Investigation of the Use of Histone Deacetylase Inhibitors for the Treatment of Inherited Disorders of the Glycolytic Pathway

by

Kalliopi Makarona

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy

Centre for Haematology
Department of Medicine
Imperial College London

January 2014

Declaration of originality

The research described in this thesis is the sole work of the author, except where acknowledgement is made and is not submitted in support of another degree at Imperial College London or any other educational institution.

Kalliopi Makarona

Copyright declaration

Abstract

Histone acetylation by histone acetyltransferases (HATs) and deacetylation by histone deacetylases (HDACs) regulate gene expression by activating or repressing transcription, respectively. HDAC inhibitors (HDACIs) are a diverse class of drugs used to treat haemoglobinopathies, urea cycle disorders and several types of malignancies. Recent evidence from genome-wide as well as gene-specific epigenetic studies suggest a model whereby active genes are more likely than silent genes to be hyperacetylated and increase their transcription levels in response to HDACIs, a process underpinned by the dynamic recruitment and antagonistic activities of HATs and HDACs. Based on this model and from a therapeutic perspective, I hypothesised that the ability of HDACIs to increase expression of active genes might be relevant for diseases caused by genes that encode proteins with enzymatic function. HDACI-mediated increase in gene transcription, even in the presence of missense, disease-causing mutations, might lead to increased enzymatic activity and amelioration of the cellular and clinical phenotype. I tested this hypothesis on a group of genes involved in the glycolytic and pentose phosphate pathway (GPPP) which, when mutated, cause chronic or episodic haemolytic anaemia.

Using RT-qPCR (B cell lines) and gene expression profiling (primary, in vitro generated human erythroid precursors and $CD4^+$ T cells) I found that of the 17 GPPP genes, only *Glucose-6-Phosphate Dehydrogenase* (*G6PD*) mRNA levels increased in response to HDACIs in a time-dependent manner. Epigenetic analysis in B cells by ChIP-qPCR showed that histone hyper-acetylation and increased recruitment of HATs and HDACs underpin the selective *G6PD* transcriptional activation in response to HDACIs. Pharmacological and genetic assays showed that increase in *G6PD* transcription was also dependent on Sp1, a generic transcription factor known to recruit both HDACs and HATs.

Finally, I directly tested the hypothesis that HDACIs may increase enzymatic activity in G6PD deficient cells. Using B cell lines and primary erythroid cells from patients with G6PD deficiency, I found that HDACIs induce the same epigenetic changes in the mutant as in the wild type *G6PD* gene; more importantly, they lead to increased levels of the mutant mRNA and protein, associated with an up to 3-fold increase in enzymatic activity. These findings are potentially of great therapeutic

significance for correction of G6PD deficiency in up to 300 million individuals worldwide with the polymorphic variants of G6PD deficiency (e.g., G6PDMed and G6PDA-).

Acknowledgements

First and foremost, I would like to thank my PhD supervisors, Professor Tassos Karadimitris and Dr Valentina Caputo, for giving me the opportunity to work on this project and for their help and invaluable scientific guidance throughout my project. I am also extremely grateful to Professor Irene Roberts who has supported me over the past four years and originally gave me the opportunity to work in this lab. I would also like to acknowledge the Sir Halley Stewart Trust for the funding they provided during my PhD.

Very special thanks must go to Joana Costa, for being an endless source of help and support, both professional and personal. Without her I wouldn't have made it so far and this journey definitely wouldn't have been as enjoyable! Also special thanks go out to Dr Katerina Goudevenou, Dr David O'Connor, Dr Gillian Cowan and Dr Aris Chaidos for giving up so much of their time to help me in the lab. Furthermore, I would like to thank everyone in the Roberts/Karadimitris lab for making the lab such an enjoyable place to work.

Additionally, many thanks go out to Dr Mark Layton who not only provided us with patient samples and cell lines, but also supported this project with his scientific input. On the same note, I would like to thank Dr Maria Papaioannou for her extraordinary ability to track down patients and send us vital patient material, without which I would not have been able to complete the project. For the help with enzymatic activity assays, I am very appreciative to Mr David Roper and Ms Lynn Robertson, and everyone in the Hammersmith Hospital Diagnostic Haematology lab. For the help with sequencing techniques, I would like to thank everyone in Professor Letizia Foroni's lab, especially Dr Miguel Valganon and Dr Gareth Gerrard.

Finally, special thanks must go to my David who provided me with endless support and to my friends and family for always being there for me. Importantly, a huge thanks to my parents for being the reason I am here, pursuing my dreams. This thesis is dedicated to my mum and dad.

Table of contents

Declaration of originality 2
Copyright declaration 3
Abstract 4
Acknowledgements 6
Table of Contents 7
List of Figures 13
List of Tables 17
Abbreviations 18

1 INTRODUCTION 22

1.1 Chromatin and transcriptional regulation 23
1.1.1 Histone modifications and acetylation 26
1.1.2 HATs and HDACs 27
1.1.2.1 HATs 27
1.1.2.2 HDACs 28
1.1.3 Transcriptional regulation by histone acetylation 30
1.1.4 The role of acetylation in housekeeping genes 32
1.1.4.1 The gene-specific example of inherited GPI deficiency 32
1.1.4.2 Sp1: Overview and its interactions with HATs and HDACs 33

1.2 HDAC inhibitors 37
1.2.1 Mechanism of HDAC inhibition and selectivity 37
1.2.2 Butyrate 40
1.2.3 Trichostatin A and SAHA 40
1.2.4 Impact of butyrate on gene expression 41

1.3 Erythropoiesis 42
1.3.1 Human adult haematopoiesis 42
1.3.2 Erythropoiesis: From HSCs to erythroid progenitors and RBCs 45
1.3.3 Sodium butyrate in *in vitro* and *in vivo* erythropoiesis 48

1.4 Disorders of the red cell metabolism 49
1.4.1 The glycolytic and pentose phosphate pathways 49
1.4.2 GPPP in erythroid cells 52
1.4.3 G6PD deficiency: overview 54
1.4.3.1 G6PD: Molecular Biology 55

1.4.3.2 G6PD deficiency: clinical manifestations 57
1.4.4 Other common deficiencies of the GPPP 59

1.5 Hypothesis 62

1.6 Aims 63

2 MATERIALS AND METHODS 64

2.1 Cell lines, Primary cells and patient samples 65
2.1.1 Cell lines 65
2.1.1.1 B lymphoblastoid and other cell lines 65
2.1.1.2 Cell line culture 66
2.1.2 Primary cell sources 67
2.1.3 Cryopreservation and thawing of cells 68

2.2 Primary cell selection 68
2.2.1 Preparation of mononuclear cells 68
2.2.2 Magnetic cell selection 68
2.2.2.1 $CD34^+$ cell selection 68
2.2.2.2 $CD36^+$ cell selection 69

2.3 In vitro erythroid differentiation system 69
2.3.1 Erythroid differentiation medium and cytokines 69
2.3.2 Plating and culturing 70

2.4 Drug treatments 70

2.5 Flow cytometry 71
2.5.1 Staining 71
2.5.2 Annexin V apoptosis assay 72
2.5.3 Data analysis 72

2.6 Cytospins and histologic staining 73

2.7 Gene expression assays 74
2.7.1 RT-qPCR assays 74
2.7.1.1 RNA extraction and cDNA synthesis 74
2.7.1.2 Primer design and testing 74
2.7.1.3 RT-qPCR and data analysis 76

2.7.2 Genome-wide expression arrays 77

2.8 Protein expression assay 78

2.8.1 Western blot 78

2.8.2 Protein expression quantification 80

2.9 G6PD enzymatic activity assay 80

2.9.1 Preparation of cell lysates 80

2.9.2 Spectrophotometric assay 80

2.10 Chromatin Immunoprecipitation 81

2.10.1 ChIP protocol 81

2.10.2 ChIP primers and testing 83

2.10.3 ChIP qPCR analysis 85

2.11 Reporter assays 86

2.11.1 Cloning 86

2.11.1.1 Plasmid 86

2.11.1.2 PCR amplification and purification of amplicons for cloning 86

2.11.1.3 Digestion, Ligation and bacterial cloning 88

2.11.1.4 DNA sequencing 89

2.11.2 Cell transfections 90

2.11.3 Luciferase assay 90

2.12 Mutagenesis assay 91

2.13 Dominant negative Sp1 expression plasmid 93

2.14 Bioinformatic analysis 94

2.14.1 Promoter and TF binding sites identification 94

2.14.2 Statistical analysis 94

3 RESULTS I: GLYCOLYTIC ENZYME GENE EXPRESSION IN HUMAN CELL LINES 95

3.1 Introduction 96

3.1.1 Aim of the chapter 96

3.1.2 Experimental plan 96

3.2 Control experiments 97

3.2.1 Identification of NaBu concentration 97
3.2.2 Determination of SAHA and TSA dose 99
3.2.3 RT- qPCR reference gene validation 103

3.3 GPPP gene expression upon HDAC inhibition 104

3.4 G6PD expression in normal and deficient cells 107

3.5 Impact of HDACIs on G6PD deficient cells 109

3.6 Importance of protein synthesis on the upregulation of G6PD 111

3.7 Conclusions 113

4 RESULTS II: GLYCOLYTIC ENZYME GENE EXPRESSION IN *IN VITRO* GENERATED ERYTHROID PRECURSORS 114

4.1 Introduction 115
4.1.1 Aim of the chapter 115
4.1.2 Experimental plan 116

4.2 Characterisation of the *in vitro* erythroid differentiation system 116

4.3 Effect of NaBu on erythroid differentiation *in vitro* 127
4.3.1 Effect of NaBu on cell number and differentiation of PBMCs 127
4.3.2 Effect of NaBu on cell number and differentiation of $CD34^+$ 131
4.3.3 Gene expression during erythroid differentiation 134

4.4 GPPP gene expression in normal erythroid precursors upon NaBu treatment 136

4.5 GPPP gene expression in patient erythroid cells upon NaBu treatment 138

4.6 Conclusions 148

5 RESULTS III: EPIGENETIC REGULATION OF THE GPPP GENES 149

5.1 Introduction 150
5.1.1 Aim of the chapter 151
5.1.2 Experimental plan 151

5.2 ENCODE analysis of GPPP promoters 151

5.3 The role of NEMO in the G6PD selective upregulation 159

5.4 Epigenetic status of the GPPP promoters 160

5.5 Sp1 binding on the GPPP promoters 175

5.6 Conclusions 178

6 RESULTS IV: ROLE OF THE TRANSCRIPTION FACTOR SP1 ON *G6PD* TRANSCRIPTIONAL ACTIVATION IN RESPONSE TO HDAC INHIBITORS 179

6.1 Introduction 180

6.1.1 Aim of the chapter 180

6.1.2 Experimental plan 180

6.2 Genome-wide implications of HDAC inhibition in primary proerythroblasts 181

6.2.1 Preparation of samples 181

6.2.2 Genome expression profiling 187

6.3 The role of Sp1 in the regulation of G6PD transcription 193

6.3.1 Transcriptional activity of G6PD promoter 193

6.3.2 Dependence of the G6PD transcriptional activity on Sp1 195

6.4 Conclusions 200

7 DISCUSSION 201

7.1 G6PD selectively responds to HDACIs 203

7.2 HDAC inhibition in erythropoiesis 205

7.3 Genome-wide implications of HDAC inhibition 207

7.4 Epigenetic mechanism of action of NaBu 209

7.4.1 NaBu increases histone acetylation 210

7.4.2 NaBu increases the recruitment of chromatin regulators 212

7.4.3 Sp1 is vital for the upregulation of G6PD 214

7.4.4 Epigenetic model for selective upregulation by NaBu 217

7.5 Directions for future work 217

8 BIBLIOGRAPHY 220

9 APPENDIX 241

Appendix A. List of significantly upregulated genes upon NaBu in CB-CD34$^+$ -differentiating erythroblasts. 242

Appendix B. List of significantly downregulated genes upon NaBu in CB-CD34$^+$ -differentiating erythroblasts. 255

List of Figures

Figure 1-1 The structure and covalent modifications of the mammalian nucleosome. 25

Figure 1-2 Human HDACs superfamily. 29

Figure 1-3 Dynamic regulation of gene transcription. 32

Figure 1-4 Structure, modifications and interactions of Sp1. 35

Figure 1-5 Structural features of HDACIs. 38

Figure 1-6 The haematopoietic hierarchy. 44

Figure 1-7 HSC commitment and terminal differentiation through the erythroid lineage. 46

Figure 1-8 HDACs in erythropoiesis. 48

Figure 1-9 The glycolytic pathway. 50

Figure 1-10 The pentose phosphate shunt. 52

Figure 1-11 World map distribution of G6PD deficiency. 54

Figure 1-12 The *G6PD* gene and its promoter features. 56

Figure 1-13 Most common characterised mutation sites of *G6PD*. 57

Figure 2-1 Flow cytometric analysis strategy. 73

Figure 2-2 pGL3-basic plasmid vector. 86

Figure 2-3 Transactivation assay design. 88

Figure 2-4 Mutagenesis assay strategy. 92

Figure 3-1 NaBu concentration titration in the P277 WT cell line. 98

Figure 3-2 NaBu efficiency tested by GPI expression of the PIGM deficient cell line YK. 99

Figure 3-3 TSA concentrations tested by GPI expression of the PIGM deficient cell line YK. 101

Figure 3-4 SAHA concentrations tested by GPI expression of the PIGM deficient cell line YK. 102

Figure 3-5 Effect of NaBu treatment of candidate qPCR reference genes. 104

Figure 3-6 GPPP expression upon NaBu treatment. 105

Figure 3-7 *TPI* and *GPI* expression in WT and deficient cell lines in response to HDAC inhibitors. 106

Figure 3-8 *G6PD* expression and enzymatic activity at baseline. 108

Figure 3-9 NaBu increases *G6PD* mRNA and protein expression and restores enzymatic activity to normal in cell lines. 110

Figure 3-10 SAHA increases *G6PD* mRNA and protein expression and restores enzymatic activity to normal in cell lines. 111

Figure 3-11 Translation-independent upregulation of *G6PD* by NaBu treatment. 112

Figure 4-1 Erythroid differentiation of PBMCs. 120

Figure 4-2 Erythroid differentiation of G-CSF mobilised PB $CD34^+$ cells. 122

Figure 4-3 Erythroid differentiation of CB-$CD34^+$ cells. 124

Figure 4-4 Comparison between same-donor fresh and frozen PBMCs in erythroid differentiation cultures. 126

Figure 4-5 Cell count on day 0 and day 7 of PBMC-differentiating erythroid cultures upon NaBu treatment. 128

Figure 4-6 PBMC-differentiating cells treated with a series of different NaBu concentrations. 130

Figure 4-7 Cell count on day 0 and day 7 of CB-$CD34^+$ erythroid cultures upon HDACI treatment. 132

Figure 4-8 CB-$CD34^+$-differentiating cells treated with different HDACI concentrations. 133

Figure 4-9 GPPP gene expression during erythroid differentiation in mice. 135

Figure 4-10 Same-stage differentiation upon NaBu in CB-$CD34^+$-differentiating cells. 136

Figure 4-11 NaBu increases *G6PD* gene expression and enzymatic activity in CB-$CD34^+$-derived erythroid precursors. 137

Figure 4-12 Erythroid differentiation of PBMCs isolated from G6PD Brighton. 139

Figure 4-13 NaBu treatment of WT and G6PD Brighton PBMC-differentiating erythroid cells. 140

Figure 4-14 *G6PD* expression in normal and G6PD Brighton PBMC-derived erythroid cells upon NaBu treatment. 142

Figure 4-15 *G6PD* expression in normal and G6PD Serres PBMC-derived erythroid cells upon NaBu treatment. 143

Figure 4-16 *G6PD* expression in normal and G6PD Harilaou PBMC-derived erythroid cells upon NaBu treatment. 144

Figure 4-17 *G6PD* expression in primary human normal and class I G6PD deficient erythroid cells. 145

Figure 4-18 *G6PD* expression in normal and G6PD Mediterranean PBMC-differentiating erythroid cells upon NaBu treatment. 146

Figure 4-19 *G6PD* expression in normal and G6PD African PBMC-derived erythroid cells upon NaBu treatment. 147

Figure 5-1 Transcriptional and epigenetic landscape of the *G6PD* promoter based on the ENCODE project. 154

Figure 5-2 Transcriptional and epigenetic landscape of the *GPI* promoter based on the ENCODE project. 156

Figure 5-3 Transcriptional and epigenetic landscape of the *TPI* promoter based on the ENCODE project. 158

Figure 5-4 *NEMO* mRNA expression upon NaBu treatment. 159

Figure 5-5 ChIP primers on the *G6PD* promoter. 160

Figure 5-6 Histone acetylation on GPPP gene promoters on WT cells. 161

Figure 5-7 Histone acetylation on GPPP gene promoters on G6PD deficient cells. 162

Figure 5-8 Acetylation levels on the GPPP gene promoters at baseline. 163

Figure 5-9 HAT and HDAC binding on GPPP gene promoters at baseline. 165

Figure 5-10 HAT binding on the GPPP gene promoters in WT B cells in the presence of NaBu. 166

Figure 5-11 HDAC binding on the GPPP gene promoters in WT B cells in the presence of NaBu. 168

Figure 5-12 HAT binding on the GPPP gene promoters in G6PD deficient B cells in the presence of NaBu. 170

Figure 5-13 HDAC binding on the GPPP gene promoters in G6PD deficient B cells in the presence of NaBu. 172

Figure 5-14 Polymerase II binding on the GPPP genes in response to NaBu. 174

Figure 5-15 Polymerase II recruitment on the GPPP gene promoters and *G6PD* gene body at baseline. 175

Figure 5-16 Sp1 recruitment on the GPPP gene promoters at baseline. 176

Figure 5-17 Sp1 binding on the GPPP gene promoters in response to NaBu. 177

Figure 6-1 Flow cytometric analysis of samples used for GEP. 183

Figure 6-2 RNA quality controls using the bioanalyser. 185

Figure 6-3 RT-qPCR analysis of *α-globin* and *β-globin* expression in erythroid cells of day 5 and day 7. 187

Figure 6-4 GEP analysis of the GPPP genes upon HDACI treatment. 190

Figure 6-5 Baseline mRNA expression of upregulated versus downregulated genes. 191

Figure 6-6 Motifs on the promoters of the genes that are responsive to HDAC inhibition. 192

Figure 6-7 Reporter assay to determine the promoter activity potential of *G6PD*. 194

Figure 6-8 *G6PD* mRNA expression in 293T cells. 195

Figure 6-9 Mutagenesis of Sp1 binding sites. 196

Figure 6-10 Chemical and functional elimination of Sp1 binding. 199

List of Tables

Table 1-1 Members of the HAT family. 28

Table 1-2 HDAC inhibitor classification. 39

Table 1-3 World Health Organisation (WHO) classification for G6PD deficiency. 58

Table 1-4 Enzyme disorders of the GPPP. 60

Table 2-1 Cell lines used in the study 65

Table 2-2 G6PD deficient patients. 67

Table 2-3 Magnetic $CD36^+$ cell selection antibodies and beads 69

Table 2-4 Concentration of HDACIs used in cells. 71

Table 2-5 Anti-human antibodies used for flow cytometric analysis. 72

Table 2-6 Primers used for RT-qPCR analysis. 76

Table 2-7 Polyacrylamide gel protocol. 78

Table 2-8 Antibodies used for Western blotting. 79

Table 2-9 Non-commercial buffers used for ChIP. 82

Table 2-10 Antibodies used for ChIP. 83

Table 2-11 Primers used for ChIP analysis. 84

Table 2-12 PCR primers used for *G6PD* promoter amplification. 87

Table 2-13 Primers designed to insert Sp1 binding site mutations. 93

Table 4-1 Cell numbers in two-phase liquid culture systems. 125

Table 6-1 Quality controls for microarray RNA samples. 183

Table 6-2 Number of genes upregulated and downregulated according to GEP analysis of microarray data. 188

Table 6-3 Sp1 binding sites within the 562bp *G6PD* promoter region. 197

Abbreviations

acetyl-coA	Acetyl coenzyme A
ADP	Adenosine diphosphate
ALDOA	Fructose diphosphate aldolase A
AMP	Adenosine monophosphate
ANOVA	Analysis of variance
AOE	2-amino-8-oxo-9,10-epozy-deconoyl
APC	Allophycocyanin
APS	Ammonium persulfate
ATP	Adenosine trisphosphate
Baso-EB	Basophilic erythroblasts
BFU-E	Burst forming units- erythroid
bp	Base pair
BPGM	Biphosphoglyceromutase
CA	Cytokines alone
CB	Cord blood
CBP	CREB binding protein
CD	Cluster of differentiation
Cdk2	Cyclin-dependent kinase2
cDNA	Complementary DNA
CFU-E	Colony forming units- erythroid
ChIP	Chromatin immunoprecipitation
ChIP-seq	ChIP-sequencing
CHX	Cycloheximide
CLP	Common lymphocyte progenitor
CMP	Common myeloid progenitor
CNS	Central nervous system
CNSHA	Chronic non-spherocytic haemolytic anaemia
Ct	Cycle of threshold
DAPI	4',6-diamidino-2-phenylindole
DBD	DNA binding domain
DF	Dilution factor
DHFR	Dihydrofolate reductase
DLC-1	Deleted in Liver Cancer-1
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethylsufoxide
DN	Dominant negative
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DNMT1	DNA methyltransferase 1
EACA	ε-aminocaproic acid
EBV	Epstein Barr virus

EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
EGFR	Epidermal growth factor receptor
ENCODE	Encyclopedia of DNA elements
ENO1	Enolase 1
EPO	Erythropoietin
EpoR	EPO receptor
Flt3	fms-related tyrosine kinase 3
G-CSF	Granulocyte-colony stimulating factor
G6PD	Glucose-6-phosphate dehydrogenase
GaHCl	Glutamic acid hydrochloride
GAPDH	Glyceraldehyde-3-P-dehydrogenase
GB	Gene body
GCSF-CD34$^+$	G-CSF mobilised CD34$^+$ cells
GEP	Gene expression profiling
GlyA or GPA	Glycophorin A
GMP	Granulocyte-macrophage progenitor
GNAT	Gcn5-related N-acetyltransferase
GPI	Glycosylphosphatisylinositol
GPI	Glusose phosphate isomerase
GPPP	Glycolytic and pentose phosphate pathway
F	Forward primer
FBS	Fetal bovine serum
FDA	Food and Drug Administration
FITC	Fluorescein isothiocyanate
H (1, 2A, 2B, 3, 4)	Histone (1, 2A, 2B, 3, 4)
H3K4	Histone 3 lysine 4
H3K27me3	Histone 3 lysine 27 trimethylaton
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HDACI	HDAC inhibitor
HIV	Human Immunodeficiency Virus
HK1	Hexokinase 1
HSC	Haematopoietic stem cell
hTERT	Human telomerase reverse transcriptase
IGD	Inherited glycosylphosphatisylinositol deficiency
IgG	Immunoglobulin
IL-3	Interleukin 3
IP	Immunoprecipitation
KLF	Kruppel-like factor
LMPP	Lymphoid-primed multipotential progenitor
Luc	Luciferase
MAPK	Shc/Ras/mitogen-activated kinase
MEF2	Myocyte enhancer factor-2

MEP	Megakaryocyte-erythroid progenitor
MgCl2	Magnesium chloride
MPP	Multipotential progenitor
mM	Millimolar
mRNA	Messenger RNA
MTA	Mithramycin A
MyoD	Myosin D
NaBu	Sodium butyrate
NaCl	Sodium chloride
NADPH	Nicotinamide adenine dinucleotide phosphate
NCoR	Nuclear receptor co-respressor 1
NEMO	NF-κB essential modulator
NF-κB	Nuclear factor κ-light-chain enhancer of activated B cells
NP-40	Nonyl phenoxypolyethoxylethanol-40
NTC	Non-template control
NuRD	Nucleosome remodeling deacetylase
Ortho-EB	Orthochromatic erythroblasts
PB	Peripheral blood
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PCAF	P300/CBP-associated factor
PCR	Polymerase chain reaction
PE	Phycoerythrin
PerCP	Peridinin-chlorophyll-protein complex
PFKL	Phosphofructokinase, liver
PGD	6-phosphogluconate dehydrogenase
PGK1	Phosphoglycerate kinase
PGLS	Phosphoglucolactonase
Pi	Orthophosphate
PIGM	Phosphatidylinositol glycan anchor biosynthesis, class M
PKLR	Pyruvate Kinase, Liver/RBC
Poly-EB	Polychormatic Erythroblasts
PPi	Pyrophosphate
Pro-EB	Proerythroblasts
PTM	Post-translational modification
PVDF	Polyvinylidene difluoride
RBC	Red blood cell
R	Reverse primer
RIN	RNA Integrity Number
RNA	Ribonucleic acid
RPE	Ribose-5-phosphate-3- epimerase
RPIA	Ribose-5-phosphate isomerase A

RPMI	Roswell Park Memorial Institute medium
RT	Room temperature
RT-qPCR	Reverse transcription quantitative PCR
SAHA	Suberoylanilide hydroxamic Acid
SCF	Stem cell factor
SD	Standard deviation
SDS	Sodium dodecyl sulfate
SEM	Standard error of the mean
SIRT	Sirtuin
Sp	Specificity protein
Sp1	Specificity protein 1
Stat	Signal transducer and activator of transcription
SWI/SNF	Switch/Sucrose non fermentable
TAD	Glutamin-rich reporter domains
TAF	TBP-associated factor
TALDO1	Transaldolase
TBP	TBP-binding protein
TE	Tris-EDTA
TEMED	N, N, N', N'-tetramethylethylenediamine
TF	Transcription Factor
TFIIB	Transcription factor II B
TK	Thymidine kinase
TKT	Transaketolase
TPI	Triose phosphate isomerase
Tpo	Thrombopoietin
TSA	Trichostatin A
TSS	Transcription start site
qPCR	Quantitative PCR
UV	Ultraviolet
VPA	Valproic acid
WT	Wild type
WHO	World Health Organisation

1 Introduction

The antagonistic action of two classes of enzymes, histone aceyltransferases (HATs) and histone deacetylases (HDACs), is known to regulate transcription through acetylation and deacetylation (Brownell and Allis, 2001). It has been recently shown that these enzymes co-exist in active housekeeping genes and dynamically regulate their expression (Wang et al., 2009). Inhibitors of HDACs (HDACIs) have been widely used for the treatment of cancers and more recently non-malignant diseases (Cang et al., 2009; Wagner et al., 2010; Wiech et al., 2009).

Within our lab, we have characterised inherited glycosylphosphatisylinositol (GPI) deficiency (IGD), a disease caused by a C>G transversion in the promoter of *Phosphatidylinositol glycan anchor biosynthesis, class M (PIGM)*, a housekeeping gene essential for GPI biosynthesis (Almeida et al., 2006). This mutation results in histone hypo-acetylation and repression of gene transcription via inhibition of Sp1 binding. HDACIs were shown to restore acetylation, Sp1 binding and ultimately GPI production providing a valuable treatment for patients with IGD (Almeida et al., 2007). This work provided the first direct *in vivo* evidence of the importance of histone acetylation and its relation to the transcription factor Sp1.

However, little is known about the genome-wide effects of HDACIs on housekeeping genes in cell lines or, more importantly, in primary human cells. To address this, I will investigate the effect of HDACIs on the transcriptional regulation of genes in a key housekeeping pathway, the ubiquitously active glycolytic and pentose phosphate pathway (GPPP). This pathway plays a key role in the function of red blood cells and is frequently disrupted in a number of inherited disorders. The most common of these diseases, Glucose-6-Phosphate dehydrogenase (G6PD) deficiency represents the most prevalent enzyme disorder, affecting 400 million people worldwide. Through this work I plan to explore the potential therapeutic benefits of HDACIs for the treatment of GPPP disorders.

1.1 Chromatin and transcriptional regulation

The term epigenetics is used to describe changes in the gene expression profile and overall phenotype of a cell or organism, which are controlled by mechanisms other than those causing changes in the underlying DNA sequence (Bonasio et al., 2010; Goodell, 2013); thus, the name is derived from the Greek επί- meaning over or

above and is used to emphasise the additional level of transcriptional control other than that of the DNA sequence.

In eukaryotic cells, genomic DNA is hierarchically packaged into a nucleoprotein complex called chromatin. The basic repeating unit of chromatin is that of the nucleosome, consisting of 147 base pairs of DNA wrapped around an octamer of histone proteins (made of two molecules of each H2A, H2B, H3 and H4 histones; Figure 1-1A; Campos and Reinberg, 2009; Kornberg and Thomas, 1974). Histone 1 (H1) is responsible for the stabilisation of the nucleosomal structure and small linker DNA molecules join the nucleosomes together to form a "beads-on-a-string" fiber as shown by high-resolution X-ray analysis (Luger et al., 1997). The nucleosomal unit is responsible for the compaction of the DNA sequence of a cell into chromosomes, through a dynamic folding process of these DNA-protein complexes, i.e. nucleosomes, into higher order structures (Figure 1-1A; Lenhard et al., 2012). However, not only does the nucleosome facilitate DNA compaction, but it also plays a crucial role in the regulation of the transcriptional control of gene expression, as it determines the accessibility and the binding of transcription factors (TF) to cognate DNA motifs in the gene regulatory areas including promoters, enhancers and more distal elements such as locus control regions (Bannister and Kouzarides, 2011; Thurman et al., 2012).

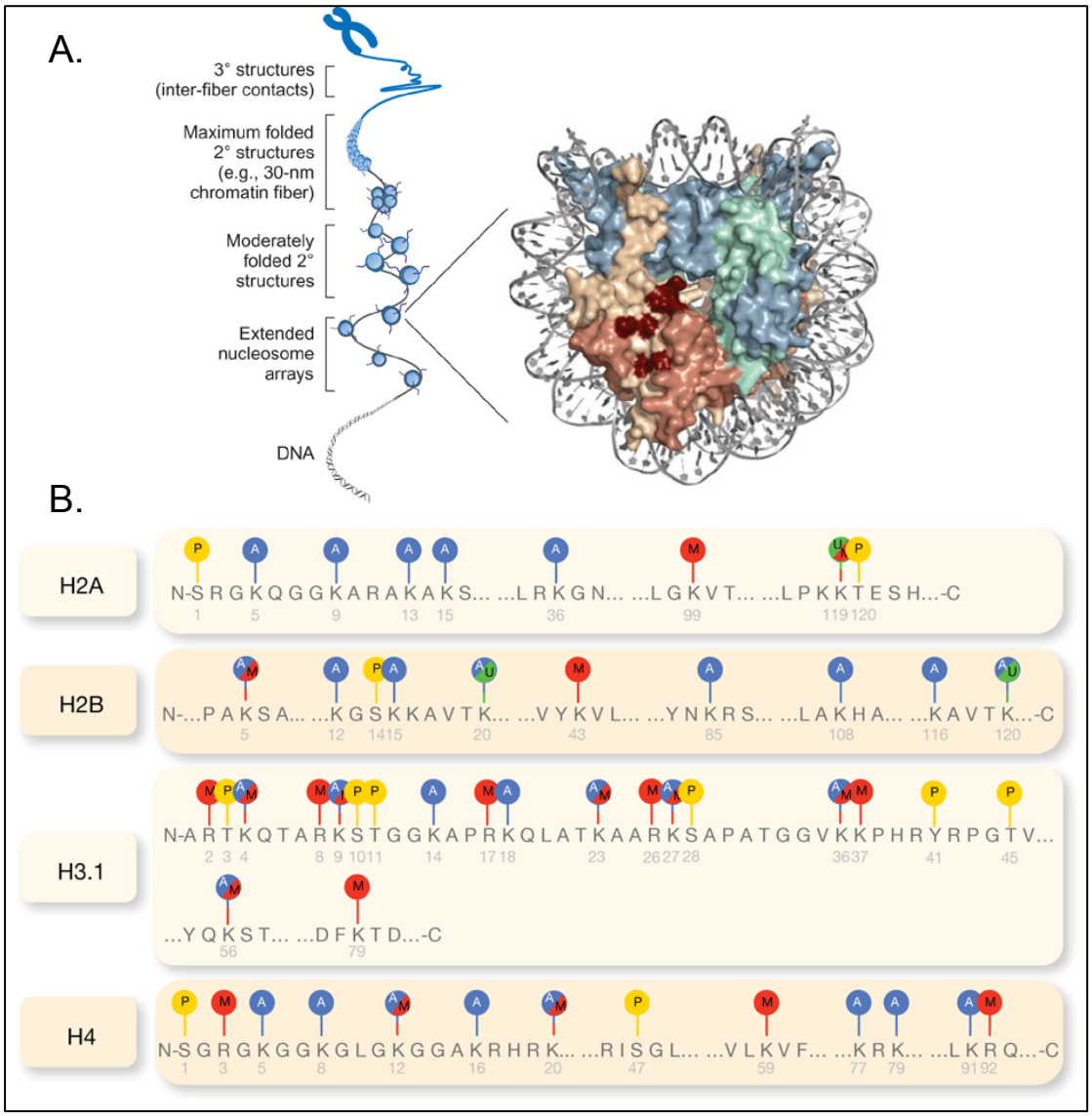


Figure 1-1 The structure and covalent modifications of the mammalian nucleosome. (A) The structure of a typical nucleosome is shown on the right. A DNA molecule is wrapped around histones H2A (yellow), H2B (light red), H3 (blue) and H4 (green), which form an octamer. The residues comprising the charged pocket are shown in dark red. Not shown here are the N-terminal histone tails protruding outside the nucleosome, which are freely accessible to post-translational modifications. These nucleosomes are packaged into higher order structures as shown on the left, resulting in the formation of the individual chromosomes. **(B)** Schematic representation of the main histone post-translational modifications: acetylation (blue), methylation (red), phosphorylation (yellow) and ubiquitination (green). Numbers in grey under the amino acids represent the position in the protein sequence. (Adapted from (Caterino and Hayes, 2007; Portela and Esteller, 2010))

Epigenetic marks (Figure 1-1B), such as DNA methylation at cytosine residues and post-translational modifications (PTMs) in the tails of histones or in the globular regions of histones, critically determine the accessibility of DNA to TFs and thus regulate transcriptional control of gene expression (Bannister and Kouzarides, 2011; Berger, 2007). Study of these epigenetic marks and their effector proteins as well as the distribution of these marks throughout the genome has resulted in a greater understanding of epigenetic regulation.

1.1.1 Histone modifications and acetylation

Histone PTMs include acetylation, methylation, phosphorylation, ubiquitination, sumoylation, Adenosine diphosphate (ADP) ribosylation, deamination and proline isomerization (Figure 1-1B; Berger, 2007; Bernstein et al., 2007; Canzio et al., 2013; Kouzarides, 2007; Li and Reinberg, 2011; Margueron and Reinberg, 2010; Weake and Workman, 2008; Zhu and Reinberg, 2011). Every histone PTM is reversible; consequently there are mechanisms in place for their removal as well as their addition (Berger, 2007). Histone PTMs have various functions; they may alter the charge of a residue to disrupt DNA-protein, protein-protein and nucleosome-nucleosome interactions and they may also form surfaces to which other proteins can bind (Bannister and Kouzarides, 2011). PTMs enhance binding to chromatin by a variety of protein domains: bromodomains can preferentially bind acetylated lysines, forkhead domains can recognise phosphorylated threonines and serines and tudor domains can bind methylated lysines and arginines (Taverna et al., 2007). In this PhD thesis, I will focus on the transcriptional control of gene expression through the action of histone acetylation.

Allfrey and colleagues (Allfrey et al., 1964) were the first to describe histone acetylation in 1964. Since then it has been established that histone acetylation is dynamically regulated by the antagonising actions of two classes of enzymes, HATs and HDACs (Brownell and Allis, 2001). Nuclear HATs use acetyl coenzyme A (acetyl-coA) to catalyse the transfer of an acetyl group to the ε-amino group of selective lysine residues in the extruding amino-terminal tails of the core histone proteins (Allfrey, 1966; Hodawadekar and Marmorstein, 2007). The negatively charged acetyl groups neutralise the highly positive charge of the histones and stabilise the influence of the electrostatic contacts between the histones and the negatively charged DNA, leading to the loosening of the histone-DNA interactions

(Luger et al., 1997) thus opening up the chromatin and facilitating transcription by allowing TFs to access the promoter. In addition, acetyl-modified histones function as 'marks' that encourage binding of co-activator complexes and chromatin remodellers such as SWI/SNF, which can further facilitate access to free DNA by promoting sliding or eviction of nucleosomes (Agalioti et al., 2002; Li and Reinberg, 2011; Nightingale et al., 1998). By removing the acetyl groups, HDACs have the opposite effects, i.e., restoration of the positive charge of the lysine residues leading to chromatin compaction and transcriptional repression.

1.1.2 HATs and HDACs

1.1.2.1 HATs

There are two major classes of HATs: type A and type B, depending on the mechanism of catalysis and also on the cellular localisation (Table 1-1). The members of the HAT A family are located in the nucleus and are responsible for the transfer of an acetyl group from acetyl-coA to an $–NH_2$ of the amino-terminal tail of a histone. The A family can be further classified into the Gcn5-related N-acetyltransferase (GNAT), MYST and CREB binding protein (CBP)/p300 subclasses, based on their homology with yeast proteins (Hodawadekar and Marmorstein, 2007; Peserico and Simone, 2011). Some HATs, such as GCN5, contain bromodomains that read specific acetylated sites on core histones and form components of chromatin-remodelling complexes (Lee and Workman, 2007). Moreover, some HATs, such as P300/CBP-associated factor (PCAF) and CBP/p300 have been shown to target non-histone proteins, for example the transcription factors p53 and MyoD (Glozak et al., 2005; Zhang and Dent, 2005). Conversely, the HAT B family is primarily cytoplasmic and function to acetylate free histones that are not part of chromatin. These HATs acetylate newly synthesised histones, regulating histone deposition in the chromatin (Parthun, 2007; 2012).

Table 1-1 Members of the HAT family. (Adapted from(Peserico and Simone, 2011))

Mammalian HATs					
Class	**Subclass**	**Yeast Homology**	**Members**	**Mechanism of action**	**Cell localisation**
A	GNAT	Gcn5	GCN5 PCAF	Transfer acetyl groups to histone N-tails after assembly into chromosomes	Nucleus
	MYST	Esa1;Sas2;Sas3	Tip60 MOF HBO1		
	CBP/p300	HAT1;Elp3; Hpa2;Nut1	CBP p300 TFIIIC complex ATF2		
B		Hat1	HAT1	Transfer acetyl groups to free histone N-tails before DNA deposition	Cytoplasm

1.1.2.2 HDACs

To date, 18 human HDACs have been identified, which, based on phylogenetic analysis, fall into four classes according to their yeast counterparts (Figure 1-2; de Ruijter et al., 2003; Gregoretti et al., 2004). HDACs within the classical family depend on Zn^{2+} for deacetylate activity and comprise classes I, II and IV. Class I HDACs are homologous to the yeast Rpd3 and comprise HDAC1, HDAC2, HDAC3 and HDAC8. Class II HDACs share homology with the yeast Hda1 and are subdivided into the subclass IIa (HDAC4, HDAC5, HDAC7 and HDAC9) and the subclass IIb (HDAC6 and HDAC10). Class IV contains only HDAC11. Finally, class III HDACs consist of seven sirtuins, which require NAD^+ cofactor for deacetylase activity. Of note, HDACIs have been designed for the treatment of several diseases and these target the Zn^{2+} domains of the Zn^{2+}-containing HDACs (Smith and Workman, 2009). The work in this thesis focuses on HDAC class I, II and IV.

Class I HDACs are ubiquitously expressed and located in the nucleus (de Ruijter et al., 2003; Delcuve et al., 2012). Knockout studies have shown that these HDACs play a key role in cell proliferation and survival (Bhaskara et al., 2008; Haberland et al., 2009; Knutson et al., 2008; Montgomery et al., 2007). Except for HDAC8, which generally has low levels of expression, class I HDACs are

components of multiprotein complexes. Due to recent evolutionary gene duplication, *HDAC1* and *HDAC2* share significant sequence homology resulting in 85% similarity in their protein products (Bradner et al., 2010; Gregoretti et al., 2004). Consequently, HDAC1 and HDAC2 can form homodimers and heterodimers and form components of the same complexes (Sin3, Nucleosome remodeling deacetylase (NuRD) and CoREST transcriptional repressing complexes; Delcuve et al., 2012). However, Wang and colleagues (Wang et al., 2009) have recently shown by a genome-wide study in $CD4^+$ T cells that these two HDACs are differentially distributed along coding and regulatory regions. Furthermore, HDAC3 is part of the nuclear receptor co-respressor 1 (NCoR) and SMRT complexes (Delcuve et al., 2012).

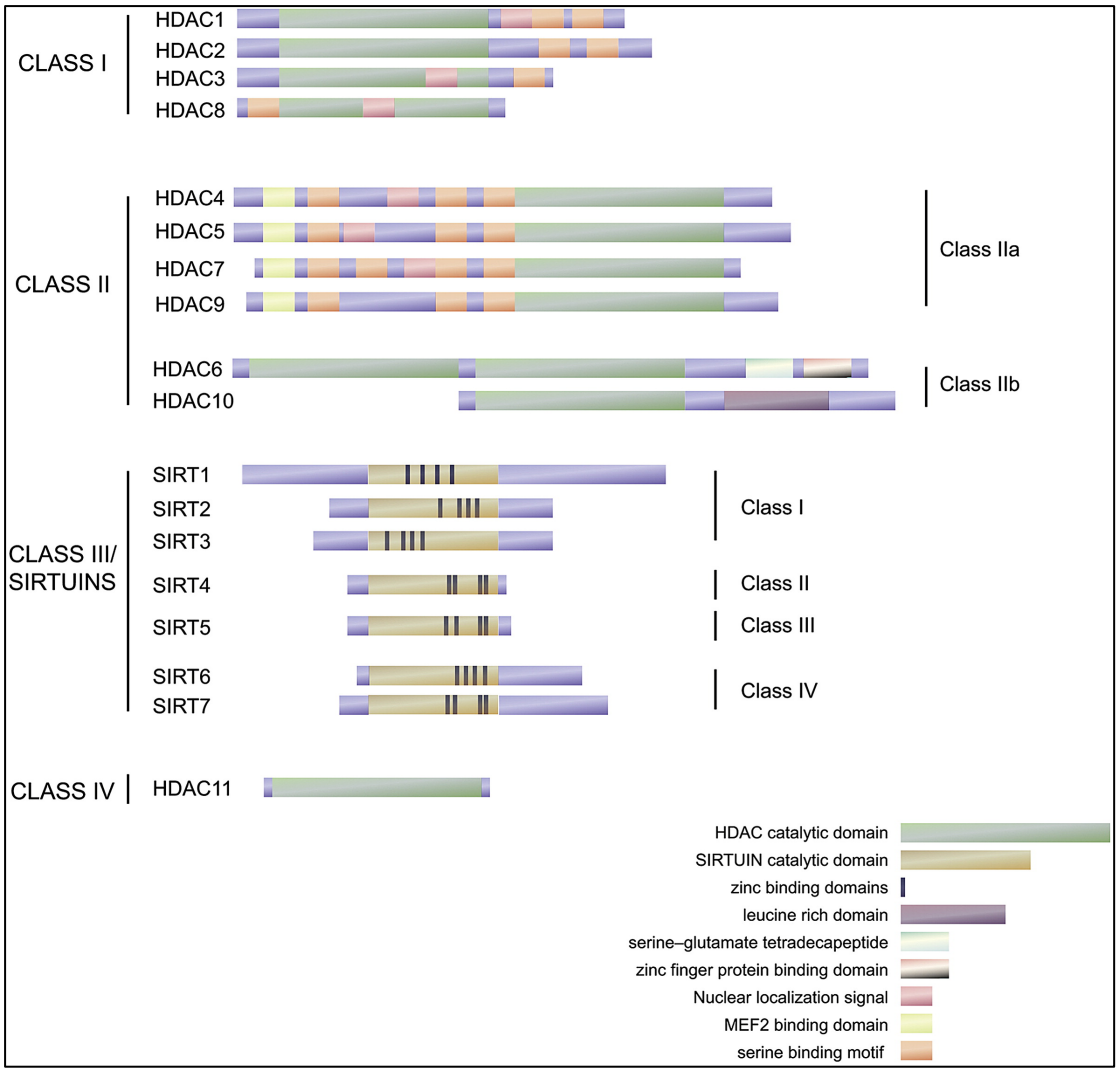


Figure 1-2 Human HDACs superfamily. HDACs fall into four classes, based on their homology to yeast proteins. Specific domains, characteristic for each class, are presented in this schematic representation. (Obtained from (Barneda-Zahonero and Parra, 2012))

Class II HDACs shuttle between the nucleus and the cytoplasm and exhibit tissue-specific expression and function (Yang and Seto, 2008). Class IIa HDACs can be found in both the nucleus and cytoplasm, as they contain intrinsic nuclear import and export signals that control the dynamic trafficking. Furthermore, they are signal transducers, containing two or three evolutionary conserved serine residues in their N-terminal domain, which are subject to reversible phosphorylation (Verdin et al., 2003; Yang and Gregoire, 2005). Once phosphorylated, they are dissociated from the genes to which they are bound and are exported from the nucleus leading to de-repression of their target genes. Class IIa HDACs have minor intrinsic deacetylase activity and usually interact with their target genes through TFs (Verdin et al., 2003), such as the Myocyte enhancer factor-2 (MEF2) family, transcriptional co-repressors, the heterochromatin protein HP1a and the NCoR and SMRT complexes in association with HDAC3 (Fischle et al., 2002). Nevertheless, it has been recently suggested that they are indeed able to bind acetylated lysines by acting as bromodomains and further recruit chromatin-modifying enzymes (Bradner et al., 2010).

Class IIb HDACs have duplicated catalytic activity domains, although duplication is only partial in HDAC10. HDAC6 shuttles in and out of the nucleus and is localised in the cytoplasm in the absence of a stimulus (Verdel et al., 2000). Similarly, HDAC10 is mainly cytoplasmic, but it can translocate to the nucleus upon stimulation (Kao, 2001). Although little is known about the function of HDAC10, it is now known that HDAC6 is able to bind to the regulatory and coding areas of genes and thus directly regulate transcription (Wang et al., 2009).

In addition, although it is known that HDAC11 has significant sequence similarity with classes I and II (de Ruijter et al., 2003; Gregoretti et al., 2004), very little is currently known about its function.

1.1.3 Transcriptional regulation by histone acetylation

Extensive work on individual genes has shown that the dynamic balance of histone acetylation and deacetylation is a critical determinant of transcriptional regulation, i.e., activation versus repression. Furthermore, HDACs have been traditionally considered to act as co-repressors that inhibit transcription through binding to gene promoters; HDACs are then replaced by co-activator HATs to drive transcription upon signal transduction (Berger, 2007; Xu et al., 1998). Genome-wide

analysis in yeast has shown that HAT binding is positively associated with gene transcription (Kurdistani et al., 2004). Nevertheless, the exact role of HDACs has been controversial as conflicting studies suggest that HDACs can be associated with both gene repression (Kadosh and Struhl, 1997) and gene activation (Wang, 2002).

A recent detailed genome-wide study by Wang and colleagues (Wang et al., 2009) that involved ChIP-seq analysis of primary $CD4^+$ T cells provided new insights into the role of histone acetylation and deacetylation of active (or housekeeping), primed (or poised) and inactive (or repressed) genes (Figure 1-3). Specifically, it was found that active genes are associated with high binding levels of both HATs and HDACs to the promoters and gene bodies; it was speculated that in this context, HDACs are required to reset the chromatin status by removing the acetyl groups after the completion of each round of transcription to prevent hyper-acetylation of the gene. Conversely, the promoters of primed genes demonstrate low levels of HATs and HDACs in association with H3K4 trimethylation, thus maintaining genes in a poised state, primed for activation. Repressed genes have no detectable HATs and HDACs and are maintained in a repressed transcriptional state primarily by the Polycomb complex and the H3K27me3 repressive mark. This study provided new perspectives on the roles of HDACs, especially for transcriptional control of active genes, suggesting that HDAC inhibition might be a simple means to influence their transcription.

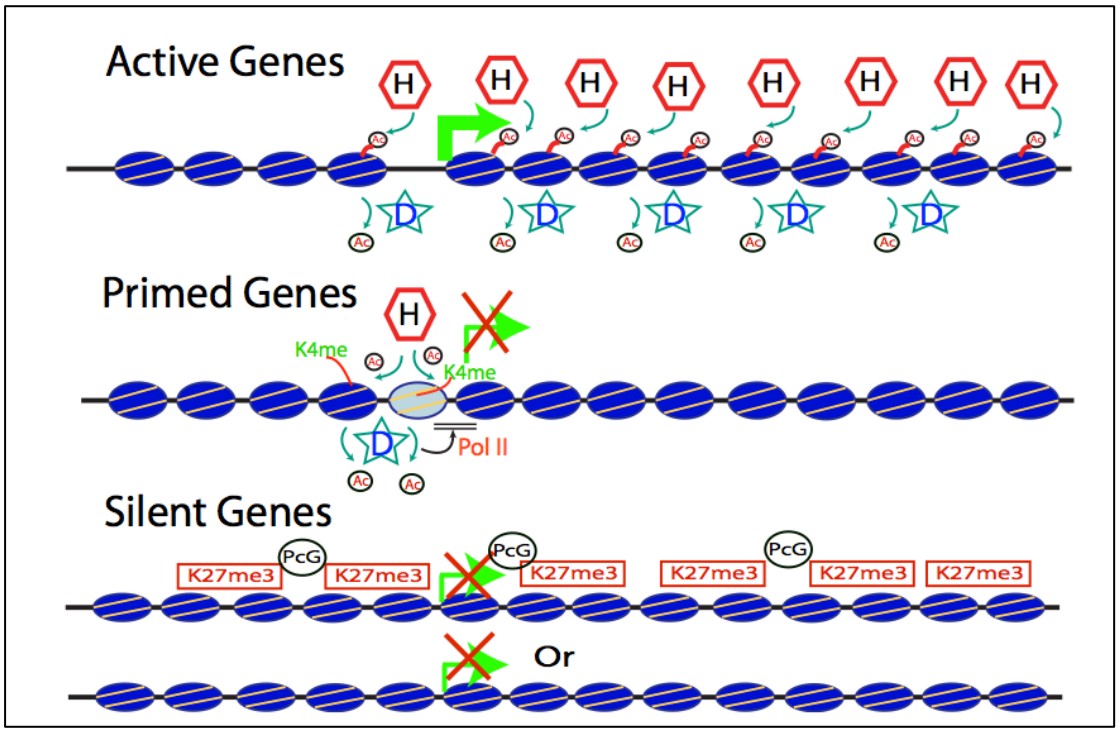


Figure 1-3 Dynamic regulation of gene transcription. High levels of HATs and HDACs are associated with active genes. HDACs remove the acetyl groups to reset chromatin state. Low levels of HATs and HDACs and the presence of H3K4 methylation show how HDACs inhibit transcription by polymerase II. Silent genes have neither HATs nor HDACs detected on their promoters. In some cases polycomb group complexes are present. (Adapted from (Wang et al., 2009))

1.1.4 The role of acetylation in housekeeping genes

1.1.4.1 The gene-specific example of inherited GPI deficiency

Although the role of histone acetylation in individual genes has been studied in different *in vitro* cellular systems, the dissection of the molecular pathogenesis of IGD provided the first direct proof of the importance of histone acetylation and its relation to the TF Sp1 in humans *in vivo*.

Specifically, in 2006 our lab demonstrated the genetic and biochemical basis of IGD, a rare autosomal recessive disorder (Almeida et al., 2006). IGD, characterised by life-threatening thrombosis and treatment-intractable epilepsy, is caused by the disruption of the biosynthesis of GPI, a glycolipid to which tens of proteins are attached before they are expressed on the cell surface as GPI-linked molecules. The genetic defect in IGD is a C>G transversion in the promoter of *PIGM*, a housekeeping gene essential for GPI biosynthesis. This point mutation disrupts binding of the TF

Sp1 to the core promoter, reducing levels of histone acetylation at the promoter and transcriptional activity of *PIGM*, therefore constituting the first example of a gene-specific histone hypo-acetylation disease.

Furthermore, it was demonstrated that HDACIs, including butyrate, which is approved for clinical use, are able to effectively restore histone acetylation, transcriptional activity and synthesis of GPI as well as the expression of GPI-linked proteins on the surface of IGD patient cells *in vitro* (Almeida et al., 2006; 2007; Caputo et al., 2013). Additionally, a clinical trial of butyrate in a child with IGD suffering from treatment-intractable epilepsy resulted in restoration of the *PIGM* promoter acetylation, transcription and GPI biosynthesis *in vivo* and importantly in prompt and complete resolution of all epileptic episodes for the first time in 12 years (Almeida et al., 2007). The findings from this work have wider ramifications for understanding gene transcriptional and epigenetic regulation and the principles of epigenetic therapy.

1.1.4.2 Sp1: Overview and its interactions with HATs and HDACs

Sp1 (Specificity protein 1) is a TF and member of the Specificity Protein/Kruppel-like Factor (Sp/KLF) family, whose members share a highly conserved DNA binding domain (DBD) with three Cys_2His_2- type zinc fingers that allow binding to GC- and GT-rich boxes (Black et al., 2001; Philipsen and Suske, 1999; Wierstra, 2008). The Sp/KLF family is divided into the Sp family, which favours GC-rich domain binding and the KLF family. The Sp family is subdivided into the Sp1-4 and Sp5-9 classes that are distinguished by the presence or absence of glutamin-rich reporter domains (TADs), respectively (Bouwaman and Philipsen, 2002; Davie et al., 2008; Li and Davie, 2010; Tan and Khachigian, 2009). Within the Sp1-4 class, Sp1 and Sp3 are ubiquitously expressed, whereas Sp2 and Sp4 display a tissue-specific expression pattern (Hagen et al., 1992).

Sp1 has two isoforms, with isoform α having a longer N-terminal domain than isoform b (Figure 1-4A). It is an 110kDa nuclear TF that binds to GC-rich motifs (such as the 5'-G/T-GGGCGG-G/A-G/A-C/T-3' or 5'G/T-G/A-GGCG-G/T-G/A-G/A-C/T-3'; Briggs et al. 1986) and positively or negatively regulates the transcription of both TATA-less and TATA-containing genes. Sp1 exhibits 90% DBD homology to Sp3, therefore these two members of the Sp family often compete for

binding to the same GC-boxes. Consequently the Sp1:Sp3 ratio often plays a key role in gene regulation (Black et al., 2001; Bouwaman and Philipsen, 2002; Resendes and Rosmarin, 2004; Li and Davie, 2010); on some promoters Sp1 cooperates with Sp3, whereas in others Sp3 suppresses the Sp1-mediated activation.

Sp1 is known to be a transactivator (Kadonaga et al., 1988). It binds to other Sp1 molecules to form tetramers and has the ability to transactivate via a single Sp1 binding site (Courey and Tijan, 1988; Courey et al., 1989), to transactivate synergistically with itself via two or more binding sites (Courey and Tijan, 1988; Courey et al., 1989), and finally to superactivate the Sp1-mediated transcription with one Sp1 molecule binding to the DNA and the other Sp1 molecules interacting with it (Pascal and Tjian, 1991). Four domains of Sp1 (Figure 1-4A) are involved in transcriptional activation: domains A and B contain the TADs and are required for all the above transactivation activities, domain C possesses low transactivation potential and domain D is essential for synergistic transactivation (Courey and Tijan, 1988; Courey et al., 1989; Wierstra, 2008). Of note, a significant difference between Sp1 and Sp3 is that Sp3 cannot synergistically activate transcription of promoters containing more than one Sp1/Sp3 binding sites.

Through its binding to the GC-rich motifs, Sp1 is shown to regulate transcription through protein-protein interactions with other transcription factors (Figure 1-4B), such as c-myc (Parisi et al., 2007), c-jun, Stat-1 (Canaff et al., 2008) and Ets-1 (Rosmarin et al., 1998), and with components of the basal transcriptional machinery. Indeed, the two TADs directly interact with both TBP (TATA-binding protein) and TAF4 (TBP-associated factor 4; Courey and Tijan, 1988; Courey et al., 1989; Wierstra, 2008). However, Sp1 does not interact with TAF1, TAF2 and TFIIB (Baniahmad et al., 1993; Chen et al., 1994), indicating its importance for the recruitment of the TBP/TFIID complex and therefore stimulation of transcription initiation but not elongation (Wierstra, 2008).

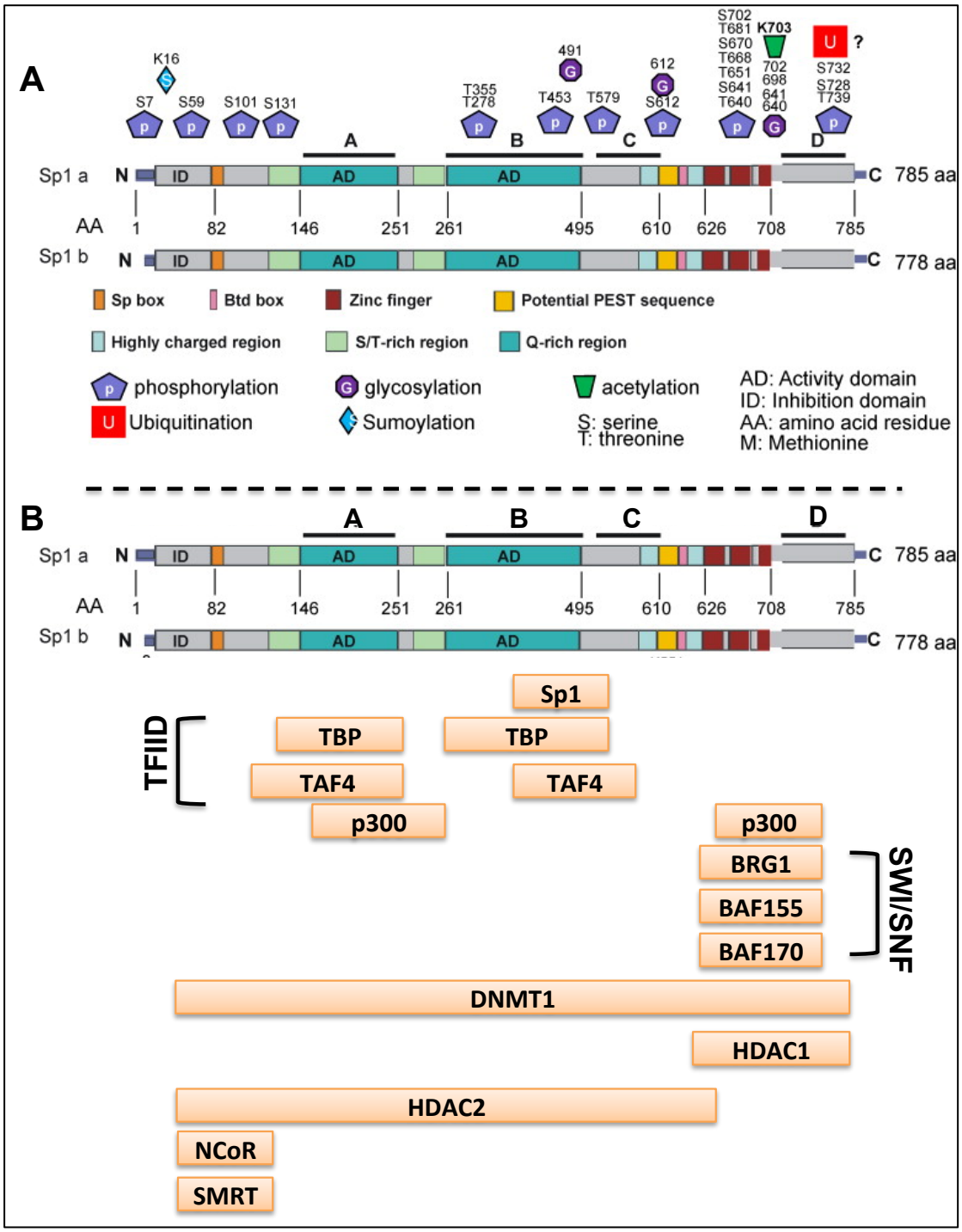


Figure 1-4 Structure, modifications and interactions of Sp1. (A) This schema shows the structural characteristics and the known post-translational modifications of human Sp1. The domains A, B, C and D that are involved in transcriptional activation are highlighted. **(B)** Interaction partners of Sp1 are shown under the corresponding site of interaction. (Adapted from (Li and Davie, 2010; Wierstra, 2008))

Sp1 has also been linked to chromatin remodelling due to its interactions with HATs and HDACs. Sp1 is shown to directly bind to the HAT p300, which acts as a co-activator for the two TADs and is required for example in the transactivation of the *p21* promoter by Sp1 (Kundu et al., 2000). Sp1 also synergises with CBP even though they do not interact directly (Kundu et al., 2000). Furthermore, Sp1 recruits the chromatin remodelling complex SWI/SNF, via direct binding to the subunits BAF155, BAF170 and BRG1 (Kadam, 2000). Sp1 can positively regulate gene expression also through binding to the nucleosomal DNA to serve as a boundary for the spreading of heterochromatin (Ishii and Laemmli, 2003). Similarly, Sp1 has been shown to protect CpG islands from methylation (Brandeis et al., 1994), although recent data from our lab shows that this may not be the case (Caputo et al., 2013). Sp1 has also been demonstrated to negatively regulate gene expression through direct binding to HDAC1 and HDAC2 though the Sin3 complex (described in 1.1.2.2) and DNA methyltransferase 1 (DNMT1) and has been shown to repress genes, among which are *human telomerase reverse transcriptase* (*hTERT)* and *thymidine kinase (TK;* Doetzlhofer et al., 1999; Hou et al., 2002).

Although Sp1 was initially thought to serve as an activator of housekeeping and other TATA-less genes, it is now known to target genes that have key roles in cell proliferation and oncogenesis, i.e. they are required for growth, anti-apoptosis, angiogenesis and metastasis. For example, Sp1 activates the transcription of the genes encoding Cyclin E, Cyclin-dependent kinase2 (Cdk2), E2F-1 and c-myc, and therefore induces the transition from the G1-phase to the S-phase in quiescent cells (Wierstra, 2008). Additionally, it activates the transcription of the seven cyclin-dependent inhibitors, $p15^{INK4B}$, $p16^{INK4A}$, $p18^{INK4C}$, $p19^{INK4D}$, $p21^{WAF1/CIP1}$, $p27^{KIP1}$ and $p57^{KIP2}$ that are responsible for cell cycle arrest and required for transient senescence and long-term quiescence (Sherr and Roberts, 1999). Sp1 has also been implicated in non-malignant diseases, such as Huntington's disease, through the disruption of Sp1 binding on the *dopamine D2 receptor* gene (Dunah, 2002) and recently in our lab, with IGD via the mutation of an Sp1 binding site, resulting in the silencing of *PIGM*, a housekeeping gene (described in 1.1.4.1).

1.2 HDAC inhibitors

HDACIs are a class of compounds shown to interfere with the activity of histone deacetylases. Inhibitors of the Zn^{2+}-dependent HDACs, i.e. classes I, II and IV, were originally discovered by their ability to promote differentiation and cell cycle arrest in cultured erythroleukaemic cells and only later were identified as histone deacetylase inhibitors (Friend et al., 1971; Marks, 2010; Richon et al., 1996). Soon after, it was recognised that HDACs are upregulated in a number of cancers or are aberrantly recruited to DNA loci upon chromosomal translocations, especially in haematological malignancies, such as non-Hodgkin's lymphoma and acute leukaemia. Consequently, the specificity of HDACIs towards tumour cells led to their development as anticancer drugs (Cang et al., 2009; Wagner et al., 2010). More recently, clinical studies that use HDACIs have been extended to non-cancer diseases, such as sickle cell anaemia, cystic fibrosis, inflammatory disorders, neurodegenerative disorders and HIV infection (Wiech et al., 2009). To date, four pharmacological HDACIs have been approved by the Food and Drug Administration (FDA) [SAHA (Zolinza, Vorinostat), Merck Research Laboratories, Romidepsin or depsipeptide or FK228 (Istodax), Gloucester Pharmaceuticals, sodium phenylbutyrate (Buphenyl), Ucyclyd Pharma and Valproic Acid (Depakene), Abbot Laboratories] and more than ten compounds are in clinical trials, such as depsipeptide and trichostatin A (Glaser, 2007; Smith and Workman, 2009).

1.2.1 Mechanism of HDAC inhibition and selectivity

The active site of Zn^{2+}-dependent HDACs consists of a tubular pocket with two histidine residues, two aspartic acid residues, one tyrosine residue (the latter is substituted by another histidine in class IIa HDACs) and a Zn^{2+} ion located at the base of the pocket (Finnin et al., 1999). HDACIs fit into the pocket and inhibit the deacetylase activity (Figure 1-5; Finnin et al., 1999). Generally, HDACIs display three common structural characteristics, which are in line with the "cap-linker-chelator" pharmacophore model (Figure 1-5). They contain a Zn^{2+}-binding moiety, an alkyl-, aryl- or vinyl- chain linker spanning the length of the tubular pocket, and a cap that blocks the active site by interacting with the external surface of the HDAC (Finnin et al., 1999; Marks, 2007; Marks and Breslow, 2007).

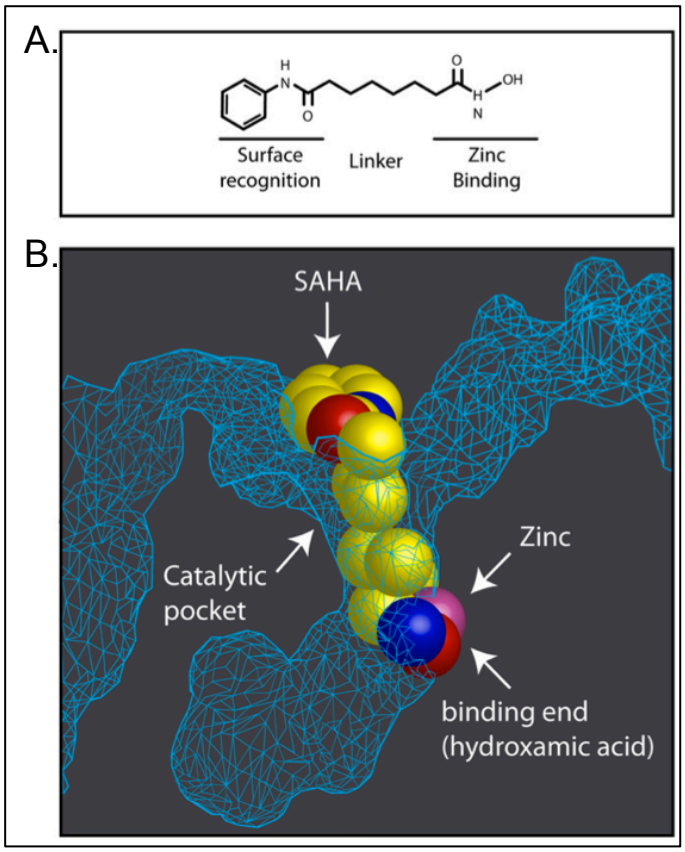


Figure 1-5 Structural features of HDACIs. (A) An example of a hydroxamic acid HDACI displaying the common "cap-linker-chelator" structure of HDACIs. **(B)** Representation of the crystal structure of the HDACI SAHA, inserted into the tubular pocket of the enzyme. (Obtained from (Marks, 2010))

Depending on their Zn^{2+} -binding group, HDACIs fall into six structurally diverse classes (Table 1-2). These are: short-chain fatty acids [e.g. the butyrates and valproic acid (VPA)], hydroxamic acids [e.g. Trichostatin A (TSA) and suberoylanilide hydroxamic acid (SAHA)], cyclic tetrapeptides containing a 2-amino-8-oxo-9,10-epozy-deconoyl (AOE) moiety [e.g. Depsipeptide FK228], epoxides not containing an AOE moiety [e.g. Depudexin], benzamides [e.g. MS-275] and electrophilic ketones [e.g. Trifluoromethyl ketones](de Ruijter et al., 2003; Delcuve et al., 2012; Marks, 2010; Marks et al., 2000).

Table 1-2 HDAC inhibitor classification. [a]

Group	Compounds	In vitro IC50 range
Short-chain fatty acids	Sodium Butyrate	mM
	Phenylbutyrate	mM
	Valproic Acid	mM
Hydroxamic acids	Trichostatin A	μM
	SAHA	nM
	Scriptaid	μM
	Oxamflatin	nM
Cyclic tetrapeptides	Depsipeptide	nM
	Trapoxin	nM
Epoxides	Depudexin	μM
Ketones	Trifluoromethyl ketones	μM
Benzamides	MS-275	μM

[a] Examples of compounds falling into each of the six structural HDACI classes and their *in vitro* concentrations used are listed in this table. (Adapted from (Bantscheff et al., 2011; Bertrand, 2010; de Ruijter et al., 2003; Dzierzak and Philipsen, 2013))

A controversial issue in the literature regarding HDACIs has been that of their isoform selectivity. HDACIs have generally been considered pan-inhibitors, implying inhibition of all Zn^{2+}-containing HDACs. However, a recent study by Badner and colleagues (Bradner et al., 2010) has shown that almost all classical inhibitors fail to target class IIa HDACs. Furthermore, a very detailed study by Bantscheff and colleagues (Bantscheff et al., 2011), which took into account the fact that HDAC activity is mostly associated with multiprotein complexes (as described in 1.1.2.2), was carried out using 16 HDACIs with different chemical structures tested in 6 human cell lines and 6 mouse tissues. Even though this study confirmed the results by Bradner et al. regarding class IIa HDACs, it conflicted with the selectivity data seen in other studies (Bertrand, 2010; Blackwell et al., 2008). Nevertheless, it confirmed the fact that different HDACIs have varied potency against each HDAC.

1.2.2 Butyrate

Butyric acid (butyrate) is a naturally occurring short-chain fatty acid produced during the natural synthesis and the breakdown of longer-chain fatty acids *in vivo*. Significant sources of butyrate in the diet include dairy products, fruit and vegetables, which are metabolised by the endogenous intestinal bacterial fragmentation of fibers in the human body (Heerdt et al., 1999; Miller, 2004b; Santillo and Albenzio, 2008).

In terms of HDAC inhibition, butyrate was the first HDAC inhibitor to be identified in 1949 (Candido et al., 1978; Sealy and Chalkley, 1978; Stadtman and Barker, 1949) and since then butyric acid and its derivatives have been used for the treatment of haemoglobinopathies (Perrine et al., 2010), urea cycle disorders (Batshaw et al., 2001), sickle cell anaemia (Dover et al., 1994) and most recently for the treatment of IGD (Almeida et al., 2007). Additionally, it is currently in advanced clinical trials for the treatment of cancers (Andriamihaja et al., 2009; Perrine et al., 2007; Yoo and Jones, 2006). The compound used for pharmaceutical purposes is sodium phenylbutyrate, which is a prodrug, rapidly metabolised to the active form phenylacetate. Despite its therapeutic potential, the form of sodium phenylbutyrate currently marketed and used for the treatment of diseases has a short half-life and is subjected to first pass hepatic clearance, explaining the milligram doses the patients are required to take in order to achieve therapeutic concentrations *in vivo* (Yoo and Jones, 2006).

Sodium butyrate (NaBu), a sodium salt of butyric acid and member of the short chain fatty acids, is the form used for laboratory research. It has been reported to have potency in the mM range (Bolden et al., 2006) and is shown to inhibit most classes of HDAC except class III (SIRT) and has very low potency for class IIb HDACs (Blackwell et al., 2008). It is now known from crystallographic data, kinetic analysis and co-immunoprecipitation studies that butyrate is a competitive and reversible inhibitor of histone deacetylases (Sekhavat et al., 2007).

1.2.3 Trichostatin A and SAHA

A class of HDACIs that is more potent than the short-chain fatty acids is that of the hydroxamic acids (Bantscheff et al., 2011), the most potent of which is Trichostatin A (TSA), which is effective at nanomolar concentrations (Bantscheff et al., 2011; Yoshida et al., 1990). Although initially used as an anti-fungal agent, it was

later shown to have proliferation inhibitory properties against cancer cells (Chang et al., 2011; Vigushin et al., 2001). TSA is shown to inhibit class I and II HDACs, with very low efficiency against class IIa (Bantscheff et al., 2011; Blackwell et al., 2008; Bradner et al., 2010).

Another HDACI of the same class is SAHA, which was the first to earn regulatory approval by the FDA in 2006 and since then has been used for the treatment of cutaneous T cell lymphoma (Mann et al., 2007). Since FDA approval, it has been extensively used in research and shown to act against other types of haematological and solid tumour cancers (Marks, 2007; Siegel et al., 2009), as well as to have therapeutic potential for other non-malignant conditions (Zhao et al., 2012). SAHA is a potent inhibitor used at micromolar concentrations and similarly to TSA it is shown to inhibit class I and II HDACs, with very low efficiency against class IIa (Bantscheff et al., 2011; Blackwell et al., 2008; Bradner et al., 2010). Interestingly, recent studies regarding SAHA, have raised concerns regarding its absorption and metabolism (Fraczek et al., 2013).

1.2.4 Impact of butyrate on gene expression

Although HDACIs, including butyrate, can lead to widespread histone hyper-acetylation, transcription of only 2-25% of genes is shown to be affected (Davie, 2003; Delcuve et al., 2012; Mitsiades et al., 2004; Sealy and Chalkley, 1978; Sekhavat et al., 2007; Van Lint et al., 1996). The time in culture, the concentration and type of HDACI used as well as the type of cells used determine the number of genes altered in transcription. The number of affected genes increases as time and concentration increase likely as a result of downstream rather than direct effects (Peart et al., 2005).

Strikingly, Mariadason and colleagues (Mariadason et al., 2000) reported for the first time that butyrate upregulates and downregulates equal number of genes in colonic epithelial cells. Following this study, other researchers showed similar results of equal number of genes being upregulated and downregulated by butyrate and other inhibitors (Joseph et al., 2004; Peart et al., 2005). To further enrich our knowledge, Rada-Iglesias and colleagues (Rada-Iglesias et al., 2007) performed a combination of Chromatin Immunoprecipitation (ChIP) with microarrays (ChIP-chip) in hepatocarcinoma (HepG2) and adenocarcinoma (HT-29) cell lines. Treatment with

HDACIs, including butyrate, in these cells showed that a number of genomic regions close to transcription start sites were deacetylated, as opposed to the global acetylation increase that was observed. Interestingly enough, these promoter regions correspond to genes that are downregulated under butyrate exposure.

It is important to highlight that to date almost all the genome-wide studies conducted to evaluate the gene expression changes in response to HDACIs involved cell lines rather than primary human cells. The exception to this, is the study by Wang and colleagues (Wang et al., 2009), also discussed in 1.1.3, in which primary human $CD4^+$ T cells were treated with NaBu and TSA to evaluate global gene expression of genes in response to HDACIs.

Butyrate responsive elements have been characterised within the promoters of the butyrate-responsive genes (Davie, 2003; Majumdar et al., 2012; Siavoshian et al., 1997). These can be categorised into two separate groups. The first group consists of genes that are either induced or repressed and have a common AGCCACCTCCA sequence, suggesting that they are bound by a common transcription factor. Examples of this group are cyclin D1 and intestinal trefoil factor genes that are repressed, as well as the metallothionein IIA and calbindin-D28k that are induced. In the 2nd group, there are genes, such as the IGF-binding protein 3 and the Cdk2 inhibitor $p21^{Waf/Cip1}$, which share an Sp1/Sp3 binding site within the butyrate responsive elements. Sp1 and Sp3 are transcription factors that are ubiquitously expressed and act as both activators and repressors of gene expression and, as discussed above, Sp1 interacts with HATs and with HDACs.

1.3 Erythropoiesis

1.3.1 Human adult haematopoiesis

Haematopoiesis is the process through which blood cells are formed. The liver forms the primary site of haematopoiesis during fetal life. From the second trimester onwards haematopoiesis begins in the bone marrow, where it remains throughout childhood and adult life. In adult life about 10^{12} haematopoietic cells are required each day. Therefore, to avoid depletion of the haematopoietic cells, a system to maintain a self-renewing stem cell pool, which can also differentiate into all types of mature blood cells is needed (Doulatov et al., 2012; Orkin and Zon, 2008).

Haematopoiesis is organised as a developmental hierarchy (Figure 1-6) with the multipotent haematopoietic stem cells (HSCs) at the apex. HSCs are rare long-lived multipotent cells occurring at a frequency of approximately 1 cell per 10^6 and can be identified based on the expression of specific cell-surface markers (Gangenahalli et al., 2005). The cell-surface marker CD34 was the first marker found to enrich HSCs and is currently known to mark HSCs as well as more committed early progenitors. The need for the characterisation of additional markers to allow for fractionation of the stem cell and progenitor pool, led to the identification of the stem cell marker CD90 (Thy1; Baum et al., 1992). Further studies identified CD45RA and CD38 as markers of more differentiated non-HSC progenitors (Bhatia et al., 1997; Conneally et al., 1997), introducing a picture of human HSCs as $CD34^+CD38^-$ $Thy1^+CD45RA^-$.

HSCs are necessary for lifelong blood production and have the capacity to self-renew, generating more HSCs or to differentiate. During their differentiation, HSCs give rise to a series of haematopoietic progenitors (Figure 1-6), which are restricted in their developmental potential and will undergo a gradual fate restriction to terminally differentiate into mature blood cells (Doulatov et al., 2012; Giebel and Punzel, 2008; Laurenti and Dick, 2012).

