

**Biochemical Analysis of the
BTG1 variants associated with
Non-Hodgkin's lymphoma**

Hibah Almasmoum, MSc

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy

April 2017

My PhD was dedicated to my wonderful Dad who passed away three years ago and whose calm mind and wise advice was my inspiration, your soul is always with me. Also, I dedicate my thesis to my nephew who was just born two weeks before my thesis submission. I know he 'll never read it, but I'll do my best to encourage him to always learn.

Abstract

Non-Hodgkin's lymphoma (NHL) is a group of lympho-proliferative disorders characterised by genetic mutations resulting in the selection of a malignant clone. Recently, mutations in the anti-proliferative B-cell translocation 1 gene (BTG1) and B-cell translocation 2 gene (BTG2) have been identified in NHL cases, which suggests a direct involvement of BTG1 and BTG2 in malignant transformation. BTG1 and BTG2 are members of the human BTG/TOB family. They are characterised by the conserved amino-terminal BTG domain, which mediates interactions with the human Caf1(hCaf1) catalytic subunit of the Ccr4-Not deadenylase complex. In addition, the BTG domain binds to the cytoplasmic poly (A)-binding protein (PABPC1). This complex plays a critical role in mRNA deadenylation and degradation as well as translational repression. It is currently unclear how, or indeed whether, mutations in BTG1 and BTG2 affect the function of the gene products. Therefore, a combination of sequence analysis and molecular modelling was used to predict the functional consequences of mutations previously identified in NHL. Sorting intolerant from tolerant (SIFT) and Suspect (Disease-Susceptibility-based SAV Phenotype Prediction) prediction tools enabled the identification of amino acid residues that would potentially interfere with the protein function, and hence may be associated with disease. In total 45 mutations in BTG1 and BTG2 were assessed. These mutations were derived from NHL samples, including diffuse large B-cell lymphoma (DLBCL), follicular lymphoma and Burkitt's lymphoma. Of the variants analysed, 15 were predicted to interfere with the function of BTG1 using SIFT analysis (Score ≤ 0.05). Only seven of these variants were predicted to be likely associated with disease using Suspect algorithm (Score ≥ 50), and an additional variant, BTG1 C149del, was predicted to interfere with protein function using PROVEN (Protein Variation Effect Analyzer). The ability of these protein variants to interact with known partners was established using yeast two hybrid assays. In addition, functional assessment of the role of the mutated proteins in cell cycle progression, translational repression and mRNA degradation was also performed. Using a yeast two-hybrid system, ten BTG1 variants were shown to affect the interaction of BTG1 with the hCaf1 (CNOT7/CNOT8) catalytic subunit of the Ccr4-Not deadenylase complex. In addition, when BTG1 variants were transfected into mammalian cells, these BTG1 variants (M11I, F25C, R27H, F40C, P58L, G66V, N73K and I115V), unlike the wild-type proteins, were not able to inhibit cell cycle progression. These results suggest that anti-proliferative BTG1 is required for hCaf1 (CNOT7/CNOT8) deadenylase activity. The remaining BTG1 variants (L37M, L94V, L104H and E117D) were not

consistent in the correlation of BTG1 interaction with hCaf1 (CNOT7/CNOT8) and inhibition of cell growth which led to the suggestion that BTG1 may require an additional factor such as PABPC1. Interestingly, several BTG1 variants (M11I, F25C, R27H, P58L, N73K I115V and E117D) did not require interaction with the hCaf1 (CNOT7/CNOT8) deadenylase enzyme to reduce reporter activity as established using 3' UTR tethering assays. This suggests that BTG1 may also have a role in regulating cell cycle progression and RNA degradation via Ccr4-Not deadenylase complex independent mechanisms. The data show that variants in BTG1 commonly found in DLBCL, are functionally significant and are likely to contribute to malignant transformation and tumour cell grow.

Acknowledgements

I must thank Allah for giving me the ability to do my PhD. I gratefully acknowledge my supervisors, Dr. Claire Seedhouse and Dr. Sebastiaan Winkler for their treasured guidance and valuable suggestions throughout this study, and I am particularly thankful for their patience, support and for having confidence in me.

My gratitude also goes out to the Saudi Arabia Ministry of Higher Education and to Umm Al-Qura University for funding my PhD scholarship.

I must also thank Lorenzo Pavanello for providing me with his plasmids (pEQ80L-PABPC1-99 and pEQ80L-PABPC1-190), as well as Dr. Laura Corbo for her gift of the BTG1 plasmid.

I would like to thank all members of the Haematology group and the Gene Regulation and RNA Biology group (GRRB), who always put up with my questions. I would also like to thank Prof David Heery for all his advice and encouragement. I am really thankful to Dr. Hilary Collins for teaching me confocal microscopy, for her support and patience in listening to me. I must thank Dr. Anna Piccinini for her encouragement and I am looking forward to seeing her lovely son. Very special thanks go out to my friends and wonderful lab mates for the lovely work atmosphere of the GRRB group, and for all the fun we have had over the last four years (Marayti, Blessing, Lorenzo, Ishwinder, Ben, Joel, Aimee, Dan, Richa, Hannah, Shoaib, Raj, Lin and Bismoay) and our lab technicians. I must thank Masar for her support and love. I would like to thank Jonny for his support, time and for encouraging me during the last two years of my PhD. Lots of thanks to my neighbor in the lab and office, Dr. (recently) Dario Leonardo Balacco; we had a lot of fun, crazy times and I want to thank you for sharing the journey with me during this research. Are we still colleagues, not friends?

I am sincerely grateful to my family. My special thanks go to my Mam, whose love and warm support were the sources of my motivation. To all my brothers and sisters for supporting me spiritually throughout this journey. I must thank Hatim, my brother, for his patience especially during my writing up stage. To my friend who has become a sister, shatha Felemban, thanks for your encouragement and for sharing my happiness and sadness. My lovely friend Hibah AbuSulayman and her lovely daughter, thanks for being a part of my journey.

Finally, I would also like to take this opportunity to thank all my friends: Sameerah, Rehab, Kholoud, Shahad, Manal, Reema, Amal, Amonah, Nada, Saja, Sheikh and Nazmeen. Thank you all for being a part of my life along the way, thank you for sharing my happiness and sadness and for being my second family in Nottingham.

List of contents

Abstract	iii
Acknowledgements	v
List of contents	vi
List of figures	xiii
List of tables	xvi
List of abbreviations	xviii
Chapter 1 Introduction	1
1.1. Overview	2
1.2. B-cell development and differentiation	2
1.2.1. Development and diversity of the B cell receptor (BCR).....	5
1.3. Cellular origin of human B cell lymphomas	7
1.4. Molecular pathogenesis of non-Hodgkin lymphoma	10
1.4.1. Chromosomal translocation	10
1.4.2. Genetic modification of tumour suppressor genes	11
1.4.3. Aberrant somatic hypermutations	11
1.5. Novel mutations identified in lymphoma using next generation sequencing (NGS) technologies	14
1.6. BTG/TOB Family	15
1.6.1. Discovery and definition	15
1.6.2. Structure of the BTG/TOB Family	15
1.6.3. BTG/TOB proteins as effectors in signalling pathways.....	19
1.6.4. BTG/TOB involvement in mRNA turnover	21
1.6.5. Role of BTG/TOB proteins in cancer	21
1.6.6. BTG1 and BTG2 proteins in haematological malignancies	25
1.7. Eukaryotic mRNA turnover	26
1.8. The Ccr4-Not complex	31
1.8.1. Overview of Ccr4-Not complex.....	31
1.8.2. Definition and structure of the Ccr4-Not complex	33
1.8.2.1. Catalytic subunits	33
1.8.2.2. Non-catalytic subunits.....	36

1.8.3.	Physiological function of the Ccr4-Not complex.....	38
1.8.4.	Ccr4-Not complex involved in cancer.....	38
1.8.5.	Role of Ccr4-Not complex in mRNA decay	39
1.8.6.	Regulation and recruitment of the Ccr4-Not complex.....	40
1.8.6.1.	The role of the Ccr4-Not complex in miRNA-mediated gene silencing	40
1.8.6.2.	Role of RNA-binding proteins	41
1.8.6.3.	Interaction with BTG/TOB proteins.....	41
1.9.	Aims of study	46
Chapter 2: Materials and Methods		49
2.1	In silico analysis of BTG1 and BTG2 mutations.....	49
2.2	Bacterial culture and transformation	49
2.2.1	Reagents, solutions and buffers used in bacterial culture	49
2.2.2	Preparation of Escherichia coli competent cells.....	50
2.2.3	Transformation of competent cells	50
2.2.4	Culture of Escherichia coli after transformation	51
2.3	Molecular biology	51
2.3.1	Reagents, stock solutions and buffers used in molecular biology methods:.....	51
2.3.2	Small scale plasmid DNA preparation.....	51
2.3.3	Large scale plasmid DNA preparation.....	52
2.3.4	Determination of DNA/RNA concentration.....	52
2.3.5	Agarose gel electrophoresis.....	52
2.3.6	DNA extraction and purification from an agarose gel.....	53
2.3.7	Restriction enzyme digestion of DNA.....	53
2.3.8	Removal of 5'phosphate group from linearised plasmid DNA.....	53
2.3.9	Ligation of DNA fragments	54
2.3.10	NEBuilder HiFi DNA Assembly Reaction.....	55
2.3.11	Polymerase chain reaction (PCR) for cloning	57
2.3.12	Site-directed mutagenesis	60
2.3.13	Genomic DNA Purification.....	62
2.3.14	Total RNA Purification.....	63
2.3.15	Reverse transcription	63
2.3.16	Genomic sequencing analysis using polymerase chain reaction (PCR)	63
2.4	Yeast two-hybrid analysis.....	65
2.4.1	Reagents, stock solutions and buffers for use in yeast culture	65

2.4.2	Culture of <i>Saccharomyces cerevisiae</i> strains	66
2.4.3	Transformation of <i>Saccharomyces cerevisiae</i> strains	66
2.4.4	Yeast two-hybrid analysis	68
2.4.5	Preparation of yeast cells for western blot analysis	69
2.5	Cell Culture	69
2.5.1	Reagents, stock solutions and buffers for use in tissue culture	69
2.5.2	Routine maintenance of lymphoma cell lines	70
2.5.3	Routine maintenance of HEK293T cells	70
2.5.4	Freezing cells for long term storage.....	71
2.5.5	Retrieving cells from liquid nitrogen storage	71
2.6	Edu labelling of S-phase cells.....	72
2.6.1	Reagents, stock solutions and buffers for use in Click-iT EdU labelling	72
2.6.2	Cell seeding and transfection for Click -iT EdU labelling	72
2.6.3	Click-iT EdU labelling.....	73
2.6.4	Calculating the number of cells in S-phase.....	74
2.6.5	Measuring cell fluorescence using ImageJ.....	75
2.6.6	Western blot analysis for proliferation assay	76
2.7	RNA tethering reporter assay	76
2.7.1	Cell seeding and transfection for mRNA analysis	76
2.7.2	RNA tethering reporter assay	77
2.7.3	Preparation of RNA.....	78
2.8	RNA analysis	78
2.8.1	Production of cDNA	78
2.8.2	Quantitative PCR analysis.....	79
2.9	Protein analysis	81
2.9.1	Reagents, stock solutions and buffers for use in protein analysis.....	81
2.9.2	Bradford assay to determine protein concentration	81
2.9.3	Preparation of sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)	82
2.9.4	Gel staining using Coomassie Brilliant Blue	83
2.9.5	Western blotting and immunodetection	83
2.10	Co-Immunoprecipitation analysis	85
2.10.1	Preparation of protein lysates for immunoprecipitation	85
2.10.2	Preparation of antibody-agarose beads and immunoprecipitation.....	86

2.11	GST and His pull down analysis	87
2.11.1	Protein expression: GST- Tagged BTG1 and BTG2	87
2.11.2	Protein expression: His Tagged-PABPC1(1-190) and PABPC1(1-99)	88
2.11.3	Large scale protein expression: His-Tagged PABPC1(1-190) and PABPC1(1-99)	88
2.11.4	GST- pull down of His-PABPC1 (1-99) or (1-190) using GST- BTG1 and GST-BTG1 and BTG2	89
2.11.5	His- pulldown of GST-BTG1 and GST BTG2 using His tagged with PABPC1 (1-190) or PABPC1 (1-99).....	90
2.12	Statistical analysis	91
Chapter 3: Identification of somatic mutations of BTG1 and BTG2 genes in non-Hodgkin's lymphoma		93
3.1	Introduction.....	93
3.2	Sequencing of BTG1 and BTG2 genes in lymphoblastoid cell lines ..	96
3.2.1	Amplification and analysis of BTG1 sequence in lymphoblastoid cell lines.....	96
3.2.2	Amplification and analysis of BTG2 sequence in lymphoblastoid cell lines and variants detected in KM-H2 cell line.....	97
3.3	Data collection and bioinformatics analysis of mutations in BTG1 and BTG2	101
3.3.1	Identification of BTG1 and BTG2 variants according to type of NHL.....	101
3.3.2	BTG1 and BTG2 variants found in lymphoma analysed using predictive algorithms	104
3.3.2.1	BTG1 and BTG2 variants found in lymphoma analysed using Sorting intolerant from tolerant (SIFT)	104
3.3.2.2	BTG1 and BTG2 variants found in lymphoma analysed using Suspect analysis	106
3.4	Discussion.....	113
3.4.1	BTG1 and BTG2 variants in lymphoma cell lines	113
3.4.2	BTG1 and BTG2 variants in lymphoma patients	114
3.4.3	BTG1 and BTG2 variants evaluated by SIFT and Suspect analysis	115
Chapter 4: Lymphoma BTG1 variants impact on the interaction with human Caf1 (CNOT7/CNOT8)		119

4.1	Introduction.....	119
4.2	Structural analysis of BTG1 variants	121
4.3	Generation of plasmids for yeast two-hybrid interaction analysis	128
4.4	Evaluation of the interaction between human Caf1 (CNOT7/CNOT8) and BTG1 variants through yeast two-hybrid analysis	128
4.4.1	Interaction analysis of BTG1 variants located in the α -helical regions and human Caf1	131
4.4.2	Interaction analysis of BTG1 variants located in the β -sheet regions and human Caf1	134
4.4.3	Interaction analysis of BTG1 variants located in the loop regions and human Caf1	136
4.4.4	The effect of the deletion of the C-terminal 21 amino acids of BTG1 on the interaction with human Caf1	138
4.5	Discussion.....	140
4.5.1	BTG1-hCaf1/CNOT7 complex	140
4.5.2	BTG1 variants located in box A and box B impact on the interaction with human Caf1	141
4.5.3	BTG1 variant residues located outside of box A and box B regions impact on the interaction with human Caf1	142
Chapter 5: The effect of lymphoma-derived BTG1 mutations on the regulation of cell proliferation		145
5.1	Introduction	145
5.1	Generation of plasmids for the proliferation assay	149
5.2	Optimisation of 5-ethynyl-2'-deoxyuridine (EdU) staining for HEK293T cells.....	151
5.3	Regulation of cell proliferation by BTG1 variants	155
5.3.1	Effect of BTG1 variants mapped to α -helices on the anti-proliferative function	155
5.3.2	Effect of BTG1 variants mapped to β -sheets on the anti-proliferative function..	156
5.3.3	Effect of the BTG1 variants mapped to loop regions on the anti-proliferative function.....	156

5.4 Investigation of the anti-proliferative activity of BTG1 variants that are able to interact with human Caf1 (CNOT7/CNOT8)	161
5.5 Discussion	163
5.5.1 Effect of BTG1 variants on the anti-proliferative function and disruption of the interaction with human Caf1	163
5.5.2 Investigation of the impact on protein functions of BTG1 mutations not affecting the interaction with human Caf1.....	165
Chapter 6: Regulation of mRNA degradation and translation by BTG1 variants	169
6.1 Introduction	169
6.2 Generation of plasmids for the RNA tethering assay	173
6.3 Effect of tethering the BTG1 variants to the 3' UTR of the luciferase reporter mRNA on mRNA translation and degradation	175
6.3.1 BTG1 variants mapping onto the α -helical structure	175
6.3.2 BTG1 variants mapped onto β -sheet structures	176
6.3.3 BTG1 variants mapped in the loop regions	177
6.4 BTG1 variants able to regulate mRNA decay via binding with human Caf1 (CNOT7/CNOT8)	181
6.5 Discussion	183
Chapter 7: An investigation of the interaction between BTG1 and PABPC1	187
7.1 Introduction	187
7.2 Generation of plasmids for the yeast two-hybrid analysis	189
7.3 Generation of plasmids for the pull-down assay	193
7.4 Investigating the interaction between BTG1 and PABPC1 using the yeast two-hybrid system	195
7.4.1 Yeast two-hybrid analysis based on the Gal4 system	195
7.4.2 Yeast two-hybrid analysis based on the LexA system	197
7.4.2.1 L40 yeast strain.....	197
7.4.2.2 EGY48 yeast strain.....	198

7.5 Investigating the interaction between BTG1 and PABPC1 using immunoprecipitation and pull-down assays	203
7.5.1 Co-immunoprecipitation analysis using mammalian cells.....	203
7.5.2 GST pull-down assay using bacterial expression.....	203
7.5.3 His pull-down assay.....	204
7.6 Discussion	210
7.6.1 Immunoprecipitation analysis.....	210
7.6.2 Yeast two-hybrid analysis.....	211
7.6.3 GST and Ni ²⁺ -NTA pulldown experiments	212
Chapter 8: Concluding Remarks and Future Outlook	214
8.1 Overview	214
8.2 BTG1 and BTG2 variants associated with non-Hodgkin's lymphoma... 214	
8.3 Truncated BTG1 with 21 amino acids at the C-terminal affects proliferation activity.....	215
8.4 The anti-proliferative activity of BTG1 requires human Caf1 (CNOT7/CNOT8) interaction.....	216
8.4.1 BTG1 variants located in the α -helical regions.....	216
8.4.2 BTG1 variants located in the β -sheet regions	217
8.4.3 BTG1 variants located in the loop regions.....	217
8.5 Does BTG1 require only the interaction with human Caf1 (CNOT7/CNOT8) to inhibit cell proliferation?.....	218
8.6 Do the BTG1 variants stimulate mRNA abundance via binding with human Caf1 (CNOT7/CNOT8)?	219
8.7 Future work.....	222
References	224
Appendix.....	246

List of figures

Chapter 1

Figure 1-1 B- cell differentiation in the germinal Centre.....	4
Figure 1-2 The B cell expressed Immunoglobulins (Igs) during the maturation and differentiation.....	6
Figure 1-3 Cellular origin of human B cell lymphomas.....	8
Figure 1-4 Schematic of the molecular pathogenesis of NHLs.....	12
Figure 1-5 Schematic view of the BTG/TOB protein family.....	17
Figure 1-6 Structure of the BTG domain.....	18
Figure 1-7 Schematic diagram of the canonical eukaryotic mRNA degradation pathway.....	30
Figure 1-8 Schematic overview of human catalytic subunits of Ccr4-Not complex ...	35
Figure 1-9 Model structure of the mammalian Ccr4-Not complex.....	37
Figure 1-10 Model for recruitment the Ccr4-Not complex.....	44
Figure 1-11 Model of the nuclease sub-complex in association with BTG domain of TOB1.....	45
Figure 1-12 Scheme of the experimental approach used.....	47

Chapter 2

Figure 2-1: Counting the cell nuclei using ImageJ.....	75
Figure 2-2: Measuring the fluorescence intensity using ImageJ.....	76

Chapter 3

Figure 3-1 Amplification of BTG1 in lymphoblastoid cell lines.....	98
Figure 3-2. Analysis of BTG1 sequence variants detected in the Raji cell line.....	99
Figure 3-3. Amplification and analysis of BTG2 sequence variant detected in KM-H2 and lymphoblastoid cell lines.....	100
Figure 3-4 Somatic mutations in the BTG1 and BTG2 genes in different types of lymphoma.....	103
Figure 3-5: Number of somatic mutations in the BTG1 and BTG2 in different types of lymphoma and evaluation by SIFT.....	108
Figure 3-6: Schematic representation of the BTG1 and BTG2 mutations mapped onto their protein sequences.....	109

Chapter 4

Figure 4-1 Protein sequence alignment of the BTG domain in TOB1, BTG1 and BTG2.	124
Figure 4-2 Structural models of the BTG1-hCaf1/CNOT7 and BTG2-hCaf1/CNOT7 complexes.....	125
Figure 4-3 Mapping of variant residues in the BTG1 structural model.....	127
Figure 4-4 Generation of vectors for yeast two-hybrid interaction analysis.	129
Figure 4-5 Schematic of yeast two-hybrid interaction analysis of BTG1 and human Caf1.....	130
Figure 4-6 Differential interactions between human Caf1 and BTG1 variants located in α -helical regions.	133
Figure 4-7 Differential interactions between human Caf1 and BTG1 variants located in β -sheets regions.....	135
Figure 4-8 Differential interactions between human Caf1 and BTG1 variants located in C-loop regions.....	137
Figure 4-9 Evaluation of interaction between BTG1 lacking the C-terminal 21 amino acids and human Caf1.....	139
Chapter 5	
Figure 5-1 Cell cycle progression and EdU labeling of dividing cells.	147
Figure 5-2 Generation of vectors for the proliferation assay.....	150
Figure 5-3 Optimisation of staining of HEK293T cells with EdU.	153
Figure 5-4 Comparison of the difference in fluorescence intensity of cells stained with EdU, in both cells expressing WT-BTG1 and the control.	154
Figure 5-5 Effect of BTG1 α -helix variants on the anti-proliferative function.....	158
Figure 5-6 Effect of BTG1 β -sheet variants on the anti-proliferative function.....	159
Figure 5-7 Effect of BTG1 loop region variants on the anti-proliferative function..	160
Figure 5-8 Investigation of the impact of BTG1 variants on protein function.....	162
Chapter 6	
Figure 6-1 Schematic model of the recruitment of BTG1 to the reporter mRNA for the tethering assay.	171
Figure 6-2 Generation of vectors for the tethering assay.....	174
Figure 6-3 BTG1 variants located in the α -helical regions and were able to reduce the mRNA abundance of a tethered mRNA	178

Figure 6-4 BTG1 variants located in the β -sheets and able to reduce the mRNA abundance of a tethered mRNA.....	179
Figure 6-5 BTG1 variants located in the loop regions and able to reduce the mRNA abundance of a tethered mRNA.....	180
Figure 6-6 BTG1 variants able to reduce the mRNA abundance of a tethered mRNA	182
Figure 6-7 Schematic model of the recruitment of BTG1 to the reporter mRNA for the tethering assay required for PABPC1 interaction.....	185
Chapter 7	
Figure 7-1 Generation of the yeast expression vector encoding LexA BTG1.	191
Figure 7-2 Generation of yeast expression vectors encoding PABPC1 for yeast two hybrid analysis.	192
Figure 7-3 Generation of bacterial expression plasmids for pulldown assays.....	194
Figure 7-4 Investigation of BTG1-PABPC1 interactions using yeast two-hybrid analysis based upon the Gal4 system.....	196
Figure 7-5 Investigation of BTG1 and PABPC1 interactions using a LexA yeast two-hybrid system and L40 cells.	200
Figure 7-6 Detection of PABPC1 and BTG1 interactions using a LexA yeast two hybrid system and EGY48 cells.	201
Figure 7-7 Interaction analysis of BTG1 variants and PABPC1.....	202
Figure 7-8 Detecting the interaction between BTG1 with PABPC1 using GST-pull down assays.....	206
Figure 7-9 Detecting the interaction between BTG2 and PABPC1 using GST-pull down assays.	207
Figure 7-10 Detecting the interaction between BTG1/BTG2 and the RRM1 domain of PABPC1 using His-pull down assays.	208
Figure 7-11 Detecting the interaction between BTG1/BTG2 and the RRM1 and RRM2 domains of PABPC1 using His-pull down assays.	209

List of tables

Chapter 1

Table 1-1 List of most common of types non-Hodgkin's lymphoma.	9
Table 1-2 Mutations involved in B-cell non-Hodgkin's lymphoma pathogenesis	13
Table 1-3 Relationship of BTG/TOB expression and cancer	24
Table 1-4 Enzymes involved in cytoplasmic mRNA degradation.....	29
Table 1-5 Ccr4-Not components in yeast and human cells.....	32

Chapter 2

Table 2-1 Standard ligation reaction.....	54
Table 2-2 DNA Assembly Reaction.....	55
Table 2-3 List of the primers used for HiFi DNA assembly	56
Table 2-4 Standard reaction set up of PCR using the Phusion DNA polymerase.....	58
Table 2-5 Standard reaction set up of PCR using Taq DNA polymerase.	58
Table 2-6 List of primers used for DNA cloning.	59
Table 2-7 List of primers used for site-directed mutagenesis.....	62
Table 2-8 List of primers used for gene expression and sequencing.....	64
Table 2-9 Plasmids used for yeast transformation	68
Table 2-10 Composition of Yeast transformation mixtures	68
Table 2-11 Transfection set up used for calcium phosphate precipitation method. .	77
Table 2-12 RNA tethering assay: cDNA synthesis reaction mix 1.....	79
Table 2-13 RNA tethering assay: cDNA synthesis reaction mix 2.....	79
Table 2-14 List of primers used for qPCR	80
Table 2-15 qPCR reaction setup.....	80
Table 2-16 Preparation of SDS-PAGE gels	82
Table 2-17 Primary antibodies used for western blotting.....	84
Table 2-18 Secondary antibodies used for western blotting.....	84
Table 2-19 Transfection setp used for co- immunoprecipitation.	85
Table 2-20 List of antibodies used for co-immunoprecipitation.....	87
Table 2-21 List of the samples tested in His-pull down experiments	91

Chapter 3

Table 3-1: list of BTG1 variants evaluated with SIFT and Suspect in different types of lymphoma.....	110
--	-----

Table 3-2: list of BTG2 mutations evaluated with SIFT and Suspect in different types of lymphoma.....	111
Table 3-3: list of BTG1 and BTG2 mutations in lymphoma cell lines evaluated with SIFT.	112
Chapter 4	
Table 4-1 Percentage identity for protein sequences used for the alignment of the BTG domain in TOB1, BTG1 and BTG2.....	123
Table 4-2 List of variant residues grouped according to their location in the secondary structure of BTG1.....	126
Chapter 5	
Table 5-1 Data analysis of BTG1 variants using bioinformatics and yeast-two-hybrid interaction analysis.	148
Chapter 6	
Table 6-1 Data analysis in BTG1 variants using bioinformatics and yeast two-hybrid interaction and proliferation analysis.	172
Chapter 8	
Table 8-1 Data analysis in BTG1 variants using bioinformatics and experimental analysis.....	221

List of abbreviations

7-methylguanosine	m ⁷ G
ABC	Activated B-cell
AD	Activation domain
Ago	Argonaute
AICDA or AID	Activation-induced cytidine deaminase
ALL	Acute lymphoblastic leukaemia
AML	Acute myeloid leukaemia
BCL-6	B-cell lymphoma 6
BCP-ALL	B-cell precursor- ALL
BCR	B-cell receptor
BL	Burkitt's lymphoma
BLAST	Basic Local Alignment Search Tool
BMP	Bone morphogenetic protein
BTG1	B-cell translocation 1
BTG2	B-cell translocation 2
Caf1	Ccr4-associated factor1
Ccr4	Carbon catabolite repression 4
CD	Circular dichroism
CDK4	Cyclin-dependent kinase 4
CLL	Chronic lymphocytic leukaemia
COSMIC	Catalogue of Somatic Mutations in Cancer
CPEB3	Cytoplasmic polyadenylation element binding 3
CSR	Class switch recombination
DBD	DNA binding domain
Disease-Susceptibility-based Single Amino acid Variants (SAVs)	Suspect
DLBCL	Diffuse large B-cell lymphoma
DSF	Differential scanning fluorimetry
EBV	Epstein-barr virus
EdU	5-ethynyl-2'-deoxyuridine

EEP	Exonuclease-endonuclease-phosphatase
eIF4E	Eukaryotic initiation factor 4E
eRF3	Eukaryotic release factor
FL	Follicular lymphoma
GC	Germinal centre
GCB	Germinal center B-cell
GEP	Gene expression profiling
GFP	Green fluorescent protein
GR	Glucocorticoid receptor
hCaf1	Human Caf1
HL	Hodgkin's lymphoma
IgH	Heavy immunoglobulin
IgL	Light immunoglobulin
Igs	Immunoglobulins
LRR	Leucine-rich repeat
MCL	Mental cell lymphoma
NGS	Next generation sequencing
NHL	Non-Hodgkin's lymphoma
NK	Natural killer
PABPC1	Cytoplasmic poly(A) binding protein
PAM2	PABP-interacting motif 2
PARN	Poly (A) specific ribonuclease
Phyre	Protein Homology/analogy Recognition Engine
PPIs	Protein-protein interactions
PRMT1	Protein Arginine Methyl Transferase
PROVEAN	Protein Variation Effect Analyzer
RA	Retinoic acid
RB	Retinoblastoma protein
RNA-seq	RNA sequencing
RRM1	RNA-binding domain 1
RSS	Recombination signal sequence

SAVs	Single Amino acid Variants
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SHM	Somatic hypermutation
SIFT	Sorting Tolerant From Intolerant
SNVs	Single nucleotide variants
SSCP	Single-strand conformation polymorphism
TBST	Tris-buffered saline supplemented with 0.05% Tween-20
TPA	12-O-tetradecanoylphorbol-13-acetate-inducible sequence 2
WGS	Whole genome exome sequencing
WHO	World Health Organization

Chapter 1
Introduction

Chapter 1 Introduction

1.1. Overview

Malignant lymphomas derived from a monoclonal proliferation of malignant lymphocytes or natural killer cells (NK), can occur in the lymph nodes or in extra-nodal sites. These lymphomas can be divided into two major groups: Hodgkin's lymphoma (HL) and non-HL (NHL) (Coupland, 2013). NHL are a large heterogeneous group of neoplasms; about 85-90% of NHLs are derived from B-cells whereas the remaining are derived from T-cells and NK cells (Coupland, 2013). These are classified according to World Health Organization (WHO) guidelines with the recent 2016 revision building on the 2008 classification, incorporating information from clinical findings, and the morphological, immunophenotyping and genetic features of each type of lymphoma (Swerdlow et al., 2008, Swerdlow et al., 2016). An increase in the availability of high-throughput sequencing techniques, in both research and diagnostic laboratories, has significantly advanced the understanding of the molecular pathogenesis of B-cell NHL and identified specific factors that influence and contribute to the initiation and maintenance of lymphogenesis and disease progression (Coupland, 2013, Morin and Gascoyne, 2013, Swerdlow et al., 2016). Many genetic mutations have been identified in lymphomas, including ones in B cell translocation gene 1 (BTG1); the functional consequence of these is not always fully understood. Therefore, understanding the function of BTG1 and how aberrations in this gene are associated with the haematological disorder will be addressed.

1.2. B-cell development and differentiation

The B lymphocyte is a capable, adaptive and long lasting immune system component, as it produces a high affinity interaction between immunoglobulins (Igs) and antigens. These cells play a role in the humoral immune response and in immunological memory (Seifert et al., 2013). The B lymphocytes originate in the bone marrow and express diversity in genes for

the B-cell receptor (BCR) during their development. Upon the expression of a specific BCR and once differentiated, the mature B lymphocyte cells leave the bone marrow as naive B-cells; any that are unable to produce a BCR-ligand complex undergo apoptosis. Mature B-cells circulate and migrate into the secondary lymphoid organs such as lymph nodes, where they reside until a specific B cell becomes activated when its BCR binds with an antigen. The antigen is presented to the B cell in the presence or absence of T-cells, which leads to T-cell dependent or independent activation, respectively. Some of the the B cells differentiation into short lived plasma cells, and some develop into memory B cells (Kurosaki et al., 2015). Alternatively, the activated B cells under go proliferation to form germinal centre (GC) site. In the dark zone of the GC clone expansion and proliferation of B-cells with specific BCR through genetic modification and diversity called somatic hypermutation (SHM) process. Then, the B-cells exit the cell cycle and relocated to the light zone, where affinity selection take place through interaction with immune complex coated with follicular dendritic cells and antigen specific T -folcuillar helper cells (Kurosaki et al., 2015). The cells undergo class switch recombination (CSR) to create specific immunoglobulins; those that produce high affinity antibodies are either differentiated into memory B cells or plasma cells, whereas the cells with low affinity immunoglobulin receptors are eliminated through apoptosis (Figure 1-1). The expression of the B-cell lymphoma 6 (BCL-6) is one of the crucial transcription factor initiates and maintains the GC reaction (Schneider et al., 2011).

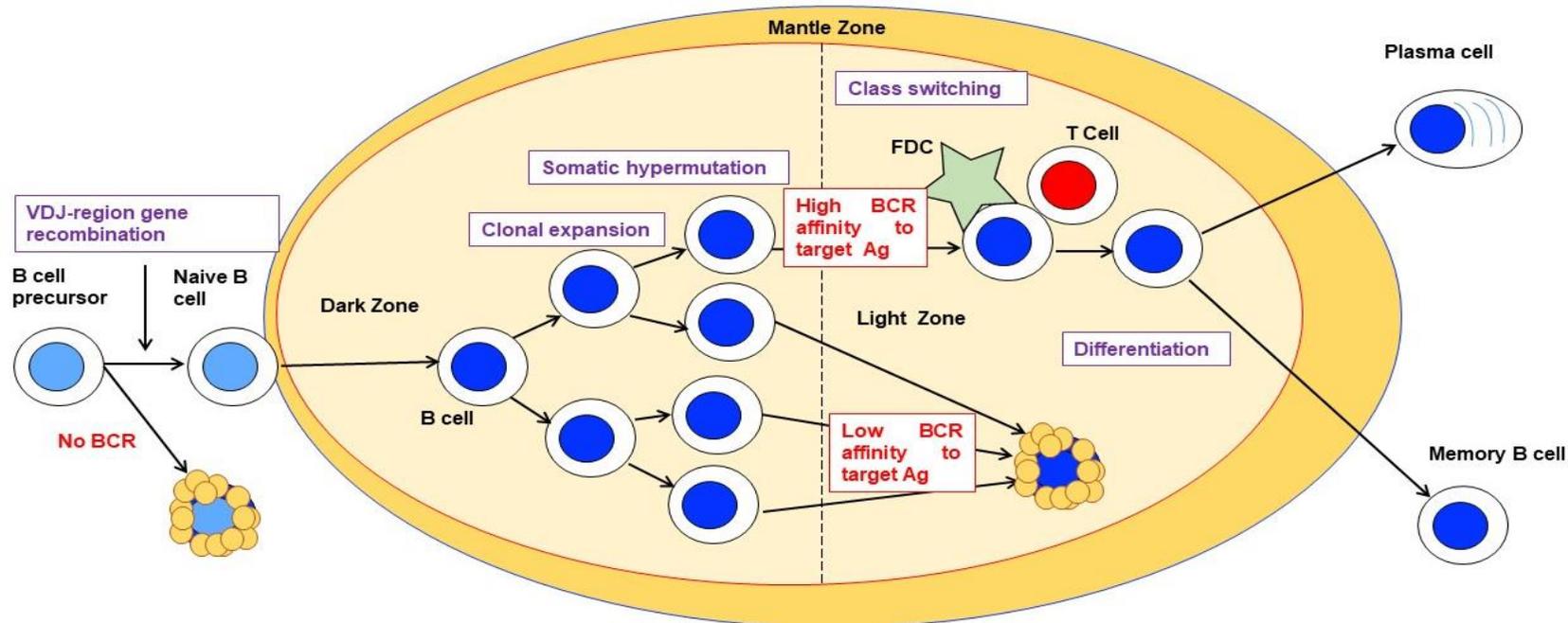


Figure 1-1 B- cell differentiation in the germinal Centre.

B cells are produced in the bone marrow and form B cell receptors (BCR) through a VDJ region gene recombination process. The antigen-activated naive B cells migrate into the primary B-cell follicles in secondary lymphoid organs such as lymph nodes, where they are driven into germinal centre (GC). The GC B cells undergo massive clonal expansion and proliferation in the dark zone of the GC and activate the process of somatic hypermutation to produce B cells with high affinity; cells with low affinity are programmed for apoptosis. In the light zone, positive selection of mutated B cells with high affinity occurs, where the GC B cells are in close contact with CD4⁺T cells and follicular dendritic cells (FDCs). Fractions of these GC B cells undergo class-switch recombination. Finally, GC B cells differentiate into memory B cells or plasma cells and leave the GC microenvironment. The majority of GC B cells will acquire disadvantageous mutations and undergo apoptosis (Küppers, 2005).

1.2.1. Development and diversity of the B cell receptor (BCR)

All mature B cells express a membrane-bound Ig or an antibody which harbours specificity for antigen. The BCR is formed with the cofactors immunoglobulin alpha and beta (Ig α /Ig β), which participate in signal transduction of this surface receptor (Seifert et al., 2013). The membrane-bound antibody or Ig is composed of two identical heavy (IgH) and two identical light (IgL) chains that are linked by disulphide bridges. The IgH chain genetic locus is located on chromosome 14, while the IgL genes (of which there are two types, either kappa or lambda) are located on chromosome 2 and 22, respectively (Küppers, 2005, Nogai et al., 2011, Blombery et al., 2015). During B cell development, these immunoglobulin genes are recombined by V(D)J recombination. This process separates the germline DNA sequence and then recombines to create the constant region of variable (V), diversity (D) and joining (J) regions in the heavy chain, and the two constant regions of the V and J at the IgL loci. The V(D)J recombination, mediated by the RAG enzyme complex, is guided into the specific recombination signal sequence (RSS). The RAG complex brings together two gene segments via associated RSSs, and then a non-homologous end-joining process repairs the DNA structure (Ma et al., 2002, Blombery et al., 2015) (Figure 1-2 A and B). SHM processing is mediated by activation-induced cytidine deaminase (AID), also known as AICDA. The AID enzyme functions to deaminate, or to induce a mutation in a cytosine (C) in the single strand DNA resulting in a uracil (U) residue, this is preferentially guided by hotspot motifs (WRCY motifs W=adenine or thymine, R=purine, C=cytosine, Y=pyrimidine, or the inverse RGYW G=guanine). Consequently, the mismatch U:G is then repaired by various DNA repair pathways (Rogozin and Kolchanov, 1992, Di Noia and Neuberger, 2002, Ramiro et al., 2003, Küppers, 2005). Indeed, AID mediates the CSR process by changing the class of heavy chain isotypes from IgM into either IgG, IgA, IgE, or IgD, with retained antigen specificity but varying effector function at the IgH locus. After the GC

reaction, B-cells develop into memory B-cells or plasma cells (Manis et al., 2002, Nogai et al., 2011)(Figure 1-2 C and D).

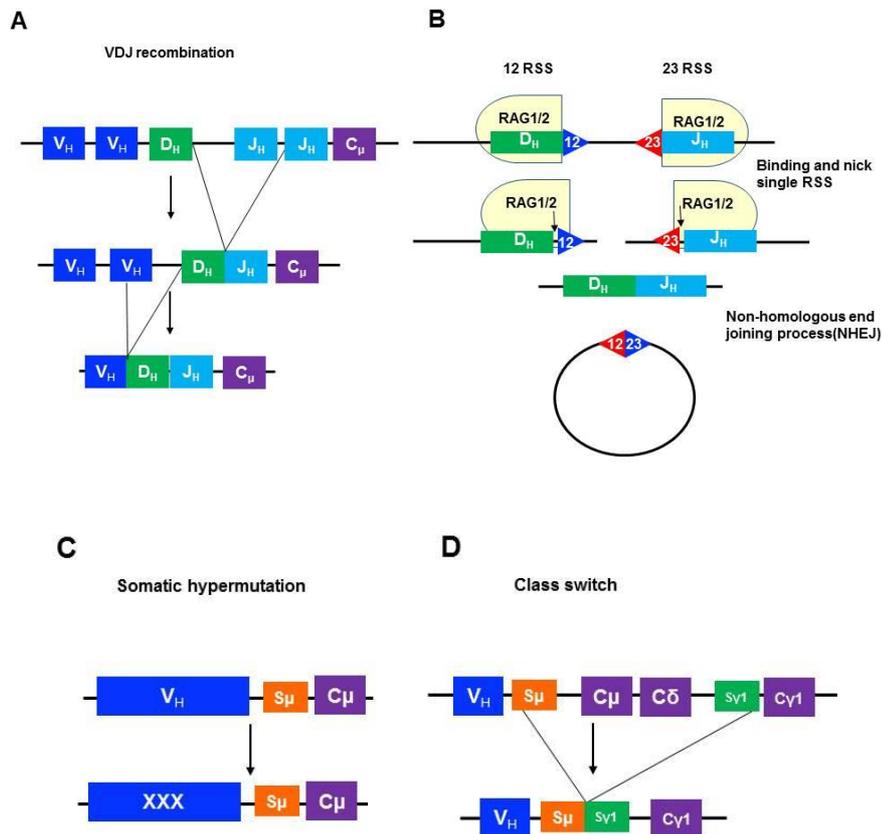


Figure 1-2 The B cell expressed Immunoglobulins (Igs) during the maturation and differentiation.

(A)The rearrangement of the DNA segments encoding either heavy (H) chain or light-chain regions of the variable gene, this process is called VDJ recombination. Three gene segments variable(V), diversity (D), and joining (J) are joined to encode the H-chain region, and two gene segments V and J genes join for the light chain (not shown). (B) The V(D)J recombination process is initiated by double strand DNA breaks catalyzed by recombinase components (RAG1/2) at specific sites called recombination signal sequences (RSSs). The RSSs composed of conserved heptamer or nonamer sequences, V(D)J rearrangement occurs between RSS with dissimilar spacer lengths (known as the 12/23 rule) followed by the non-homologous end joining (NHEJ) process to repair a double strand DNA breaks. (C) Somatic hypermutation process is when the activated B cell reach GC, and mutations are introduced to rearrange the V- region of Ig genes to produce a high affinity B cell. (D) Class switch recombination (CSR) from IgM or IgD into another isotype of Igs. In the diagram, the C-region for IgM and IgD are exchanged for the C-region of IgG to produce antibody with different effector functions but with the same antigen-binding domain (Küppers, 2005, Seifert et al., 2013).

1.3. Cellular origin of human B cell lymphomas

The majority of malignant B-cell lymphomas arise either during the GC reaction or after. The classification of B-cell lymphomas is not only based on the cell origin, but also on the specific characteristics of the B-cell differentiation stage of the lymphoma precursor. Histological and immunohistochemistry studies are required to determine the origin of the malignant cells. Follicular lymphoma (FL) is morphologically similar to GC B-cells, grows in the follicular structures and represents 20% of cases out of all lymphoma types (de Jong, 2005). Burkitt's lymphoma (BL) is a highly aggressive disorder caused by the disruption of normal lymph node structure, and growth into a diffuse pattern. These malignant cells are morphologically similar to centroblast cells derived from the GCB cells (Küppers et al., 1999, Küppers, 2005). Genome expression profiling (GEP) with a histological analysis can be used to identify the diffuse large B cell lymphoma (DLBCL) with subtypes, which represents the most common lymphoma about 30-40%. The first subtype is the germinal centre DLBCL (GCB-DLBCL) which often occurs in the SHM process. The activated B-DLBCL (ABC-DLBCL) lacks most of the features of the GC B-cells, and has a highly activated phenotype (Alizadeh et al., 2000, Rosenwald et al., 2002). Other types of NHL are summarised in Table (1-1). (Figure 1-3).

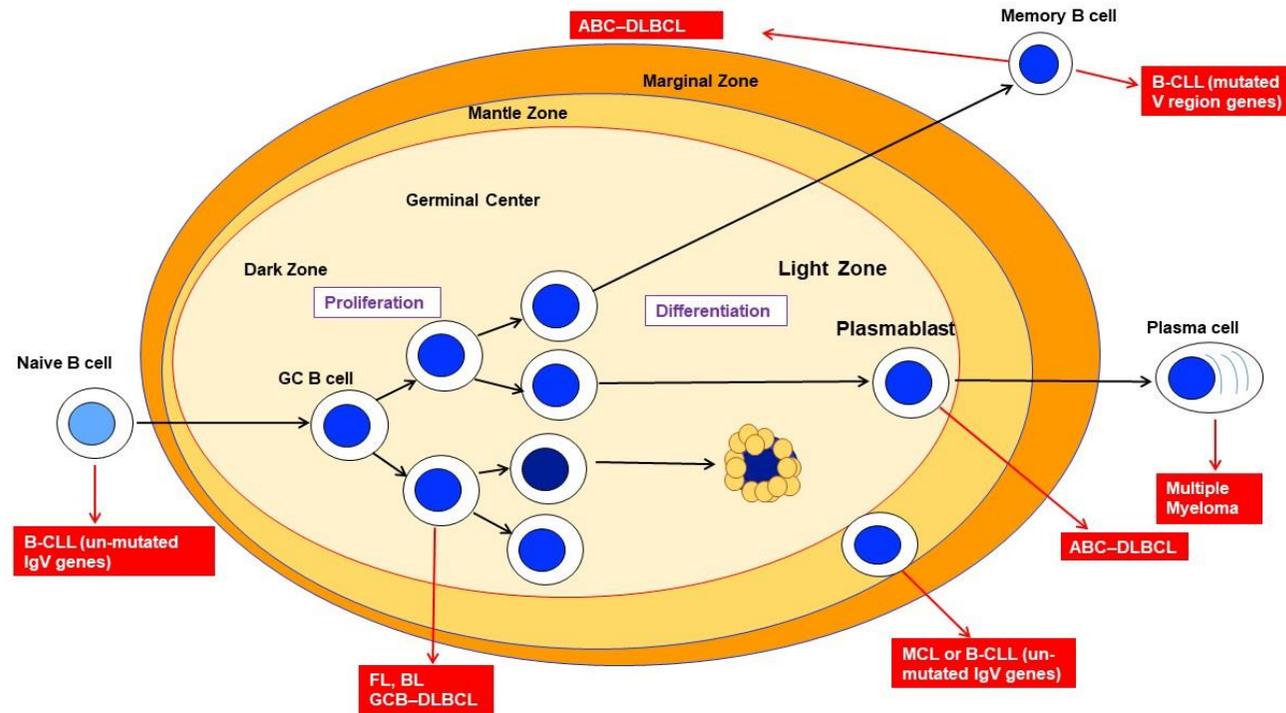


Figure 1-3 Cellular origin of human B cell lymphomas.

Most lymphomas are derived from a germinal centre (GC), arising from cells which are blocked at specific differentiation stages. Such lymphomas include Burkitt's lymphoma (BL), follicular lymphoma (FL) and GC B-cell-like diffuse large B-cell lymphoma (GCB-DLBCL). Alternatively, activated B-cell-like DLBCL (ABC-DLBCL) shows characteristics of late GC B-cells as memory B-cells, or plasmablasts, that are committed to plasma cell differentiation. Also, the majority of mantle cell lymphoma (MCL), derived from the B-cell (CD5+) in the mantle zone, carries unmutated IgV genes. However, 20-30% of MCL lymphomas are mutated in IgV genes, indicating that they have passed through the GC. About 50% of B-cell chronic lymphocytic leukaemia (B-CLL) cases carry mutations in IgV genes. B-cell chronic lymphocytic leukaemia has been proposed to derive either from CD5+ B-cells, marginal zone B-cells (unmutated IgV genes) or plasma cells (mutated IgV genes) (Küppers, 2005, Seifert et al., 2013).

Lymphoma cellular	Features	Frequency among Proposed (%)	Origin lymphomas
Diffuse large B-cell	Heterogenous group of lymphomas characterized by large B cells. Several subtypes are recognized. Morphological variants include centroblasts and immunoblasts.	30–40	GC or post-GC B cell
Follicular lymphoma	A nodal lymphoma with a follicular growth pattern. Lymphoma cells morphologically and phenotypically resemble GC B cells.	20	GC B cell
Mantle-cell lymphoma	Lymphoma arises from cells that populate the mantle zone of follicles, express CD5 and show aberration in cyclin-D1 expression.	5	CD5+ mantle-zone B cell
Multiple myeloma	Neoplastic proliferation of plasma cells in the bone marrow.	10	plasma cell
Extranodal marginal-zone B-cell (MALT) lymphoma	Develops mostly in lymphoid structures.	7	Marginal-zone B cell
Burkitt's lymphoma	Fast growing. Mostly extranodal. Characterized by a MYC–Ig translocation.	2	GC B cell
Primary mediastinal B-cell lymphoma	Subtype of diffuse large B-cell lymphoma located in the mediastinum. Tumour cells are large B cells but also show a number of similarities to Reed–Sternberg cells of classical Hodgkin's lymphoma.	2	Thymic B cell
Lymphoplasmacytic lymphoma	Involves lymph nodes, bone marrow and spleen. The tumour-cell population is composed of small B cells, plasmacytoid lymphocytes and plasma cells. Most patients present with a serum monoclonal protein, usually of the IgM type.	2	Thymic B cell
B-cell chronic lymphocytic	Leukaemia of small B cells that express the CD5 antigen, involving peripheral-blood and bone-marrow cells. Common in elderly patients. Called 'small lymphocytic lymphoma' when lymph-node cells are predominantly involved. Patients with leukaemia cells that lack variable (V)-region gene mutations have a worse prognosis than patients with mutations in V-region genes.	7	Memory B cell? Naive B cell? Marginal-zone B cell?

This table is adapted from (Küppers, 2005).

Table 1-1 List of most common types of non-Hodgkin's lymphoma.

1.4. Molecular pathogenesis of non-Hodgkin lymphoma

Genetic alterations such as chromosomal translocations, mutations or amplifications can occur in a proto-oncogene or tumour suppressor gene. If the affected gene is important in the regulation of cell proliferation mechanisms or cell differentiation, the cell may become malignantly transformed (Gouveia et al., 2012) (Table 1-2 and Figure 1-4). Other factors involved in lymphogenesis include viruses and cell microenvironments. Most of the BL cases are associated with Epstein–Barr virus (EBV) in geographical areas where the virus is endemic (90%) (Küppers, 2003). In FL, the interaction of tumour cells with other cells in the microenvironment has been shown to play an essential role in both their survival, and in their proliferation (Johnson et al., 1993, Seifert et al., 2013).

1.4.1. Chromosomal translocation

Chromosomal translocations are one of the hallmarks of B-cell lymphomas (Biomberly et al., 2015). A translocation can be produced during VDJ recombination, SHM or class switching and frequently occurs at the IgH locus (Küppers and Dalla-Favera, 2001, Pasqualucci et al., 2001). *t*(14;18) translocations where the anti-apoptotic gene B-cell lymphoma 2 (BCL2), located on chromosome 18, becomes fused with the IgH locus on chromosome 14 have been observed in more than 85-90% of FL cases (Tsujiimoto et al., 1985, Jäger et al., 2000). The juxtaposed BCL2 gene is brought under the control of the immunoglobulin heavy-chain enhancer, leading to deregulation of BCL2 expression (McDonnell et al., 1989, Nogai et al., 2011). Another example occurs in 95% of mantle cell lymphoma (MCL) where the *t*(11;14) translocation juxtaposes the CCND1 gene (which encodes the cyclin D1 protein) with the IgH locus, and causes overexpression of cyclin D1. Cyclin D1 plays a role in the regulation of the cell cycle, and is normally not highly expressed in B-cells (Vaandrager et al., 1996, Seifert et al., 2013). Moreover, the same mechanism occurs in *t*(8;14) translocations, a hallmark of BL, where

the MYC gene on chromosome 8 (encodes a transcription factor) is translocated to a location close to the enhancers of the IgH locus on chromosome 14 (KuÈppers and Dalla-Favera, 2001, Burmeister et al., 2013). Furthermore, translocations in DLBCL involve the BCL6 gene in 30-40% of cases, which is located in the chromosome band 3q27. BCL6 is translocated with the IgH or IgL gene locus, or alternatively with more than 20 non-Ig locus genes, with the result being an active promoter fusing with BCL6 leading to constitutive BCL6 expression (Ye et al., 1993, Wlodarska et al., 1995, Chaganti et al., 1998, Keller et al., 2006). Deregulation of BCL6 causes failure of B-cell differentiation and continuous cell proliferation. Subsequently, it promotes a higher cell survival rate and genetic instability, that both contribute to malignancy (Dierlamm et al., 1999, Gouveia et al., 2012). Others are summarised in (Table 1-2).

1.4.2. Genetic modification of tumour suppressor genes

In the cases of genetic mutations that promote lymphomagenesis, 20% of post-GCB cell lymphoma cases occur through an inactivation mutation of CD95, also known as FAS. CD95 mediates apoptosis, a negative selection of B-cells within the GC, and acts as a tumour suppressor gene to prevent malignancies of B-cells (Müschen et al., 2002). Furthermore, a deletion mutation in the TP53 gene occurs in 20% of DLBCL cases and is associated with a more aggressive disease (Phan and Dalla-Favera, 2004, Lossos, 2005). Inactivation or deletion of PRDM1 in ABC-DLBCL and subsequent loss of Blimp-1 expression occurs in about 25% of cases (Pasqualucci et al., 2006, Tam et al., 2006, Nogai et al., 2011).

1.4.3. Aberrant somatic hypermutations

In the normal GC process, aberrant SHMs occur due to the off target activity of AID in the 5' regions of non-Ig genes such as BCL6, FAS/CD95 and CD79a (a component of BCR). This frequently occurs in DLBCL (Pasqualucci et al., 1998, Müschen et al., 2000, Khodabakhshi et al., 2012). Recently, in approximately

50% of DLBCL cases, aberrant SHMs were shown to occur in additional proto-oncogenes that are involved in cell proliferation, differentiation and signal transduction (MYC, PAX5 and PIMI). These are common exclusive features of DLBCL, which are absent in other B-NHLs (Lossos, 2005, Nogai et al., 2011, Seifert et al., 2013, Blombery et al., 2015).

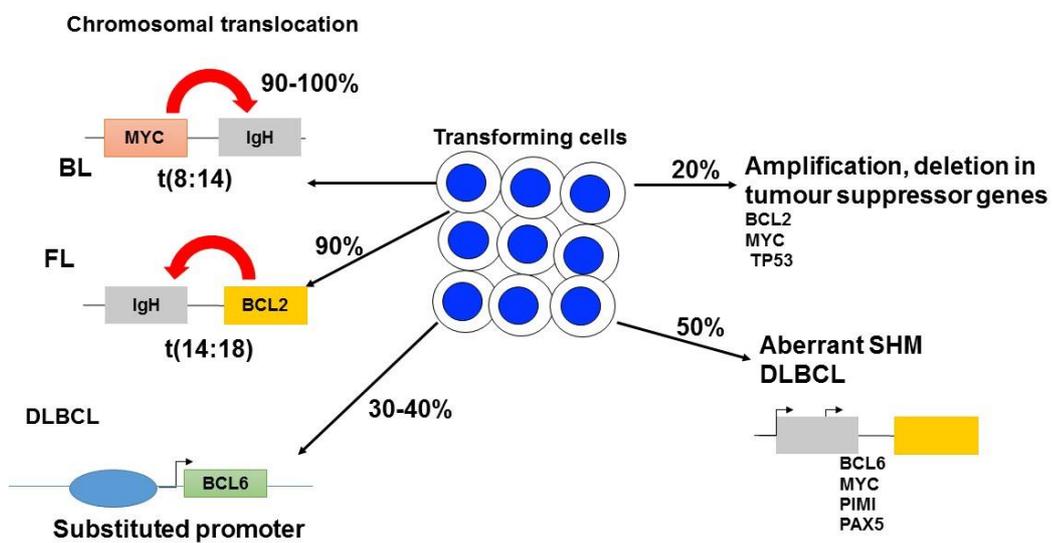


Figure 1-4 Schematic of the molecular pathogenesis of NHLs.

Lymphoma	Chromosomal translocation	Tumour suppressor gene mutations	Other alterations
MCL	CCND1-IgH (95%)	ATM (40%)	Deletion on 13q14 (50-70%)
FL	BCL2-IgH (90%)		
DLBCL	BCL6-various (35%) BCL2-IgH (15-30%) MYC-IgH or MYC-IgL (15%)	CD95 (10-20%), ATM (15%), TP53 (25%) BLIMPI (20% for ABC, EP300 10%, MLL2 (32%), TNFAIP3 (38%), CDKN2A(35%)	Aberrant hypermutation of multiple proto-oncogenes (50%)
FL	BCL2-IgH (90%)	MLL2 (89%) CREBBP (33%) EP300 (9%)	
BL	MYC-IgH or MYC-IgL (90-100%)	TP53 (40%), RB2 (20-80%)	
MALT lymphoma	API2-MALT1 (30%), BCL10-IgH (5%), MALT1-IgH (15-20%), FOXP1-IgH (10%)	CD95 (5-80%)	
Multiple myeloma	CCND1-IgH (15-20%), FGFR3-IgH (10%), MAF-IgH (5-10%)	CD95 (10%)	Various MYC alterations (40%), RAS mutations (40%), deletion on 13q14 (50%)

Mantle-cell lymphoma (MCL), follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL), burkitt's lymphoma (BL) and mucosa associated lymphoid tissue (MALT) lymphoma, this table adapted from (Küppers, 2005, Seifert et al., 2013).

Table 1-2 Mutations involved in B-cell non-Hodgkin's lymphoma pathogenesis.

1.5. Novel mutations identified in lymphoma using next generation sequencing (NGS) technologies

The next generation sequencing (NGS) technologies were developed in 2005, and have accelerated the understanding of the genetic alterations present in lymphoma disorders. Reasons for this better understanding include more global mutational surveys like the whole genome exome sequencing (WGS), as well as exome sequencing and RNA sequencing (RNA-seq) (Morin et al., 2008). Whilst some lymphomas, such as MCL, BL and FL, have long been characterised by the presence of certain chromosomal translocations, this is not the case for DLBCL which is a heterogeneous disorder (Morin and Gascoyne, 2013, Swerdlow et al., 2016). Using GEP and DNA copy number analysis helps us to understand the difference between the two molecular subtypes of DLBCL (Alizadeh et al., 2000, Morin and Gascoyne, 2013, Swerdlow et al., 2016). Using RNA-seq to screen a large sample population, in 29% of ABC-DLBCL cases, the amino acid substitution L265P in the myeloid differentiation primary response 88 (MYD88) gene was identified as being pathogenic, while this pathogenicity was absent in other types of NHL and GCB-DCBCL (Ngo et al., 2011, Morin and Gascoyne, 2013). Seventy-one somatic mutations are common in both DLBCL subtypes, including inactivation mutations in TP53, alterations of epigenetic regulators such as CREBBP/EP300 and oncogenic activation of BCL6 (Morin et al., 2011, Swerdlow et al., 2016). Also, in BL, Love et al., (2012) discovered novel mutations in ID3, a negative regulator gene and multiple other genes that are important in GC B-cells. Recently, studies have analysed genomic exome sequencing and RNA-seq of lymphoma samples from 50–100 patients and cell lines. Novel mutations were discovered in genes involved in either chromatin modification or transcriptional regulation. Also, novel point mutations were discovered in the anti-proliferative genes B-cell translocation 1 (BTG1) and B-cell translocation 2 (BTG2) (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013).

1.6. BTG/TOB Family

1.6.1. Discovery and definition

The anti-proliferative BTG/TOB family is composed of six members in mammalian cells: BTG1, BTG2, BTG3, BTG4, TOB1 and TOB2 (Winkler, 2010). Overexpression of BTG/TOB proteins can arrest cell cycle progression at the G1/S phase and inhibit cell proliferation in a variety of cell types (Rouault et al., 1992, Matsuda et al., 1996, Montagnoli et al., 1996, Ikematsu et al., 1999, Buanne et al., 2000). BTG2 was first reported in 1991 as a member of the BTG/TOB family, in two different laboratories, gene symbols approved by the HUGO Gene Nomenclature Committee. Their work found that the immediate early PC3 gene response in the rat neuronal cell line PC12, stimulated nerve growth factors, while in the mouse 3T3 fibroblast cell line, expression of the Tis21 gene was stimulated during treatment with 12-O-tetradecanoylphorbol-13-acetate-inducible (TPA) (Bradbury et al., 1991, Fletcher et al., 1991). The following year, the human B-cell translocation 1 gene (BTG1) was identified in a chromosomal translocation $t(8;12)(q24;q22)$ in chronic lymphocytic leukaemia (CLL) (Rouault et al., 1992). Several years later, transducer Of ERB2 (TOB1) was discovered because of its ability to bind the growth factor receptor P185erbB2 tyrosine-kinase (Matsuda et al., 1996). The other three members of the BTG/TOB family BTG3 (ANA), BTG4 (PC3B) and TOB2 were discovered later, based on their conserved sequence homology in the N-terminal domain (Guehenneux et al., 1997, Yoshida et al., 1998, Ikematsu et al., 1999, Tirone, 2001, Jia and Meng, 2007).

1.6.2. Structure of the BTG/TOB Family

The BTG/TOB family members share sequence homology in the BTG domain that is located in the N-terminal, 104-106 amino acids region. The domain contains two short conserved regions (box A and box B). Box A is known as growth regulatory, and box B plays a key role in anti-proliferative activity and

mediates the binding with molecular targets (Rouault et al., 1998, Guardavaccaro et al., 2000, Yang et al., 2008). The best-characterised role of the BTG domain is in protein-protein interactions, where it mediates and interacts with transcription factors and post-transcriptional gene regulation (Rouault et al., 1998, Ikematsu et al., 1999, Prévôt et al., 2001, Yoshida et al., 2001, Busson et al., 2005, Ou et al., 2007, Winkler, 2010). The C-terminal regions are less conserved among the family members, allowing their classification into three distinct subfamilies: the BTG1/BTG2 subfamily, the BTG3/BTG4 subfamily, and the TOB1 and TOB2 subfamily (Figure 1-5) (Winkler, 2010). BTG1 and BTG2 have a conserved domain known as box C, located in the C-terminal region and close to the BTG domain. This conserved domain allows interaction with the methyltransferase known as Protein Arginine Methyl Transferase 1 (PRMT1) (Lin et al., 1996). TOB1 and TOB2 have the largest C-terminal regions of the BTG/TOB family and also contain a motif known as the PABP-interacting motif 2 (PAM2), which mediates the interaction with cytoplasmic poly (A) binding protein 1 (PABPC1) (Ezzeddine et al., 2007, Funakoshi et al., 2007, Mauxion et al., 2008). Furthermore, BTG1 contains a single copy of the L XX LL motif, while BTG2 contains two copies of the L XX LL motif. This motif mediates binding with nuclear receptors and is involved in the regulation of estrogen receptor- α (ER α)-mediated activation (Heery et al., 1997, Prévôt et al., 2001).

The human BTG domain crystal structure has been published for BTG2 and TOB1 (Yang et al., 2008, Horiuchi et al., 2009), with high similarity in the conserved BTG domain (Winkler, 2010) (Figure 1-5). The crystal structure of human BTG2 is composed of five α -helices and four β -strands that form two anti-parallel β -sheets. The three long α -helices (α 1, α 2 and α 4) are a bundle followed by four β -strands, and two small α -helices (α 3 and α 5) inserted between strands β 1 and β 2. Box A is composed of helix α 3, part of the α 2 helix, strand β 1 and loops between them. Box B is composed of two anti-parallel β -strands (β 2 and β 3). Box C is composed of strand β 4 and the extended C-

terminal loop (Yang et al., 2008) (Figure 1-6).

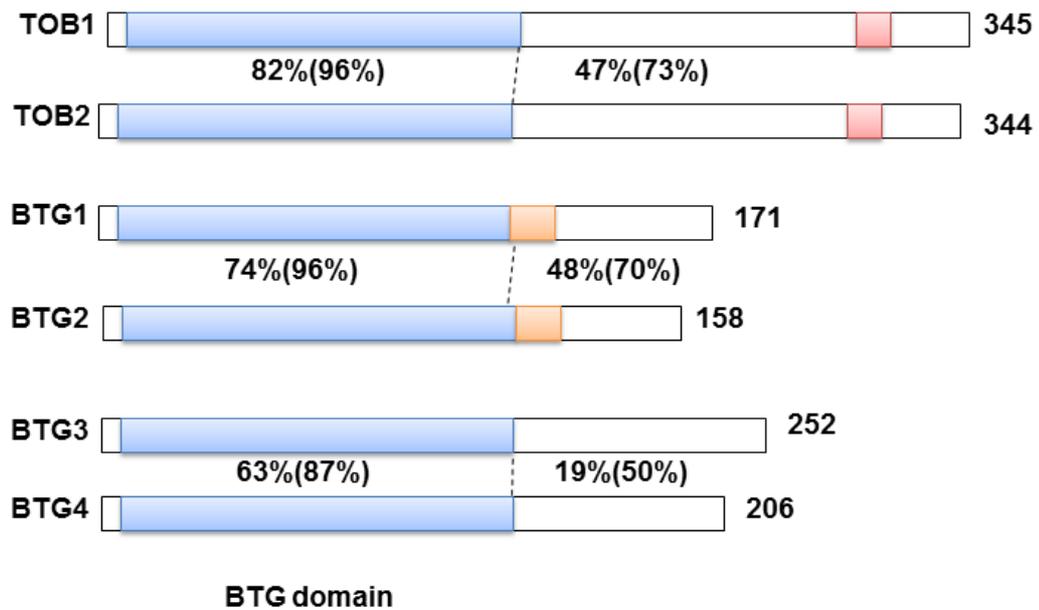


Figure 1-5 Schematic view of the BTG/TOB protein family.

The BTG/TOB family contains a conserved BTG domain in the amino terminus. The figure shows the percentage of identity (similarity) between the BTG domains and the C-terminal regions. The BTG domain, PAM2 motifs and box C are coloured in blue, red and orange, respectively (Winkler, 2010).

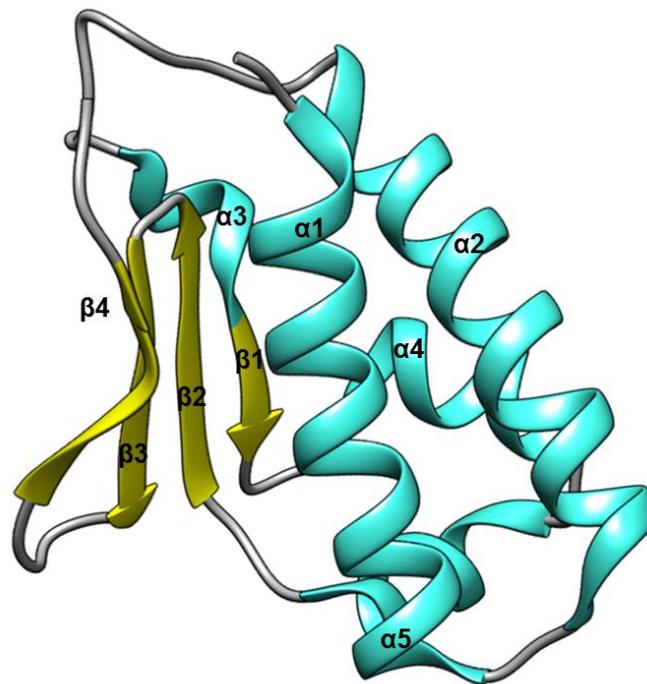


Figure 1-6 Structure of the BTG domain.

Schematic representation of the crystal structure of BTG2 (PDB accession number 3DJU)(Yang et al., 2008). Indicated are α -helix structures coloured sky-blue, the β -sheets coloured yellow and connecting loops which are coloured grey. UCSF Chimera (version 1.11.2) was used for visualization (Meng et al., 2006).

1.6.3. BTG/TOB proteins as effectors in signalling pathways

The anti-proliferative BTG/TOB members participate in a variety of signalling pathways that affect cell cycle progression, positively and negatively. In breast cancer cells (MCF-7) responding to estrogen or estradiol (E2) treatment, repression of ER α or estrogen receptor activity suppresses the expression of BTG2. As a result, the expression of cyclin D1 is upregulated and this induces the proliferation of breast cancer cells (Frasor et al., 2003, Karmakar et al., 2009). Conversely, cells treated with retinoic acid (RA) increased the transcriptional levels of BTG2 and inhibit the cell cycle progression (Passeri et al., 2006).

A conserved domain at the C-terminus of BTG1 and BTG2 interacts with PRMT1 (Lin et al., 1996). PRMT1 is involved in post-translational modifications and chromatin remodelling via the methylation of target proteins and the arginine 3 residue of histone 4 (Wang et al., 2001, Bedford and Clarke, 2009). Cells treated with retinoic acid (RA) induce the expression of BTG2, and results showed an increase in the methylation and acetylation levels of histone 4 at the retinoic acid receptor- β (RAR β) promoter, suggesting increased levels of the PRMT1 protein. Thus, it can be suggested that the BTG2-PRMT1 complex is released from the RAR containing this transcriptional complex and the nucleus exports this, which induces cellular differentiation (Passeri et al., 2006). Additionally, the interaction between BTG1 and BTG2 with PRMT1, contributes to inhibit the proliferation of the murine B-cell lymphoma cell line. Deletion of the box C regions in BTG1 impaired or arrested cell growth, which was deemed likely due to the absence of the interaction with PRMT1 (Hata et al., 2007).

It has been shown that BTG1 is a co-activator of positive regulators of myoblast differentiation, through mediating the transcriptional activity of the nuclear receptors such as TR α 1 and RAR α , and through direct interactions with the myogenic factor MyoD (Rodier et al., 1999, Busson et al., 2005). In line with

this, BTG1 expression was induced in primary mouse bone marrow by Forkhead box O3 (FoxO3A), which is involved in erythroid differentiation and blocked the outgrowth of erythroid colonies (Bakker et al., 2004). BTG1 and BTG2 interact directly with Hoxb9 in the conserved N-terminus that has been mapped (1-38 amino acid residues) and confirmed using the yeast two hybrid analysis (Prévôt et al., 2000). BTG2 enhances Hoxb9 binding with target DNA sequences to form the DNA-Hoxb9-BTG2 complex that was detected with gel-mobility shift experiments. This complex is involved in the cell's proliferation and differentiation capacity (Prévôt et al., 2000).

The BTG/TOB proteins regulate cellular proliferation by affecting the expression of cell cycle transcription factors (Guardavaccaro et al., 2000, Suzuki et al., 2002, Donato et al., 2007, Ou et al., 2007, Lim and Kaldis, 2013). BTG3 interacts directly with the E2F1 transcription factor and plays a critical role in the cell cycle during the S-phase, by activating genes important for G1/S progression, such as Cyclin E (Bracken et al., 2004, Ou et al., 2007). An overexpression of PC12 or PC3/BTG2 in mouse fibroblast NIH 3T3 and MCF7 cells, reduced the transcriptional level of cyclin D1 and the activity of cyclin-dependent kinase 4 (CDK4), leading to hypo-phosphorylation of the retinoblastoma protein (RB). This suggests that BTG2 inhibits cell cycle progression, dependent on the RB phosphorylation pathways (Guardavaccaro et al., 2000, Donato et al., 2007, Lim and Kaldis, 2013). Conversely, the absence of TOB1 protein activity caused a stimulation of cyclin D1 transcription, and allowed cells to enter S-phase (Suzuki et al., 2002). Fibroblast cells treated with growth factors induced the activity of mitogenic factors, such as Erk1 and Erk2, both of which phosphorylate TOB1 on three serine residues, which results in the inactivation of TOB1 and increase cyclin D1 levels (Suzuki et al., 2002).

