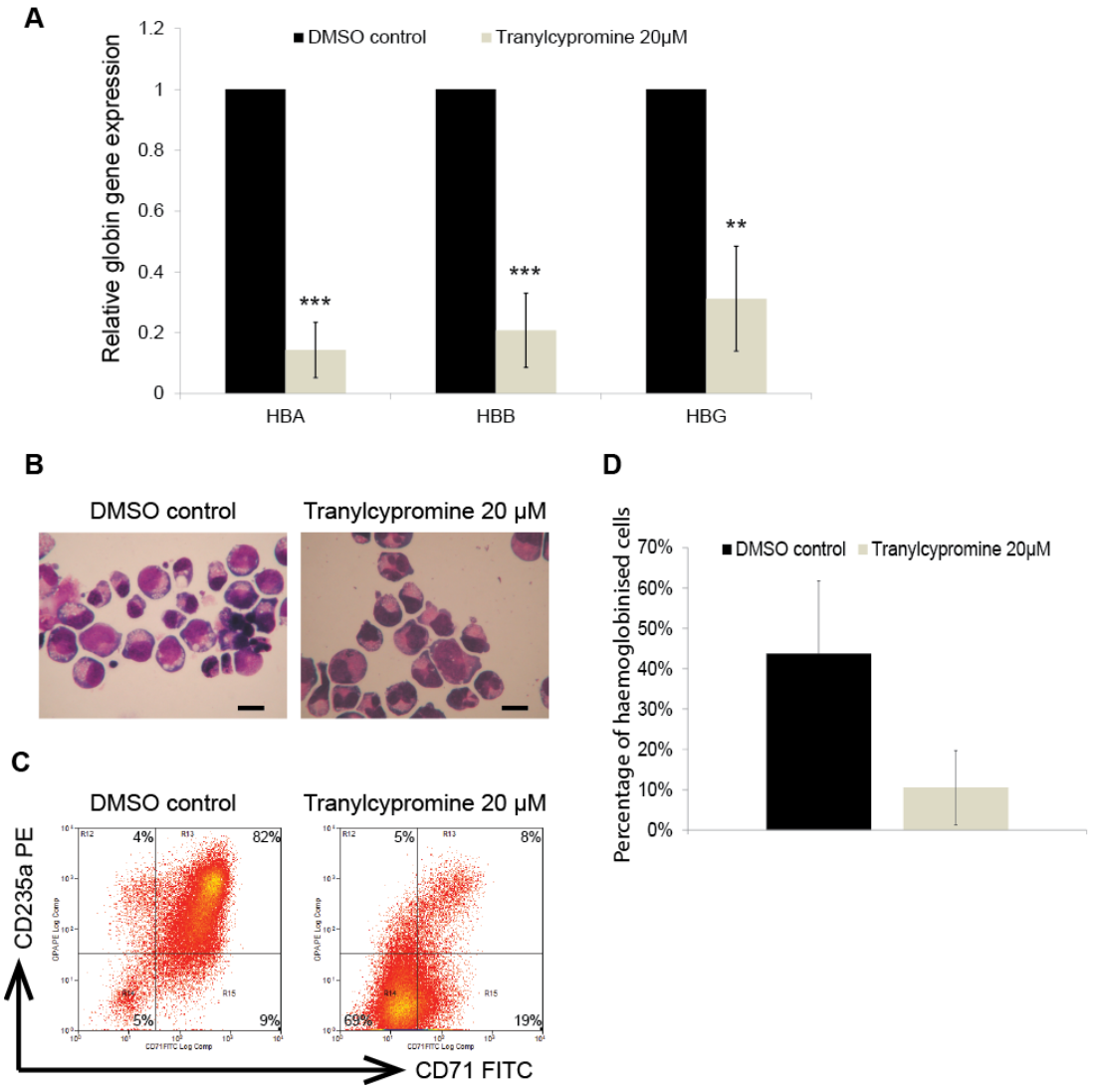


Figure 4.4 – Effects of tranylcypromine in erythroid cells. Erythroid cells treated for 72 hours with tranylcypromine 20μM and DMSO control were analysed on day 10 of erythroid differentiation. (A) Mean relative expression of α-, β- and γ-globin normalised to RPL13A and referenced to a DMSO control from 5 independent biological repeats; error bars represent SD; **p<0.01 and ***p<0.001 relative to the DMSO control. (B) Representative cytospins of cells stained by modified Wright stain; scale bar – 10μm. (C) Representative flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-CD235a antibodies. (D) Mean haemoglobinised cell percentages from 2 independent biological repeats are shown; error bars represent SD. Abbreviations: HBA, α-globin; HBB, β-globin.

4.2.5 HDAC inhibitors induce foetal and embryonic globin gene expression in the β-globin gene cluster

Because of the purposeful inclusion of all globin genes in the experimental design, in addition to data on the α-globin expression, the small molecule screen provided useful information on other globin genes. The small molecule library contained a number of compounds inhibiting HDAC enzymes, and the erythroid cells treated with several of these demonstrated significant increases in γ/β mRNA ratios (Figure 4.5). This was, however, not surprising as previous studies have shown that inhibition of HDAC1 and HDAC2 results in γ-globin induction (Bradner et al., 2010). Similarly, most of the HDAC inhibitors significantly increased ε/β ratio, confirming the vital role of HDAC in haemoglobin and globin gene switch in the β-globin cluster.

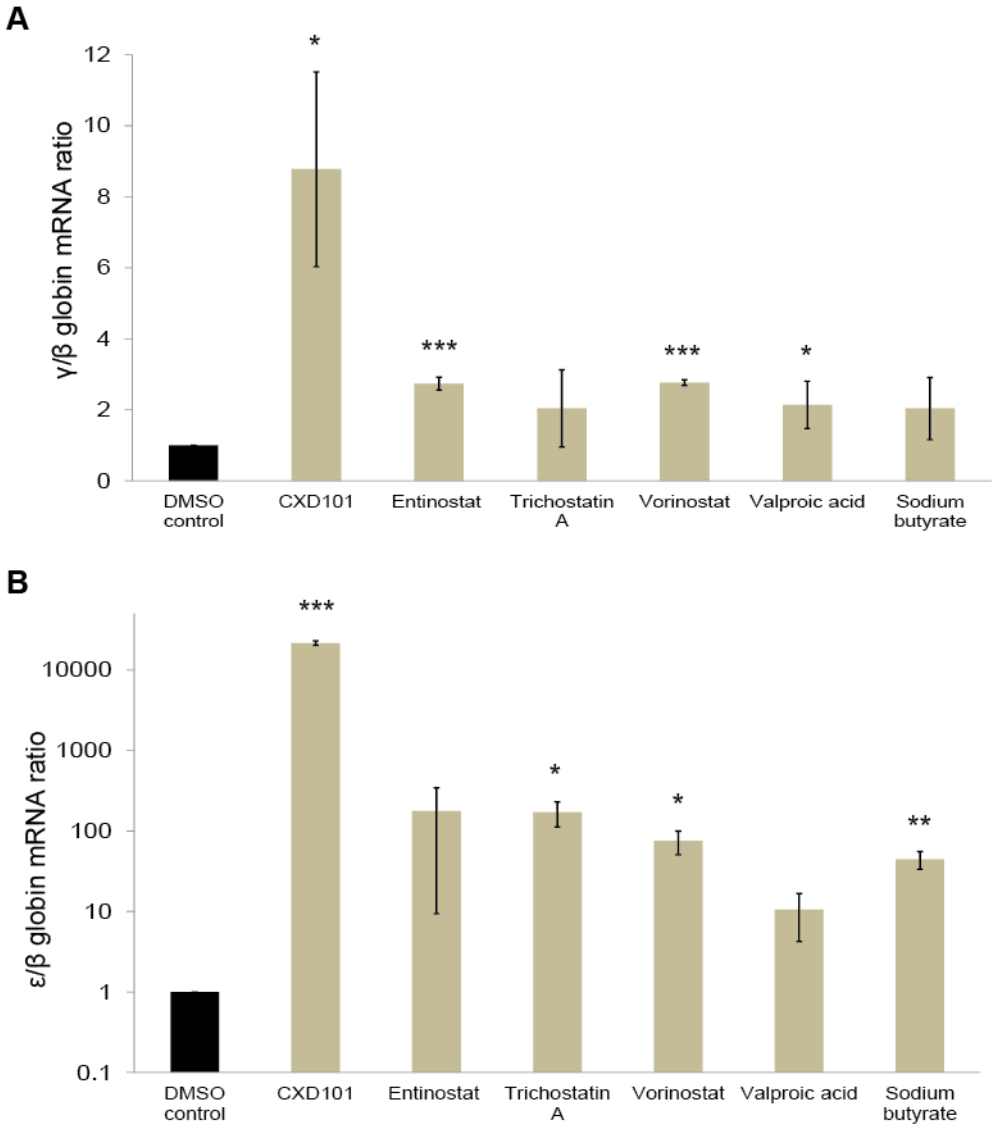


Figure 4.5 – Effects of HDAC inhibition on the β-globin cluster. γ/β (A) and ε/β (B) mRNA ratios of erythroid cells treated for 72 hours with HDAC inhibitors. Mean ratios referenced to DMSO control from 3 independent biological repeats are shown; error bars represent SD; *p<0.05, **p<0.01 and ***p<0.001.

4.2.6 DNMT inhibitor induces embryonic globin gene expression

The DNA methyltransferase (DNMT) inhibitor, 5-azadeoxycytidine (decitabine) demonstrated significant effects on the globin gene expression. Erythroid cells treated with this compound showed a marked increase in the expression of two embryonic globin genes, ζ and ε (Figure 4.6). This effect was seen at both α- and β-globin gene clusters suggesting an important role of DNMT in suppressing embryonic globin gene expression.

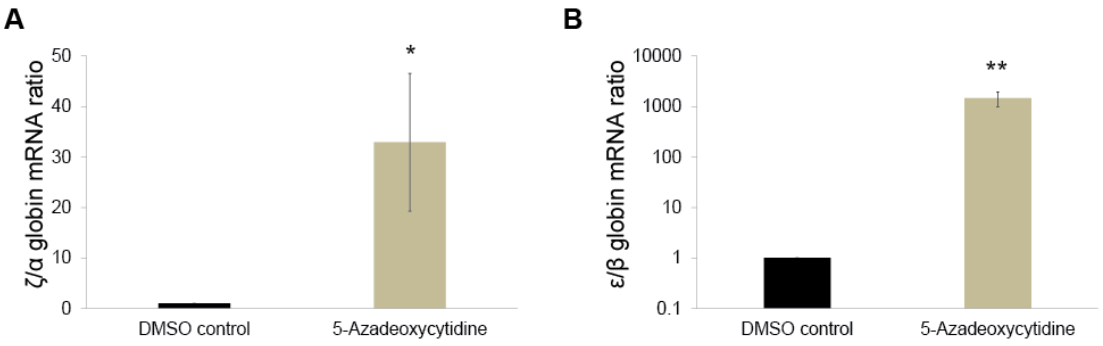


Figure 4.6 – Effect of DNMT inhibition on embryonic globin gene expression. ζ/α (A) and ε/β (B) mRNA ratios of erythroid cells treated with 5-azadeoxycytidine for 72 hours. Mean ratio referenced to a DMSO control from 3 independent biological repeats is shown; error bars represent SD; *p<0.05 and **p<0.01.

4.3 Discussion

The high-throughput screening (HTS) approaches of drug discovery involve laboratory automation and collection of a large amount of experimental data and test hundreds of thousands of compounds in a relatively short period of time to identify suitable candidates (Carnero, 2006). Although great success has been reported in some areas of drug development, the implementation of HTS is technologically demanding, labour intensive and vastly expensive hence, is not suitable for all types of research (Bleicher et al., 2003). In fact, it has been postulated to be a contributory factor to the decline in productivity in the pharmaceutical industry and stifling the creativity of drug discovery (Janzen, 2014; Macarron et al., 2011). In the search for small molecules selectively down-regulating α-globin expression, rather than performing a HTS of many thousands of compounds we decided to adopt a targeted approach. As stated before, the largest difference in the regulation of α- and β-globin genes, that has been identified thus far, lies in their epigenetic environment. Therefore our search was focused on altering this part of gene expression and I performed a medium-throughput screening of a carefully selected collection of epigenetically active and cell-permeable small molecules with favourable medicinal properties.

To mimic the natural process, in this screening, I used primary human erythroid cells, the exact cell type that ultimately requires to be targeted *in vivo*. Primary human erythroid cells were generated using the miniature differentiation system which I developed and the globin gene expression was assessed using the Fluidigm Biomark HD high throughput qRTPCR system which was validated independently. The compounds were added to the cells on day seven, when a

reasonable number of cells could be generated without compromising cell viability. Furthermore, the globin gene expression was minimal before day seven and it is important to treat the cells at a point where the background expression levels of globin genes are low. Taken together, the culture system and the Fluidigm platform aid the execution of this screen in an extremely efficient and cost effective fashion. This screen identified four compounds with desirable changes in human α-globin expression. One compound, the histone methyltransferase inhibitor, chaetocin, had to be dropped at the early stages due to its extremely cytotoxic properties. However, two others, IOX1 and vorinostat were studied in detail and are discussed in chapter 5 and 6 respectively.

The sensitivity of this assay was high compared to previously used assays for similar purposes. In a small molecule screen designed to detect γ-globin inducers, Bradner and colleagues used a ligation mediated amplification method combined with a flow cytometric detection system. This method involved several cumbersome steps including ligation of annealed probes, biotinylation, amplification followed by hybridisation of capture probes which were then quantified by flow cytometry. The complexity of the assay introduced several biases hindering validity and reproducibility (Bradner et al., 2010; Peck et al., 2006). Again to identify γ-globin inducers, Chan and others developed a fluorescent-based cellular reporter assay by inserting two reporter genes into the human β–globin locus. Although feasible in cell lines, this method is not amenable to primary human erythroid cells (Chan et al., 2012).

Tranylcypromine, a previously widely used anti-depressant was found to down-regulate α-globin expression during our screen. Tranylcypromine is well known for its monoamine oxidase inhibitor properties however, was recently reported to inhibit the epigenetic enzyme LSD1 (Lee et al., 2006). LSD1 is a demethylase enzyme, which removes methyl groups from mono- and dimethyl H3K4, both of which are activating epigenetic signatures (Shi, 2007; Shi et al., 2004). Interestingly, in a recent publication, tranylcypromine was also found to induce γ-globin mRNA (measured as γ/[γ+β] ratio) and foetal haemoglobin in human erythroid cells and was proposed as a therapy for β-thalassaemia (Shi et al., 2013). However, our data suggests that, tranylcypromine retards erythroid differentiation markedly as evidenced by immature cell morphology, lack of expression of erythroid specific cell surface proteins and reduced haemoglobinisation. The apparent knock down effect on α-globin was likely a result of a global reduction in globin gene expression due to poor erythroid differentiation. Our data on the γ/β mRNA ratio was consistent with the results presented by Shi and colleagues, but again the rise in the γ/β ratio is most likely a secondary effect of poor erythroid differentiation reducing expression of β-globin relative to γ-globin. Taken together, it is extremely unlikely that tranylcypromine is a clinically useful treatment for β-haemoglobinopathies.

Because of the ability of the Fluidigm platform to perform multiple qRTPCR reactions to test the expression of up to 96 genes in 96 samples, in a single experiment, we were able to test the effects of small molecules on the other globin genes. Consistent with the already published data (Bradner et al., 2010), most of the HDAC inhibitors increased γ/β and ε/β mRNA ratios confirming a

critical role of HDAC in silencing embryonic and foetal globin expression at the β-globin locus. In addition, the DNMT inhibitor, 5-azadeoxycytidine, significantly up-regulated the expression of the embryonic globin genes in both α- and β-globin gene clusters. Although, DNMT inhibition has been previously reported to induce γ- and ε-globin expression (Saunthararajah et al., 2003) (Roosjen et al., 2014), the reactivation of ζ-globin by this pathway has not been well documented. Our data suggests that DNMT is likely to play a significant role in the embryonic to adult haemoglobin switch at the α-globin locus, similar if not identical to the β-globin locus.

In conclusion, the data presented in this chapter confirms the successful utilisation of the miniature erythroid differentiation system and Fluidigm Biomark HD high throughput qRTPCR platform to perform a highly sensitive, targeted medium-throughput screen of an epigenetic inhibitor library. This screen, not only identified small molecules that produce desirable effects on the α-globin expression, but also provided useful insights into the regulation of globin genes, especially the embryonic globins.

Chapter 5: Selective silencing of human α-globin by histone demethylase inhibitor, IOX1

5.1 Introduction

Histone methylation is an important epigenetic mechanism regulating gene expression. Methylation of the lysine residues of histone proteins can either activate or repress transcription depending on the site of modification. Methylation at H3K4, H3K36 and H3K79 sites are associated with transcriptional activation, whereas methylation of H3K9 and H3K27 are signatures of transcriptional repression (Sims et al., 2003). Although, histone methylation was once considered to be irreversible, since the discovery of H3K4 demethylation by LSD1 (Shi et al., 2004), several histone demethylase (KDM) enzymes which actively demethylase these modifications have now been identified (Trojer and Reinberg, 2006). These KDM enzymes fall into two major families- the LSD family and the family of JumonjiC (JmjC) domain-containing demethylases (Hojfeldt et al., 2013). JmjC demethylases belong to the superfamily of 2-oxoglutarate (2OG) oxygenases which catalyses hydroxylation and demethylation reactions on multiple substrates including nucleic acids, proteins, and lipids. (Shi, 2007).

The targeted small molecule screen which was described in the previous chapter identified IOX1 as one of the promising candidate compounds which down-regulate α-globin expression. IOX1 (also known as 5-carboxy-8-hydroxyquinoline) is a cell-permeable and broad-spectrum JmjC histone

demethylase inhibitor which was first identified by a quantitative high-throughput screening performed using a fluorescence-based assay (King et al., 2010). Since the discovery of JmjC enzymes as mediators of histone demethylation (Tsukada et al., 2006), over 20 JmjC enzymes have been characterised with confirmed histone substrates and several more are being investigated to establish their endogenous targets (Johansson et al., 2014). IOX1 was found to inhibit many of these JmjC family KDM enzymes with variable affinity and in cellular assays it was active against both cytosolic and nuclear enzymes confirming its broad spectrum of activity and cellular and nuclear membrane permeability (Hopkinson et al., 2013).

In this chapter, I have further evaluated the effects of IOX1 on globin gene expression in human erythroid cells. Also, I examined the effects of IOX1 on cellular proliferation and viability as well as erythroid differentiation and transcriptome in detail. Furthermore, I have attempted to dissect out the exact mechanism by which IOX1 exerts the desirable changes in human globin gene expression.

5.2 Results

5.2.1 Down-regulation of α-globin by IOX1 is dose-dependent

During the small molecule screen using Fluidigm BioMark high-throughput qRTPCR platform, IOX1 (40μM) down-regulated α-globin expression by over 50% and the α/β globin mRNA ratio was 0.37 compared to a control. To study the effect of IOX1 on globin gene expression further, erythroid cells were incubated with a dose range of IOX1. In erythroid cells differentiated from $CD34^+$ cells from UCB, IOX1 demonstrated a dose-dependent decrease in α-globin expression with minimal or no effect on β-globin expression (Figure 5.1A). Similarly, α/β globin mRNA ratios were significantly reduced in the IOX1 treated cells in the dose range of 5-40μM (Figure 5.1B). This effect appears to plateau between the doses 20μM and 40μM suggesting 20μM dose may be as equally effective as 40μM.

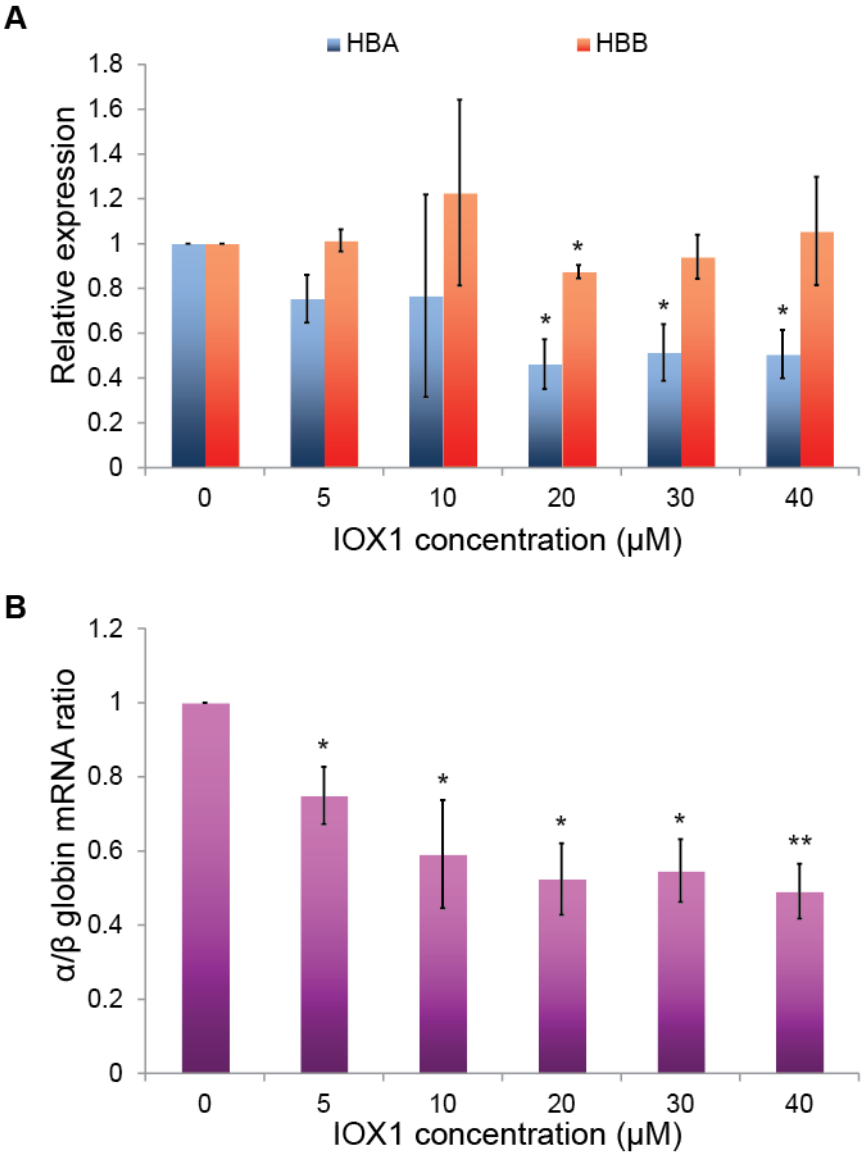


Figure 5.1 – Effects of IOX1 on α- and β-globin gene expression in human erythroid cells. Erythroid cells were incubated with a dose range (0μm-40μM) of IOX1 for 72 hours on day 7 of the differentiation and globin gene expressions were analysed using qRTPCR. (A) α- and β-globin expressions relative to RPL13A expression in cells differentiated from UCB. (B) α/β globin mRNA ratios in cells differentiated from UCB. In both graphs mean expression values from 3 independent biological repeats are shown; error bars represent SD; *p<0.05 and **p<0.01 relative to the 0μM concentration. Abbreviations: HBA, α-globin; HBB, β-globin.

5.2.2 Effect of IOX1 on globin mRNA is predominantly seen at the α-globin locus

The observations made by Fluidigm BioMark and qRTPCR were further verified using Nanostring technique (Nanostring Technologies). The mRNA levels of all

globin genes in erythroid cells treated with IOX1 were quantified using Nanostring nCounter digital analyser, which provides accurate and direct quantification of mRNA without reverse transcription or amplification. In erythroid cells, IOX1 significantly down-regulated α, γ, μ and ζ-globin expression, whereas the expression levels of β, δ and ε-globins were unaffected (Figure 5.2). Interestingly, with the exception of γ-globin, IOX1 down-regulated α- and other α-like globin genes (μ and ζ) situated at the α-globin locus in chromosome 16, whereas the expression levels of β-like globin genes (β, δ and ε) located at the β-globin locus in chromosome 11 were unaffected, suggesting that IOX1 acts selectively at the α-globin locus.

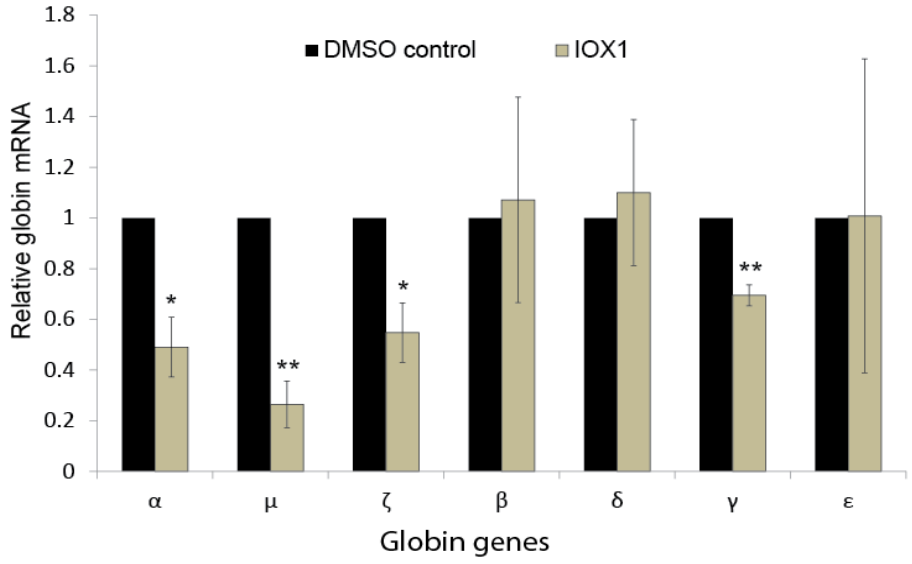


Figure 5.2 – Effects of IOX1 on globin gene expression in erythroid cell. Erythroid cells were incubated with IOX1 (40μM) for 72 hours on day 7 of the differentiation and the globin mRNA levels were quantified using Nanostring nCounter digital analyser. Mean Nanostring counts of globin genes normalized to the mean of multiple housekeeping genes (RPL13A, RPL18, GAPDH, PABPC1, CA2, FTH1, PAIP2 and LAPTM4A) from 3 independent biological repeats is shown; error bars represent SD; *p<0.05, **p<0.01 relative to DMSO control.

5.2.3 Effects of IOX1 on erythroid cell proliferation and viability

The next step was to assess the toxic and off-target effects of IOX1. Evaluation of cell proliferation by determining the number of cells at different time points of culture, revealed that treatment of IOX1 reduced cell expansion by about 40% (fold expansion dropped from 18-fold to 11-fold) at 40µM concentration in erythroid cells (Figure 5.3A). However, the percentages of viable cells determined by trypan blue test were unchanged in treatment and control groups (Figure 5.3B). This suggests that IOX1 has an *in vitro* inhibitory action on erythroid cell proliferation, although it does not adversely affect cellular viability.

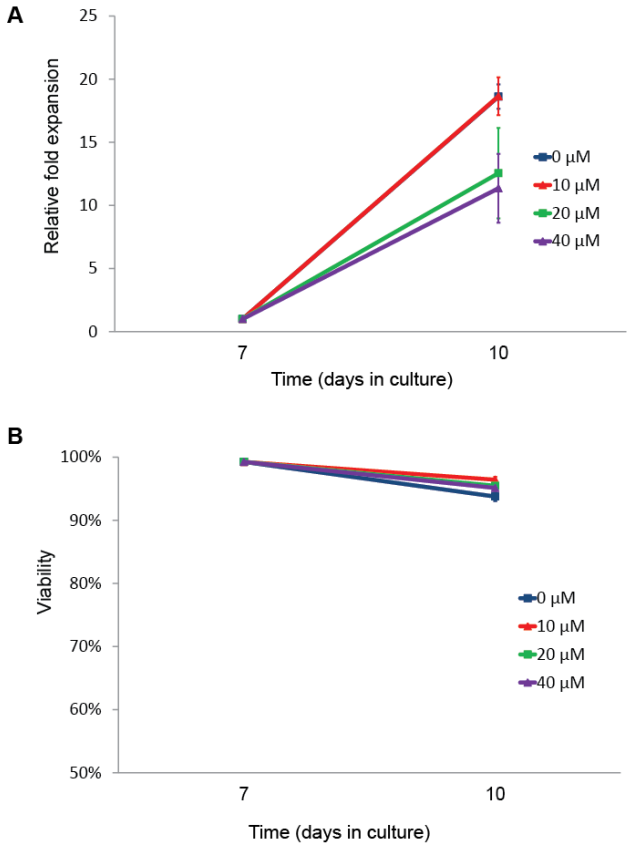


Figure 5.3 – Expansion and viability of erythroid cells treated with IOX1. Erythroid cells were incubated with a dose range (0µm-40µM) of IOX1 for 72 hours on day 7 of differentiation. The 0nM concentration refers to a DMSO control. (A) Cell growth shown as mean fold expansion relative to the number of cells on day 7 (B) Percentages of viable cells analysed by trypan blue test. In both figures mean values from 3 independent biological repeats are shown; error bars represent SD.

5.2.4 IOX1 does not alter erythroid differentiation

Morphological analysis by stained cytospins revealed that the erythroid cells treated with a dose-range of IOX1 were at a similar stage of differentiation to the untreated cells, suggesting that IOX1 treatment does not alter erythroid differentiation (Figure 5.4A). This was further confirmed immunologically by cell surface marker expression. Expression levels of CD71 and CD235a, the two most commonly used markers of erythroid differentiation, were unaltered in cells treated with IOX1 (Figure 5.4B). Percentages of double-positive cells for both these markers were similar in IOX1 treated and control cells confirming cells at similar stages of differentiation in the two groups (Figure 5.4C). The expression levels of CD34, a cell surface marker present in early progenitor cells were very low in both groups, further verifying the normal progression of erythroid differentiation of IOX1 treated cells (Figure 5.4D&E).

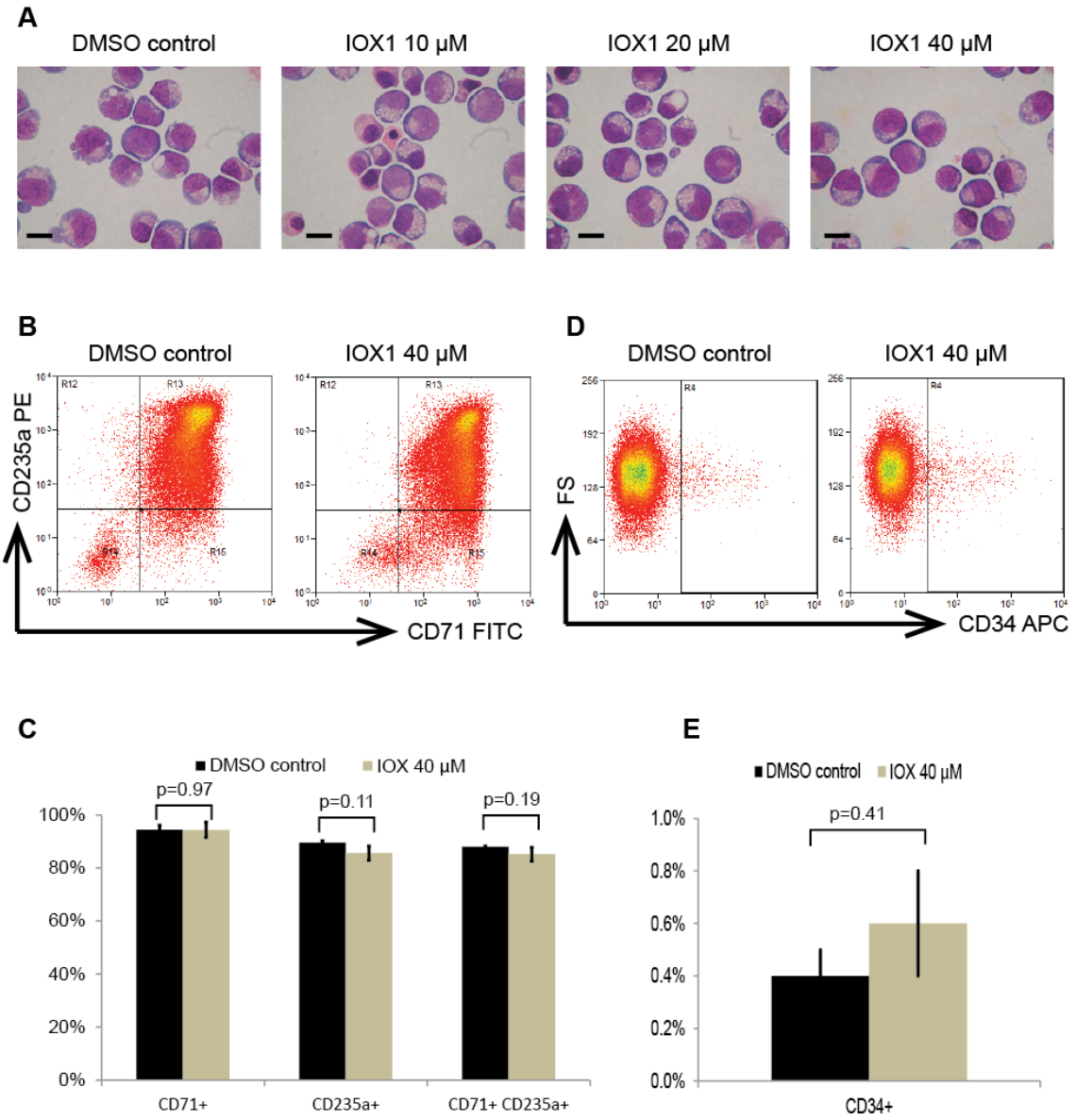


Figure 5.4 – Effect of IOX1 on erythroid differentiation. Erythroid cells were incubated with a dose range (0μm-40μM) of IOX1 for 72 hours on day 7 of differentiation and analysed on day 10. (A) Representative cytospins of cells stained by modified Wright stain; scale bar – 10μm. (B) Representative flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-CD235a antibodies. (C) Percentages of cells expressing CD71 and CD235a in IOX1 treated and control groups; mean of 3 independent biological repeats are shown; error bars represent SD. (D) Representative flow cytometry plots of cells stained with APC-conjugated anti-CD34 plotted against forward scatter (FS). (E) Percentage of cells expressing CD34 in IOX1 treated and control groups; mean of 3 independent biological repeats are shown; error bars represent SD.

5.2.5 IOX1 exerts relatively small effects on the erythroid transcriptome

To evaluate the effects of IOX1 on global erythroid gene expression, a microarray analysis was carried out on IOX1 treated cells. This was performed at the Wellcome Trust Centre for Human Genetics, University of Oxford using the Illumina human HT12v4.0 expression BeadChip which assays over 47,000 transcripts. (Data were analysed by Stephen Taylor of the Computational Biology Research Group, Radcliffe Department of Medicine, University of Oxford.)