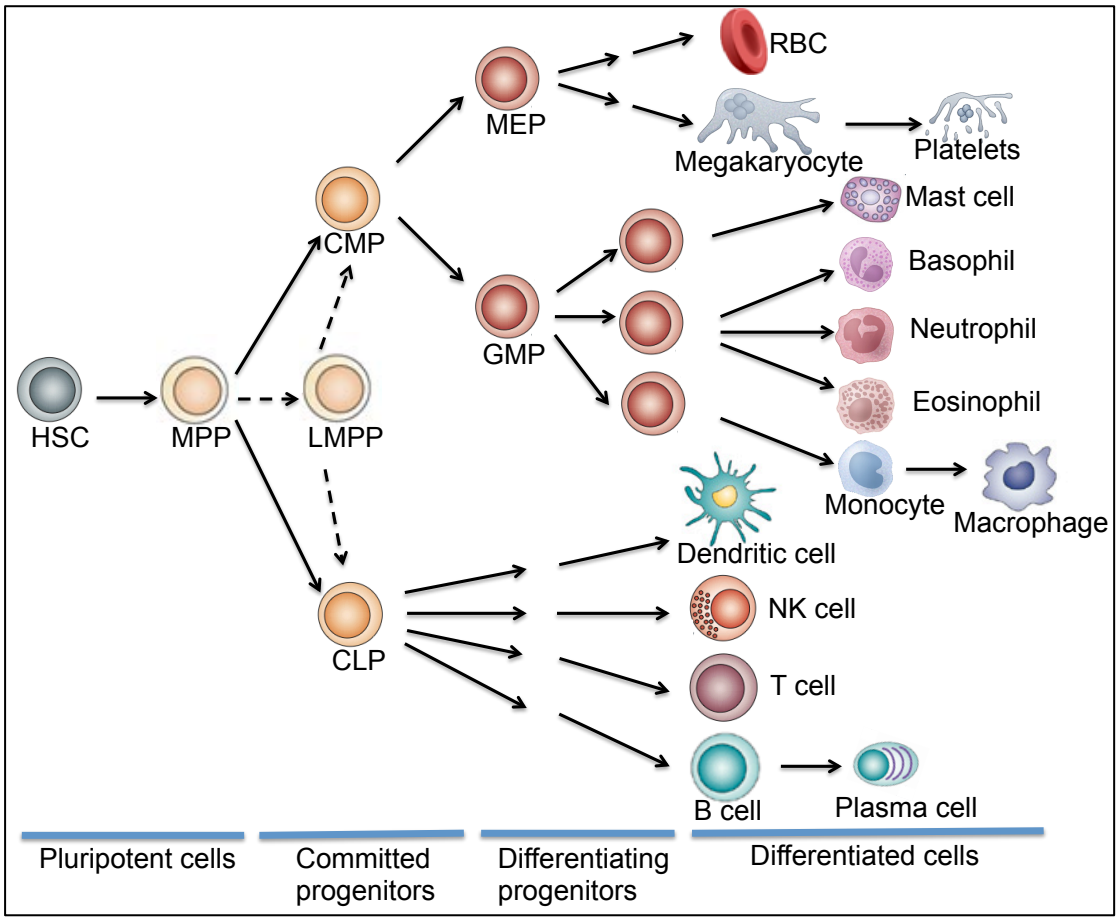


Figure 1-6 The haematopoietic hierarchy. Pluripotent haematopoietic stem cells (HSCs) give rise to multipotent progenitors (MPP), which will then divide symmetrically to self-renew or asymmetrically to differentiate into either a common lymphocyte progenitor (CLP) or a common myeloid progenitor (CMP). The CMP in turn gives rise to either a granulocyte-macrophage progenitor (GMP), or a megakaryocyte-erythroid progenitor (MEP). These progenitors further differentiate through multiple stages into mature haematopoietic cells. Recently, it has been accepted that there is one more cell type, the lymphoid-primed multipotential progenitor (LMPP) that can give rise to both CLPs and CMPs. (Reproduced and adapted from (Dzierzak and Philipsen, 2013))

At present, the mechanisms underlying self-renewal and lineage specification are not well understood. A dominant model of haematopoiesis posits that lineage transitions are controlled by a small number of TFs that are sequentially expressed and restricted to specific lineages (Iwasaki and Akashi, 2007). However, recent genome-wide studies suggest that there is a more complex architecture in regulatory circuits, which involves a complex interplay of TFs that are expressed at varying levels across multiple lineages (Novershtern et al., 2011; Suzuki et al., 2009). This

differential expression of TFs seems to play a key role in developmental decisions during haematopoiesis.

1.3.2 Erythropoiesis: From HSCs to erythroid progenitors and RBCs

Erythropoiesis is the process that generates mature erythrocytes from HSCs through intermediary forms of progenitor and precursor cells that undergo terminal differentiation and maturation to give rise to red blood cells (RBCs; Hoffbrand et al., 2005). During human ontogeny, erythropoiesis first appears in the embryonic yolk sac (primitive erythropoiesis), transfers to the fetal liver and then to the bone marrow before birth (definitive erythropoiesis; Palis and Segel, 1998; Palis, 2008). As the site of erythropoiesis changes, the globin genes being expressed change as well from embryonic to fetal and then adult. Production of haemoglobin is the hallmark of erythropoiesis and RBCs and is responsible for transferring oxygen to the tissues (Hoffbrand et al., 2005).

As mentioned in 1.3.1, HSCs differentiate to generate progenitors that are committed to specific blood lineages, which can also be defined by a combination of cell surface markers. The erythroid lineage is derived from the bipotent MEPs, which can further differentiate and commit to either the erythroid or megakaryocytic lineages. Downstream erythroid progenitors functionally correspond to the erythroid bursts (BFU-E) that have the ability to produce large colonies *in vitro* and the more differentiated late erythroid progenitors corresponding to colony forming units-erythroid (CFU-E; Hattangadi et al., 2011; Lodish et al., 2010; Wu et al., 1995). In Figure 1-7, surface markers that can be used to identify BFU-E/CFU-E and the morphologically distinct erythroid precursors are shown.

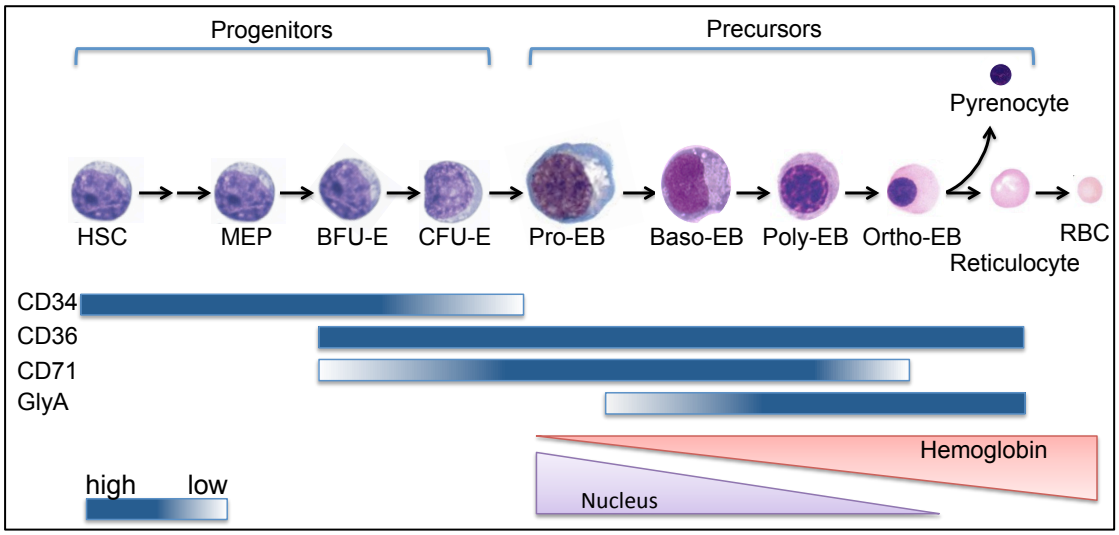


Figure 1-7 HSC commitment and terminal differentiation through the erythroid lineage. In the human bone marrow, the HSC self-renews and differentiates from the multipotent state, though common and erythroid-specific progenitors, to erythroid precursors and terminally matures to form the RBC. The stages are those of the megakaryocyte-erythroid progenitor (MEP), burst-forming units (BFU-E), colony-forming units (CFU-E), proerythroblasts (Pro-EB), basophilic erythroblasts (Baso-EB), polychormatic erythroblasts (Poly-EB) and orthochromatic erythroblasts (Ortho-EB). The latter gives rise to the reticulocyte and the pyrenocyte upon enucleation. The reticulocyte then enters the bloodstream and terminally matures to the RBC. The expression levels of cell surface markers used to distinguish the different cell types are shown. (Reproduced and adapted from (Dzierzak and Philipsen, 2013; Hoffbrand et al., 2005; Sinclair and Elliott, 2012))

Upon erythroid specification, the final phase of erythropoiesis leads to the maturation of committed erythroid progenitors in order to produce the fully differentiated RBCs. As erythroblasts mature, they decrease in size, undergo chromatin condensation, synthesise more haemoglobin and show altered gene expression patterns (Figure 1-7). The earliest recognisable erythroid precursor in the bone marrow is the proerythroblast (or pronormoblast), a large cell with a non-granular, deep-blue cytoplasm and a large nucleus occupying a large proportion of the cell. This cell is characterised by expression of CD71 (transferin receptor 1), but lacks CD235-GlycophorinA (GlyA or GPA; Constantinescu et al., 1999; Jelkmann and Metzen, 1996; Sinclair and Elliott, 2012) i.e. it is a $CD34^{low}CD71^{+}GlyA^{-}$ cell. After further divisions the proerythroblast differentiates into the basophilic normoblast, a similar cell to its ancestor, but with a more condensed heterochromatic nucleus ($CD34^{-}CD71^{+}GlyA^{low}$ cell). Further divisions form early polychromatic (expressing both CD71 and GlyA surface markers) and then late orthochromatic normoblasts

(these have lost expression of CD71), with increasing development of a condensed pink cytoplasm nucleus due to the production of haemoglobin by the cell. The late orthochromatic normoblasts are non-dividing cells with deeply stained nuclei(Lodish et al., 2010; Ronzoni et al., 2008).

During terminal differentiation, the nucleoli disappear, the nucleus further condenses and it is eventually extruded, through a procedure called enucleation(Ji et al., 2011; Migliaccio, 2010). The extruded nuclei (pyrenocytes) are phagocytosed and degraded by the macrophages of the bone marrow. The erythroid cells are now called reticulocytes, migrate outside the bone marrow, entering the bloodstream. Although lacking a nucleus, the mature reticulocytes do still possess mitochondria and ribosomes. These cells further mature to form the RBCs, which have no organelles and nucleus, but carry haemoglobin in very high concentrations. Over the 120 days of a RBC's life, the activities of the various enzymes decrease, contributing to the ageing process of the cell (Hattangadi et al., 2011).

Erythropoiesis takes place in specialised niches in the bone marrow, containing a macrophage, which is surrounding by maturing erythroid cells, a formation called erythroblastic island. The microenvironment in which erythroid cells lie plays a key role in red cell production, as it is a process tightly regulated by the expression of cytokines (Hattangadi et al., 2011; Sinclair and Elliott, 2012). Erythropoietin (EPO), a cytokine produced in the kidney in response to low oxygen pressure, is the principal regulator of erythropoiesis. Its binding to the EPO receptor (EpoR) allows signalling through a number of signalling pathways, including the signal transducer and activator of transcription 5 (Stat5), phosphoinositide-3 kinase/Akt, and Shc/Ras/mitogen-activated kinase (MAPK) pathways (Hattangadi et al., 2011). Erythropoiesis is EPO-dependent from the stages of CFU-E differentiation up to the formation of the orthochromatic normoblasts (England et al., 2011; Hattangadi et al., 2011; Koury, 2011). Another very important regulator of erythropoiesis and HSC expansion is stem cell factor (SCF) or c-kit, which is a cytokine that binds to the c-kit receptor and supports expansion of HSCs and production of erythroid progenitors up to the stage of the formation of proerythroblasts (Hattangadi et al., 2011; Koury, 2011). SCF binding to its receptor signals through the PI3 Kinase pathway (Huddleston, 2003).

1.3.3 Sodium butyrate in *in vitro* and *in vivo* erythropoiesis

HDACIs have pleiotropic functions in human erythropoiesis. HDACIs TSA and VPA were shown to block enucleation, without affecting differentiation and proliferation of erythroblasts (Ji et al., 2010; Migliaccio, 2010), whilst VPA is also known to promote the decision for erythroid/megakaryocytic differentiation (Zini et al., 2012). Knocking down of HDAC 1, 2 and 3 also inhibited terminal enucleation (Figure 1-8). In a recent study, combination of TSA and SAHA with cytokines (100ng/ml SCF, 100ng/ml fms-related tyrosine kinase 3 (Flt3) ligand, 100ng/ml Tpo and 50ng/ml IL-3) resulted in a considerably higher number of $CD36^{+}$ cells generated from $CD34^{+}$ HSCs than cytokines alone (Chaurasia et al., 2011), whereas in another study using primary G1ER cells SAHA permitted EPO-independent erythroid differentiation (Delehanty et al, 2012). On the other hand, Romidepsin inhibited the generation of $CD36^{+}GlyA^{high}$ mature erythroblasts from $CD34^{+}$ cells and induced apoptosis of both $CD36^{+}GlyA^{high}$ and $CD36^{+}GlyA^{low/-}$ i.e., immature erythroblasts (Yamamura et al., 2006).

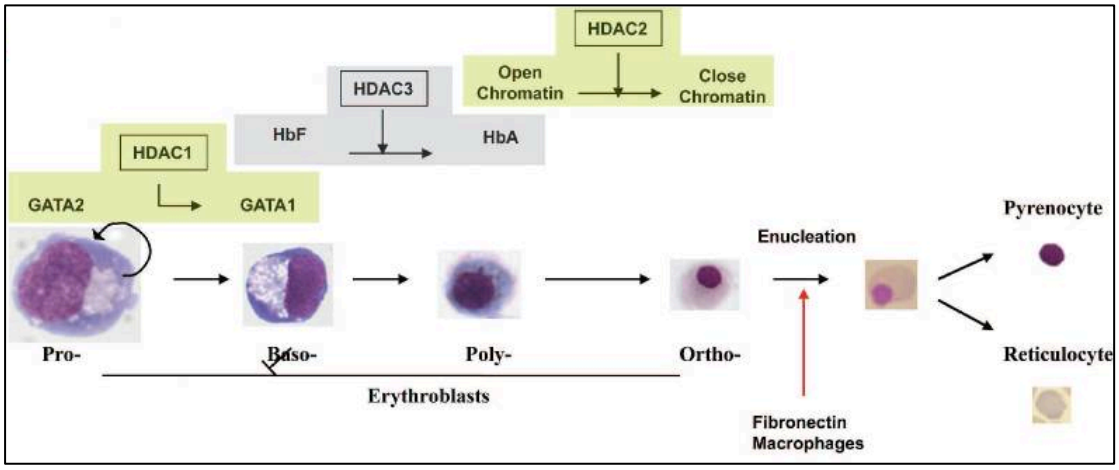


Figure 1-8 HDACs in erythropoiesis. HDAC isoforms involved in erythroid maturation are shown in this graph. HDAC1, HDAC3 and HDAC2 regulate the decision between self-replication and maturation, the switch from γ- to β- globin expression and chromatin condensation prior to enucleation, respectively. These HDACs are the ones affecting erythropoiesis upon treatment with HDAC inhibitors. (Obtained from (Migliaccio, 2010))

Unlike the other inhibitors that have been studied so far, NaBu has diverse effects on metabolism and cell morphology *in vitro*, inhibits cells proliferation and in some cell lines causes differentiation followed by apoptosis. Studies in the erythroid

J2E cell line, show that treatment with NaBu induces erythroid differentiation, coupled with increase of the haemoglobin expression, but blocks cell proliferation (Jaster et al., 1996). In human cervix tumour cell lines, NaBu causes cell death at concentrations higher than 0.5mM, whilst lower concentrations reduce proliferation without inducing apoptosis (Dyson et al., 1992).

Despite the *in vitro* studies showing that erythroid differentiation is inhibited by HDACIs, their use *in vivo* and in clinical practice suggests otherwise. *In vivo* studies in anaemic mice and primates (baboons) have shown that short-chain fatty acid derivatives, including butyrate, stimulate γ globin gene expression and erythropoiesis by increasing the BFU-E and reticulocyte counts (Cao et al., 2005; Pace, 2002). Patients with sickle cell anaemia treated with sodium-4-phenylbutyrate (plasma concentration 0.5mM-1mM) experience increased reticulocyte counts (therefore, no differentiation block) and improvement of their anaemia (Dover et al., 1994).

1.4 Disorders of the red cell metabolism

1.4.1 The glycolytic and pentose phosphate pathways

Metabolism is broadly defined as a set of more than 8,700 biochemical reactions that either produce or consume energy and are responsible for maintaining the living state of cells and consequently organisms. Metabolic pathways can be classified into three categories: those that synthesise or polymerise molecules to form more complex macromolecules (anabolism); those that degrade molecules to release energy (catabolism); and those that are involved in the elimination of toxic waste (toxic disposal). In the mid-19th century the majority of such pathways were identified, including glycolysis, respiration, urea cycle and oxidative phosphorylation (Deberardinis and Thompson, 2012; Metallo and Vander Heiden, 2013).

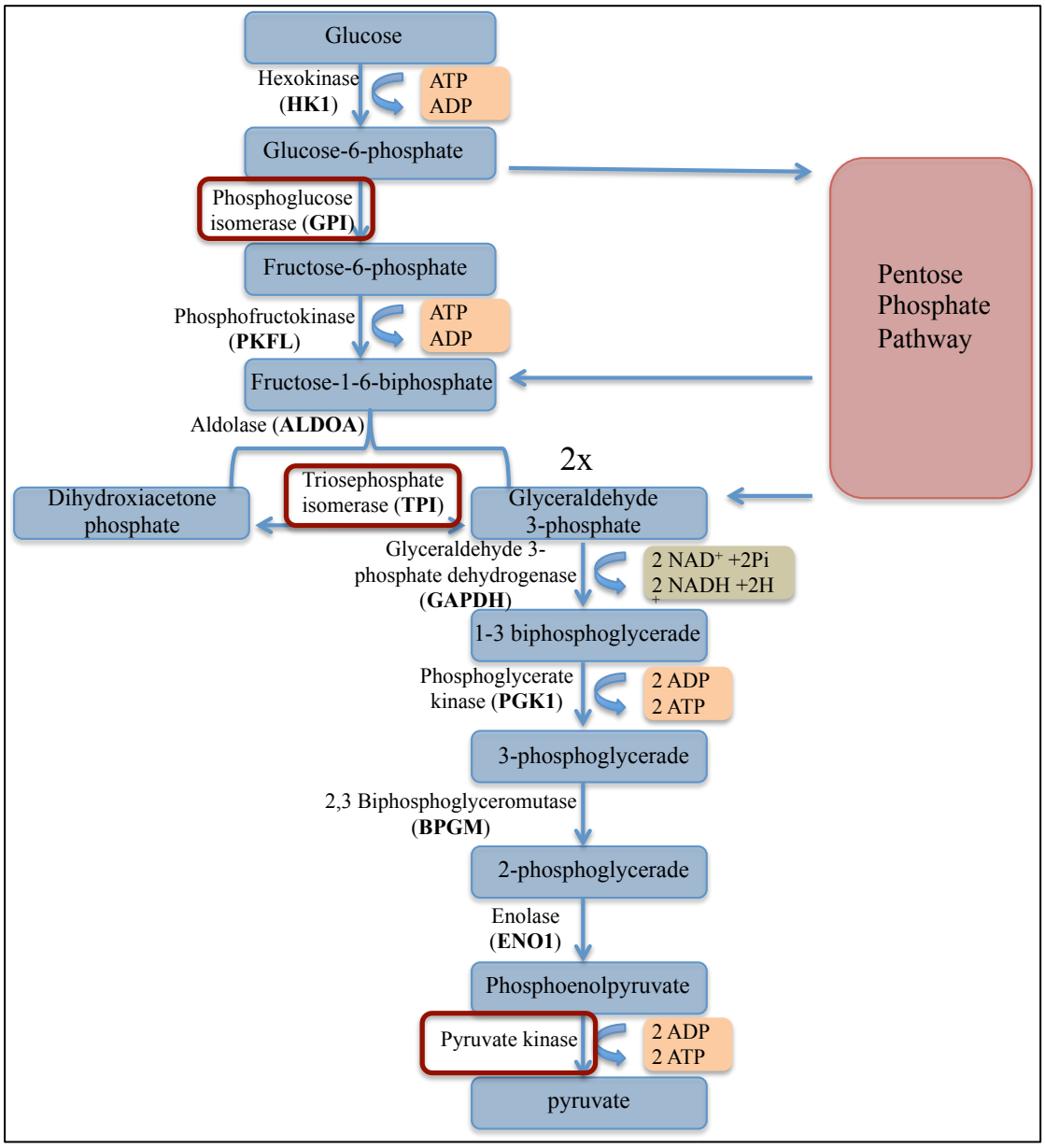


Figure 1-9 The glycolytic pathway. The ten steps of the Embden-Meyerhof pathway of glycolysis are shown. For each glucose molecule entering the glycolytic pathway two ATP and two NADH molecules are produced. The enzymes involved in the most common disorders of the red cell metabolism are highlighted in red circles. (Reproduced and adapted from (Berg et al., 2002; Varki et al., 2009))

The glycolytic (Embden-Meyerhof pathway; Figure 1-9) and its closely associated pentose phosphate (Figure 1-10) pathways (GPPP) are the ubiquitous metabolic processes responsible for the generation of adenosine triphosphate (ATP), the cellular free energy donor, especially under conditions in which the mitochondria, the main source of ATP production are either unable to function (e.g., in the absence of sufficient O_2 supplies) or absent (e.g., in mature RBCs; Berg et al., 2002; Varki et al., 2009). Energy is released once ATP is hydrolysed to adenosine diphosphate (ADP) and orthophosphate (Pi) or to adenosine monophosphate (AMP) and pyrophosphate (PPi). Similar to ATP that acts as an activated carrier of phosphoryl groups, electron carriers have a very important role in metabolic reactions, especially in anaerobic processes during which they assist in oxidation in the absence of oxygen. A major electron carrier for fuel oxidation is NADH, the reduced form of nicotinamide adenine dinucleotide (NAD^+), whereas NADPH, the reduced form of nicotinamide adenine dinucleotide phosphate ($NADP^+$) is an activated carrier for reductive biosynthesis. While NADPH is redundant in biosynthetic (mainly anabolic metabolism) processes, NADH is oxidised for the generation of ATP.

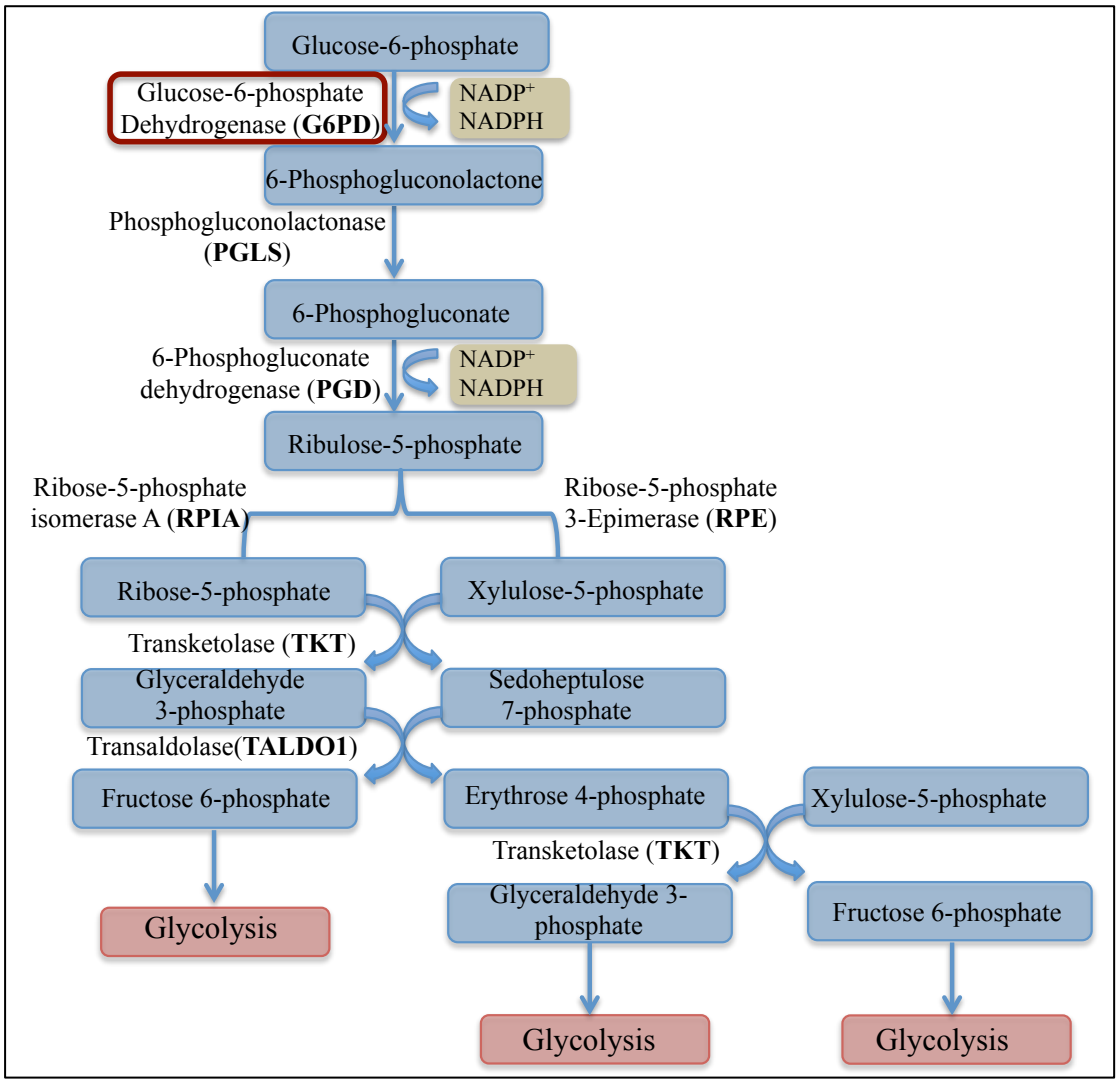


Figure 1-10 The pentose phosphate shunt. The pentose phosphate shunt is connected to the glycolytic pathway through the G6PD enzyme and results in the production of two NADPH molecules in total. G6PD, the enzyme causing the most common enzymatic defect is highlighted. (Reproduced and adapted from (Berg et al., 2002; Varki et al., 2009))

1.4.2 GPPP in erythroid cells

The main function of the RBCs is to carry haemoglobin in the bloodstream in high concentrations, facilitating gas exchange in the lungs and the tissue capillaries. Haemoglobin is an iron-containing oxygen-transport metalloprotein, which contains a heme group (consisting of an organic molecule having an iron in its structure) that binds oxygen in the lungs and in peripheral tissues and it exchanges it for carbon dioxide in order to bring it back to the lungs. For this purpose, the RBC needs a supply of energy and also a source of reducing power.

The mature RBC has lost DNA and RNA due to nuclear extrusion. Moreover, it is incapable of protein synthesis and all the mitochondria present in the reticulocytes have now been lost. Therefore, the only source of energy in the form of ATP, needed for both the maintenance of the RBC complex three-layer membrane and for the regulation of ion and water exchange, comes from either the GPPP (Figure 1-9 and Figure 1-10), or the glutathione cycle. Glycolysis takes place in the cytoplasm and generates ATP in anaerobic conditions through processing of glucose. Glycolysis is tightly connected with the pentose phosphate pathway through the enzyme glucose-6-phosphate dehydrogenase (G6PD; Figure 1-10; Castagnola et al., 2010). Under physiological conditions, the glycolytic pathway metabolises about 90% of glucose and only 10% enters the pentose phosphate shunt. However, in cases of oxidative stress the contribution of the pentose phosphate pathway is significantly increased. The same applies for the RBCs, which lack mitochondria.

For the RBC to ensure oxygen delivery, the cell has to maintain reducing power in the form of NADH over NAD to reduce methaemoglobin back to the functional state of deoxyhaemoglobin and to deal with the oxidative stresses that may occur during the circulation of the oxygen. Maintenance of higher levels of NADPH to NADP (produced by the pentose phosphate shunt) is also needed for the reduction of glutathione and protection against oxidative damage by the free oxygen molecules and hydrogen peroxide produced. The main role of the pentose phosphate shunt is to produce NADPH and the reduced form of glutathione, the major defence of the RBC against oxidative stress (Castagnola et al., 2010; Metallo and Vander Heiden, 2013). Moreover, optimal oxygen delivery is ensured by the high concentration of 2,3bis-phosphoglyceric acid produced by the Rapoport-Luebering shunt, an additional pathway closely related to glycolysis (Hoffbrand et al., 2005; van Wijk, 2005).

Mutations in most of the enzymes of the GPPP have been described and are generally associated with haemolytic anaemia. The most common disorders are caused by missense mutations on the coding regions of genes encoding GPPP enzymes.

1.4.3 G6PD deficiency: overview

G6PD deficiency was first described in the 1950s and is now known to be the most common human enzyme deficiency, affecting more than 400 million people worldwide (Beutler, 2007; Hoffbrand et al., 2005). The global distribution of the deficiency (Figure 1-11) is highly equated with that of *Plasmodium falciparum* and *Plasmodium vivax* malaria, leading to the so-called malaria protection hypothesis, as it is shown to protect against lethal malaria, for female heterozygotes and male hemizygotes (46% and 58%, respectively), particularly in childhood (Ruwende et al., 1995). The highest prevalence rates are found predominantly in Africa, the Middle East, Asia, the Mediterranean and South America. From the haematological point of view, G6PD is by far the most important enzyme in the GPPP. G6PD catalyses the first step of the pentose phosphate shunt in which glucose is converted into pentose sugars providing reducing power in the form of NADPH and is the one controlling the flux through this pathway. In addition, G6PD is essential for the regeneration of the reduced form of glutathione through the glutathione cycle (Cappellini and Fiorelli, 2008; Tsai et al., 1998).

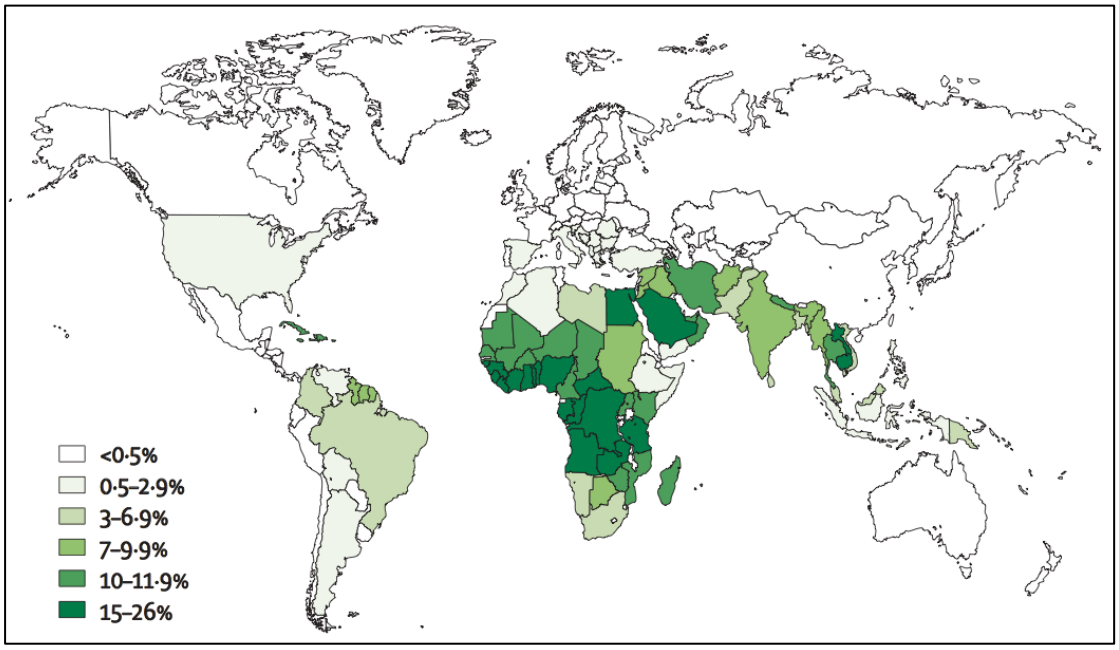


Figure 1-11 World map distribution of G6PD deficiency. Heat map showing the G6PD deficiency distribution, which is equated with that of malaria. (Obtained from (Cappellini and Fiorelli, 2008))

1.4.3.1 G6PD: Molecular Biology

G6PD is a housekeeping gene, spanning 15.9kb, expressed in all cell types at varying levels. *G6PD* gene is X-linked, residing at the telomeric region Xq28, has 13 exons and 12 introns and was cloned in 1986 (Martini et al., 1986; Persico et al., 1986). There are two characterised mRNA variants; the larger variant 1 that is inactive, as exon 1 is non-coding (Galgoczy et al., 2001) and the shorter variant 2, which encodes for exons 2-13. Two transcription start sites (TSS) drive the expression of these variants; one of them - located upstream of exon 1 - has been recently identified by the Encyclopedia of DNA elements (ENCODE) project and the other one - located within exon 1 - was bioinformatically characterised in the past (Philippe et al., 1994).

G6PD is evolutionary highly conserved and in humans is arranged in a 'head to head' configuration with *NEMO* (NFκB essential modulator; Figure 1-12), a gene that encodes a non-catalytic subunit of the cytokine-dependent IκB kinase, involved in the activation of the T Nuclear factor κ-light-chain enhancer of activated B cells (F NF-κB) (Jin and Jeang, 1999). Bioinformatics analysis has shown that four *NEMO* transcripts are transcribed due to alternative splicing of the 5' exons 1A, 1B and 1C under the influence of 2 promoters (promoter A and B). Promoter B (868bp; including a 192bp core promoter (Philippe et al., 1994)) is housekeeping and has strong bidirectional activity driving the transcription of both *G6PD* and *NEMO* genes.

The bidirectional *G6PD* gene promoter (Figure 1-12) is highly GC-rich (70%) and is embedded within a CpG island (1245bp; Fusco et al., 2006). It has been reported to contain two GC-boxes that drive its expression and also a non-canonical TATA-box (ATTAAAT) which is required for correct start of transcription but not for determining the level of gene expression (Ursini et al., 1990). However, a more recent bioinformatics analysis found that there is no TATA-box in the promoter (Galgoczy et al., 2001). Furthermore, 12 binding sites for the transcription factor Sp1 are predicted within a region of 1327bp surrounding exon 1 of the *G6PD* gene (Fusco et al., 2006; Galgoczy et al., 2001; Philippe et al., 1994). Two of these Sp1 sites (located ~100bp upstream of the TSS driving the exon 1 expression and overlapping with the two GC boxes) were previously experimentally validated and shown to be required for promoter activity (Franzè et al., 1998; Philippe et al., 1994).

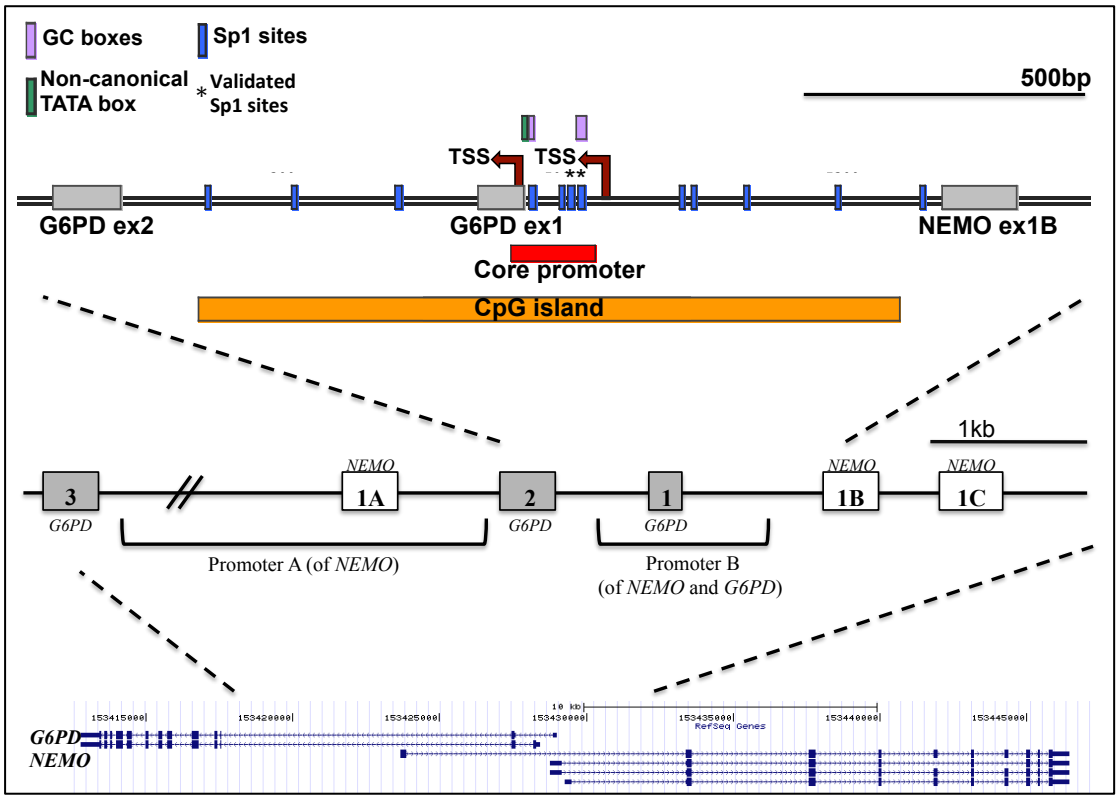


Figure 1-12 The *G6PD* gene and its promoter features. This is a schematic representation of the "head-to-head" configuration of *G6PD* and *NEMO* genes. The overlapping region of the genes corresponds to their shared bidirectional promoter B. The transcription start sites (TSS) of G6PD, promoter A (*NEMO* only) and B (*NEMO* and *G6PD*), the CpG island and Sp1 binding sites (the 2 validated ones are circled) are shown. (Reproduced from (Franzè et al., 1998; Galgoczy et al., 2001; Philippe et al., 1994))

G6PD consists of 515 amino acids and has a molecular weight of 59kDa. Inside the cell, the enzyme is active as a tetramer or dimer, in a pH-dependent equilibrium. Studies on the three-dimensional and the crystal structure of G6PD revealed a $NADP^+$- binding domain in each monomer and a second larger domain with the active catalytic domain between the two.

1.4.3.2 G6PD deficiency: clinical manifestations

To date, more than 180 mutations have been identified in the coding region of the *G6PD* gene (exons 2-13; Figure 1-13), suggesting genetic heterogeneity (Minucci et al., 2012). Approximately 85% of those are single nucleotide substitutions, producing missense variants; the rest of the mutations are either multiple mutations (8%), or small in-frame deletions (5%), or mutations that affect the introns (2%). All mutations reported so far affect only the coding region, as no association has been reported with the promoter region (Menounos et al., 2003). The wild type gene is referred to as *G6PD* type B.

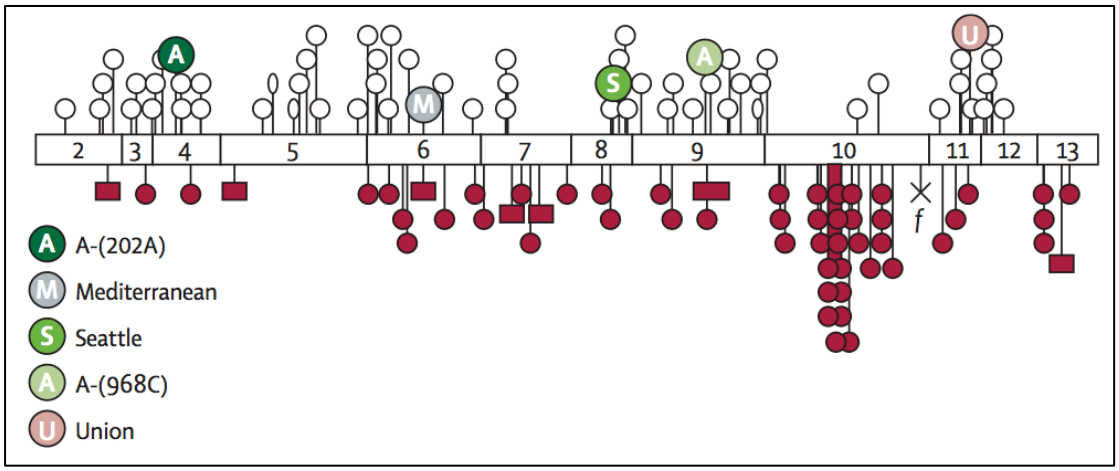


Figure 1-13 Most common characterised mutation sites of *G6PD*. The numbered boxes in the gene map correspond to the *G6PD* gene exons. The white circles refer to classes II and III and white ellipses are mutations causing class IV variants. Red circles refer to class I mutations, whereas red squares are small in-frame deletions causing class I deficiency. The "f" indicate a splice site mutation. (Obtained from (Cappellini and Fiorelli, 2008))

G6PD deficiency is a typical X-linked trait, usually affecting males who are hemizygous for the gene. On the other hand, females who have two copies of the gene due to the presence of two X chromosomes can be homozygous for the normal gene, heterozygous, or homozygous for the G6PD-deficient allele in areas in which the frequency of the deficiency is high. Due to X chromosome inactivation, heterozygous females are genetic mosaics and this is the reason why they have less severe phenotype than affected males.

The *G6PD* variants have been classified by the World Health Organisation (WHO) into five groups according to their activity relative to type B (Table 1-3; working group, 1989). In the unstressed normal erythrocyte, the G6PD activity is only about 2% of the total expression capacity, which can be greatly increased to meet challenges such as oxidative stress. However, the G6PD deficient erythrocyte has significantly reduced G6PD activity. Class I and II include patients with severe deficiency (below 10% of normal enzymatic activity), associated with chronic non-spherocytic haemolytic anaemia (CNSHA) and acute haemolytic anaemia, respectively. In class III there are the individuals with moderate deficiency (10-60% of normal) who usually present with neonatal jaundice, favism and drug-induced haemolytic anaemia. Examples of drugs that have been shown to cause drug-induced haemolytic anaemia are the anti-malarials primaquine and pamaquine, dapsone, which is used in leprosy and the analgesic acetanilide. Lastly, in class IV are the normal individuals and in class V those with increased activity (working group, 1989). The majority of class I mutations affect exons 6, 10 and 13, encoding the parts of the enzyme that bind the substrate, dimer interface and $NADP^+$ site, respectively (Minucci et al., 2012).

Table 1-3 World Health Organisation (WHO) classification for G6PD deficiency. (Adapted from (working group, 1989))

Class	Enzymatic activity	Examples	Clinical features
I	Severe (1–10% residual activity)	Harilaou Serres Brighton	CNSHA
II	Severe (1–10% residual activity)	Mediterranean Canton Orissa	Favism, Drug-induced acute haemolysis, Neonatal jaundice
III	Moderate (10–60% residual activity)	A- Mediterranean	Favism, Drug-induced acute haemolysis
IV	Normal activity (60–150%)	B (wild type)	None
V	Increased activity (>150%)	B (wild type)	None

The most common variant of G6PD deficiency is G6PD A- (African), which results in class III deficiency caused by mutations at 202A and 968C positions on the exons 4 and 9, respectively (Table 1-3 and Figure 1-13). This variant accounts for approximately 90% of G6PD deficiency in Africa and is also very frequent in America, in the West Indies and other areas where people of African origin are present. G6PD A- is also present, albeit less frequent, in Italy, Spain, Portugal, Egypt, the Middle East and Lebanon. The second most common variant is G6PD Mediterranean (G6PD Med), which is frequent in all the countries surrounding the Mediterranean Sea, such as Spain, Greece, Italy, Egypt, although it is not frequent in the Middle East, including Israel (Bayoumi et al., 1996). G6PD Med is caused by a 563T mutation on exon 6 (Figure 1-13) and belongs to class II or III depending on the severity of the deficiency (Cappellini and Fiorelli, 2008). Favism, one of the clinical features associated with G6PD deficiency, caused by the consumption of fava beans, is now believed to be most frequently connected to G6PD Med, amongst all variants. However, both G6PD A- and G6PD Med variants have been associated with a very significant reduction in the risk of severe malaria caused by *Plasmodium falciparum* and *Plasmodium vivax*, as discussed in 1.4.3.

The diagnosis of G6PD deficiency is based on the estimation of enzyme activity, by the quantification of the rate of NADPH production from $NADP^+$ by spectrophotometric analysis. Furthermore, the development of PCR, direct sequencing and denaturating gradient electrophoresis have allowed the detection of specific common mutations e.g. the Mediterranean type, population screening and in some severe cases prenatal diagnosis (Mason, 1996).

1.4.4 Other common deficiencies of the GPPP

Enzymatic disorders associated with point mutations on the exons of genes encoding enzymes of the GPPP have been characterised for almost all the GPPP enzymes (Table 1-4), with the exception of 6-phosphogluconate dehydrogenase (PGD), ribose-5-phosphate-3- epimerase (RPE) and transketolase (TKT).

Table 1-4 Enzyme disorders of the GPPP.

Enzyme	Pathway	Genetics	Haematologic phenotype	References
Hexokinase (**HK1**)	Glycolytic	10q11.2 autosomal recessive	CNSHA	(de Vooght et al., 2009)
Glucose phosphate isomerase (**GPI**)	Glycolytic	19q13.1 autosomal recessive	CNSHA	(Climent et al., 2009; Kugler and Lakomek, 2000; Repiso et al., 2006)
Phosphofructokinase, liver (**PFKL**)	Glycolytic	21q.22.3 autosomal recessive	Erythrocytosis, Minimal haemolysis	(García et al., 2009; Ronquist et al., 2001)
Fructose diphosphate aldolase (**ALDOA**)	Glycolytic	16p11.2 autosomal recessive	CNSHA	(Esposito et al., 2004; Kish et al., 1987; Yao, 2004)
Triose phosphate isomerase (**TPI**)	Glycolytic	12p13 autosomal recessive	CNSHA, Susceptibility to infections	(Oláh et al., 2005; Orosz et al., 2009; Ralser et al., 2006)
Phosphoglycerate kinase (**PGK1**)	Glycolytic	Xq13.3 X-linked	CNSHA	(Pey et al., 2013; Svaasand et al., 2007)
Biphosphoglycero mutase (**BPGM**)	Glycolytic	7q33 autosomal recessive	Erythrocytosis	(Hoyer et al., 2004; Lemarchandel et al., 1992)
Glyceraldehyde-3-P-dehydrogenase (**GAPDH**)	Glycolytic	12p13 autosomal dominant	None reported	(Phadke et al., 2009; Pretsch and Favor, 2007)
Enolase (**ENO1**)	Glycolytic	1p36.2 autosomal dominant	CNSHA	(Stefanini, 1972)
Pyruvate kinase, liver/ RBC (**PK**)	Glycolytic	1q21 autosomal recessive	CNSHA	(Zanella et al., 2007a; 2007b)
Glucose-6-phosphate dehydrogenase (**G6PD**)	Pentose phosphate	Xq28 X-linked	CNSHA, Drug-induced HA, Favism	(Cappellini and Fiorelli, 2008; Minucci et al., 2012; Verhoeven et al., 2001; vulliamy et al., 1998; Wamelink et al., 2010)
Phosphoglucolactonase (**PGLS**)	Pentose phosphate	19p13.2 autosomal dominant	CNSHA	(Beutler et al., 1985)
Ribose-5-phosphate isomerase A (**RPIA**)	Pentose phosphate	2p11.2 autosomal recessive	None reported (single patient with encephalopathy)	(Balasubramanyam, 2004; Valayannopoulos et al., 2006)
Transaldolase (**TALDO1**)	Pentose phosphate	11p15.5 autosomal recessive	Infantile liver failure	(Verhoeven et al., 2001; Wamelink et al., 2010)

Another common disorder of the red cell metabolism and the most common of the glycolytic pathway disorders is the pyruvate kinase (PK) deficiency, which is also associated with CNSHA and inherited as an autosomal recessive trait. Two housekeeping genes, one on chromosome 15 and the other one on chromosome 1 have been shown to cause PK deficiency. These are responsible for the expression of four isoenzymes due to alternative splicing. The *PK* gene on chromosome 15 encodes for PKM1, present in skeletal muscles and PKM2, present in leukocytes. The second *PK* gene on chromosome 1 gives rise to the PKL in the liver and the PKR in the RBCs isoforms. Mutations on these genes indicate genetic heterogeneity as in G6PD deficiency; therefore the clinical features vary from severe anaemia and jaundice, severe CNSHA, moderate haemolysis to asymptomatic haemolysis or anaemia (Zanella et al., 2007a; 2007b).

Compared to PK deficiency, the other disorders of the glycolytic pathway are very rare. Deficiency of glucose phosphate isomerase (GPI) is one of the commonest causes of CNSHA after G6PD and PK deficiencies. More than 20 mutations have been identified and affect protein stability (Repiso et al., 2006). Additionally, triose phosphate isomerase (TPI) is an enzyme that in its mutated form causes not only CNSHA but also neuromuscular defects and can even cause sudden death. These two genes are located on the chromosomes 19 and 12, respectively and are inherited as autosomal recessive genes. It is of great interest that these genes are expressed in a housekeeping manner and all of them contain several Sp1 binding sites on their promoters (Orosz et al., 2009).

Children with inherited glycolytic enzyme deficiencies often suffer from life-long, transfusion-dependent anaemia and those with TPI or GPI deficiency may have progressive, severely debilitating neuromuscular and central nervous system (CNS) complications, which begin in the first few years of life (Orosz et al., 2009; Repiso et al., 2006). Whilst anaemia can be treated with blood transfusion, this causes tissue iron overload and major end-organ damage particularly of the heart and liver, and no specific treatment is available for the neuromuscular or CNS complications. There is therefore, a pressing need for effective treatment, which can be initiated before the onset of the severe complications in order to prevent the severe disability associated with glycolytic enzyme deficiencies.

1.5 Hypothesis

According to the lessons learned from *PIGM* and IGD (as described in 1.1.4.1) and given that *PIGM* is a housekeeping gene and part of a ubiquitous biosynthetic pathway, I hypothesised that other genes required for GPI biosynthesis might be subjected to the same mechanism of Sp1-dependent transcriptional control. This notion was extended; therefore, I further hypothesised that Sp1-dependent control of histone acetylation and transcriptional activation may also apply for genes within other enzymatic biosynthetic pathways regulated in a housekeeping manner. If this is correct, characterisation of the Sp1-dependent epigenetic control of other genes with such characteristics could offer innovative therapeutic opportunities for inherited disorders of ubiquitous biosynthetic pathways such as that of the GPPP.

My hypothesis further postulates that if increased transcription is achieved by HDAC inhibition, then this may result in increased production of the mutant protein, which will lead to increased enzymatic activity. Once the enzymatic activity exceeds a certain threshold, it could ameliorate the cellular and clinical phenotype. For example, research in our lab has shown in the past that in IGD an increase of the *PIGM* mRNA in the deficient B cell lines from 1% to as little as 5-10% of normal is associated with complete restoration of GPI expression on the surface of the cell (unpublished data). It should be noted that for most enzymes of the glycolytic pathway, red cell activity levels as low as 10-20% of normal are sufficient for abrogation of haemolysis. Hence, even a small increase in the overall activity after HDAC inhibition might have significant therapeutic value.

1.6 Aims

The aims of this PhD project are to:

I. Test *in vitro* and *in vivo* the ability of HDACIs to enhance the gene expression and the enzymatic activity of GPPP genes and proteins, respectively. This will be tested in B cell lines and primary erythroid cells derived from patients with inherited glycolytic enzyme deficiencies and healthy donors.

II. Explore the epigenetic events at baseline and in response to HDACIs to identify the epigenetic mechanism that drives gene expression.

III. Explore the genome-wide implications of HDAC inhibition in primary erythroid cells and potentially identify novel targets of HDACIs outside the GPPP that could be of therapeutic value.

2 Materials and Methods

2.1 Cell lines, primary cells and patient samples

2.1.1 Cell lines

2.1.1.1 B lymphoblastoid and other cell lines

Epstein Barr Virus (EBV)-transformed B lymphoblastoid cell lines, derived from patients with glycolytic enzyme disorders were used for the purposes of this project. To control these experiments, B lymphoblastoid cell lines derived from normal individuals not affected by glycolytic enzyme deficiencies were also used. Additionally, a B lymphoblastoid cell line derived from a patient with IGD was utilised as a control of NaBu treatment.

Table 2-1 Cell lines used in the study.

Cell line	Type	Culture features	Individual Genotype	Individual sex
PP0007 (or P7)	B cell line	Suspension	Class I deficiency (G6PD Brighton - in frame exon 13 deletion)	Male
PP0065 (or P65)	B cell line	Suspension	Homozygous for TPI mutation	Male
PP0054 (or P54)	B cell line	Suspension	Homozygous for TPI mutation	Male
PP0112 (or P112)	B cell line	Suspension	Homozygous for TPI mutation	Male
PP0267 (or P267)	B cell line	Suspension	Homozygous for TPI mutation	Female
PP0091 (or P91)	B cell line	Suspension	TPI mutation carrier	Female
PP0092 (or P92)	B cell line	Suspension	TPI mutation carrier	Male
PP0277 (or P277)	B cell line	Suspension	Normal B cell line	Male
PP0015 (or P15)	B cell line	Suspension	Normal B cell line	Female
YK	B cell line	Suspension	Homozygous for PIGM mutation	Male
K562	Erythroid progenitors	Suspension	Erythroleukaemia (Chronic Myeloid Leukaemia patient in blast crisis)	Female
293T	Embryonic Kidney	Adherent	Normal Embryonic Kidney	Female

The human erythroleukaemia cell line K562 was used to determine the glycolytic enzyme gene expression in an erythroid environment. Furthermore, Human Embryonic Kidney 293 (293T) cells were utilised to transfect with the plasmid constructs. Information regarding the characteristics of all the cell lines used is listed in Table 2-1.

2.1.1.2 Cell line culture

The suspension cell lines, as listed in Table 2-1, were grown in Roswell Park Memorial Institute medium (RPMI) 1640 medium (Sigma-Aldrich Company Ltd., Dorset, UK) supplemented with 10% Fetal Bovine Serum (FBS) (Sigma-Aldrich Company Ltd., Dorset, UK), 2mM L-Glutamine and 10ml/L Penicillin-Streptomycin (Stem Cell Technologies, Vancouver, Canada). Culturing took place in a moist incubator with 5% CO_2 at 37°C. While in culture, the cells were maintained at a concentration of $4x10^5$-$5x10^5$cells/ml in order to be in the exponential phase of their growth and were fed every three days.

Adherent 293T cells were used for efficient plasmid construct transfections. They were grown in 75cm^2 culture flasks in Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich Company Ltd., Dorset, UK), supplemented with 10% Fetal Bovine Serum (FBS; Sigma-Aldrich Company Ltd., Dorset, UK), 2mM L-Glutamine and 10ml/L Penicillin-Streptomycin (Stem Cell Technologies, Vancouver, Canada). The cells were maintained at 75% confluency by splitting every three to four days, depending on confluency. 293T cells were detached by replacing DMEM with phosphate buffered saline (PBS) for an initial wash to remove the excess of cations and trypsin inhibitory proteins in the medium. This was followed by incubation of the cells with 1x trypsin- Ethylenediaminetetraacetic acid (EDTA) solution (Sigma-Aldrich Company Ltd., Dorset, UK) at 37°C for 10 minutes to detach the cells. Cells were then re-plated at the desired concentration in full DMEM medium.

The trypan blue (Sigma-Aldrich Company Ltd., Dorset, UK) exclusion method was used to count live cells on a Neubaeur haemocytometer to determine concentrations. For counting, three squares (1mm^2) were used to count cells and the average was then calculated for the final concentration. Dead cells, stained blue, were excluded from counting.

2.1.2 Primary cell sources

The National Research Ethics Service approved this study (record reference 1/LO/1050), which allowed for sample collection from healthy and deficient individuals. Site-specific approval was obtained from the consultants taking the samples.

For this study, whole blood was obtained from healthy and deficient adult individuals in order to isolate peripheral blood mononuclear cells (PBMCs). Table 2-2 shows the genotype, phenotype and enzymatic activity in the blood (as measured in Hammersmith Hospital at the time of collection of the blood samples) of the recruited patients. Furthermore, cord blood and GCSF-mobilised peripheral blood were also obtained for the isolation of mononuclear cells, which was then followed by the magnetic isolation of $CD34^+$ stem and progenitor cells. Cord blood was provided by Queen Charlotte's Hospital and GCSF-mobilised peripheral blood was provided by the John Goldman Stem Cell lab at Hammersmith Hospital.

Table 2-2 G6PD deficient patients.

Patient	Genotype	Clinical Phenotype	Enzymatic activity in blood (our measurement)	Hb (g/dl)	Retics (%)	Reference
G6PD Brighton	Exon 13; 1488-1490 deletion of GAA; Lys	Class I; CNSHA	1.1 U/gr Hb	12.4 (8.1 pre-splenecto-my)	9.7	(McGonigle et al., 1998; vulliamy et al., 1988)
G6PD Serres	Exon 10; 1082C->T; Ala->Val	Class I; CNSHA	1.7 U/gr Hb	10.5 (6.6 pre-)	16.6	(Vulliamy et al., 1998)
G6PD Harilaou	Exon 7; 648T->G; Phe->Leu	Class I; CNSHA	0.5 U/gr Hb	9 (8.2 pre-)	17.1	(Poggi et al., 1990)
G6PD Mediterranean (Med)	Exon6; 563C->T; Ser->Phe	Class II/III; Favism, Drug-induced acute haemolysis	3.7 U/gr Hb	12	2	(Vulliamy et al., 1988)
G6PD African (A- (202A))	Exon 4; 202G->A; Val->Met	Class III; Favism, Drug-induced acute haemolysis	4.0 U/gr Hb	13.8		(Cappelini et al., 1996)
Normal adult			7-10.4 U/gr Hb	13-16.8		

2.1.3 Cryopreservation and thawing of cells

When necessary, cells were cryopreserved in freezing medium that is FBS supplemented with 10% dimethyl sulfoxide (DMSO; Sigma-Aldrich Company Ltd., Dorset, UK) at a maximum concentration of 10^7 cells per vial in 1ml of freezing medium. Immediately after resuspension in the freezing medium, they were frozen in -80°C in an isopropanol-containing Nalgene Cell Freezing Container (Sigma-Aldrich Company Ltd., Dorset, UK) for up to two days before long-term storage in liquid nitrogen.

In order to thaw the cryopreserved cells, they were incubated for 1 minute in a 37°C waterbath. They were then resuspended in appropriate warm medium.

2.2 Primary cell selection

2.2.1 Preparation of mononuclear cells

Mononuclear cells were isolated by gradient density centrifugation from peripheral blood or cord blood using Lymphoprep™ (Axis-Shield, Oslo, Norway). Blood samples were first diluted 1:2 in PBS (Gibco, Invitrogen, Paisley, UK) and then layered above the density gradient medium in Falcon tubes followed by centrifugation at 1500rpm for 30 minutes (break off). Finally, mononuclear cells were recovered from the plasma-lymphoprep interface, washed with PBS, counted and then either immediately placed in culture or frozen in liquid nitrogen (as described in 2.1.3).

2.2.2 Magnetic cell selection

2.2.2.1 $CD34^+$ cell selection

$CD34^+$ cell selection was carried out using the human CD34 MicroBead kit (Miltenyi Biotec Ltd., Surrey, UK) as per manufacturer's instructions. Briefly, up to 10^8 cells are resuspended in 300µl Robosep buffer and incubated with 100µl CD34 Microbeads in the presence of 100µl FcR blocking reagent. Following a 30min incubation at 4°C in the dark, they were washed in Robosep buffer, resuspended in 500µl of Robosep and then selected through a primed column attached to a magnet. The cells that were going to be selected were held in the column by the magnet and were washed three times with Robosep buffer before. To isolate them, the column

was detached from the magnet, 1ml of Robosep was added in the column and the cells were collected with the help of a plunger.

2.2.2.2 $CD36^+$ cell selection

CD36+ cell separation was performed using the MACS® technology (Miltenyi Biotec Ltd., Surrey, UK). Positive magnetic selection using anti-PE or anti-Allophycocyanin (APC) beads (Miltenyi Biotec Ltd., Surrey, UK) was used. The protocol followed was as per manufacturer's instructions. In brief, up to 10^7 cells were resuspended in 100µl Robosep buffer (Miltenyi Biotec Ltd., Surrey, UK) and incubated with 10µl antibody for 10min at 4°C in the dark to stain. After washing with PBS, the cells were resuspended in 80µl Robosep buffer and 20µl of the beads against the specific antibody used, were added (Table 2-3). After incubating for 15min at 4°C in the dark and washing with PBS, the CD36+ cells were separated using magnetic columns (Miltenyi Biotec Ltd., Surrey, UK) that are under the control of the MACS separator's magnetic field.

Table 2-3 Magnetic $CD36^+$ cell selection antibodies and beads.

Antibody	Beads
anti-CD36-PE (eBioscience Ltd., Hatfield, UK)	anti-PE microbeads (Miltenyi Biotec Ltd., Surrey, UK)
anti-CD36-APC (eBioscience Ltd., Hatfield, UK)	anti-APC microbeads (Miltenyi Biotec Ltd., Surrey, UK)

2.3 In vitro erythroid differentiation system

The erythroid differentiation system used was adapted from previous publications (Ohene-Abuakwa, 2005; Ronzoni et al., 2008). It is a two-phase liquid culture system that models human adult erythropoiesis. Each phase lasts for seven days. At the end of Phase 1, $CD36^+$ erythroid cell were selected as described in 2.2.2.2 and replated in Phase 2 medium. Various sources were used for erythroid cell differentiation, including PBMCs and $CD34^+$ cells isolated from cord blood (CB) and peripheral blood (PB) after Granulocyte-colony stimulating factor (G-CSF) mobilisation.

2.3.1 Erythroid differentiation medium and cytokines

Erythroid differentiation was performed using the serum-free stem cell maintenance and differentiation medium, StemSpan® SFEM (Stem Cell

Technologies, Vancouver, Canada), supplemented with a cytokine cocktail and 10ml/L Penicillin-Streptomycin (Stem Cell Technologies, Vancouver, Canada). The cytokine cocktail used consisted of 10ng/ml IL-3, 100ng/ml SCF and EPO. The optimised EPO concentration was 0.5U/ml during Phase 1 of the culture and then 4U/ml during Phase 2 to achieve later stages of differentiation in this highly EPO-dependent stage. EPO was supplied from R&D (R&D Systems Europe, Abingdon, UK) and all other cytokines from Peprotech (Peprotech Inc, Rocky Hill, NJ, USA). The erythroid medium was prepared fresh before use.

2.3.2 Plating and culturing

Before plating the cells in the freshly made erythroid medium, they were counted using the trypan blue exclusion method. After optimisation, the PBMCs were plated (day 0) at a concentration of $2x10^6$ cells/ml in 48-well plates for 7 days (Phase 1 of culture). Every 2-3 days of culture half of the erythroid medium was replaced with fresh one. The PBMCs are floating cells, which tend to settle at the bottom of the well, facilitating the process of removing medium from the top. Usually, on day 2 non-adherent cells were washed and split 1:2 in fresh medium. This process allowed removal of monocytes, which adhere and tend to be toxic for the culture. Additionally, on day 7 of the erythroid culture, the $CD36^+$ cells were selected as described in 2.2.2.2 and placed in culture at a concentration of $5x10^5$ cells/ml for another 7 days (Phase 2 of culture).

For the erythroid differentiation of $CD34^+$ cells, those were plated at a concentration of $5x10^5$ cells/ml at day 0. The rest of the differentiation proceeded as for the PBMCs except when the cultures reached 80% or more of differentiation at day 7 (assessed by flow cytometric analysis), in which case the cells were re-plated in Phase 2 medium without the $CD36^+$ selection.

2.4 Drug treatments

The HDACIs used in this study were purchased from Sigma-Aldrich Company Ltd., Dorset, UK. They were all received in powder form and resuspended according to manufacturer's advice. In particular, NaBu (98% purity; ref. 303410) was reconstituted fresh before use in phosphate buffered saline (PBS; Sigma-Aldrich Company Ltd., Dorset, UK) at a 600mM concentration and kept at -20°C for a maximum of a week to avoid loss of activity after long storage. SAHA (>98% purity;

ref SML0061) was reconstituted in DMSO at a 15mg/ml concentration and subsequently was diluted in PBS at a 600μM concentration and stored at -20°C. TSA (>98% purity; ref T8552) was initially reconstituted in ethanol at a 2mg/ml concentration, which was then diluted in PBS and stored at -20°C at a 10μM stock concentration. Table 2-4 shows the final concentrations the HDACIs are used in the cells.

Table 2-4 Concentration of HDACIs used in cells.

HDACI	Concenantration used in cells
NaBu	3mM in cell lines 1mM in primary cells
SAHA	4μM in cell lines 1μM in primary cells
TSA	100nM in cell lines 1nM in primary cells

Other than the HDACIs, chemical treatments using cycloheximide (CHX; ref C4859; Sigma-Aldrich Company Ltd., Dorset, UK) and mithramycin A (MTA; ref M6891; Sigma-Aldrich Company Ltd., Dorset, UK) were employed. Both chemicals were reconstituted in DMSO and further diluted in PBS. Final concentrations used in the cells were 10μg/ml and 1mM for CHX and MTA, respectively.

2.5 Flow cytometry

2.5.1 Staining

Expression of extracellular markers was performed by flow cytometric analysis. Samples were analysed using a four-laser BD LSRFortessa™ cell analyser (Beckton-Dickinson, Oxford, UK). Each sample was stained with antigen-specific monoclonal antibodies and an unstained control of the cells was used to determine the negative expression of each cell surface marker. Compensation settings were set using single-stained samples. To stain the cells, they were resuspended in 100μl of PBS and upon adding 2μl of FcR blocking reagent (Miltenyi Biotec Ltd., Surrey, UK) they were incubated for 5min at room temperature. Without washing, the antibody cocktail

(Table 2-5) was added and staining took place for 20min at 4°C in the dark. After washing, resuspending in PBS and adding 1µl of 4',6-diamidino-2-phenylindole (DAPI), the cells were loaded on the flow cytometer for analysis.

Table 2-5 Anti-human antibodies used for flow cytometric analysis.

Antibody specificity	Conjugate	Supplier	µl per test
CD34	PerCP-Cy5.5	BioLegend, San Diego, USA	0.5µl
CD36	PE	BD Biosciences, Oxford, UK	1µl
CD36	APC	BD Biosciences, Oxford, UK	2µl
CD71	PE	eBioscience Ltd., Hatfield, UK	1µl
GlyA	eFluor450 (Pacific Blue channel)	eBioscience Ltd., Hatfield, UK	1µl
CD33	PE-Cy7	eBioscience Ltd., Hatfield, UK	1µl
CD61	FITC	BD Biosciences, Oxford, UK	1µl
FLAER	Alexa® 488 (FITC channel)	Pinewood Scientific Services Inc., Victoria, Canada	2.5µl

2.5.2 Annexin V apoptosis assay

To estimate the apoptosis rates of the cells treated by NaBu, the AnnexinV protocol of eBiosciences was followed as per manufacturer instructions (eBioscience Ltd., Hatfield, UK). In brief, the cells were washed with PBS and then washed again in 1x binding buffer. These were then spun down and resuspended in 1x binding buffer at a concentration of 1-5 x 10^6 cells/ml. Only 100µl of these were used and stained with 1µl of AnnexinV –APC (eBioscience Ltd., Hatfield, UK) for 10min at room temperature in the dark. Next, the cells were washed in 1x binding buffer and loaded on the flow cytometer.

2.5.3 Data analysis

Data was analysed (Figure 2-1) using the FlowJo software version 9.3.2 (Tree Star Inc., Ashland, USA). Initial gating was set on live cells based on the forward (FSC-A) and side scatter (SSC-A) parameters. Following that, single cells were

selected and then DAPI$^+$ dead cells were excluded from the analysis. This method was initially performed in an unstained cell sample, to set the negative expression for every cell marker. Then, these settings were applied to the stained samples for analysis.

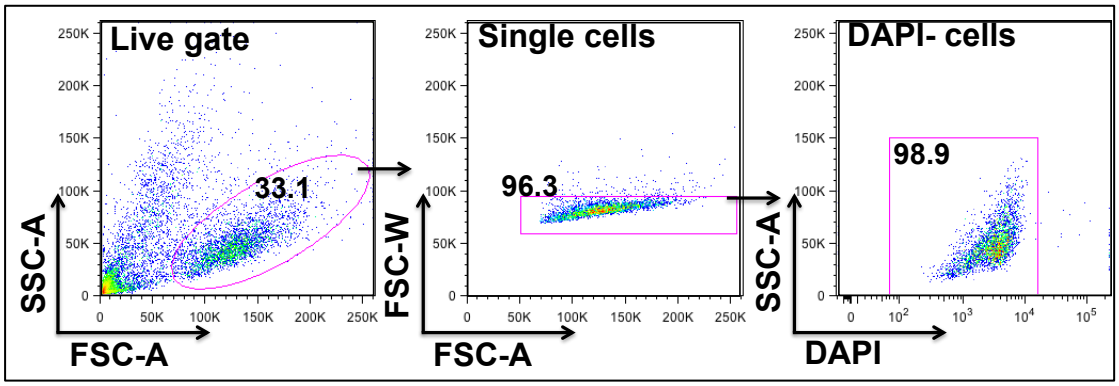


Figure 2-1 Flow cytometric analysis strategy. Live cells are selected on the forward (FSC-A) and side scatter (SSC-A). Gating on the forward scatters FSC-W and FSC-A allows exclusion of the duplets and out of those, the DAPI$^+$ cells are excluded from the analysis too. The remaining cells are used for analysis of cell surface markers.

2.6 Cytospins and histologic staining

To determine the morphological characteristics of primary cells, cytospin slides were prepared, stained and observed under an optical microscope. Briefly, 15.000 cells were resuspended in 200µl of full medium and then loaded into the cytospin cartridges. These were spun down for 5min at 400rpm in a Shandon Cytospin2 centrifuge (Block Scientific Inc, NY, USA). SuperFrost Ultra Plus (VWR International Ltd, Leicestershire, UK) slides were used. Upon centrifugation, the slides were removed from the cartridges and air-dried. Then they were fixed in methanol for one minute and stained in the May-Grunwald (Merck&Co, Hertfordshire, UK) solution for 7 min, followed by an additional staining in the Giemsa (Merck&Co, Hertfordshire, UK) solution for 20min. After May-Grunwald and Giemsa (MGG) staining, the slides were washed in distilled water and after air-drying they were mounted with PTX mountant (Sigma-Aldrich Company Ltd., Dorset, UK) and covered with a 50x25mm coverslip (VWR International Ltd, Leicestershire, UK).

Staining solutions were optimally prepared 30min before use. Preparation involves filtering the dyes through a 0.2μm bore needle gauge and then dilution 1:2 of May-Grunwald and 1:10 of Giemsa in distilled water.

The cytospin slides were visualised under a Nikon eclipse E400 microscope (Nikon, UK) and photos were taken using 40x and 100x magnification lenses.

2.7 Gene expression assays

2.7.1 RT-qPCR assays

2.7.1.1 RNA extraction and cDNA synthesis

Harvested cells, either primary or cell lines, were washed with PBS, pelleted down and then usually stored at -80°C prior to RNA extraction. When ready for extraction, the pellets were thawed and extraction was carried out using the GeneJET™ RNA Purification kit (Fermentas, York, UK) to process large samples. To process small samples, usually coming from primary cells (less than 10^5 cells), the RNeasy Plus Micro kit (Qiagen, West Sussex, UK) was used instead. When using the GeneJET™ RNA Purification kit, a DNase treatment step was added to the protocol, to eliminate any remaining DNA, as per manufacturer's advise (Qiagen, West Sussex, UK). Upon RNA extraction, spectroscopic photometry using a NanoDrop machine (Thermo Scientific, USA) was carried out to estimate RNA concentration. Smaller samples went straight to cDNA synthesis. All RNA was stored at -80°C.

Directly after RNA extraction, reverse transcription of up to 1μg of each sample was carried out in order to produce functional cDNA. The RevertAid™ Reverse Transcriptase kit (Fermentas, York, UK) was used following manufacturer's instructions and using $Oligo(dT)_{18}$ primers. cDNA synthesis samples were placed at -20°C for long term storage.

2.7.1.2 Primer design and testing

Primers (Table 2-6) were purchased from Sigma-Aldrich Company Ltd. (Dorset, UK). They were designed using the Primer Express Software (Applied Biosystems, UK) and tested for the dimers and hairpins formation using the OligoCalc software (http://www.basic.northwestern.edu/biotools/OligoCalc.html). Primers were designed to span two exons, amplify sequences shared by all splicing

variants, have annealing temperature of 60°C and have GC content 40-60%. The amplicons size was between 90bp and 150bp. Each primers pair was predicted to be specific for the target gene, as tested with BLAST (Basic Local Alignment Search Tool).

Further primer testing was performed once the primers were received. Firstly, the annealing temperature was accurately determined by running a PCR using these primers, cDNA (as the primers are for reverse transcription quantitative PCR (RT-qPCR), as opposed to ChIP primers that are tested in the same manner but using DNA template) and running the program at different annealing temperatures, such as 58°C, 60°C and 63°C. The PCR was run with RT-qPCR conditions, but the end product was run on a 2% agarose gel to ensure that the product is unique, no dimers are formed and that the selected temperature forms a distinct band of the predicted size. The melting curve of the product was monitored as well by RT-qPCR. Secondly, the linearity of the primers was tested. For this step, RT-qPCR was run using the temperature determined at the previous step and using serial dilutions of the cDNA used, eg 10x, 1x, 0.1x and 0.01x. The threshold cycle (Ct) values obtained were plotted against the concentration and the linear regression formula was calculated on microsoft excel. If the $0.9<R^2\leq1$, then the primers are amplifying in a linear manner, meaning that the Ct value obtained is dependent on the concentration of cDNA used in the reaction. Finally, although the primers were originally designed not to amplify genomic DNA, they were checked for genomic DNA amplification, running a RT-qPCR reaction with template that was produced with or without the addition of reverse transcriptase during the cDNA synthesis.

Table 2-6 Primers used for RT-qPCR analysis.

Human Gene	Forward Primer	Reverse Primer	Annealing Temperature
β-actin	ACTCTTCCAGCCTTCCTTCCTTCC	GTTGGCGTACAGGTCTTTGC	60°C
B2M	ATGAGTATGCCTGCCGTGTGA	GGCATCTTCAAACCTCCATG	60°C
G6PD	CGTCACCAAGAACATTCACG	ACAGGGAGGAGATGTGGTTG	60°C
GPI	CACACGCCATGCTGCCCTATG	TGGTGGTCCACACGGGTTCCA	60°C
TPI	CGCCTGCATTGGGGAGAAGCT	ACCAATGGCCCACACAGGCT	60°C
PK	GCTGCAACTTGGCGGGCA	AGTCAGCCCCATCCAGCAC	60°C
HK1	CACAGTCAAGATGTTGCCAAC	TGTGAACATTCTGGTTTTTCTCATG	60°C
PFKL	CTGGCTGTTCATCCCCGAG	CAATGGCACCCTCAGCGATG	60°C
ALDOA	GATTGCCATGGCGACCGTCA	TGTTAATGGCATTGAGGTTGATG	60°C
PGK1	CCGCTTTCATGTGGAGGAAG	ATAGACATCCCCTAGCTTGGAA	60°C
PBGM	TATGATGTCCCACCACCTCC	TGGCAATAGTATCCTTCAGACTC	60°C
ENO1	TGGGGCGTCATGGTGTCTCA	TGGGGCGTCATGGTGTCTCA	60°C
PGLS	TCGATCACGCCGAGAGCAC	TTCTTGGCGTAGTCCTCAGC	60°C
RPIA	CCGAGGAGGCCAAGAAGCT	GGACAATTGTAGAACCACTTCC	60°C
RPE	ATATAGAGGTCGATGGTGGAGTA	ATTGATCACAGATCTGGGGTCTT	60°C
TKT	GGCCAACCGCCTACGTATC	ACTTGTAGCGCATGGTGTGGA	60°C
PGD	ACACCACAAGACGGTGCCGAG	GAGCGATGGGCCATACCGGG	60°C
TALDO1	GATGCCCGCTTACCAGGAG	CTGCTCCAAACAACACAAAAAGTT	60°C
NEMO	TCGCTTGGAGGCTGCCACT	GCTCACTCTCCAGCTGCC	60°C
α-globin	GAGGCCCTGGAGAGGATGTTCC	ACAGCGCGTTGGGCATGTCGTC	60°C
β-globin	TACATTTGCTTCTGACACAAC	ACAGATCCCCAAAGGAC	60°C

2.7.1.3 RT-qPCR and data analysis

The expression of *G6PD* and other GPPP genes was assessed using SYBR Green-based RT-qPCR. The Maxima® SYBR Green/ROX qPCR Master Mix (2X) (Fermentas, York, UK) was used and the reactions run in the Applied Biosystems 7500 Real Time PCR Systems. The annealing temperature of the reaction was

adjusted based on primers' annealing temperature. The RT-qPCR run in three stages: stage 1: 95°C 10min, stage 2: 42 cycles of 94°C 15sec, annealing temp (60°C) 30sec, 72°C 33sec, stage 3: 95°C 15sec, 60°C 1min, 95°C 15sec, 60°C 15sec. Data acquisition was performed in the 3rd Step of Stage 2.