TOB1 inhibits bone morphogenetic protein (BMP) signalling, which is a member of the TGF β family, by interacting with receptor-regulated Smads in osteoblasts, such as Smad 1, 5 and 8. This was observed in TOB1-null mice,

which had increased osteoblast activity leading to increased bone mass, in response to BMP signalling (Yoshida et al., 2000, Yoshida et al., 2003b). In contrast, overexpression of BTG2 enhanced the Smad-dependent signalling of BMP and mice harbouring a double or single deletion of BTG1 and BTG2 showed impaired BMP/Smad signalling, causing posterior homeotic transformations of the axial skeleton vertebrae (Park et al., 2004, Tijchon et al., 2015).

1.6.4. BTG/TOB involvement in mRNA turnover

The best characterised role of the BTG/TOB proteins is their involvement in the regulation of gene expression, mediated by the binding of the conserved BTG domain with human Caf1(hCaf1)(CNOT7/CNOT8) subunits of the Ccr4-Not complex. Experimentally, all BTG/TOB proteins can interact with high similarity paralogues hCaf1/CNOT7 and hCaf1/CNOT8, except BTG4 (Rouault et al., 1998, Ikematsu et al., 1999, Prévôt et al., 2001, Yoshida et al., 2001, Doidge et al., 2012a, Ezzeddine et al., 2012). In mammalian cells, the deadenylase activity of the Ccr4-Not complex is one of the fundamental steps involved in mRNA degradation, via the removal of the poly A tail at the 3' end of mRNA (Goldstrohm and Wickens, 2008, Collart, 2016). Direct interaction of the BTG/TOB proteins with hCaf1(CNOT7/CNOT8) suggests that this interaction recruits the human Caf1-Ccr4 Not complex onto particular mRNAs (Doidge et al., 2012a, Ezzeddine et al., 2012). Furthermore, TOB1, TOB2, BTG1 and BTG2 play additional roles through their interaction with PABPC1 (Ezzeddine et al., 2007, Funakoshi et al., 2007, Mauxion et al., 2008, Stupfler et al., 2016). This is discussed later in detail in section (1.8.6.3 interactions of BTG/TOB proteins).

1.6.5. Role of BTG/TOB proteins in cancer

The BTG/TOB proteins are involved in several tumours, with absent or low level expression is found in several types of tumours such as lung, liver, kidney,

and breast (Yoshida et al., 2003a, Struckmann et al., 2004, Ito et al., 2005, Kawakubo et al., 2006, Majid et al., 2009). In response to DNA damage, p53 increases the transcription of BTG2 and BTG3, suggesting a dependence on the p53 pathway (Rouault et al., 1996, Boiko et al., 2006, Ou et al., 2007). BTG2 expression was observed to arrest tumour cells at the G1/S transition checkpoint through an accumulation of RB hypo-phosphorylation, suggesting that BTG2 acts as a tumour suppressor that links p53 and RB pathways in human tumorigenesis (Cortes et al., 2000, Guardavaccaro et al., 2000, Boiko et al., 2006). The second mechanism of BTG2 acting as a tumour suppressor, is via p53-independent mechanisms, including NF- κ B activation (Duriez et al., 2002). Tumour necrosis factor- α (TNF- α) is a cytokine that strongly activates NF-kappa B signalling and consequently activates the expression of BTG2 in breast cancer cells (Kawakubo et al., 2004, Imran and Lim, 2013). The third mechanism is when Ras mediated transformation is impaired by TOB1 activity. Erk1 and Erk2, which are Ras signalling components, inactivate TOB1 via phosphorylation, leading to a loss of proliferative activity (Kawakubo et al., 2004). Mice that lack TOB1 develop spontaneous carcinogenesis at a higher rate than wild-type mice, and p53 and TOB1 knockout mice show accelerated tumour formation in comparison with single null mice, suggesting that both work synergistically to prevent tumour formation (Yoshida et al., 2003a).

Low transcription levels of BTG2 and BTG3 found in primary renal carcinoma tissues and cell lines, suggests that the absence of BTG2 and BTG3 is associated with tumour development (Struckmann et al., 2004, Majid et al., 2009). Moreover, low levels of BTG4 messenger RNA (mRNA) were found in colorectal cancer tissues, and upregulation of BTG4 expression reduced the colony formation ability of colorectal cancer (Toyota et al., 2008). In addition, low expression levels of TOB1 and BTG1 were found in clinical gastric cancer samples (Yu et al., 2011, Kanda et al., 2015).

In breast cancer cells, a loss of BTG2 expression was observed in ER α positive human breast tumours; an inverse correlation between BTG2 expression and

the size of the breast tumour was seen (Kawakubo et al., 2006). Also, a strong TOB1 expression in breast cancer patients showed a significantly shorter distant metastasis-free survival (Helms et al., 2009). Recently, using orthotopic metastasis assays, knockdown of TOB1 in mice was found to be associated with lung metastasis and suppressed lung colonization by 6DTT1 cells. Absence of TOB1 as an adaptor, which when present recruits hCaf1/CNOT7, exerts post-transcriptional gene regulation that promotes metastatic disease which requires deadenylase activity (Faraji et al., 2016) (Table 1-3).

Gene	Cancer	Causes /Alteration	Reference
TOB1	Lung	Spontaneous tumour formation§	(Yoshida et al., 2003a)
		Decreased expression/increased phosphorylation†	(Iwanaga et al., 2003)
TOB1	Lymph node	Spontaneous tumour formation§	(Yoshida et al., 2003a)
	Liver	Spontaneous tumour formation§	(Yoshida et al., 2003a)
	Breast	Reduced expression in human breast cancer cell lines	(Helms et al., 2009)
	Thyroid	Decreased mRNA expression†	(Ito et al., 2005)
	Pancreatic	Induced expression inhibits tumorigenesis in nude mice	(Yanagie et al., 2009)
BTG1	Gastric	Low expression associate with poor prognosis†	(Kanda et al., 2015)
BTG2	Breast	Reduced expression and relocalization (nuclear to cytoplasm) †	(Kawakubo et al., 2006)
	Prostate	Low/undetectable mRNA levels†	(Ficazzola et al., 2001)
	Brain	Induced expression inhibits medulloblastomas (transgenic mice)	(Farioli-Vecchioli et al., 2007)
BTG3	Lung	Increased lung tumour formation§	(Yoneda et al., 2009)
		Reduced expression in adenocarcinoma samples†	(Yoneda et al., 2009)
	Renal	Reduced mRNA expression†	(Majid et al., 2009)
BTG4	Colon	Reduced mRNA expression†	(Toyota et al., 2008)

§ Observations made using mouse knock-out models. † Observation made using human clinical cancer samples and biopsies. To date, no changes in TOB2 expression are reported in human cancer samples.

Table 1-3 Relationship of BTG/TOB expression and cancer.

1.6.6. BTG1 and BTG2 proteins in haematological malignancies

BTG1 was first identified in a patient with CLL harbouring the translocation $t(8;12)(q24;q22)$ (Rouault et al., 1992). Also, a high frequency of BTG1 deletions occur in cases of acute lymphoblastic leukaemia (ALL) (Cho et al., 2004, Kuiper et al., 2007, Tsuzuki et al., 2007, Waanders et al., 2012, Xie et al., 2014, La Starza et al., 2016, Scheijen et al., 2016, Bhandari et al., 2017, Patel et al., 2017). Moreover, deletions in the BTG1 gene are a biomarker of paediatric B-cell precursor- ALL (BCP-ALL) and these occur predominantly in two cytogenetic subgroups: ETV6-RUNX1 and BCR-ABL1 (Tsuzuki et al., 2007, Waanders et al., 2012, Xie et al., 2014). One result found a deletion clustered in the second and last exon of the BTG1 gene, suggesting that illegitimate RAG-mediated recombination plays a role in BTG1 deletions, and further suggesting that the RSS's hotspot similarities are present in exon2 of BTG1 (Waanders et al., 2012). Recently, it has been found that deletion alone in the BTG1 gene had no impact on a patient's prognosis; however, a combination of a deletion of BTG1 and IKZF-1 was associated with poor prognosis and increased relapse risk in BCP-ALL patients (Scheijen et al., 2016). IKZF-1 is one of the key regulatory transcription factors in B-cell development, along with EBF1 and PAX5 (Busslinger, 2004). Studies over a five year period found that genetic alteration of these factors is predominant in BCP-ALL cases (Mullighan et al., 2007, Somasundaram et al., 2015). In addition, deletion of the BTG1 gene affects the co-activation of nuclear receptors, causing resistance to glucocorticoid (GC) gene expression. BTG1 interaction with PRMT1 increases expression of the glucocorticoid receptor (GR). PRMT1 is recruited to the GC-responsive promoter that regulates GR expression, in a BTG1-dependent manner (van Galen et al., 2010). Detection of BTG1 has been suggested as a biomarker for acute myeloid leukaemia (AML) patients in complete remission, and undergoing treatment with chemotherapy (Cho et al., 2004).

Over the last seven years, somatic mutations in BTG1 and BTG2 have been discovered in lymphoma. Studies were performed on lymphoma patient

samples and cell lines, using whole genome sequencing or exome DNA-seq and RNA-seq (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013, Fukumura et al., 2016). Gene mutations affecting either chromatin modification or transcriptional regulation were identified, including novel point mutations in BTG1 and BTG2. However, the nature of those mutations did not clearly point out the effectiveness of their functions, nor mechanisms for gains or losses of function (Morin et al., 2011). Zhang et al., (2013) suggested that AID-related mutations act as a major mechanism in BTG1, CD79B and PIM1 genes, and thus may be a lymphoma-specific mechanism. Another paper found that mutations in BTG2, CD79B, MYD88 and PIM1 are major targets of aberrant SHM in primary central nervous system lymphoma (Fukumura et al., 2016). This suggests that BTG1 and BTG2 are required for proper B-cell development, and the loss of BTG1 function is largely related to a defect in the growth of B lymphoid progenitors (Tijchon et al., 2016).

Furthermore, low expression of BTG1 coupled with the down regulation or loss of function of CNOT6, has been suggested to be implicated in the pathogenesis of a subset of HOXA positive T-ALL cases with terminal 5q deletions (La Starza et al., 2016). CNOT6 encodes a member of the Ccr4-Not complex and BTG1 is involved with the Ccr4-Not complex, by interacting with the hCaf1(CNOT7/CNOT8) subunit (Prévôt et al., 2001, Collart and Panasenko, 2012).

1.7. Eukaryotic mRNA turnover

The modulation of mRNA function within the cytoplasm plays a critical role in controlling gene expression. The function is primarily determined by the balance between transcript degradation and translation. mRNA decay is an important and fundamental step for the modulation of gene expression and quality control over mRNA biogenesis (Wu and Brewer, 2012). The stability of mRNA is influenced by the presence of the 7-methylguanosine (m^7G) cap at

the 5' terminus, and the poly (A) tail at the 3' terminus. The 5' m⁷G cap binds with a cytoplasmic protein called eukaryotic initiation factor 4E (eIF4E), while the 3' poly (A) tail interacts with PABPC1. Both binding proteins protect the transcript from exonucleases and promote translation initiation through the recruitment of mRNA to ribosomes (Garneau et al., 2007). Eukaryotic mRNA turnover follows two general pathways, starting with shortening of the poly (A) tail by a variety of mechanisms, including mRNA deadenylases. Then, Lsm 1-7 proteins bind to the 3' end of the deadenylated products and assist the decapping enzyme, which is composed of two subunits: Dcp1 and Dcp2. Both enzymes remove the 5' cap structures and the transcript is exposed to 5' → 3' mRNA exonuclease such as Xrn1. Alternatively, the mRNA body can be degraded after the deadenylation step in the 3' → 5' direction by a cytoplasmic exosome (Parker and Song, 2004, Garneau et al., 2007, Goldstrohm and Wickens, 2008, Wu and Brewer, 2012, Wahle and Winkler, 2013) (Figure 1-7). There are three classes of mRNA degradation enzymes: endonucleases that cut RNA internally, the 5' end exonucleases that hydrolyse RNA from 5' → 3' (such as Xrn1) and 3' end exonucleases which degrade mRNA as an exosome (Table 1-4). In the case of the endonuclease enzymes, the cleavage occurs either by sequence-specific endonucleases, or in response to microRNAs (miRNAs), or small interfering RNAs (siRNAs) (Parker and Song, 2004, Houseley and Tollervey, 2009).

In the deadenylation process, the poly (A) tail of an mRNA is degraded by the deadenylase activity of Mg²⁺-dependent exoribonucleases. These exoribonucleases use the poly (A) tail as their main substrate and hydrolyse RNA in the 3' → 5' direction, producing 5'-adenosine monophosphate (Goldstrohm and Wickens, 2008). There are two main groups based on the characteristics of the nuclease domain. The first group, called DEDD-type nucleases, encompasses the Asp and Glu residues in the exonuclease motif, and two metal ions are required for their activity. There are several subfamilies under this group, such as RNase D, RNase T and oligoribonuclease. The most

common deadenylase enzymes in this group include POP2 (also known Caf1/CNOT7, Caf1/CNOT8, TOE1/Caf1z), poly (A) specific ribonuclease (PARN), and Pan2 (Zuo and Deutscher, 2001, Goldstrohm and Wickens, 2008). The second group of deadenylases belongs to the exonuclease-endonuclease-phosphatase (EEP) deadenylase family, which contains the conserved catalytic residues Asp and His in their nuclease domains. They cooperate with metal ions for their catalytic activity and includes Ccr4/CNOT6, Ccr4/CNOT6L, Nocturnin and Angel (Dlakić, 2000, Zuo and Deutscher, 2001)(Table 1-4).

In mammals, there are two major and distinct enzymatic complexes involved in the deadenylation process: the initial step is performed by the Pan2-Pan3 complex, followed by the Ccr4-Not complex (Parker and Song, 2004, Goldstrohm and Wickens, 2008) (Figure 1-7). The recruitment of the Pan2-Pan3 complex to the target the mRNA occurs via the binding of Pan3 with PABPC1 through the PAM2 motif, and this interaction stimulates the activity of Pan2 (Garneau et al., 2007, Wahle and Winkler, 2013).

Protein	Function	Human names
Deadenylase enzyme	GroupI (DEDD-type deadenylases)	Caf1(CNOT7/CNOT8) TOE1 (Caftz) PARN Pan2
	GroupII (EEP-type deadenylases)	Ccr4(CNOT6/CNOT6 L) Nocturnin Angel Angel 2
Decapping enzyme	Members of decapping enzyme	Dcp1a, Dcp1b Dcp2
	Lsm RNA binding complex containing Lsm subunits 1 to 7	Lsm 1-7
Exoribunuclease enzymes	Cytoplasmic 5' exoribonuclease	Xrn1
	Multi subunit exonuclease containing two enzymatic subunits: Exosome 3' hydrolytic exonuclease	Exosome

Table 1-4 Enzymes involved in cytoplasmic mRNA degradation.

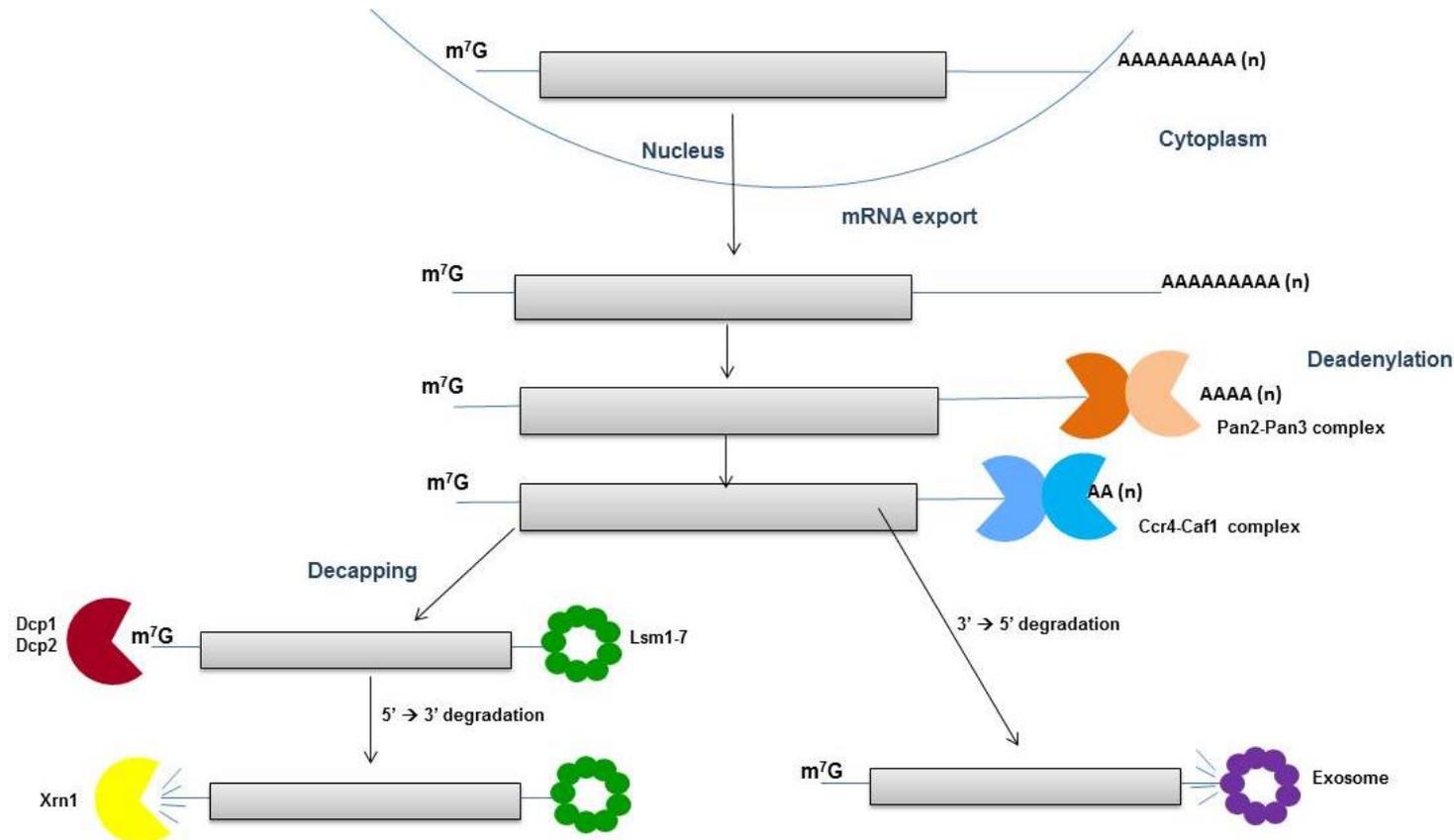


Figure 1-7 Schematic diagram of the canonical eukaryotic mRNA degradation pathway.

Deadenylation is biphasic with the first step catalyzed by the Pan2– Pan3 complex (orange); the second step is mediated by the Ccr4–Not complex (blue). After deadenylation, the mRNA is both decapped by removal of the 7-methylguanylate (m⁷G) group and then degraded rapidly in the 5' to 3' direction by Xrn1, or mRNA can be degraded in the 3' to 5' direction by the exosome.

1.8. The Ccr4-Not complex

1.8.1. Overview of Ccr4-Not complex

The Ccr4-Not complex consists of several subunits: the catalytic subunits and non-catalytic subunits. The complex is conserved from yeast to humans. In yeast; first, the catalytic module contains two components that are associated with deadenylase activity: carbon catabolite repression 4 (Ccr4) and Ccr4-associated factor (Caf) 1. Second, non –catalytic module is composed of five subunits: NOT1, NOT2, NOT3, NOT4 and NOT5. Two additional subunits: Caf40 and Caf1130; Caf1130 is only present in yeast (Tucker et al., 2001, Denis and Chen, 2003). The structure of the complex in yeast has an estimated molecular weight of 1.MDa for the complex, and it is L- shaped with Caf1 and Cc4 subunits located in the hinge connection between the two arms (Nasertorabi et al., 2011).

In humans, the catalytic subunits CNOT6 and CNOT6L are homologues of Ccr4, and CNOT7 and CNOT8 are orthologues of Caf1. The non-catalytic subunits are CNOT1 (NOT1), CNOT2 (NOT2), CNOT3 (NOT3/NOT5), CNOT4 (NOT4) and CNOT9 or RQCD1 (Caf40) (Liu et al., 1998, Chen et al., 2001, Collart, 2003, Lau et al., 2009, Collart and Panasenko, 2012, Winkler and Balacco, 2013). CNOT4 is present in a separate complex, and is not associated with the human Ccr4-Not complex (Lau et al., 2009). Two additional subunits are missing in yeast and are specific to humans: CNOT10 and CNOT11 (C2orf29) (Lau et al., 2009) (Table 1-5).

Yeast	Human	Domain/motifs	Other names
NOT 1	CNOT1	LxxLL, HEAT repeats, MIF4G and DUF	Human Not1
NOT 2	CNOT2	NOT box	Human Not2
NOT 3	CNOT3	NOT box	Human Not3
NOT 4	CNOT4	RING	Human Not4
NOT 5	CNOT5	Homology with Not3	
Ccr4	CNOT6/CNOT6L	LRR; EEP	Human Ccr4
Caf1	CNOT7/CNOT8	DEDD	Human Caf1
Caf1/Pop2	CNOT7/CNOT8	DEDD	Calif, human Pop2
Caf40	RQCD1	Armadillo repeats	Rcd1, CNOT9
Caf130	-	-	-
	CNOT10	-	
	CNOT11	-	C2ORF29

Table 1-5 Ccr4-Not components in yeast and human cells.

1.8.2. Definition and structure of the Ccr4-Not complex

1.8.2.1. Catalytic subunits

The nuclease module of the Ccr4-Not complex contains two distinct deadenylase enzymes, known in yeast as Ccr4 and Caf1 (Pop2). Two orthologues of each enzyme exist in humans: Ccr4 homologues are the two subunits CNOT6 and CNOT6L (78% identity and 88% similarity at the amino acid sequence level), whereas the Caf1 homologues are the two subunits CNOT7 and CNOT8 (76% identity and 89% similarity at the amino acid sequence level) (Dupressoir et al., 2001, Collart, 2003, Aslam et al., 2009, Winkler and Balacco, 2013) (Figure 1-8).

Caf1 contains an RNase D domain, which belongs to the Asp-Glu-Asp-Asp (DEDD) super family of proteins, associated with ribonuclease activity (Collart and Panasenko, 2012, Winkler and Balacco, 2013). Caf1 binds directly to the Ccr4-Not complex, but only one of these DEDD-type nucleases can interact at a time (Lau et al., 2009). In 2003, the crystal structure of Caf1 was reported from *Saccharomyces cerevisiae* (Thore et al., 2003), and several years later was discovered in *Schizosaccharomyces pombe* (Jonstrup et al., 2007, Andersen et al., 2009). In the crystal structure from *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and humans, the Caf1 tertiary structures take a kidney shape, and the central core is composed of β sheets surrounded by α -helices. In the active site, Caf1 contains two Mg^{2+} metal ions, which are required for catalytic activity, and are coordinated with glutamate and three aspartate residues. Amino acid substitutions of these active residues abolished the enzymatic activity (Thore et al., 2003, Jonstrup et al., 2007, Horiuchi et al., 2009). The model of hCaf1/CNOT7-RNA was proposed using the superposition of the PARN enzyme with poly (A) and DEDD domains of hCaf1/CNOT7 (Winkler, 2010). The poly A tail binding site of Caf1 has a different shape and is wider compared to the substrate binding pocket of Ccr4 (Winkler, 2010, Winkler and Balacco, 2013). The crystal structure of the

human CNOT1 (middle domain of (eukaryotic initiation factor 4G (MIF4G)) domain has been reported, in complex with hCaf1/CNOT7. This crystal structure shows that the MIF4G domain of CNOT1 binds on the opposite site of the catalytic site of hCaf1/CNOT7, which allows RNA substrates to bind (Petit et al., 2012).

Ccr4 (CNOT6/CNOT6L) deadenylase enzyme belongs to the EEP super family. The enzymatic activity of Ccr4, analysed in humans and in yeast, indicated that it strongly preferred poly A tail residues (Chen et al., 2002, Wang et al., 2010). The protein structure contains two highly conserved domains: an amino-terminal leucine-rich repeat (LRR) and C-terminal deadenylase domains. The five tandem copies of LRR domains provide surface interaction with Caf1 subunits, and this interaction is required to stabilize the binding with the Ccr4-Not complex. The ribonuclease activity is located in the C-terminal region that contains the EEP domain (Chen et al., 2002, Collart, 2016). The human crystal structure of the Ccr4 nuclease domains in the C-terminus, form a two layer α/β sandwich fold, which has significant similarity with another hydrolyse called apurinic/aprimidinic endonuclease (APE) 1 (Wang et al., 2010). Two metal Mg^{+2} ions located at the active site, coordinate with the five conserved active residues: asparagine, glutamate, two aspartates and histidine. Substitution of the active residues caused abolishment of Ccr4 catalytic activity (Chen et al., 2002, Wang et al., 2010).

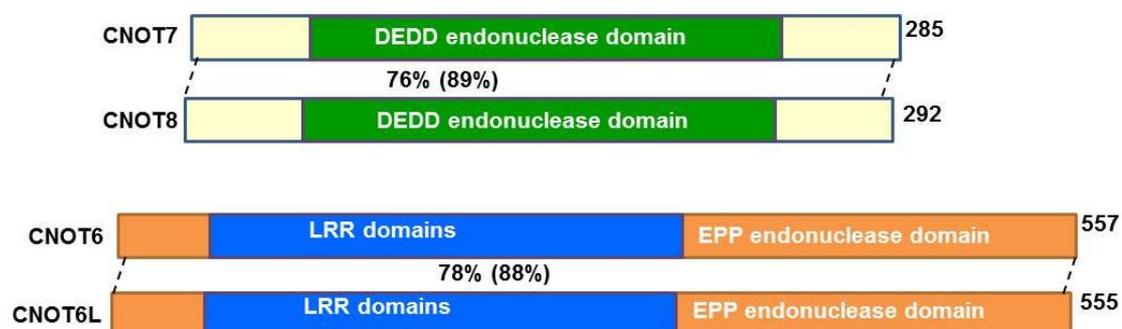


Figure 1-8 Schematic overview of human catalytic subunits of Ccr4-Not complex.

The schematic representation of the Ccr4-Not deadenylases highlights conserved domains, DEDD endonuclease domain (green), EPP endonuclease domain (orange), LRR domains (blue). Also indicated are the size of the proteins and percentage of identical (similar) amino acids is also shown.

1.8.2.2. Non-catalytic subunits

The large Not1 (CNOT1) subunit is a scaffold protein located in the centre of the complex, to which several modules of Ccr4-Not subunits are attached (Bai et al., 1999). CNOT1 contains three main regions: the amino terminus, carboxyl terminus and the middle region (Bai et al., 1999, Boland et al., 2013, Xu et al., 2014) (Figure 1-9).

In the middle region, CNOT1 contains two domains: the MIF4G and a domain of unknown function called (DUF3819). The nuclease module is anchored to the central MIF4G domain of NOT1, and Caf1 acts as an adaptor to stabilize the interaction of the LRR domain of Ccr4 to the complex (Bai et al., 1999, Dupressoir et al., 2001, Basquin et al., 2012, Petit et al., 2012). Another module of the Ccr4-Not complex, termed CNOT9/RQCD1/Caf40, is attached to the DUF3819 domain of CNOT1 in the central region. This domain is also termed the CNOT9 binding domain (CN9BD) (Bawankar et al., 2013).

In the amino terminal region, CNOT1 largely contains α -helices and HEAT (Huntingtin, elongation factor 3 (EF3), protein phosphatase 2A and yeast kinase Tor1 repeats) domains that are attached to CNOT10 and CNOT11 (Bawankar et al., 2013, Mauxion et al., 2013). CNOT4 (Not4) has similar structures in yeast and in humans. However, there is a difference in protein stability and interactions with CNOT1. An *in vitro* study confirmed that the function of CNOT4 is as an E3 ligase in ubiquitination (Lau et al., 2009).

The C-terminal region of CNOT1 contains the NOT box domain, which is attached to the NOT domains of CNOT2 (Not2) and CNOT3 (NOT3/NOT5) (Bai et al., 1999, Boland et al., 2013). Recently, the human crystal structures of CNOT1, CNOT2 and CNOT3 have been reported. CNOT2 and CNOT3 interact with each other through their conserved domains, called NOT boxes, present in the C-terminal regions and interact via their NAR (Not Anchor region) with CNOT1 (Boland et al., 2013).

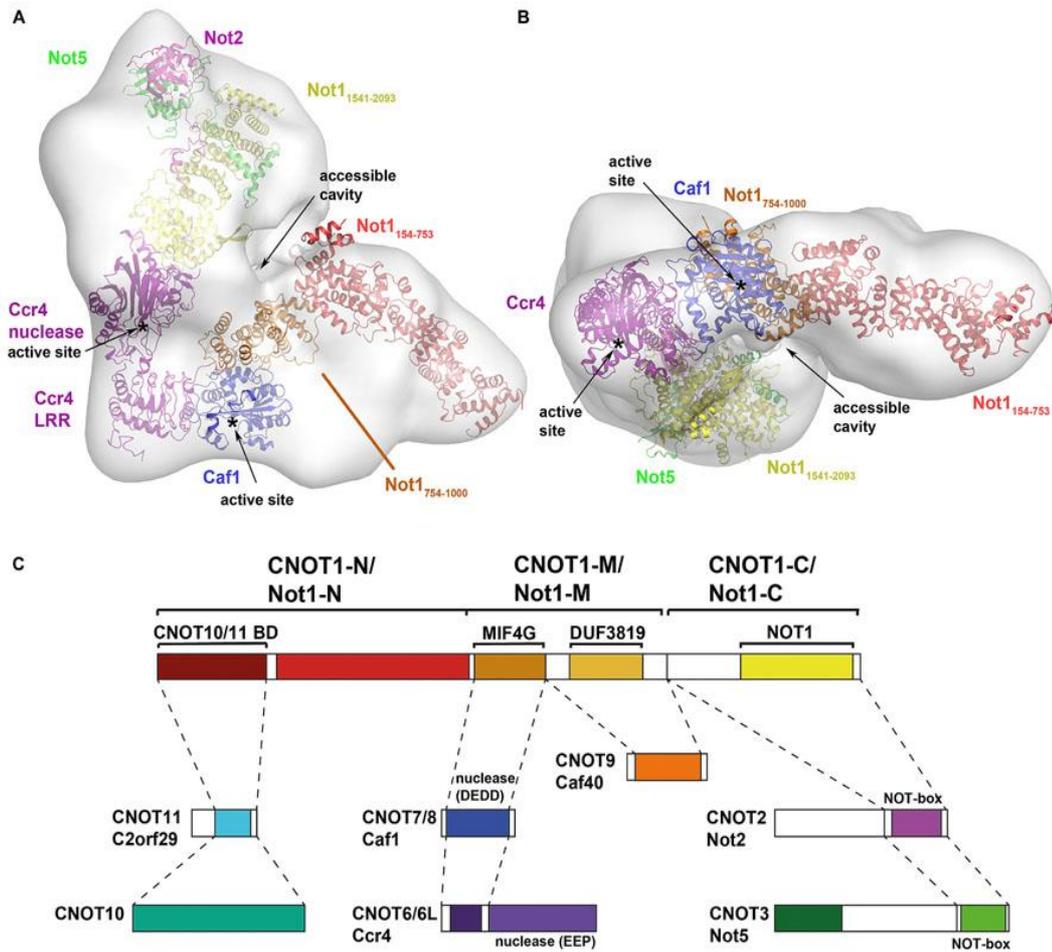


Figure 1-9 Model structure of the mammalian Ccr4-Not complex.

(A) and (B) Ccr4-Not complex using electron microscopy (EM) and image reconstruction. (C) Interactions maps between CNOT1 with other subunits of Ccr4-Not complex. Figure as originally published in Xu K, Bai Y, Zhang A, Zhang Q and Bartlam MG (2014) Insights into the structure and architecture of the CCR4-NOT complex *Front. Genet.* 5:137. doi: 10.3389/fgene.2014.00137. The figure was reproduced under the terms of the Creative Commons Attribution License, (CC-BY 3.0).

1.8.3. Physiological function of the Ccr4-Not complex

The multifunctional role of the Ccr4-Not complex is to regulate homeostasis in mammals, and to disrupt functions associated with disease (Shirai et al., 2014). Depletion of the central scaffold (CNOT1) of the Ccr4-Not complex destabilizes the complex, as it has been shown that knockout of CNOT1 in HeLa cells compromises deadenylase activity. Results showed that the number of P-bodies within cells is reduced where mRNA decay takes place (Ito et al., 2011b). Indeed, the interaction between CNOT1 and Nanos2 plays an important role in the development of the murine male germ cell, via regulation of the mRNA decay of a specific target gene (Suzuki et al., 2012). Heterozygous deletion of CNOT3 in mice showed lean body types, and resistance to a high fat diet, while obesity was induced in the wildtype mice. Also, the reduction of the expression of CNOT3 affects the stabilization of an mRNA coding for energy and metabolism related components, suggesting that CNOT3 recruits the Ccr4-Not deadenylase to the target mRNA (Morita et al., 2011). It was proposed that the Ccr4-Not complex including CNOT1, CNOT2, Ccr4 (CNOT6/CNOT6L) plays a role in cellular growth and survival (Morita et al., 2007, Ito et al., 2011a, Ito et al., 2011b, Mittal et al., 2011). Bone mass was reduced in CNOT7 knockout mice, suggesting that the Ccr4-Not subunits play a role in the regulation of bone formation (Washio - Oikawa et al., 2007).

1.8.4. Ccr4-Not complex involved in cancer

Several reports showed the importance of the expression of the Ccr4-Not complex in cancer cell lines, patient samples, and as a regulator of metastasis (Aslam et al., 2009, Maragozidis et al., 2012, De Keersmaecker et al., 2013, Faraji et al., 2014, Vicente et al., 2015, Faraji et al., 2016, La Starza et al., 2016). In breast cancer cells, CNOT2 is the central network of the transcriptional module, and down regulation of CNOT2 was reported to enhance metastasis (Faraji et al., 2014). Moreover, CNOT7 and CNOT8 are required for cell cycle

progression in the human breast cancer cell line MCF-7, via the repression of an anti-proliferative gene, namely PMP22 by deadenylase activity (Aslam et al., 2009). Through a bioinformatics analysis of hCaf1/CNOT8 and TOB1 as potential modulators of metastatic disease, the role of the Ccr4-Not complex in mammary tumour metastasis suggests that the deadenylation activity of Ccr4-Not complex might be a critical determinant. The hCaf1/CNOT8 and TOB1 were associated with distant metastasis free survival in human patients (Faraji et al., 2014). A recent study showed that a high expression of hCaf1/CNOT7 controlled metastasis in tumour cells. Furthermore, the deadenylase activity of hCaf1/CNOT7 is promoted in the presence of the interaction with CNOT1 and TOB1. Collectively the RNA-binding proteins, CNOT7, TOB1 and CNOT1 control a metastasis suppressive transcriptional program to drive tumour cell metastasis by using post-transcriptional (Faraji et al., 2016).

In haematological malignancies, frequent mutations or down regulation of expression were identified in the Ccr4-Not subunit genes. For example, mutations have been identified in the CNOT3 and human Ccr4/CNOT6 genes in T-ALL samples (De Keersmaecker et al., 2013, Vicente et al., 2015, La Starza et al., 2016) and the expression of the deadenylase subunits hCaf1/CNOT7 and human Ccr4/CNOT6L, were reported to be down regulated in ALL and AML samples (Maragozidis et al., 2012).

1.8.5. Role of Ccr4-Not complex in mRNA decay

The catalytic activity of deadenylase enzymes is directed onto their substrate poly (A) tail, which stimulates mRNA degradation. *In vitro*, the catalytic activity of deadenylase enzymes is independent of other Ccr4-Not components (Chen et al., 2002, Wang et al., 2010, Maryati et al., 2015). In humans, the Ccr4-Not complex contains one EEP-type subunit, either CNOT6 or CNOT6L, and one DEDD-type subunit, either CNOT7 or CNOT8. In human cells, two deadenylases are found, one from each type (Lau et al., 2009). In human cell lines where

knockdown of one of the subunits of the subfamily of either human Caf1/CNOT7 or hCaf1/CNOT8 had been successfully achieved, microarray results were largely identical when compared to each other, for both the positively and negatively regulated gene expression sets. In addition, the knockdown of either the human Ccr4/CNOT6 or Ccr4/CNOT6L, resulted in an identical overall gene expression set. In comparison, the results of the types of genes regulated in both subfamilies show little overlap in the regulation of mRNA abundance. Therefore, the mechanisms of both enzymes do not cooperate in the regulation of mRNA, nor in targeting a different mRNA (Aslam et al., 2009). Furthermore, the deadenylase activity of both Caf1 and Ccr4 is cooperative, being more active together than either in isolation. Interestingly, trimeric BTG2-Caf1-Ccr4 complexes display even higher activity compared with Caf1-Ccr4 (Maryati et al., 2015).

1.8.6. Regulation and recruitment of the Ccr4-Not complex

The deadenylase activity of the Ccr4-Not complex does not extensively degrade the cytoplasmic poly (A) tails. Several proteins interact with Ccr4-Not subunits, and recruit its activity to particular mRNA molecules. Thus, mediating the Ccr4-Not complex into specific targeted mRNA can be accomplished via the RNA binding proteins and components involved in miRNA mediated repression (Doidge et al., 2012b).

1.8.6.1. The role of the Ccr4-Not complex in miRNA-mediated gene silencing

miRNA induces the silencing complexes that repress translation and degrade the mRNA targets. The Ccr4-Not deadenylase subunits are associated with the miRNA silencing complex through deadenylation of the target mRNA, in human cells (Piao et al., 2010). To utilize the miRNA function, it must associate with Argonaute (Ago) proteins that interact with the GW182/TNRC6 protein (Zekri et al., 2013). In human and *Drosophila melanogaster* cells, the

GW182/TNRC6-mediated recruitment of the CNOT1-Ccr4-NOT complex occurs through conserved GW or WG motifs (either Gly/Ser/Thr-Trp or Trp-Gly/Ser/Thr) that are located in the N-terminal and C-terminal effector domains of GW182/TNRC6 proteins (Chekulaeva et al., 2011, Fabian et al., 2011) (Figure 1-10 A). In addition, RQCD1/CNOT9 contains GW-binding motifs that interact with GW182, and assist in the recruitment of the CCR4-NOT complex to miRNA targets (Chen et al., 2014).

1.8.6.2. Role of RNA-binding proteins

Tristetraprolin mediates a rapid degradation of RNA containing adenylate-uridylylate-rich elements (AU-rich elements; AREs). Tristetraprolin acts as bridge to recruit the Ccr4-Not complex via interacting with the scaffold protein CNOT1, which mediates a rapid decay to ARE-containing mRNAs (Sandler et al., 2011) (Figure 1-10 B).

TOB1 recruits hCaf1(CNOT7/CNOT8) on specific sequences of mRNA for degradation processing. The carboxyl terminal of TOB1 interacts with cytoplasmic polyadenylation element binding 3 (CPEB3) proteins. The CPEB3 protein binds directly to a U-rich RNA sequence. Recruitment of CPEB3 to the 3' UTR of mRNA results in deadenylation and mRNA turnover, which is dependent of both the presence of TOB1 and hCaf1(CNOT7/CNOT8). This suggests that hCaf1(CNOT7/CNOT8) is mediated into the target mRNA and stimulates their degradation via RNA binding proteins and TOB1 (Hosoda et al., 2011) (Figure 1-10 C).

1.8.6.3. Interaction with BTG/TOB proteins

The anti-proliferative BTG/TOB proteins directly interact with hCaf1(CNOT7/CNOT8) via the BTG domain (Rouault et al., 1998, Ikematsu et al., 1999, Prévôt et al., 2001, Yoshida et al., 2001). TOB/BTG is implicated as a general activator of deadenylation, by mediating the Ccr4-Not complex

(Ezzeddine et al., 2007, Funakoshi et al., 2007, Mauxion et al., 2008, Doidge et al., 2012a, Ezzeddine et al., 2012). Additionally, TOB1 and TOB2 contain conserved PAM2 motifs in the C-terminal, which allows them to bind directly with the RNA binding protein PABPC1. This suggests a role for TOB1 as an adaptor protein that mediates the recruitment of the Ccr4-Not complex to mRNA (Ezzeddine et al., 2007, Funakoshi et al., 2007). Stimulation of the deadenylation process occurs through competitive interactions between the PAM2 motifs of eukaryotic release factor (eRF3), Pan3 and TOB1, with PABPC1 (Uchida et al., 2002, Fabian et al., 2011). Moreover, BTG1 and BTG2 were shown to interact with PABPC1 through the BTG domain, specifically in the box C region, suggesting this interaction is required for recruitment of hCaf1(CNOT7/CNOT8) onto the target mRNA (Stupfler et al., 2016) (Figure 1-10 D). It has been shown that when TOB1 and TOB2 lack PAM2 motifs, the deadenylation function of hCaf1(CNOT7/CNOT8) is compromised (Ezzeddine et al., 2007, Funakoshi et al., 2007).

In 2009, the structure of the BTG domain of TOB1-hCaf1/CNOT7 complex was reported. The complex contains box A and box B in the N-terminal 138 residues of TOB, named TobN138 (PDB accession number 2D5R) (Horiuchi et al., 2009) (Figure 1-11). By examining the crystal structures of the TOB1-hCaf1/CNOT7 complex, and the related DEDD-type deadenylase PARN bound to RNA, it can be predicted that the N-terminal domain of TOB1 binds to hCaf1/CNOT7 away from its active site (Winkler, 2010). Presence of the BTG2 proteins increased the Caf1-Ccr4 deadenylation activity more than when BTG2 proteins are absent, which demonstrates the importance of the BTG2-Caf1 interaction (Maryati et al., 2015). Also, experimental studies identified the important residues of BTG2, TOB1 and TOB2, in interaction with hCaf1(CNOT7/CNOT8), such as BTG2W103 and TOB1 (W93, L32, L35, and L36). They are required for mRNA decay via binding with hCaf1(CNOT7/CNOT8) (Doidge et al., 2012b, Ezzeddine et al., 2012). Moreover, BTG/TOB proteins are required for the hCaf1(CNOT7/CNOT8)

interaction to inhibit cell proliferation, suggesting that the BTG/TOB proteins recruit hCaf1(CNOT7/CNOT8) into target mRNA, and suppress the gene expression of the transcription factors that are involved in cell proliferation (Doidge et al., 2012b, Ezzeddine et al., 2012).

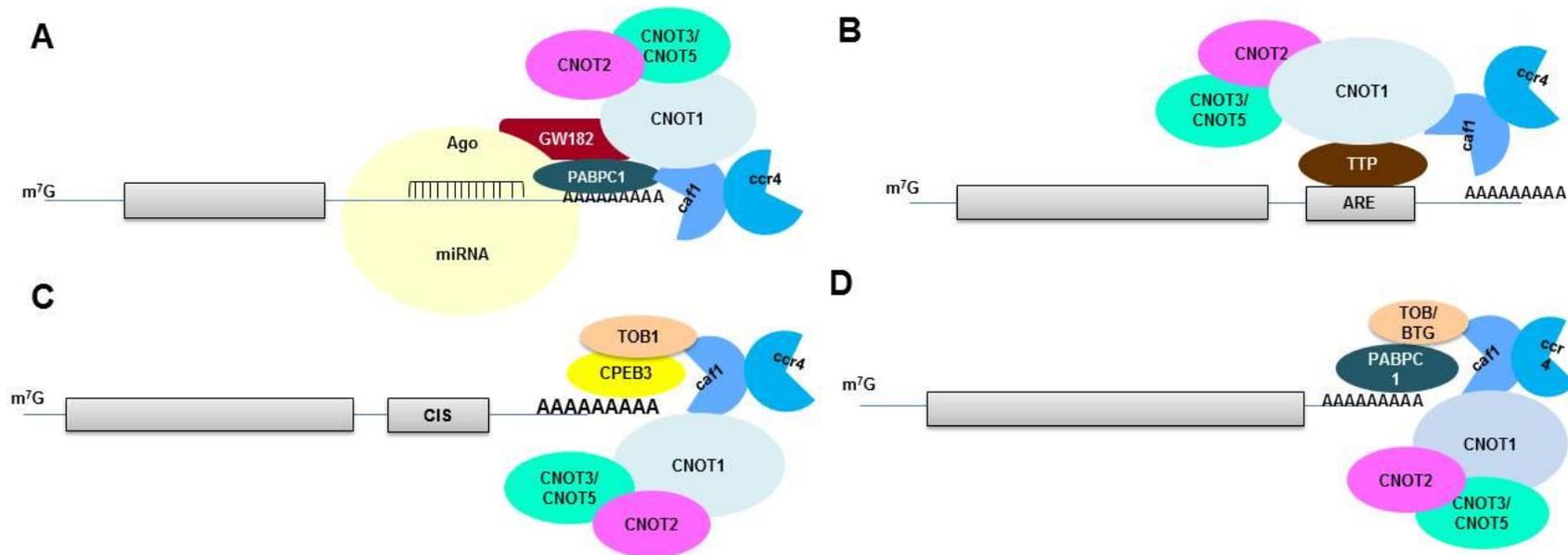


Figure 1-10 Model for recruitment the Ccr4-Not complex.

(A) MicroRNA (miRNA)-mediated recruitment of the Ccr4-Not complex to mRNA. miRNA comes together with GW182 and one of the Argonaute(Ago) proteins to form a complex called miRNA-Induced Silencing Complex (miRISC). GW182 interacts with PABPC1 and recruits the target mRNA. CNOT1 binds to GW182 and mediates the Ccr4-Not complex to facilitate deadenylation of miRNA-targeted mRNAs. (B) Adenylate-uridylate-rich elements (ARE) mediates mRNA decay by binding with the tristetraprolin (TTP) protein. CNOT1, the scaffold protein of Ccr4-Not complex, binds with RNA binding protein TTP that acts as a bridge to recruit the mRNA decay enzyme. (C) TOB1 recruits the Ccr4-Not complex to a sequence-specific RNA-binding protein called cytoplasmic Polyadenylation Element Binding Protein 3 (CPEB3) via binding with hCaf1 subunits. (D) TOB1, TOB2, BTG1 and BTG2 facilitates the recruitment of the Ccr4-Not complex to specific RNA, via binding with PABPC1.

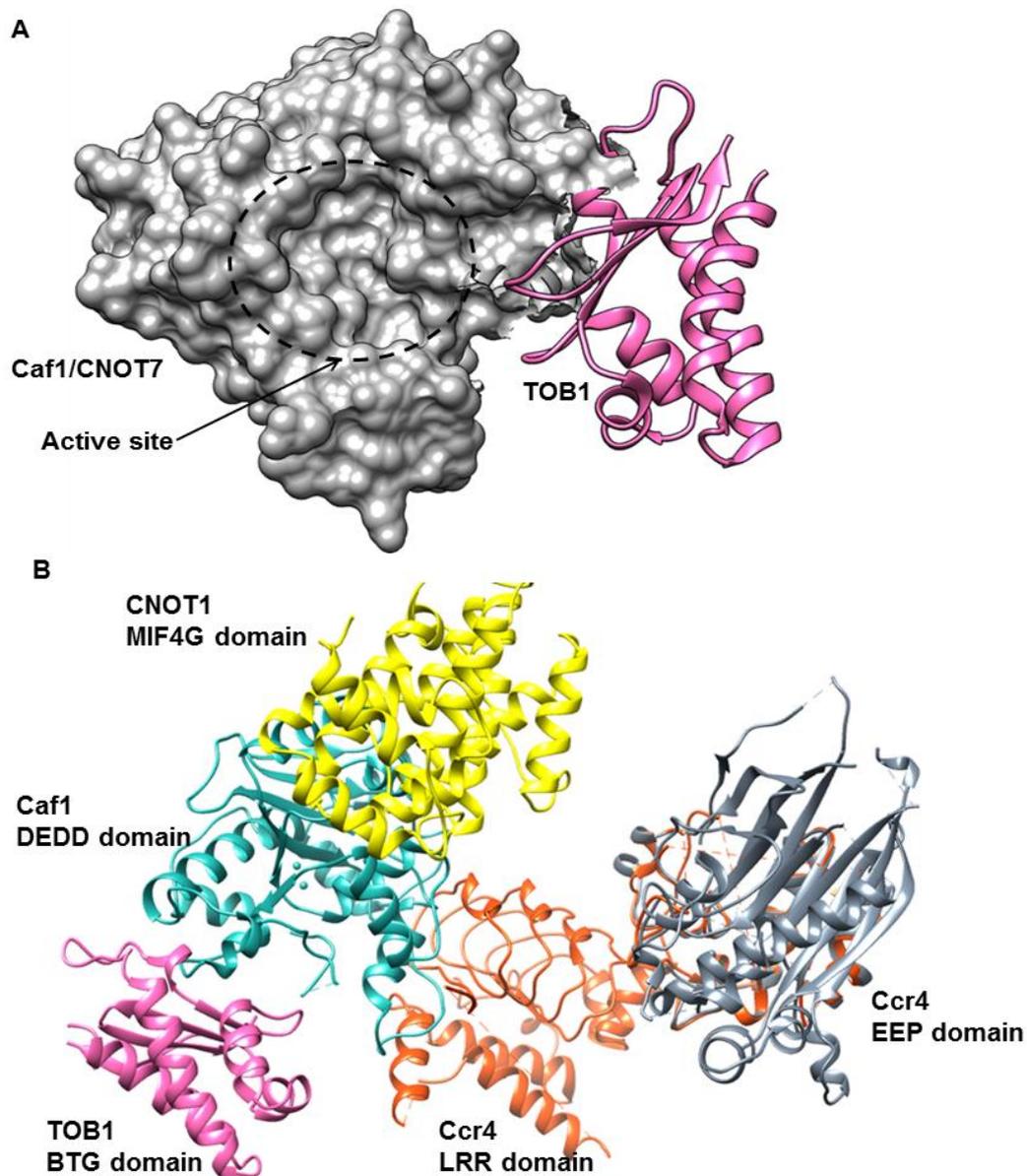


Figure 1-11 Model of the nuclease sub-complex in association with BTG domain of TOB1.

(A) A model of BTG domain interacts with hCaf1/CNOT7. Human crystal structure of the complex TOBN138-hCaf1/CNOT7 (PDB accession number 2D5R), TOB1 is coloured pink and hCaf1/CNOT7 coloured grey. (B) Indicated are the CNOT1 MIF4G domain (yellow), Ccr4 EEP domain (grey), Ccr4 LRR domain (orange), Caf1 DEDD domain (turquoise) and TOB1 BTG domain (pink). The model was provided by alignment of the Tob-hCaf1 structure (PDB accession number 2D5R), the MIF4G domain in complex with hCaf1/CNOT7 (PDB accession number 4GMJ), and the EEP nuclease domain of human Ccr4/CNOT6L (PDB accession number 3NGO) on the structure of the yeast Not1-Caf1 Ccr4 complex (PDB accession number 4B8C). Structure was visualized using the UCSF Chimera package (version 1.11.2) (Meng et al., 2006).

1.9. Aims of study

The human anti-proliferative BTG/TOB family shares a conserved BTG domain that mediates interactions with the hCaf1 (CNOT7/CNOT8) deadenylase subunit of the Ccr4-Not complex and is involved in the regulation of post-transcriptional mechanisms. Caf1 is one of the deadenylase subunits of the Ccr4-Not complex, which induces mRNA degradation via removal of the poly (A) tail of the target mRNA. Previous studies concluded that the BTG/TOB proteins inhibit cell proliferation by recruiting hCaf1(CNOT7/CNOT8) to the target mRNA with subsequent mRNA degradation. Recently, BTG1 and BTG2 variants were identified in NHL samples however, it is unclear whether or how these mutations effect protein function.

Hypothesis: understanding the effects of the BTG1 variants on proliferation activity, and on the interaction with the hCaf1(CNOT7/CNOT8) deadenylase subunit, is important. It is imperative to identify whether the anti-proliferative activity of BTG1 requires the deadenylase activity of hCaf1(CNOT7/CNOT8) to regulate cell proliferation. Therefore, five inter-related questions will be addressed:

1. Do the lymphoma-associated BTG1 or BTG2 variants that have been reported likely affect protein function? Also, are BTG1 or BTG2 mutations present in lymphoma cell lines? To this end, data related to BTG1 and BTG2 variants will be collected from genome browsers and databases and analysed using predictive algorithms of protein function, and selecting variants for experimental analysis (Figure 1-12). Also, sequence analysis of BTG1 and BTG2 in the lymphoblastoid cell lines will be performed.
2. Do BTG1 variants have the ability to interact with hCaf1 (CNOT7/CNOT8)? To investigate this question, a series of BTG1 variants will be generated using site-directed mutagenesis and interactions between BTG1 variants and hCaf1(CNOT7/CNOT8) will be studied using yeast two hybrid analysis (Figure 1-12).

3. Do the BTG1 variants have the capacity to inhibit cell cycle progression? This will be investigated by transfecting BTG1 variants into a mammalian cell line to evaluate protein function using a proliferation assay based on the incorporation of the thymidine analogue EdU (5-ethynyl-2'-deoxyuridine) to label newly synthesised DNA (Figure 1-12).
4. Are BTG1 variants able to induce mRNA degradation in a cellular context? This will be studied using a luciferase reporter mRNA to which BTG1 variants are artificially recruited using an RNA binding peptide. The subsequent effect on mRNA translation and degradation will be studied using luciferase reporter assays and reverse transcriptase-quantitative PCR (Figure 1-12).
5. Do amino acid substitutions found in BTG1 variants impact on the interaction with the polyA-binding protein PABPC1? This will be addressed by studying PABPC1-BTG1/BTG2 protein-protein interactions using the yeast two-hybrid system or GST pulldown analysis.

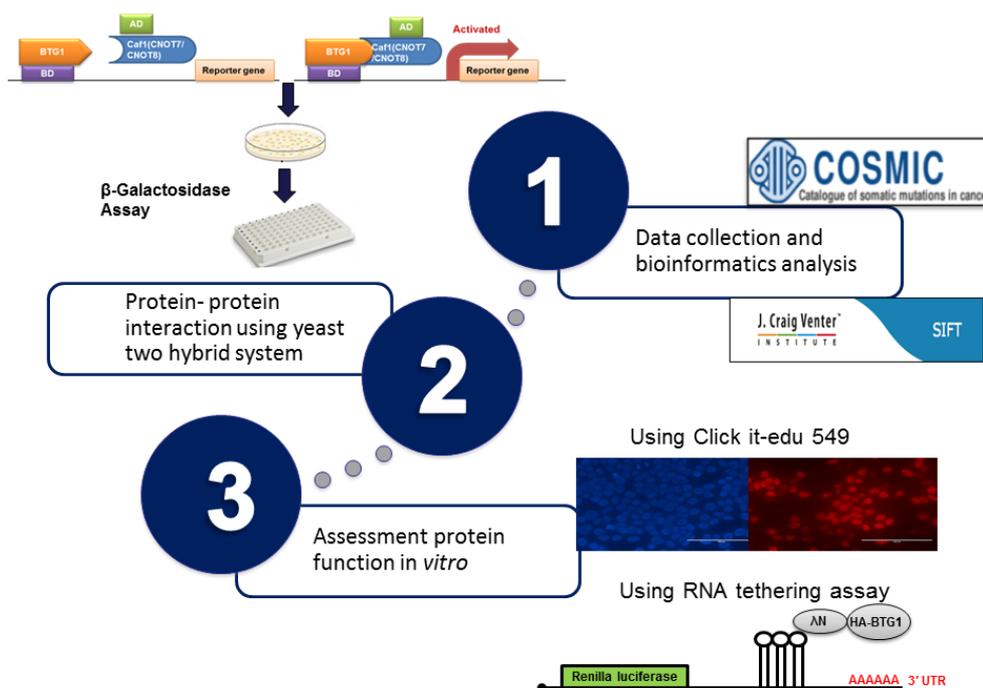


Figure 1-12 Scheme of the experimental approach used.

Chapter 2
Materials and Methods

Chapter 2: Materials and Methods

2.1 In silico analysis of BTG1 and BTG2 mutations

The Catalogue of Somatic Mutations in Cancer (COSMIC) v65 and the Ensembl v70 genomic browser were used to identify the somatic variants in BTG1, BTG2 and genes encoding Ccr4-Not subunits (Forbes et al., 2010, Flicek et al., 2012). The COSMIC browser provides information of the somatic variation in human cancers, from the Cancer Genome project and the Sanger Institute in the United Kingdom (Forbes et al., 2010). Furthermore, novel BTG1 and BTG2 variants were reported in the literature (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Zhang et al., 2013). The Sorting Tolerant From Intolerant (SIFT) algorithm was applied to classify amino acid substitutions into 'tolerated' or 'damaging' using a cut-off score value =0.05 (Kumor et al., 2009) (as described in the Appendix A Figures 1). The prediction values provided were determined through calculating the probability of all 20 amino acids at that position and alignment of the protein family sequence. Also, the Suspect (Disease-Susceptibility-based Single Amino acid Variants (SAVs)) algorithm was used. It not only calculates the conserved protein sequence but also the protein-protein interactions (PPIs) and associated disease (Yates et al., 2014), with a score above 50 indicated as more associated with disease (as described in the Appendix A Figure 2 to 5).

2.2 Bacterial culture and transformation

2.2.1 Reagents, solutions and buffers used in bacterial culture

The chemicals used to make lysogeny Broth (LB) media were: 10 g/L tryptone, 5 g/L yeast extract, 5 g/L NaCl and made to pH 7.2 with NaOH. For LB agar, 15g/L bacteriological agar was included and both forms of media were made up to a total volume of 400 ml with sterile water. Sterilization was achieved by using an autoclave at 121°C for 15 mins, then media was stored at room temperature.

Ampicillin (1000× stock solution): 100 mg/ml in 50% ethanol. Solutions were sterilised by filtration (0.22 µm pore size) and stored at -20°C.

Choramphenicol (1000× stock solution): 34 mg/ml in absolute ethanol. Solutions were sterilised by filtration (0.22 µm pore size).and stored at -20°C.

2.2.2 Preparation of *Escherichia coli* competent cells

Escherichia coli strain DH5 or BL21 (DE3) was prepared for transformation using calcium chloride. A single colony was inoculated in 5 ml of LB medium and grown in the shaker at 37°C, 200 rpm overnight. Then, the culture was diluted in LB media (1:60) and grown in the shaker at 37°C, 200 rpm until the OD₆₀₀ reached 0.5. The culture was then placed on ice for 10 mins in 50 ml Falcon tubes. Then, cells were harvested by centrifugation at 1100 g for 10 mins at 4°C. The supernatant was discarded and cell pellets were re-suspended in 50 ml of 0.1M MgCl₂ and centrifuged at 1100 g for 10 mins at 4°C. After that, the supernatant was removed and the cell pellets were resuspended in 8.6 ml of 0.1 M CaCl before 1.4 ml glycerol was added. The suspension of cells was then distributed into 100 µl aliquots in sterile microfuge tubes before being stored at -80°C after snap freezing in liquid nitrogen.

2.2.3 Transformation of competent cells

For each transformation, approximately 500 ng of plasmid DNA was mixed gently with 50 µl of competent cells in a 1.5 ml tube on ice. Then, the transformation mix was heat shocked by incubation in a water bath at 42°C for 90 seconds, followed by incubation on ice for 2 minutes. 1 ml of LB media was added to the cells followed by incubation at 37°C, 200 rpm for one hour. 100 µl of cell suspension was spread on a selective agar plate (containing the appropriate antibiotic(s)) and incubated overnight at 37°C to obtain single colonies the following day. For ligations or mutagenesis, transformation cultures were spun down for 30 sec at 3000 g, 650 µl of the supernatant was removed and the cell pellet was re-suspended in the remaining supernatant.

150 µl of this cell suspension was then streaked on LB agar plates (containing the appropriate antibiotic(s)) and incubated overnight at 37°C.

2.2.4 Culture of *Escherichia coli* after transformation

Escherichia coli DH5α or BL21 (DE3) was used for all DNA manipulations. Transformed bacteria were streaked onto agar plates with selective antibiotic(s) and incubated at 37°C. Plates were subsequently stored at 4°C for up to four weeks. Liquid cultures of LB media and selective antibiotic(s) were inoculated with single colonies and grown in sterile tubes at 37°C, 200 rpm for 24 hours.

2.3 Molecular biology

2.3.1 Reagents, stock solutions and buffers used in molecular biology methods:

TE buffer: 10 mM Tris-HCl pH 8.0, 1mM EDTA.

Oligonucleotides: Primers were dissolved in TE buffer to make 100 µM solutions and were stored at -20°C.

6 × Ficoll loading dye: 16% Ficoll, 0.2% orange G in H₂O and stored at room temperature.

5 × TBE: 40 mM Tris base, 40 mM boric acid, 1 mM EDTA, pH 8.0, stored at room temperature.

Ethidium bromide (10mg/ml): diluted to 0.5 µg/ml in 0.5 × TBE buffer.

2.3.2 Small scale plasmid DNA preparation

Plasmid Mini prep kits (Macherey-Nagel) were used to prepare plasmid DNA, A single bacterial colony was picked and used to inoculate 3 ml of LB media containing the appropriate antibiotic(s) and was grown overnight in a rotary shaker at 37°C, 200 rpm. Cells were harvested by centrifugation in a 1.5 ml microfuge tube at 11,000 g for 30 seconds and the manufacturer's protocol

was then followed. The DNA concentration was determined by optical density measurements at 260 and 280 nm wavelengths using a Nanodrop ND-1000 UV-Vis spectrophotometer. Purified plasmids were stored at -20°C.

2.3.3 Large scale plasmid DNA preparation

For production of a large amount of plasmid DNA for mammalian transfections, the Sigma GenElute HP maxi prep kit was used. A single bacterial colony was picked and used to inoculate 5 ml of LB containing the appropriate antibiotic(s) and this was grown for 4-6 hours in a rotary shaker at 37°C, 220 rpm. The whole 5 ml culture was then used to inoculate a flask containing 200 ml of LB media with the appropriate antibiotic(s) and grown overnight in a rotary shaker at 37°C, 220 rpm. All of the overnight culture was pelleted by centrifugation at 4000 g for 15 mins. The supernatant was discarded and the cell pellet was re-suspended following the manufacturers protocol. The plasmid was eluted in 1 ml of elution buffer and stored at -20°C.

2.3.4 Determination of DNA/RNA concentration

The Nanodrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies) was used to determine all the DNA and RNA concentrations through spectrophotometry. This calculated the value of the A_{260}/A_{280} wavelength ratio to estimate the purity of the samples. A ratio between 1.8 to 2.0 was considered acceptable for DNA samples, while for RNA the acceptable ratio was between 2.0 and 2.2.

2.3.5 Agarose gel electrophoresis

DNA was resolved on 0.8-2% agarose gels (depending on the size of DNA fragments). The agarose gel (w/v) was prepared by dissolving agarose (g) in 50 ml of 0.5 × TBE. Ethidium bromide (5 µg/ml) was added to the gel before it began to solidify and this mixture was poured into a gel tray equipped with an appropriate comb. Each DNA sample was loaded into the relevant well in 1×Ficoll sample buffer. DNA gels were run at 80V for 45-60 minutes and

visualised by UV transillumination. The 2-log ladder was used (New England Biolabs) as a reference of DNA fragment size.

2.3.6 DNA extraction and purification from an agarose gel

To purify individual DNA bands, DNA was visualised using ultraviolet light after electrophoresis and excised with a scalpel. Extracted gel was dissolved in binding buffer (100 µg per 100 µl of buffer) and incubated at 65°C for 5-10 mins. Samples were applied to a spin column in a collection tube and were spun down at 11,000 g for 1 min. The spin column was washed twice with wash buffer containing ethanol and after each time the flow-through was discarded. The DNA was eluted in 30µl of pre-heated elution buffer at 50°C. This procedure was performed following the manufacturer's protocol from the PCR clean-up and Gel extraction kit (Macherey-Nagel)

2.3.7 Restriction enzyme digestion of DNA

In order to generate compatible ends of DNA for cloning and to confirm the newly generated plasmids, restriction endonucleases were used. 0.5-1 µg of cDNA (PCR products) or plasmid was digested with 1 µl of each restriction enzyme and 5 µl of the appropriate buffer (depending on the type of the restriction enzymes), in a final volume of 50 µl. The mixture of DNA/restriction enzymes was incubated at 37°C for 2 hours, followed by incubation at 65°C for 5 mins to deactivate the enzyme, before analysis by agarose gel electrophoresis. If large DNA quantities (4 µg) were used for cloning, it was divided into two reactions containing 2 µg of DNA in each, and incubated longer for 3 hours at 37°C. All the restriction enzymes were supplied by NEB and the manufacturer's protocol was followed.

2.3.8 Removal of 5'phosphate group from linearised plasmid DNA

Removing the 5'phosphate group from restriction enzyme digested plasmid was used to prevent self-ligation of linear plasmid in ligation reactions. Using Antarctic phosphatase (NEB), 1 µl of enzyme was added to the digested

plasmid (2-4 μg of DNA), incubated at 37°C for 30 mins and then this incubation was repeated again by adding another 1 μl of enzyme. This was then followed by incubation at 65°C for 5 mins to deactivate the enzyme activity.

2.3.9 Ligation of DNA fragments

Recombinant plasmids were generated by ligation of DNA fragments using T4 DNA ligase (NEB). Ligation of vector to insert used about 80 ng of digested vector DNA and a 3-fold more molar ratio of digested insert DNA as shown in table 2-1. The reaction was incubated overnight at room temperature.

Component/reagent	Volume
80 ng vector DNA	-
3-fold ratio more insert DNA	-
10 \times T4 DNA Ligase buffer	1.5 μl
1 \times T4 DNA ligase enzyme	1 μl
H ₂ O	Up to 15 μl

Table 2-1 Standard ligation reaction.

2.3.10 NEBuilder HiFi DNA Assembly Reaction

Restriction enzyme-treated vectors can have 5' overhangs, 3' overhangs or blunt ends. A vector was linearized by digestion with the appropriate restriction enzyme. The DNA insert fragment was amplified with oligonucleotides designed to overlap the 5' region of the insert with the 5' end of the vector and extend the DNA sequence for at least 30 nucleotide bases, including the restriction enzyme site. The corresponding 3' oligo was designed in the same way but in the 5' end regions; the list of the oligonucleotides that were used are in table 2-3. The reaction set up was prepared in a total volume of 3 μ l (Table 2-2) in PCR tubes. Then, DNA mixtures were incubated in a thermocycler at 50°C for 15 minutes, then stored on ice or at -20°C for subsequent transformation, following the manufacturer's protocol (NEBuilder® HiFi DNA Assembly Cloning Kit). The primers used are listed in Table 2-3.

Component/ reagent	Volume
Insert DNA (20 ng/ μ l)	-
Vector DNA (30 ng/ μ l)	-
NEBuilder HiFi DNA Assembly Master Mix	1 μ l
H ₂ O DNAase free	Up to 3 μ l

Table 2-2 DNA Assembly Reaction.

Name	Sequence (5`-3`)
BTG1_EcoRI FW	AAA AAA GAA TTC ATG CAT CCC TTC TAC ACC CGG GCC
pCI-HA-XhoI-BTG1 FW	CCACTGCTGGGCCTGGACAGCACCCCTCGAGTACCCAT ACGATGTTCCAG
pCI-HA-XhoI-BTG1 RV	GGTCGACTCTAGAGGTACCACGCGTGAATTTTAACCT GATACAGTCATC
pCI-HA-XhoI- BTG1C149del RV	GGTCGACTCTAGAGGTACCACGCGTGAATTTTAACAG CTGATTCCGGCTG
pEG202-LEXAEcoRI- BTG1 FW	GTTATTGCGAACGGCGACTGGCTGGAATTCATGCATC CCTTCTACACCC
pEG202-LEXA-Sall- BTG1 RV	CGAATTAGCTTGGCTGCAGGTCGACTCGAGTTAACCT GATACAGTCATC
pJG45-B42-XhoI- PABPC1-FW	GATTATGCCTCTCCCGAATTCGGCCGACTCGAGTGAA CCCCAGTGCCCC
pJG45-B42-XhoI- PABPC1-146-RV	CAAACCTCTGGCGAAGAAGTCCAAAGCTTTTACTCAA AGTGTACAAATC
pJG45-B42-XhoI- PABPC1-190-RV	CAAACCTCTGGCGAAGAAGTCCAAAGCTTTTAGAATT CTTTTGCCCTAGC
pJG45-B42-XhoI- PABPC1-99-RV	CAAACCTCTGGCGAAGAAGTCCAAAGCTTTTAGCCTA CTCCACTTTTGCG

Table 2-3 List of the primers used for HiFi DNA assembly.

2.3.11 Polymerase chain reaction (PCR) for cloning

PCR was performed using a peqlab primus 96 advanced thermal cycler. PCR was used for the generation of DNA fragments for cloning and screening for recombinant plasmids.

The BTG1 cDNA was ordered from Source Bioscience (NCBI accession number BQ229731). For synthesis of blunt end DNA fragments for cloning, PCR reactions were performed in 50 μ l reactions at a standard concentration of phusion DNA polymerase, as shown in Table 2-4. All reagents were supplied from NEB. PCR amplifications were performed for 30 cycles (denaturation 95°C for 15 seconds, followed by annealing at 50°C for 30 seconds and elongation at 68°C for 90 seconds). The PCR products obtained were resolved by 1-2% agarose gel electrophoresis.

For screening of the new recombination plasmids, colony PCR was performed. A single colony from a ligation reaction was used to inoculate 1 ml of LB with the appropriate antibiotic and this was grown for 4 h in a shaker at 37°C, 200 rpm. A standard Taq Polymerase PCR reaction was set up as shown in Table 2-5; all reagents were supplied from NEB. Samples were initially denatured at 95°C for 3 mins, followed by 35 cycles each consisting of a denaturing step of 95°C for 20 secs, annealing at 55°C for 30 secs, then elongation at 72°C for 1 min/kb. A final elongation step for 10 mins at 72°C was included. The PCR products obtained were resolved by 0.8-2% agarose gel electrophoresis, to check for the presence of the insert.

For screening mycoplasma contamination in the cell lines, PCR was performed. 10 μ l of cell media was collected, and amplified using a standard Taq polymerase, at 50°C annealing temperature. The PCR products obtained were resolved by 2% agarose gel electrophoresis, to check for the presence of the insert. All primers used are listed in Table 2-6.

Reagent	Volume
10 mM dNTPs	1 μ l
Phusion enzyme	0.5 μ l
5 \times Phusion Buffer HF	10 μ l
10 μ M forward primer (Sigma), 100 μ M stock in TE	1 μ l
10 μ M reverse primer (Sigma), 100 μ M stock in TE	1 μ l
20 ng DNA template	-
Nuclease-free H ₂ O	Up to 50 μ l

Table 2-4 Standard reaction set up of PCR using the Phusion DNA polymerase.