In the microarray, expression levels of most of the genes were similar in cells treated with two concentrations of IOX1 (20μM and 40μM) as compared to a DMSO control, with very high correlation coefficients (r=0.99 in both instances) (Figure 5.5). In total, only 197 transcripts were differentially expressed in cells treated with IOX1 (40μM) compared to a DMSO control (Table 5.1, Supplementary table 1&2). There were no differentially expressed transcripts between the groups treated with 20μM and 40μM concentrations of IOX1.

Next, I analysed the expression levels and fold differences of 52 genes which were reported as essential for erythroid physiology. This list was adopted from the publicly available online database, Hembase (http://hembase.niddk.nih.gov/) and included genes coding for proteins which function as RBC cytoskeleton components, heme synthesis enzymes and essential cytoplasmic enzymes. Expression levels in IOX1 treated and untreated cells were not significantly different in all but one of the 52 genes, further confirming minimal effects of IOX1 on erythroid physiology (Table 5.2).

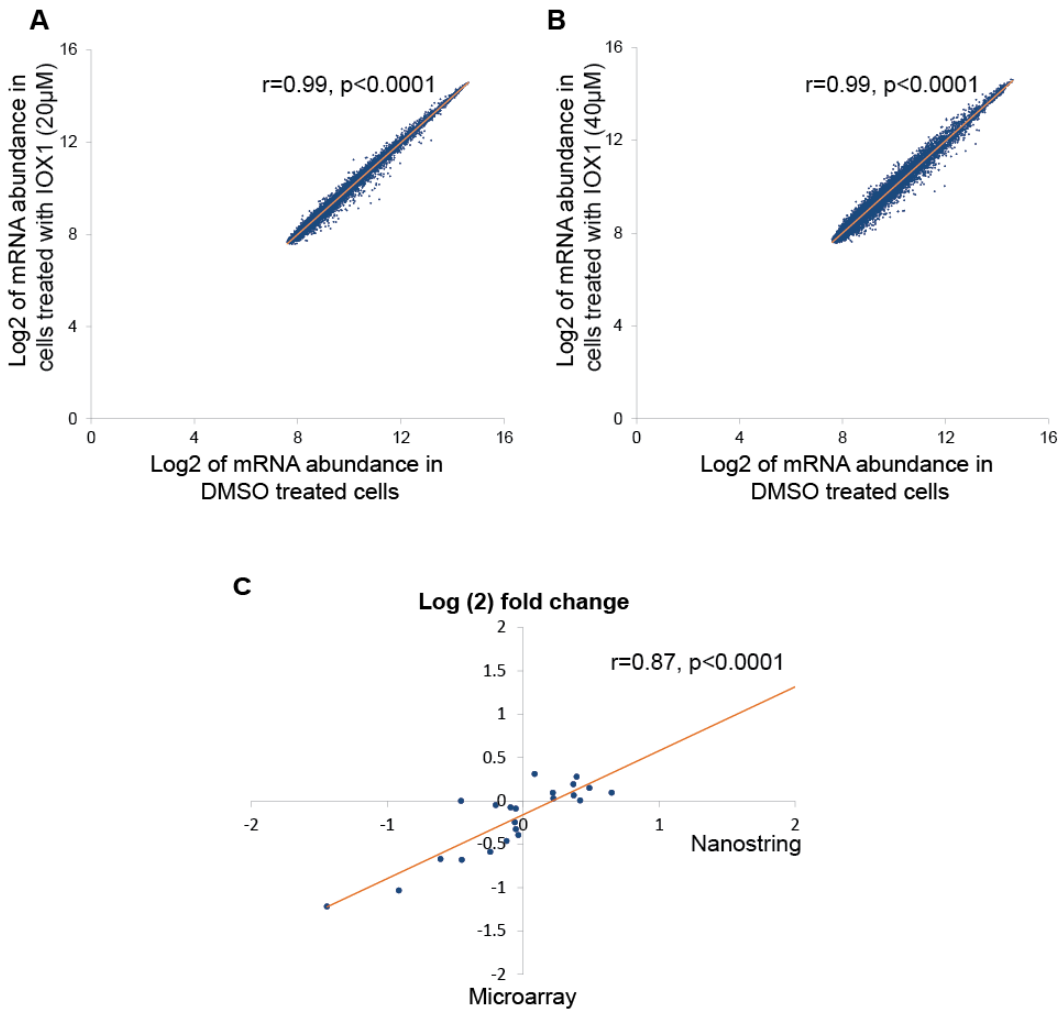


Figure 5.5 – Effects of IOX1 on erythroid transcriptome. Erythroid cells were incubated with IOX1 (20μM or 40μM) for 72 hours on day 7 and the microarray analysis was performed on day 10 of erythroid differentiation (n=4). Low expressing genes were filtered out when comparing the expression levels. (A) Scatter plot of Log_2 mRNA abundance of all the transcripts in IOX1 (20μM) or DMSO treated cells demonstrating very high (r=0.99) and statistically significant (p<0.0001) correlations. (B) Scatter plot of Log_2 mRNA abundance of all the transcripts in IOX1 (40μM) or DMSO treated cells demonstrating very high (r=0.99) and statistically significant (p<0.0001) correlations. (C) Validation of microarray data by nanostring; mRNA fold changes of several genes in the same RNA samples used in microarray were independently quantified by nanostring; high correlation (r=0.87) between data sets validates microarray.

Table 5.1 – Number of differentially expressed transcripts in IOX1 treated cells compared to DMSO control cells (n=4)[2]

IOX1 concentration	*Number of transcripts (%)*		
	Up-regulated	*Down-regulated*	*Total differentially expressed*
20µM concentration	5 (<0.1%)	17 (<0.1%)	22 (<0.1%)
40µM concentration	110 (0.2%)	87 (0.2%)	197 (0.4%)
Common to 20µM and 40µM concentrations	4 (<0.1%)	16 (<0.1%)	20 (<0.1%)

Table 5.2 – Expression levels of genes essential for erythrocyte physiology in IOX1 treated cells compared to DMSO control cells (n=4)[3]

Gene symbol	*Gene name*	*Log_2 of mRNA abundance*		*Fold change*
		DMSO	*IOX1 40µM*	
RHCE	Rhesus blood group, CcEe antigens	10.16	8.94	0.43
RHD	Rhesus blood group, D antigen	9.29	8.16	0.46
TMOD1	Tropomodulin 1	10.40	9.42	0.51
SLC4A1	Solute carrier family 4	11.45	10.65	0.58
EPB42	Erythrocyte membrane protein band 4.2	11.61	10.93	0.63
GYPB	Glycophorin B	12.36	11.69	0.63
HMBS	Hydroxymethylbilane synthase	12.20	11.71	0.71
GYPA	Glycophorin A	10.71	10.24	0.72
UROS	Uroporphyrinogen III synthase	12.52	12.05	0.72
PPOX	Protoporphyrinogen oxidase	11.31	10.85	0.73
ALAS2	Aminolevulinate, delta-, synthase 2	8.77	8.34	0.74
GPI	Glucose phosphate isomerase	11.26	10.85	0.75
CPOX	Coproporphyrinogen oxidase	12.61	12.21	0.76
GYPE	Glycophorin E	10.87	10.50	0.77
RHAG	Rhesus blood group-associated glycoprotein	12.51	12.17	0.79
STOM	Stomatin	11.85	11.53	0.80
PKLR	Pyruvate kinase, liver and RBC	10.65	10.36	0.82
TPM1	Tropomyosin 1	10.60	10.33	0.83
ALDOA	Aldolase A, fructose-bisphosphate	12.59	12.33	0.84
SPTA1	Spectrin, alpha	12.04	11.80	0.85
NT5C3	5'-nucleotidase, cytosolic III	10.63	10.41	0.86
FECH	Ferrochelatase	8.73	8.52	0.87

[2] Cut-off adjusted P value of <0.05 was used to identify differentially expressed genes.

[3] The mRNA abundance determined by microarray in cells treated with IOX1 (40µM) and DMSO (control) for 72 hours on day 7 of erythroid differentiation is shown here. The list of genes is adopted from Hembase (http://hembase.niddk.nih.gov/), a database of genes with specific and essential roles in erythrocyte physiology. Fold change represents the fold difference of mRNA abundance between IOX1 and DMSO treated cells. Only one (highlighted in yellow) out of 52 genes was differentially expressed in IOX1 treated cells compared to control.

Gene symbol	Gene name	Log_2 of mRNA abundance		Fold change
		DMSO	IOX1 40μM	
ANK1	Ankyrin 1	10.79	10.59	0.87
UROD	Uroporphyrinogen decarboxylase	13.10	12.91	0.88
GSTT1	Glutathione S-transferase theta 1	9.27	9.10	0.89
EPB49	Erythrocyte membrane protein band 4.9	9.34	9.17	0.89
ADD1	Adducin 1	8.47	8.31	0.89
CD47	CD47 antigen	9.68	9.53	0.90
ADD3	Adducin 3	9.38	9.24	0.91
GYPC	Glycophorin C	12.40	12.26	0.91
SPTB	Spectrin, beta	8.37	8.24	0.91
BPGM	2,3-bisphosphoglycerate mutase	8.84	8.73	0.92
GCLC	Glutamate-cysteine ligase	8.01	7.92	0.94
PGK1	Phosphoglycerate kinase 1	11.03	10.97	0.96
GSS	Glutathione synthetase	9.09	9.03	0.96
ALAD	Aminolevulinate, delta-, dehydratase	8.03	7.98	0.97
AQP3	Aquaporin 3	7.76	7.73	0.98
HK1	Hexokinase 1	13.17	13.14	0.98
TPI1	Triosephosphate isomerase 1	12.96	12.93	0.98
GSR	Glutathione reductase	7.82	7.80	0.99
AQP1	Aquaporin 1	7.73	7.71	0.99
HMOX1	Heme oxygenase (decycling) 1	7.93	7.92	0.99
ENO1	Enolase 1	13.13	13.14	1.00
GPX1	Glutathione peroxidase 1	10.93	10.93	1.00
PRDX2	Peroxiredoxin 2	7.82	7.83	1.01
PFKM	Phosphofructokinase, muscle	8.77	8.81	1.03
LDHB	Lactate dehydrogenase B	12.71	12.81	1.07
EPB41	Erythrocyte membrane protein band 4.1	8.07	8.18	1.08
G6PD	Glucose-6-phosphate dehydrogenase	8.30	8.47	1.12
PGD	Phosphogluconate dehydrogenase	10.71	10.90	1.14
ACTB	Actin, beta	13.49	13.73	1.18
ADA	Adenosine deaminase	8.59	8.84	1.18

To examine the possible mechanism by which IOX1 exerts the desirable effects on globin gene expression in erythroid cells, we performed a gene ontology enrichment analysis on differentially regulated gene sets obtained by microarray experiment. This did not allow a simple interpretation of how IOX1 treatment has its effects specifically on the α-globin expression however, this was not surprising as IOX1 probably acts through altering chromatin marks and its action may not be specific to a recognised cellular signalling and metabolic pathway. Similar results have been reported before when testing epigenetic inhibitor compounds affecting chromatin marks suggesting that the available pathway maps may not suitable to dissect out mechanisms of actions of epigenetic compounds (Shi et al., 2013).

5.2.6 IOX1 increases the H3K27me3 repressive chromatin modification globally and locally at the α-globin locus

IOX1 is a broad-spectrum KDM inhibitor that can inhibit large number of KDM enzymes, which in theory would increase methylation of several lysine residues of histone proteins. Therefore, to understand the mechanism of action of IOX1 on globin gene expression, I examined the changes of histone methylation patterns in erythroid cells treated with IOX1. Histone proteins were extracted from erythroid cells treated with IOX1 (40μM) and western blots were performed using antibodies to two repressive chromatin modifications, H3K27me3 and H3K9me3. Treatment of IOX1 increased the levels of both these chromatin modifications in erythroid cells (Figure 5.6A).

I then looked at the changes of these chromatin modifications at the α- and β-globin loci using chromatin immunoprecipitation (ChIP)-PCR assays. In untreated erythroid cells, the abundance of H3K27me3 was high at the α-globin promoter, however this was low (similar to the level observed at the negative control region) at the β-globin promoter (Figure 5.6B). Treatment of IOX1 increased H3K27me3 abundance at both the α- and β-globin promoters but more pronounced changes were observed at the α-promoter. In contrast, H3K9me3 was more abundant at the β-globin promoter compared to the α-globin promoter in untreated cells which further increased after IOX1 treatment (Figure 5.6C). These observations suggest that the selective down-regulation of α-globin by IOX1 is more likely to be associated with the increased methylation of H3K27 rather than H3K9.

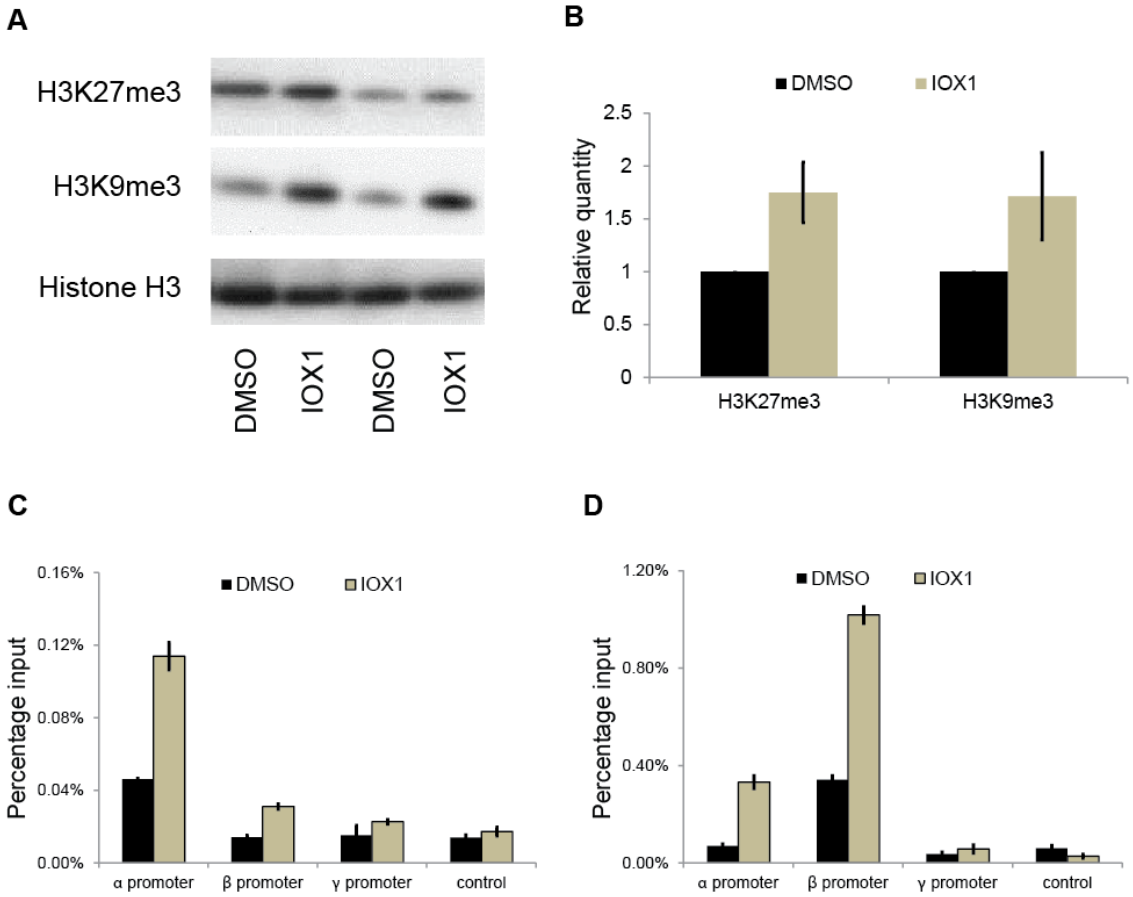


Figure 5.6 – Effect of IOX1 on histone methylation. Erythroid cells were incubated with IOX1 (40μM) or DMSO for 72 hours on day 7 of erythroid differentiation. (A) Western blot of histone extracts from erythroid cells treated with IOX1 showing abundance of H3K27me3 and H3K9me3 histone modifications and histone H3 (internal control). Two technical replicates (different dilutions) of one of the two biologically independent experiments are shown. (B) Relative quantification of abundance of H3K27me3 and H3K9me3 histone modifications analysed by western blot. (C&D) ChIP-PCR assay demonstrating abundance of H3K27me3 (C) and H3K9me3 (D) histone modifications at the α, β and γ-globin promoters in erythroid cells treated with IOX1 compared to a DMSO control. An intergenic region between the ε- and γ-globin genes was used as the negative control as described previously (Shi et al., 2013). Result of one of two biologically independent experiments is shown; error bars represent SD of technical repeats.

5.2.7 shRNA mediated knockdown of KDM6B

Considering the changes noted in the repressive chromatin signatures in IOX1 treated cells, we hypothesized that the α-globin silencing effect of IOX1 is mediated via inhibition of KDM enzymes responsible for removal of H3K27me3 methylation marks at the locus. KDM6A and KDM6B are the only enzymes

which are known to catalyse demethylation of H3K27me3. Previously work from our lab has demonstrated recruitment of KDM6B to the α-globin locus in erythroid cells (Vernimmen et al., 2011) suggesting that it might have a role in the silencing of α-globin. Therefore, I examined the effects of KDM6B knockdown in erythroid cells. Primary human erythroid cells were infected with either of two shRNAs targeting KDM6B and the effect on globin gene expression was evaluated. Contrary to our expectations, over 70% knockdown of KDM6B resulted in reductions in both α- and β-globin expression and did not produce the desired changes in the α/β-globin ratios.

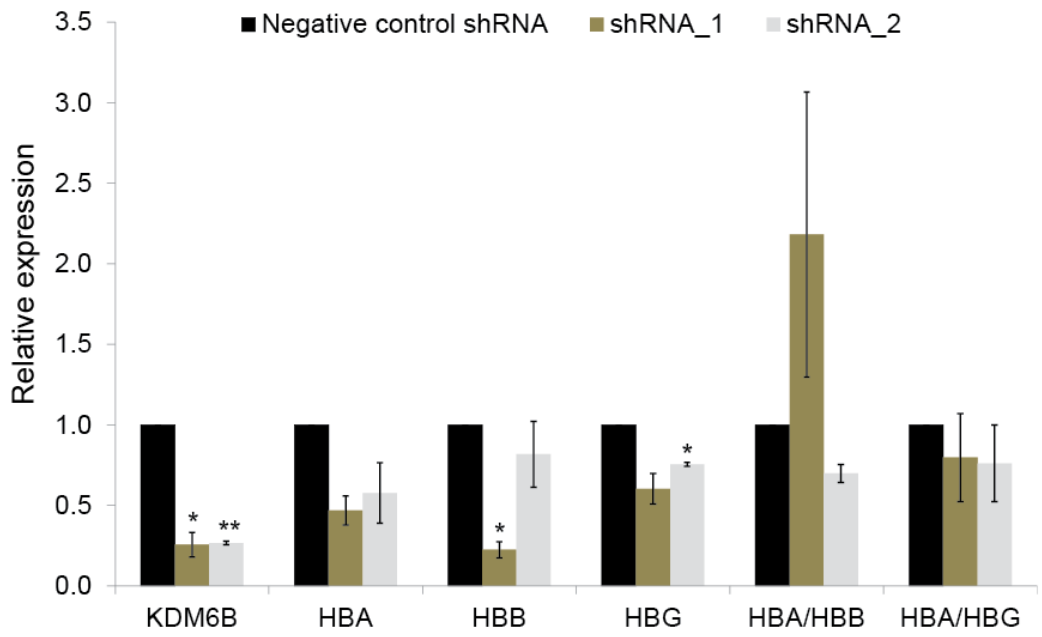


Figure 5.7 – KDM6B knockdown by shRNA in erythroid cells. Erythroid cells were transfected with either of two shRNA targeting KDM6B on day 4 of culture. Transfected cells were then selected using puromycin and differentiated further down the erythroid lineage. Gene expression was analysed on day 14 using qRTPCR. Mean expression level from 3 independent biological repeats normalised to housekeeping gene RPL13A are shown; error bars represent SD; *p<0.05 and **p<0.01 relative to negative control shRNA. Abbreviations: KDM6B, histone lysine demethylase 6B; HBA, α-globin; HBB, β-globin; HBG, γ-globin.

5.2.8 Effect of IOX1 on human and mouse α-globin expression in humanised mouse erythroid cells

Finally, with a view to testing IOX1 *in vivo* in mice, we first tested its effect on mouse primary erythroid cells *in vitro*. As there are marked differences in the regulation of human and mouse α-globin gene expression (Lynch et al., 2012), for this, we used a humanized mouse model generated previously in our lab (Wallace et al., 2007). In the humanized mouse, the mouse α-globin locus has been replaced by a 117kb segment of DNA containing the human α-globin gene cluster and its *cis*-acting regulatory elements (Figure 5.8).

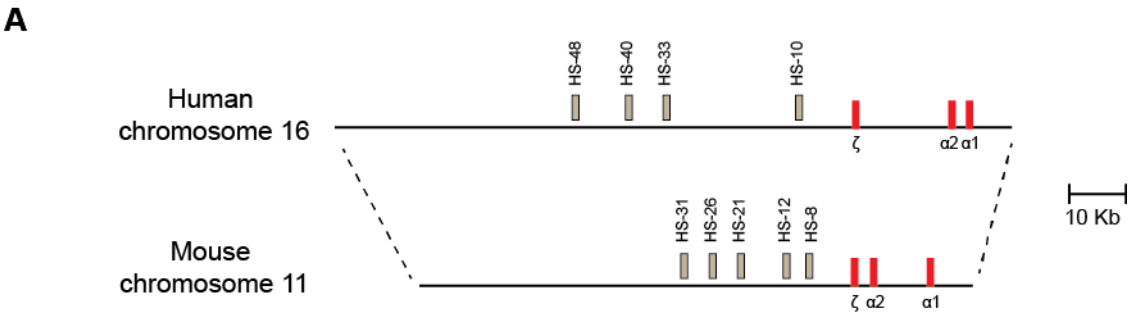


Figure 5.8 – Schematic diagram of human and mouse α-globin loci. In the humanised mouse, 117kb-long human α-globin locus with all the important *cis*-acting regulatory elements replaced the 85kb genome segment in the corresponding locus in the mouse chromosome 11. Genes are shown in red boxes and the *cis*-acting regulatory elements are shown in grey boxes.

Foetal livers from embryos of 12.5 day pregnant wild type (WT) mice were obtained after mating with male mice heterozygous for the human α-globin cluster (Het), and the embryos were screened using a PCR strategy previously designed and validated in the lab, to distinguish WT and Het embryos (Figure 5.9A&B) (Wallace et al., 2007). Erythroid progenitors from foetal livers were expanded and the immature erythroid precursors were then selected by depleting mature erythroid cells expressing Ter119 (Figure 5.9C). Ter119 is a

cell surface antigen expressed in mature mouse erythroid cells, but is absent in early erythroid progenitors. After enrichment for immature erythroid progenitors, synchronous erythroid differentiation was induced using differentiation medium which also contained IOX1 (Murine experiments including dissection of mouse foetal livers were done by Jacqueline Sharpe and Jackie Sloane-Stanley of Weatherall Institute of Molecular Medicine, University of Oxford).

Contrary to our findings on human primary erythroid cells, two concentrations of IOX1 (20μM and 40μM) failed to produce significant changes in human or mouse α-globin expression in the mouse primary erythroid cells (Figure 5.9D&E). This was confirmed by analysing cells at three different time points (8 hours [data not shown], 24 hours [data not shown] and 36 hours) corresponding to three distinct stages of erythroid differentiation. This observation may have many explanations, however, it may be that the human α-globin cluster behaves differently in the context of mouse erythroid cells and mouse may not be a suitable model to test IOX1 *in vivo*. It is also possible that there is too much globin mRNA produced before the drug is added and the effect is masked, because the mouse foetal liver erythroid differentiation is rapid and accelerated compared to the human system.

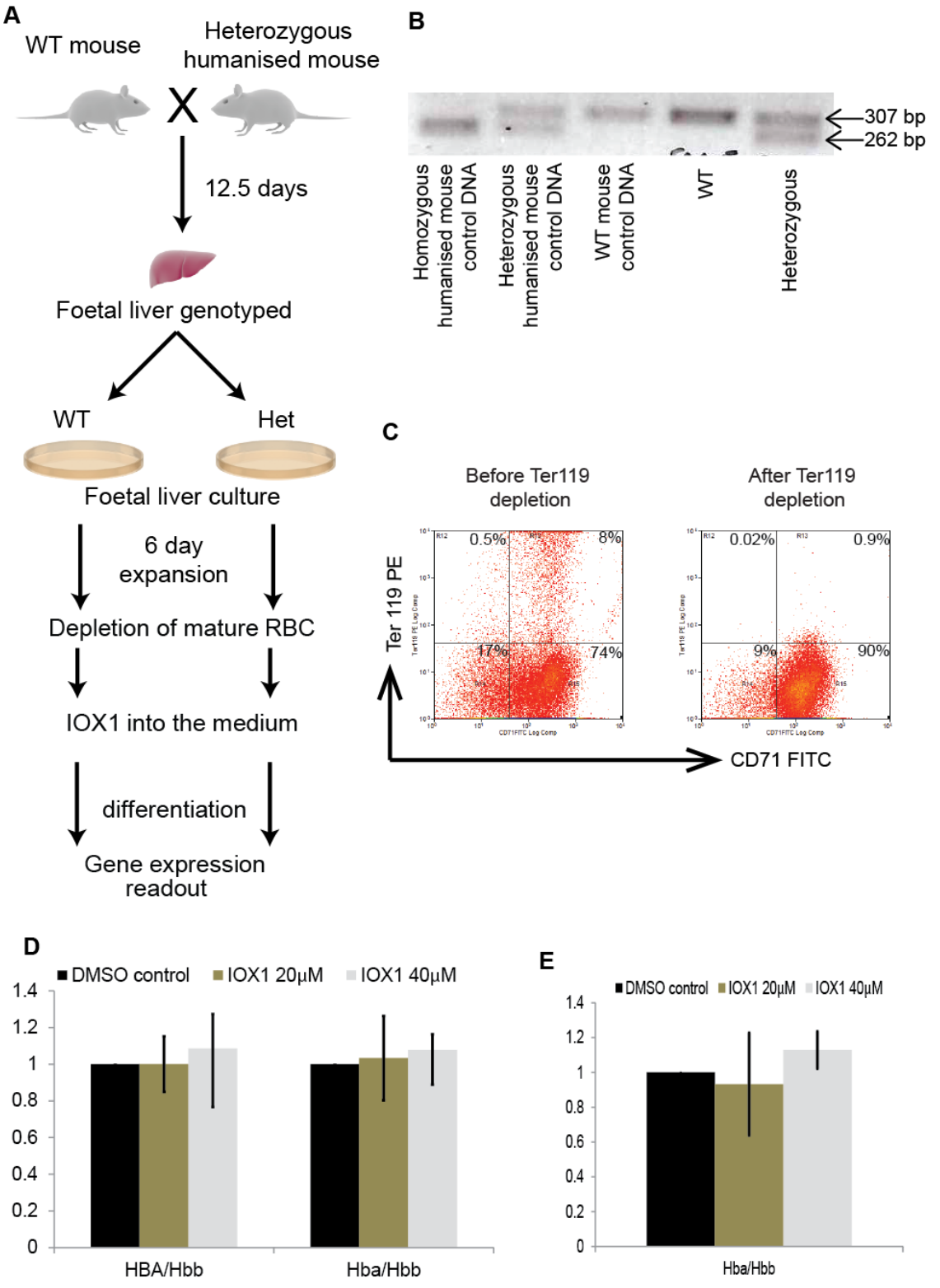


Figure 5.9 – Effect of IOX1 on human and mouse α-globin expression in humanised mouse primary erythroid cells. (A) A schematic diagram of the work flow for testing IOX1 in humanised mouse primary erythroid cells. (B) Representative gel electrophoresis image of PCR product used to screen mouse genotype. In the presence of a human allele 262bp-amplicon is produced and the presence of mouse allele gives rise to an amplicon of 307bp. First three lanes represent respective controls and the last two lanes show test WT and heterozygous foetal liver cells. (C) Representative flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-Ter119 antibodies before and after depletion for $Ter119^{+}$ mature

erythroid cells. (D) Human α/mouse β-globin and mouse α/mouse β-globin mRNA ratios of erythroid cells differentiated from heterozygous humanized mouse foetal livers treated with two concentrations of IOX1 for 36 hours. Mean of 3 independent biological repeats is shown; error bars represent SD. (E) Mouse α/mouse β-globin mRNA ratios of erythroid cells differentiated from WT mouse foetal livers treated with two concentrations of IOX1 for 36 hours. Mean of 3 independent biological repeats is shown; error bars represent SD. Abbreviations: HBA, human α-globin; Hba, mouse α-globin; Hbb, mouse β-globin.

5.3 Discussion

Selective silencing of α-globin expression without altering β-like globin expression is a plausible strategy to ameliorate β-thalassaemia. However, the α- and β-globin genes are expressed in a similar developmental-stage-specific manner, hence selective silencing of one locus in erythroid cells has long being a challenge (Higgs et al., 2005). Here, for the first time, I have demonstrated that the broad-spectrum JmjC KDM inhibitor, IOX1 selectively down regulates α-globin expression *in vitro* to levels known to be beneficial in β-thalassaemia, with minimal effects on β-globin expression.

Notably, at least *in vitro*, IOX1 does not have undesirable effects on erythroid cell viability or differentiation. This is of importance as some other compounds, in particular tranylcypromine, that are currently being tested for treatment of β-thalassaemia through up-regulation of γ-globin and foetal haemoglobin (HbF), have been shown to alter erythroid cell differentiation (chapter 4; section 4.2.4) (Shi et al., 2013). Furthermore, the transcriptome analysis showed that IOX1 has relatively small effects on global erythroid gene expression with less than 200 genes being differentially regulated and only one of those being essential in erythroid physiology. However, this should be considered with caution and given the fact that IOX1 has an inhibitory action on a broad range of KDM enzymes, it could have greater effects on transcriptomes of other cell types.

Our attempt to dissect out the exact mechanism by which IOX1 exerts its action in primary human erythroid cells was only partially successful. A single experiment performed on un-spliced mRNA nascent transcripts (data not

shown) suggested that the action of IOX1 is at the level of RNA transcription rather than at RNA stability. Furthermore, the western blot and ChIP assays revealed that IOX1 increased methylation at H3K27. Garrick and colleagues previously demonstrated that the silencing of α-globin in non-erythroid cells is mediated via recruitment of the PRC2 to the CpG islands associated with the α-globin promoter and methylation of H3K27 (Garrick et al., 2008). This is not the case at β-globin, which is not silenced through PRC2 or H3K27me3-mediated pathways. Vernimmen and colleagues demonstrated recruitment of KDM6B, a KDM responsible for demethylation of H3K27me3, to the α-globin locus in erythroid cells when the gene is activated (Vernimmen et al., 2011). Among other KDM enzymes, IOX1 inhibits KDM6B with high potency (half-maximum inhibitory concentration [IC_{50}] – 0.14μM), therefore, we hypothesised that the selective down-regulation of α-globin by IOX1 is mediated through inhibition of KDM6B (Hopkinson et al., 2013). However, KDM6B knock down using RNAi did not recapitulate the results I observed with IOX1. Therefore, at present, I am unable to conclude on the mechanism of action of IOX1. However, in my opinion, this should not preclude the use of IOX1 as a lead compound or the progress of this compound to *in vivo* studies, as we still do not know the mechanisms of action of many efficacious drugs which are currently being used in patients.

Another approach to identify the exact cellular target of IOX1 in erythroid cells is to develop a chemical probe by introducing a tag (eg: fluorophore, biotin) to the small molecule. Then, the protein targets interacting with IOX1 can be identified using various methods including affinity pull-down (Tashiro and Imoto, 2012).

However, introducing a tag to an active compound is time-consuming and difficult (or sometimes totally impossible) since any modification has the potential to cause a loss of activity (Kim and Chang, 2007). Although a photo-reactive version of IOX1 has been developed and used in cell extracts before (Rotili et al., 2011), it has not been cell permeable. The development of a cell permeable probe was considered not feasible within the time frame of my project; therefore, I did not proceed with this approach.

A major drawback of IOX1 is the lack of specificity in its inhibitory action. IOX1 has been shown to inhibit a wide array of KDM enzymes, including KDM4A (IC_{50}-0.2μM), KDM4C (IC_{50}-0.6μM), KDM6A (IC_{50}-1.1μM) and KDM6B (IC_{50}-0.14μM) with variable affinities (Hopkinson et al., 2013). To overcome this limitation, more selective inhibitors of KDM enzymes are currently being developed (Schiller et al., 2014; Wang et al., 2013). Of particular relevance, Kruidenier and others developed an inhibitor (GSK-J4), which was reported to inhibit the H3K27me3-specific KDM6 subfamily of demethylases selectively (Kruidenier et al., 2012). However, when I tested this in primary human erythroid cells, GSK-J4 was cytotoxic even at very low concentrations. Furthermore, a recent publication revealed that the action of GSK-J4 is not specific to the KDM6 proteins, but it also inhibits KDM5 enzymes that demethylase the H3K4me3 chromatin signature, which might explain its untoward effects observed in erythroid cells (Heinemann et al., 2014).

Another challenge uncovered by our results is the finding that the mouse may not be a suitable model to test whether compounds alter α-globin expression *in*

vivo. Given the contrasting and unique epigenetic architectures of human and mouse, this may well be true for most epigenetically active compounds. In fact this has been reported before (Best and Carey, 2010) and future research on epigenetic drug development may have to rely on expensive alternatives like lower primates.

In conclusion, I have demonstrated that selective down regulation of α-globin expression without affecting the β-like globin expression or erythroid differentiation is pharmacologically feasible. The histone demethylase inhibitor IOX1 exerts these desirable changes in erythroid cells by a mechanism which is still unclear. However this can be used as a lead compound to develop a novel therapy for β-thalassaemia, which is still a life-limiting condition in a majority of patients.

Chapter 6: Histone deacetylase inhibitor, vorinostat down-regulates α-globin expression whilst up-regulating γ-globin

6.1 Introduction

Acetylation of histone proteins at multiple lysine residues is a dynamic and reversible process which plays a key role in regulating gene expression in eukaryotes. In general, histone acetylation signals for transcriptional activation (Ivanov et al., 2014). This post-translational modification is regulated by two groups of enzymes with opposing effects. Histone acetyl transferases (HAT) are the 'writers' of this epigenetic signature through catalysing the transfer of acetyl groups on to the side chains of lysine residues in histone proteins. Histone deacetylases (HDACs) provide counterbalancing activity through deacetylating lysine residues and erasing these epigenetic marks (Khan and La Thangue, 2008).

The HDAC family consists of 18 different enzymes which are grouped into four classes based on homology to yeast proteins, subcellular location and enzymatic activities (Micelli and Rastelli, 2015). Classes I, II, and IV, which require zinc ions as cofactors, are called "classical HDAC classes", whereas class III HDACs are NAD^+-dependent enzymes which are also known as sirtuins (New et al., 2012).