It is very important when running RT-qPCR assays to include a housekeeping gene in the analysis, which serves as an internal control, called reference gene. The reference gene is used to normalise differences in cell numbers, experimental treatment or RNA extraction efficiency. Since I very often treat the cells with pharmacological agents, it is crucial to ensure that the reference gene used remains stably expressed during the treatment. Upon validation, discussed in detail in 3.2.3, the reference gene used for this study is β-actin.

The ABI software was used to analyse the RT-qPCR data and calculate the Ct of each reaction. Samples were run in triplicates and the mean average of the Ct values was used. For the relative quantification of expression the ΔCt equation was used. Specifically, the difference of Ct values between the target and reference gene was calculated $\Delta Ct = Ct_x - Ct_{\beta\text{-}actin}$. Finally, the relative quantification of expression was calculated in the equation $2^{-(\Delta Cqx - \Delta Cq0)}$. Data is presented in error bars, showing the mean average and the standard error of mean, of n=3 independent experiments, unless stated otherwise.

2.7.2 Genome-wide expression arrays

For genome-wide gene expression analysis, RNA was isolated using the RNeasy Plus Micro kit (Qiagen, West Sussex, UK), as per manufacturer's advice. The extracted RNA was subjected to quality control analysis, which involved firstly Nanodrop analysis to assess the concentration and quality (OD 260/280 ratio > 1.9) and secondly control analysis using an Agilent Bioanalyser 2100, to assess the quality of the RNA and to ensure it is not degraded.

The extracted RNAs were used to run Affymetrix GeneChip Gene ST 2.0 Array (Affymetrix UK Ltd., High Wycombe UK), which was performed in collaboration with Dr. Robert Geffers' laboratory at the Helmholtz Centre for Infection Research, Braunschweig, Germany. For each run, 200ng of total RNA were used.

For data analysis, I used the Affymetrix Expression Console Software (Affymetrix UK Ltd., High Wycombe UK) to extract the data from the machine and then the GenePattern Software (Broad Institute, USA) for further analysis. For motif analysis, the MEME Suite (http://meme.nbcr.net/meme/) was used.

2.8 Protein expression assay

2.8.1 Western blot

Using the trypan blue exclusion method, 10^5 cells were counted and harvested, washed with PBS and resuspended in 50µl 2x loading buffer (0.25M Tris-HCl, pH:6.8, 20% β-mercaptoethanol, 40% glycerol, 16% SDS, and bromophenol blue). Protein was not quantified prior to western blotting, as the HDACI treatment alters the expression of only particular genes; therefore loading a particular number of cells in each well was decided to be a more accurate approach. After boiling the samples in 90°C for 5min to lyse and denature them, 20µl were loaded on a 12% polyacrylamide gel (Table 2-7) and run at 150V in 1x SDS-running buffer (5x SDS-running buffer: 15,1gr Trizma, 72gr Glycine, 5gr SDS in 1L water). Along with the samples, 6µl of a prestained broad-range protein marker (Cell signaling technology, New England Biolabs Ltd, Herts, UK) was run.

Table 2-7 Polyacrylamide gel protocol.

Reagents	Stacking gel (3%)	Resolving gel (12%)
Acrylamide (40%)	1.25ml	3ml
Lower buffer **(1.5M Tris-HCl pH 8.8, 0.4%SDS)**	-	2.5ml
Upper buffer **(0.5Tris-HCl pH 6.8, 0.4%SDS)**	1.25ml	-
Water	2.5ml	4.3ml
Ammonium persulfate (APS; 10%)	75µl	200µl
N, N, N', N'tetramethylethylenediamine (TEMED)	10µl	20µl

Once the gel had run, the proteins were transferred onto a PVDF membrane that was pre-wet in 100%methanol. A semi-dry approach (BioRad, UK) was used in which a transfer stack (2 pre-wet in transfer buffer Whatman papers, the PDVF membrane, the gel and then another 2 pre-wet Whatman papers) was produced. The proteins were transferred at 15V for 40 min in 1x transfer buffer (5.92g Trizma, 2.93g Glycine in 1L water). Then, the membrane was removed from the transfer-blot and was blocked in 5% skim milk (Sigma-Aldrich Company Ltd., Dorset, UK) in PBS+0.1%Tween20 for 1hour at room temperature (or 4°C overnight). Following that, the membrane was rinsed in PBS+0.1%Tween20 twice for 5min and then incubated for 1hour at room temperature (or 4°C overnight) in the primary antibody (Table 2-8) at a 1:2000 dilution in PBS+0.1%Tween20. The membrane was washed three times in PBS+0.5%Tween20 for 5min and then incubated for 30min in the secondary antibody (Table 2-8) at a 1:2000 dilution in PBS+0.1%Tween20. After washing the membrane three times for 10min each in PBS+0.5%Tween20, it was incubated for 1min with ECL detection reagent (GE healthcare, UK) and then developed in the dark room on X-ray films (Kodak).

Table 2-8 Antibodies used for Western blotting.

Primary Antibodies	**Species**	**Dilution**	**Supplier**
α-human G6PD (ab993)	Rabbit polyclonal	1:2000	Abcam, Cambridge, UK
α- human β-actin (sc-1616)	Goat polyclonal	1:1000	Santa Cruz Biotechnology Inc., Heidelberg, Germany
Secondary Antibodies	**Species**	**Dilution**	**Supplier**
α -rabbit HRP	Goat	1:2000	Dako, Cambridgeshire, UK
α -goat HRP	Rabbit	1:2000	Dako, Cambridgeshire, UK

To re-blot the membrane with the loading control β-actin antibody, the membrane was stripped using a stripping buffer (Tris-HCl 20mM, pH 7.5, 6M Glutamic acid hydrochloride (GaHCl), 0.2% Nonyl phenoxypolyethoxylethanol-40 (NP-40) and 0.1M β-mercaptoethanol added fresh) for 5min twice. Then, the membrane was washed for 1min with water. The primary antibody was added at a dilution of 1:1000 in blocking solution (5% milk in PBS+ 0.1%Tween20) and then the rest of the procedure was followed as with the first blotting.

2.8.2 Protein expression quantification

To quantify the bands and estimate the fold difference in the expression of G6PD relative to the loading control β-actin, I used the Adobe photoshopCS3 and ImageJ software (http://rsbweb.nih.gov/ij/). Specifically, Adobe photoshopCS3 was used to select the bands. Following the selection, the ImageJ software quantified the background and each band's intensity to provide numerical values. The estimated values of the bands were substracted from the background and then the G6PD values were normalised to β-actin loading control values. In that way the fold changes were estimated.

2.9 G6PD enzymatic activity assay

2.9.1 Preparation of cell lysates

Cultured cells were harvested, washed in PBS and then the cell lysates were prepared by four cycles of flash-freezing-thawing of the cell pellets in freshly-made G6PD lysis buffer (Tris-HCl pH 7.4 10 mM, EDTA 1 mM, ε-aminocaproic acid (EACA) 1 mM, NaCl 10 mM, MgCl2 3 mM, NADP 20 mM). The freezing-thawing cycles took place in liquid nitrogen and then in a room temperature waterbath. Cell extracts were centrifuged (15.000 rpm, 10 min, 4°C) and the clear supernatants were assayed for G6PD activity. The protein lysates were stored at -80°C until assayed. An aliquot of each lysate (30μl) was used to measure total protein concentration by the Pierce BCA protein assay (Thermo Scientific, USA), as per manufacturer's instructions.

2.9.2 Spectrophotometric assay

Total G6PD enzymatic activity was measured in duplicates by a spectrophotometric assay in the Hammersmith Hospital Diagnostic Haematology lab by Mr David Roper and Ms Lynn Robertson. In brief, cell extracts were added at the reaction solution (Tris-HCl pH 8.0 8.7 mM, MgCl2 8.7 mM, NADP 0.35 mM, Glucose-6-phosphate 1.33 mM). Then, the activity of G6PD is assayed by following the rate of production of NADPH, which unlike NADP has a peak of UV light absorption at 340 nm. The measured rate was converted to International Units, and divided by the amount of protein measured with the colorimetric Pierce BCA protein assay (Thermo Fisher Scientific, Rockford, IL), as per manufacturer's instructions.

2.10 Chromatin Immunoprecipitation

2.10.1 ChIP protocol

To investigate the epigenetic status of the GPPP gene promoters I employed ChIP. The first step of the protocol involved the crosslinking of the chromatin and its sonication to obtain small fragments of chromatin (with the preserved protein status on the DNA). This step was conducted using the Shearing ChIP kit (Diagenode, Cambridge, UK) as per manufacturer's advice and was subjected to further optimisation to match my experimental conditions. Briefly, the cells were crosslinked in 1% final concentration of formaldehyde for 10min at room temperature, whilst the samples were rotating. Crosslinking was performed in 9ml PBS per 10^7 cells with the addition of 630µl buffer A. Crosslinking was quenched with the addition of glycine at a 0.125M final concentration and an extra incubation for 5min at room temperature. The cells were washed in ice-cold PBS (working on ice throughout the protocol is crucial in order to maintain the crosslinked chromatin and reduce the degradation of proteins) and incubated sequentially with buffers B and C (3ml per 10^7 cells), each followed by 10min incubation at 4°C. Incubation with buffers B and C ensured cell lysis in order to obtain cell nuclei. These are then resuspended in buffer D (10^7 cells /300µl), in which proteinase inhibitors have been added at a 1:100 concentration. Samples were sonicated using a standard Bioruptor® sonicator (Diagenode, Cambridge, UK), for 20min (0.5ON, 0.5OFF on high power) in with the aim of getting fragments of around 300bp. The sonicated sample was spun down to precipitate the debris and thus obtained a clear chromatin suspension, which was used for the immunoprecipitation. A 10µl sample was kept to assess sonication efficiency by running on an agarose gel.

Table 2-9 Non-commercial buffers used for ChIP.

Dilution buffer	Wash buffer A	Wash buffer B	Elution buffer
50 mM Hepes pH 7.9 140 mM NaCl 1mM EDTA 1% Triton X-100 0.1% Na-deoxycholate *+add proteinase inhibitors before use*	50 mM Hepes pH 7.9 500 mM NaCl 1mM EDTA 1% Triton X-100 0.1% Na-deoxycholate 0.1% SDS	20 mM Tris pH 8.0 1 mM EDTA 250 mM LiCl 0.5% NP-40 0.5% Na-deoxycholate	50 mM Tris pH 8.0 50mM NaCl 1 mM EDTA *+add* 1% SDS *on the day* *+add DNase-free RNaseA to 20µg/ml final on the day*

The next step of the ChIP protocol was the immunoprecipitation (IP). The chromatin was diluted 10 times in dilution buffer (Table 2-9), so that the amount of SDS coming from the lysis buffer would be low and would not affect the IP. For each IP, the diluted chromatin used was estimated to be that of 2million cells. Also, 1:10 of that amount was used as input control and temporarily stored at -80°C. Next, the antibody-bead complex was prepared by the incubation of 15µl protein G magnetic beads (Invitrogen, Paisley, UK) with 2.5µg antibody (Table 2-10) in 800µl dilution buffer for 2h at 4°C rotating. Following that, the supernatant was removed using a magnet and the chromatin was added to the antibody-bead complex and incubated in 800µl dilution buffer overnight at 4°C rotating. The day after, the antibody-bead-chromatin complex was washed sequentially 150µl in dilution buffer, wash buffer A, wash buffer B (Table 2-9) and TE buffer for 5 min at 4 °C on rotating wheel. From this stage on, the stored input control was processed together with the eluted chromatin. The washed chromatin was eluted from the beads and the crosslinks were reverted in 150 µl elution buffer (Table 2-9) for 4h at 68°C (shaking at 1400rpm).

To extract the DNA, proteinase K was added to a final concentration of 200µg/ml to the samples, which were incubated at 45°C for 2h. The DNA was recovered by phenol:chloroform extraction and ethanol precipitation. Equal volume of phenol:chloroform:isoamyl alcohol (Sigma-Aldrich Company Ltd., Dorset, UK) was added to the samples, they were spun down at full speed and the top aqueous fraction was collected and precipitated with 2x 100%ethanol, 1/10 Sodium Acetate and 1µl

glycogen. The pellet was finally washed in 70% ethanol, air dried and then resuspended in TE buffer prior to RT-qPCR analysis.

Table 2-10 Antibodies used for ChIP.

Antibody	Species	Supplier	Antibody reference
H4 Acetylation	Rabbit polyclonal anti-human	Millipore, Watford, UK	06-598
H3 Acetylation	Rabbit polyclonal anti-human	Millipore, Watford, UK	06-599
Sp1	Rabbit polyclonal anti-human	Millipore, Watford, UK	07-645
Pol II N-terminal	Rabbit polyclonal anti-human	Santa Cruz Biotechnology	sc-899
CBP	Rabbit polyclonal anti-human	Santa Cruz Biotechnology, Heidelberg Germany	sc-369
p300	Rabbit polyclonal anti-human	Santa Cruz Biotechnology, Heidelberg Germany	sc-584
GCN5	Rabbit polyclonal anti-human	Santa Cruz Biotechnology, Heidelberg Germany	sc-20698
HDAC1	Rabbit polyclonal anti-human	Santa Cruz Biotechnology, Heidelberg Germany	sc-6299
HDAC3	Rabbit polyclonal anti-human	Santa Cruz Biotechnology, Heidelberg Germany	sc-11417
HDAC4/5/7	Rabbit polyclonal anti-human	Santa Cruz Biotechnology, Heidelberg Germany	sc-11421
HDAC6	Rabbit polyclonal anti-human	Millipore, Watford, UK	07-732
IgG	Rabbit polyclonal	Santa Cruz Biotechnology, Heidelberg Germany	sc-2027

2.10.2 ChIP primers and testing

Primers (Table 2-11) were purchased from Sigma-Aldrich Company Ltd. (Dorset, UK). They were designed and tested for the dimers and hairpins formation using the OligoCalc software (http://www.basic.northwestern.edu/

biotools/OligoCalc.html). Primers were designed to amplify part of the gene's promoter and in some cases the open reading frame or gene body. They have annealing temperature of 60°C and GC content 40-60%. The amplicons size was between 100bp and 150bp. Each primers pair was predicted to be specific for the target gene, as tested with BLAST (Basic Local Alignment Search Tool).

Table 2-11 Primers used for ChIP analysis.

Human Gene	Forward Primer	Reverse Primer	Annealing Temperature
G6PD.1	GGGAGCGGCGGACTGTGAAC	GGGGGCGGGGCTTGTGTTTT	60°C
G6PD.2	AGGCGGGGAAACCGGACAGT	ATCCCCAATTCCGGCGGGC	60°C
G6PD.3	TGTTGTGCTTGAGAACCGAGCA	TTGCCAAGCTGGGTGACCC	60°C
G6PD.4	TGGCAAGGGGAGGGCTGG	TGGAACTGCGTGCCCAGGA	60°C
G6PD.GB1	CCCAAGCCCATCCCCTATAT	CCACTTGTAGGTGCCCTCAT	60°C
G6PD.GB2	ACCCACGTGAGAGAATCTGC	CTGCTGCGTCTGCTTTTCTTA	60°C
GPI	CCAGCCCCAGAGTTCTTACA	AGGTGAAGACTGCAGTGAGC	60°C
TPI	GCCGGAGCTCACAGGTCT	GCACTGTTCCGACGTTCCC	60°C
GAPDH	TACTAGCGGTTTTACGGGCG	TCGAACAGGAGGAGCAGAGAGCGA	60°C
PGLS	CTGCGCATGTGCCAAAGACAA	GAGGAGGAAGCGCTCCCTA	60°C
RPIA	AGCAAGACCCGCGCAGCAG	TCCGGCCTCCGCTGAAGTC	60°C
DHFR	TCGCCTGCACAAATAGGGAC	AGAACGCGCGGTCAAGTTT	60°C
p21	GTGGCTCTGATTGGCTTTCTG	CTGAAAACAGGCAGCCCAAG	60°C
GW10	GGCTAATCCTCTATGGGAGTCTGTC	CCAGGTGCTCAAGGTCAACATC	60°C

Further primer testing was performed once the primers were received to establish the correct annealing temperature and the linearity of the primers. The process followed is the same as the one described in 2.7.1.2 for the RT-qPCR primers. The only difference to that protocol is that the input DNA used for testing is genomic DNA rather than cDNA, as the primers are not designed to amplify cDNA.

2.10.3 ChIP qPCR analysis

To analyse the results from the ChIP experiments, the isolated DNA was used to run qPCR using the same conditions as described in 2.7.1.3. The DNA used was the genomic DNA purified at the last step of the ChIP protocol (DNA that resulted from IP against the specific antibodies and IgG, as well as the genomic DNA input sample) and the primers used are those listed in Table 2-11. The primers GAPDH, Dihydrofolate reductase (DHFR) and p21 were used as positive controls for certain modifications at baseline. Specifically, GAPDH is used as a positive control for acetylation and polymerase II binding, DHFR is known to be bound by Sp1 on its promoter and *p21* promoter is bound by all the HATs and HDACs tested. The positive controls for the modifications were chosen based on ENCODE analysis of the promoters. Primers designed for the region named as GW10 (Table 2-11), amplify an intergenic region on chromosome 10 (AL392045); GW10 is not a gene. Therefore, GW10 is used as a negative control for promoter modifications.

To analyse the qPCR results I used the percentage of input method. In this method, signals (measured as Ct values) obtained from the ChIP are divided by signals obtained from the input sample. Typically, 1% of starting chromatin is used as input. The formula used is the following: %input=(2^-IP/2^-input)*(DF/100), where DF is the dilution factor of the input, i.e. 100 when 1% of starting chromatin is used, IP is the Ct value obtained by either an antibody against a modification or against IgG and finally input is the Ct value obtained by the input control reaction. The %input obtained by an IP against a modification must be above the one obtained from the IgG sample that is the background in order to show presence of a given modification at the region tested by the specific primers.

2.11 Reporter assays

2.11.1 Cloning

2.11.1.1 Plasmid

The pGL3-basic plasmid vector (Figure 2-2; Promega, Southampton, UK) was used for cloning of promoter fragments, as it is a luciferase gene-containing plasmid that is frequently used for reporter assays. It also contains a multi-cloning site upstream the luciferase gene.

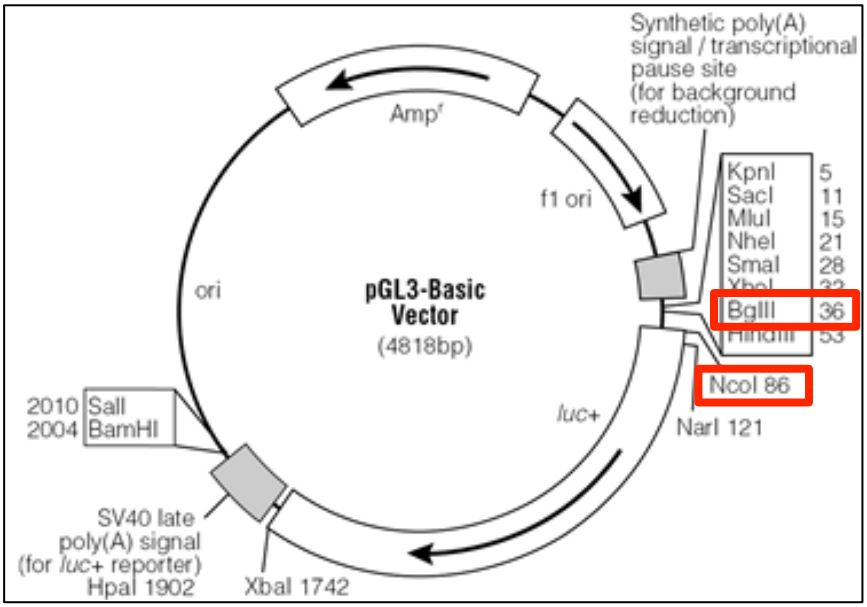


Figure 2-2 pGL3-basic plasmid vector. The plasmid (Promega, Southampton, UK) contains a multi-cloning site upstream the luciferase gene allows cloning of promoter fragments. Boxed in red are the BglII and NcoI restriction enzymes used to clone the promoter DNA. The plasmid is ampicillin-resistant.

2.11.1.2 PCR amplification and purification of amplicons for cloning

G6PD promoter DNA fragments were amplified using specific primers (Table 2-12 and Figure 2-3; Sigma-Aldrich Company Ltd., Dorset, UK). A common reverse primer and different forward primers were used. During primer designing, restriction sites were added to the 5' ends of primers (An NcoI site was added to the R primer and BglII sites were added to the F primers). The fragments were amplified by end-

point PCR, using Ready-made Dream Taq Polymerase mix (Fermentas, York, UK), according to manufacturer's instructions. The PCR cycle was as follows: 95°C 2min 1 cycle; 95°C 30sec, 60°C 30sec, 72°C 30sec (F1, F2, F3) or 90sec (F4, F5, F6) 35 cycles; 72°C 10min 1 cycle. Negative template controls were run for all different primers used.

Table 2-12 PCR primers used for *G6PD* promoter amplification.[a]

Primer	Sequence	Annealing Temperature	Amplicon size amplified with R primer
R	ATCACCATGGACGCTGTCTGG TGGAAGA	60°C	
F1	ATCAAGATCTGAGCCCAGAG CCAGCAGT	60°C	200bp
F2	ATCAAGATCTAGGGCTGGAG CTGAACTC	60°C	415bp
F3	ATCAAGATCTGACGAAGCGC AGGTAACC	60°C	645bp
F4	ATCAAGATCTGGTATGGCAGG CAGCCGG	60°C	908bp
F5	ATCAAGATCTCTCCGGGGGAG GAATCAAG	60°C	1207bp
F6	ATCAAGATCTGGTGTCATAGC TGTGGGATC	60°C	1464bp

[a] In red are highlighted the restriction sites added to the 5' end of each primers. NcoI sites for R and BglII sites for F.

The amplified DNA was run on 1.5% agarose gel and the bands of interest were extracted using a Gel extraction kit (Fermentas, York, UK), according to manufacturer's instructions.

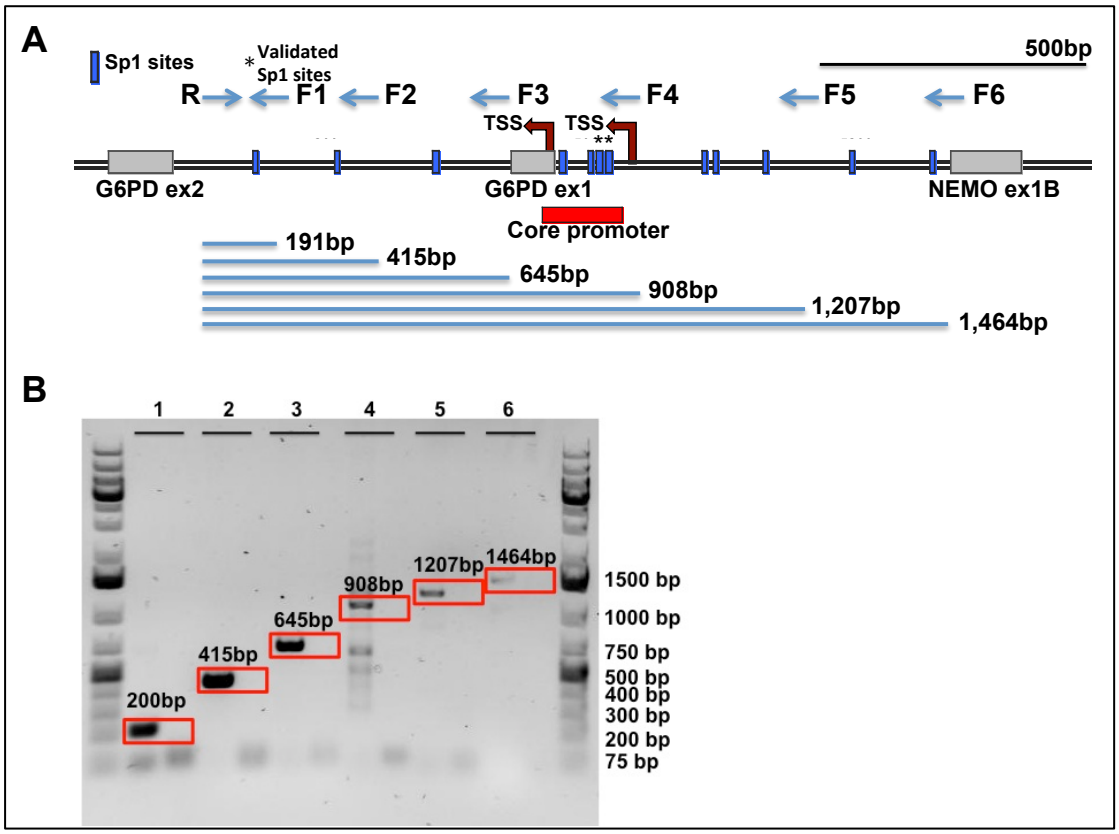


Figure 2-3 Transactivation assay design. (A) Representation of the wider *G6PD* promoter B, spanning between *G6PD* exon 2 and *NEMO* exon 1B. The core promoter is also highlighted. In blue are the predicted Sp1 binding sites and marked with an asterisk are the two previously validated sites. The common reverse (R) and the different forward (F1-6) primers that were used to amplify six increasing in length parts of the promoter. **(B)** Picture of a 1.5% agarose gel showing the amplified fragments of the promoter and the NTC control. The fragments were gel extracted, digested with BglII/NcoI and cloned into a digested with BglII/NcoI and gel extracted pGL3-basic vector (vector shown in **Figure 2-2**).

2.11.1.3 Digestion, Ligation and bacterial cloning

Inserts and plasmid (Figure 2-2 and Figure 2-3) were digested with 1unit of BglII and NcoI (Fermentas, York, UK) for 1h at 37°C. Following the digestion and gel extraction, the inserts were ligated to the plasmid using T4 ligase (Fermentas, York, UK). Vector and insert amounts were determined based on the following formula: ng insert= (ng vector*kb insert*3)/ kb vector.

To transform DH5a competent cells (Zymo Research, Freiburg, Germany) with the plasmid constructs, 1µl of contruct was mixed with 50µl DH5a competent cells and were incubated on ice for 30min. After a heat shock incubation at 42°C for

45sec and then an incubation on ice for 1min, 1ml of SOC medium (Invitrogen, Paisley, UK) was added to the cells and they were subsequently incubated at 37°C for 45 min shaking. Following the incubation 100µl of cell suspension was plated on LB+Amp (100µg/ml) plates and there were incubated at 37°C overnight. The following day single colonies were picked and further grown in liquid LB+Amp (100µg/ml) overnight to prepare miniprep cultures.

Plasmid DNA was purified using Qiagen mini prep kit according to manufacturer's instructions. After purification, the clones were screened for the presence of the specific insert of interest by BglII/NcoI restriction enzyme digestion. DNA from positive minipreps was sequenced directly by using BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, UK) and run on ABI 3730 Genetic Analyzer, as described in DNA sequencing. Once the clones were confirmed by sequencing, larger plasmid quantities were isolated using the Qiagen maxi prep kit according to manufacturer's instructions.

2.11.1.4 DNA sequencing

The sequencing reaction was performed using a BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems, UK). The reaction mix was prepared by adding 1µl the purified PCR product (10 to 50ng DNA), 1µl of 8pmol/µl of template- specific primer, 0.5µl of BigDye Terminator ready reaction mix (containing the fluorescent dideoxy terminator nucleotides) and 2µl of 5x BigDye Terminator buffer to 5.5µl of distilled water in a 96-well plate. The cycle sequencing reaction was performed in an ABI 3130 thermal cycler (Applied Biosystems, UK). The reaction consisted of an initial denaturation step of 94 °C for 1 minute followed by 25 cycles of 96°C for 10seconds, 50°C for 5 seconds and 60°C for 4 minutes. Reactions were held at 16°C until further use. DNA from the sequencing reaction was precipitated by adding 50µl of ice-cold 100% ethanol, 2µl of 3M Sodium acetate and 2µl of 125mM EDTA to each well, followed by centrifugation at 3400rpm for 30 minutes. A further precipitation step was performed by the addition of 50µl of ice-cold 70% ethanol and centrifugation at 3400rpm for 15 minutes. The alcohol was then removed by brief centrifugation and the precipitated DNA was resuspended in 10µl Hi-Di Formamide, a highly deionised formamide used for electrokinetic injection on capillary electrophoresis systems. Capillary electrophoresis was then performed on an ABI3700 DNA Analyzer.

The chromatograms resulting from the sequencing reactions were visualised with the 4peaks software (mekentosj.com/4peaks) and sequences were aligned using the NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.11.2 Cell transfections

The 293T cells were used for transient transfections. On day 1, cells were seeded at concentration of 1.5-2x10^5 cells/well on a 24 well plate. On day 2, confluence of approximately 70% was obtained. One well was left untranfected, one well was transfected with pGL3-basic alone and the rest of the well were transfected with the plasmid constructs. The cells were co-transfected with the renilla plasmid control, which is essential for transfection efficiency control. Furthermore, they were co-transfected with a pEGFP-c1 plasmid in order to estimate transfection efficiency either visually with light microscopy or with flow cytometry. Plasmid DNA (450μg construct + 450μg pEGFP-c1 + 50μg renilla control plasmid) was diluted in 50μl Optimem medium (Gibco, Invitrogen, Paisley, UK). Additionally lipofectamine (2μl/well) was diluted in Optimem medium (50μl/transfection) and was incubated in RT for 5min before adding 50μl of it to the diluted plasmid. The diluted plasmid was incubated for 30 min at RT in the presence of lipofectamine. The media of the cells was replenished with 400μl antibiotics-free medium and then 100μl of the plasmid mixture was added to them. After 2h incubation at 37°C, the lipofection was removed by replenishing the medium with fresh full medium (containing antibiotics). The cells were incubated for 48h (when treating with NaBu, that was done during the last 24h of the incubation) and then transfection efficiency was estimated by flow cytometric analysis. GFP^+ cells appear as $FITC^+$. Transfection was above 90% for all replicate experiments.

2.11.3 Luciferase assay

Luciferase was measured using the Dual-Glo Luciferase assay system (Promega, Southampton, UK), as per manufacturer's instructions. Briefly, 1x passive lysis buffer is added to the cells, which are incubated for 15min at RT shaking. Following stages were conducted on ice. Cells were spun at 14000 rpm 4°C for 5min to remove cell debris. 10μl of the supernatant were transferred to the Luciferase reaction plate (done in technical triplicates). 35μl of LARII reagent were added to each well, which acts as a substrate for luciferase, producing a luminescent signal measured by a fluoroscan (Qiagen, West Sussex, UK). Then, 35μl of Stop&Glo

reagent were added to each well, which quenches the luminescence from the construct and provides a substrate for renilla luminescence that is measured by the same machine.

To analyse the results obtained by the luminometer, the average of each construct luciferase measurement was normalised against the equivalent renilla measurement. Results are shown relative to pGL3-basic vector.

2.12 Mutagenesis assay

To mutagenise Sp1 binding sites, I used the pGL3-Fr5 plasmid construct produced in Cloning using the primers R and F5 (Table 2-12). I then, as shown in Figure 2-4, conducted two rounds of PCR to insert the mutations. Specifically, I designed primers (Table 2-13) on the mutation region to insert more than three mismatches on each direction. These primers were first used in combination with the external primers R and F5 (Table 2-12). For example, to mutagenise site 1, primers R (Table 2-12) + 1F (Table 2-13) and primers F5 (Table 2-12) + 1R (Table 2-13) were used to amplify two fragments of the insert. For PCR amplification DreamTaq Polymerase (Fermentas, York, UK) was used because it does not have proofreading activity and therefore allows inserting mutations. For this purpose, the annealing temperature was also set as low as 55°C. Then a 2nd PCR was performed using primers R and F5 (Table 2-12) and the 2 gel extracted PCR products from the 1st round as template to obtain the whole mutated insert. This was further cloned into the pGL3-basic vector, sequenced and luciferase assays were performed, as described in 2.11.

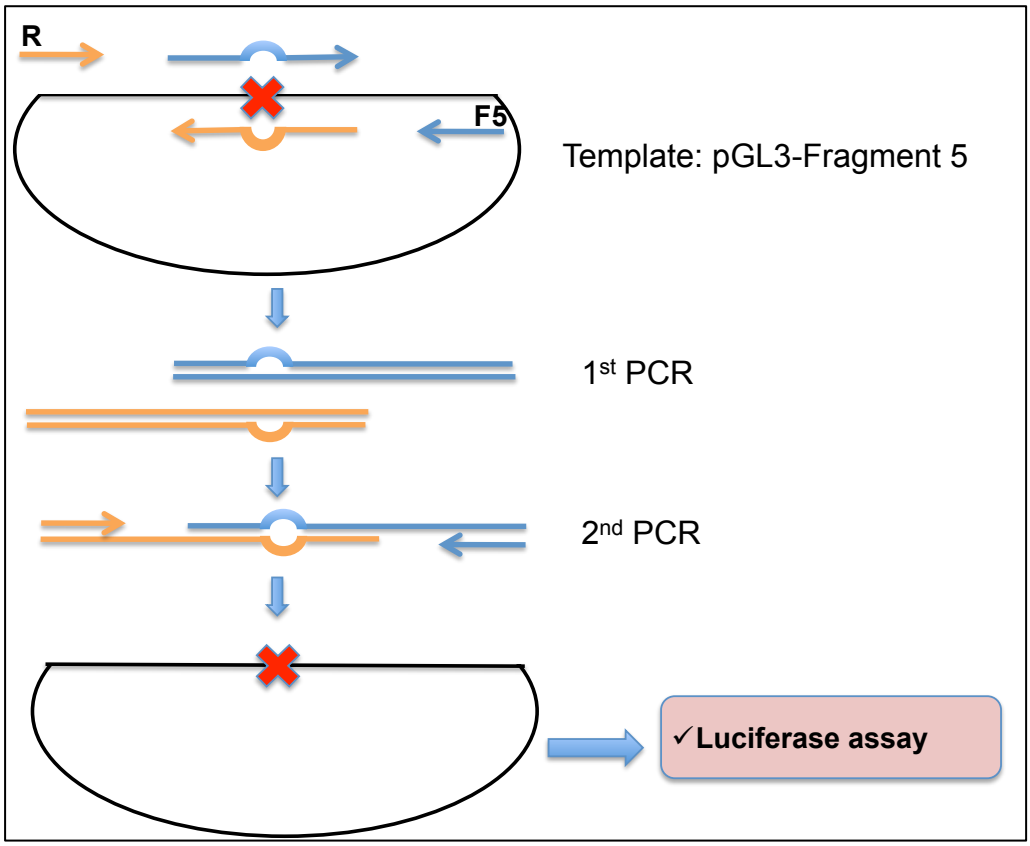


Figure 2-4 Mutagenesis assay strategy. pGL3-Fr5 construct produced in 2.11.1 is used as a template. Primers designed to insert mutations were used in combination with primers F5 and R to amplify the insert in two fragments. A 2nd PCR was performed using the external primers to connect the two fragments. The mutated fragment was then cloned in pGL3-basic vector. Luciferase assay was then carried out to measure the relevant promoter activity.

Table 2-13 Primers designed to insert Sp1 binding site mutations. [a]

Primer	Sequence
1F	GCC CAG GCG CCC GCC AAC GAT CCC GCC GAT TAA ATG GGC C
1R	GGC CCA TTT AAT CGG CGG GAT CGT TGG CGG GCG CCT GGG C
2F	GGG TGG TGG CCG AGG CTA CGA CAC GCA CGC CTC GCC TGA G
2R	CTC AGG CGA GGC GTG CGT GTC GTA GCC TCG GCC ACC ACC C
3F	CCG CCC CGC CCG CAC GAG AAG TGG TGG CCG AGG CCC CGC C
3R	GGC GGG GCC TCG GCC ACC ACT TCT CGT GCG GGC GGG GCG G
4F	CGC ACC TGC CCT CGC ATC GAT CCG CCC GCA CGA GGG GTG G
4R	CCA CCC CTC GTG CGG GCG GAT CGA TGC GAG GGC AGG TGC G
5F	GGA AAC CGG ACA GTA GGT ACT AGG CCT GGC CGG CGA TGG G
5R	CCC ATC GCC GGC CAG GCC TAG TAC CTA CTG TCC GGT TTC C
6F	CGA GGC CGC CGG GGC AAT CGA AGA AAC CGG ACA GTA GGG G
6R	CCC CTA CTG TCC GGT TTC TTC GAT TGC CCC GGC GGC CTC G
7F	GTG GCG CGG CAG AAG GCT ATG CAC AGG AGC CGA GGG ACA G
7R	CTG TCC CTC GGC TCC TGT GCA TAG CCT TCT GCC GCG CCA C

[a] In red are highlighted the sites were Sp1 binding was disrupted.

2.13 Dominant negative Sp1 expression plasmid

Prof. Gerald Thiel kindly provided us with the pEBGV and pEBGV-Sp1.DN plasmids (Al-Sarraj et al., 2004). The pEBGV is an empty of insert control plasmid. The pEBGV-Sp1.DN encodes for a fusion protein consisting of GST-NLS and the zinc finger domain of Sp1. Therefore the expressed protein is able to bind to Sp1 binding sites but lacks of transactivation domain. The plasmids were transfected into 293T cells (2.11.2) and mRNA expression was determined using RT-qPCR (2.7.1.3).

2.14 Bioinformatic analysis

2.14.1 Promoter and TF binding sites identification

Core and extended promoter regions were obtained from the NCBI Gene tool (http://www.ncbi.nlm.nih.gov/gene) and were verified by analysis on the UCSC ENCODE genome browser (http://genome.ucsc.edu/ENCODE/). TF binding sites were predicted using the TF binding site prediction browsers TFSearch (http://www.cbrc.jp/research/db/TFSEARCH.html) and CONSITE (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/).

2.14.2 Statistical analysis

Statistical analysis was conducted using the GraphPad Prism software. The p values were obtained from either one-way analysis of variance (ANOVA) or student's t-test. The asterisks correlate to the p values as follows: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$. Error bars represent either standard deviation (SD) for microarray analysis or standard error of the mean (S.E.M.) for all other analyses.

3 Results I: Glycolytic enzyme gene expression in human cell lines

3.1 Introduction

Currently, there is no specific treatment for haemolytic anaemias caused by mutations on the genes expressing enzymes of the GPPP. Symptomatic patients with chronic anaemia require regular blood transfusion, which carries the risk of transfusion-induced iron overload. Each red cell concentrate used for transfusion contains 250mg of iron; however, under physiological conditions, only 1-2mg of iron can be excreted each day. This leads to deposition of iron in essential organs, particularly the liver and heart, which can result in potentially life-threatening organ dysfunction. New treatments for such conditions are therefore urgently required.

HDACIs lead to histone hyper-acetylation and therefore can increase gene transcription. In a specific example studied in our lab, it was shown that a point mutation in the proximal promoter of *PIGM* abrogates binding of the TF Sp1, causes histone hypo-acetylation and transcriptional repression that is specific to *PIGM.* NaBu was shown to restore Sp1 binding to the core promoter, restore levels of histone acetylation and activate *PIGM* transcription (Almeida et al., 2006; 2007; almeida et al., 2009). *PIGM* is part of a housekeeping pathway, i.e. the GPI biosynthetic pathway; therefore, I hypothesised that NaBu could have therapeutic effects in other pathways with housekeeping functions, such as that of glycolysis.

3.1.1 Aim of the chapter

This chapter aims to determine the effect of HDAC inhibition on the GPPP in human cell lines. This was investigated at the levels of RNA, protein and enzymatic activity. Ultimately, I intend to identify genes within the GPPP that are upregulated upon treatment with HDACIs.

3.1.2 Experimental plan

To address the hypothesis that HDACIs would increase mRNA expression followed by an increase in protein levels and enzymatic activity, I treated normal wild type (WT) B cell lines as well as B cell lines from patients with specific GPPP deficiencies (Table 2-1) with HDACIs and assessed the effect at different time points.

First, I performed a series of control experiments aimed at a) determining the appropriate HDACI concentration, b) ensuring that NaBu used in individual experiments was active, and c) validating the reference genes that would be more

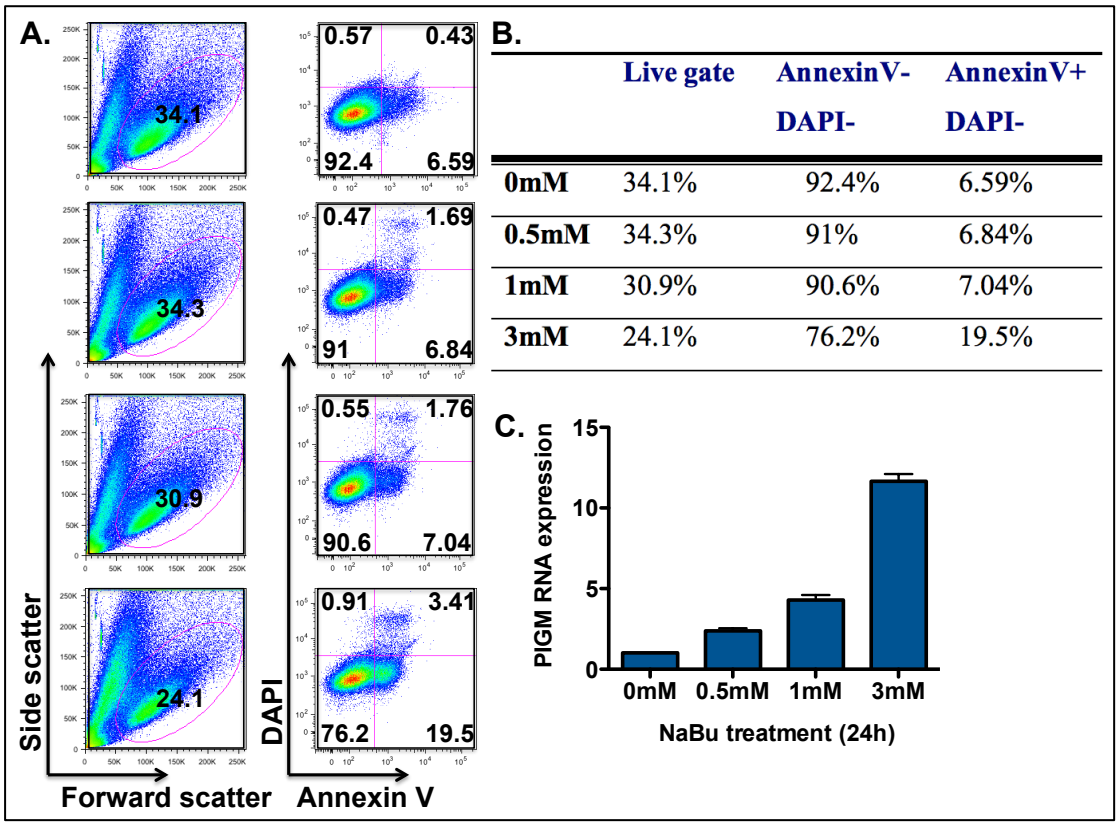
appropriate to use in RT-qPCR assays in samples treated with HDAC inhibitors, including NaBu.

3.2 Control experiments

3.2.1 Identification of NaBu concentration

To date, various studies have been published using NaBu to increase transcriptional expression of genes. The concentration typically used to treat cell lines, including B cell lines, with NaBu by our lab (Almeida et al., 2007; Caputo et al., 2013) and others (Bordonaro et al., 2011; deFazio et al., 2001; Miki et al., 2007; Robinson et al., 2012; Yu et al., 1998) is 3mM, as this was determined to be the lowest concentration to maintain cell viability whilst allowing HDAC inhibition. To establish that this concentration is appropriate for the purposes of my project and the particular cell line used, I titrated the concentration and determined the apoptosis rates as well as the effect on *PIGM* expression levels, which is known to be upregulated by NaBu(Almeida et al., 2007).

The WT B cell line P277 was treated with 0.5mM, 1mM, 3mM NaBu for 24h. Cell death was assessed by flow cytometry after staining with AnnexinV and DAPI (Figure 3-1). $AnnexinV^{-}DAPI^{-}$ events represent live cells, $AnnexinV^{+}DAPI^{-}$ events represent early apoptotic cells, whereas $AnnexinV^{+}DAPI^{+}$ cells are late apoptotic cells. As shown and summarised in Figure 3-1, NaBu has a dose-dependent effect on the viability: in particular, 3mM NaBu results in the highest apoptotic and lowest survival cell rates amongst the concentrations tested (Figure 3-1A, B). Nevertheless, the 3mM concentration resulted in the highest increase in *PIGM* expression mRNA levels (Figure 3-1C), which was used as a control gene that is known to be upregulated by NaBu; therefore, I performed the rest of my experiments in B cell lines using 3mM NaBu. Further concentration optimisation was performed when NaBu was used in primary erythroid cells and adherent 293T cell lines as shown in Figure 4-5, Figure 4-7 and Figure 6-8.



	Live gate	AnnexinV- DAPI-	AnnexinV+ DAPI-
0mM	34.1%	92.4%	6.59%
0.5mM	34.3%	91%	6.84%
1mM	30.9%	90.6%	7.04%
3mM	24.1%	76.2%	19.5%

Figure 3-1 NaBu concentration titration in the P277 WT cell line. (A) Flow cytometric analysis of the Annexin V and DAPI-stained P277 normal cell line upon 24h treatment with 0.5mM, 1mM and 3mM NaBu. **(B)** The table shows the percentage of live cells, estimated by the forward/side scatter (circled in 3-1A), the percentage of live cells ($AnnexinV^{-}DAPI^{-}$) and the percentage of apoptotic cells ($AnnexinV^{+}DAPI^{-}$). **(C)** Graph showing *PIGM* mRNA expression normalised to β-actin and relative to the untreated cells. Mean and S.E.M. are shown for n=3.

In order to confirm that NaBu was active in every single experiment, it was freshly reconstituted each week and was regularly tested for its ability to restore GPI expression on the surface of *PIGM* mutant, GPI deficient B cell lines, such as YK. The reason for regularly testing the effectiveness of NaBu arose from the observation that NaBu is unstable after long storage at -20°C. This can be assessed using flow cytometric analysis to measure staining with FLAER, an Alexa® 488-labeled variant of aerolysin, a unique protein that binds tightly and specifically to mammalian GPI anchors. Therefore, by treating *PIGM* mutant cell lines and observing restoration of surface GPI expression, I was able to confirm that NaBu used in my experiments was biologically active and efficient at inhibiting HDACs (Figure 3-2).

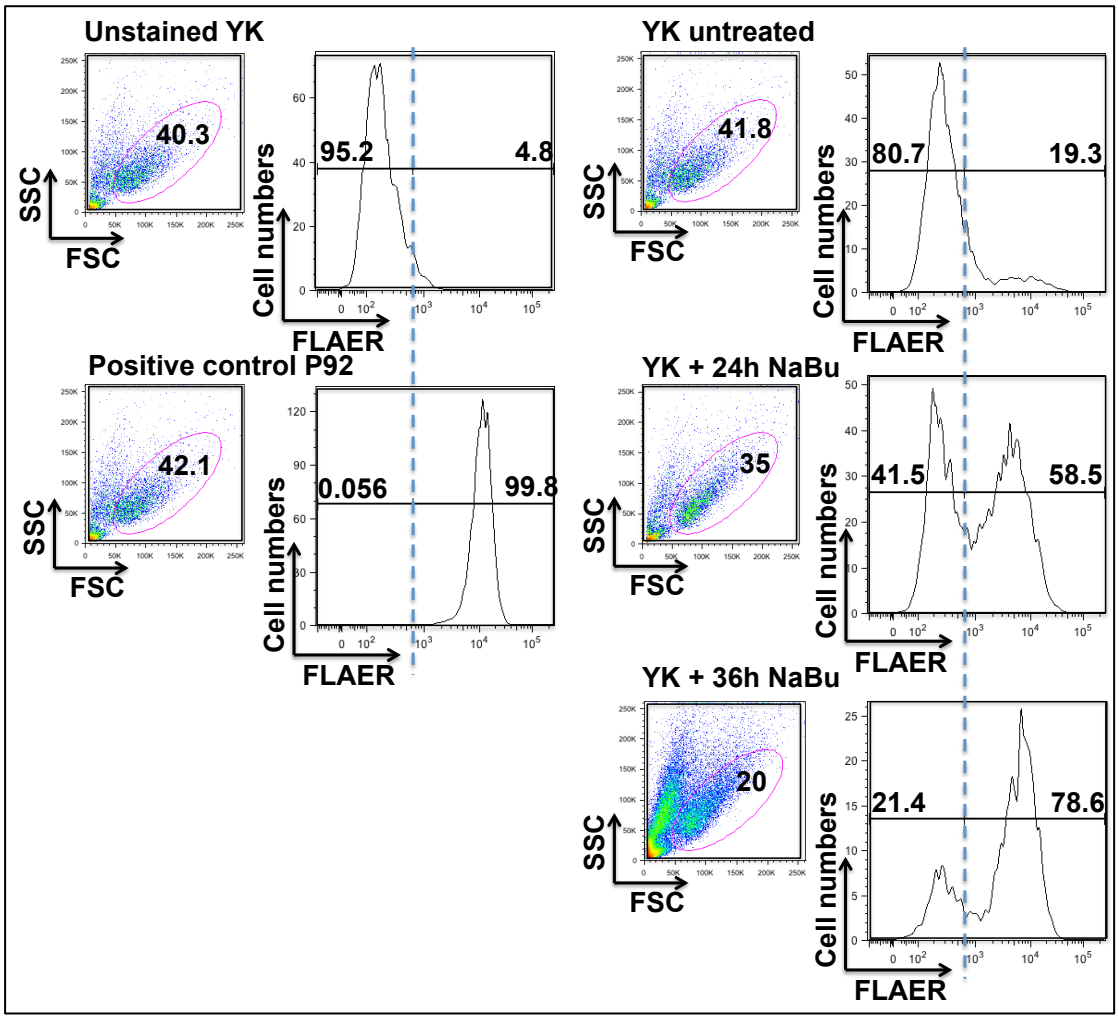


Figure 3-2 NaBu efficiency tested by GPI expression of the PIGM deficient cell line YK. The *PIGM* mutant and GPI deficient, YK cell line was treated with NaBu for 24h and 36h. Unstained YK and the GPI^{+} cell line P92 were also used as negative and positive controls, respectively. The gradual restoration of GPI expression by FLAER staining as shown by flow cytometry indicates active NaBu.

3.2.2 Determination of SAHA and TSA dose

Similarly to NaBu, I also optimised the concentrations of SAHA and TSA used for my experiments. To do so, I performed a flow cytometric assay using different concentrations of SAHA and TSA to restore GPI expression on the surface of the *PIGM* mutant GPI-deficient B cell line YK. Both the cell viability estimated by the live gate on a forward/side scatter during flow cytometric analysis and the upregulation of the GPI protein expression assessed by FLAER expression were taken into account to determine the appropriate drug concentrations.

YK cells were treated for 24h with 2.5nM, 10nM and 100nM TSA (Figure 3-3) or with 2μM, 4μM, 6μM and 8μM SAHA (Figure 3-4), which are concentrations within the suggested by the manufacturer range. As a positive control, the WT GPI-expressing P92 cell line was used, whilst unstained YK cells and DMSO-treated YK cells were used as negative controls. DMSO-treated cells were used instead of untreated cells because both SAHA and TSA are dissolved in DMSO. Based on the cell viability resulting from the addition of the drug as well as the upregulation of the GPI expression, I chose the 4μM and 100nM concentrations for SAHA and TSA, respectively for subsequent experiments. The findings of this experiment are consistent with our prior knowledge from published research in our lab (Almeida et al., 2007) and unpublished work that has shown that at 24h post-treatment GPI expression is partly restored by approximately 30-60%, whereas for complete restoration of the GPI expression treatment for 36h is required.

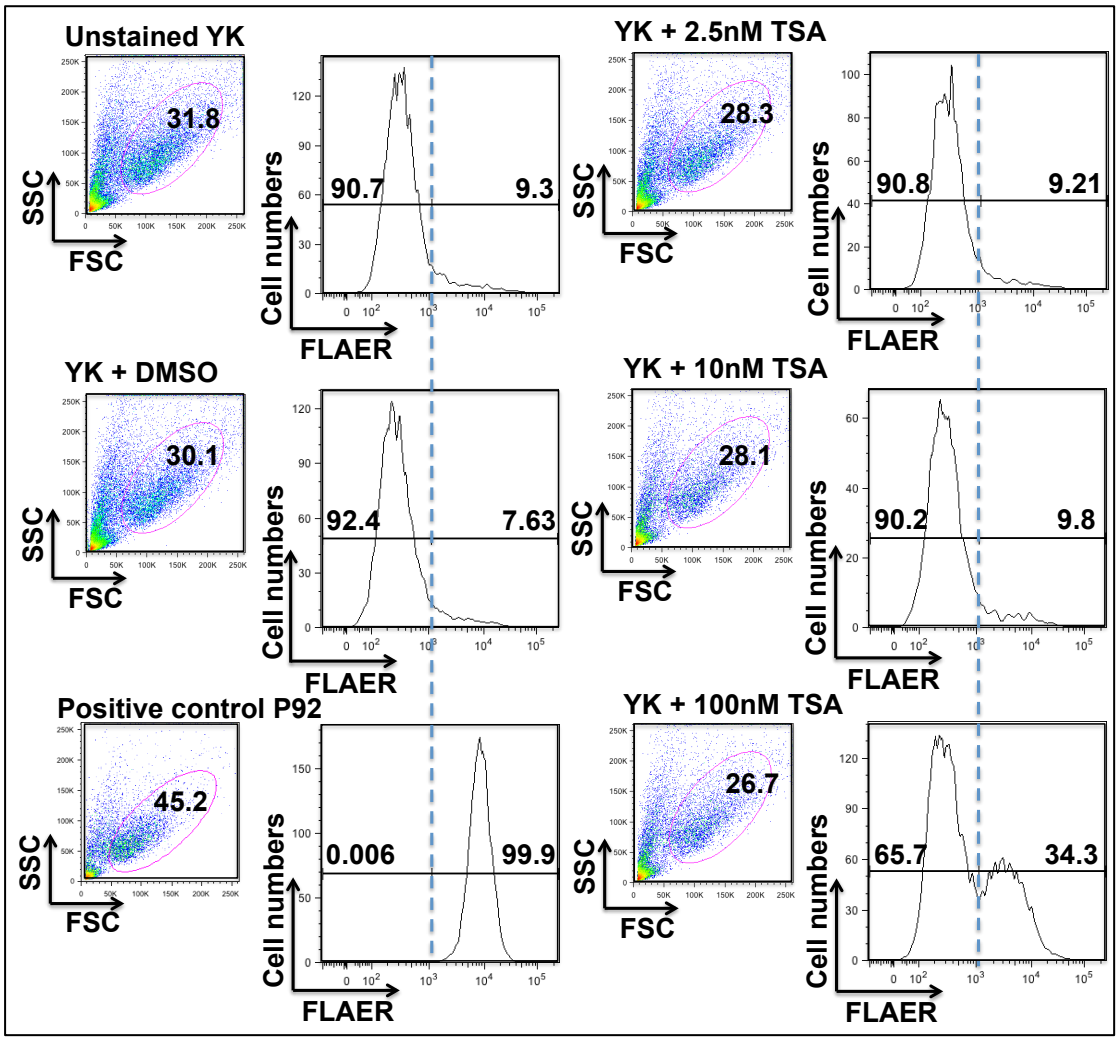


Figure 3-3 TSA concentrations tested by GPI expression of the PIGM deficient cell line YK. The *PIGM* mutant and GPI deficient, YK cell line was treated with 2.5nM, 10nM and 100nM TSA for 24h. Unstained YK and also YK treated with DMSO, which is what TSA is dissolved in are used as negative controls. The GPI^{+} cell line P92 is used as a positive control. The 100nM concentration was chosen for TSA treatments in cell lines.

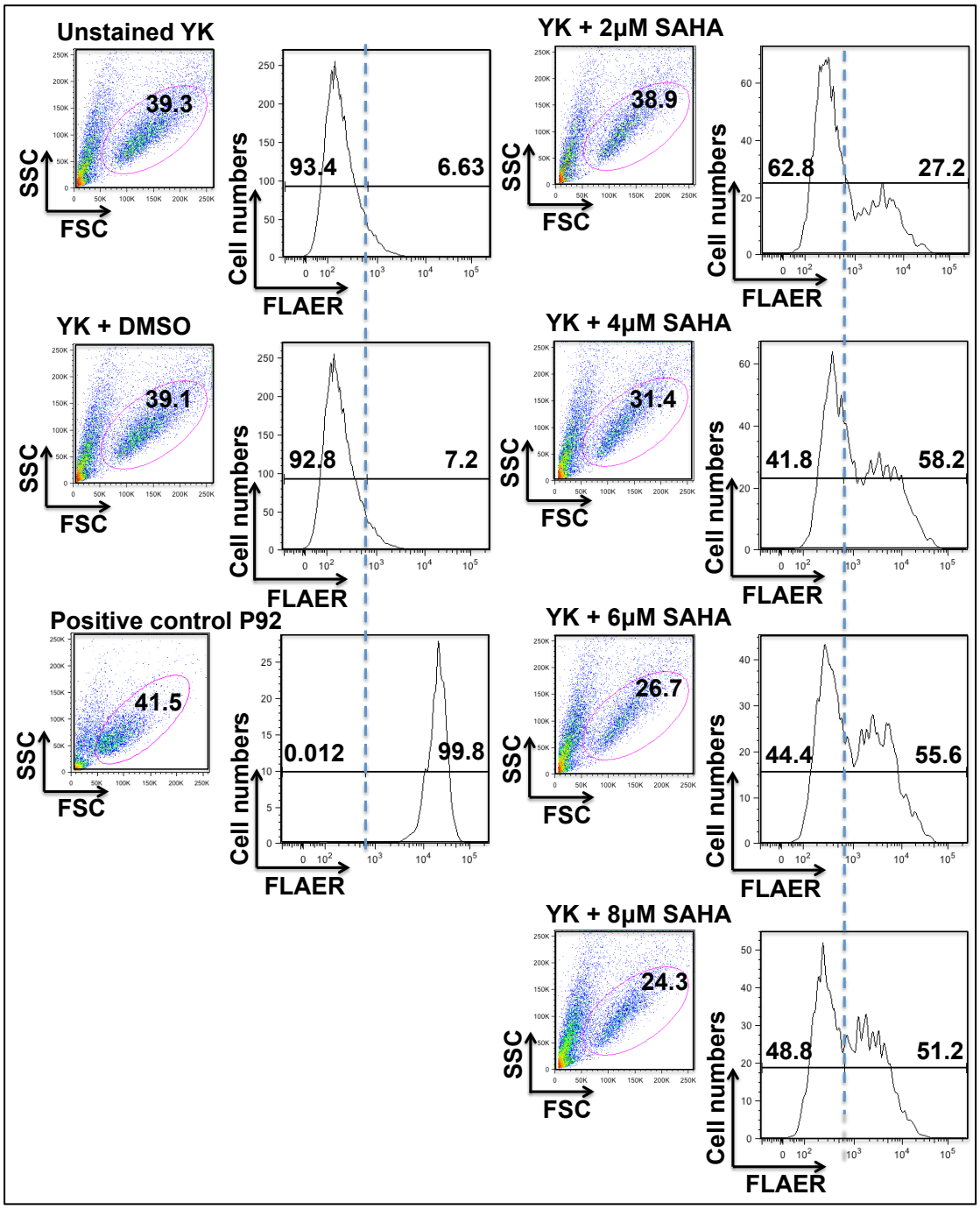


Figure 3-4 SAHA concentrations tested by GPI expression of the PIGM deficient cell line YK. The *PIGM* mutant and GPI deficient YK cell line was treated with 2μM, 4μM, 6μM and 8μM SAHA for 24h. Unstained YK and also YK treated with DMSO, which is what SAHA is dissolved in are used as negative controls. The GPI^{+} cell line P92 is used as a positive control. The 4μM concentration was chosen for SAHA treatments in cell lines.

3.2.3 RT- qPCR reference gene validation

RT-qPCR typically produces relative quantification of the target gene expression in reference to an internal control gene (reference gene). The amount of RNA assayed may fluctuate due to differences in cell numbers, experimental treatment or RNA extraction efficiency. However, the conditions of the experiment should not influence the expression of the reference gene in order to ensure accurate results. To ensure accuracy of my future RT-qPCR experiments it was essential that I chose a reference gene that was not affected by NaBu treatment. I therefore tested whether NaBu treatment impacts on the expression of commonly used housekeeping reference genes: *GAPDH, β-actin, 18S rRNA, β2-microglobulin* and *G6PD*, with *G6PD* and *GAPDH* coincidentally being part of the GPPP.

Total RNA was extracted after treatment of the wild type cell line P277 with 3mM NaBu for 0h, 8h and 24h. Equal amounts were used for reverse transcription and cDNA synthesis. For each time point, three samples were prepared with serial dilutions (1:10, 1:100 and 1:1000) and each sample was tested in triplicate by qPCR. The ratio of Ct values at each time point to the Ct value of the untreated cells was calculated for each sample (fold change of expression in time). The mean average Ct ratios ± S.E.M. for each time point is shown in Figure 3-5.

The mRNA levels of the candidate reference genes *GAPDH*, *β-actin*, *β-2-microtubulin* and *18S rRNA* did not change significantly over time ($p>0.05$) and therefore could be used for qPCR analysis. From these, I chose to use *β-actin* as my reference gene in all following qPCR assays discussed in this thesis as it showed the most consistent expression ($p = 0.91$) and it is not involved in the GPPP, suggesting it is an impartial reference gene.

This experiment also showed a significant increase in *G6PD* mRNA levels ($p<0.05$) with NaBu treatment (Figure 3-5). Furthermore, it shows that NaBu affects expression of only a few and not all housekeeping genes.

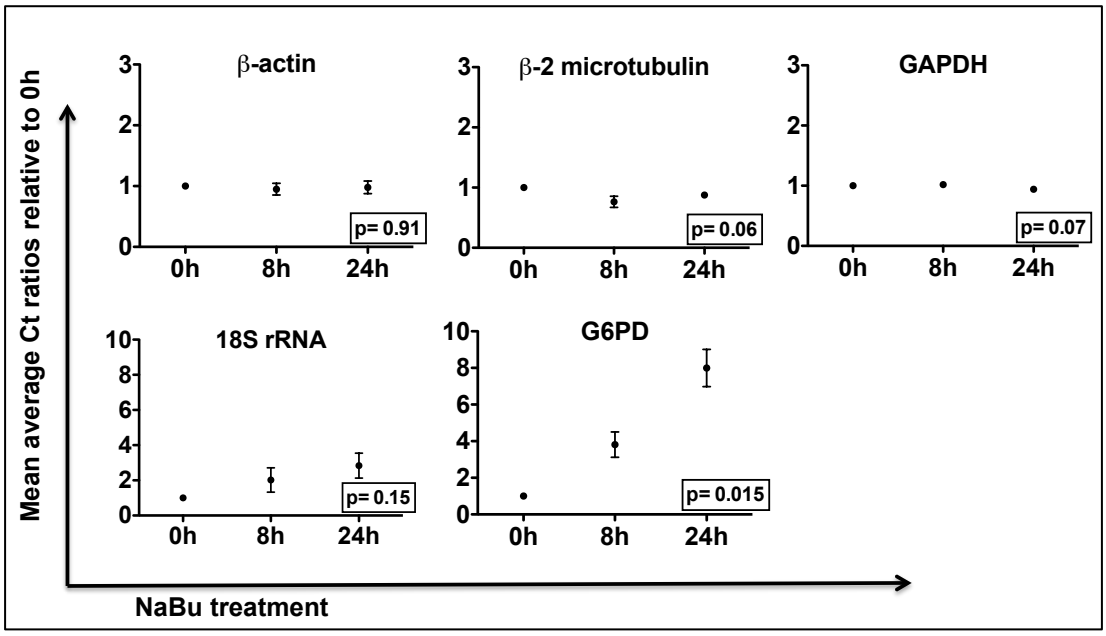


Figure 3-5 Effect of NaBu treatment of candidate qPCR reference genes. The P277 WT cell line was treated with NaBu for 8 and 24 hours. cDNA synthesised for equal amounts of total RNA were used in each time point. Mean and S.E.M. are shown for n=3. Significance values between differences in gene expression at different time points were tested by one-way ANOVA.

3.3 GPPP gene expression upon HDAC inhibition

To assess the responsiveness of the glycolytic pathway genes to HDAC inhibition and test my initial hypothesis that HDACIs can be used for the treatment of disorders of the GPPP, I treated the WT cell line P277 with 3mM NaBu in a time course of 8h, 16h, 24h and 36h. Then, I employed RT-qPCR to examine the expression of all the GPPP genes. Student's t-test was performed to examine the significance of each time point in comparison to its untreated condition.

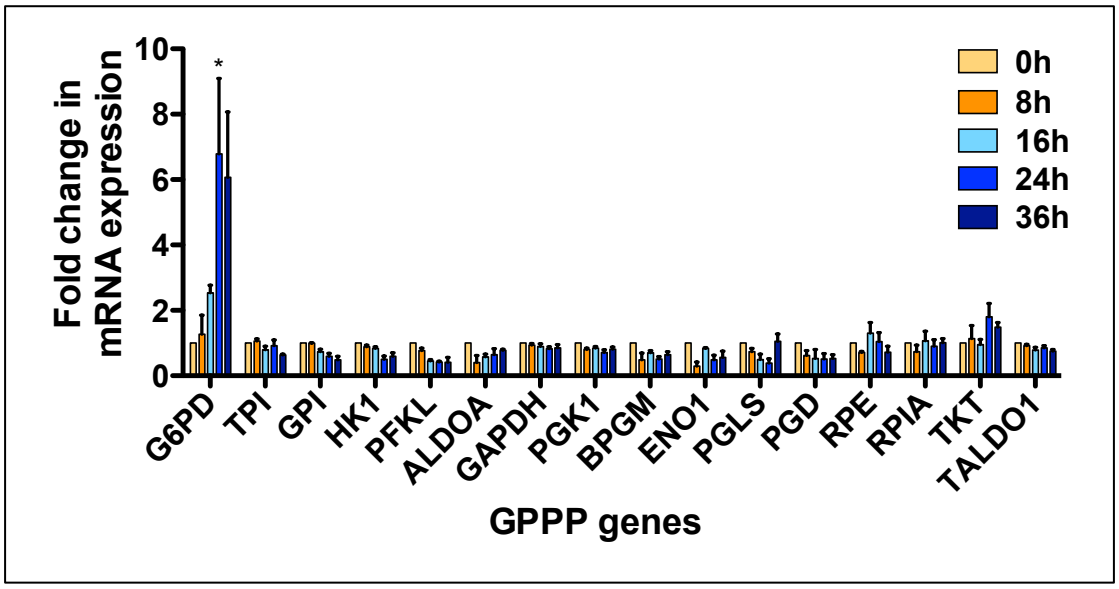


Figure 3-6 GPPP expression upon NaBu treatment. The WT cell line P277 was treated with 3mM NaBu in a time course of 8h, 16h, 24h and 36h. Fold change in gene expression relative to the untreated cells' expression for each gene is shown, as obtained by RT-qPCR analysis. Mean and S.E.M. are shown for n=3. Student's t-test was employed to show the significance of each time point compared to its untreated condition.

Figure 3-6 presents the effect of NaBu on the mRNA expression of all the GPPP genes, except *PKLR*, which is expressed only in erythroid and not B cells. It is apparent from this figure that NaBu selectively upregulates the mRNA expression of *G6PD* among 16 genes of the GPPP. In fact, there is a significant increase of *G6PD* after 24h of treatment, whereas no significant change was observed for the other genes.

In order to confirm the *G6PD*-specific increase upon NaBu treatment, I decided to study the expression of two of the genes that remain unaffected in more detail. Therefore, I performed treatments of the normal B cell line P277 not only with 3mM NaBu, but also with 4μM SAHA and 100nM TSA in a time course of 8h, 16h, 24h and 36h. As shown in Figure 3-7A, no significant change of either the *TPI* or *GPI* mRNA expression was shown by RT-qPCR analysis. Moreover, since the GPPP is primarily important in erythroid cells, I repeated the same experiment using the erythroleukaemia cell line K562 (Figure 3-7B), in order to determine whether these genes could be upregulated in a cell type-specific manner. However, neither NaBu,

SAHA, nor TSA were able to induce changes in the mRNA expression of *TPI* and *GPI.*

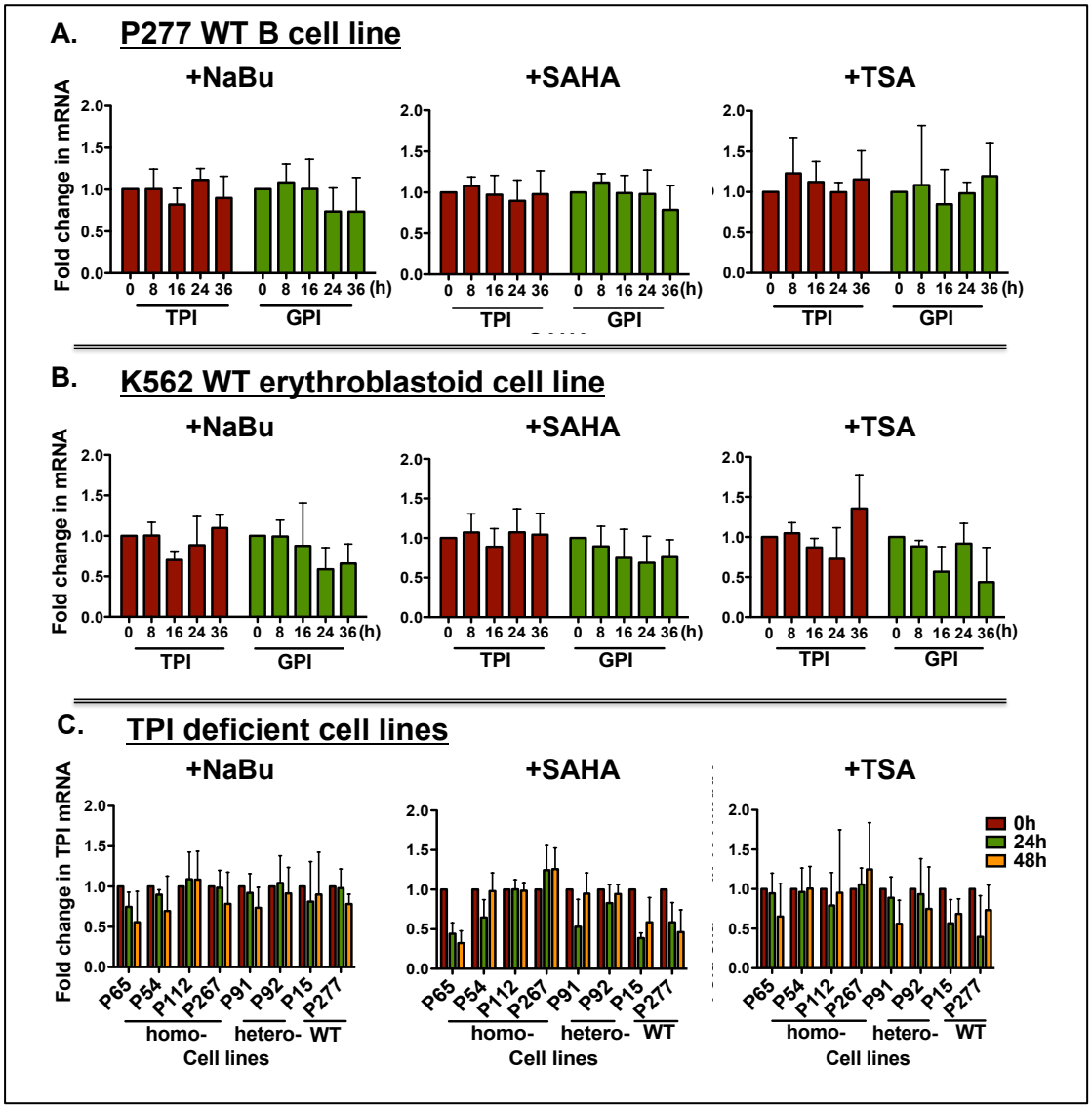


Figure 3-7 *TPI* and *GPI* expression in WT and deficient cell lines in response to HDAC inhibitors. (A) The normal B cell line P277 and **(B)** the erythroleukaemia cell line K562 were treated with 3mM NaBu, 4µM SAHA and 100nM TSA for 8h, 16h, 24h and 36h. Following the treatments, the mRNA expression of the genes *TPI* and *GPI* was assessed by RT-qPCR. **(C)** TPI deficient cells lines derived from homozygote deficient (P65, P54, P112 and P267) and heterozygote deficient (P91 and P92) individuals as well as the normal B cell lines P15 and P277 we treated with NaBu, SAHA and TSA for 24h and 48h. The mRNA expression of *TPI* was assessed by RT-qPCR. The mRNA expression is shown relative to the untreated cells expression. Mean and S.E.M are shown for n=3. Student's t-test was employed to show the significance of each time point compared to its untreated condition.

Furthermore, I explored the possibility that genes that are not affected by HDAC inhibition under physiological conditions might be more susceptible to HDAC inhibition under pathological conditions. Therefore, I took advantage of the TPI deficient cell lines that were available in the lab to test the *TPI* mRNA expression upon NaBu, SAHA and TSA treatment after 24h and 48h. For this experiment I used B cell lines derived from homozygous TPI deficient patients (P65, P54, P112 and P267) and heterozygous TPI mutant individuals (P91 and P92, who are the parents of P65) and compared these TPI deficient cell lines with the P277 and the P15 normal B cell lines. Treatment of these cell lines followed by RT-qPCR analysis (Figure 3-7C) showed that HDACIs do not affect the *TPI* mRNA expression levels in TPI deficient cell lines.

Taken together, the results of the experiments presented above provide significant evidence that HDACIs increase mRNA expression of *G6PD* specifically amongst the 16 genes of the GPPP that were tested. This data generates further questions regarding the mechanistic basis of the marked difference between *G6PD* and the other GPPP genes. In addition, it raises therapeutic implications for *G6PD* deficiency, which will be studied later in this thesis.

3.4 G6PD expression in normal and deficient cells

Having established the *G6PD*-selective effect of HDAC inhibition in the GPPP and in order to understand the significance and therapeutic relevance of my findings, I next focused on the impact of HDACIs on G6PD deficient cell lines. First, I compared baseline mRNA and protein levels as well as enzymatic activity of G6PD deficient and normal cells. As discussed in 1.4.3.2 and shown in Table 1-3, according to criteria, the severity of G6PD deficiency has been classified based upon the G6PD enzymatic activity levels. However, little is known about the *G6PD* mRNA and protein expression in deficient cells as opposed to normal cells. To address this issue, I compared *G6PD* mRNA and protein expression of the G6PD deficient P7 cell line (derived from the patient with G6PD Brighton patient, previously characterised by Dr Mark Layton, Hammersmith Hospital (McGonigle et al., 1998)) to several B (P277, P91, P92, P54, P267, P15, P174 and C1R) and erythroid (K562) cell lines, which express normal G6PD activity.

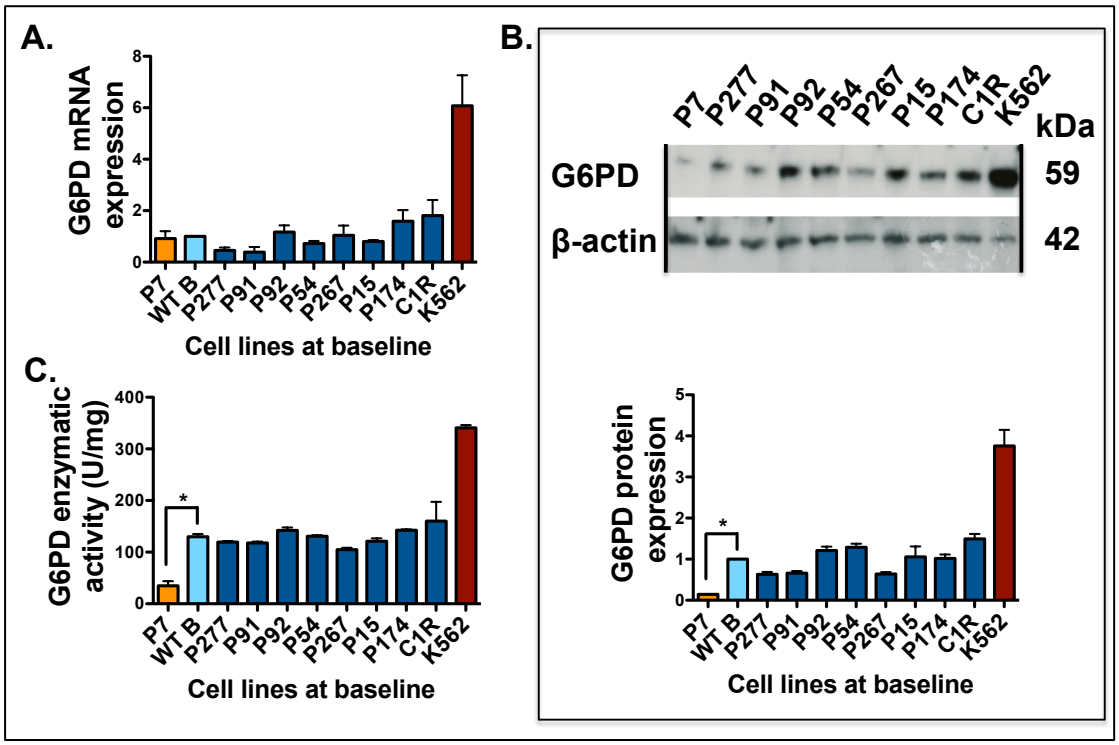


Figure 3-8 *G6PD* expression and enzymatic activity at baseline. The G6PD deficient B cell line P7 (shown in orange), the G6PD wild type B cell lines P277, P91, P92, P54, P267, P15, P174 and C1R (shown in blue) and the erythroid cell line K562 (shown in red) were used to assess *G6PD* expression and enzymatic activity at baseline. The average of the wild type cell lines is shown in light blue and is labelled WT B. *G6PD* mRNA and protein expression levels are shown here normalised to the average of the WT. **(A)** RT-qPCR was employed to assess *G6PD* mRNA expression and **(B)** western blotting (top) was employed to assess the G6PD protein expression. G6PD protein levels were normalised against β-actin and subsequently against the average WT B; the ImageJ software was used for quantification (bottom). **(C)** The same cells were also tested for G6PD enzymatic activity, which were measured in units per milligram of total protein. Mean and S.E.M are shown for n=3. Student's t-test was employed to show the significance of the P7 cell line expression and activity compared to the average of the wild type cell lines.