Reagent	Volume
10 mM dNTPs	0.5 μ l
Taq DNA polymerase	0.25 μ l
10 \times Taq polymerase Buffer HF	2.5 μ l
10 μ M forward primer (Sigma), 100 μ M stock in TE	0.5 μ l
10 μ M reverse primer (Sigma), 100 μ M stock in TE	0.5 μ l
Culture	2 μ l
Nuclease-free H ₂ O	Up to 25 μ l

Table 2-5 Standard reaction set up of PCR using Taq DNA polymerase.

Name	Sequence (5`- 3`)
pBD-Gal4 FW	GTGCGACATCATCATCGGAAG
M13RV	CACACAGGAAACAGCTATGACCAT
T7FW	TAATACGACTCACTATAGGG
T7RV	GCTAGTTATTGCTCAGCGG
EcoRI-BTG1 FW	AAA AAA GAA TTC ATG CAT CCC TTC TAC ACC CGG GCC
BTG1-SalI RV	AAA AAA GTC GAC TTA ACC TGA TAC AGT CAT CAT ATT G
BglII-BTG1 FW	AAA AAA AGA TCT ATG CAT CCC TTC TAC ACC CGG GCC
BTG1 C149del RV	AAA AAA GTC GAC TTA TCA GCT GAT TCG GCT GTC TAC C
BTG1 H2Y PBD FW	CCCGGGAATTCATGTATCCCTTCTACACC
BTG1 H2Y PBD RV	GGTGTAGAAGGGATACATGAATTCCCGGG
Δ10N BTG1 FW	AAAAGAATTCATGGATAGGCGATA
MCGpF1	ACACCATGGGAGCTGGTAAT
MCGpR1	CTTCATCGACTTTCAGACCCAAGGCAT
MCGpF2	GTTCTTTGAAAACCTGAAT
MCGpR2	GCATTCCACCATATACTCT

Table 2-6 List of primers used for DNA cloning.

2.3.12 Site-directed mutagenesis

To introduce point mutations in plasmids, primers for site directed mutagenesis were designed using primerx (<http://www.bioinformatics.org/primerx>). The primers are listed in Table 2. Following the Stratagene Quickchange protocol, plasmids were amplified using Phusion DNA polymerase in a final total reaction volume of 50 μ l. The PCR reaction began with a hot start at 95°C for 5 minutes. Subsequently, denaturation occurred at 95 °C for 15 seconds for 30 cycles and annealing at 50°C for 30 seconds. Finally, the length of elongation was calculated to the size of the DNA: 1 min per 0.5 kb at 72°C each cycle and then a final elongation of 15 minutes. The template DNA was digested using 1 μ l DpnI enzyme (NEB) and incubated at 37°C for 2 hours, before analysis via 1% agarose gel electrophoresis.

The remaining sample was then used to transform *E.coli* DH5 α before DNA isolation and sequencing to confirm the presence of the desired mutation(s). The plasmid used was pBD-Gal4HA-BTG2 and was created by Doidge et al., 2012. All primers used for this work are listed in Table 2-7.

Name	Sequence (5`- 3`)
BTG1_M11I FW	GGCCGCCACCATAATAGGCGAGATC
BTG1_M11I RV	GATCTCGCCTATTATGGTGGCGGCC
BTG1_F25C FW	CTTCATCTCCAAGTGTCTCCGCACCAAG
BTG1_F25C RV	CTTGGTGCGGAGACACTTGGAGATGAAG
BTG1_R27H FW	CCAAGTTTCTCCACACCAAGGGGCTC
BTG1_R27H RV	GAGCCCCTTGGTGTGGAGAACTTGG
BTG1_Q36H FW	GAGCGAGCGACATCTGCAGACCTTC
BTG1_Q36H RV	GAAGGTCTGCAGATGTCGCTCGCTC
BTG1_L37M FW	GAGCGAGCGACAGATGCAGACCTTCAG
BTG1_L37M RV	CTGAAGGTCTGCATCTGTCGCTCGCTC
BTG1_F40C FW	CAGCTGCAGACCTGCAGCCAGAGCCTG
BTG1_F40C RV	CAGGCTCTGGCTGCAGGTCTGCAGCTG
BTG1_P58L FW	AACATCACTGGTTCCTAGAAAAGCCATGCAAG
BTG1_P58L RV	CTTGCAATGGCTTTTCTAGGAACCAGTGATGTT
BTG1_E59D FW	CACTGGTTCCAGATAAGCCATGCAAGG
BTG1_E59D RV	CCTTGCATGGCTTATCTGGGAACCAGTG
BTG1_G66V FW	CATGCAAGGGATCGGTTTACCGTTGTATTCG
BTG1_G66V RV	CGAATACAACGGTAAACCGATCCCTTGCATG
BTG1_N73K FW	CGTTGTATTCGCATCAAGCATAAAATGGATCCTC
BTG1_N73K RV	GAGGATCCATTTTATGCTTGATGCGAATACAACG
BTG1_L94V FW	GAGCAGTCAGGAGGTGTTTCAGGCTTC
BTG1_L94V RV	GAAGCCTGAACACCTCCTGACTGCTC
BTG1_L104 FW	GTGAACTCACACACTGGGTTGACCCC
BTG1_L104H RV	GGGGTCAACCCAGTGTGTGAGTTCAC
BTG1_I115V FW	GAAGTGTCCCTACAGAGTTGGAGAGGATGGC
BTG1_I115V RV	GCCATCCTCTCCAACCTCTGTAGGACACTTC
BTG1_E117D FW	CTACAGAATTGGAGATGATGGCTCCATCTG
BTG1_E117D RV	CAGATGGAGCCATCATCTCCAATTCTGTAG
AD-Gal4-PABPC1 FW	TAAGTCGACTCTAGAGCC
AD-Gal4-PABPC1-190RV	GAATTCTTTTGGCCCTAGC
AD-Gal4-PABPC1-146RV	CTCAAAGGTGTACAAATCCATAG
AD-Gal4-PABPC1-99RV	TAAGTCGACTCTAGAGCC
BTG2_K36Q FW	GAGCAGAGGCTTCAGGTCTTCAGCG
BTG2_K36Q RV	CGCTGAAGACCTGAAGCCTCTGCTC
BTG2_A45E FW	CGCTCCAGGAGGAACTCACAGAGCA

Name	Sequence (5' - 3')
BTG2_A45E RV	TGCTCTGTGAGTTCCTCCTGGAGCG
BTG2_A45T FW	CGCTCCAGGAGACACTCACAGAGCA
BTG2_A45T RV	TGCTCTGTGAGTGTCTCCTGGAGCG
BTG2_H49Y FW	GAGGCACTCACAGAGTACTACAAACACCAC
BTG2_H49Y RV	GTGGTGTGTTGTAGTACTCTGTGAGTGCCTC
BTG2_L46F FW	GCTCCAGGAGGCATTACAGAGCACTAC
BTG2_L46F RV	GTAGTGCTCTGTGAATGCCTCCTGGAGC
BTG2_G64S FW	GTCCAAGGGCTCCAGCTACCGCTGC
BTG2_G64S RV	GCAGCGGTAGCTGGAGCCCTTGGAC
BTG2_I70M FW	CGCTGCATTTCGCATGAACCACAAGATGG
BTG2_I70M RV	CCATCTTGTGGTTCATGCGAATGCAGCG
BTG2_L100P FW	CTGCCCAGCGAGCCAACCCTGTGGGT
BTG2_L100P RV	ACCCACAGGGTTGGCTCGCTGGGCAG
BTG2_I70M RV	CCATCTTGTGGTTCATGCGAATGCAGCG
BTG2_S158C FW	CTACGTGATGGCAGTCTCCTGCTAG
BTG2_S158C RV	CTAGCAGGAGACTGCCATCACGTAG
AD-Gal4-PABPC1 FW	TAAGTCGACTCTAGAGCC
AD-Gal4-PABPC1(1-190)RV	GAATTCTTTTGCCCTAGC
AD-Gal4-PABPC1(1-146)RV	CTCAAAGTGTACAAATCCATAG
AD-Gal4-PABPC1(1-99)RV	TAAGTCGACTCTAGAGCC

Table 2-7 List of primers used for site-directed mutagenesis.

2.3.13 Genomic DNA Purification

Genomic DNA was purified using a Blood and cell culture DNA mini kit and was extracted from mammalian cells using the manufacturer's protocol (Sigma). DNA was harvested from approximately 2×10^6 of lymphoblastoid cells. Cells were pelleted for 5 minutes at 4000 g and the culture medium was removed. Then, the pellet was re-suspended in 200 μ L of resuspension solution. After that, 20 μ L of the Proteinase K solution was added to the suspension cells, followed by 200 μ L of Lysis Solution C (B8803). This mixture was vortexed thoroughly (about 15 seconds), and then incubated at 70°C for 10 minutes. These homogeneous mixtures were applied to spin columns which purified and decontaminated the genomic DNA from cell debris and solution. The total genomic DNA was determined by optical density measurement at 260 and 280

nm using the Nanodrop ND-1000 UV-Vis spectrophotometer. DNA was stored at -20°C.

2.3.14 Total RNA Purification

Typically, the Macherey-Nagel RNeasy kit was used for isolation of total RNA from 1×10^6 cells of lymphoblastoid cell lines. Cells were pelleted for 5 minutes at 4000 g and the culture medium was removed. Cells were lysed by the addition of 350µl of RLT lysis buffer and re-suspended by pipetting up and down. After transfer of the lysate to a 1.5ml microfuge tube, RNA extraction was conducted as stated in the manufacturer's protocol. Lysed cells were applied to spin columns and treated with ethanol and DNase enzyme, before purified RNA was eluted with RNA-free water. The total RNA concentration of the eluate was determined by optical density measurements at 260 and 280 nm using the Nanodrop ND-1000 UV-Vis spectrophotometer. RNA was stored at -80°C.

2.3.15 Reverse transcription

The first strand of cDNA was synthesised by adding 0.5 µg of total RNA in a sterile microfuge tube with buffer containing 1µl dNTPs and 1µl of oligo. This mixture was incubated at 70°C for 5 minutes. Synthesis of cDNA was done using the M-MuLV Reverse Transcriptase (200 units/µL), which was incubated at 42°C for 30 minutes. Following this, the enzyme was denatured by incubation at 85°C for 5 minutes. cDNA was stored at -20°C.

2.3.16 Genomic sequencing analysis using polymerase chain reaction (PCR)

Genomic DNA or cDNA was amplified using Taq polymerase and specific primers as listed in Table 2-8. PCR amplifications were performed in 25 µl reactions at a standard concentration (200 µM dNTP, 10× reaction buffer, 0.2 µM primers, 1.25 units/50 µl PCR Taq polymerase and 2 µL of cDNA). All reagents were from NEB. For the analysis of gene expression, the PCR

reactions were run for 30 cycles (denaturation at 95°C for 15 seconds followed by annealing at 60°C for 30 seconds and elongation at 68°C for 60 seconds). The PCR products obtained were resolved on a 2% agarose gel by electrophoresis. For the analysis of BTG2 expressing cell lines, the PCR reactions were run for 30 cycles (denaturation at 95°C for 15 seconds followed by annealing at 55°C for 30 seconds and elongation at 68°C for 60 seconds). All PCR products were cleaned up before DNA sequence analysis. Amplification of exons 1 and 2 of BTG1 was carried out using Taq polymerase (2.5 units/25 µl PCR Taq). For exon 2, 100 ng of genomic DNA was used in PCR. For exon 1, 500 ng of genomic DNA was used. To optimize the PCR reaction, 2mM of MgCl₂ and 2% DMSO were added to the reactions. The PCR products obtained were resolved by 2% agarose gel electrophoresis.

Name	Sequence (5`- 3`)
BTG2 Fw	AGGGTAACGCTGTCTTGTGG
BTG2-Rv	TACAGTTCCCCAGGTTGAGG
BTG1 Exon1 Fw	GACTCTGACCCAGGGATGTG
BTG1 EXon1 Rv	CATCACGCTCCAGCTACG
BTG1 Exon2 Fw	TCCATAATCCATCCCCAAGA
BTG1 EXon2 Rv	GGATGCAATCCTGGACATTT

Table 2-8 List of primers used for gene expression and sequencing.

2.4 Yeast two-hybrid analysis

2.4.1 Reagents, stock solutions and buffers for use in yeast culture

YPD medium (400mL): 4g yeast extract, 8g glucose, 8g bacteriological peptone. For YPD agar, 8g bacteriological agar was also added and made up to the total volume with distilled water. The medium was dissolved and sterilised by autoclaving at 121°C for 15 minutes before storing at 4°C. All chemicals were from Sigma.

Yeast selective (YS) complete medium (400ml) for YRG2 and L40 yeast strains: 2.67g yeast nutrient broth, 8g glucose, 0.62 g amino acid drop out mix (without tryptophan and leucine). For selective agar plates, 8g bacteriological agar was added. The medium was sterilised and dissolved by autoclaving at 121°C for 15 minutes before storing at 4°C.

Yeast selective (YS) complete medium (400ml) for EGY48 yeast strains: The same as for yeast selective complete medium for YRG2 and L40 yeast strains, with the exception of the amino acid drop out mix, where 0.4 g amino acid drop out mix (without tryptophan, leucine, histidine and uracil) was added instead. For selective agar plates, 8g bacteriological agar was added. The medium was dissolved and sterilised by autoclaving at 121°C for 15 minutes before storing at 4°C.

Yeast selective (YS) complete medium (500ml): 1.67g yeast nutrient broth, 10g galactose, 5g sucrose, 5g ammonium sulfate, 0.4 g amino acid drop out mix (without tryptophan, leucine, histidine and uracil). For selective agar plates, 10g bacteriological agar was added. The medium was sterilised and dissolved by autoclaving at 121°C for 15 minutes before storing at 4°C.

Leucine solution: stock 22mg/ml in H₂O (add 2.6 ml per 500 ml of medium). Sterilised by filtration (0.22 µm pore size). Stored at room temperature.

Single-stranded DNA (2mg/ml): Herring Sperm single-stranded DNA (10mg/ml, Abcam, ab46666) was heated to 98°C for 8 minutes and placed on ice before use. Stored at -20°C.

Lithium acetate solutions: 1M LiAc solution (Sigma) was prepared using 1.25g of LiAc dissolved in 25ml distilled water, sterilised by filtration (0.22 µm pore size) and stored at room temperature. 0.1M LiAc solution (0.1M Sigma) was prepared by diluting 1M LiAc solution 1:10 in 50ml of distilled water, sterilised by filtration (0.22 µm pore size) and stored at room temperature.

PEG-3350 solution: a solution containing 50% (w/v) polyethylene glycol (PEG-3350; Sigma) was prepared with 12.5 g of PEG and 12.5 ml of distilled water, following sterilising by filtration (0.22 µm pore size). The solution was stored at room temperature for two to three months.

SDS loading buffer (4×): 200 mM Tris-Cl (pH 6.8), 40% glycerol, 8% SDS, 40% glycerol and 0.4% bromophenol blue.

2.4.2 Culture of *Saccharomyces cerevisiae* strains

The yeast YRG2, L40 and EGY48 were stored in YPD containing 50% glycerol at -80°C. Before starting yeast two hybrid analysis, yeast was streaked on YPD agar and left growing for 2-4 days at 30°C in the incubator. Once the colonies had formed, the plate was stored at 4°C for up to three months. For further analysis, one large colony was used to inoculate 10 ml of YPD culture medium (in a universal tube), shaking at 30°C 200 rpm overnight. The overnight culture was split into volumes of 25ml (2-4 ml in 100ml, using 50ml Falcon tubes) with an absorbance of OD₆₀₀ 0.2. The cultures were incubated at 30°C, 200 rpm for 5-6 hours until the OD₆₀₀ reached 0.8 (enough cells for 10-20 transformations depending on the experiment requirement).

2.4.3 Transformation of *Saccharomyces cerevisiae* strains

Different plasmids were used for yeast two-hybrid assays, according to the yeast strain as shown in Table 2-9. The competent yeast cells were heat shock

treated for transformations. Once the OD₆₀₀ of the incubated culture reached 0.8, the yeast cells were pelleted by centrifugation at 3000 g for 5 mins and the culture media was discarded. The cells were washed with 50 ml distilled water, centrifuged again and the water discarded. Cells were then re-suspended in 2 ml of 0.1M LiAc and transferred into two 1.5 ml microfuge tubes, followed by centrifugation at 11,000 g for 30 sec and the supernatant discarded. Then, cells were re-suspended in a final volume of 450 µl of 0.1M LiAc and 50 µl of the suspension was transferred into each transformation tube. The centrifugation step was repeated to remove the LiAc from the transformation tubes and 240 µl of PEG was added to protect the cells from the LiAc toxicity. Each tube contained the relevant plasmids and mixture of reagents used for yeast transformation (Table 2-10). Tubes were vortexed to disrupt the pellets, before being incubated in the shaker for 30 min at 30°C. Then, tubes were incubated in a water bath at 42°C for 20-25 min. The cells were centrifuged at 4,000 g for 30 sec to remove the supernatant, which was discarded. Following this, the pellets were re-suspended in 500 µl of distilled water. 200 µl of each suspension of cells were streaked on the selective agar plates for YRG2 and L40 yeast strains (lacking leucine and tryptophan). Alternatively, for the EGY48 strain the selective agar plates used lacked histidine, uracil and tryptophan. All plates were incubated at 30°C for 3-4 days.

Yeast strain	Plasmids
YRG2	pBD-GAL4 Cam and pAD-GAL4-2.1
L40	pAD-GAL4-2.1 and pBmod116LexA
EGY48	pEG202Nls, pJG45 and pSH18-34

Table 2-9 Plasmids used for yeast transformation.

Component/Reagent	Volume
250ng of DNA plasmid diluted in dH ₂ O	50 μ l
50% polyethylene glycol (PEG-3350)	240 μ l
1M lithium acetate	36 μ l
Single-stranded DNA (2 mg/ml)	25 μ l

Table 2-10 Composition of Yeast transformation mixtures.

2.4.4 Yeast two-hybrid analysis

After transformation, three yeast clones were tested independently. Yeast colonies were streaked on to small rectangles on selective agar plates and incubated for two days in the incubator at 30°C. Different selective media were used according to the yeast strain. Glucose selective media were used for YRG2 and L40 strains, while for the EGY48 strain galactose/sucrose YS media was used due to the additional reporter gene that was added to the transformation process (pSH18-34). Then, the growing yeast were inoculated in 2 ml YS media and incubated with shaking at 200 rpm at 30°C. After 24 hours, 100 μ l of each yeast culture was diluted 10 fold with YS media and the OD₆₀₀ value was measured to determine the culture density. The remaining amount of each culture was used for β -Galactosidase assays. For this, an equal amount of the culture was added with Beta-Glo® Reagent (Promega) (about 25 μ l) in a white 96 well flat bottom plate and incubated with shaking at 150 rpm for 5 min, followed by 25 min in the dark at room temperature. β -Galactosidase luminescence was measured for 1 sec using a Berthold Orion microplate luminometer and Simplicity 4.02 software. For each yeast culture, the β -galactosidase activity was tested. The β -galactosidase activities were

normalised using the OD₆₀₀ values. Three independent experiments were performed.

2.4.5 Preparation of yeast cells for western blot analysis

Protein lysates were produced and subjected to western blot analysis to confirm protein expression. From each stored selective agar plate, a single colony was re-streaked on another selective agar plate and incubated for two days at 30°C. Then a single re-streaked colony was inoculated in 2 ml of YS media and cultured at 200 rpm at 30°C overnight. After that, the culture was split into 1.5 ml tubes and centrifuged at 3000 g for 2 mins. The supernatant was removed and the pellet re-suspended in 100 µl of distilled water. Cells were lysed by adding 100 µl of 0.2M NaOH, vortexed and incubated for 5 mins at room temperature. Then, the lysed cells were centrifuged at 3000 g for 2 mins and the supernatant discarded. This was followed by the addition of 50 µl of 1×SDS loading dye and vortexing. Finally, proteins were extracted by boiling at 95°C for 5 mins, before running SDS-PAGE. Protein extracts were stored at -20°C.

2.5 Cell Culture

2.5.1 Reagents, stock solutions and buffers for use in tissue culture

Phosphate buffer saline: Sigma

DMEM complete: DMEM (sigma, D6546), 10% fetal calf serum (FCS; Biosera S1900-500, lot-S0611951900), 100 units/ml penicillin and 100 units/ml streptomycin (PAA,P11-010), glutamine (200mM;Lonza,BE17-60SE).

RPMI-1640 complete: RPMI-1640 (sigma, R0883), 10% fetal calf serum (FCS; Biosera S1900-500, lot-S0611951900), 100units/ml penicillin and 100units/ml streptomycin (PAA,P11-010), glutamine (200mM; Lonza, BE17-60SE).

Trypsin/EDTA solutions: 10× concentrated solutions of phosphate-buffered saline containing 0.5% trypsin and 0.2% EDTA were purchased from PAA (L11-003) and stored at -20°C. 1× solutions were obtained by the addition of PBS and were stored at 4°C.

2.5.2 Routine maintenance of lymphoma cell lines

To maintain lymphoblastoid suspension cell lines, all work was performed in a class II biological safety cabinet. Two human cell lines, GM03201 and GM01953, were derived from B cell lymphocytes of healthy individuals and were purchased from the NIGMS Human Genetic Cell Repository. Cell lines Raji, KM-H2, and DB were isolated from B-cell lymphoma tissues and purchased from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ).

The lymphoblastoid suspension cell lines were maintained in RPMI-1640 complete media supplemented with 10% FCS, 2mM glutamine, 100 units/ml penicillin and 100 units/ml streptomycin. Cells were cultured in a humidified incubator at 37°C and 5% CO₂. When cells were 70-80% confluent, they were collected in a 15 ml Falcon tube and centrifuged at 300 g for 5 mins. The supernatant was then removed and the pellet re-suspended in complete RPMI media. Cells were passaged and seeded at a density of 3×10^5 cells/ml.

2.5.3 Routine maintenance of HEK293T cells

Working in a class II biological safety cabinet, the human embryonic kidney (HEK293T cells) were maintained in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10% FCS, 2mM glutamine, 100 units/ml penicillin and 100 units/ml streptomycin (complete DMEM), in a humidified incubator at 37°C and 5% CO₂. When cells were 70-80% confluent, cells were washed with phosphate buffered saline (PBS, pre-warmed at 37°C) and incubated with 1-2 ml of trypsin/EDTA for 2 mins at 37°C. Trypsin was deactivated by the addition of 9 ml of complete medium and the cells were re-

suspended and transferred into a 15 ml Falcon tube. The cell suspension was centrifuged at 250 g for 5 mins to pellet the cells and the supernatant was removed; the pellet was re-suspended in complete DMEM for passaging and seeding.

2.5.4 Freezing cells for long term storage

To place suspension cell lines from a confluent 75cm² flask into liquid nitrogen storage and working in a class II biological safety cabinet, cells were collected and transferred into a 15 ml Falcon tube. Cells were centrifuged at 250 g for 5 mins. Then, the supernatant was discarded and the cell pellet was re-suspended in 1 ml of fetal calf serum (FCS) and 10% dimethylsulfoxide (DMSO), and transferred into cryovials. For adherent cells, they were detached from a confluent 75cm² flask using trypsin and transferred to a 15 ml Falcon tube, before being centrifuged at 250 g for 5 mins. After centrifugation, the supernatant was removed and the cell pellet was re-suspended in 2 ml of complete media (DMEM) containing 10% DMSO and 1 ml aliquots of the resulting cell suspension were transferred into cryovials. The cryovials were then placed at -80°C in a Mr. Frosty cryo-container (NALGENE Labware), which allows slow freezing of -1°C per minute. After one to several days, cryovials were placed in liquid nitrogen for long term storage.

2.5.5 Retrieving cells from liquid nitrogen storage

To retrieve cells from liquid nitrogen storage, cells were thawed quickly by incubation at 37°C in a water bath. Working in a class II biological safety cabinet, cells were transferred into a 15 ml Falcon tube containing 9 ml of complete media pre-warmed to 37°C before use. Cells were pelleted by centrifugation at 250 g for 5 mins. After that, the supernatant was discarded to remove the DMSO and the cell pellet re-suspended in 10 ml of complete media. The cell suspension was transferred into a 75cm² flask and placed in a humidified incubator (5% CO₂ and 37°C). After 24 h, the medium was changed

and the cells were passaged between 48 and 96 h after retrieval from liquid nitrogen.

2.6 Edu labelling of S-phase cells

To identify the effect of BTG1 variants on cell proliferation, the Click-iT Edu Imaging kit (Invitrogen) was used to label cells in S-phase.

2.6.1 Reagents, stock solutions and buffers for use in Click-iT Edu labelling

Lysis buffer contained 1% NP40: 50 mM 1M Tris-HCl pH 8.0, 150 mM NaCl, 5 mM MgCl₂, 0.5 mM EDTA, 5% glycerol, 1% NP-40 and 1 mM DTT. The lysis buffer was stored at -20°C after the addition of one tablet of Roche complete protease inhibitor cocktail into 50ml of buffer.

4% paraformaldehyde (PFA): 4% PFS w/v in H₂O, stored at -20°C.

0.5% Triton X-100: 0.5% Triton X-100 in PBS, stored at 4°C.

3% Bovine serum albumin (BSA): 3% BSA in PBS, stored at 4°C.

Mounting medium: 90% glycerol, 10% PBS.

Phosphate buffered saline (PBS) without Ca²⁺, Mg²⁺ : 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.47mM KH₂PO₄ (PAA, H15-002)

2× Hebes-buffer saline (HBS): 280 mM NaCl, 50 mM HEPES pH 7.0, 10 mM KCl, 12 mM dextrose, 1.5 mM Na₂PO₄.

2.6.2 Cell seeding and transfection for Click -iT Edu labelling

Working in the safety cabinet, the round glass cover slips (22 mm x 22 mm, 0.13-0.17 mm thick) were placed into a 6-well culture plate. HEK293T cells were seeded at a density of 2×10⁵/ml in 2 ml DMEM onto the coverslips. After 24 h, cells were at 60-70% confluency and transfected with 1 µg of plasmid DNA using the calcium phosphate precipitation method. This was performed

by mixing 1 µg of plasmid DNA with 12 µl of 2M CaCl₂ and made up to a total volume of 100 µl with dH₂O in 1.5 ml tube. 100 µL of 2×HBS were added dropwise into the DNA/CaCl₂ mixture while vortexing and incubated for 20-30 mins at room temperature before adding to the cells. To monitor the transfection efficiency, a GFP plasmid was transfected in triplicate with each experiment. 16 h post-transfection, cells were washed with PBS after removing the cell media and this was replaced with fresh media to reduce cell toxicity. Forty-eight hours after transfection, cells were subjected to Click-iT EdU labelling. Each transfection was carried out in triplicate.

2.6.3 Click-iT EdU labelling

Forty-eight hours after transfection, the GFP plasmid-transfected cells had their media replaced with PBS and were stored at 4°C until ready to be fixed. For other cells which had been transfected with plasmids (either empty vector or BTG1 (wildtype or variants)), 1ml of medium was removed from each well and replaced with 1ml of fresh medium containing 20µM EdU, to give a final working concentration of 10 µM EdU per well. HEK293T cells were incubated with EdU for either 30, 60 or 120 mins, in a humidified incubator at 37°C and 5% CO₂. After labelling, the medium containing EdU was removed and the cover slips were washed twice with 1ml of PBS. Cells were fixed by a 10 min incubation with 1ml of 4% PFA and washed twice with PBS; they could then be stored at 4°C for 24 hours. Following this, the cover slips were washed with 3% BSA in PBS, before being incubated with 1ml of ice cold 0.5% Triton X-100 (diluted in PBS) for 20 mins to permeabilise the cells followed by two washes with 1ml of 3% BSA in PBS. During the 20 mins incubation, the Click-iT reaction cocktail was prepared as described by the manufacturer in the Click-IT protocol (Click-iT® EdU Alexa Fluor® 594 Imaging Kit C10339) and all component volumes were reduced by a factor of ten to allow 50 µl of cocktail per coverslip. Cells were stained by spotting 50µl of Click-iT reaction cocktail onto parafilm in a humidification chamber and the coverslips were placed face

down onto the solution. Following a 30 min incubation at room temperature in the dark the coverslips were returned into 6 well culture plates and washed twice with 1 ml 3% BSA in PBS. After that, cells were washed with 1 ml PBS before staining the nucleus (DNA) using Hoechst (5 μ g/ml) which was added to each coverslip and incubated in the dark at room temperature for 10 mins. Again, cells were washed twice with 1 ml PBS before the mounting step (can be stored at 4°C for a week), protected from light. To mount the coverslips, a drop of mounting solution (90% glycerol with 10% PBS) was placed on a glass slide and the coverslip was lowered onto it face down. The coverslips were sealed with clear nail varnish and slides were kept in the dark until imaging. For each slide, at least three images were acquired, each containing >50 cells using the EVOS® FL Color Imaging System inverted microscope (Catalog number: AMEFC4300, supplier Thermo Fisher Scientific).

2.6.4 Calculating the number of cells in S-phase

To count the S-phase nuclei and total numbers of nuclei in cells, the ImageJ package was used url: web address <https://imagej.nih.gov/ij/download.html>. The Alexa Fluor® 594 EdU proliferation signal was clearly displayed as red fluorescence. The total number of nuclei were counted for each area using DAPI to detect Hoechst-stained nuclei. Before analysing using ImageJ, the format of the image was changed into TIF-8 black and white to enable the cells to be counted in an automated manner. Firstly, the background was reduced using the subtract background 50-80% option. Secondly, the threshold was used to convert the image into black and white and to detect every single cell; the threshold was reduced to make all cells appear. Then, to fill the gaps for cells or to get a border between cells, the binary function was used and finally, the measurement for cell counting was selected as shown in the Figure 2-1. On average 50-150 nuclei were counted for each slide. The percentage of nuclei in S-phase as a fraction of the total number of nuclei was calculated for each slide. For each sample, the average number of cells in S-phase and the standard error of the mean were calculated from the biological triplicates. The GFP plasmid-

transfected samples were counted via the same method to monitor transfection efficiency, which typically varied between 75-80%.

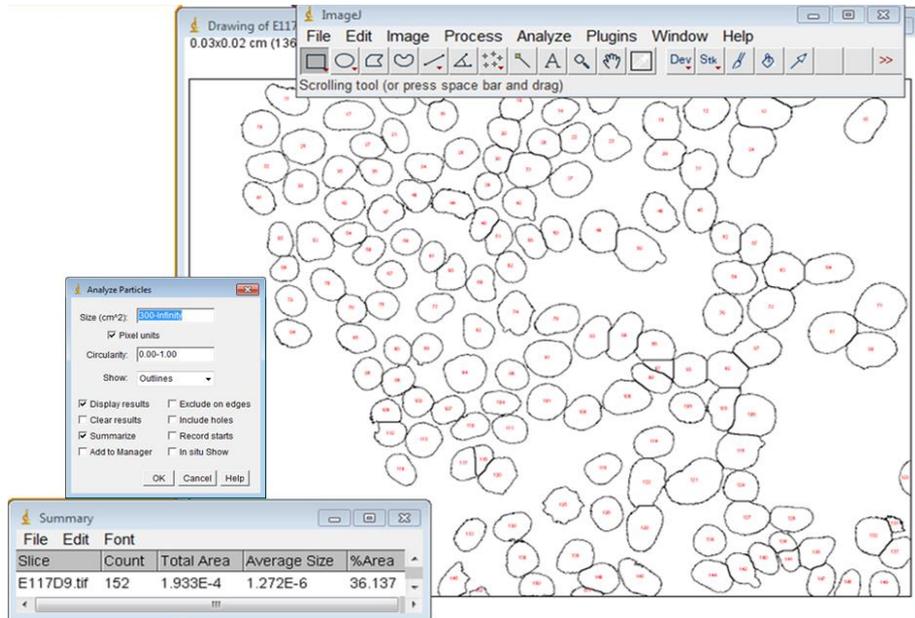


Figure 2-1: Counting the cell nuclei using ImageJ.

2.6.5 Measuring cell fluorescence using ImageJ

To determine the fluorescence intensity of cells pulsed with EdU following different incubation times (20 mins-120 mins) ImageJ software was used. The cell of interest was selected using the drawing/selection tools (using magic point), and then the area integrated intensity and mean grey value were selected to be measured. Three images without cells were selected to monitor the background fluorescence. Fluorescence cells from the three different conditions were counted, measuring about 50 cells from each image. The formula below was finally used to calculate the corrected total cell fluorescence (CTCF).

CTCF = Integrated Density - (Area of selected cell X Mean fluorescence of background readings)

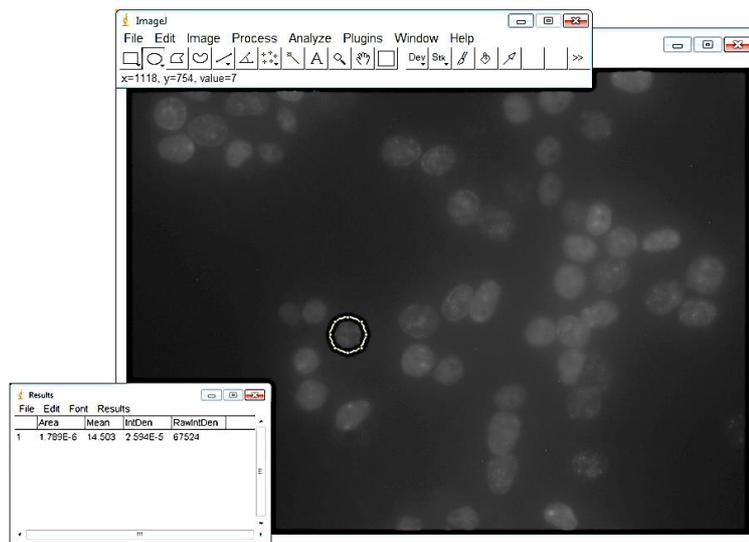


Figure 2-2: Measuring the fluorescence intensity using ImageJ.

2.6.6 Western blot analysis for proliferation assay

To analyse protein expression, cells were seeded and transfected as stated previously for the Click-iT EdU protocol. After 48 h, media was removed before washing the cells with 1 ml of cold PBS. The cells were re-suspended in 500 μ l of cold PBS to detach them, before being transferred to a pre-chilled 1.5 ml microfuge tube. This cell suspension was centrifuged at 16,000 g for 5 mins at 4°C and the supernatant was discarded. The cell pellet was re-suspended in 100 μ l of 1% NP-40 lysis buffer. The cell lysate was freeze-thawed (-80°C), before a second centrifugation at 16,000 g for 5 mins at 4°C. The supernatant was collected in a new 1.5 ml microfuge tube and either frozen at -80°C for storage, or used immediately for western blot analysis.

2.7 RNA tethering reporter assay

2.7.1 Cell seeding and transfection for mRNA analysis

HEK293T cells were seeded into 12-well culture plates at 80,000 cells per well. After 24 hours, cells were at 70% confluency and plasmids were co-transfected using the CaPO₄ precipitation transfection method as described in section (2.6.2 Cell seeding and transfection for Click -iT EdU labelling). In total, 1 μ g of plasmid DNA was transfected per well (Table 2-11). After 20-30 mins

incubation the mixture was added drop wise to the cells in the 12-well culture plate. Sixteen hours after transfection, the cell medium was removed. Cells were washed with PBS and then this was replaced with fresh medium to reduce cell toxicity. Twenty-four hours after transfection, cells were harvested for the luciferase assay.

Component/Reagent	Volume
500 ng of each DNA plasmid (pRL-5boxB and pCI λ N)	x μ l
2 M CaCl ₂	6 μ l
dH ₂ O	Up to 50 μ l
2 x HBS	50 μ l

Table 2-11 Transfection set up used for calcium phosphate precipitation method.

2.7.2 RNA tethering reporter assay

For the luciferase assay, cells were harvested 24 hours post-transfection. Cells were washed with 1 ml of PBS, which was removed before adding 100 μ l of 1x cell lysis buffer (luciferase cell lysis buffer, NEB B3321S) directly to the cells in the 12-well culture plate. The plates were placed at -80°C for 15 mins until completely frozen, or were kept at -80°C until the next day. Then, samples were thawed and mixed by pipetting to aid the lysis process before being transferred into 1.5 ml microfuge tubes. Samples were then centrifuged at 16,000g for 5 mins at 4°C and stored at -80°C or used immediately.

The Gluc assay solution (BioLux Gaussia Luciferase Assay Kit, NEB E3300S) was prepared according to the manufacturer's instructions. The BioLux GLuc substrate was diluted (1:100) with BioLux GLuc assay buffer and mixed by gentle inversion. In a white, opaque, flat bottomed 96-well plate, 10 μ l of cell lysate was added and the plate was placed into a GLO max microplate luminometer. For each well, after injection of 50 μ l of GLuc assay solution, the chemiluminescence was measured for 2 secs after a 5 secs delay.

A Bradford assay was conducted for all samples (section 2.9.2) to determine total protein concentrations, which were then used to normalise the luciferase activity.

2.7.3 Preparation of RNA

Total RNA was extracted using the Macherey-Nagel RNeasy kit, which included treatment with RNase-free DNase. Twenty-four hours after transfection, cells were washed with 1 ml PBS and then dislodged from the culture surface by pipetting. The cells were transferred into a 1.5 ml microfuge tube and then pelleted by centrifugation at 11,000 g for 5 mins. The supernatant was removed and the pellet re-suspended in a mixture of 350 μ l RLT lysis buffer + 1.5 μ l β -mercaptoethanol. RNA extraction was then conducted as stated in the manufacturer's protocol; the RNase-free DNase was added directly to the column and incubated at room temperature for 15 mins before RNA elution was performed using 50 μ l of RNA-free water. RNA samples were stored at -80°C.

2.8 RNA analysis

2.8.1 Production of cDNA

For the production of cDNA free from plasmid DNA contamination, QuantiTect Reverse Transcriptase was used (Qiagen). The reaction mixture was set up as shown in Table 2-12. Then, this mixture was incubated in a Peqlab Primus 96 advanced thermal cycler at 42°C for 2 mins, followed by incubation on ice for 1 min. The second master mix was prepared containing the reverse transcriptase and buffers, as shown in Table 2-13, which was then added to the sample. This mixture was placed in a Peqlab Primus 96 advanced thermal cycler and incubated at 42°C for 15 mins. After that, the mixture was incubated at 95°C for 3 mins to deactivate the enzyme. The resulting cDNA was diluted 1:5 with RNA-free water and stored at -20°C until required.

Volume	Component/Reagent
--------	-------------------

x μ l	300 ng RNA template
1.0 μ l	gDNA wipeout buffer x7
up to 7.0 μ l	RNase-free H ₂ O

Table 2-12 RNA tethering assay: cDNA synthesis reaction mix 1.

Volume	Component/Reagent
7 μ l	Reaction mix 1
2 μ l	Reverse transcriptase buffer x5
0.5 μ l	Reverse transcriptase
0.5 μ l	RT primer mix

Table 2-13 RNA tethering assay: cDNA synthesis reaction mix 2.

2.8.2 Quantitative PCR analysis

All qPCR reactions were performed using TaqMan® Advanced Master Mix (Applied Biosystems). Samples were tested in technical triplicate using primers specific for the gene being tested and separately with primers for a housekeeping gene (GAPDH (Hs02758991-g) which were used for normalisation. A master mix was produced containing the appropriate primers (Table 2-14) and the TaqMan® Advanced Master Mix (Table 2-15).

Primers	Sequence
RT GAPDH FW	5' TCTGGTAAAGTGGATATTCTTG 3'
RT GAPDH RV	5' CATGGTGATGGGATTTCC 3'
RT-Luciferase FW (1)	5' TCGTCCATGCTGAGAGTGTC 3'
RT-Luciferase RV (1)	5' CTAACCTCGCCCTTCTCCTT 3'
RT-Luciferase FW (2)	5' CGAGCACCAAGACAAGATCA 3'
RT-Luciferase RV (2)	5' GTAGGCAGCGAACTCCTCAG 3'
Luciferase probe	FAM 5'-agtTcgCtgCctAcctgga-3'NFQ-MGB

Table 2-14 List of primers used for qPCR.

Volume	Component/Reagent
1.0 µl	cDNA template
5.0 µl	TaqMan® Fast Advanced Master Mix
1.0µl	TaqMan® probe (2.5 µM)
1.0µl	Forward primer (10 µM)
1.0µl	Reverse primer (10 µM)
1.0µl	RNase-free H ₂ O

Table 2-15 qPCR reaction setup.

2.9 Protein analysis

2.9.1 Reagents, stock solutions and buffers for use in protein analysis

4x Upper buffer: 0.5 M Tris base, 0.4% SDS, pH6.8; stored at room temperature

4x Lower Buffer: 1.5 M Tris base, 0.4% SDS, pH8.8; stored at room temperature

10x Running buffer: 0.25 M Tris Base, 1.0% SDS, 1.92 M glycine; stored at room temperature

10x Transfer buffer: 0.25 M Tris Base, 1.92 M glycine; stored at room temperature

4x SDS loading buffer: 2.4 ml 1 M Tris buffer pH 6.8, 4 ml 100% glycerol, 0.5ml β -mercaptoethanol, 0.8 g SDS, 4 mg bromophenol blue and 3.1 ml distilled water. Stored at -20°C

Tris-buffered saline supplemented with 0.05% Tween-20 (TBST): 50mM Tris-HCl pH 7.8, 150mM NaCl , 0.1% Tween-20. Stored at room temperature.

2.9.2 Bradford assay to determine protein concentration

A Bradford assay was used to determine protein concentration using Bradford reagent from Sigma (catalog number B6916). BSA (NEB 20 mg/ml) was diluted to 10 μ g/ μ l with lysis buffer used in the protein isolation method. Standard curves were made using different concentration of BSA (0, 2, 4, 6, 8, 10) μ g/ μ l. Ten microlitres of protein sample was added to 200 μ l of Bradford reagent and made up to a total volume of 1 ml with water, then vortexed and incubated for 5-20 minutes at room temperature. Samples were transferred to 1ml cuvettes and the absorbance read at 600 nm. The standard curve was obtained by linear regression analysis using Microsoft Excel and this was used to calculate the protein concentration of the samples.

2.9.3 Preparation of sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE)

The Invitrogen X-Cell SureLock Mini Cell system was used for SDS-PAGE analysis. The separating gels were prepared at concentrations of 12% and 14% as shown in Table 2-16, depending on the molecular mass of the protein of interest. Firstly, the separating gel was cast in a gel cassette, (Invitrogen) leaving ~1.5 cm of room at the top for the stacking gel to be cast. Immediately after pouring into the cassette, 500 μ l of isopropanol was added on the top and the gel was left to set for 15-20 mins. Once complete, the isopropanol was removed by inverting the gel and draining onto tissue paper. Then the stacking gel was cast on top and the appropriate comb was inserted immediately after pouring. The gel was left to stand at room temperature for 30 mins before being used or stored at 4°C. The comb was removed before use and the gel was washed with 1 \times running buffer. Protein extracts were denatured by boiling for 5 mins in 1x SDS sample buffer immediately before loading. The protein ladder that was used was from NEB (Color Prestained Protein Standard, Broad Range 11–245 kDa). Gels were run in 1x running buffer at 180V for approximately 1.5 h.

	Resolving Gel		Stacking gel
	12%	14%	4%
Protein MW range, kDa	20-150	10-80	
4x Lower Buffer	2 ml	2 ml	-
4x Upper Buffer	-	-	750 μ l
40%acrylamide: Bisacrylamide (29:1)	2.4 ml	2.8 ml	300 μ l
10% APS	80 μ l	80 μ l	60 μ l
TEMED	8 μ l	8 μ l	6 μ l
H ₂ O	3.6 ml	3.2 ml	1950 μ l

Table 2-16 Preparation of SDS-PAGE gels.

2.9.4 Gel staining using Coomassie Brilliant Blue

To stain SDS-PAGE gels, Coomassie Brilliant Blue G was used (250 mg/ml dissolved in 2 L H₂O and 200µl of HCl at least 3 hours before use). Firstly, the SDS-PAGE gel was removed from the cassette and rinsed in dH₂O in a suitable container with a lid, then heated in the microwave on high power for 30 secs and agitated on a rocking table for 5 mins. The washing step was repeated again. The gel was then covered with Coomassie Brilliant Blue stain and heated on high power in the microwave for 10 secs. Following this, the gel was incubated in the Coomassie stain for 45 minutes on a rocking table, before being destained with dH₂O. During the destaining, the water was changed twice every 10 mins and after the last change of water, it was left on the rocking table overnight before being imaged using the Fujifilm LAS-4000 system. Image analysis was carried out using ImageJ.

2.9.5 Western blotting and immunodetection

For western blotting, proteins were separated on SDS-PAGE gels and transferred to nitrocellulose membranes (Whatman Protran 0.45µm), using the Invitrogen X-Cell SureLock Mini Cell system for 1 h at 25V in transfer buffer. Then, membranes were blocked with 5% dried milk dissolved in TBST for 1 hour at room temperature. After that, membranes were transferred into 50 ml Falcon tubes containing 5% dried milk with TBST and the specific primary antibody. These were incubated on a rotator overnight at 4°C. Before incubating the membranes with secondary antibody, they were washed three times with TBST buffer for 5 mins each time. Then, the membranes were incubated in 5% dried milk containing TBST and the specific secondary antibody for one hour on a rotator at room temperature. Again, membranes were washed three times with TBST. Finally, for immunodetection, membranes were incubated with 600µl of ECL western blotting substrate (Pierce) for 1 min prior to imaging. The Fujifilm LAS-4000 system and image reader computer software (Fujifilm, Japan) were used for imaging. Image

analysis was carried out using ImageJ. All antibodies used are listed in Table 2-17 and Table 2-18).

Primary Antibodies	Dilution	Clone	Isotype	Supplier
Mouse anti-Gal4 TA	1/1000 ^a	C-10	IgG _{2b}	Santa Cruz (sc-1663)
Rat anti-HA	1/500 ^a	3F10	IgG ₁	Roche (1 867 423)
Goat anti-γ tubulin	1/1000 ^a	C-20	Polyclonal	Santa Cruz (sc-7396)
Mouse anti-Flag	1/1000 ^a	M2	IgG ₁	Sigma (F1805)
Mouse anti His	1/2000 ^b	H-8	IgG _{2b}	Thermo Fisher Scientific (MA1-21315)
Mouse anti-GST	1/5000 ^a	B-14	IgG ₁	Santa Cruz (SC-138)
Rabbit anti-LexA	1/1000 ^a	-	Polyclonal	Merck Millipore(06-719)

^aAntibodies were diluted in TBST containing 5% dried milk powder.

^bAntibodies were diluted in TBST containing 5% BSA.

Table 2-17 Primary antibodies used for western blotting.

Secondary Antibodies	Dilution	Isotype	Supplier
Chicken anti-Mouse HRP	1/1000	All isotypes of IgG	Santa Cruz (sc-2954)
Goat anti-Rat HRP	1/1000	All isotypes of IgG	Santa Cruz (sc-2006)
Donkey anti-Goat HRP	1/1000	All isotypes of IgG	Santa Cruz (sc-2020)

Table 2-18 Secondary antibodies used for western blotting.

2.10 Co-Immunoprecipitation analysis

2.10.1 Preparation of protein lysates for immunoprecipitation

HEK293T cells were seeded at 80% confluency in culture dishes (6 cm diameter). After 24 hours, cells were transfected with two appropriate expression vectors using the calcium phosphate precipitation method (section 2.6.2 Cell seeding and transfection for Click -iT EdU labelling). In each culture dish, a total of 5 μg of plasmid DNA was transfected (2.5 μg of each plasmid) as shown in Table 2-19. 16 hours post-transfection, cells were washed twice with PBS before being replaced with warm medium. Cells were harvested 48 hours post-transfection by pipetting the culture medium up and down to dislodge the cells. Then the cell suspensions were transferred into 15 ml Falcon tubes and centrifuged at 4000 g, 4°C for 5 mins, after which the media was discarded. Cell pellets were washed twice with cold PBS and the centrifugation step was repeated. After that, the cell pellets were re-suspended in 500 μl of cold cell lysis buffer (0.2% NP-40) and transferred into 1.5ml tubes. Cells were then lysed by freeze/thawing twice at -80°C, followed by centrifugation at 16,000 g, 4°C for 10 mins. The supernatants were transferred into new 1.5 ml tubes and either stored at -80°C or used directly for the assay.

Component/Reagent	Volume
2.5 μg of each DNA plasmids	x μl
2 M CaCl ₂	24 μl
dH ₂ O	Up to 200 μl

Table 2-19 Transfection setp used for co- immunoprecipitation.

2.10.2 Preparation of antibody-agarose beads and immunoprecipitation

For immunoprecipitation using A and G agarose beads (Roche), the following protocol was followed and scaled up depending on the number of samples. 10 μ l of protein A and 10 μ l of protein G agarose beads were mixed together in a 1.5 ml tube, after which ethanol was removed by centrifugation at 3000 g for 30 secs. Then, the beads were washed twice with 500 μ l of ice cold-PBS and centrifugation was repeated; between each centrifugation, the tubes were placed on ice for 1 min to allow the beads to settle down, before removing the supernatant. After that, either anti-HA or anti-FLAG antibody was added, depending upon the experiment requirements: approximately 1 μ g of antibody was added to 10 μ l of beads (2 μ g per immunoprecipitation), as shown in Table 2-20. This volume was then increased with 500 μ l of ice cold-PBS and the beads were incubated with the antibody for 1 h at room temperature on a rotator. After that, the mixture of beads and antibody was centrifuged at 3000 g for 30 secs. Following this, the tube was placed on ice for 1 min before removing the supernatant with a gel loading tip. The previously mentioned washing step was repeated three times and in the final wash the beads were re-suspended with lysis buffer (NP-40 0.2%). The two antibodies coated with agarose beads (protein A/protein G) were added to a 1.5 ml microfuge tube, along with 450 μ l of protein lysate. This microfuge tube was then packed into a 50 ml Falcon tube and the mixture was incubated overnight at 4°C on a rotator. The lysate/bead mixtures were centrifuged at 3000 g for 30 secs and the supernatant carefully discarded using a gel loading tip. Beads were then washed three times with 500 μ l of lysis buffer containing 0.2% NP-40. After the final wash, the beads were re-suspended in 10 μ l PBS, followed by the addition of 10 μ l of 4 x SDS loading buffer. Each sample was boiled for 5 mins to denature the protein-antibody complexes. The samples were then subjected to western blot analysis or stored at -80°C until required.

Primary Antibody	Dilution	Supplier
Anti-Flag M2	1 µg per 10 µl of agarose beads	Sigma (F1805)
Anti-HA	1 µg per 10 µl of agarose beads	Roche (1 867 423)

Table 2-20 List of antibodies used for co-immunoprecipitation.

2.11 GST and His pull down analysis

2.11.1 Protein expression: GST- Tagged BTG1 and BTG2

Plasmids pGEX-4T1, pGEX4T1-BTG1 and pGEX4T1-BTG2 were transformed into *Escherichia coli* strain BL21 (D3) by heat shock, after which transformants were plated on ampicillin agar plates. A single colony was grown as a starter culture in 5 ml of LB containing 100 µg/ml ampicillin at 37°C overnight. 2.5 ml of the overnight culture was used to inoculate 200 ml of LB containing 100 µg/ml ampicillin, which was then incubated at 37°C, 160 rpm until the OD₆₀₀ reading was between 0.6 and 0.8. The expression of the proteins of interest (IPTG) at a final concentration 0.2 mM for 3 hours at 30°C. After that, the cells were harvested by centrifugation using an Eppendorf 5810R centrifuge at 3000 g 4°C for 30 mins. The supernatant was discarded and the bacterial pellet was re-suspended in 1.5 ml of chilled lysis buffer (20 mM Tris-HCl pH 7.8, 300 mM NaCl, 10% glycerol and 1 mM β-mercaptoethanol). Cells were frozen and kept at -80°C until further use. When ready, the bacterial suspension was thawed and then lysed by sonication on ice using a Qsonica XL2000, with five cycles of 10secs on/10secs off until the solution appeared clear. This lysate was then centrifuged at 16,000 g for 10 mins at 4°C to remove insoluble material. The supernatant was collected into a fresh 1.5 ml microfuge tube and stored until further use at -80°C.

2.11.2 Protein expression: His Tagged-PABPC1(1-190) and PABPC1(1-99)

Plasmids pQE80L-PABPC1(1-190) and pQE80L-PABPC1(1-99) (a gift of L.Pavanello, school of pharmacy) were transformed into *Escherichia coli* strain BL21 (D3) by heat shock, after which transformants were plated on ampicillin agar plates. 5 ml of LB containing 100 µg/ml ampicillin was inoculated with a single colony and cultured at 37°C overnight. 2.5 ml of the starter culture was used to inoculate 200 ml of LB containing 100 µg/ml ampicillin, which was then cultured shaking at 160 rpm, 37°C, until the OD₆₀₀ was between 0.6 and 0.8. The expression of proteins of interest was induced and lysed as the previous section (2.12 Protein expression: GST- Tagged BTG1 and BTG2) using different lysis buffer (20 mM Tris-HCl pH 7.8, 300 mM NaCl, 5% glycerol, 0.1% NP-40, 1 mM β-mercaptoethanol, and 40mM imidazole). This lysate was then centrifuged at 16,000 g for 10 mins at 4°C to remove insoluble material. The supernatant was collected to a fresh 1.5ml microfuge tube and stored until further use at -80°C.

2.11.3 Large scale protein expression: His-Tagged PABPC1 (1-190) and PABPC1(1-99)

Plasmids pQE80L-PABPC(1-190) and pQE80L-PABPC1(1-99) were transformed into *Escherichia coli* strain BL21 (DE3). A single colony was used to inoculate 1 ml of LB containing 100 µg/ml ampicillin and was grown for 6 hours at 37°C. Then the starter culture was used to inoculate 50 ml of LB containing 100 µg/ml ampicillin, which was incubated at 37°C, overnight. This overnight pre-culture was diluted in 1L of LB and cultured at 37°C until the OD₆₀₀ was between 0.6 and 0.8. Expression of the proteins of interest was then induced by the addition of IPTG at a final concentration of 0.2 mM and culturing at 18°C overnight. Once complete, the cells were harvested by centrifugation at 3000 g 4°C for 60 mins. The supernatant was discarded and the bacterial pellet was re-suspended in 3-5 ml of ice cold extraction buffer for

His-PABPC1 (20 mM Tris-HCl pH 7.8, 300 mM NaCl, 5% glycerol, 1 mM β -mercaptoethanol and 40 mM imidazole). The cell suspension was frozen and kept at -80°C until further use. When ready, the bacterial suspension was thawed and then lysed by sonication on ice using a Qsonica XL2000, with five cycles of 30secs on/30secs off. The crude lysate was then transferred into 1.5 ml tubes and centrifuged at 16,000 g, 4°C for 30 mins to remove insoluble material. The supernatant was collected and stored until further use at -80°C . For purification, this protein extract was thawed before being applied to the His-trap column (GE healthcare) in the AKTA machine. Proteins were eluted using an elution buffer containing 20 mM Tris-HCl pH 7.8, 500 mM NaCl, 10% glycerol, 1 mM β -mercaptoethanol & 200 mM imidazole.

2.11.4 GST- pull down of His-PABPC1 (1-99) or (1-190) using GST- BTG1 and GST-BTG1 and BTG2

For GST pulldown, 20 μl of glutathione-agarose beads (Catalog Number G4510, Sigma), were added to a 1.5 ml tube and equilibrated with 500 μl of chilled binding buffer by centrifugation at 3000 g for 30 secs. The tubes were placed on ice for 1 min to allow the beads to settle, before removing the supernatant with gel loading tips. Following this, up to 1 ml of the lysate was added to the beads and incubated on the rotator overnight at 4°C . After that, the mixture of lysate and beads was washed three times with 500 μl of chilled binding buffer (20 mM Tris-HCl pH 7.8, 150 mM NaCl, 10% glycerol, 0.1% NP-40 and 1 mM β -mercaptoethanol), each time being centrifuged at 3000 g for 30 secs. The tubes were placed on ice between the washing steps and the buffer was removed by gel loading tips to avoid aspiration of the beads. Then, 100 μg of purified protein tagged with His was added to the mixture of cell lysate and beads; this was made up to a final volume of 500 μl with binding buffer. This mixture was mixed on the rotator for 2-3 hours at 4°C or overnight. Once complete, the beads were washed three-four times with 500 μl of chilled binding buffer as described above. Finally, 20 μl of 2 \times SDS loading buffer was

added and the sample was boiled for 5 mins at 95°C. Samples were analysed by SDS-PAGE (14% gels) and immunoblotting.

2.11.5 His- pulldown of GST-BTG1 and GST BTG2 using His tagged with PABPC1 (1-190) or PABPC1 (1-99)

For His-tag pulldown, 20µl of Ni-NTA agarose beads (Catalog number 30210, Qiagen) were placed in a 1.5 ml microfuge tube. Then, the beads were equilibrated with 500 µl of lysis buffer (20 mM Tris-HCl pH 7.8, 300 mM NaCl, 5% glycerol, 0.1% NP-40, 40mM imidazole and 1 mM β-mercaptoethanol) by centrifugation at 3000 g for 30 secs. After this, the tube was placed on ice for 1 min to allow the beads to settle before removing the supernatant with gel loading tips. 750 µl of the lysates from two cultures which expressed proteins tagged with His and GST was added to the beads, as shown in Table 2-21. The mix of lysates and beads was incubated on the rotator for one hour at 4°C. Then, the beads were washed three times with 500 µl of chilled lysis buffer, followed by centrifugation at 3000 g for 30 secs after each wash. The tube were placed on the ice between each washing step and the buffer was removed with gel loading tips to avoid aspiration of the beads. After that, the beads were boiled with 20µl of 2× SDS loading buffer for 5 mins at 95°C, before 10 µl was analysed by SDS-PAGE (14%) or 2 µl by immunoblotting.

	His-PABPC1(1-190)	GST	GST-BTG1	GST-BTG2
1	+	-	-	-
2	+	+	-	-
3	+	-	+	-
4	+	-	-	+
5	-	-	+	-
6	-	-	-	+

Table 2-21 List of the samples tested in His-pull down experiments.

2.12 Statistical analysis

One way Anova and student paired t-test were carried out using Graph Pad Prism version 7.01. All experiments were carried out in triplicate, repeated at least two to three times independently and the results are presented as the mean \pm standard error. A one-way ANOVA (dunnett's post-hoc method) was conducted yeast two hybrids experiment and one-way ANOVA (tukey's post-hoc method) was conducted in both proliferation assay and RNA tethering reporter assay. Both methods calculate the significance of the values (*p <0.05, **p <0.01, ***p <0.001 and ****p <0.0001) compared to wildtype BTG1 and empty vector (control). An unpaired t-test was used in the optimization the proliferation assessment, and calculates the significance of the values (*p<0.05, **p<0.01 and ****p<0.0001) when comparing BTG1-expressing cells to the empty vector cells (control).

Chapter 3

Identification of somatic mutations in BTG1 and BTG2 in non-Hodgkin's lymphoma

Chapter 3: Identification of somatic mutations of BTG1 and BTG2 genes in non-Hodgkin's lymphoma

3.1 Introduction

Non-Hodgkin's lymphoma (NHL) represents a highly heterogeneous group of lympho-proliferative disorders, characterized by the occurrence of multiple genetic mutations. The three most common types are follicular lymphoma (FL), Burkitt's lymphoma (BL), and diffuse large B-cell lymphoma (DLBCL). Based on the 2008 WHO classification of lymphoid neoplasms, the two molecular subtypes of DLBCL are germinal centre B-cell (GCB) and activated B-cell (ABC) lymphoma (Swerdlow et al., 2008, Blombery et al., 2015). Certain molecular genetic abnormalities of NHLs are related to specific lymphoma subtypes and are associated with the patient's prognosis and the choice of therapeutic target. For example, DLBCL is closely associated with BCL6, BCL2, and MYC gene translocations (Guo et al., 2016, Miles et al., 2016). In recent years, multiple research groups have used either exome or RNA-seq to determine the genes targeted by somatic variants in NHLs including those involved in transcription, histone modification, and tumour suppression (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013). Furthermore, groups have identified novel BTG1 and BTG2 somatic variants. However, it is not clear whether these variations result in gain or loss of function (Morin et al., 2011).

BTG1 and BTG2 are part of the anti-proliferative BTG/TOB family, sharing similarities within the conserved BTG and C-terminal domains (Winkler, 2010). Loss of expression of BTG2 has been found in breast cancer (Kawakubo et al., 2006) and loss of expression of BTG1 in solid tumours results has been shown to result in poor clinical outcomes (Sheng et al., 2014, Sun et al., 2014). In addition, 9% of paediatric BCP-ALL cases commonly show small deletions of BTG1 suggesting a deletion may contribute to leukaemogenesis (Kuiper et al., 2007). BTG1 and BTG2 expression is required in multiple stages of normal

early B-cell development and differentiation and this provides an important link to their role as tumour suppressors in B-cell malignancies (Tijchon et al., 2016). Furthermore, it has been suggested that low expression of BTG1 with downregulation or loss function of CNOT6, plays a role in the pathogenesis of a subset of HOXA- positive T ALL- with terminal 5q deletions (La Starza et al., 2016). In addition, CNOT3 loss of function mutations were found in 7% in T-ALL patients (De Keersmaecker et al., 2013). Both CNOT3 and CNOT6 encode members of the Ccr4-Not complex and BTG1 is involved with the Ccr4-Not complex by interacting with the hCaf1(CNOT7/CNOT8) subunit (Prévôt et al., 2001, Collart and Panasenko, 2012).

The Catalogue of Somatic Mutations in Cancer (COSMIC) is a resource on somatic mutations for genes known to be involved in human cancers. It provides scientific literature on tumour resequencing data from the cancer genome project and the Sanger Institute in the United Kingdom (Forbes et al., 2010). Ensembl is a genome browser but not specifically based on cancer type. It is based on the alignment of biological sequences, including complementary DNAs (cDNAs), proteins, and RNA-seq reads, to the target genome in order to construct candidate transcript models (Flicek et al., 2011). Many single nucleotide variants (SNVs) found in human samples and human cell lines are annotated in the genome browser, which is therefore a useful resource for the identification of common variants. However, variants need to be prioritised to identify those that may potentially affect protein function and are associated with disease. Thus, algorithms such as Sorting Tolerant From Intolerant (SIFT) or Suspect (Disease-Susceptibility-based single amino acid variants (SAVs) Phenotype Prediction) can help in this respect (Kumar et al., 2009, Yates et al., 2014). SIFT works by predicting the potential impact of amino acid substitutions on protein function by calculating the probability for all 20 amino acids at that position, and alignment of the protein family sequence. Amino acids in highly conserved positions are scored by SIFT as intolerant for most substitutions, while, amino acids in poorly conserved positions may

tolerate substitution (Kumar et al., 2009). The SIFT tool has shown 69% accuracy (Kumar et al., 2009). Alternately, the Suspect method calculates not only the conserved protein sequence but also the protein-protein interactions (PPIs) or network feature and associated diseases resulting in 82% accuracy. Studying the effect of SAVs in the network feature depends on the protein structure (Yates et al., 2014).

Recently, BTG1 and BTG2 variants have been reported in NHL samples. However, it is unclear whether these variants affected protein function. Therefore, data of BTG1 and BTG2 variants that have recently been reported in the genomic browsers (COSMIC v65 and Ensembl v70) or literature reports were collected. Analysis of the BTG1 and BTG2 sequence in several different lymphoblastoid cell lines was also performed to identify possible novel variants. Finally, the algorithms for software prediction (SIFT and Suspect) were used to analyse the effect of these variants on protein function. Based on the results presented in this chapter, a number of variants were selected for further study.

3.2 Sequencing of BTG1 and BTG2 genes in lymphoblastoid cell lines

To identify possible novel mutations of BTG1 and BTG2 in cell lines, the coding regions of BTG1 and BTG2 were screened using DNA sequencing. The five human cell lines used were derived from Burkitt's lymphoma (Raji), human Hodgkin's lymphoma (KM-H2), and B-cell lymphoma (DB); the other two cell lines were derived from disease-free individuals (GM03201 and GM1953).

3.2.1 Amplification and analysis of BTG1 sequence in lymphoblastoid cell lines

To analyse the coding sequences of BTG1, amplification of BTG1 cDNA was attempted. Because the amplification was unsuccessful, the genomic DNA was isolated from the cell lines and amplified using exon 1 and exon 2 primers described before (Waanders et al., 2012).

The BTG1 exon sequences were then amplified using flanking intron-based primers. BTG1 contains two exons; exon 1 has 150 base-pair nucleotides (the encoding first 50 amino acids), and exon 2 contains the major body of the open reading frame (466 amino acids). To screen for exon 1 and 2 of BTG1 in all five cell lines, the PCR products were resolving on a 2% agarose gel. The size was as expected (0.6 kb) and Sanger sequencing was carried out (Source Biosciences) (Figure 3-1). In order to analyse the sequencing results, the Basic Local Alignment Search Tool (BLAST) was used to align and compare all cell-line nucleotide sequences with reference wild-type BTG1 transcript (BTG1: Chromosome 12: 92, 140, 278-92, 145, 897).

The results for exon 1 showed that three cell lines, GM03201, KM-H2, and DB are 100% identical to the reference gene. The result of the GM01953 cell line was 93% identical to the reference gene, because the sequence quality was poor the remaining of DNA sequence could not be identified. Two variants (A16P and I22L) were discovered in the Raji cell line, however, the quality of the sequencing was poor and the presence of mutations was not confirmed (Figure 3-2). When analysing exon 2 BTG1 sequences, all cell lines had

sequence homology (100% identical with reference gene using BLAST), and no mutations or deletions were detected.

3.2.2 Amplification and analysis of BTG2 sequence in lymphoblastoid cell lines and variants detected in KM-H2 cell line

The sequence analysis of BTG2 used cDNA amplification instead of genomic DNA as used in the BTG1 analysis. Total RNA was isolated from the five lymphoma cell lines and used to prepare cDNA. The primers were designed to amplify the open reading frame of BTG2, based on the protein sequence of 158 amino acids, and resolved on a 2% agarose gel resulting in a band of the expected size of 0.7 kb (Figure 3-3 A and B). PCR products were analysed by Sanger sequencing (Source Biosciences). To analyse the sequence results, BLAST was used to align and compare the nucleotide sequences with the reference wild-type BTG2 transcript (NCBI reference sequence: NP_006754.1). The results showed that all cell lines contain BTG2 identical to the reference sequence, apart from the KM-H2 lymphoma cell line which contained a single-base substitution at the end of the coding sequence. This variation occurs at DNA sequence position 461 with substitution of the G nucleotide to A. As a result of this variant, the valine (GTG) residue at position 153 is altered to methionine (ATG) c.461G>A; p.V153M. Notably, the sequencing trace of this cell line shows two peaks in the position of the mutation instead of one. These peaks correlate with two alleles, A and G, suggesting the mutation mono-allelic (Figure 3-3 C).

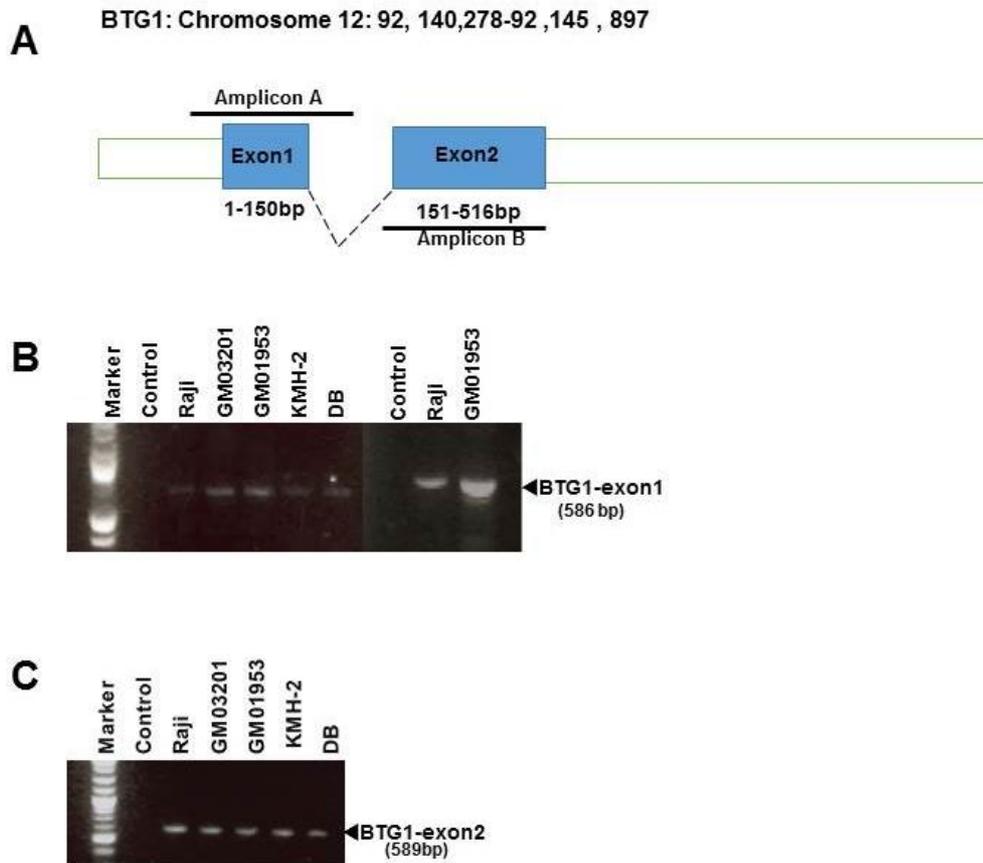


Figure 3-1 Amplification of BTG1 in lymphoblastoid cell lines.

(A) Schematic of the wild-type human BTG1 gene composed by two exons coloured blue, the dotted lines, and the white box in white colour indicated intron. Exon 1 covers 150 bp of the codon region, while exon 2 covers 151–516 bp. The PCR primers used to detect expression of the wild-type BTG1 transcript are indicated as primers A for exon 1 and primers B for exon 2. (B) Agarose gel showing the amplification of isolated genomic DNA by PCR of BTG1-exon1 in five different cell lines including Raji, GM03201, GM01953, KM-H2, and DB. PCR was repeated in two cell lines, GM01953 and Raji, under the same conditions. (C) Agarose gel showing amplification of isolated genomic DNA by PCR of BTG1-exon2 in the same cell lines mentioned above. DNA 2-log ladder (marker) and no DNA template (negative control) were used. All experiments were performed at least three times.