A vast array of structurally diverse natural and synthetic chemical compounds functions as HDAC inhibitors (HDACi). HDACis are classified into several groups according to the chemical structure and include hydroxamic acids, short-chain fatty acids, benzamides, cyclic peptides and epoxyketones (West and Johnstone, 2014). The conventional HDACis, inhibit HDACs of classes I, II, and IV with variable specificity but, in general, are not active against sirtuins.

Vorinostat (also known as suberoylanilide hydroxamic acid or SAHA), a hydroxamic acid-group pan-HDACi, was one of the first epigenetic therapies approved by the US FDA. In 2006, it was licenced for the treatment of patients with advanced cutaneous T-cell lymphoma. It is also currently being investigated in phase-3 clinical trials for number of haematological malignancies including multiple myeloma, acute myeloid leukaemia and myelodysplastic syndrome as well as for solid tumours including brain gliomas and pleural mesotheliomas (NIH, 2015).

HDAC inhibition has recently been considered as a possible therapeutic option for β-thalassaemia and sickle cell disease (SCD), through induction of γ-globin and foetal haemoglobin (HbF) (Sankaran and Weiss, 2015). HDACs are suggested to be involved in repression of HbF through the modification of chromatin and association with TFs and other epigenetic enzymes that regulate HbF (Sankaran and Orkin, 2013; Suzuki et al., 2014). More precisely, inhibition of HDAC1 and HDAC2 has been shown to induce HbF (Bradner et al., 2010). A number of *in vitro* studies found that a variety of HDACis including sodium butyrate, vorinostat and givinostat are potent inducers of HbF (Bradner et al.,

2010; Hebbel et al., 2010; Ronzoni et al., 2014), which was also confirmed in the small molecule screen which I performed (chapter 4; section 4.2.5). Intriguingly, results of a phase 1/2 clinical trial evaluating the efficacy and safety of vorinostat to induce HbF in patients with SCD have recently been published (Okam et al., 2015). Although only one out of five patients demonstrated significant induction of HbF at the doses tested, this study demonstrated that vorinostat is safe and tolerable.

Remarkably, the epigenetic inhibitor small molecule screen, which I performed to identify compounds that down-regulate α-globin expression revealed vorinostat as such a compound. This was extremely interesting, as it suggests that vorinostat has the ability to reduce the excess free α-globin pool in β-thalassaemia RBCs through two independent, but synergistic pathways. In this chapter, I have further evaluated the effects of vorinostat in human primary erythroid cells to confirm the desirable changes in globin expression.

6.2 Results

6.2.1 Effects of vorinostat on globin gene expression

At the doses (1-5μM) tested during the initial screening, vorinostat reduced cell expansion, and was therefore used at lower concentrations (<500nM) in subsequent experiments. In low nanomolar (0-500nM) concentrations, vorinostat demonstrated a dose-dependent decrease in the α-globin expression in primary human erythroid cells (Figure 6.1A). The expression of β-globin was unaltered, but that of γ-globin was up-regulated in a dose-dependent manner. Furthermore, the treatment of vorinostat resulted in a modest decrease in the α/β globin mRNA ratio and a pronounced, dose-dependent and statistically-significant reduction in the α/γ globin mRNA ratio (Figure 6.1B). Also, as expected, vorinostat demonstrated up-regulation of the γ/β globin mRNA ratio. Overall, these findings suggest, that in addition to its effects on γ-globin induction, vorinostat suppresses α-globin expression relative to β-like globins, which could be an added benefit to the patients with β-thalassaemia.

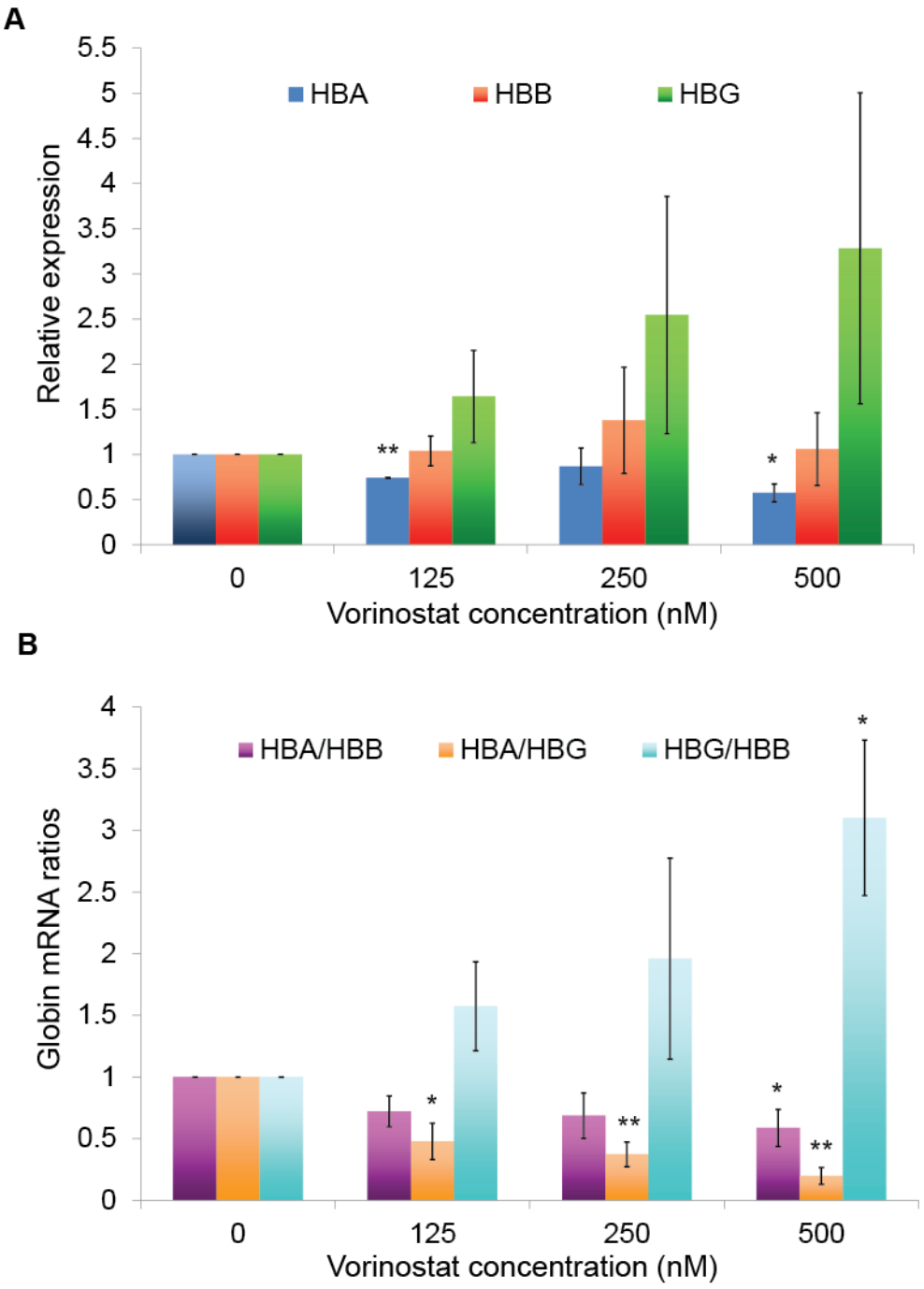


Figure 6.1 – Effects of vorinostat on globin gene expression in human erythroid cells. Erythroid cells differentiated from adult $CD34^+$ cells were incubated with a dose range (0-500nM) of vorinostat for 72 hours on day 7 of the differentiation and globin gene expressions were analysed using qRTPCR. (A) α, β and γ-globin expressions relative to RPL13A expression (B) α/β, α/γ and γ/β-globin mRNA ratios. In both graphs mean expression levels from 3 independent biological repeats are shown; error bars represent SD; *p<0.05 and **p<0.01 relative to the 0nM concentration.

These observations were verified by the Nanostring technique. Total mRNA quantification using this more sensitive assay confirmed that treatment with vorinostat resulted in a significant reduction in α-globin expression by about 35% in erythroid cells (Figure 6.2). γ-globin was up-regulated 2-fold and the expression of β-globin was unaltered. Although a 7-fold rise in the expression of ε-globin was observed after treatment with vorinostat, the absolute levels of this embryonic globin mRNA were very low compared to that of β- and γ- globin, suggesting that the changes are not clinically significant.

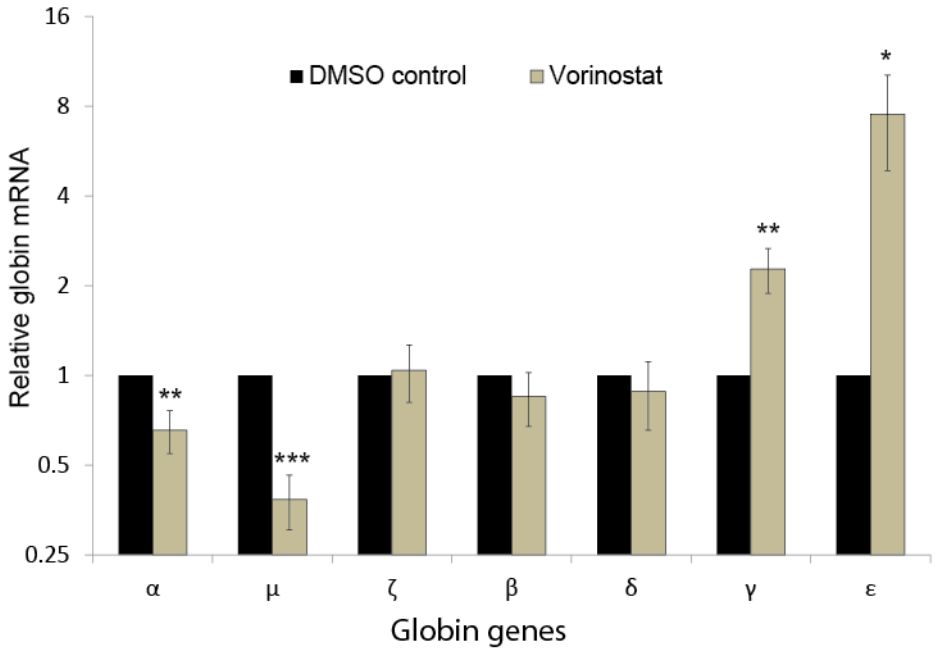


Figure 6.2 – Nanostring quantification of globin gene expression in erythroid cell after treatment with vorinostat. Erythroid cells differentiated from adult $CD34^+$ cells were incubated with vorinostat (500nM) or DMSO (control) for 72 hours on day 7 of the differentiation and the globin mRNA levels were quantified using Nanostring nCounter digital analyser. Mean Nanostring counts of globin genes normalized to the mean of multiple housekeeping genes (RPL13A, RPL18, GAPDH, PABPC1, CA2, FTH1, PAIP2 and LAPTM4A) from 3 independent biological repeats is shown; error bars represent SD; *p<0.05, **p<0.01, ***p<0.001 relative to DMSO control. Scale of the y-axis is logarithmic (base-2).

6.2.2 Effects of vorinostat on erythroid cell proliferation and viability

As mentioned previously, vorinostat was toxic to the erythroid cells at 1-5µM concentrations, however, at lower concentrations (0-500nM), it did not produce significant effects on cell proliferation; the fold expansions were similar to those of controls (Figure 6.3A). Similarly, no significant difference in cellular viability in treated and untreated cells was observed (Figure 6.3B).

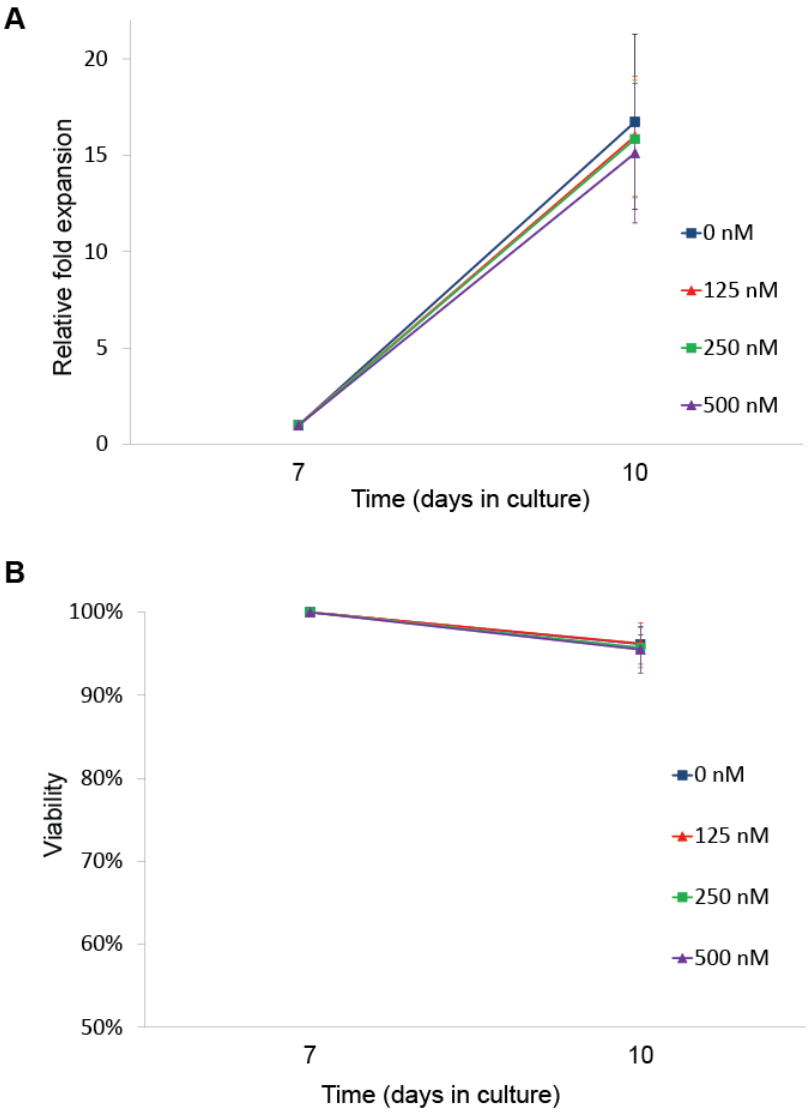


Figure 6.3 – Expansion and viability of erythroid cells treated with vorinostat. Erythroid cells were incubated with a dose range (0-500nM) of vorinostat for 72 hours on day 7 of the differentiation. The 0nM concentration refers to a DMSO control. (A) Cell growth shown as mean fold expansion relative to the number of cells on day 7 (B) Percentages of viable cells analysed by trypan blue test. In both figures mean values from 3 independent biological repeats are shown; error bars represent SD.

6.2.3 Effects of vorinostat on erythroid differentiation

Treatment of vorinostat did not have significant effects on the morphology of the erythroid cells in the doses tested and were similar to the control cells on day 10 of culture (Figure 6.4A). However, immunophenotypical analyses showed that, treatment of vorinostat (500nM) have effects on erythroid differentiation (Figure 6.4B). Percentages of cells positive for CD71 and double positive for CD71 and CD235a were lower in the vorinostat treated group compared to the controls; (Figure 6.4C). Furthermore, the expression of CD34, an early progenitor cell marker, was significantly higher (2.6% vs 0.9%) in the vorinostat treatment group. Overall, these findings suggest that vorinostat has significant effects on erythroid differentiation *in vitro* (Figure 6.4D&E).

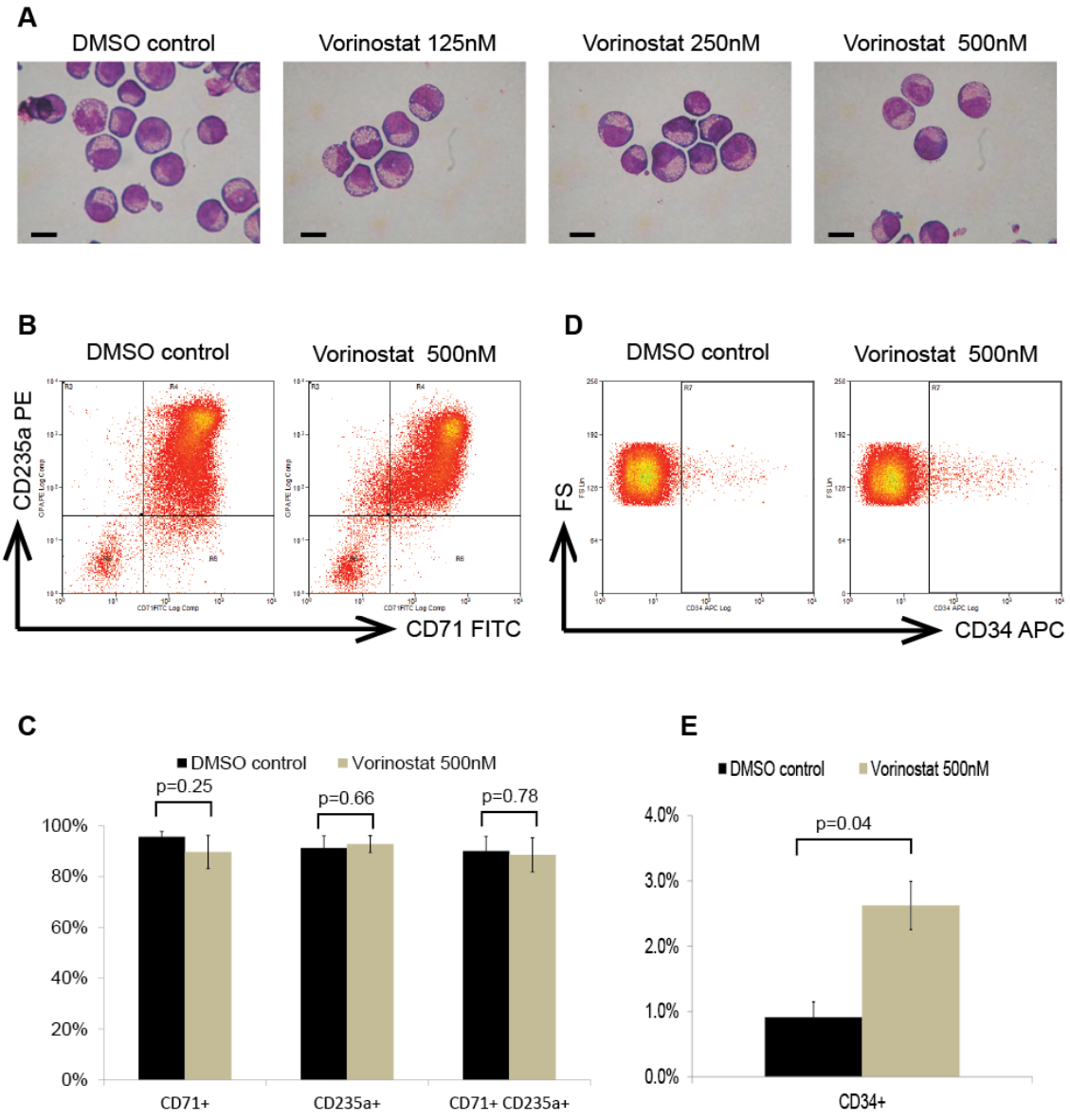


Figure 6.4 – Effects of vorinostat on erythroid differentiation. Erythroid cells were incubated with a dose range (0-500nM) of vorinostat for 72 hours on day 7 of the differentiation and analysed on day 10. (A) Representative cytospins of cells stained by modified Wright stain; scale bar – 10μm. (B) Representative flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-CD235a antibodies. (C) Percentages of cells expressing CD71 and CD235a in vorinostat treated and control groups; mean of 3 independent biological repeats is shown; error bars represent SD. (D) Representative flow cytometry plots of cells stained with APC-conjugated anti-CD34 plotted against forward scatter (FS). (E) Percentage of cells expressing CD34 in vorinostat treated and control groups; mean of 3 independent biological repeats is shown; error bars represent SD.

6.2.4 Effects of vorinostat on erythroid transcriptome

Then, I tested the effect of vorinostat on global erythroid gene expression by microarray. This was performed using the Illumina human HT12v4.0 expression BeadChip (Data were analysed by Stephen Taylor of the Computational Biology Research Group, Radcliffe Department of Medicine, University of Oxford). Similar to the observations made with IOX1 (chapter 5; section 5.2.5), the expression levels of the majority of transcripts were similar in vorinostat treated and untreated cells (Figure 6.5A&B). Two-hundred and four transcripts were found to be differentially regulated between the two groups when a cut-off adjusted p value of <0.05 was used; 128 (0.3%) transcripts were up-regulated, whereas 76 (0.2%) were down-regulated in vorinostat treated cells (Supplementary table 3&4).

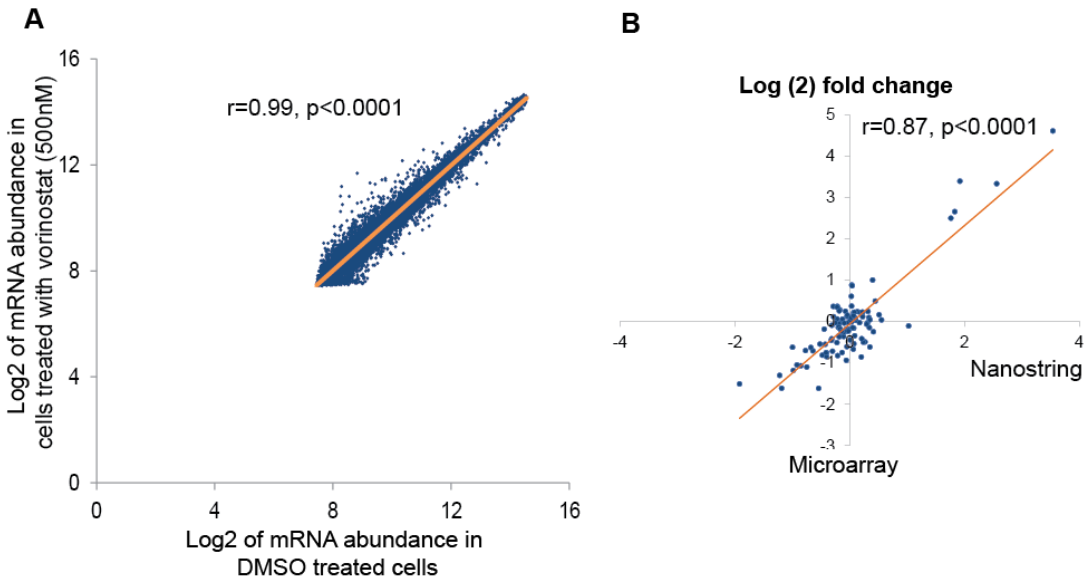


Figure 6.5 – Effect of vorinostat on the erythroid transcriptome. Erythroid cells were incubated with vorinostat (500nM) for 72 hours on day 7 and the microarray analysis was performed on day 10 of erythroid differentiation (n=3). Low expressing genes were filtered out when comparing the expression levels. (A) Scatter plot of Log_2 mRNA abundance of all the transcripts in vorinostat or DMSO treated cells demonstrating very high (r=0.99) and statistically significant (p<0.0001) correlations. (B) Validation of microarray data by nanostring; mRNA fold changes of several genes in the same RNA samples used in microarray were independently quantified by nanostring; high correlation (r=0.87) between data sets validates microarray.

6.2.5 Combination of IOX1 and vorinostat

Finally, I examined the effects of a combination of IOX1 and vorinostat on human primary erythroid cells. The combination of these two compounds demonstrated significant effects on globin gene expression in erythroid cells. IOX1 and vorinostat in combination down-regulated α-globin expression by over 60%, however, simultaneous modest but significant reductions in β-globin expression were also noted. Up-regulation of γ-globin was significant but not very striking, probably due to the counterbalancing actions of two drugs; vorinostat up-regulates γ-globin whereas IOX1 down-regulates the same. Analysis of more meaningful globin ratios showed that the drug combination has synergistic effects which could be beneficial to the patients with β-thalassaemia. In the treated cells, α-to β-like globin mRNA ratios (α/β and α/γ) were significantly lowered while γ/β mRNA ratios were increased.

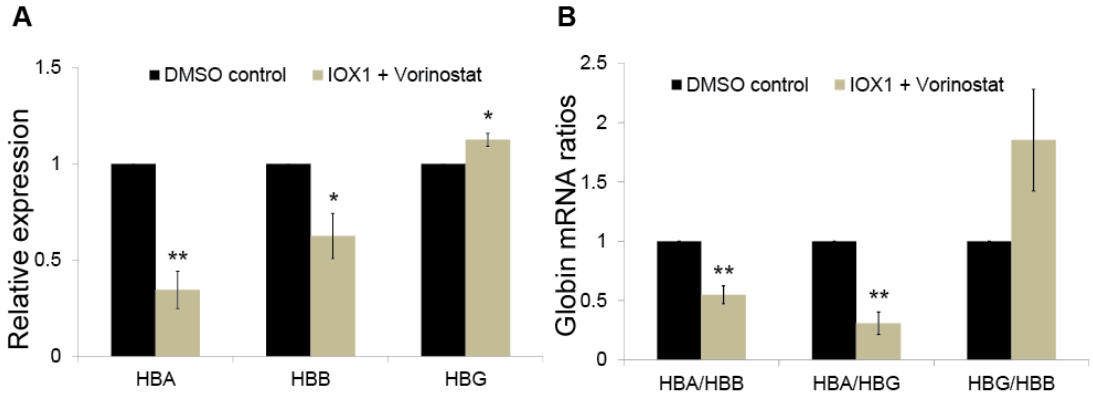


Figure 6.6 – Effects of a combination of IOX1 and vorinostat on globin gene expression in human erythroid cells. Erythroid cells differentiated from UCB $CD34^+$ cells were incubated with IOX1 (20μM) and vorinostat (500nM) for 72 hours on day 7 of the differentiation and globin gene expressions were analysed using qRTPCR. (A) α, β and γ-globin expressions relative to RPL13A expression (B) α/β, α/γ and γ/β-globin mRNA ratios. In both graphs mean expression levels from 3 independent biological repeats are shown; error bars represent SD; *p<0.05 and **p<0.01 relative to a DMSO control.

6.3 Discussion

Vorinostat is one of the first epigenetic inhibitor drugs approved by the US FDA and has recently been tested in a clinical trial for treatment of β-haemoglobinopathies (Okam et al., 2015). In addition to its previously described γ-globin and HbF inducing capabilities (Sankaran and Weiss, 2015), here for the first time, we have shown that it has a potential to down-regulate α-globin expression, making it an impressive candidate in the treatment of β-thalassaemia. In our assays, vorinostat demonstrated a dose-dependent down-regulation of α-globin expression at low nanomolar concentrations without significant changes in erythroid cell growth or viability and with only mild changes in erythroid differentiation. Our results on the γ-globin induction potential of vorinostat were comparable to the results previously reported using K562 cell lines (Hebbel et al., 2010) and human primary erythroid cells (Bradner et al., 2010) as well as to the γ-globin induction potential of givinostat (a hydroxamic acid group pan-HDACi similar to vorinostat) in human erythroid cells (Ronzoni et al., 2014).

Vorinostat has several advantages over IOX1 as an α-globin repressor and a potential therapy for β-thalassaemia. It is already a licenced medication for other indications with an available oral dosage form, of which, the safety, tolerability and clinical pharmacology profile have been characterised (Duvic et al., 2007; Iwamoto et al., 2013). Therefore, it could be used in human subjects and patients without the necessity for additional pre-clinical or animal studies. In addition, vorinostat would benefit patients with β-thalassaemia through two independent pathways which act synergistically to reduce the excess free α-

globin chains in the RBCs. In clinical studies among patients with β-thalassaemia, the reduction of α-globin and upregulation of γ-globin and HbF by genetic modifiers have shown to provide additive effects in ameliorating the disease phenotype (Galanello et al., 2009).

Another advantage of vorinostat, as already being investigated (Okam et al., 2015), is that it could be a useful therapy for SCD as well. Pharmacological induction of γ-globin and HbF is beneficial to the patients with SCD (Fathallah and Atweh, 2006) and in fact, the HbF inducing agent hydroxyurea is used as a therapy in patients with moderate to severe SCD (Platt, 2008). In addition, reduction of α-globin decreases the amount of HbS in RBCs, reduces sickling, ameliorates haemolysis and is associated with lower incidence of some of the complications of SCD which include stroke, priapism, and leg ulcers (Steinberg and Sebastiani, 2012). Therefore, vorinostat might be a useful therapy not only for β-thalassaemia, but also for SCD and could benefit over 300,000 new patients born each year (Weatherall, 2010).

Our results show that vorinostat might have an influence on the differentiation of human erythroid cells. We observed significantly higher percentages of $CD34^+$ cells in the vorinostat treatment groups compared to controls suggesting retardation of differentiation. Similar observations were made by others using closely related HDACis in primary human erythroid cells. In an *in vitro* study done by Ronzoni and colleagues, a pan-HDACi givinostat, decreased the expression of erythroid cell surface markers and delayed erythroid differentiation at higher concentrations. (Ronzoni et al., 2014).

Another disadvantage of vorinostat is its side effect profile identified during clinical trials. The most common adverse reactions associated with vorinostat treatment were diarrhoea, fatigue, nausea, anorexia and thrombocytopenia (Duvic et al., 2007; Iwamoto et al., 2013). However, the desirable effects of vorinostat on the globin gene expression were observed at low nanomolar doses, hence it could be efficacious *in vivo* at much lower doses that might not cause significant side effects.

One limitation of our study is that we did not explore the mechanism by which vorinostat exerts the beneficial effects on globin gene expression. Although, the γ-globin induction is most likely to be due to inhibition of HDACs, silencing of α-globin may be through a different or an indirect pathway. Interestingly, there is evidence to suggest that vorinostat has inhibitory activity against histone demethylases (KDMs) in addition to HDACs (Rose et al., 2008; Thinnes et al., 2014). As histone methylation is a recognised epigenetic mechanism regulating α-globin (Mettananda et al., 2015) and the broad-range KDM inhibitor, IOX1, down-regulates α-globin expression (chapter 5), it is possible that the α-globin down-regulation by vorinostat is also mediated through inhibition of histone demethylases. This is an important research area which needs to be explored in the future.

Okam and colleagues recently reported a phase 1/2 clinical trial which examined the effects of vorinostat in five patients with severe SCD who were either intolerant or non-responding to hydroxyurea therapy (Okam et al., 2015).

They analysed the HbF percentages and expression of β, γ and ε-globin mRNA levels in blood during treatment of vorinostat but did not look at the expression of α-globin. We planned to obtain samples from patients who are recruited into a phase-2 clinical trial evaluating the effects of a combination therapy which included vorinostat in relapsed multiple myeloma patients (ClinicalTrials.gov Identifier NCT01720875 [MUK four trial]). Our intended strategy was to obtain blood samples before and after treatment from patients and to analyse basic haematology, haemoglobin H inclusion bodies, globin mRNA levels and HbF levels. However, due to the low patient recruitment for this trial from Oxford University Hospitals NHS trust, this was not successful. Undoubtedly, in a clinical trial testing vorinostat in the future, it will be useful to include analysis of α-globin mRNA levels in blood as a secondary outcome measure.

In conclusion, I have demonstrated that the US FDA approved HDACi vorinostat down-regulates α-globin expression whilst inducing γ-globin expression in human primary erythroid cells and is a potential therapy for β-thalassaemia. This significant novel finding would strengthen the knowledge base of this epigenetic drug, which has already made significant progress into the clinical trials on treatment of β-haemoglobinopathies.

Chapter 7: CRISPR/Cas9 mediated deletion of the α-globin enhancer in human CD34^{+} cells

7.1 Introduction

7.1.1 Cis-acting regulatory elements of the α-globin gene

The human α-globin gene cluster is located on the short arm of chromosome 16 (16p13.3) very close (~150kb) to the telomere (chapter 1; figure 1.2). It is known that the α-globin expression is controlled by four distant enhancers located upstream to the genes. Distal regulators of α-globin first emerged in the early 1990s, following a report of a family with α-thalassaemia due to a large (62kb) deletion upstream of the α-globin locus, but leaving the α-globin genes intact (Hatton et al., 1990; Higgs et al., 1990). Since then, over 20 similar mutations causing α-thalassaemia have been reported (Harteveld and Higgs, 2010). The mapping of positions of DNAse1 hypersensitive sites to the MCSs confirmed there are four erythroid specific distant regulatory elements (enhancers) situated 10 (MCS-R4), 33 (MCS-R3), 40 (MCS-R2) and 48 (MCS-R1) kilo bases upstream of the gene locus enhancing human α-globin expression (Figure 7.1) (Higgs and Wood, 2008; Hughes et al., 2005).

7.1.2 MCS-R2 is the most critical enhancer of α-globin

A variety of early experiments suggested that MCS-R2 (previously known as HS-40) is the most critical regulatory element capable of enhancing α-globin

expression on its own (Higgs et al., 1990; Sharpe et al., 1992). Multispecies analysis revealed that MCS-R2 is the most highly conserved regulatory element and it could be detected in all mammals (Hughes et al., 2005). In a humanised mouse model, in which the mouse α-globin locus was replaced by a 117kb segment of DNA containing the human α-globin gene cluster and regulatory elements, the deletion of a 1.1kb segment covering MCS-R2 in the humanised chromosome resulted in a reduction of the human α-globin expression to almost undetectable levels (~2% compared to wild type humanised chromosome) (Wallace et al., 2007). Similarly, all naturally occurring human deletions which cause α-thalassaemia by removing the upstream regulatory elements but leaving α-globin genes intact, include MCS-R2 (Harteveld and Higgs, 2010). These observations confirm that MCS-R2 is crucial for the expression of the human α-globin genes.

Careful inspection of MCS regions has revealed that the most conserved sequences in these elements correspond to the binding sites for TFs that play a major role in the erythroid transcriptional network (Higgs and Wood, 2008; Hughes et al., 2005). The regulatory element MCS-R2 behaves as a classical enhancer. The functional domain of this element is restricted to a ~260bp core fragment, in which several well-conserved nuclear protein binding sites have been identified. These include four potential binding sites for the erythroid-specific factor GATA1 and two binding sites for the erythroid-factor NF-E2 (Figure 7.1) (Chen et al., 1997; Jarman et al., 1991). De Gobbi and others analysed the patterns of TF binding across the α-globin locus in human erythroid cells, using ChIP-chip technology. The four upstream regulatory

elements (MCSR1-4) were shown to bind transcription factors GATA1 and SCL however, the enrichment of NF-E2 was only observed at MCS-R2 and to a lesser extent MCS-R1 (De Gobbi et al., 2007).