The G6PD deficient and wild type cell lines were tested for mRNA expression by RT-qPCR and protein expression by western blotting that was quantified using the ImageJ software. As shown in Figure 3-8A, the *G6PD* mRNA levels are comparable between G6PD Brighton and normal B cell lines; interestingly, *G6PD* mRNA levels are 6-fold higher in the erythroid K562 cells compared to B cells, possibly reflecting the essential role of G6PD in erythroid cells for protection against oxidative stress. Furthermore, these findings suggest that the specific coding mutation in G6PD Brighton does not affect production or stability of mRNA. However, looking at the

protein expression levels (Figure 3-8B), the G6PD protein levels in the G6PD deficient cells are 5-10 fold lower than those of the normal B cells and 25 fold lower than the erythroid cells, suggesting G6PD Brighton is associated with either ineffective protein translation or decreased protein stability. The higher expression of the G6PD protein in the K562 erythroblastoid cell line, as mentioned above, is expected as G6PD plays a key role in the erythroid lineage. Finally, the G6PD protein enzymatic activity was measured by a spectrophotometric assay based on the rate of production of NADPH, which unlike NADP has a peak of UV light absorption at 340 nm. As expected, G6PD enzymatic activity is lower (approximately 3-fold) in G6PD Brighton than in their normal counterparts (Figure 3-8C). It should be noted that G6PD enzymatic activity in G6PD Brighton red cells is about 10% of normal (1.7U/gr Hb as measured in Hammersmith Hospital diagnostic lab (normal range 7-10U/gr H); Table 2-2), consistent with WHO class I, i.e. severely deficient variant. The higher activity that I have found in the G6PD Brighton B cell line is likely to be due to the fact that B cells are nucleated cells and thus have the ability to constantly produce G6PD. By contrast, it is well established that in normal enucleated mature red cells, G6PD activity declines with time reaching low levels in aged red cells (Jansen et al., 1985; Piomelli et al., 1968).

3.5 Impact of HDACIs on G6PD deficient cells

In the next set of experiments I tested in more detail the effect of NaBu on *G6PD* mRNA, protein and enzymatic activity in the deficient G6PD Brighton cell line and the wild type (P277) B cell line. I found that *G6PD* mRNA levels, as assessed by RT-qPCR, increased in the WT line in a time-dependent manner, as shown in Figure 3-9A. Importantly, the same time-dependent effect was also observed for the G6PD deficient cell line with the first increase detected at 8h and a 10-fold increase observed at 24h. Analysis of the G6PD protein levels by western blot and using an antibody specific against G6PD, showed that consistent with the mRNA, G6PD protein expression increases in both cell lines in a time-dependent manner showing a 5-fold increase at 36h in the G6PD deficient cell line (Figure 3-9B).

Furthermore, following NaBu treatment for 8h, 16h, 24h and 36h, the enzymatic activity of the G6PD protein was assessed in G6PD deficient and wild type cell lines. A time-dependent increase in the protein activity of both cell lines was

observed. The most striking observation is that the G6PD deficient line achieved normal levels of activity at 24h (Figure 3-9C).

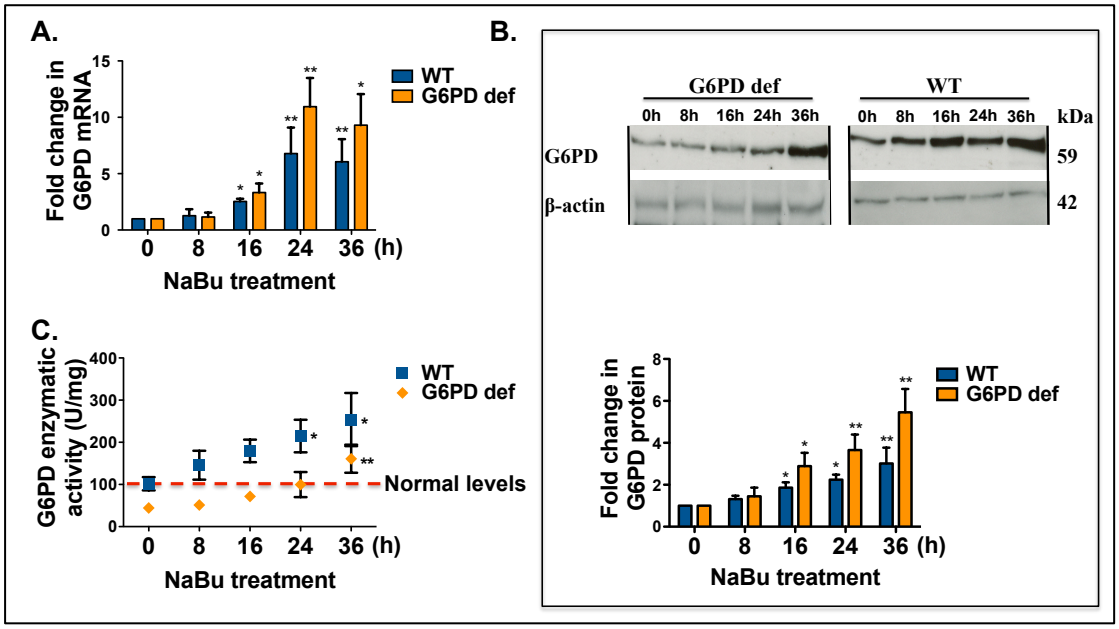


Figure 3-9 NaBu increases *G6PD* mRNA and protein expression and restores enzymatic activity to normal in cell lines. (A) Treatment of the G6PD deficient (P7) and wild type (P277) cell line with 3mM NaBu gradually increases *G6PD* mRNA levels as assessed by RT-qPCR and **(B)** G6PD protein levels assessed by western blotting (top; representative of 1 out of 3 independent experiments is shown) and quantified using the ImageJ software (bottom) show a time-dependent increase of the G6PD protein. **(C)** Treatment of the cell lines also increases G6PD enzymatic activity, which is restored to normal levels in the G6PD-deficient cell line. Mean and S.E.M. are shown for n=3. Student's t-test was performed.

To confirm these results, I repeated the same experiment using an alternative HDACI. In this case I treated the cells over the same time course with 4μM SAHA, instead of NaBu. Figure 3-10 presents the effect of SAHA on *G6PD* expression and enzymatic activity. Similarly to NaBu, treatment with SAHA increases the mRNA (Figure 3-10A) and protein (Figure 3-10B) levels of G6PD in both the wild type and the G6PD deficient cell lines. Specifically for the G6PD deficient cell line, *G6PD* mRNA and protein expression increase by 15-fold and 6-fold, respectively after 36h of treatment. In terms of the G6PD enzymatic activity (Figure 3-10C), this also increases in a time-dependent manner in both cell lines and indeed, in G6PD Brighton cells, it exceeds the normal levels after 36h of treatment with SAHA.

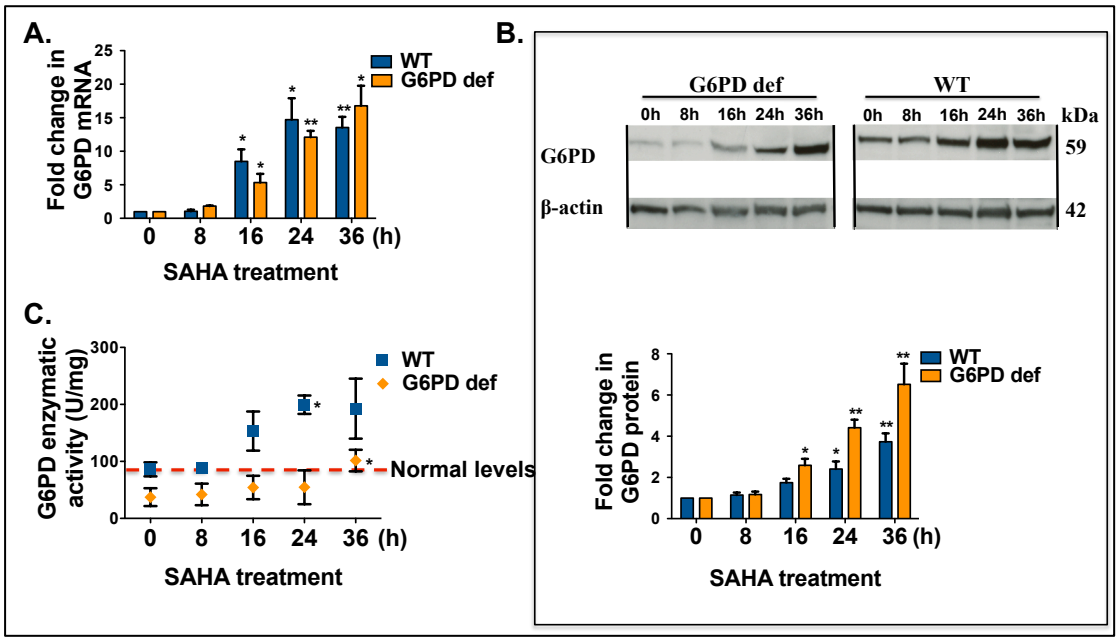


Figure 3-10 SAHA increases *G6PD* mRNA and protein expression and restores enzymatic activity to normal in cell lines. (A) Treatment of the G6PD deficient (P7) and wild type (P277) cell line with 4μM SAHA gradually increases the *G6PD* mRNA expression assessed by RT-qPCR and **(B)** G6PD protein levels assessed by western blotting (top; representative of 1 out of 3 independent experiments is shown) and quantified using the ImageJ software (bottom) show a time-dependent increase of the G6PD protein. **(C)** Treatment of the cell lines also increases the G6PD enzymatic activity and is restored to normal levels in the deficient cell line. Mean and S.E.M. are shown for n=3. Student's t-test was performed.

3.6 Importance of protein synthesis on the upregulation of G6PD

Having determined (Figure 3-6) and described (Figure 3-9 and Figure 3-10) the selective effect of HDACIs on *G6PD* gene expression in detail, it is necessary to establish if HDACIs act on *G6PD* directly. Specifically, I next aimed to answer whether *G6PD* is upregulated by NaBu due to direct effect of the drug on *G6PD* transcription, or through an indirect pathway, e.g. upregulation of a TF that secondarily regulates *G6PD* transcription. Therefore, I used CHX, a protein synthesis inhibitor that interferes with translational elongation at the ribosome. In the presence of CHX, although protein production is blocked, the cell is still able to produce mRNA; thus, in the presence of both NaBu and CHX any changes in mRNA levels would reflect a direct effect of HDAC inhibition.

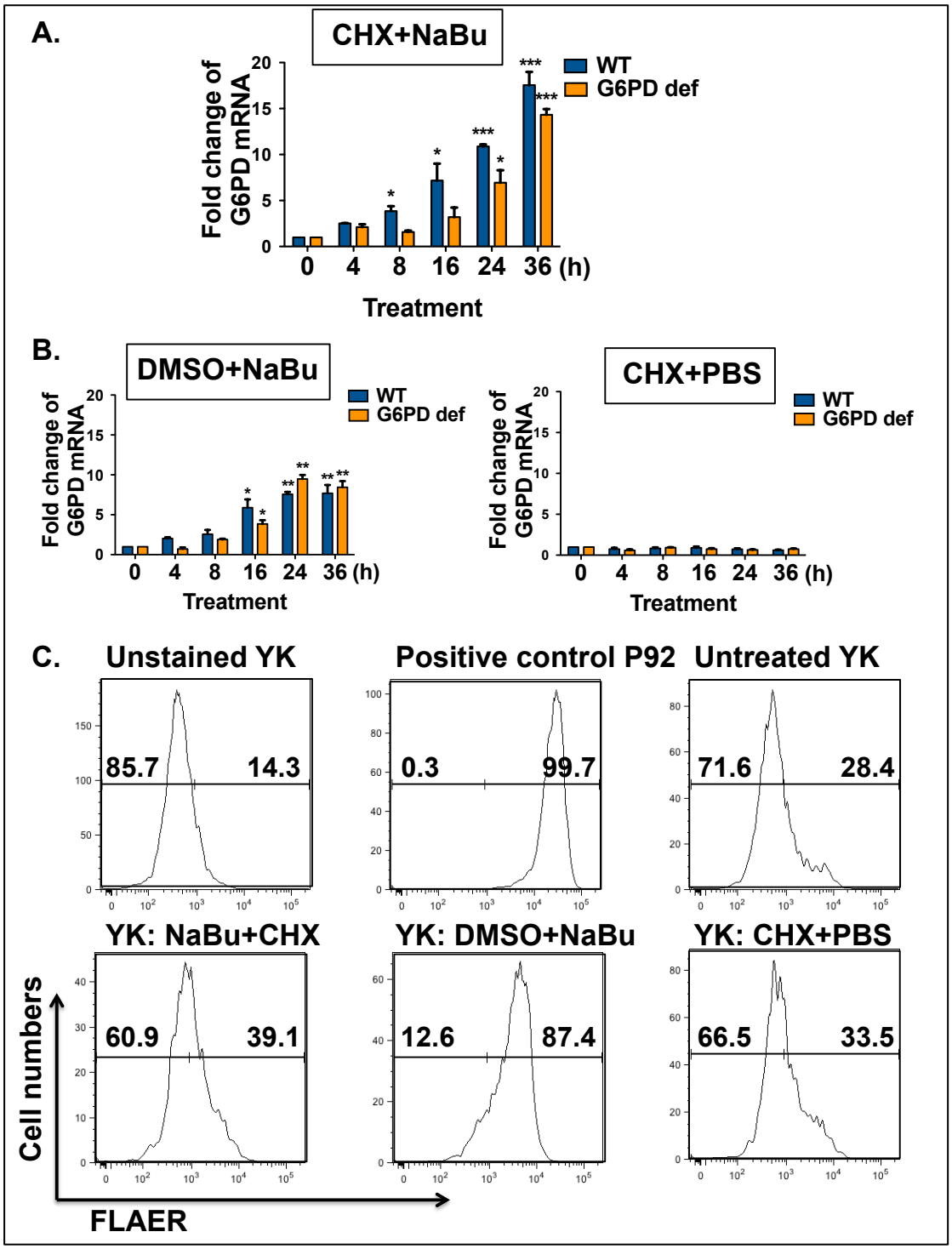


Figure 3-11 Translation-independent upregulation of *G6PD* by NaBu treatment. **(A)** The wild type (P277) and G6PD deficient (P7) cell lines were co-treated with 3mM NaBu together with 10µg/ml CHX and the mRNA levels of *G6PD* were assessed by RT-qPCR. **(B)** Control experiments treating with NaBu and DMSO or CHX and PBS are shown. Mean and S.E.M. are shown for n=3. Student's t-test has been performed. **(C)** As a positive control for CHX activity, flow-cytometric analysis of GPI expression on the PIGM-deficient cells YK, which were either co-treated or treated with one of the two drugs. One out of three experiments is shown.

Co-treatment of the wild type (P277) and G6PD deficient (P7) cell lines with CHX and NaBu during a time course of 4h, 8h, 16h, 24h and 36h (Figure 3-11A) showed a similar effect to that of the treatment with NaBu alone (Figure 3-11B). The fact that *G6PD* expression increase persists after incubation of the cells with CHX provides evidence that NaBu directly upregulates *G6PD* in a translation-independent manner. This finding is also supported by the observation in this experiment that *G6PD* upregulation happens as early as 4h post-treatment (Figure 3-11A), eliminating the window for secondary effects to take place. Control experiments were also conducted and involved treatment with NaBu and DMSO, instead of CHX and treatment with CHX and PBS, instead of NaBu, to show that NaBu was active and that CHX alone does not change the expression of *G6PD,* respectively (Figure 3-11B). To also establish active CHX, I co-treated the PIGM-deficient cell line YK with NaBu and CHX to show that FLAER expression on the cell surface is inhibited (Figure 3-11C).

3.7 Conclusions

Taken together, the evidence of this study shows that HDACIs can selectively upregulate *G6PD* expression among all the genes of the GPPP. Using wild type B cells and a cell line derived from a G6PD deficient individual (G6PD Brighton), I was able to show that HDACIs upregulate *G6PD* expression at the mRNA and protein levels and that most importantly they increase the G6PD enzymatic activity, which is restored to normal levels after 24h and 36h with NaBu and SAHA, respectively in the G6PD deficient cell line. The novel findings of this study have demonstrated that this upregulation takes place as early as 4h after treatment at the mRNA level and it is the result of a direct effect by the HDACIs. Contrary to my initial hypothesis, the other genes of the GPPP are not significantly affected by treatment with HDACIs, including NaBu and SAHA, in both normal and deficient cell lines.

4 Results II: Glycolytic enzyme gene expression in *in vitro* generated erythroid precursors

4.1 Introduction

Erythropoiesis is the process that generates mature erythrocytes. HSCs differentiate through intermediary forms of progenitor and precursor cells that undergo terminal differentiation and maturation to give rise to RBCs (Ji et al., 2011; Lodish et al., 2010). As described in 1.3.2 in greater detail, the erythroid lineage is derived from the MEPs, which can further differentiate and commit to either the erythroid or megakaryocytic lineages. Downstream erythroid progenitors are the BFU-Es, which differentiate into CFU-Es (Hattangadi et al., 2011; Lodish et al., 2010; Wu et al., 1995). In Figure 1-7, surface markers that can be used to identify BFU-E/CFU-E and the morphologically distinct erythroid precursors are shown. As erythroblasts mature, they pass through the precursor stages of proerythroblasts, basophilic normoblasts and early and late polychromatic normoblasts. These precursors decrease in size, undergo chromatin condensation, synthesise more haemoglobin and show altered gene expression patterns (Figure 1-7). To terminally differentiate, late polychromatic normoblasts lose their nucleus and give rise to mature RBCs.

The main function of the RBCs is to carry haemoglobin in the bloodstream in high concentrations, facilitating gas exchange in the lungs and the tissue capillaries. For this purpose, the RBC needs a supply of energy and also a source of reducing power, which is generated by the GPPP (Castagnola et al., 2010). Consequently, mutations in most of the enzymes of the GPPP are associated with haemolytic anaemia. Since G6PD deficiency is associated with a clinical phenotype, i.e., haemolytic anaemia, which is restricted in the erythroid lineage it is important to investigate the effect of HDACIs in an erythroid system.

4.1.1 Aim of the chapter

The aim of this chapter is to further investigate the effect of HDACIs on the expression of GPPP enzyme genes and in particular of *G6PD*, using an *in vitro* erythroid differentiation system that recapitulates *in vivo* late erythroid differentiation. Furthermore, this system is used to assess the effect of HDACIs on *in vitro*- generated primary erythroid precursors from individuals with G6PD deficiency.

4.1.2 Experimental plan

An *in vitro* two-phase liquid culture differentiation system was optimised to recapitulate the erythroid differentiation of mature erythrocytes starting from either PBMCs, CB-$CD34^+$ cells or G-CSF-mobilised PB haematopoietic stem and progenitor $CD34^+$ cells (GCSF-$CD34^+$). The protocol used was modified from previous published protocols (Ohene-Abuakwa, 2005; Ronzoni et al., 2008) and is described in 2.3.

Upon optimisation of the two-phase liquid culture erythroid differentiation system, I used it to treat purified proerythroblasts with HDACIs and determined their effects on GPPP expression. This system was used to determine the effects on both normal and G6PD deficient human cells obtained from patients.

4.2 Characterisation of the *in vitro* erythroid differentiation system

The results obtained from studying B cells, as described in Chapter 3, demonstrate that HDACIs increase *G6PD* mRNA and protein expression in WT and in G6PD deficient cell lines, resulting in restoration of the enzymatic activity of the G6PD Brighton B cell line to WT levels. However, since G6PD deficiency is associated with a haematological clinical phenotype, restricted to the erythroid lineage, it is important to show that the same effect can be achieved in erythroid cells. For this purpose, an *in vitro* two-phase liquid culture differentiation system was used to recapitulate erythroid development of mature erythrocytes from progenitor cells. The system was optimised for the use of either whole adult PBMCs, GCSF-$CD34^+$ cells or CB-$CD34^+$ cells.

Erythroid differentiation of whole PBMCs ($2x10^6$/ml) or $CD34^+$ cells ($5x10^5$ cells/ml), as described in 2.3, was performed using serum-free medium supplemented with a cytokine cocktail of IL-3 (10ng/ml), SCF (100ng/ml) and EPO. This system allows the gradual differentiation from $CD34^+$ progenitor cells ($CD34^+CD36^-CD71^-GlyA^-$) to BFU-E/CFU-E ($CD34^+CD36^+CD71^{low}GlyA^-$) and then sequentially to the precursors: proerythroblasts ($CD34^{low}CD36^+CD71^+GlyA^-$), basophilic normoblasts ($CD34^-CD36^+CD71^+GlyA^{low}$), polychromatic normoblasts ($CD34^-CD36^+CD71^+GlyA^+$) and finally orthochromatic normoblasts ($CD34^-CD36^+CD71^-GlyA^+$). The system consists of two phases, each lasting for 7 days. Phase 1 uses a low EPO concentration (0.5U/mL) to allow initial erythroid cell expansion and

differentiation. At day 7 $CD36^+$ cell selection is performed by immunomagnetic bead selection to isolate all erythroid progenitors. Cells are then replated into a high EPO (4U/mL) media to allow further differentiation. Cells are analysed by flow cytometry and cytospin staining at both day 7 and day 14.

The system was first used with whole adult PBMCs. At baseline only a very small proportion of the whole PBMCs display erythroid specific cell surface markers (approximately 2.7% $CD71^+GlyA^+$; Figure 4-1A). However, by the end of phase 1, approximately 60% of the live cells are $CD36^+$ erythroid cells. Of these, 33% are $CD36^+CD71^+GlyA^-$ corresponding to proerythroblasts and 62% are characterised as $CD36^+CD71^+GlyA^{low}$ basophilic normoblasts (Figure 4-1). After phase 2, the $CD36^+$ cells give rise to more mature erythroid stages; the majority are $CD36^+CD71^+GlyA^+$ polychromatic normoblasts and only 20% achieve differentiation as late as the orthochromatic normoblast's stage (Figure 4-1). During the erythroid differentiation of PBMCs only a very small minority of cells display myeloid or megakaryocytic characteristics as assessed by staining with anti-CD11b/CD14 and anti-CD61 respectively (Figure 4-1A).

In an attempt to produce greater numbers of erythroid cells, I used the culture system to differentiate GCSF-$CD34^+$ cells, as this is an easily accessible source of $CD34^+$ cells. Although I was able to induce erythroid differentiation on some occasions (Figure 4-2), most cultures led to the differentiation of myeloid cells (data not shown). Due to the lack of consistent erythroid differentiation I chose to not use the GCSF-$CD34^+$ cells in further experiments.

As an alternative to GCSF-$CD34^+$cells, I next tried normal CB, as this also offers an easily obtainable source of $CD34^+$ cells. The erythroid differentiation of $CD34^+$ cells demonstrated differences to the differentiation of whole PBMCs. At day 7, more than 80% of the live cells are $CD36^+$ erythroid cells compared to only 60% in the PBMC culture, suggesting that the $CD34^+$ cells have greater erythroid potential. Of note, as shown in Figure 4-2 and Figure 4-3, the GCSF-$CD34^+$ cells show more rapid erythroid differentiation than the CB-$CD34^+$ cells, with more rapid loss of CD34. After phase 2 of the $CD34^+$ (both CB and GCSF) erythroid differentiation culture, the majority of cells are basophilic normoblasts, whilst the other cells are mainly pronormoblasts and polychromatic normoblasts. Compared to the PBMC cultures, the $CD34^+$-differentiating cell cultures do not give rise to orthochromatic normoblasts by

day 14. This is likely to be because they represent a more immature starting population and therefore may require longer to achieve late-stage differentiation.

Overall, the system allows effective erythroid differentiation of both PBMCs and $CD34^+$ cells. In the case of the $CD34^+$ cultures, either originating from CB (Figure 4-3) or from G-CSF mobilised PB (Figure 4-2), the cell types obtained at day 7 and day 14 are at earlier stages than those obtained during the differentiation of PBMCs (Figure 4-1). This may be explained by the fact that in the case of the PBMCs cultures the differentiating cells are already committed to the erythroid lineage and thus enter differentiation earlier.

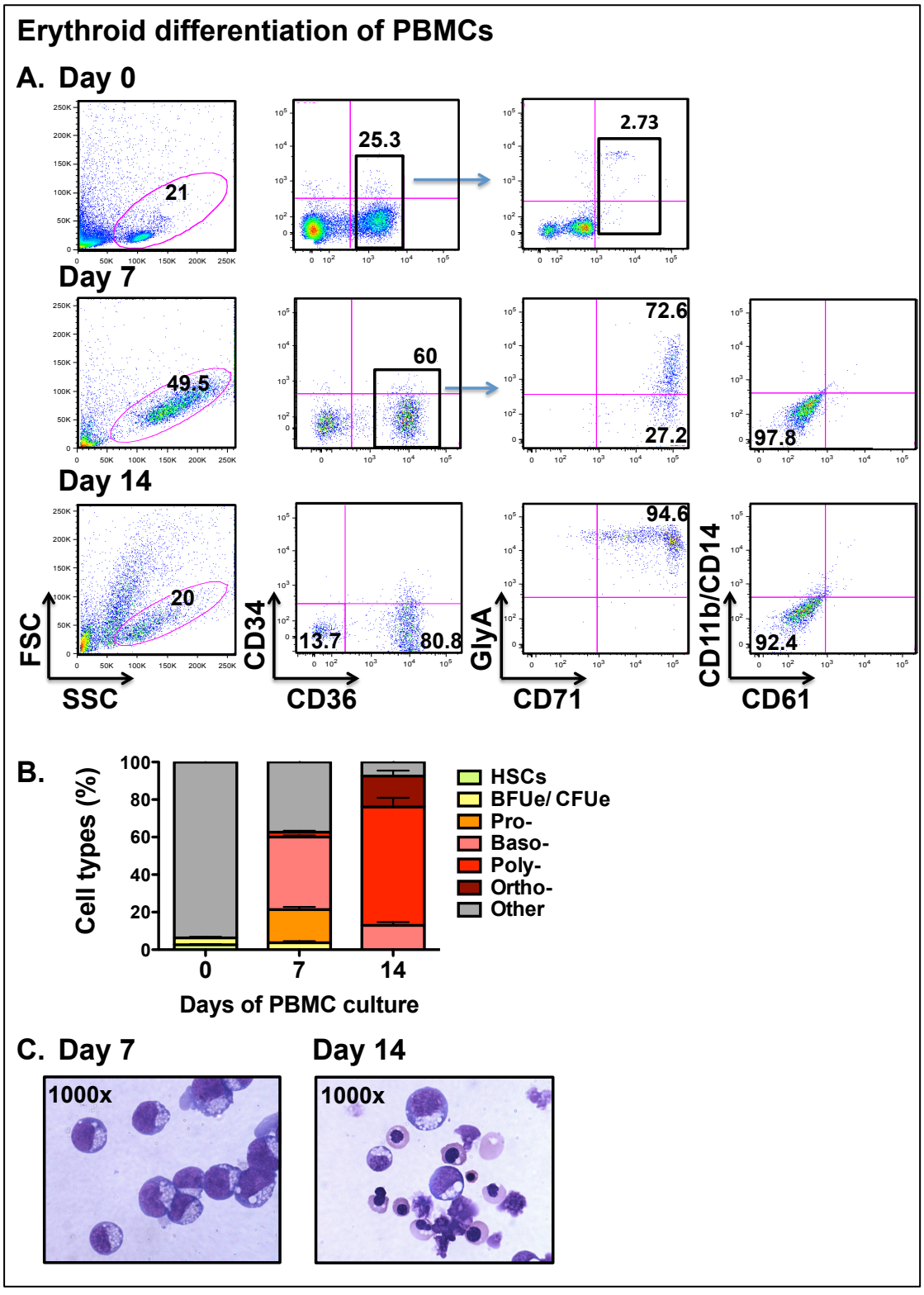

Erythroid differentiation of PBMCs
A. Day 0
21
25.3
2.73
Day 7
49.5
60
72.6
27.2
97.8
Day 14
20
13.7
80.8
94.6
92.4
FSC
SSC
CD34
CD36
GlyA
CD71
CD11b/CD14
CD61
B.
Cell types (%)
100
80
60
40
20
0
0
7
14
Days of PBMC culture
HSCs
BFUe/ CFUe
Pro-
Baso-
Poly-
Ortho-
Other
C. Day 7
Day 14
1000x
1000x

Figure 4-1 Erythroid differentiation of PBMCs. (A) Flow cytometric analysis and **(B)** quantification of the cell types obtained during PBMC-erythroid differentiation at days 0, 7 and 14 (n=6). Day 0: PBMCs are plated at a concentration of $2x10^6$ cells/ml and undergo the first phase of the erythroid differentiation, lasting for seven days. The erythroid progenitors present at day 0 constitute only 2.7% of the live gate. Day 7: At the end of Phase 1, approximately 60% of the live gate cells are erythroid ($CD36^+$), of which 33% are $CD36^+CD71^+GlyA^-$, i.e. proerythroblasts and 62% are $CD36^+CD71^+GlyA^{low}$, i.e. basophilic normoblasts. Upon $CD36^+$ selection (cells boxed), they are placed back in culture for the second phase of differentiation. Day 14: $CD36^+$ cells have further differentiated into polychromatic and orthochromatic normoblasts. **(C)** MGG staining of cytospins confirms the stage of differentiation and determines the cells' identity.

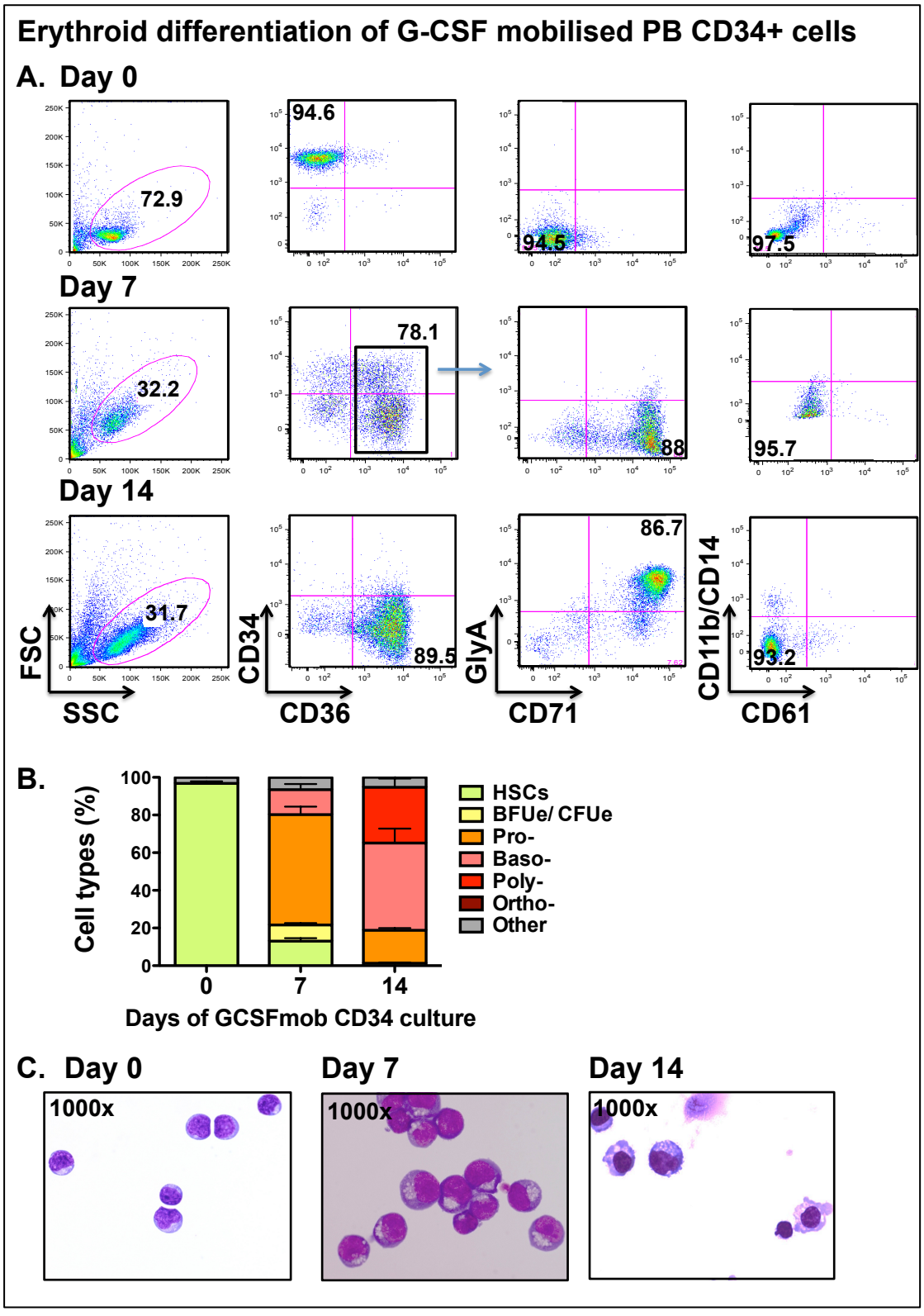

Erythroid differentiation of G-CSF mobilised PB CD34+ cells
A. Day 0
72.9
94.6
94.5
97.5
Day 7
32.2
78.1
88
95.7
Day 14
31.7
89.5
86.7
93.2
FSC
SSC
CD34
CD36
GlyA
CD71
CD11b/CD14
CD61
B.
Cell types (%)
100
80
60
40
20
0
0
7
14
Days of GCSFmob CD34 culture
HSCs
BFUe/ CFUe
Pro-
Baso-
Poly-
Ortho-
Other
C. Day 0
Day 7
Day 14
1000x
1000x
1000x

Figure 4-2 Erythroid differentiation of G-CSF mobilised PB $CD34^+$ cells. (A) Flow cytometric analysis and **(B)** quantification of the different cell types obtained during G-CSF mobilised $CD34^+$-erythroid differentiation at days 0, 7 and 14 (n=6). Day 0: The G-CSF mobilised $CD34^+$ cells are plated at a concentration of $5x10^5$ cells/ml and undergo the first phase of erythroid differentiation, following the same protocol as the PBMC erythroid cultures. A high proportion (>90%) is $CD34^+CD36^-$, indicating the high purity and potential to differentiate. Day 7: At the end of Phase 1, approximately 80%-90% of the live gate cells are $CD36^+$ erythroid cells. Approximately 5% of these $CD36^+$ cells are $CD34^+CD36^+$ BFU-Es and CFU-Es, 80% and 15% represent more differentiated forms of proerythroblasts and basophilic normoblasts, respectively. $CD36^+$ cells (boxed) are plated for Phase 2. Day 14: $CD36^+$ cells have further differentiated until the stage of polychromatic normoblasts. **(C)** MGG staining of cytospins confirms the stage of differentiation and determines the cells' identity.

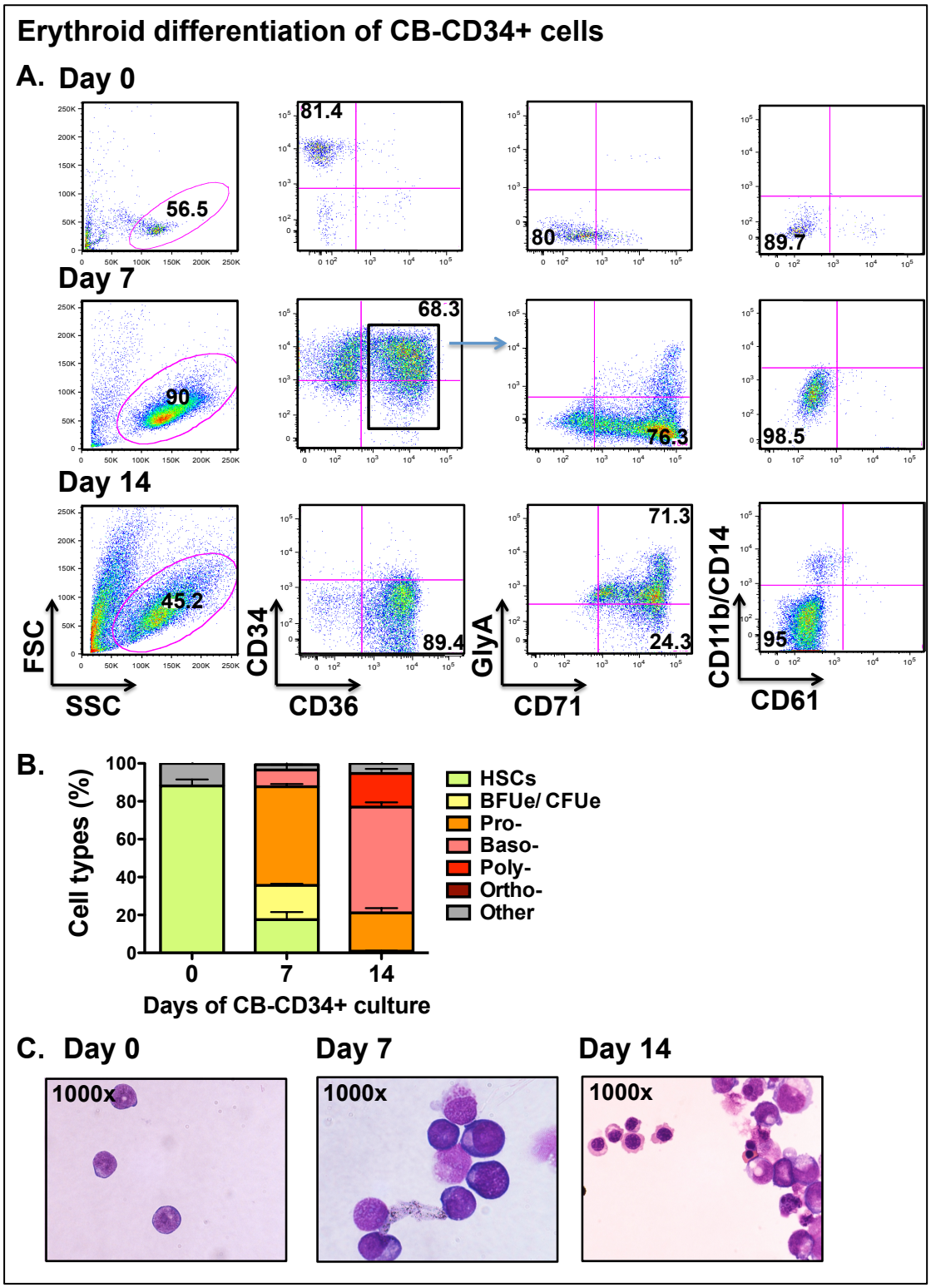

Erythroid differentiation of CB-CD34+ cells
A.
Day 0
Day 7
Day 14
56.5
81.4
80
89.7
90
68.3
76.3
98.5
45.2
89.4
71.3
24.3
95
FSC
SSC
CD34
CD36
GlyA
CD71
CD11b/CD14
CD61
B.
Cell types (%)
100
80
60
40
20
0
0
7
14
Days of CB-CD34+ culture
HSCs
BFUe/ CFUe
Pro-
Baso-
Poly-
Ortho-
Other
C.
Day 0
Day 7
Day 14
1000x
1000x
1000x

Figure 4-3 Erythroid differentiation of CB-CD34$^+$ cells. (A) Flow cytometric analysis and **(B)** quantification of the different cell types obtained during CB-CD34$^+$ - erythroid differentiation at days 0, 7 and 14 (n=6). Day 0: The CB-CD34$^+$ cells are plated at a concentration of 5x10^5 cells/ml and undergo the first phase of erythroid differentiation, following the same protocol as the PBMCs. A high proportion (>80%) is CD34$^+$CD36$^-$, indicating the high purity and potential to differentiate. Day 7: At the end of Phase 1, 70%-90% of the live gate cells are CD36$^+$ erythroid cells. Approximately 20% of these CD36$^+$ cells are CD34$^+$CD36$^+$ BFU-Es and CFU-Es, whereas 70% represent more differentiated proerythroblasts and 10% are basophilic normoblasts. CD36$^+$ cells (boxed) are plated for Phase 2. Day 14: CD36$^+$ cells have further differentiated until the stage of polychromatic normoblasts. **(C)** MGG staining of cytospins confirms the stage of differentiation and determines the cells' identity.

In terms of cell expansion, the number of erythroid cells obtained from the $CD34^+$ cultures is higher than the PBMCs cultures, as a greater proportion of $CD34^+$ cells have the potential for erythroid commitment. Overall, the expansion of $CD34^+$ cells is 10-fold by day 7 and 100-fold expansion by day 14. In contrast, PBMC expansion is only 80-fold from day 7 to day 14. It should be noted that during the PBMC differentiation it is not feasible to measure expansion of PBMCs as the initial culture contains a large proportion of cells that are committed to other lineages and do not give rise to erythroid cells.

Table 4-1 Cell numbers in two-phase liquid culture systems. [a]

	Day 0	Day 7: pre-selection	Day 7: post-selection	Day 14
PBMCs cultures	$1x\ 10^6$	$5x\ 10^5 \pm 0.37x10^5$	$1x\ 10^5 \pm 0.1x10^5$	$8x10^6 \pm 0.55x10^6$
CB-CD34 cultures	$5x\ 10^5$	$5x\ 10^6 \pm 0.45x10^6$	$4x\ 10^6 \pm 0.28x10^6$	$4x10^8 \pm 0.13x10^8$
GCSFmob-CD34 cultures	$5x\ 10^5$	$7x\ 10^6 \pm 0.23x10^6$	$6x\ 10^6 \pm 0.16x10^6$	$5x10^8 \pm 0.19x10^8$

[a] The values represent the average of n=6 experiments. Mean±SD is shown.

Further optimisation of the erythroid culture system involved comparison of fresh and frozen PBMCs (Figure 4-4). Frozen PBMCs maintained their ability to differentiate to erythroid cells. However differentiation proceeded more slowly in the frozen PBMCs; at day 7 approximately 55% of the $CD36^+$ cells in the fresh PBMCs erythroid culture were $CD71^+GlyA^{+/low}$ whilst only 6% of the $CD36^+$ cells obtained from the frozen PBMCs erythroid cultures were $CD71^+GlyA^{+/low}$. It is possible that amongst the circulating erythroid progenitors and precursors those further differentiated are more sensitive to the freezing-thawing procedure. Nevertheless, these results show that it is feasible to use frozen PBMCs in the two-phase erythroid culture system.

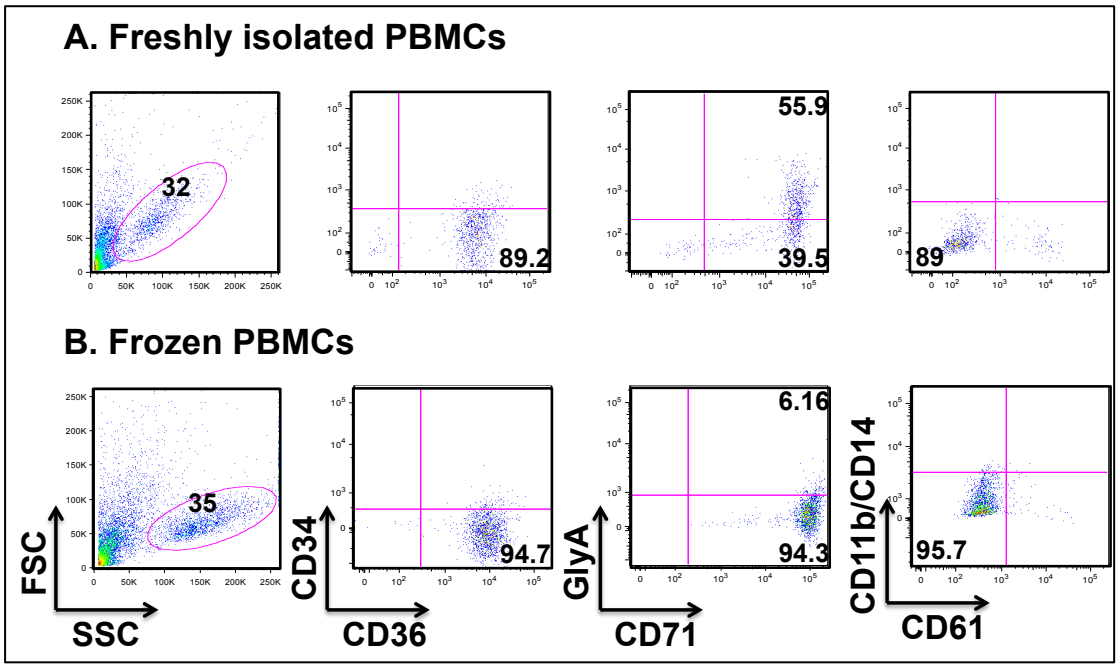


Figure 4-4 Comparison between same-donor fresh and frozen PBMCs in erythroid differentiation cultures. After cultured for seven days, erythroid differentiation of PBMC was assessed by flow cytometry. Frozen PBMCs maintain their ability to differentiate to erythroid cells, although differentiation proceeds slower than in fresh cell cultures.

Taken together, in this study I optimised an *in vitro* erythroid differentiation system, which allows differentiation of PBMCs and $CD34^+$ cells of different origins. Using this system I can successfully obtain erythroid precursor cells by the end of the first phase of culture and differentiated erythroid cells after the second phase of culture. Although I found differences in differentiation between alternate cell sources, these differences can be effectively controlled using control cells from equivalent sources.

This system is of great importance, as it allows me to establish the effect of HDACIs on primary human erythroid cells. In the subsequent studies in this thesis I have used adult PBMCs and CB-$CD34^+$ cells, which were treated with HDACIs and then assessed for the GPPP expression.

4.3 Effect of NaBu on erythroid differentiation *in vitro*

4.3.1 Effect of NaBu on cell number and differentiation of PBMCs

Having established conditions for erythroid differentiation of PBMCs and CB-$CD34^+$ cells, I proceeded to study the effect of NaBu on erythroid differentiation starting with cell number. Two alternate treatment schedules were used, both of which lasted for 7 days. In both cases PBMCs were initially plated in media supplemented with cytokines. Cells were then either treated at 16h with 0.5mM, 1mM or 3mM NaBu followed by cytokine and NaBu supplementation every 3 days, or treated at day 5 with 1mM or 3mM NaBu.

Assessment on day 7 showed that treatment with NaBu decreased cell numbers (Figure 4-5) and delayed differentiation under both schedules (Figure 4-6). In the untreated control cultures, approximately 41% of the cells at day 7 are $CD36^+$, of which 92% are $CD36^+CD71^+$. NaBu results in a decreasing proportion of $CD36^+CD71^+$ cells in a dose-dependent manner, consistent with delayed erythroid differentiation (Figure 4-6). It should be noted that decreased cell number cannot be attributed to reduced viability or reduced proliferation, as due to limiting cell numbers, appropriate experiments could not be conducted for this purpose.

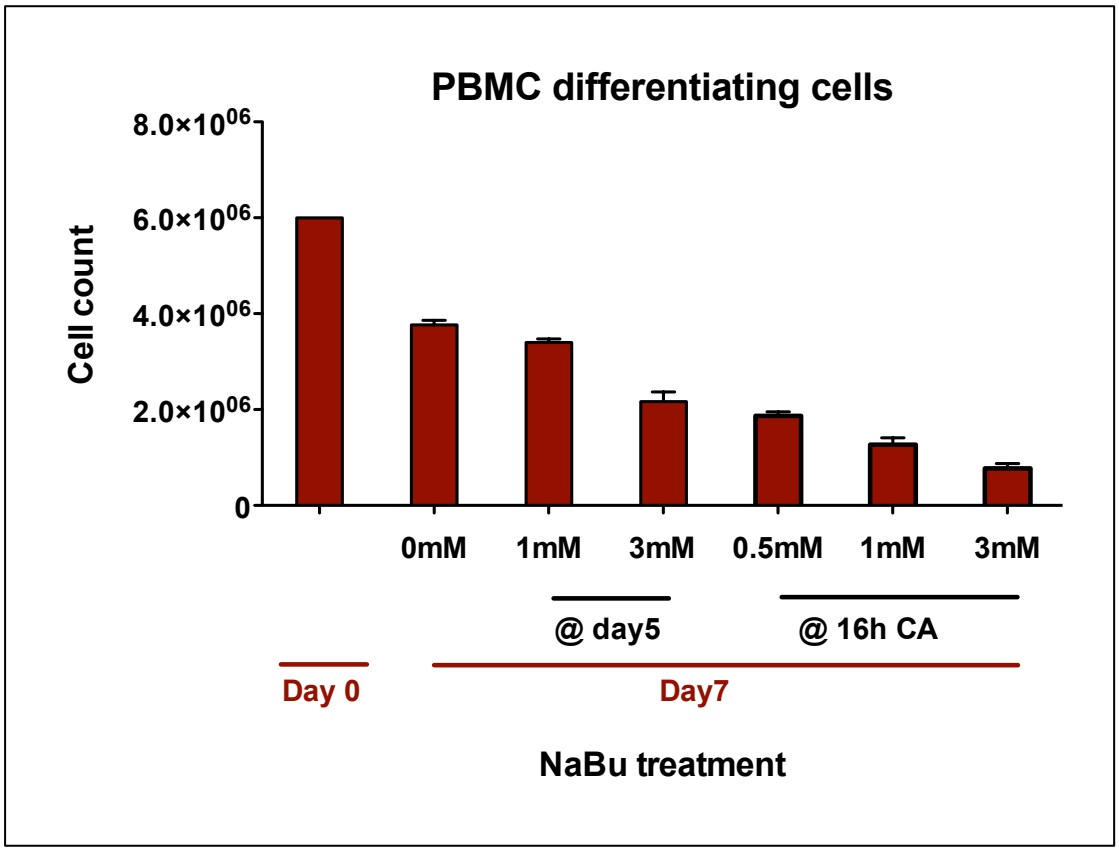


Figure 4-5 Cell count on day 0 and day 7 of PBMC-differentiating erythroid cultures upon NaBu treatment. Number of cells plated at day 0 and those harvested at day 7 before $CD36^+$ cell selection. The 2 different NaBu treatment methods involve treatment with 0.5mM, 1mM and 3mM NaBu after plating the cells for 16h with cytokines alone or, with 1mM and 3mM NaBu at the fifth day of the cultures. Mean and S.E.M. are shown for n=3.

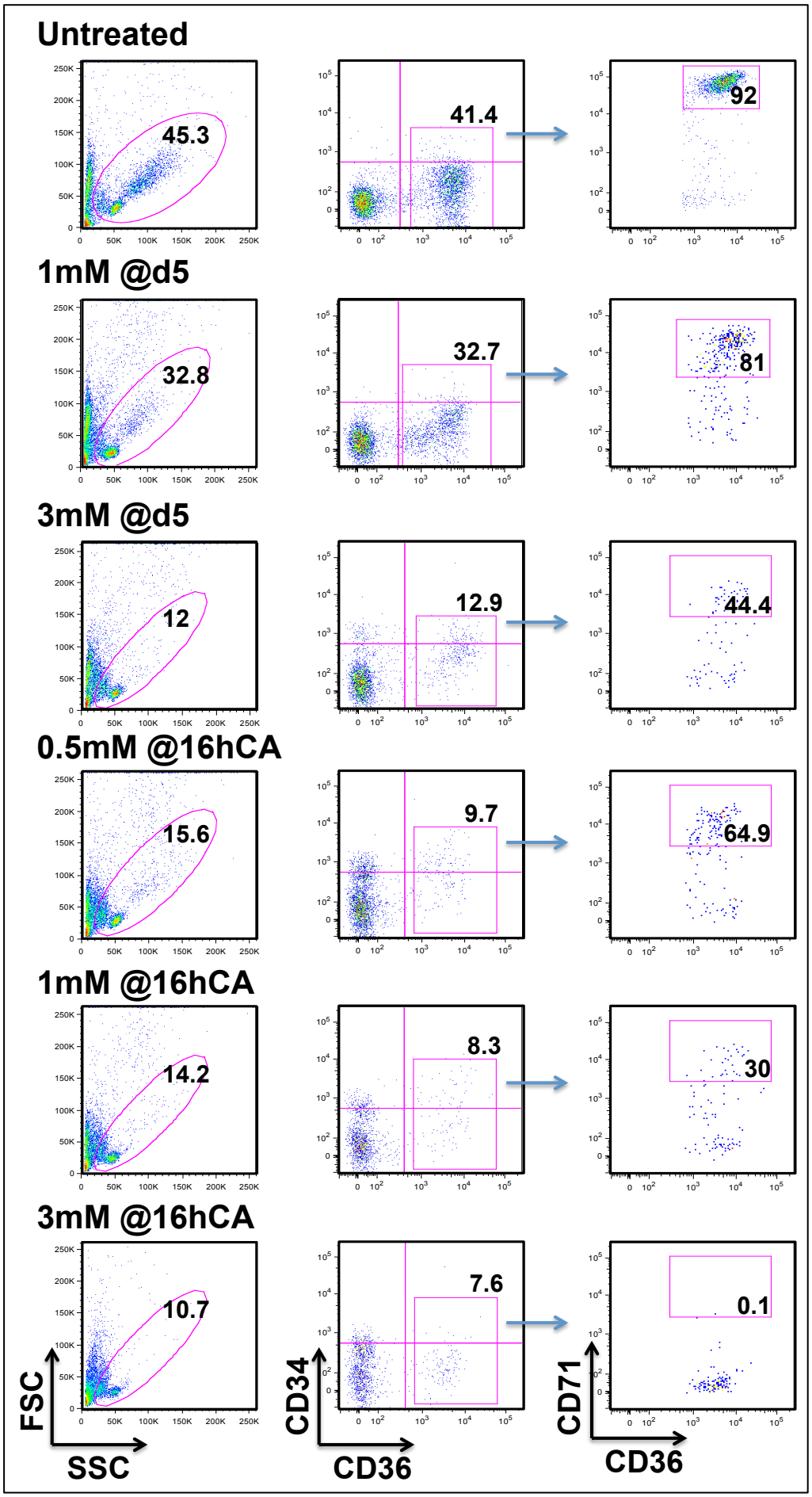

Untreated
45.3
41.4
92
1mM @d5
32.8
32.7
81
3mM @d5
12
12.9
44.4
0.5mM @16hCA
15.6
9.7
64.9
1mM @16hCA
14.2
8.3
30
3mM @16hCA
10.7
7.6
0.1
FSC
SSC
CD34
CD36
CD71
CD36

Figure 4-6 PBMC-differentiating cells treated with a series of different NaBu concentrations. Flow cytometric analysis of PBMC differentiating cells at day 7 of differentiation. The cultures are treated with 0.5mM, 1mM and 3mM NaBu after plating the cells for 16h with CA or, with 1mM and 3mM NaBu at day 5 of the cultures.

4.3.2 Effect of NaBu on cell number and differentiation of CD34$^+$

Similar to the PBMC erythroid differentiation cultures, the CB-CD34$^+$ erythroid differentiation cultures were used to generate CD36$^+$ erythroid cells to be tested for their *G6PD* expression upon treatment with NaBu at variable concentrations. Before doing so, I tested the effect of HDACIs, including NaBu, on the erythroid differentiation of CB-CD34$^+$ cells.

Treatment with 1mM NaBu on day 5 of the culture for a total of 48h did not have any effect on cell number (Figure 4-7) and erythroid differentiation as assessed on day 7 of the phase 1 culture (Figure 4-8). Prolonged treatment with 1mM NaBu throughout the 7 days of the culture (added 16 hours after plating the cells with CA) resulted in decreased cell number, although this was less pronounced than in the PBMC cultures (Figure 4-7). NaBu also promoted erythroid differentiation towards more differentiated cells, i.e. approximately 44% proerythroblasts and 49% basophilic normoblasts (Figure 4-8).

In addition, I treated the CB-CD34$^+$ cultures with SAHA and TSA to establish the effect that these have compared to NaBu. I treated with 1μM SAHA and 1nM TSA after plating the cells for 16h with cytokines alone, as described by Chaurasia and colleagues (Chaurasia et al., 2011). Although prolonged treatment with SAHA and TSA in this culture system resulted in reduced cell number (Figure 4-7), it did not seem to affect erythroid differentiation (Figure 4-8).

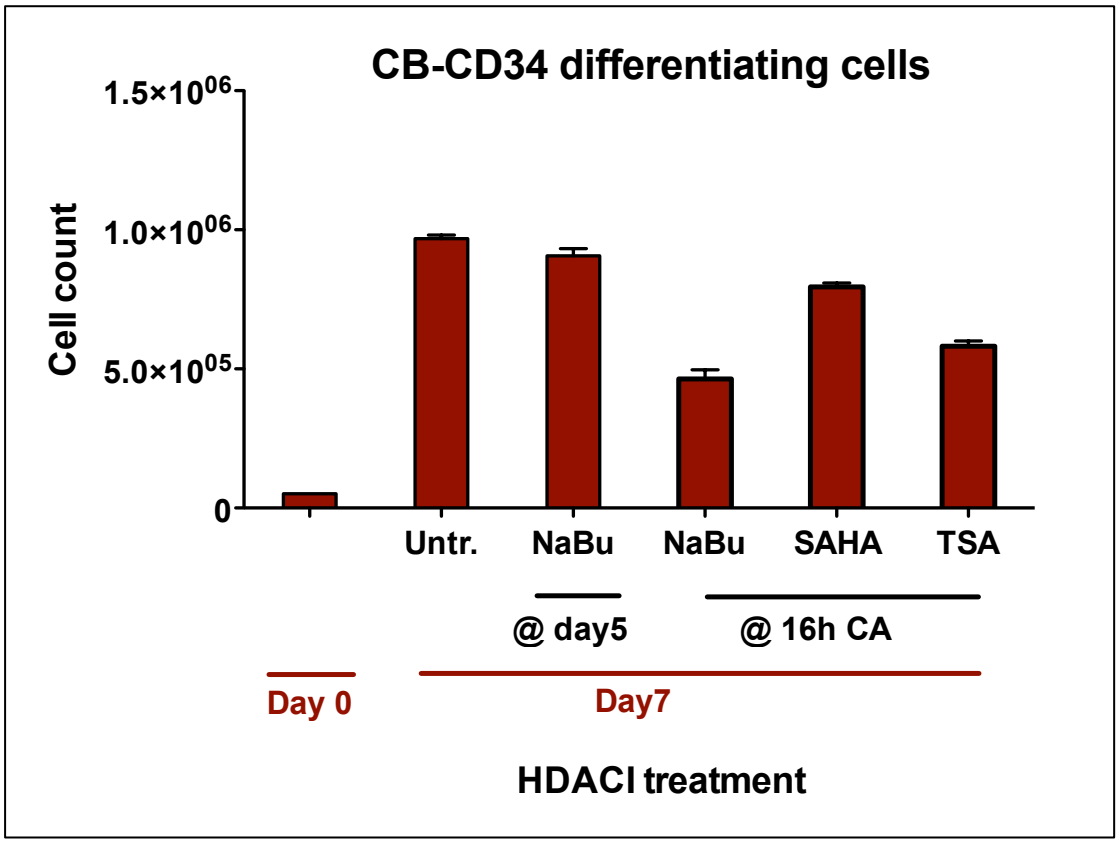


Figure 4-7 Cell count on day 0 and day 7 of CB-CD34$^+$ erythroid cultures upon HDACI treatment. Number of cells plated at day 0 and those harvested at day 7 before CD36$^+$ cell selection. The 2 different treatment methods involve treatment with 1mM NaBu at the fifth day of the cultures for 48h, or 1mM NaBu, 1µM SAHA and1nM TSA after plating the cells for 16h with cytokines alone. Mean and S.E.M. are shown for n=3.

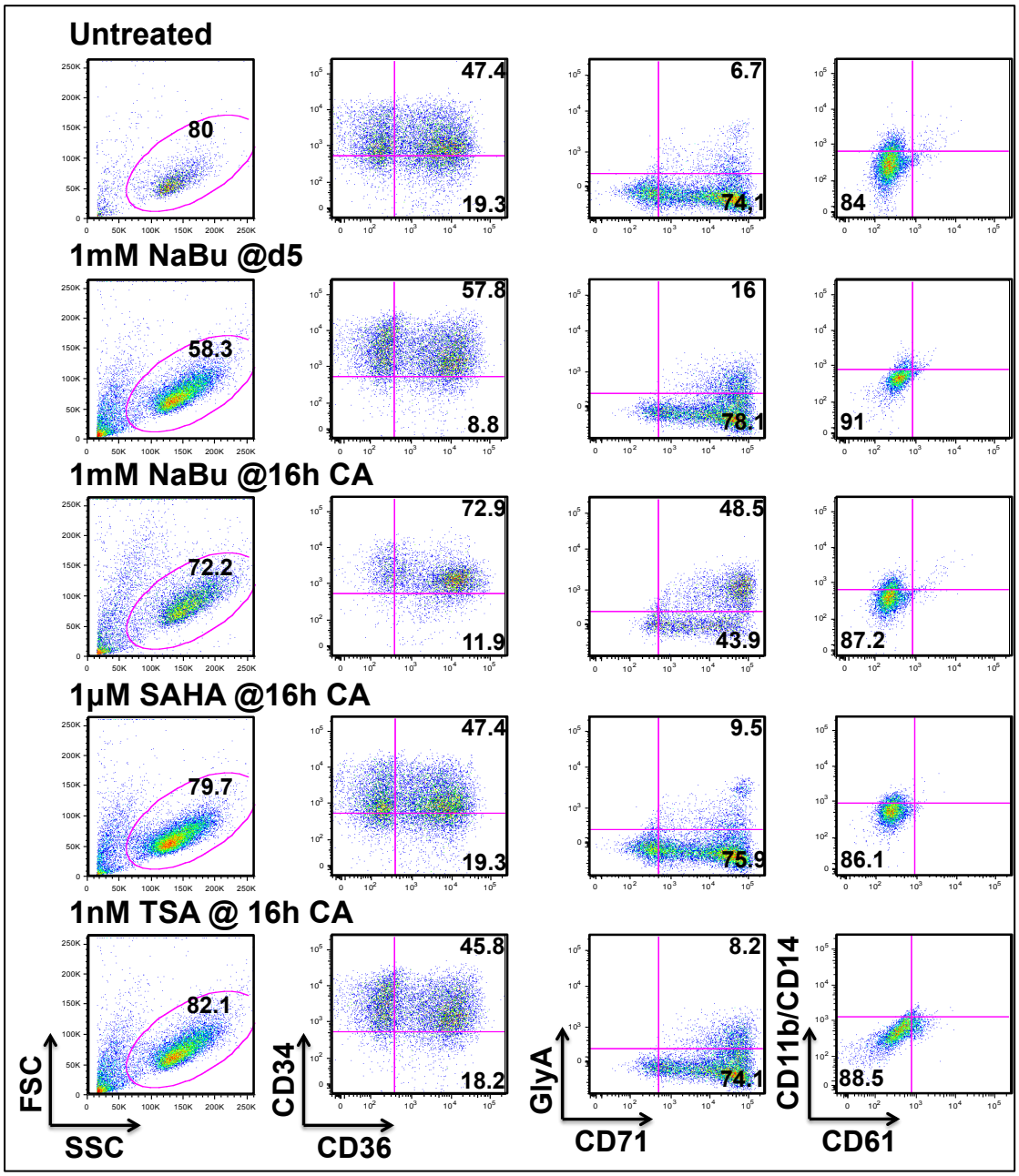


Figure 4-8 CB-CD34^{+}-differentiating cells treated with different HDACI concentrations. Flow cytometric analysis of CD34^{+}-differentiating cells at day 7 of differentiation. The cultures are treated with either 1mM NaBu at the fifth day of the cultures for 48h, or 1mM NaBu, 1µM SAHA and1nM TSA after plating the cells for 16h with CA.

4.3.3 Gene expression during erythroid differentiation

A recent RNA-sequencing study conducted by Wong and colleagues (Wong et al., 2011) on primary mouse fetal liver erythroid cells, gave insights into the baseline expression of genes during erythroid differentiation. In this study, researchers isolated erythroid cells from mice and then FACS-sorted different populations, representing the stages of erythroid cell differentiation. The populations were FACS-sorted based on their CD71 and Ter119 (equivalent of GlyA in mice) expression (Figure 4-9A); R1 corresponds to BFU-E/CFU-E, R2 to proerythroblasts and basophilic normoblasts, R3 to early polychromatic normoblasts, R4 to late polychromatic normoblasts and R5 orthochromatic normoblasts. The sorted cells were then subjected to RNA sequencing to examine changes in gene expression during erythroid differentiation. Of the 6,929 genes that were found to be differentially expressed between the R2 and R5 stage, the vast majority (6,455) were downregulated. This finding is consistent with the fact that during erythropoiesis erythroid cells undergo condensation of the nucleus, prior to enucleation.

Taking advantage of these data, which have been deposited on the Gene Expression Omnibus (GEO) database under the accession number GSE27893, I assessed the expression of the GPPP genes. Figure 4-9B shows that the expression of 5 GPPP genes (*G6PD*, *TPI*, *GPI*, *RPIA* and *PGLS*) decreases during erythroid differentiation, reaching a minimum level at the R4 late polychromatic normoblast differentiation stage and then marginally, yet consistently, increasing at the R5 orthochromatic normoblast stage. It is also shown how the expression of *β –actin*, the housekeeping gene used as an RT-qPCR reference gene, changes in comparison to *G6PD*, *TPI*, *GPI*, *RPIA* and *PGLS* GPPP genes that are regularly tested in this thesis.

Bioinformatic analysis of the available data therefore suggests that the gene expression of both the GPPP genes and also the reference gene *β-actin* changes during the different stages of differentiation. One of the issues that emerge from this finding is that comparison of gene expression upon treatment of erythroid cultures with HDACIs should be assessed only when cultures at the same-stage have been produced with and without HDACIs. As shown in Figure 4-6 and Figure 4-8, long exposures or high concentration treatment with NaBu results in altered erythroid differentiation patterns. For this reason, I decided to use short exposures (i.e. treatment with HDACI

on day 5 for 48h) and low concentrations (1mM of NaBu) to assess expression of *G6PD* and other GPPP genes in primary human erythroid cells.

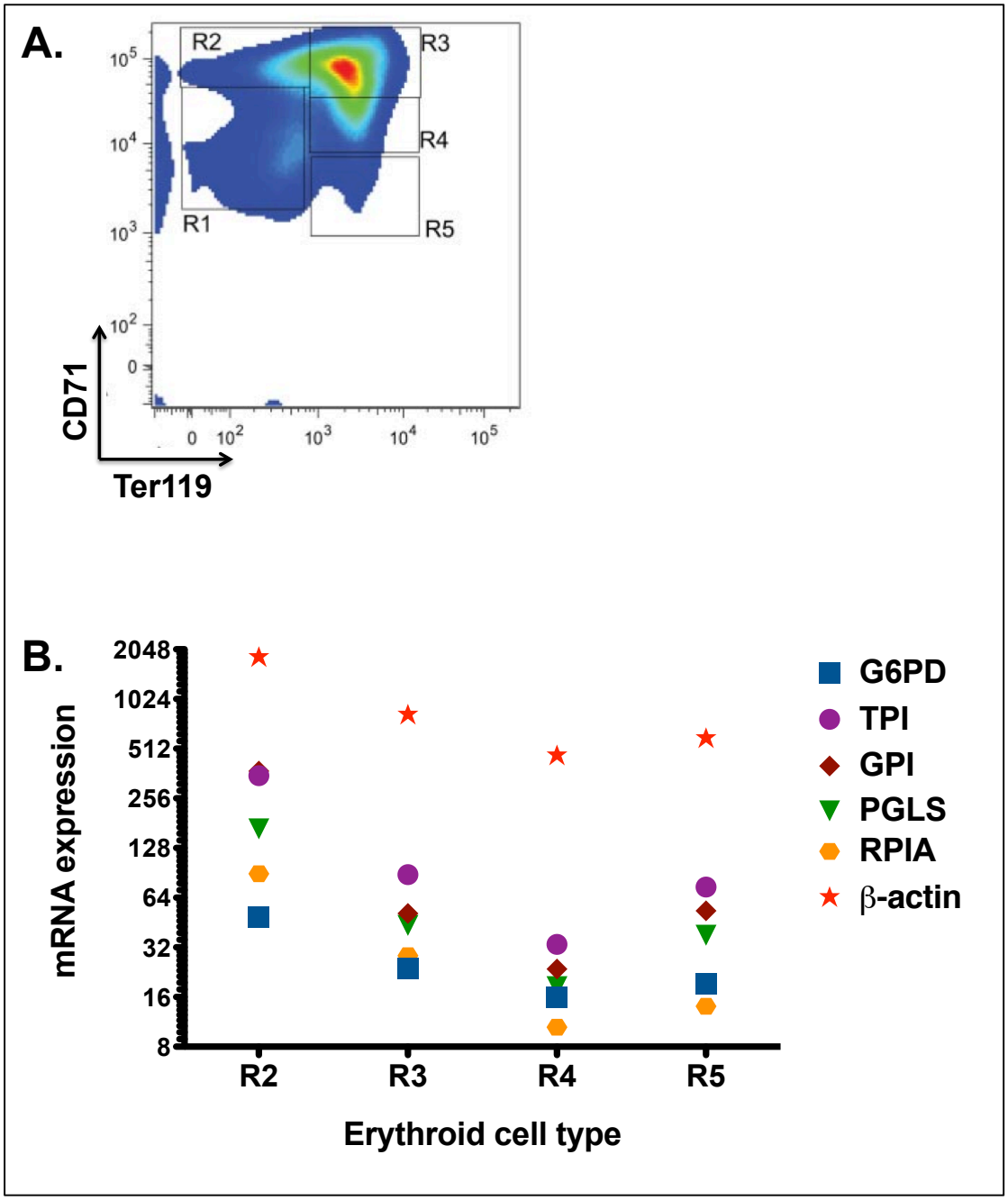


Figure 4-9 GPPP gene expression during erythroid differentiation in mice. (A) Flow cytometric analysis shows the five erythroid differentiation stages, as used by Wong et al (Wong et al., 2011) for further RNA sequencing analysis (GEO database accession number GSE27893). **(B)** Gene expression analysis during erythroid differentiation from stages R2 to R5 of the reference gene β-actin, in comparison to 5 genes of the GPPP (*G6PD*, *TPI*, *GPI*, *PGLS* and *RPIA*).

4.4 GPPP gene expression in normal erythroid precursors upon NaBu treatment

To test the effect of NaBu on normal primary erythroid cells, $CD36^+$ cells were purified from CB-$CD34^+$ erythroid differentiation cultures as shown in Figure 4-10, ensuring that erythroid cells obtained from NaBu-treated and control untreated cultures were at the same stage of differentiation.

$CD36^+$ cells were selected (post-selection purity >95% as assessed by flow cytometry) and then tested for *G6PD* mRNA expression (Figure 4-11A). Similarly to the B cell line data (Figure 3-6), as compared to untreated control, of all 17 GPPP genes (including the erythroid-specific *PKLR* gene), treatment with NaBu selectively increased *G6PD* mRNA expression levels to as much as 3.2-fold. In addition to mRNA, G6PD protein and enzymatic activity levels also increased by 3-fold and 2.5-fold, respectively (Figure 4-11B and C).

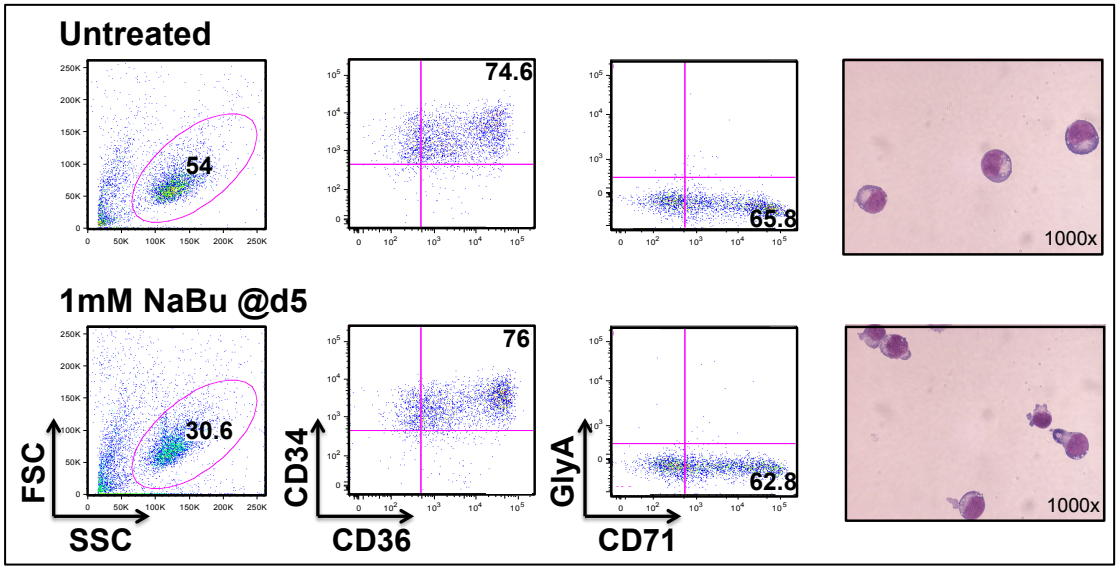


Figure 4-10 Same-stage differentiation upon NaBu in CB-$CD34^+$-differentiating cells. Flow cytometric and MGG staining analysis on CB-$CD34^+$ -differentiating cells at day 7 of the culture. Untreated and treated cells with 1mM NaBu for 48h are shown. Treatment with NaBu is shown not to affect the erythroid differentiation under these conditions.

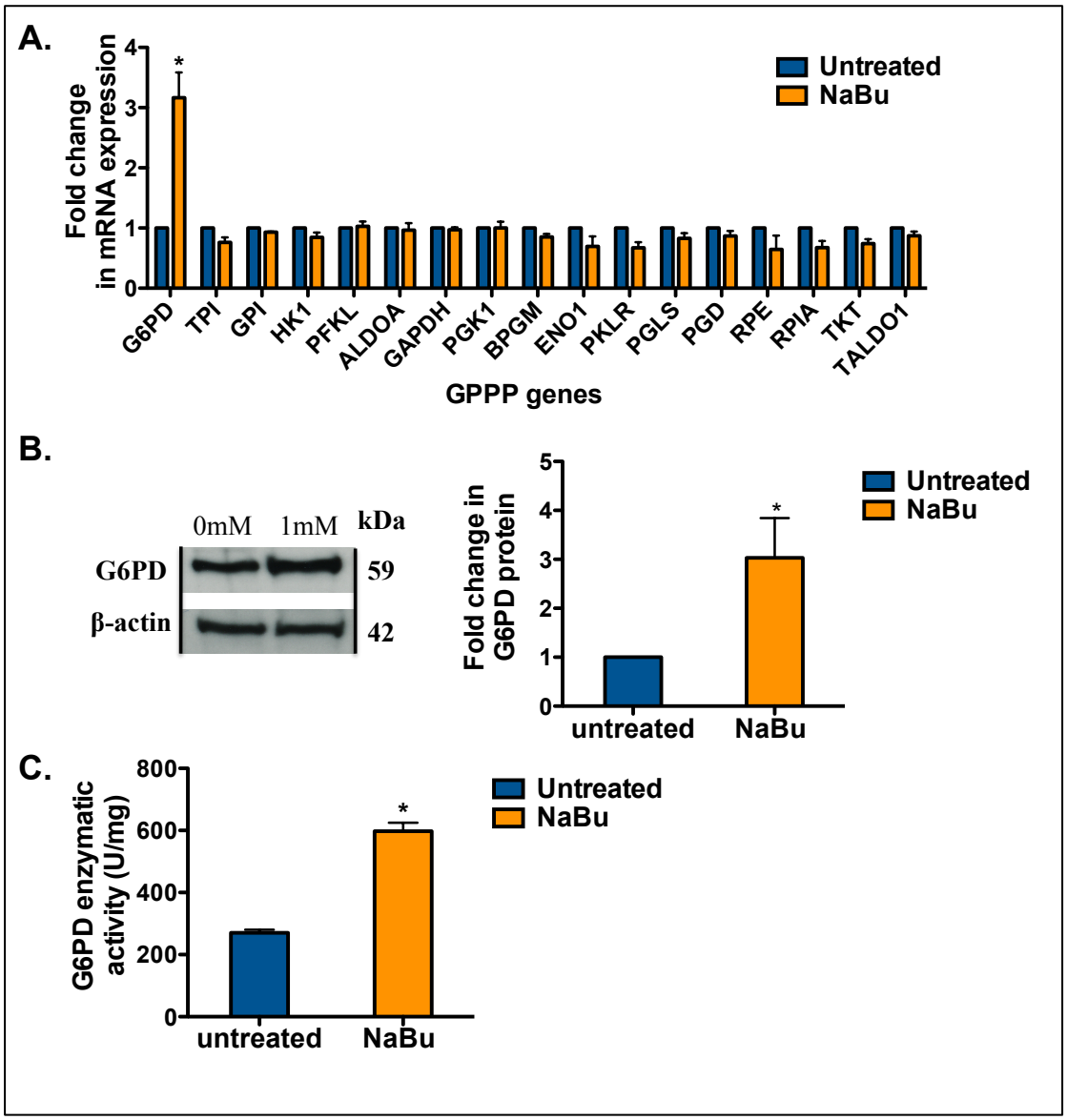


Figure 4-11 NaBu increases *G6PD* gene expression and enzymatic activity in CB-CD34$^+$-derived erythroid precursors. (A) NaBu treatment of CB-CD34$^+$-differentiating cells for 48h selectively increases *G6PD* mRNA as assessed by RT-qPCR at day 7 of the culture. **(B)** G6PD protein levels are assessed by Western blot and quantified by the ImageJ software. G6PD protein is shown to increase 3-fold upon NaBu treatment. **(C)** Treatment of the erythroid cells also increases G6PD enzymatic activity. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to the treated condition.