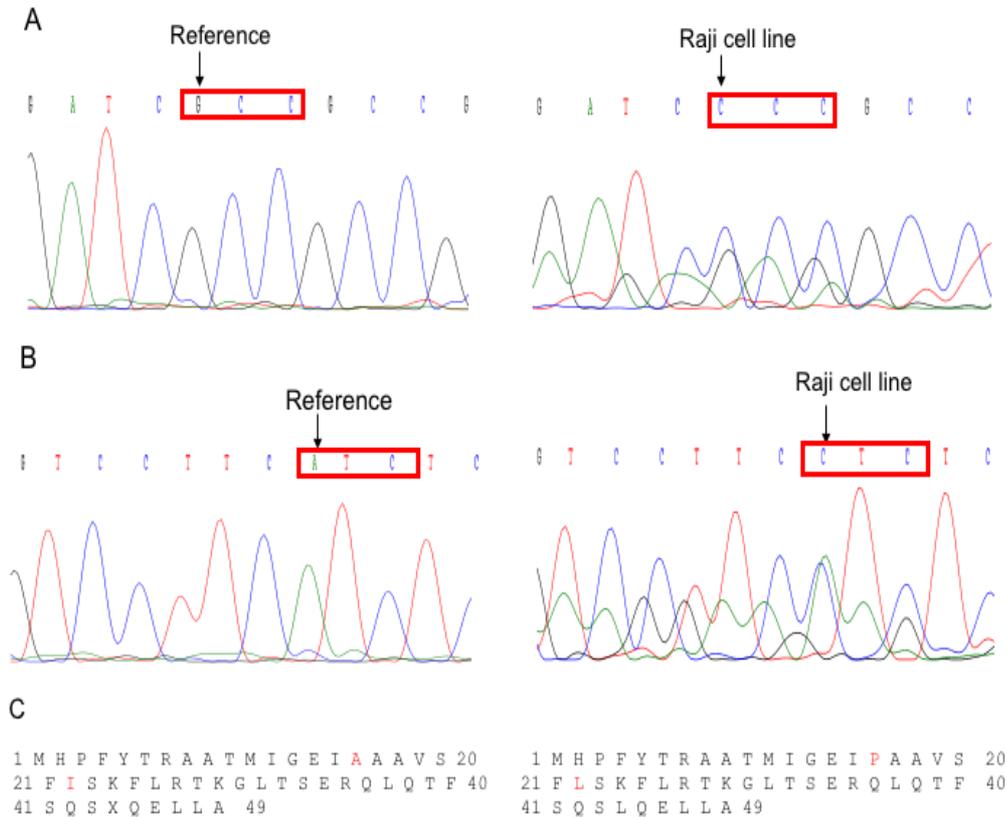


Figure 3-2. Analysis of BTG1 sequence variants detected in the Raji cell line.

Chromatograms are shown representing the genomic BTG1 sequences of two variants detected in the Raji cell line. (A) The first amino acid substitution identified was alanine (GCC) to proline (CCC) in position 16. (B) The second amino acid substitution identified was isoleucine (ATC) to leucine (CTC) in position 22. (C) The amino acid sequence of BTG1 (exon 1) comparing the reference (wild type) to the mutated residues found in the Raji cell line. All experiments were performed at least three times.

NCBI Reference Sequence: NP_006754.1

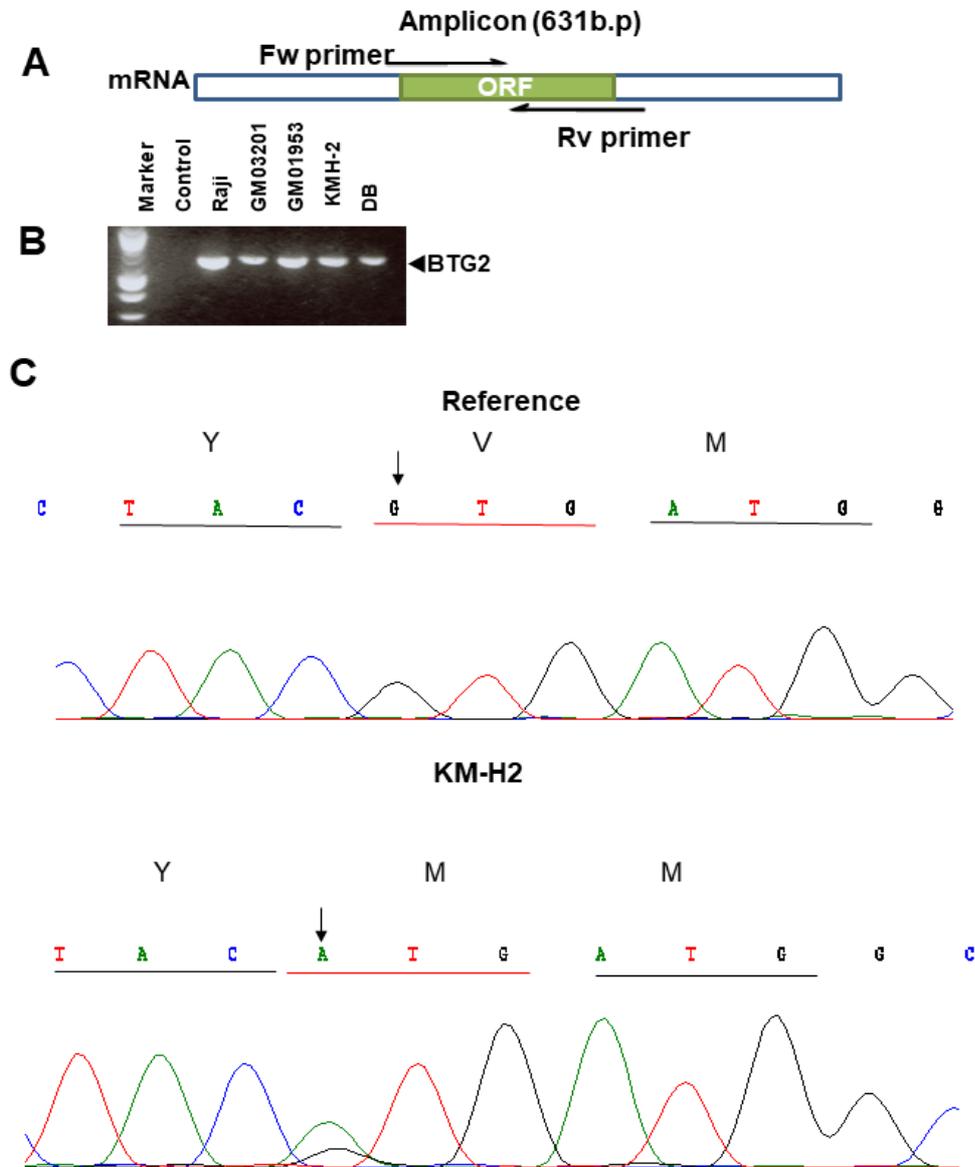


Figure 3-3. Amplification and analysis of BTG2 sequence variant detected in KM-H2 and lymphoblastoid cell lines.

(A) A schematic of wild-type human BTG2 contains the open reading frame (ORF), primers designed to cover the coding region of BTG2 (477bp) and the additional nucleotides from the intron region, about 100 bp from each side. (B) Agarose gel showing the amplification of BTG2 by PCR in five different cell lines including Raji, GM03201, GM01953, KM-H2, and DB. No DNA template as negative control and DNA 2-log ladder (marker) were used. (C) Chromatogram corresponding, to BTG2 transcript sequence variant detected in KM-H2 cell line. The amino acid substitution identified was from alanine (CTG) to methionine (ATG) in position 153. The letters above the sequence indicate the codon. . All experiments were performed at least three times.

3.3 Data collection and bioinformatics analysis of mutations in BTG1 and BTG2

Recently, the literature reported novel mutations in BTG1 and BTG2 in different types of NHL (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013). However, the effect of these mutations on protein function is unknown. In order to select sequence variants for this project, the Ensembl genome browser and the COSMIC database were used. Ensembl is a genome browser for vertebrate genomes that supports annotation of sequence variations, predicts regulatory function and collects disease data (Flicek et al., 2011). The COSMIC database is a resource of somatic mutations for genes known to be involved in human cancers (Forbes et al., 2010). Both resources were used in a combination with bioinformatics analysis tools such as SIFT and Suspect (Kumar et al., 2009, Yates et al., 2014).

3.3.1 Identification of BTG1 and BTG2 variants according to type of NHL

In order to identify BTG1 and BTG2 variants in lymphomas, data was collected using the Ensembl v70 genome browser and the COSMIC v65 database as well as the literature that had reported somatic mutations in BTG1 and BTG2 in NHL samples. Out of a total of 76 mutations identified, 33 were of non-specified tissue origin or in cancers such as lung, large intestinal, ovary, kidney, and others (Appendix A). Mutations of Ccr4-Not genes were also collected and reported in different types of cancers such lung, kidney, ovary and large intestine and others. Only three variants were identified in haematopoietic tissues; namely CNOT2 (R137G) and CNOT3 (E20K, E70K) and frequent mutations in CNOT3 identified in T-ALL samples (De Keersmaecker et al., 2013) (Appendix A).

In this chapter, 45 somatic mutations in BTG1 or BTG2 found in lymphoma tissue were analysed. These mutations were originally found using whole exome sequencing and RNA-seq on over 100 NHL samples (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013). All

the somatic mutations were identified via RNA-seq and were validated using the PCR amplification and Sanger sequencing in tumour and normal DNA (Morin et al., 2011). In the whole exome sequencing, validation of selected mutations was targeted and re-sequenced using microfluidic PCR (Access array system, Fluidigm) and the MiSeq sequencing system (Illumina). Moreover, tumour and matched normal samples were selected based on the presence of the indicated mutations using whole exome sequencing. The common generic variants that occur in normal exomes were eliminated (Lohr et al., 2012, Zhang et al., 2013).

The majority of the mutations were found in samples with DLBCL, followed by BL and FL. 32 of the somatic variants (72.7 %) were discovered in DLBCL; 23 of these variants were in BTG1. Specifically, 15 variants were identified only as DLBCL without subtyping; others were identified as GCB (n=5) and ABC (n=2), and the final mutation was identified in both subtype GCB and ABC as reported by Morin et al (2011). In BTG2; ten variants were found in DLBCL with their subtypes; 3 in ABC and 7 in GCB. The remaining twelve somatic variants were mostly in BTG2; BL (n=7), FL (n=2). Two BTG1 variants were identified in lymphoid neoplastic tissue (Figure 3-4). Finally, three variants were reported as NHL-lymphoma without specified the type, in BTG1 (n=2) and in BTG2 (n=1) (Figure 3-4). Evaluation of the effect of all somatic mutations on protein function was undertaken regardless of whether these were missense or non-missense mutations.

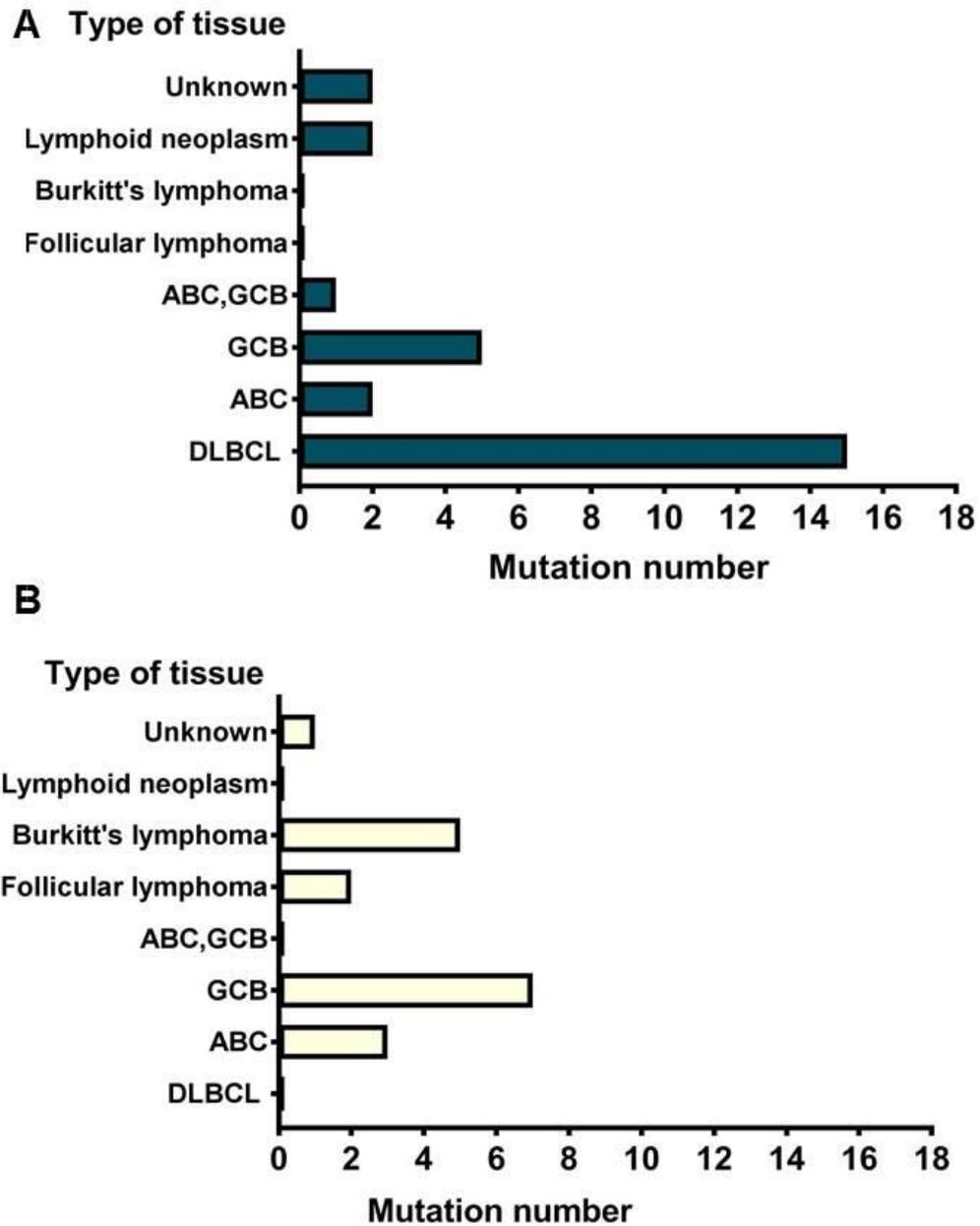


Figure 3-4 Somatic mutations in the BTG1 and BTG2 genes in different types of lymphoma.

In total, 45 mutations were identified in lymphoma. (A) The number of BTG1 variants according the type of lymphoma. (B) The number of BTG2 variants according the type of lymphoma.

3.3.2 BTG1 and BTG2 variants found in lymphoma analysed using predictive algorithms

To distinguish whether these SNVs may lead to functional variation and therefore potentially contribute to lymphoma pathogenesis, two algorithms (SIFT and Suspect) were used to analyse and predict the effects of coding non-synonymous variants on protein function.

3.3.2.1 BTG1 and BTG2 variants found in lymphoma analysed using Sorting intolerant from tolerant (SIFT)

The SIFT algorithm predicts whether a predicted amino acid substitution affects protein function based on sequence homology and the physical properties of amino acids. Based on the protein sequence, SIFT aligns the related proteins sequence with the query sequence. Then, SIFT calculates the probability score of changing the amino acid at position that appears in the alignment from 0 to 1, with cut-off value=0.05. The amino acid substitution is predicted damaging if the score ≤ 0.05 and tolerated if the score is >0.05 (Kumar et al., 2009).

Overall, half of the somatic mutations studied predicted an effect on either BTG1 or BTG2 protein function (score ≤ 0.05) using SIFT (Figure 3-5 A).

In BTG1, 15 variants in total had a predicted effect on protein function (SIFT score ≤ 0.05). Interestingly, the majority of these variants were identified in DLBCL, and two mutations were identified as lymphoma without determining the type of NHL. Of these, seven variants were reported as DLBCL without determination of the subtype. The remaining were either in DLBCL-GCB (n=3) or DLBCL-ABC (n=2), and only one mutation was DLBCL-ABC and GCB (Figure 3-5 B). The majority of the somatic variants were clustered in the BTG domain based on the protein sequence which is characterised as a protein-protein interaction domain (Figure 3-6).

G66V and N73K were identified in NHL of unknown origin. Both were predicted to be damaging to protein function (score=0.001). Three mutations,

R27H, L37M, and L94V, were identified in GCB-DLBCL and two mutation M11I and Q36H were identified in ABC-DLBCL. The mutation (H2Y) was identified in both subtypes of DLBCL (ABC and GCB). The last six mutations resulting in amino acid substitutions F25C, F40C, P58L, L104H, I115V, and E117D were identified in DLBCL without specified subtypes. All point mutations mentioned above were predicted by SIFT to be likely to be damaging to protein function (score ≤ 0.05).

Two mutations were predicted to cause deletions of ten or more amino acids located in either the N- or C- terminal of BTG1. Both of these mutations were found in DLBCL cases. The first variant occurred at the start of the coding region of BTG1; this single nucleotide base change altered the initiator methionine codon from ATG to ATA. The absence of the first starting codon means translation will not start until the second AUG at position M11 of BTG1. Thus, this base substitution causes the deletion of the first ten amino acids of BTG1 in the N-terminus ($\Delta 10N$ BTG1). The second missense mutation occurs at position 149 of BTG1, changing the residue from cysteine to a stop codon. This would result in deletion of the 21 C-terminal amino acids (BTG1 C149del). Because both SIFT and Suspect are designed for point mutations, and not truncated proteins, a different algorithm called Protein Variation Effect Analyzer (PROVEAN) was used to assess the possible effect of the C-terminal deletion. PROVEAN predicts the functional effect of protein sequence variations, including single amino acid substitutions and small insertions and deletions. Based on the query protein sequence to the related protein sequence, and predicting the changing caused by a given variation the cut-off score is -2.5. If the score ≤ -2.5 the protein variant is predicted to have a deleterious effect and if its > -2.5 the variant is predicted to have a "neutral" effect (Choi et al., 2012). The deletion was predicted likely to affect the protein function, with a score of -7.75. The details of the BTG1 variants are summarized in Table 3-1.

60% of the BTG2 variants (n=11) were predicted to be unlikely to affect protein function (SIFT score >0.05). The majority of these mutations were found in DLBCL in subtypes GCB (n=5) and ABC (n=3). Of the three remaining mutations, two were identified in FL and one was identified as lymphoma without specified the type (Figure 3-5 C). Notably, all the mutations that were considered to be tolerated in the SIFT prediction were in positions 5 to 45 of the protein sequence, and the majority of mutated residues were clustered in the BTG domain (Figure 3-6). Seven BTG2 variants were predicted to affect protein function (score \leq 0.05). Four of these mutations were identified in BL; I70M, L100P, S158C, and S158I. The remaining were found in DLBCL-GCB (L46F and H49Y), or identified as lymphoma of unknown type (G117D). Other mutations are summarized in Table 3-2.

Finally, the mutation identified in the KM-H2 cell line BTG2 V153M was predicted to be unlikely to affect protein function using SIFT (score=0.115) (Table 3-3).

3.3.2.2 BTG1 and BTG2 variants found in lymphoma analysed using Suspect analysis

Using the Suspect algorithm, the prediction values scored from 1 to 100 are calculated based on the protein sequence and structures. The human structure of BTG2 (PDB accession number 3DJU), published by (Yang et al., 2008) was used and the BTG1 model was generated using the Protein Homology/analogy Recognition Engine (Phyre) V 2.0 webserver (Kelley et al., 2015). This approach modelled a single sequence of human BTG1 (P62324) and used multiple alignments of homologues to predict the template of the secondary structure. The resulting BTG1 model was prepared with high-confidence (100%) covering 72% of protein sequence with BTG domain-like secondary structure (ID: d3e9va1). The recommended cut-off value is 50; thus, a value above 50 indicated a point mutation is more associated with disease. A score less than 50 suggests a point mutation is likely to be neutral.

In contrast to the SIFT analysis, only seven somatic variants in BTG1 (M11I, F25C, Q36H, F40C, P58L, G66V, and L104H) and in BTG2 (L100P) were predicted to be more associated with disease (score ≥ 50) by Suspect. The majority of the BTG2 variants, predicted to be likely neutral, scored < 25 . Of note, three BTG2 variants scored close to the cut-off (40-49), including S39N, L46V and I70M (Table 3-1 and Table 3-2), suggesting they might be associated with disease.

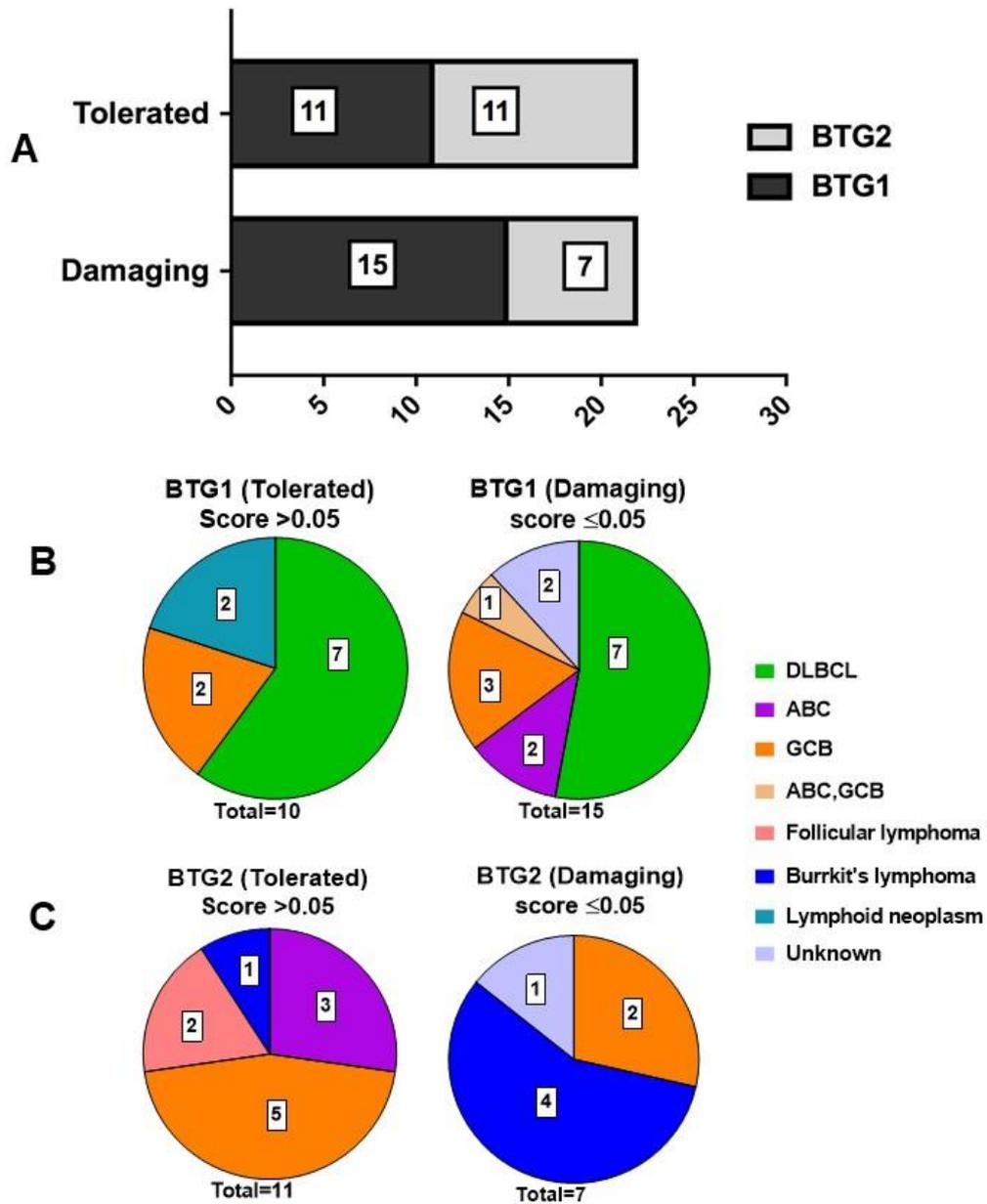


Figure 3-5: Number of somatic mutations in the *BTG1* and *BTG2* in different types of lymphoma and evaluation by SIFT.

(A) Overview of mutations in *BTG1* and *BTG2* identified by RNA-seq or exome sequencing data from over 100 NHLs. (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013) according to SIFT prediction. Variants of *BTG1* gene (B) or *BTG2* gene (C) found in different types of lymphoma and evaluated by SIFT prediction. Mutations predicted with SIFT affecting protein function at cut-off =0.05 (tolerated at score>0.05 and damaging at score ≤0.05).

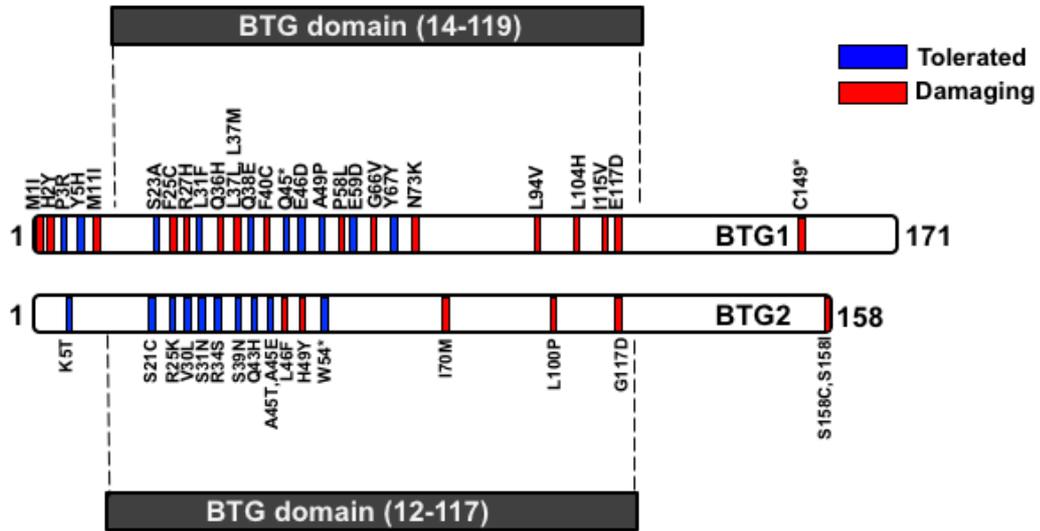


Figure 3-6: Schematic representation of the BTG1 and BTG2 mutations mapped onto their protein sequences.

Mutations in BTG1 and BTG2 reported in the literature). Both BTG1 and BTG2 share the first 120 amino acids in the N-terminus, or BTG domain. These mutations were evaluated using SIFT to predict how an amino acid substitution affects protein function at a cut-off =0.05 (tolerated at score >0.05 and damaging at score ≤0.05). (*) means amino acids substitution into stop codon caused deletion mutation.

No	Position	Residue (reference)	Residue (alteration)	Mutation cds ¹	Codons	SIFT score*	Prediction (≤0.05)*	Suspect score	Type of cancer†	Zygoty	Reference
1	1	M	I	G>A	ATG,ATA	0	Damaging	/	DLBCL	Homozygote	b
2	2	H	Y	C>T	CAT, TAT	0.037	Damaging	/	ABC,GCB	Heterozygote	a
3	3	P	R	C>G	CCC, CGC	0.069	Tolerated	/	DLBCL	Homozygote	b
4	5	Y	H	T>C	TAC,CAC	0.409	Tolerated	/	DLBCL	Homozygote	b
5	11	M	I	G>A	ATG,ATA	0.028	Damaging	87	DLBCL	Homozygote	b
6	23	S	A	T>G	TCC, GCC	0.427	Tolerated	13	DLBCL	Homozygote	b
7	25	F	C	T>G	TTT,TGT	0.001	Damaging	68	DLBCL	Homozygote	b
8	27	R	H	G>A	CGC, CAC	0.039	Damaging	20	GCB	Heterozygote	a
9	31	L	F	C>T	CTC,TTC	0.726	Tolerated	13	DLBCL	Homozygote	b
10	36	Q	H	G>C	CAG, CAC	0.006	Damaging	53	ABC	Heterozygote	a
11	37	L	M	C>A	CTG,ATG	0.02	Damaging	31	GCB	U	a/d
12	37	L	L	C>T	CTG,TTG	1	Tolerated	2	Lymphoid neoplasm, MM	Homozygote	b/d
13	38	Q	E	C>G	CAG, GAG	0.444	Tolerated	7	GCB	Heterozygote	a
14	40	F	C	T>G	TTC,TGC	0	Damaging	71	DLBCL	Homozygote	b
15	46	E	D	G>C	GAG, GAC	0.239	Tolerated	11	GCB	Heterozygote	a
16	46	E	Q	G>C	GAG,CAG	0.155	Tolerated	11	DLBCL	Homozygote	b
17	58	P	L	C>T	CCA, CTA	0	Damaging	53	ABC	U	a
18	59	E	D	A>T	GAA,GAT	0.541	Tolerated	2	DLBCL	Homozygote	b
19	66	G	V	G>T	GGT,GTT	0.001	Damaging	70	U	U	c
20	67	Y	Y	C>T	TAC,TAT	1	Tolerated	2	Lymphoid neoplasm, MM	U	d
21	73	N	K	C>A	AAC,AAA	0.001	Damaging	24	U	U	c
22	94	L	V	C>G	CTG, GTG	0.002	Damaging	28	GCB	Heterozygote	a
23	104	L	H	T>A	CTC,CAC	0	Damaging	89	DLBCL	U	c
24	115	I	V	A>G	ATT,GTT	0.003	Damaging	29	DLBCL	Homozygote	b
25	117	E	D	G>T	GAG,GAT	0.022	Damaging	14	DLBCL	Homozygote	b
26	149	C	*	T>A	TGT, TGA	/	/	/	DLBCL	U	c
27	165	N	S	A>G	AAT,AGT	0.46	Tolerated	/	DLBCL	Homozygote	b

(a)Morin et al., 2012, (b) Lohr et al., 2011, (c) Zhang et., 2013, (d) Walker et al., 2012 and (e) Love et al., 2012. Follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL), the germinal centre B-cell (GCB), Activated B-cell(ABC) and Multiple myeloma (MM). ¹CDS coding sequences, (*) Using SIFT prediction software v 1. 03, (†) U=Unknown

Table 3-1: list of BTG1 variants evaluated with SIFT and Suspect in different types of lymphoma.

No	Position	Residue (reference)	Residue (alteration)	Mutation CDS1	Codons	SIFT Score*	Prediction (≤0.05)*	SUSPECT score	Type of cancer†	Reference
1	5	K	T	A>C	AGG,CGG	0.168	Tolerated	/	GCB	a
2	21	S	C	C>G	TCC,TGC	0.156	Tolerated	11	GCB	a
3	25	R	K	G>A	AGG,AAG	0.209	Tolerated	16	FL	a
4	30	V	L	G>T	GTG,TTG	1	Tolerated	10	ABC	a
5	31	S	N	G>A	AGC,AAC	0.48	Tolerated	6	GCB	a
6	34	R	S	G>C	AGG,AGC	0.077	Tolerated	28	ABC	a
7	36	K	Q	A>G	AAG,CAG	1	Tolerated	4	Burkitt's	e
8	39	S	N	G>A	AGC,AAC	0.156	Tolerated	43	GCB	a
9	43	Q	H	G>C	CAG,CAC	0.052	Tolerated	23	ABC	a
10	45	A	E	C>A	GCA, GAA	0.218	Tolerated	4	GCB	a
11	45	A	T	G>A	GCA, ACA	0.287	Tolerated	14	FL	a
12	46	L	F	C>T	CTC,TTC	0.009	Damaging	43	GCB	a
13	49	H	Y	C>T	CAC,TAC	0.003	Damaging	41	GCB	a
14	70	I	M	C>G	ATC,ATG	0.054	Damaging	40	Burkitt's	e
15	100	L	P	T>C	CTG,CCG	0	Damaging	54	Burkitt's	e
16	117	G	D	G>A	GGC,GAC	0.002	Damaging	38	U	a
17	158	S	C	A>T	AGC,TGC	0	Damaging	/	Burkitt's	e
18	158	S	I	G>T	AGC,ATC	0	Damaging	/	Burkitt's	e

(a)Morin et al., 2012, (b) Lohr et al., 2011, (c) Zhang et., 2013, (d) Walker et al., 2012 and (e) Love et al., 2012. Follicular lymphoma (FL), diffuse large B-cell lymphoma (DLBCL), the germinal centre B-cell (GCB) and Activated B-cell(ABC). ¹CDS coding sequences, (*) Using SIFT prediction software v 1.03, (†) U=Unknown. Note, the zygosity was not identified in all BTG2 mutations.

Table 3-2: list of BTG2 mutations evaluated with SIFT and Suspect in different types of lymphoma.

Gene	Position	Residue (reference)	Residue (alteration)	Mutation Cds ¹	Codons	SIFT Score*	Prediction (score <0.05)*	Cell Line
BTG1	16	A	P	G>C	GCC,,CCC	0.151	Tolerated	Raji
BTG1	22	I	L	A>C	ATC,CTC	1.000	Tolerated	Raji
BTG2	153	V	M	G>A	GTG,ATG	0.115	Tolerated	KMH-2

¹CDS coding sequences, (*) Using SIFT prediction software v 1.03, Raji cell line (Burkitt lymphoma) and KMH-2(Hodgkin's lymphoma).

Table 3-3: list of BTG1 and BTG2 mutations in lymphoma cell lines evaluated with SIFT.

3.4 Discussion

Recently, the tumour suppressor genes BTG1 and BTG2 have been shown to be mutated in lymphomas (Morin et al., 2011) and leukaemias (Waanders et al., 2012). In fact, BTG1 was first identified in a patient with chronic lymphocytic leukaemia (Rouault et al., 1992). A number of studies have now found mutations in BTG1 and BTG2 in haematological malignancies (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013). In this project, sequencing analysis of human lymphoblastoid cell lines was conducted, to identify whether there are novel variants in the codon regions of the BTG1 or BTG2 genes. Also, data on BTG1 and BTG2 mutations were collected and analysed using prediction algorithms.

3.4.1 BTG1 and BTG2 variants in lymphoma cell lines

Short deletion in BTG1 exon 2 have been identified in leukaemia patients, and the deletions of this region were shown to be associated with a poor risk prognosis in BCP-ALL paediatric patients (Waanders et al., 2012). Sequence analysis of the lymphoblastoid cell lines was conducted. BTG1-exon2 in the lymphoblastoid cell lines showed no variants or deletions; exon 2 covers approximately 70% of the BTG1 coding region. The BTG1 exon 1 region also matched the reference gene sequence (BTG1: Chromosome 12: 92, 140, 278-92, 145, 897) in the three cell lines DB, KM-H2 and GM03021. A limited number of papers have identified variants in the BTG1 or BTG2 genes in lymphoma cell lines. One study conducted genome sequencing in Raji cell line and identified 654 variants in EBV genes but not in BTG1 (Xiao et al., 2016). While Morin et al. (2013) identified 74 genes with mutations including BTG1 and BTG2 mutations in samples derived from DLBCL patients, they also did whole genome sequencing in 13 DLBCL cell lines including the DB cell line. A number of mutations were identified in the DB cell line such as TP53, LRRN3, MLL2, but not in the BTG1 or BTG2 genes.

In the KM-H2 Hodgkin's lymphoma cell line, a mutation resulting in amino acid change V153M was identified in *BTG2*. This amino acid substitution was

due to a change in the nucleotide base from G to A. The sequencing trace indicates likely monoallelic expression (Figure 3-3 C). However, further experiments are required to validate this mutation. Single-strand conformation polymorphism (SSCP) is one method to detect point mutations based on fluorescent primers that detect mutated DNA using their mobility in electrophoresis gel (Konstantinos et al., 2008). Based on the evaluation of SIFT prediction, amino acid substitutions from valine into methionine do not affect protein function.

3.4.2 BTG1 and BTG2 variants in lymphoma patients

DLBCL and FL constitute the majority of non-Hodgkin's lymphomas. Recently, efforts have been undertaken to use a molecular classification alongside morphology and immunophenotyping to identify the type of lymphoma. Genetic alterations may be due to chromosomal translocations, mutations caused by aberrant SHMs and sporadic somatic mutations and copy number alterations, as denoted by deletions and amplifications (Coupland, 2013). Data showed BTG1 and BTG2 variants are common in DLBCLs and BL. Zhang et al. (2013) reported a significant activation-induced cytidine deaminase (AICDA) known as AID-related mutation as a major mechanism underlying genetic mutations in the BTG1 gene and others such as the CD79B and PIM1 genes. The AICDA enzyme introduces mutations by deamination of cytosine, targeting specific domains including the WRCY motif or the inverse RGYW (W=adenine or thymine, R=purine, C=cytosine, Y=pyrimidine, G=guanine), and helps to produce antibody diversity in the lymph nodes. This is involved in the somatic hyper-mutation and class-switch recombination of immunoglobulin genes in B cells (Yu et al., 2004). Therefore, the mutations in the WRCY motifs of AICDA are significant contributors to the somatic alterations in the BTG1, CD79B and PIM1 genes, thus explaining the presence of these variants in more matured B-lineage malignancies, as well as their absence in solid tumours (Zhang et al., 2013).

It is also worth mentioning that some mutations were identified in DLBCL without subtypes, while others were identified as ABC or GCB. In Lohr et al. (2012) and Zhang et al. (2013), whole genome/DNA sequencing was performed and variants were found in BTG1. However the work did not specify the type of DLBCL. Meanwhile, Morin et al. (2011) reported the subtype of DLBCL of NHL samples. This was based on RNA-seq data that identified expression of genes related to germinal centre B cells (DLBCL-GCB), and other subtypes. For example, ABC has an expression pattern related to mitogenically activated B cells close to cells with a secretory function (Alizadeh et al., 2000, Rosenwald et al., 2002).

3.4.3 BTG1 and BTG2 variants evaluated by SIFT and Suspect analysis

Forty-five BTG1 and BTG2 variants are clustered in the BTG domains close to the N terminus (Figure 3-6). All these variants were analysed using the bioinformatic tools SIFT and Suspect. Both are used to predict the effect of the amino acid substitutions on the protein function. However, when using the Suspect analysis based on protein structure, eight variants of BTG1 and BTG2 could not be analysed. In the human BTG2 structure (PDB accession number 3DJU), the first six N-terminal amino acids are not observed, because the BTG2 structure only includes residues 7 to 127 (Yang et al., 2008). Therefore, BTG2 variants K5T, S158C and S158I could not be analysed with Suspect. Because the BTG1 model structure which was generated using the Phyre V 2.0 webserver, it only included residues 9 to 128 (Kelley et al., 2015). Thus, BTG1 variants (M1I (Δ 10N BTG1), H2Y, P3R, Y5H and N165S) could not be analysed with Suspect.

Analysis using SIFT and Suspect showed different results. In total, 22 BTG1 and BTG2 variants were predicted to be damaging using SIFT. However, using Suspect analysis, only eight of these variants were predicted to be associated with disease, including BTG1 M11I, F25C, Q36H, F40C, P58L, G66V and L104H, and BTG2 L100P. The remaining variants were predicted to affect protein function using SIFT but not with Suspect: BTG1 variants R27H, L37M, N73K,

L94V, I115V and E117D and BTG2 variants L46F, H49Y, I70M and G117D. Interestingly, no variant predicted with Suspect as likely to be neutral (score < 50) was then predicted with SIFT to be damaging to protein function (score ≤ 0.05). It has been shown that the Suspect has higher accuracy compared with SIFT about 82% and 62%, respectively. As the Suspect algorithm predicts the SAVs that associated with disease and high selectively (Yates et al., 2014). In our data analysis, the following may explain the differences observed with SIFT and Suspect analysis. It may be that the SIFT algorithm has a high-sensitivity and predicts a greater number of variants as compared with Suspect at the expense of a greater number of false positive results. Likewise, Suspect may have a lower sensitivity combined with a lower false-positive rate.

SIFT and Suspect algorithms both predict the SAV in conserved sequences, the Suspect algorithm involves an additional evaluation in comparison with SIFT. The Suspect algorithm considers PPIs, which helps to distinguish between disease-associated and tolerated SAVs, as opposed to the prediction of protein functions (Yates et al., 2014). Notably, losing protein function does not necessarily lead to the development of a disease (Yates et al., 2014). Neither SIFT or Suspect could analyse the predicted effects of the BTG1 C149del mutation, because both programmes are designed to predict the effect of amino acid substitutions, not deletions. For this reason, another software package was used; PROVEAN. This program can be used to predict the impact of many types of protein sequence variations including deletions (Choi et al., 2012). The results predicted that the BTG1 C149del variant was deleterious and scored -7.754 less than the cut-off (-2.5), which indicated a likely effect on protein function. Interestingly, microdeletion of BTG1 is common to find in the BCP-ALL cases, suggesting might be play role in the leukaemogenesis (Kuiper et al., 2007, Waanders et al., 2012, Xie et al., 2014, Scheijen et al., 2016).

In summary, in total 45 BTG1 and BTG2 variants were identified in NHLs. The majority of BTG1 and BTG2 variants identified in DLBCL, FL and BL are clustered in the BTG domain. According to SIFT prediction analysis, 22 BTG1 and BTG2 variants were predicted to likely affect protein function. Moreover, one deletion mutation was identified in DLBCL (BTG1 C149del) and was predicted to be damaging using PROVEAN analysis.

In this chapter, 16 BTG1 variants were identified in NHL, and predicted to be damaging using bioinformatics analysis. The additional variant (E59D) was selected as negative control, because it is a conservative change which was predicted to be neutral using SIFT and Suspect.

Chapter 4

Lymphoma BTG1 variants impact on the interaction with human Caf1 (CNOT7/CNOT8)

Chapter 4: Lymphoma BTG1 variants impact on the interaction with human Caf1 (CNOT7/CNOT8)

4.1 Introduction

The BTG/TOB protein family is characterised by the BTG domain located in the N-terminus, which contains 104-106 amino acids. This domain contains two short conserved motifs: box A and box B. The C-terminal regions are less conserved among the family members, which allows them to be classified into three distinct subfamilies: the BTG1/BTG2 subfamily, the BTG3/BTG4 subfamily and the TOB subfamily. The role of the BTG domain is to allow protein–protein interactions (Winkler, 2010). Specifically, the BTG/TOB proteins interact with the hCaf1 deadenylase subunit of the Ccr4-Not complex, which is encoded by CNOT7 or CNOT8 (Prévôt et al., 2001, Yang et al., 2008, Aslam et al., 2009, Horiuchi et al., 2009). The hCaf1/CNOT7 and hCaf1/CNOT8 proteins play a role in shortening the poly (A) tail of cytoplasmic mRNA (Maryati et al., 2015). Both proteins regulate largely identical gene sets as determined using wide-genome expression analysis (Aslam et al., 2009). Apart from BTG4, all BTG/TOB proteins have been shown to be involved in post-transcriptional regulation of gene expression (Prévôt et al., 2001, Mauxion et al., 2008, Collart and Panasenko, 2012, Doidge et al., 2012a). TOB1 and TOB2 have PAM2 motifs in the C-terminal region that mediate the interaction with the PABPC1. This motif is not present in the BTG proteins (Ezzeddine et al., 2007, Funakoshi et al., 2007).

Horiuchi et al. (2009) published the crystal structure complex of the conserved BTG domain in the N-terminal 138 residues of TOB1 with hCaf1/CNOT7. Also, the crystal structure of human BTG2 was established by Yang et al. (2008). They identified BTG2 residues that are important for the interaction with hCaf1/CNOT7 through the superposition of BTG2 with the BTG domain of TOB1 in the TOB1–hCaf1/CNOT7 complex crystal structure (Yang et al., 2008). In addition, experiments confirmed that Y65, D75, W103 and D105 residues in BTG2 and F55, D65, W93 and D95 residues in TOB1 are required for the

interaction with hCaf1 (CNOT7/CNOT8) (Doidge et al., 2012a). It has been confirmed that the residue K203 in hCaf1/CNOT7 is also necessary in the interaction with TOB1 (Horiuchi et al., 2009). Moreover, both residue E247 and Y260 in hCaf1/CNOT7 are required in the interaction with BTG2 (Aslam et al., 2009). All amino acids mentioned above were experimentally changed to alanine as uncharged residues to examine the interaction. Using the overexpression of BTG2 and TOB1 variants, it was concluded that the physical interactions between hCaf1(CNOT7/CNOT8) and BTG2/TOB1 are required for the anti-proliferative activity of BTG2 and TOB1 (Doidge et al., 2012a, Ezzeddine et al., 2012). Recently, frequent mutations in BTG1 and BTG2 were reported in lymphoma samples (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013). At the amino acid level, the similarity and identity of the BTG domain in BTG1 and BTG2 is 96% and 74%, respectively (Winkler, 2010). However, in contrast to BTG2, very few studies have investigated which residues of BTG1 are required for the interaction with hCaf1(CNOT7/CNOT8) (Prévôt et al., 2001). It is therefore interesting to discover whether lymphoma BTG1 variants are affected in their interactions with hCaf1(CNOT7/CNOT8). In this study, 17 variants were selected; 16 variants were predicted to have an impact on the protein function (as described in Chapter 3) and one additional mutation was selected as a conservative amino acid change and therefore unlikely to impact protein function.

Yeast two-hybrid analysis was used to evaluate the interactions between the BTG1 variants and hCaf1(CNOT7/CNOT8) *in vitro*. The principle of the yeast two-hybrid method is based on the interaction between a protein fused to the DNA-binding domain of the Gal4 transcriptional activator and a second protein fused to the Gal4 activation domain. Expression of a β -Galactosidase reporter can be used as a measure for the physical interaction between the proteins (Perkins et al., 2010).

4.2 Structural analysis of BTG1 variants

Sixteen variant residues, which were identified in NHL and predicted to affect protein function using SIFT were selected. To understand the location of these variant residues, a structural model of the BTG1-hCaf1/CNOT7 complex was required. To this end, the protein homology/analogy recognition engine (Phyre) V 2.0 webserver was used to obtain the BTG1 model (Kelley et al., 2015) and the UCSF Chimera analyser was used to visualise the structural model of the BTG1-hCaf1/CNOT7 complex (Meng et al., 2006).

Phyre2 modelled a single sequence of human BTG1 (P62324) to generate a high-confidence (100%) model covering 72% of the protein sequence with the BTG domain-like secondary structure (ID: d3e9va1). In order to obtain the BTG1-hCaf1/CNOT7 complex, two additional structures were used. The first structure used, was the highly similar paralogue human BTG2 (PDB accession number: 3DJU) (Yang et al., 2008). The second structure was the BTG domain of TOB1 in complex with hCaf1/CNOT7 (PDB accession number: 2D5R). The latter structure confirmed that TOB1 interacts with hCaf1/CNOT7 through the BTG domain composed of box A and box B (Horiuchi et al., 2009). All structures were analysed using the Match Maker extension of the UCSF Chimera package to generate structure-guided sequence alignments (Meng et al., 2006). The identity of the amino acids of TOB1 and BTG2 is 42.2%; the identity between TOB1 and BTG1 45.7% (Table 4-1). The most highly conserved residues in the BTG domains of TOB1, BTG1 and BTG2 are located in the box A and box B regions (Figure 4-1). To prepare a structural model of the BTG1-hCaf1/CNOT7 complex, the BTG2 structure (PDB accession number 3DJU) was first aligned with the BTG domain of TOB1 in the TOB1-hCaf1/CNOT7 crystal structure (PDB accession number 2D5R). Then, the structure of the BTG1 model was superposed on the aligned BTG2-hCaf1/CNOT7 complex. Because BTG1 and BTG2 are highly similar based on the level of the amino acid identity (72.7%), the secondary structure of BTG1 and BTG2 is likely the same and composed of a bundle of three α -helices followed by four β -strands with two α -helices

inserted between strands $\beta 1$ and $\beta 2$ (Figure 4-2 A and B). The 17 selected BTG1 mutations were grouped based on their location in the model of the BTG1 structure. Seven mutations were located in α -helices; two were located in β -sheets and five were present in turns (loops). The remaining mutations (three) were not present in the BTG1 structure model (Figure 4-3) and (Table 4-2). Notably, one variant (G66V) is located in the small $\alpha 3$ helix and very closely to a loop. Therefore, this variant was included in Group III containing residues located in loops.

Protein (1)	*Residues	¹PDB	Protein (2)	**Residues	²PDB	Identity (%)
TOB1	24-138	2D5R	BTG2	7-128	3DJU	42.2
TOB1	24-138	2D5R	BTG1	11-129	-	45.7
BTG2	7-128	3DJU	BTG1	11-129	-	72.7

Protein sequences of the three structures (TOB1, BTG2, and BTG1) were aligned using the Match Maker extension of the UCSF Chimera software v.1.11.2. The BTG1 model was generated by Phyre2, using BTG domain-like secondary structure (ID: d3e9va1) with 100% confidence. Using (*) and (**) identified the range of residues for protein 1 and protein 2, respectively. (¹PDB) and (²PDB) used for the accession number of protein 1 and protein 2, respectively.

Table 4-1 Percentage identity for protein sequences used for the alignment of the BTG domain in TOB1, BTG1 and BTG2.

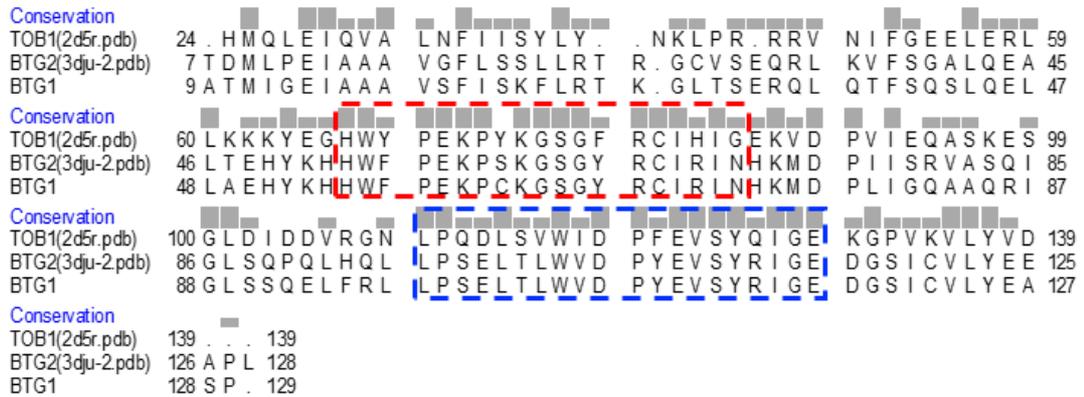


Figure 4-1 Protein sequence alignment of the BTG domain in TOB1, BTG1 and BTG2.

Structural alignment of the BTG domain of the TOB1, BTG2 and BTG1. TOB1 structure (PDB accession number 2D5R) and the human BTG2 structure (PDB accession number 3DJU) were used. BTG1 model was generated by Phyre2, using BTG domain-like secondary structure (ID: d3e9va1) with 100% confidence. As shown in this figure; the majority of the conserved residues are located in the box A and box B regions that are highlighted in the red and blue boxes, respectively. Conservation is also indicated as histogram bars (grey) above the alignment.

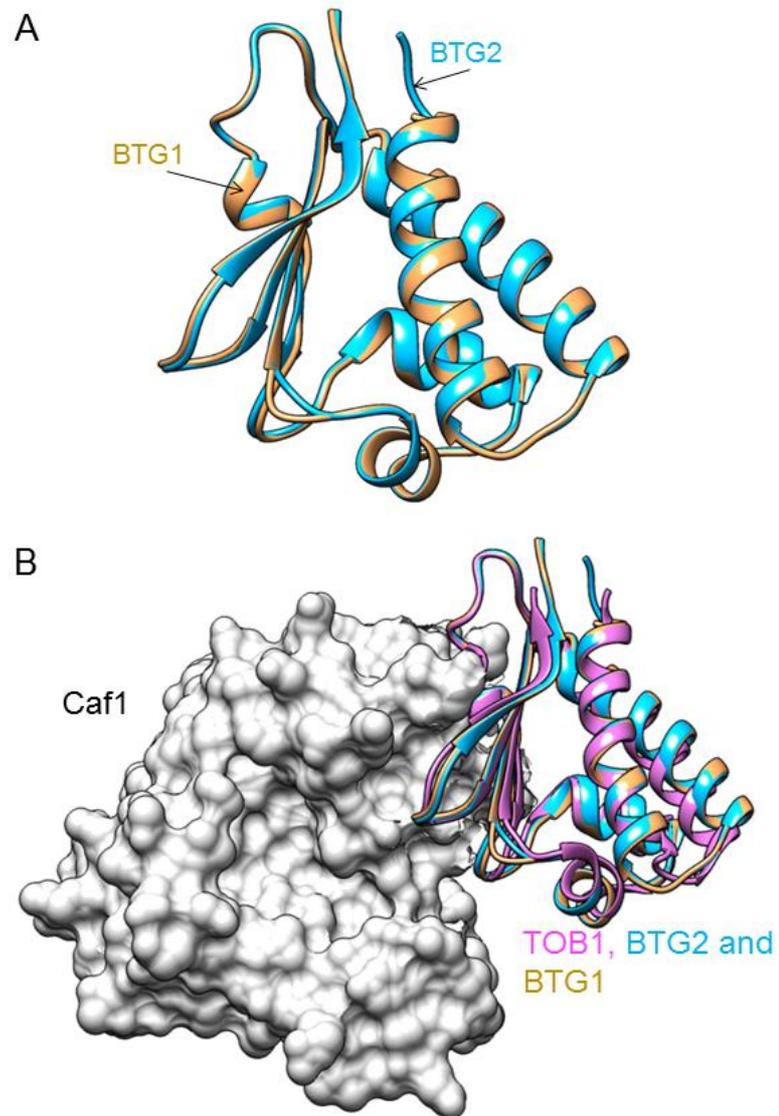


Figure 4-2 Structural models of the BTG1-hCaf1/CNOT7 and BTG2-hCaf1/CNOT7 complexes.

(A) Alignment of the BTG1 model with the human BTG2 crystal structure (PDB accession number: 3DJU), coloured sandy brown and blue, respectively. BTG1 model was generated by phyre2, using BTG domain-like secondary structure (ID: d3e9va1) with 100% confidence (B) Superimposition of the BTG1 and BTG2 structures with the structure of the TOB-hCaf1/CNOT7 complex (PDB accession number: 2D5R). UCSF Chimera (version 1.11.2) was used for visualization (Meng et al., 2006).

Variant residue	Location	Group
Δ 10N BTG1	-	Group III
H2Y	-	Group III
M11I	L1	Group III
F25C	α 1	Group I
R27H	α 1	Group I
Q36H	α 2	Group I
L37M	α 2	Group I
F40C	α 2	Group I
P58L	L1 (box A)	Group III
E59D	L2 (box A)	Group III
G66V	α 3 (box A)	Group III
N73K	L4 (box A)	Group III
L94V	α 5	Group I
L104H	β 2 (box B)	Group II
I115V	β 3 (box B)	Group II
E117D	L8 (box B)	Group III

Location of variant residues in BTG1. The variant residues were divided into three group based on their location in the BTG1 secondary structure; GroupI (α -helix), GroupII (β -sheets) and Group III (loops(L)).

Table 4-2 List of variant residues grouped according to their location in the secondary structure of BTG1.

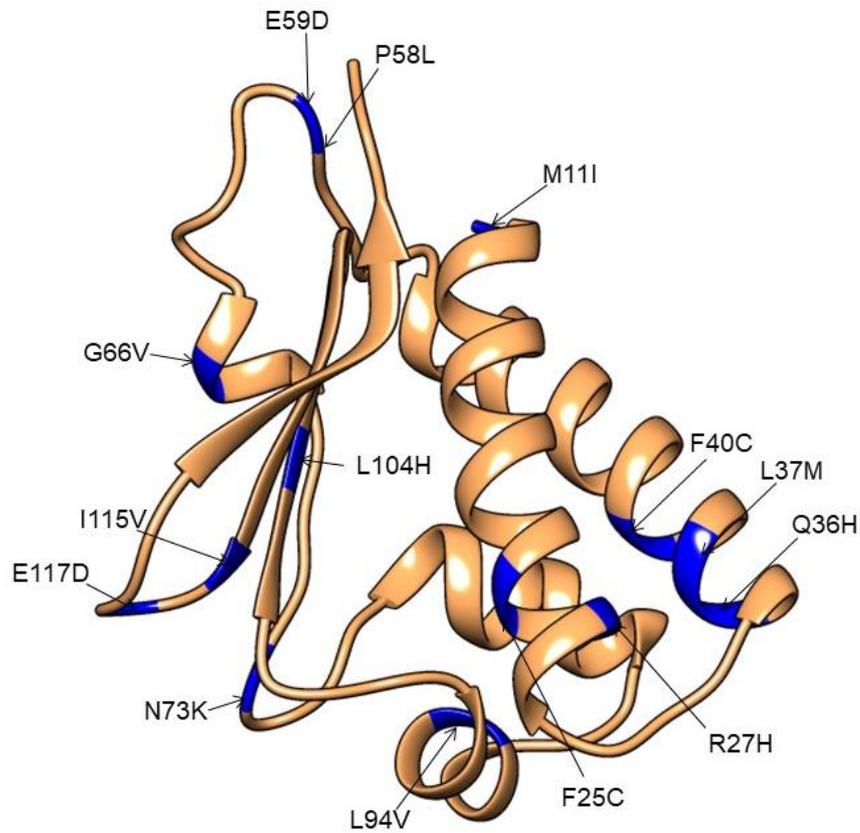


Figure 4-3 Mapping of variant residues in the BTG1 structural model.

BTG1 model was generated by Phyre 2, using BTG domain-like secondary structure (ID: d3e9va1) with 100% confidence. Amino acid substitutions predicted to be damaging to BTG1 protein function (SIFT score ≤ 0.05) were mapped on a structural model of BTG1 except E59D. UCSF Chimera (version 1.11.2) was used for visualisation.

4.3 Generation of plasmids for yeast two-hybrid interaction analysis

To evaluate the effect of the BTG1 variants on the interaction with hCaf1(CNOT7/CNOT8), plasmids were prepared for yeast two-hybrid analysis. The Gal4 two-hybrid system requires two separate plasmids for the expression of the Gal4 activation domain (Gal4-AD) fusion protein (pAD-GAL4-2.1) and the Gal4-DNA binding domain (Gal4-DBD) fusion protein (pBD-GAL4 Cam).

Plasmid pBD-Gal4-HA-BTG1 was prepared to produce the Gal4-DNA binding domain fused to BTG1. To this end, PCR was used to amplify BTG1 cDNA to introduce an EcoRI restriction site at the 5' end and a Sall restriction site at the 3' end. After restriction enzyme digestion and ligation into a modified pBD-Gal4 Cam yeast expression vector, the plasmid was created. The modified vector contained a HA tag inserted between the N-terminal Gal4-DBD and the fusion protein and facilitated the detection of the Gal4-DBD-BTG1 fusion protein (Figure 4-4 A). Plasmids containing the BTG1 variants were generated using site-directed mutagenesis of the pBD-Gal4-HA-BTG1 wild type vector.

The two deletion mutations, Δ 10N BTG1 and BTG1 C149del, were made by designing alternative 5' and 3' oligonucleotides, respectively. For Δ 10N BTG1, the first 10 amino acids of the BTG1 coding region was excluded. For BTG1 C149del, a specific oligonucleotide was designed to exclude the last 21 C-terminal amino acids of the coding region of BTG1. All BTG1 sequences were confirmed by DNA sequencing. The pAD-Gal4-hCaf1/CNOT7 and pAD-Gal4-hCaf1/CNOT8 plasmids have been described previously (Doidge et al., 2012a) and are depicted in (Figure 4-4 B).

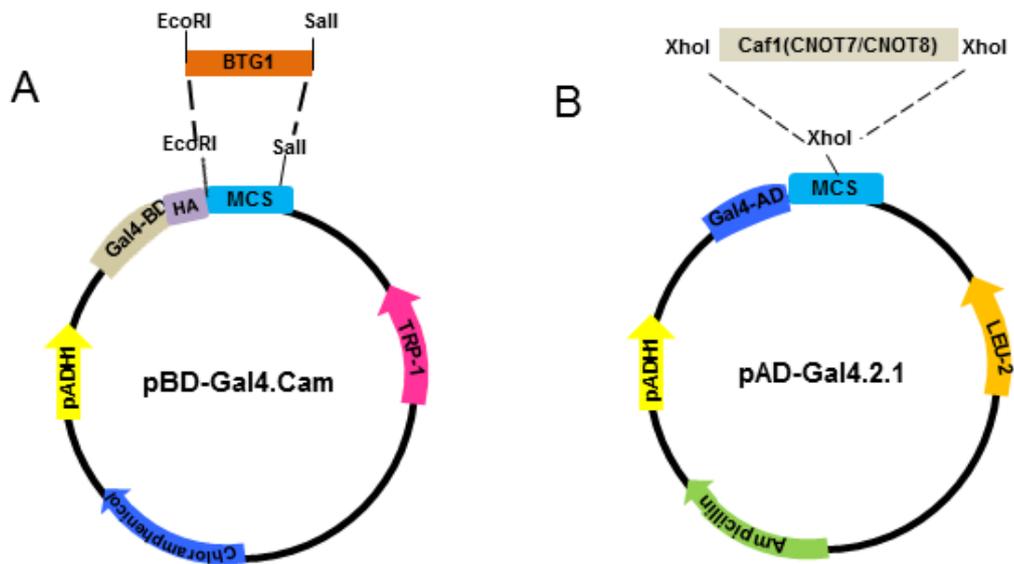


Figure 4-4 Generation of vectors for yeast two-hybrid interaction analysis.

(A) Gal4 DNA binding domain hybrid vector. BTG1 was amplified by PCR. The 5' primer included an EcoRI restriction site and the 3' primer contained a stop codon followed by a Sall restriction site. The cDNA fragments were ligated into the EcoRI and Sall restriction sites of a modified yeast expression vector pBD- Gal4 Cam containing an HA epitope coding region in frame with the multiple cloning site. (B) Gal4 activation domain hybrid vectors. The CNOT7 and CNOT8 cDNAs were ligated into the yeast expression vector pAD-Gal4-2.1 as described before (Doidge et al., 2012a).

4.4 Evaluation of the interaction between human Caf1 (CNOT7/CNOT8) and BTG1 variants through yeast two-hybrid analysis

To assess whether BTG1 variants were able to interact with the hCaf1(CNOT7/CNOT8) deadenylase subunit, the yeast two-hybrid assay was used as a semi-quantitative test. The system is based on the expression of two proteins that are fused to the Gal4-HA-DBD and Gal4-AD, respectively. A physical interaction between the fusion proteins will bring the Gal4-HA-DBD and Gal4-AD in close proximity, which results in expression of a reporter gene in the yeast host cell. Thus, vectors expressing Gal4-DBD-HA-BTG1 and Gal4-AD-hCaf1 were transformed into the yeast host (YRG2) strain, and were grown in selective media without leucine or tryptophan. Three colonies were subjected to a β -galactosidase assay for each BTG1 variant and three independent experiments were conducted (Figure 4-5).

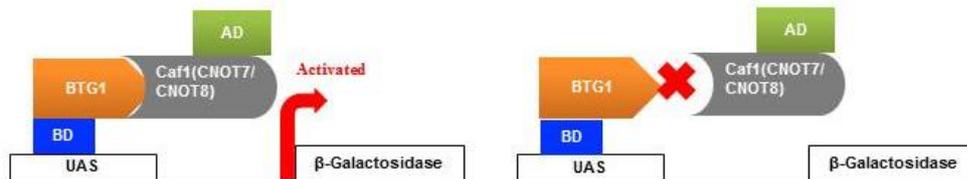


Figure 4-5 Schematic of yeast two-hybrid interaction analysis of BTG1 and human Caf1.

The Gal4 activation domain (AD) was fused to hCaf1(CNOT7/CNOT8); the Gal4 DNA binding domain (BD) was fused to BTG1. Interactions between the Gal4-BD-BTG1 and Gal4-AD-hCaf1 proteins induce β -Galactosidase reporter activity (left panel). In the absence of an interaction, no expression of the β -Galactosidase is seen (right panel).

4.4.1 Interaction analysis of BTG1 variants located in the α -helical regions and human Caf1

Six variant residues, which were identified in NHL and predicted to affect protein function using SIFT, were located in α -helices. Using the structural model of the BTG1-hCaf1/CNOT7 complex, three variant residues, F25C (α 1), R27H (α 1) and L94V (α 5), were shown to be located close to the hCaf1/CNOT7 interaction surface, while Q36H (α 2), L37M (α 2) and F40C (α 2) were found to be located on the opposite side to the interaction surface (Figure 4-6 A).

When yeast was transformed with the empty vectors (pAD-Gal4 and pBD-HA-Gal4) or with pAD-Gal4 and pBD-Gal4-HA-BTG1, no β -galactosidase activity was observed. This confirmed that BTG1 fused to the DNA binding domain did not contain a cryptic activation domain. Transformation of yeast with pBD-Gal4-HA-BTG1 and either pAD-Gal4-hCaf1/CNOT7 or pAD-Gal4-hCaf1/CNOT8 induced β -galactosidase expression in agreement with the known ability of BTG1 to interact with hCaf1(CNOT7/CNOT8) (Prévôt et al., 2001) (Figure 4-6 B).

Using BTG1 variants, R27H and F40C, did not produce β -galactosidase activity and were therefore strongly impaired in their interaction with hCaf1(CNOT7/CNOT8) ($p < 0.0001$). By contrast, the interactions between variants F25C and L94V and hCaf1(CNOT7/CNOT8) were only marginally disrupted. Q36H and L37M were able to stimulate the β -galactosidase activity to similar levels as the wild type BTG1 and hence were able to interact with hCaf1/CNOT7. However, the results showed a marginal reduction for the interaction with hCaf1/CNOT8 in case of these two variants ($p < 0.05$ and $p < 0.001$, respectively). Interestingly, both Q36H and L37M were located away from the hCaf1/CNOT7 interaction surface.

To confirm expression of the BTG1 variants, proteins were extracted from yeast cultures and subjected to western blot analysis using the anti-HA antibody for detection of Gal4 DBD-HA-BTG1 with an expected size 30kDa,

however unspecific bands were detected below the DBD-HA-BTG1 bands. Also, the anti-AD-Gal4 antibody were used for detection Gal4 AD-hCaf1/CNOT7 and Gal4-AD-hCaf1/CNOT8 fusion proteins with expected size 58kDa(Figure 4-6 C).

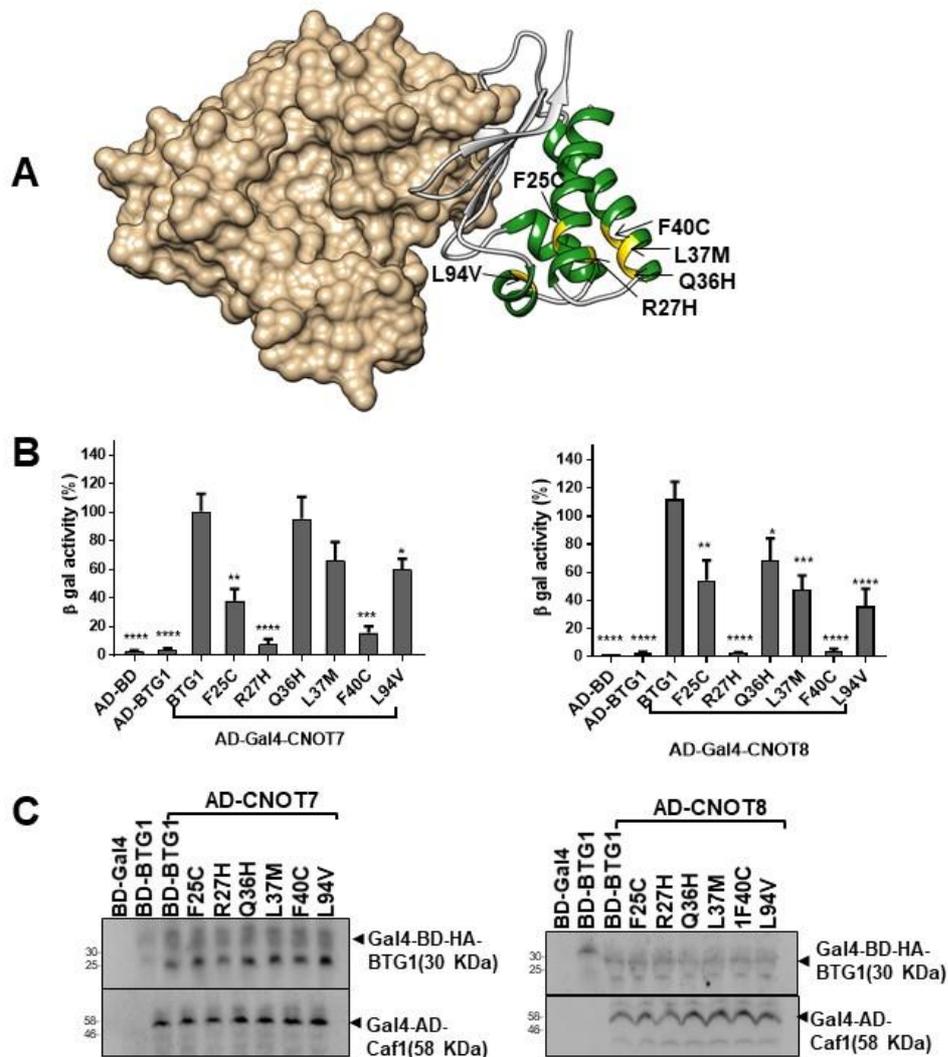


Figure 4-6 Differential interactions between human Caf1 and BTG1 variants located in α -helical regions.

(A) Structural model of the BTG1-hCaf1/CNOT7 complex. α -helices are highlighted in green. BTG1 variant residues predicted to be damaging to protein function by SIFT (score ≤ 0.05) are coloured yellow. UCSF Chimera was used for visualization. (B) Interactions between Caf1 with BTG1 variants located in α -helices. YRG2 yeast cells were transformed with vectors pAD-Gal4 or pAD-Gal4-hCaf1 and pBD-Gal4-HA or pBD-Gal4-HA-BTG1. β -Galactosidase activity was normalised using cell density (O.D. measured at 600 nm). Error bars indicate the standard error of the mean (n=3). One way Anova (dunnett post-hoc method) was used to calculate P values (* p<0.05, ** p<0.01, *** p<0.001 and **** p<0.0001). Three independent experiments were obtained. (C) Protein expression was confirmed by western blot analysis using Anti-HA and anti-AD antibodies.

4.4.2 Interaction analysis of BTG1 variants located in the β -sheet regions and human Caf1

Two variant residues, which were identified in NHL and predicted to affect protein function using SIFT, were located in the β -sheets. Using the structural model of the BTG1-hCaf1/CNOT7 complex, L104H and I115V were shown to be located close to the hCaf1/CNOT7 interaction surface in β 2 and β 3, respectively (Figure 4-7 A).

The ability of BTG1 variants to interact with hCaf1(CNOT7/CNOT8) was assessed by semi-quantitative interaction analysis using a yeast two-hybrid system. As previously described, when the yeast was transformed with empty vectors (pAD-Gal4 and pBD-HA-Gal4) or with pAD-Gal4 and pBD-Gal4-HA-BTG1, the β -galactosidase activity was absent. Once the yeast was transformed with pBD-Gal4-HA-BTG1 and either pAD-Gal4-hCaf1/CNOT7 or pAD-Gal4-hCaf1/CNOT8 the β -galactosidase activity was induced in agreement with the known ability of BTG1 to interact with hCaf1(CNOT7/CNOT8) (Prévôt et al., 2001) (Figure 4-7 B).

When these BTG1 variants were assessed in the yeast two-hybrid system, the results showed a similar pattern in the interaction with hCaf1/CNOT7 and hCaf1/CNOT8. Neither L104H and I115V produced β -galactosidase activity, and are therefore strongly implied to be impaired in their interaction with hCaf1(CNOT7/CNOT8) ($p < 0.0001$ and $p < 0.01$), respectively.