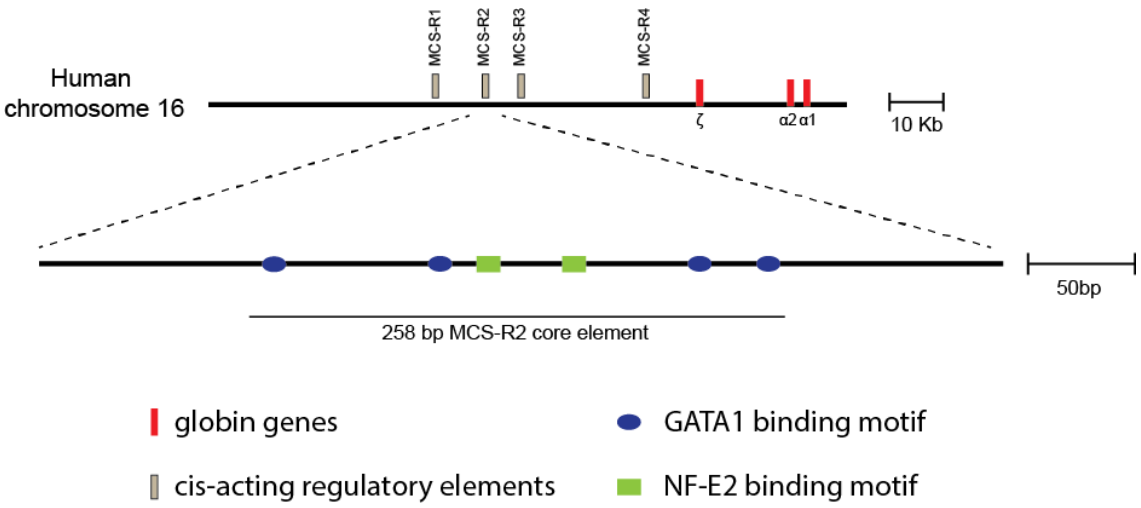


Figure 7.1 – Schematic representation of the α-globin gene cluster and the TF binding motifs in the MCS-R2 core element.

Vernimmen and colleagues analysed the changes of TF binding and chromosomal looping in wild type and MCS-R2 knockout (ΔMCS-R2) humanised mice. During normal erythroid differentiation, key transcription factors are recruited to the remote regulatory elements followed by recruitment of the pre-initiation complex to both the enhancers and promoters thereby initiating transcription. During this process a physical interaction between enhancers and promoters are believed to occur through a mechanism of chromatin looping. In the ΔMCS-R2 mutant, the absence of α-globin expression was associated with a lack of RNA polymarase II (PolII) binding and impaired loop formation (Vernimmen et al., 2007; Vernimmen et al., 2009).

In non-erythroid cells, α-globin is normally silenced through a mechanism involving the PRC2 and the associated H3K27me3 chromatin modification. In erythroid cells, when the α-globin gene is activated, PRC2 is detached and the H3K27me3 chromatin mark is erased through passive and active mechanisms including demethylation by the specific histone lysine demethylase enzyme KDM6B. However, in ΔMCS-R2, PRC2 is no longer cleared from the α-globin promoter and KDM6B is not recruited (Vernimmen et al., 2011). In summary, it is evident that MCS-R2 has multiple roles in facilitating α-globin gene expression, which include, recruitment of PolII and key TFs to the promoter, formation of a looped structure involving remote regulatory elements and the promoter, removal of the repressive PRC2 complex, recruitment of KDM6B and loss of H3K27me3 (Vernimmen, 2014).

7.1.3 Naturally occurring deletion limited to MCS-R2

A rare patient (referred to as MC from now onwards) with a naturally occurring deletion of the MCS-R2 region reported by Coelho and colleagues provides very interesting observations which are somewhat inconsistent with the perceived pre-eminence of MCS-R2, derived from *in vitro* and animal studies (Coelho et al., 2010). MC presented with Haemoglobin H (HbH) disease, and was found to be homozygous for a 3.3kb deletion (named $(\alpha\alpha)^{ALT}$) that removes only MCS-R2, leaving the α-globin genes and the other MCSs intact.

MC was further characterised by a past member of our lab, Deborah Hay. Basic haematology, peripheral blood smear and HbH inclusion body preparation from MC were consistent with HbH disease. Surprisingly, the α-globin mRNA level in

erythroid cells was approximately 30% of the normal level (figure 7.2) (Hay, 2014). This was in contrast to the previous findings in hybrid cell-lines and the ΔMCS-R2 humanised mouse which reported extremely low (~2%) mRNA levels of α-globin associated with the loss of MCS-R2 (Hatton et al., 1990; Higgs et al., 1990; Wallace et al., 2007). The α-globin mRNA level of the daughter of MC, who was heterozygous for the $(\alpha\alpha)^{ALT}$ mutation was 55%. Interestingly, in the $(\alpha\alpha)^{ALT}$ mutation, the 3.3kb deletion is associated with the insertion of a 39 nucleotide (39nt) orphan segment, whose sequence was not found in the human genome. Further work by Deborah Hay showed that this 39nt sequence neither has binding motifs for general or erythroid specific TFs nor shows binding of TFs and suggested that this sequence is probably not responsible for the unexpectedly high α-globin mRNA levels seen in MC. Furthermore, chromosome conformation capture in cultured erythroid cells from MC demonstrated normal interactions between the residual regulatory elements and the α-globin promoter despite the absence of MCS-R2 (Hay, 2014). In summary, these results suggest that in the natural human chromosome environment, deletion of MCS-R2 reduces but does not completely abolish the expression of the α-globin genes.

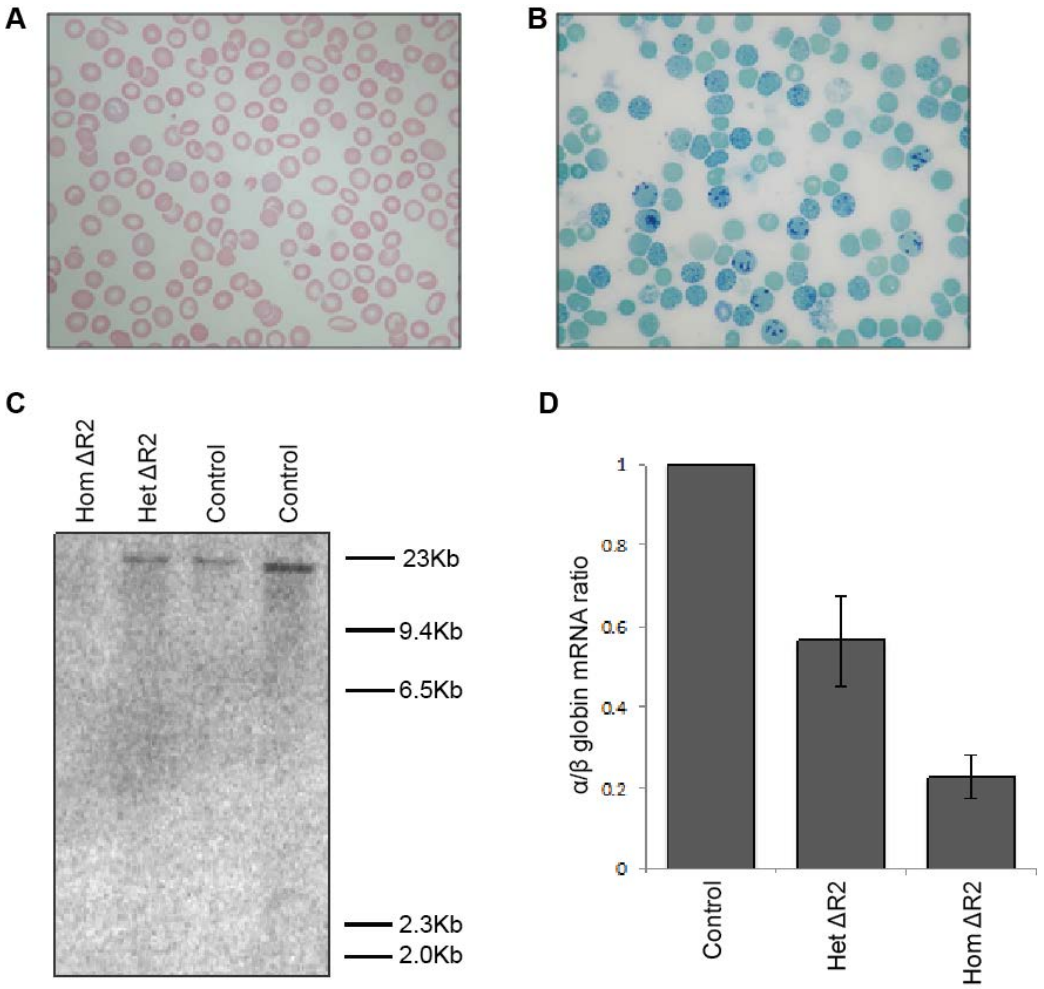


Figure 7.2 – Clinical data from individuals with the $(\alpha\alpha)^{ALT}$ (ΔMCS-R2) mutation. (A) Peripheral blood smear, Giemsa stained from a patient with a homozygous $(\alpha\alpha)^{ALT}$ mutation. (B) Peripheral blood HbH preparation, brilliant cresyl blue stained, from a patient with a homozygous $(\alpha\alpha)^{ALT}$ mutation. (C) A probe for the core of MCS-R2 which gives a 19kb band with BglII digested genomic DNA from normal controls and heterozygous $(\alpha\alpha)^{ALT}$ but not in homozygous $(\alpha\alpha)^{ALT}$. (D) α/β-globin mRNA ratio in heterozygous $(\alpha\alpha)^{ALT}$ and homozygous $(\alpha\alpha)^{ALT}$ mutation. Adapted from (Hay, 2014).

7.1.4 Implications of deletion of MCS-R2 in β-thalassaemia

From the evidence presented, it is clear that the MCS-R2 segment is critical in enhancing α-globin expression but, in its natural human chromosomal environment, is not essential. Deletion of MCS-R2 (ΔMCS-R2) down-regulates α-globin expression but does not completely silence it. In the heterozygous state, α-globin production is 55% of the normal level and in the homozygous state is approximately 30%. Notably, as clinically evident in MC, the

homozygous ΔMCS-R2 results in a mild non-transfusion dependent anaemia but, does not result in any other abnormalities. Therefore, we hypothesised, if a ΔMCS-R2 mutation is engineered in the patients with β-thalassaemia, in the heterozygous state, it would down-regulate α-globin expression to a degree that would ameliorate the clinical phenotype of β-thalassaemia; even in the homozygous state, it would not be harmful and in fact, could still be beneficial. In summary, selective deletion of MCS-R2 is a potential effective, safe and novel approach to down-regulate α-globin expression to ameliorate β-thalassaemia.

7.1.5 Genome editing in HSCs

As detailed in chapter 1 (section 1.7.3), genome editing using sequence-specific programmable artificially engineered nucleases has provided a realistic approach for the treatment of many human genetic diseases including the haemoglobinopathies. These nucleases create double strand breaks at specific chosen locations in the genome, which when repaired, create mutations at the targeted sites and, depending on their position, result in either loss of gene function or disruption of non-coding regulatory sequences. The zinc-finger nuclease (ZFN) was the first such technology to be developed. Although, this generated lots of promise, many drawbacks including cytotoxicity, low-targeting efficiency, difficulties in engineering proteins for each target and limitations in target sites (one per ~100bp) precluded its wide spread use (Kim and Kim, 2014). The second group, transcription activator-like effector nucleases (TALENs), were popular as they dramatically increased the availability of target sites allowing almost any sequence to be targeted. However, the construction

of TALENs is challenging and time-consuming (Cheng and Alper, 2014). Both ZFNs and TALENs use DNA-binding proteins to tether endonuclease catalytic domains to specific genomic loci. By contrast, in the recently developed RNA-guided CRISPR-Cas system, the endonuclease Cas9 is guided by small RNAs through Watson-Crick base pairing with target DNA (Garneau et al., 2010). Therefore it represents a system that is markedly easier to design, highly specific, efficient and well-suited for high-throughput and multiplexed gene editing for a variety of cell types and organisms (Ran et al., 2013b).

The MCS-R2 segment fulfils many of the requirements for an excellent candidate genomic region targeted by genome editing technology and disruption of this in human HSCs is likely to be feasible. Firstly, the segment for the desired deletion is relatively short (~250bp) and therefore, is easy to manipulate. Secondly, evidence from patient studies confirms that the removal of this segment does not lead to untoward effects. Thirdly, the target cell type (HSCs) is readily obtainable using previously optimised technologies and ample clinical facilities and experience are available for transplanting (bone marrow transplantation) them back to patients.

If this strategy is successful, translation into clinical practice would broadly involve three steps (Figure 7.3). Firstly, the CD34$^+$ HSPC are harvested from patients with β-thalassaemia after mobilisation using granulocyte colony-stimulating factor. Then the harvested HSPCs are edited *ex vivo* using programmable nucleases and transplanted back to the same patient after myeloablation.

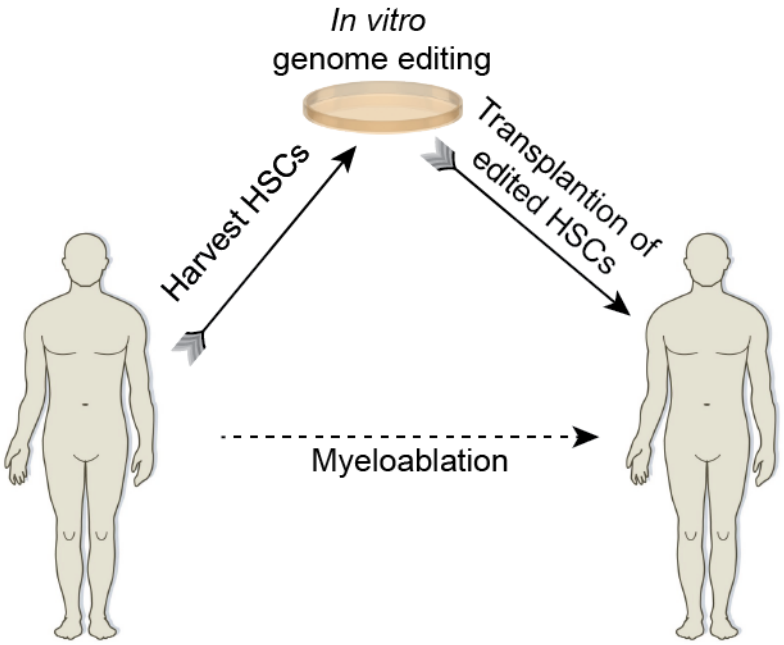


Figure 7.3 – Schematic of workflow of genome editing treatment for β-thalassaemia. HSCs are harvested from patients, edited *ex vivo* to introduce desired effects and transplanted back in to the same patient after myeloablation.

The objective of this part of my project was to knockout the MCS-R2 regulatory element by genome editing in normal human CD34$^+$ HSPCs to achieve reductions in α-globin gene expression to levels beneficial for the patients with β-thalassaemia. Due to the easier design and higher efficiency, I have utilised the CRISPR-Cas9 system for genome editing.

7.2 Results

7.2.1 Design of guide RNAs

Several guide RNAs (gRNAs) were designed to target Cas9 either at the 5′ or 3′ end of the MCS-R2 core element using an online CRISPR design tool (crispr.mit.edu). Four gRNAs targeting the 5′ end and three gRNAs targeting 3′ end were selected based on the quality score derived from the software and were commercially synthesised (Figure 7.4).

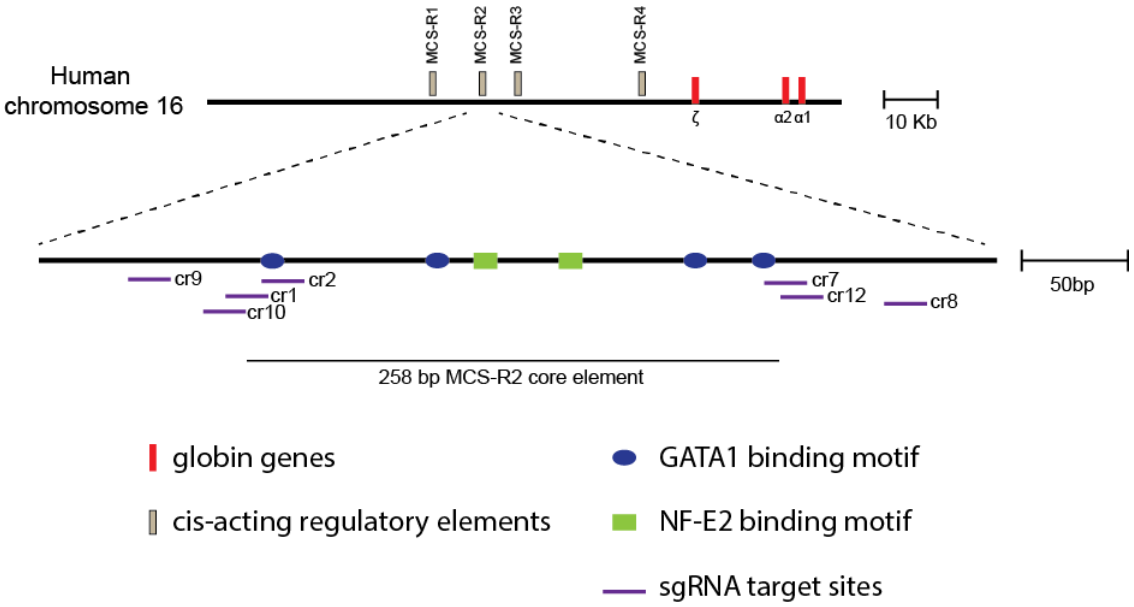


Figure 7.4 – Schematic of gRNA target sites. Four gRNAs (cr1, cr2, cr9 and cr10) were designed to target the 5′ end of the MCS-R2 core element whereas three gRNAs (cr7, cr8 and cr12) target the 3′ end.

7.2.2 Preparation of the gRNA expression construct

Expression plasmids for gRNAs were constructed by cloning oligonucleotides into a backbone vector which contains the green fluorescent protein (GFP) gene fused to a Cas9 gene (Addgene plasmid 48138; pSpCas9[BB]-2A-GFP [pX458]) (Figure 2.1 and Figure 7.5). (Cloning of expression constructs and sequence validation of the CRISPR/Cas9 plasmids were performed by Philip Hublitz in the Genome Engineering Facility, Weatherall Institute of Molecular Medicine, University of Oxford)

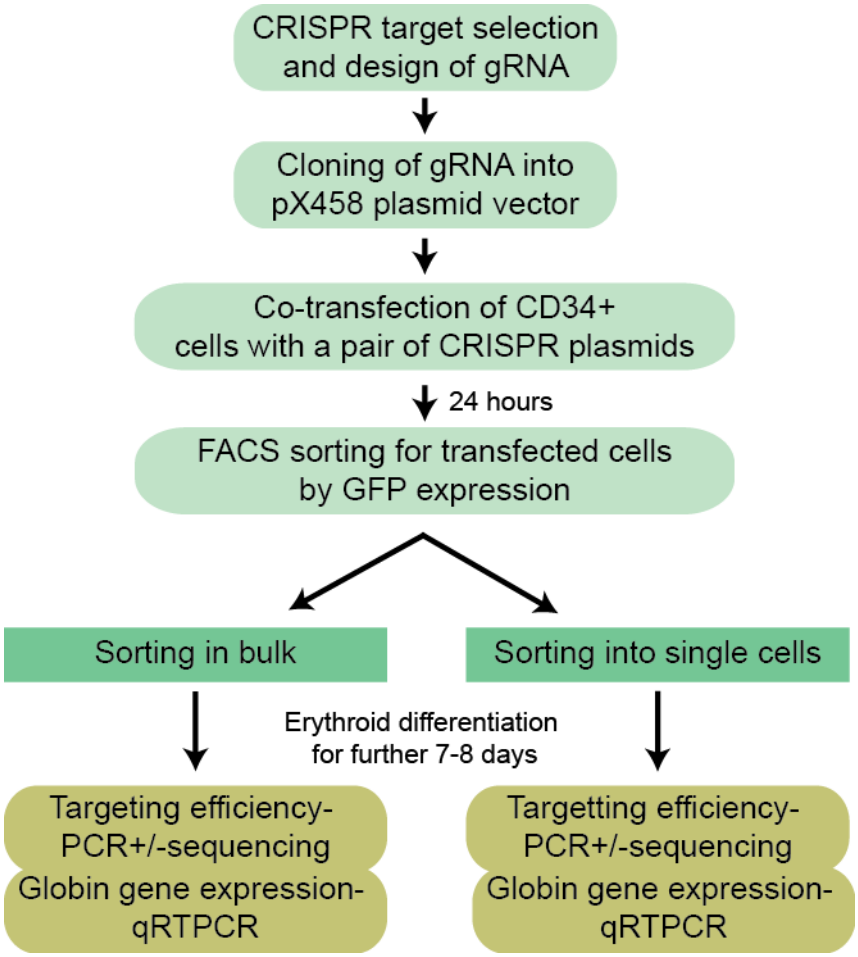


Figure 7.5 – Summary of experimental workflow to generate deletions of the MCS-R2 region in human CD34+ cells using the CRISPR/Cas9 system.

7.2.3 Transfection of human CD34$^+$ cells

Efficient transfection of plasmids into primary human cells is considered a challenge. I decided to use the Amaxa nucleofector technology for transfections because it provided an optimised protocol for transfection of human CD34$^+$ cells. The plasmid vector I used, co-expressed GFP with Cas9 protein therefore, determination of GFP expression by flow cytometry provided a surrogate marker for transfected and Cas9 expressing cells. Initial results with the Amaxa protocol were disappointing with a very low percentage (8%) of GFP expressing cells. However, after incorporating some recently described optimisation steps (Meissner et al., 2014) which included incubating cells for ten minutes in the transfection buffer before transferring to culture medium and growing cells in an antibiotic-free medium for 24 hours, improved the efficiency of transfection to 75% (Figure 7.6).

Then I tested the optimal number of cells required for efficient transfection within a range of 200,000 to 1 million cells (Figure 7.7). The use of 1 million cells produced the best results. Lower cell numbers were associated with reduced cell viability as well as reduced transfection efficiencies and the use of fewer than 500,000 cells resulted in poor and unacceptable results.

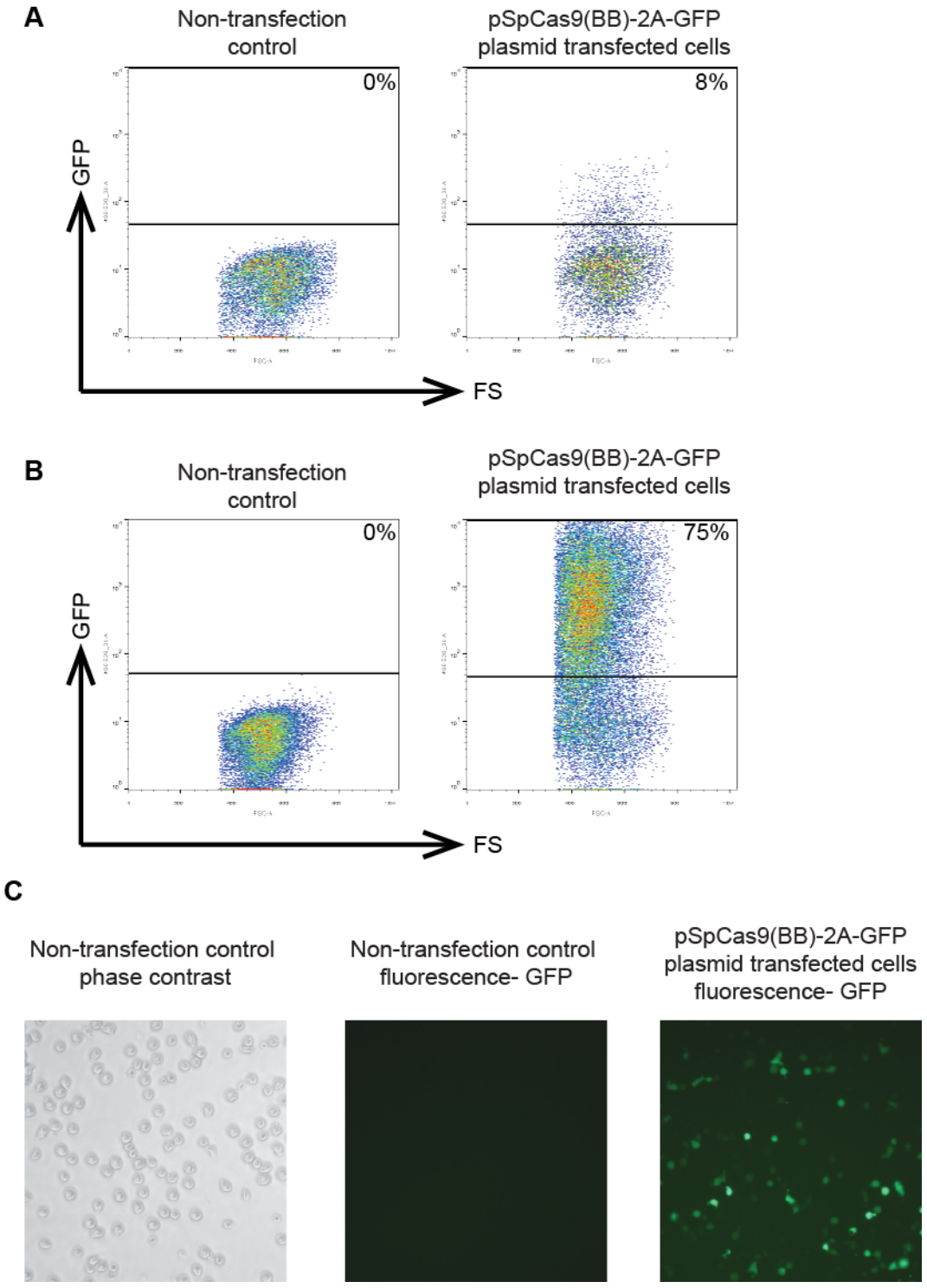


Figure 7.6 – Optimisation and efficiency of the transfection protocol. Human primary CD34$^+$ cells were transfected with pSpCas9[BB]-2A-GFP plasmids and transfection efficiencies were measured by flow cytometry after 24 hours. Flow cytometry plots show GFP expression against forward scatter after gating for live cells. (A) Flow cytometry plots of non-transfection control and plasmid transfected cells using the standard Amaxa protocol. (B) Flow cytometry plots of non-transfection control and plasmid transfected cells using the optimised protocol. (C) Fluorescent microscopy image of non-transfection control and pSpCas9[BB]-2A-GFP plasmid transfected cells using the optimised protocol showing green fluorescence in transfected cells. Exposure times of fluorescence images of control and plasmid transfected cells were constant.

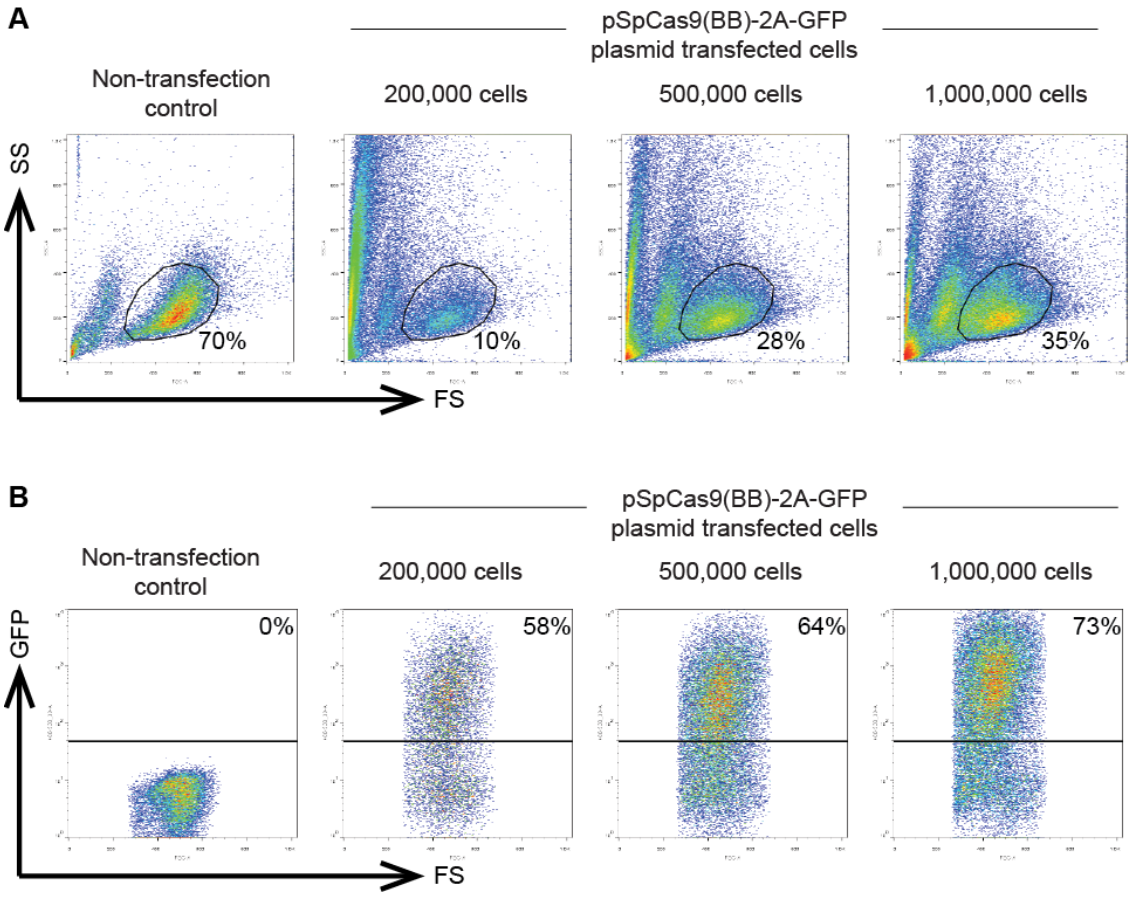


Figure 7.7 – Cell number optimisation for transfection reactions. Human primary CD34+ cells were transfected with pSpCas9[BB]-2A-GFP plasmid and the transfection efficiencies were measured by flow cytometry after 24 hours. (A) Flow cytometry plots showing forward and side scatter of non-transfection control and plasmid transfected cells with different cell numbers. Live cells lie within the polygonal shaped gate and the percentages of cell within that gate are shown. (B) Flow cytometry plots showing GFP expression and forward scatter after gating for live cells in non-transfection control and plasmid transfected cells with different cell numbers.

7.2.4 Erythroid differentiation of CD34+ cells after nucleofection

Nucleofection involves transmission of an electric pulse which might alter the differentiation potential of the CD34+ cells. To examine this, nucleofected CD34+ cells were differentiated along the erythroid lineage using the protocol described in chapter 3 (Figure 7.8). Morphologically, cells which underwent nucleofection were at similar differentiation stages to the non-nucleofection control after 10 and 14 days. This was confirmed by immunophenotyping which

showed similar expression profiles for cell surface markers, CD71, CD235a and CD34 in nucleofected and non-nucleofection control cells. These observations confirm that the nucleofection does not alter the erythroid differentiation potential of the CD34$^+$ cells.

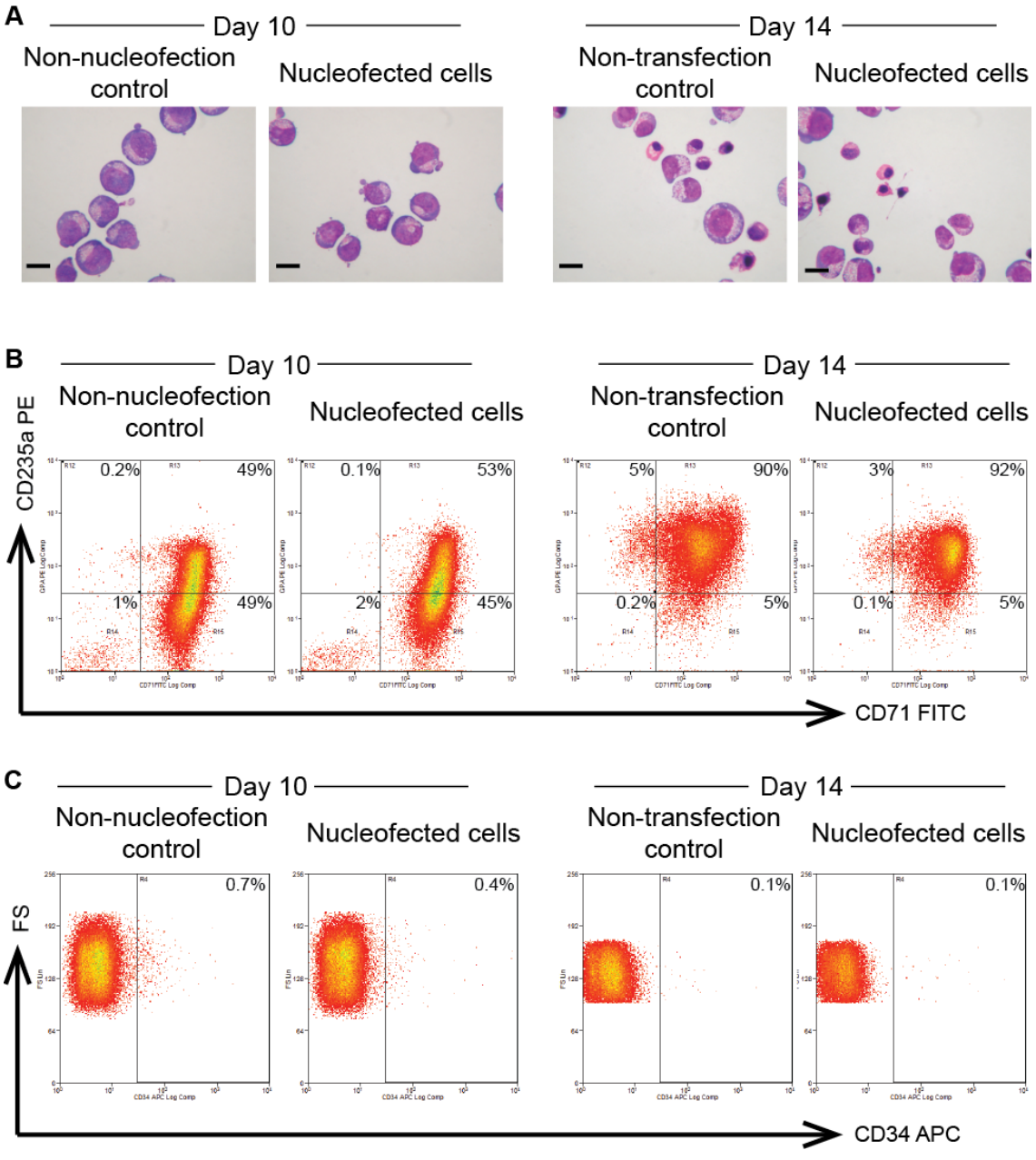


Figure 7.8 – Erythroid differentiation potential of CD34$^+$ cells after nucleofection. CD34$^+$ cells were nucleofected on day 1 and then differentiated for 14 days along the erythroid lineage. (A) Representative cytospins of cells stained by modified Wright stain; scale bar – 10μm. (B) Flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-CD235a antibodies. (C) Flow cytometry plots of cells stained with APC-conjugated anti-CD34 plotted against forward scatter (FS).