4.5 GPPP gene expression in patient erythroid cells upon NaBu treatment

Having established that NaBu increases *G6PD* gene expression and enzymatic activity in normal erythroid cells derived from CB-CD34^{+} cells, I next aimed to evaluate the potential clinical significance of this finding. Consequently, I tested the effect of NaBu on erythroid precursors derived from PBMCs of individuals with G6PD deficiency.

Firstly, I employed the erythroid differentiation assay using PBMCs from a patient with G6PD Brighton, a class I G6PD variant caused by an in-frame deletion on exon 13 and clinically characterised by CNSHA. As shown in Figure 4-12, G6PD Brighton PBMCs generate proerythroblasts and basophilic normoblasts by day 7 and polychromatic and orthochromatic normoblasts by day 14 as assessed by flow cytometry (Figure 4-12A) and MGG staining (Figure 4-12B). Overall, the differentiation of G6PD deficient samples is the similar to normal PBMCs as previously described in Figure 4-1.

PBMC-differentiating cells from G6PD Brighton as well as a healthy donor were treated with NaBu for 48h and analysed at day 7 (Figure 4-13). NaBu treatment is shown here not to affect the erythroid differentiation irrespective of the origin of the PBMCs. This finding also implies that the pathogenic mutation on the *G6PD* promoter does not affect the differentiation of erythroid cells. In fact, it is known that the defect is associated with the mature RBCs, which haemolyse due to the low G6PD enzymatic activity and not with earlier erythroid progenitors (Beutler, 2007; Cappellini and Fiorelli, 2008).

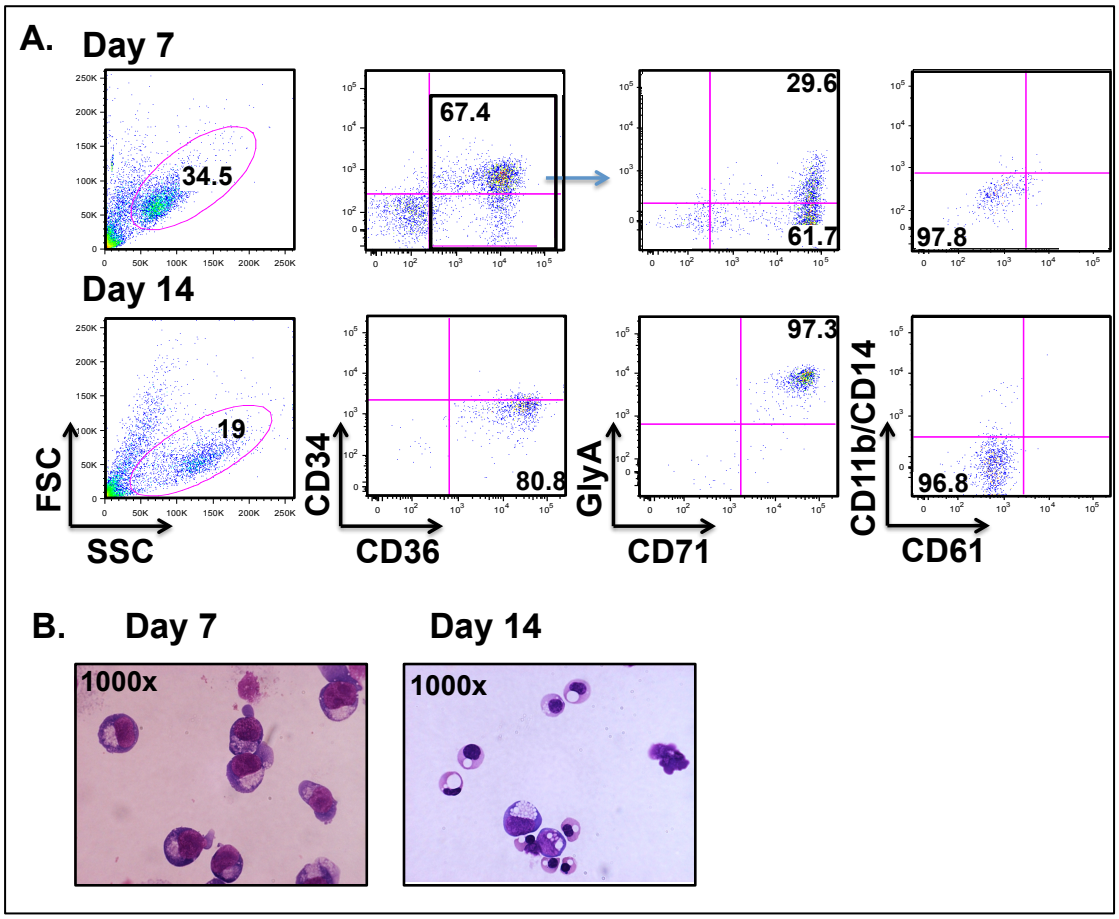


Figure 4-12 Erythroid differentiation of PBMCs isolated from G6PD Brighton. **(A)** Flow cytometric analysis during PBMC erythroid differentiation from a G6PD deficient patient (G6PD Brighton) at days 7 and 14. Day 7: At the end of Phase 1, approximately 60% -70% of the live gate cells are $CD36^{+}$ erythroid differentiated at the stages of proerythroblasts and basophilic normoblasts. Day 14: $CD36^{+}$ cells have further differentiated until the stage of polychromatic and orthochromatic normoblasts. **(C)** MGG staining of cytospins confirms the stage of differentiation and determines the cells' identity.

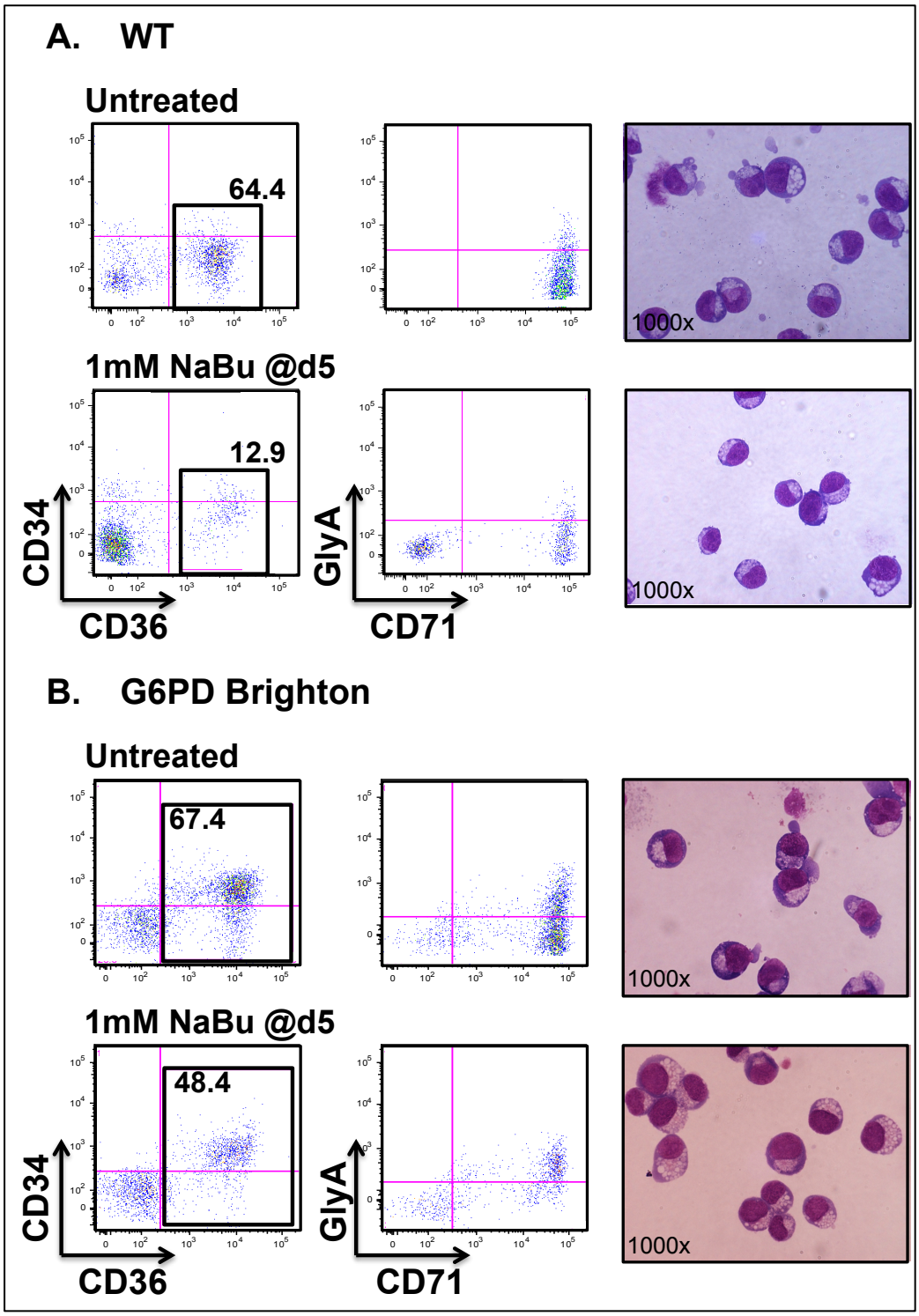


Figure 4-13 NaBu treatment of WT and G6PD Brighton PBMC-differentiating erythroid cells. (A) Normal and **(B)** G6PD Brighton PBMC-differentiating erythroid cells were treated at day 5 of phase 1 for 48h. Flow cytometric analysis and MGG staining confirms the stage of differentiation (proerythroblasts and basophilic normoblasts).

Having established that PBMCs isolated from G6PD deficient individuals can be differentiated into erythroid progenitors (Figure 4-12) and that their differentiation is not affected by short (48h) and low concentration (1mM) incubations with NaBu (Figure 4-13), I next sought to investigate *G6PD* mRNA and protein expression and enzymatic activity levels after treatment of these cells. For this purpose, I obtained PBMC-derived erythroid precursors from 3 class I deficient patients (G6PD Brighton, G6PD Serres and G6PD Harilaou) and from individuals with class III G6PD deficiency (G6PD Med and G6PD A- mutations; Table 2-2).

Figure 4-14, Figure 4-15 and Figure 4-16 show *G6PD* mRNA and protein levels and enzymatic activity results for G6PD Brighton, G6PD Serres and G6PD Harilaou, respectively, obtained by treating the erythroid cells for 48h (from day 5 to day 7) with 1mM NaBu. Although the healthy donor (n=4) sample results are summarised in each figure, the results from each class I patient are shown in individual figures as they represent different pathogenic mutations. Figure 4-17 summarises the cumulative results of all 3 class I G6PD deficient patients.

In G6PD Brighton cells (Figure 4-14), *G6PD* mRNA and protein expression increases by 3.4-fold and 4-fold in response to NaBu treatment, respectively. Importantly, G6PD enzymatic activity is also increased by 3.1-fold, consistent with the results from CB-$CD34^+$ erythroid cultures (Figure 4-11) as well as B cell lines (Figure 3-9). It is also very interesting to note that the absolute expression of G6PD protein as shown in Figure 4-14B by western blotting is lower in G6PD Brighton in comparison to the healthy donor sample. This is consistent with my findings in cell lines, which were presented in Figure 3-8. Similarly, *G6PD* mRNA and protein expression, as well as enzymatic activity increased 6.8-fold, 5.7-fold and 2.5-fold in G6PD Serres erythroid progenitors upon NaBu treatment (Figure 4-15). Furthermore, treated erythroid cells from G6PD Harilaou showed 10-fold, 4.7-fold and 4.2-fold increases in the *G6PD* mRNA expression, protein expression and enzymatic activity, respectively (Figure 4-16). Altogether, class I G6PD deficient cells showed a statistically significant ($p<0.05$) 6.7-fold increase in *G6PD* mRNA expression and a 4.8-fold increase in G6PD protein expression as well as a 3-fold increase in G6PD enzymatic activity in response to NaBu (Figure 4-17).

It should be noted that during the *in vitro* erythroid differentiation culture baseline G6PD enzymatic activity in class I patients appears to around 50% that of

normal cells, as opposed to approximately 10% that it is in RBCs *in vivo* (See Table 2-2 with patient characteristics). This reflects the fact that erythroid precursors are nucleated cells with preserved ability to continuously transcribe *G6PD*, while mature RBCs have lost this capacity because they are anuclear.

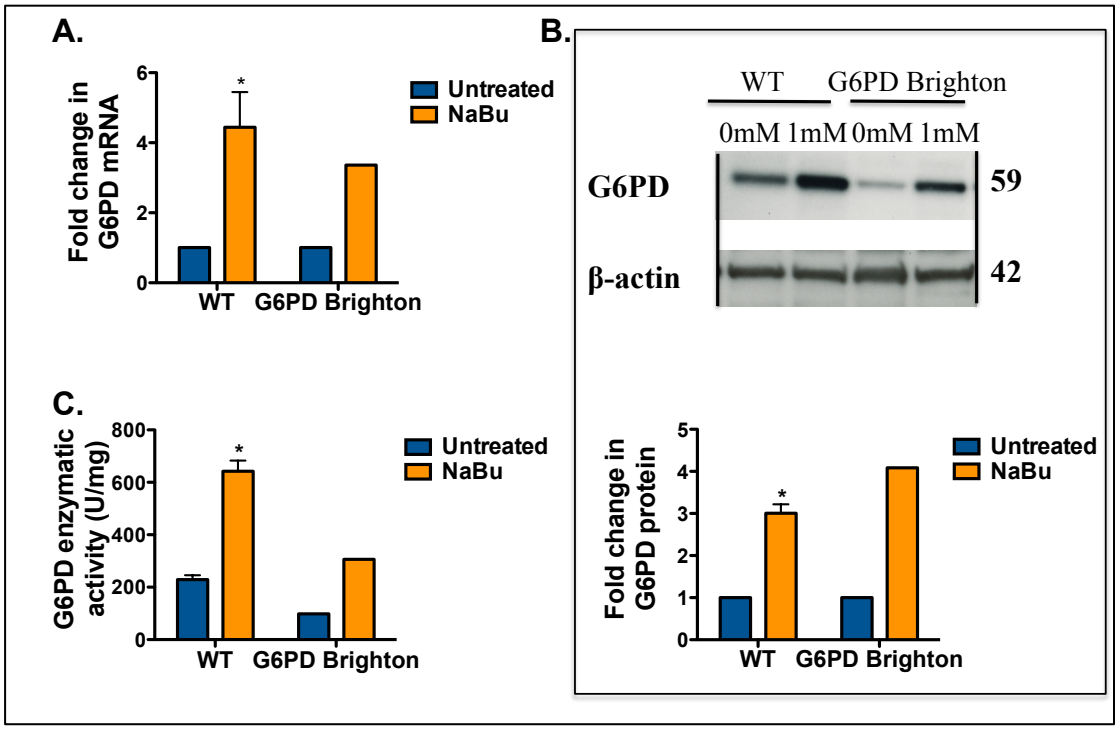


Figure 4-14 *G6PD* expression in normal and G6PD Brighton PBMC-derived erythroid cells upon NaBu treatment. PBMC cells were differentiated into erythroid precursors from healthy donors and the class I G6PD Brighton patient. **(A)** *G6PD* mRNA expression, assessed by RT-qPCR, increased 4.4-fold and 3.4-fold in healthy donors and G6PD Brighton, respectively upon 1mM NaBu treatment for 48h. **(B)** Similarly, Western blot assessed G6PD protein expression, which is shown to increase 3-fold and 4-fold in healthy donors and G6PD Brighton, respectively. **(C)** G6PD enzymatic activity increases 2.8-fold and 3.1-fold in healthy donors and G6PD Brighton, respectively upon NaBu treatment. Mean and S.E.M. are shown for n=4 for healthy donors and n=1 for G6PD Brighton. Student's t-test was performed to compare the untreated to the treated condition of the healthy donors.

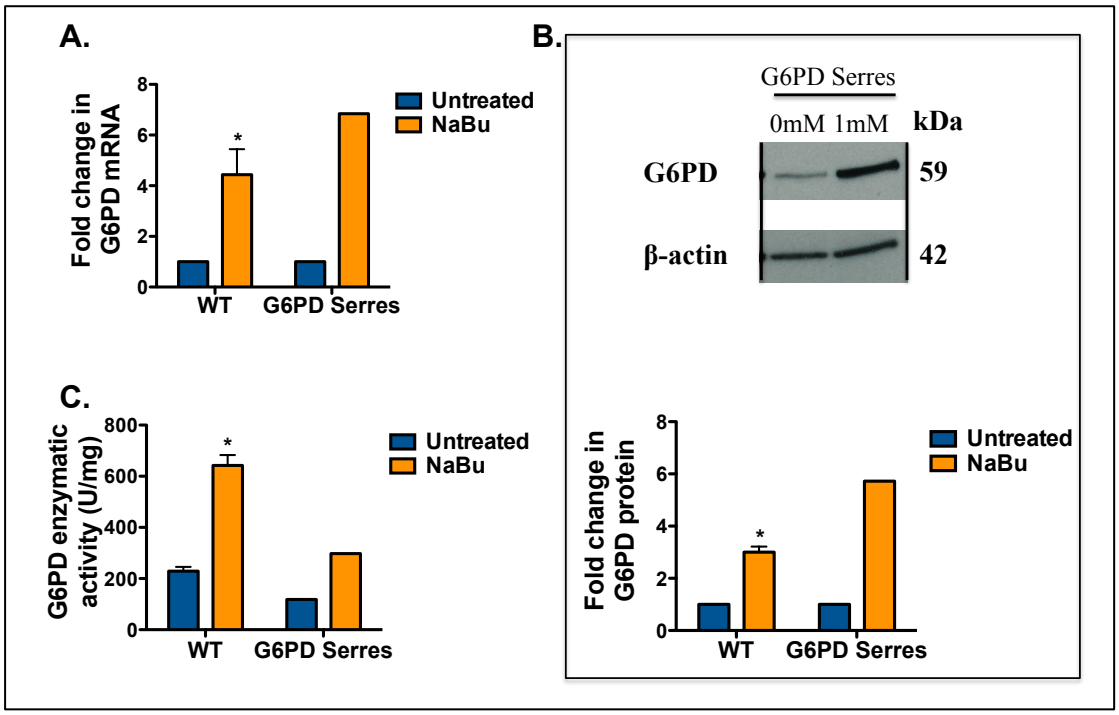


Figure 4-15 *G6PD* expression in normal and G6PD Serres PBMC-derived erythroid cells upon NaBu treatment. PBMC cells were differentiated into erythroid precursors from healthy donors and the class I G6PD Serres patient. **(A)** *G6PD* mRNA expression, assessed by RT-qPCR, increased 4.4-fold and 6.8-fold in healthy donors and G6PD Serres, respectively upon 1mM NaBu treatment for 48h. **(B)** Similarly, Western blot assessed G6PD protein expression, which is shown to increase 3-fold and 5.7-fold in healthy donors and G6PD Serres, respectively. **(C)** G6PD enzymatic activity increases 2.8-fold and 2.5-fold in healthy donors and G6PD Serres, respectively. Mean and S.E.M. are shown for n=4 for healthy donors and n=1 for G6PD Serres. Student's t-test was performed to compare the untreated to the treated condition of the healthy donors.

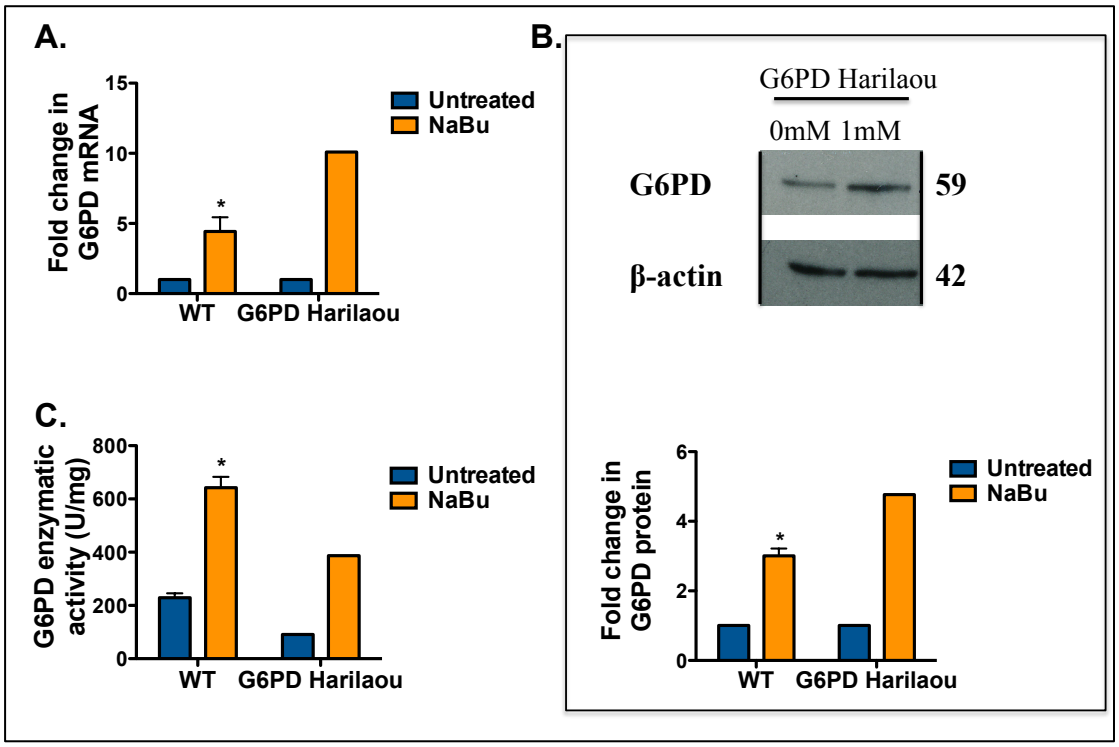


Figure 4-16 *G6PD* expression in normal and G6PD Harilaou PBMC-derived erythroid cells upon NaBu treatment. PBMC cells were differentiated into erythroid precursors from healthy donors and the class I G6PD Harilaou patient. **(A)** *G6PD* mRNA expression, assessed by RT-qPCR, increased 4.4-fold and 10-fold in healthy donors and G6PD Harilaou, respectively upon 1mM NaBu treatment for 48h. **(B)** Similarly, Western blot assessed G6PD protein expression, which is shown to increase 3-fold and 4.7-fold in healthy donors and G6PD Harilaou, respectively. **(C)** G6PD enzymatic activity increases 2.8-fold and 4.2-fold in healthy donors and G6PD Harilaou, respectively. Mean and S.E.M. are shown for n=4 for healthy donors and n=1 for G6PD Harilaou. Student's t-test was performed to compare the untreated to the treated condition of the healthy donors.

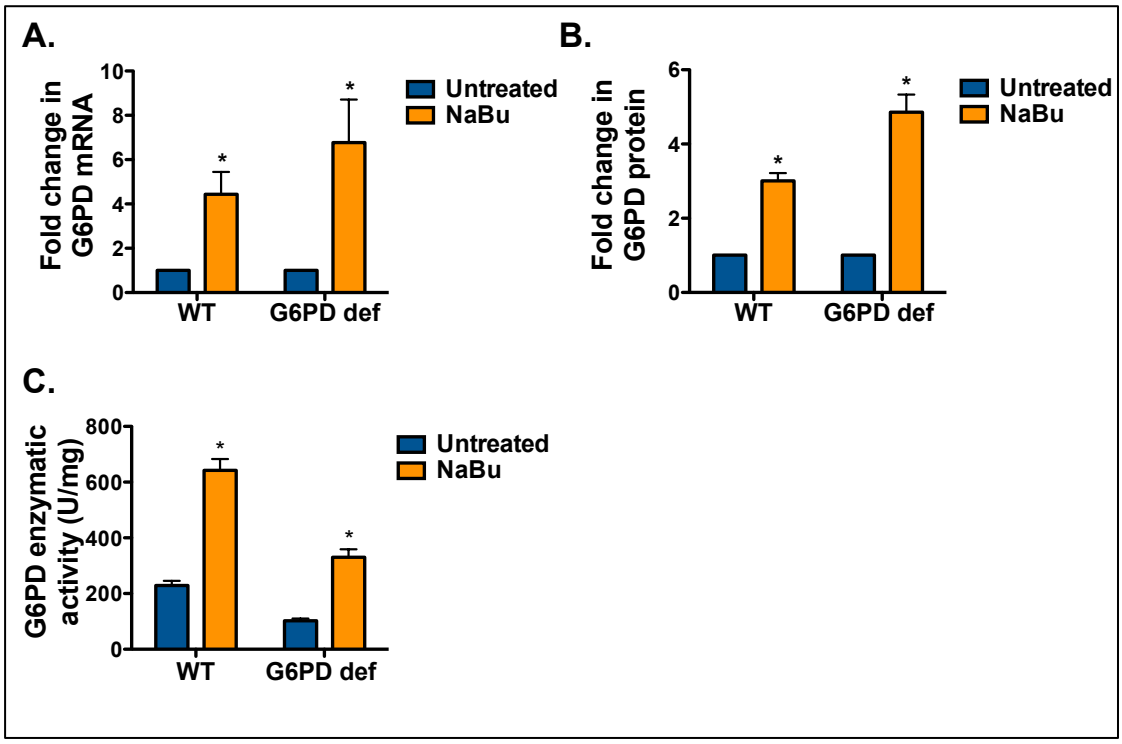


Figure 4-17 *G6PD* expression in primary human normal and class I G6PD deficient erythroid cells. PBMC cells were differentiated into proerythroblasts from normal and class I G6PD deficient individuals. **(A)** *G6PD* mRNA expression, assessed by RT-qPCR, increased 4.4-fold and 6.7-fold in normal and G6PD deficient samples, respectively upon 1mM NaBu treatment for 48h. **(B)** Similarly, Western blot assessed G6PD protein expression, which is shown to increase 3-fold and 4.8-fold in normal and G6PD deficient cells, respectively. **(C)** G6PD enzymatic activity increases 2.8-fold and 3.2-fold in normal and G6PD deficient cells, respectively upon NaBu treatment. Mean and S.E.M. are shown for n=4 for normal and n=3 for class I deficient samples (i.e. G6PD Brighton, G6PD Serres and G6PD Harilaou). Student's t-test was performed to compare the untreated to the treated condition.

To extend the studies conducted on class I G6PD deficient patients, two class III G6PD deficient individuals carrying the G6PD Med and G6PD A- polymorphic mutations were also recruited and isolated PBMCs were subjected to the same erythroid differentiation process as well as NaBu treatment. Figure 4-18 shows that treated erythoid cells from a G6PD Med individual have 2.5-fold increased expression of *G6PD* mRNA and protein and display a 3.5-fold increase in G6PD enzymatic activity. Similarly, erythroid cells produced from G6PD A- PBMCs had increased G6PD mRNA and protein expression 15-fold and 3.5-fold, respectively as well as 3.3-fold increase in G6PD enzymatic activity upon NaBu treatment.

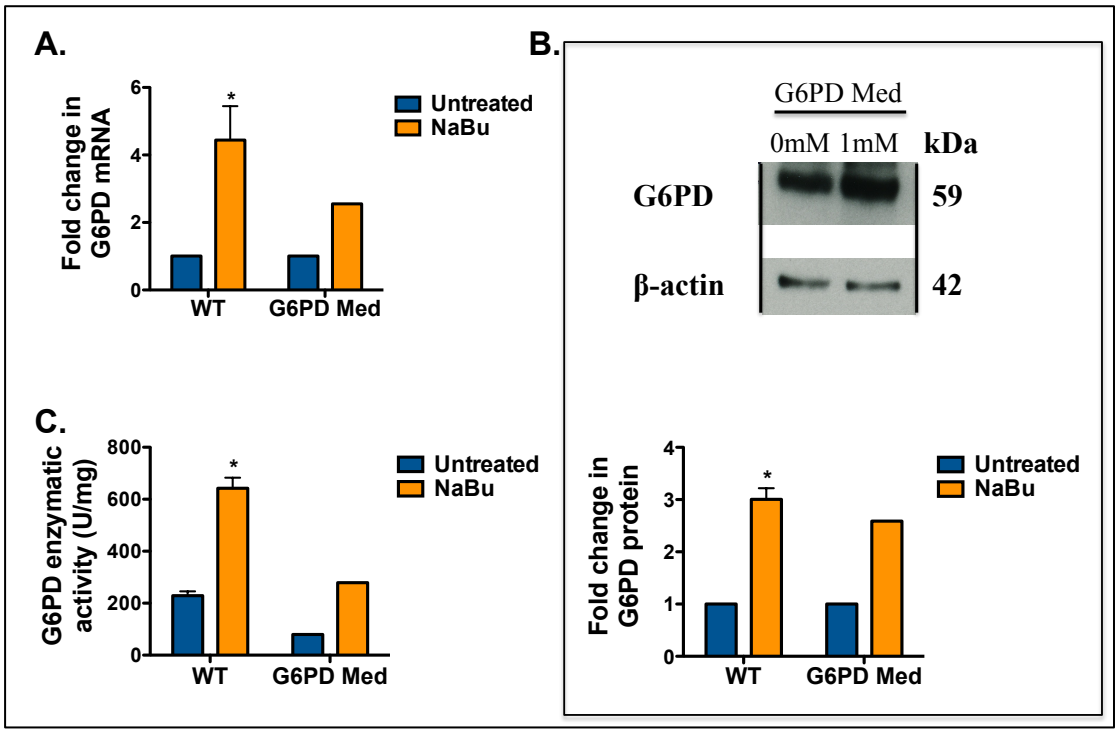


Figure 4-18 *G6PD* expression in normal and G6PD Mediterranean PBMC-differentiating erythroid cells upon NaBu treatment. PBMC cells were differentiated into proerythroblasts from healthy donors and the class III G6PD Med. **(A)** *G6PD* mRNA expression, assessed by RT-qPCR, increased 4.4-fold and 2.5-fold in healthy donors and G6PD Med, respectively upon 1mM NaBu treatment for 48h. **(B)** Similarly, Western blot assessed G6PD protein expression, which is shown to increase 3-fold and 2.5-fold in healthy donors and G6PD Med, respectively. **(C)** G6PD enzymatic activity increases 2.8-fold and 3.5-fold in healthy donors and G6PD Med, respectively. Mean and S.E.M. are shown for n=4 for healthy donors and n=1 for G6PD Med. Student's t-test was performed to compare the untreated to the treated condition of the healthy donors.

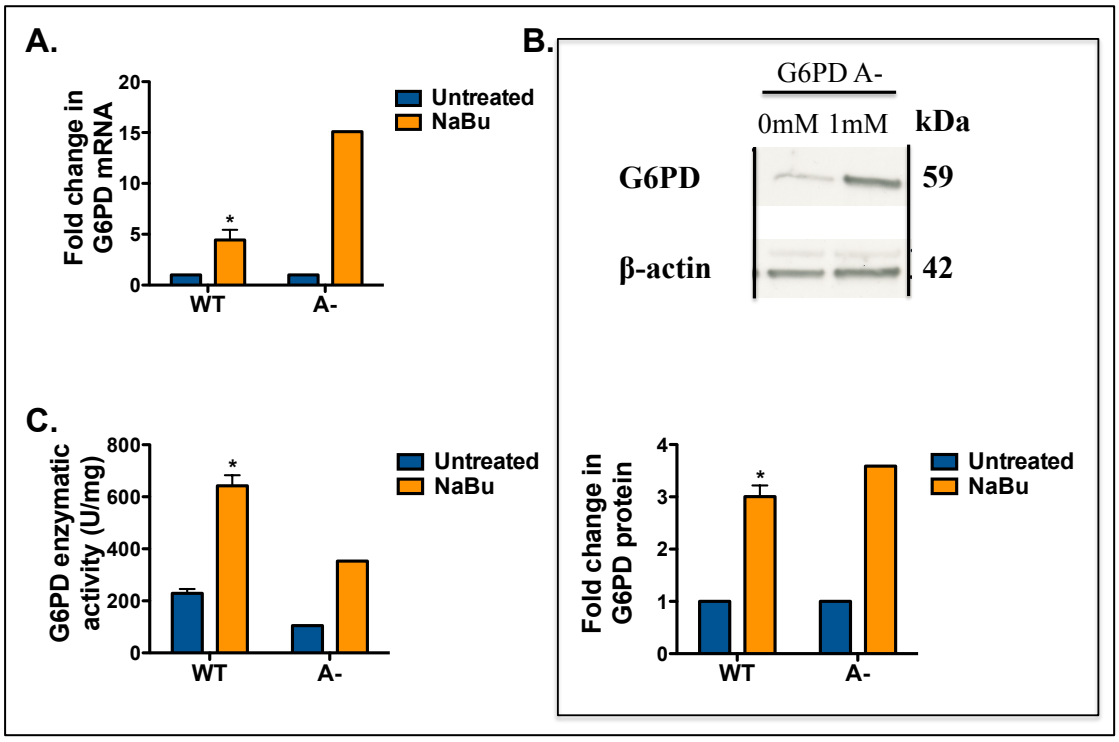


Figure 4-19 *G6PD* expression in normal and G6PD African PBMC-derived erythroid cells upon NaBu treatment. PBMC cells were differentiated into proerythroblasts from healthy donors and the class III G6PD A-. **(A)** *G6PD* mRNA expression, assessed by RT-qPCR, increased 4.4-fold and 15-fold in healthy donors and G6PD A-, respectively upon 1mM NaBu treatment for 48h. **(B)** Similarly, Western blot assessed G6PD protein expression, which is shown to increase 3-fold and 3.5-fold in healthy donors and G6PD A-, respectively. **(C)** G6PD enzymatic activity increases 2.8-fold and 3.3-fold in healthy donors and G6PD A-, respectively. Mean and S.E.M. are shown for n=4 for healthy donors and n=1 for G6PD A-. Student's t-test was performed to compare the untreated to the treated condition of the healthy donors.

In conclusion, treatment with NaBu selectively increases *G6PD* mRNA expression, protein expression and enzymatic activity in both healthy donors and G6PD deficient patients. Strikingly, the upregulation of G6PD occurs irrespective of the type of mutation on the G6PD exons, as I show here that the same trend is observed in 3 class I and two class III patients, who carry independent mutations.

4.6 Conclusions

In this chapter, I aimed to test the effect of NaBu on the GPPP gene expression in primary human erythroid cells. Therefore, I set up a two-phase liquid culture system for erythroid differentiation starting from either PBMCs or CB-CD34$^+$ cells, which recapitulates erythropoiesis *in vitro* and produces erythroid precursors in 7 days. However, treatment of the generated erythroid precursors with NaBu (but not other HDACIs) is shown to affect erythroid differentiation and to reduce cell numbers. In fact, PBMC-derived erythroid cells display delayed erythroid maturation, whereas in the case of the CB-CD34$^+$-derived cells erythropoiesis seems to be promoted. SAHA and TSA are shown to reduce the number of erythroid cells produced, without affecting the differentiation of CB-CD34$^+$-differentiating cells. Acknowledging the importance of producing same-stage cells upon NaBu treatment in order to compare them with the untreated cells in the context of GPPP expression, I established the conditions required for NaBu: low concentration of 1mM NaBu and short exposure of 48h from day 5 to day 7 of phase 1.

The results show that *G6PD* expression is selectively upregulated in CB-CD34$^+$-derived erythroid cells amongst all the 17 genes of the GPPP. Further investigation showed that during erythroid differentiation, irrespective of the origin of the erythroid cells, i.e. PBMCs or CD34$^+$ cells, *G6PD* mRNA and protein expression as well as G6PD enzymatic activity increase. Importantly, this study reveals that the expression and enzymatic activity of G6PD can be increased by NaBu in human primary erythroid G6PD deficient cells, which were isolated from 5 different G6PD deficient individuals, including 3 class I and 2 class II G6PD variants, representing 5 distinct characterised mutations.

Returning to my original hypothesis posed at the beginning of this study, it is now possible to state that NaBu selectively upregulates *G6PD* expression and enzymatic activity in normal and G6PD deficient cells, suggesting new therapeutic potential for individuals with G6PD deficiency.

5 Results III: Epigenetic regulation of the GPPP genes

5.1 Introduction

The dynamic balance of histone acetylation and deacetylation plays a critical role in transcriptional regulation and determines activation versus repression. HDACs have been traditionally regarded as co-repressors that bind to gene promoters, instead of HATs and consequently inhibit transcription through histone deacetylation (Berger, 2007; Xu et al., 1998). However, a recent detailed genome-wide study (Wang et al., 2009) that involved ChIP-seq analysis of primary $CD4^+$ T cells provided new insights into the role of histone acetylation and deacetylation of active, primed and inactive genes. Specifically high binding levels of both HATs and HDACs were found at the promoters and gene bodies of active, housekeeping genes, whilst at primed genes there are high binding levels of both HATs and HDACs only at the gene promoters. In the case of inactive genes, HATs and HDACs were not recruited. This finding has led to the speculation that HDAC activity is required to reset the chromatin status by removing the acetyl groups after the completion of each round of transcription so that the gene does not hyper-acetylate (Wang et al., 2009).

HDACIs are a class of compounds that inhibit the activity of histone deacetylases (Marks, 2010). One of the commonest used HDACIs is butyric acid and its derivatives (including NaBu), which have been previously used for the treatment of haemoglobinopathies (Perrine et al., 2010), urea cycle disorders (Batshaw et al., 2001), sickle cell anaemia (Dover et al., 1994) and IGD (Almeida et al., 2007).

Butyrate responsive elements have been characterised within the promoters of the butyrate-responsive genes (Davie, 2003; Siavoshian et al., 1997) and can be categorised into two separate groups. The first group consists of genes that are either induced or repressed and have a common AGCCACCTCCA sequence, suggesting that they are bound by a common transcription factor. In the other group, there are genes, which share an Sp1/Sp3 binding site within the butyrate responsive elements. Sp1 and Sp3 are transcription factors that are ubiquitously expressed and act as both activators and repressors of gene expression and have the ability to interact with HATs and HDACs.

5.1.1 Aim of the chapter

The aim of this chapter is to dissect the transcriptional and epigenetic basis of the selective responsiveness of *G6PD* to HDACIs amongst the genes of the GPPP. By comparing the epigenetic and transcriptional landscape of the *G6PD* promoter to other genes of the GPPP.

5.1.2 Experimental plan

To explore the epigenetic and transcriptional status of the *G6PD* promoter, I performed bioinformatic analysis of the ENCODE project genome-wide ChIP-Seq data provided by the UCSC genome browser (http://genome.ucsc.edu/ENCODE/). Following this, ChIP analysis was employed to assess the acetylation status as well as the binding of HATs, HDACs and the TF Sp1 on the *G6PD* promoter compared to other GPPP promoters at baseline and in response to NaBu.

5.2 ENCODE analysis of GPPP promoters

As described in 1.4.3.1, *G6PD* is a housekeeping gene, which as a result of 2 alternative TSSs encodes two variants: a short variant that consists of exons 2-13 and a longer inactive variant consisting of exons 1-13. Human *G6PD* is arranged in a 'head to head' configuration with *NEMO,* a gene that encodes a non-catalytic subunit of the cytokine-dependent IκB kinase, involved in the activation of the TF NFkB (Jin and Jeang, 1999). *NEMO* is transcribed under the influence of 2 promoters (promoter A and B). Promoter B (868bp; including a 192bp core promoter(Philippe et al., 1994)) is housekeeping and has strong bidirectional activity driving the transcription of both *G6PD* and *NEMO* genes. The bidirectional *G6PD* gene promoter (Figure 1-12) is embedded within a CpG island (1245bp; Fusco et al., 2006) and contains two GC-boxes that drive its expression (Ursini et al., 1990). Furthermore, 12 binding sites for the transcription factor Sp1 are predicted within a region of 1327bp surrounding exon 1 of the *G6PD* gene (Fusco et al., 2006; Galgoczy et al., 2001; Philippe et al., 1994).

Analysis of ENCODE-derived data (Figure 5-1) confirmed expression of *G6PD* in 9 different human cell lines (B lymphoblastoid cell line GM12878, stem cell line H1-hESC, cervical cancer cell line HeLa, hepatic cell line HepG2, myoblastic cell line HSMM, umbilical vein endothelial cell line HUVEC, erythroleukaemia cell line K562, epidermal keratinocytes NHEK and lung fibroblasts NHLF) and showed a considerable enrichment of Sp1 covering the length of the shared, housekeeping

promoter. Furthermore, consistent with the features of housekeeping gene promoters, H3 acetylation marks but not silencing marks such as H3K27me3 are readily identified. As well as HATs, HDACs are shown to bind to the promoter at low level, thus confirming the validity of the results obtained by Wang et al.(Wang et al., 2009) i.e., that HATs and HDACs are both present in housekeeping gene promoters. Taken together these results suggest that Sp1, HATs, HDACs and histone acetylation are important structural and functional components of the *G6PD* promoter.

Further bioinfomatic analysis sought to compare the *G6PD* promoter with other promoters of the GPPP, such as those of *TPI* and *GPI*, which are genes that remain unaffected by NaBu. This was conducted using both the ENCODE browser and the transcription factor binding site prediction browsers TFSearch (http://www.cbrc.jp/research/db/TFSEARCH.html) and CONSITE (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite). Similarly to *G6PD*, the promoters of *TPI* and *GPI* present a typical housekeeping gene structure. They are embedded within CpG islands, contain GC boxes and are enriched for acetylation, HDACs, HATs, Sp1 and other common TFs (Figure 5-2 and Figure 5-3).

Taken together, bioinformatics analysis of the promoters of *G6PD*, *TPI* and *GPI* shows that the key features of these promoters are very similar and consistent to those of housekeeping gene promoters. The only difference evident from this abalysis between *G6PD* and all the other GPPP genes is the fact that *G6PD* shares a bidirectional promoter with another gene, i.e. *NEMO*.

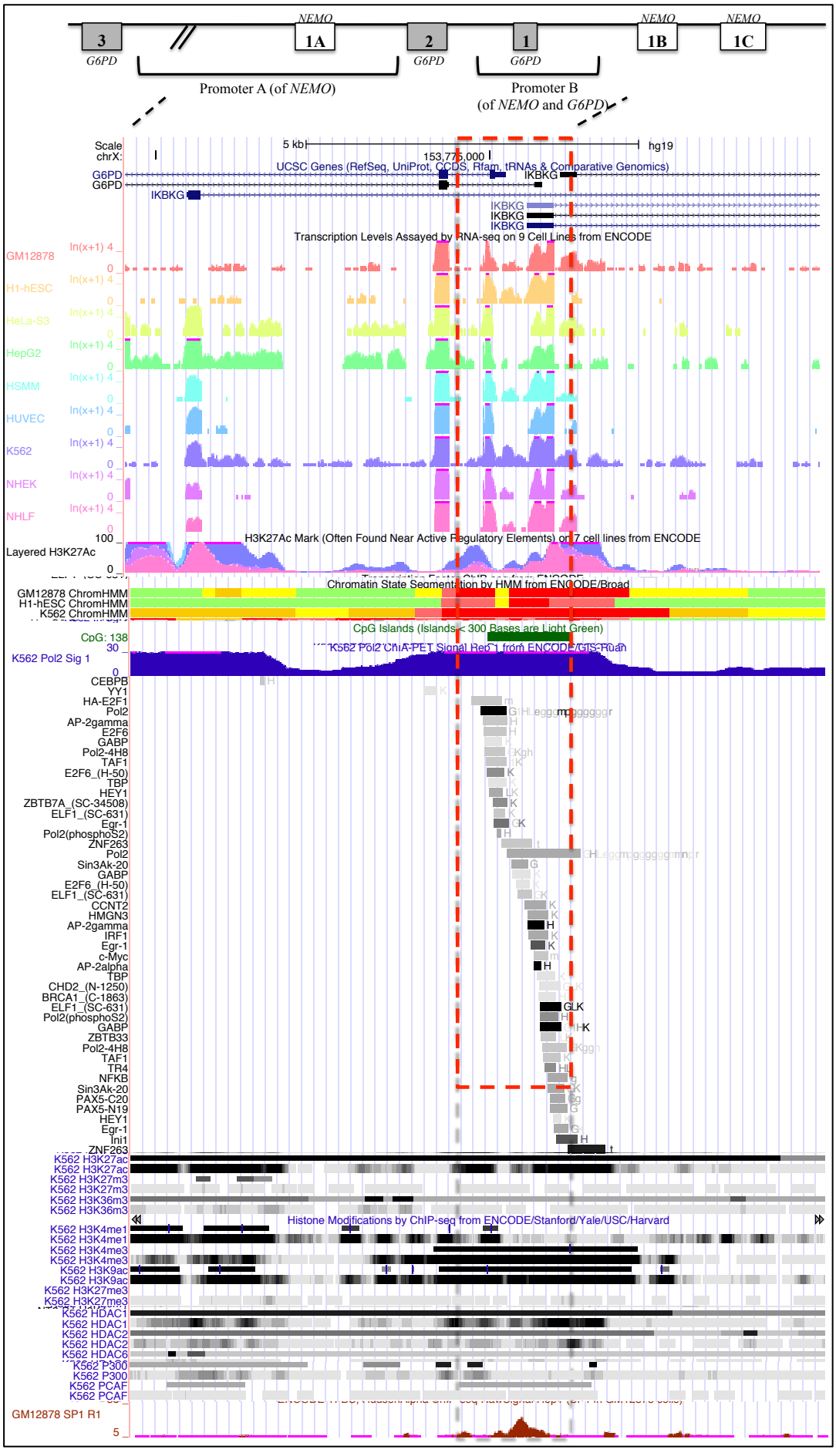

NEMO
3
1A
2
1
1B
1C
G6PD
G6PD
G6PD
Promoter A (of NEMO)
Promoter B
(of NEMO and G6PD)
Scale
chrX:
5 kb
153,775,000
hg19
UCSC Genes (RefSeq, UniProt, CCDS, Rfam, tRNAs & Comparative Genomics)
G6PD
G6PD
IKBKG
IKBKG
IKBKG
IKBKG
IKBKG
Transcription Levels Assayed by RNA-seq on 9 Cell Lines from ENCODE
GM12878
H1-hESC
HeLa-S3
HepG2
HSMM
HUVEC
K562
NHEK
NHLF
ln(x+1) 4
0
100
Layered H3K27Ac
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Chromatin State Segmentation by HMM from ENCODE/Broad
GM12878 ChromHMM
H1-hESC ChromHMM
K562 ChromHMM
CpG Islands (Islands < 300 Bases are Light Green)
CpG: 138
K562 Pol2 ChIA-PET Signal Rep1 from ENCODE/GIS-Ruan
30
K562 Pol2 Sig 1
0
CEBPB
YY1
HA-E2F1
Pol2
AP-2gamma
E2F6
GABP
Pol2-4H8
TAF1
E2F6_(H-50)
TBP
HEY1
ZBTB7A_(SC-34508)
ELF1_(SC-631)
Egr-1
Pol2(phosphoS2)
ZNF263
Pol2
Sin3Ak-20
GABP
E2F6_(H-50)
ELF1_(SC-631)
CCNT2
HMGN3
AP-2gamma
IRF1
Egr-1
c-Myc
AP-2alpha
TBP
CHD2_(N-1250)
BRCA1_(C-1863)
ELF1_(SC-631)
Pol2(phosphoS2)
GABP
ZBTB33
Pol2-4H8
TAF1
TR4
NFKB
Sin3Ak-20
PAX5-C20
PAX5-N19
HEY1
Egr-1
Ini1
ZNF263
K562 H3K27ac
K562 H3K27ac
K562 H3K27m3
K562 H3K27m3
K562 H3K36m3
K562 H3K36m3
Histone Modifications by ChIP-seq from ENCODE/Stanford/Yale/USC/Harvard
K562 H3K4me1
K562 H3K4me1
K562 H3K4me3
K562 H3K4me3
K562 H3K9ac
K562 H3K9ac
K562 H3K27me3
K562 H3K27me3
K562 HDAC1
K562 HDAC1
K562 HDAC2
K562 HDAC2
K562 HDAC6
K562 P300
K562 P300
K562 PCAF
K562 PCAF
GM12878 SP1 R1
5

Figure 5-1 Transcriptional and epigenetic landscape of the *G6PD* promoter based on the ENCODE project. Schema obtained from the ENCODE project via the UCSC genome browser, showing the *G6PD* and *NEMO* genes in a bidirectional promoter conformation. Boxed in red is the bidirectional promoter B. The expression of the housekeeping genes is shown in 9 different cell lines. The enrichment for positive and negative chromatin marks is shown, as well as the enrichment for HATs (p300, PCAF), HDACs (HDAC1, 2 and 6) and Sp1 at baseline.

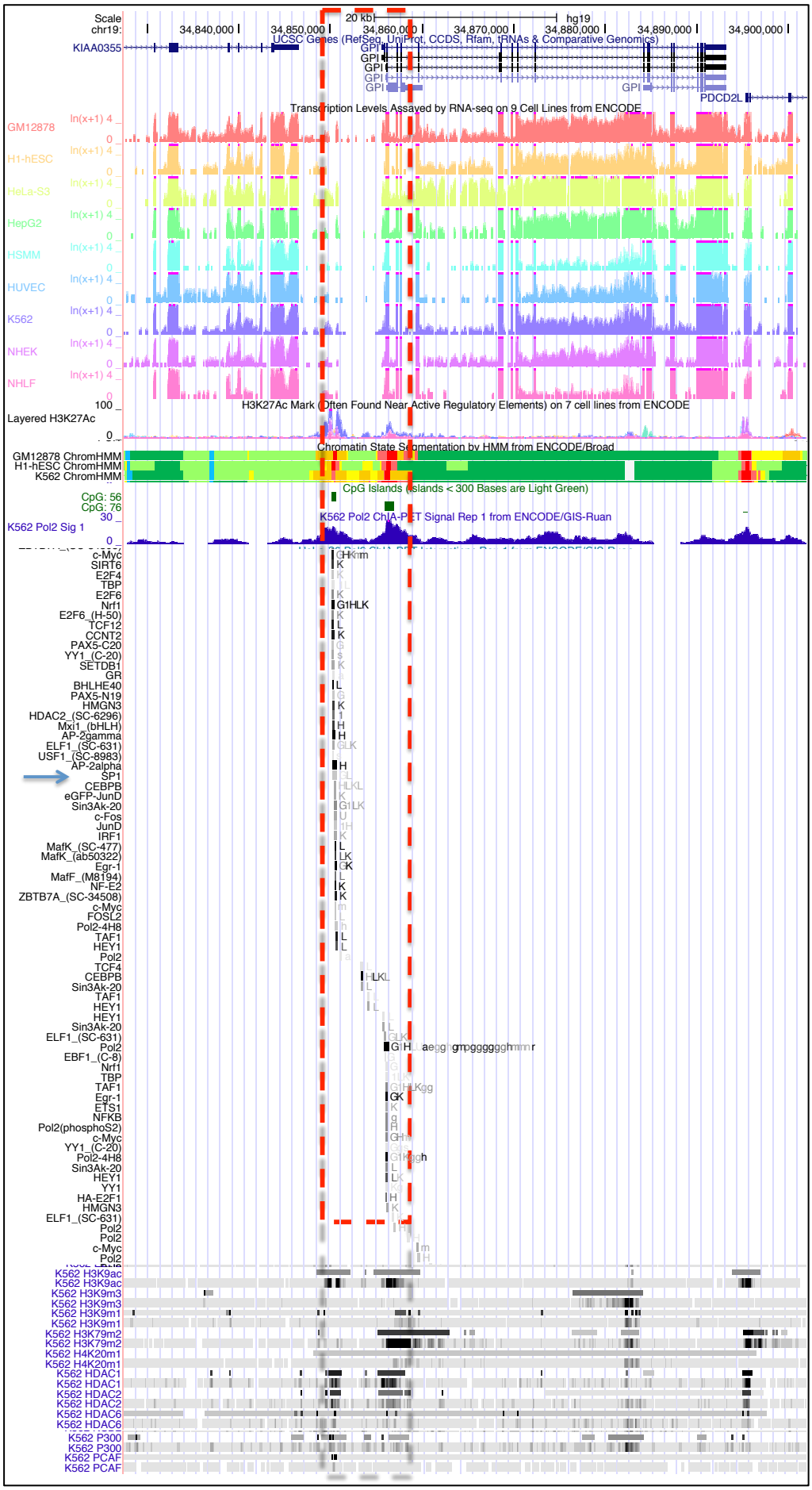

Scale
chr19:
20 kb
hg19
34,840,000
34,850,000
34,860,000
34,870,000
34,880,000
34,890,000
34,900,000
UCSC Genes (RefSeq, UniProt, CCDS, Rfam, tRNAs & Comparative Genomics)
KIAA0355
GPI
PDCD2L
Transcription Levels Assayed by RNA-seq on 9 Cell Lines from ENCODE
GM12878
H1-hESC
HeLa-S3
HepG2
HSMM
HUVEC
K562
NHEK
NHLF
ln(x+1) 4
0
100
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Layered H3K27Ac
Chromatin State Segmentation by HMM from ENCODE/Broad
GM12878 ChromHMM
H1-hESC ChromHMM
K562 ChromHMM
CpG Islands (Islands < 300 Bases are Light Green)
CpG: 56
CpG: 76
30
K562 Pol2 ChIA-PET Signal Rep 1 from ENCODE/GIS-Ruan
K562 Pol2 Sig 1
c-Myc
SIRT6
E2F4
TBP
E2F6
Nrf1
E2F6_(H-50)
TCF12
CCNT2
PAX5-C20
YY1_(C-20)
SETDB1
GR
BHLHE40
PAX5-N19
HMGN3
HDAC2_(SC-6296)
Mxi1_(bHLH)
AP-2gamma
ELF1_(SC-631)
USF1_(SC-8983)
AP-2alpha
SP1
CEBPB
eGFP-JunD
Sin3Ak-20
c-Fos
JunD
IRF1
MafK_(SC-477)
MafK_(ab50322)
Egr-1
MafF_(M8194)
NF-E2
ZBTB7A_(SC-34508)
c-Myc
FOSL2
Pol2-4H8
TAF1
HEY1
Pol2
TCF4
CEBPB
Sin3Ak-20
TAF1
HEY1
HEY1
Sin3Ak-20
ELF1_(SC-631)
Pol2
EBF1_(C-8)
Nrf1
TBP
TAF1
Egr-1
ETS1
NFKB
Pol2(phosphoS2)
c-Myc
YY1_(C-20)
Pol2-4H8
Sin3Ak-20
HEY1
YY1
HA-E2F1
HMGN3
ELF1_(SC-631)
Pol2
Pol2
c-Myc
Pol2
K562 H3K9ac
K562 H3K9ac
K562 H3K9m3
K562 H3K9m3
K562 H3K9m1
K562 H3K9m1
K562 H3K79m2
K562 H3K79m2
K562 H4K20m1
K562 H4K20m1
K562 HDAC1
K562 HDAC1
K562 HDAC2
K562 HDAC2
K562 HDAC6
K562 HDAC6
K562 P300
K562 P300
K562 PCAF
K562 PCAF

Figure 5-2 Transcriptional and epigenetic landscape of the *GPI* promoter based on the ENCODE project. Schema obtained from the ENCODE project via the UCSC genome browser, showing the *GPI* gene. Its promoter region is boxed in red. The expression of the housekeeping gene is shown in 9 different cell lines. The enrichment for positive and negative chromatin marks is shown, as well as the enrichment for HATs (p300, PCAF), HDACs (HDAC1, 2 and 6) and Sp1 (indicated with arrow) at baseline.

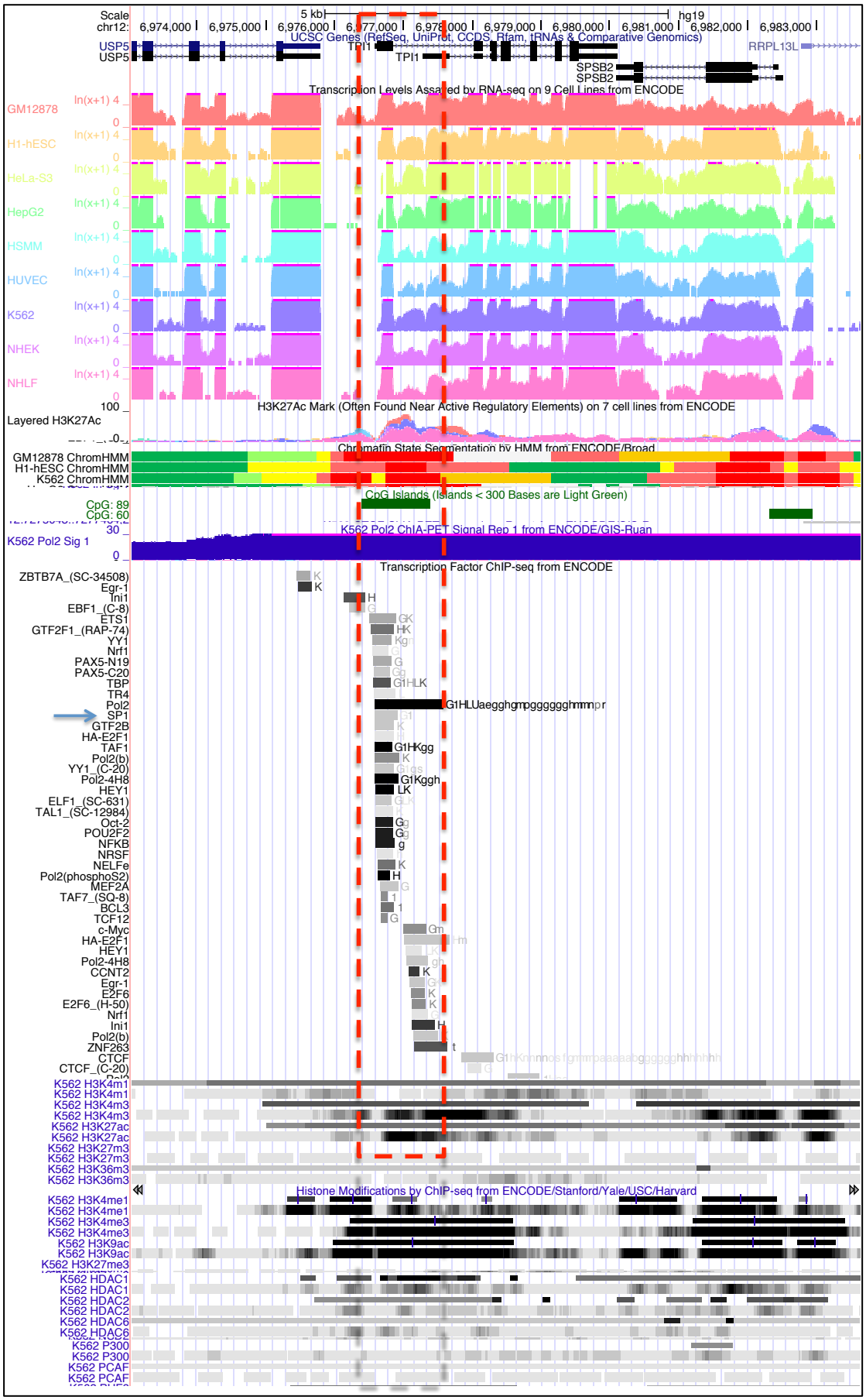

Scale
chr12:
6,974,000
6,975,000
6,976,000
6,977,000
6,978,000
6,979,000
6,980,000
6,981,000
6,982,000
6,983,000
5 kb
hg19
UCSC Genes (RefSeq, UniProt, CCDS, Rfam, tRNAs & Comparative Genomics)
USP5
TPI1
TPI1
RRPL13L
SPSB2
Transcription Levels Assayed by RNA-seq on 9 Cell Lines from ENCODE
GM12878
H1-hESC
HeLa-S3
HepG2
HSMM
HUVEC
K562
NHEK
NHLF
ln(x+1)
H3K27Ac Mark (Often Found Near Active Regulatory Elements) on 7 cell lines from ENCODE
Layered H3K27Ac
Chromatin State Segmentation by HMM from ENCODE/Broad
GM12878 ChromHMM
H1-hESC ChromHMM
K562 ChromHMM
CpG Islands (Islands < 300 Bases are Light Green)
CpG: 89
CpG: 60
K562 Pol2 ChIA-PET Signal Rep 1 from ENCODE/GIS-Ruan
K562 Pol2 Sig 1
Transcription Factor ChIP-seq from ENCODE
ZBTB7A_(SC-34508)
Egr-1
Ini1
EBF1_(C-8)
ETS1
GTF2F1_(RAP-74)
YY1
Nrf1
PAX5-N19
PAX5-C20
TBP
TR4
Pol2
SP1
GTF2B
HA-E2F1
TAF1
Pol2(b)
YY1_(C-20)
Pol2-4H8
HEY1
ELF1_(SC-631)
TAL1_(SC-12984)
Oct-2
POU2F2
NFKB
NRSF
NELFe
Pol2(phosphoS2)
MEF2A
TAF7_(SQ-8)
BCL3
TCF12
c-Myc
HA-E2F1
HEY1
Pol2-4H8
CCNT2
Egr-1
E2F6
E2F6_(H-50)
Nrf1
Ini1
Pol2(b)
ZNF263
CTCF
CTCF_(C-20)
K562 H3K4m1
K562 H3K4m1
K562 H3K4m3
K562 H3K4m3
K562 H3K27ac
K562 H3K27ac
K562 H3K27m3
K562 H3K27m3
K562 H3K36m3
K562 H3K36m3
Histone Modifications by ChIP-seq from ENCODE/Stanford/Yale/USC/Harvard
K562 H3K4me1
K562 H3K4me1
K562 H3K4me3
K562 H3K4me3
K562 H3K9ac
K562 H3K9ac
K562 H3K27me3
K562 HDAC1
K562 HDAC1
K562 HDAC2
K562 HDAC2
K562 HDAC6
K562 HDAC6
K562 P300
K562 P300
K562 PCAF
K562 PCAF

Figure 5-3 Transcriptional and epigenetic landscape of the *TPI* promoter based on the ENCODE project. Schema obtained from the ENCODE project via the UCSC genome browser, showing the *TPI* gene. Its promoter region is boxed in red. The expression of the housekeeping gene is shown in 9 different cell lines. The enrichment for positive and negative chromatin marks is shown, as well as the enrichment for HATs (p300, PCAF), HDACs (HDAC1, 2 and 6) and Sp1 (indicated with arrow) at baseline.

5.3 The role of NEMO in the G6PD selective upregulation

Having established the *G6PD*-selective effect of HDACIs on the genes of the GPPP and taking into account the main structural difference between *G6PD* and the other genes is that it shares a bidirectional promoter with *NEMO*, I sought to understand the role *NEMO* might have in the upregulation of *G6PD*. I hypothesised that *NEMO* might be a target of NaBu and therefore might be upregulated whilst also driving the upregulation of *G6PD*.

Normal and G6PD deficient cell lines were treated with NaBu at a time course of 8h, 16h, 24h and 36h and were tested for *NEMO* mRNA expression with RT-qPCR. Figure 5-4A shows that *NEMO* expression is not affected by NaBu in cell lines. Furthermore, CB-CD34^{+}-differentiating erythroid cells were produced as described in 4.2 and were treated with NaBu at the fifth day of the culture for 48h. In this case, RT-qPCR analysis (Figure 5-4B) confirmed that *NEMO* is stably expressed in the presence of NaBu. I therefore concluded that the *G6PD*-selective upregulation is not driven by *NEMO*.

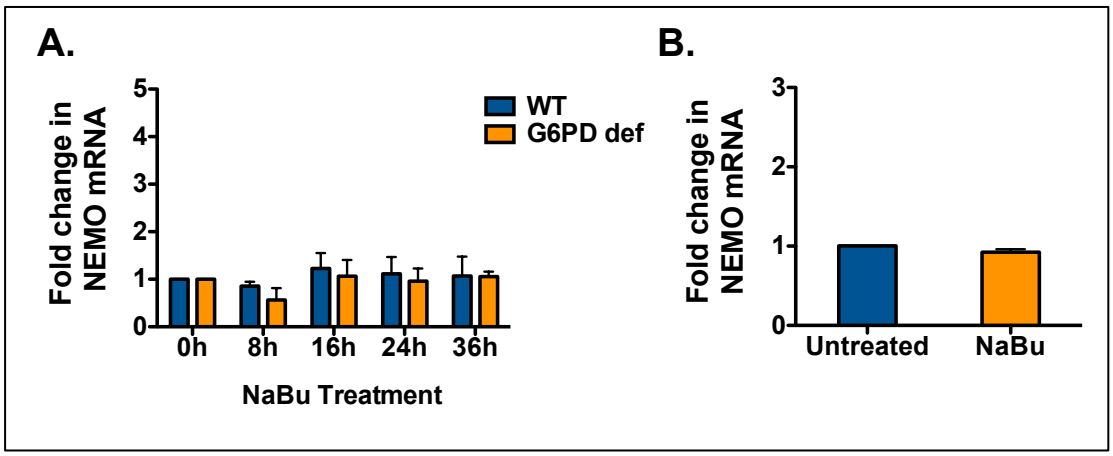


Figure 5-4 *NEMO* mRNA expression upon NaBu treatment. RT-qPCR was employed to assess *NEMO* mRNA expression in **(A)** WT and G6PD deficient cell lines at a time course of 3mM NaBu treatment as well as in **(B)** CB-CD34^{+}-differentiating erythroid cells after 1mM NaBu treatment for 48h. In both cases, *NEMO* expression remains unaffected. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to the treated conditions.

5.4 Epigenetic status of the GPPP promoters

To gain insights into the transcriptional and epigenetic basis of the selective responsiveness of *G6PD* to HDACIs, I assessed the levels of histone acetylation, a modification that is associated with transcriptional activation at the promoters of genes of the GPPP. For this purpose I employed ChIP assays using four sets of G6PD ChIP primers (Table 2-11 and Figure 5-5), designed to span the full length of the *G6PD* promoter including a primer pair that amplifies part of the core promoter (primer set 2). Furthermore, these modifications were tested in other genes of the GPPP, particularly *TPI*, *GPI*, *PGLS* and *RPIA* for which I designed primers to amplify regions within their core promoters (Table 2-11).

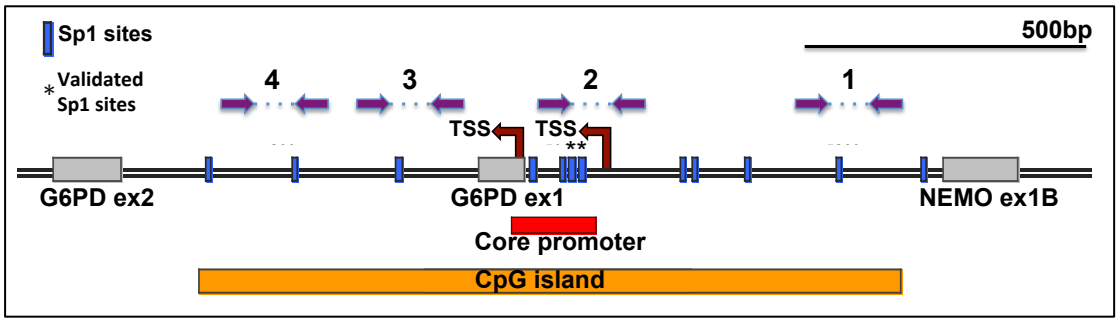


Figure 5-5 ChIP primers on the *G6PD* promoter. Schematic representation of the 4 sets of the primers used for ChIP analysis of *G6PD*.

Figure 5-6 shows that in WT B cell lines there are increased levels of H3 and H4 acetylation in the core promoters of all GPPP genes at baseline, consistent with their status as active, housekeeping genes. However, histone hyper-acetylation in response to NaBu is only observed in *G6PD* and in none of the other GPPP genes and was evident as early as 1 hr post NaBu treatment (Figure 5-6). Similarly, in G6PD deficient B cells (Figure 5-7), NaBu treatment leads to histone hyper-acetylation of *G6PD* but not of any other genes of the GPPP tested, i.e *TPI*, *GPI*, *PGLS*, *RPIA* and *GAPDH*. *GAPDH* is also used here as a positive control for acetylation at baseline. As a negative control I have used primers against an intergenic region (GW10). Figure 5-8 shows acetylation levels at baseline for *G6PD* and other GPPP gene promoters. In this figure, since I am not showing the enrichment relative to the untreated, it is clear that the negative control GW10 is indeed negative.

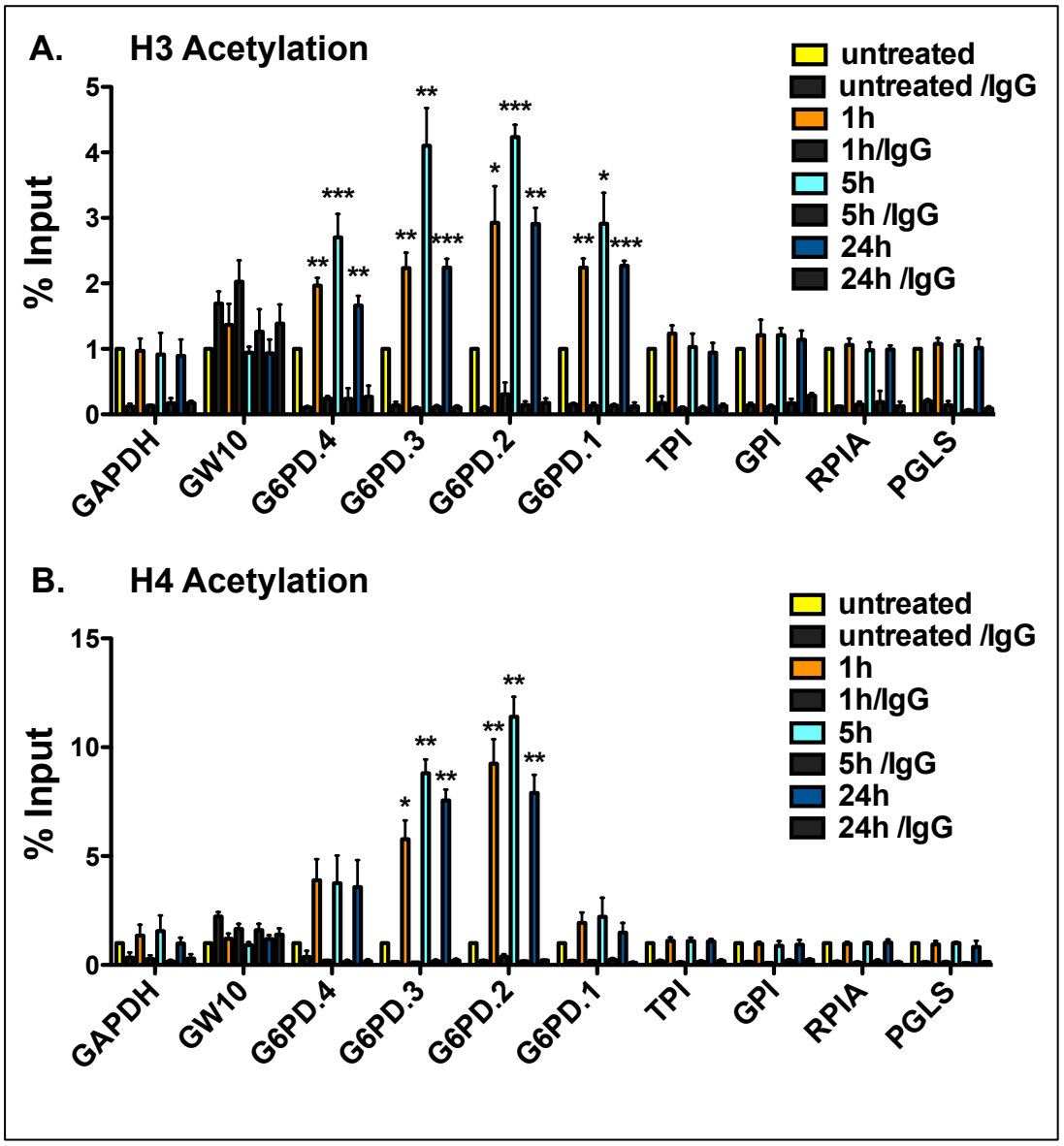


Figure 5-6. Histone acetylation of GPPP gene promoters in WT cells. (A) Histone 3 and **(B)** histone 4 acetylation assessed by ChIP in the WT B cell line P277 at a time course NaBu treatment of 1h, 5h and 24h. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *GAPDH* is used as a positive control for enrichment and GW10 is the negative control for acetylation at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to the treated conditions.

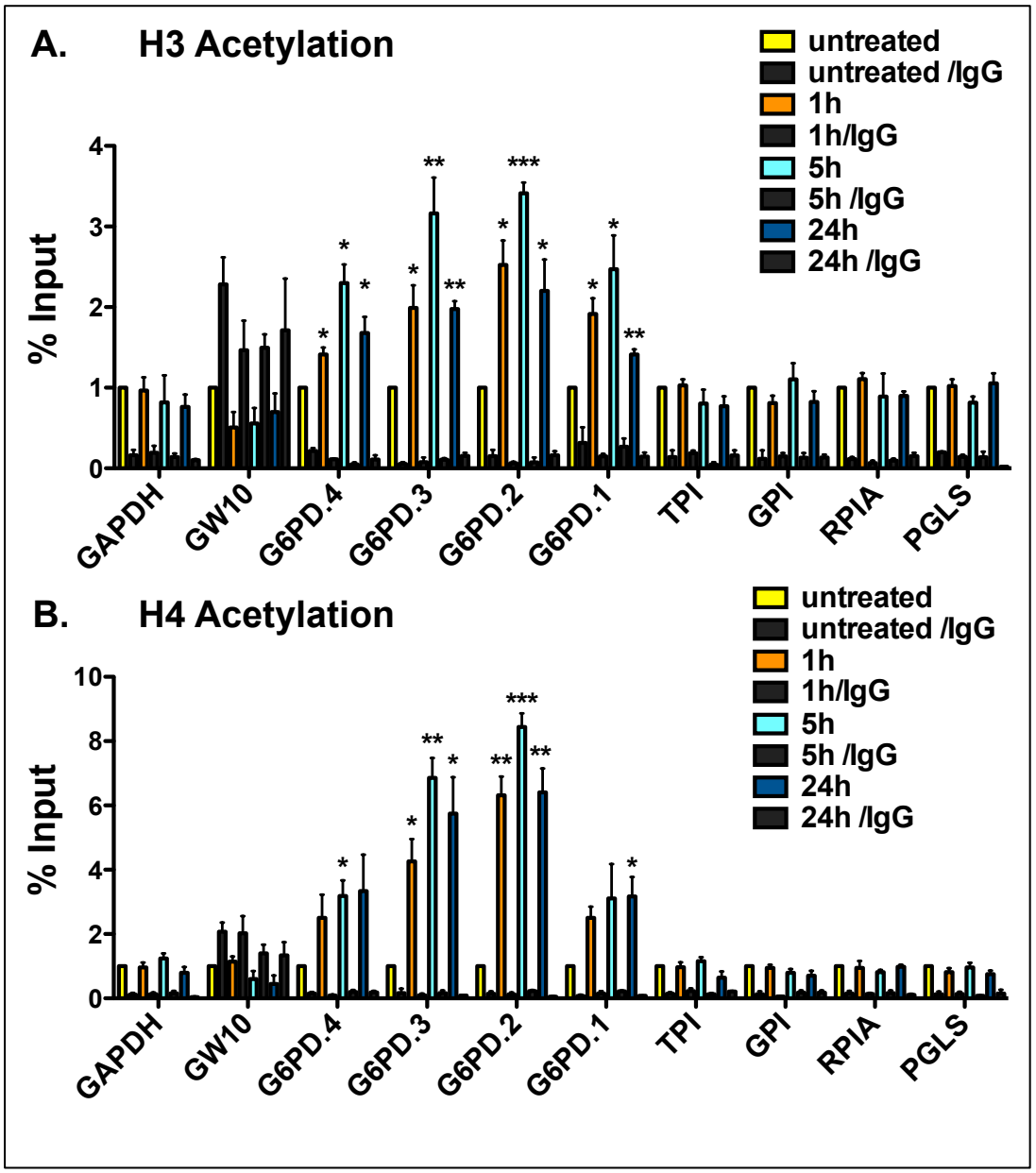


Figure 5-7 Histone acetylation of GPPP gene promoters in G6PD deficient cells. **(A)** Histone 3 and **(B)** histone 4 acetylation assessed by ChIP in the G6PD deficient B cell line P7 at a time course NaBu treatment of 1h, 5h and 24h. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *GAPDH* is used as a positive control for enrichment and GW10 is the negative control for acetylation at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to the treated conditions.