To confirm expression of the BTG1 variants, proteins were extracted from yeast cultures and subjected to western blot analysis using the anti-HA antibody for detection of Gal4 DBD-HA-BTG1 with an expected size 30kDa, however unspecific bands were detected below the DBD-HA-BTG1 bands. Also, the anti-AD-Gal4 antibody were used for detection Gal4 AD-hCaf1/CNOT7 and Gal4-AD-hCaf1/CNOT8 fusion proteins with expected size 58kDa (Figure 4-7 C).

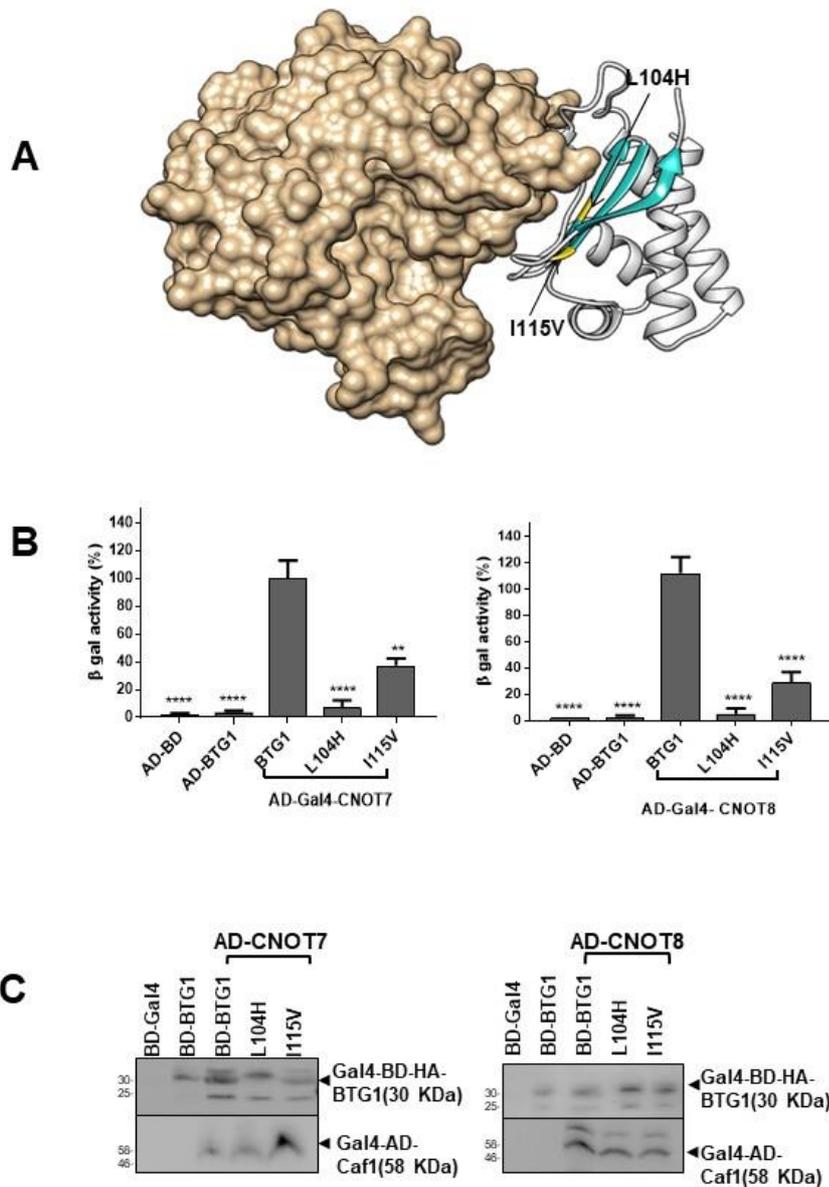


Figure 4-7 Differential interactions between human Caf1 and BTG1 variants located in β -sheets regions.

(A) Structural model of the BTG1-hCaf1/CNOT7 complex. β -sheets are highlighted in blue. BTG1 variant residues predicted to be damaging to protein function by SIFT (score ≤ 0.05) are coloured yellow. UCSF Chimera was used for visualization. (B) Interactions between hCaf1 with BTG1 variants located in β -sheets. YRG2 yeast cells were transformed with vectors pAD-Gal4 or pAD-Gal4-hCaf1 and pBD-Gal4-HA or pBD-Gal4-HA-BTG1. β -Galactosidase activity was normalised using cell density (O.D. measured at 600 nm). Error bars indicate the standard error of the mean (n=3). One way Anova (dunnett post-hoc method) was used to calculate P values (** p<0.01 and **** p<0.0001). Three independent experiment were obtained. (C) Protein expression was confirmed by western blot analysis using Anti-HA and anti-AD antibodies.

4.4.3 Interaction analysis of BTG1 variants located in the loop regions and human Caf1

Eight variant residues present in the loop regions had been identified in NHL. Seven of these were predicted to affect the protein function using SIFT; E59D was predicted to have no effect on the function using SIFT (score >0.05). Using the structural model of the BTG1-hCaf1/CNOT7 complex, variant residues were located, P58L (L1), E59D (L2), G66V (α 3), N73K (L4) and E117D (L8) and shown to be close to the interaction surface. Only M11I (L1) is located away from the hCaf1/CNOT7 surface interaction. The Δ 10N BTG1 and H2Y were not resolved in the structural model of the BTG1 protein (Figure 4-8 A).

Using semi-quantitative interaction analysis using a yeast two-hybrid system, the ability of BTG1 variants to interact with hCaf1(CNOT7/CNOT8) was assessed. Using the BTG1 loop variants in yeast two-hybrid analysis, it was found that M11I, P58L, G66V and N73K, did not show β -galactosidase activity and therefore strongly impaired the interaction with hCaf1(CNOT7/CNOT8). By contrast, Δ 10N BTG1, H2Y and E117D were able to stimulate β -galactosidase activity similarly to the wildtype BTG1 and were therefore able to interact with hCaf1(CNOT7/CNOT8) although the E117D was marginally disrupted in the interaction with hCaf1/CNOT8 ($p < 0.001$). Interestingly, E59D, which was selected due the conservative amino acid change, was located close to the surface of the hCaf1/CNOT7 interaction and the yeast two-hybrid result showed a marginal disrupted in the interaction with hCaf1/CNOT7 and hCaf1/CNOT8 ($p < 0.05$ and $p < 0.01$), respectively (Figure 4-8 B). To confirm expression of the BTG1 variants, proteins were extracted from yeast cultures and subjected to western blot analysis using the anti-HA antibody for detection of Gal4 DBD-HA-BTG1 with an expected size 30kDa, however unspecific bands were detected below the DBD-HA-BTG1 bands. Also, the anti-AD-Gal4 antibody were used for detection Gal4 AD-hCaf1/CNOT7 and Gal4-AD-hCaf1/CNOT8 fusion proteins with expected size 58kDa (Figure 4-8 C).

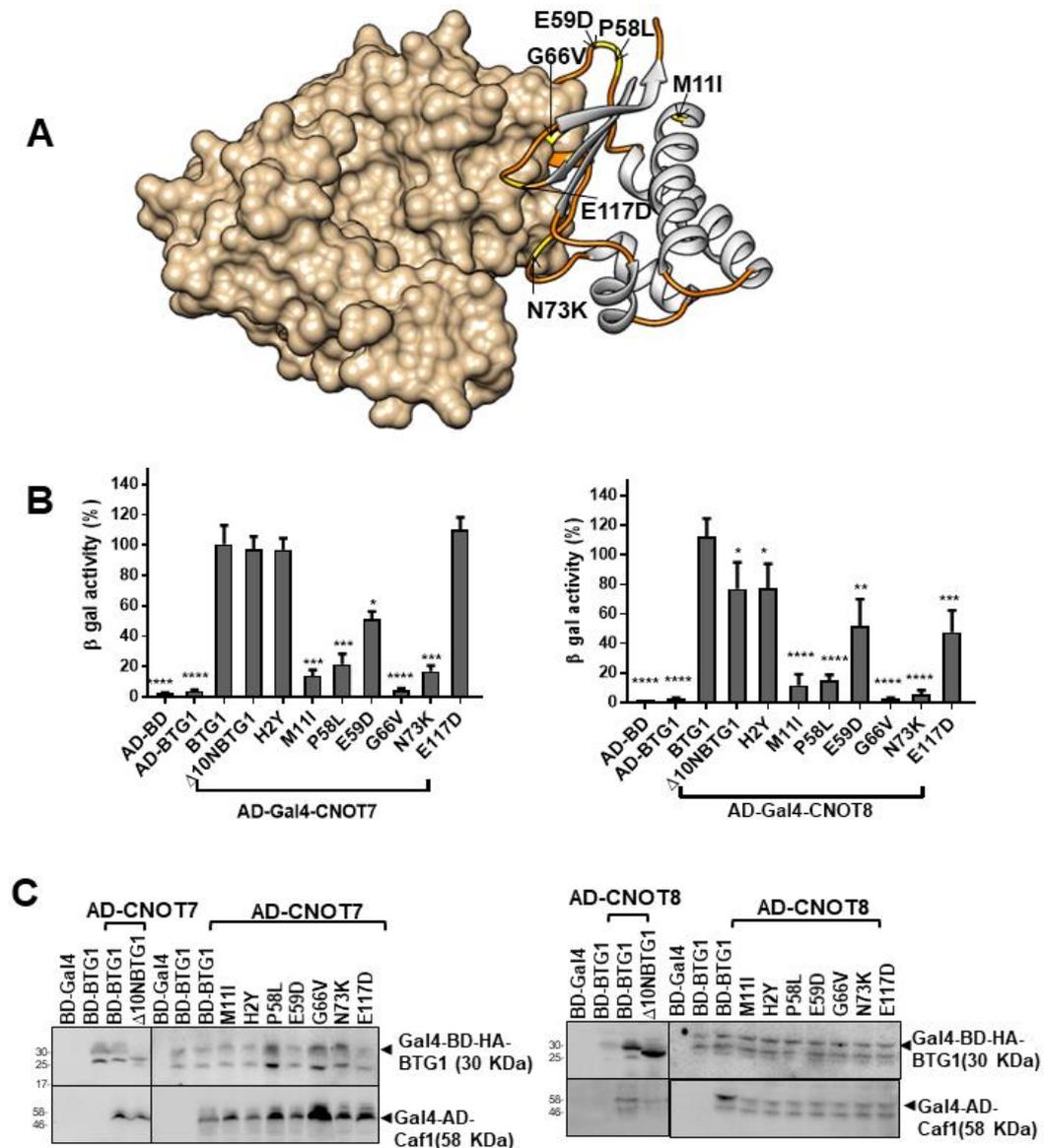


Figure 4-8 Differential interactions between human Caf1 and BTG1 variants located in C-loop regions.

(A) Structural model of the BTG1-hCaf1/CNOT7 complex. Loops are highlighted in orange. BTG1 variant residues predicted to be damaging to protein function by SIFT (score ≤ 0.05) are coloured yellow. UCSF Chimera was used for visualization. (B) Interactions between hCaf1 with BTG1 variants located in loops. YRG2 yeast cells were transformed with vectors pAD-Gal4 or pAD-Gal4-hCaf1 and pBD-Gal4-HA or pBD-Gal4-HA-BTG1. β -Galactosidase activity was normalised using cell density (O.D. measured at 600 nm). Error bars indicate the standard error of the mean (n=3). One way Anova (dunnett post-hoc method) was used to calculate P values (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$). Three independent experiment were obtained. (C) Protein expression was confirmed by western blot analysis using Anti-HA and anti-AD antibodies.

4.4.4 The effect of the deletion of the C-terminal 21 amino acids of BTG1 on the interaction with human Caf1

The semi-quantitative yeast two-hybrid analysis was used to assess the effect of the BTG1 C149del in the interaction with hCaf1/CNOT7 and hCaf1/CNOT8. When yeast was transformed with empty vector (pAD-Gal4 and pBD-HA-Gal4) or with pAD-Gal4 and pBD-Gal4-HA-BTG1, β -galactosidase activity was absent. This confirmed that BTG1 fused to the DNA binding domain did not contain a cryptic activation domain. However, when yeast was transformed with pBD-Gal4-HA-BTG1 and either pAD-Gal4- hCaf1/CNOT7 or pAD-Gal4- hCaf1/CNOT8, β -galactosidase expression was induced (Figure 4-9 A), in agreement with the known ability of BTG1 to interact with hCaf1(CNOT7/CNOT8) (Prévôt et al., 2001). Unexpectedly, BTG1 C149del was able to stimulate the β -galactosidase activity 100-fold higher than the wild type BTG1 and was therefore inferred to interact with hCaf1(CNOT7/CNOT8). Proteins were extracted from the yeast cultures and used for western blot analysis using anti-HA antibody. The protein analysis showed a higher expression of BTG1 C149del compared to the wild type BTG1, as shown in the (Figure 4-9 B).

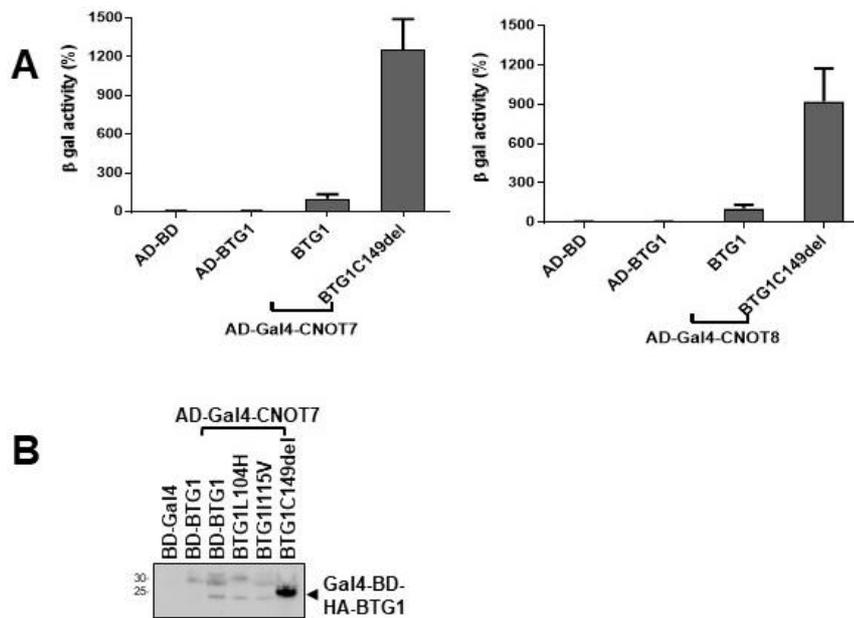


Figure 4-9 Evaluation of interaction between BTG1 lacking the C-terminal 21 amino acids and human Caf1.

(A) Interactions between hCaf1 with BTG1 C149del. YRG2 yeast cells were transformed with vectors pAD-Gal4 or pAD-Gal4-hCaf1 and pBD-Gal4-HA or pBD-Gal4-HA-BTG1. β -Galactosidase activity was normalised using cell density (O.D. measured at 600 nm). Error bars indicate the standard error of the mean (n=3). Three independent experiments were obtained. (B) Protein expression was confirmed by western blot analysis using Anti-HA antibody.

4.5 Discussion

The BTG/TOB protein family regulates cell proliferation, and it is involved in mRNA degradation through the interaction with hCaf1(CNOT7/CNOT8). Caf1 plays a role in translational repression and mRNA degradation by shortening the poly (A) tail of cytoplasmic mRNA (Ezzeddine et al., 2007, Funakoshi et al., 2007, Mauxion et al., 2008). Based on the protein sequences, hCaf1/CNOT7 and hCaf1/CNOT8 are 89% similar and 76% identical at the amino acid level (Winkler and Balacco, 2013). It has previously been concluded that BTG/TOB protein interactions with hCaf1(CNOT7/CNOT8) are required for the deadenylase activity (Doidge et al., 2012a, Ezzeddine et al., 2012). However, it is unclear whether BTG1 variants identified in lymphoma samples affect the interactions with hCaf1(CNOT7/CNOT8). In the previous chapter, 16 BTG1 variants identified in NHL samples were predicted to have impact on protein function. In this chapter, these BTG1 mutations, as well as an additional variant with a conservative amino acids substitution (E59D), were assessed in the interaction with hCaf1(CNOT7/CNOT8) using the yeast two-hybrid analysis.

4.5.1 BTG1-hCaf1/CNOT7 complex

To understand the role of BTG1 variants in the interaction with hCaf1(CNOT7/CNOT8), a structural model was required. However, the BTG1–Caf1 complex and BTG1 secondary structure are not yet published. By contrast, the structure of the domain encompassing the N-terminal 138 residues of TOB1 and the human BTG2 structure, are available (Yang et al., 2008, Horiuchi et al., 2009). The BTG/TOB proteins family share the conserved BTG domain. The primary protein sequence of the BTG domain of BTG1 and BTG2 which is highly conserved (96% similarity and 74% identity) (Winkler, 2010). Therefore, a model of BTG1 was created based on similarity with BTG2. The BTG1 model was generated using phyre2 (Kelley et al., 2015). This tool used a single sequence of human BTG1 (P62324) to generate a high-confidence model (100%), covering 72% of protein sequence with BTG domain-like secondary structure (ID: d3e9va1) and aligned with BTG2 (PDB

accession number 3DJU) and TOB1-hCaf1/CNOT7. The model structure of the BTG1-hCaf1/CNOT7 complex helped us to understand the location of the variant residues in relation to the interaction surface with hCaf1(CNOT7/CNOT8).

4.5.2 BTG1 variants located in box A and box B impact on the interaction with human Caf1

Previously, it has been reported that the residues located specifically in box A and box B of the BTG domain regions of BTG/TOB proteins are required for interaction with hCaf1(CNOT7/CNOT8) (Prévôt et al., 2001, Yang et al., 2008, Horiuchi et al., 2009). Consistent with this notion, the results presented show that the L104 and I115 residues, located in the box B region of BTG1, are important for the interaction with hCaf1(CNOT7/CNOT8). Changing the amino acids to histidine and valine, respectively, impacted the interaction with hCaf1(CNOT7/CNOT8). In addition, both variant residues, located in β 2 (L104H) and β 3 (I115V), are close to the hCaf1/CNOT7 surface suggestive of an effect on the interaction (Figure 4-7 A). Also, it has been demonstrated that conserved residues located in box A such as BTG2 residues Y65 and D75 or TOB1 residues F55 and D65 are required for the interaction with hCaf1(CNOT7/CNOT8) (Doidge et al., 2012a). In agreement with this notion, P58, G66 and N73 of BTG1, which are localised in the box A region, were required for the interaction with hCaf1(CNOT7/CNOT8). Interestingly, the variant residues localised in L1 (P58L), α 3(G66V) and L4 (N73K) are also close to the Caf1 surface, thus explaining their effect on the interaction with hCaf1(CNOT7/CNOT8) (Figure 4-8 A). In contrast, the variant residue E117D located in L8, whilst close the hCaf1/CNOT7 surface, still allows the interaction with hCaf1/CNOT7. This is in agreement with Doidge et al., (2012), who demonstrated that the corresponding residues E115 in BTG2 and E105 in TOB1 are not necessary in the interaction. Although, it has been suggested that K51 and E105 of TOB1 stabilised the interaction by binding hCaf1/CNOT7 residues E247 and Y260, respectively (Horiuchi et al., 2009). Finally, the E59D

variant residue was selected due to the conservative amino acid change. Unexpectedly, this mutation is marginally impaired in the interaction with hCaf1/CNOT7 and hCaf1/CNOT8 ($p < 0.05$ and $p < 0.01$, respectively). One of the explanation could be that the E59D is located in L2, which is also close to the interaction surface (Figure 4-8 A).

4.5.3 BTG1 variant residues located outside of box A and box B regions impact on the interaction with human Caf1

Six variant residues of BTG1 were located to be outside box A and B. Two of these variants located to $\alpha 2$ (Q36H and L37M) and away from the interaction surface and, as expected, were shown to interact with hCaf1 (CNOT7/CNOT8). By contrast, F25C and R27H are located in the $\alpha 1$ structure, close to the interaction surface, and were shown to affect binding with hCaf1(CNOT7/CNOT8) (Figure 4-6 A). Interestingly, the conserved residue F40 is also located in $\alpha 2$, but the opposite side of the hCaf1/CNOT7 binding surface (Figure 4-6 A), and this variant abrogated the hCaf1(CNOT7/CNOT8) interaction. A possible explanation for this is that F40 is a conserved hydrophobic residue located in the long helical structure corresponding to TOB1 F28, which has been suggested as being important in supporting the β -strands of box B in the interaction (Horiuchi et al., 2009). However, the identity of BTG1 and TOB1 at the level of amino acids is 45% (Table 4-1) so this would need further work to confirm. Surprisingly, the variant M11 is required for binding with hCaf1(CNOT7/CNOT8), although is located in the L1 region away from the interaction surface (Figure 4-8 A and B). Finally, the variant with a deletion of 21 amino acids in the C-terminal of BTG1, showed this region was not essential for the interaction with hCaf1(CNOT7/CNOT8). The high protein expression level of the BTG1 C149del might explain the unexpected result in the β -galactosidase activity, it is 100-fold higher than observed with wild type of BTG1 (Figure 4-9), one of the possibility is that the absence of the last amino acids of BTG1, increased the protein stability. This result confirmed by another study suggesting that the C-terminal regions of BTG/TOB proteins act as

protein degradation signals (Sasajima et al., 2002). Further experiments are required to confirm protein stability and folding structure, such as circular dichroism (CD) spectroscopy, which allows for the characterisation of secondary and tertiary protein structure, or differential scanning fluorimetry (DSF), which provides information about the stability of the folded protein; increased stability is reflected in a higher melting temperature (T_m).

One of the limitations of using the yeast two hybrid analysis, that fused yeast reporter proteins or anchors may cause steric hindrance that impedes interaction. Alternatively, the presence or absence of different post-translational protein modifications in yeast compared to that of higher eukaryotes may influence the ability of proteins to interact (Brückner et al., 2009).

In summary, nine somatic BTG1 variants were found to interfere with the interaction with the hCaf1(CNOT7/CNOT8) deadenylase subunit of the CCR4-NOT complex. Five of these variants are located in the box A and box B regions, which is consistent with the notion that these regions are important for the interaction with hCaf1(CNOT7/CNOT8). Surprisingly, M11I, F25C, R27H and F40C variants, which are not localised in those regions, also impacted on the interaction with hCaf1(CNOT7/CNOT8).

Chapter 5
**The effect of lymphoma-derived BTG1 mutations on
the regulation of cell proliferation**

Chapter 5: The effect of lymphoma-derived BTG1 mutations on the regulation of cell proliferation

5.1 Introduction

Expression of BTG/TOB protein family members inhibit cell cycle progression. The expression of BTG/TOB proteins in cell lines has been shown to increase the number of cells in the G1 phase, which suggests that the BTG/TOB proteins inhibit the transition from G1 to S phase (Rouault et al., 1992, Ikematsu et al., 1999, Hata et al., 2007). BTG1 expression acts as a negative regulator of in several cell types including fibroblasts, myoblasts, macrophages and T-lymphocytes (Rouault et al., 1992, Suk et al., 1997, Rodier et al., 1999). Additionally, BTG1 is involved in the regulation of cell differentiation by acting as a coactivator of transcription factors in myoblast cells (Busson et al., 2005). In erythroid cells, BTG1 expression is induced by the FoxO3A transcription factor (Bakker et al., 2004). Interestingly, the overexpression of either BTG1 or BTG2 in the murine B-lymphoma cell line WEH-231, prevents cell cycle progression during the G1/S phase and reduces the number of S-phase cells (Hata et al., 2007). Moreover, the overlapping functions of BTG1 and BTG2 have been described in B-lymphopoiesis; the number of progenitor B-cells decreases significantly by 40% in BTG1 and BTG2 double knockout mice, whereas the number of these progenitors was only reduced by 20% in mice deficient in BTG1 or by 10% in mice deficient in BTG2 (Tijchon et al., 2016). Although the expression of BTG1 and BTG2 regulates the early stage of B-cell development, the expression of critical regulators of B-cell lineage commitment, such as PAX5 and EBF1, is not inhibited in BTG1 and BTG2 double knockout mice (Nechanitzky et al., 2013, Tijchon et al., 2016). However, in these double knockout mice, the expression of certain genes in the non-fate B cell lineage are significantly increased. These genes included *CD4*, *IKZF2*, *TCF7*, *GATA3* and *NOTCH1*, which suggests that BTG1 and BTG2 are implicated in regulating B-cell lineage commitment by suppressing the

expression of T-cell lineage genes (Tijchon et al., 2016). Deletions of BTG1 have been found to be common in BCP-ALL cases (Xie et al., 2014, Scheijen et al., 2016). Moreover, somatic mutations of both BTG1 and BTG2 were found to occur frequently in B-cell lymphoma, which provides an important indication of their role as a possible tumour suppressor in B-cell malignancies (Tijchon et al., 2016). It also appears that the anti-proliferative activity of BTG/TOB proteins requires their interaction with hCaf1(CNOT7/CNOT8): TOB1/TOB2 and BTG2 variants unable to interact with hCaf1(CNOT7/CNOT8), do not inhibit cell cycle progression (Doidge et al., 2012a, Ezzeddine et al., 2012). Conversely, hCaf1/CNOT7 requires the ability to interact with TOB1 in order to inhibit cell proliferation. Cells that were transfected with mutant hCaf1/CNOT7 E247A and Y260A, were unable to interact with BTG2 and the rate of cell proliferation in mutant-expressing MCF7 cell lines was slightly reduced, compared to that of cells with endogenous hCaf1/CNOT7 knockdown by siRNA (Aslam et al., 2009). In this chapter, BTG1 variants identified in NHL samples (Morin et al., 2011, Lohr et al., 2012, Zhang et al., 2013), were selected for further analysis.

In Chapter 3, SIFT predicted that 15 BTG1 variants were likely to affect protein function and these mutations are further studied in this thesis with an additional conservative variant predicted not to affect function, namely E59D, also included. With the evidence summarised above, it was important to understand the effect of the BTG1 variants on proliferation function. Work presented in Chapter 4 concluded that ten BTG1 variants were affected in their ability to interact with hCaf1(CNOT7/CNOT8) (Table 5-1). Based on previous work, it can be predicted that these variants are not able to inhibit cell proliferation. To address this issue, HEK293T cells were transfected with the BTG1 variants in order to conduct a proliferation assay. The S-phase cell cycle progression was detected using 5-ethynyl-2'-deoxyuridine (EdU), a nucleoside analogue of thymidine, which is incorporated during DNA replication and can be labelled with fluorophores using click chemistry (Figure 5-1).

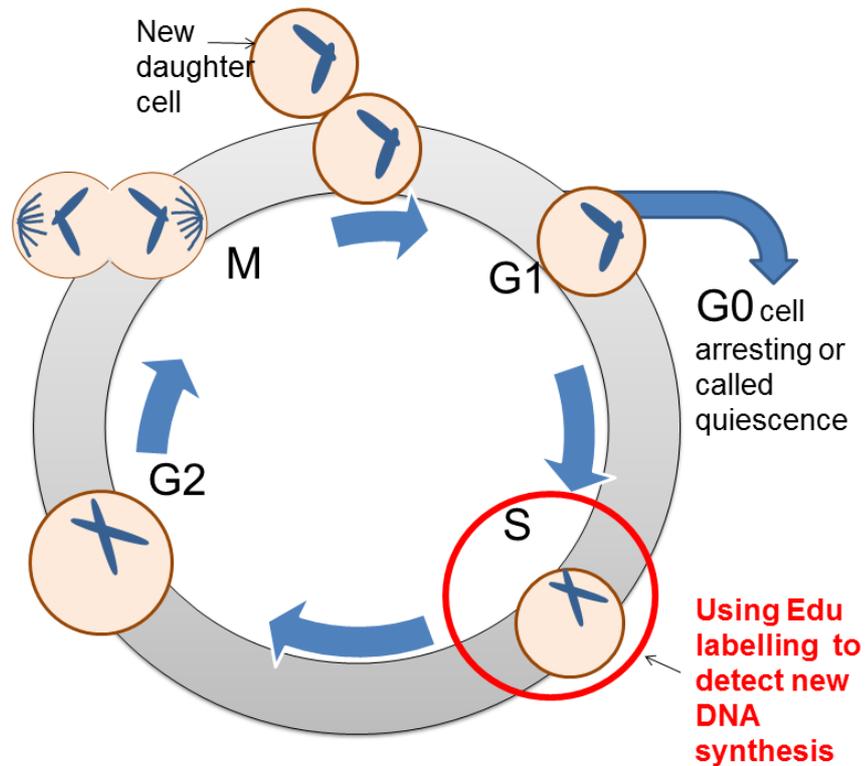


Figure 5-1 Cell cycle progression and EdU labeling of dividing cells.

There are four distinct phases in the cell cycle., the first phase is the period of cell growth called G1, then new DNA is synthesised and duplicated in S phase. After that, the cell prepares for division (G2 phase). The cell enters the mitotic (M) phase of the cell cycle and the chromosome pairs condenses and attaches to fibres that pull sister chromatids to opposite sides of the cell, producing two identical daughter cells. Sometimes, cells can enter a quiescence state, going from the G1 phase to a G0 state. The detection of dividing cells in vitro is via the incorporation of the thymidine analogue 5-ethynyl-2'-deoxyuridine (EdU) into newly synthesised DNA during S-phase.

Variant residue	Location	SIFT (score≤0.05)	Suspect (score>50)	Protein-protein interaction analysis	
				hCaf1/CNOT7	hCaf1/CNOT8
F25C	α1	0.001	68	+	++
R27H	α1	0.039	20	-	-
F40C	α2	0	71	-	-
L104H	β2	0	89	-	-
I115V	β3	0.003	29	+	+
M11I	L1	0.028	87	-	-
P58L	L2	0	53	+	-
G66V	α3	0.001	70	-	-
N73K	L4	0.001	24	-	-
E117D	L8	0.022	14	++++	++
Q36H	α2	0.006	53	++++	+++
L37M	α2	0.02	31	+++	++
L94V	α5	0.002	28	+++	+
Δ10N BTG1	-	0	4	++++	+++
H2Y	-	0.037	17	++++	+++
E59D	L2	0.0541	2	++	++

Table 5-1 Data analysis of BTG1 variants using bioinformatics and yeast-two-hybrid interaction analysis.

Summary of the data analysis obtained in Chapters 3 and 4. Variant residues are coloured based on their location in the BTG1 structure, green (α-helical), blue (β-sheet) and orange (L-loops). Bioinformatics analysis (SIFT and Suspect) was used to predict the effect of the variants on protein function and yeast two-hybrid analysis was used to evaluate the effect of the BTG1 variants on their interaction with hCaf1(CNOT7/CNOT8). The relative values of β-Gal activity to BTG1 are indicated in different ranges such as 0-20% (-), 21-40% (+), 41-60% (++) , 61-80% (+++) and 81-100% (++++).

5.1 Generation of plasmids for the proliferation assay

To identify the effect of the BTG1 variants on the regulation of the cell cycle *in vitro*, mammalian expression plasmids were prepared.

PCR was used to amplify the BTG1 cDNA to include an EcoRI restriction site at the 5' end and a Sall restriction site at the 3' end. Then, the BTG1 cDNA was digested with the EcoRI and Sall restriction enzymes and ligated into the mammalian expression plasmid pCMV5-HA (Figure 5-2 A). The BTG1 variant cDNAs were obtained by two rounds of PCR amplification. The first amplification was to achieve BTG1 fused with the DNA binding domain of Gal4, using a 5' primer incorporating the Gal4-DNA binding domain, in conjunction with a 3' primer containing a BTG1 stop codon followed by a Sall restriction site. The second amplification performed introduced the EcoRI and Sall restriction sites at the 5' and 3' ends, respectively. Finally, the DNA fragments were digested with EcoRI and Sall and ligated into the pCMV5-HA vector (Figure 5-2 B and C).

In order to clone the two deletion mutations, Δ 10N BTG1 and BTG1 C149del, primers were designed as described in section 4.3 (Generation of plasmids for yeast two-hybrid interaction analysis).

The plasmid contains an HA tag inserted at the N-terminus, which facilitates the detection of BTG1 expression by western blotting. All the BTG1 cDNA sequences were confirmed by DNA sequencing.

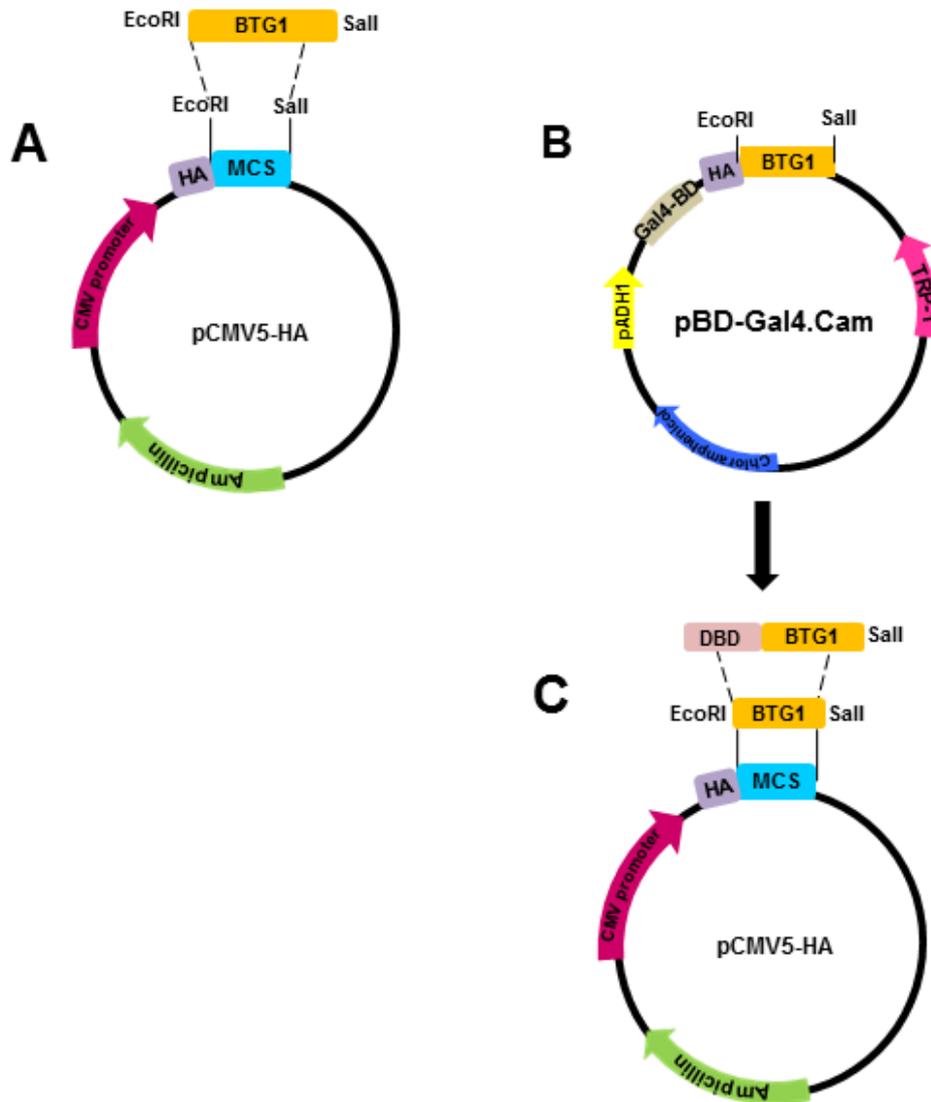


Figure 5-2 Generation of vectors for the proliferation assay.

(A). The BTG1 cDNA was amplified by PCR. The 5' primer included an EcoRI restriction site, and the 3' primer contained a stop codon followed by a Sall restriction site. The cDNA fragments were ligated into the EcoRI and Sall restriction sites of mammalian expression vector pCMV5-HA. The cDNAs of BTG1 variants were prepared by two step amplifications. (B) In the first round of PCR, BTG1 variants were amplified using template pBD-HA-BTG1 and the 5' primer containing the DBD and the 3' primer contained a stop codon followed by a Sall restriction site. (C) In the second round of PCR, DBD-BTG1 fragments were used as a template and amplified to introduce the EcoRI and Sall restriction sites at the 5' and 3' ends, respectively. Then, the DNA fragments were cloned into the EcoRI and Sall sites of expression vector pCMV5-HA.

5.2 Optimisation of 5-ethynyl-2'-deoxyuridine (EdU) staining for HEK293T cells.

To evaluate the effect of the BTG1 variants on the regulation of the cell cycle, a proliferation assay was performed. In HEK293T cells, the extent and intensity of the DNA labelling following a pulse of EdU thymidine analogue during S-phase correlates with the length of time of EdU labelling. Therefore, optimisation of the duration of the EdU pulse applied to the cells was required. The cells were transfected with a green fluorescent protein (GFP)-expressing plasmid in order to determine the transfection efficiency. Furthermore, the cells were transfected with either pCMV-HA-BTG1 or empty vector as a control using the calcium phosphate precipitation method. Forty-eight hours after transfection, the cells were harvested to assess the transfection efficiency and cell proliferation was determined by EdU incorporation. The calculation of the percentage of cells that expressed GFP (75–90%) with nuclei stained with Hoechst (DAPI) determined the transfection efficiency. The thymidine analogue EdU (10 μ M) was added to the HEK293T cell medium for 20, 60 and 120 minutes to label the cells in S-phase. The percentage of cells in S-phase was calculated using ImageJ by counting the nuclei labelled with the EdU thymidine analogue, compared to the nuclei stained with Hoechst (DAPI).

The results showed that the number of S-phase cells was reduced by 20% in cells expressing BTG1, when compared to the control (empty vector) in all three conditions ($p < 0.05$ in the 60-minute condition and $p < 0.01$ in 20- and 120-minute conditions) (Figure 5-3 A). This confirmed that BTG1 expression was able to inhibit cell proliferation. Thus, the duration of the labelling with EdU did not affect the difference between the control and BTG1 in terms of the percentage of S-phase cells. However, it showed a significant difference between the control and BTG1-transfected cells when the fluorescence intensity measurement was used. The cells pulsed with EdU for 60 or 120 minutes were very intensely stained in the nucleus following the reaction with Alexa 488 azide. In contrast, the cells pulsed with EdU for a short period of 20

minutes displayed only a low staining intensity and they were not easy to count (Figure 5-4 A and B). Moreover, the fluorescence intensity increased proportionately with the length of time of EdU labelling. When comparing the cells expressing BTG1 with the control, the fluorescence intensity demonstrated a two-fold reduction in the 20- and 60-minute conditions ($p < 0.0001$). Interestingly, when the cells were pulsed for longer (120 minutes), the difference between the cells expressing BTG1 and the control in terms of the fluorescence intensity, was the exhibition a significant four-fold decline ($p < 0.0001$) (Figure 5-3 B). Thus, the most appropriate condition to use for the HEK293T cells in the proliferation assay was labelling with 10 μ M EdU for 120 minutes, because the extended time would allow the detection of smaller differences between BTG1 variants.

It is important to note that the Click-iT EdU Alexa 488 azide reagent was changed to Click-iT EdU Alexa 549 azide due to the failure of an alternative experiment (Figure 5-4 C and D). That experiment was the co-transfection of two plasmids, namely GFP and pCMV5-HA-BTG1 due to the assumption that cells that expressed GFP would also express BTG1. However, the fluorescent staining was destroyed due to the copper used in the Click-iT reaction setup. An attempt was made to chelate the copper with EDTA but the experiment still did not work (Appendix B).

To confirm the expression of BTG1, the cells were harvested after 48 hours of transfection. They were then subjected to western blot analysis using the anti-HA antibody for the detection of HA-BTG1 and the γ -tubulin antibody for the detection of housekeeping protein expression (Figure 5-3 C).

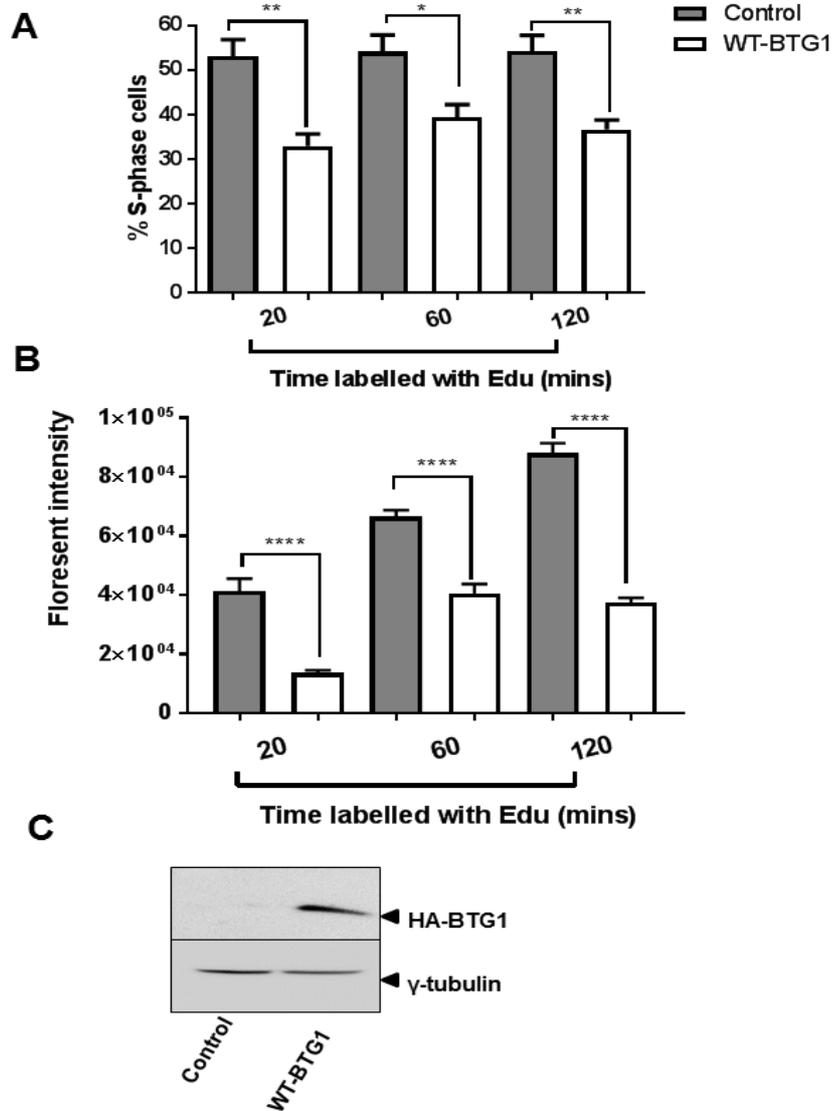


Figure 5-3 Optimisation of staining of HEK293T cells with EdU.

HEK293T cells were seeded onto glass coverslips prior to transfection with pCMVHA-BTG1. An empty vector was used as the control. After 48 hours, the cells were pulse-labelled for 20, 60 or 120 minutes with the nucleoside analogue EdU (10 μ M) and then detected using a fluorescence microscope. (A) Comparison of cells transfected with WT-BTG1 and the control (empty vector) in terms of the percentage of S-phase cells following different labelling times. (B) Comparison of the difference in fluorescence intensity of cells stained with EdU in cells expressing BTG1 and control cells. All the samples were tested in biological triplicates and the nuclei were counted for three images from each slide. ImageJ was used to count and measure the fluorescence intensity. The mean values and standard error of the mean (n=3) are indicated. An unpaired t-test was used to calculate the significance of the values (*p<0.05, **p<0.01 and ****p<0.0001) when comparing BTG1-expressing cells to the control cells. (C) Western blots confirming the protein expression of the HA tag which is fused to BTG1 in the HA-BTG1 expressing cells, with the housekeeper anti- γ -tubulin expressed in both BTG1 and control cells.

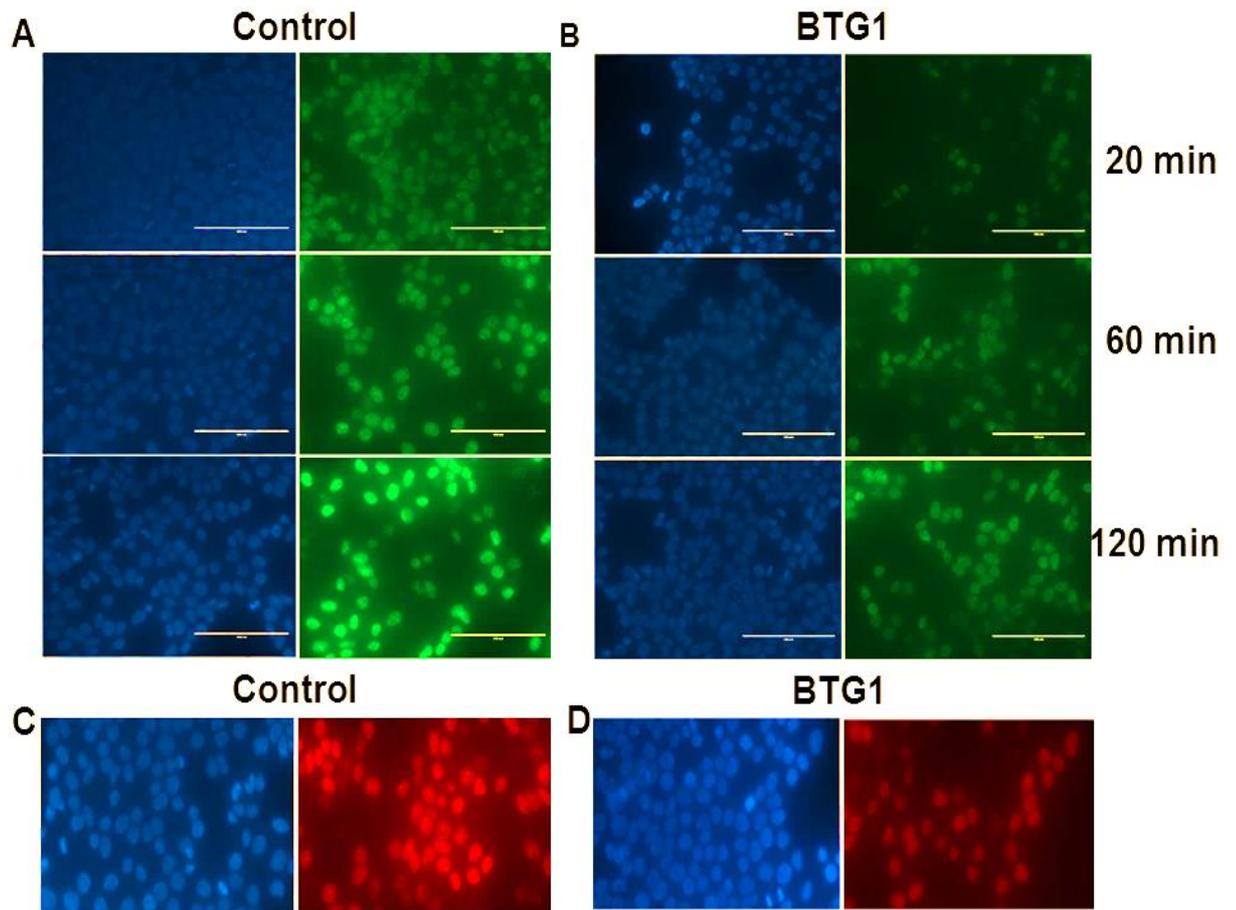


Figure 5-4 Comparison of the difference in fluorescence intensity of cells stained with EdU, in both cells expressing WT-BTG1 and the control.

HEK293T cells were seeded onto glass coverslips prior to transfection with the control (empty vector) and pCMV5-HA-BTG1. After 48 hours, cell nuclei were pulse-labelled for 20, 60 or 120 minutes with 10 μ M of the nucleoside analogue EdU and detected using Alexa Fluor® 488. (A) Cells expressing control, (B) Cell expressing BTG1. (C) and (D) The nuclei of cells expressing control or BTG1, were pulse-labelled for 120 minutes with 10 μ M of the nucleoside analogue EdU and detected using Alexa Fluor® 549. All cell nuclei were stained with Hoechst staining (DAPI). The fluorescent cell nuclei were detected using an Evos FL fluorescence microscope. The nuclei were counted for three images from each slide. ImageJ was used to count and measure the fluorescence intensity.

5.3 Regulation of cell proliferation by BTG1 variants

Ten BTG1 variants were shown to be impaired in terms of their interaction with hCaf1(CNOT7/CNOT8) (Chapter 4). All these variants were identified in NHL, and were predicted to affect protein function using SIFT. Three variant residues were located in α -helices, two in β -sheets and the remaining variant residues were located in loop regions. In order to determine whether the impaired hCaf1(CNOT7/CNOT8) interaction of the BTG1 variants affected the ability of BTG1 to inhibit cell cycle progression, HEK293T cells were transfected with pCMV5-HA-BTG1 variants or an empty vector, which acted as the control, using the calcium phosphate precipitation method. Additionally, the cells were transfected with a GFP expression plasmid for assessment of transfection efficiency. Each sample was transfected in triplicate in three different wells. After 48 hours of transfection, the cells transfected with GFP expression plasmid were harvested and stained with Hoechst. Their transfection efficiency was found to be 75–90%. For the proliferation assay, the percentage of cells in S-phase was determined.

5.3.1 Effect of BTG1 variants mapped to α -helices on the anti-proliferative function

Three variant residues, which were impaired in terms of their interaction with hCaf1(CNOT7/CNOT8), were located in α -helices. The results showed that the number of S-phase cells was significantly reduced by approximately 20% in the cells expressing wild type BTG1 when compared to the control (empty vector) cells, confirming the anti-proliferative effect of this protein. However, when the BTG1 variants were expressed, it was found that the R27H and F40C variants exhibited no alteration in the number of S-phase cells as compared to the empty vector (control), and they were unable to inhibit cell proliferation ($p < 0.0001$ and $p < 0.001$, respectively). In the case of cells expressing BTG1 F25C, a reduction in proliferative activity was seen when compared to the empty vector cells ($p < 0.05$), however this was to a lesser extent than the inhibition seen with wild type BTG1 ($p < 0.01$). This result indicates that whilst

the F25C variant has altered activity, it does not totally abolish its function (Figure 5-5 A).

To confirm BTG1 protein expression, the cells were harvested after 48 hours of transfection and the proteins were subjected to western blot analysis using an anti-HA antibody. The proteins were shown to be expressed, except in the cells where BTG1 F40C was only weakly expressed (Figure 5-5 B). Three independent experiments were conducted and confirmed the weak expression. This suggests that this variant may be less stable than the wild type protein.

5.3.2 Effect of BTG1 variants mapped to β -sheets on the anti-proliferative function

Two variant residues, which were impaired in terms of their interaction with hCaf1(CNOT7/CNOT8), were located in β -sheets. Cells expressing wild type BTG1 showed a reduction in the number of cells in S-phase ($\geq 20\%$). This was not seen with BTG1 I115V expression, indicating that this variant results in loss of anti-proliferation activity ($p < 0.0001$). In addition, cells expressing BTG1 L104H showed the number of S-phase cells as being between the BTG1 wild type and control (empty vector) cells, suggesting that this mutation diminishes the anti-proliferative activity of BTG1, but does not result in its loss entirely ($p < 0.01$ and $p < 0.0001$, respectively) (Figure 5-6 A).

To confirm the expression of the BTG1 variants, cells were subjected to western blot analysis. All the protein variants were similarly expressed (Figure 5-6 B).

5.3.3 Effect of the BTG1 variants mapped to loop regions on the anti-proliferative function

Four BTG1 variant residues impaired the interaction of this protein with hCaf1(CNOT7/CNOT8), namely M11I, P58L, G66V and N73K. Also, the E117D mutant marginally impaired the interaction with hCaf1/CNOT8 (Table 5-1). All these variants were located in loop regions, with the exception of G66V,

which is located in helix α 3. Cells expressing wild type BTG1 were able to reduce the number of S-phase cells by 20% compared to the control ($p < 0.0001$), indicating that BTG1 is able to regulate cell proliferation. In contrast to wild type BTG1, cells expressing BTG1 M11I, P58L or G66V were completely unable to inhibit their proliferation ($p < 0.0001$). The same results were observed in cells expressing BTG1 N73K or E117D, indicating that these BTG1 mutants lost their anti-proliferative function ($p < 0.0001$) (Figure 5-7 A and B).

The protein expression of BTG1 variants was confirmed using western blot analysis. All the protein variants were similarly expressed (Figure 5-7 C and D).

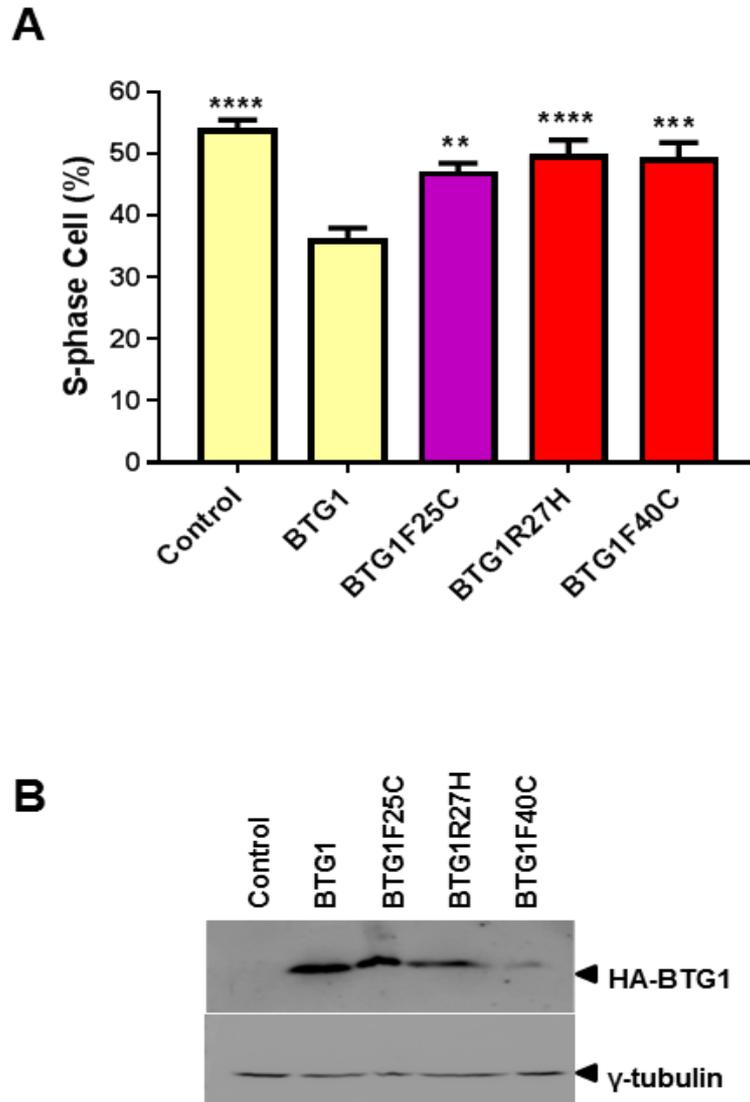


Figure 5-5 Effect of BTG1 α -helix variants on the anti-proliferative function.

(A) Percentage S phase cells in cells transfected with BTG1 variants which are located in α -helices. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with red indicating an inability to interact, purple indicating moderate interaction and yellow for the control and wild type BTG1. All samples were tested in biological triplicates and the number of nuclei were counted for three images from each slide. The error bars indicate the standard error of the mean ($n=3$). One way Anova (tukey's post-hoc method) was used to calculate P values (** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$) compared to cells transfected with pCMV5-HA-BTG1. Three independent experiment were obtained. (B) Western blots confirming protein expression of HA-BTG1 using an anti-HA antibody. Anti- γ -tubulin was used as a loading control.

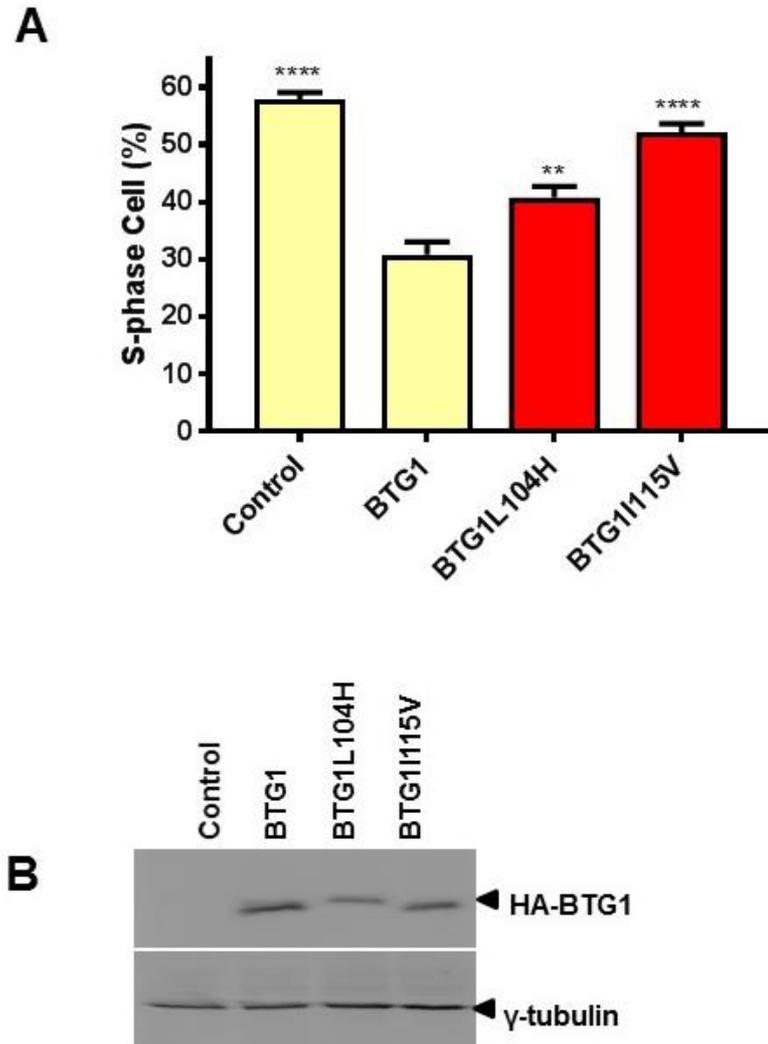


Figure 5-6 Effect of BTG1 β -sheet variants on the anti-proliferative function.

(A) Percentage S phase cells in cells transfected with BTG1 variants which are located in β -sheets. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with red indicating an inability to interact, and yellow for the control and wild type BTG1. All samples were tested in biological triplicates and the number of nuclei were counted for three images from each slide. The error bars indicate the standard error of the mean (n=3). One way Anova (tukey's post-hoc method) was used to calculate P values (**p <0.01 and ****p <0.0001) compared to cells transfected with pCMV5-HA-BTG1. Three independent experiment were obtained. (B) Western blots confirming protein expression of BTG1 using an anti-HA antibody. Anti- γ -tubulin was used as a loading control.

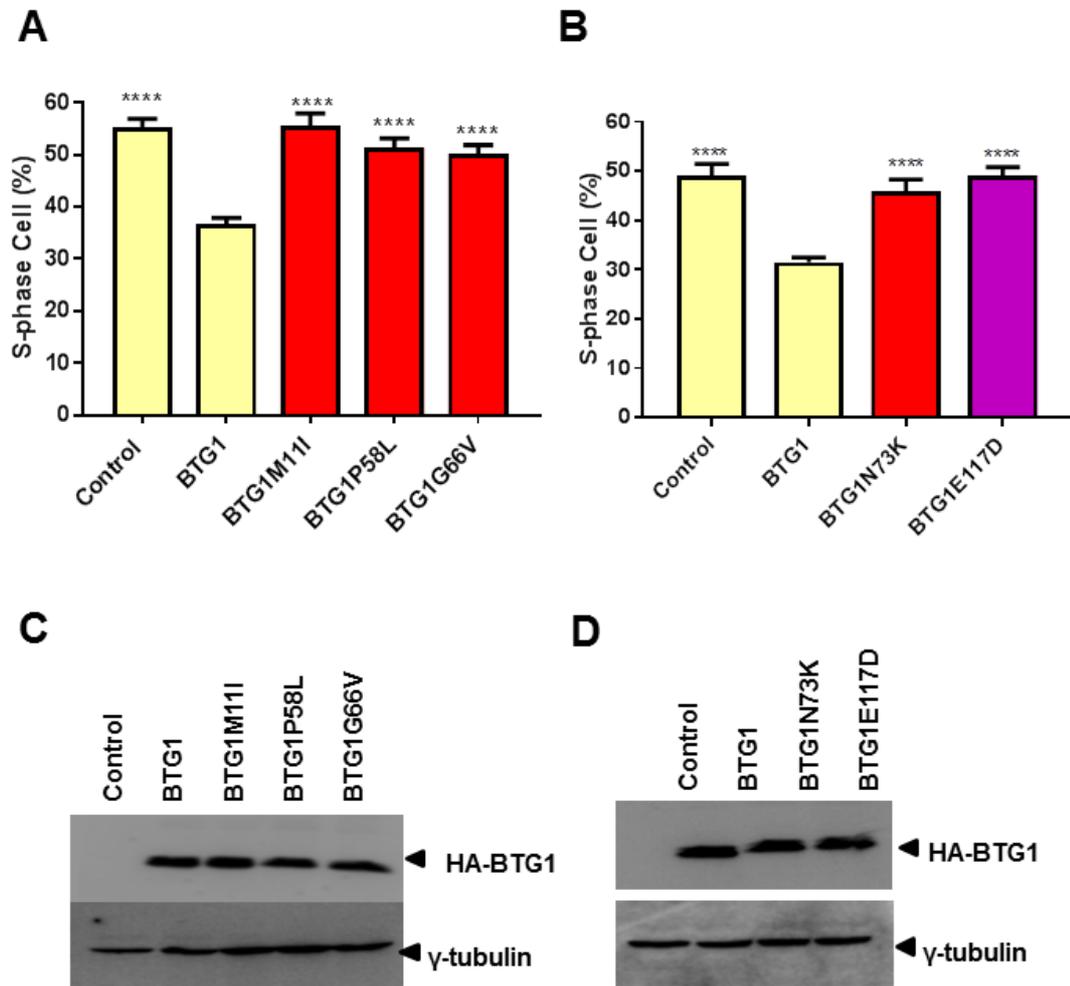


Figure 5-7 Effect of BTG1 loop region variants on the anti-proliferative function.

(A) and (B) Percentage S phase cells in cells transfected with BTG1 variants located in loop regions. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with red indicating an inability to interact, purple indicating moderate interaction and yellow for the control and BTG1. All samples were tested in biological triplicates and the number of nuclei were counted for three images from each slide. The error bars indicate the standard error of the mean ($n=3$). One way Anova (tukey's post-hoc method) was used to calculate P values (**** $p < 0.0001$) compared to cells transfected with pCMV5-HA-BTG1. Three independent experiments were obtained. (C) and (D) western blots confirming protein expression of BTG1 using an anti-HA antibody. Anti- γ -tubulin was used as a loading control.

5.4 Investigation of the anti-proliferative activity of BTG1 variants that are able to interact with human Caf1 (CNOT7/CNOT8)

Three BTG1 variants identified in NHL were able to interact with hCaf1(CNOT7/CNOT8), while a further three only had a modest effect on the interaction ($\geq 50\%$ able to interact with hCaf1(CNOT7/CNOT8)). The SIFT analysis predicted all these variants would have an impact upon protein function (score ≤ 0.05), except for E59D. One additional variant, namely BTG1 C149del, showed an unexpected result in the interaction with hCaf1(CNOT7/CNOT8), and using PROVEN software was predicted to alter the protein function. However, the effect of all of these mutations on anti-proliferative activity is unknown. The BTG1 variants were analysed using a proliferation assay to determine whether the anti-proliferative activity of BTG1 requires an interaction with hCaf1(CNOT7/CNOT8). The results showed that the cells expressing the H2Y and Q36H variants were able to reduce the number of S-phase cells by 20%, the same extent as wild type BTG1. In contrast, the cells expressing BTG1 C149del exhibited no alteration in the number of S-phase cells, suggesting an inability to inhibit cell proliferation (Figure 5-8 A). The L37M and E59D BTG1 variants reduced the number of S-phase cells, therefore these variants did not completely lose their anti-proliferative activity ($p < 0.05$ and $p < 0.01$, respectively). Furthermore, the $\Delta 10N$ BTG1 and BTG1L94V variants both displayed a reduction in proliferative activity when compared to the control (empty vector) cells ($p < 0.001$ and $p < 0.05$, respectively). However this was to a lesser extent than the inhibition seen with wild type BTG1 (Figure 5-8 B).

Finally, to confirm comparable BTG1 protein expression, the cells were harvested after 48 hours of transfection and subjected to western blot analysis using an anti-HA antibody. All the protein variants were similarly expressed (Figure 5-8 C and D).

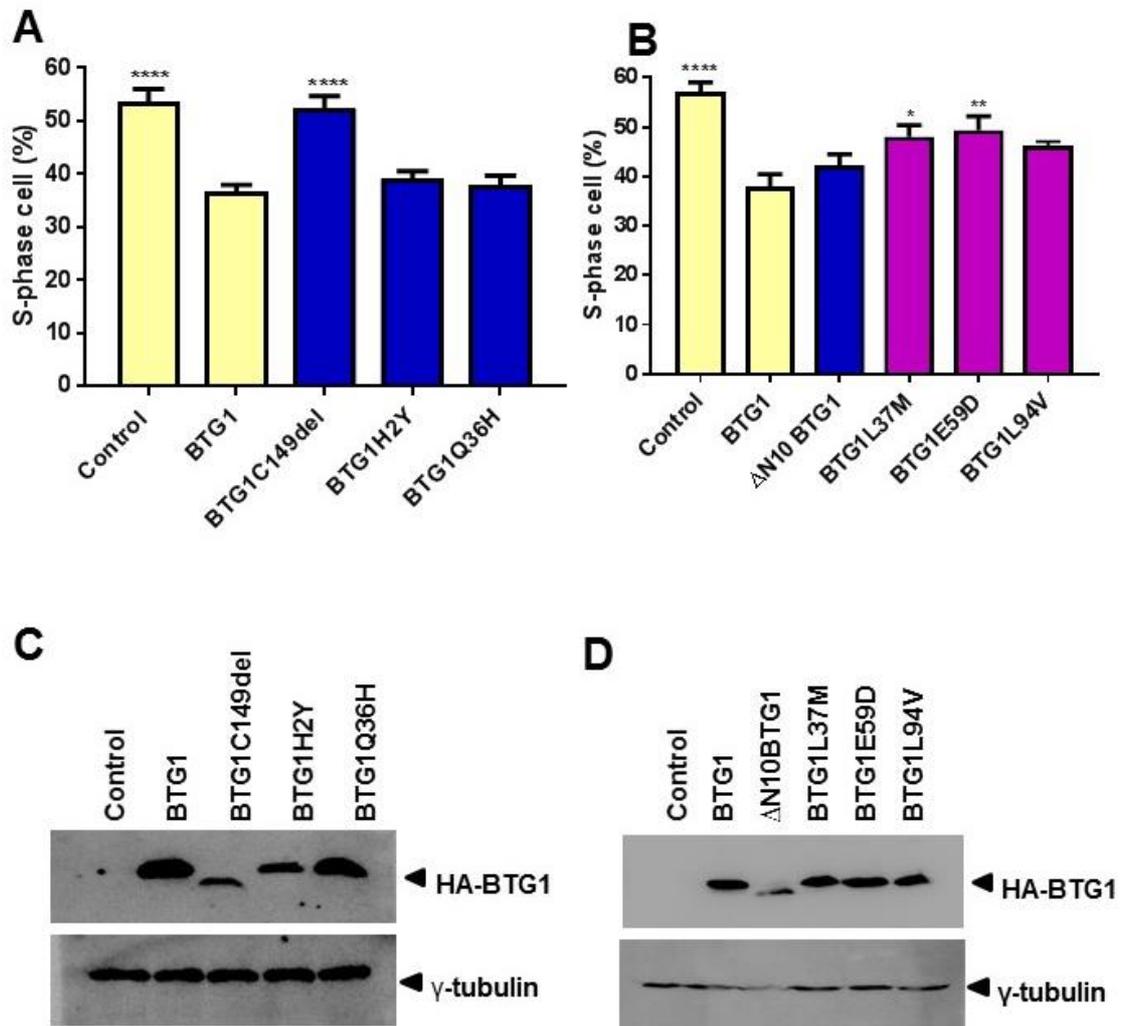


Figure 5-8 Investigation of the impact of BTG1 variants on protein function.

(A) and (B) percentage S phase cells in cells transfected with BTG1 variants located in α -helices and loop regions. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with purple indicating a moderate interaction, blue indicating an ability to interact and yellow for the control and BTG1 wild type. All samples were tested in biological triplicates and the number of nuclei were counted for three images from each slide. The error bars indicate the standard error of the mean ($n=3$). One way Anova (tukey's post-hoc method) was used to calculate P values (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$) compared to cells transfected with pCMV5-HA-BTG1. Three independent experiment were obtained. (C) and (D) Western blots confirming protein expression of HA-BTG1 using an anti-HA antibody. Anti- γ -tubulin was used as a loading control.

5.5 Discussion

BTG1 plays a physiological role in the regulation of cell cycle progression in different types of tissue (Rouault et al., 1992, Prévôt et al., 2000, Nahta et al., 2006, Hata et al., 2007). Overexpression of BTG1 suppresses cell cycle progression during the G1/S phase in a murine lymphoma cell line, decreasing the number of S-phase cells (Hata et al., 2007). BTG1 is a member of the BTG/TOB protein family, members of which share the BTG domain. BTG1 has been found to regulate the cell cycle by binding with transcription factors, via histone modifications or binding with the post-transcriptional factor Caf1 (Lin et al., 1996, Rodier et al., 1999, Prévôt et al., 2000, Prévôt et al., 2001, Winkler, 2010). Recently, it has been found that the anti-proliferative activity of BTG2 and TOB1 requires hCaf1(CNOT7/CNOT8) interaction (Doidge et al., 2012a, Ezzeddine et al., 2012). In the previous chapter, it was shown that ten of the lymphoma BTG1 variants abrogated the interaction with hCaf1(CNOT7/CNOT8), while the six remaining variants were still able to interact with hCaf1(CNOT7/CNOT8) (Table 5-1). Hence, a question remained regarding whether the lack of a hCaf1(CNOT7/CNOT8) interaction in turn, affected the anti-proliferative activity of these BTG1 variants. Moreover, all variants except for E59D were predicted to likely damage protein function (score ≤ 0.05) using SIFT analysis. To the best of our knowledge, none of these BTG1 variants had been experimentally analysed in terms of their proliferation function. Therefore, HEK293T cells were transfected with the BTG1 variants and S-phase cells were labelled with EdU staining, before being quantified to assess cell proliferation.