7.2.5 Selection of the most efficient CRISPR/Cas9 plasmid pairs

Using CRISPR/Cas9 plasmids with gRNAs targeting either ends of the MCS-R2 region (Figure 7.4), several different combinations of plasmid pairs capable of generating the required deletion could be generated. I tested nine of these pairs to identify the most efficient combinations. CD34$^+$ cells were co-transfected with a CRISPR/Cas9 plasmid pair and sorted by flow cytometry to select for the cells expressing GFP. Sorted cells were expanded for 7 days and then screened using a pair of PCR primers amplifying a region across the MCS-R2 segment. The non-mutated (wild type) allele produced an amplicon of 613bp and depending on the gRNA target site, the mutated alleles produced amplicons of lengths ranging between 302-390bp (Figure 7.9). This screening revealed the variable efficiency of deletion of the different CRISPR pairs which ranged from 5% to 68%. Variable mutation frequencies among different CRISPR plasmids targeting sites even few base pairs apart have been reported before for reasons which are still unclear (Mandal et al., 2014; Ran et al., 2013b). Four CRISPR pairs which demonstrated efficiencies over 35% were selected for the subsequent experiments.

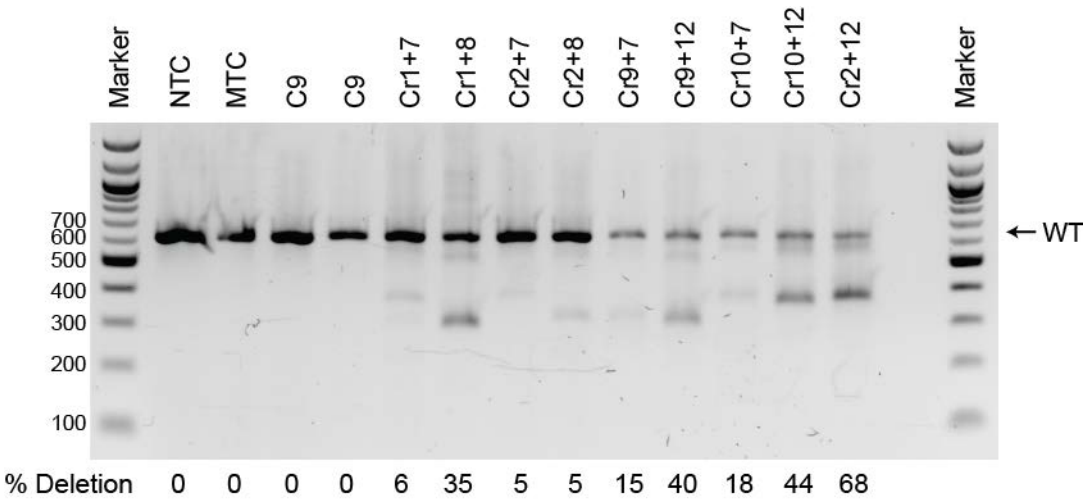


Figure 7.9 – Gel electrophoresis image of populations of CD34+ cells targeted by different CRISPR/Cas9 plasmid pairs analysed by PCR. Wild type (WT) amplicon is 613bp and depending on the gRNA target site, deletions produced amplicons with lengths between 302-390bp. Percentages of mutated alleles determined by band intensity is shown below each lane. Abbreviations: NTC, non-transfection control; MTC, mock transfection control; C9, Cas9-only control.

7.2.6 Deletion of MCS-R2 regulatory element selectively down-regulates α-globin expression

Then, I tested the effects of deletion of the MCS-R2 region on globin gene expression in human erythroid cells. Human CD34+ cells were transfected with four different CRISPR/Cas9 plasmid pairs and the transfected cells were differentiated into erythroid cells. Analysis of genomic DNA from CRISPR/Cas9 targeted cells by PCR showed predominantly mutated alleles with amplicon lengths corresponding to the predicted deletions. Sanger sequencing of the extracted DNA from PCR gel bands verified the exact deletions as predicted by the target sites. Deletion efficiencies ranged between 70%-81% in the four CRISPR pairs tested (Figure 7.10A&B). Protocol optimisation (as described in section 7.2.3) and use of optimal cell numbers for transfections were thought to be the reasons behind improved deletion efficiencies seen in these experiments,

compared to the results shown in the previous section using the same CRISPR pairs.

As expected, deletion of the MCS-R2 region resulted in significant down-regulation of α-globin expression; the mean expression levels for CRISPR/Cas9 targeted cells were between 32%-42% compared to a control (Figure 7.10C). The expression of β-globin gene was unaltered. Consequently, the α/β globin mRNA ratios of CRISPR/Cas9 targeted cells decreased to levels between 28%-40% of normal (Figure 7.10D). These results show that efficient mutation of MCS-R2 region results in significant and selective down-regulation of α-globin expression in human erythroid cells.

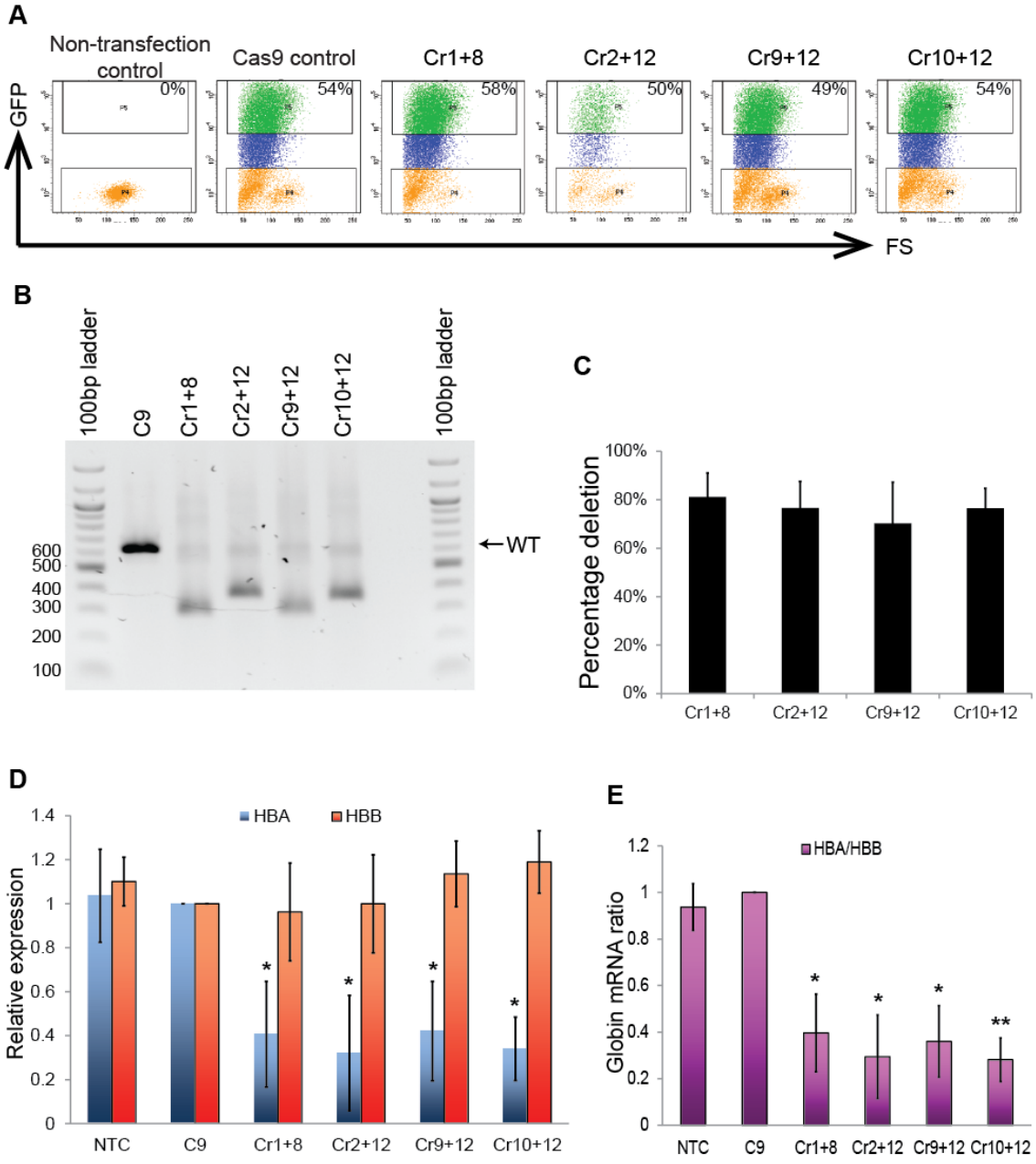


Figure 7.10 – Deletion efficiencies and the effects of MCS-R2 deletion on human globin gene expression. CD34+ cells were nucleofected with CRISPR/Cas9 plasmid pairs on day 1 and sorted transfected cells were differentiated for 9-10 days along the erythroid lineage. (A) Representative flow cytometry plots showing GFP expression and forward scatter (FS) after gating for live cells in non-transfection control and CRISPR/Cas9 plasmid transfected cells. (B) Representative gel electrophoresis image of genomic DNA extracted from cells targeted by four CRISPR/Cas9 plasmid pairs analysed by PCR. (C) Mean percentages of mutated alleles determined by band intensity from 3 independent biological repeats; error bars represent SD. (D) α and β-globin gene expression normalised to the expression of RPL13A and relative to Cas9 control (C9). Mean of 3 independent biological repeats is shown; error bars represent SD; *p<0.05 relative to C9. (E) α/β-globin mRNA ratios analysed by qPCR. Mean of 3 independent biological repeats is shown; error bars represent SD; *p<0.05 and **p<0.01 relative to C9.

7.2.7 Deletion of MCS-R2 regulatory element does not have significant effects on erythroid differentiation

In order to determine whether the CD34$^+$ cells that had undergone deletion of MCS-R2 region differentiate normally into erythroid cells, I analysed the cell morphology and immunophenotypic characteristics of cells that were differentiated for 9-10 days. Cells edited by CRISPR/Cas9 were morphologically indistinguishable to non-transfection and Cas9 controls cells (Figure 7.11A). Furthermore, expression levels of erythroid specific cell surface markers, CD71 and CD235a, were not significantly different in CRISPR/Cas9 transfected cells, Cas9 control cells and non-transfection control cells (Figure 7.11B). These results confirm that the deletion of MCS-R2 region does not have significant adverse effects on human erythroid differentiation.

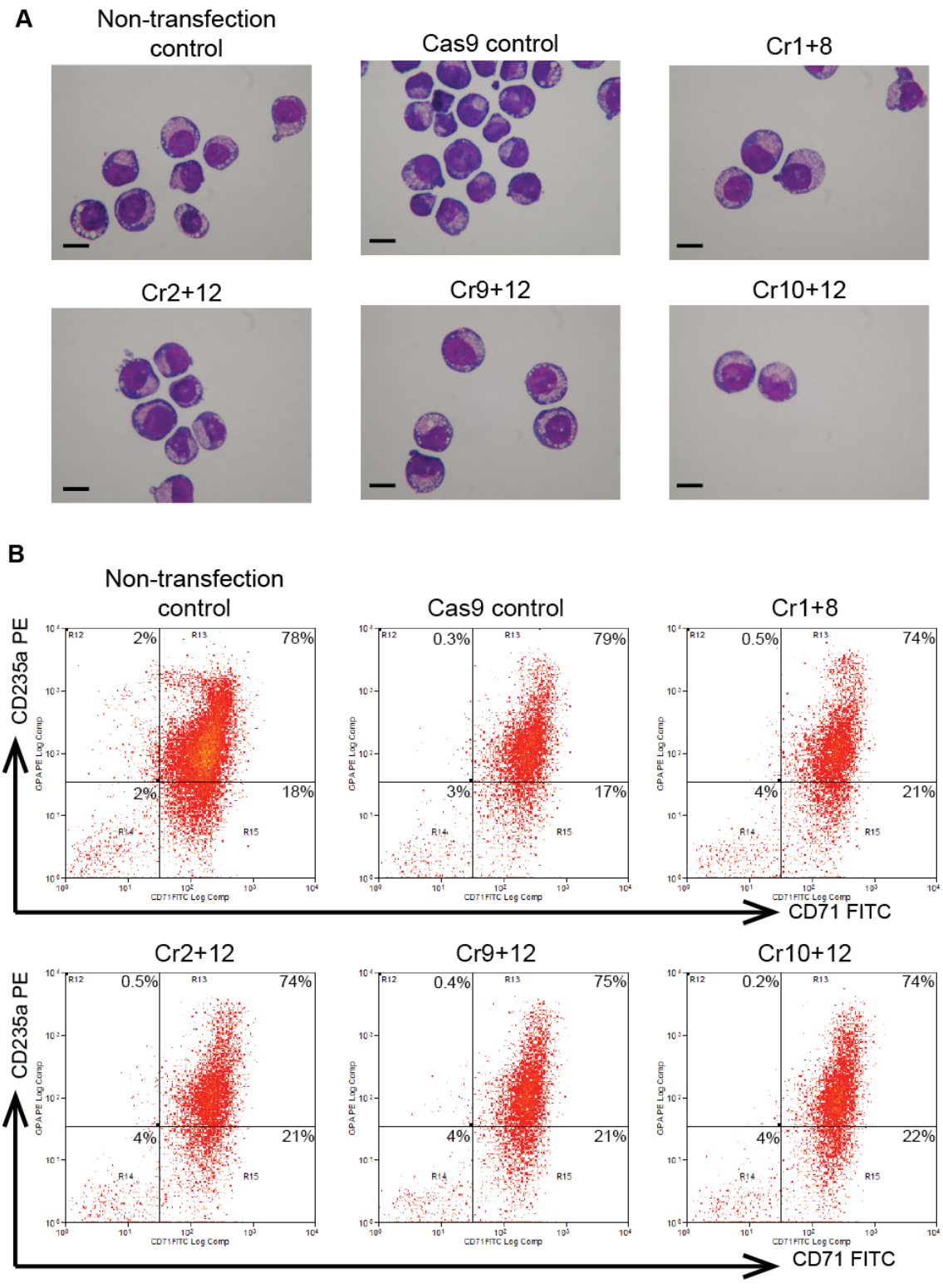


Figure 7.11 – Effects of deletion of MCS-R2 region in human erythroid differentiation. Characterisation of same cell populations which were presented in figure 7.10 is shown here. (A) Representative cytospins of cells stained by modified Wright stain; scale bar – 10μm. (B) Flow cytometry plots of cells stained with FITC-conjugated anti-CD71 and PE-conjugated anti-CD235a antibodies.

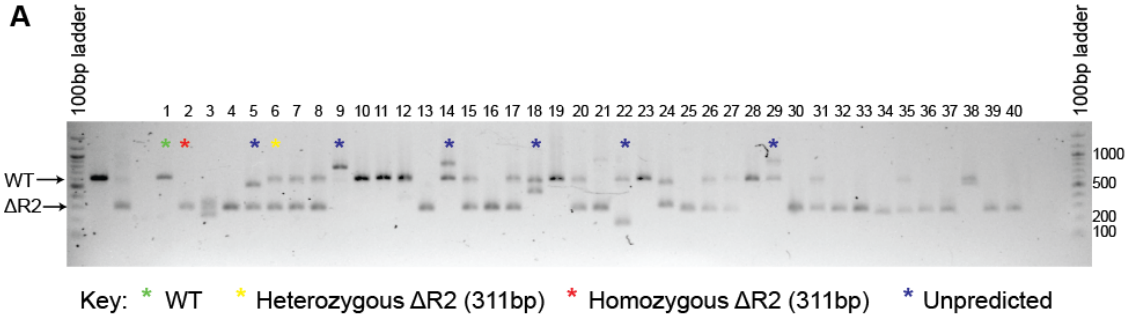
7.2.8 Analysis of effects of deletion of the MCS-R2 regulatory element at the single cell level

Finally, to examine the precise consequences of the MCS-R2 deletion at the single cell level, the alterations in globin gene expression in erythroid cells generated from single CRISPR/Cas9 transfected cells were evaluated. Human CD34$^+$ cells were co-transfected with a pair of CRISPR/Cas9 plasmids and the GFP (and Cas9) expressing cells were sorted into Terasaki multi-well plates as single cell per well. Erythroid cells were generated from the sorted single cells using the protocol described in chapter 3 (section 3.2.9). After 10 days of differentiation, cells from each clone were split into two; half of the cells were used to screen for the mutaion by PCR after uniform amplification of genomic DNA, while the remainder was used for gene expression analysis. Two CRISPR/Cas9 plasmid pairs (cr1+8 and cr2+12) were tested in this manner (Figure 7.12 and 7.13).

Results from both CRISPR/Cas9 pairs were approximately the same and a mixture of clones with different genotypes was found. A higher proportion of clones with homozygous deletions were seen in one CRISPR/Cas9 pair (cr2+12) compared to the other (cr1+8) (47%vs38%). However, the percentages of clones with heterozygous mutations were similar (28%) in both pairs.

One pair (cr1+8) resulted in a higher occurance (15%) of unpredicted bands (marked with blue asterisks in Figure 7.12A) in screening by PCR than the other. Some of these unpredicted bands were further analysed by Sanger

sequencing (Sequencing was performed at the DNA sequencing facility and the sequencing reads were analysed by Chris Fisher of the Weatherall Institute of Molecular Medicine, University of Oxford). The unexpected higher molecular weight band seen in clone 9 (Figure 7.12A) was found to be due to an insertion of an 179bp sequence. Surprisingly this 179bp orphan sequence did not show homology to anywhere in the human genome. The smaller band in clone 22 (Figure 7.12A) was confirmed to be due to a larger 416bp deletion. In the second CRISPR/Cas9 pair (cr2+12), only two (4%) unpredicted bands were seen (marked with blue asterisks in Figure 7.13A). This suggest that CRISPR/Cas9 pair cr2+12 is more selective and specific in generating the desired deletion in human CD34$^+$ cells.



B

Mutation type	Frequency (%) (N=40)
WT	8 (20%)
Heterozygous Δ R2 (311bp)	11 (28%)
Homozygous ΔR2 (311bp)	15 (38%)
Unpredicted	6 (15%)

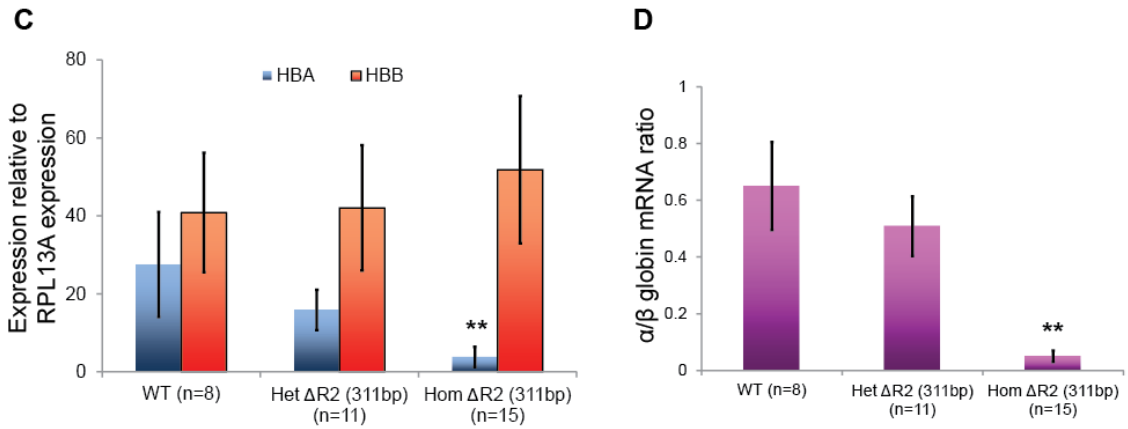
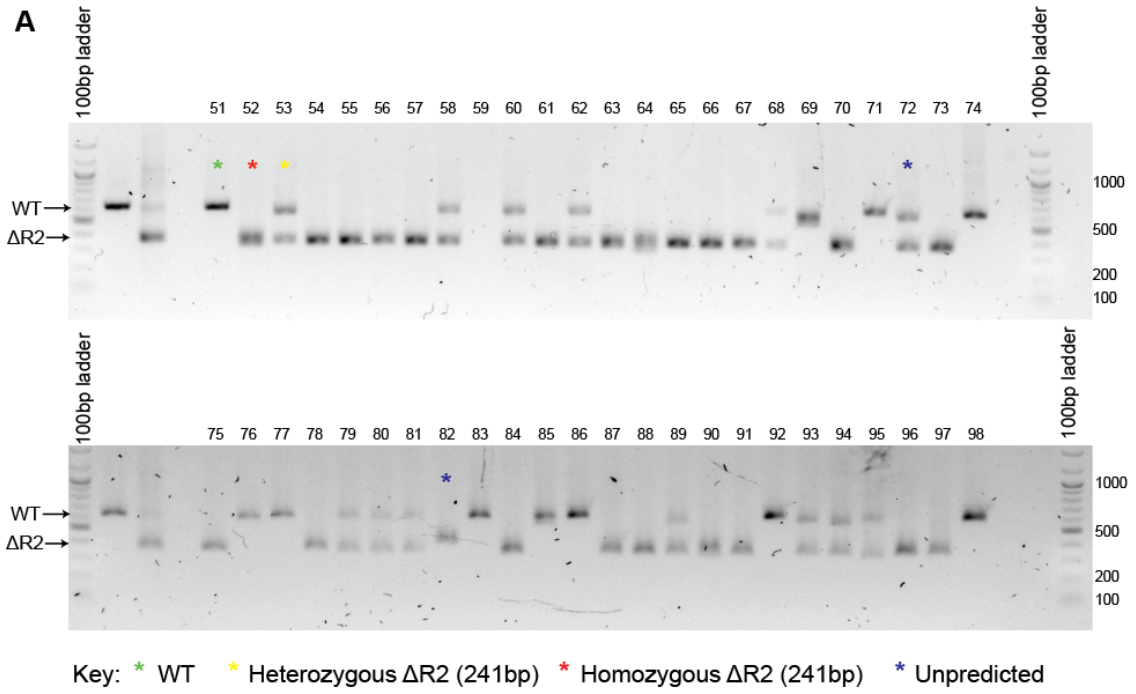


Figure 7.12 –Deletion of MCS-R2 region using CRISPR/Cas9 plasmid pair cr1+8; single cell clone analysis. (A) Gel electrophoresis image of genomic DNA from 40 individual single cell clones (from 3 biological donors) analysed by PCR. The amplicon from the wild type allele is 613bp and the mutated amplicon is 302bp. Clones are numbered 1-40. (B) Frequency of different types of mutations generated. (C) α and β-globin gene expression relative to the expression of RPL13A in erythroid cells which are wild type (WT), heterozygous (HetΔR2) or homozygous (HomΔR2) for a 311bp deletion of MCS-R2 analysed by qPCR; error bars represent standard error of mean (SEM); **p<0.01 relative to WT control. (D) α/β-globin mRNA ratios of erythroid cells which are wild type (WT), heterozygous (HetΔR2) or homozygous (HomΔR2) for a 311bp deletions of MCS-R2 region analysed by qPCR; error bars represent SEM; **p<0.01 relative to WT control.



B

Mutation type	Frequency (%) (N=47)
WT	10 (21%)
Heterozygous ΔR2 (241bp)	13 (28%)
Homozygous ΔR2 (241bp)	22 (47%)
Unpredicted	2 (4%)

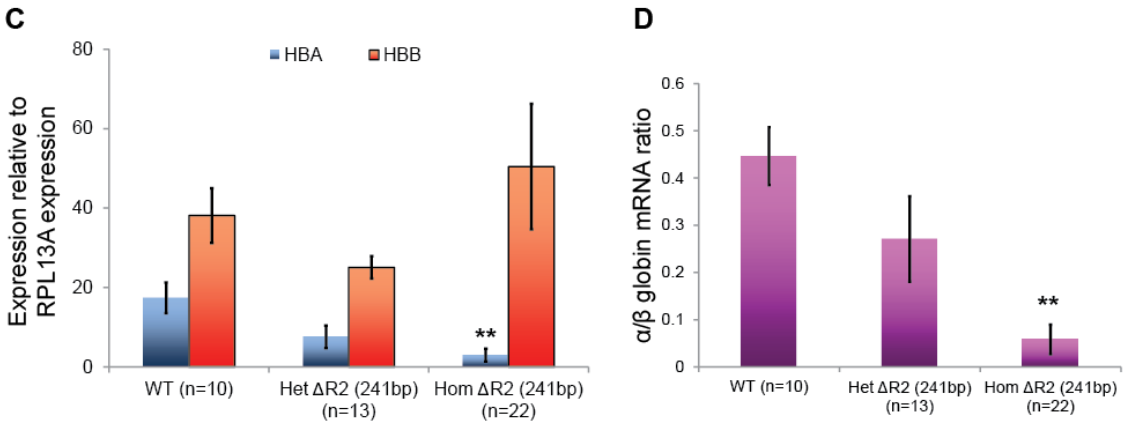


Figure 7.13 – Deletion of MCS-R2 region using CRISPR/Cas9 plasmid pair cr2+12; single cell clone analysis. (A) Gel electrophoresis image of genomic DNA from 48 individual single cell colonies (from 3 biological donors) analysed by PCR. The amplicon from the wild type allele is 613bp and the mutated amplicon is 372bp. Clones are numbered 51-98; DNA from clone 59 failed to amplify. (B) Frequency of different types of mutations generated. (C) α and β-globin gene expression relative to the expression of RPL13A in erythroid cells which are wild type (WT), heterozygous (HetΔR2) or homozygous (HomΔR2) for a 241bp deletion of MCS-R2 analysed by qPCR; error bars represent SEM; **p<0.01 relative to WT control. (D) α/β-globin mRNA ratios of erythroid cells which are wild type (WT), heterozygous (HetΔR2) or homozygous (HomΔR2) for a 241bp deletions of MCS-R2 region analysed by qPCR; error bars represent SEM; **p<0.01 relative to WT control.

Then I analysed the α- and β-globin gene expression levels in each clone using qRTPCR. Mean levels of α-globin expression in clones with heterozygous deletions of MCS-R2 were 57% (Cr1+8) and 43% (Cr2+12) compared to the means of WT clones (Figure 7.12C and Figure 7.13C). In the homozygous state, the α-globin expressions were significantly reduced and were 14% (Cr1+8) and 17% (Cr2+12) compared to WT controls. As expected the expression levels of the β-globin did not show significant changes. Although the α-globin down regulations in homozygous state were more pronounced than predicted from patient data, these findings confirm that the selective silencing of α-globin can be effectively achieved in human CD34$^+$ cells by deletion of MCS-R2 segment using the CRISPR/Cas9 genome editing technique.

7.3 Discussion and future work

The successful utilisation of a long-known bacterial adaptive immune mechanism, CRISPR/Cas9 system for genome editing in early 2013 revolutionised many aspects of biomedical research and is considered to be one of the biggest game changes in science (Cong et al., 2013; Jinek et al., 2012; Ledford, 2015; Mali et al., 2013). It is a powerful tool for precise editing of the genome in a convenient, cheap and highly efficient way (Ran et al., 2013b). Here, we have demonstrated the utilisation of this new technique to disrupt a non-coding regulatory element of the genome in human CD34$^+$ cells in order to produce predictable and selective modifications in gene expression.

The results of this work are important in several ways. Firstly, we have demonstrated editing of primary human CD34$^+$ cells, which is the same type of cells that needs to be targeted in patients. This is of particular importance as most proof of principle studies are being carried out in immortalised cell lines, which behave differently to primary cells and the results obtained by these studies pose challenges for clinical translation (Cox et al., 2015). Secondly, we achieved high transfection efficiencies (75%) in human primary CD34$^+$ cells, which have been traditionally considered as ‘difficult to transfect’ by non-viral methods. Our transfection efficiencies were higher than what has been reported previously using similar protocols (Meissner et al., 2014). This could be due to a better vector selection, superior culture conditions pre- and post-transfection, technical factors related to nucleofection or a combination of all.

Thirdly, we achieved very high rates of mutation frequencies in these cells. Robust and reproducible results were obtained using four different CRISPR/Cas9 plasmid pairs which gave rise to mutations at comparable high rates (70-81%). This was also confirmed by single cell assays, which showed that 80% of the clones are mutated. This was much higher than the deletion efficiencies achieved by Mandal and colleague, describing a similar dual gRNA strategy to disrupt the CCR5 gene in human CD34$^+$ cells (Mandal et al., 2014). In that study, the rate of heterozygous mutations was between 13%-19% and the homozygous deletions was between 19%-26% using three different CRISPR/Cas9 plasmid pairs. In their approach, gRNAs and Cas9 were expressed in two different plasmids which were co-transfected into the cells whereas in our study, we cloned gRNA in to the same plasmid which contained Cas9. This might explain the improved deletion rates seen in our study. However, variable efficiency rates between CRISPR/Cas9 plasmids are well documented for reasons still poorly understood, despite the proposal of several models by a number of groups (Ren et al., 2014; Xu et al., 2015).

Taken together, a transfection efficiency of 75% and a deletion efficiency of 80%, provide an overall mutation rate of 60% in the live CD34$^+$ cells. This is higher than most of the recent genome editing studies involving primary human cells. Studies utilising ZFN based approaches to disrupt CCR5 gene, reported allele mutation frequencies of ~50% in primary human T lymphocytes (Perez et al., 2008) and just 17% in human CD34$^+$ cells (Holt et al., 2010). In β-thalassaemia, Persons and colleagues showed that chimerism levels of only 10% to 20% of normal HSCs are sufficient to result in nearly complete

haematological correction of β-thalassaemia mice (Arumugam and Malik, 2010; Persons et al., 2001); erythroid cells generated from normal HSCs had a selective advantage over the diseased cells, probably due to less ineffective erythropoiesis. Therefore, using our strategy, it is likely that we will be able to provide sufficient numbers of genome edited CD34$^+$ cells that would produce clinically significant beneficial effects in patients with β-thalassaemia.

As hypothesised, deletion of the MCS-R2 enhancer resulted in down-regulation of α-globin gene expression. Although this has been previously demonstrated in knock-out mice (Wallace et al., 2007) and in patients with naturally occurring mutations (Coelho et al., 2010), this is the first time that these results were recapitulated therapeutically in primary human erythroid cells. The deletion of this critical enhancer element resulted in selective and significant down-regulation of α-globin expression albeit, more pronouncedly than expected from the patient with homozygous MCS-R2 deletion. Importantly, neither nucleofection nor ΔMCS-R2 resulted in unwanted effects on erythroid differentiation.

However, several challenges still remain if CRISPR/Cas9 mediated genome editing techniques are to be used in patients. The specificity of CRISPR/Cas9 and the potential risks of off-target effects are the main safety concerns for clinical application (Corrigan-Curay et al., 2015). In fact, at present, no consensus opinion is available even about the techniques which should be used to evaluate off-targets. A common approach used is, to screen genomic sites with high similarity to the target sequences. However, the scope of

possible off-target sites evaluated by these is limited to computationally predicted sites (Cox et al., 2015). More recently, Tsai and others described a novel method, GUIDE-seq for unbiased assessment of genome-wide off-target editing by labelling double strand breaks through the incorporation of a “reporter” sequence (Tsai et al., 2015). Surprisingly, the majority of sites identified by this method were not predicted by existing computational methods. Although, I have not evaluated the off-target effects of editing of MCS-R2, we plan to undertake an un-biased approach to test the off-target effects in the future.

Another outstanding issue before any clinical application is the determination of long-term repopulating potential of genome edited CD34$^+$ cells. In patients with β-thalassaemia, to achieve transfusion-independency and sustainable clinical benefits through genome editing, a significant proportion of long-term repopulating HSCs (LT-HSCs) should undergo editing. Although, the engraftment and self–renewal potential of CD34$^+$ cells after genome editing by ZNF has been demonstrated before (Genovese et al., 2014), to my knowledge, the same has not yet been shown in CD34$^+$ cells edited using CRISPR/Cas9 system. To examine this, we generated mouse xenografts, by transplanting genome edited human CD34$^+$ cells to immuno-compromised irradiated mice. These mice are awaiting analysis in the near future.

Another strategy that we aim to undertake in the future, is to narrow down the sequence of disruption within the MCS-R2 region. In erythroid cells, the core element of MCS-R2 demonstrates binding of essential erythroid transcription factors, GATA1 and NF-E2 (De Gobbi et al., 2007). Therefore, we hypothesise

that disruption of a single TF binding site could down-regulate α-globin expression similar to what was observed in the larger deletion. To test this, we have already designed, synthesised and cloned six CRISPR/Cas9 plasmids to target the TF binding sites individually. If this strategy is successful, we would be able to down-regulate the α-globin expression to levels beneficial in β-thalassaemia, using a single edit, rather that two double strand breaks created in the current approach. This would also decrease the potential off-target effects and might ameliorate the slight overshooting in α-globin knockdown which we observed at the homozygous ΔMCS-R2 state.

In conclusion, thus far I have demonstrated that highly efficient targeted mutation of the MCS-R2 regulatory element in human CD34$^+$ cells is feasible using the CRISPR/Cas9 system and as seen in a patient study, this mutation results in a significant and selective down-regulation of α-globin expression.

Chapter 8: General discussion and conclusions

8.1 Selective silencing of human α-globin is therapeutically feasible

Despite being one of the first molecular diseases identified and its pathophysiology characterised, β-thalassaemia still remains an essentially fatal condition without an effective cure in a large proportion of patients with severe forms. Cellular and molecular data established over many years, without doubt, points towards the un-balanced excess production of α-globin chains as the main factor which triggers a cascade of events that leads to anaemia in patients with β-thalassaemia (chapter 1; section 1.3.3). Similarly, a significant number of high-quality clinical studies have highlighted how a natural reduction of α-globin output in the form of α-thalassaemia improves disease severity in patients with β-thalassaemia (Mettananda et al., 2015). Despite the availability of clear clinical, genetic and molecular data for over three decades, with the exception of very few studies (Voon et al., 2008; Xie et al., 2011), the feasibility of therapeutic down-regulation of α-globin as a treatment option for β-thalassaemia has not been evaluated before. The emphasis has been centred on various mechanisms involved in the reactivation of γ-globin gene and the alternative strategy of silencing the α-globin gene has been overlooked. In fact, none of the recent reviews which discuss alternative therapeutic approaches for β-thalassaemia have mentioned this pathway (Rivella, 2015; Sankaran and Weiss, 2015; Vichinsky, 2012).

In this thesis, I have explored two pathways which could selectively down-regulate α-globin expression in human erythroid cells, in order to facilitate the conversion of this long-standing clinical observation into an effective therapy towards β-thalassaemia. The available clinical data suggests that a reduction in α-globin expression to 75%-50% of the normal level provides the optimal beneficial effects for the majority of these patents (chapter 1; section 1.5.4). Furthermore, reductions down to 25% of the normal levels (as seen in patients with α-thalassaemia due to three α-globin gene deletions) do not result in significant morbidity due to anaemia and can be considered as a safe tolerable lower limit. With these therapeutic levels and contrasting regulation of α-and β-globin genes in mind, I explored two novel and promising pathways that could bring about selective down-regulation of α-globin expression. Firstly, the epigenetic inhibitor small molecule screen identified two compounds which produced favourable results; a broad range KDM inhibitor, IOX1 and an HDAC inhibitor, vorinostat. Secondly, the engineering of a series of deletions involving the MCS-R2 enhancer (ΔMCS-R2) using highly efficient, targeted genome editing by CRISPR/Cas9 system in human erythroid cells resulted in selective down-regulation of α-globin expression to the desirable levels. In this way, I have demonstrated that selective silencing of α-globin expression in human erythroid cells to the levels beneficial in patients with β-thalassaemia is therapeutically feasible (figure 8.1).