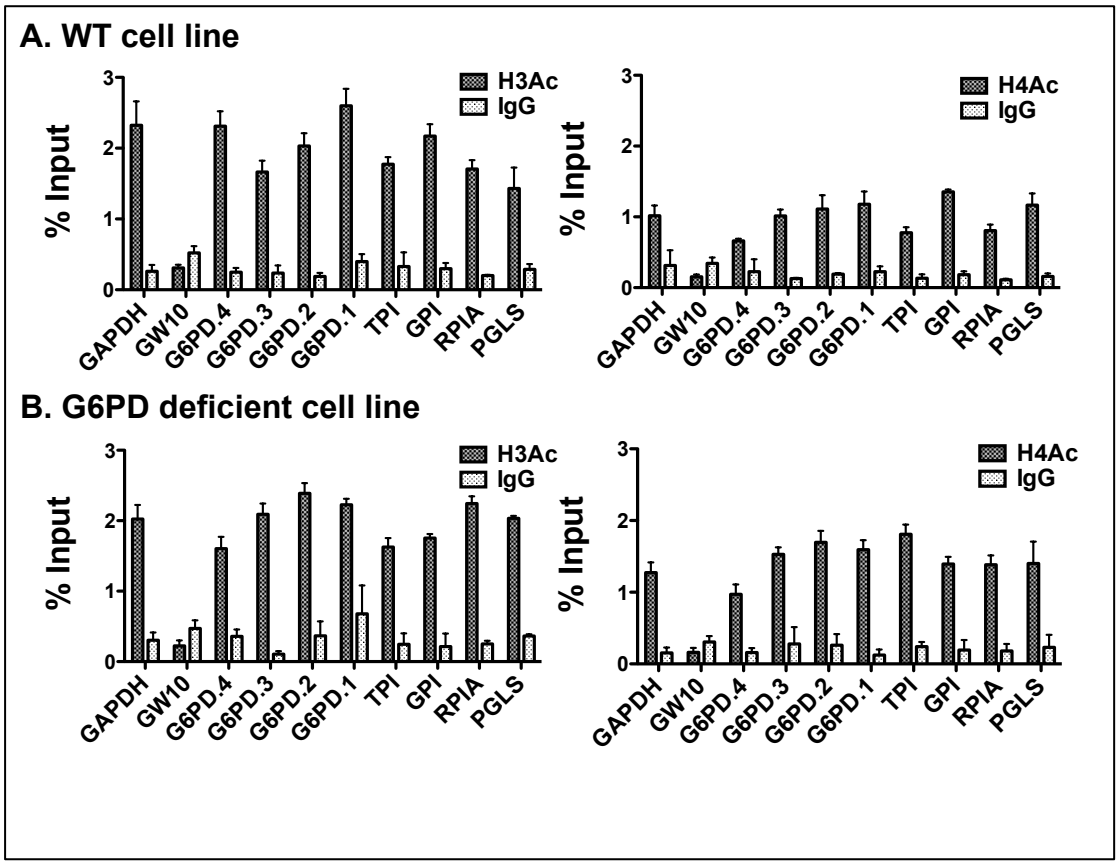


Figure 5-8 Acetylation levels on the GPPP gene promoters at baseline. H3 and H4 acetylation assessed by ChIP on the promoters of *G6PD* and other GPPP genes (*TPI, GPI, RPIA* and *PGLS*). Baseline acetylation is shown in the **(A)** WT B cell line and **(B)** G6PD deficient cell line. Enrichment on the promoters is calculated as % of input. *GAPDH* is used as a positive control for enrichment and GW10 is the negative control for acetylation at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3.

Next, I measured the binding of HATs and HDACs on the promoters of the GPPP genes. Specifically, I performed ChIP assays using antibodies against the HATs p300, CBP and GCN5 as well as the HDACs 1, 3, 4/5/7 and 6. The gene *p21* is used as a positive control for HAT and HDAC binding (it is also shown in Appendix A that *p21* is upregulated by NaBu), whereas GW10 is used as a negative control.

I first compared the baseline binding of several HATs and HDACs on the core promoter of genes of the GPPP. It is evident from the results shown in Figure 5-9 that all HDACs and HATs tested bind on the core promoters of the GPPP genes. This is consistent with the notion that levels of histone acetylation and consequent transcriptional activity are determined by the dynamic recruitment and antagonistic

activity of HATs and HDACs. Furthermore, baseline binding levels of HATs and HDACs on the *G6PD* promoter are similar to those of other GPPP genes. Therefore it cannot be suggested that higher HAT or HDAC binding on the *G6PD* promoter drives its selective transcriptional upregulation.

Following this, I treated WT B cells with NaBu for 5h, in order to capture the early epigenetic events in response to NaBu and performed ChIP to assess the binding of HATs (Figure 5-10) and HDACs (Figure 5-11). It is apparent from this data that HDAC and HAT occupancy in response to HDAC inhibition was altered only at the promoter of *G6PD*, but not the other GPPP genes that were tested. Whilst HAT binding is significantly increased for all 3 HATs tested, HDAC binding increase is statistically significant only in the case of HDAC1 and HDAC6.

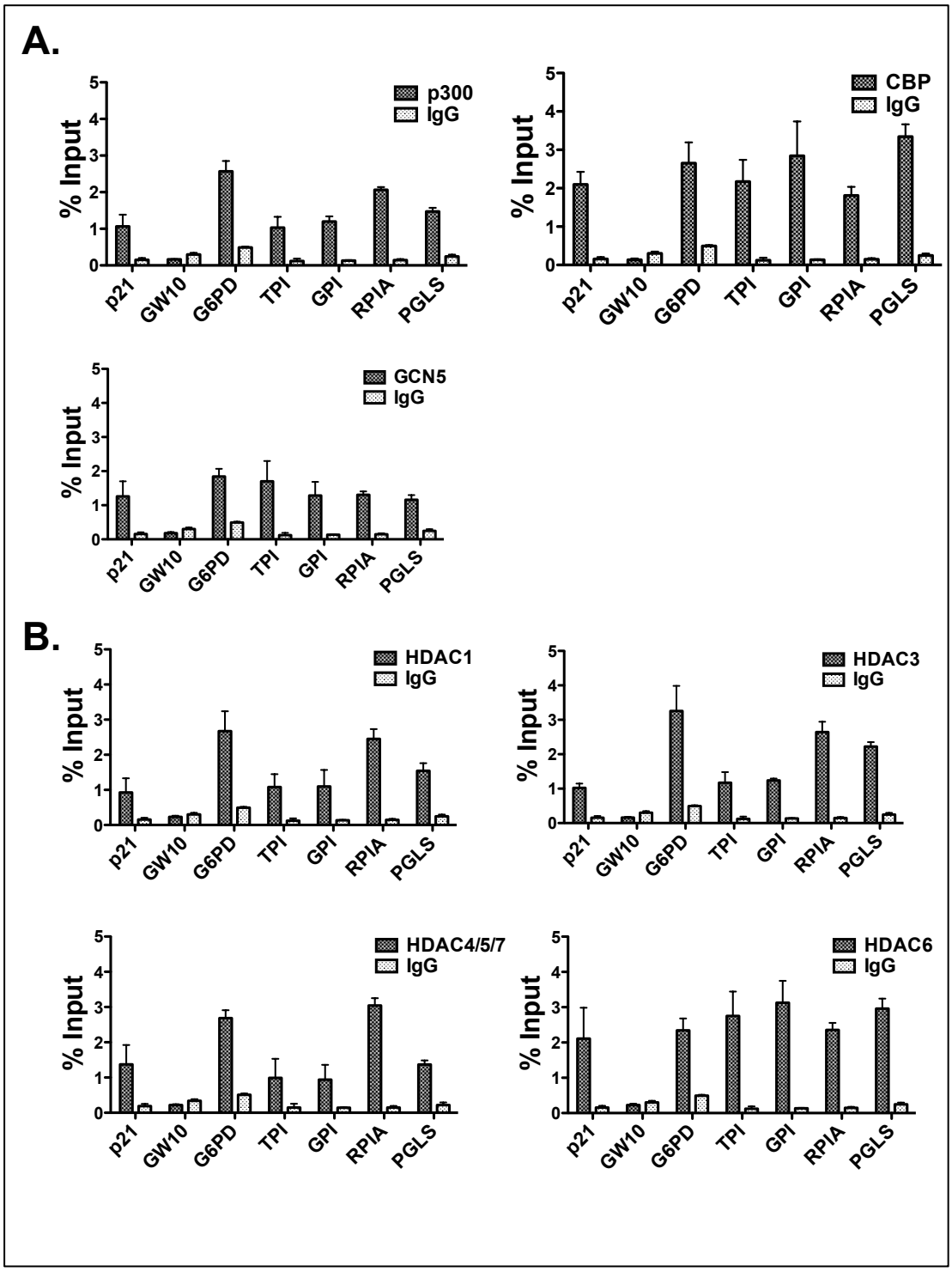


Figure 5-9 HAT and HDAC binding on GPPP gene promoters at baseline. (A) HAT and **(B)** HDAC binding assessed by ChIP in the WT B cell line P277 at baseline followed by RT-qPCR using primers amplifying the core promoters of the GPPP genes (In the case of G6PD, that is primer set 2 as shown in **Figure 5-5**). Enrichment on the promoters is calculated as % of input. *p21* is used as a positive control for enrichment and GW10 is the negative control at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3.

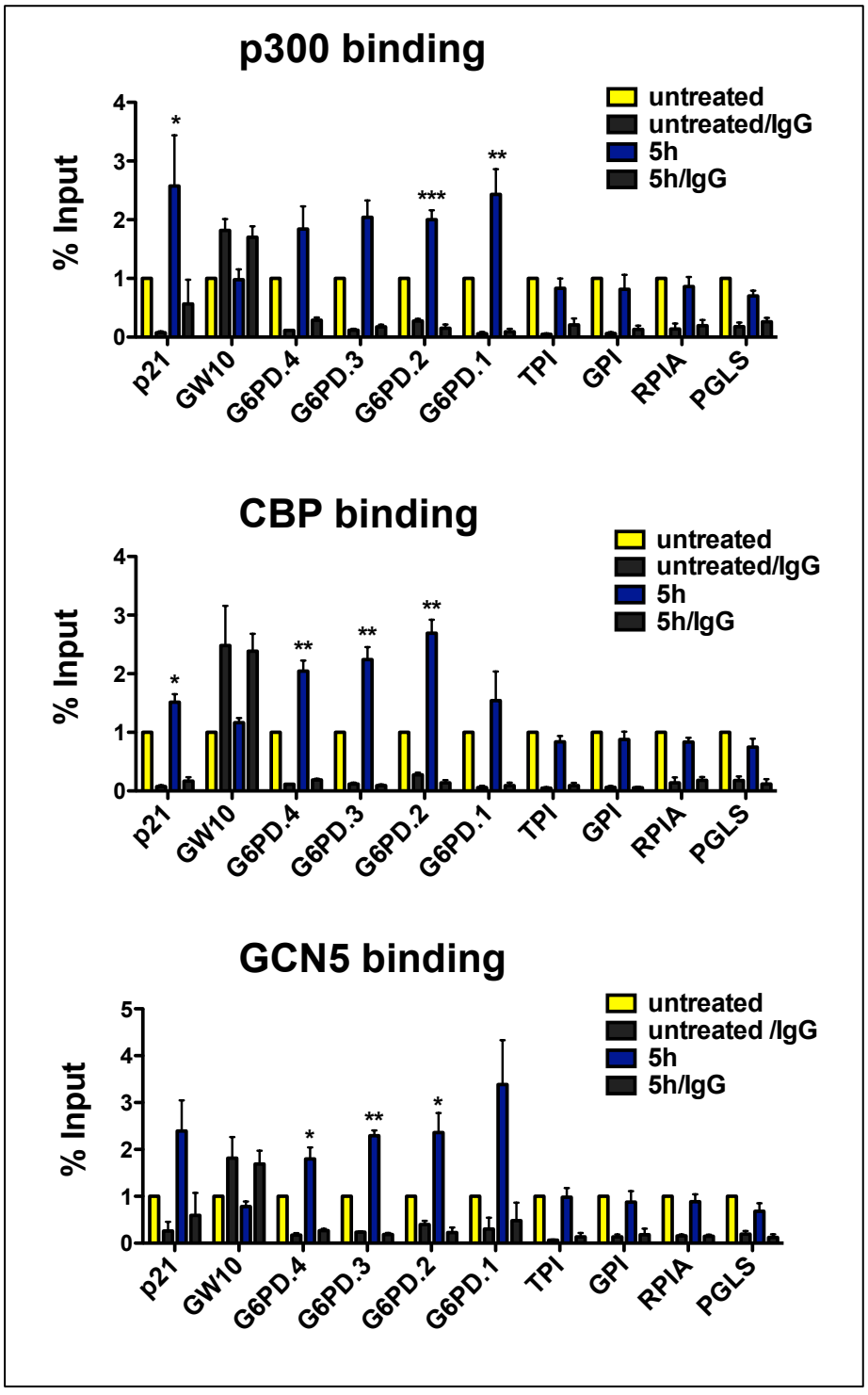


Figure 5-10 HAT binding on the GPPP gene promoters in WT B cells in the presence of NaBu. p300, CBP and GCN5 binding assessed by ChIP in the WT B cell line P277 upon 5h NaBu treatment. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *p21* is used as a positive control for enrichment and GW10 is the negative control for binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to 5h treatment.

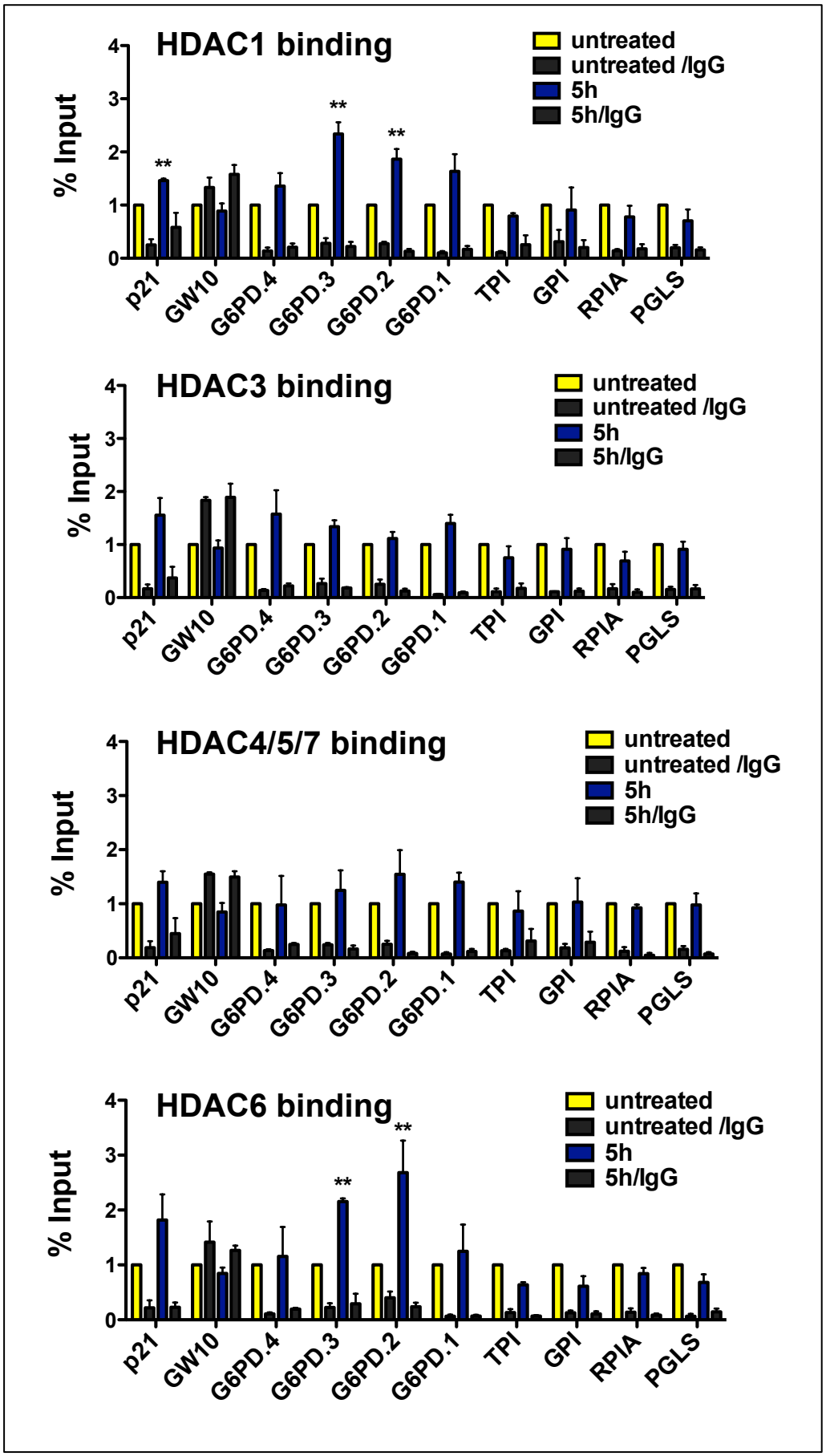

HDAC1 binding
% Input
untreated
untreated /IgG
5h
5h/IgG
**
**
**
p21
GW10
G6PD.4
G6PD.3
G6PD.2
G6PD.1
TPI
GPI
RPIA
PGLS
HDAC3 binding
% Input
untreated
untreated /IgG
5h
5h/IgG
HDAC4/5/7 binding
% Input
untreated
untreated /IgG
5h
5h/IgG
HDAC6 binding
% Input
untreated
untreated /IgG
5h
5h/IgG
**
**

Figure 5-11 HDAC binding on the GPPP gene promoters in WT B cells in the presence of NaBu. HDAC1, 3, 4/5/7 and 6 binding assessed by ChIP in the WT B cell line P277 upon 5h NaBu treatment. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *p21* is used as a positive control for enrichment and GW10 is the negative control for binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to 5h treatment.

I next determined HAT and HDAC binding on the promoters of GPPP genes in G6PD deficient cells in response to NaBu. I therefore treated the G6PD Brighton deficient cell line for 5h with NaBu and then immunoprecipitated against the HATs p300, CBP and GCN5 as well as the HDACs 1, 3, 4/5/7 and 6. As shown in Figure 5-12 and Figure 5-13, as with the normal cells, GPPP gene promoters are co-occupied by HATs and HDACs at baseline, which would be expected to dynamically regulate transcription. HAT binding is significantly upregulated on the G6PD promoter, but not any other GPPP gene promoters; HDAC1 and HDAC6 binding is also increased on the *G6PD* promoter, whereas the rest of the HDACs show the same degree of binding after treatment. Of note, *p21* is a gene that is upregulated upon HDACI treatment (Gui et al., 2004) and here shows the same characteristics as *G6PD* in response to NaBu.

It should be also noted here that because the sonication fragments obtained for the purposes of the ChIP assays are between 200-400bp, it is very challenging to distinguish binding in regions that are very close to each other; however, it still allows us to understand the general binding pattern in the promoter region. Consequently, the use of more than one primer pairs top cover the *G6PD* promoter and this aims at getting a better understanding of the features of the promoter as a whole, rather than to separately characterise fragments of it.

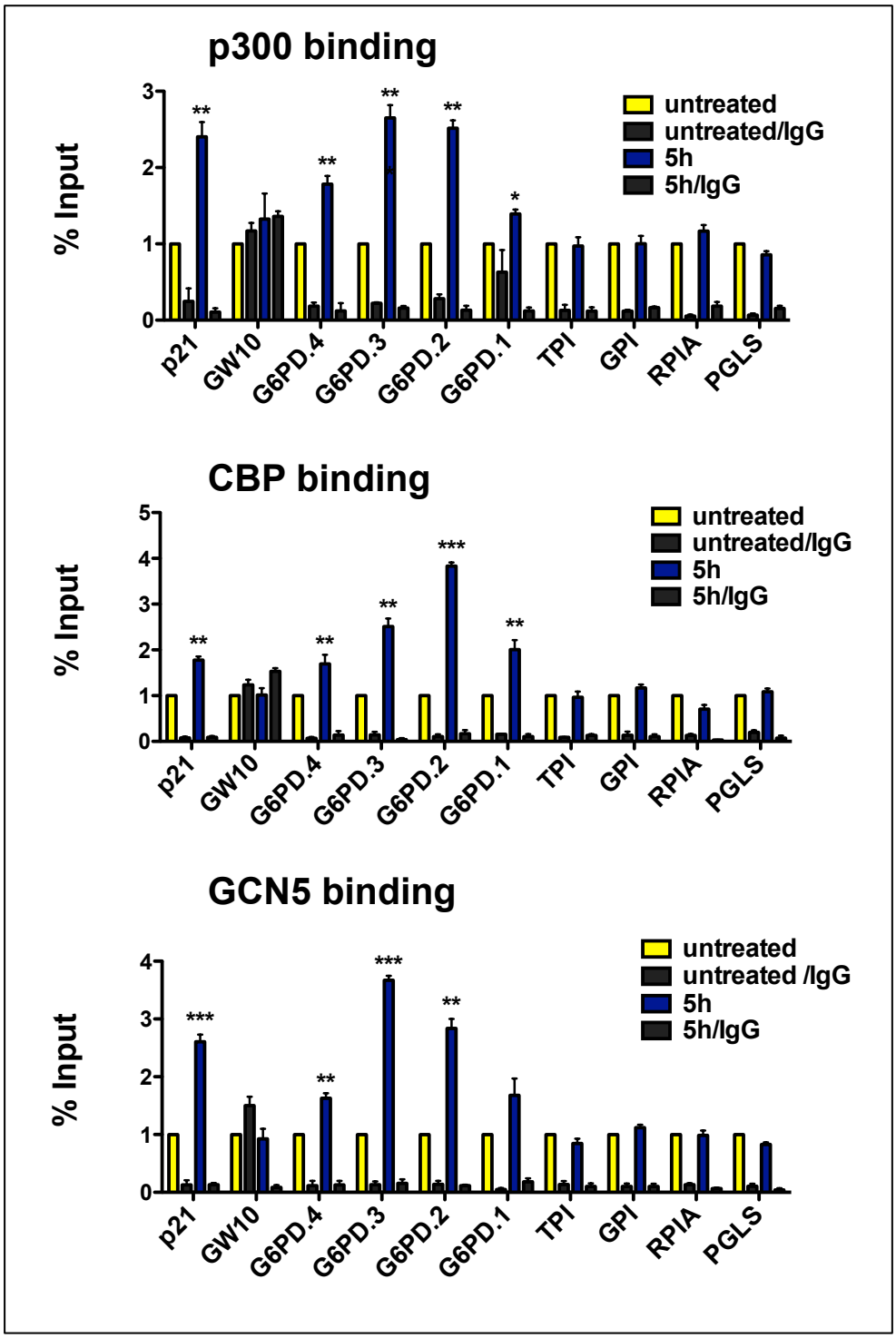


Figure 5-12 HAT binding on the GPPP gene promoters in G6PD deficient B cells in the presence of NaBu. p300, CBP and GCN5 binding assessed by ChIP in the G6PD deficient B cell line P7 upon 5h NaBu treatment. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. p21 is used as a positive control for enrichment and GW10 is the negative control for binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to 5h treatment.

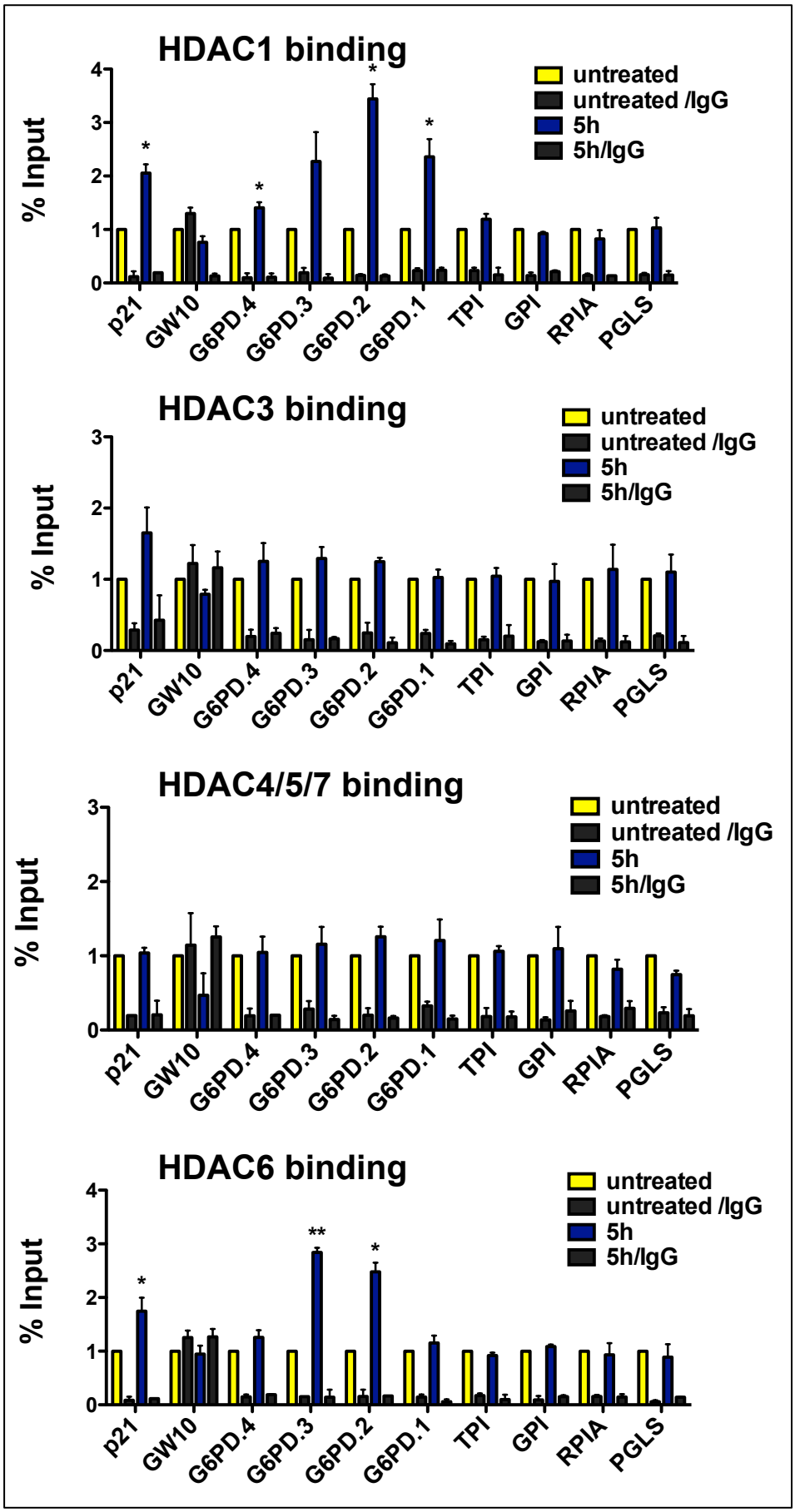

HDAC1 binding
% Input
untreated
untreated /IgG
5h
5h/IgG
p21
GW10
G6PD.4
G6PD.3
G6PD.2
G6PD.1
TPI
GPI
RPIA
PGLS
HDAC3 binding
HDAC4/5/7 binding
HDAC6 binding

Figure 5-13 HDAC binding on the GPPP gene promoters in G6PD deficient B cells in the presence of NaBu. HDAC1, 3, 4/5/7 and 6 binding assessed by ChIP in the G6PD deficient B cell line P7 upon 5h NaBu treatment. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *p21* is used as a positive control for enrichment and GW10 is the negative control for binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to 5h treatment.

Linking these selective epigenetic events with the increased *G6PD* mRNA levels following HDACI treatment (described in Chapters 3 and 4), I decided to investigate polymerase II (PolII) binding. Hence, I employed ChIP analysis and used an antibody against the N-terminus of polymerase II in order to capture binding of both its transcription initiating and elongating forms.

ChIP analysis revealed enhancement of PolII binding to the core and wider promoter of *G6PD*, as well as within the gene body (primers GB1 and GB2), an increase that was not observed in other GPPP genes (Figure 5-14). This finding indicates that amongst genes of the GPPP, the selective increase of *G6PD* transcription in response to HDACIs is underpinned by increased recruitment of HATs, HDACs, histone hyper-acetylation and increased PolII recruitment. Figure 5-15 shows PolII recruitment at baseline for *G6PD* and other GPPP gene promoters as well as the *G6PD* gene body, which appears to be at similar levels in all genes tested in both cell lines.

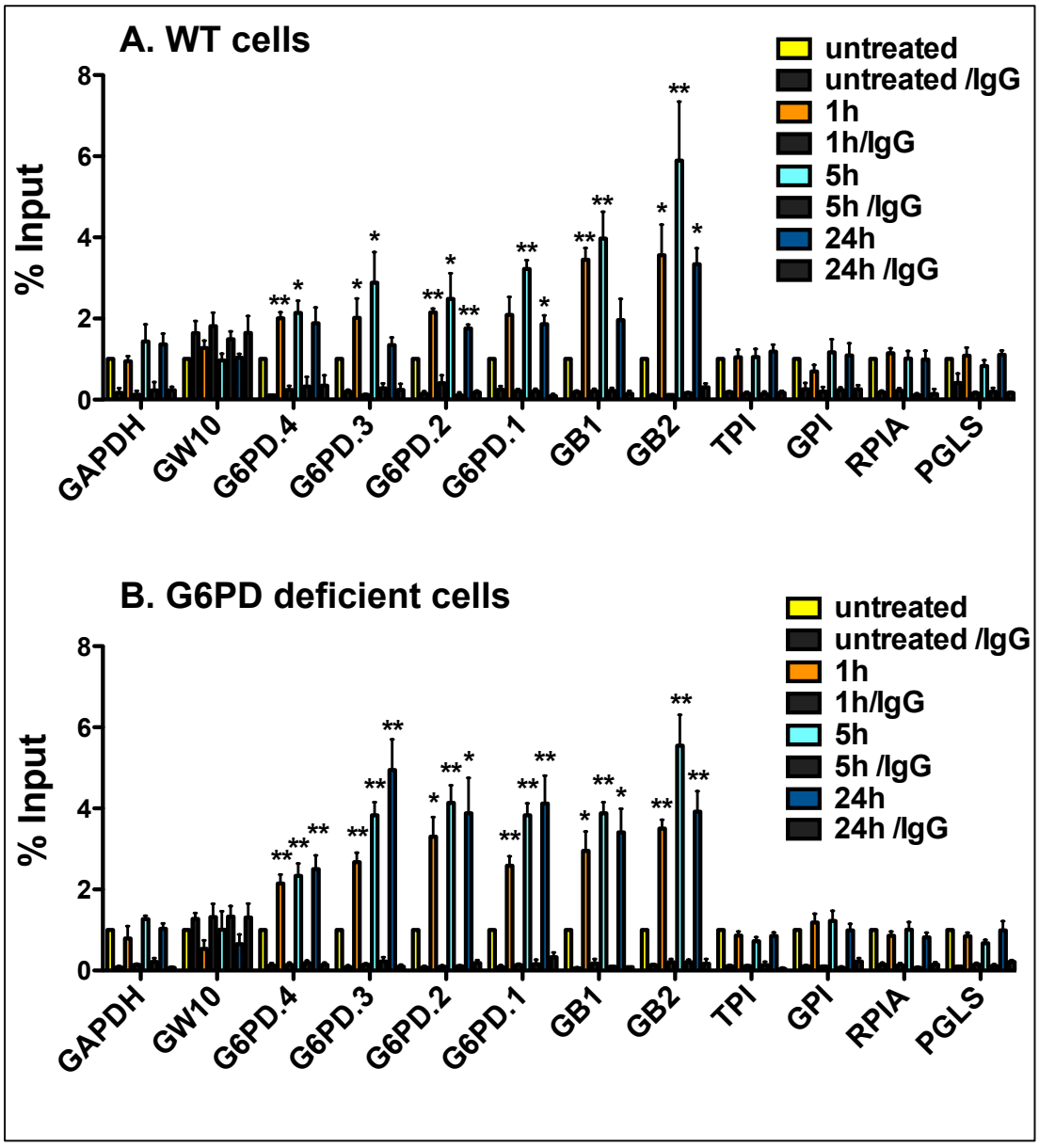


Figure 5-14 Polymerase II binding on the GPPP genes in response to NaBu. Polymerase II binding assessed by ChIP in the **(A)** WT and **(B)** G6PD deficient B cell lines P277 and P7, respectively upon 5h NaBu treatment. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *GAPDH* is used as a positive control for enrichment and GW10 is the negative control for binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to 5h treatment.

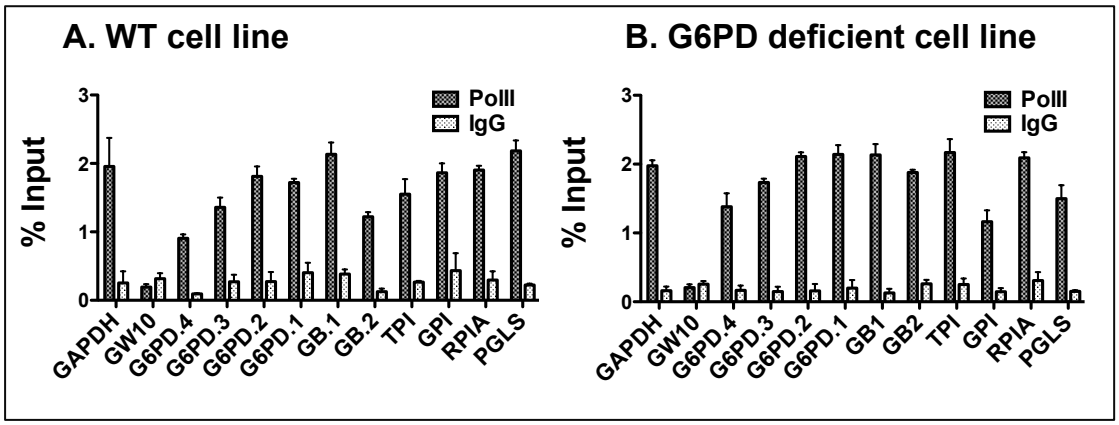


Figure 5-15 Polymerase II recruitment on the GPPP gene promoters and *G6PD* gene body at baseline. PolII recruitment assessed by ChIP on the promoters of *G6PD* and other GPPP genes (*TPI, GPI, RPIA* and *PGLS*) as well as the *G6PD* gene body. Baseline acetylation is shown in the **(A)** WT B cell line and **(B)** G6PD deficient cell line. Enrichment on the promoters is calculated as % of input. *GAPDH* is used as a positive control for enrichment and GW10 is the negative control for PolII at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3.

5.5 Sp1 binding on the GPPP promoters

To test my initial hypothesis that the Sp1-dependent control of histone acetylation and transcriptional activation characterised in IGD (Almeida et al., 2006; 2007; Caputo et al., 2013) can be extended to the genes of the GPPP, I performed ChIP analysis to assess Sp1 binding on the GPPP gene promoters upon treatment with NaBu. Figure 5-16 shows that Sp1 promoter occupancy is similar amongst all 5 GPPP genes tested. However, Figure 5-17 shows that in response to NaBu, Sp1 binding increased only at the promoter of *G6PD* of both normal (Figure 5-17A) and G6PD deficient cells (Figure 5-17B). This upregulation in the Sp1 recruitment is *G6PD*-specific. These findings suggest that the selective increase in Sp1 binding at the *G6PD* promoter in response to HDACIs is required for transcriptional upregulation of *G6PD*.

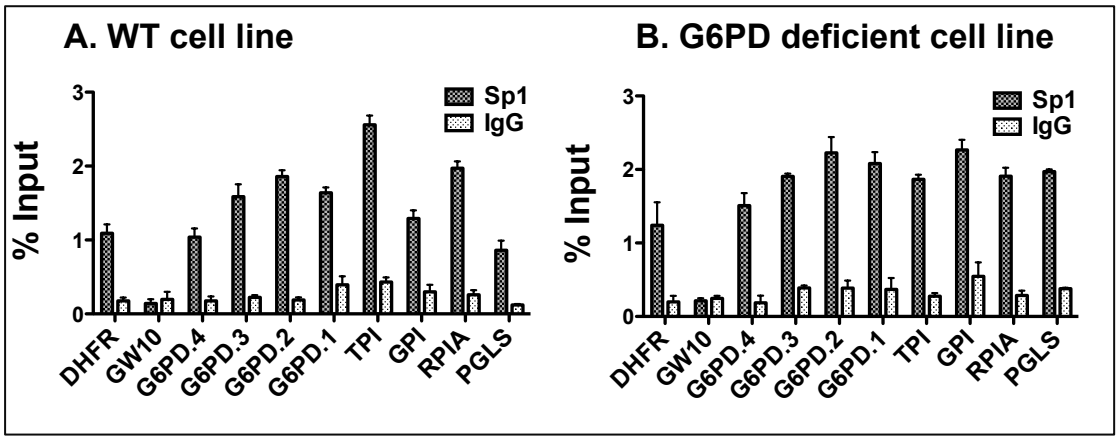


Figure 5-16 Sp1 recruitment on the GPPP gene promoters at baseline. Sp1 recruitment assessed by ChIP on the promoters of *G6PD* and other GPPP genes (*TPI, GPI, RPIA* and *PGLS*). Baseline acetylation is shown in the **(A)** WT B cell line and **(B)** G6PD deficient cell line. Enrichment on the promoters is calculated as % of input. *DHFR* is used as a positive control for enrichment and GW10 is the negative control for Sp1 binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3.

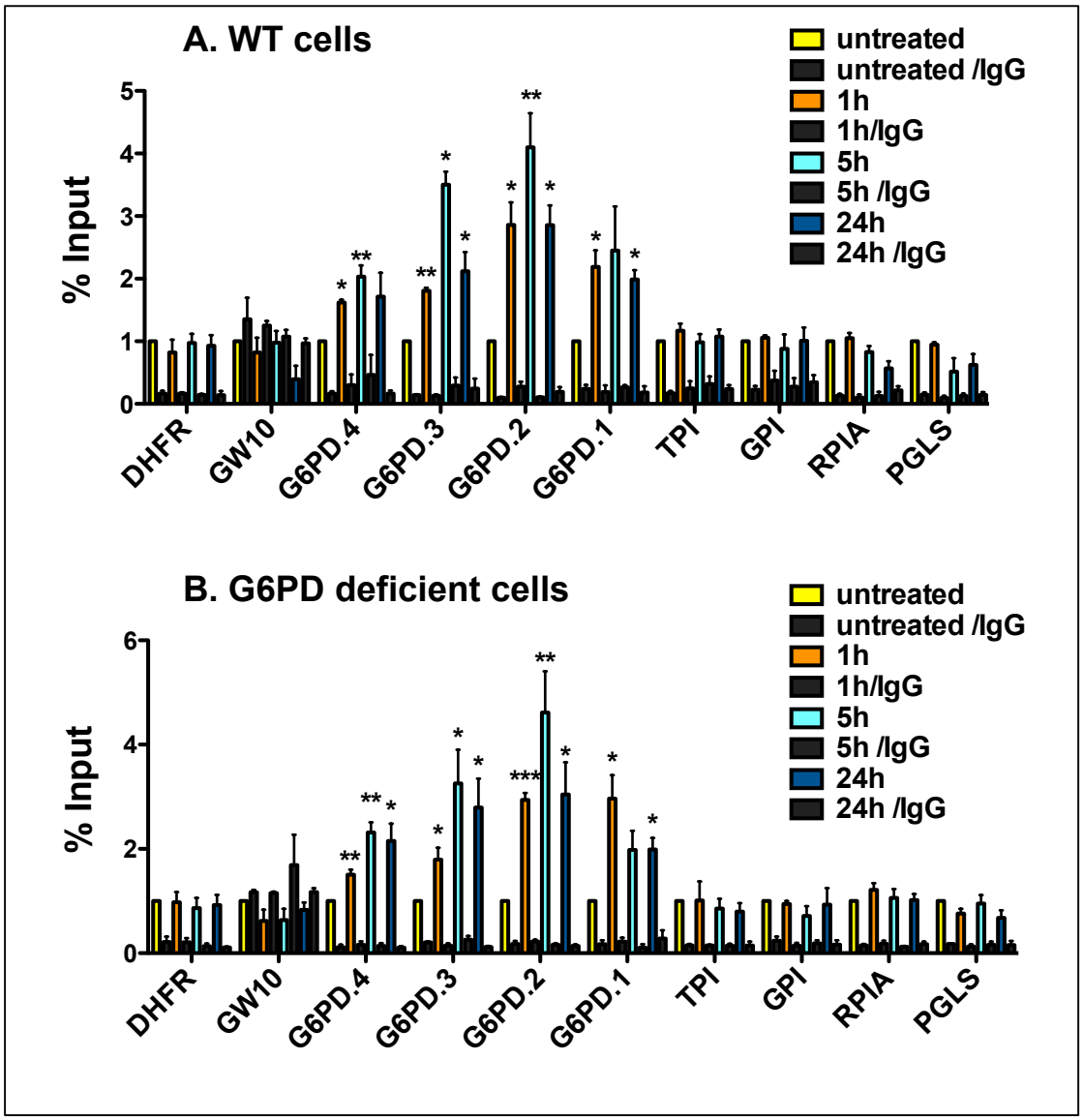


Figure 5-17 Sp1 binding on the GPPP gene promoters in response to NaBu. Sp1 binding assessed by ChIP in the **(A)** WT and **(B)** G6PD deficient B cell lines P277 and P7, respectively upon 5h NaBu treatment. Enrichment on the promoters is calculated as % of input relative to the untreated for each primer set. *DHFR* is used as a positive control for enrichment and GW10 is the negative control for binding at baseline. Immunoprecipitation against IgG is employed to set the enrichment background. Mean and S.E.M. are shown for n=3. Student's t-test was performed to compare the untreated to 5h treatment.

5.6 Conclusions

In this chapter my aim was to dissect the transcriptional and epigenetic basis of the selective responsiveness of *G6PD* to HDACIs amongst the genes of the GPPP. Therefore, I compared the transcriptional and epigenetic status of the *G6PD* promoter to other genes of the GPPP.

Bioinformatic analysis of the GPPP gene promoters using ENCODE, TFsearch and CONSITE databases showed that the *G6PD* and other GPPP gene promoters contain features characteristic of housekeeping genes. Specifically, they are highly acetylated on histones 3 and 4, they lack silencing chromatin marks and are co-occupied by HATs and HDACs indicating that they are dynamically regulated. Additionally, all the genes tested showed considerable enrichment for Sp1 binding. However, even though Sp1 binding sites are normally abundant on butyrate-responsive promoters, only *G6PD* out of the 17 GPPP genes responds to HDACIs by transcriptional upregulation. According to the bioinformatics analysis the most striking difference between *G6PD* and the other GPPP gene promoters is the fact that *G6PD* shares a bidirectional promoter with *NEMO*. Nevertheless, *NEMO* mRNA expression remains unaffected by NaBu treatment.

ChIP analysis was also employed to determine the epigenetic status of the GPPP gene promoters at baseline and in response to NaBu. In agreement with the bioinformatics analysis, the ChIP experiments showed high levels of histone acetylation and PolII, HAT, HDAC and Sp1 binding at baseline. However, the epigenetic status of the *G6PD* promoter was the only one to change in response to HDAC inhibition. Consistent with increased transcription, the *G6PD* promoter was hyper-acetylated in response to NaBu treatment and increased PollII was recruited to the promoter and gene body of *G6PD*. Furthermore, binding of the HATs CBP, p300 and GCN5 was increased as well as binding of HDAC1 and HDAC6. The selective increase of these features underpins the increased transcriptional activity on the *G6PD* promoter. It is interesting to note that HDAC inhibition led to a commensurate increase of both HATs and HDACs suggesting that, similar to baseline conditions, their dynamic balance is required to sustain increased transcription following HDACI treatment. Finally Sp1 binding is selectively increased on the *G6PD* promoter in response to NaBu, suggesting that Sp1 is central to transcriptional upregulation in response to HDACIs. This will be further examined in Chapter 6.

6 Results IV: Role of the transcription factor Sp1 on *G6PD* transcriptional activation in response to HDAC inhibitors

6.1 Introduction

It is well known that HDAC inhibition causes widespread histone hyper-acetylation, yet affects the transcription of only 2-20% of genes (Davie, 2003; Mitsiades et al., 2004; Sealy and Chalkley, 1978; Van Lint et al., 1996). Factors that determine the number of genes altered in transcription involve the length of exposure, the concentration, the type of HDACI and the type of cells used (Peart et al., 2005). For example, in one study, it was reported (Daly and Shirazi-Beechey, 2006) that butyrate upregulates and downregulates equal number of genes in colonic epithelial cells. In line with this observation, another study (Rada-Iglesias et al., 2007) showed that as opposed to the global increase in chromatin acetylation, a number of genomic regions close to transcription start sites corresponding to genes that were downregulated under butyrate exposure, were deacetylated.

To date, very little is known about the global gene expression of erythroid cells upon HDACI treatment. In fact, almost all the genome-wide studies conducted so far to evaluate gene expression changes in response to HDACIs involved cell lines rather than primary human cells. The sole study that has been conducted in primary cells was published by Wang and colleagues (Wang et al., 2008, 2009; 1.1.3), in which primary human $CD4^+$ T cells were treated with NaBu and TSA to evaluate global chromatin changes and gene expression in response to HDACIs.

6.1.1 Aim of the chapter

The aim of this chapter is to identify the genome-wide gene expression changes in response to NaBu in primary human erythroblasts. Analysis of the promoters of upregulated versus downregulated and stably expressed genes might allow identifying genetic and epigenetic features that are responsible for the selective upregulation of G6PD amongst all the 17 genes of the GPPP. Furthermore, identification of genes of medical interest that are upregulated by NaBu might offer new therapeutic opportunities.

6.1.2 Experimental plan

To study the impact of HDAC inhibition on global gene expression in primary human erythroid cells, I first generated proerythroblasts from CB-$CD34^+$ cells, as described in 4.2. These were treated with NaBu and then subjected to gene expression profiling (GEP) analysis. The results were compared with other recent genome-wide

studies conducted in human primary $CD4^+$ T cells (Wang et al., 2009) and mouse pancreatic cell lines (Kubicek et al., 2012) which, were treated with HDACIs. Subsequently, *de novo* motif discovery was performed on the promoters of the upregulated versus the downregulated and genes that do not change in order to identify the *in cis* DNA features driving the selective upregulation of specific genes, including *G6PD*. Upon identification of a candidate motif, further transactivation and mutagenesis assays were performed to establish its functional importance.

6.2 Genome-wide implications of HDAC inhibition in primary proerythroblasts

6.2.1 Preparation of samples

To dissect the genome-wide implications of HDAC inhibition induced by NaBu in primary erythroid cells, I performed GEP analysis on CB-$CD34^+$-differentiating proerythroblasts. Specifically, three independent CB samples were processed and $CD34^+$ cells were differentiated down the erythroid lineage as described in 4.3.2. At day 5 of the culture (Figure 6-1), flow cytometric analysis was employed to identify the stage of differentiation by gating the $CD36^+$ cells and assessing CD71 and GlyA staining. For each sample $5x10^5$ out of 10 $x10^5$ of the cells were treated for 6h with 1mM NaBu and the other half remained in culture untreated. Both the untreated and treated cells were then CD36-selected.

RNA was extracted from the six samples and the RNA quality and integrity were assessed using the Nanodrop and the Agilent 2100 Bioanalyser. The Nanodrop is used to calculate RNA concentration and the 260/280 purity ratio. Nucleic acids and proteins absorb at 260nm and 280nm wavelengths, respectively. Traditionally, a ratio of ~2.0 is accepted as "pure" RNA. All six samples (Table 6-1) had acceptable 260/280 ratios. Furthermore, the Agilent 2100 bioanalyser was used to calculate the RNA Integrity Number (RIN) algorithm, which is an additional quality control for eukaryotic total RNA. The RIN is calculated based on the fluorescence of the 18S and 28S rRNA and values equal to 10 indicate intact RNA. As shown in Table 6-1 and Figure 6-2, all six samples used for gene expression profiling have RIN equal or above 8, which lies within the acceptable range. Importantly, in Figure 6-2B, which visualises the samples run in an acrylamide gel, it is shown clearly that all samples are of good quality and are not degraded.

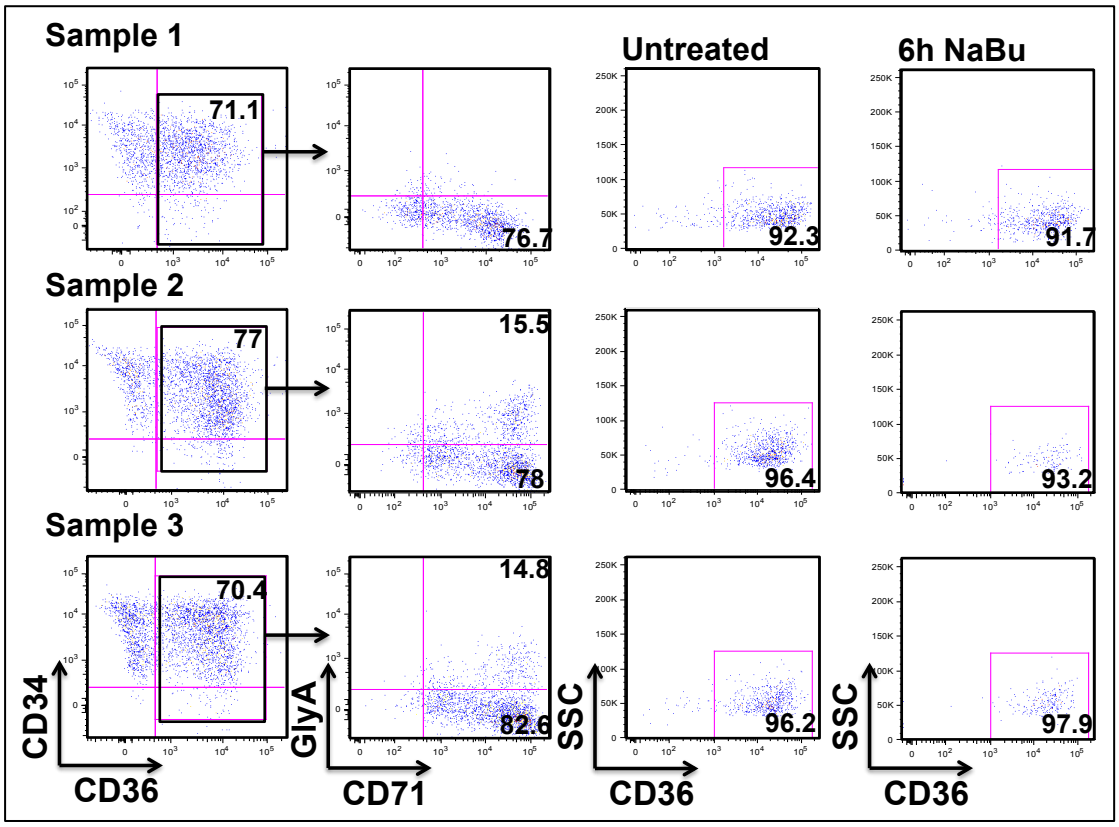


Figure 6-1 Flow cytometric analysis of samples used for GEP. Three CB-CD34^{+} samples were used and subjected to erythroid differentiation. At day 5 flow cytometry assay was performed to assess the stage of differentiation by anti -CD34, -CD36, -CD71 and -GlyA staining. Each sample was split in half; half of the cells were treated for 6h with 1mM NaBu and the other half remained in culture untreated. Both were then CD36-selected and the selection purity was again tested by flow cytometric analysis.

Table 6-1 Quality controls for microarray RNA samples.

	Nanodrop		Bioanalyser
	RNA concentration (ng/µl)	RNA 260/280 ratio	RNA integrity number
Sample 1 untreated	74.5	2.00	8.00
Sample 1 NaBu	68.6	2.07	8.70
Sample 2 untreated	112.8	1.94	9.60
Sample 2 NaBu	96.7	1.99	9.90
Sample 3 untreated	97	2.00	9.90
Sample 3 NaBu	89	2.17	9.80

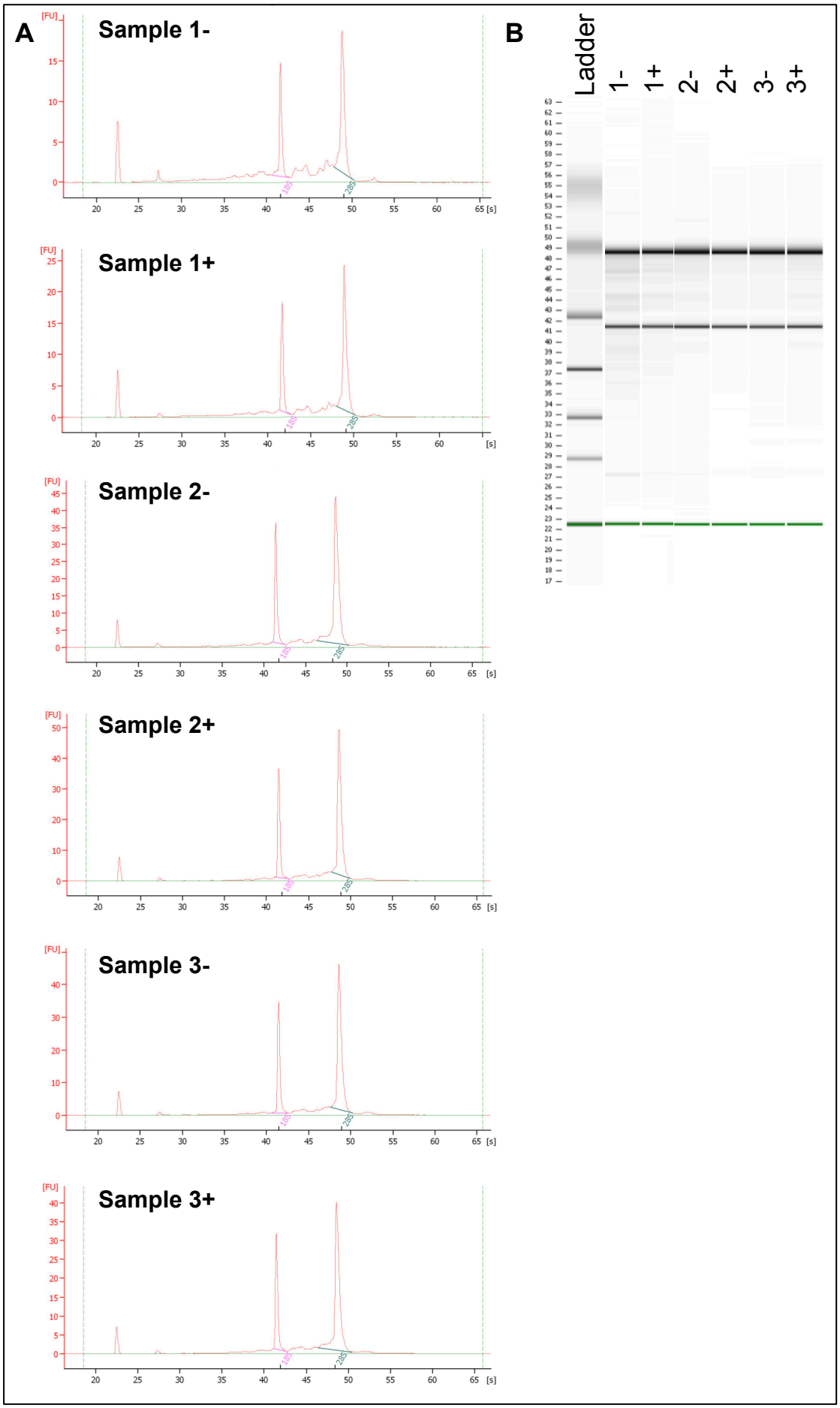

A
Sample 1-
Sample 1+
Sample 2-
Sample 2+
Sample 3-
Sample 3+
[FU]
[s]
18S
28S
B
Ladder
1-
1+
2-
2+
3-
3+

Figure 6-2 RNA quality controls using the bioanalyser. Plots as taken from the Agilent 2100 bioanalyser showing the quality of the RNA used for microarray analysis. **(A)** Graphs represent the fluorescence against time (in seconds) for the marker (1^{st} peak), the 18S rRNA (2^{nd} peak) and the 28S rRNA (3^{rd} peak). **(B)** Visualisation of the ladder and samples run on an acrylamide gel.

Furthermore, expression of *α-* and *β-globin* was tested by RT-qPCR to decide whether it was necessary to eliminate the expressed globin before further processing the samples. It is common when conducting GEP in erythroid cells to eliminate the *globin* mRNA, as it might interfere with the accuracy of the results due to its very high expression at the later stages of erythropoiesis. However, in this set of experiments, the erythroid cells obtained at day 5 are early erythroblasts and therefore do not yet express high levels of haemoglobin. In Figure 6-3, the Ct values obtained by RT-qPCR for the adult *α-globin* and *β-globin* are shown in comparison to *G6PD* and *GAPDH*. Conventionally, the Ct value is the number of cycles for the fluorescent signal to cross the threshold cycle and therefore exceed the background levels. Higher Ct values correspond to lower expression. The Ct values of *α-globin* and *β-globin* shown in Figure 6-3 are higher than those of *GAPDH* indicating that the expression of the globin genes at this stage of erythroid differentiation is relatively low thus negating the need to remover *globin* gene mRNA.

Once the RNA quality was established, 1µg total RNA per sample was used to hybridise onto Affymetrix GeneChIP 2.0 ST Arrays in collaboration with Dr. Robert Geffers at the Helmholtz Centre for Infection Research, Braunschweig, Germany.

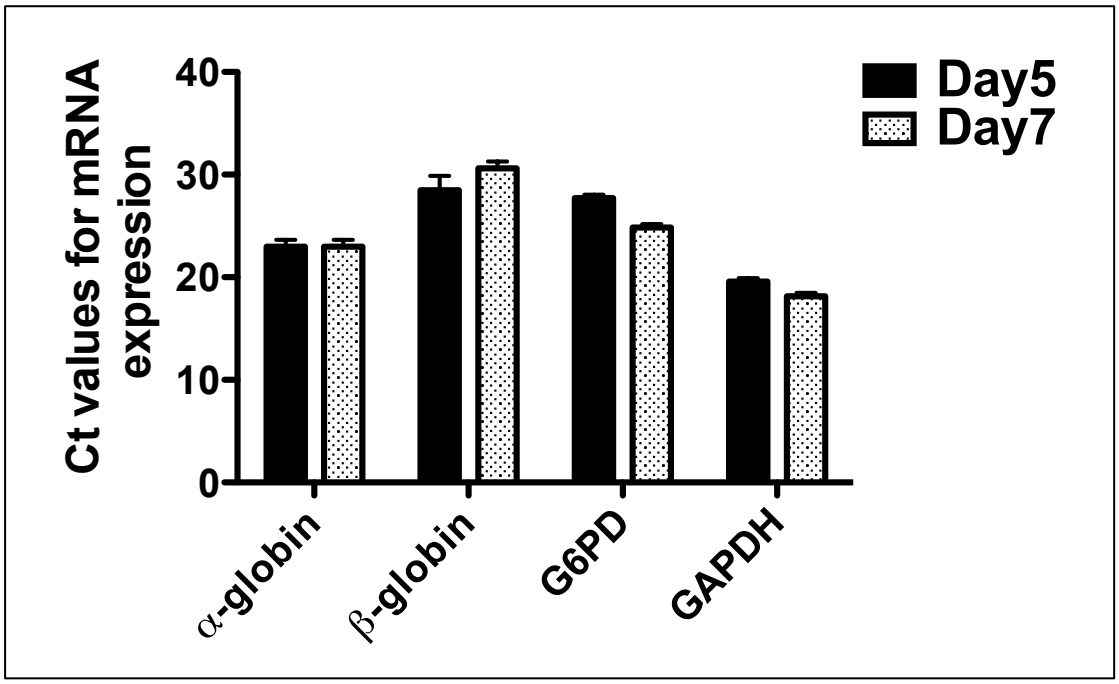


Figure 6-3 RT-qPCR analysis of *α-globin* and *β-globin* expression in erythroid cells of day 5 and day 7. Ct values obtained by RT-qPCR for the expression of *α-* and *β-globin* in comparison to *G6PD* and *GAPDH*. The Ct values of *α-* and *β-globin* are comparable to *G6PD* and *GAPDH* and do not exhibit higher expression that them. Mean and S.E.M. are shown for n=3.

6.2.2 Genome expression profiling

GEP analysis of the microarray data revealed an equal number of upregulated versus downregulated genes (Table 6-2; Appendix A and B). Out of 19,137 annotated genes, only 607 and 528 genes were found to be significantly upregulated and downregulated, respectively. To estimate the number of upregulated and downregulated genes statistical significance was set to $p \leq 0.05$ and fold change of ≥ 1.5 was used as a threshold. Similar to previously reported data (Daly and Shirazi-Beechey, 2006; Rada-Iglesias et al., 2007; Wang et al., 2009) I found that 3.17% were upregulated and 2.76% of genes were downregulated, whereas the majority of genes did not show significantly altered expression after 6 hours of NaBu treatment of phase 1 proerythroblasts.

Table 6-2 Number of genes upregulated and downregulated according to GEP analysis of microarray data. [a]

	Number of genes	P value and Fold Change	Percentage
Upregulated genes	607	p≤0.05 FC≥1.5	3.17%
Downregulated genes	528	p≤0.05 FC≥1.5	2.76%
Total genes	19,137		

[a] GEP analysis was conducted in human primary proerythroblasts derived from CD-CD34+ cells, which were treated for 6h with 1mM NaBu. Out of a total of 19,137 genes, 2,283 and 2,039 genes were significantly upregulated and downregulated, respectively.

I next utilised the microarray data to analyse the expression of the GPPP genes upon NaBu treatment. As I have already shown in 3.3 and 4.4, against my initial hypothesis, experiments employing RT-qPCR that were conducted in B cell lines and CB-$CD34^+$ - differentiating cells revealed that amongst the 17 genes of the GPPP only mRNA levels of G6PD increased in a time-dependent fashion in response to NaBu. This finding was confirmed in GEP analysis of proerythroblasts generated from CB-$CD34^+$ cells that were treated with 1mM NaBu for 6h (Figure 6-4A). Out of 17 genes of the GPPP, microarray analysis showed that only *G6PD* is significantly upregulated 1.6-fold post-treatment. This finding was also confirmed by bioinfomatic analysis I conducted on previously published data by Wang and colleagues (Wang et al., 2009), who treated primary human $CD4^+$ T cells with a combination of 100ng/ml TSA and 2mM NaBu for 12 hours (Figure 6-4B). These expression data sets can be found in the GEO database under the accession number GSE15735 (Wang et al., 2009). I next performed bioinfomatic analysis on data published by Kubicek and colleagues (Kubicek et al., 2012), who treated the mouse pancreatic cell line a-TC1 with either 1.3μM SAHA, or 0.015μM panobinostat or 0.22μM TSA for 6h and then performed Affymetrix microarray gene expression analysis (Figure 6-4C). Figure 6-4C shows that even in mouse cells, *G6pdx* expression significantly is upregulated upon SAHA and panobinostat treatment. However, in mice the expression of *Hk1, Bpgm, Pklr, Pgd* and *Taldo1* is also upregulated in the presence of certain HDACIs.

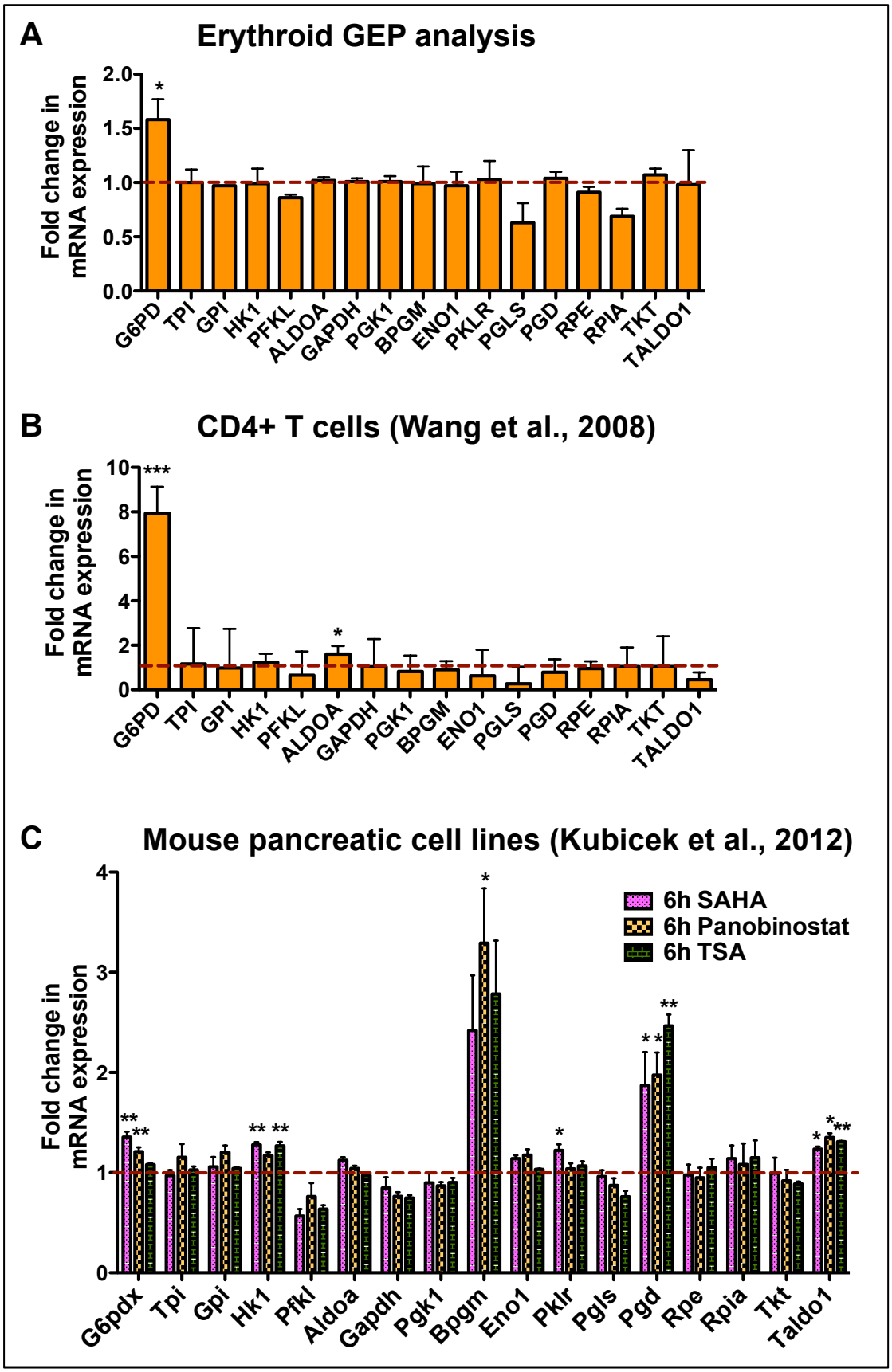

A
Erythroid GEP analysis
Fold change in mRNA expression
2.0
1.5
1.0
0.5
0.0
G6PD
TPI
GPI
HK1
PFKL
ALDOA
GAPDH
PGK1
BPGM
ENO1
PKLR
PGLS
PGD
RPE
RPIA
TKT
TALDO1
B
CD4+ T cells (Wang et al., 2008)
Fold change in mRNA expression
10
8
6
4
2
0
G6PD
TPI
GPI
HK1
PFKL
ALDOA
GAPDH
PGK1
BPGM
ENO1
PGLS
PGD
RPE
RPIA
TKT
TALDO1
C
Mouse pancreatic cell lines (Kubicek et al., 2012)
Fold change in mRNA expression
6h SAHA
6h Panobinostat
6h TSA
4
3
2
1
0
G6pdx
Tpi
Gpi
Hk1
Pfkl
Aldoa
Gapdh
Pgk1
Bpgm
Eno1
Pklr
Pgls
Pgd
Rpe
Rpia
Tkt
Taldo1

Figure 6-4 GEP analysis of the GPPP genes upon HDACI treatment. (A) Fold change in mRNA expression of GPPP genes shown by GEP analysis of microarray data generated from CB-$CD34^+$-differentiating proerythroblasts after 6h treatment with 1mM NaBu. **(B)** Bioinformatic analysis on data published by Wang et al., 2008. $CD4^+$ T cells were treated with 100ng/ml TSA and 2mM NaBu for 12h. Fold change in mRNA expression of GPPP genes is shown. **(C)** Bioinformatic analysis on data published by Kubicek et al., 2012. Mice pancreatic cell lines were treated with 1.3μM SAHA, 0.015μM Panobinostat and 0.22μM TSA for 6h. Fold change in mRNA expression of GPPP genes is shown. Mean and S.E.M. are shown for n=3 independent microarray experiments. One-way ANOVA has been performed to compare the pre- and the post- treatment values.

I next aimed to use the microarray data I generated from primary human proerythroblats to identify factors that selectively drive upregulation or downregulation of specific genes, such as *G6PD* amongst the GPPP genes. For this purpose, I first compared the baseline expression of the genes that were significantly upregulated and downregulated (as described in Table 6-2) as well as of 2,000 more stably expressed genes after the treatment with NaBu. As shown in Figure 6-5A, the baseline expression of genes that are upregulated by NaBu is lower than the baseline expression of the genes that are downregulated. This could explain why some genes have the potential to further increase their expression as opposed to other genes, albeit the genes that are unaffected display the lower baseline expression. This interesting finding is confirmed by also analyzing the baseline expression in $CD4^+$ T cells (Figure 6-5B).

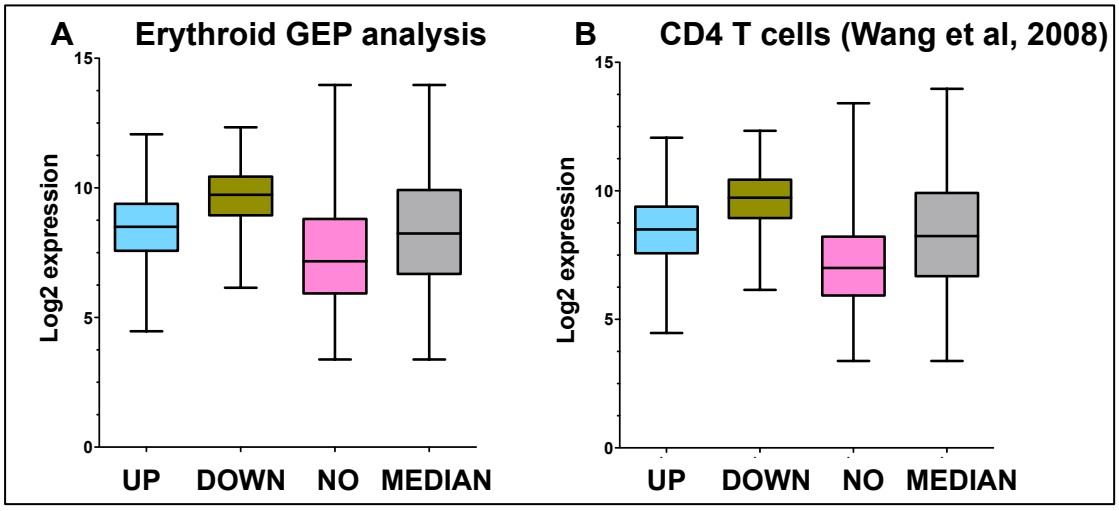


Figure 6-5 Baseline mRNA expression of upregulated versus downregulated genes. Box and whisker plots of data obtained from GEP analysis of **(A)** CB-$CD34^+$-differentiating proerythroblasts ($p<0.0001$ for comparisons between UP vs DOWN, UP vs NO, UP vs MEDIAN, DOWN vs NO, DOWN vs MEDIAN and NO vs MEDIAN) and **(B)** $CD4^+$ T cells (Wang et al., 2008; $p<0.001$ for comparisons between UP vs DOWN, UP vs NO, UP vs MEDIAN, DOWN vs NO, DOWN vs MEDIAN and NO vs MEDIAN). Boxes are shown for the upregulated (UP, n=14.25%; FC≥1.5, p≤0.05) and downregulated genes (DOWN, n=12.25%; FC≥1.5, p≤0.05), as well as 2,000 genes that remain the most unaffected (NO) by HDACI treatment. The median value represents the middle value. One-way ANOVA has been performed.

Furthermore, to search for *in cis* DNA elements that determine increased transcription in response to HDACIs, I performed *de novo* motif discovery using MEME Suite, a motif-based sequence analysis tool (http://meme.nbcr.net/meme/). For this purpose, I compared the promoter sequences of the 50 top ranked in fold change genes from my microarray data, i.e. those that were the most upregulated, including *G6PD*, the 50 lower ranked genes, i.e. those that were most dramatically dowregulated and finally 50 mid ranged genes that were identified to be the most stably expressed ones. This analysis identified two motifs that are highly enriched in the overexpressed but not in the other 2 groups of genes (Figure 6-6). Interestingly, one of the two motifs (Figure 6-6B) corresponds to a bona fide Sp1 binding sequence (Sp1 binding sequences are 5'-G/T-GGGCGG-G/A-G/A-C/T-3' or 5'G/T-G/A-GGCG-G/T-G/A-G/A-C/T-3' as mentioned in 1.1.4.2).

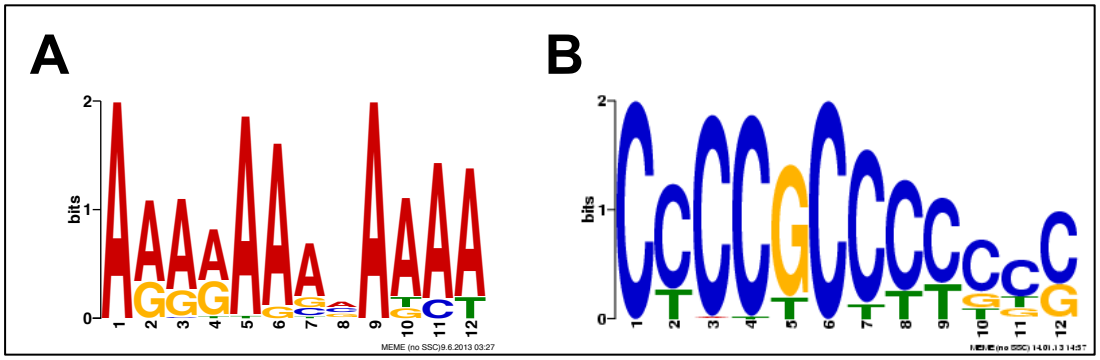


Figure 6-6 Motifs on the promoters of the genes that are responsive to HDAC inhibition. MEME Suite analysis was performed on the promoters of genes that are upregulated versus those that are downregulated or stably expressed upon NaBu treatment in primary human proerythroblasts. Motifs **(A)** and **(B)** are present on the promoter of genes that are upregulated with the latter corresponding to a bona fide Sp1 binding sequence. According to MEME instructions "the height of the motif block is proportional to –log(p value), truncated at the height for a motif with a p value of 1e-10 ". Here, p= 8.72e-09 for motif A and p=3.17e-07 for motif B.

The *G6PD* promoter has been shown to have 12 putative binding sites for the transcription factor Sp1 within its 1327bp-long wider promoter region, two of which have been previously validated (1.4.3.1 and Figure 1-12; Franzè et al., 1998; Fusco et al., 2006; Galgoczy et al., 2001; Philippe et al., 1994). In addition, I have already shown (6.3 and Figure 5-17) that NaBu selectively increases the binding of Sp1 on the *G6PD* promoter amongst other GPPP genes, which suggests a potential role of Sp1 in

the selective upregulation of *G6PD* expression. Taking this information along with the fact that Sp1 binding sites are key components of some butyrate-responsive elements, I decided to focus the next part of my research on defining the role of Sp1 on the NaBu effect.

6.3 The role of Sp1 in the regulation of G6PD transcription

6.3.1 Transcriptional activity of G6PD promoter

I next aimed to assess the importance of Sp1 in the selective effect of HDACIs on transcriptional upregulation of *G6PD*. To determine the significance of core, proximal and distal promoter in the transcriptional activity of G6PD, I performed reporter assays with amplified regions of the promoter. For this purpose, I amplified (Table 2-12 and Figure 2-3) promoter fragments of increasing length (Figure 6-7A), which were used to transfect 293T cells and measure luciferase activity of promoter regions using the approach described in 2.11. Prior to measurement, the cells were treated for 24h with 3mM NaBu or PBS. In a separate set of experiments I established that the treatment of 293T cells with 3mM NaBu for 24h is sufficient to significantly increase *G6PD* mRNA expression (Figure 6-8).

Measurement of luciferase activity driven by the amplified *G6PD* promoter fragments revealed that both at baseline and post- NaBu treatment, *G6PD* promoter activity is dependent on a 562bp promoter fragment. As shown in Figure 6-7B, constructs 4 and 5 showed the highest activity both at baseline and after treatment. On the contrary, constructs 1, 2 and 3 did not show any activity implying the promoter sequences they contain are not important for *G6PD* transcription. Construct 6 showed similar activity to construct 5, which indicates that the additional upstream DNA sequence included in construct 6 is dispensable for transcriptional control of *G6PD in vitro*. The 562bp promoter region that appears to be crucial for *G6PD* promoter activity contains 7 potential Sp1 bindings sites, including the 2 previously validated Sp1 binding motifs (Figure 6-7A, B).