5.5.1 Effect of BTG1 variants on the anti-proliferative function and disruption of the interaction with human Caf1

Ten of the BTG1 variants displayed impaired interactions with hCaf1(CNOT7/CNOT8); all these variants had been identified in NHL. In this chapter, the variants were divided into three groups based upon where the variant residues are located in the BTG1 structure. The first group comprised

of three variant residues that are located in α -helical regions. Although these residues are not located in the box A or box B domains, they were still necessary for the interaction with hCaf1(CNOT7/CNOT8) (Chapter 4) and therefore expected to be important in the regulation of cell proliferation. As the results show, the anti-proliferation function of BTG1 was lost with some of the variant residues such as R27H and F40C. This indicates that the hCaf1(CNOT7/CNOT8) interaction is likely to be required for the anti-proliferative activity of BTG1, with results in agreement with other studies (Doidge et al., 2012a, Ezzeddine et al., 2012). The last variant in this group, BTG1 F25C did not demonstrate a total loss of anti-proliferative activity unlike BTG1 F40C and BTG1 R27H (Figure 5-5 A). Interestingly, in terms of the interaction analysis, the BTG1 F25C variant was only marginally impaired in relation to its interaction with hCaf1(CNOT7/CNOT8) (Table 5-1). Notably, the weak expression of BTG1 F40C in the western blot analysis (Figure 5-5 B), directed us to another possible functional consequence of the amino acid substitution loss of BTG1 protein stability. However, further experiments are required to establish this possibility.

In the second group, five BTG1 variant residues located in loop regions were analysed. Four of them, namely M11I, P58L, G66V and N73K were impaired in their interaction with hCaf1(CNOT7/CNOT8), however a fifth, E117D was only marginally impacted in terms of its interaction with hCaf1/CNOT8. None of these variants were able to inhibit cell proliferation, suggesting a loss of BTG1 activity in these variant proteins. When comparing the experimental functional assessment with the Suspect prediction analysis, all the variants in group 1 and group 2 were predicted to have impaired protein function and be associated with disease, except for R27H and N73K. However, these two variants were strongly impaired in terms of their interaction with hCaf1(CNOT7/CNOT8) ($p < 0.0001$) (Table 5-1).

The third group comprised of variant residues located in β -sheets. Variant I115V displayed partial loss of interaction with hCaf1 (CNOT7/CNOT8), but

the L104H variant resulted in a total abrogation of this interaction. In terms of the protein function assessment, BTG1 I115V was unable to inhibit cell proliferation. By contrast, the BTG1 L104H variant did not completely reduce the proliferation activity and still maintained the ability to reduce the number of S-phase cells significantly, compared to cells transfected with empty vector (control) ($p < 0.0001$). These results were the opposite of the Suspect analysis, as I115V was predicted as a neutral substitution, but L104H was not. Both residues are located in the box B region of the BTG domain. A previous study demonstrated that deletions of box B of BTG1 impaired protein function in a murine lymphoma cell line. However, it was hypothesised that BTG1 was required to interact with PRMT1 in order to suppress cell proliferation (Hata et al., 2007). Interestingly, the majority of these results are in agreement with the notion that the anti-proliferative activity of BTG/TOB is dependent on the hCaf1(CNOT7/CNOT8) interaction (Doidge et al., 2012a, Ezzeddine et al., 2012). However, this was demonstrated using different protein family members: BTG2, TOB1 and TOB2. In one of these studies, when the MCF-7 breast cancer cell line was transfected with either BTG2 W103A or TOB1 W93A, both variant proteins resulted in a loss of proliferation activity, suggesting the requirement of the hCaf1(CNOT7/CNOT8) interaction (Doidge et al., 2012a).

5.5.2 Investigation of the impact on protein functions of BTG1 mutations not affecting the interaction with human Caf1.

Three BTG1 variant residues were still able to interact with hCaf1(CNOT7/CNOT8) (Δ N10 BTG1, H2Y and Q36H). In term of proliferation, cells expressing BTG1 H2Y and Δ N10 BTG1 were able to inhibit cell proliferation, similar to wild type BTG1, although all these variants were predicted to have defective damaging protein function using SIFT (score ≤ 0.05). Moreover, the Q36H mutation did not affect the activity of BTG1 in terms of proliferative function, although Suspect analysis predicted this variant to be associated with disease (Table 5-1). These results provide

evidence that the anti-proliferative activity of BTG1 requires hCaf1(CNOT7/CNOT8) interaction, which is in agreement with the findings of previous studies (Doidge et al., 2012a, Ezzeddine et al., 2012). BTG1 variants L37M and E59D induced a decrease in the number of S-phase cells, but this was not to the same extent as that seen with wild type protein ($p < 0.05$ and $p < 0.01$, respectively). Interestingly, both variants are marginally impaired in their interaction with hCaf1(CNOT7/CNOT8) ($\geq 50\%$). Notably, the E59D variant was selected as a conservative amino acid substitution, and was predicted to be neutral using SIFT and Suspect. BTG1 L94V is able to inhibit cell proliferation, but to a lesser extent than BTG1 wild type. Statistically, the difference in the percentage of S-phase cells between BTG1 L94V and the control (empty vector) was significant ($p < 0.05$).

During the course of this work, it has been discovered that BTG1 and BTG2 interact with PABPC1 through the BTG domain, specifically box C. When U2OS cells are transfected with a BTG2 mutant at box C, which is able to interact with hCaf1(CNOT7/CNOT8) but not with PABPC1, it is unable to inhibit cell proliferation (Stupfler et al., 2016). Interestingly, the same result is obtained when the cells are transfected with a BTG2 mutant (+71) comprising of an insertion of five amino acids after N71; this mutant is unable to interact with hCaf1(CNOT7/CNOT8) but can still bind to PABPC1. It has therefore been suggested that the anti-proliferative activity of BTG2 requires both hCaf1(CNOT7/CNOT8) and PABPC1 interactions (Stupfler et al., 2016). This might explain some of the results discussed in this chapter. Some variants are able to interact with hCaf1(CNOT7/CNOT8), but there is a reduction in their anti-proliferation activity (L37M, E59D and L94V). Other mutations have a defect in their proliferation function (BTG1 E117D), but could still interact with hCaf1/CNOT7.

Finally, the results of the proliferation assessment showed that BTG1 requires the last 21 amino acids in the C-terminus to be able to control cell proliferation. Contrary to this finding, a previous study found that a BTG1 deletion mutant

(Δ C137 BTG1) was able to reduce the cell growth of WEHI-231 cells stimulated with anti-IgM 48 hours after transfection. However, their hypothesis was different from that of our study, namely that the anti-proliferative activity of BTG1 requires PRMT1 interaction at the binding site of box C in BTG1 (location of residues 120–128) (Hata et al., 2007).

In summary, the majority of the BTG1 variants affected protein function, as measured by cell proliferation and were also predicted to have defective protein function using SIFT. Six out of the seven BTG1 variants that lost their activity were predicted to have defective protein function and to be associated with disease using Suspect analysis. Interestingly, the variant Q36H is predicted to damage protein function, using SIFT and Suspect, however BTG1 did not lose its activity with this variant and was still able to interact with hCaf1(CNOT7/CNOT8). In contrast, the two variants R27H and N73K were predicted to be neutral, yet they lost their anti-proliferation activity and ability to bind with hCaf1(CNOT7/CNOT8). Finally, the E59D variant was also predicted to be neutral in both SIFT and Suspect. This variant showed a marginally reduced anti-proliferative activity and was partially impaired in its interaction with hCaf1(CNOT7/CNOT8). Overall, the BTG1 variants were generally impaired in both their interaction with hCaf1(CNOT7/CNOT8), as well as in their anti-proliferative function. This provides evidence that the anti-proliferative activity of BTG1 requires hCaf1(CNOT7/CNOT8) interaction.

Chapter 6
Regulation of mRNA degradation and translation by BTG1
variants

Chapter 6: Regulation of mRNA degradation and translation by BTG1 variants

6.1 Introduction

The anti-proliferative BTG/TOB protein interacts directly with hCaf1(CNOT7/CNOT8) deadenylase enzyme through BTG domain which allows them to be involved in post-transcriptional gene regulation (Horiuchi et al., 2009, Mauxion et al., 2009, Winkler, 2010). Caf1 is associated with Ccr4 to form a complex that plays a key role in shortening the poly (A) tail of the cytoplasmic mRNA in the eukaryotic cells. Both Caf1 and Ccr4 are catalytic subunits of the Ccr4-Not complex, meaning that they exhibit ribonuclease activity that displays selectivity for poly (A) residues (Collart, 2003, Parker and Song, 2004, Collart, 2016). It has recently been demonstrated that human Caf1(CNOT7/CNOT8) and human Ccr4 (CNOT6/CNOT6L) cooperate in mRNA deadenylation activity *in vitro*, while the presence of the BTG2 interaction increases the enzymatic activity (Maryati et al., 2015).

The deadenylation process is necessary for the regulation of gene expression in eukaryotic cells through the effects on mRNA degradation and the rate of mRNA translation (Parker and Song, 2004, Goldstrohm and Wickens, 2008, Wahle and Winkler, 2013, Collart, 2016). The BTG/TOB proteins share the N-terminal domain that allows interaction with Caf1 and, based on the structure of the hCaf1/CNOT7-TOB1 complex, TOB1 interacts away from the surface of the RNA nuclease activity (DEDD) site (Horiuchi et al., 2009, Winkler, 2010). Additionally, it has been demonstrated that BTG2 and TOB1 impact on deadenylation and mRNA decay via binding with hCaf1(CNOT7/CNOT8) (Mauxion et al., 2008, Horiuchi et al., 2009, Doidge et al., 2012a, Ezzeddine et al., 2012). TOB1 and TOB2 contain the PAM2 motif in the C-terminus, which is absent from the BTG proteins. The PAM2 motif allows interaction with PABPC1 (an RNA-binding protein) and, consequently, it recruits hCaf1(CNOT7/CNOT8) to the target mRNA (Ezzeddine et al., 2007, Horiuchi et

al., 2009, Ezzeddine et al., 2012). Further, BTG2 expression serves as a general activator for mRNA deadenylation through binding with hCaf1(CNOT7/CNOT8) (Mauxion et al., 2008). When using mutant BTG2 (five amino acids inserted at position 66), the lack of interaction with hCaf1(CNOT7/CNOT8) meant that deadenylation could not occur, although the mechanism behind this remained unclear (Mauxion et al., 2008). It has recently been found that BTG1 and BTG2 bind with PABPC1 via the BTG domain, specifically in the box C region, with the BTG2-PABPC1 interaction stimulating deadenylase activity *in vitro*, which suggests that BTG2 bridges the PABPC1 RNA-binding protein and hCaf1(CNOT7/CNOT8) (Stupfler et al., 2016). Previous studies have concluded that TOB1, TOB2 and BTG2 regulate cell proliferation by recruiting hCaf1(CNOT7/CNOT8) to the target mRNA and thereby causing an alteration in the target mRNA/protein expression (Doidge et al., 2012a, Ezzeddine et al., 2012). Such studies have demonstrated that the use of the mutant BTG/TOB proteins BTG2 W103A or TOB1 (L32G, L35G, L36G and W93A) cannot regulate mRNA decay and translation due to their inability to bind with hCaf1(CNOT7/CNOT8) (Doidge et al., 2012a, Ezzeddine et al., 2012).

Chapter Five discussed how the present study suggests that the anti-proliferative activity of BTG1 is required for hCaf1(CNOT7/CNOT8) interaction. Nine BTG1 variants were found to be impaired in terms of the interaction with hCaf1(CNOT7/CNOT8) and hence defective with regards to their proliferation activity (F25C, R27H, F40C, L104H, I115V, M11I, P58L, G66V and N73K). Additionally, three BTG1 mutations were found to be able to bind with Caf1 and thereby inhibit cell proliferation, namely Δ 10N BTG1, H2Y and Q36H (Table 6-1). Therefore, two unresolved questions remain. First, does BTG1 impact on mRNA translation and degradation via hCaf1(CNOT7/CNOT8) binding? Second, does BTG1 regulate cell proliferation through the recruitment of the Ccr4-Not complex to the target mRNA and thus regulate mRNA translation and abundance? To answer these questions, BTG1 and the

BTG1 variants were fused to the λ N peptide that binds to the specific RNA hairpin loop containing the five box B sequences present in the 3' UTR of renilla luciferase mRNA. Then, the λ N-BTG1 is able to recruit accessory factors, such as the hCaf1(CNOT7/CNOT8) deadenylase, onto the reporter mRNA and induce mRNA decay and/or repress translation. The HEK293T cells were co-transfected with the reporter plasmid pRL-5boxB and plasmids expressing either the λ N peptide alone (as a control) or the λ N peptide fused with HA-BTG1 or the variants. After 24 hours of transfection, the proteins were extracted in order to measure the luciferase activity and to perform western blot analysis. In addition, the RNA was isolated and obtained for luciferase qPCR analysis using a probe-based system with the values normalised using GAPDH to measure mRNA levels (Figure 6-1).

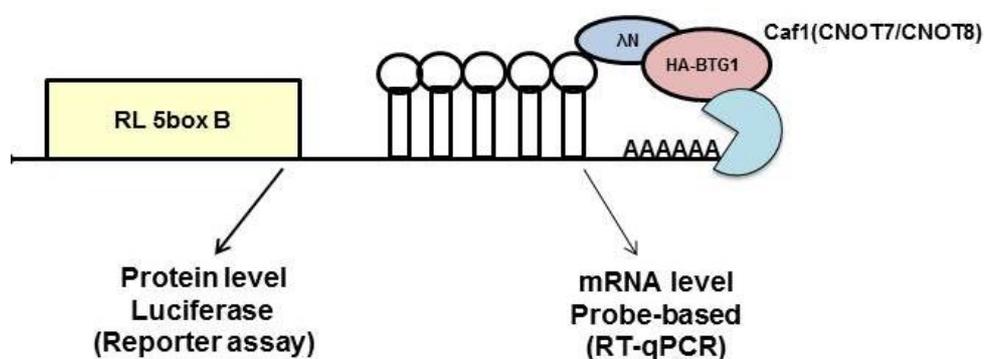


Figure 6-1 Schematic model of the recruitment of BTG1 to the reporter mRNA for the tethering assay.

A schematic representation of the reporter mRNA containing the coding sequences for the renilla luciferase enzyme and five box B sequences in the 3' UTR. The recruitment of wildtype or mutant BTG1 to the 3' UTR was mediated by the fused λ N peptide, which has a high affinity for the box B sequences. HEK293T cells were transfected with plasmids pCI- λ N-HA-BTG1 or empty vector as control and the luciferase reporter plasmid pRL-5boxB. 24 hours after of transfection, protein was extracted and luciferase activity was assessed and normalized to total protein. RNA was also isolated and subjected to RT-qPCR analysis for the determination of luciferase mRNA levels using GAPDH mRNA levels as a reference.

Variant residue	SIFT score ≤ 0.05	Suspect score > 50	protein-protein interaction		Proliferation function $p < 0.05$	
			hCaf1/CNOT7	hCaf1/CNOT8	BTG1	Control
F25C	0.001	68	+	++	**	*
R27H	0.039	20	-	-	****	NS
F40C	0	71	-	-	****	NS
L104H	0	89	-	-	**	****
I115V	0.003	29	+	+	****	NS
M11I	0.028	87	-	-	****	NS
P58L	0	53	+	-	****	NS
G66V	0.001	70	-	-	****	NS
N73K	0.001	24	-	-	****	NS
E117D	0.022	14	++++	++	****	NS
Q36H	0.006	53	++++	+++	NS	****
L37M	0.02	31	+++	++	*	NS
L94V	0.002	28	+++	+	NS	*
$\Delta 10N$ BTG1	0	4	++++	+++	NS	***
H2Y	0.037	17	++++	+++	NS	****
E59D	0.054	2	++	++	**	NS

Table 6-1 Data analysis in BTG1 variants using bioinformatics and yeast two-hybrid interaction and proliferation analysis.

Summary of the data analysis obtained in Chapter 3, 4 and 5. Variant residues are coloured based on the location in the BTG1 structure, green (α -helical), blue (β -sheet) and orange (L-loops). The bioinformatics software (SIFT and Suspect) were used to predict the effect of the variant on protein function and yeast two hybrid analysis was used to evaluate the effect of the BTG1 variant on the interaction with hCaf1(CNOT7/CNOT8). The relative values of β -Gal activity to BTG1 are indicated in different ranges such as 0-20 (-), 21-40(+), 41-60(++), 61-80(+++) and 81-100 (++++). A one-way ANOVA was used to calculate the significance of the values (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$) compared to the cells transfected with pCMV5-HA-BTG1 and control (empty vector), NS=not significant.

6.2 Generation of plasmids for the RNA tethering assay

In order to explore the effect of BTG1 variants on mRNA stability, the BTG1 variants were recruited to the target luciferase mRNA and their functions were determined using an RNA tethering assay. Therefore, the necessary mammalian expression plasmids required to produce the λ N peptide fused to BTG1 variants was prepared. The function of the λ N peptide is to tether the desired protein to the corresponding RNA hairpin loop present in the luciferase reporter mRNA. The reporter plasmid pRL-5boxB, expressing the renilla luciferase reporter mRNA was described before (Pillai et al., 2004).

The plasmid pBD-HA-BTG1 was used as a template to obtain the BTG1 cDNA sequence and introduce the sequence of the left arm of the pCI- λ N vector, including the XhoI restriction site, followed by the HA epitope of the AUG start codon at the 5' end. Further, at the 3' end, the stop codon of the BTG1 was introduced, followed by the right arm sequence of the pCI- λ N vector, including the XhoI restriction site. Then, the DNA fragments were ligated into the digested pCI- λ N modified vector at the XhoI site using a HiFi DNA Assembly Cloning Kit (NEBuilder®) (Figure 6-2). Plasmids containing the cDNA BTG1 variants were generated in the same way as the wild type BTG1, while the deletion mutation BTG1 C149del was generated by designing an alternative 3' oligonucleotide by means of deleting the last 21 amino acids of the C-terminal in the codon region of BTG1. The HA epitope was introduced to allow for the detection of protein expression using western blot analysis. All BTG1 sequences were confirmed by DNA sequencing.

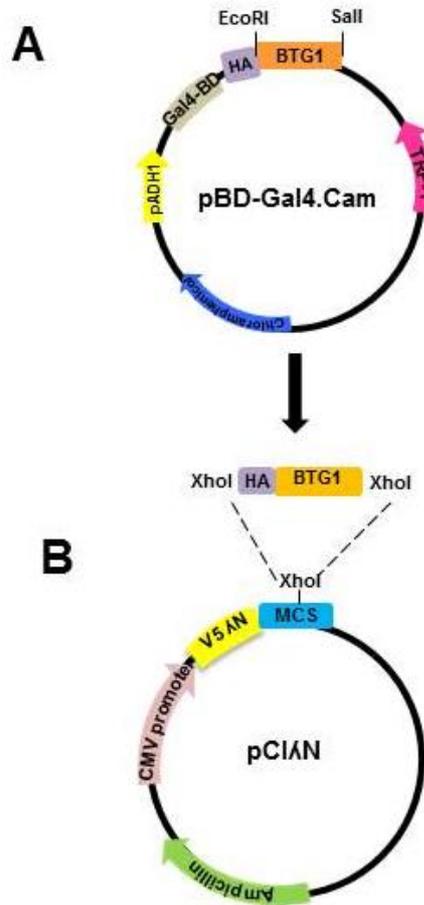


Figure 6-2 Generation of vectors for the tethering assay.

(A) The HA-BTG1 cDNA was amplified using the pBD-HA-BTG1 plasmid as a template. The 5' primer was designed to include the 5' DNA sequence of the vector and the XhoI restriction site at the 5' end, while the 3' primer was designed to include the stop codon followed by the 3' DNA sequence of the vector and the XhoI restriction site at the 3' end. (B) The PCR products were ligated into the digested, modified mammalian expression vector pCI-λN in the XhoI restriction site using the HiFi DNA Assembly method.

6.3 Effect of tethering the BTG1 variants to the 3' UTR of the luciferase reporter mRNA on mRNA translation and degradation

In order to determine whether BTG1 regulates mRNA translation and degradation via binding with hCaf1(CNOT7/CNOT8). Previously, it had been concluded that nine BTG1 variants exhibited impaired interaction with hCaf1(CNOT7/CNOT8). Additionally, one BTG1 variant, namely E117D, was impacted in terms of the interaction with hCaf1/CNOT8 (Table 6-1). Therefore, an RNA tethering assay was conducted to investigate if the presence of BTG1 variants at the 3' UTR of the reporter mRNA impacts on mRNA degradation and translation. Therefore, HEK293T cells were co-transfected with the reporter plasmid pRL-5boxB and plasmids expressing either the λ N peptide alone (as a control) or the λ N peptide fused with HA-BTG1 or the variants.

6.3.1 BTG1 variants mapping onto the α -helical structure

Three BTG1 variants were located in the α -helical structure, which exhibited an impaired interaction with hCaf1(CNOT7/CNOT8) (Table 6-1). The results showed that the cells expressing λ N-HA-BTG1 reduced the luciferase activity. In contrast, in cells expressing the λ N peptide alone (control), the level of luciferase activity did not alter, which confirmed that BTG1 is involved in the regulation of mRNA translation. In terms of the BTG1 variants, the tethering of λ N-HA-BTG1 F25C and λ N-HA-BTG1 R27H to the reporter mRNA caused a reduction in luciferase activity to the same extent as the wild type BTG1. However, this result was not observed in the case of the BTG1 F40C, since the level of luciferase activity was only reduced by 50%, which indicated that this mutation affected the BTG1 regulation of mRNA translation ($p < 0.0001$) (Figure 6-3 A).

To determine whether the reduced luciferase reporter activity was due to repression of translation or the activation of mRNA degradation, luciferase mRNA levels were determined using RT-qPCR. In terms of the mRNA levels,

the amount of luciferase mRNA did not alter when the cells expressed the λ N peptide alone. In contrast, the luciferase mRNA reduced by 70% when the cell expressed λ N-HA-BTG1. As for the BTG1 variants, the tethering of F25C and R27H decreased the amount of luciferase mRNA to the same extent as BTG1. However, the tethering of BTG1 F40C proved incapable of inducing mRNA degradation, compared with wild type BTG1 ($p < 0.001$) (Figure 6-3 B). To confirm equivalent BTG1 protein expression, the cells were harvested after 48 hours of transfection, the secreted proteins were subjected to western blot analysis using an anti-HA antibody (Figure 6-3 C).

6.3.2 BTG1 variants mapped onto β -sheet structures

Two BTG1 variants located in the β -sheet structures exhibited an impaired interaction with hCaf1(CNOT7/CNOT8). Cells expressing λ N-HA-BTG1 showed significantly reduced luciferase activity, which was not the case in the cells expressing only the λ N peptide. In terms of the BTG1 variants, tethering BTG1 L104H to the reporter mRNA decreased the luciferase activity by less than 50% when compared to the wild type BTG1 ($p < 0.0001$). In contrast, BTG1 I115V was able to reduce luciferase activity to the same extent as the wild type BTG1 (Figure 6-4 A). Similar results were found in relation to measuring the mRNA levels, since the tethering of BTG1 and BTG1 I115V caused a significant reduction in the amount of luciferase mRNA when compared with the tethering of the λ N peptide alone. In contrast, tethering BTG1 L104H to the reporter mRNA did not significantly reduce the amount of luciferase mRNA, which indicated that it is incapable of inducing mRNA degradation ($p < 0.01$) (Figure 6-4 B). Interestingly, this mutation totally abrogated the interaction with hCaf1(CNOT7/CNOT8) (Table 6-1).

The protein expression of BTG1 variants were confirmed using western blot analysis. All the BTG1 variants were equally expressed (Figure 6-4 C).

6.3.3 BTG1 variants mapped in the loop regions

Six BTG1 variants were located in the loop regions and, except for E59D and E117D, exhibited an impaired interaction with hCaf1(CNOT7/CNOT8). Notably, variant E117D was marginally impaired in terms of the interaction with hCaf1/CNOT8 (Table 6-1). When using the BTG1 variants, the tethering of λ N-HA-BTG1 variants M11I, P58L, N73, E59D and E117D to the luciferase mRNA caused a reduction in the luciferase activity to the same extent as BTG1. In contrast, when tethering λ N-HA-BTG1 G66V to the reporter mRNA, the luciferase activity was reduced by only 20% ($p < 0.0001$) (Figure 6-5 A and B). The results showed the same pattern relating to the regulation of mRNA levels. Tethering λ N-HA-BTG1 variants M11I, P58L, N73K, E117D and E59D to the luciferase promoted degradation of the reporter mRNA to the same extent as the wild type BTG1. In contrast, tethering BTG1 G66V to the reporter mRNA resulted in the same amount of luciferase mRNA as tethering the λ N peptide alone (control), which indicated that this mutation was unable to induce mRNA degradation ($p < 0.0001$) (Figure 6-5 C and D).

The protein expression of BTG1 variants were confirmed using western blot analysis. All the protein variants were found to have a similar level of BTG1 expression (Figure 6-5 E and F).

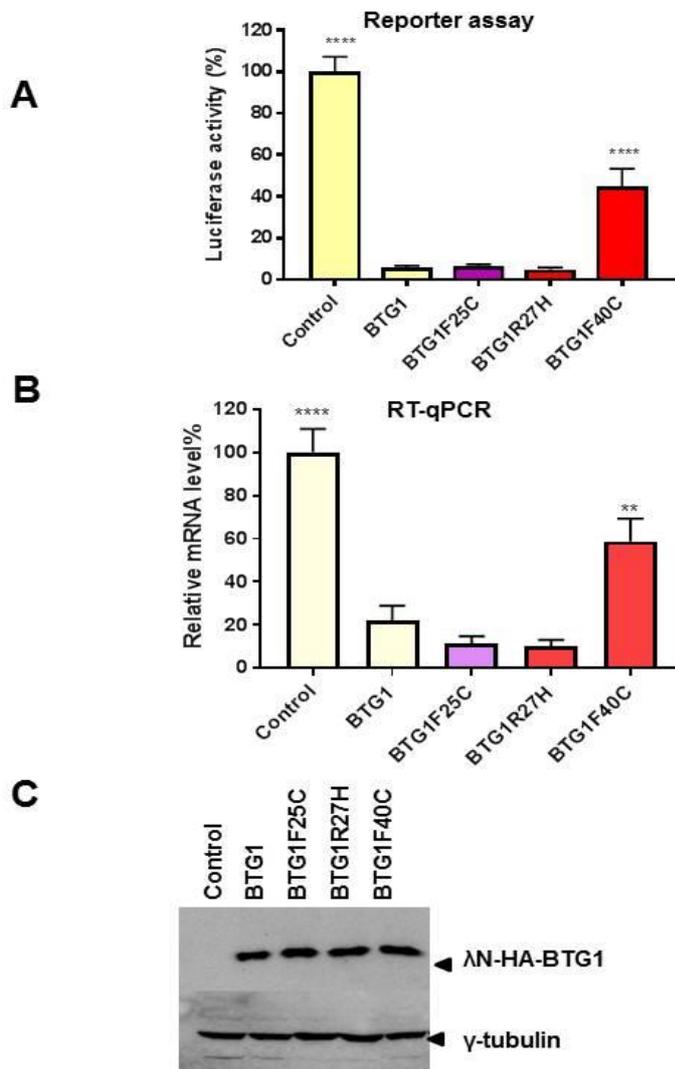


Figure 6-3 BTG1 variants located in the α -helical regions and were able to reduce the mRNA abundance of a tethered mRNA.

BTG1 variants that mapped onto the α -helices and unable to interact with hCaf1(CNOT7/CNOT8) were tethered to 3' UTR of luciferase mRNA. (A) Measurements of the level of luciferase activity affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. (B) Measuring the amount of the mRNA luciferase affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with red indicating unable to interact, purple indicating moderate and yellow for the control and BTG1. The error bars indicate the standard error of the mean ($n=3$). A one-way ANOVA (tukey's post-hoc method) was used to calculate the significance of the value ** $p < 0.01$ **** $p < 0.0001$ compared to the cells expressed λ N-HA-BTG1. Three independent experiment were obtained. (C) Western blots confirming the protein expression of anti-HA to detect λ N -BTG1 and anti- γ -tubulins as loading control.

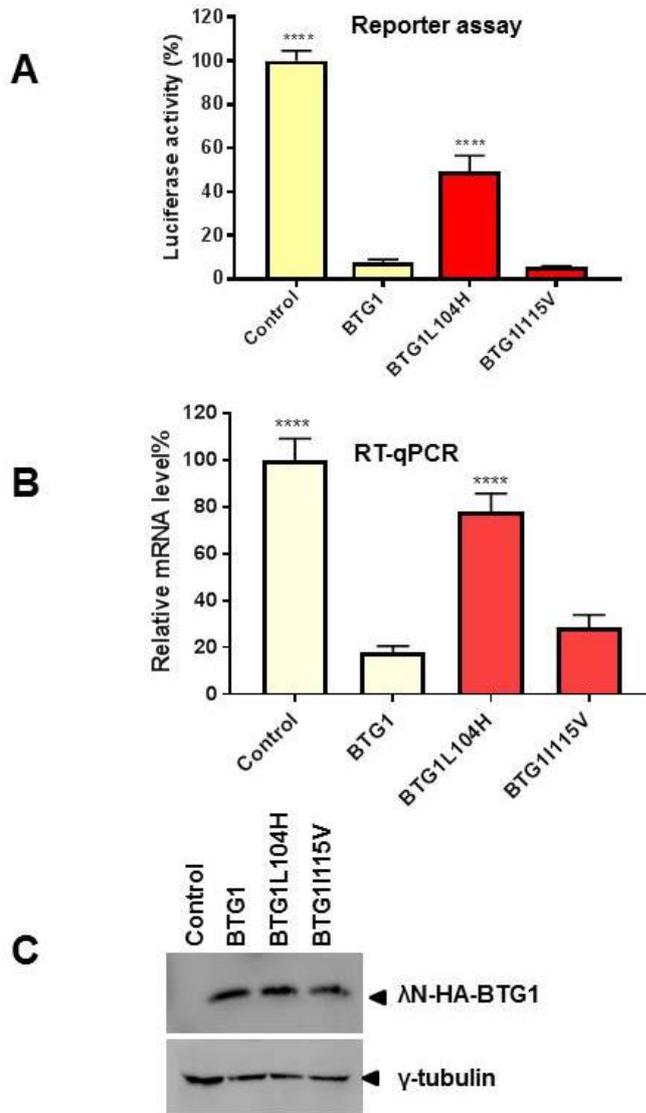


Figure 6-4 BTG1 variants located in the β -sheets and able to reduce the mRNA abundance of a tethered mRNA.

BTG1 variants that mapped onto the β -sheets and unable to interact with hCaf1(CNOT7/CNOT8) were tethered to 3' UTR of luciferase mRNA. (A) Measurements of the level of luciferase activity affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. (B) Measuring the amount of the mRNA luciferase affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with red indicating unable to interact and yellow for the control and BTG1. The error bars indicate the standard error of the mean (n=3). A one-way ANOVA (tukey's post-hoc method) was used to calculate the significance of the value ****p < 0.0001 compared to the cells expressed λ N-HA-BTG1. Three independent experiment were obtained. (C) Western blots confirming the protein expression of anti-HA to detect λ N -BTG1 and anti- γ -tubulins as loading control.

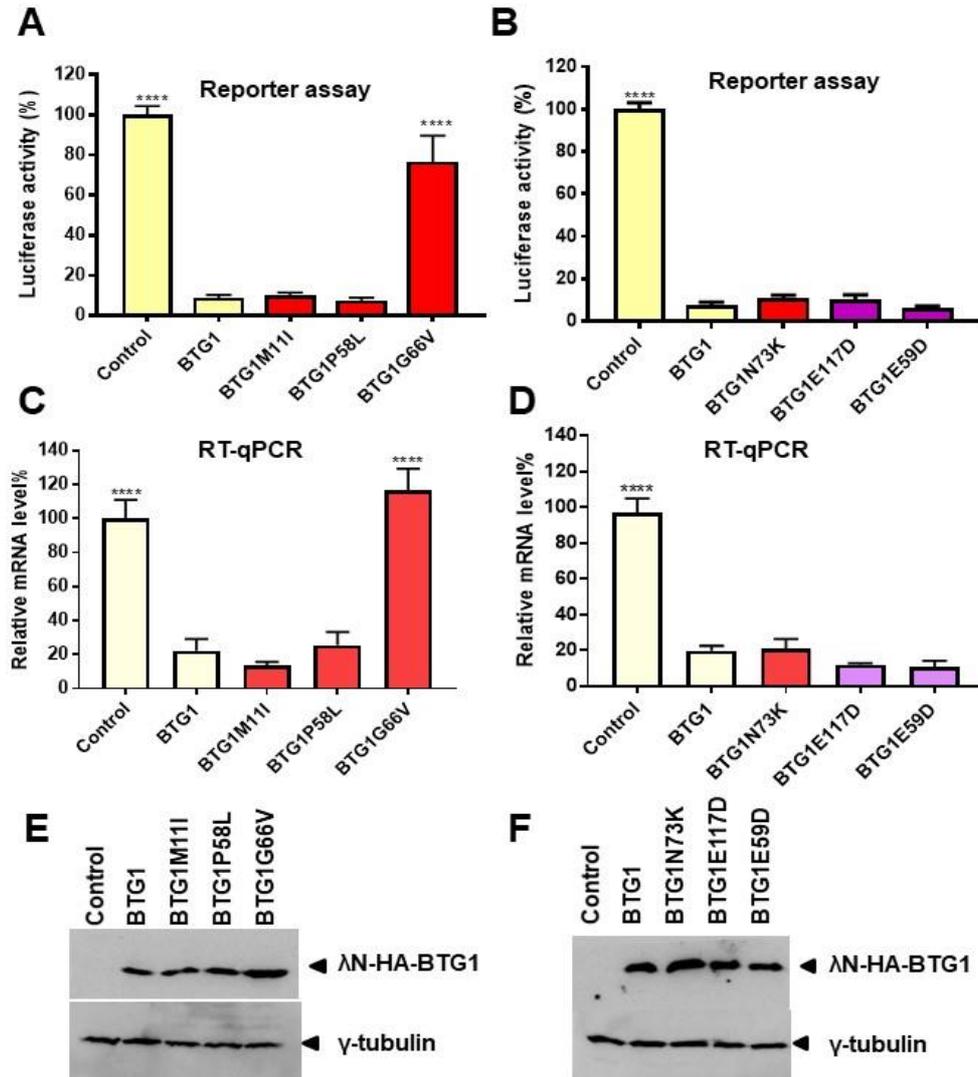


Figure 6-5 BTG1 variants located in the loop regions and able to reduce the mRNA abundance of a tethered mRNA.

BTG1 variants that mapped onto the loop regions and unable to interact with hCaf1(CNOT7/CNOT8) were tethered to 3' UTR of luciferase mRNA. (A) and (B) Measurements of the level of luciferase activity affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. (C) and (D) Measuring the amount of the mRNA luciferase affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, with red indicating unable to interact, purple indicating moderate in the interaction, and yellow for the control and BTG1. The error bars indicate the standard error of the mean (n=3). A one-way ANOVA (tukey's post-hoc method) was used to calculate the significance of the value ****p < 0.0001 compared to the cells expressed λN-HA-BTG1. Three independent experiment were obtained. (E) and (F) Western blots confirming the protein expression of anti-HA to detect λN -BTG1 and anti-γ-tubulins as loading control.

6.4 BTG1 variants able to regulate mRNA decay via binding with human Caf1 (CNOT7/CNOT8)

The remaining variant residues were either able to interact with hCaf1(CNOT7/CNOT8), for example, Δ 10N BTG1, H2Y and Q36H, or the interaction was only marginally impaired (>50%), for example, L37M and L94V (Table 6-1). This left the unresolved question of whether or not BTG1 variant regulation of mRNA translation and degradation is dependent on the interaction with hCaf1(CNOT7/CNOT8). Additionally, the BTG1 C149del residues showed a high affinity in terms of their binding, although this mutation disrupted the proliferation activity. The results showed that the tethering of λ N-HA-BTG1 variants such as H2Y, Q36H and BTG1 C149del to the reporter mRNA decreased the level of luciferase activity to the same extent as the wild type BTG1. Further, similar results were found for the remaining variant residues (Δ 10N BTG1, L37M and L94V), which indicated that these mutations did not influence the ability of BTG1 to repress mRNA translation (Figure 6-6 A and B). The results concerning the measurement of the mRNA luciferase levels were as expected. The tethering of λ N BTG1 or these variants to the reporter mRNA decreased the amount of luciferase mRNA by more than 70% when compared with the control (Figure 6-6 C and D).

The protein expression of BTG1 variants were confirmed using western blot analysis. All the protein variants were found to have a similar level of BTG1 expression (Figure 6-6 E and F).

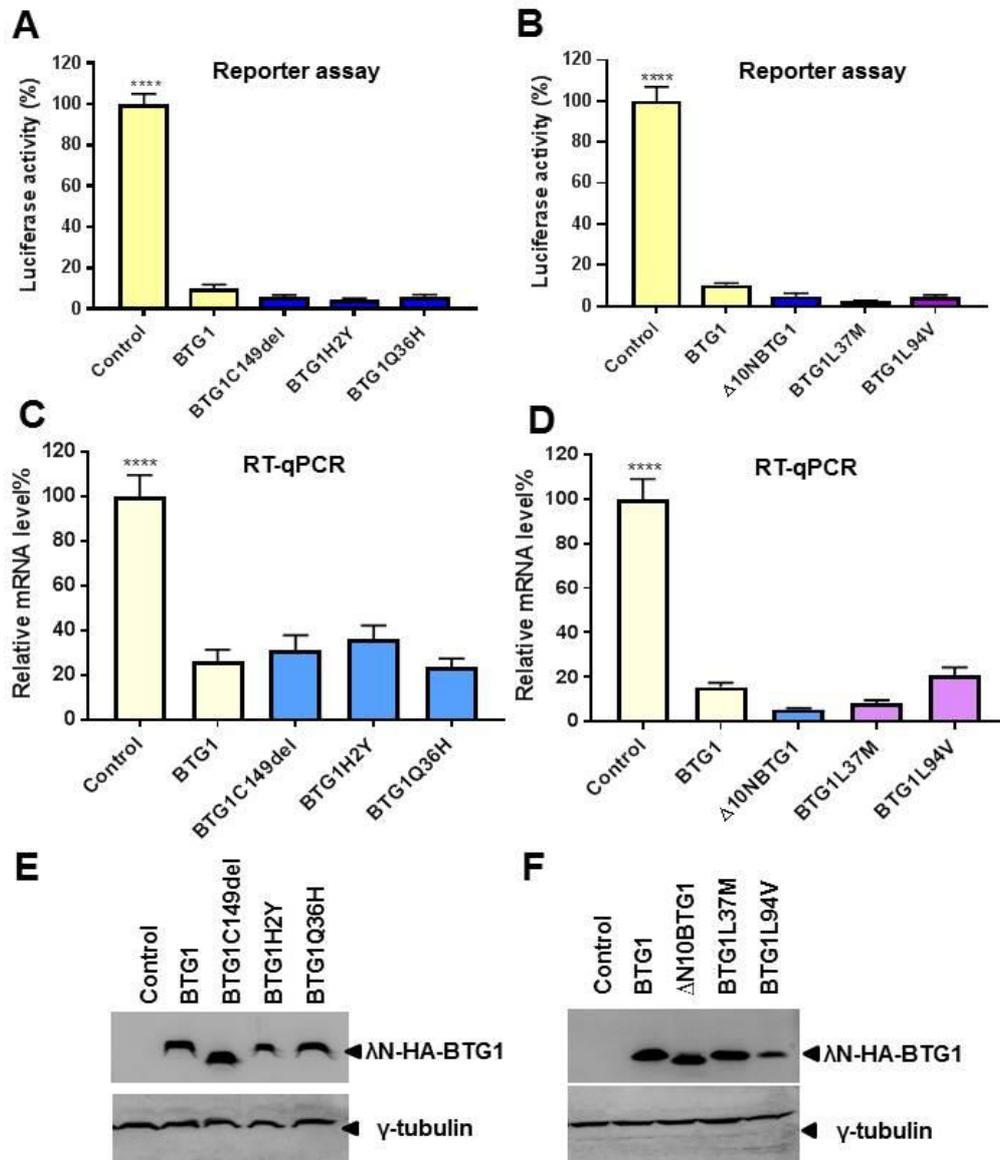


Figure 6-6 BTG1 variants able to reduce the mRNA abundance of a tethered mRNA.

Tethering BTG1 variants mapped on the loop and α -helical regions and able to interact with hCaf1(CNOT7/CNOT8) to 3' UTR of luciferase mRNA. (A) and (B) Measurements of the level of luciferase activity affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. (C) and (D) Measuring the amount of the mRNA luciferase affected by tethering BTG1 variants to the luciferase mRNA, and compared with BTG1. The graphs are coloured based on the results of the hCaf1(CNOT7/CNOT8) interaction, purple indicating moderate in the interaction, blue able to interact and yellow for the control and BTG1. The error bars indicate the standard error of the mean (n=3). A one-way ANOVA (tukey's post-hoc method) was used to calculate the significance of the value ****p < 0.0001 compared to the cells expressed λ N-HA-BTG1. Three independent experiment were obtained. (E) and (F) Western blots confirming the protein expression of anti-HA to detect λ N -BTG1 and anti- γ -tubulins as loading control.

6.5 Discussion

The anti-proliferative BTG/TOB proteins are involved in post-transcriptional regulation through the interaction with the hCaf1(CNOT7/CNOT8) deadenylase subunit of the Cc4-Not complex (Horiuchi et al., 2009, Mauxion et al., 2009, Winkler, 2010). The Caf1-Ccr4 complex is involved in the regulation of gene expression through removing the poly (A) tail of the targeted mRNA (Collart, 2003, Parker and Song, 2004, Maryati et al., 2015). It has previously been shown that the BTG2 and TOB proteins are involved in the regulation of mRNA deadenylation and, further, that the absence of the BTG/TOB proteins binding with hCaf1(CNOT7/CNOT8) effects the rate of the deadenylation process. This suggests that the BTG/TOB proteins inhibit cell proliferation through recruiting hCaf1(CNOT7/CNOT8) to the target mRNA and thereby causing mRNA degradation and/or translation (Doidge et al., 2012a, Ezzeddine et al., 2012). Previously, nine of the BTG1 variants had been found to be affected in terms of their interaction with hCaf1(CNOT7/CNOT8) and they were also shown to exhibit impaired proliferation activity. The remaining variants were able to either interact with hCaf1(CNOT7/CNOT8) or their interaction was only marginally impaired (Table 6-1). Therefore, two key questions needed to be addressed. Is BTG1 involved in mRNA decay via the interaction with hCaf1(CNOT7/CNOT8)? Further, does BTG1 regulate cell proliferation through mRNA degradation? Therefore, the HEK293T cells were transfected so they expressed the luciferase reporter and the BTG1 variants were tethered to the 3' UTR of the reporter mRNA.

Only a limited number of studies have demonstrated the regulation of BTG1 in relation to mRNA decay. In the present study, the observation concerning the tethering of BTG1 is consistent with the published data, which establishes that the tethering of BTG2 and the TOB proteins to the 3' UTR of the reporter mRNA caused a significant reduction in reporter activity, as well as mRNA degradation via binding with hCaf1(CNOT7/CNOT8) (Ezzeddine et al., 2007, Funakoshi et al., 2007, Mauxion et al., 2008, Doidge et al., 2012a, Ezzeddine et

al., 2012). Also, the reduction of the luciferase activity and suppression of the mRNA translation being caused by the tethering of BTG2 to the 3'UTR was confirmed using the reporter plasmid pRL-TK that does not contain the BoxB motifs. When comparing the expression of λ N-HA-BTG2 or λ N-HA-BTG2 W103A to cells transfected with empty vectors there was no significant difference between the samples expressing the various λ N fusion proteins, indicating that tethering was indeed responsible for reduced luciferase activity when a reporter containing BoxB sequences was used (Doidge et al., 2012a).

This finding is consistent with the current hypothesis that tethering the BTG1 variants to the target mRNA on the 3' UTR is able to reduce luciferase activity and induce mRNA decay via binding with hCaf1(CNOT7/CNOT8). The relevant variants to which this applied to were Δ 10N BTG1, H2Y, Q36H, L37M, E59D, L94V and BTG1 C149del (Figure 6-5 and Figure 6-6). Surprisingly, the results differ from the study hypothesis in that the BTG1 variants which were unable to bind with hCaf1(CNOT7/CNOT8) were still able to induce mRNA degradation (M11I, F25C, R27H, P58L, N73K, I115V and E117D). Only three variants were entirely unable to interact with hCaf1(CNOT7/CNOT8) and hence lost their ability to regulate mRNA decay, namely F40C, L104H and G66V (Figure 6-3, Figure 6-4 and Figure 6-5). Most previous studies have suggested that the mechanism of the TOB proteins differs from that of the BTG2 protein due to the absence of the PAM2 motif, which allows for the interaction with PABPC1 (an RNA-binding protein) and recruits hCaf1(CNOT7/CNOT8) to the target mRNA for both deadenylation and decay (Ezzeddine et al., 2007, Funakoshi et al., 2007, Mauxion et al., 2008, Doidge et al., 2012a, Ezzeddine et al., 2012). While the present study was being conducted, a paper was published suggesting that BTG1 and BTG2 interact with PABPC1 through the BTG domain, specifically in box C, which indicates that BTG2 serves as a bridge between the PABPC1 and hCaf1(CNOT7/CNOT8) (Stupfler et al., 2016). It was shown that the mutant BTG2 in the box C region

of the BTG domain (substituted from DGSIC to KGPVK as in TOB1) lacked the ability to interact with PABPC1 but not with hCaf1(CNOT7/CNOT8). *In vitro*, they demonstrated that the presence of the BTG2-PABPC1 interaction strongly stimulates the deadenylation activity of hCaf1/CNOT7, with the opposite result being reported in the absence of this interaction using the mutant box C BTG2 (Stupfler et al., 2016). A further study showed that, in relation to TOB proteins and PABPC1, hCaf1(CNOT7/CNOT8) reduces the deadenylation activity in the absence of the TOB-PABPC1 interaction, although the TOB proteins interact with a different domain of PABPC1, namely MLL (Ezzeddine et al., 2007). Additionally, the results in relation to the knockdown of the hCaf1/CNOT7 or hCaf1/CNOT8, indicate that tethering of BTG2 to the 3' UTR of mRNA is still able to reduce luciferase activity even in the absence of hCaf1(CNOT7/CNOT8) activity (Doidge et al., 2012a). One of the possibilities is that, BTG2 does not only require the hCaf1(CNOT7/CNOT8) interaction, but that the presence of PABPC1 binding is also necessary to recruit for mRNA decay (Stupfler et al., 2016).

In summary, it was hypothesised that BTG1 regulates mRNA decay via binding with hCaf1(CNOT7/CNOT8). However, based on the results of this study, BTG1 likely requires the presence of another factor, such as PABPC1, to regulate mRNA degradation (Figure 6-7). Further experiments are required to determine the effect of the BTG1 variants on the interaction with PABPC1 (Chapter 7).

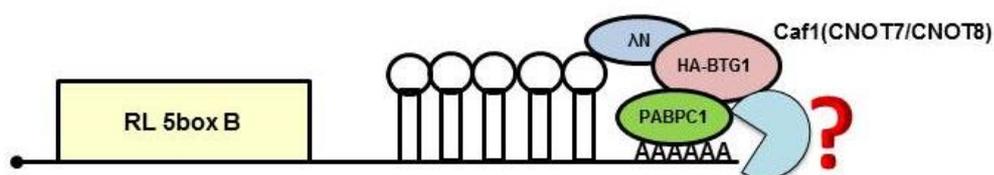


Figure 6-7 Schematic model of the recruitment of BTG1 to the reporter mRNA for the tethering assay required for PABPC1 interaction.

Chapter 7

An investigation of the interaction between BTG1 and PABPC1

Chapter 7: An investigation of the interaction between BTG1 and PABPC1

7.1 Introduction

The cytoplasmic poly(A) binding protein 1 (PABPC1) is an essential eukaryotic translational initiation factor, and its covered the long poly(A) tails at the 3' of Eukaryotic mRNA. The interaction between PABPC1 and the eukaryotic initiation factor (eIF4F), a cap-binding complex, is important for the initiation of translation (Mauxion et al., 2009). The poly(A) tails help to protect the mRNA from degradation when the poly(A) is shortened by deadenylase enzymes such as the Pan2-Pan3 complex and the Caf1-Ccr4 deadenylase complex (Meyer et al., 2004, Collart, 2016). Interaction with Pan2-Pan3 complex is required by PABPC1 in order to facilitate the deadenylation activity; Pan3 contains the PAM2 motif, which allows the interaction with PABPC1, and through this interaction the Pan2 deadenylation process is activated (Uchida et al., 2002, Mangus et al., 2004, Wolf and Passmore, 2014). The TOB1/TOB2 proteins also contain the PAM2 motif, which binds to the C-terminal domain (termed PABC or MLLE) of PABPC1. Also, TOB1 and TOB2 recruit the Ccr4-Not complex via hCaf1(CNOT7/CNOT8) interaction. Suggesting that the TOB proteins serve as an adaptor between hCaf1(CNOT7/CNOT8) and PABPC1. Subsequently, the PABPC1 binding stimulates the hCaf1(CNOT7/CNOT8) deadenylase activity (Ezzeddine et al., 2007, Funakoshi et al., 2007). Previous studies have found that the conserved residues in TOB1 (F139 and F247) and TOB2 (F140 and F260) represent a necessary part of the interaction with PABPC1. Consequently, this interaction enhances the deadenylation process (Ezzeddine et al., 2007). In addition, it has been found that the BTG2 protein plays a general activator role in mRNA deadenylation (Mauxion et al., 2008). While the present study was conducted, new findings were published suggesting BTG1 and BTG2 interact with PABPC1 though the box C region in the BTG domain and the RNA Recognition Motif 1

(RRM1) of PABPC1 (Stupfler et al., 2016). This interaction was confirmed using three different systems, namely a yeast two-hybrid analysis, co-immunoprecipitation in a mammalian cell line, and a pull-down assay using bacterial expression hosts (Stupfler et al., 2016). Additionally, Stupfler et al. (2016) found that the presence of the first two RRM domains of PABPC1 and the BTG domain of BTG2 is sufficient to stimulate hCaf1(CNOT7/CNOT8) deadenylase activity *in vitro* in the absence of other subunits of the CCR4-NOT complex. The absence of the BTG2-PABPC1 interaction caused a reduction in deadenylase activity, as demonstrated using BTG2 containing a variant box C sequence. Interestingly, this BTG2 variant was able to interact with hCaf1(CNOT7/CNOT8) but not with PABPC1.

The majority of the lymphoma-associated BTG1 variants with single amino acid substitutions (except for F40C, L104H and G66V) were able to regulate mRNA decay independently of the hCaf1(CNOT7/CNOT8) interaction (Chapter 6). This indicates the involvement of another factor, which may correspond to PABPC1. Therefore, this chapter aimed to determine the most appropriate assay for monitoring the interaction between BTG1 and PABPC1 by testing via yeast two-hybrid analysis and pull-down assays.

7.2 Generation of plasmids for the yeast two-hybrid analysis

In order to investigate the interaction between BTG1 and PABPC1, plasmids were prepared for use in both the Gal4 and LexA yeast two-hybrid systems. For the Gal4 hybrid, the system requires two separate plasmids: (i) the plasmid containing the activation domain, which contained Gal4-AD fused to PABPC1, and (ii) the DNA-binding domain plasmid, which contained Gal4-DBD fused to BTG1. Three different PABPC1 constructs were prepared using pAD-Gal4-PABPC1 (full length) as the template (Doidge et al., 2012a): plasmids containing Gal4-AD fused to the RRM1 domain (amino acids 1–99), the RRM1 and RRM2 domains (amino acids 1–190), or RRM-1 and a part of the RRM2 domain (amino acids 1–146).

For the LexA yeast two-hybrid system, two separate plasmids were required, namely a vector expressing the LexA DNA-binding domain, which was fused to BTG1, and a second vector expressing the B42 activation domain, which was fused to PABPC1. To generate the LexA-DBD-BTG1 plasmid, the cDNA of BTG1 was amplified by PCR using pBD-Gal4-HA-BTG1 as a template. Oligonucleotides were designed at the 5' end in order to introduce the left arm sequence of pEG202NLS, including the EcoRI restriction site. The 3' oligonucleotide included the stop codon of BTG1, followed by the right arm sequence of vector pEG202NLS, including the Sall restriction site. The pLexA-BTG1 plasmids were then assembled using the PCR products and the pEG202NLS plasmid digested with the EcoRI and Sall restriction enzyme nucleases following a modified Gibson assembly protocol (Figure 7-1). All BTG1 variant plasmids were prepared as described for the BTG1 wildtype.

Three yeast vectors expressing the B42-AD domain fused with the RRM1 domain (amino acids 1-99), the RRM1 and RRM2 domains (amino acids 1-190), or RRM1 and a part of the RRM2 domain (amino acids 1-146) were generated from the corresponding pAD-Gal4-PABPC1 plasmids. To generate these constructs, PCR was used to amplify the cDNA using oligonucleotides

designed to introduce the sequence of the pJG4-5 vector at the 5' end, including an XhoI restriction site. In addition, the 3' end oligonucleotide was designed based on the plasmid template used, followed by a stop codon and the sequence of the 3' end of the vector, including an XhoI restriction site. Then, the PCR products were assembled with the pJG4-5 vector at the XhoI restriction site using the modified Gibson Assembly kit (Figure 7-2).

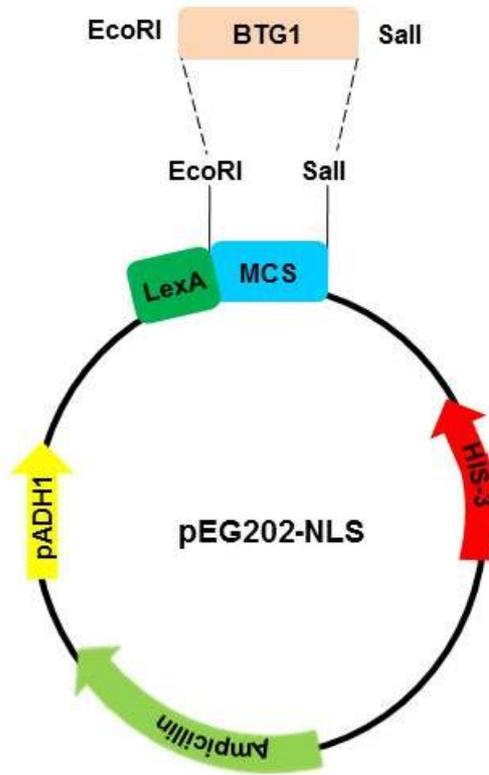


Figure 7-1 Generation of the yeast expression vector encoding LexA BTG1.

The cDNAs of wildtype BTG1 and lymphoma-associated variants of BTG1 were generated using PCR to include the 5' EcoRI and 3' Sall restriction sites. The cDNA fragments were then combined with pEG202-NLS digested with EcoRI and Sall restriction enzymes using a modified Gibson assembly kit.

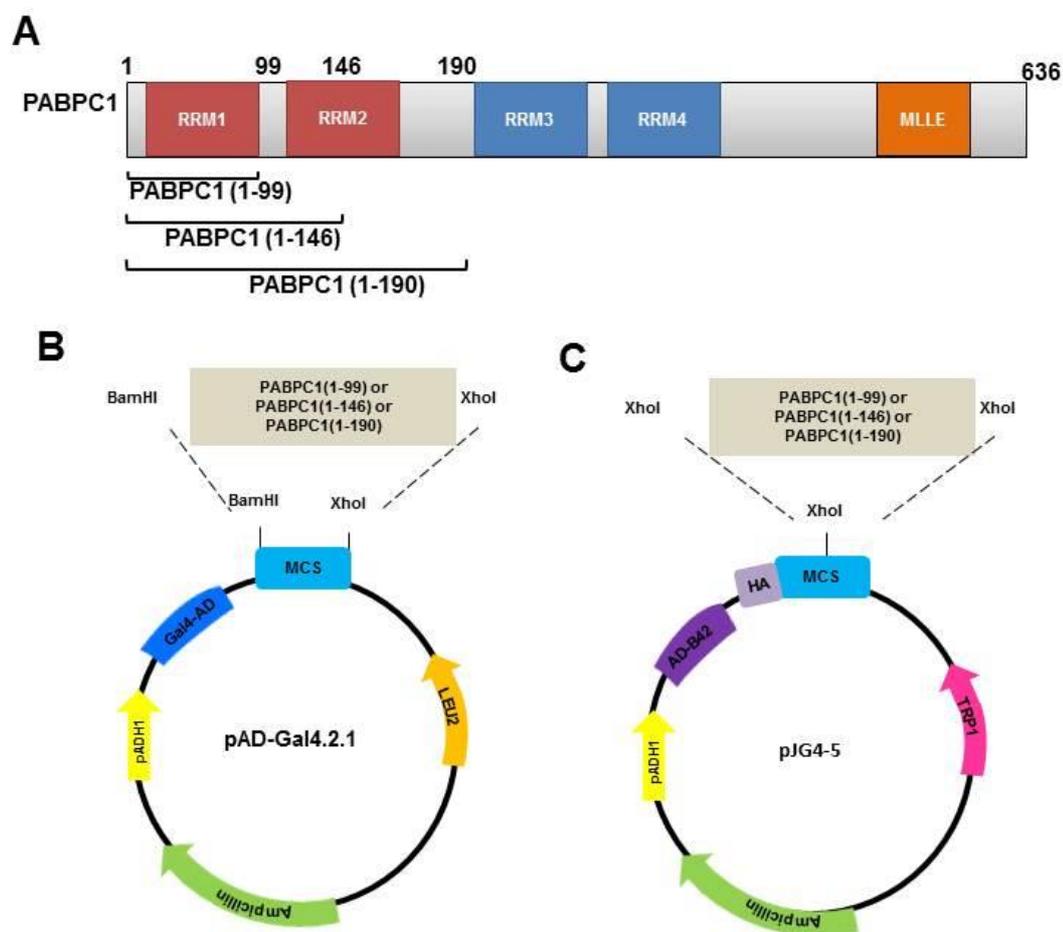


Figure 7-2 Generation of yeast expression vectors encoding PABPC1 for yeast two hybrid analysis.

(A) PABPC1 protein structure. The human PABPC1 is characterized by four highly conserved amino-terminal RNA recognition motifs (RRM1–RRM4) and a conserved carboxy-terminal domain (MLLE). The first RRM1 domain is composed of amino acids 1-99, RRM1 and a part of the RRM2 domain is composed of amino acids 1-146 and the RRM1 and RRM2 domains are composed of amino acids 1-190. (B) Three plasmids were generated which encode the AD-Gal4 domain fused PABPC1 fragments using site-directed mutagenesis and the pAD-Gal4 plasmid containing full-length PABPC1 as a template. (C) Generation of plasmids containing the LexA DNA-binding domain fused to PABPC1 fragments. PABPC1 cDNA fragments were amplified by PCR using plasmids pAD-Gal4-PABPC1 (1-99), pAD-Gal4-PABPC1 (1-146), and pADGal4-PABPC1 (1-190) as a template. Plasmids were assembled using pJG4-5 digested with XhoI and a modified Gibson assembly kit.

7.3 Generation of plasmids for the pull-down assay

A bacterial expression vector containing the glutathione S-transferase (GST) protein fused to the coding region of BTG1 was generated. To this end, PCR was used to amplify the cDNA of BTG1 containing EcoRI and SalI restriction sites at the 5' and 3' ends, respectively. Then, the PCR products were digested with the appropriate restriction enzymes and ligated into the corresponding sites of vector pGEX-4T-1 (Figure 7-3).

The bacterial expression vector encoding GST-tagged BTG2 (pGEX-4T1-BTG2) was prepared and described before (Maryati et al., 2015).

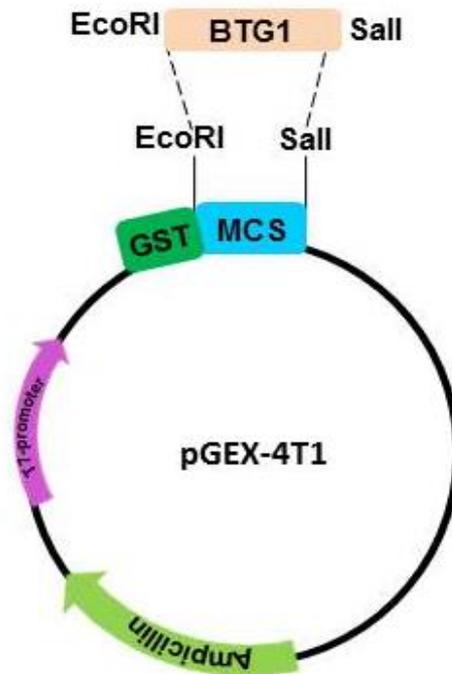


Figure 7-3 Generation of bacterial expression plasmids for pulldown assays.

A plasmid containing the Glutathione S-transferase coding region was fused to BTG1 cDNAs. BTG1 sequences were amplified by PCR to include EcoRI and Sall restriction sites. PCR products were digested and ligated into pGX4T-1 using the EcoRI and Sall restriction enzyme sites at the 5' and 3' ends respectively.

7.4 Investigating the interaction between BTG1 and PABPC1 using the yeast two-hybrid system

7.4.1 Yeast two-hybrid analysis based on the Gal4 system

The YRG2 yeast strain was transformed with two plasmids, namely PABPC1 fused to the activation domain of Gal4 and BTG1 fused to the DNA-binding domain of Gal4. Using a semi-quantitative analysis, no β -galactosidase activity was observed when the yeast was transformed with the empty vectors (pAD-Gal4 and pBD-HA-Gal4). In addition, the same result was observed when the yeast were transformed with either pAD-Gal4 and pBD-Gal4-HA-BTG1 or pBD-Gal4-HA and pAD-Gal4-PABPC1. Unexpectedly, when the yeast was transformed with pAD-Gal4-PABPC1 and pBD-Gal4-HA-BTG1, there was an absence of β -galactosidase expression, indicating that there was no interaction between BTG1 and PABPC1 (Figure 7-4 A). Based on this result, an alternative method was required. An interaction analysis using short fragments of PABPC1 rather than the full length protein previously demonstrated that BTG1 interacts with the RRM1 domain of PABPC1 (Stupfler et al., 2016). Therefore, three constructs of AD-Gal4 fused to short fragments of PABPC1, namely either the RRM1 domain only (1-99), the RRM1 and RRM2 domains (1-190), or RRM1 and a part of the RRM2 domain (1-146), were used. Notably, the interaction of BTG1 and PABPC1 1-190 fragment could not be evaluated because there were no colonies produced after the yeast transformation. However, when plasmids pAD-Gal4-PABPC1 (1-99) or pAD-Gal4-PABPC1 (1-146) were transformed into yeast with pBD-Gal4-HA-BTG1, no β -galactosidase activity was observed (Figure 7-4 B and C). This result was unexpected and to confirm whether it was correct, or whether the YRG2/Gal4 system was unsuitable for detection of the BTG1-PABPC1 interaction a second system, the LexA yeast two-hybrid system was utilised.

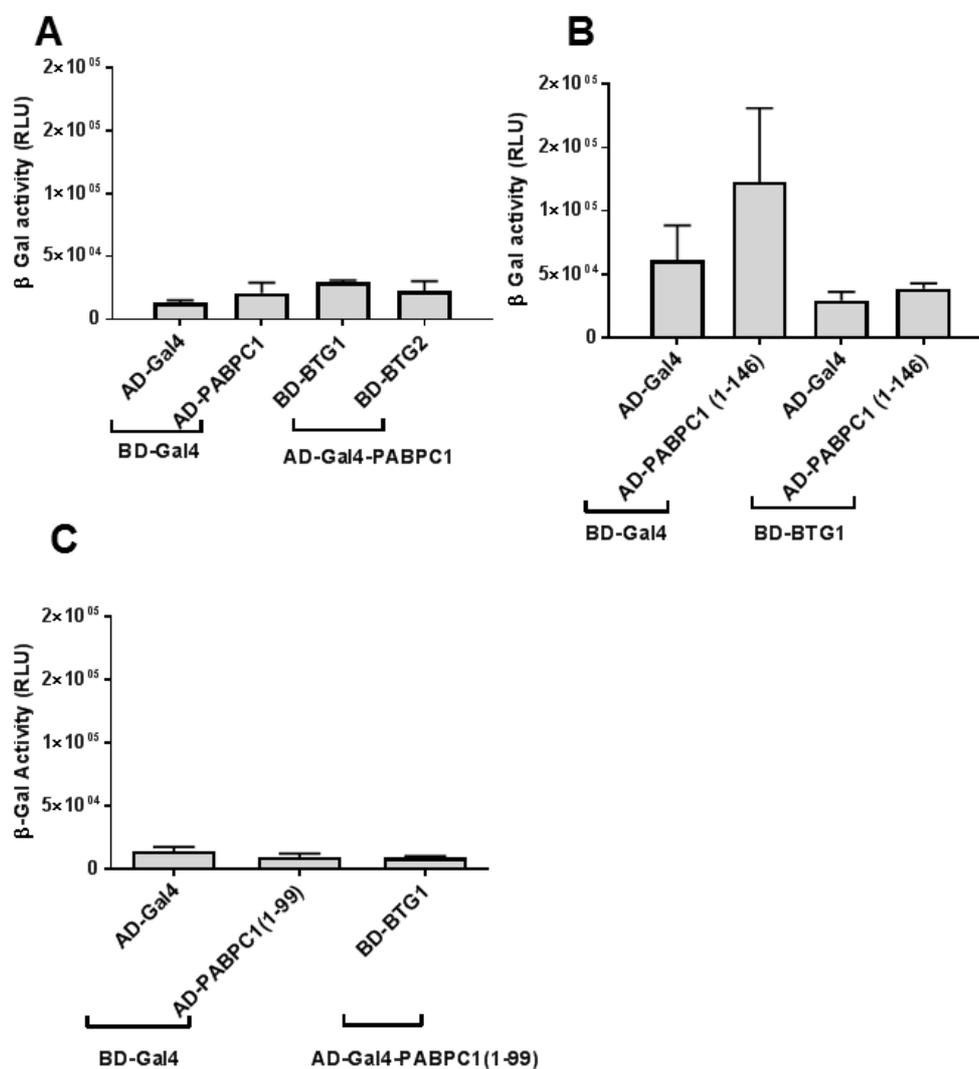


Figure 7-4 Investigation of BTG1-PABPC1 interactions using yeast two-hybrid analysis based upon the Gal4 system.

Interactions between PABPC1 and BTG1. (A) YRG2 yeast cells were transformed with vectors pAD-Gal4 or pAD-Gal4-PABPC1 (full length) and pBD-Gal4-HA or pBD-Gal4-HA-BTG1. YRG2 yeast cells transformed with the same plasmids as above except in the case of PABPC1, where plasmids encoding short fragments of PABPC1 were transformed instead: either (B) RRM1 and a part of RRM2 (pAD-Gal4-PABPC1 (1-146)), or (C) RRM1 only (pAD-Gal4-PABPC1 (1-99)). β -Galactosidase activity was normalised using cell density (O.D. measured at 600 nm). Error bars indicate the standard error of the mean (n=3).

7.4.2 Yeast two-hybrid analysis based on the LexA system

The LexA approach involves two domains, namely the LexA DNA binding domain and either the Gal4 activation domain or the B42 activation domain, depending on the yeast strain involved. Proteins to be tested for interactions were fused to these domains and the presence of protein-protein interactions induce expression of the β -galactosidase (LacZ) reporter gene. Two yeast strains were available for this assay system, namely the L40 and EGY48 strains.

7.4.2.1 L40 yeast strain

One yeast two-hybrid approach used to detect the interaction between BTG1 and PABPC1 involved a system based on the expression of two proteins that are fused to the LexA-DBD and Gal4-AD, respectively (Stupfler et al., 2016). A physical interaction between the fused proteins will bring the LexA-DBD and Gal4-AD into close proximity, which results in the expression of a reporter gene in the host L40 yeast strain (Figure 7-5 A). Thus, vectors expressing LexA-DBD-BTG1 and Gal4-AD-PABPC1 (1-146), Gal4-AD-PABPC1 (1-99), or Gal4-AD-CNOT7 were transformed into the host yeast strain (L40). They were then grown in selective media without leucine or tryptophan. Notably, when the yeast was transformed with PABPC1 (1-190), similar to the Gal4 system described in 7.4.1, it did not produce colonies. The results showed that, as expected, when the yeast was transformed with empty vectors or pEG202NLS-BTG1 with pAD-Gal4, β -galactosidase activity was not induced. Thus, the results confirmed that the LexA-DBD-BTG1 construct alone did not induce β -galactosidase expression. When plasmids AD-Gal4-PABPC1 (1-99) or AD-Gal4-PABPC1 (1-146) were used with LexA-DBD-BTG1, β -galactosidase activity was not observed, which indicated that no interaction could be detected. In contrast, when the yeast was expressed with LexA-DBD-BTG1 and AD-Gal4-hCaf1/CNOT7 (as positive control), β -galactosidase expression was induced due to the protein-protein interactions (Figure 7-5 B).

7.4.2.2 EGY48 yeast strain

The EGY48 yeast strain was transformed in the same manner as described for the L40 strain with vectors expressing LexA-DBD-BTG1, although using different vectors to express the B42 activation domain fused to fragments of PABPC1 or CNOT7. These plasmids were pEG202NLS-BTG1, pJG4-5-PABPC1 (1-146), pJG4-5-PABPC1 (1-99), or pJG4-5-PABPC1 (1-190) and pJG4-5-CNOT7, respectively (Figure 7-6 A). The EGY48 yeast strain does not contain the reporter gene, therefore an additional plasmid was transformed, namely the pSH18-34 reporter. Once transformed, yeast cells were grown on selective glucose media without uracil, tryptophan and histidine.

The results showed an absence of β -galactosidase activity when yeast were transformed with empty vectors or pJG4-5 and pEG202NLS-BTG1, which indicated that the LexA-DBD-BTG1 did not induce β -galactosidase expression. In contrast, when the yeast was transformed using pEG202NLS-BTG1 and pJG4-5-CNOT7, β -galactosidase expression was induced due to the BTG1-hCaf1/CNOT7 interaction. Furthermore, when yeast were transformed with pJG4-5-PABPC1 (1-146) and pEG202NLS-BTG1, β -galactosidase activity was seen, although it was less than the level achieved with the BTG1-hCaf1/CNOT7 interaction (Figure 7-6 B and C).

PABPC1 interactions with BTG1 variants harbouring mutations located in the α -helical structure such as F25 (α 1), R27H (α 1) and F40C (α 2) were also assessed. The results showed that two variant residues, namely F25C and R27H, were able to induce β -galactosidase expression to the same extent as wildtype BTG1. This indicated that these two variants were able to interact with PABPC1 (1-146). In contrast, the final variant, namely F40C, was significantly impaired in terms of the interaction with PABPC1 (1-146) ($p < 0.001$) (Figure 7-7 A). Another pair of BTG1 variants, L104H and I115V, which are located in the beta strands β 2 and β 3, respectively were analysed. The results showed that L104H was impaired in terms of its interaction with

PABPC1, but this was not the case for I115V, which indicated that L104H was unable to interact with PABPC1 (Figure 7-7 B).