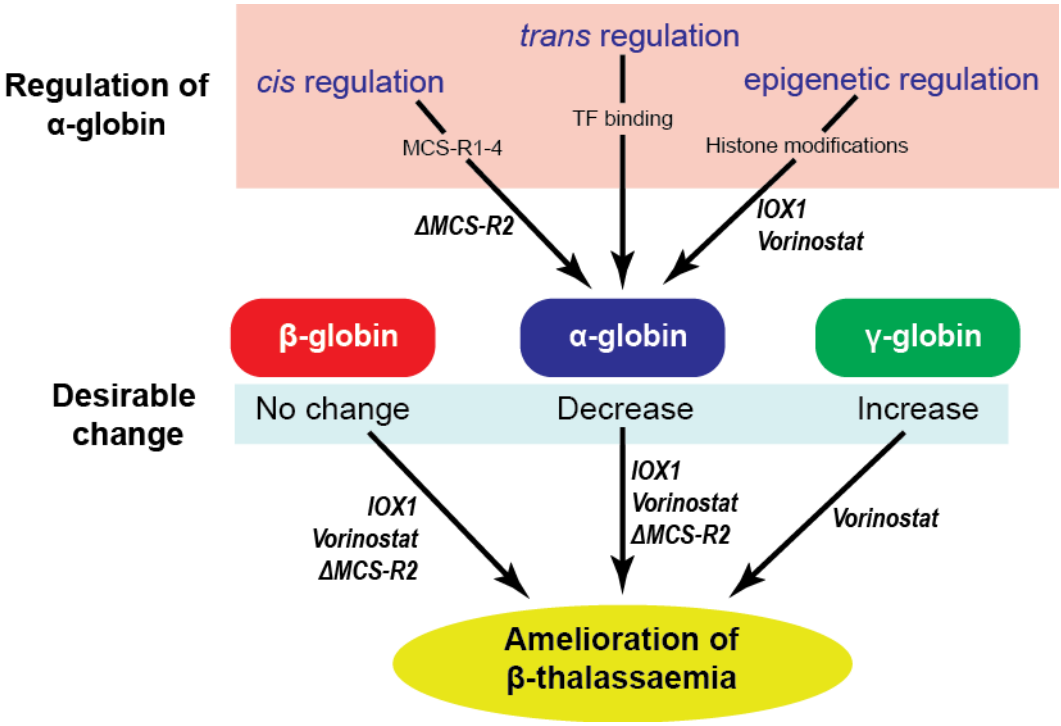


Figure 8.1 – A summary of pharmacological and genome editing approaches, which aid in amelioration of β-thalassaemia uncovered in this thesis. Possible sites/mechanisms of action and the desirable changes in globin gene expression employed by each approach are shown.

8.2 Limitations of the study

One important limitation of this study is that the down-regulation of α-globin expression has been demonstrated only at the mRNA level and not tested at the protein level. In erythroid cells, α-, β-, and γ-globin chains are found as tetramers in haemoglobin molecules and the analysis of individual globin chains is challenging. Most of the routine methods used in the analysis of dissociated globin chains which include cellulose acetate electrophoresis, polyacrylamide gel electrophoresis and capillary electrophoresis, provide qualitative or semi-quantitative information but an accurate quantification of individual globin chains using these techniques is difficult (Wajcman and Riou, 2009). In a high

background level of globin proteins, it would be extremely hard to assess the changes in α-globin expression precisely using these assays, immunoblotting, immunofluorescence or flow cytometry based assays.

This is in contrast to the situation in studies which evaluate the small molecules and mechanisms inducing γ-globin expression. Up-regulation of γ-globin is accurately reflected at protein level as a relative change in HbF level, which can be quantified with precision and ease using cation-exchange high-performance liquid chromatography (CE-HPLC) (Steinberg and Nagel, 2009). Conversely, all the main haemoglobin isoforms which are detected by CE-HPLC (HbA, HbF and HbA_2) contain α-globin, and changes in α-globin expression would not be reflected by the relative levels of these isoforms.

The standard globin chain synthesis analysis using carboxymethyl cellulose chromatography described by Weatherall and Clegg can provide very reliable quantification of globin chain ratios. However, it requires large numbers of cells producing high amounts of globin proteins, hence is not feasible for *in vitro* cultured erythroblasts (Weatherall et al., 1965). Reversed-phase high-performance liquid chromatography (RP-HPLC) is another technique that is used to quantify individual globin chains and has successfully been utilised in the past to assess the relative levels of γ-globin isoforms and screen for thalassaemia syndromes (Galanello et al., 1998; Wajcman and Riou, 2009). We attempted the optimisation of an RP-HPLC protocol to accurately quantify the levels of α-globin chains in cultured erythroblast; however, despite several attempts this was not successful.

Although the effects were not studied at protein level, in most of the experiments quantification of globin mRNA was done by at least two methods which included Fluidigm Biomark, standard qRTPCR and the sensitive Nanostring technology. Data obtained using these independent methods were comparable, thus validating our results. In patients with thalassaemias, α/β globin mRNA ratios reflect the changes seen in α/β globin chain protein ratios, haematological indices and disease severity (Chaisue et al., 2007; Higgs, 2009). Furthermore there is no known regulatory mechanism which controls mRNA translation which is specific to α-globin and which might confound our RNA based analysis. Therefore it is reasonable to postulate that the effects seen at mRNA level reflect the changes at protein level.

Another important limitation of our strategy is that achieving the exact optimal balance of α/non-α globin chain ratios can be challenging therefore in most of the patients this treatment would result in a milder phenotype and not a complete ‘cure’. In other words, patients who have transfusion-dependent β-thalassaemia major would be converted to thalassaemia intermedia states. Although thalassaemia intermedia is a markedly less severe disease compared to thalassaemia major, number of long-term medical complications which include iron overload, hypercoagulability, pulmonary hypertension, gall stones and bone deformities are associated with this form of the disease (Musallam et al., 2013a; Viprakasit et al., 2014). However in my opinion our strategy would aid a significant proportion of patients with β-thalassaemia to live longer better quality lives until a permanent cure to this disease is developed.

8.3 Epigenetic inhibitor therapy to down-regulate α-globin: therapeutic prospects and challenges

Epigenetic therapies in the treatment of haemoglobinopathies have been investigated in recent years and the results from the very first clinical trial which tested the effect of vorinostat on SCD were recently published (Okam et al., 2015). The epigenetic enzymes which are currently being targeted include HDACs, LSD1, and DNMTs, all of which are supposed to work through increasing HbF production via de-repressing the silenced γ-globin. The primary aim of these therapies is to ameliorate SCD, in which an induction of HbF has been proven to be clinically beneficial (Charache et al., 1995; Platt et al., 1984). In fact, the licenced HbF inducing drug, hydroxyurea is a recognised pharmacological treatment modality for SCD at present (Platt, 2008).

However, clinical evidence of disease amelioration of β-thalassaemia by the therapeutic induction of γ-globin and HbF is controversial (Musallam et al., 2013b). The largest body of evidence is provided by studies which evaluated the effectiveness of hydroxyurea in β-thalassaemia. The results from these studies are erratic and taken together they failed to demonstrate consistent beneficial effects either by increasing total haemoglobin levels or decreasing transfusion requirements. Of note, in a large proportion of these failures in improvement in haematological outcomes were in spite of increases in HbF levels (Banan, 2013; Musallam et al., 2013b). Therefore, convincing evidence is still lacking to back the idea that novel epigenetic therapies which induce HbF will provide beneficial effects in patients with β-thalassaemia.

The therapeutic aim of inducing γ-globin in β-thalassaemia is to increase the availability of β-like globins to mop up the excess free α-globin chains which promote haemolysis and ineffective erythropoiesis (Musallam et al., 2013b). In contrast, we tested a more direct approach of decreasing the pool of excess α-globin chains by down-regulating α-globin itself. Our epigenetic inhibitor small molecule screen identified IOX1 and vorinostat as promising compounds. These two compounds have different mechanisms of actions, alter globin gene expression differently and have been shown to have synergistic effects. IOX1 exhibited the best effect in down-regulating α-globin expression but vorinostat, in addition to its action of down-regulating α-globin, also up-regulated γ-globin. In this way, vorinostat may be more efficient *in vivo* at decreasing the unwanted effects of free α-globin chains.

The widely accepted optimal chemical properties of drugs intended for oral dosing include, high potency against targets (<100nM), low molecular weight (<450Da), high solubility (>100μM) and lipophilicity (measured as partition coefficient; LogP<4) (Yusof et al., 2014). The favourable outcomes in human globin gene expression of IOX1 were seen at a dose range of 20μM-40μM suggesting low potency. Although the potency was sub-optimal, IOX1 possesses favourable biochemical properties which include low molecular weight (189Da), high solubility (up to 10mM) and low LogP values (1.29) (SGC, 2015). Similarly, even at higher concentrations (40μM), IOX1 did not affect erythroid cell viability or differentiation nor did transcriptome analysis suggest off target effects confirming the low toxicity of this agent in erythroid cells *in vitr*o. These observations suggest that IOX1 is a suitable lead compound to

synthesize derivatives with higher potency and better medicinal properties for use as oral medicines in patients with β-thalassaemia to selectively down-regulate α-globin expression.

In contrast, vorinostat is already a licenced medication for other indications and has an effective oral dosage form. Therefore, the progress of vorinostat to clinic to treat patients with β-thalassaemia would be quicker and easier. In fact, as already mentioned, it has been tested in a phase-1 clinical trial for SCD (Okam et al., 2015). In this trial, only 1 out of 5 patients demonstrated an induction of HbF, however, the expression levels of α-globin have not been tested. Therefore, with the new evidence presented in this thesis, it is highly justifiable to carry-out phase 1/2 clinical trials to test vorinostat in patients with β-thalassaemia to evaluate the effects on α-globin mRNA levels, percentages of HbF, haemoglobin levels, other haematological parameters and transfusion requirements.

Despite tremendous advances, many general challenges still remain in epigenetic drug targeting. Most of the epigenetic proteins are associated with thousands of genes throughout the genome and hence could have numerous off-target effects including tumourgenesis (Helin and Dhanak, 2013). For instance, mutations in components of PRC2 promote lymphoid transformation, myelodysplastic syndrome and leukaemia (Helin and Dhanak, 2013). Furthermore, PRC2 has been shown to suppresses genes involved in differentiation, cell-cycle, self-renewal and apoptosis in HSCs and when components of PRC2 are mutated, HSCs failed to differentiate to mature blood

cells and were prone to cell death (Xie et al., 2014). Therefore, any short-term beneficial effects demonstrated by epigenetic drugs should be interpreted with caution and widespread use of these drugs in non-malignant conditions should be backed by meticulous safety data obtained by careful and long-term follow-up in well-controlled phase 3/4 clinical trials.

8.4 Genome editing to down regulate α-globin: therapeutic prospects and challenges

The recent advances in the development of targeted genome editing technologies based on programmable nucleases have substantially improved the ability to make precise changes in the genome. In fact, during the past few years genome editing has been in pre-clinical stage experimentation for β-thalassaemia. Again this was solely to explore mechanisms which upregulate γ-globin and is predominantly through targeting BCL11A. A well-known biotechnology company, Sangamo biosciences, is currently developing ZFN-based genome editing tools to precisely disrupt the expression of the transcription factor BCL11A (SangamoBiosciences, 2015). However, BCL11A is a transcription factor which has essential functions in number of non-erythroid cells including lymphocytes and neurones therefore, the disruption of this could be problematic (Sankaran and Weiss, 2015).

Targeted gene therapy is another feasible approach that can be used in patients with β-thalassaemia. This involves the utilisation of programmable nucleases to introduce a copy of the wild-type β-globin gene into a

predetermined locus of the genome, identified as a ‘safe harbour’, whose disruption does not lead to discernible phenotypic effects (Cox et al., 2015). Many problems encountered by traditional gene therapy approaches related to random integration of the vector DNA, including development of leukaemia by activation of proto-oncogenes, can be avoided by using this method (Chandrakasan and Malik, 2014). However, gene insertion by this approach utilises the homology directed repair (HDR) pathway of DNA repair, which is much less efficient and occurs at a lower frequency than the non-homologous end joining (NHEJ) mediated repair which is used in gene deletion or disruption.

The genome editing approach used in this thesis is distinct to the above approaches. We engineered a targeted deletion at the MCS-R2 regulatory element of the human α-globin gene to achieve a selective down-regulation of α-globin expression. Our approach reproducibly achieved very high deletion efficiencies (70%-81%) in HSPCs, at rates which are not possible by approaches utilising HDR, at least with the technologies available at present. Furthermore, contrary to the situation of BCL11A, a large amount of data, including a similar deletion in individuals only affected by mild anaemia, suggest that the MCS-R2 element is an erythroid specific enhancer of α-globin with no known functions in other tissues or genes hence, the disruption of this element is unlikely to have off-target effects. Therefore, ΔMCS-R2 in HSCs is superior in efficiency, selectivity and safety to the other current experimental genome editing approaches for β-thalassaemia.

However, the lack of specificity and the potential for off-target effects of CRISPR/Cas9 genome editing tools still remains a major challenge (Calero-Garcia and Gaspar, 2014). Genetic modifications are permanent, and deleterious off-target mutations might create cells with oncogenic potential or functional impairment (Cox et al., 2015). To overcome this limitation, several novel and sensitive strategies have been developed to screen for the potential off-targets effects which are poorly predicted by computational methods (Crosetto et al., 2013; Frock et al., 2015; Tsai et al., 2015). In addition, use of Cas9 nickase that primarily generates double-strand breaks by creating two separate single-strand breaks on opposite DNA strands via the expression of two separate gRNAs is known to decrease the off-target effects (Ran et al., 2013a) . Although many issues remain outstanding, given the current momentum of CRISPR/Cas9 research, the move to clinical application of genome editing technologies in patients with β-thalassaemia is unlikely to be far away.

8.5 Conclusions

In this thesis, I have demonstrated that selective silencing of α-globin expression in human erythroid cells to the levels beneficial in patients with β-thalassaemia is therapeutically feasible. Utilisation of several pathways was uncovered. Pharmacologically, epigenetic inhibitor compounds, IOX1 and vorinostat demonstrated beneficial and synergistic effects. Similarly, targeted genome editing of the MCS-R2 enhancer of the human α-globin gene provided equally favourable results. Further progress on these pathways including lead

optimisation of IOX1, early phase clinical trials with vorinostat and preclinical evaluation of genome editing techniques undoubtedly have the potential to add strength to the armoury of clinicians treating patients with β-thalassaemia or in fact lead to a ‘cure’. This provides a very clear example of *bedside to bench to bedside* research, demonstrating how the clinical data helps to understand the basic physiology of gene regulation and the pathophysiology of a disease and the improved knowledge gained from basic research can be translated back to the clinic.

References

Amid, A., Saliba, A.N., Taher, A.T., and Klaassen, R.J. (2015). Thalassaemia in children: from quality of care to quality of life. Arch Dis Child Published Online First: doi:10.1136/archdischild-2014-308112.

Angelucci, E. (2010). Hematopoietic stem cell transplantation in thalassemia. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2010*, 456-462.

Angelucci, E., and Baronciani, D. (2008). Allogeneic stem cell transplantation for thalassemia major. Haematologica *93*, 1780-1784.

Angelucci, E., Matthes-Martin, S., Baronciani, D., Bernaudin, F., Bonanomi, S., Cappellini, M.D., Dalle, J.H., Di Bartolomeo, P., de Heredia, C.D., Dickerhoff, R.*, et al.* (2014a). Hematopoietic stem cell transplantation in thalassemia major and sickle cell disease: indications and management recommendations from an international expert panel. Haematologica *99*, 811-820.

Angelucci, E., Srivastava, A., and Usai, S. (2014b). Haematopoietic Stem Cell Transplantation. In Guidelines for the management of transfusion dependent thalassaemia, M. Cappellini, A. Cohen, J. Porter, A. Taher, and V. Viprakasit, eds. (Cyprus: Thalassaemia International Federation).

Arlet, J.B., Ribeil, J.A., Guillem, F., Negre, O., Hazoume, A., Marcion, G., Beuzard, Y., Dussiot, M., Moura, I.C., Demarest, S.*, et al.* (2014). HSP70 sequestration by free alpha-globin promotes ineffective erythropoiesis in beta-thalassaemia. Nature *514*, 242-246.

Arumugam, P., and Malik, P. (2010). Genetic therapy for beta-thalassemia: from the bench to the bedside. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2010*, 445-450.

Aydinok, Y., Kattamis, A., and Viprakasit, V. (2014). Current approach to iron chelation in children. Br J Haematol *165*, 745-755.

Badens, C., Joly, P., Agouti, I., Thuret, I., Gonnet, K., Fattoum, S., Francina, A., Simeoni, M.C., Loundou, A., and Pissard, S. (2011). Variants in genetic modifiers of beta-thalassemia can help to predict the major or intermedia type of the disease. Haematologica *96*, 1712-1714.

Banan, M. (2013). Hydroxyurea treatment in beta-thalassemia patients: to respond or not to respond? Annals of hematology *92*, 289-299.

Bandyopadhyay, S., Roychowdhury, K., Chandra, S., Das, M., and Dasgupta, U.B. (2001). Variable severity of beta-thalassemia patients of eastern India: effect of alpha-thalassemia and xmnI polymorphism. Clinical and experimental medicine *1*, 155-159.

Beris, P., Solenthaler, M., Deutsch, S., Darbellay, R., Tobler, A., Bochaton-Pialat, M.L., and Gabbiani, G. (1999). Severe inclusion body beta-thalassaemia with haemolysis in a patient double heterozygous for beta(0)-thalassaemia and quadruplicated alpha-globin gene arrangement of the anti-4.2 type. Br J Haematol *105*, 1074-1080.

Best, J.D., and Carey, N. (2010). Epigenetic therapies for non-oncology indications. Drug Discov Today *15*, 1008-1014.

Bleicher, K.H., Bohm, H.J., Muller, K., and Alanine, A.I. (2003). Hit and lead generation: beyond high-throughput screening. Nature reviews Drug discovery *2*, 369-378.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. Science *330*, 612-616.

Borg, J., Papadopoulos, P., Georgitsi, M., Gutierrez, L., Grech, G., Fanis, P., Phylactides, M., Verkerk, A.J., van der Spek, P.J., Scerri, C.A.*, et al.* (2010). Haploinsufficiency for the erythroid transcription factor KLF1 causes hereditary persistence of fetal hemoglobin. Nature genetics *42*, 801-805.

Borg, J., Patrinos, G.P., Felice, A.E., and Philipsen, S. (2011). Erythroid phenotypes associated with KLF1 mutations. Haematologica *96*, 635-638.

Borgna-Pignatti, C., Rugolotto, S., De Stefano, P., Zhao, H., Cappellini, M.D., Del Vecchio, G.C., Romeo, M.A., Forni, G.L., Gamberini, M.R., Ghilardi, R.*, et al.* (2004). Survival and complications in patients with thalassemia major treated with transfusion and deferoxamine. Haematologica *89*, 1187-1193.

Bourantas, K., Economou, G., and Georgiou, J. (1997). Administration of high doses of recombinant human erythropoietin to patients with beta-thalassemia intermedia: a preliminary trial. European journal of haematology *58*, 22-25.

Bradai, M., Abad, M.T., Pissard, S., Lamraoui, F., Skopinski, L., and de Montalembert, M. (2003). Hydroxyurea can eliminate transfusion requirements in children with severe beta-thalassemia. Blood *102*, 1529-1530.

Bradner, J.E., Mak, R., Tanguturi, S.K., Mazitschek, R., Haggarty, S.J., Ross, K., Chang, C.Y., Bosco, J., West, N., Morse, E.*, et al.* (2010). Chemical genetic strategy identifies histone deacetylase 1 (HDAC1) and HDAC2 as therapeutic targets in sickle cell disease. Proceedings of the National Academy of Sciences of the United States of America *107*, 12617-12622.

Brown, K.E., Amoils, S., Horn, J.M., Buckle, V.J., Higgs, D.R., Merkenschlager, M., and Fisher, A.G. (2001). Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. Nature cell biology *3*, 602-606.

Calero-Garcia, M., and Gaspar, H.B. (2014). Gene-ectomy: gene ablation with CRISPR/Cas9 in human hematopoietic cells. Cell stem cell *15*, 529-530.

Camaschella, C., Mazza, U., Roetto, A., Gottardi, E., Parziale, A., Travi, M., Fattore, S., Bacchiega, D., Fiorelli, G., and Cappellini, M.D. (1995). Genetic interactions in thalassemia intermedia: analysis of beta-mutations, alpha-genotype, gamma-promoters, and beta-LCR hypersensitive sites 2 and 4 in Italian patients. American journal of hematology *48*, 82-87.

Cantor, A.B., and Orkin, S.H. (2002). Transcriptional regulation of erythropoiesis: an affair involving multiple partners. Oncogene *21*, 3368-3376.

Cao, A., and Galanello, R. (2010). Beta-thalassemia. Genetics in medicine : official journal of the American College of Medical Genetics *12*, 61-76.

Carnero, A. (2006). High throughput screening in drug discovery. Clinical & translational oncology : official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico *8*, 482-490.

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K.*, et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. Nature *467*, 318-322.

Chaisue, C., Kitcharoen, S., Wilairat, P., Jetsrisuparb, A., Fucharoen, G., and Fucharoen, S. (2007). alpha/beta-Globin mRNA ratio determination by multiplex quantitative real-time reverse transcription-polymerase chain reaction as an indicator of globin gene function. Clinical biochemistry *40*, 1373-1377.

Chan, K.S., Xu, J., Wardan, H., McColl, B., Orkin, S., and Vadolas, J. (2012). Generation of a genomic reporter assay system for analysis of gamma- and beta-globin gene regulation. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *26*, 1736-1744.

Chandrakasan, S., and Malik, P. (2014). Gene therapy for hemoglobinopathies: the state of the field and the future. Hematology/oncology clinics of North America *28*, 199-216.

Chandy, M. (2008). Stem cell transplantation in India. Bone marrow transplantation 42 Suppl 1, S81-S84.

Chang, K.H., Fang, X., Wang, H., Huang, A., Cao, H., Yang, Y., Bonig, H., Stamatoyannopoulos, J.A., and Papayannopoulou, T. (2013). Epigenetic modifications and chromosome conformations of the beta globin locus throughout development. Stem cell reviews *9*, 397-407.

Charache, S., Terrin, M.L., Moore, R.D., Dover, G.J., Barton, F.B., Eckert, S.V., Mcmahon, R.P., Bonds, D.R., Orringer, E., Jones, S*., et al.* (1995). Effect of Hydroxyurea on the Frequency of Painful Crises in Sickle-Cell-Anemia. New Engl J Med *332*, 1317-1322.

Chen, H., Lowrey, C.H., and Stamatoyannopoulos, G. (1997). Analysis of enhancer function of the HS-40 core sequence of the human alpha-globin cluster. Nucleic acids research *25*, 2917-2922.

Cheng, J.K., and Alper, H.S. (2014). The genome editing toolbox: a spectrum of approaches for targeted modification. Current opinion in biotechnology *30C*, 87-94.

Cheung, J.O.P., Casals-Pascual, C., Roberts, D.J., and Watt, S.M. (2007). A small-scale serum-free liquid cell culture model of erythropoiesis to assess the effects of exogenous factors. J Immunol Methods *319*, 104-117.

Coates, T.D. (2014). Physiology and pathophysiology of iron in hemoglobin-associated diseases. Free radical biology & medicine *72*, 23-40.

Coelho, A., Picanco, I., Seuanes, F., Seixas, M.T., and Faustino, P. (2010). Novel large deletions in the human alpha-globin gene cluster: Clarifying the HS-40 long-range regulatory role in the native chromosome environment. Blood cells, molecules & diseases *45*, 147-153.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A*., et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Cooley, T.B., and Lee, P. (1925). A Series of Cases of Splenomegaly in Children, with Anemia and Peculiar Bone Changes. Transactions of the American Pediatric Society *37*, 29-30.

Cooley, T.B., Witwer, E.R., and Lee, P. (1927). Anemia in children with splenomegaly and peculiar changes in bones; report of cases. The American journal of diseases of children *34*, 347.

Corrigan-Curay, J., O'Reilly, M., Kohn, D.B., Cannon, P.M., Bao, G., Bushman, F.D., Carroll, D., Cathomen, T., Joung, J.K., Roth, D*., et al.* (2015). Genome editing technologies: defining a path to clinic. Molecular therapy : the journal of the American Society of Gene Therapy *23*, 796-806.

Cox, D.B., Platt, R.J., and Zhang, F. (2015). Therapeutic genome editing: prospects and challenges. Nature medicine *21*, 121-131.

Crosetto, N., Mitra, A., Silva, M.J., Bienko, M., Dojer, N., Wang, Q., Karaca, E., Chiarle, R., Skrzypczak, M., Ginalski, K.*, et al.* (2013). Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. Nature methods *10*, 361-365.

Danjou, F., Anni, F., and Galanello, R. (2011). Beta-thalassemia: from genotype to phenotype. Haematologica *96*, 1573-1575.

Davidson, B.L., and McCray, P.B., Jr. (2011). Current prospects for RNA interference-based therapies. Nature reviews Genetics *12*, 329-340.

De Gobbi, M., Anguita, E., Hughes, J., Sloane-Stanley, J.A., Sharpe, J.A., Koch, C.M., Dunham, I., Gibbons, R.J., Wood, W.G., and Higgs, D.R. (2007). Tissue-specific histone modification and transcription factor binding in alpha globin gene expression. Blood *110*, 4503-4510.

De Gobbi, M., Garrick, D., Lynch, M., Vernimmen, D., Hughes, J.R., Goardon, N., Luc, S., Lower, K.M., Sloane-Stanley, J.A., Pina, C.*, et al.* (2011). Generation of bivalent chromatin domains during cell fate decisions. Epigenetics & chromatin *4*, 9.

Deng, Y., Wang, C.C., Choy, K.W., Du, Q., Chen, J., Wang, Q., Li, L., Chung, T.K., and Tang, T. (2014). Therapeutic potentials of gene silencing by RNA interference: principles, challenges, and new strategies. Gene *538*, 217-227.

Dong, A., Rivella, S., and Breda, L. (2013). Gene therapy for hemoglobinopathies: progress and challenges. Translational research : the journal of laboratory and clinical medicine *161*, 293-306.

Dorn, I., Lazar-Karsten, P., Boie, S., Ribbat, J., Hartwig, D., Driller, B., Kirchner, H., and Schlenke, P. (2008). In vitro proliferation and differentiation of human CD34+ cells from peripheral blood into mature red blood cells with two different cell culture systems. Transfusion *48*, 1122-1132.

Driscoll, M.C., Dobkin, C.S., and Alter, B.P. (1989). Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. Proceedings of the National Academy of Sciences of the United States of America *86*, 7470-7474.

Dussiot, M., Maciel, T.T., Fricot, A., Chartier, C., Negre, O., Veiga, J., Grapton, D., Paubelle, E., Payen, E., Beuzard, Y.*, et al.* (2014). An activin receptor IIA ligand trap corrects ineffective erythropoiesis in beta-thalassemia. Nature medicine *20*, 398-407.

Duvic, M., Talpur, R., Ni, X., Zhang, C., Hazarika, P., Kelly, C., Chiao, J.H., Reilly, J.F., Ricker, J.L., Richon, V.M.*, et al.* (2007). Phase 2 trial of oral vorinostat (suberoylanilide hydroxamic acid, SAHA) for refractory cutaneous T-cell lymphoma (CTCL). Blood *109*, 31-39.

Dzierzak, E. (2011). Stem cells and haemopoiesis. In Postgraduate Haematology, A. Hoffbrand, D. Catovsky, E. Tuddenham, and A. Gree, eds. (Wiley).

Efremov, D.G., Dimovski, A.J., Sukarova, E., Schiliro, G., Zisovski, N., Efremov, G.D., Burrone, O.R., and Huisman, T.H. (1994). gamma-mRNA and Hb F levels in beta-thalassaemia. Br J Haematol *88*, 311-317.

Fathallah, H., and Atweh, G.F. (2006). Induction of fetal hemoglobin in the treatment of sickle cell disease. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program, 58-62.

Fessas, P. (1963). Inclusions of hemoglobin erythroblasts and erythrocytes of thalassemia. Blood *21*, 21-32.

Fessas, P., Loukopoulos, D., and Kaltsoya, A. (1966). Peptide analysis of the inclusions of erythroid cells in beta-thalassemia. Biochimica et biophysica acta *124*, 430-432.

Fessas, P., Loukopoulos, D., and Thorell, B. (1965). Absorption Spectra of Inclusion Bodies in Beta-Thalassemia. Blood *25*, 105-109.

Fessas, P., Stamatoyannopoulos, G., and Karaklis, A. (1961). Hereditary persistence of foetal haemoglobin and its combination with alpha and beta-thalassaemia. Paper presented at: 8th Congress of European Society of Haematology (Karger, Basel, Vianna).

Fibach, E., Burke, L.P., Schechter, A.N., Noguchi, C.T., and Rodgers, G.P. (1993). Hydroxyurea increases fetal hemoglobin in cultured erythroid cells derived from normal individuals and patients with sickle cell anemia or beta-thalassemia. Blood *81*, 1630-1635.

Fibach, E., Manor, D., Oppenheim, A., and Rachmilewitz, E.A. (1989). Proliferation and Maturation of Human Erythroid Progenitors in Liquid Culture. Blood *73*, 100-103.

Fibach, E., and Prus, E. (2005). Differentiation of human erythroid cells in culture. Current protocols in immunology / edited by John E Coligan [et al] *Chapter 22*, Unit 22F 27.

Finotti, A., Breda, L., Lederer, C.W., Bianchi, N., Zuccato, C., Kleanthous, M., Rivella, S., and Gambari, R. (2015). Recent trends in the gene therapy of beta-thalassemia. Journal of blood medicine *6*, 69-85.

Frock, R.L., Hu, J., Meyers, R.M., Ho, Y.J., Kii, E., and Alt, F.W. (2015). Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat Biotechnol *33*, 179-186.

Fucharoen, S., and Weatherall, D.J. (2012). The hemoglobin E thalassemias. Cold Spring Harbor perspectives in medicine *2*.

Fujisawa, T., Takeda, K., and Ichijo, H. (2007). ASK family proteins in stress response and disease. Molecular biotechnology *37*, 13-18.

Galanello, R., Dessi, E., Melis, M.A., Addis, M., Sanna, M.A., Rosatelli, C., Argiolu, F., Giagu, N., Turco, M.P., Cacace, E.*, et al.* (1989). Molecular analysis of beta zero-thalassemia intermedia in Sardinia. Blood *74*, 823-827.

Galanello, R., and Origa, R. (2010). Beta-thalassemia. Orphanet journal of rare diseases *5*, 11.

Galanello, R., Sanna, S., Perseu, L., Sollaino, M.C., Satta, S., Lai, M.E., Barella, S., Uda, M., Usala, G., Abecasis, G.R.*, et al.* (2009). Amelioration of Sardinian beta0 thalassemia by genetic modifiers. Blood *114*, 3935-3937.

Galanello, R., Satta, S., Pirroni, M.G., Travi, M., and Maccioni, L. (1998). Globin chain synthesis analysis by high performance liquid chromatography in the screening of thalassemia syndromes. Hemoglobin *22*, 501-508.

Galarneau, G., Palmer, C.D., Sankaran, V.G., Orkin, S.H., Hirschhorn, J.N., and Lettre, G. (2010). Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. Nature genetics *42*, 1049-1051.

Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67-71.

Garrick, D., De Gobbi, M., Samara, V., Rugless, M., Holland, M., Ayyub, H., Lower, K., Sloane-Stanley, J., Gray, N., Koch, C.*, et al.* (2008). The role of the polycomb complex in silencing alpha-globin gene expression in nonerythroid cells. Blood *112*, 3889-3899.

Gavrilov, K., and Saltzman, W.M. (2012). Therapeutic siRNA: principles, challenges, and strategies. The Yale journal of biology and medicine *85*, 187-200.

Geiss, G.K., Bumgarner, R.E., Birditt, B., Dahl, T., Dowidar, N., Dunaway, D.L., Fell, H.P., Ferree, S., George, R.D., Grogan, T.*, et al.* (2008). Direct multiplexed measurement of gene expression with color-coded probe pairs. Nat Biotechnol *26*, 317-325.

Genovese, P., Schiroli, G., Escobar, G., Di Tomaso, T., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C.*, et al.* (2014). Targeted genome editing in human repopulating haematopoietic stem cells. Nature *510*, 235-240.

Giarratana, M.C., Kobari, L., Lapillonne, H., Chalmers, D., Kiger, L., Cynober, T., Marden, M.C., Wajcman, H., and Douay, L. (2005). Ex vivo generation of fully mature human red blood cells from hematopoietic stem cells. Nat Biotechnol *23*, 69-74.

Gringras, P., Wonke, B., Old, J., Fitches, A., Valler, D., Kuan, A.M., and Hoffbrand, V. (1994). Effect of alpha thalassaemia trait and enhanced gamma chain production on disease severity in beta thalassaemia major and intermedia. Archives of disease in childhood *70*, 30-34.

Haldane, J.B.S. (1949). The rate of mutation of human genes. Hereditas *35*, 7.

Hannon, G.J. (2002). RNA interference. Nature *418*, 244-251.

Harteveld, C.L., and Higgs, D.R. (2010). Alpha-thalassaemia. Orphanet journal of rare diseases *5*, 13.

Hatton, C.S., Wilkie, A.O., Drysdale, H.C., Wood, W.G., Vickers, M.A., Sharpe, J., Ayyub, H., Pretorius, I.M., Buckle, V.J., and Higgs, D.R. (1990). Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human alpha globin gene cluster. Blood *76*, 221-227.

Hay, D. (2014). The role of distal regulatory elements in α-globin transcription (University of Oxford), pp. 206.

Hebbel, R.P., Vercellotti, G.M., Pace, B.S., Solovey, A.N., Kollander, R., Abanonu, C.F., Nguyen, J., Vineyard, J.V., Belcher, J.D., Abdulla, F.*, et al.* (2010). The HDAC inhibitors trichostatin A and suberoylanilide hydroxamic acid exhibit multiple modalities of benefit for the vascular pathobiology of sickle transgenic mice. Blood *115*, 2483-2490.

Heinemann, B., Nielsen, J.M., Hudlebusch, H.R., Lees, M.J., Larsen, D.V., Boesen, T., Labelle, M., Gerlach, L.O., Birk, P., and Helin, K. (2014). Inhibition of demethylases by GSK-J1/J4. Nature *514*, E1-E2.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A.*, et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics *39*, 311-318.