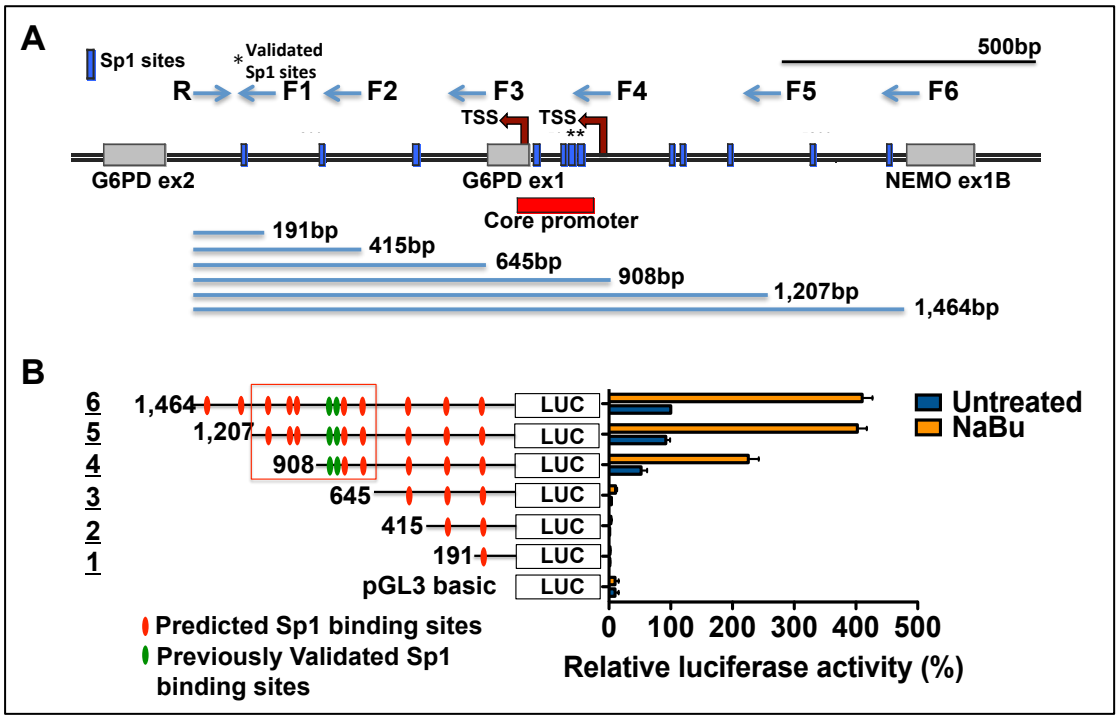


Figure 6-7 Reporter assay to determine the promoter activity potential of *G6PD*. **(A)** Representation of the wider *G6PD* promoter B, spanning between *G6PD* exon 2 and *NEMO* exon 1B. The core promoter is also highlighted. In blue are the predicted Sp1 binding sites and marked with an asterisk are the two previously validated sites. The common reverse (R) and the different forward (F1-6) primers that were used to amplify six increasing in length parts of the promoter. **(C)** Plasmid constructs containing promoter fragments were transfected into 293T cells and luciferase activity driven by these promoter parts was measured at baseline and upon 3mM NaBu treatment for 24h. Activity is normalised against pGL3 basic (without insert) and is shown relative to the baseline levels of the long fragment (6). Boxed in red is the 562bp fragment that appears to be responsible for baseline and upon HDACI activity.

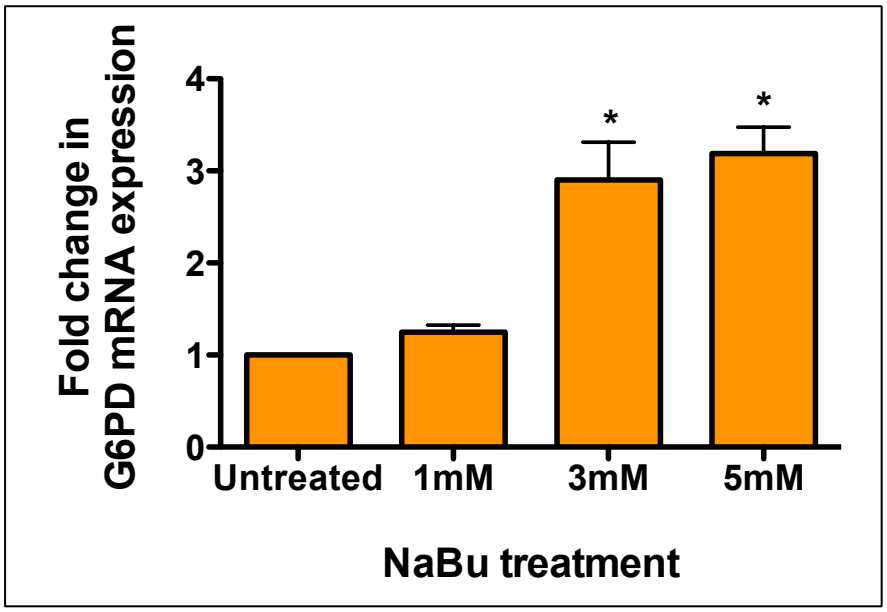


Figure 6-8 *G6PD* mRNA expression in 293T cells. RT-qPCR performed on 293T cells upon treatment with 1mM, 3mM and 5mM NaBu for 24h. Fold expression is shown. Mean and S.E.M. are shown for n=3 experiments. Student's t-test has been performed to compare untreated with treated values normalised to β-actin.

6.3.2 Dependence of the G6PD transcriptional activity on Sp1

I next proceeded to confirm the functional importance of the Sp1 motifs that are contained in the 562bp promoter region in the regulation of baseline G6PD activity and its enhancement following NaBu treatment (Figure 6-9A). For this purpose, I mutagenized the Sp1 motifs on the *G6PD* promoter region of interest.

For the mutagenesis assays, I utilised construct 5, which was generated for the reporter assay described above and also contains the 562bp responsive *G6PD* promoter region. Using pGL3-Fr5 as a template, I separately mutagenised each one of the 7 Sp1 binding sites (Figure 6-9A) as described in 2.12. After the mutated plasmid constructs were confirmed by DNA sequencing analysis, they were transfected into 293T cells followed by luciferase activity measurement 48h later. Prior to measurement, the cells were treated for 24h with 3mM NaBu or PBS control.

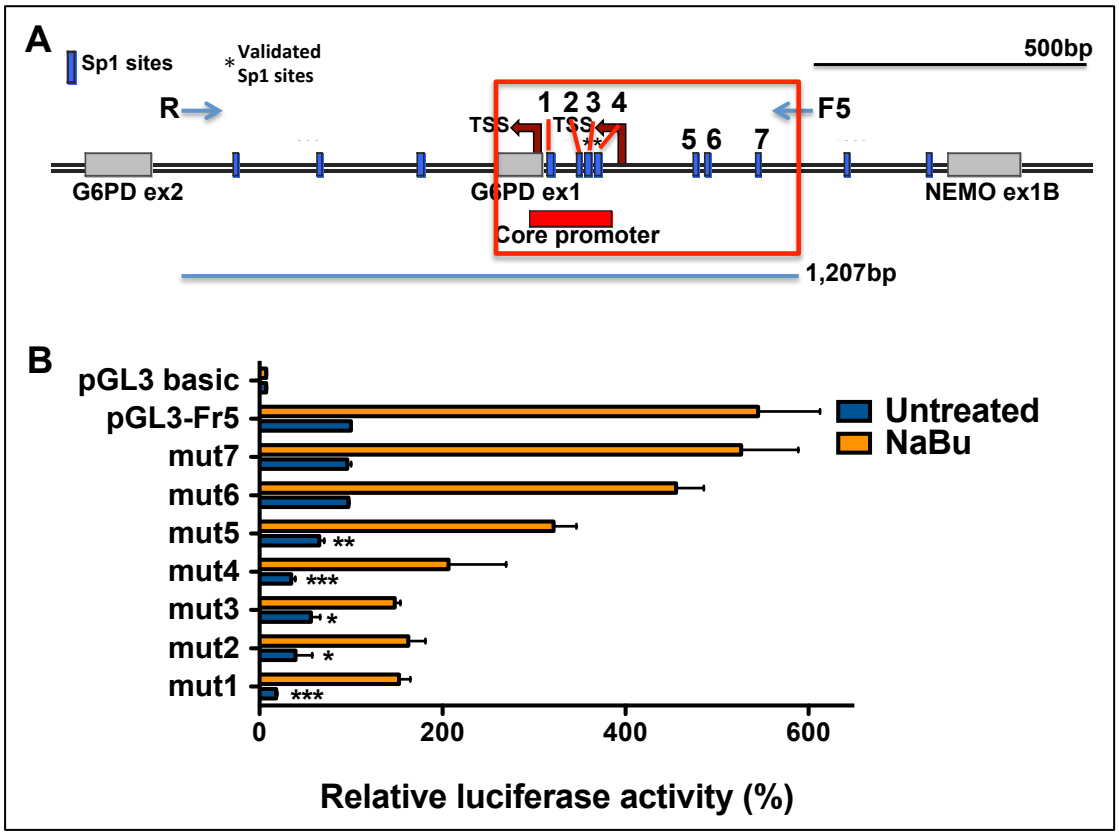


Figure 6-9 Mutagenesis of Sp1 binding sites. (A) Schematic representation of the promoter region of *G6PD*. Boxed in red is the 562bp region responsible for baseline and post-treatment activity. R and F5 are the primers that were previously used to amplify and clone the promoter region containing the 562bp region into a pGL3-basic vector and create construct 5. Numbered 1-7 are the Sp1 binding sites that are included in the 562bp promoter region. **(B)** Luciferase assay conducted on 293T cells after transfection with the constructs containing the mutated Sp1 binding sites at baseline and upon 3mM NaBu treatment for 24h. Activity is normalised against pGL3 basic (without insert) and is shown relative to the baseline levels of the WT pGL3-Fr5 (construct 5). Mean and S.E.M. are shown for n=3 experiments. Student's t-test has been performed to compare pGL3-Fr5 activity to the mutated construct activities at baseline.

Figure 6-9B shows luciferase activity relative to the WT pGL3-Fr5 activity (set as 100% activity) and normalised against the activity of pGL3-basic. Sp1 mutants 1,2,3,4 and 5 significantly reduce the promoter activity to 18% ($p \leq 0.0001$), 39% ($p \leq 0.05$), 56% ($p \leq 0.05$), 34% ($p \leq 0.0001$) and 65% ($p \leq 0.001$), respectively at baseline (Figure 6-9 C and Table 6-3). Mutation of more than one site simultaneously might be needed to completely abrogate activity due to the transactivation ability of Sp1. It is very interesting to note that mutant 1, which corresponds to the Sp1 site with 100% homology to the motif identified to be unique amongst the upregulated genes by motif analysis, is the one that results in the most dramatic effect with only 18% promoter activity (Table 6-3). Furthermore, treatment with 3mM NaBu for 24h restores the promoter activity and exceeds the baseline activity of all mutant constructs. However, mutations that have resulted in a great reduction in the promoter activity are not able to reach activity levels after treatment as those reached by the WT promoter or mutations that do not affect the activity, i.e. Sp1 sites 6 and 7.

Table 6-3 Sp1 binding sites within the 562bp *G6PD* promoter region. [a]

Sp1 sites	DNA sequence	Similarity to -CCCCGCCCCC- motif	Promoter activity after mutagenesis (at baseline)
Site 1	CCCCGCCCCC	100%	18%
Site 2	GCCCCGCCCC	90%	39%
Site 3	GAGGGGTGGT	40%	56%
Site 4	GCCCCGCCCC	90%	34%
Site 5	GGGGCGGGGC	90%	65%
Site 6	CAGGCGGGGA	70%	97%
Site 7	GCCCCGCCCA	80%	96%

[a] The DNA sequence and the % of similarity to the –CCCCGCCCCC- motif identified by microarray analysis of the Sp1 sites within the 562bp *G6PD* promoter region are shown.

To further highlight and confirm the importance of Sp1 in the transcriptional regulation of *G6PD*, the wild type (P277) and G6PD deficient (P7) B cell lines were co-treated with 3mM NaBu and 1mM mithramycin A (MTA) during a time course of 4 - 36h (Figure 6-10A). MTA binds to GC-rich DNA and as such it inhibits binding of Sp family to their cognate motifs (Sleiman et al., 2011). Co-treatment of the two cell lines with MTA and NaBu prevented the increase in *G6PD* transcription confirming that the enhanced transcriptional activity upon NaBu treatment is Sp1-dependent.

Furthermore, by adopting a genetic approach, I transfected a dominant negative (DN; contains the zinc finger but nor the transactivation domain) form of Sp1 (Al-Sarraj et al., 2004) into 293T cells and assessed *G6PD* mRNA expression As a control, the pEBGV plasmid without an insert is used, therefore allows for the endogenous Sp1 to bind to DNA. As shown in Figure 6-10B, the DN form of Sp1 nearly abrogates transcription of *G6PD* 48h post-transfection. Interestingly, NaBu treatment is not able to increase *G6PD* transcription in the presence of the DN Sp1.

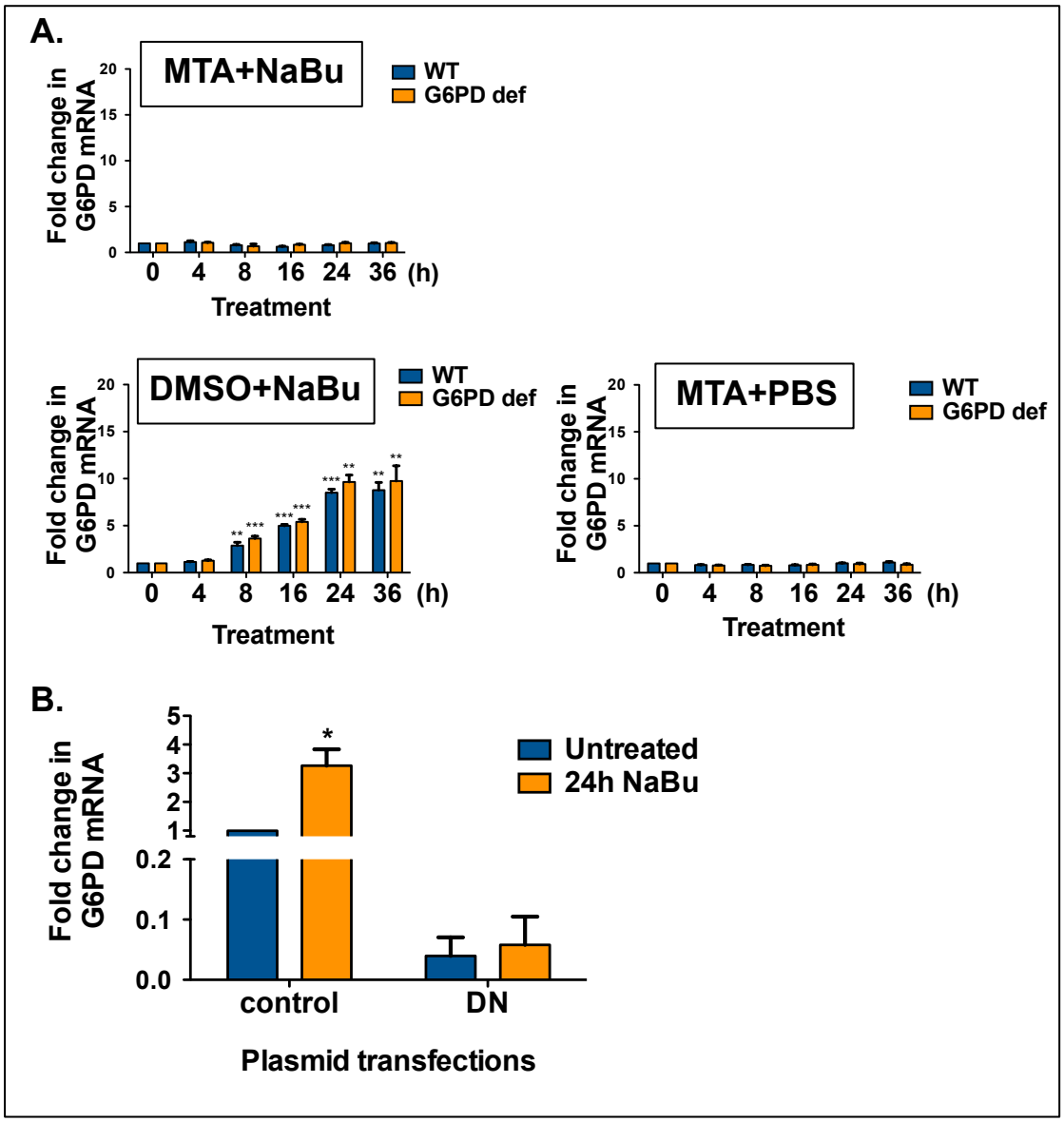


Figure 6-10 Chemical and functional elimination of Sp1 binding. (A) The wild type (P277) and G6PD deficient (P7) cell lines were co-treated with 3mM NaBu together with 1mM MTA and the mRNA levels of *G6PD* were assessed by RT-qPCR. Control experiments treating with NaBu and DMSO or MTA and PBS are shown. Mean and S.E.M. are shown for n=3. **(B)** 293T cells were transfected with a plasmid containing a gene expressing a dominant negative form of Sp1, which lacks the transactivation domain, or with a control plasmid that does not contain any insert. 48h after the tranfection the cells were collected and RT-qPCR was conducted to evaluate the *G6PD* mRNA expression. Mean and S.E.M. are shown for n=3.

6.4 Conclusions

In this final chapter I aimed to describe the genome-wide gene expression changes in response to NaBu in primary human proerythroblasts generated from CB-$CD34^+$ cells. Microarray analysis showed that 3.17% and 2.76% of genes are upregulated and downregulated, respectively. Analysis of the expression of the GPPP genes, confirmed the selective increase of *G6PD* amongst the 17 genes of the GPPP. This result was also confirmed by bioinformatic analysis I performed on published data by Wang and colleagues (Wang et al., 2009) on $CD4^+$ T cells, although in this set of data *ALDOA*, another PPP gene, appears to be marginally upregulated by a combination of NaBu and TSA. Bioinformatic analysis of microarray data obtained on mouse cells lines by Kubicek and colleagues (Kubicek et al., 2012) also showed increased expression of *G6pdx*, upon treatment with SAHA or panobinostat.

De novo DNA motif analysis on the data I obtained from the microarray assays on CB-$CD34^+$-differentiating proerythroblasts, identified a unique Sp1 motif (-CCCCGCCCCC-) amongst the upregulated genes. Reporter assays on the *G6PD* promoter revealed a small 562bp-long promoter region that is responsible for basal and in response to NaBu activity. This region contains 7 Sp1 binding sites, out of which one is the CCCCGCCCCC- motif site. Mutation of the CCCCGCCCCC- motif and other 4 Sp1 binding sites significantly reduces the promoter activity, with the CCCCGCCCCC-motif mutation resulting in the most dramatic decrease down to 18% activity. It was further confirmed that the *G6PD* upregulation is Sp1-dependent by the overexpression of a dominant negative form of Sp1 into 293T cells, which almost completely abrogated *G6PD* expression. Co-treatment of B cell lines with NaBu and MTA, an Sp1 binding inhibitor, also prevented the *G6PD* upregulation. Taken together, these findings clearly show that *G6PD* upregulation is Sp1-dependent. Furthermore, the CCCCGCCCCC- motif appears to be essential for upregulation, although it does not appear to be the only Sp1 motif needed on a given promoter. This is indicated by the fact that mutation of other Sp1 motifs is also shown to reduce promoter activity. However, the presence of the CCCCGCCCCC- motif only on the *G6PD* promoter and not on the other 17 GPPP enzyme genes could be the reason for the selective responsiveness of *G6PD* to NaBu.

7 Discussion

HATs and HDACs antagonistically regulate transcription through acetylation and deacetylation of the histone tails (Brownell and Allis, 2001). Active genes are associated with high binding levels of both HATs and HDACs to the promoters and gene bodies, which dynamically regulate transcription of genes (Wang et al., 2009). HDAC inhibition, using drugs such as NaBu, dysregulates this process resulting in altered expression of a large number of genes. Both HATs and HDACs are recruited to butyrate-responsive genes by Sp1 (Davie, 2003; Wierstra, 2008). Disruption of Sp1 binding in IGD, a rare inherited disease caused by a mutation in the Sp1 binding site of the promoter of the housekeeping gene *PIGM,* results in reduced *PIGM* expression (Almeida et al., 2006; 2007; Caputo et al., 2013). The finding that HDACIs restore Sp1 binding and histone acetylation in IGD led to the main hypothesis of my research, that Sp1-dependent control of histone acetylation and transcriptional activation might also apply for genes that are part of other enzymatic biosynthetic pathways with housekeeping function. Therefore, characterisation of Sp1-dependent epigenetic control of genes with these characteristics could offer new therapeutic opportunities for inherited disorders of ubiquitous biosynthetic pathways such as that of glycolysis.

In summary, the work I have presented in this thesis provides direct evidence that *G6PD* is selectively upregulated by NaBu in primary erythroid cells and B cell lines derived from normal and G6PD deficient individuals. The upregulation of *G6PD* mRNA expression leads to the increased protein expression and enzymatic activity. In fact, G6PD enzymatic activity is restored to normal levels in G6PD deficient cells, indicating the therapeutic potential of phenylbutyrate in the treatment of G6PD deficiency. Epigenetic analysis has shown that the upregulation of *G6PD* is associated with increased acetylation of the gene promoter, recruitment of HATs, HDACs and the TF Sp1. *De novo* motif analysis identified a specific Sp1 binding motif that is present only in genes that are upregulated by NaBu in primary erythroid cells. Further work presented in this thesis shows that this motif is crucial for the upregulation of a gene due to the action of HDACIs.

7.1 G6PD selectively responds to HDACIs

My work provides evidence for the first time that, within the GPPP, HDAC inhibition selectively upregulates *G6PD* expression. The use of wild type B cells, as well as a cell line derived from a patient with G6PD Brighton (class I G6PD deficiency), enabled me to demonstrate that HDACIs (NaBu and SAHA) upregulate *G6PD* mRNA, protein expression and enzymatic activity. Importantly, enzymatic activity in the deficient cells was restored to normal levels after 24h and 36h with NaBu and SAHA, respectively (3.3 and 3.5).

The selective upregulation of *G6PD* over the other GPPP genes was confirmed using bioinformatic analysis of the published data by Wang and colleagues (Figure 6-4; Wang et al., 2009). Although Wang et al did find upregulation of one other GPPP gene (*ALDOA*) on HDACI treatment, the increase was considerably lower than the upregulation of *G6PD*. I did not find any effect of HDACI treatment on *ALDOA* levels in my experiments, which suggests this may be a cell type-specific effect.

Contrary to my expectations, the other 16 genes of the GPPP were not significantly affected by treatment with HDACIs in either normal or deficient cell lines (3.3). My results therefore demonstrate that active genes sharing the same pathway or other common mechanism, such as regulation by a particular TF, are not necessarily regulated in the same manner. This is contrary to the notion that they would share the same specialised transcription factory, as suggested by Schoenfelder and colleagues who showed that globin genes regulated by the common TF Klf1, are transcribed in the same factory (Schoenfelder et al., 2009). Transcription factories are considered to be "ever-changing and self-organising structures that contain DNA or chromatin loops tethered to active transcription units through the transcription machinery" (Xu and Cook, 2008), i.e. a polymerase (active or inactive) and/or its transcription factors (activators or repressors; Dillon, 2008; Osborne et al., 2004; Papantonis et al., 2010; Schoenfelder et al., 2009; Sexton et al., 2007; Xu and Cook, 2008). This occurs in order to achieve more efficient transcription by concentrating relevant machinery and raw materials in one location. Therefore, I had initially hypothesised that transcription of the genes of the GPPP at baseline and in response to NaBu would be co-regulated and be part of a transcription factory structurally and functionally organised by a TF, such as Sp1, as all GPPP genes are bound by Sp1

based on ENCODE analysis (Figure 5-1, Figure 5-2 and Figure 5-3). However, the results of this study clearly demonstrate that, at least in the case of the GPPP, genes that share a common pathway are not necessarily regulated by a common epigenetic mechanism, as has previously been suggested (Schoenfelder et al., 2009).

From an evolutionary perspective, the selective upregulation of *G6PD* might reflect its critical and indispensible role as the main source of reductive potential in the form of NADPH (Castagnola et al., 2010; Metallo and Vander Heiden, 2013). A rapid increase in *G6PD* transcription in response to oxidative stress challenges would result in increased NADPH production and protection of the cell from oxidative damage and death.

The cell line results shown in this thesis suggested that HDACI treatment could be used to overcome G6PD deficiency. However, it was important to confirm that NaBu produced a similar increase in G6PD levels in erythroid cells. To address this, I optimised a two-phase liquid culture system that allows the production of large numbers of erythroblasts starting from either PBMCs or $CD34^+$ cells (4.2). Although PBMCs showed more rapid differentiation by day 7 than $CD34^+$- cells (Figure 4-1, Figure 4-2 and Figure 4-3), both cell sources exhibited efficient erythroid differentiation. The differences are likely due to the fact that PBMCs contain cells that are already committed to the erythroid lineage and therefore enter differentiation earlier. Interestingly, this result has not previously been described. I used this system to assess the effect of HDACI on *in vitro* generated erythroblasts produced from either normal CB-CD34+ cells (4.4), or PBMCs from normal and G6PD deficient individuals (4.4, 4.5) carrying various mutations. Consistent with the cell line results, *G6PD* mRNA and protein expression, as well as G6PD enzymatic activity, were selectively upregulated, whilst all other genes of the GPPP remained unaffected by HDACI treatment. This confirmed that *G6PD* upregulation is neither cell type nor HDACI specific, as it occurs in human B, T, erythroid and embryonic kidney cells and mouse pancreatic cells, using either NaBu, SAHA or panobinostat (Figure 3-6, Figure 3-9, Figure 3-10 and Figure 6-4). Importantly, these results strongly suggest that NaBu could be effective in the treatment of G6PD deficiency.

7.2 HDAC inhibition in erythropoiesis

HDAC inhibitors have been used since the 1970s, when they were first discovered, for the treatment of cancer, haemoglobinopathies, sickle cell anaemia and most recently IGD. Compared to chemotherapeutics, short-chain fatty acids, including sodium phenylbutyrate, are not mutagenic and therefore do not carry carcinogenic risk long-term. However, they do have limitations, including their rapid metabolism, which requires administration of large doses and the inhibition of cell proliferation, which limits the pool of early erythroid progenitors (Perrine, 2008).

In view of my results showing that HDACIs can restore G6PD enzymatic activity to normal levels, I wanted to further explore the effect of HDACIs on cell viability and differentiation. To investigate the effect of HDAC inhibition on cell viability and differentiation, I conducted experiments on cell lines and primary cells. I demonstrated that NaBu induces apoptosis in B lymphoblastoid cell lines, assessed by annexinV staining (Figure 3-1) and reduced cell numbers of primary erythroid cells generated by either PBMC (Figure 4-5) or $CD34^+$ (Figure 4-7) erythroid differentiation cultures, in a dose-dependent manner. Annexin V staining was not performed in primary erythroid cultures, due to cell number limitations, and it was therefore not possible to confirm that the reduced cell numbers were due to increased apoptotic rates. My findings are consistent with other studies conducted on human cervix tumour cell lines (Dyson et al., 1992) and on a multilineage haematopoietic cell line (Ikuta et al., 1998) that were treated with butyrate. In fact, concentrations below 0.5mM were shown to decrease cell proliferation without inducing cell death for up to 5 days. In concentrations above 0.5mM cell growth was arrested and apoptosis was accelerated in a dose-dependent manner (Dyson et al., 1992). In contrast, a more recent study (Chaurasia et al., 2011) in which cord blood $CD34^+$ cells were expanded to differentiate down the erythroid lineage, treatment with other chromatin modifying agents, i.e. SAHA, TSA and VPA, showed that these HDACIs could actually promote expansion of erythroid cells. This finding led the authors to suggest that HDACIs could be used to enhance *in vitro* generation of erythroid cells for the production of red cell concentrates for use in transfusion. However, the results of this study could not be confirmed by my work in which I have shown that, as with NaBu, both SAHA and TSA, decrease erythroid cell numbers at day 7 of the primary cell cultures (Figure 4-7). A possible explanation for this might be the fact that,

although the concentration and time of exposure to the HDACIs are the same, the culture system is different. Chaurasia et al are using a serum-containing two-phase liquid culture system consisting of IMDM implemented with 30% FBS, 100ng/ml SCF, 100ng/ml Flt3, 100ng/ml Tpo and 50ng/ml IL-3, whereas I am using a serum-free system.

Together, the data presented in this thesis show that treatment of the generated erythroid precursors with NaBu (but not other HDACIs) not only reduces cell number but also alters erythroid differentiation. In fact, PBMC-derived erythroid cells display delayed erythropoiesis (Figure 4-6), whereas in CB-$CD34^+$-derived cells, erythropoiesis appears to be promoted after long exposure (treatment after 16h CA) to NaBu, whereas short-term exposures of 48h do not alter differentiation (Figure 4-8). On the other hand, although SAHA and TSA are shown to reduce the number of erythroid cells produced (Figure 4-7), they do not seem to affect the differentiation of CB-$CD34^+$-differentiating cells even after long periods of treatment (treatment after 16h CA; Figure 4-8). Other groups have conducted studies that have also shown variable results. Effects similar to those that I have observed for PBMC-derived erythroblasts have been characterised by Yamamura and colleagues (Yamamura et al., 2006) who showed that the HDACI, romidepsin, inhibited the generation and proliferation of $CD36^+GlyA^{high}$ mature erythroblasts, which were expanded from $CD34^+$ cells. In agreement with the pleiotropic effect of HDACIs, and in contrast to romidepsin, TSA and VPA are shown to block enucleation at later stages of differentiation, but do not affect early differentiation and proliferation of erythroblasts (Ji et al., 2010; Migliaccio, 2010). In addition, two newly characterised HDACIs are shown to restore the impaired *in vitro* maturation of β-thalassaemic erythroblasts, by producing terminally differentiated erythroid cells with restored β-globin levels (Mai et al., 2007). On the other hand and in agreement with my findings in CB-$CD34^+$ cells, studies in the erythroid J2E cell line show that treatment with NaBu induces erythroid differentiation, coupled with increased haemoglobin expression, but blocks cell proliferation (Busfield et al., 1993; Jaster et al., 1996). My findings may be explained by the fact that cells originating from different tissues respond differently to HDACIs and in particular NaBu. Irrespective of which cell type is used, altered erythroid differentiation in response to HDACIs underpins the importance of HDACs in the progression of every stage of erythoid differentiation (Migliaccio, 2010).

The rapid metabolism of NaBu and phenylbutyrate (serum half-life in humans is 0.8h; Kramer et al., 2001) and the fact that I have found that they inhibit cell proliferation and alter differentiation *in vitro* may set limitations regarding their potential clinical usefulness in the treatment of anaemias associated with glycolytic enzyme deficiencies. However, *in vivo* studies using HDACIs in mice and primates (baboons) with thalassaemia have not shown worsening of anaemia (Cao et al., 2005; Pace, 2002). In fact, short-chain fatty acid derivatives, including butyrate, stimulate *γ globin* gene expression and erythropoiesis by increasing the BFU-E and reticulocyte counts. Furthermore, sickle cell anaemia patients treated with butyrate experienced increased reticulocyte levels (no differentiation block), showing that this is a safe drug to use in patients with haematological disease (Dover et al., 1994). Another advantage of butyrate is that it is a naturally occurring substance and has little or no toxic effect (Daniel et al., 1989; Fraczek et al., 2013; Miller, 2004a). In contrast to butyrate, other HDAC inhibitors have a wider spectrum of side effects. For instance, panobinostat, romidepsin and SAHA have been shown to block platelet production and cause thrombocytopenia (Bishton et al., 2011). More importantly, their long-term safety profile has not yet been established, placing limitations on their clinical use. Therefore, I suggest sodium phenylbutyrate would be the most suitable for the treatment of severe G6PD deficiency and other inherited diseases that might be caused by enzyme defects and respond to butyrate treatment, as shown by our genome-wide studies. However, it is important to note that given the potential therapeutic effects of sodium phenylbutyrate on G6PD deficiency, shown in this study, and other diseases, such as sickle cell anaemia, there is need for the production of a more patient-friendly form of sodium phenylbutyrate, which requires less frequent dosage.

7.3 Genome-wide implications of HDAC inhibition

To further investigate the effects of NaBu, I used microarray to assess global gene expression. GPPP enzyme deficiencies are manifest in the erythroid lineage; therefore, I chose to assess the genome-wide implications of HDAC inhibition in primary erythroblasts. Almost all prior studies conducted to evaluate gene expression in response to HDACIs, including butyrate, involved cell lines rather than primary human cells. Wang and colleagues (Wang et al., 2008) were the first to conduct gene

expression profiling in human primary cells, specifically primary human $CD4^+$ T cells, which were treated with a combination of NaBu and TSA.

The gene expression profiling data on primary human proerythroblasts generated by the *in vitro* erythroid differentiation system from CB-$CD34^+$ cells, which I presented in this thesis, showed that only 5.93% of the genes tested are significantly affected ($FC \geq 1.5$, $p \leq 0.05$) by NaBu treatment with 3.17% and 2.76% of genes being upregulated and downregulated, respectively. These findings are in line with studies conducted by other researchers in various cell lines, using different HDACIs, which came to the general conclusion that only 2-25% of genes are affected by HDAC inhibition (Davie, 2003; Delcuve et al., 2012; Mitsiades et al., 2004; Sealy and Chalkley, 1978; Sekhavat et al., 2007; Van Lint et al., 1996). The time in culture, the concentration, the specific HDACI used and also the type of cells determine the number of genes altered in transcription. The number of affected genes increases as time and concentration increase likely as a result of secondary rather than direct effects (Peart et al., 2005). For example, treatment of a colonic epithelial cell line (SW620) with 5mM NaBu for 48h showed that 7% of the genes showed altered gene expression, with 3% upregulated and 4% downregulated (Mariadason et al., 2000). Similarly, treatment of a non-small lung carcinoma cell line (H460) with 1mM of NaBu for 24h, provided evidence that 4% of the tested genes were upregulated and 8% were downregulated (Joseph et al., 2004). Gene-specific studies have also confirmed that HDACIs can decrease the expression of genes, such as that of the Epidermal Growth Factor Receptor (EGFR) in colorectal cell lines treated with TSA and SAHA independently (Chou et al., 2011) and that of cholesterol biosynthesis in enterocyte cell lines treated with NaBu (Alvaro et al., 2008). The results arising from these studies are in line with my results, showing a similar number of upregulated versus downregulated genes, the exact numbers of which appears to depend on the cell type, type of HDACI used and the HDACI concentration.

Genome-wide experiments that have assessed the effect of HDACIs at different time points have provided further information. Peart and colleagues (Peart et al., 2005) treated T cell leukaemia cell lines (CEM) with either 2.5μM SAHA or 1ng/ml depsipeptide and showed that the number of HDAC-regulated genes increased over time (SAHA: from 1.29% after 1h to 22% after 16h. Depsipeptide: from 0.27% after 1h to 24.8% after 16h.). Similarly, Tabuchi and collegues(Tabuchi et al., 2006)

treated colonic epithelial cells (MCE301) with 2mM NaBu over a 6, 12, 24h time course and showed that 2.7% (out of which 70% upregulated and 30% downregulated), 6.3% (out of which 53% upregulated and 47% downregulated) and 11.5% (out of which 41% upregulated and 59% downregulated) of genes showed altered levels of expression, respectively. This dynamic change in gene expression during HDAC inhibition emphasises the fact that longer treatments may result in downstream rather than direct effects (although some genes might simply respond faster than others). It is also interesting to note that this information suggests that shorter HDACI treatment increases the number of genes upregulated rather than downregulated, whereas longer incubations with HDACIs result in a larger number of dowregulated genes, which is possibly the result of secondary effects.

7.4 Epigenetic mechanism of action of NaBu

As discussed previously, HDACIs restore Sp1 binding and histone acetylation in IGD, which led to the main hypothesis of my work, i.e., that Sp1-dependent control of histone acetylation and transcriptional activation could also apply for genes that are part of other enzymatic biosynthetic pathways with housekeeping function. Therefore, characterisation of Sp1-dependent epigenetic control of genes with these characteristics could offer new therapeutic opportunities for inherited disorders of ubiquitous biosynthetic pathways such as glycolysis.

However, my finding that HDAC inhibition selectively upregulates *G6PD* led me to explore the mechanism behind this selective upregulation and therefore determine why *G6PD* differs from the other non-responsive genes. During the initial stages of this project, I showed that *G6PD* responsiveness is seen with a range of HDACIs, including NaBu, TSA and SAHA (Figure 3-7). Furthermore, this effect is cell type-independent, as it is present in human B cells (Figure 3-6), erythroid (Figure 4-6, Figure 4-8 and Figure 6-4), T cells (Figure 6-4) and embryonic kidney cells (Figure 6-8), as well as mouse pancreatic cells (Figure 6-4). Importantly, this meant that I was able to study the epigenetic changes in cell lines, overcoming the potential limitations with cell number that I may have encountered had I needed to use primary cells.

In this thesis, I have established that the epigenetic landscapes of *G6PD* and the other genes of the GPPP at baseline are that of typical housekeeping genes.

ENCODE analysis of the *G6PD*, *TPI* and *GPI* promoters (Figure 5-1, Figure 5-2 and Figure 5-3) and ChIP assays that assessed the *G6PD*, *TPI, GPI, PGLS* and *RPIA* promoters at baseline (Figure 5-8, Figure 5-9, Figure 5-15 and Figure 5-16) showed that they are all highly acetylated and enriched for positive chromatin marks, HATs, HDACs and Sp1. Binding of Polymerase II on the promoters and gene bodies of these genes is in line with their housekeeping expression pattern. A notable structural difference between *G6PD* and the other genes of the GPPP is the presence of a bidirectional promoter shared with *NEMO*. However, the hypothesis that the *G6PD* upregulation is NEMO-driven was disproved, as *NEMO* is not upregulated by NaBu (Figure 5-4). Having established all the information discussed above and based on evidence from CHX treatment (Figure 3-11) that *G6PD* is directly upregulated by NaBu in a translation-independent manner, I next sought to understand the epigenetic changes that occur as a result of the NaBu treatment.

7.4.1 NaBu increases histone acetylation

Prior studies have shown that HDAC inhibition causes widespread histone hyper-acetylation (Davie, 2003; Delcuve et al., 2012; Mitsiades et al., 2004; Sealy and Chalkley, 1978; Sekhavat et al., 2007; Van Lint et al., 1996). An extremely well studied gene-specific model of this is the *p21* gene, which encodes the cyclin-dependent kinase inhibitor *p21* and mediates cell cycle arrest and thus apoptosis. HDACIs have been shown to upregulate the expression of *p21* whilst causing hyper-acetylation of both the proximal and distal promoter regions (Gui et al., 2004; also shown in Appendix A in primary proerythroblasts upon NaBu treatment). In line with those studies, ChIP analysis of the *G6PD* promoter showed increased acetylation on H3 and H4 in both normal and G6PD deficient cells (Figure 5-8). My findings are in disagreement with a study conducted by Halsall and colleagues (Halsall et al., 2012) who compared the genome-wide histone hyper-acetylation caused by VPA in the promyelocytic leukaemia cell line HL60 and transcriptional responses of selected genes. The main observation arising from this study was that promoter acetylation of individual genes was not increased in response to VPA, even at genes showing enhanced transcription. However, Halsall et al might have failed to identify an increase in acetylation due to the use of antibodies against only three acetylation marks (H3K9Ac, H4K8Ac and H4K16Ac), whereas I have used antibodies against

total H3 and H4 acetylation, which allows me to capture changes of all acetylation marks.

Interestingly, the H4 acetylation increase is significantly stronger than the H3 acetylation increase in both cell lines. In fact, H3 and H4 acetylation enrichment of the core promoter (G6PD.2) increased 4.2% and 12% in WT cells, respectively, and 3% and 8% in G6PD deficient cells, respectively. A possible explanation might be that H4 acetylation is favoured due to the lower acetylation on H4 rather than H3 at baseline (2-fold; Figure 5-8). Another explanation may be that H4 acetylation is not regulated in the same manner as H3 acetylation and it might rely on some HDACs more than others. Although NaBu is a pan-HDAC inhibitor, it has been shown to have different levels of efficiency against target HDACs, which may influence acetylation of specific histones. Blackwell et al (Blackwell et al., 2008) have previously shown that butyrate is most potent against HDACs 1, 2, 3 and 6. More recently, chemoproteomic profiling of HDACIs (Bantscheff et al., 2011; Bradner et al., 2010) revealed different degrees of efficiency of a variety of HDACIs, including SAHA, TSA and VPA, against each HDAC examined. To test this assumption I could inhibit each HDAC separately either chemically or with the use of shRNAs. However, it should be noted that in this study I have only focused on general H3 and H4 acetylation and haven't examined specific acetylation marks that might be specifically increased on the *G6PD* promoter. Further studies looking at specific acetylation marks might provide more evidence to describe a detailed mechanism of action of NaBu.

The most interesting finding was that NaBu did not hyper-acetylate other genes of the GPPP. Specifically, H3 and H4 acetylation of the *TPI*, *GPI*, *RPIA* and *PGLS* gene core promoters did not change significantly in either of the two cell lines used (Figure 5-6 and Figure 5-7). The fact that mRNA expression of these genes was not significantly altered in response to NaBu, suggests that histone acetylation may be required for upregulation of *G6PD* expression. This complements the findings of Rada-Iglesias' study (Rada-Iglesias et al., 2007) in HepG2 and HT-29 cell lines. Rada-Iglesias has shown that treatment with HDACIs, including butyrate, in these cells caused deacetylation of a number of promoter regions, as opposed to the global acetylation increase that was observed. These promoter regions corresponded to genes that are downregulated under butyrate exposure, as opposed to unaffected genes like

those of the GPPP. It should be noted that in Rada-Iglesias' study, upregulated genes display increased acetylation on their promoters, a finding, which is in agreement with the increased acetylation I observed on the *G6PD* promoter.

7.4.2 NaBu increases the recruitment of chromatin regulators

In this study, ChIP analysis of HAT and HDAC binding on the core promoters of *G6PD* and other genes of the GPPP at baseline demonstrated co-occupancy of those two classes of chromatin regulators for these genes. As shown in Figure 5-10 and Figure 5-12 for HATs (CBP, p300 and GCN5) as well as Figure 5-11 and Figure 5-13 for HDACs (HDAC1, 3, 4/5/7 and 6), recruitment of both these two classes of enzymes on the GPPP gene promoters is enriched.

This finding supports previous research that has given further complexity to the general model of transcription, providing evidence for a dynamic cycle of acetylation and deacetylation by the simultaneous binding of HATs and HDACs. For instance, the genome-wide study by Wang (Wang et al., 2009) showed that active genes are associated with high binding levels of both HATs and HDACs. This suggested that HDACs are required to reset the chromatin status by removing the acetyl groups after the completion of each round of transcription so that the gene does not hyper-acetylate. Furthermore, Ram and colleagues (Ram et al., 2011) developed a novel ChIP method (ChIP-string) to map the genome-wide binding of 29 chromatin regulators, including several HATs and HDACs in K562 cells and H1 ESCs. Among other findings in this study, it was confirmed that HATs and HDACs dynamically regulate the expression of active genes. A further genome-wide study (Johnsson et al., 2009) conducted in *Schizosaccharomyces pombe,* mapped the recruitment of the HAT Gcn5 to gene promoters and the distribution of H3K14ac. Gcn5 was found to localise in high concentration on coding regions of highly transcribed genes and to collaborate antagonistically with Clr3, a class II HDAC. This interplay between Gcn5 and Clr3 was shown to be crucial for the regulation of stress-response genes. Several gene-specific studies have also confirmed the co-occupancy of active gene promoters by both HATs and HDACs (Huang, 2005; Huang et al., 2007; Park, 2001). Confirmation of the co-recruitment of HATs and HDACs to the promoters of GPPP genes at baseline, through my work, is of great importance as it underpins the dynamic regulation of their transcription. Therefore, looking closer at the epigenetic status of the dynamically regulated GPPP gene promoters before and after NaBu treatment will

provide insights into the mechanism of action of NaBu and importantly the selective upregulation of *G6PD*.

One of the most important findings of the present study was that HAT and HDAC recruitment on *G6PD* is enriched upon NaBu treatment. In fact, in both normal and G6PD deficient cell lines, 5h treatment with NaBu increased the recruitment of all HATs tested, i.e. p300, CBP and GCN5, as well as selected HDACs, i.e. HDAC1 and HDAC6. HDAC 3, 4, 5 and 7 binding is not significantly affected by NaBu, thus these HDACs remain bound - or constantly present through rapid cycles of recruitment - to the *G6PD* promoter. The fact that only the recruitment of HDAC1 and 6 is upregulated might be related to the differential potency of NaBu against different HDACs. It should be noted that HAT and HDAC binding in all the other GPPP gene promoters is not significantly altered. Although there are few studies in the literature showing what happens to the recruited HDACs upon HDAC inhibition, the limited results that are available are controversial. Specifically, Huang and colleagues (Huang, 2005) have previously shown that TSA increased the recruitment of the HATs CBP and p300, as well as the TF Sp1 on the promoter of TSA-responsive TGFβ type II receptor in human pancreatic cancer cell lines (BxPC-3, PANC-1, CFPAC-1, and MIA PaCa-2). Moreover, the same study provided evidence that HDAC1 recruitment on the promoter of those genes decreases after TSA treatment. Similarly, Nunes and colleagues (Nunes et al., 2010) treated the human neuroblastoma cell line SH-SY5Y and looked at the promoter status of the Sp1-dependent HDAC-responsive *CYP46A1*. Although in this publication the authors are examining the recruitment of a small number of HATs (CBP and p300) and HDACs (HDAC1 and HDAC2), they show that TSA treatment induces a slight yet significant decrease in HDAC2 occupancy whereas it increases p300 occupancy. This was also followed by an increase in Sp1 binding. In contrast to these findings, Sekhavat and colleagues (Sekhavat et al., 2007) performed a combination of immunoprecipitation and immunoblotting in the human breast cancer cell line MCF-7 and showed that inhibition of HDAC1 and HDAC2 by TSA does not disturb their binding to the chromatin, without however investigating specific TSA-responsive genes. The same study showed that the association of HDAC2 with Sp1 is also not affected by TSA. These studies show the variability of the limited results available in the literature. However, the limitation of these studies is that they only investigate the binding of

some HDACs, whereas my work explores the binding of almost all HDACs and therefore provides a clear picture of the GPPP gene promoters upon HDAC inhibition.

My data shed light on the controversy regarding the binding of HDACs after HDACI treatment and suggest a method of regulation of NaBu-responsive genes through the dynamic recruitment of both HATs and HDACs. My results suggest that NaBu treatment encourages the preferential action of HATs over the inhibited HDACs. This leads to further recruitment of additional HATs, which result in the increase of H3 and H4 acetylation and thus PolII binding resulting in the increased expression of *G6PD*. Furthermore, the inhibited HDACs do not dissociate from the transcription complex. In fact, more HDACs are recruited with the aim of balancing the increased HATs and keeping the system under control.

7.4.3 Sp1 is vital for the upregulation of G6PD

As a final step for this thesis I sought to understand what factor(s) determine the increased recruitment of HATs and HDACs on the *G6PD* promoter and also why *G6PD* is selectively upregulated by NaBu. Reporter assays of fragments of the G6PD promoter indicated that a 562bp region of the promoter is responsible for basal activity as well as responsiveness of *G6PD* to NaBu. This indicates that this region must contain key binding sites for whichever factor is responsible for the upregulation of *G6PD* upregulation. Importance of this particular fragment for basal activity has also been previously shown by Philippe and colleagues (Philippe et al., 1994) who conducted reporter assays in HepG2 and K562 cells.

Since Sp1 has previously been associated with HDACI-responsive genes, I studied the importance of Sp1 within this region. Specifically, as discussed in the introduction it is known that a group of butyrate-responsive genes contain Sp1 sites within their promoters (Davie, 2003; Majumdar et al., 2012; Siavoshian et al., 1997). Additionally, Sp1 has been shown to interact with HATs and HDACs (Hou et al., 2002; Kundu et al., 2000; Nunes et al., 2010; Wierstra, 2008). Furthermore, Franze and colleagues (Franzè et al., 1998) have previously conducted transactivation assays in HeLa cells using co-transfection with an Sp1-expression plasmid and reporter plasmids containing different fragments of the -126 to +16 core promoter region, which is included in the 562-long region that I identified using transactivation assays. In this study, they showed that Sp1 is able to activate all Sp1 binding site-containing

fragments and lead to induction of promoter activity. This study showed the important role of Sp1 in the regulation of the *G6PD* core promoter at baseline.

So far various studies have shown increased recruitment of Sp1 on the promoters of HDACI-responsive genes that are upregulated (Huang, 2005; Nunes et al., 2010; Zhang et al., 2006). Importantly, more recent studies have proven that gene expression upregulation through HDAC inhibition is dependent on Sp1. Specifically, a study that employed mutant construct transfections and MTA treatment of the human colon adenocarcinoma cell line BCS-TC2 proved that Sp1 is essential for the upregulation of the metalloprotein-expressing *MMP11* gene upon NaBu or TSA treatment (Barrasa et al., 2012). In a similar approach, it was demonstrated by Yang et al (Yang et al., 2012) that MTA blocked the TSA- and SAHA- induced *CD1d* mRNA expression in human alveolar epithelial A549 cells and mouse melanoma B16/F0 cells. In the same study, co-transfection assays using Gal4-Sp1 and Fc-luciferase reporters shown that both TSA and SAHA induced *CD1d* promoter luciferase activity by enhanced Sp1 transactivation activity, which results in increased recruitment of Sp1 on the *CD1d* promoter. Interestingly, a study by Law and colleagues(Law et al., 2011) demonstrated through the knockdown of Sp1 expression in human colon adenocarcinoma HT-29 cells that Sp1 is also responsible for the downregulation of stanniocalcin-1 (*STC-1*) upon TSA treatment. This latter study underpins the diverse yet crucial role of Sp1 during HDAC inhibition.

In this study, I aimed to not only determine the importance of Sp1 in *G6PD* upregulation, but also to understand why *G6PD* amongst all GPPP genes is upregulated. ENCODE (Figure 5-1, Figure 5-2 and Figure 5-3) and ChIP (Figure 5-16) analysis at baseline showed that all the genes of the GPPP are bound by Sp1, yet Sp1 recruitment is increased only on the promoter of *G6PD* (Figure 5-17) in response to NaBu. *De novo* DNA motif discovery analysis on the data obtained from the microarray assays on CB-CD34^{+}-differentiating proerythroblasts, identified a unique Sp1 motif (-CCCCGCCCCC-) amongst the upregulated genes, thus providing independent evidence supporting the critical role of Sp1 in regulating transcription of *G6PD* in response to NaBu. Interestingly, this Sp1 binding motif is one of the 7 Sp1 sites that fall within the 562bp region, which is responsible for basal and in response to NaBu activity.

Although it has been previously shown that Sp1 is able to transactivate genes through its binding to multiple Sp1 binding sites on a given promoter (Davie et al., 2008), it might be possible that a specific Sp1 binding site is responsible for the preferential responsiveness upon HDAC inhibition. In fact, mutation of the CCCCGCCCCC- motif and other 4 Sp1 binding sites significantly reduces the promoter activity, with the CCCCGCCCCC-motif mutation resulting in the most dramatic decrease down to 18% activity. Similarly, Kim and colleagues (Kim et al., 2008) have previously shown in gastric cancer cell lines the dependence of *Deleted in Liver Cancer-1 (DLC-1)* upregulation on Sp1 occupancy. In this study, luciferase reporter assays with mutant Sp1 sites were conducted to show the abrogation of TSA-induced upregulation through the mutation of specific Sp1 sites. Indeed, mutation of 2 out of 8 Sp1 binding sites had the most dramatic effect in the prevention of the *DLC-1* upregulation. However, in this study the researchers did not further investigate the functional difference in these 2 sites in comparison to the other Sp1 sites. In my study, it was further confirmed that *G6PD* transcriptional upregulation is Sp1-dependent by the overexpression of a dominant negative form of Sp1 into 293T cells, which almost completely abrogated *G6PD* expression. Co-treatment of B cell lines with NaBu and MTA also prevented *G6PD* upregulation. Taken together, these findings clearly show that *G6PD* upregulation is Sp1-dependent and for the first time it is proven that the CCCCGCCCCC sequence is the essential motif for upregulation. Further proof of the importance of this motif could be given if reporter assays were conducted using a synthetic basic promoter with and without the Sp1 binding motif in order to show its importance for basal and in response to NaBu promoter activity.

Although in this study I have focused on the role of Sp1, further research could determine the importance of other factors in the responsiveness to HDACIs. A potential approach would be the combination of immunoprecipitation of specific parts of the genome (e.g. *G6PD* versus other genes of the GPPP) with mass spectrometry assays to determine what other factors bind on these genes. This approach has been used previously for the analysis of factors bound on telomeric DNA (Sperry et al., 2008).

7.4.4 Epigenetic model for selective upregulation by NaBu

In this study, I suggest a novel mechanism for the selective upregulation of genes by HDAC inhibition. Although my studies focused on the GPPP in response to NaBu, I suggest that this mechanism might apply globally in response to HDAC inhibition. Specifically, I suggest that HDACIs increase the recruitment of Sp1 in promoters containing the CCCCGCCCCC- motif. Furthermore, our lab (unpublished data) and others (Yang et al., 2012) have confirmed that HDACIs upregulate Sp1 expression, which is likely to further enhance Sp1 binding. Increased Sp1 binding in the promoters containing the CCCCGCCCCC- motif leads to increased recruitment of HATs and thus H3 and H4 hyper-acetylation. Hyper-acetylation of these promoters leads to increased binding of PolII and therefore increased mRNA expression. At the same time, the inhibited HDACs do not dissociate from the transcription complex; more HDACs are in fact recruited with the aim of balancing the increased HATs and keeping the system under control.

7.5 Directions for future work

This work strongly suggests that the use of butyrate could be of therapeutic use for the treatment of severe G6PD deficiency. This has been suggested based on findings in G6PD deficient and normal cell lines and primary erythroblasts. In these cell types, treatment with NaBu increased *G6PD* mRNA and protein expression and most importantly restored G6PD enzymatic activity to normal levels in the G6PD deficient cells. However, a question that remains unanswered is whether the same effect could be observed *in vivo*.

To confirm the *in vivo* upregulation of *G6PD* expression and activity, a G6PD deficient mouse model could be used. In fact, Pretsch and colleagues (Pretsch et al., 1988) first described a G6PD deficient mouse strain that results in severely reduced G6PD levels (15% of WT) in hemizygous male or homozygous female mice. The genetic defect in these mice is an in-frame deletion on exon 1 that causes lower expression of the G6PD protein. The G6pdx-^{m1Neu} strain is available to order on the EMMA database (http://www.emmanet.org/) under the code EM:00073. Phenylbutyrate injections could be administered to this mildly deficient mouse to assess the effect on G6PD enzymatic activity. Additionally, to test the effect of phenylbutyrate in severely deficient mice, naphthalene injections could be administered to the mice prior to phenylbutyrate treatment, as naphthalene is known

to cause acute haemolysis and mimic drug-incuded haemolytic anaemia in Class II/III G6PD deficient individuals. Such an experiment is likely to succeed as it has been previously shown (Kubicek et al., 2012) that *G6pdx* responds in the same manner as *G6PD* to NaBu (Figure 6-4).

Alternatively, *in vivo* studies could be conducted in G6PD deficient patients, particularly those with severe chronic haemolytic anaemia. Confirmation of the effectiveness of NaBu in humans *in vivo* may be crucial for class I G6PD deficient patients who currently rely on regular blood transfusion.

The overlap between the geographical distribution of G6PD deficiency and malaria infection presents a further potential use of NaBu. As mentioned in 1.4.3.2, amongst the drugs that have been shown to cause drug-induced haemolytic anaemia in G6PD deficiency are the anti-malarials primaquine and pamaquine. This places limitations on the treatment of malaria in patients with G6PD deficiency as treatment can induce severe haemolysis. However, co-treatment with phenylbutyrate could potentially increase expression and enzymatic activity of G6PD and prevent occurrence of acute haemolysis allowing the use of a wider range of anti-malarial drugs.

Phenylbutyrate could also be of therapeutic use in other diseases that are caused by mutations that lead to reduced enzymatic activity. Appendix A shows the list of genes that were significantly upregulated (FC≥1.5, p≤0.05) and Appendix B those that are significantly downregulated (FC≥1.5, p≤0.05). Amongst the upregulated genes is coagulation factor VIII (*FVIII*) that shows a 1.77-fold increase on NaBu treatment (p≤0.001). Mutations in the *FVIII* gene promoter cause haemophilia A, a hereditary X-linked recessive disorder, associated with low levels of FVIII protein(Bolton-Maggs and Pasi, 2003; Mannucci and Tuddenham, 2001). Depending on the FVIII activity in patient plasma, haemophilia A is classified as severe (<1%), moderate (1–5%) or mild (>5% to <40%; Franchini and Mannucci, 2013; Kulkarni and Soucie, 2011). To test this hypothesis, *FVIII* expression upon NaBu treatment could be measured in WT hepatic cell lines, such as HepG2, FVIII deficient cell lines and in primary cells *in vitro*. As a next step, *FVIII* promoter – containing plasmids that also contain a luciferase gene could be transfected into HepG2 cells in order to measure the promoter activity upon NaBu treatment.

Zimmermann and colleagues (Zimmerman et al., 2012) have cloned the WT and several patient mutant variations of the *FVIII* core promoter into a luciferase gene-containing plasmid construct, which they then used to transfect into HepG2 cells. These constructs could also be used to transfect HepG2 cells whilst treating them with NaBu in order to measure the changes in the FVIII promoter activity. Finally, if these experiments confirm our hypothesis, we would suggest conducting clinical trials in haemophilia A patients to test the administration of phenylbutyrate as an alternative to the current treatment methods.

The future work suggested here underpins the therapeutic potential of phenylbutyrate beyond its current use. Despite its therapeutic potential, the form of sodium phenylbutyrate currently marketed and used for the treatment of diseases has a short half-life and is subjected to first pass hepatic clearance, explaining the milligram doses the patients are required to take in order to achieve therapeutic concentrations *in vivo* (Yoo and Jones, 2006). This often causes mild yet unpleasant side affects, such as muscle pain, swelling of the legs and tiredness. However, as mentioned in 7.2, the advantage of butyrate is that it is a naturally occurring substance and has little or no toxic effect (Daniel et al., 1989; Fraczek et al., 2013; Miller, 2004a). In contrast to butyrate, other HDAC inhibitors have many side effects, for instance panobinostat, romidepsin and SAHA were shown to block platelet production and cause thrombocytopenia (Bishton et al., 2011); thus, their long-term safety profile has not been established as yet. To conclude, a more patient-friendly form of butyrate should be developed for the treatment of G6PD deficiency and possibly other enzymatic deficiencies.

8 Bibliography

Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the Transcriptional Histone Acetylation Code for a Human Gene. Cell *3,* 381–392.

Al-Sarraj, A., Day, R.M., and Thiel, G. (2004). Specificity of transcriptional regulation by the zinc finger transcription factors Sp1, Sp3, and Egr-1. J. Cell. Biochem. *94*, 153–167.

Allfrey, V.G. (1966). Structural modifications of histones and their possible role in the regulation of ribonucleic acid synthesis. Proc Can Cancer Conf *6*, 313–335.

Allfrey, V.G., Faulkner, R., and Mirsky, A.E. (1964). Acetylation and Methylation of Histones and Their Possible Role in The Regulation of RNA Synthesis. Proc. Natl. Acad. Sci. U.S.A. 786–794.

Almeida, A.M., Murakami, Y., Baker, A., Maeda, Y., Roberts, I., Kinoshita, T., Layton, D.M., and Karadimitris, A. (2007). Targeted Therapy for Inherited GPI Deficiency. The New England Journal of Medicine *356,* 1641–1647.

Almeida, A.M., Murakami, Y., Layton, D.M., Hillmen, P., Sellick, G.S., Maeda, Y., Richards, S., Patterson, S., Kotsianidis, I., Mollica, L., et al. (2006). Hypomorphic promoter mutation in PIGM causes inherited glycosylphosphatidylinositol deficiency. Nat Med *12*, 846–851.

Almeida, A., Layton, M., and Karadimitris, A. (2009). Inherited glycosylphosphatidyl inositol deficiency: A treatable CDG. BBA - Molecular Basis of Disease *1792*, 874–880.

Alvaro, A., Solà, R., Rosales, R., Ribalta, J., Anguera, A., Masana, L., and Vallvé, J.C. (2008). Gene expression analysis of a human enterocyte cell line reveals downregulation of cholesterol biosynthesis in response to short-chain fatty acids. IUBMB Life *60*, 757–764.

Andriamihaja, M., Chaumontet, C., Tome, D., and Blachier, F. (2009). Butyrate metabolism in human colon carcinoma cells: Implications concerning its growth-inhibitory effect. Journal of Cellular Physiology *218*, 58–65.

Balasubramanyam, K. (2004). Curcumin, a Novel p300/CREB-binding Protein-specific Inhibitor of Acetyltransferase, Represses the Acetylation of Histone/Nonhistone Proteins and Histone Acetyltransferase-dependent Chromatin Transcription. J Biol Chem *279*, 51163–51171.

Baniahmad, A., Ha, I., Reinberg, D., Tsai, S., Tsai, M.-J., and O'Malley, B. (1993). Interaction of human thyroid hormone receptor 1Bwith transcription factor TFIIB may mediate target. Proc. Natl. Acad. Sci. U.S.A. *90,* 8832–8836.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Bantscheff, M., Hopf, C., Savitski, M.M., Dittmann, A., Grandi, P., Michon, A.-M., Schlegl, J., Abraham, Y., Becher, I., Bergamini, G., et al. (2011). Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. Nature Biotechnology *29*, 255–265.

Barneda-Zahonero, B., and Parra, M. (2012). Histone deacetylases and cancer. Molecular Oncology *6*, 579–589.

Barrasa, J.I., Olmo, N., Santiago-Gómez, A., Lecona, E., Anglard, P., Turnay, J., and Lizarbe, M.A. (2012). Histone deacetylase inhibitors upregulate MMP11 gene expression through Sp1/Smad complexes in human colon adenocarcinoma cells. BBA - Molecular Cell Research *1823*, 570–581.

Batshaw, M.L., MacArthur, R.B., and Tuchman, M. (2001). Alternative pathway therapy for urea cycle disorders: twenty years later. Journal of Pediatrics. *140,* 46-54.

Baum, C., Weissman, I., Tsumakoto, A., Buckle, A.-M., and Peault, B. (1992). Isolation of a candidate human hematopoietic stem-cell population. Proc. Natl. Acad. Sci. U.S.A. *89,* 2804–2808.

Bayoumi, R.A., Nurpe-Kamal, M.S., Tadayyon, M., Mohamed, K.K., Mahhoob, B.H., Qureshi, M.M., Lakhani, M.S., Awaad, M.O., Kaeda, J.S., Vulliamy, T.J., et al. (1996). Molecular characterization of erythrocyte glucose-6-phosphate dehydrogenase deficiency in Al-Ain District, United Arab Emirates. Human Heredity *46*, 136–141.

Berg, J., Tymoczko, J., and Stryer, L. (2002). Biochemistry. W H Freeman.

Berger, S.L. (2007). The complex language of chromatin regulation during transcription. Nature *447*, 407–412.

Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The Mammalian Epigenome. Cell *128*, 669–681.

Bertrand, P. (2010). Inside HDAC with HDAC inhibitors. European Journal of Medicinal Chemistry *45*, 2095–2116.

Beutler, E. (2007). Glucose-6-phosphate dehydrogenase deficiency: a historical perspective. Blood *111*, 16–24.

Beutler, E., Kuhl, W., and Gelbart, T. (1985). 6-Phosphogluconolactonasedeficiency, a Hereditary Erythrocyte Enzyme Deficiency: Possible Interaction with Glucose-6-Phosphate Dehydrogenase Deficiency. Pnas *82*, 3876–3878.

Bhaskara, S., Chyla, B.J., Amann, J.M., Knutson, S.K., Cortez, D., Sun, Z.-W., and Hiebert, S.W. (2008). Deletion of Histone Deacetylase 3 Reveals Critical Roles in S Phase Progression and DNA Damage Control. Mol. Cell *30*, 61–72.

Bhatia, M., Wang, J., Kapp, U., Bonnet, D., and Dick, J.E. (1997). Purification of primitive human hematopoietic cells capable of repopulating immune-deficient mice. Proc. Natl. Acad. Sci. U.S.A. *94*, 5320–5325.

Bishton, M.J., Harrison, S.J., Martin, B.P., McLaughlin, N., James, C., Josefsson, E.C., Henley, K.J., Kile, B.T., Prince, H.M., and Johnstone, R.W. (2011). Deciphering the molecular and biologic processes that mediate histone deacetylase inhibitor-induced thrombocytopenia. Blood *117*, 3658–3668.

Black, A., Black, J., and Azizkhan-Clifford, J. (2001). Sp1 and Kru¨ppel-Like Factor Family of Transcription Factors in Cell Growth Regulation and Cancer. Journal of Cellular Physiology *188,* 143–160.

Blackwell, L., Norris, J., Suto, C.M., and Janzen, W.P. (2008). The use of diversity profiling to characterize chemical modulators of the histone deacetylases. Life Sciences *82*, 1050–1058.

Bolden, J.E., Peart, M.J., and Johnstone, R.W. (2006). Anticancer activities of histone deacetylase inhibitors. Nat Rev Drug Discov *5*, 769–784.

Bolton-Maggs, P., and Pasi, J. (2003). Haemophilias A and B. Lancet *361*, 1801–1809.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular Signals of Epigenetic States. Science *330*, 612–616.

Bordonaro, M., Tewari, S., Cicco, C.E., Atamna, W., and Lazarova, D.L. (2011). A Switch from Canonical to Noncanonical Wnt Signaling Mediates Drug Resistance in Colon Cancer Cells. PLoS ONE *6*, e27308.

Bouwaman, P., and Philipsen, S. (2002). Regulation of the activity of Sp1-related transcription factors. Molecular and Cellular Endocrinology *195*, 27–38.

Bradner, J.E., West, N., Grachan, M.L., Greenberg, E.F., Haggarty, S.J., Warnow, T., and Mazitschek, R. (2010). Chemical phylogenetics of histone deacetylases. Nat. Methods *6*, 238–243.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. Nature *371*, 435–438.

Briggs, M.R., Kadonaga, J.T., Bell, S.P., and Tijan, R. (1986) Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. Science. *234,* 47-52.

Brownell, J.E., and Allis, D.C. (2001). Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. Current Opinion in Genetics & Development *6,* 176–184.

Busfield, S.J., Meyer, G.T., and Klinken, S.P. (1993). Erythropoietin induced ultrastructural alterations to J2E cells and loss of proliferative capacity with terminal differentiation. Growth Factors *9*, 317–328.

Campos, E.I., and Reinberg, D. (2009). Histones: Annotating Chromatin. Annu. Rev. Genet. *43*, 559–599.

Canaff, L., Zhou, X., and Hendy, G. (2008). The Proinflammatory Cytokine, Interleukin-6, Up-regulates Calcium-sensing Receptor Gene Transcription viaStat1/3 and Sp1/3. J Biol Chem *283,* 13586–13600.

Candido, P., Reeves, R., and Davie, J.R. (1978). Sodium Butyrate Inhibits Histone Deacetylation in Cultured Cells. Cell *14,* 105–133.

Cang, S., Ma, Y., and Liu, D. (2009). New clinical developments in histone deacetylase inhibitors for epigenetic therapy of cancer. J Hematol Oncol *2*, 22-33.

Canzio, D., Liao, M., Naber, N., Pate, E., Larson, A., Wu, S., Marina, D.B., Garcia, J.F., Madhani, H.D., Cooke, R., et al. (2013). A conformational switch in HP1 releases auto-inhibition to drive heterochromatin assembly. Nature 1–8.

Cao, H., Jung, M., and Stamatoyannopoulos, G. (2005). Hydroxamide derivatives of short-chain fatty acid have erythropoietic activity and induce γ gene expression in vivo. Experimental Hematology *33*, 1443–1449.

Cappelini, M.D., Martiniez di Montemuros, F., De Bellis, G., Debernardi, S., Dotti, C., and Fiorelli, G. (1996). Multiple G6PD mutations are associated with a clinical and biochemical phenotype similar to that of G6PD Mediterranean. Blood *87*, 3953–3958.

Cappellini, M.D., and Fiorelli, G. (2008). Glucose-6-phosphate dehydrogenase deficiency. Lancet *371,* 64–74.

Caputo, V., Costa, J., Makarona, K., Georgiou, E., Layton, M., Roberts, I., and Karadimitris, A. (2013). Mechanism of Polycomb recruitment to CpG islands revealed by inherited disease-associated mutation. Human Molecular Genetics *22*, 3187–3194.

Castagnola, M., Messana, I., Sanna, M.T., and Giardina, B. (2010). Oxygen-linked modulation of erythrocyte metabolism: state of the art. Blood Transfusion *3,* 53-58.

Caterino, T., and Hayes, J. (2007). Chromatin structure depends on what's in the Nucleosome's pocket. Nature Structural & Molecular Biology *14*, 1056–1058.

Chang, J., Varghese, D.S., Gillam, M.C., Peyton, M., Modi, B., Schiltz, R.L., Girard, L., and Martinez, E.D. (2011). Differential response of cancer cells to HDAC inhibitors trichostatin A and depsipeptide. British Journal of Cancer *106*, 116–125.

Chaurasia, P., Berenzon, D., and Hoffman, R. (2011). Chromatin-modifying agents promote the ex vivo production of functional human erythroid progenitor cells. Blood *117*, 4632–4641.

Chen, J.-L., Attardi, L., Verrijzer, P., Yokomori, K., and Tjian, R. (1994). Assembly of Recombinant TFIIDReveals Differential Coactivator Requirements for Distinct Transcriptional Activators. Cell *79,* 93–105.

Chou, C.-W., Wu, M.-S., Huang, W.-C., and Chen, C.-C. (2011). HDAC Inhibition Decreases the Expression of EGFR in Colorectal Cancer Cells. PLoS ONE *6*, e18087.

Climent, F., Roset, F., Repiso, A., and Perez de la Ossa, P. (2009). Red cell glycolytic enzyme disorders caused by mutations: an update. Cardiovasc Hematol Disord Drug Targets *9*, 95–106.

Conneally, E., Cashman, J., Petzer, A., and Eaves, C. (1997). Expansion in vitro of transplantable human cord blood stem cells demonstrated using a quantitative assay of their lympho-myeloid repopulating activity in nonobese diabetic–scid. Proc. Natl. Acad. Sci. U.S.A. *94,* 9836–9841.

Constantinescu, S., Ghaffari, S., and Lodish, H. (1999). The Erythropoietin Receptor: Structure, Activation and Intracellular Signal Transduction. TEM *10,* 18–23.

Courey, A.J., and Tijan, R. (1988). Analysis of Spl In Vivo Reveals Mutiple Transcriptional Domains, Including a Novel Glutamine-Rich Activation Motif. Cell *55,* 887–898.

Courey, A.J., Holtzman, D., Jackson, S., and Tijan, R. (1989). Synergistic Activation by the Glutamine-Rich Domains of Human Transcription Factor Spl. Cell *59,* 827–836.

Daly, K., and Shirazi-Beechey, S. (2006). Microarray Analysis of Butyrate Regulated Genes in Colonic Epithelial Cells. DNA and Cell Biology *25*, 49–62.

Daniel, P., Brazier, M., Cerutti, I., Pieri, F., Tardivel, I., Desmet, G., Baillet, J., and Chany, C. (1989). Pharmacokinetic study of butyric acid administered in vivo as sodium and arginine butyrate salts. Clin Chim Acta *181*, 255–263.

Davie, J.R. (2003). Inhibition of Histone Deacetylase Activity by Butyrate. Nutritional Proteomics in Cancer Prevention 2485–2494.

Davie, J.R., He, S., Li, L., Sekhavat, A., Espino, P., Drobic, B., Dunn, K.L., Sun, J.-M., Chen, H.Y., Yu, J., et al. (2008). Nuclear organization and chromatin dynamics – Sp1, Sp3 and histone deacetylases. Advances in Enzyme Regulation *48*, 189–208.

de Ruijter, A.J.M., van Gennip, A.H., Caron, H.N., Kemp, S., and van Kuilenburg, A.B.P. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem. J. *370*, 737–749.

de Vooght, K.M.K., van Solinge, W.W., van Wesel, A.C., Kersting, S., and van Wijk, R. (2009). First mutation in the red blood cell-specific promoter of hexokinase combined with a novel missense mutation causes hexokinase deficiency and mild chronic hemolysis. Haematologica *94*, 1203–1210.

Deberardinis, R.J., and Thompson, C.B. (2012). Cellular metabolism and disease: what do metabolic outliers teach us? Cell *148*, 1132–1144.

deFazio, A., Chiew, Y.-E., Donoghue, C., Lee, C., Suther, and Sutherland, R. (2001). Effect of Sodium Butyrate on Estrogen Receptor and Epidermal GrowthFactorReceptorGenEexpressioninHumanBreast Cancer Cell Lines. J Biol Chem *267*, 18008–18012.

Delehanty L., Bullock G., and Goldfarb A. (2012). Protein kinase D-HDAC5 signaling regulates erythropoiesis and contributes to erythropoietin crosstalk with GATA1. Blood *120*, 4219-4228.

Delcuve, G.P., Khan, D.H., and Davie, J.R. (2012). Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. Clinical Epigenetics *4*, 5-18.

Dev, A., Fang, J., Sathyanarayana, P., Pradeep, A., Emerson, C., and Wojchowski, D.M. (2010). During EPO or anemia challenge, erythroid progenitor cells transit

through a selectively expandable proerythroblast pool. Blood *116*, 5334–5346.

Dillon, N. (2008). The Impact of Gene Location in the Nucleus on Transcriptional Regulation. Developmental Cell *15*, 182–186.

Doetzlhofer, A., Rotheneder, H., Lagger, G., Koranda, M., Kurtev, V., Brosch, G., Wintersberger, E., and Seiser, C. (1999). Histone Deacetylase 1 can repress transcription by binding to Sp1. Molecular and Cellular Biology *19,* 5504–5511.

Doulatov, S., Notta, F., Laurenti, E., and Dick, J.E. (2012). Hematopoiesis: A Human Perspective. Stem Cell *10*, 120–136.

Dover, G., Brusilow, S., and Charache, S. (1994). Induction of fetal hemoglobin production in subjects with sickle cell anemia by oral sodium phenylbutyrate. Blood *84,* 339–343.

Dunah, A.W. (2002). Sp1 and TAFII130 Transcriptional Activity Disrupted in Early Huntington's Disease. Science *296*, 2238–2243.

Dyson, J.E.D., Daniel, J., and Surrey, C.R. (1992). The effect of sodium butyrate on the growth characteristics of human cervix tumour cells. British Journal of Cancer *65,* 803–808.

Dzierzak, E., and Philipsen, S. (2013). Erythropoiesis: Development and Differentiation. Cold Spring Harbor Perspectives in Medicine *3,* a011601.

England, S.J., McGrath, K.E., Frame, J.M., and Palis, J. (2011). Immature erythroblasts with extensive ex vivo self-renewal capacity emerge from the early mammalian fetus. Blood *117*, 2708–2717.

Esposito, G., Vitagliano, L., Constanzo, P., Borrelli, L., Barone, R., Pavone, L., Izzo, P., Zagari, A., and Salvatore, F. (2004). Human aldolase A natural mutants: relationship between flexibility of the C-terminal region and enzyme function. Journal of Biochemistry *380*, 51–56.

Finnin, M., Doniglan, J., Cohen, A., Richon, V., Rifkind, R., Marks, P.A., Breslow, R., and Pavletich, N. (1999). Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. Nature *401,* 188–193.

Fischle, W., Dequiedt, F., Hendzel, M., Guenther, M.G., Lazar, M., Voelter, W., and Verdin, E. (2002). Enzymatic Activity Associated with Class II HDACs Is Dependent on a Multiprotein Complex Containing HDAC3 and SMRT/N-CoR. Mol. Cell *9,* 45–57.

Fraczek, J., Vanhaecke, T., and Rogiers, V. (2013). Toxicological and metabolic considerations for histone deacetylase inhibitors. Expert Opin. Drug Metab. Toxicol. *9,* 441–457.

Franchini, M., and Mannucci, P.M. (2013). Hemophilia A in the third millennium. Yblre *27*, 179–184.

Franzè, A., Ferrante, M.I., Fusco, F., Santoro, A., Sanzari, E., Martini, G., and Ursini, M.V. (1998). Molecular anatomy of the human glucose 6-phosphate dehydrogenase

core promoter. FEBS Lett. *437*, 313–318.

Friend, C., Scher, W., Holland, J.G., and Sato, T. (1971). Hemoglobin Synthesis in Murine Virus-Induced Leukemic Cells In Vitro: Stimulation of Erythroid Differentiation byDimethyl Sulfoxide. Proc. Natl. Acad. Sci. U.S.A. *68,* 378–382.

Fusco, F., Mercadante, V., Miano, M.G., and Ursini, M.V. (2006). Multiple regulatory regions and tissue-specific transcription initiation mediate the expression of NEMO/IKKgamma gene. Gene *383*, 99–107.

Galgoczy, P., Rosenthal, A., and Platzer, M. (2001). Human-mouse comparative sequence analysis of the NEMO gene reveals an alternative promoter within the neighboring G6PD gene. Gene *271,* 93–98.

Gangenahalli, G., Gupta, P., Saluja, D., Verma, Y., Kishore, V., Chandra, R., Sharma, R.K., and Ravindranath, T. (2005). Stem Cell Fate Specification: Role of Master Regulatory Switch Transcription Factor PU.1 in Differential Hematopoiesis. Stem Cells Dev *14,* 140–152.

García, M., Pujol, A., Ruzo, A., Riu, E., Ruberte, J., Arbós, A., Serafín, A., Albella, B., Felíu, J.E., and Bosch, F. (2009). Phosphofructo-1-Kinase Deficiency Leads to a Severe Cardiac and Hematological Disorder in Addition to Skeletal Muscle Glycogenosis. PLoS Genet *5*, e1000615.

Giebel, B., and Punzel, M. (2008). Lineage development of hematopoietic stem and progenitor cells. Biological Chemistry *389,* 318-824.

Glaser, K.B. (2007). HDAC inhibitors: Clinical update and mechanism-based potential. Biochemical Pharmacology *74*, 659–671.

Glozak, M.A., Sengupta, N., Zhang, X., and Seto, E. (2005). Acetylation and deacetylation of non-histone proteins. Gene *363*, 15–23.