To confirm equal protein expression, proteins were extracted from yeast cell cultures and subjected to western blot analysis using an anti-HA primary antibody to detect B42-HA-PABPC1 (1-146), as well as an anti-LexA primary antibody to detect the expression of LexA-BTG1 variants (Figures 7-6 D and 7-7 C).

Despite these initial results, it proved difficult to repeat the experiments and assess the interactions between PABPC1 and all BTG1 variants. The experiments shown in Figure 7-6 and 7-7 were repeated more than three times in both variant groups using the same experimental conditions. However, the results were inconsistent with several experiments failing to show β -galactosidase expression. Similar results were found when the yeast were transformed with pEG202NLS -BTG1 and either pJG4-5-PABPC1 (1-99) or pJG4-5-PABPC1 (1-190). The results were not consistent and it is hence difficult to reach a conclusion based on these findings. Therefore, alternative methods to detect the BTG1-PABPC1 interaction were evaluated.

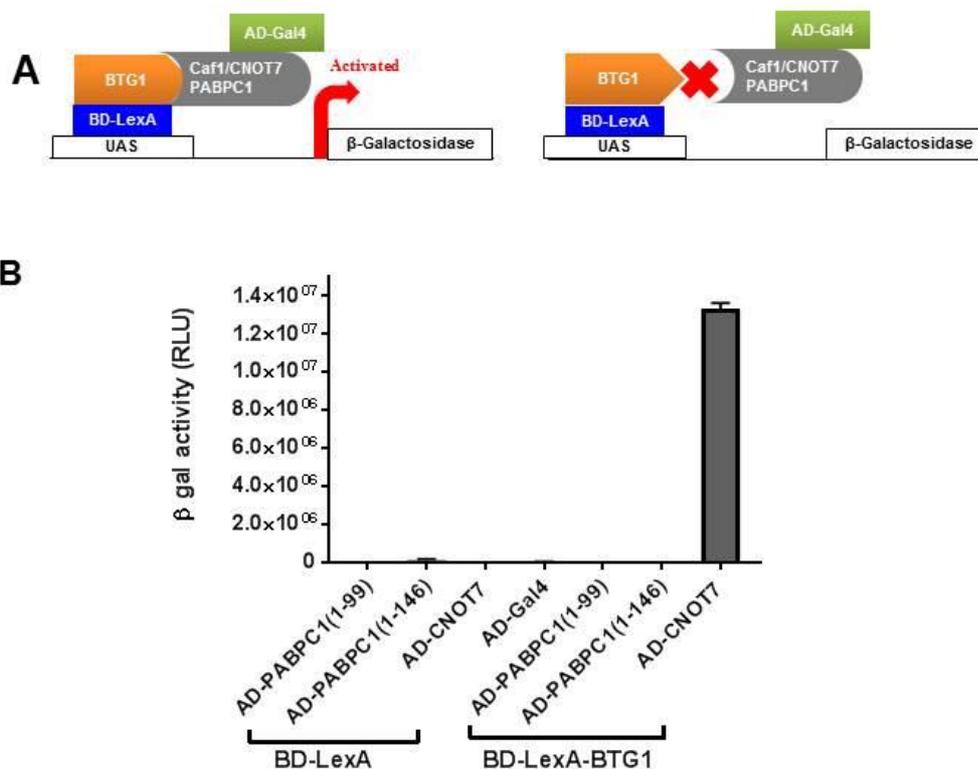


Figure 7-5 Investigation of BTG1 and PABPC1 interactions using a LexA yeast two-hybrid system and L40 cells.

(A) Schematic of yeast two-hybrid interaction analysis of BTG1 and PABPC1. The Gal4 activation domain (AD) was fused to PABPC1 and the LexA DNA binding domain (DBD) was fused to BTG1. Interactions between the DBD-LexA-BTG1 and Gal4-AD-PABPC1 proteins induces the expression of a β -Galactosidase reporter (left panel). In the absence of an interaction, no expression of β -Galactosidase is seen (right panel). (B) Interactions between PABPC1 and BTG1. L40 yeast cells were transformed with vectors pAD-Gal4, pAD-Gal4-PABPC1 (1-146), pAD-Gal4-PABPC1 (1-99) or pAD-Gal4-CNOT7 and either pEG202NLS or pEG202NLS-BTG1. β -Galactosidase activity was normalised using cell density. Error bars indicate the standard error of the mean (n=3).

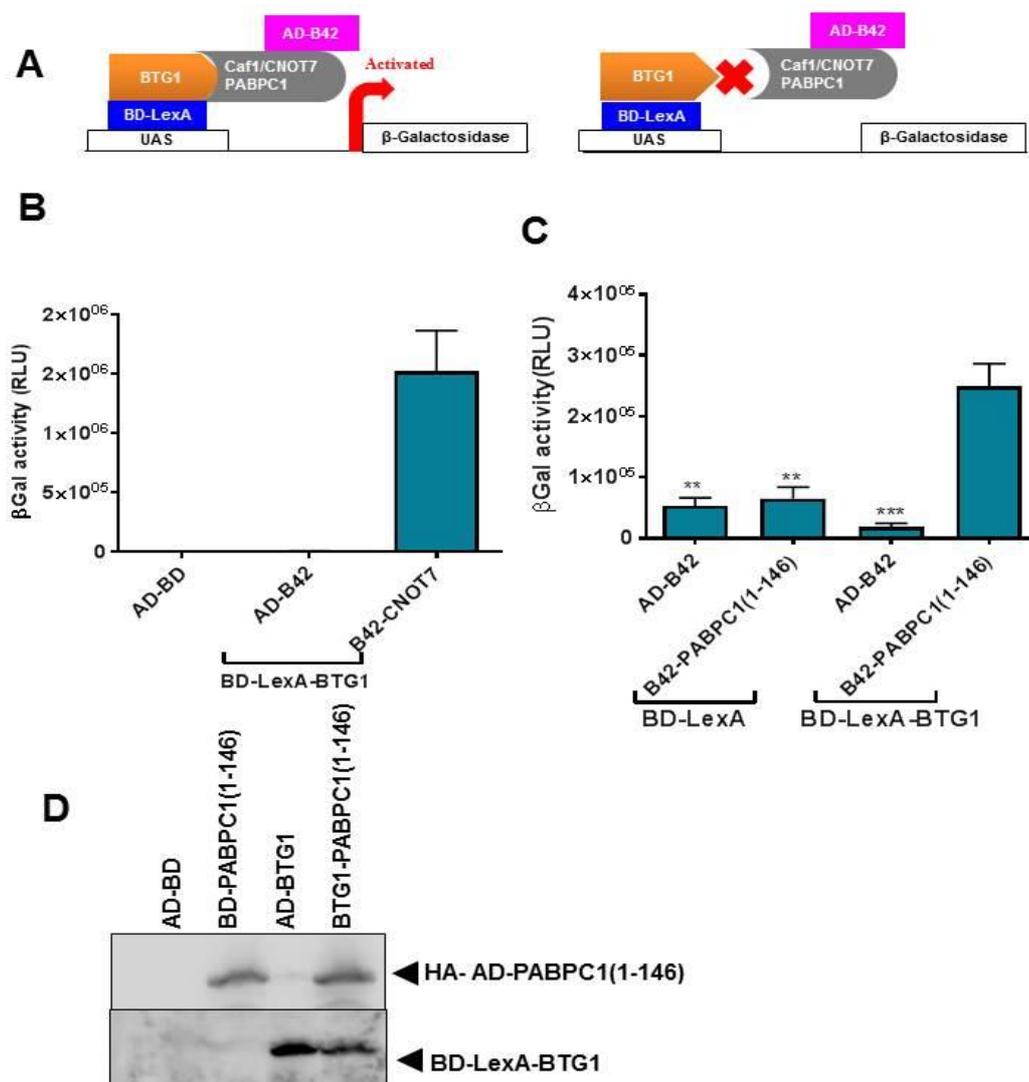


Figure 7-6 Detection of PABPC1 and BTG1 interactions using a LexA yeast two hybrid system and EGY48 cells.

(A) Schematic of yeast two-hybrid interaction analysis of BTG1 and PABPC1. The B42 activation domain (AD) was fused to PABPC1 and the LexA DNA-binding domain (DBD) was fused to BTG1. Interactions between LexA-DBD-BTG1 and B42-AD-HA-PABPC1 (1-146) proteins induces β -Galactosidase reporter expression (left panel). In the absence of an interaction, no expression of the β -Galactosidase gene is seen (right panel). EGY48 yeast cells were transformed with vectors pJG4-5, pJG4-5-CNOT7 or pJG4-5-PABPC1 (1-146) and either pEG20NLS or pEG20NLS-BTG1. (B) Interaction between BTG1-hCaf1/CNOT7 and (C) BTG1-PABPC1 interactions. β -Galactosidase activity was normalised using cell density. Error bars indicate the standard error of the mean ($n=3$). One way Anova (dunnett post-hoc method) was used to calculate P values (** $p<0.01$ and *** $p<0.001$). (D) Protein expression was confirmed by western blot analysis using an Anti-HA primary antibody to detect AD-B42 HA-PABPC1-146 and an anti-LexA primary antibody to detect DBD-LexA-BTG1.

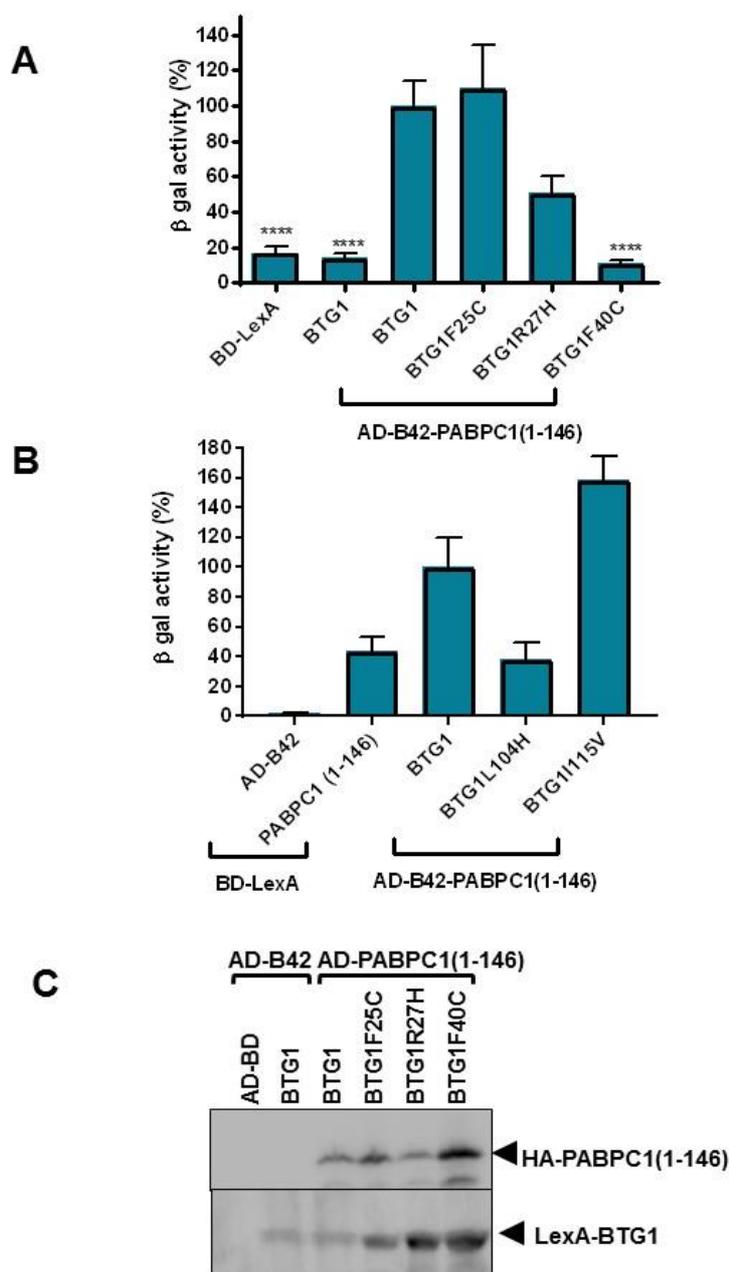


Figure 7-7 Interaction analysis of BTG1 variants and PABPC1.

EGY48 yeast cells were transformed with vectors pJG4-5, pJG4-5-CNOT7 or pJG4-5-PABPC1 (1-46) and pEG202NLS containing wild type BTG1 or lymphoma-associated-BTG1 variants. (A) BTG1 variants located in α -helical regions F25C (α 1), R27H (α 1) and F40C(α 2). (B) BTG1 variants located in β -sheets L104H (β 2) and I115V (β 3). β -Galactosidase activity was normalised using cell density. The error bars indicate the standard error of the mean of one or two independent experiments. One way Anova (dunnnett post-hoc method) was used to calculate P values (**** $p < 0.0001$). (C) Protein expression was confirmed by western blot analysis using an Anti-HA primary antibody to detect AD-B42-HA-PABPC1-146 and an anti-LexA primary antibody to detect DBD-LexA-BTG1.

7.5 Investigating the interaction between BTG1 and PABPC1 using immunoprecipitation and pull-down assays

7.5.1 Co-immunoprecipitation analysis using mammalian cells

Co-immunoprecipitation was performed using the mammalian cell line HEK293T. Cells were co-transfected with HA-tagged BTG1 and either Flag-tagged hCaf1/CNOT7 or PABPC1. After 48 hours, the cells were harvested. Proteins were then precipitated using the anti-HA antibody bound to agarose beads and the presence of immunoprecipitated proteins revealed by western blot analysis using anti-HA and anti-Flag primary antibodies. HA-BTG1 was able to precipitate hCaf1/CNOT7, but not Flag-PABPC1, which indicates that BTG1-hCaf1/CNOT7 interaction could be detected, although the BTG1-PABPC1 interaction could not.

7.5.2 GST pull-down assay using bacterial expression

Next, GST-pulldown assays were designed to detect the interaction between BTG1 and the N-terminus of PABPC1. Stupfler et al. (2016) utilised GST-tagged BTG2, but not the equivalent for BTG1. For this reason, BTG2 was analysed as well as BTG1. The expression of GST-BTG1 or GST-BTG2 recombinant proteins were analysed in terms of their ability to pull down purified His-tagged PABPC1 consisting of the RRM1 and RRM2 domains or only the RRM1 domain. To this end, GST-BTG1, GST-BTG2 and His-PABPC1 fragments were expressed in bacterial cells. After cell lysis, lysates were mixed and incubated with glutathione-agarose beads. Bound proteins were separated by 14% SDS-PAGE that were stained with Coomassie brilliant blue or analysed via western blot analysis. The results showed that purified GST was unable to bind with either His-PABPC1 (1-190) or His-PABPC1 (1-99), as shown in Figure 7-8, lanes 6 and 13, respectively. Unexpectedly, the purified GST-BTG1 was unable to pull down His tagged PABPC1 (1-190) or PABPC1 (1-99), as shown in Figure 7-8, lanes 7 and 14, respectively. The same scenario was found in the cases of solely

GST and GST-tagged BTG2, which both proved unable to bind to either His-tagged PABPC1 (1-190) or PABPC1 (1-99) (Figure 7-9).

These results were confirmed by western blot analysis, which provides a more sensitive detection method. The anti-His antibody was used to detect His tagged with PABPC1 (1-190) and His-PABPC1 (1-99), while the anti-GST antibody was used to detect GST tagged with BTG1 (Figure 7-8 B).

7.5.3 His pull-down assay

A His pull-down experiment was then conducted using recombinant proteins consisting of His tagged PABPC1 RRM1-RRM2 domain (1-190) or PABPC1-RRM1 domain (1-99) and GST tagged BTG1 or BTG2. To this end, His-PABPC1 fragments, GST-BTG1 and GST-BTG2 were expressed in bacterial cells. After cell lysis, lysates were mixed and incubated with Ni²⁺-NTA-agarose beads. Bound proteins were separated by 14% SDS-PAGE that were stained with Coomassie brilliant blue or analysed via western blot analysis. The results showed that purified His-PABPC1 (1-99) was unable to bind to GST alone, which confirmed the specificity in terms of their interaction. However, purified His-PABPC1 (1-99) was also incapable of interacting with the purified GST-BTG1, as shown in lanes 6 and 7, respectively, of Figure 7-10. However, purified His-PABPC1 (1-99) was able to bind with the GST-BTG2 (lane 8, Figure 7-10), which confirmed that BTG2 interacts with the RRM1 domain of PABPC1. Although there is a high degree of similarity between BTG1 and BTG2 in their BTG domain, it was not possible to detect a BTG1 interaction with PABPC1, possibly due to the lower expression levels of BTG1.

When His tagged RRM1 and RRM2 of PABPC1 was used, the results showed the same finding as for the pull down of His-PABPC1 (1-99). The purified His-PABPC1 (1-190) was incapable of binding with solely GST or GST-BTG1 (lanes 6 and 7, respectively, of Figure 7-11). In contrast, the RRM1 and RRM2 domains of PABPC1 were able to bind to GST-BTG2, as shown in lane 8 of Figure 7-11. In both experiments, the lysed GST tagged BTG1 or BTG2 was

incubated with Ni²⁺-NTA-agarose beads in order to confirm the specificity of the interaction and analyse any background. Importantly, the appearance of GST-BTG2 in this His pull-down assay was due to the interaction with PABPC1 and not because of non-specific binding with the beads (lanes 9 and 10, in Figures 7-10 and 7-11).

All the mixtures with beads were subjected to Western blot analysis using an anti-His primary antibody for the detection of His-tagged PABPC1 (1-99) and PABPC1 (1-190), or an anti-GST primary antibody for the detection of GST tagged BTG1 or BTG2.

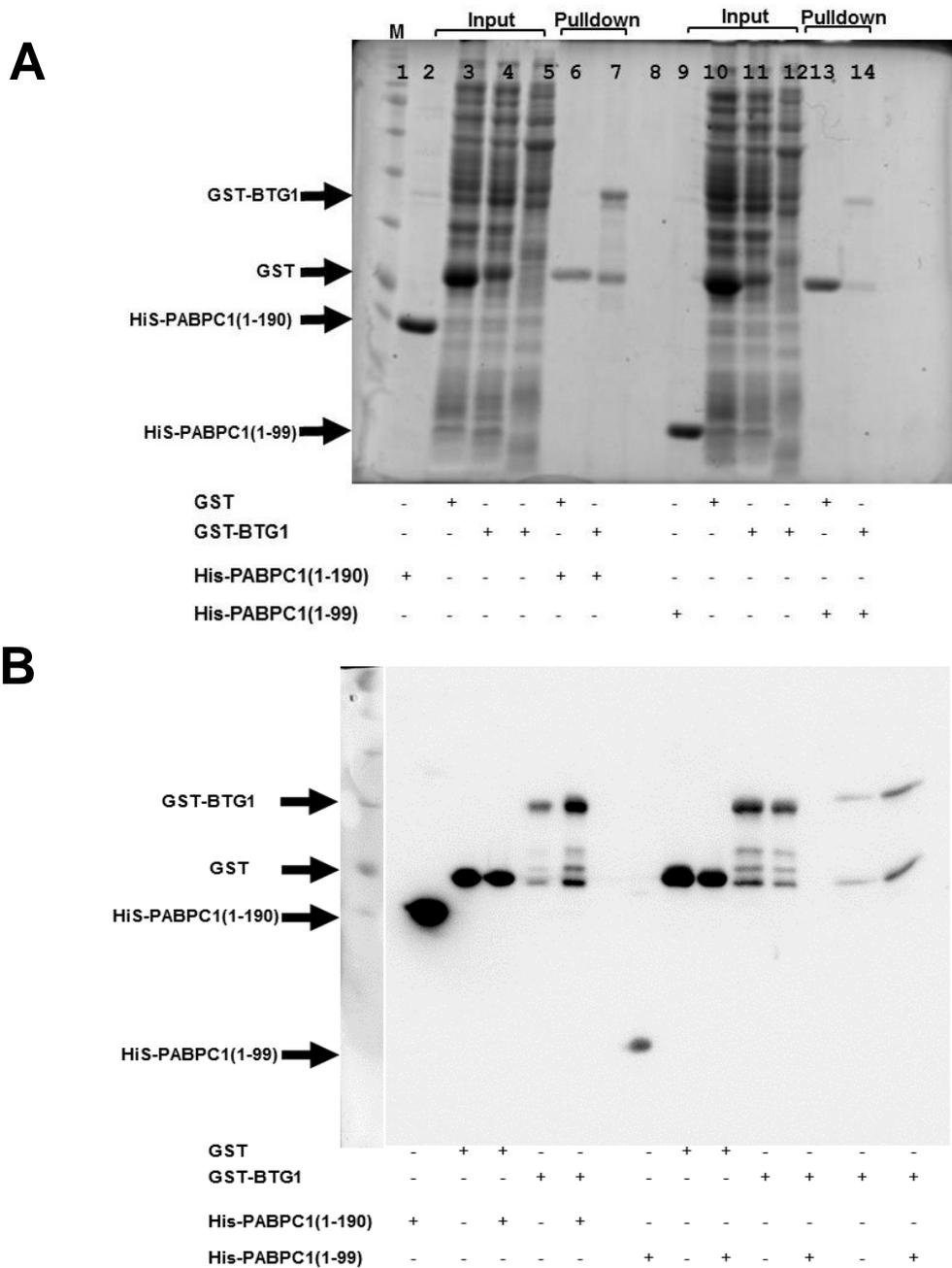


Figure 7-8 Detecting the interaction between BTG1 with PABPC1 using GST-pull down assays.

(A) Pull down assay performed with glutathione *S*-transferase (GST) beads. GST only or GST-BTG1 was incubated with 100 μ g of purified His tagged RRM1 and RRM2 domains of PABPC1 (1-190) or His-tagged RRM1 domain of PABPC1 (1-99). After pull down using glutathione-agarose, bound proteins were resolved by 14% SDS-PAGE and stained with Coomassie Blue.(B) Western blot analysis. An anti-GST antibody and anti-His antibody were used to detect GST-tagged BTG1 and His tagged PABPC1, respectively. Input represents 10% of the sample volume.

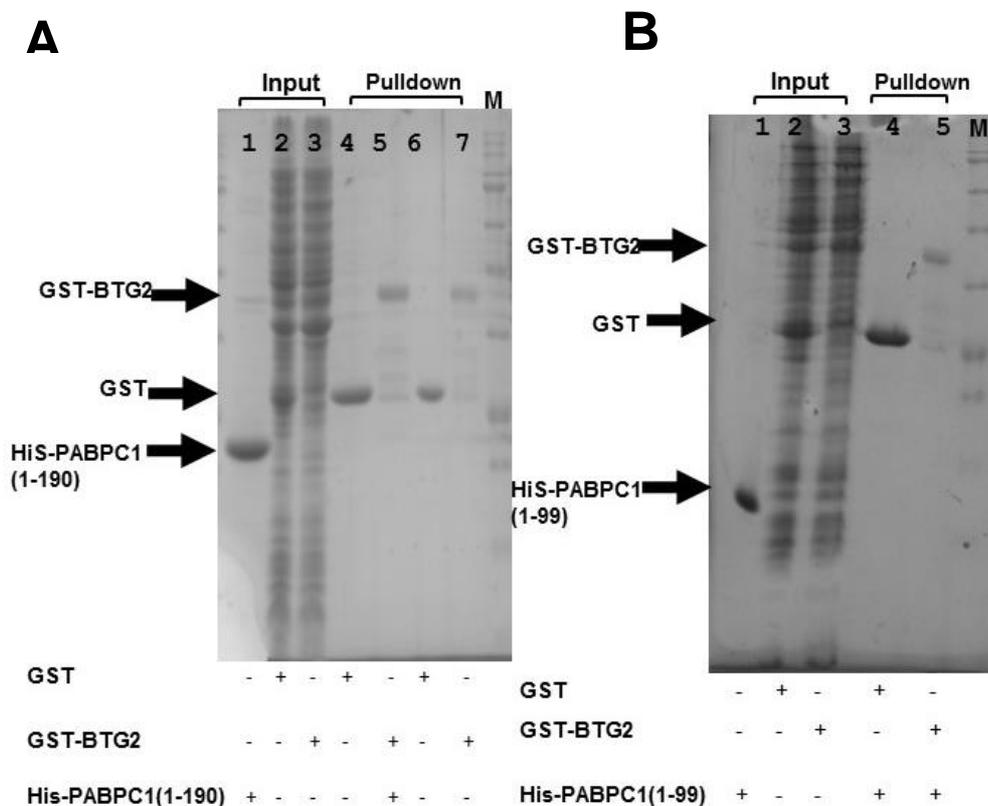


Figure 7-9 Detecting the interaction between BTG2 and PABPC1 using GST-pull down assays.

(A) Pulldown of GST-BTG2 and PABPC1 RRM1-RRM2 (1-190). GST-pull down assays were performed with glutathione agarose beads. GST only or GST-BTG2 was incubated with 100 μ g of purified His tagged PABPC1 (1-190). (B) Pulldown of GST-BTG2 and PABPC1 RRM1 (1-99). Samples were resolved by 14% SDS-PAGE and stained with Coomassie Blue.

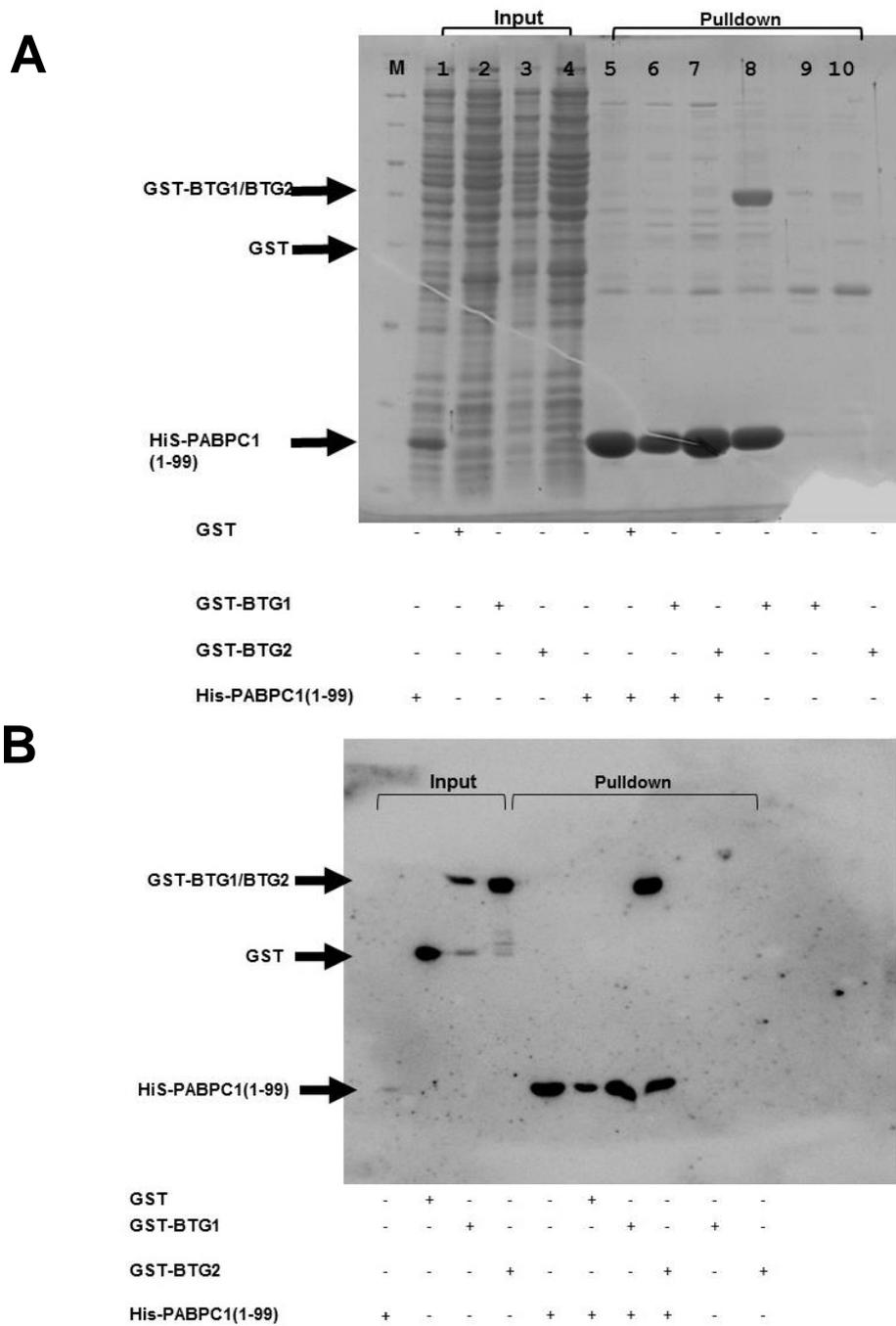


Figure 7-10 Detecting the interaction between BTG1/BTG2 and the RRM1 domain of PABPC1 using His-pull down assays.

(A) Pull-down of GST-BTG1/GST-BTG2 and PABPC1 RRM1 (1-99). Pull down assays were performed using Ni²⁺-NTA agarose beads. His-PABPC1 (1-99) was incubated with purified GST-BTG1, GST-BTG2 or GST alone. Samples were resolved by 14% SDS-PAGE and stained with Coomassie Blue. (B) Western blot analysis of pull-down of GST-BTG1/GST-BTG2 and PABPC1 RRM1 (1-99). An anti-GST antibody was used to detect GST-BTG1 and GST-BTG2, while an anti-His antibody was used to detect His-PABPC1 (1-99). Input represents 10% of the sample volumes.

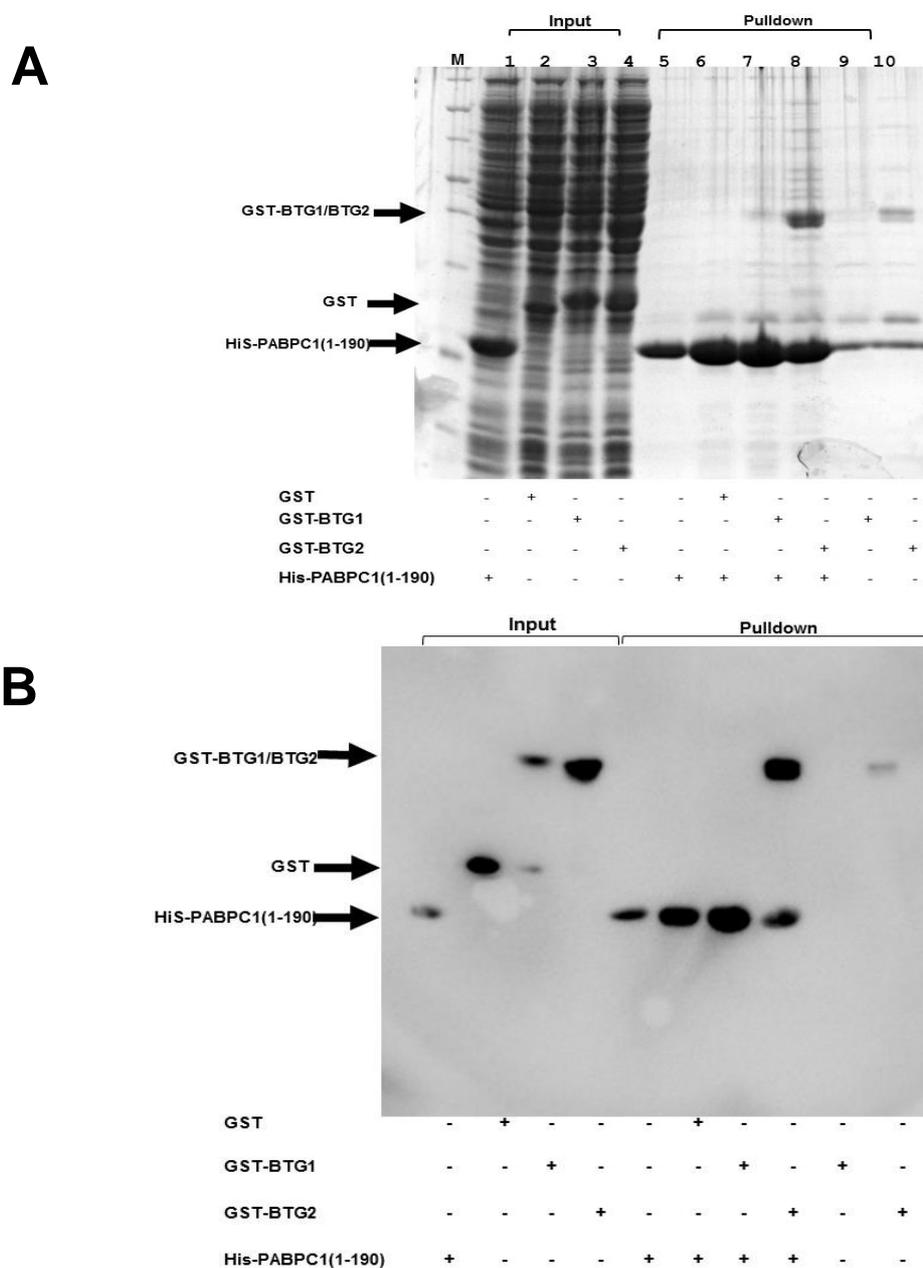


Figure 7-11 Detecting the interaction between BTG1/BTG2 and the RRM1 and RRM2 domains of PABPC1 using His-pull down assays.

(A) Pull-down of GST-BTG1/GST-BTG2 and PABPC1 RRM1-RRM2 (1-190). Pull down assays were performed using Ni²⁺-NTA agarose beads. His-PABPC1 (1-190) was incubated with purified GST-BTG1, GST-BTG2 or GST alone. Samples were resolved by 14% SDS-PAGE and stained with Coomassie Blue.(B) Western blot analysis of pull-down of GST-BTG1/GST-BTG2 and PABPC1 RRM1 (1-190). An anti-GST antibody was used to detect GST-BTG1 and GST-BTG2, while an anti-His antibody was used to detect His-PABPC1 (1-190). Input represents 10% of the sample volumes.

7.6 Discussion

In the last decade, it has been shown that TOB proteins recruit the Ccr4-Not complex to target mRNA by interacting with the RNA-binding protein PABPC1 through the PAM2 motif. However, this motif is absent in the BTG1/BTG2 proteins (Funakoshi et al., 2007, Ezzeddine et al., 2007), which suggests that BTG1 and BTG2 recruit the Ccr4-Not complex to target mRNA using a different mechanism (Mauxion et al., 2008). Recently, Stupfler and colleagues (2016) discovered that BTG1 and BTG2 bind to PABPC1 through their BTG domain, specifically in the box C region. BTG1 and BTG2 interact with the RRM1 domain of PABPC1, rather than the C-terminal PABC domain which is how TOB proteins bind to PABPC1. This suggests that BTG2 acts as an adaptor between PABPC1 and hCaf1(CNOT7/CNOT8), and that this interaction is required in the deadenylation process (Stupfler et al., 2016). Previously, through tethering the BTG1 variants using artificial mRNA, BTG1 has been shown to regulate mRNA decay partially independent of hCaf1(CNOT7/CNOT8) binding, which suggests that BTG1 requires another factor to recruit hCaf1(CNOT7/CNOT8) to target mRNA. For this reason, it was decided to use three different approaches, namely co-immunoprecipitation, yeast two-hybrid analysis, and pull-down assays, to detect the BTG1-PABPC1 interaction.

7.6.1 Immunoprecipitation analysis

First, co-immunoprecipitation in a mammalian cell line was performed by transfecting HEK293T cells with HA-BTG1 and either Flag-tagged CNOT7 or PABPC1 expression vectors. The interaction between BTG1 and CNOT7 has previously been confirmed (Prévôt et al., 2001), and the results of the present study confirm this interaction. However, the interaction between BTG1 and PABPC1 was not observed using this method. In the study by Stupfler et al., (2016), HEK293T cells were treated with an amine-reactive crosslinker known as Lomant's reagent (dithiobis succinimidyl propionate, DSP) after transfection. This reagent is normally used for weak protein-protein

interactions because it captures such interactions. It might be that this interaction is weak and thus covalent crosslinking is required.

7.6.2 Yeast two-hybrid analysis

As an alternative approach to study the BTG1-PABPC1 interaction, yeast two-hybrid analysis was used. However, analysis using the Gal4 (YRG2 yeast strain) or LexA (L40 yeast strain) system could not detect the interaction. There are many reasons why a false negative results might be obtained in yeast two-hybrid experiments, including the fact that fused yeast reporter proteins or anchors may cause steric hindrance that impedes interaction. Alternatively, the presence or absence of different post-translational protein modifications in yeast compared to that of higher eukaryotes may influence the ability of proteins to interact (Brückner et al., 2009). In the study by Stupfler et al. (2016), the diploid yeast strain Y187/L40 was used for their protein interaction analysis, which was based on the yeast mating system. The principle behind this analysis involves using two haploid yeast strains of opposite mating types (a and α), one containing, for example a DBD fusion construct and the other containing an AD fusion construct. These are then simply mated in order to obtain a diploid yeast strain containing both constructs (Causier and Davies, 2002). One of the advantages of using the diploid system in cases where the expression of the bait interferes with the yeast's viability, is that they can better tolerate the expression of toxic proteins (Golemis et al., 2011).

In terms of the experiments performed using the EGY48 yeast strain, the findings were not consistent (Figures 7-6 and 7-7). The interaction between BTG1 and PABPC1 was only confirmed in some experiments, which precluded the analysis of all BTG1 variants. Interestingly, out of both groups of BTG1 variants, the F40C and L104H mutations resulted in a loss of the ability of BTG1 to regulate mRNA decay and to interact with PABPC1. In contrast, the BTG1 variants F25C, R27H and I115V could induce mRNA

degradation to the same extent as wildtype BTG1. They were also still able to interact with PABPC1, however further experiments are required to confirm these findings.

7.6.3 GST and Ni²⁺-NTA pulldown experiments

Finally, the His pull-down assays proved to be a promising means of detecting the BTG2-PABPC1 interaction. Yet, surprisingly it could not detect an interaction between BTG1 and PABPC1. The reason for this could be the weak expression of GST-BTG1 compared to GST-BTG2 or GST alone (Figures 7-8, 7-10 and 7-11). Further optimisation of the expression of GST-BTG1 in bacteria is therefore required. In addition, expression of GST tagged human BTG1 cDNA may be improved by using a codon-optimised variant. Another way to help the human BTG1 expression is to transform a different bacterial strain such as *E.coli. Rosetta* that contains additional tRNA genes optimised for eukaryotic codon usage.

Chapter 8

Concluding Remarks and Future Outlook

Chapter 8: Concluding Remarks and Future Outlook

8.1 Overview

BTG1 is a member of the anti-proliferative BTG/TOB proteins that share a conserved BTG domain, which allows binding with hCaf1(CNOT7/CNOT8) (Prévôt et al., 2001, Funakoshi et al., 2007, Horiuchi et al., 2009, Mauxion et al., 2009, Winkler, 2010). Caf1 is one of the deadenylase subunits of the Ccr4-Not complex, which induces mRNA degradation via removal of the poly (A) tail of the target mRNA (Goldstrohm and Wickens, 2008, Collart, 2016). Previous studies concluded that the BTG/TOB proteins inhibit cell proliferation by recruiting hCaf1(CNOT7/CNOT8) to the target mRNA with subsequent mRNA degradation (Doidge et al., 2012a, Ezzeddine et al., 2012). Over the last seven years, studies have identified novel somatic variants in *BTG1* and *BTG2* in NHLs using high-throughput sequencing; however, the nature of these mutations and their effects on protein function remain unknown (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013, Fukumura et al., 2016). Understanding the effects of the BTG1 variants on proliferation activity, and on the interaction with the hCaf1(CNOT7/CNOT8) deadenylase subunit, is important. Indeed, it is imperative to identify whether the anti-proliferative activity of BTG1 requires the deadenylase activity of hCaf1(CNOT7/CNOT8) to regulate cell proliferation.

8.2 BTG1 and BTG2 variants associated with non-Hodgkin's lymphoma

Using next generation sequencing provides information that helps to reveal features exclusive to specific types of B-cell lymphomas (Morin and Gascoyne, 2013, Swerdlow et al., 2016). Recently, several reports identified variants in BTG1 and BTG2 in NHLs (Morin et al., 2011, Lohr et al., 2012, Love et al., 2012, Walker et al., 2012, Zhang et al., 2013). Therefore, data was collected regarding BTG1 and BTG2 variants from the literature and genomic resources, such as the Ensembl and COMSIC databases (Forbes et al., 2010, Flicek et al., 2011). In

total, 45 variants of BTG1 and BTG2 were identified, mostly represented in DLBCL, FL and BL. The majority of these variants are clustered in the BTG domain (Figure 3-4 and Figure 3-6). It has been suggested that aberrant SHM is significantly correlated with mutations in BTG1 and BTG2 genes, evidenced by samples drawn primarily from DLBCL and central nervous system lymphoma (Jiang et al., 2012, Zhang et al., 2013, Fukumura et al., 2016). All variants were considered, using prediction algorithms to identify those that potentially affect protein function. In total, 21 variants (15 BTG1 and 7 BTG2) were predicted by SIFT analysis to be damaging (Figure 3-5). However, when using the Suspect algorithms additionally analysed protein-protein interaction, the results showed only eight of these variants were predicted to be associated with disease: in BTG1, these were the variants (M11I, F25C, Q36H, F40C, P58L, G66V and L104H) and in BTG2 only the variant L100P met these criteria (Table 3-1 and Table 3-2). It has been shown that the Suspect algorithms have higher accuracy or selectively compared with SIFT, about 82% and 62%, respectively (Yates et al., 2014). Differences observed with SIFT and Suspect analysis from this data analysis. It may be that the SIFT algorithm has a higher sensitivity and predicts a greater number of variants, as compared with Suspect at the expense of a greater number of false positive results. As the majority of variants that are predicted to affect protein function using SIFT analysis were represented in BTG1. Therefore, 16 BTG1 variants were identified in NHL and predicted to be damaging using bioinformatics analysis. The additional variant (E59D) was selected as a negative control, because it is a conservative change which was predicted to be neutral using SIFT and Suspect.

8.3 Truncated BTG1 with 21 amino acids at the C-terminal affects proliferation activity

BTG1 deletions occur in a proportion of BCP-ALL cases, suggesting that incidence of recombination hotspots occur in the second exon of *BTG1*,

resulting in the deletion of amino acids due to illegitimate RAG-mediated recombination (Waanders et al., 2012). The BTG1 C149 deletion was evaluated with PROVEAN algorithms designed for these cases, and the result showed predicted that this mutation affects protein function. A proliferation assessment showed that the last 21 amino acids at in the C-terminal of BTG1 are required to suppress cell proliferation.

8.4 The anti-proliferative activity of BTG1 requires human Caf1 (CNOT7/CNOT8) interaction

BTG/TOB proteins are required for the hCaf1(CNOT7/CNOT8)deadenylase activity that inhibits cell proliferation (Doidge et al., 2012a, Ezzeddine et al., 2012). Based on the protein sequences, hCaf1/CNOT7 and hCaf1/CNOT8 are 89% similar and 76% identical at the amino acid level (Aslam et al., 2009, Winkler and Balacco, 2013). All BTG1 variants were evaluated in their interaction with hCaf1(CNOT7/CNOT8) using a yeast two-hybrid analysis, before assessing their proliferation activity in HEK293T cells.

8.4.1 BTG1 variants located in the α -helical regions

Three BTG1 variants are located in the α -helical structures. F25C (α 1) impaired the interaction with hCaf1(CNOT7/CNOT8) <50%, but did not totally abrogate it, as occurred with R27H (α 1) and F40C (α 2) (Figure 4-6 B). Although these variant residues are not located in box A or box B in the BTG domain, they are still required for the interaction with hCaf1 (CNOT7/CNOT8). This suggests that both F25C and R27H variant residues are located close to the hCaf1/CNOT7 surface interaction site (Figure 4-6 A). Consistent with this notion, a proliferation assessment showed that BTG1 R27H causes a complete loss of proliferation activity, while BTG1 F25C also caused a loss in the proliferation activity, however this was not a total loss, which correlates to the partial impairment of the interaction with hCaf1(CNOT7/CNOT8) (Figure 5-5). In the case of BTG1 F40C, which is located away from the surface of the interaction, BTG1 is unable to bind with hCaf1(CNOT7/CNOT8) (Figure 4-6 A

and B). Consequently, BTG1 F40C is unable to inhibit cell proliferation. However, the western blot analysis showed only a weak expression of BTG1 F40C compared with the BTG1 wild type, strongly suggesting that the variant protein is unstable (Figure 5-5 B). Further experiments are required to confirm protein stability and folding structure, such as circular dichroism (CD) spectroscopy, which allows for the characterisation of secondary and tertiary protein structure, or differential scanning fluorimetry (DSF), which provides information about the stability of the folded protein; increased stability is reflected in a higher melting temperature (T_m). Thus, BTG1 F40C may have a less stable folding ability, which is reflected in a lower T_m as compared to the wildtype protein.

8.4.2 BTG1 variants located in the β -sheet regions

The two variant residues L104H and I115V impair the interaction with hCaf1(CNOT7/CNOT8), that both are located in the box B domain, and close to the surface of interaction in β_2 and β_3 , respectively (Figure 4-7). This is consistent with (Yang et al., 2008, Doidge et al., 2012a), which found that BTG/TOB proteins interact with hCaf1(CNOT7/CNOT8) via box B in the BTG domain. Using a proliferation assay, BTG1 I115V was unable to inhibit cell growth, although this mutation also does not totally abrogate the interaction with hCaf1(CNOT7/CNOT8) (20-40%) (Table 8-1). In contrast, BTG1 L104H is entirely unable to bind with hCaf1(CNOT7/CNOT8) ($p < 0.0001$), with a corresponding partial loss of anti-proliferation activity ($p < 0.01$) (Figure 4-7 and Figure 5-6). This result is not consistent with our hypothesis, suggesting that BTG1 is not the only requirement for hCaf1(CNOT7/CNOT8) interaction, and that the presence of another factor may also mediate the suppression of cell proliferation.

8.4.3 BTG1 variants located in the loop regions

Three variant residues that suggested to impair the interaction with hCaf1(CNOT7/CNOT8) suggested are located closed to the interaction surface

interaction and are in box A in the BTG domain; P58L (L1), G66V (α 3) and N73K (L4) in agreement with (Doidge et al., 2012a, Horiuchi et al., 2009, Yang et al., 2008) (Figure 4-8). Using a proliferation assay, BTG1 requires hCaf1(CNOT7/CNOT8) interaction to suppress cells proliferation (Figure 5-7). The M11I variant is located on the opposite side of the surface of interaction, however, it causes a total abrogation in the binding with hCaf1(CNOT7/CNOT8), and is consequently unable to inhibit cell growth (Figure 4-8 and Table 8-1). The last variant residue in this group, E117D, is able to interact with hCaf1/CNOT7 but marginally impairs the interaction with hCaf1/CNOT8. Using the proliferation assay, BTG1 was unable to inhibit the cell progression due to this mutation, suggesting the presence of another factor as discussed in the next section.

8.5 Does BTG1 require only the interaction with human Caf1 (CNOT7/CNOT8) to inhibit cell proliferation?

The first 10 amino acids of BTG1 are not necessary in the interaction with hCaf1(CNOT7/CNOT8), as shown with results from the variants Δ 10N BTG1 and BTG1 H2Y. Consistent with this, both variants are able to suppress cell cycle progression, suggesting via hCaf1(CNOT7/CNOT8) deadenylase activity. To support this hypothesis, BTG1 E59D consistently correlated results in a reduction of anti-proliferative activity ($p < 0.01$), and interaction with hCaf1(CNOT7/CNOT8) ($\geq 50\%$). Of the last three variant residues, Q36H (α 2) and L37M (α 2) are able to bind with hCaf1(CNOT7/CNOT8), suggesting that both amino acid residues are located away from the surface of interaction. In contrast, the variant residue L94V (α 5) is close to the surface of interaction. However, it only marginally impaired the interaction with hCaf1(CNOT7/CNOT8) ($\geq 50\%$) (Figure 4-6 A and B). However, the result of the proliferation function analysis does not consistently correlate with the hCaf1(CNOT7/CNOT8) interaction results. BTG1 Q36H and BTG1 L94V are able to inhibit cell proliferation, but less effectively than the BTG1 wild type,

while BTG1 L37M does not cause a total loss of anti-proliferative activity ($p < 0.05$) (Figure 5-8). These findings suggest the presence of another factor involved in the anti-proliferation activity of BTG1. Recently, it was discovered that BTG1 and BTG2 interact with PABPC1 through the box C region (Stupfler et al., 2016). Moreover, BTG2 requires the PABPC1 interaction to suppress cell growth in the U2SO cell line. This was discovered through analysing mutant BTG2 (in box C region), which is unable to bind with PABPC1 but can bind hCaf1/CNOT7 (Stupfler et al., 2016). This provides strong evidence suggesting the importance of the PABPC1 interaction for BTG1 anti-proliferation activity (Stupfler et al., 2016).

8.6 Do the BTG1 variants stimulate mRNA abundance via binding with human Caf1 (CNOT7/CNOT8)?

The hCaf1 (CNOT7/CNOT8) interaction via the conserved BTG domain allows for the recruitment of BTG1 to the 3'UTR of mRNA. Results have shown a reduction in luciferase activity upon the tethering of BTG1 to reporter mRNA, which is due to translational repression and/or mRNA degradation; this is consistent with data from experiments upon BTG2, TOB1 and TOB2 (Doidge et al., 2012a, Ezzeddine et al., 2012). *In vitro*, the trimeric complex BTG2-Caf1-Ccr4 strongly stimulates deadenylation activity, compared with an absence of BTG2 (Maryati et al., 2015). Moreover, analysis of BTG2 W105 and TOB1 (W93, L32, L35 and L36) showed that these residues are required for hCaf1(CNOT7/CNOT8) interaction, and stimulate mRNA degradation by recruiting hCaf1(CNOT7/CNOT8) to the target mRNA (Doidge et al., 2012, Ezzeddine et al., 2012). In agreement with this, BTG1 variants which are able to bind hCaf1(CNOT7/CNOT8) such as $\Delta 10N$ BTG1, H2Y, Q36H, L37M, E59D and L94V, are able to stimulate mRNA degradation (Figure 6-5 and Figure 6-6). Conversely, three variants (F40C, L104H and G66V) totally abrogated the interaction with hCaf1(CNOT7/CNOT8) and consequently, this negatively impacted their ability to induce translational repression and/or mRNA

degradation (Figure 6-4 to 6-6). However, the remaining seven variants (M11I, F25C, R27H, P58L, N73K, I115V and E117D) stimulated mRNA decay with the same effectiveness as the BTG1 wild type, although they were unable to bind with hCaf1(CNOT7/CNOT8) (Table 8-1). Taking these results together, implies a different perspective: that BTG1 induces mRNA degradation independently of hCaf1(CNOT7/CNOT8) binding. During the course of this work, the BTG2 box C mutant which was unable to bind with PABPC1, was co-expressed in HEK293 Tet-Off cells in conjunction with the β -globin reporter. The length of the poly (A) tails remained unchanged in the absence of BTG2 or with mutant BTG2, compared with BTG2 wild type, suggesting hCaf1(CNOT7/CNOT8) deadenylase activity requires the BTG2-PABPC1 complex (Stupfler et al., 2016). For this reason, there is a need to understand the effects of BTG1 variants on the interaction with PABPC1. Yeast two-hybrid, Co-IP and pulldown were attempted, but none proved successful.

BTG1 acts as tumour suppressor gene in B-cell malignancies. This suggests BTG1 is required to repress alternative non B-cell fates, specifically repressing the transcription factor of the T-cell lineage, however the molecular mechanism for this remains unclear (Tijchon et al., 2016). In fact, BTG1 and BTG2 are unable to regulate gene expression directly, which suggests that the PABPC1-BTG1-hCaf1(CNOT7/CNOT8) complex is required to induce mRNA degradation of target genes.

In summary, the data shows that mutations in BTG1, commonly found in DLBCL, are functionally significant and are likely to contribute to malignant transformation and tumour cell growth. The anti-proliferative activity of BTG1 variants is diminished or eliminated, which is deemed to be due to the absence of their interaction with hCaf1(CNOT7/CNOT8). However, other factors may be involved in the proliferation activity and degradation of mRNA, including PABPC1. Further investigation into this will be an exciting next step in understanding the effects of BTG1 variants on the interaction with PABPC1.

Variant residue	Location	SIFT score ≤ 0.05	Suspect score > 50	protein-protein interaction		Proliferation function $p < 0.05$		Tethering assay
				hCaf1/CNOT7	hCaf1/CNOT8	BTG1	Control	
F25C	$\alpha 1$	0.001	68	+	++	**	*	✓
R27H	$\alpha 1$	0.039	20	-	-	****	NS	✓
F40C	$\alpha 2$	0	71	-	-	****	NS	/
L104H	$\beta 2$ (box B)	0	89	-	-	**	****	/
I115V	$\beta 3$ (box B)	0.003	29	+	+	****	NS	✓
M11I	L1	0.028	87	-	-	****	NS	✓
P58L	L1(box A)	0	53	+	-	****	NS	✓
G66V	$\alpha 3$ (box A)	0.001	70	-	-	****	NS	/
N73K	L4(box A)	0.001	24	-	-	****	NS	✓
E117D	L8(box B)	0.022	14	++++	++	****	NS	✓
Q36H	$\alpha 2$	0.006	53	++++	+++	NS	****	✓
L37M	$\alpha 2$	0.02	31	+++	++	*	NS	✓
L94V	$\alpha 5$ (box B)	0.002	28	+++	+	NS	*	✓
$\Delta 10$ NBTG1	-	0	4	++++	+++	NS	***	✓
H2Y	-	0.037	17	++++	+++	NS	****	✓
E59D	L2(box A)	0.0541	2	++	++	**	NS	✓

Table 8-1 Data analysis in BTG1 variants using bioinformatics and experimental analysis

Summary of the data analysis obtained in bioinformatics and experimental analysis including effect on yeast two-hybrid interaction, proliferation analysis and tethering assay (from chapters 3 to 6). Variant residues are coloured based on their location in the BTG1 structure, green (α -helical), blue (β -sheet) and orange (L-loops). The bioinformatics software (SIFT and Suspect) were used to predict the effect of the variant on protein function and yeast two hybrid analysis was used to evaluate the effect of the BTG1 variant on the interaction with hCaf1(CNOT7/CNOT8). The relative values of β -Gal activity to BTG1 are indicated in different ranges such as 0-20 (-), 21-40 (+), 41-60 (++) , 61-80 (+++) and 81-100 (++++). The proliferation assay compared the S-phase cell numbers of BTG1 variants with BTG1 wildtype and control (empty vector). A one-way ANOVA was used to calculate the significance of the values (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$), NS=not significant. . For tethering assays with the variant residues, (/) indicates an impact upon the ability to induce mRNA degradation and (✓) indicates they are able to stimulate mRNA degradation as BTG1 wildtype does.

8.7 Future work

- **BTG1-PABPC1 interaction:** To identify the effect of the BTG1 variants on the interaction with PABPC1, as the RNA tethering assay results show that BTG1 induces mRNA degradation independently of hCaf1 (CNOT7/CNOT8) binding. Also, using the *lexA* system of yeast two hybrid, BTG1 F40C and L104H were defective in inducing mRNA degradation because they were unable to bind with PABPC1. However, this experiment was not deemed conclusive because the analysis was unable to be repeated. In the pulldown assay of the bacterial expression experiment, this was more promising because a BTG2-PABPC1 interaction was detected, although a similar interaction was not detected for BTG1 due to weak expression of GST-BTG1 in bacteria. Optimisation of the expression of GST-tagged human BTG1 cDNA may be improved by using a codon-optimised variant. Another way to encourage human BTG1 expression would be to transform a different bacterial strain, such as *E. coli Rosetta* which contains additional tRNA genes optimised for eukaryotic codon usage.
- **Analysing protein stability:** Understanding the effect of the amino acid substitutions on the stability and folding of the BTG1 structure is needed. This is particularly true for the variant residues that are located away from the hCaf1/CNOT7 interaction surface, or that show weak expression in the western blot analysis. This could be done by using differential scanning fluorimetry (DSF), a screening method that identifies low-molecular-weight ligands that bind and stabilize purified proteins. The temperature at which a protein unfolds is measured by an increase in the fluorescence of a dye with affinity for hydrophobic parts of the protein, which are exposed as the protein unfolds. Another more sensitive method for analysing the secondary structure of polypeptides and proteins is circular dichroism (CD) spectroscopy.

- **Using different cell lines:** As BTG1 function is important in B-cell lymphogenesis, it will be interesting to assess the effect of BTG1 variants in the growth and development of B cells. Although initial attempts were made to use the Raji cell line and transfection using the Neon® Transfection system, further optimisation was needed, even though the recommended voltage and settings for this cell line were used. Because of time constraints, other cell line models were used to proceed with protein function assessment.
- **Using animal models:** As the effects of BTG1 variants were assessed on protein function *in vitro*, it would be illuminating to understand whether the BTG1 variants still confer protein activity as a tumour suppressor or if they accelerate tumour formation *in vivo*. As the NHL is either a heterogeneous malignancy or multistep tumorigenesis, using transgenic mice with double- or triple-hit mutations (e.g., MYC, BCL2 and BCL6) and one BTG1 variant (e.g., F40C, P58L, G66V and L104H) will be useful to understand the effect of BTG1 variants in cancer progression,

References

References

- ALIZADEH, A. A., EISEN, M. B., DAVIS, R. E., MA, C., LOSSOS, I. S., ROSENWALD, A., BOLDRICK, J. C., SABET, H., TRAN, T. & YU, X. 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*, 403, 503-511.
- ANDERSEN, K. R., JONSTRUP, A. T., VAN, L. B. & BRODERSEN, D. E. 2009. The activity and selectivity of fission yeast Pop2p are affected by a high affinity for Zn²⁺ and Mn²⁺ in the active site. *RNA*, 15, 850-861.
- ASLAM, A., MITTAL, S., KOCH, F., ANDRAU, J.-C. & WINKLER, G. S. 2009. The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation. *Molecular biology of the cell*, 20, 3840-3850.
- BAI, Y., SALVADORE, C., CHIANG, Y.-C., COLLART, M. A., LIU, H.-Y. & DENIS, C. L. 1999. The CCR4 and CAF1 proteins of the CCR4-NOT complex are physically and functionally separated from NOT2, NOT4, and NOT5. *Molecular and Cellular Biology*, 19, 6642-6651.
- BAKKER, W. J., BLÁZQUEZ-DOMINGO, M., KOLBUS, A., BESOOYEN, J., STEINLEIN, P., BEUG, H., COFFER, P. J., LÖWENBERG, B., VON LINDERN, M. & VAN DIJK, T. B. 2004. FoxO3a regulates erythroid differentiation and induces BTG1, an activator of protein arginine methyl transferase 1. *The Journal of cell biology*, 164, 175-184.
- BASQUIN, J., ROUDKO, V. V., RODE, M., BASQUIN, C., SÉRAPHIN, B. & CONTI, E. 2012. Architecture of the nuclease module of the yeast Ccr4-not complex: the Not1-Caf1-Ccr4 interaction. *Molecular cell*, 48, 207-218.
- BAWANKAR, P., LOH, B., WOHLBOLD, L., SCHMIDT, S. & IZAURRALDE, E. 2013. NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain. *RNA biology*, 10, 228-244.
- BEDFORD, M. T. & CLARKE, S. G. 2009. Protein arginine methylation in mammals: who, what, and why. *Molecular cell*, 33, 1-13.
- BHANDARI, P., AHMAD, F. & DAS, B. R. 2017. Molecular profiling of gene copy number abnormalities in key regulatory genes in high-risk B-lineage acute lymphoblastic leukemia: frequency and their association with clinicopathological findings in Indian patients. *Medical Oncology*, 34, 92.
- BLOMBERY, P. A., WALL, M. & SEYMOUR, J. F. 2015. The molecular pathogenesis of B - cell non - Hodgkin lymphoma. *European journal of haematology*, 95, 280-293.
- BOIKO, A. D., PORTEOUS, S., RAZORENOVA, O. V., KRIVOKRYSENKO, V. I., WILLIAMS, B. R. & GUDKOV, A. V. 2006. A systematic search for

- downstream mediators of tumor suppressor function of p53 reveals a major role of BTG2 in suppression of Ras-induced transformation. *Genes & development*, 20, 236-252.
- BOLAND, A., CHEN, Y., RAISCH, T., JONAS, S., KUZUOĞLU-ÖZTÜRK, D., WOHLBOLD, L., WEICHENRIEDER, O. & IZAURRALDE, E. 2013. Structure and assembly of the NOT module of the human CCR4–NOT complex. *Nature structural & molecular biology*, 20, 1289-1297.
- BOURGE, M., FORT, C., SOLER, M. N., SATIAT - JEUNEMAÎTRE, B. & BROWN, S. C. 2015. A pulse - chase strategy combining click - EdU and photoconvertible fluorescent reporter: tracking Golgi protein dynamics during the cell cycle. *New Phytologist*, 205, 938-950.
- BRACKEN, A. P., CIRO, M., COCITO, A. & HELIN, K. 2004. E2F target genes: unraveling the biology. *Trends in biochemical sciences*, 29, 409-417.
- BRADBURY, A., POSSENTI, R., SHOOTER, E. M. & TIRONE, F. 1991. Molecular cloning of PC3, a putatively secreted protein whose mRNA is induced by nerve growth factor and depolarization. *Proceedings of the National Academy of Sciences*, 88, 3353-3357.
- BRÜCKNER, A., POLGE, C., LENTZE, N., AUERBACH, D. & SCHLATTNER, U. 2009. Yeast two-hybrid, a powerful tool for systems biology. *International journal of molecular sciences*, 10, 2763-2788.
- BUANNE, P., CORRENTE, G., MICHELI, L., PALENA, A., LAVIA, P., SPADAFORA, C., LAKSHMANA, M. K., RINALDI, A., BANFI, S. & QUARTO, M. 2000. Cloning of PC3B, a novel member of the PC3/BTG/TOB family of growth inhibitory genes, highly expressed in the olfactory epithelium. *Genomics*, 68, 253-263.
- BURMEISTER, T., MOLKENTIN, M., SCHWARTZ, S., GÖKBUGET, N., HOELZER, D., THIEL, E. & REINHARDT, R. 2013. Erroneous class switching and false VDJ recombination: molecular dissection of t (8; 14)/MYC-IGH translocations in Burkitt-type lymphoblastic leukemia/B-cell lymphoma. *Molecular oncology*, 7, 850-858.
- BUSSLINGER, M. 2004. Transcriptional control of early B cell development 1. *Annu. Rev. Immunol.*, 22, 55-79.
- BUSSON, M., CARAZO, A., SEYER, P., GRANDEMANGE, S., CASAS, F., PESSEMESSE, L., ROUAULT, J.-P., WRUTNIAK-CABELLO, C. & CABELLO, G. 2005. Coactivation of nuclear receptors and myogenic factors induces the major BTG1 influence on muscle differentiation. *Oncogene*, 24, 1698-1710.
- CAUSIER, B. & DAVIES, B. 2002. Analysing protein-protein interactions with the yeast two-hybrid system. *Plant molecular biology*, 50, 855-870.

- CHAGANTI, S. R., CHEN, W., PARSA, N., OFFIT, K., LOUIE, D. C., DALLA - FAVERA, R. & CHAGANTI, R. 1998. Involvement of BCL6 in chromosomal aberrations affecting band 3q27 in B - cell non - Hodgkin lymphoma. *Genes, Chromosomes and Cancer*, 23, 323-327.
- CHEKULAEVA, M., MATHYS, H., ZIPPRICH, J. T., ATTIG, J., COLIC, M., PARKER, R. & FILIPOWICZ, W. 2011. miRNA repression involves GW182-mediated recruitment of CCR4–NOT through conserved W-containing motifs. *Nature structural & molecular biology*, 18, 1218-1226.
- CHEN, J., CHIANG, Y. C. & DENIS, C. L. 2002. CCR4, a 3' –5' poly (A) RNA and ssDNA exonuclease, is the catalytic component of the cytoplasmic deadenylase. *The EMBO journal*, 21, 1414-1426.
- CHEN, J., RAPPSILBER, J., CHIANG, Y.-C., RUSSELL, P., MANN, M. & DENIS, C. L. 2001. Purification and characterization of the 1.0 MDa CCR4–NOT complex identifies two novel components of the complex. *Journal of molecular biology*, 314, 683-694.
- CHEN, Y., BOLAND, A., KUZUOĞLU-ÖZTÜRK, D., BAWANKAR, P., LOH, B., CHANG, C.-T., WEICHENRIEDER, O. & IZAURRALDE, E. 2014. A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. *Molecular cell*, 54, 737-750.
- CHO, J. W., KIM, J. J., PARK, S. G., LEE, D. H., LEE, S. C., KIM, H. J., PARK, B. C. & CHO, S. 2004. Identification of B - cell translocation gene 1 as a biomarker for monitoring the remission of acute myeloid leukemia. *Proteomics*, 4, 3456-3463.
- CHOI, Y., SIMS, G. E., MURPHY, S., MILLER, J. R. & CHAN, A. P. 2012. Predicting the functional effect of amino acid substitutions and indels. *PloS one*, 7, e46688.
- COLLART, M. A. 2003. Global control of gene expression in yeast by the Ccr4-Not complex. *Gene*, 313, 1-16.
- COLLART, M. A. 2016. The Ccr4 - Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdisciplinary Reviews: RNA*.
- COLLART, M. A. & PANASENKO, O. O. 2012. The Ccr4–not complex. *Gene*, 492, 42-53.
- CORTES, U., MOYRET - LALLE, C., FALETTE, N., DURIEZ, C., GHISSASSI, F. E., BARNAS, C., MOREL, A. P., HAINAUT, P., MAGAUD, J. P. & PUISIEUX, A. 2000. BTG gene expression in the p53 - dependent and - independent cellular response to DNA damage. *Molecular carcinogenesis*, 27, 57-64.

- COUPLAND, S. 2013. Molecular pathology of lymphoma. *Eye*, 27, 180-189.
- DE JONG, D. 2005. Molecular pathogenesis of follicular lymphoma: a cross talk of genetic and immunologic factors. *Journal of Clinical Oncology*, 23, 6358-6363.
- DE KEERSMAECKER, K., ATAK, Z. K., LI, N., VICENTE, C., PATCHETT, S., GIRARDI, T., GIANFELICI, V., GEERDENS, E., CLAPPIER, E. & PORCU, M. 2013. Exome sequencing identifies mutation in CNOT3 and ribosomal genes RPL5 and RPL10 in T-cell acute lymphoblastic leukemia. *Nature genetics*, 45, 186-190.
- DENIS, C. L. & CHEN, J. 2003. The CCR4–NOT complex plays diverse roles in mRNA metabolism. *Progress in nucleic acid research and molecular biology*, 73, 221-250.
- DI NOIA, J. & NEUBERGER, M. S. 2002. Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature*, 419, 43-48.
- DIERLAMM, J., BAENS, M., WLODARSKA, I., STEFANOVA-OUZOUNOVA, M., HERNANDEZ, J. M., HOSSFELD, D. K., DE WOLF-PEETERS, C., HAGEMEIJER, A., VAN DEN BERGHE, H. & MARYNEN, P. 1999. The apoptosis inhibitor gene API2 and a novel 18q gene, MLT, are recurrently rearranged in the t (11; 18)(q21; q21) associated with mucosa-associated lymphoid tissue lymphomas. *Blood*, 93, 3601-3609.
- DLAKIĆ, M. 2000. Functionally unrelated signalling proteins contain a fold similar to Mg²⁺-dependent endonucleases. *Trends in biochemical sciences*, 25, 272-273.
- DOIDGE, R., MITTAL, S., ASLAM, A. & WINKLER, G. S. 2012a. The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex. *PLoS One*, 7, e51331.
- DOIDGE, R., MITTAL, S., ASLAM, A. & WINKLER, G. S. 2012b. Deadenylation of cytoplasmic mRNA by the mammalian Ccr4–Not complex. Portland Press Limited.
- DONATO, L. J., SUH, J. H. & NOY, N. 2007. Suppression of mammary carcinoma cell growth by retinoic acid: the cell cycle control gene Btg2 is a direct target for retinoic acid receptor signaling. *Cancer research*, 67, 609-615.
- DUPRESSOIR, A., MOREL, A.-P., BARBOT, W., LOIREAU, M.-P., CORBO, L. & HEIDMANN, T. 2001. Identification of four families of yCCR4- and Mg²⁺-dependent endonuclease-related proteins in higher eukaryotes, and characterization of orthologs of yCCR4 with a conserved leucine-rich repeat essential for hCAF1/hPOP2 binding. *BMC genomics*, 2, 9.

- DURIEZ, C., FALETTE, N., AUDOYNAUD, C., MOYRET-LALLE, C., BENSAAD, K., COURTOIS, S., WANG, Q., SOUSSI, T. & PUISIEUX, A. 2002. The human BTG2/TIS21/PC3 gene: genomic structure, transcriptional regulation and evaluation as a candidate tumor suppressor gene. *Gene*, 282, 207-214.
- EZZEDDINE, N., CHANG, T.-C., ZHU, W., YAMASHITA, A., CHEN, C.-Y. A., ZHONG, Z., YAMASHITA, Y., ZHENG, D. & SHYU, A.-B. 2007. Human TOB, an antiproliferative transcription factor, is a poly (A)-binding protein-dependent positive regulator of cytoplasmic mRNA deadenylation. *Molecular and cellular biology*, 27, 7791-7801.
- EZZEDDINE, N., CHEN, C.-Y. A. & SHYU, A.-B. 2012. Evidence providing new insights into TOB-promoted deadenylation and supporting a link between TOB's deadenylation-enhancing and antiproliferative activities. *Molecular and cellular biology*, 32, 1089-1098.
- FABIAN, M. R., CIEPLAK, M. K., FRANK, F., MORITA, M., GREEN, J., SRIKUMAR, T., NAGAR, B., YAMAMOTO, T., RAUGHT, B. & DUCHAINE, T. F. 2011. miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. *Nature structural & molecular biology*, 18, 1211-1217.
- FARAJI, F., HU, Y., WU, G., GOLDBERGER, N. E., WALKER, R. C., ZHANG, J. & HUNTER, K. W. 2014. An integrated systems genetics screen reveals the transcriptional structure of inherited predisposition to metastatic disease. *Genome research*, 24, 227-240.
- FARAJI, F., HU, Y., YANG, H. H., LEE, M. P., WINKLER, G. S., HAFNER, M. & HUNTER, K. W. 2016. Post-transcriptional control of tumor cell autonomous metastatic potential by CCR4-NOT deadenylase CNOT7. *PLoS Genet*, 12, e1005820.
- FARIOLI-VECCHIOLI, S., TANORI, M., MICHELI, L., MANCUSO, M., LEONARDI, L., SARAN, A., CIOTTI, M. T., FERRETTI, E., GULINO, A. & PAZZAGLIA, S. 2007. Inhibition of medulloblastoma tumorigenesis by the antiproliferative and pro-differentiative gene PC3. *The FASEB Journal*, 21, 2215-2225.
- FICAZZOLA, M. A., FRAIMAN, M., GITLIN, J., WOO, K., MELAMED, J., RUBIN, M. A. & WALDEN, P. D. 2001. Antiproliferative B cell translocation gene 2 protein is down-regulated post-transcriptionally as an early event in prostate carcinogenesis. *Carcinogenesis*, 22, 1271-1279.
- FLETCHER, B., LIM, R., VARNUM, B., KUJUBU, D., KOSKI, R. & HERSCHMAN, H. 1991. Structure and expression of TIS21, a primary response gene induced by growth factors and tumor promoters. *Journal of Biological Chemistry*, 266, 14511-14518.