Heinz, S., Romanoski, C.E., Benner, C., and Glass, C.K. (2015). The selection and function of cell type-specific enhancers. Nature reviews Molecular cell biology *16*, 144-154.

Helin, K., and Dhanak, D. (2013). Chromatin proteins and modifications as drug targets. Nature *502*, 480-488.

Higgs, D., and Wood, W. (2011). Erythropoiesis. In Postgraduate Haematology, A. Hoffbrand, D. Catovsky, E. Tuddenham, and A. Green, eds. (NJ: Wiley-Blackwell), pp. 12-25.

Higgs, D.R. (2009). The Pathophysiology and Clinical Features of alpha-thalassemia. In Disorders of Hemoglobin, M.H. Steinberg, B.G. Forget, D.R. Higgs, and D.J. Weatherall, eds. (New York: Cambridge University Press), pp. 266-295.

Higgs, D.R., Engel, J.D., and Stamatoyannopoulos, G. (2012). Thalassaemia. Lancet *379*, 373-383.

Higgs, D.R., Garrick, D., Anguita, E., De Gobbi, M., Hughes, J., Muers, M., Vernimmen, D., Lower, K., Law, M., Argentaro, A.*, et al.* (2005). Understanding alpha-globin gene regulation: Aiming to improve the management of thalassemia. Annals of the New York Academy of Sciences *1054*, 92-102.

Higgs, D.R., Sharpe, J.A., and Wood, W.G. (1998). Understanding alpha globin gene expression: a step towards effective gene therapy. Seminars in hematology *35*, 93-104.

Higgs, D.R., and Wood, W.G. (2008). Long-range regulation of alpha globin gene expression during erythropoiesis. Current opinion in hematology *15*, 176-183.

Higgs, D.R., Wood, W.G., Jarman, A.P., Sharpe, J., Lida, J., Pretorius, I.M., and Ayyub, H. (1990). A major positive regulatory region located far upstream of the human alpha-globin gene locus. Genes & development *4*, 1588-1601.

Ho, P.J., Hall, G.W., Luo, L.Y., Weatherall, D.J., and Thein, S.L. (1998). Beta-thalassaemia intermedia: is it possible consistently to predict phenotype from genotype? Br J Haematol *100*, 70-78.

Hojfeldt, J.W., Agger, K., and Helin, K. (2013). Histone lysine demethylases as targets for anticancer therapy. Nature reviews Drug discovery *12*, 917-930.

Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C.*, et al.* (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. Nat Biotechnol *28*, 839-847.

Hopkinson, R.J., Tumber, A., Yapp, C., Chowdhury, R., Aik, W., Che, K.H., Li, X.S., Kristensen, J.B.L., King, O.N.F., Chan, M.C.*, et al.* (2013). 5-Carboxy-8-hydroxyquinoline is a broad spectrum 2-oxoglutarate oxygenase inhibitor which causes iron translocation. Chem Sci *4*, 3110-3117.

Hughes, J.R., Cheng, J.F., Ventress, N., Prabhakar, S., Clark, K., Anguita, E., De Gobbi, M., de Jong, P., Rubin, E., and Higgs, D.R. (2005). Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. Proceedings of the National Academy of Sciences of the United States of America *102*, 9830-9835.

Hughes, J.R., Lower, K.M., Dunham, I., Taylor, S., De Gobbi, M., Sloane-Stanley, J.A., McGowan, S., Ragoussis, J., Vernimmen, D., Gibbons, R.J.*, et al.* (2013). High-resolution analysis of cis-acting regulatory networks at the alpha-globin locus. Philosophical transactions of the Royal Society of London Series B, Biological sciences *368*, 20120361.

Iskander, D., Psaila, B., Gerrard, G., Chaidos, A., En Foong, H., Harrington, Y., Karnik, L.C., Roberts, I., de la Fuente, J., and Karadimitris, A. (2015). Elucidation of the EP defect in Diamond-Blackfan anemia by characterization and prospective isolation of human EPs. Blood 125, 2553-2557.

Italia, K.Y., Jijina, F.J., Merchant, R., Panjwani, S., Nadkarni, A.H., Sawant, P.M., Nair, S.B., Ghosh, K., and Colah, R.B. (2009). Response to hydroxyurea in beta thalassemia major and intermedia: experience in western India. Clinica chimica acta; international journal of clinical chemistry *407*, 10-15.

Ivanov, M., Barragan, I., and Ingelman-Sundberg, M. (2014). Epigenetic mechanisms of importance for drug treatment. Trends in pharmacological sciences *35*, 384-396.

Iwamoto, M., Friedman, E.J., Sandhu, P., Agrawal, N.G., Rubin, E.H., and Wagner, J.A. (2013). Clinical pharmacology profile of vorinostat, a histone deacetylase inhibitor. Cancer chemotherapy and pharmacology *72*, 493-508.

Janzen, W.P. (2014). Screening technologies for small molecule discovery: the state of the art. Chem Biol *21*, 1162-1170.

Jarman, A.P., Wood, W.G., Sharpe, J.A., Gourdon, G., Ayyub, H., and Higgs, D.R. (1991). Characterization of the major regulatory element upstream of the human alpha-globin gene cluster. Molecular and cellular biology *11*, 4679-4689.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816-821.

Johansson, C., Tumber, A., Che, K., Cain, P., Nowak, R., Gileadi, C., and Oppermann, U. (2014). The roles of Jumonji-type oxygenases in human disease. Epigenomics *6*, 89-120.

Kan, Y.W., and Nathan, D.G. (1970). Mild thalassemia: the result of interactions of alpha and beta thalassemia genes. The Journal of clinical investigation *49*, 635-642.

Kanavakis, E., Traeger-Synodinos, J., Lafioniatis, S., Lazaropoulou, C., Liakopoulou, T., Paleologos, G., Metaxotou-Mavrommati, A., Stamoulakatou, A., and Papassotiriou, I. (2004). A rare example that coinheritance of a severe form of beta-thalassemia and alpha-thalassemia interact in a "synergistic" manner to balance the phenotype of classic thalassemic syndromes. Blood cells, molecules & diseases 32, 319-324.

Kanasty, R., Dorkin, J.R., Vegas, A., and Anderson, D. (2013). Delivery materials for siRNA therapeutics. Nature materials *12*, 967-977.

Kautz, L., Jung, G., Du, X., Gabayan, V., Chapman, J., Nasoff, M., Nemeth, E., and Ganz, T. (2015). Erythroferrone contributes to hepcidin suppression and iron overload in a mouse model of beta-thalassemia. Blood.

Kautz, L., Jung, G., Valore, E.V., Rivella, S., Nemeth, E., and Ganz, T. (2014). Identification of erythroferrone as an erythroid regulator of iron metabolism. Nature genetics *46*, 678-684.

Khan, O., and La Thangue, N.B. (2008). Drug Insight: histone deacetylase inhibitor-based therapies for cutaneous T-cell lymphomas. Nature clinical practice Oncology *5*, 714-726.

Khandros, E., Thom, C.S., D'Souza, J., and Weiss, M.J. (2012). Integrated protein quality-control pathways regulate free alpha-globin in murine beta-thalassemia. Blood *119*, 5265-5275.

Kim, A., and Dean, A. (2012). Chromatin loop formation in the beta-globin locus and its role in globin gene transcription. Molecules and cells *34*, 1-5.

Kim, H., and Kim, J.S. (2014). A guide to genome engineering with programmable nucleases. Nature reviews Genetics *15*, 321-334.

Kim, Y.K., and Chang, Y.T. (2007). Tagged library approach facilitates forward chemical genetics. Molecular bioSystems *3*, 392-397.

Klose, R.J., Cooper, S., Farcas, A.M., Blackledge, N.P., and Brockdorff, N. (2013). Chromatin sampling--an emerging perspective on targeting polycomb repressor proteins. PLoS genetics *9*, e1003717.

Kruidenier, L., Chung, C.W., Cheng, Z.J., Liddle, J., Che, K.H., Joberty, G., Bantscheff, M., Bountra, C., Bridges, A., Diallo, H.*, et al.* (2012). A selective jumonji H3K27 demethylase inhibitor modulates the proinflammatory macrophage response. Nature *488*, 404-408.

Ledford, H. (2015). CRISPR, the disruptor. Nature *522*, 20-24.

Lee, M.G., Wynder, C., Schmidt, D.M., McCafferty, D.G., and Shiekhattar, R. (2006). Histone H3 lysine 4 demethylation is a target of nonselective antidepressive medications. Chem Biol *13*, 563-567.

Letvin, N.L., Linch, D.C., Beardsley, G.P., McIntyre, K.W., and Nathan, D.G. (1984). Augmentation of fetal-hemoglobin production in anemic monkeys by hydroxyurea. The New England journal of medicine *310*, 869-873.

Ley, T.J., DeSimone, J., Anagnou, N.P., Keller, G.H., Humphries, R.K., Turner, P.H., Young, N.S., Keller, P., and Nienhuis, A.W. (1982). 5-azacytidine selectively increases gamma-globin synthesis in a patient with beta+ thalassemia. The New England journal of medicine *307*, 1469-1475.

Li, B., Ding, L., Li, W., Story, M.D., and Pace, B.S. (2012). Characterization of the transcriptome profiles related to globin gene switching during in vitro erythroid maturation. BMC genomics *13*, 153.

Li, H., Rybicki, A.C., Suzuka, S.M., von Bonsdorff, L., Breuer, W., Hall, C.B., Cabantchik, Z.I., Bouhassira, E.E., Fabry, M.E., and Ginzburg, Y.Z. (2010). Transferrin therapy ameliorates disease in beta-thalassemic mice. Nature medicine *16*, 177-182.

Liakopoulou, E., Blau, C.A., Li, Q., Josephson, B., Wolf, J.A., Fournarakis, B., Raisys, V., Dover, G., Papayannopoulou, T., and Stamatoyannopoulos, G. (1995). Stimulation of fetal hemoglobin production by short chain fatty acids. Blood *86*, 3227-3235.

Libani, I.V., Guy, E.C., Melchiori, L., Schiro, R., Ramos, P., Breda, L., Scholzen, T., Chadburn, A., Liu, Y., Kernbach, M.*, et al.* (2008). Decreased differentiation of erythroid cells exacerbates ineffective erythropoiesis in beta-thalassemia. Blood *112*, 875-885.

LifestyleProTheme (2015), pp. http://antranik.org/blood-components-hemoglobin-typerh-factor-agglutination/.

Liu, D., Zhang, X., Yu, L., Cai, R., Ma, X., Zheng, C., Zhou, Y., Liu, Q., Wei, X., Lin, L.*, et al.* (2014). KLF1 mutations are relatively more common in a thalassemia endemic region and ameliorate the severity of beta-thalassemia. Blood *124*, 803-811.

Locatelli, F., Kabbara, N., Ruggeri, A., Ghavamzadeh, A., Roberts, I., Li, C.K., Bernaudin, F., Vermylen, C., Dalle, J.H., Stein, J.*, et al.* (2013). Outcome of patients with hemoglobinopathies given either cord blood or bone marrow transplantation from an HLA-identical sibling. Blood *122*, 1072-1078.

Lucarelli, G., Isgro, A., Sodani, P., and Gaziev, J. (2012). Hematopoietic stem cell transplantation in thalassemia and sickle cell anemia. Cold Spring Harbor perspectives in medicine *2*, a011825.

Lynch, M.D., Smith, A.J., De Gobbi, M., Flenley, M., Hughes, J.R., Vernimmen, D., Ayyub, H., Sharpe, J.A., Sloane-Stanley, J.A., Sutherland, L.*, et al.* (2012). An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. The EMBO journal *31*, 317-329.

Macarron, R., Banks, M.N., Bojanic, D., Burns, D.J., Cirovic, D.A., Garyantes, T., Green, D.V., Hertzberg, R.P., Janzen, W.P., Paslay, J.W.*, et al.* (2011). Impact of

high-throughput screening in biomedical research. Nature reviews Drug discovery *10*, 188-195.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823-826.

Manchinu, M.F., Marongiu, M.F., Poddie, D., Casu, C., Latini, V., Simbula, M., Galanello, R., Moi, P., Cao, A., Porcu, S., et al. (2014). In vivo activation of the human delta-globin gene: the therapeutic potential in beta-thalassemic mice. Haematologica 99, 76-84.

Mandal, P.K., Ferreira, L.M., Collins, R., Meissner, T.B., Boutwell, C.L., Friesen, M., Vrbanac, V., Garrison, B.S., Stortchevoi, A., Bryder, D.*, et al.* (2014). Efficient ablation of genes in human hematopoietic stem and effector cells using CRISPR/Cas9. Cell stem cell *15*, 643-652.

Meissner, T.B., Mandal, P.K., Ferreira, L.M., Rossi, D.J., and Cowan, C.A. (2014). Genome editing for human gene therapy. Methods in enzymology *546*, 273-295.

Menzel, S., Garner, C., Gut, I., Matsuda, F., Yamaguchi, M., Heath, S., Foglio, M., Zelenika, D., Boland, A., Rooks, H.*, et al.* (2007). A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nature genetics *39*, 1197-1199.

Merryweather-Clarke, A.T., Atzberger, A., Soneji, S., Gray, N., Clark, K., Waugh, C., McGowan, S.J., Taylor, S., Nandi, A.K., Wood, W.G.*, et al.* (2011). Global gene expression analysis of human erythroid progenitors. Blood *117*, E96-E108.

Mesa, R.A., and Cortes, J. (2013). Optimizing management of ruxolitinib in patients with myelofibrosis: the need for individualized dosing. Journal of hematology & oncology *6*, 79.

Mettananda, S., Gibbons, R.J., and Higgs, D.R. (2015). alpha-Globin as a molecular target in the treatment of beta-thalassemia. Blood *125*, 3694-3701.

Micelli, C., and Rastelli, G. (2015). Histone deacetylases: structural determinants of inhibitor selectivity. Drug Discov Today *20*, 718-735.

Michlitsch, J., Azimi, M., Hoppe, C., Walters, M.C., Lubin, B., Lorey, F., and Vichinsky, E. (2009). Newborn screening for hemoglobinopathies in California. Pediatric blood & cancer *52*, 486-490.

Migliaccio, G., Di Pietro, R., di Giacomo, V., Di Baldassarre, A., Migliaccio, A.R., Maccioni, L., Galanello, R., and Papayannopoulou, T. (2002). In vitro mass production of human erythroid cells from the blood of normal donors and of thalassemic patients. Blood cells, molecules & diseases *28*, 169-180.

Muller, S., and Brown, P.J. (2012). Epigenetic Chemical Probes. Clin Pharmacol Ther *92*, 689-693.

Musallam, K.M., Rivella, S., Vichinsky, E., and Rachmilewitz, E.A. (2013a). Non-transfusion-dependent thalassemias. Haematologica 98, 833-844.

Musallam, K.M., Taher, A.T., Cappellini, M.D., and Sankaran, V.G. (2013b). Clinical experience with fetal hemoglobin induction therapy in patients with beta-thalassemia. Blood *121*, 2199-2212; quiz 2372.

Nagababu, E., Fabry, M.E., Nagel, R.L., and Rifkind, J.M. (2008). Heme degradation and oxidative stress in murine models for hemoglobinopathies: thalassemia, sickle cell disease and hemoglobin C disease. Blood cells, molecules & diseases *41*, 60-66.

Nagababu, E., and Rifkind, J.M. (2004). Heme degradation by reactive oxygen species. Antioxidants & redox signaling *6*, 967-978.

Nathan, D.G., and Gunn, R.B. (1966). Thalassemia: the consequences of unbalanced hemoglobin synthesis. The American journal of medicine *41*, 815-830.

Neildez-Nguyen, T.M., Wajcman, H., Marden, M.C., Bensidhoum, M., Moncollin, V., Giarratana, M.C., Kobari, L., Thierry, D., and Douay, L. (2002). Human erythroid cells produced ex vivo at large scale differentiate into red blood cells in vivo. Nat Biotechnol *20*, 467-472.

New, M., Olzscha, H., and La Thangue, N.B. (2012). HDAC inhibitor-based therapies: can we interpret the code? Molecular oncology *6*, 637-656.

Nienhuis, A.W., and Nathan, D.G. (2012). Pathophysiology and Clinical Manifestations of the beta-Thalassemias. Cold Spring Harbor perspectives in medicine *2*, a011726.

NIH, U.S. (2015). ClinicalTrials.gov.

Okam, M.M., Esrick, E.B., Mandell, E., Campigotto, F., Neuberg, D.S., and Ebert, B.L. (2015). Phase 1/2 trial of vorinostat in patients with sickle cell disease who have not benefitted from hydroxyurea. Blood *125*, 3668-3669.

Olivieri, N.F., Saunthararajah, Y., Thayalasuthan, V., Kwiatkowski, J., Ware, R.E., Kuypers, F.A., Kim, H.Y., Trachtenberg, F.L., Vichinsky, E.P., and Thalassemia Clinical Research, N. (2011). A pilot study of subcutaneous decitabine in beta-thalassemia intermedia. Blood *118*, 2708-2711.

Origa, R., Sollaino, M.C., Borgna-Pignatti, C., Piga, A., Feliu Torres, A., Masile, V., and Galanello, R. (2013). alpha-Globin Gene Quadruplication and Heterozygous beta-Thalassemia: A Not So Rare Cause of Thalassemia Intermedia. Acta haematologica *131*, 162-164.

Orkin, S.H., Kazazian, H.H., Jr., Antonarakis, S.E., Ostrer, H., Goff, S.C., and Sexton, J.P. (1982). Abnormal RNA processing due to the exon mutation of beta E-globin gene. Nature *300*, 768-769.

Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003). The beta-globin nuclear compartment in development and erythroid differentiation. Nature genetics *35*, 190-194.

Panzenbock, B., Bartunek, P., Mapara, M.Y., and Zenke, M. (1998). Growth and differentiation of human stem cell factor/erythropoietin-dependent erythroid progenitor cells in vitro. Blood *92*, 3658-3668.

Peck, D., Crawford, E.D., Ross, K.N., Stegmaier, K., Golub, T.R., and Lamb, J. (2006). A method for high-throughput gene expression signature analysis. Genome biology *7*, R61.

Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L*., et al.* (2008). Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol *26*, 808-816.

Perrine, S.P., Ginder, G.D., Faller, D.V., Dover, G.H., Ikuta, T., Witkowska, H.E., Cai, S.P., Vichinsky, E.P., and Olivieri, N.F. (1993). A short-term trial of butyrate to stimulate fetal-globin-gene expression in the beta-globin disorders. The New England journal of medicine *328*, 81-86.

Persons, D.A. (2010). Gene therapy: Targeting beta-thalassaemia. Nature *467*, 277-278.

Persons, D.A., Allay, E.R., Sabatino, D.E., Kelly, P., Bodine, D.M., and Nienhuis, A.W. (2001). Functional requirements for phenotypic correction of murine beta-thalassemia: implications for human gene therapy. Blood *97*, 3275-3282.

Piel, F.B., and Weatherall, D.J. (2014). The alpha-Thalassemias. New Engl J Med *371*, 1908-1916.

Platt, O.S. (2008). Hydroxyurea for the treatment of sickle cell anemia. New Engl J Med *358*, 1362-1369.

Platt, O.S., Orkin, S.H., Dover, G., Beardsley, G.P., Miller, B., and Nathan, D.G. (1984). Hydroxyurea Enhances Fetal Hemoglobin Production in Sickle-Cell-Anemia. Journal of Clinical Investigation *74*, 652-656.

Pope, S.H., Fibach, E., Sun, J., Chin, K., and Rodgers, G.P. (2000). Two-phase liquid culture system models normal human adult erythropoiesis at the molecular level. European journal of haematology *64*, 292-303.

Porter, J., and Viprakasit, V. (2014). Iron overload and chelation. In Guidelines for the management of transfusion dependent thalassaemia, M. Cappellini, A. Cohen, J. Porter, A. Taher, and V. Viprakasit, eds. (Cyprus: Thalassaemia International Federation).

Premawardhena, A., Fisher, C.A., Olivieri, N.F., de Silva, S., Arambepola, M., Perera, W., O'Donnell, A., Peto, T.E., Viprakasit, V., Merson, L.*, et al.* (2005a). Haemoglobin E beta thalassaemia in Sri Lanka. Lancet *366*, 1467-1470.

Premawardhena, A., Fisher, C.A., Olivieri, N.F., de Silva, S., Sloane-Stanley, J., Wood, W.G., and Weatherall, D.J. (2005b). A novel molecular basis for beta thalassemia intermedia poses new questions about its pathophysiology. Blood *106*, 3251-3255.

Rachmilewitz, E.A., and Aker, M. (1998). The role of recombinant human erythropoietin in the treatment of thalassemia. Annals of the New York Academy of Sciences *850*, 129-138.

Rachmilewitz, E.A., and Giardina, P.J. (2011). How I treat thalassemia. Blood *118*, 3479-3488.

Ramos, E., Ruchala, P., Goodnough, J.B., Kautz, L., Preza, G.C., Nemeth, E., and Ganz, T. (2012). Minihepcidins prevent iron overload in a hepcidin-deficient mouse model of severe hemochromatosis. Blood *120*, 3829-3836.

Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y.*, et al.* (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380-1389.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. Nature protocols *8*, 2281-2308.

Ren, X., Yang, Z., Xu, J., Sun, J., Mao, D., Hu, Y., Yang, S.J., Qiao, H.H., Wang, X., Hu, Q.*, et al.* (2014). Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in Drosophila. Cell reports *9*, 1151-1162.

Ribeil, J.A., Arlet, J.B., Dussiot, M., Moura, I.C., Courtois, G., and Hermine, O. (2013). Ineffective erythropoiesis in beta -thalassemia. TheScientificWorldJournal *2013*, 394295.

Rich, A. (1952). Studies on the Hemoglobin of Cooley's Anemia and Cooley's Trait. Proceedings of the National Academy of Sciences of the United States of America *38*, 187-196.

Riddihough, G., and Zahn, L.M. (2010). Epigenetics. What is epigenetics? Introduction. Science *330*, 611.

Rivella, S. (2009). Ineffective erythropoiesis and thalassemias. Current opinion in hematology *16*, 187-194.

Rivella, S. (2015). beta-thalassemias: paradigmatic diseases for scientific discoveries and development of innovative therapies. Haematologica *100*, 418-430.

Ronzoni, L., Bonara, P., Rusconi, D., Frugoni, C., Libani, I., and Cappellini, M.D. (2008). Erythroid differentiation and maturation from peripheral CD34+ cells in liquid culture: cellular and molecular characterization. Blood cells, molecules & diseases *40*, 148-155.

Ronzoni, L., Sonzogni, L., Fossati, G., Modena, D., Trombetta, E., Porretti, L., and Cappellini, M.D. (2014). Modulation of gamma globin genes expression by histone deacetylase inhibitors: an in vitro study. Br J Haematol *165*, 714-721.

Roosjen, M., McColl, B., Kao, B., Gearing, L.J., Blewitt, M.E., and Vadolas, J. (2014). Transcriptional regulators Myb and BCL11A interplay with DNA methyltransferase 1 in developmental silencing of embryonic and fetal beta-like globin genes. FASEB journal : official publication of the Federation of American Societies for Experimental Biology *28*, 1610-1620.

Rose, N.R., Ng, S.S., Mecinovic, J., Lienard, B.M., Bello, S.H., Sun, Z., McDonough, M.A., Oppermann, U., and Schofield, C.J. (2008). Inhibitor scaffolds for 2-oxoglutarate-dependent histone lysine demethylases. J Med Chem *51*, 7053-7056.

Rotili, D., Altun, M., Kawamura, A., Wolf, A., Fischer, R., Leung, I.K., Mackeen, M.M., Tian, Y.M., Ratcliffe, P.J., Mai, A.*, et al.* (2011). A photoreactive small-molecule probe for 2-oxoglutarate oxygenases. Chem Biol *18*, 642-654.

Rund, D., and Rachmilewitz, E. (2005). Beta-thalassemia. The New England journal of medicine *353*, 1135-1146.

SangamoBiosciences (2015). Hemoglobinopathies (Beta-thalassemia & Sickle Cell Disease) (Sangamo Biosciences), pp. http://www.sangamo.com/pipeline/hemoglobinopathies.html.

Sankaran, V.G. (2011). Targeted therapeutic strategies for fetal hemoglobin induction. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2011*, 459-465.

Sankaran, V.G., Menne, T.F., Xu, J., Akie, T.E., Lettre, G., Van Handel, B., Mikkola, H.K., Hirschhorn, J.N., Cantor, A.B., and Orkin, S.H. (2008). Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. Science *322*, 1839-1842.

Sankaran, V.G., and Orkin, S.H. (2013). The switch from fetal to adult hemoglobin. Cold Spring Harbor perspectives in medicine *3*, a011643.

Sankaran, V.G., and Weiss, M.J. (2015). Anemia: progress in molecular mechanisms and therapies. Nature medicine *21*, 221-230.

Sauntharajah, Y., Hillery, C.A., Lavelle, D., Molokie, R., Dorn, L., Bressler, L., Gavazova, S., Chen, Y.H., Hoffman, R., and DeSimone, J. (2003). Effects of 5-aza-2'-deoxycytidine on fetal hemoglobin levels, red cell adhesion, and hematopoietic differentiation in patients with sickle cell disease. Blood *102*, 3865-3870.

Schiller, R., Scozzafava, G., Tumber, A., Wickens, J.R., Bush, J.T., Rai, G., Lejeune, C., Choi, H., Yeh, T.L., Chan, M.C.*, et al.* (2014). A Cell-Permeable Ester Derivative of the JmjC Histone Demethylase Inhibitor IOX1. Chemmedchem *9*, 566-571.

Schmidt, P.J., and Fleming, M.D. (2014). Modulation of hepcidin as therapy for primary and secondary iron overload disorders: preclinical models and approaches. Hematology/oncology clinics of North America *28*, 387-401.

Schrier, S.L. (1997). Pathophysiology of the thalassemias. The Albion Walter Hewlett Award presentation. The Western journal of medicine *167*, 82-89.

Schrier, S.L. (2002). Pathophysiology of thalassemia. Current opinion in hematology *9*, 123-126.

SGC (2015). IOX1: A Cell-Penetrating Broad-Spectrum 2-Oxoglutarate Oxygenase Inhibitor (The Structural Genome Consortium), pp. http://www.thesgc.org/chemical-probes/IOX1.

Shaeffer, J.R. (1988). ATP-dependent proteolysis of hemoglobin alpha chains in beta-thalassemic hemolysates is ubiquitin-dependent. The Journal of biological chemistry *263*, 13663-13669.

Sharma, S.K., Choudhary, D., Gupta, N., Dhamija, M., Khandelwal, V., Kharya, G., Handoo, A., Setia, R., and Arora, A. (2014). Cost of hematopoietic stem cell transplantation in India. Mediterranean journal of hematology and infectious diseases 6, e2014046.

Sharma, V., and Saxena, R. (2009). Effect of alpha-gene numbers on phenotype of HbE/beta thalassemia patients. Annals of hematology *88*, 1035-1036.

Sharpe, J.A., Chan-Thomas, P.S., Lida, J., Ayyub, H., Wood, W.G., and Higgs, D.R. (1992). Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. The EMBO journal *11*, 4565-4572.

Sherman, M.L., Borgstein, N.G., Mook, L., Wilson, D., Yang, Y., Chen, N., Kumar, R., Kim, K., and Laadem, A. (2013). Multiple-dose, safety, pharmacokinetic, and pharmacodynamic study of sotatercept (ActRIIA-IgG1), a novel erythropoietic agent, in healthy postmenopausal women. Journal of clinical pharmacology *53*, 1121-1130.

Shi, L., Cui, S., Engel, J.D., and Tanabe, O. (2013). Lysine-specific demethylase 1 is a therapeutic target for fetal hemoglobin induction. Nature medicine *19*, 291-294.

Shi, Y. (2007). Histone lysine demethylases: emerging roles in development, physiology and disease. Nature reviews Genetics *8*, 829-833.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A., and Shi, Y. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shinar, E., and Rachmilewitz, E.A. (1993). Haemoglobinopathies and red cell membrane function. Bailliere's clinical haematology *6*, 357-369.

Shyamala, V. (2014). Transfusion transmitted infections in thalassaemics: need for reappraisal of blood screening strategy in India. Transfusion medicine *24*, 79-88.

Simon, J.A., and Kingston, R.E. (2013). Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. Molecular cell *49*, 808-824.

Sims, R.J., 3rd, Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. Trends in genetics : TIG *19*, 629-639.

Smetanina, N.S., Molchanova, T.P., and Huisman, T.H. (1997). Analysis of mRNA from red cells of patients with thalassemia and hemoglobin variants. Hemoglobin *21*, 437-467.

Socolovsky, M., Nam, H., Fleming, M.D., Haase, V.H., Brugnara, C., and Lodish, H.F. (2001). Ineffective erythropoiesis in Stat5a(-/-)5b(-/-) mice due to decreased survival of early erythroblasts. Blood *98*, 3261-3273.

Sollaino, M.C., Paglietti, M.E., Perseu, L., Giagu, N., Loi, D., and Galanello, R. (2009). Association of alpha globin gene quadruplication and heterozygous beta thalassemia in patients with thalassemia intermedia. Haematologica *94*, 1445-1448.

Souma, T., Suzuki, N., and Yamamoto, M. (2015). Renal erythropoietin-producing cells in health and disease. Frontiers in physiology *6*, 167.

Sripichai, O., Munkongdee, T., Kumkhaek, C., Svasti, S., Winichagoon, P., and Fucharoen, S. (2008). Coinheritance of the different copy numbers of alpha-globin gene modifies severity of beta-thalassemia/Hb E disease. Annals of hematology *87*, 375-379.

Steinberg, M.H., and Nagel, R.L. (2009). Hemoglobins of the Embryo, Fetus and Adult. In Disorders of hemoglobin, M.H. Steinberg, B.G. Forget, D.R. Higgs, and D.J. Weatherall, eds. (New York: Cambridge University Press), pp. 119-136.

Steinberg, M.H., and Sebastiani, P. (2012). Genetic modifiers of sickle cell disease. American journal of hematology *87*, 795-803.

Suragani, R.N., Cawley, S.M., Li, R., Wallner, S., Alexander, M.J., Mulivor, A.W., Gardenghi, S., Rivella, S., Grinberg, A.V., Pearsall, R.S*., et al.* (2014). Modified activin receptor IIB ligand trap mitigates ineffective erythropoiesis and disease complications in murine beta-thalassemia. Blood *123*, 3864-3872.

Suresh, S., Fisher, C., Ayyub, H., Premawardhena, A., Allen, A., Perera, A., Bandara, D., Olivieri, N., and Weatherall, D. (2013). Alpha thalassaemia and extended alpha globin genes in Sri Lanka. Blood cells, molecules & diseases *50*, 93-98.

Suzuki, M., Yamamoto, M., and Engel, J.D. (2014). Fetal globin gene repressors as drug targets for molecular therapies to treat the beta-globinopathies. Molecular and cellular biology *34*, 3560-3569.

Tanno, T., Bhanu, N.V., Oneal, P.A., Goh, S.H., Staker, P., Lee, Y.T., Moroney, J.W., Reed, C.H., Luban, N.L., Wang, R.H*., et al.* (2007). High levels of GDF15 in thalassemia suppress expression of the iron regulatory protein hepcidin. Nature medicine *13*, 1096-1101.

Tashiro, E., and Imoto, M. (2012). Target identification of bioactive compounds. Bioorganic & medicinal chemistry *20*, 1910-1921.

Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G*., et al.* (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. The New England journal of medicine *370*, 901-910.

Thein, S.L. (2008). Genetic modifiers of the beta-haemoglobinopathies. Brit J Haematol *141*, 357-366.

Thein, S.L. (2012). The emerging role of fetal hemoglobin induction in non-transfusion-dependent thalassemia. Blood reviews *26 Suppl 1*, S35-39.

Thein, S.L. (2013). The molecular basis of beta-thalassemia. Cold Spring Harbor perspectives in medicine *3*, a011700.

Thein, S.L., Hesketh, C., Wallace, R.B., and Weatherall, D.J. (1988). The molecular basis of thalassaemia major and thalassaemia intermedia in Asian Indians: application to prenatal diagnosis. Br J Haematol *70*, 225-231.

Thinnes, C.C., England, K.S., Kawamura, A., Chowdhury, R., Schofield, C.J., and Hopkinson, R.J. (2014). Targeting histone lysine demethylases - progress, challenges, and the future. Biochimica et biophysica acta *1839*, 1416-1432.

Thompson, A.A., Rasko, J.E., Hongeng, S., Kwiatkowski, J.L., Schiller, G., von Kalle, C., Cavazzana, M., Leboulch, P., Petrusich, A., Soni, S*., et al.* (2014). Initial Results from the Northstar Study (HGB-204): A Phase 1/2 Study of Gene Therapy for beta-Thalassemia Major Via Transplantation of Autologous Hematopoietic Stem Cells Transduced Ex Vivo with a Lentiviral beta(A-T87Q) Globin Vector (LentiGlobin BB305 Drug Product). Blood *124*.

Trojer, P., and Reinberg, D. (2006). Histone lysine demethylases and their impact on epigenetics. Cell *125*, 213-217.

Trompeter, S., and Cohen, A. (2014). Blood transfusion. In Guidelines for the management of transfusion dependent thalassaemia, C. MD, C. A, P. J, T. A, and V. V, eds. (Cyprus: Thalassaemia International Federation), pp. 28-42.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P*., et al.* (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol *33*, 187-197.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811-816.

Uda, M., Galanello, R., Sanna, S., Lettre, G., Sankaran, V.G., Chen, W., Usala, G., Busonero, F., Maschio, A., Albai, G*., et al.* (2008). Genome-wide association study shows BCL11A associated with persistent fetal hemoglobin and amelioration of the phenotype of beta-thalassemia. Proceedings of the National Academy of Sciences of the United States of America *105*, 1620-1625.

Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., and Gregory, P.D. (2010). Genome editing with engineered zinc finger nucleases. Nature reviews Genetics *11*, 636-646.

van den Akker, E., Satchwell, T.J., Pellegrin, S., Daniels, G., and Toye, A.M. (2010). The majority of the in vitro erythroid expansion potential resides in CD34(-) cells, outweighing the contribution of CD34(+) cells and significantly increasing the erythroblast yield from peripheral blood samples. Haematol-Hematol J *95*, 1594-1598.

van den Heuvel, A., Stadhouders, R., Andrieu-Soler, C., Grosveld, F., and Soler, E. (2015). Long-range gene regulation and novel therapeutic applications. Blood *125*, 1521-1525.

Varawalla, N.Y., Old, J.M., and Weatherall, D.J. (1991). Rare beta-thalassaemia mutations in Asian indians. Br J Haematol *79*, 640-644.

Vernimmen, D. (2014). Uncovering Enhancer Functions Using the alpha-Globin Locus. PLoS genetics *10*, e1004668.

Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G., and Higgs, D.R. (2007). Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. The EMBO journal *26*, 2041-2051.