Goodell, M.A. (2013). Epigenetics in hematology: introducing a collection of reviews. Blood *121*, 3059–3060.

Gregoretti, I., Lee, Y.-M., and Goodson, H.V. (2004). Molecular Evolution of the Histone Deacetylase Family: Functional Implications of Phylogenetic Analysis. Journal of Molecular Biology *338*, 17–31.

Gui, C., Ngo, L., Xu, W.S., Richon, V.M., and Marks, P.A. (2004). Histone deacetylase (HDAC) inhibitor activation of p21WAF1 involves changes in promoter-associated proteins, including HDAC1. Proc. Natl. Acad. Sci. U.S.A. *101*, 1241–1246.

Haberland, M., Montgomery, R.L., and Olson, E.N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. Nature Publishing Group *10*, 32–42.

Hagen, G., Muller, S., Beato, M., and Suske, G. (1992). Cloning by recognition site screening of two novel GT box binding proteins:a family of Sp1 related genes. Nucleic Acids Research *20,* 5519–5525.

Halsall, J., Gupta, V., O'Neill, L.P., Turner, B.M., and Nightingale, K.P. (2012). Genes Are Often Sheltered from the Global Histone Hyperacetylation Induced by HDAC Inhibitors. PLoS ONE *7*, e33453.

Hattangadi, S.M., Wong, P., Zhang, L., Flygare, J., and Lodish, H.F. (2011). From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. Blood *118*, 6258–6268.

Heerdt, B., Houston, M., Anthony, G., and Augenlicht, L. (1999). Initiation of Growth Arrest and Apoptosis of MCF-7 Mammary Carcinoma Cells by Tributyrin, a Triglyceride Analogue of the Short-Chain Fatty Acid Butyrate, Is Associated with Mitochondrial Activity. *59,* 1584–1591.

Hodawadekar, S.C., and Marmorstein, R. (2007). Chemistry of acetyl transfer by histone modifying enzymes: structure, mechanism and implications for effector design. Oncogene *26*, 5528–5540.

Hoffbrand, V., Catovsky, D., and Tuddenham, E. (2005). Postgraduate Haematology. Blackwell Publishing 1–1087.

Hou, M., Wang, X., Popov, N., Zhang, A., Zhao, X., Zhou, R., Zetterberg, A., Bjorkholm, M., and Xu, D. (2002). The Histone Deacetylase Inhibitor Trichostatin A Derepresses the Telomerase Reverse Transcriptase (hTERT) Gene in Human Cells. Experimental Cell Research *274*, 25–34.

Hoyer, J.D., Allen, S.L., Beutler, E., Kubik, K., West, C., and Fairbanks, V.F. (2004). Erythrocytosis due to bisphosphoglycerate mutase deficiency with concurrent glucose-6-phosphate dehydrogenase (G-6-PD) deficiency. Am. J. Hematol. *75*, 205–208.

Huang, W. (2005). Trichostatin A Induces Transforming Growth Factor Type II Receptor Promoter Activity and Acetylation of Sp1 by Recruitment of PCAF/p300 to a Sp1-NF-Y Complex. J Biol Chem *280*, 10047–10054.

Huang, X., Cho, S., and Spangrude, G.J. (2007). Hematopoietic stem cells: generation and self-renewal. Cell Death and Differentiation *14*, 1851–1859.

Huddleston, H. (2003). Functional p85 gene is required for normal murine fetal erythropoiesis. Blood *102*, 142–145.

Ikuta, T., Atweh, G., Boosalis, V., White, G.L., Da Fonseca, S., Boosalis, M., Faller, D.V., and Perrine, S.P. (1998). Cellular and molecular effects of a pulse butyrate regimen and new inducers of globin gene expression and hematopoiesis. Annals of the New York Academy of Sciences *850*, 87–99.

Ishii, K., and Laemmli, U. (2003). Structural and Dynamic Functions Establish Chromatin Domains. Mol. Cell *11,* 237–248.

Iwasaki, H., and Akashi, K. (2007). Hematopoietic developmental pathways: on cellular basis. Oncogene *26*, 6687–6696.

Jansen, G., Koenderman, L., Rijksen, G., Cats, B.P., and Staal, G.E.J. (1985). Characteristics of hexokinase, pyruvate kinase, and glucose-6-phosphate dehydrogenase during adult and neonatal reticulocyte maturation. Am. J. Hematol. *20*, 203–215.

Jaster, R., Bittorf, T., Klinken, P., and Brock, J. (1996). Inhibition of Proliferation but not Erythroid Differentiation of JZE Cells by Rapamycin. Biochemical Pharmacology *51,* 1181–1185.

Jelkmann, W., and Metzen, E. (1996). Erythropoietin in the control of red cell production. Annals of Anatomy *178*, 391–403.

Ji, P., Yeh, V., Ramirez, T., Murata-Hori, M., and Lodish, H.F. (2010). Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts. Haematologica *95*, 2013–2021.

Ji, P., Murata-Hori, M., and Lodish, H.F. (2011). Formation of mammalian erythrocytes: chromatin condensation and enucleation. Trends in Cell Biology *21*, 409–415.

Jin, D.Y., and Jeang, K.T. (1999). Isolation of full-length cDNA and chromosomal localization of human NF-kappaB modulator NEMO to Xq28. J. Biomed. Sci.*6,* 115-120.

Johnsson, A., Durand-Dubief, M.E.L., n, Y.X.-F.E., nnerblad, M.R.O., Ekwall, K., and Wright, A. (2009). HAT-HDAC interplay modulates global histone H3K14 acetylation in gene-coding regions during stress. Embo J *10*, 1009–1014.

Joseph, J., Mudduluru, G., Antony, S., Vashistha, S., Ajitkumar, P., and Somasundaram, K. (2004). Expression profiling of sodium butyrate (NaB)-treated cells: identification of regulation of genes related to cytokine signaling and cancer metastasis by NaB. Oncogene *23*, 6304–6315.

Kadam, S. (2000). Functional selectivity of recombinant mammalian SWI/SNF subunits. Genes Dev. *14*, 2441–2451.

Kadonaga, J.T., Courey, A.J., Ladika, J., and Tijan, R. (1988). Distinct regions of Sp1 modulate DNA binding and transcriptional activation. Science.*242,* 1566-1570.

Kadosh, D., and Struhl, K. (1997). Repression by Ume6 Involves Recruitment ofa Complex Containing Sin3 Corepressor and Rpd3 Histone Deacetylase to Target Promoters. Cell *89,* 365–371.

Kao, H.Y. (2001). Isolation and Characterization of Mammalian HDAC10, a Novel Histone Deacetylase. Journal of Biological Chemistry *277*, 187–193.

Kim, T.Y., Kim, I.S., Jong, H.-S., Lee, J.W., Kim, T.-Y., Jung, M., and Bang, Y.-J. (2008). Transcriptional induction of DLC-1 gene through Sp1 sites by histone deacetylase inhibitors in gastric cancer cells. Exp Mol Med *40*, 639-645.

Kish, H., Mukai, T., Hirono, A., Fujii, H., Miwa, S., and Hori, K. (1987). Human aldolaseA deficiencyassociatedwithahemolyticanemia: Thermolabile aldolase due to a single base mutation. Proc. Natl. Acad. Sci. U.S.A. *87*, 8623–8627.

Knutson, S.K., Chyla, B.J., Amann, J.M., Bhaskara, S., Huppert, S.S., and Hiebert, S.W. (2008). Liver-specific deletion of histone deacetylase 3 disrupts metabolic transcriptional networks. Embo J *27*, 1017–1028.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science.*184,* 865-868.

Koury, M. (2011). Self-renewal in late-stage erythropoiesis. Blood *117*, 2562–2564.

Kouzarides, T. (2007). Chromatin Modifications and Their Function. Cell *128*, 693–705.

Kramer, O., Gottlicher, M., and Heinzel, T. (2001). Histone deacetylase as a therapeutic target. Trends in Endocrinology and Metabolism *12*, 294–300.

Kubicek, S., Gilbert, J., Fomina-Yadlin, D., Gitlin, A., Yuan, Y., Wagner, F., Holson, E., Luo, T., Lewis, T., Taylor, B., et al. (2012). Chromatin-targeting small molecules cause class-specific transcriptional changes in pancreatic endocrine cells. Proc. Natl. Acad. Sci. U.S.A.*109*, 5364–5369.

Kugler, W., and Lakomek, M. (2000). Glucose-6-phosphate isomerase deficiency. Baillieres Best Pract Res Clin Haematol *13*, 89–101.

Kulkarni, R., and Soucie, J.M. (2011). Pediatric Hemophilia: A Review. Semin Thromb Hemost *37*, 737–744.

Kundu, T.K., Palhan, V.B., Wang, Z., An, W., Cole, P.A., and Roeder, R.G. (2000). Activator-Dependent Transcription from Chromatin In Vitro Involving Targeted Histone Acetylation by p300. Mol. Cell *6,* 551–561.

Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping Global Histone Acetylation Patterns to Gene Expression. Cell *117,* 721–733.

Laurenti, E., and Dick, J.E. (2012). Molecular and functional characterization of early human hematopoiesis. Annals of the New York Academy of Sciences *1266*, 68–71.

Law, A.Y.S., Yeung, B.H.Y., Ching, L.Y., and Wong, C.K.C. (2011). Sp1 is a transcription repressor to stanniocalcin-1 expression in TSA-treated human colon cancer cells, HT29. J. Cell. Biochem. *112*, 2089–2096.

Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol *8*, 284–295.

Lemarchandel, V., Joulin, V., Valentin, C., Rosa, R., Galacteros, F., J, R., and Cohen-Solal, M. (1992). Compound heterozygosity in a complete erythrocyte bisphosphoglycerate mutase deficiency. Blood *80*, 2643–2649.

Lenhard, B., Sandelin, A., and Carninci, P. (2012). Regulatory elements: Metazoan promoters: emerging characteristics and insights into transcriptional regulation. Nature Publishing Group *13*, 233–245.

Li, G., and Reinberg, D. (2011). Chromatin higher-order structures and gene regulation. Current Opinion in Genetics & Development *21*, 175–186.

Li, L., and Davie, J.R. (2010). The role of Sp1 and Sp3 in normal and cancer cell biology. Annals of Anatomy *192*, 275–283.

Lodish, H., Flygare, J., and Chou, S. (2010). From stem cell to erythroblast: Regulation of red cell production at multiple levels by multiple hormones. IUBMB Life *62*, 492–496.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389* 251–260.

Mai, A., Jelicic, K., Rotili, D., Di Noia, A., Alfani, E., Valente, S., Altucci, L., Nebbioso, A., Massa, S., Galanello, R., et al. (2007). Identification of Two New Synthetic Histone Deacetylase Inhibitors That Modulate Globin Gene Expression in Erythroid Cells from Healthy Donors and Patients with Thalassemia. Molecular Pharmacology *72*, 1111–1123.

Majumdar, G., Adris, P., Bhargava, N., Chen, H., and Raghow, R. (2012). Pan-histone deacetylase inhibitors regulate signaling pathways involved in proliferative and pro-inflammatory mechanisms in H9c2 cells. BMC Genomics *13*, 708–730.

Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R., and Pazdur, R. (2007). FDA Approval Summary: Vorinostat for Treatment of Advanced Primary Cutaneous T-Cell Lymphoma. The Oncologist *12*, 1247–1252.

Mannucci, P.M., and Tuddenham, E. (2001). THe HEmophilias — FRom ROyal GEnes to GEne THerapy. The New England Journal of Medicine *344*, 1773–1779.

Margueron, R., and Reinberg, D. (2010). Chromatin structure and the inheritance of epigenetic information. Nature Publishing Group *11*, 285–296.

Mariadason, J., Corner, G., and Augenlicht, L. (2000). Genetic Reprogramming in Pathways of Colonic Cell Maturation Induced by Short Chain Fatty Acids: Comparison with Trichostatin A, Sulindac, and Curcumin and Implications for Chemoprevention of Colon Cancer. Cancer Research *60*, 4561–4572.

Marks, P.A. (2007). Discovery and development of SAHA as an anticancer agent. Oncogene *26*, 1351–1356.

Marks, P.A. (2010). Histone deacetylase inhibitors: A chemical genetics approach to understanding cellular functions. BBA - Gene Regulatory Mechanisms *1799*, 717–725.

Marks, P.A., and Breslow, R. (2007). Dimethyl sulfoxide to vorinostat: development of this histone deacetylase inhibitor as an anticancer drug. Nature Biotechnology *25*, 84–90.

Marks, P.A., Richon, V., and Rifkind, R. (2000). Histone Deacetylase Inhibitors: Inducers of Differentiation or Apoptosis of Transformed Cells. Journal of National Cancer Institute *92,* 1210–1218.

Martini, G., Toniolo, D., vulliamy, T.J., Luzzatto, L., Dono, R., Viglietto, G., Paonessa, G., D'Urso, M., and Persico, G.M. (1986). Structural analysis of the X-linked gene encoding human glucose 6-phosphate dehydrogenase. Embo J *5,* 1849–1855.

Mason, P. (1996). New insights into G6PD deficiency. Br. J. Haematol.*62,* 9-10.

McGonigle, D.P., Lalloz, M., Wild, B.J., and Layton, D.M. (1998). G6PD-Brighton: A new class I glucose-6-phosphate dehydrogenase variant due to a deletion in exon 13. Br. J. Haematol. *1*, 51–58.

Menounos, P.G., Garinis, G.A., and Patrinos, G.P. (2003). Glucose-6-phosphate dehydrogenase deficiency does not result from mutations in the promoter region of the G6PD gene. J. Clin. Lab. Anal. *17*, 90–92.

Metallo, C.M., and Vander Heiden, M.G. (2013). Understanding Metabolic Regulation and Its Influence on Cell Physiology. Mol. Cell *49*, 388–398.

Migliaccio, A.R. (2010). Erythroblast enucleation. Haematologica *95*, 1985–1988.

Miki, Y., Mukae, S., Murakami, M., Ishikawa, Y., Ishii, T., Ohki, H., Matsumoto, M., and Komiyama, K. (2007). Butyrate Inhibits Oral Cancer Cell Proliferation and Regulates Expression of Secretory Phospholipase A2-X and COX-2. Anticancer Research *27*, 1493–1502.

Miller, J.L. (2004a). A genome-based approach for the study of erythroid biology and disease. Blood Cells Mol. Dis. *32*, 341–343.

Miller, S.J. (2004b). Cellular and physiological effects of short-chain fatty acids. Mini Reviews in Medical Chemistry.*4,* 839- 845.

Minucci, A., Moradkhani, K., Hwang, M.J., Zuppi, C., Giardina, B., and Capoluongo, E. (2012). Glucose-6-phosphate dehydrogenase (G6PD) mutations database: Review of the "old" and update of the new mutations. Blood Cells Mol. Dis. *48*, 154–165.

Mitsiades, C., Mitsiades, N., McMullan, C., Poulaki, V., Shringarpure, R., Hideshima, T., Akiyama, M., Chauhan, D., Munshi, N., Gu, X., et al. (2004). Transcriptional signature of histone deacetylaseinhibition in multiple myeloma: Biological andclinical implications. Proc. Natl. Acad. Sci. U.S.A. *101,* 540–545.

Montgomery, R.L., Davis, C.A., Potthoff, M.J., Haberland, M., Fielitz, J., Qi, X., Hill, J.A., Richardson, J.A., and Olson, E.N. (2007). Histone deacetylases 1 and 2 redundantly regulate cardiac morphogenesis, growth, and contractility. Genes Dev. *21*, 1790–1802.

Nightingale, K.P., Wellinger, R.E., Sogo, J.M., and Bechker, P.B. (1998). Histone acetylation facilitates RNA polymerase II transcription of the Drosophila hsp26 gene in chromatin. EMBO J *17,* 2865–2877.

Novershtern, N., Subramanian, A., Lawton, L.N., Mak, R.H., Haining, W.N., McConkey, M.E., Habib, N., Yosef, N., Chang, C.Y., Shay, T., et al. (2011). Densely Interconnected Transcriptional Circuits Control Cell States in Human Hematopoiesis. Cell *144*, 296–309.

Nunes, M.J., Milagre, I., Schnekenburger, M., Gama, M.J., Diederich, M., and Rodrigues, E. (2010). Sp proteins play a critical role in histone deacetylase inhibitor-mediated derepression of CYP46A1gene transcription. Journal of Neurochemistry *113*, 418–431.

Ohene-Abuakwa, Y. (2005). Two-phase culture in Diamond Blackfan anemia: localization of erythroid defect. Blood *105*, 838–846.

Oláh, J., Orosz, F., Puskás, L.G., Hackler, L., Jr, Horányi, M., Polgár, L., Hollán, S., and Ovádi, J. (2005). Triosephosphate isomerase deficiency: consequences of an inherited mutation at mRNA, protein and metabolic levels. Biochem. J. *392*, 675.

Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: An Evolving Paradigm for Stem Cell Biology. Cell *132*, 631–644.

Orosz, F., Oláh, J., and Ovádi, J. (2009). Triosephosphate isomerase deficiency: New insights into an enigmatic disease. BBA - Molecular Basis of Disease *1792*, 1168–1174.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet *36*, 1065–1071.

Pace, B.S. (2002). Short-chain fatty acid derivatives induce fetal globin expression and erythropoiesis in vivo. Blood *100*, 4640–4648.

Palis, J., and Segel, G.B. (1998). Developmental biology of erythropoiesis. Blood *12, 1* 106–114.

Palis, J. (2008). Ontogeny of erythropoiesis. Current Opinion in Hematology *15,* 155–161.

Papantonis, A., Larkin, J.D., Wada, Y., Ohta, Y., Ihara, S., Kodama, T., and Cook, P.R. (2010). Active RNA Polymerases: Mobile or Immobile Molecular Machines? PLoS Biol *8*, e1000419.

Parisi, F., Wirapati, P., and Naef, F. (2007). Identifying synergistic regulation involving c-Myc and sp1 in human tissues. Nucleic Acids Research *35*, 1098–1107.

Park, S.H. (2001). Transcriptional Regulation of the Transforming Growth Factor beta Type II Receptor Gene by Histone Acetyltransferase and Deacetylase Is Mediated by NF-Y in Human Breast Cancer Cells. Journal of Biological Chemistry *277*, 5168–5174.

Parthun, M.R. (2007). Hat1: the emerging cellular roles of a type B histone acetyltransferase. Oncogene *26*, 5319–5328.

Parthun, M.R. (2012). Histone acetyltransferase 1: More than just an enzyme? BBA - Gene Regulatory Mechanisms *1819*, 256–263.

Pascal, E., and Tjian, R. (1991). Different activation domains of Sp1 govern formation of multimers and mediate transcriptional synergism. Genes Dev. *5*, 1646–1656.

Peart, M.J., Smyth, G., van Laar, R., Bowtell, D., Richon, V., Marks, P.A., Holloway, A., and Johnstone, R.W. (2005). Identification and functional significance of genesregulated by structurally different histonedeacetylase inhibitors. *Proc*. Natl. Acad. Sci. U.S.A. *102,* 3697–3702.

Perrine, S.P., Hermine, O., Small, T., Suarez, F., O'Reilly, R., Boulad, F., Fingeroth, J., Askin, M., Levy, A., Mentzer, S.J., et al. (2007). A phase 1/2 trial of arginine butyrate and ganciclovir in patients with Epstein-Barr virus-associated lymphoid malignancies. Blood *109*, 2571–2578.

Perrine, S.P. (2008). Fetal globin stimulant therapies in the beta-hemoglobinopathies: principles and current potential. Pediatric Annual *37*, 339–346.

Perrine, S.P., Castaneda, S.A., Chui, D.H.K., Faller, D.V., Berenson, R.J., Siritanaratku, N., and Fucharoen, S. (2010). Fetal globin gene inducers: novel agents and new potential. Annals of the New York Academy of Sciences *1202*, 158–164.

Persico, G.M., Viglietto, G., Martini, G., Toniolo, D., Paonessa, G., Moscatelli, C., Dono, R., vulliamy, T.J., Luzzatto, L., and D'Urso, M. (1986). Isolation of human glucose-6-pbosphate debydrogenase (G6PD) cDNA clones: primary structure of the protein and unusual 5' non-coding region. Nucleic Acids Research *14,* 2511–2522.

Peserico, A., and Simone, C. (2011). Physical and Functional HAT/HDAC Interplay Regulates Protein Acetylation Balance. Journal of Biomedicine and Biotechnology *2011*, 1–10.

Pey, A.L., Mesa-Torres, N., Chiarelli, L.R., and Valentini, G. (2013). Structural and Energetic Basis of Protein Kinetic Destabilization in Human Phosphoglycerate Kinase 1 Deficiency. Biochemistry *52*, 1160–1170.

Phadke, M.S., Krynetskaia, N.F., Mishra, A.K., and Krynetskiy, E. (2009). Glyceraldehyde 3-Phosphate Dehydrogenase Depletion Induces Cell Cycle Arrest and Resistance to Antimetabolites in Human Carcinoma Cell Lines. Journal of Pharmacology and Experimental Therapeutics *331*, 77–86.

Philippe, M., Larondelle, Y., Lemaigre, F., Mariamé, B., Delhez, H., Mason, P., Luzzatto, L., and Rousseau, G.G. (1994). Promoter function of the human glucose-6-phosphate dehydrogenase gene depends on two GC boxes that are cell specifically controlled. European Journal of Biochemistry / FEBS *226*, 377–384.

Philipsen, S., and Suske, G. (1999). A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. Nucleic Acids Research *27,* 2991–3000.

Piomelli, S., Corash, L.M., Davenport, D.D., Miraglia, J., and Amorosi, E.L. (1968). In vivo lability of glucose-6-phosphate dehydrogenase in G6PDA- and G6PD Mediterranean deficiency. J. Clin. Invest. *47*, 940–948.

Poggi, V., Town, M., Foulkes, N., and Luzzatto, L. (1990). Identification of a single base change in a new human mutant glucose-6-phosphate dehydrogenase gene by polymerase-chain-reaction amplification for the entire coding region from genomic DNA. Biochem. J. *271*, 157–160.

Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. Nature Biotechnology *28*, 1057–1068.

Pretsch, W., and Favor, J. (2007). Genetic, biochemical, and molecular characterization of nine glyceraldehyde-3-phosphate dehydrogenase mutants with reduced enzyme activity in Mus musculus. Mamm Genome *18*, 686–692.

Pretsch, W., Charles, D.J., and Merkle, S. (1988). X-linked glucose-6-phosphate dehydrogenase deficiency in Mus musculus. Biochem Genet. *18,* 89–103.

Rada-Iglesias, A., Enroth, S., Ameur, A., Koch, C.M., Clelland, G.K., Respuela-Alonso, P., Wilcox, S., Dovey, O.M., Ellis, P.D., Langford, C.F., et al. (2007). Butyrate mediates decrease of histone acetylation centered on transcription start sites and down-regulation of associated genes. Genome Research *17*, 708–719.

Ralser, M., Heeren, G., Breitenbach, M., Lehrach, H., and Krobitsch, S. (2006). Triose Phosphate Isomerase Deficiency Is Caused by Altered Dimerization–Not Catalytic Inactivity–of the Mutant Enzymes. PLoS ONE *1*, e30.

Ram, O., Goren, A., Amit, I., Shoresh, N., Yosef, N., Ernst, J., Kellis, M., Gymrek, M., Issner, R., Coyne, M., et al. (2011). Combinatorial Patterning of Chromatin Regulators Uncovered by Genome-wide Location Analysis in Human Cells. Cell *147*, 1628–1639.

Repiso, A., Oliva, B., Vives-Corrons, J.-L., Beutler, E., Carreras, J., and Climent, F. (2006). Red cell glucose phosphate isomerase (GPI): a molecular study of three novel mutations associated with hereditary nonspherocytic hemolytic anemia. Human Mutation *27*, 1159–1159.

Resendes, K.K., and Rosmarin, A.G. (2004) Sp1 control of gene expression in myeloid cells. Crit Rev Eukaryot Gene Expr. *14,* 171-181.

Richon, V.M., Webb, Y., Merger, R., Sheppard, T., Jursic, B., Ngo, L., Civoli, F., Breslow, R., Rifkinf, R.A., and Marks, P.A. (1996). Second generation hybrid polar compounds are potent inducers of transformed cell differentiation. Proc. Natl. Acad. Sci. U.S.A. *93,* 5707–5708.

Robinson, A.R., Kwek, S.S., and Kenney, S.C. (2012). The B-Cell Specific Transcription Factor, Oct-2, Promotes Epstein-Barr Virus Latency by Inhibiting the Viral Immediate-Early Protein, BZLF1. PLoS Pathog *8*, e1002516.

Ronquist, G., Rudolphi, O., Engstrom, I., and Waldenstrom, A. (2001). Familial phosphofructokinase deficiency is associated with a disturbed calcium homeostasis in erythrocytes. Journal of Internal Medicine *249*, 85–95.

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 Mol. Dis. *40*, 148–155.

Rosmarin, A.G., Luo, M., Caprio, D., Shang, J., and Simkevits, K. (1998). Sp1 Cooperates with the ets Transcription Factor, GABP, to Activate the CD18 (beta2 leukocyte integrin) promoter.. J Biol Chem *273*, 13097–13103.

Ruwende, C., Khoo, S., Snow, R., Yates, S., Kwlatkowski, D., Gupta, S., Warn, P., Allsopp, C., Gilbert, S., Peschu, N., et al. (1995). Natural selection of hemi- and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. Nature *376*, 246–249.

Santillo, A., and Albenzio, M. (2008). Influence of Lamb Rennet Paste Containing Probiotic on Proteolysis and Rheological Properties of Pecorino Cheese. Journal of Dairy Science *91*, 1733–1742.

Schoenfelder, S., Sexton, T., Chakalova, L., Cope, N.F., Horton, A., Andrews, S., Kurukuti, S., Mitchell, J.A., Umlauf, D., Dimitrova, D.S., et al. (2009). Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet *42*, 53–61.

Sealy, L., and Chalkley, R. (1978). The Effect of Sodium Butyrate,on Histone Modification. Cell *14,* 115–121.

Sekhavat, A., Sun, J.-M., and Davie, J.R. (2007). Competitive inhibition of histone deacetylase activity by trichostatin A and butyrate. Biochem. Cell Biol. *85*, 751–758.

Sexton, T., Umlauf, D., Kurukuti, S., and Fraser, P. (2007). The role of transcription factories in large-scale structure and dynamics of interphase chromatin. Seminars in Cell & Developmental Biology *18*, 691–697.

Sherr, C., and Roberts, J. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev. *13,* 1501–1512.

Siavoshian, S., Blottiere, H., Cherbut, C., and Galmiche, J.-P. (1997). Butyrate Stimulates Cyclin D and p21 and Inhibits Cyclin-Dependent Kinase 2 Expression in HT-29 Colonic Epithelial Cells. Biochem. Biophys. Res. Commun. *232,* 169–172.

Siegel, D., Hussein, M., Belani, C., Robert, F., Galanis, E., Richon, V.M., Garcia-Vargas, J., Sanz-Rodriguez, C., and Rizvi, S. (2009). Vorinostat in solid and hematologic malignancies. J Hematol Oncol *2*, 31-42.

Sinclair, A., and Elliott, S. (2012). The effect of erythropoietin on normal and neoplastic cells. Btt *6,* 163-189.

Sleiman, S.F., Langley, B.C., Basso, M., Berlin, J., Xia, L., Payappilly, J.B., Kharel, M.K., Guo, H., Marsh, J.L., Thompson, L.M., et al. (2011). Mithramycin Is a Gene-Selective Sp1 Inhibitor That Identifies a Biological Intersection between Cancer and Neurodegeneration. Journal of Neuroscience *31*, 6858–6870.

Smith, K.T., and Workman, J.L. (2009). Histone deacetylase inhibitors: Anticancer compounds. The International Journal of Biochemistry & Cell Biology *41*, 21–25.

Sperry, J.B., Shi, X., Rempel, D.L., Nishimura, Y., Akashi, S., and Gross, M.L. (2008). A Mass Spectrometric Approach to the Study of DNA-Binding Proteins: Interaction of Human TRF2 with Telomeric DNA †. Biochemistry *47*, 1797–1807.

Stadtman, E.R., and Barker, H.A. (1949). Fatty acid synthesis by enzyme preperations of Clostrodium Kluyveri. J Biol Chem 221–236.

Stefanini, M. (1972). Chronic hemolytic anemia associated with erythrocyte enolase deficiency exacerbated by ingestion of nitrofurantoin. Am J Clin Pathol *58*, 408–414.

Suzuki, H., Forrest, A.R.R., van Nimwegen, E., Daub, C.O., Balwierz, P.J., Irvine, K.M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M.J.L., et al. (2009). The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat Genet *41*, 553–562.

Svaasand, E.K., Aasly, J., Landsem, V.M., and Klungland, H. (2007). Altered expression of PGK1in a family with phosphoglycerate kinase deficiency. Muscle Nerve *36*, 679–684.

Tabuchi, Y., Takasaki, I., Doi, T., Ishii, Y., Sakai, H., and Kondo, T. (2006). Genetic networks responsive to sodium butyrate in colonic epithelial cells. FEBS Lett. *580*, 3035–3041.

Tan, N.Y., and Khachigian, L.M. (2009). Sp1 Phosphorylation and Its Regulation of Gene Transcription. Molecular and Cellular Biology *29*, 2483–2488.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nature Structural & Molecular Biology *14*, 1025–1040.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *488*, 75–82.

Tsai, K.-J., Hung, I.-J., Chow, C.K., Stern, A., Chao, S.S., and Chiu, D.T.-Y. (1998). Impaired production of nitric oxide, superoxide, and hydrogen peroxide in glucose 6-phosphate-dehydrogenase-deficient granulocytes. FEBS Lett. *436,* 411–414.

Ursini, M.V., Scalera, L., and Martini, G. (1990). High levels of transcription driven by a 400 bp segment of the human G6PD promoter. Biochem. Biophys. Res. Commun. *170*, 1203–1209.

Valayannopoulos, V., Verhoeven, N.M., Mention, K., Salomons, G.S., Sommelet, D., Gonzales, M., Touati, G., de Lonlay, P., Jakobs, C., and Saudubray, J.-M. (2006). Transaldolase deficiency: A new cause of hydrops fetalis and neonatal multi-organ disease. The Journal of Pediatrics *149*, 713–717.

Van Lint, C., Emiliani, S., and Verdin, E. (1996). The expression of a small fraction of cellular genes is changed in response to histone hyperacetylation. Gene. *5,* 245-256.

Van Wijk, R. (2005). The energy-less red blood cell is lost: erythrocyte enzyme abnormalities of glycolysis. Blood *106*, 4034–4042.

Varki, A., Cummings, R., Esko, J., Freeze, H., Stanley, P., Bertozzi, C., Hart, G., and Etzler, M. (2009). Essentials of Glycobiology. Cold Spring Harbor Laboratory Press.

Verdel, A., Curtet, S., Brocard, M.-P., Rousseaux, S., Lemercier, C., Yoshida, M., and Khochbin, S. (2000). Active maintenance of mHDA2/mHDAC6 histone-deacetylase in the cytoplasm. Current Biology *10*, 747–749.

Verdin, E., Dequiedt, F., and Kasler, H.G. (2003). Class II histone deacetylases: versatile regulators. Trends in Genetics *19*, 286–293.

Verhoeven, N., Huck, J., Roos, B., Struys, E., Salomons, G., Douwes, A., van der Knaap, M., and Jakobs, C. (2001). Transaldolase Deficiency: Liver Cirrhosis Associated with a New Inborn Error in the Pentose Phosphate Pathway. Am J Hum Genet *68,* 1086–1092.

Vigushin, D., Ali, S., Pace, P., Mirsaidi, N., Ito, K., Adcock, I., and Coombes, C. (2001). Trichostatin A Is a Histone Deacetylase Inhibitor with Potent Antitumor Activity against Breast Cancer in vivo. Clinical Cancer Research *7,* 971–977.

Vulliamy, T.J., D'Urso, M., Battistuzzi, G., Estrada, M., Foulkes, N.S., Martini, G., Calabro, V., Poggi, V., Giordano, R., Town, M., et al. (1988). Diverse point mutations in the human glucose-6-phosphate dehydrogenase gene cause enzyme deficiency and mild or severe hemolytic anemia. Pnas *85*, 5171–5175.

Vulliamy, T.J., Kaeda, J.S., Ait-Chafa, D., Mangerini, R., Roper, D., Barbot, J., Mehta, A.B., Athanassiou-Metaxa, M., Luzzatto, L., and Mason, P.J. (1998). Clinical and haematological consequences of recurrent G6PD mutations and a single new mutation causing chronic nonspherocytic haemolytic anaemia. Br. J. Haematol. *101*, 670–675.

Wagner, J.M., Hackanson, B., Lübbert, M., and Jung, M. (2010). Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. Clinical Epigenetics *1*, 117–136.

Wamelink, M.M.C., Grüning, N.-M., Jansen, E.E.W., Bluemlein, K., Lehrach, H., Jakobs, C., and Ralser, M. (2010). The difference between rare and exceptionally rare: molecular characterization of ribose 5-phosphate isomerase deficiency. J Mol Med *88*, 931–939.

Wang, A. (2002). Requirement of Hos2 Histone Deacetylase for Gene Activity in Yeast. Science *298*, 1412–1414.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009). Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. Cell *138*, 1019–1031.

Weake, V.M., and Workman, J.L. (2008). Histone Ubiquitination: Triggering Gene Activity. Mol. Cell *29*, 653–663.

Wiech, N.L., Fisher, J.F., Helquist, P., and Wiest, O. (2009). Inhibition of histone deacetylases: a pharmacological approach to the treatment of non-cancer disorders. Current Topics in Medical Chemistry. *9,* 257-271.

Wierstra, I. (2008). Sp1: Emerging roles—Beyond constitutive activation of TATA-less housekeeping genes. Biochem. Biophys. Res. Commun. *372*, 1–13.

Wong, P., Hattangadi, S.M., Cheng, A.W., Frampton, G.M., Young, R.A., and Lodish, H.F. (2011). Gene induction and repression during terminal erythropoiesis are mediated by distinct epigenetic changes. Blood *118*, 128–138.

working group, W. (1989). Glucose-6-phosphate dehydrogenase deficiency. Bulletin of the World Health Organisation 1–11.

Wu, H., Liu, X., Jaenisch, R., and Lodish, H. (1995). Generation of Committed ErythroidBFU-E and CFU-E Progenitors Does Not Require Etythropoietin or the Etythropoietin Receptor. Cell *83,* 59–67.

Xu, L., Lavinsky, R., Dasen, J., Flynn, S., McInerney, E., Tina-Marie, M., Heinzel, T., Szeto, D., Korzus, E., Kurokawa, R., et al. (1998). Signal-specific co-activatordomain requirements for Pit-1 activation. Nature *395,* 301–306.

Xu, M., and Cook, P.R. (2008). Similar active genes cluster in specialized transcription factories. The Journal of Cell Biology *181*, 615–623.

Yamamura, K., Ohishi, K., Katayama, N., Yu, Z., Kato, K., Masuya, M., Fujieda, A., Sugimoto, Y., Miyata, E., Shibasaki, T., et al. (2006). Pleiotropic role of histone deacetylases in the regulation of human adult erythropoiesis. Br. J. Haematol. *135*, 242–253.

Yang, P.-M., Lin, P.-J., and Chen, C.-C. (2012). CD1d induction in solid tumor cells by histone deacetylase inhibitors through inhibition of HDAC1/2 and activation of Sp1. Epigenetics *7*, 390–399.

Yang, X.-J., and Gregoire, S. (2005). Class II Histone Deacetylases: from Sequence to Function, Regulation, and Clinical Implication. Molecular and Cellular Biology *25*, 2873–2884.

Yang, X.-J., and Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. Nat Rev Mol Cell Biol *9*, 206–218.

Yao, D.C. (2004). Hemolytic anemia and severe rhabdomyolysis caused by compound heterozygous mutations of the gene for erythrocyte/muscle isozyme of aldolase, ALDOA(Arg303X/Cys338Tyr). Blood *103*, 2401–2403.

Yoo, C.B., and Jones, P.A. (2006). Epigenetic therapy of cancer: past, present and future. Nat Rev Drug Discov *5*, 37–50.

Yoshida, M., Kijima, M., Akita, M., and Beppu, T. (1990). Potent and Specific Inhibition of Mammalian Histone Deacetylase Both. J Biol Chem 1–6.

Yu, Y., Black, J., Goldsmith, C., Browning, P., Bhalla, K., and Offermann, M. (1998). Induction of human herpesvirus-8 DNA replication and transcription by butyrate and TPA in BCBL-1 cells. Journal of General Virology *80,* 83–90.

Zanella, A., Bianchi, P., and Fermo, E. (2007a). Pyruvate kinase deficiency. Haematologica *21,* 217–231.

Zanella, A., Fermo, E., Bianchi, P., Chiarelli, L.R., and Valentini, G. (2007b). Pyruvate kinase deficiency: The genotype-phenotype association. Blood Reviews *21*, 217–231.

Zhang, K., and Dent, S.Y.R. (2005). Histone modifying enzymes and cancer: Going beyond histones. J. Cell. Biochem. *96*, 1137–1148.

Zhang, Y., Liao, M., and Dufau, M.L. (2006). Phosphatidylinositol 3-Kinase/Protein Kinase C -Induced Phosphorylation of Sp1 and p107 Repressor Release Have a Critical Role in Histone Deacetylase Inhibitor-Mediated Depression of Transcription of the Luteinizing Hormone Receptor Gene. Molecular and Cellular Biology *26*, 6748–6761.

Zhao, L., Chen, C.N., Hajji, N., Oliver, E., Cotroneo, E., Wharton, J., Wang, D., Li, M., McKinsey, T.A., Stenmark, K.R., et al. (2012). Histone Deacetylation Inhibition in Pulmonary Hypertension: Therapeutic Potential of Valproic Acid and Suberoylanilide Hydroxamic Acid. Circulation *126*, 455–467.

Zhu, B., and Reinberg, D. (2011). Epigenetic inheritance: Uncontested? Cell Res. *21*, 435–441.

Zimmerman, M.A., Singh, N., Martin, P.M., Thangaraju, M., Ganapathy, V., Waller, J.L., Shi, H., Robertson, K.D., Munn, D.H., and Liu, K. (2012). Butyrate suppresses colonic inflammation through HDAC1-dependent Fas upregulation and Fas-mediated apoptosis of T cells. AJP: Gastrointestinal and Liver Physiology *302*, G1405–G1415.

Zini, R., Norfo, R., Ferrari, F., Bianchi, E., Salati, S., Pennucci, V., Sacchi, G., Carboni, C., Ceccherelli, G.B., Tagliafico, E., et al. (2012). Valproic acid triggers erythro/megakaryocyte lineage decision through induction of GFI1B and MLLT3 expression. Experimental Hematology *40,* 1043–1054.

9 Appendix

Appendix A. List of significantly upregulated genes upon NaBu in CB-CD34$^+$ - differentiating erythroblasts.

Gene	Fold Change	P value
HIST1H1T	5.95	0.01
DENND2C	5.90	0.00
SERPINI1	5.83	0.01
ENPP5	5.47	0.00
MAPRE3	5.20	0.01
PFN2	4.71	0.01
TMOD2	4.33	4.73E-04
KIF5A	3.97	0.02
ROPN1L	3.82	0.00
PLEKHH2	3.69	0.00
ETNK2	3.61	0.01
MYBL1	3.49	0.01
KIF5A	3.49	0.03
TUBB4A	3.47	0.00
SERPINB9	3.46	6.72E-04
PEG10	3.37	2.98E-04
KIF5C	3.33	0.00
GBP5	3.20	0.00
CCDC68	3.08	0.00
RBM11	3.02	3.14E-04
IRF6	3.02	0.03
TCP11L2	2.98	1.78E-04
BMP6	2.98	0.01
OCLN	2.89	0.01
MAP1B	2.87	0.00
FAM49A	2.84	0.00
HSPA2	2.80	4.50E-05
EFNA4	2.78	4.06E-04
ZNF204P	2.68	0.00
FBXO16	2.66	0.01
SESN3	2.64	0.02
TUFT1	2.61	0.02
MERTK	2.60	0.01
SLC7A8	2.58	0.00
STXBP1	2.55	0.01
STAT4	2.55	0.00
KIAA1161	2.54	5.34E-04
SV2A	2.54	0.05
PTCH1	2.53	0.05
H1FX	2.52	7.53E-04
C5orf4	2.51	0.00
ABCB9	2.51	0.00
IER3	2.49	0.00
AHNAK2	2.47	0.02
HIST1H1A	2.44	0.00

HDAC11	2.44	0.00
HIST2H4A // HIST2H4B	2.42	0.04
HIST2H4A // HIST2H4B	2.42	0.04
RAB3A	2.42	0.00
LPPR2	2.40	0.01
TSPAN13	2.40	0.05
VCL	2.40	0.00
ENDOD1	2.40	0.03
RAB30	2.40	6.10E-04
TESK2	2.38	0.00
IER3	2.37	8.99E-04
IER3	2.37	8.99E-04
GALM	2.37	0.01
CLDN10	2.36	0.01
ABCB1	2.34	0.01
FAM171A2	2.33	0.01
ST3GAL5	2.32	0.01
SCAMP5	2.31	0.00
PBX1	2.27	0.00
CCR7	2.26	7.42E-04
C17orf104	2.26	0.01
TPST1	2.25	0.04
DPF1	2.25	1.10E-04
TSPAN15	2.24	0.02
ZC2HC1A	2.24	0.02
SORT1	2.23	0.02
PPM1J	2.22	6.17E-04
FAM69A	2.22	0.00
KIF3A	2.21	0.01
SPATA6	2.20	0.01
UBE2H	2.19	9.89E-05
ETS1	2.18	0.03
ALDH1A1	2.18	0.01
REEP6	2.17	0.04
DHRS1	2.16	1.36E-04
TNFAIP6	2.16	0.05
SMARCA1	2.16	0.02
MOSPD1	2.16	0.03
AKTIP	2.16	0.02
DOK4	2.15	3.27E-04
ABHD4	2.15	0.00
SCD5	2.15	0.01
SMOX	2.14	0.00
HPSE	2.13	0.00
ETV5	2.13	8.21E-04
CDKL5	2.13	0.00
TMEM63C	2.12	0.03

ZNF774	2.12	0.01
CXADR	2.11	0.05
KHK	2.11	0.01
GABARAPL1	2.11	4.25E-05
WDR31	2.11	0.03
CAPN5	2.10	0.01
C16orf45	2.09	4.15E-04
EXPH5	2.09	0.03
ENPP2	2.08	0.01
C11orf67	2.08	0.01
DNAJC28	2.08	0.01
P2RX7	2.08	0.00
CPEB4	2.08	6.21E-04
ASMTL	2.08	0.00
IFT80	2.06	0.00
SORBS1	2.06	0.04
MAST3	2.05	0.00
FGD6	2.05	0.00
ARHGEF26	2.05	0.00
NBEAL1	2.05	3.82E-04
TBC1D2	2.04	0.01
RECK	2.04	0.00
GLCE	2.03	0.01
TSPAN2	2.03	0.03
PLXNA3	2.03	5.54E-04
AGPHD1	2.03	0.02
ASPHD2	2.02	0.01
C1orf116	2.02	3.73E-04
NBEAL1	2.01	0.01
EFHC1	2.01	0.02
ATL1	2.00	0.00
BIRC3	2.00	0.00
SGTB	2.00	0.00
RND2	2.00	0.00
BACE1	2.00	1.13E-04
ATP6V1G2	1.99	0.01
ATP6V1G2	1.99	0.01
CREB3	1.99	6.10E-06
TIPARP	1.99	0.02
RBMS2	1.99	0.01
SPIRE1	1.98	0.01
KIAA0513	1.98	0.00
KIAA1324L	1.98	0.01
NCOA3	1.98	0.00
DNM3	1.98	0.00
JUP	1.98	0.02
ATP1A3	1.97	0.01
GUCY1B3	1.97	4.77E-04
ARL4D	1.97	0.00

OSBPL6	1.96	0.00
PRKACA	1.96	5.83E-06
LRCH2	1.96	0.00
LPHN1	1.95	0.03
SH3PXD2B	1.95	0.01
STARD9	1.95	0.00
CD22	1.94	0.00
FBXO48	1.94	0.04
CTTNBP2NL	1.94	6.36E-04
PLA2G15	1.93	0.01
RCOR2	1.93	0.01
SLC25A37	1.93	0.03
PCYOX1	1.93	9.17E-04
WBP5	1.92	0.00
RRAS	1.92	0.01
GNA11	1.92	0.01
MARK4	1.92	7.43E-04
KRTAP6-1	1.92	0.01
TRIP10	1.92	0.00
GNG12	1.92	0.01
CDC42EP3	1.92	0.00
RASGEF1B	1.92	0.01
DZIP3	1.91	0.01
PGM2L1	1.91	0.02
PAIP2B	1.91	0.01
MT1G	1.91	0.05
AGPAT4	1.90	6.73E-04
POLB	1.90	0.01
CYB5R1	1.90	0.00
RHPN2	1.90	0.01
SEMA4D	1.90	0.02
PCGF2	1.90	0.00
NXN	1.89	0.02
CEP19	1.89	0.01
EMP2	1.89	0.01
ADD3	1.89	0.01
TTLL7	1.88	0.02
HIST3H2BB	1.88	0.01
RAP1GAP	1.88	0.01
RAB3D	1.88	0.02
FAM102A	1.87	0.00
SEMA4G	1.87	5.17E-05
KIF3C	1.87	0.01
GATSL3	1.87	8.50E-04
CDC42BPB	1.87	0.01
IQCK	1.87	0.01
MAMLD1	1.86	0.01
EPB41L5	1.86	9.31E-04
TJP2	1.86	0.01

PITPNM1	1.86	0.02
CTNS	1.86	0.01
IL6ST	1.86	0.01
FAM89B	1.86	0.01
ABCD1	1.85	0.00
BMPR1A	1.85	0.03
FGFR1	1.85	2.22E-04
HERC1	1.84	0.00
FAM126B	1.84	0.01
IFT57	1.84	0.00
ANKRD6	1.83	0.00
CHRM4	1.83	1.35E-04
TCTEX1D1	1.83	0.02
RBKS	1.83	0.05
PFKM	1.83	5.39E-04
YPEL2	1.83	0.00
FMNL2	1.82	0.00
SIAE	1.82	0.05
ITGA6	1.82	0.03
ARL3	1.82	0.01
RFX2	1.82	3.70E-04
NFKB2	1.82	6.01E-04
CASP10	1.82	0.00
FAM190B	1.82	0.00
MAST1	1.81	0.03
FAM171B	1.81	0.00
TPMT	1.81	0.01
CACNB3	1.80	0.03
ARRDC4	1.80	0.03
PPP1R21	1.80	0.01
SCML1	1.80	0.01
RTN4RL2	1.80	0.00
SLC29A4	1.80	0.02
CAMSAP2	1.80	0.00
SCPEP1	1.80	0.01
SEMA4D	1.79	0.00
SMAD3	1.79	0.01
NKIRAS1	1.79	0.00
TSC22D3	1.79	0.02
RNF122	1.78	0.02
SRGAP2	1.78	0.01
KIAA0895	1.78	0.03
DYRK3	1.78	0.03
MT1P1	1.78	0.02
DNAJB5	1.78	0.00
SALL2	1.77	5.40E-05
GPR160	1.77	0.01
SBF2	1.77	0.01
SAT1	1.77	0.01

MCTP1	1.77	0.03
PKP2	1.77	0.01
DNMT3B	1.77	0.00
F8	1.77	0.00
ACSF2	1.76	0.01
RWDD2A	1.76	0.00
MT1E	1.76	0.02
GOLGB1	1.76	0.00
TEAD3	1.75	3.51E-04
CYTH3	1.75	0.05
CYTH1	1.75	3.90E-04
STK17B	1.75	0.01
CCDC92	1.75	6.42E-04
SWAP70	1.75	0.00
PRKD2	1.75	8.25E-04
KDELC1	1.75	0.01
TTLL1	1.75	0.01
LGR4	1.75	0.03
EXTL2	1.75	0.02
MKNK1	1.75	0.00
ACOX1	1.75	0.00
ITSN1	1.74	0.01
PTPN13	1.74	0.01
MAP3K12	1.74	5.44E-04
ACSL1	1.74	0.01
TTYH2	1.74	0.00
SERPINE1	1.74	0.01
ANG	1.73	4.02E-04
CYP2U1	1.73	0.01
GRK5	1.73	0.05
MKNK2	1.73	5.97E-05
ATP8B1	1.73	0.02
CEP70	1.73	5.04E-04
RHOC	1.73	0.01
C9orf89	1.73	1.69E-05
KCTD21	1.73	0.02
FOXO4	1.73	1.15E-04
TTC26	1.73	0.03
IFT81	1.72	0.01
PGAP1	1.72	0.01
AFMID	1.72	0.00
IGSF3	1.72	0.00
COQ10A	1.72	0.01
TXK	1.72	0.03
ATP8B2	1.72	0.02
ZNF483	1.72	0.00
FAM43A	1.71	0.01
RPP25	1.71	0.01
NMT2	1.71	0.00

NEU1	1.71	0.02
NEU1	1.71	0.02
GATSL2	1.71	0.04
GATSL2	1.71	0.04
C7orf23	1.71	0.00
RCN3	1.71	0.01
TOM1L2	1.71	0.03
AK1	1.71	0.03
FAM220A	1.71	0.01
TMEM241	1.71	0.00
NEU1	1.70	0.02
FAM117A	1.70	0.01
TMEFF1	1.70	0.03
XRRA1	1.70	0.00
TNFRSF11A	1.70	0.00
GMCL1	1.70	6.62E-04
ATP6V1G2	1.70	0.01
APLF	1.70	0.03
POLN	1.70	0.01
NEK3	1.70	0.03
MGAT5	1.70	0.04
BMPR2	1.69	0.01
SDE2	1.69	0.02
WASF3	1.69	0.01
ARID3B	1.69	0.03
DNAJC18	1.69	0.03
BACH2	1.68	0.05
TRERF1	1.68	0.04
MITF	1.68	0.01
BMF	1.68	0.00
SLC45A4	1.68	0.01
CBX7	1.68	0.03
DYNC2H1	1.68	0.03
UBTD2	1.68	0.03
ADHFE1	1.68	0.05
AGTPBP1	1.68	3.42E-04
CHRNA5	1.67	0.01
MFSD12	1.67	0.01
SLC25A30	1.67	0.00
TMEM198B	1.67	0.02
KIAA0040	1.67	0.01
BBS7	1.67	0.01
INPP5K	1.67	0.00
PNPLA8	1.67	3.07E-04
KLHL24	1.67	0.00
MLLT4	1.67	0.01
ZSCAN16	1.67	0.03
REEP2	1.67	4.09E-04
KSR1	1.67	0.02

USP2	1.66	0.01
TMEM205	1.66	0.00
USP44	1.66	0.02
ENAH	1.66	0.01
FAM220BP	1.66	0.01
PLCG1	1.66	0.02
HIBCH	1.66	0.01
HABP4	1.66	0.00
ARG2	1.66	0.03
IL1A	1.66	0.04
KCND1	1.66	0.01
CYP2R1	1.66	0.01
TC2N	1.65	0.03
BAMBI	1.65	0.00
DDAH2	1.65	0.02
DDAH2	1.65	0.02
C15orf39	1.65	0.02
C14orf37	1.65	0.00
ATG2A	1.65	0.02
TMEM169	1.64	0.02
TAC3	1.64	0.00
RHOF	1.64	0.03
ARHGAP18	1.64	0.00
KIF1B	1.64	9.73E-04
KATNB1	1.64	8.93E-04
ST6GALNAC4	1.64	0.00
MEGF9	1.63	0.00
ATP6V1D	1.63	0.03
CCDC113	1.63	0.03
DMXL1	1.63	0.00
FAM122C	1.63	0.02
C17orf65	1.63	0.01
CBL	1.63	0.05
DYNLT3	1.63	1.84E-04
PLCB3	1.63	4.58E-04
KCNN1	1.63	0.01
MPZL1	1.63	0.01
ATP8A1	1.62	0.01
PPP2R5B	1.62	0.01
TMEM180	1.62	0.00
MAP3K3	1.62	0.01
TMEM25	1.62	0.03
CMTM8	1.62	0.01
ZNF555	1.62	0.04
FGFR3	1.61	0.02
FZD5	1.61	0.02
PSMB9	1.61	0.02
PSMB9	1.61	0.02
PSMB9	1.61	0.02

NIT1	1.61	2.94E-04
PLS1	1.61	0.00
IFT88	1.61	0.02
KBTBD8	1.61	0.03
NRGN	1.61	0.01
UBE2S	1.61	0.00
MAP2K3	1.61	0.03
TSNAX	1.61	0.02
SYP	1.61	0.01
LLGL2	1.61	0.00
PKIA	1.61	0.00
DHTKD1	1.61	0.01
SAMD15	1.61	0.03
BCOR	1.60	0.01
RAB3IP	1.60	0.01
TCEA3	1.60	0.04
C14orf129	1.60	0.02
SLC45A4	1.60	0.03
PI4K2A	1.60	0.01
PRRG1	1.60	0.00
WWTR1	1.60	0.05
SLC2A4	1.60	0.03
ANO10	1.60	0.01
RABL5	1.60	0.05
PPP2R3A	1.60	0.02
PORCN	1.60	0.05
PARD6B	1.59	0.03
PELI1	1.59	0.03
AAK1	1.59	0.01
CASZ1	1.59	0.01
PPARD	1.59	0.00
SNAI1	1.59	0.01
STAT2	1.59	0.01
IGF2R	1.59	0.01
SYAP1	1.59	0.01
MYO10	1.59	0.02
CYP2S1	1.59	0.00
CNNM4	1.59	0.00
MICA	1.59	0.01
RTKN2	1.59	0.01
SLC4A11	1.59	0.01
PLCD1	1.59	0.01
CTSL2	1.59	0.02
WDR26	1.59	4.45E-04
EPG5	1.59	0.02
DDAH2	1.59	0.02
AHI1	1.58	0.00
G6PD	1.58	0.02
RCBTB1	1.58	0.01

ZNF396	1.58	0.03
TMEM232	1.58	0.05
PGBD1	1.58	0.02
MED20	1.58	0.00
PPP1R13B	1.58	0.02
STIM1	1.58	2.70E-06
C4orf34	1.58	0.01
ECE1	1.58	0.01
WDR47	1.58	0.00
CABLES2	1.58	0.00
WASF1	1.58	0.02
CD86	1.58	0.04
PLEKHM1	1.58	0.01
PRAF2	1.57	0.00
NR6A1	1.57	0.01
AJUBA	1.57	0.02
RPS6KA2	1.57	0.01
RYBP	1.57	0.01
ANXA4	1.57	0.01
HEXIM1	1.57	0.00
ERBB3	1.57	0.03
IRAK2	1.57	5.19E-04
ZNF117	1.57	0.01
KDM8	1.57	0.01
MAGI3	1.57	0.05
PLAT	1.57	0.05
HDHD3	1.57	0.02
RAP1GAP2	1.57	0.02
BCORL1	1.57	0.00
RABGAP1	1.57	0.01
RALGAPA2	1.57	3.90E-04
TNNI3	1.57	9.02E-04
DZIP1	1.57	0.00
FAM107B	1.57	0.04
MT2A	1.57	0.02
HOXA1	1.56	0.01
TRIM26	1.56	0.00
MT2A	1.56	0.02
CRYZ	1.56	0.00
FKBP7	1.56	0.02
KLHL5	1.56	0.01
ALS2	1.56	0.04
EPB41L2	1.56	0.03
APLP1	1.56	0.00
BCL2L11	1.56	0.00
MICB	1.56	0.00
NR4A1	1.56	0.02
MAP4K2	1.56	0.01
HOOK2	1.56	0.03

CGRRF1	1.56	0.02
NEK1	1.56	0.01
EHD3	1.56	0.05
CELSR2	1.56	0.01
ATP6V1A	1.56	2.78E-04
RASGEF1A	1.55	0.02
SNAI3	1.55	0.00
STX5	1.55	0.00
UBE2S	1.55	0.00
SEC61A2	1.55	0.01
TRGV3	1.55	0.00
KBTBD3	1.55	0.01
GRHL1	1.55	0.01
CLEC16A	1.55	2.36E-04
TRIM62	1.55	0.00
CNPY4	1.55	0.01
ASAP2	1.55	0.02
ARL13B	1.55	0.01
APBB2	1.55	0.05
TCEA2	1.55	0.01
TPT1-AS1	1.55	0.00
TMEM184B	1.55	0.00
TMCC3	1.55	0.00
ZNF217	1.55	0.00
C5orf45	1.55	0.04
PRRT3	1.54	0.01
SLC9A9	1.54	0.03
SP4	1.54	2.53E-04
GSTA4	1.54	0.00
SOAT1	1.54	8.95E-05
MYLIP	1.54	0.00
GNAZ	1.54	0.00
AVPI1	1.54	0.01
ICK	1.54	0.01
MDM1	1.54	0.03
LIMA1	1.54	0.02
ZER1	1.54	0.03
PRCP	1.54	0.01
PAQR3	1.54	6.73E-04
GTF2IRD1	1.54	7.53E-04
LMBRD2	1.54	0.01
GCC2	1.54	0.00
FAM108A3P	1.54	8.67E-05
FAM108A3P	1.54	8.67E-05
PFKFB2	1.54	0.01
LYST	1.54	0.05
INPP5F	1.54	0.01
P4HA2	1.54	0.01
PQLC1	1.53	0.01

ENTHD1	1.53	0.02
MT1X	1.53	0.05
PITPNC1	1.53	7.63E-04
KANK2	1.53	0.03
DCUN1D3	1.53	0.04
CRIP2	1.53	0.01
LRFN3	1.53	0.02
SNAP29	1.53	0.01
FHL1	1.53	0.05
HES1	1.53	0.00
ZDHHC8	1.53	0.01
RGS9	1.53	0.00
MAP2	1.53	0.02
GGCX	1.53	8.99E-05
ANKRD13A	1.53	0.00
U2AF1L4	1.53	0.03
GNAI1	1.53	0.01
GNG7	1.53	0.02
USP28	1.53	0.01
MLF1	1.53	0.01
WIPI1	1.53	4.57E-04
TRIM26	1.53	0.00
STK38L	1.53	0.01
SEZ6L2	1.52	0.00
MPRIP	1.52	0.00
LPCAT3	1.52	0.01
MAP3K9	1.52	0.00
HSD17B6	1.52	0.05
PRKAR2B	1.52	0.03
MICB	1.52	0.00
BET1L	1.52	0.00
CPNE3	1.52	0.01
BST2	1.52	1.74E-04
FADS3	1.52	0.01
ARMC9	1.52	0.01
FYN	1.52	0.02
ZNF608	1.52	0.02
ZNF277	1.52	0.01
GLS	1.52	0.02
PIGX	1.52	3.92E-04
PLXNA2	1.52	0.01
ATXN1	1.51	0.02
KIAA1467	1.51	0.00
ARL6IP6	1.51	0.01
TOR1AIP2	1.51	0.00
AMOT	1.51	0.01
TBC1D7	1.51	0.05
TNRC6B	1.51	9.33E-04
CAMLG	1.51	0.01

SLFN14	1.51	0.04
NRARP	1.51	0.01
SSBP3	1.51	6.14E-04
GDI1	1.51	0.02
SUFU	1.51	0.00
SLC25A23	1.51	0.01
CALCOCO2	1.51	0.00
NFAT5	1.51	0.04
ARL6	1.51	0.02
AHR	1.51	0.04
FNBP1L	1.50	0.01
TUBB2B	1.50	0.02
EFNB2	1.50	0.01
PPM1N	1.50	0.03
FRRS1	1.50	0.01
ATP6V0A1	1.50	0.00
CALCOCO1	1.50	0.01
LRRC1	1.50	0.00
SNRNP48	1.50	2.89E-04
C12orf51	1.50	0.01
FAM108A4P	1.50	3.60E-05
SLC27A1	1.50	0.01
ITPR2	1.50	0.02
RHOQP2	1.50	0.01
ACER3	1.50	0.01

Appendix B. List of significantly downregulated genes upon NaBu in CB-CD34$^+$ -differentiating erythroblasts.

Genes	Fold Change	P value
LANCL2	3.82	0.00
PLA2G7	3.36	0.03
ANGPT1	3.35	0.03
IGSF6	3.32	0.01
SLC7A11	3.32	0.01
KMO	3.15	0.05
UTP20	3.11	0.01
IL7R	3.05	0.04
C3orf14	3.03	0.04
NAIP	3.01	0.04
NAIP	2.99	0.03
KCNQ5	2.97	0.00
FGL2	2.92	0.04
GPR141	2.90	0.01
AVEN	2.84	0.00
SNORD77	2.83	0.03
RCN1	2.80	0.04
CD84	2.79	0.03
NIPAL2	2.75	0.01
FPR3	2.75	0.04
ZMYND11	2.75	0.01
GTF2IRD2B	2.74	0.00
MS4A6A	2.74	0.00
SLAMF8	2.73	0.01
CARD6	2.71	0.02
PHF19	2.65	0.00
MS4A2	2.65	0.01
CLEC5A	2.61	0.05
C15orf41	2.57	0.00
LOC100128816	2.56	0.00
ZBTB2	2.56	0.02
GTF2IRD2B	2.55	0.00
KDELC2	2.49	0.00
CD1E	2.47	0.02
PHF15	2.46	0.00
GPR171	2.42	0.02
CXorf26	2.42	0.02
DIS3L	2.42	0.00
TMEM117	2.40	0.02
PARM1	2.40	0.02
ZFP64	2.38	0.00
NRP1	2.37	0.01
IL12RB2	2.36	0.02
LOC339524 // HS2ST1	2.35	0.02

EPC1	2.34	0.01
USP13	2.33	0.00
FAM57A	2.33	0.04
SNORD34	2.33	0.03
TMCO7	2.32	0.00
PHF17	2.32	0.01
METAP1D	2.30	0.02
GINS3	2.29	0.02
STAG1	2.29	0.00
MTHFS	2.27	0.03
PTPN7	2.27	0.05
MRTO4	2.27	0.00
STS	2.26	0.00
FDXACB1	2.25	0.01
CTPS1	2.25	0.00
MRM1	2.25	0.02
NETO2	2.24	0.04
CLEC4A	2.24	0.01
FAIM	2.23	0.03
ARSB	2.21	0.01
GALNT14	2.21	0.03
UBXN8	2.19	0.00
GTF2I	2.18	0.01
IL27RA	2.18	0.02
CABLES1	2.18	0.05
BEND3	2.17	0.00
HPDL	2.17	0.00
NMI	2.16	0.00
TFAP4	2.16	0.01
ARMC6	2.16	0.00
SLTM	2.15	0.02
CASD1	2.15	0.04
ADAMTS3	2.15	0.00
ADAT2	2.13	0.01
GPR34	2.13	0.01
CD1A	2.13	0.04
NSMAF	2.12	0.02
LRMP	2.12	0.01
EPC2	2.11	0.01
BDH1	2.10	0.01
AMICA1	2.10	0.04
STAT5A	2.10	0.00
HCFC1	2.09	0.01
SMYD5	2.08	0.01
ZBTB1	2.08	0.04
ALG14	2.08	0.01
UBTF	2.07	0.00
FAM136A	2.07	0.05
ADI1	2.06	0.01

FAS	2.05	0.02
SNTB1	2.04	0.00
ZNF804A	2.04	0.04
CCL1	2.04	0.04
PRDM10	2.04	0.01
TOE1	2.03	0.02
PHF2	2.03	0.00
NOP16	2.03	0.02
NFKB1	2.03	0.00
MCM9	2.03	0.01
ING5	2.02	0.00
PATZ1	2.02	0.01
CCR6	2.02	0.01
GPATCH4	2.01	0.01
HERC6	2.01	0.02
CYP7B1	2.01	0.03
CD1C	2.00	0.05
ALDH5A1	2.00	0.03
RPP40	1.99	0.00
ING5	1.98	0.00
GPR125	1.98	0.01
KCTD15	1.98	0.00
MTHFD1L	1.97	0.00
ANAPC13	1.97	0.03
TMEM87A	1.97	0.01
DUS3L	1.96	0.00
CSF1	1.95	0.03
FAM217B	1.95	0.01
GRAP2	1.94	0.04
GGTA1P	1.94	0.01
PDCL3	1.94	0.01
NBN	1.94	0.02
U2AF2	1.93	0.02
PPRC1	1.93	0.02
MSL1	1.93	0.02
CD93	1.93	0.02
FERMT3	1.92	0.03
ATF5	1.91	0.01
SNORA24	1.91	0.01
KIAA0368	1.91	0.00
PON2	1.90	0.01
NAGPA	1.90	0.00
C14orf102	1.90	0.00
BANK1	1.90	0.00
C10orf57	1.89	0.00
INO80D	1.89	0.01
EOGT	1.89	0.02
C1QTNF9B-AS1	1.89	0.03
VPS52	1.88	0.01

VPS52	1.88	0.01
NLRC4	1.88	0.01
PRKCQ	1.87	0.01
ZNF366	1.87	0.01
ACPP	1.87	0.03
SNORA4	1.87	0.02
RNASE6	1.87	0.05
WT1	1.86	0.01
NOL6	1.86	0.00
VPS52	1.86	0.01
P2RY14	1.86	0.02
SLC39A10	1.85	0.02
ANKLE1	1.85	0.02
SUSD1	1.85	0.02
METTL8	1.84	0.00
PUS7	1.84	0.02
NCOA2	1.84	0.00
EFCAB4B	1.84	0.01
LIF	1.84	0.04
MINA	1.83	0.02
FJX1	1.83	0.01
IL9R	1.83	0.01
IL9R	1.83	0.01
FMNL1	1.83	0.00
TYW3	1.83	0.01
CLEC10A	1.83	0.04
PMF1	1.83	0.00
SLC38A5	1.83	0.02
C8orf33	1.83	0.01
PRKAR1B	1.83	0.01
SNAPC5	1.82	0.05
TMEM192	1.82	0.00
CAB39L	1.82	0.01
WARS2	1.81	0.02
STAMBPL1	1.81	0.01
RRS1	1.81	0.01
LARP7	1.81	0.00
GCFC2	1.81	0.02
RABL3	1.81	0.01
KEAP1	1.81	0.00
SPR	1.81	0.00
DTX3L	1.81	0.00
FABP5	1.80	0.02
PEF1	1.80	0.01
ZNF114	1.80	0.05
FBXO4	1.80	0.02
SPRYD4	1.79	0.01
CCDC94	1.79	0.03
DDX26B	1.79	0.00

SPAG7	1.79	0.00
KAT5	1.79	0.01
ZFAT	1.79	0.01
FST	1.79	0.01
CWF19L2	1.79	0.01
S1PR4	1.78	0.01
SPATS2L	1.78	0.05
LRP8	1.78	0.01
SELRC1	1.78	0.03
TSR1	1.78	0.00
FABP5	1.78	0.02
ITFG2	1.78	0.01
TMPO	1.78	0.00
PRMT6	1.77	0.01
SLC5A6	1.77	0.01
N6AMT2	1.77	0.01
QRSL1	1.77	0.00
GTPBP6	1.77	0.00
LACE1	1.77	0.04
GTF3C6	1.77	0.00
PPIL1	1.77	0.00
ZBTB38	1.76	0.00
ASTE1	1.76	0.03
SLC25A13	1.76	0.01
SMARCB1	1.76	0.00
C12orf26	1.76	0.00
KAT2A	1.76	0.01
FAM124B	1.76	0.02
KCNH2	1.75	0.01
ZFP161	1.75	0.00
MTX3	1.75	0.03
B3GALTL	1.75	0.01
KAT6B	1.75	0.00
ZC3HC1	1.75	0.01
ACSL5	1.75	0.02
HAVCR2	1.75	0.05
E2F8	1.74	0.00
CAD	1.74	0.00
SMYD2	1.74	0.00
MEPCE	1.74	0.00
GPR125	1.74	0.00
PILRA	1.74	0.02
CUL1	1.73	0.03
APOBR	1.73	0.00
POLE4	1.73	0.01
RPRD2	1.73	0.02
HSD17B7P2	1.73	0.01
EPM2A	1.73	0.01
ZNF827	1.73	0.01

C6orf223	1.72	0.02
SLC6A9	1.72	0.03
TMBIM4	1.72	0.00
ALG13	1.72	0.04
WDFY4	1.72	0.05
TMX4	1.72	0.00
LRRC41	1.72	0.00
C9orf91	1.72	0.04
NSD1	1.72	0.01
LHFPL2	1.71	0.01
USP41	1.71	0.03
ZNF692	1.71	0.00
MANEA	1.71	0.00
PCBD2	1.71	0.00
LMNB1	1.71	0.04
ZNF362	1.71	0.01
MB21D1	1.70	0.05
URGCP	1.70	0.01
ZNF573	1.70	0.01
BLNK	1.70	0.02
CHST10	1.70	0.01
UTP15	1.70	0.02
BRPF3	1.69	0.01
PBLD	1.69	0.03
YDJC	1.69	0.01
SNUPN	1.69	0.05
SLC16A7	1.68	0.02
HUS1	1.68	0.01
RRP12	1.68	0.01
MPDU1	1.68	0.01
SDC2	1.68	0.01
ADRBK1	1.68	0.02
ZAK	1.67	0.05
ZNF33A	1.67	0.00
RN5S242	1.67	0.02
SNORA74A	1.67	0.03
DCPS	1.67	0.04
ZC3HAV1L	1.67	0.03
HSD17B7	1.67	0.01
ALKBH8	1.67	0.00
ARHGAP9	1.67	0.05
GNE	1.67	0.01
STEAP3	1.66	0.01
LOC100131190	1.66	0.02
SMG9	1.66	0.00
MYBBP1A	1.66	0.01
CHRAC1	1.66	0.05
NUP205	1.66	0.00
COQ7	1.66	0.00

FAM216A	1.66	0.03
PTS	1.66	0.02
CHCHD4	1.66	0.02
TIMM21	1.66	0.01
DHX33	1.66	0.00
THRB	1.66	0.03
TNFAIP8L2	1.66	0.00
ZNF383	1.66	0.02
TRMT2B	1.65	0.00
GNA15	1.65	0.00
ZNF407	1.65	0.02
TTC27	1.65	0.01
THOC6	1.65	0.00
UCK2	1.65	0.01
COMMD10	1.65	0.03
SLC39A3	1.65	0.03
NLRX1	1.64	0.00
TYMS	1.64	0.01
TOX	1.64	0.03
SNORD35A	1.64	0.02
MRPS28	1.64	0.01
ARID1B	1.64	0.00
ZNF639	1.64	0.01
ZNF37A	1.63	0.01
CDC7	1.63	0.01
CLK4	1.63	0.04
KIAA1598	1.63	0.02
NUBP1	1.63	0.01
MATK	1.63	0.01
SLC25A26	1.63	0.04
C3orf75	1.63	0.01
CNTF	1.63	0.03
DPH2	1.63	0.01
CNNM1	1.63	0.05
ST7	1.63	0.03
IVNS1ABP	1.63	0.01
IDNK	1.62	0.03
KCNE3	1.62	0.05
PARP8	1.62	0.00
CNTLN	1.62	0.00
IPO8	1.62	0.00
CRISPLD1	1.62	0.01
H2AFY2	1.62	0.02
ADCK1	1.62	0.02
PSMB10	1.62	0.01
VPS36	1.62	0.00
PISD	1.62	0.01
LRFN4	1.62	0.00
LIN9	1.62	0.04

LTV1	1.62	0.01
COIL	1.62	0.00
SLC29A2	1.61	0.03
RAB11FIP2	1.61	0.01
SLC19A1	1.61	0.01
L3MBTL2	1.61	0.00
NUDT15	1.61	0.02
FASTKD2	1.61	0.02
ATP11C	1.61	0.01
CCDC101	1.60	0.02
RRP15	1.60	0.01
SLC25A15	1.60	0.03
CREBBP	1.60	0.00
ADNP	1.60	0.04
GEMIN4	1.60	0.02
KDM8	1.60	0.00
SGK196	1.60	0.03
TMEM168	1.60	0.01
ITPA	1.60	0.02
IFRD2	1.59	0.00
ENTPD1	1.59	0.02
BCL6	1.59	0.02
TRRAP	1.59	0.02
LTB4R2	1.59	0.02
CTCF	1.59	0.00
CUL4A	1.59	0.00
AKT1S1	1.59	0.00
SUV420H1	1.59	0.00
GBE1	1.59	0.00
BAG2	1.58	0.01
GRB10	1.58	0.02
SLC35F2	1.58	0.04
EPDR1	1.58	0.00
TIMELESS	1.58	0.01
TCERG1	1.58	0.02
GIMAP6	1.58	0.00
FAM172A	1.58	0.03
KRT79	1.58	0.00
SIRPA	1.58	0.05
UNG	1.58	0.00
SLC8A1	1.58	0.01
NCF2	1.58	0.04
ARHGAP21	1.58	0.01
LIG3	1.58	0.00
WDFY4	1.58	0.04
WDR74	1.58	0.00
BCLAF1	1.58	0.02
NQO1	1.57	0.03
DNAJC2	1.57	0.03

TXNDC15	1.57	0.00
ITGA4	1.57	0.02
DPP3	1.57	0.01
PTPRCAP	1.57	0.03
SLC7A1	1.57	0.00
RNLS	1.57	0.02
AEN	1.57	0.02
WDR4	1.57	0.03
SLC20A2	1.57	0.00
EMB	1.57	0.03
C1orf31	1.57	0.00
SLC9B2	1.57	0.03
CYP20A1	1.57	0.02
XPO4	1.57	0.00
GNL3L	1.57	0.01
ZNF582	1.57	0.00
INTS6	1.57	0.01
CWC22	1.57	0.01
STAG2	1.57	0.00
LOC390940	1.56	0.02
DNAJB12	1.56	0.02
EPB49	1.56	0.01
COG7	1.56	0.01
COX17	1.56	0.02
ELOVL6	1.56	0.01
YARS2	1.56	0.02
KIAA0020	1.56	0.00
ITPKB	1.56	0.03
CCNJ	1.56	0.02
STK39	1.56	0.02
COX17	1.56	0.02
ZNF32	1.56	0.01
SEPHS1	1.56	0.02
HDAC7	1.56	0.02
MFSD2B	1.55	0.00
C10orf32	1.55	0.02
MSH3	1.55	0.00
ZNF618	1.55	0.00
MON1B	1.55	0.00
MTAP	1.55	0.04
TRMT11	1.55	0.01
PIWIL3	1.55	0.04
CST7	1.55	0.05
C10orf2	1.55	0.02
ALG13	1.55	0.01
DHX37	1.54	0.00
RPL7L1	1.54	0.00
STXBP3	1.54	0.00
PACRGL	1.54	0.03

SUMO3	1.54	0.01
MRPL3	1.54	0.00
SERF1A	1.54	0.01
SERF1B	1.54	0.01
SERF1A	1.54	0.01
SENP6	1.54	0.05
SAFB	1.54	0.01
AKAP8	1.54	0.03
QTRTD1	1.54	0.01
KDM3B	1.54	0.00
BTBD6	1.54	0.02
METTL21A	1.53	0.01
SLC43A1	1.53	0.03
ACAT2	1.53	0.05
TMEM204	1.53	0.00
MRPL15	1.53	0.01
SMEK1	1.53	0.04
DTYMK	1.53	0.03
DTYMK	1.53	0.03
CCL23	1.53	0.03
WDR6	1.53	0.01
SPG20	1.53	0.01
GAR1	1.53	0.05
PPCS	1.53	0.02
TSEN2	1.53	0.02
PRDM15	1.53	0.00
ZBTB40	1.53	0.02
SRPRB	1.53	0.02
TSPAN32	1.52	0.00
NEFH	1.52	0.02
OTUD4	1.52	0.03
SIPA1L1	1.52	0.00
EXOSC10	1.52	0.00
TTC32	1.52	0.02
PICK1	1.52	0.00
LIMD2	1.52	0.00
GRWD1	1.52	0.00
DDX56	1.52	0.00
PLD6	1.52	0.03
TIMM44	1.52	0.00
ZNF565	1.52	0.03
XIAP	1.52	0.02
ALG1	1.52	0.02
ZNF660	1.52	0.03
KIAA0240	1.52	0.01
ARAF	1.52	0.01
ZZZ3	1.52	0.00
NUP35	1.52	0.01
EXOG	1.52	0.03

ATAD2B	1.52	0.00
IFNGR2	1.51	0.02
ICA1L	1.51	0.01
MMAB	1.51	0.01
PNISR	1.51	0.04
GPN1	1.51	0.00
FUCA2	1.51	0.05
FAM165B	1.51	0.00
SREK1	1.51	0.00
METTL16	1.51	0.02
ARID2	1.51	0.01
DNAJC11	1.51	0.01
WIBG	1.51	0.03
ZNF566	1.51	0.05
C7orf29	1.51	0.01
MNF1	1.51	0.02
PTPRO	1.51	0.00
ELAC2	1.51	0.00
MANEAL	1.51	0.01
AP5S1	1.51	0.00
SLC35G1	1.51	0.01
HIRIP3	1.51	0.02
NSUN4	1.51	0.02
SMG5	1.51	0.02
PHF14	1.50	0.00
UMPS	1.50	0.01
CTPS2	1.50	0.03
IAH1	1.50	0.03
CLOCK	1.50	0.00
TRMT1	1.50	0.02
ZNF644	1.50	0.03
SELPLG	1.50	0.04
PCK2	1.50	0.03
CPNE8	1.50	0.05
MTMR1	1.50	0.01
RYR3	1.50	0.00
ZC3H7B	1.50	0.00
RBM23	1.50	0.04
TUBGCP5	1.50	0.00
NR3C1	1.50	0.00
CDK6	1.50	0.00
PLK1	1.50	0.01
MBNL3	1.42	0.00