Vernimmen, D., Lynch, M.D., De Gobbi, M., Garrick, D., Sharpe, J.A., Sloane-Stanley, J.A., Smith, A.J., and Higgs, D.R. (2011). Polycomb eviction as a new distant enhancer function. Genes & development *25*, 1583-1588.

Vernimmen, D., Marques-Kranc, F., Sharpe, J.A., Sloane-Stanley, J.A., Wood, W.G., Wallace, H.A., Smith, A.J., and Higgs, D.R. (2009). Chromosome looping at the human alpha-globin locus is mediated via the major upstream regulatory element (HS -40). Blood *114*, 4253-4260.

Vichinsky, E. (2012). Emerging 'A' therapies in hemoglobinopathies: agonists, antagonists, antioxidants, and arginine. Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program *2012*, 271-275.

Viprakasit, V., Ekwattanakit, S., Riolueang, S., Chalaow, N., Fisher, C., Lower, K., Kanno, H., Tachavanich, K., Bejrachandra, S., Saipin, J.*, et al.* (2014). Mutations in Kruppel-like factor 1 cause transfusion-dependent hemolytic anemia and persistence of embryonic globin gene expression. Blood *123*, 1586-1595.

Viprakasit, V., Tyan, P., Rodmai, S., and Taher, A.T. (2014). Identification and key management of non-transfusion-dependent thalassaemia patients: not a rare but potentially under-recognised condition. Orphanet journal of rare diseases 9, 131.

Voon, H.P., and Vadolas, J. (2008). Controlling alpha-globin: a review of alpha-globin expression and its impact on beta-thalassemia. Haematologica *93*, 1868-1876.

Voon, H.P., Wardan, H., and Vadolas, J. (2007). Co-inheritance of alpha- and beta-thalassaemia in mice ameliorates thalassaemic phenotype. Blood cells, molecules & diseases *39*, 184-188.

Voon, H.P., Wardan, H., and Vadolas, J. (2008). siRNA-mediated reduction of alpha-globin results in phenotypic improvements in beta-thalassemic cells. Haematologica *93*, 1238-1242.

Wainscoat, J.S., Old, J.M., Weatherall, D.J., and Orkin, S.H. (1983). The molecular basis for the clinical diversity of beta thalassaemia in Cypriots. Lancet *1*, 1235-1237.

Wajcman, H., and Riou, J. (2009). Globin chain analysis: An important tool in phenotype study of hemoglobin disorders. Clinical biochemistry *42*, 1802-1806.

Wallace, H.A.C., Marques-Kranc, F., Richardson, M., Luna-Crespo, F., Sharpe, J.A., Hughes, J., Wood, W.G., Higgs, D.R., and Smith, A.J.H. (2007). Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. Cell *128*, 197-209.

Wang, L., Chang, J.J., Varghese, D., Dellinger, M., Kumar, S., Best, A.M., Ruiz, J., Bruick, R., Pena-Llopis, S., Xu, J.J.*, et al.* (2013). A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth (vol 4, pg 2035, 2013). Nature communications *4*.

Weatherall, D.J. (2008). Genetic variation and susceptibility to infection: the red cell and malaria. Br J Haematol *141*, 276-286.

Weatherall, D.J. (2010). The inherited diseases of hemoglobin are an emerging global health burden. Blood *115*, 4331-4336.

Weatherall, D.J. (2011). The challenge of haemoglobinopathies in resource-poor countries. Br J Haematol *154*, 736-744.

Weatherall, D.J. (2012). The definition and epidemiology of non-transfusion-dependent thalassemia. Blood reviews *26 Suppl 1*, S3-6.

Weatherall, D.J., and Clegg, J.B. (1996). Thalassemia--a global public health problem. Nature medicine *2*, 847-849.

Weatherall, D.J., and Clegg, J.B. (2001). The thalassaemia syndromes, 4 edn (Oxford: Blackwell Science).

Weatherall, D.J., Clegg, J.B., and Naughton, M.A. (1965). Globin synthesis in thalassaemia: an in vitro study. Nature *208*, 1061-1065.

West, A.C., and Johnstone, R.W. (2014). New and emerging HDAC inhibitors for cancer treatment. The Journal of clinical investigation *124*, 30-39.

Winichagoon, P., Fucharoen, S., Chen, P., and Wasi, P. (2000). Genetic factors affecting clinical severity in beta-thalassemia syndromes. Journal of pediatric hematology/oncology *22*, 573-580.

Xie, H., Xu, J., Hsu, J.H., Nguyen, M., Fujiwara, Y., Peng, C., and Orkin, S.H. (2014). Polycomb repressive complex 2 regulates normal hematopoietic stem cell function in a developmental-stage-specific manner. Cell stem cell *14*, 68-80.

Xie, S.Y., Li, W., Ren, Z.R., Huang, S.Z., Zeng, F., and Zeng, Y.T. (2011). Correction of beta654-thalassaemia mice using direct intravenous injection of siRNA and antisense RNA vectors. International journal of hematology *93*, 301-310.

Xie, S.Y., Ren, Z.R., Zhang, J.Z., Guo, X.B., Wang, Q.X., Wang, S., Lin, D., Gong, X.L., Li, W., Huang, S.Z.*, et al.* (2007). Restoration of the balanced alpha/beta-globin gene expression in beta654-thalassemia mice using combined RNAi and antisense RNA approach. Human molecular genetics *16*, 2616-2625.

Xu, H., Xiao, T., Chen, C.H., Li, W., Meyer, C., Wu, Q., Wu, D., Cong, L., Zhang, F., Liu, J.S.*, et al.* (2015). Sequence determinants of improved CRISPR sgRNA design. Genome research.

Yin, H., Kanasty, R.L., Eltoukhy, A.A., Vegas, A.J., Dorkin, J.R., and Anderson, D.G. (2014). Non-viral vectors for gene-based therapy. Nature reviews Genetics *15*, 541-555.

Yu, X., Kong, Y., Dore, L.C., Abdulmalik, O., Katein, A.M., Zhou, S., Choi, J.K., Gell, D., Mackay, J.P., Gow, A.J.*, et al.* (2007). An erythroid chaperone that facilitates folding of alpha-globin subunits for hemoglobin synthesis. The Journal of clinical investigation *117*, 1856-1865.

Yuan, J., Kannan, R., Shinar, E., Rachmilewitz, E.A., and Low, P.S. (1992). Isolation, characterization, and immunoprecipitation studies of immune complexes from membranes of beta-thalassemic erythrocytes. Blood *79*, 3007-3013.

Yusof, I., Shah, F., Hashimoto, T., Segall, M.D., and Greene, N. (2014). Finding the rules for successful drug optimisation. Drug Discov Today *19*, 680-687.

Appendix

Appendix 1 – Supplementary Tables

Supplementary table 1 – List of genes significantly up-regulated in IOX1 treated cells (in alphabetical order)

Gene Symbol	*Gene Name*	*Fold change*
AAGAB	Alpha- and gamma-adaptin binding protein	1.28
ABCE1	ATP-binding cassette, sub-family E (OABP), member 1	1.28
ACTL6A	Actin-like 6A	1.28
AHSA1	Activator of heat shock 90kDa protein ATPase homolog 1	1.57
ANXA2	Annexin A2	1.49
BCCIP	BRCA2 and CDKN1A interacting protein	1.44
BRI3BP	BRI3 binding protein	1.50
C17ORF79	Chromosome 17 open reading frame 79	1.47
C18ORF55	Chromosome 18 open reading frame 55	1.39
C19ORF2	Chromosome 19 open reading frame 2	1.28
CACYBP	Calcyclin binding protein	1.77
CDC123	Cell division cycle 123 homolog (S. cerevisiae)	1.25
CEP78	Centrosomal protein 78kDa	1.35
CLP1	Cleavage and polyadenylation factor I subunit	1.20
CRKRS	Cdc2-related kinase, arginine/serine-rich	1.18
CSTF3	Cleavage stimulation factor, 3' pre-RNA, subunit 3	1.24
CYB5B	Cytochrome b5 type B	1.30
CYP1B1	Cytochrome P450, family 1, subfamily B, polypeptide 1	1.92
DCUN1D5	Defective in cullinneddylation 1, domain containing 5	1.34
DHX15	DEAH (Asp-Glu-Ala-His) box polypeptide 15	1.39
DYRK2	Dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 2	1.37
EBNA1BP2	EBNA1 binding protein 2	1.49
EI24	Etoposide induced 2.4 mRNA	1.37
EIF4G2	Eukaryotic translation initiation factor 4 gamma 2	1.43
EMG1	EMG1 nucleolar protein homolog (S. cerevisiae)	1.39
EMP3	Epithelial membrane protein 3	1.48
ENOPH1	Enolase-phosphatase 1	1.18
FABP5L2	Fatty acid binding protein 5-like 2	1.49
FAM92A1	Family with sequence similarity 92, member A1	1.26
FJX1	Four jointed box 1 (Drosophila)	1.34
FREQ	Frequenin homolog (Drosophila)	1.30
FTL	Ferritin, light polypeptide	1.17

Gene Symbol	Gene Name	Fold change
G3BP1	GTPase activating protein (SH3 domain) binding protein 1	1.36
G3BP2	GTPase activating protein (SH3 domain) binding protein 2	1.27
GART	Phosphoribosylglycinamide formyltransferase	1.36
GNL3	Guanine nucleotide binding protein-like 3	1.54
GPATCH4	G patch domain containing 4	1.27
GPN3	GPN-loop GTPase 3 (GPN3)	1.22
GTPBP4	GTP binding protein 4	1.51
HES5	Hairy and enhancer of split 5 (Drosophila)	1.45
HMGN2	High-mobility group nucleosomal binding domain 2	1.26
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	1.73
HNRNPD	Heterogeneous nuclear ribonucleoprotein D	1.40
HNRNPM	Heterogeneous nuclear ribonucleoprotein M	1.48
HSP90AA1	Heat shock protein 90kDa alpha (cytosolic), class A member 1	1.47
HSPC111	Hypothetical protein HSPC111	1.62
LOC100130003	Misc_RNA	1.56
LOC389816	Cytokeratin associated protein	1.29
LOC389873	Misc_RNA	1.24
LOC642031	Hypothetical protein LOC642031	1.34
LOC647150	Misc_RNA	1.43
LOC654244	Similar to mitochondrial carrier protein MGC4399	1.29
LRRC26	Leucine rich repeat containing 26	1.25
LSM12	LSM12 homolog (S. cerevisiae)	1.27
MED27	Mediator complex subunit 27	1.30
MFAP1	Microfibrillar-associated protein 1	1.19
MIF	Macrophage migration inhibitory factor	1.38
MORF4L2	Mortality factor 4 like 2	1.35
MPP6	Membrane protein, palmitoylated 6	1.33
NDUFS5	NADH dehydrogenase (ubiquinone) Fe-S protein 5	1.31
NHP2L1	NHP2 non-histone chromosome protein 2-like 1 (S. cerevisiae)	1.32
NIP7	Nuclear import 7 homolog (S. cerevisiae)	1.38
NME1	Non-metastatic cells 1, protein	1.47
NOLC1	Nucleolar and coiled-body phosphoprotein 1	1.39
NOP16	NOP16 nucleolar protein homolog (yeast)	1.61
NSUN2	NOP2/Sun domain family, member 2	1.35
NXT1	NTF2-like export factor 1	1.44
PAK1IP1	PAK1 interacting protein 1	1.42
PDCL3	Phosducin-like 3	1.43
PGRMC1	Progesterone receptor membrane component 1	1.43
PIAS2	Protein inhibitor of activated STAT, 2	1.24

Gene Symbol	Gene Name	Fold change
PIGW	Phosphatidylinositol glycan anchor biosynthesis, class W	1.37
POLR2A	Polymerase (RNA) II (DNA directed) polypeptide A	1.40
POLR2F	Polymerase (RNA) II (DNA directed) polypeptide F	1.42
POLR3K	Polymerase (RNA) III (DNA directed) polypeptide K	1.39
PPPDE1	PPPDE peptidase domain containing 1	1.30
PRDX1	Peroxiredoxin 1	1.44
PRPF38A	Pre-mRNA processing factor 38 (yeast) domain containing A	1.37
PSMA3	Proteasome (prosome, macropain) subunit, alpha type, 3	1.46
PSMA4	Proteasome (prosome, macropain) subunit, alpha type, 4	1.31
PTPLAD1	Protein tyrosine phosphatase-like A domain containing 1	1.32
PTRH2	Peptidyl-tRNA hydrolase 2	1.37
RABEPK	Rab9 effector protein with kelch motifs	1.37
RANBP1	RAN binding protein 1	1.52
RARS	Arginyl-tRNA synthetase	1.48
RBBP4	Retinoblastoma binding protein 4	1.20
RBM12	RNA binding motif protein 12	1.23
RCC1	Regulator of chromosome condensation 1	1.25
RPL36A	Rribosomal protein L36a	1.54
RPL6	Ribosomal protein L6	1.89
RRP15	Ribosomal RNA processing 15 homolog (S. cerevisiae)	1.36
SETMAR	SET domain and mariner transposase fusion gene	1.39
SNRPA1	Small nuclear ribonucleoprotein polypeptide A	1.26
SNRPF	Small nuclear ribonucleoprotein polypeptide F	1.33
TBRG4	Transforming growth factor beta regulator 4	1.34
TFB2M	Transcription factor B2, mitochondrial	1.32
THOC4	THO complex 4	1.41
TIMM23	Translocase of inner mitochondrial membrane 23 homolog (yeast)	1.39
TIPIN	TIMELESS interacting protein	1.35
TSPAN3	Tetraspanin 3	1.27
TTC4	Tetratricopeptide repeat domain 4	1.27
TXNRD1	Thioredoxin reductase 1	1.42
UCK2	Uridine-cytidine kinase 2	1.38
UCRC	Ubiquinol-cytochrome c reductase complex	1.22
USP22	Ubiquitin specific peptidase 22	1.17
YWHAZ	Tyrosine 3-monooxygenase	1.33
ZBED1	Zinc finger, BED-type containing 1	1.30
ZNF275	Zinc finger protein 275	1.17
ZNF326	Zinc finger protein 326	1.27
ZNF410	Zinc finger protein 410	1.18

Supplementary table 2 – List of genes significantly down-regulated in IOX1 treated cells (in alphabetical order)

Gene Symbol	*Gene Name*	*Fold change*
AADACL1	Arylacetamide deacetylase-like 1	0.63
AARS	Alanyl-tRNA synthetase	0.64
ACSBG1	Acyl-CoA synthetase bubblegum family member 1	0.55
AMT	Aminomethyltransferase	0.74
AMY1A	Amylase, alpha 1A (salivary)	0.78
AMY1C	Amylase, alpha 1C (salivary)	0.76
ANKRA2	Ankyrin repeat, family A	0.67
ARHGEF2	Guanine nucleotide exchange factor	0.78
C10ORF33	Chromosome 10 open reading frame 33	0.69
C14ORF93	Chromosome 14 open reading frame 93	0.73
C17ORF90	Chromosome 17 open reading frame 90	0.75
C1ORF59	chromosome 1 open reading frame 59	0.67
C4ORF14	Chromosome 4 open reading frame 14	0.72
C6ORF192	Chromosome 6 open reading frame 192	0.76
CAPRIN2	Caprin family member 2	0.73
CARS	Cysteinyl-tRNA synthetase	0.62
CCDC26	Coiled-coil domain containing 26	0.80
CD37	CD37 antigen	0.72
CD53	CD53 molecule	0.81
CIDEB	Cell death-inducing DFFA-like effector b	0.87
CITED2	Cbp/p300-interacting transactivator	0.71
CTH	Cystathionase	0.77
CTSH	Cathepsin H	0.65
CXORF12	Chromosome X open reading frame 12	0.81
DECR1	2,4-dienoyl CoA reductase 1	0.71
DENND2D	DENN/MADD domain containing 2D	0.72
DLK1	Delta-like 1 homolog (Drosophila)	0.66
DNAJB9	DnaJ (Hsp40) homolog, subfamily B, member 9	0.83
DPEP2	Dipeptidase 2	0.81
FAM113B	Family with sequence similarity 113, member B	0.56
FAM178B	Family with sequence similarity 178, member B	0.26
FAM83A	Family with sequence similarity 83, member A	0.55
FBXO11	F-box protein 11	0.77
FCGR2A	Fc fragment of IgG	0.68
FHL2	Four and a half LIM domains 2	0.54
GALC	Galactosylceramidase	0.76
GALT	Galactose-1-phosphate uridylyltransferase	0.72

Gene Symbol	Gene Name	Fold change
GBGT1	Globoside alpha-1,3-N-acetylgalactosaminyltransferase 1	0.66
IFNAR2	Interferon (alpha, beta and omega) receptor 2	0.79
IGSF3	Imunoglobulin superfamily, member 3	0.78
IRF9	Interferon regulatory factor 9	0.68
ITGA5	Integrin, alpha 5	0.75
ITGB1	Integrin, beta 1	0.68
KLF6	Kruppel-like factor 6	0.68
LMO2	LIM domain only 2	0.62
LOC130773	Similar to 60S ribosomal protein L23a	0.79
LOC284023	Hypothetical protein LOC284023	0.81
LOC339970	Misc_RNA	0.79
LOC387841	Similar to ribosomal protein L13a	0.70
LOC390940	Similar to R28379_1	0.72
LOC400464	Similar to FLJ43276 protein	0.68
LOC440348	Similar to nuclear pore complex interacting protein	0.73
LOC440353	Nuclear pore complex interacting protein pseudogene	0.72
LOC441013	Misc_RNA	0.75
LOC642299	Hypothetical protein	0.71
LOC653907	Similar to complement receptor related protein isoform 1	0.58
LOC729642	Hypothetical LOC729642	0.72
LOC91561	Similar to ribosomal protein S2	0.76
LRG1	Leucine-rich alpha-2-glycoprotein 1	0.80
MAP7	Microtubule-associated protein 7	0.65
MGC4677	Hypothetical protein MGC4677	0.77
MST1	Macrophage stimulating 1	0.57
MT2A	Metallothionein 2A	0.69
NPIP	Nuclear pore complex interacting protein	0.75
PARP3	Poly (ADP-ribose) polymerase family, member 3	0.76
PCK2	Phosphoenolpyruvate carboxykinase 2	0.53
PEAR1	Platelet endothelial aggregation receptor 1	0.81
PIP5K1B	Phosphatidylinositol-4-phosphate 5-kinase, type I, beta	0.80
PLEKHM2	Pleckstrin homology domain containing, family M	0.80
PLSCR4	Phospholipid scramblase 4	0.76
PPM1M	Protein phosphatase 1M	0.67
PRG2	Proteoglycan 2, bone marrow	0.40
PTGS1	Prostaglandin-endoperoxide synthase 1	0.67
PTPN6	Protein tyrosine phosphatase, non-receptor type 6	0.61
RENBP	Rrenin binding protein	0.81
RHCE	Rh blood group, CcEe antigens	0.43

Gene Symbol	*Gene Name*	*Fold change*
RINL	Ras and Rab interactor-like	0.73
SKAP1	Src kinase associated phosphoprotein 1	0.59
SLC11A1	Solute carrier family 11 , member 1	0.56
SLC16A9	Solute carrier family 16, member 9	0.38
SLC1A5	Solute carrier family 1, member 5	0.68
SLC2A10	Solute carrier family 2, member 10	0.73
SLC7A1	Solute carrier family 7, member 1	0.69
TUBAL3	Tubulin, alpha-like 3	0.65
UCA1	Urothelial cancer associated 1	0.43
VEGFA	Vascular endothelial growth factor A	0.68
ZNF419	Zinc finger protein 419	0.76

Supplementary table 3 – List of genes significantly up-regulated in vorinostat treated cells (in alphabetical order)

Gene Symbol	*Gene Name*	*Fold change*
ABCB1	ATP-binding cassette, sub-family B	1.39
ABHD6	Abhydrolase domain containing 6	2.07
ALOX5AP	Arachidonate 5-lipoxygenase-activating protein	2.07
ANXA1	Annexin A1	3.31
AP1S2	Adaptor-related protein complex 1, sigma 2 subunit	2.00
APOE	Apolipoprotein E	2.09
ARRB1	Beta 1	1.68
ATP2A3	ATPase, Ca++ transporting, ubiquitous	1.89
BEND4	BEN domain containing 4	1.29
BEX1	Brain expressed, X-linked 1	2.85
C11ORF60	Chromosome 11 open reading frame 60	1.36
C12ORF45	Chromosome 12 open reading frame 45	1.50
C14ORF112	Chromosome 14 open reading frame 112	1.40
C15ORF39	Chromosome 15 open reading frame 39	1.99
C16ORF45	Chromosome 16 open reading frame 45	1.31
C1ORF183	Chromosome 1 open reading frame 183	1.25
C20ORF3	Chromosome 20 open reading frame 3	1.56
C21ORF63	Chromosome 21 open reading frame 63	1.49
C2ORF18	Chromosome 2 open reading frame 18	1.53
CBX6	Chromobox homolog 6	1.87
CCND2	Cyclin D2	2.15
CD52	CD52 molecule	4.17
CD53	CD53 molecule	1.67
CD74	CD74 molecule	7.26
CENPBD1	CENPB DNA-binding domains containing 1	1.32
CHD7	Chromodomain helicase DNA binding protein 7	1.68
CHST7	Carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 7	1.41
CLDN10	Claudin 10	1.94
CLDN11	Claudin 11	3.09
CORO1A	Coronin, actin binding protein, 1A	1.70
CSF2RB	Colony stimulating factor 2 receptor, beta	1.45
CUGBP2	CUG triplet repeat, RNA binding protein 2	1.99
CYFIP2	Cytoplasmic FMR1 interacting protein 2	1.70
DBI	Diazepam binding inhibitor	1.54
DEPDC7	DEP domain containing 7	1.25
EIF3E	Eukaryotic translation initiation factor 3, subunit E	1.21
EPB41L2	Erythrocyte membrane protein band 4.1-like 2	1.69

Gene Symbol	Gene Name	Fold change
EXTL2	Exostoses (multiple)-like 2	1.53
FADS1	Fatty acid desaturase 1	1.42
FAM190B	Family with sequence similarity 190, member B	1.23
FAM27A	Family with sequence similarity 27, member A	1.40
FCGR2A	Fc fragment of IgG, low affinity IIa, receptor	1.90
FKSG30	Actin-like protein	1.83
GALM	Galactose mutarotase (aldose 1-epimerase)	1.84
HLA-A	Major histocompatibility complex, class I, A	2.31
HLA-B	Major histocompatibility complex, class I, B	2.64
HLA-DMA	Major histocompatibility complex, class II, DM alpha	3.27
HLA-DPA1	Major histocompatibility complex, class II, DP alpha 1	4.54
HLA-DRA	Major histocompatibility complex, class II, DR alpha	7.40
HLA-DRB1	Major histocompatibility complex, class II, DR beta 1	2.21
HLA-DRB6	Major histocompatibility complex, class II, DR beta 6 (pseudogene)	1.98
HS.133181	Soares_parathyroid_tumor_NbHPA	1.62
HS.193406	cDNA FLJ34755 fis	1.66
ID2	Inhibitor of DNA binding 2	2.10
IGSF3	Immunoglobulin superfamily, member 3	1.72
IL1B	Interleukin 1, beta	2.30
ITM2C	Integral membrane protein 2C	1.81
KIAA1522	KIAA1522	1.56
KLHL5	Kelch-like 5 (Drosophila)	1.90
LCP1	Lymphocyte cytosolic protein 1 (L-plastin)	2.17
LEF1	Lymphoid enhancer-binding factor 1	1.83
LIMA1	LIM domain and actin binding 1	1.33
LITAF	Lipopolysaccharide-induced TNF factor	3.42
LOC100132948	Similar to Protein FAM27A/B/C, transcript variant 1	1.54
LOC401076	Misc_RNA	1.31
LOC606724	Actin binding protein, 1A pseudogene	1.70
LOC643319	Similar to Transgelin-2	1.51
LOC644760	Hypothetical protein LOC644760	2.38
LOC645553	Hypothetical LOC645553	2.56
LOC730415	Hypothetical LOC730415, transcript variant 2	3.94
LRRC20	Leucine rich repeat containing 20	1.56
LY6E	Lymphocyte antigen 6 complex, locus E	2.41
LYN	Yamaguchi sarcoma viral related oncogene homolog	1.70
MAP3K7	Mitogen-activated protein kinase kinase kinase 7	1.33
MEF2C	Myocyte enhancer factor 2C	1.50
MFSD1	Major facilitator superfamily domain containing 1	1.45

Gene Symbol	Gene Name	Fold change
MGC71993	Similar to DNA segment, Chr 11,	1.54
MLLT11	Myeloid/lymphoid or mixed-lineage leukemia	2.37
MSN	Moesin	2.31
MTM1	Myotubularin 1	1.28
MYC	Myelocytomatosis viral oncogene homolog	1.62
NME7	Non-metastatic cells 7, protein	1.47
NUDT11	Nudix (nucleoside diphosphate linked moiety X)-type motif 11	1.67
P76	Mannose-6-phosphate protein p76	1.35
PA2G4	Proliferation-associated 2G4	1.49
PAQR8	Progestin and adipoQ receptor family member VIII	1.40
PELI2	Pellino homolog 2 (Drosophila)	1.82
PHC1	Polyhomeotic homolog 1 (Drosophila)	1.31
PNKD	Paroxysmal nonkinesiogenic dyskinesia	1.73
POU4F1	POU class 4 homeobox 1	1.69
PRKCB	Protein kinase C, beta	3.02
PRKCB1	Protein kinase C, beta 1	3.50
PRNP	Prion protein	1.45
PSMB9	Proteasome (prosome, macropain) subunit, beta type, 9	1.51
PTDSS1	Phosphatidylserine synthase 1	1.36
RAB11FIP5	RAB11 family interacting protein 5 (class I)	1.53
RAB31	RAB31, member RAS oncogene family	2.66
RAB37	RAB37, member RAS oncogene family	1.72
RAB38	RAB38, member RAS oncogene family	2.44
RNASEK	Ribonuclease, RNase K	1.82
RRAGD	Ras-related GTP binding D	1.46
SAMSN1	SAM domain, SH3 domain and nuclear localization signals 1	1.98
SDPR	Serum deprivation response (phosphatidylserine binding protein)	2.32
SEPT6	Septin 6	1.46
SKAP1	Src kinase associated phosphoprotein 1	2.18
SLC16A10	Solute carrier family 16, member 10	1.61
SLC27A5	Solute carrier family 27	1.64
SLC44A1	Solute carrier family 44, member 1	1.72
SLC8A3	Solute carrier family 8 , member 3	1.72
SNPH	Syntaphilin	1.34
SPI1	Spleen focus forming virus proviral integration oncogene spi1	1.64
STAT3	Signal transducer and activator of transcription 3	1.70
STXBP5	Syntaxin binding protein 5	2.09
SWAP70	SWAP switching B-cell complex 70kDa subunit	1.65
TAC3	Tachykinin 3	2.16

Gene Symbol	Gene Name	Fold change
TCEAL8	Transcription elongation factor A (SII)-like 8	1.59
TCN1	Transcobalamin I	2.25
TMEM206	Transmembrane protein 206	1.48
TMSL3	Thymosin-like 3	1.70
TRDMT1	tRNA aspartic acid methyltransferase 1	1.26
TSC22D1	TSC22 domain family, member 1	1.55
TSPAN33	Tetraspanin 33	2.06
TUBA1A	Tubulin, alpha 1a	1.99
TUBB4	Tubulin, beta 4	1.35
TUFM	Tu translation elongation factor, mitochondrial	1.44
VAT1	Vesicle amine transport protein 1 homolog	2.02
WBP5	WW domain binding protein 5	2.26
ZFP36L2	Zinc finger protein 36, C3H type-like 2	1.89

Supplementary table 4 – List of genes significantly down-regulated in vorinostat treated cells (in alphabetical order)

Gene Symbol	*Gene Name*	*Fold change*
ADD2	Adducin 2 (beta)	0.61
AFG3L2	AFG3 ATPase family gene 3-like 2 (yeast)	0.70
ANKRD9	Ankyrin repeat domain 9	0.43
ASNS	Asparagine synthetase	0.49
ASPSCR1	Alveolar soft part sarcoma chromosome region	0.61
ATF4	Activating transcription factor 4	0.72
ATF5	Activating transcription factor 5	0.30
B9D2	B9 protein domain 2	0.72
BTBD6	BTB (POZ) domain containing 6	0.49
C12ORF24	Chromosome 12 open reading frame 24	0.72
CCDC121	Coiled-coil domain containing 121	0.77
CENPM	Centromere protein M	0.75
CENPV	Centromere protein V	0.66
CLCN3	Chloride channel 3	0.72
CREBBP	CREB binding protein	0.74
DDX59	DEAD (Asp-Glu-Ala-Asp) box polypeptide 59	0.55
DNAJB2	DnaJ (Hsp40) homolog, subfamily B, member 2	0.57
DOCK7	Dedicator of cytokinesis 7	0.67
EIF4EBP3	Eukaryotic translation initiation factor 4E binding protein 3	0.66
EPCAM	Epithelial cell adhesion molecule	0.45
EPRS	Glutamyl-prolyl-tRNA synthetase	0.68
FBXL20	F-box and leucine-rich repeat protein 20	0.75
FGFR3	Fibroblast growth factor receptor 3	0.56
FOXRED2	FAD-dependent oxidoreductase domain containing 2	0.74
FREQ	Frequenin homolog (Drosophila)	0.42
GOLSYN	Golgi-localized protein	0.69
GPR137	G protein-coupled receptor 137	0.64
GTF2IRD2B	GTF2I repeat domain containing 2B	0.59
HS.352549	cDNA clone IMAGE:30406177 5, mRNA sequence	0.69
IL18BP	Interleukin 18 binding protein	0.61
INO80D	INO80 complex subunit D	0.71
KAT2B	K(lysine) acetyltransferase 2B	0.53
KBTBD2	Kelch repeat and BTB (POZ) domain containing 2	0.72
KLHL22	Kelch-like 22 (Drosophila)	0.69
LIME1	Lck interacting transmembrane adaptor 1	0.61
LOC440498	Hypothetical gene supported by AK001829	0.67
LOC643008	PP12104, transcript variant 1	0.58

Gene Symbol	Gene Name	Fold change
LOC729446	Smilar to AT rich interactive domain 1B (SWI1-like) isoform 1	0.73
LOC729920	Notch1-induced protein	0.80
MESP1	Mesoderm posterior 1 homolog (mouse)	0.66
MKLN1	Muskelin 1, intracellular mediator containing kelch motifs	0.68
MKNK2	MAP kinase interacting serine/threonine kinase 2	0.52
MST1	Macrophage stimulating 1	0.43
MTX1	Metaxin 1	0.67
MUC1	Mucin 1	0.36
NT5DC1	5'-nucleotidase domain containing 1	0.75
OPLAH	5-oxoprolinase	0.54
PARP3	Poly (ADP-ribose) polymerase family, member 3	0.64
PCK2	Phosphoenolpyruvate carboxykinase 2	0.57
PEX19	Peroxisomal biogenesis factor 19	0.73
PHF2	PHD finger protein 2	0.73
PHPT1	Phosphohistidine phosphatase 1	0.69
PLEKHH3	Pleckstrin homology domain containing, family H member 3	0.54
PMM1	Phosphomannomutase 1	0.55
POLE4	Polymerase (DNA-directed), epsilon 4 (p12 subunit)	0.57
PQLC1	PQ loop repeat containing 1	0.52
PSAT1	Phosphoserine aminotransferase 1	0.34
PSMB10	Proteasome (prosome, macropain) subunit, beta type, 10	0.69
RCN1	Reticulocalbin 1, EF-hand calcium binding domain	0.50
SFMBT1	Scm-like with four mbt domains 1	0.78
SIX5	SIX homeobox 5	0.74
SLC11A2	Solute carrier family 11 , member 2	0.65
SLC2A4RG	SLC2A4 regulator	0.61
SLC3A2	Solute carrier family , member 2	0.55
STAU1	RNA binding protein, homolog 1 (Drosophila)	0.76
TACSTD1	Tumor-associated calcium signal transducer 1	0.39
TMEM18	Transmembrane protein 18	0.70
TMEM87A	Transmembrane protein 87A	0.71
TMPRSS9	Transmembrane protease, serine 9	0.83
TPM1	Tropomyosin 1 (alpha)	0.66
TRMT5	TRM5 tRNA methyltransferase 5 homolog (S. cerevisiae)	0.72
TXNRD2	Thioredoxin reductase 2	0.50
UFC1	Ubiquitin-fold modifier conjugating enzyme 1	0.61
URG4	Up-regulated gene 4	0.72
VPS37C	Vacuolar protein sorting 37 homolog C (S. cerevisiae)	0.70
ZNF428	Zinc finger protein 428	0.65

Appendix 2 – Publications, presentations and awards arising from this thesis

Publications

1. **Sachith Mettananda**, Richard Gibbons and Douglas Higgs. α-Globin as a molecular target in the treatment of β-thalassemia. *Blood* 2015; 125(24): 3694-701.
2. Cheng-Tao Yang, Anna French, Pollyanna Agnes Goh, Alistair Pagnamenta, **Sachith Mettananda**, Jenny Taylor, Sam Knight, Amit Nathwani, David J. Roberts, Suzanne M. Watt and Lee Carpenter. Human induced pluripotent stem cell derived erythroblasts can undergo definitive erythropoiesis and co-express gamma and beta globins. *British Journal of Haematology* 2014; 166 (3): 435-48.

Presentations

1. **Sachith Mettananda**. CRISPR-Cas9 mediated deletion of α-globin regulatory element in human haematopoietic progenitor cells: a novel pathway of therapy for β-thalassaemia. 12th Annual Medical Sciences DPhil Day, Oxford, UK, 2015 (oral presentation).
2. **Sachith Mettananda**, Chris Fisher, Philip Hublitz, Kevin Clark, Richard Gibbons, Doug Higgs. Ameliorating β-thalassaemia by manipulating expression of the α-globin gene. Gordon Research Conferences – Red Cells, New Hampshire, USA, 2015 (poster presentation).

3. **Sachith Mettananda**, Philip Hublitz, Kevin Clark, Chris Fisher, Richard Gibbons, Doug Higgs. CRISPR-Cas9 mediated deletion of α-globin regulatory element in human haemopoietic progenitor cells: a novel pathway of therapy for β-thalassaemia. WIMM Day, Oxford, UK, 2015 (poster presentation).
4. **Sachith Mettananda**, Chris Fisher, Jackie Sloane-Stanley, Jacqueline Sharpe, Udo Oppermann, Richard Gibbons, Doug Higgs. Ameliorating β-thalassaemia by manipulating expression of the α-globin gene. 19th Haemoglobin Switching Conference, Oxford, UK, 2014 (poster presentation).
5. **Sachith Mettananda.** Ameliorating β-thalassaemia by down regulating α-globin: a search for target pathways. MHU Away Day, Oxford, UK, 2013 (oral presentation).

Awards

1. Elected as a Senior Hulme Scholar of the Brasenose College, University of Oxford (2013/14 and 2014/15).