- FLICEK, P., AMODE, M. R., BARRELL, D., BEAL, K., BRENT, S., CARVALHO-SILVA, D., CLAPHAM, P., COATES, G., FAIRLEY, S. & FITZGERALD, S. 2011. Ensembl 2012. *Nucleic acids research*, gkr991.
- FORBES, S. A., BINDAL, N., BAMFORD, S., COLE, C., KOK, C. Y., BEARE, D., JIA, M., SHEPHERD, R., LEUNG, K. & MENZIES, A. 2010. COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic acids research*, gkq929.
- FRASOR, J., DANES, J. M., KOMM, B., CHANG, K. C., LYTTLE, C. R. & KATZENELLENBOGEN, B. S. 2003. Profiling of estrogen up-and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. *Endocrinology*, 144, 4562-4574.
- FUKUMURA, K., KAWAZU, M., KOJIMA, S., UENO, T., SAI, E., SODA, M., UEDA, H., YASUDA, T., YAMAGUCHI, H. & LEE, J. 2016. Genomic characterization of primary central nervous system lymphoma. *Acta neuropathologica*, 131, 865-875.
- FUNAKOSHI, Y., DOI, Y., HOSODA, N., UCHIDA, N., OSAWA, M., SHIMADA, I., TSUJIMOTO, M., SUZUKI, T., KATADA, T. & HOSHINO, S.-I. 2007. Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. *Genes & development*, 21, 3135-3148.
- GARNEAU, N. L., WILUSZ, J. & WILUSZ, C. J. 2007. The highways and byways of mRNA decay. *Nature reviews Molecular cell biology*, 8, 113-126.
- GOLDSTROHM, A. C. & WICKENS, M. 2008. Multifunctional deadenylase complexes diversify mRNA control. *Nature reviews Molecular cell biology*, 9, 337-344.
- GOLEMIS, E. A., SEREBRIISKII, I., FINLEY, R. L., KOLONIN, M. G., GYURIS, J. & BRENT, R. 2011. Interaction trap/two - hybrid system to identify interacting proteins. *Current protocols in cell biology*, 17.3. 1-17.3. 42.
- GOUVEIA, G. R., SIQUEIRA, S. A. C. & PEREIRA, J. 2012. Pathophysiology and molecular aspects of diffuse large B-cell lymphoma. *Revista brasileira de hematologia e hemoterapia*, 34, 447-451.
- GUARDAVACCARO, D., CORRENTE, G., COVONE, F., MICHELI, L., D'AGNANO, I., STARACE, G., CARUSO, M. & TIRONE, F. 2000. Arrest of G1-S progression by the p53-inducible gene PC3 is Rb dependent and relies on the inhibition of cyclin D1 transcription. *Molecular and cellular biology*, 20, 1797-1815.
- GUEHENNEUX, F., DURET, L., CALLANAN, M., BOUHAS, R., HAYETTE, S., BERTHET, C., SAMARUT, C., RIMOKH, R., BIROT, A. & WANG, Q. 1997. Cloning of the mouse BTG3 gene and definition of a new gene family

- (the BTG family) involved in the negative control of the cell cycle. *Leukemia*, 11, 370-375.
- GUO, M., MAO, X. & DING, X. 2016. Molecular genetics related to non-Hodgkin lymphoma. *Open Life Sciences*, 11, 86-90.
- HATA, K., NISHIJIMA, K. & MIZUGUCHI, J. 2007. Role for Btg1 and Btg2 in growth arrest of WEHI-231 cells through arginine methylation following membrane immunoglobulin engagement. *Experimental cell research*, 313, 2356-2366.
- HEERY, D. M., KALKHOVEN, E., HOARE, S. & PARKER, M. G. 1997. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature*, 387, 733.
- HELMS, M. W., KEMMING, D., CONTAG, C. H., POSPISIL, H., BARTKOWIAK, K., WANG, A., CHANG, S.-Y., BUERGER, H. & BRANDT, B. H. 2009. TOB1 is regulated by EGF-dependent HER2 and EGFR signaling, is highly phosphorylated, and indicates poor prognosis in node-negative breast cancer. *Cancer research*, 69, 5049-5056.
- HORIUCHI, M., TAKEUCHI, K., NODA, N., MUROYA, N., SUZUKI, T., NAKAMURA, T., KAWAMURA-TSUZUKU, J., TAKAHASI, K., YAMAMOTO, T. & INAGAKI, F. 2009. Structural basis for the antiproliferative activity of the Tob-hCaf1 complex. *Journal of Biological Chemistry*, 284, 13244-13255.
- HOSODA, N., FUNAKOSHI, Y., HIRASAWA, M., YAMAGISHI, R., ASANO, Y., MIYAGAWA, R., OGAMI, K., TSUJIMOTO, M. & HOSHINO, S. I. 2011. Anti - proliferative protein Tob negatively regulates CPEB3 target by recruiting Caf1 deadenylase. *The EMBO journal*, 30, 1311-1323.
- HOUSELEY, J. & TOLLERVEY, D. 2009. The many pathways of RNA degradation. *Cell*, 136, 763-776.
- IKEMATSU, N., YOSHIDA, Y., KAWAMURA-TSUZUKU, J., OHSUGI, M., ONDA, M., HIRAI, M., FUJIMOTO, J. & YAMAMOTO, T. 1999. Tob2, a novel anti-proliferative Tob/BTG1 family member, associates with a component of the CCR4 transcriptional regulatory complex capable of binding cyclin-dependent kinases. *Oncogene*, 18.
- IMRAN, M. & LIM, I. K. 2013. Regulation of Btg2/TIS21/PC3 expression via reactive oxygen species-protein kinase C-NFκB pathway under stress conditions. *Cellular signalling*, 25, 2400-2412.
- ITO, K., INOUE, T., YOKOYAMA, K., MORITA, M., SUZUKI, T. & YAMAMOTO, T. 2011a. CNOT2 depletion disrupts and inhibits the CCR4-NOT deadenylase complex and induces apoptotic cell death. *Genes to Cells*, 16, 368-379.

- ITO, K., TAKAHASHI, A., MORITA, M., SUZUKI, T. & YAMAMOTO, T. 2011b. The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability. *Protein & cell*, 2, 755-763.
- ITO, Y., SUZUKI, T., YOSHIDA, H., TOMODA, C., URUNO, T., TAKAMURA, Y., MIYA, A., KOBAYASHI, K., MATSUZUKA, F. & KUMA, K. 2005. Phosphorylation and inactivation of Tob contributes to the progression of papillary carcinoma of the thyroid. *Cancer letters*, 220, 237-242.
- IWANAGA, K., SUEOKA, N., SATO, A., SAKURAGI, T., SAKAO, Y., TOMINAGA, M., SUZUKI, T., YOSHIDA, Y., JUNKO, K. & YAMAMOTO, T. 2003. Alteration of expression or phosphorylation status of tob, a novel tumor suppressor gene product, is an early event in lung cancer. *Cancer letters*, 202, 71-79.
- JÄGER, U., BÖCSKÖR, S., LE, T., MITTERBAUER, G., BOLZ, I., CHOTT, A., KNEBA, M., MANNHALTER, C. & NADEL, B. 2000. Follicular lymphomas' BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of t (14; 18) translocation. *Blood*, 95, 3520-3529.
- JIA, S. & MENG, A. 2007. Tob genes in development and homeostasis. *Developmental Dynamics*, 236, 913-921.
- JOHNSON, P., WATT, S., BETTS, D., DAVIES, D., JORDAN, S., NORTON, A. & LISTER, T. 1993. Isolated follicular lymphoma cells are resistant to apoptosis and can be grown in vitro in the CD40/stromal cell system. *Blood*, 82, 1848-1857.
- JONSTRUP, A. T., ANDERSEN, K. R., VAN, L. B. & BRODERSEN, D. E. 2007. The 1.4-Å crystal structure of the *S. pombe* Pop2p deadenylase subunit unveils the configuration of an active enzyme. *Nucleic acids research*, 35, 3153-3164.
- KANDA, M., OYA, H., NOMOTO, S., TAKAMI, H., SHIMIZU, D., HASHIMOTO, R., SUEOKA, S., KOBAYASHI, D., TANAKA, C. & YAMADA, S. 2015. Diversity of clinical implication of B-cell translocation gene 1 expression by histopathologic and anatomic subtypes of gastric cancer. *Digestive diseases and sciences*, 60, 1256-1264.
- KARMAKAR, S., FOSTER, E. A. & SMITH, C. L. 2009. Estradiol downregulation of the tumor suppressor gene BTG2 requires estrogen receptor - α and the REA corepressor. *International journal of cancer*, 124, 1841-1851.
- KAWAKUBO, H., BRACHTEL, E., HAYASHIDA, T., YEO, G., KISH, J., MUZIKANSKY, A., WALDEN, P. D. & MAHESWARAN, S. 2006. Loss of B-cell translocation gene-2 in estrogen receptor-positive breast

- carcinoma is associated with tumor grade and overexpression of cyclin d1 protein. *Cancer research*, 66, 7075-7082.
- KAWAKUBO, H., CAREY, J. L., BRACHTEL, E., GUPTA, V., GREEN, J. E., WALDEN, P. D. & MAHESWARAN, S. 2004. Expression of the NF- κ B-responsive gene BTG2 is aberrantly regulated in breast cancer. *Oncogene*, 23, 8310-8319.
- KELLER, C. E., NANDULA, S., VAKIANI, E., ALOBEID, B., MURTY, V. V. & BHAGAT, G. 2006. Intrachromosomal rearrangement of chromosome 3q27: an under recognized mechanism of BCL6 translocation in B-cell non-Hodgkin lymphoma. *Human pathology*, 37, 1093-1099.
- KELLEY, L. A., MEZULIS, S., YATES, C. M., WASS, M. N. & STERNBERG, M. J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, 10, 845-858.
- KHODABAKHSHI, A. H., MORIN, R. D., FEJES, A. P., MUNGALL, A. J., MUNGALL, K. L., BOLGER-MUNRO, M., JOHNSON, N. A., CONNORS, J. M., GASCOYNE, R. D. & MARRA, M. A. 2012. Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget*, 3, 1308-1319.
- KONSTANTINOS, K. V., PANAGIOTIS, P., ANTONIOS, V. T., AGELOS, P. & ARGIRIS, N. V. 2008. PCR-SSCP: A method for the molecular analysis of genetic diseases. *Molecular biotechnology*, 38, 155-163.
- KUÈPPERS, R. & DALLA-FAVERA, R. 2001. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene*, 20, 5580.
- KUIPER, R., SCHOENMAKERS, E., VAN REIJMERSDAL, S., HEHIR-KWA, J., VAN KESSEL, A. G., VAN LEEUWEN, F. & HOOGERBRUGGE, P. 2007. High-resolution genomic profiling of childhood ALL reveals novel recurrent genetic lesions affecting pathways involved in lymphocyte differentiation and cell cycle progression. *Leukemia*, 21, 1258-1266.
- KUMAR, P., HENIKOFF, S. & NG, P. C. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols*, 4, 1073-1081.
- KÜPPERS, R. 2003. B cells under influence: transformation of B cells by Epstein-Barr virus. *Nature Reviews Immunology*, 3, 801-812.
- KÜPPERS, R. 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nature Reviews Cancer*, 5, 251-262.
- KÜPPERS, R., KLEIN, U., HANSMANN, M.-L. & RAJEWSKY, K. 1999. Cellular origin of human B-cell lymphomas. *New England Journal of Medicine*, 341, 1520-1529.
- KUROSAKI, T., KOMETANI, K. & ISE, W. 2015. Memory B cells. *Nature Reviews Immunology*, 15, 149-159.

- LA STARZA, R., BARBA, G., DEMEYER, S., PIERINI, V., DI GIACOMO, D., GIANFELICI, V., SCHWAB, C., MATTEUCCI, C., VICENTE, C. & COOLS, J. 2016. Deletions of the long arm of chromosome 5 define subgroups of T-cell acute lymphoblastic leukemia. *Haematologica*, haematol. 2016.143875.
- LAU, N.-C., KOLKMAN, A., VAN SCHAIK, F. M., MULDER, K. W., PIJNAPPEL, W. P., HECK, A. J. & TIMMERS, H. T. M. 2009. Human Ccr4–Not complexes contain variable deadenylase subunits. *Biochemical Journal*, 422, 443-453.
- LIM, S. & KALDIS, P. 2013. Cdks, cyclins and CKIs: roles beyond cell cycle regulation. *Development*, 140, 3079-3093.
- LIN, W.-J., GARY, J. D., YANG, M. C., CLARKE, S. & HERSCHMAN, H. R. 1996. The mammalian immediate-early TIS21 protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *Journal of Biological Chemistry*, 271, 15034-15044.
- LIU, H. Y., BADARINARAYANA, V., AUDINO, D. C., RAPPSILBER, J., MANN, M. & DENIS, C. L. 1998. The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively. *The EMBO journal*, 17, 1096-1106.
- LOHR, J. G., STOJANOV, P., LAWRENCE, M. S., AUCLAIR, D., CHAPUY, B., SOUGNEZ, C., CRUZ-GORDILLO, P., KNOECHEL, B., ASMANN, Y. W. & SLAGER, S. L. 2012. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proceedings of the National Academy of Sciences*, 109, 3879-3884.
- LOSSOS, I. S. 2005. Molecular pathogenesis of diffuse large B-cell lymphoma. *Journal of Clinical Oncology*, 23, 6351-6357.
- LOVE, C., SUN, Z., JIMA, D., LI, G., ZHANG, J., MILES, R., RICHARDS, K. L., DUNPHY, C. H., CHOI, W. W. & SRIVASTAVA, G. 2012. The genetic landscape of mutations in Burkitt lymphoma. *Nature genetics*, 44, 1321-1325.
- MA, Y., PANNICKE, U., SCHWARZ, K. & LIEBER, M. R. 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V (D) J recombination. *Cell*, 108, 781-794.
- MAJID, S., DAR, A. A., AHMAD, A. E., HIRATA, H., KAWAKAMI, K., SHAHRYARI, V., SAINI, S., TANAKA, Y., DAHIYA, A. V. & KHATRI, G. 2009. BTG3 tumor suppressor gene promoter demethylation, histone modification and cell cycle arrest by genistein in renal cancer. *Carcinogenesis*, 30, 662-670.

- MANGUS, D. A., EVANS, M. C., AGRIN, N. S., SMITH, M., GONGIDI, P. & JACOBSON, A. 2004. Positive and negative regulation of poly (A) nuclease. *Molecular and Cellular Biology*, 24, 5521-5533.
- MANIS, J. P., TIAN, M. & ALT, F. W. 2002. Mechanism and control of class-switch recombination. *Trends in immunology*, 23, 31-39.
- MARAGOZIDIS, P., KARANGELI, M., LABROU, M., DIMOULOU, G., PAPASPYROU, K., SALATAJ, E., POURNARAS, S., MATSOUKA, P., GOURGOULIANIS, K. I. & BALATSOS, N. A. 2012. Alterations of deadenylase expression in acute leukemias: evidence for poly (a)-specific ribonuclease as a potential biomarker. *Acta haematologica*, 128, 39-46.
- MARYATI, M., AIRHIHEN, B. & WINKLER, G. S. 2015. The enzyme activities of Caf1 and Ccr4 are both required for deadenylation by the human Ccr4-Not nuclease module. *Biochemical Journal*, 469, 169-176.
- MATSUDA, S., KAWAMURA-TSUZUKU, J., OHSUGI, M., YOSHIDA, M., EMI, M., NAKAMURA, Y., ONDA, M., YOSHIDA, Y., NISHIYAMA, A. & YAMAMOTO, T. 1996. Tob, a novel protein that interacts with p185erbB2, is associated with anti-proliferative activity. *Oncogene*, 12, 705-713.
- MAUXION, F., CHEN, C.-Y. A., SÉRAPHIN, B. & SHYU, A.-B. 2009. BTG/TOB factors impact deadenylases. *Trends in biochemical sciences*, 34, 640-647.
- MAUXION, F., FAUX, C. & SÉRAPHIN, B. 2008. The BTG2 protein is a general activator of mRNA deadenylation. *The EMBO Journal*, 27, 1039-1048.
- MAUXION, F., PRÈVE, B. & SÉRAPHIN, B. 2013. C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex. *RNA biology*, 10, 267-276.
- MCDONNELL, T. J., DEANE, N., PLATT, F. M., NUNEZ, G., JAEGER, U., MCKEARN, J. P. & KORSMEYER, S. J. 1989. bcl-2-immunoglobulin transgenic mice demonstrate extended B cell survival and follicular lymphoproliferation. *Cell*, 57, 79-88.
- MENG, E. C., PETTERSEN, E. F., COUCH, G. S., HUANG, C. C. & FERRIN, T. E. 2006. Tools for integrated sequence-structure analysis with UCSF Chimera. *BMC bioinformatics*, 7, 339.
- MEYER, S., TEMME, C. & WAHLE, E. 2004. Messenger RNA turnover in eukaryotes: pathways and enzymes. *Critical reviews in biochemistry and molecular biology*, 39, 197-216.
- MILES, R. R., SHAH, R. K. & FRAZER, J. K. 2016. Molecular genetics of childhood, adolescent and young adult non - Hodgkin lymphoma. *British journal of haematology*, 173, 582-596.

- MITTAL, S., ASLAM, A., DOIDGE, R., MEDICA, R. & WINKLER, G. S. 2011. The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4–Not complex contribute to the prevention of cell death and senescence. *Molecular biology of the cell*, 22, 748-758.
- MONTAGNOLI, A., GUARDAVACCARO, D., STARACE, G. & TIRONE, F. 1996. Overexpression of the nerve growth factor-inducible PC3 immediate early gene is associated with growth inhibition. *Cell growth & differentiation: the molecular biology journal of the American Association for Cancer Research*, 7, 1327-1336.
- MORIN, R. D., BAINBRIDGE, M., FEJES, A., HIRST, M., KRZYWINSKI, M., PUGH, T. J., MCDONALD, H., VARHOL, R., JONES, S. J. & MARRA, M. A. 2008. Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *Biotechniques*, 45, 81.
- MORIN, R. D. & GASCOYNE, R. D. Newly identified mechanisms in B-cell non-Hodgkin lymphomas uncovered by next-generation sequencing. *Seminars in hematology*, 2013. Elsevier, 303-313.
- MORIN, R. D., MENDEZ-LAGO, M., MUNGALL, A. J., GOYA, R., MUNGALL, K. L., CORBETT, R. D., JOHNSON, N. A., SEVERSON, T. M., CHIU, R. & FIELD, M. 2011. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature*, 476, 298-303.
- MORITA, M., OIKE, Y., NAGASHIMA, T., KADOMATSU, T., TABATA, M., SUZUKI, T., NAKAMURA, T., YOSHIDA, N., OKADA, M. & YAMAMOTO, T. 2011. Obesity resistance and increased hepatic expression of catabolism - related mRNAs in Cnot3+/- mice. *The EMBO journal*, 30, 4678-4691.
- MORITA, M., SUZUKI, T., NAKAMURA, T., YOKOYAMA, K., MIYASAKA, T. & YAMAMOTO, T. 2007. Depletion of mammalian CCR4b deadenylase triggers elevation of the p27Kip1 mRNA level and impairs cell growth. *Molecular and cellular biology*, 27, 4980-4990.
- MULLIGHAN, C. G., GOORHA, S., RADTKE, I., MILLER, C. B., COUSTAN-SMITH, E., DALTON, J. D., GIRTMAN, K., MATHEW, S., MA, J. & POUNDS, S. B. 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*, 446, 758-764.
- MÜSCHEN, M., RAJEWSKY, K., KRÖNKE, M. & KÜPPERS, R. 2002. The origin of CD95-gene mutations in B-cell lymphoma. *Trends in immunology*, 23, 75-80.
- MÜSCHEN, M., RE, D., JUNGNICHEL, B., DIEHL, V., RAJEWSKY, K. & KÜPPERS, R. 2000. Somatic mutation of the CD95 gene in human B cells as a side-effect of the germinal center reaction. *Journal of Experimental Medicine*, 192, 1833-1840.
- NAHTA, R., YUAN, L. X., FITERMAN, D. J., ZHANG, L., SYMMANS, W. F., UENO, N. T. & ESTEVA, F. J. 2006. B cell translocation gene 1 contributes to

- antisense Bcl-2-mediated apoptosis in breast cancer cells. *Molecular cancer therapeutics*, 5, 1593-1601.
- NASERTORABI, F., BATISSE, C., DIEPHOLZ, M., SUCK, D. & BÖTTCHER, B. 2011. Insights into the structure of the CCR4 - NOT complex by electron microscopy. *FEBS letters*, 585, 2182-2186.
- NECHANITZKY, R., AKBAS, D., SCHERER, S., GYÖRY, I., HOYLER, T., RAMAMOORTHY, S., DIEFENBACH, A. & GROSSCHEDL, R. 2013. Transcription factor EBF1 is essential for the maintenance of B cell identity and prevention of alternative fates in committed cells. *Nature immunology*, 14, 867-875.
- NGO, V. N., YOUNG, R. M., SCHMITZ, R., JHAVAR, S., XIAO, W., LIM, K.-H., KOHLHAMMER, H., XU, W., YANG, Y. & ZHAO, H. 2011. Oncogenically active MYD88 mutations in human lymphoma. *Nature*, 470, 115-119.
- NOGAI, H., DÖRKEN, B. & LENZ, G. 2011. Pathogenesis of non-Hodgkin's lymphoma. *Journal of clinical oncology*, 29, 1803-1811.
- OU, Y. H., CHUNG, P. H., HSU, F. F., SUN, T. P., CHANG, W. Y. & SHIEH, S. Y. 2007. The candidate tumor suppressor BTG3 is a transcriptional target of p53 that inhibits E2F1. *The EMBO journal*, 26, 3968-3980.
- PARK, S., LEE, Y. J., LEE, H.-J., SEKI, T., HONG, K.-H., PARK, J., BEPPU, H., LIM, I. K., YOON, J.-W. & LI, E. 2004. B-cell translocation gene 2 (Btg2) regulates vertebral patterning by modulating bone morphogenetic protein/smud signaling. *Molecular and cellular biology*, 24, 10256-10262.
- PARKER, R. & SONG, H. 2004. The enzymes and control of eukaryotic mRNA turnover. *Nature structural & molecular biology*, 11, 121-127.
- PASQUALUCCI, L., COMPAGNO, M., HOULDSWORTH, J., MONTI, S., GRUNN, A., NANDULA, S. V., ASTER, J. C., MURTY, V. V., SHIPP, M. A. & DALLA-FAVERA, R. 2006. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *Journal of Experimental Medicine*, 203, 311-317.
- PASQUALUCCI, L., MIGLIAZZA, A., FRACCHIOLLA, N., WILLIAM, C., NERI, A., BALDINI, L., CHAGANTI, R., KLEIN, U., KÜPPERS, R. & RAJEWSKY, K. 1998. BCL-6 mutations in normal germinal center B cells: evidence of somatic hypermutation acting outside Ig loci. *Proceedings of the National Academy of Sciences*, 95, 11816-11821.
- PASQUALUCCI, L., NEUMEISTER, P., GOOSSENS, T., NANJANGUD, G., CHAGANTI, R., KÜPPERS, R. & DALLA-FAVERA, R. 2001. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature*, 412, 341-346.

- PASSERI, D., MARCUCCI, A., RIZZO, G., BILLI, M., PANIGADA, M., LEONARDI, L., TIRONE, F. & GRIGNANI, F. 2006. Btg2 enhances retinoic acid-induced differentiation by modulating histone H4 methylation and acetylation. *Molecular and cellular biology*, 26, 5023-5032.
- PATEL, S., MASON, C. C., GLENN, M. J., PAXTON, C. N., SOUTH, S. T., CESSNA, M. H., ASCH, J., COBAIN, E. F., BIXBY, D. L. & SMITH, L. B. 2017. Genomic analysis of adult B-ALL identifies potential markers of shorter survival. *Leukemia Research*, 56, 44-51.
- PERKINS, J. R., DIBOUN, I., DESSAILLY, B. H., LEES, J. G. & ORENGO, C. 2010. Transient protein-protein interactions: structural, functional, and network properties. *Structure*, 18, 1233-1243.
- PETIT, A.-P., WOHLBOLD, L., BAWANKAR, P., HUNTZINGER, E., SCHMIDT, S., IZAURRALDE, E. & WEICHENRIEDER, O. 2012. The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4–NOT deadenylase complex. *Nucleic acids research*, gks883.
- PHAN, R. T. & DALLA-FAVERA, R. 2004. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature*, 432, 635-639.
- PIAO, X., ZHANG, X., WU, L. & BELASCO, J. G. 2010. CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells. *Molecular and cellular biology*, 30, 1486-1494.
- PILLAI, R. S., ARTUS, C. G. & FILIPOWICZ, W. 2004. Tethering of human Ago proteins to mRNA mimics the miRNA-mediated repression of protein synthesis. *Rna*, 10, 1518-1525.
- PRÉVÔT, D., MOREL, A.-P., VOELTZEL, T., ROSTAN, M.-C., RIMOKH, R., MAGAUD, J.-P. & CORBO, L. 2001. Relationships of the Antiproliferative Proteins BTG1 and BTG2 with CAF1, the Human Homolog of a Component of the Yeast CCR4 Transcriptional Complex INVOLVEMENT IN ESTROGEN RECEPTOR α SIGNALING PATHWAY. *Journal of Biological Chemistry*, 276, 9640-9648.
- PRÉVÔT, D., VOELTZEL, T., BIROT, A.-M., MOREL, A.-P., ROSTAN, M.-C., MAGAUD, J.-P. & CORBO, L. 2000. The leukemia-associated protein Btg1 and the p53-regulated protein Btg2 interact with the homeoprotein Hoxb9 and enhance its transcriptional activation. *Journal of Biological Chemistry*, 275, 147-153.
- RAMIRO, A. R., STAVROPOULOS, P., JANKOVIC, M. & NUSSENZWEIG, M. C. 2003. Transcription enhances AID-mediated cytidine deamination by exposing single-stranded DNA on the nontemplate strand. *Nature immunology*, 4, 452-456.

- RODIER, A., MARCHAL-VICTORION, S., ROCHARD, P., CASAS, F., CASSAR-MALEK, I., ROUAULT, J.-P., MAGAUD, J.-P., MASON, D. Y., WRUTNIAK, C. & CABELLO, G. 1999. BTG1: a triiodothyronine target involved in the myogenic influence of the hormone. *Experimental cell research*, 249, 337-348.
- ROGOZIN, I. B. & KOLCHANOV, N. A. 1992. Somatic hypermutagenesis in immunoglobulin genes: II. Influence of neighbouring base sequences on mutagenesis. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1171, 11-18.
- ROSENWALD, A., WRIGHT, G., CHAN, W. C., CONNORS, J. M., CAMPO, E., FISHER, R. I., GASCOYNE, R. D., MULLER-HERMELINK, H. K., SMELAND, E. B. & GILTANNE, J. M. 2002. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *New England Journal of Medicine*, 346, 1937-1947.
- ROUAULT, J.-P., PRÉVÔT, D., BERTHET, C., BIROT, A.-M., BILLAUD, M., MAGAUD, J.-P. & CORBO, L. 1998. Interaction of BTG1 and p53-regulated BTG2 Gene Products with mCaf1, the Murine Homolog of a Component of the Yeast CCR4 Transcriptional Regulatory Complex. *Journal of Biological Chemistry*, 273, 22563-22569.
- ROUAULT, J., FALETTE, N., GUÉHENNEUX, F., GUILLOT, C., RIMOKH, R., WANG, Q., BERTHET, C., MOYRET-LALLE, C., SAVATIER, P. & PAIN, B. 1996. Identification of BTG2, an antiproliferative p53-dependent component of the DNA damage cellular response pathway. *Nature genetics*, 14, 482.
- ROUAULT, J., RIMOKH, R., TESSA, C., PARANHOS, G., FFRENCH, M., DURET, L., GAROCCIO, M., GERMAIN, D., SAMARUT, J. & MAGAUD, J. 1992. BTG1, a member of a new family of antiproliferative genes. *The EMBO journal*, 11, 1663.
- SANDLER, H., KRETH, J., TIMMERS, H. T. M. & STOECKLIN, G. 2011. Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraprolin. *Nucleic acids research*, gkr011.
- SASAJIMA, H., NAKAGAWA, K. & YOKOSAWA, H. 2002. Antiproliferative proteins of the BTG/Tob family are degraded by the ubiquitin - proteasome system. *The FEBS Journal*, 269, 3596-3604.
- SCHEIJEN, B., BOER, J. M., TIJCHON, E., VAN INGEN SCHENAU, D., VAN EMST, L., VAN DER MEER, L. T., PIETERS, R., ESCHERICH, G., HORSTMANN, M. A. & SONNEVELD, E. 2016. Tumor suppressors BTG1 and IKZF1 cooperate during mouse leukemia development and increase relapse risk in B-cell precursor acute lymphoblastic leukemia patients. *Haematologica*, haematol. 2016.153023.

- SCHNEIDER, C., PASQUALUCCI, L. & DALLA-FAVERA, R. Molecular pathogenesis of diffuse large B-cell lymphoma. *Seminars in diagnostic pathology*, 2011. Elsevier, 167-177.
- SEIFERT, M., SCHOLTYSIK, R. & KÜPPERS, R. 2013. Origin and pathogenesis of B cell lymphomas. *Lymphoma: Methods and Protocols*, 1-25.
- SHENG, S., ZHAO, C. & SUN, G. 2014. BTG1 expression correlates with the pathogenesis and progression of breast carcinomas. *Tumor Biology*, 35, 3317-3326.
- SHIRAI, Y.-T., SUZUKI, T., MORITA, M., TAKAHASHI, A. & YAMAMOTO, T. 2014. Multifunctional roles of the mammalian CCR4–NOT complex in physiological phenomena. *Frontiers in genetics*, 5, 286.
- SOMASUNDARAM, R., PRASAD, M. A., UNGERBÄCK, J. & SIGVARDSSON, M. 2015. Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia. *Blood*, 126, 144-152.
- STRUCKMANN, K., SCHRAML, P., SIMON, R., ELMENHORST, K., MIRLACHER, M., KONONEN, J. & MOCH, H. 2004. Impaired expression of the cell cycle regulator BTG2 is common in clear cell renal cell carcinoma. *Cancer research*, 64, 1632-1638.
- STUPFLER, B., BIRCK, C., SÉRAPHIN, B. & MAUXION, F. 2016. BTG2 bridges PABPC1 RNA-binding domains and CAF1 deadenylase to control cell proliferation. *Nature communications*, 7.
- SUK, K., SIPES, D. & ERICKSON, K. L. 1997. Enhancement of B - cell translocation gene - 1 expression by prostaglandin E2 in macrophages and the relationship to proliferation. *Immunology*, 91, 121-129.
- SUN, G., LU, Y., CHENG, Y. & HU, W. 2014. The expression of BTG1 is downregulated in NSCLC and possibly associated with tumor metastasis. *Tumor Biology*, 35, 2949-2957.
- SUZUKI, A., SABA, R., MIYOSHI, K., MORITA, Y. & SAGA, Y. 2012. Interaction between NANOS2 and the CCR4-NOT deadenylation complex is essential for male germ cell development in mouse. *PLoS One*, 7, e33558.
- SUZUKI, T., JUNKO, K., AJIMA, R., NAKAMURA, T., YOSHIDA, Y. & YAMAMOTO, T. 2002. Phosphorylation of three regulatory serines of Tob by Erk1 and Erk2 is required for Ras-mediated cell proliferation and transformation. *Genes & development*, 16, 1356-1370.
- SWERDLOW, S., CAMPO, E., HARRIS, N., JAFFE, E., PILERI, S., STEIN, H., THIELE, J. & VARDIMAN, J. 2008. WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues 4th Ed.(2008).

- SWERDLOW, S. H., CAMPO, E., PILERI, S. A., HARRIS, N. L., STEIN, H., SIEBERT, R., ADVANI, R., GHIELMINI, M., SALLES, G. A. & ZELENETZ, A. D. 2016. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*, 127, 2375-2390.
- TAM, W., GOMEZ, M., CHADBURN, A., LEE, J. W., CHAN, W. C. & KNOWLES, D. M. 2006. Mutational analysis of PRDM1 indicates a tumor-suppressor role in diffuse large B-cell lymphomas. *Blood*, 107, 4090-4100.
- THORE, S., MAUXION, F., SÉRAPHIN, B. & SUCK, D. 2003. X - ray structure and activity of the yeast Pop2 protein: a nuclease subunit of the mRNA deadenylase complex. *EMBO reports*, 4, 1150-1155.
- TIJCHON, E., VAN EMST, L., YUNIATI, L., VAN INGEN SCHENAU, D., HAVINGA, J., ROUAULT, J.-P., HOOGERBRUGGE, P. M., VAN LEEUWEN, F. N. & SCHEIJEN, B. 2016. Tumor suppressors BTG1 and BTG2 regulate early mouse B-cell development. *Haematologica*, haematol. 2015.139675.
- TIJCHON, E., VAN INGEN SCHENAU, D., VAN OPZEELAND, F., TIRONE, F., HOOGERBRUGGE, P. M., VAN LEEUWEN, F. N. & SCHEIJEN, B. 2015. Targeted deletion of Btg1 and Btg2 results in homeotic transformation of the axial skeleton. *PloS one*, 10, e0131481.
- TIRONE, F. 2001. The gene PC3TIS21/BTG2, prototype member of the PC3/BTG/TOB family: regulator in control of cell growth, differentiation, and DNA repair? *Journal of cellular physiology*, 187, 155-165.
- TOYOTA, M., SUZUKI, H., SASAKI, Y., MARUYAMA, R., IMAI, K., SHINOMURA, Y. & TOKINO, T. 2008. Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. *Cancer research*, 68, 4123-4132.
- TSUJIMOTO, Y., GORHAM, J., COSSMAN, J., JAFFE, E. & CROCE, C. M. 1985. The t (14; 18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. *Science*, 229, 1390-1394.
- TSUZUKI, S., KARNAN, S., HORIBE, K., MATSUMOTO, K., KATO, K., INUKAI, T., GOI, K., SUGITA, K., NAKAZAWA, S. & KASUGAI, Y. 2007. Genetic abnormalities involved in t (12; 21) TEL-AML1 acute lymphoblastic leukemia: Analysis by means of array - based comparative genomic hybridization. *Cancer Science*, 98, 698-706.
- TUCKER, M., VALENCIA-SANCHEZ, M. A., STAPLES, R. R., CHEN, J., DENIS, C. L. & PARKER, R. 2001. The transcription factor associated Ccr4 and Caf1 proteins are components of the major cytoplasmic mRNA deadenylase in *Saccharomyces cerevisiae*. *Cell*, 104, 377-386.
- UCHIDA, N., HOSHINO, S.-I., IMATAKA, H., SONENBERG, N. & KATADA, T. 2002. A novel role of the mammalian GSPT/eRF3 associating with

- poly (A)-binding protein in Cap/Poly (A)-dependent translation. *Journal of Biological Chemistry*, 277, 50286-50292.
- VAANDRAGER, J.-W., SCHUURING, E., ZWIKSTRA, E., DE BOER, C., KLEIVERDA, K., VAN KRIEKEN, J., KLUIN-NELEMANS, H., VAN OMMEN, G., RAAP, A. & KLUIN, P. 1996. Direct visualization of dispersed 11q13 chromosomal translocations in mantle cell lymphoma by multicolor DNA fiber fluorescence in situ hybridization. *Blood*, 88, 1177-1182.
- VAN GALEN, J. C., KUIPER, R. P., VAN EMST, L., LEVERS, M., TIJCHON, E., SCHEIJEN, B., WAANDERS, E., VAN REIJMERSDAL, S. V., GILISSEN, C. & VAN KESSEL, A. G. 2010. BTG1 regulates glucocorticoid receptor autoinduction in acute lymphoblastic leukemia. *Blood*, 115, 4810-4819.
- VICENTE, C., SCHWAB, C., BROUX, M., GEERDENS, E., DEGRYSE, S., DEMEYER, S., LAHORTIGA, I., ELLIOTT, A., CHILTON, L. & LA STARZA, R. 2015. Targeted sequencing identifies associations between IL7R-JAK mutations and epigenetic modulators in T-cell acute lymphoblastic leukemia. *Haematologica*, 100, 1301-1310.
- WAANDERS, E., SCHEIJEN, B., VAN DER MEER, L. T., VAN REIJMERSDAL, S. V., VAN EMST, L., KROEZE, Y., SONNEVELD, E., HOOGERBRUGGE, P. M., VAN KESSEL, A. G. & VAN LEEUWEN, F. N. 2012. The origin and nature of tightly clustered BTG1 deletions in precursor B-cell acute lymphoblastic leukemia support a model of multiclonal evolution. *PLoS Genet*, 8, e1002533.
- WAHLE, E. & WINKLER, G. S. 2013. RNA decay machines: deadenylation by the Ccr4-Not and Pan2-Pan3 complexes. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1829, 561-570.
- WALKER, B. A., WARDELL, C. P., MELCHOR, L., HULKKI, S., POTTER, N. E., JOHNSON, D. C., FENWICK, K., KOZAREWA, I., GONZALEZ, D. & LORD, C. J. 2012. Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t (4; 14) and t (11; 14) myeloma. *Blood*, 120, 1077-1086.
- WANG, H., HUANG, Z.-Q., XIA, L., FENG, Q., ERDJUMENT-BROMAGE, H., STRAHL, B. D., BRIGGS, S. D., ALLIS, C. D., WONG, J. & TEMPST, P. 2001. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science*, 293, 853-857.
- WANG, H., MORITA, M., YANG, X., SUZUKI, T., YANG, W., WANG, J., ITO, K., WANG, Q., ZHAO, C. & BARTLAM, M. 2010. Crystal structure of the human CNOT6L nuclease domain reveals strict poly (A) substrate specificity. *The EMBO journal*, 29, 2566-2576.

- WASHIO - OIKAWA, K., NAKAMURA, T., USUI, M., YONEDA, M., EZURA, Y., ISHIKAWA, I., NAKASHIMA, K., NODA, T., YAMAMOTO, T. & NODA, M. 2007. Cnot7 - Null Mice Exhibit High Bone Mass Phenotype and Modulation of BMP Actions. *Journal of Bone and Mineral Research*, 22, 1217-1223.
- WINKLER, G. S. 2010. The mammalian anti - proliferative BTG/Tob protein family. *Journal of cellular physiology*, 222, 66-72.
- WINKLER, G. S. & BALACCO, D. L. 2013. Heterogeneity and complexity within the nuclease module of the Ccr4-Not complex. *Frontiers in genetics*, 4, 296.
- WLODARSKA, I., MECUCCI, C., STUL, M., MICHAUX, L., PITTALUGA, S., HERNANDEZ, J. M., CASSIMAN, J. J., WOLF - PEETERS, D. & VAN DEN BERGHE, H. 1995. Fluorescence in situ hybridization identifies new chromosomal changes involving 3q27 in non - Hodgkin's lymphomas with BCL6/LAZ3 rearrangement. *Genes, Chromosomes and Cancer*, 14, 1-7.
- WOLF, J. & PASSMORE, L. A. 2014. mRNA deadenylation by Pan2–Pan3. Portland Press Limited.
- WU, X. & BREWER, G. 2012. The regulation of mRNA stability in mammalian cells: 2.0. *Gene*, 500, 10-21.
- XIAO, K., YU, Z., LI, X., LI, X., TANG, K., TU, C., QI, P., LIAO, Q., CHEN, P. & ZENG, Z. 2016. Genome-wide analysis of Epstein-Barr virus (EBV) integration and strain in C666-1 and Raji cells. *Journal of Cancer*, 7, 214.
- XIE, J., WANG, Q., WANG, Q., YAO, H., WEN, L., MA, L., WU, D. & CHEN, S. 2014. High frequency of BTG1 deletions in patients with BCR-ABL1–positive acute leukemia. *Cancer genetics*, 207, 226-230.
- XU, K., BAI, Y., ZHANG, A., ZHANG, Q. & BARTLAM, M. G. 2014. Insights into the structure and architecture of the CCR4–NOT complex. *Frontiers in genetics*, 5, 137.
- YANAGIE, H., TANABE, T., SUMIMOTO, H., SUGIYAMA, H., MATSUDA, S., NONAKA, Y., OGIWARA, N., SASAKI, K., TANI, K. & TAKAMOTO, S. 2009. Tumor growth suppression by adenovirus-mediated introduction of a cell growth suppressing gene tob in a pancreatic cancer model. *Biomedicine & Pharmacotherapy*, 63, 275-286.
- YANG, X., MORITA, M., WANG, H., SUZUKI, T., YANG, W., LUO, Y., ZHAO, C., YU, Y., BARTLAM, M. & YAMAMOTO, T. 2008. Crystal structures of human BTG2 and mouse TIS21 involved in suppression of CAF1 deadenylase activity. *Nucleic acids research*, 36, 6872-6881.

- YATES, C. M., FILIPPIS, I., KELLEY, L. A. & STERNBERG, M. J. 2014. SuSPect: enhanced prediction of single amino acid variant (SAV) phenotype using network features. *Journal of molecular biology*, 426, 2692-2701.
- YE, B. H., RAO, P., CHAGANTI, R. & DALLA-FAVERA, R. 1993. Cloning of bcl-6, the locus involved in chromosome translocations affecting band 3q27 in B-cell lymphoma. *Cancer research*, 53, 2732-2735.
- YONEDA, M., SUZUKI, T., NAKAMURA, T., AJIMA, R., YOSHIDA, Y., KAKUTA, S., SUDO, K., IWAKURA, Y., SHIBUTANI, M. & MITSUMORI, K. 2009. Deficiency of antiproliferative family protein Ana correlates with development of lung adenocarcinoma. *Cancer science*, 100, 225-232.
- YOSHIDA, Y., HOSODA, E., NAKAMURA, T. & YAMAMOTO, T. 2001. Association of ANA, a member of the antiproliferative Tob family proteins, with a Caf1 component of the CCR4 transcriptional regulatory complex. *Japanese journal of cancer research*, 92, 592-596.
- YOSHIDA, Y., MATSUDA, S., IKEMATSU, N., KAWAMURA-TSUZUKU, J., INAZAWA, J., UMEMORI, H. & YAMAMOTO, T. 1998. ANA, a novel member of Tob/BTG1 family, is expressed in the ventricular zone of the developing central nervous system. *Oncogene*, 16, 2687-2693.
- YOSHIDA, Y., NAKAMURA, T., KOMODA, M., SATOH, H., SUZUKI, T., TSUZUKU, J. K., MIYASAKA, T., YOSHIDA, E. H., UMEMORI, H. & KUNISAKI, R. K. 2003a. Mice lacking a transcriptional corepressor Tob are predisposed to cancer. *Genes & development*, 17, 1201-1206.
- YOSHIDA, Y., TANAKA, S., UMEMORI, H., MINOWA, O., USUI, M., IKEMATSU, N., HOSODA, E., IMAMURA, T., KUNO, J. & YAMASHITA, T. 2000. Negative regulation of BMP/Smad signaling by Tob in osteoblasts. *Cell*, 103, 1085-1097.
- YOSHIDA, Y., VON BUBNOFF, A., IKEMATSU, N., BLITZ, I. L., TSUZUKU, J. K., YOSHIDA, E. H., UMEMORI, H., MIYAZONO, K., YAMAMOTO, T. & CHO, K. W. 2003b. Tob proteins enhance inhibitory Smad-receptor interactions to repress BMP signaling. *Mechanisms of development*, 120, 629-637.
- YU, J., LIU, P., CUI, X., SUI, Y., JI, G., GUAN, R., SUN, D., JI, W., LIU, F. & LIU, A. 2011. Identification of novel subregions of LOH in gastric cancer and analysis of the HIC1 and TOB1 tumor suppressor genes in these subregions. *Molecules and cells*, 32, 47-55.
- YU, K., HUANG, F.-T. & LIEBER, M. R. 2004. DNA substrate length and surrounding sequence affect the activation-induced deaminase activity at cytidine. *Journal of Biological Chemistry*, 279, 6496-6500.
- ZEKRI, L., KUZUOĞLU - ÖZTÜRK, D. & IZAURRALDE, E. 2013. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *The EMBO journal*, 32, 1052-1065.

- ZHANG, J., GRUBOR, V., LOVE, C. L., BANERJEE, A., RICHARDS, K. L., MIECZKOWSKI, P. A., DUNPHY, C., CHOI, W., AU, W. Y. & SRIVASTAVA, G. 2013. Genetic heterogeneity of diffuse large B-cell lymphoma. *Proceedings of the National Academy of Sciences*, 110, 1398-1403.
- ZUO, Y. & DEUTSCHER, M. P. 2001. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic acids research*, 29, 1017-1026.

Appendix

Appendix A

1. Steps of using SIFT analysis

1-Input: a list of protein sequence variants.

Each variant is represented in comma-separated (or space-separated) values as the following:

<protein ID>,<position>,<reference amino acids>,<variant amino acids>,<comment(optional)>

Protein ID: Ensembl Protein ID, NCBI RefSeq ID, or UniProt Accession ID (for human) for example in the case of BTG1 ENSP00000256015.

Position: Reference position, with the 1st amino acid having position 1.

Reference amino acids. For example (ENSP00000256015, P3S), as showing in the (Figure 1).

2-PROVEAN Protein Batch - Output Format

The results are represented in tab-separated columns as showing in the Figure below . The column headers and their meanings are shown below.

VARIATION

ROW_NO. - sequential number

INPUT - protein variant provided by the user

PROTEIN_ID - protein ID

POSITION - position of amino acid residue affected

RESIDUE_REF - reference amino acid residue

RESIDUE_ALT - variant amino acid residue

SIFT PREDICTION

SCORE - SIFT score

PREDICTION (cutoff=0.05) - tolerated or damaging

MEDIAN_INFO - median sequence information used to measure the diversity of the sequences used for prediction

#SEQ - number of sequences used for prediction.

**J. Craig Venter
INSTITUTE** **SIFT Human Protein**

JCVI Home | SIFT Home | Help | Team | Contact us

+ SIFT Home | SIFT Human Protein provides SIFT predictions for all Ensembl transcripts with an assigned ENSP r substitutions are shown.

+ Help

+ Contact us

User Input

Enter your email address if you want the results through email :
Please check that your address is correct and your mailbox is not full.

Protein Ensembl ENSP IDs
One ENSP number (starting with ENSP) per line (Limit 1000 proteins, please!).
[Sample format]

Paste in ENSP numbers, and any substitutions

```
ENSP00000256015_P3S
ENSP00000256015_2HY
ENSP00000256015_S20P
ENSP00000256015_R27H
ENSP00000256015_L37M
ENSP00000256015_Q36H
ENSP00000256015_Q38E
ENSP00000256015_L44Q
ENSP00000256015_Q45E
```

-or-

Upload file containing ENSP numbers and substitutions
 No file chosen

Page last modified: August 2011



SIFT: PREDICTIONS

User Input	ENSP	Pos	Ref	Subst	Prediction	SIFT Score	Median Information Content	# Seqs
ENSP00000256015_P3S	ENSP00000256015	3	P	S	TOLERATED	0.33	3.76	27
ENSP00000256015_2HY	ENSP00000256015	2	H	H	TOLERATED	1	3.77	26
ENSP00000256015_2HY	ENSP00000256015	2	H	Y	DAMAGING	0.05	3.77	26
ENSP00000256015_S20P	ENSP00000256015	20	S	P	DAMAGING	0.05	2.78	93
ENSP00000256015_R27H	ENSP00000256015	27	R	H	DAMAGING	0.03	2.78	92
ENSP00000256015_L37M	ENSP00000256015	37	L	M	DAMAGING	0.01	2.79	94
ENSP00000256015_Q36H	ENSP00000256015	36	Q	H	DAMAGING	0.02	2.79	94
ENSP00000256015_Q38E	ENSP00000256015	38	Q	E	TOLERATED	0.28	2.79	94
ENSP00000256015_L44Q	ENSP00000256015	44	L	Q	DAMAGING	0	2.79	95
ENSP00000256015_Q45E	ENSP00000256015	45	Q	E	TOLERATED	0.49	2.79	95

Figure 1. A sample of using SIFT analysis.

Input screen for SIFT batch tool that takes as an input a list of protein identifiers (BTG1: ENSP00000256015) with corresponding amino acid substitutions (AAs). PREDICTION (cutoff=0.05) - tolerated or damaging.

2. Generation BTG1 model structure and analysing using Suspect

The main method of Phyre2 for generating a 3D model of a protein sequence is composed of four algorithmic steps, as described below (Figure 2). BTG1 model was generated with BTG domain like protein Id: d3e9va1 (100% confidence, covering 72%), as showing the Figure 3. After the BTG1 model generated by phyre 2, then the analysed using Suspect prediction analysis (www.sbg.bio.ic.ac.uk/suspect). As in the example AAs (e.g M11I), the cut off score=50. The score >50 is indicated this mutation associated with disease, and <50 is indicated neutral as showing the Figure 4.

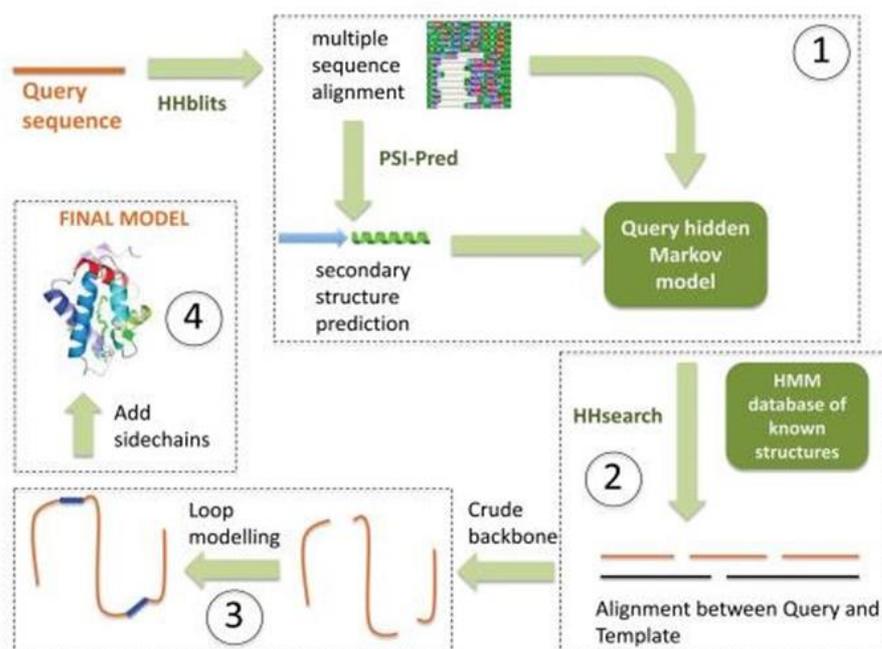


Figure 2. The method of normal mode Phyre2 for generating a 3D model of a protein sequence.

Stage 1 (gathering homologous sequences) using protein sequence database with Hhblits, the result of the multiple sequence alignment used to predict secondary structure with PSI-pred. Stage 2 (Fold library scanning): the top score used to construct crude backbone-only models. Stage 3 (loop modelling): Insertions and deletions in these models are corrected by loop modelling. Stage 4 (Side chain placement): Finally amino acid side chains are added to generate the final Phyre2 model. Adapted from Kelley et al .,2014.

www.sbg.bio.ic.ac.uk/phyre2/phyre2_output/3a0509aa8e6a2a7f/summary.html

University web site Campus News University Email Office 365 Your IT Services

Download zip of all results

Summary

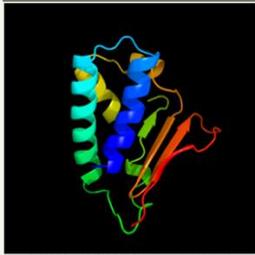


Image coloured by rainbow N → C terminus
Model dimensions (Å): X:41.125 Y:34.063 Z:47.747

Model (left) based on template **d3e9va1**

Fold:BTG domain-like
Superfamily:BTG domain-like
Family:BTG domain-like

Confidence: **100.0%** Coverage: **70%**

120 residues (70% of your sequence) have been modelled with 100.0% confidence by the single highest scoring template.

You may wish to submit your sequence to [PhyreAlarm](#). This will automatically scan your sequence every week for new potential templates as they appear in the Phyre2 library.

[Interactive 3D view in JSmol](#)
For other options to view your downloaded structure offline see the [FAQ](#)

Sequence analysis

[View PSI-Blast Pseudo-Multiple Sequence Alignment](#) [Download FASTA version](#)

Secondary structure and disorder prediction [\[Show\]](#)

Domain analysis [\[Show\]](#)

Detailed template information [\[Hide\]](#)

Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	d3e9va1 Alignment		100.0	72	Fold: BTG domain-like Superfamily: BTG domain-like Family: BTG domain-like Phyre Run Investigator

Figure 3. BTG1 model using Phyre2.

Summary results page On the left is an image of a large all-beta structure. On the right are various data regarding the model including: PDB code of the template used, information about the protein template extracted from the PDB file, confidence in the model and coverage of the query sequence (100% and 72% respectively). Using JSmol 3D viewer to visualise.

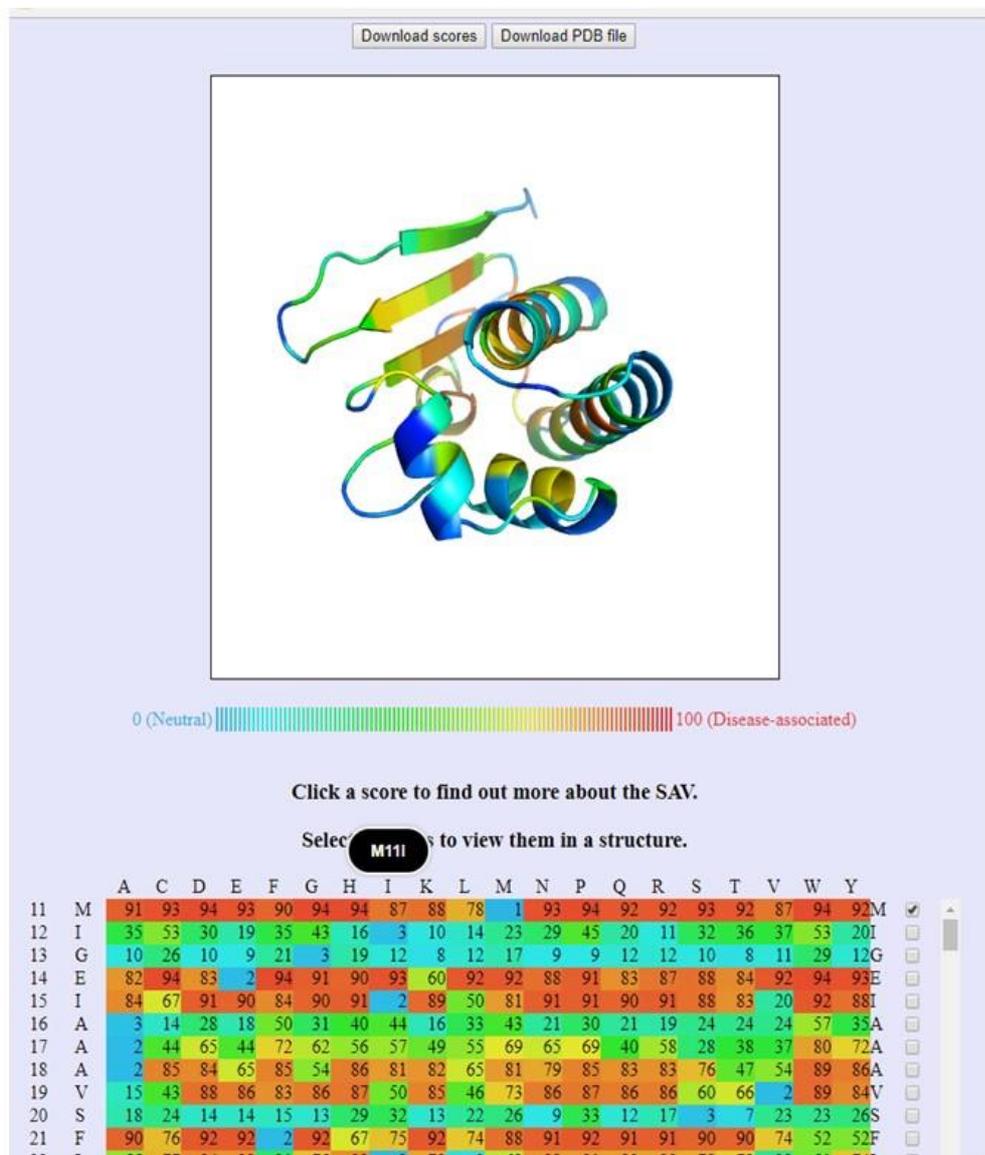


Figure 4. A sample of predicting AAs using Suspect analysis.

3. Identification of BTG1 and BTG2 variants according to type of tissues

BTG1 and BTG2 variants were collected from the Ensembl v70 genome browser and the COSMIC v65 database (Flicek et al., 2011, Forbes et al., 2010). Both resources were used in a combination with bioinformatics analysis tools such as SIFT and Suspect (Kumar et al., 2009, Yates et al., 2014). In total of 76

mutations identified, 33 were of non-specified tissue origin or in cancers such as lung, large intestinal, ovary, kidney, and others (Figure 5 A)

4. Somatic variants of the genes encoding the CCR4-NOT deadenylase complex in tissues

In this study mutations of Ccr4-Not genes were collected from the Ensembl v70 genome browser and the COSMIC v65 database, and analysed using SIFT algorithm with cut-off score=0.05.

In the deadenylase subunits a total of 141 somatic base substitutions were found in protein-coding sequences resulting in amino acid changes (non-synonymous). Of these, 66 were reported in lung, large intestine, kidney, breast, myeloma, skin and central nervous system cancers. We subsequently assessed the effect on protein function in those cases with missense mutations. In lung cancer, the CNOT7 gene carried the highest number of mutations (score ≤ 0.05) in comparison with other enzymes. Seven mutations in the coding region of deadenylase subunits were reported in large intestine (Figure 5 B). In the non-catalytic subunit, in total around 100 mutations were reported in different types of cancer. However, about 50% of these mutations remain uncharacterised or unknown. Nine out of twenty missense mutations in the CNOT2 and CNOT3 genes in lung cancer were significant and impacted protein function (Figure 5 C)

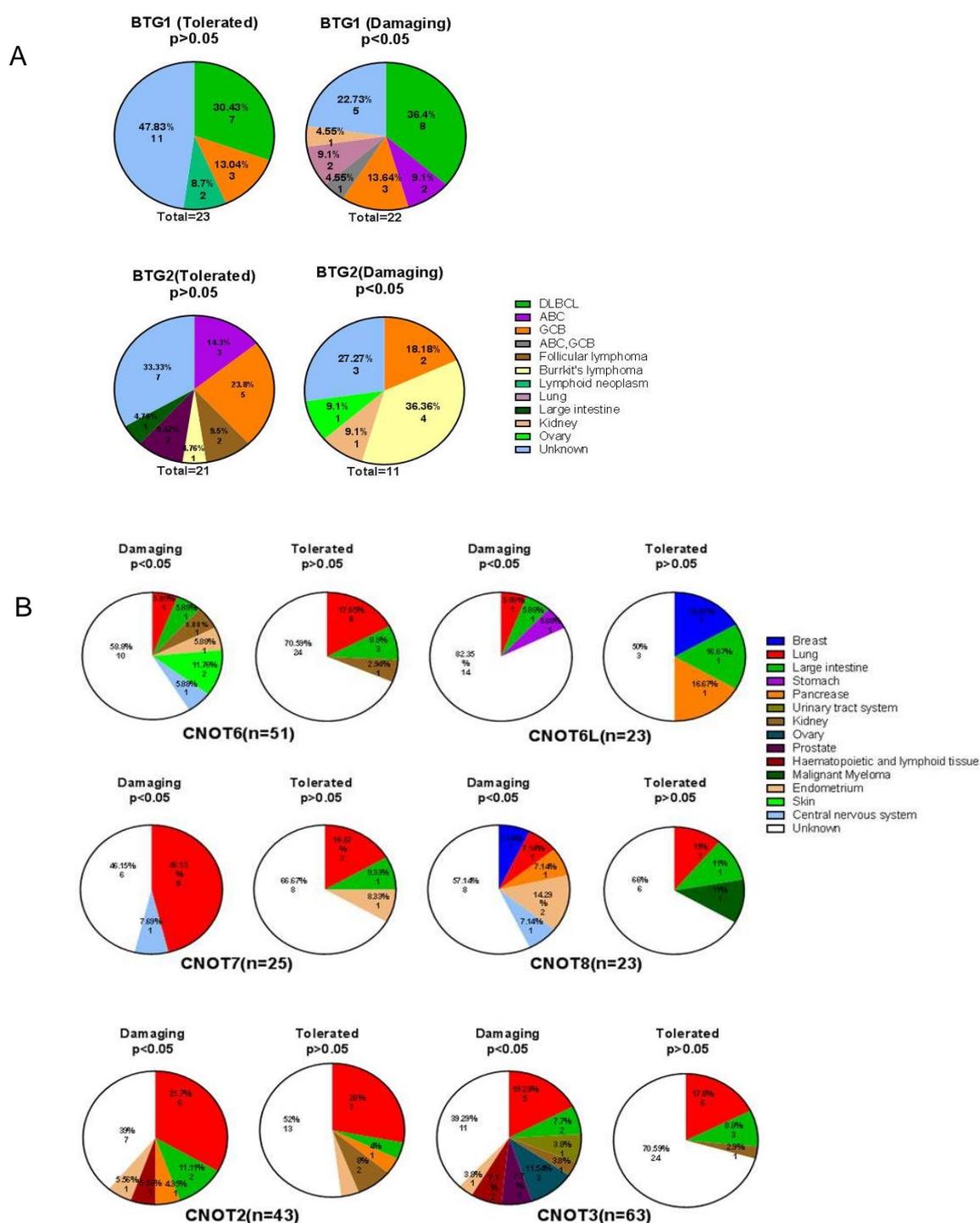


Figure 5. Number of somatic mutations in the BTG1, BTG2 and Ccr4 Not subunits in different types tissues and evaluation by SIFT.

(A) Variants of BTG1 and BTG2 genes (B) Ccr4 Not subunits genes in different types of tissue and evaluated by SIFT prediction with cut-off = 0.05 (tolerated at score > 0.05 and damaging at score \leq 0.05).

Appendix B

The HEK293T cells co-transfected with plasmid expressed GFP with HA-BTG1, assuming the cells that expressed GFP is also expressed HA-BTG1. Cells were harvested after 48 hours of transfection cells were fixed and stained with Click-iT® EdU Alexa Fluor® 594 Imaging Kit (Catalog number: C10339) as described. The click reaction quenches fluorescent proteins. Therefore, restoration of protein fluorescence was achieved by introducing acidic EDTA washes to strip the copper from these proteins which were then imaged at neutral pH. Cells were treated using the fluorescence restoration which described in (Bourge et al., 2015). Fluorescence restoration Cells were incubated for 90 min at room temperature in PMB (PBS, 10 mM MES, pH 5.0, and 30 mg ml⁻¹ BSA) with 40 mM EDTA.Na₂. In the Figure 6, showing the difference of using the EDTA and without EDTA, required for further optimization.

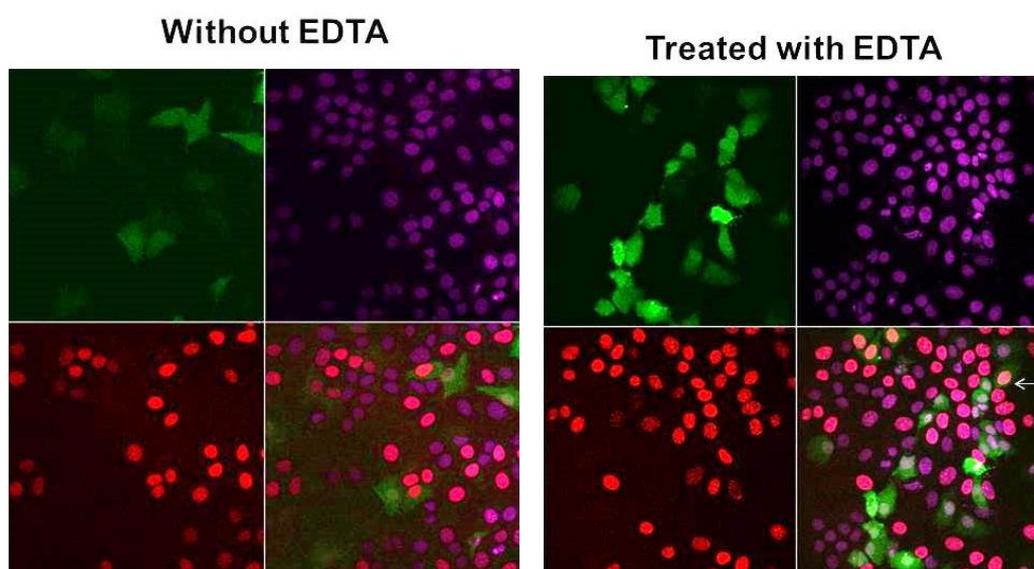


Figure 6. Efficiency of EdU restoration protocol with HEK293T cells.

Cells expressing BTG1 and GFP followed by 2 h of EdU incubation, cells were fixed, permeabilized, processed with the click reaction, restored with acidic EDTA, neutralized and imaged.

Appendix C

Primers designed for qPCR, using sequencing of renilla luciferase . The Fw and Rv primers were designed using PRIMER3. The probe was designed by Beacon designer application (Figure 7). Amplification of the cells expressed renilla luciferase using primer designed (Figure 8).

The calculation of the relative of the $2^{\Delta\Delta Ct}$ as showing below.

Calculation the relative control ratio for qPCR

- 1- Average the CT values of the housekeeping gene
- 2- Calculates the ΔCt values (CT test -CT H)
- 3- Calculates the average of the ΔCt (the reference) Control.
- 4- Calculates the $\Delta\Delta Ct = \Delta Ct$ test - average ΔCt reference.
- 5- Calculates the $2^{\Delta\Delta Ct}$

Refers the reference (control) as 100%, by calculation the average of the $2^{\Delta\Delta Ct}$ and presented as 100%.

Primers for Renilla luciferase (Co-reporter vector phRL-null, complete sequence)

```
atggettccaAGGTGTACGACCCCGAGCAACGCAAACGCATGATCACTGGGCCTCAGTGGTGGGCTCGCTGCAAGCAAATG
AACGTGCTGGACTCCTTCATCAACTACTATGATTCGAGAAGCACGCCGAGAACGCCGTGATTTTTCTGCATGGTAAAC
GCTGCCTCCAGCTACCTGTGGAGGCACGTCGTGCCTCACATCGAGCCCGTGGCTAGATGCATCATCCCTGATCTGATC
GGAatgggtaagtccggcaagaGCGGGAATGGCTCATATCGCCTCCTGGATCACTACAAGTACCTCACCGCTTGGTTCCGAGCTG
CTGAACTTCCAAAGAAAATCATCTTTGTGGGCCACGACTGGGGGGCTTGTCTGGCCTTTCACTACTCCTACGAGCAC
CAAGACAAGATCAAGGCCATCGTCCATGCTGAGAGTGTCTGTGGACGTGATCGAGTCCTGGGACGAGTGGCCTGACAT
CGAGGAGGATATCGCCCTGATCAAGAGCGAAGAGGGCGAGAAAATGGTGCTTGAGAATAACTTCTTCGTCGAGACCA
TGCTCCCAAGCAAGATCATGCGGAAACTGGAGCCTGAGGAGTTCGCTGCCTACCTGGAGCCATTCAAGGAGAAGGGC
GAGGTTAGACGGCCTACCCTCTCCTGGCCTCGCGAGATCCCTCTCGTTAAGGGAGGCAAGCCCGACGTCGTCCAGAT
TGTCGCAACTACAACGCCTACCTTCGGGCCAGCGACGATCTGCCTAAGATGTTTCATCGAGTCCGACCCTGGGTTCTT
TTCCAACGCTATTGTCGAGGGAGCTAAGAAGTTCCCTAACACCGAGTTCGTGAAGGTGAAGGGCCTCCACTTCAGCCA
GGAGGACGCTCCAGATGAAATGGGTAAGTACATCAAGAGCTTCGTGGAGCGCGTGCTGAAGAAcagcagtaa
```

RT-Leuciferase FW1: **TCGTCCATGCTGAGAGTGTCT**

RT-Leuciferase RV2: **CTAACCTCGCCCTTCTCCTT**

FAM 5'-**agtTcgCtgCctAcctgga**-3'NFQ-MGB

Produced for :		Page No.	
Beacon Designer		Project Name: DefaultProject Accession No/Name : *000001	
<pre> 567 20 586 741 19 759 767 20 786 TCCTCCATCCTCAGACTCTC agTcgCtgCctAcctgga AAGGAGAAGCCGACGTAG </pre>			
Sequence Information			
Accession No/Name	*000001	Definition	Nucleotide sequence, Length
			1,162
Reaction Conditions			
Primer Conc. μ M	200.0	Probe Conc. μ M	100.0
Monovalent Conc. mM	100.0	Free Mg ⁺⁺ Conc. mM	5.0
Temp. for Free Energy Calc. $^{\circ}$ C	25.0	Na ⁺ Conc. mM	382.84
Forward Primer - TCGTCCATGCTGAGAGTCTC			
Rating	35.9	Length bp	20
Tm $^{\circ}$ C	67.3	GC%	55
GC Clamp	1	3'End Δ G	-4.5
Self Dimer Δ G	-2.4	0 (3'End)	
Hairpin Δ G	0		
Reverse Primer - CTAACCTGCCTTCTCCTT			
Rating	39	Length bp	20
Tm $^{\circ}$ C	67.3	GC%	55
GC Clamp	2	3'End Δ G	-3.9
Self Dimer Δ G	0	0 (3'End)	
Hairpin Δ G	0		
Search Options			
Avoid Cross Homology	false	Avoid Template Sec Struct	false
Multiplexing	false		
Primer Search Parameters			
Length Range bp	18-25	Target TM Range $^{\circ}$ C	59 (+/-5)
Advance Primer Parameters			
G/C clamp - Target Consecutive G/Cs at 3' End	1	3' End Max Δ G	12
Self Dimer Max Δ G	10	10 (3'End)	
Hairpin Max Δ G	6		
Amplicon			
Pair Quality	Poor	Pair Rating	43.6
Product Len. bp	220	Product Tm $^{\circ}$ C	84.4
Product TaOpt $^{\circ}$ C	64.3		
Cross Dimer Δ G	-1	-2.4 (3'End)	
Product Parameters			

Beacon Designer, PREMIER Biosoft International. 6/22/2016

Produced for :		Page No.	
Beacon Designer		Project Name: DefaultProject Accession No/Name : *000001	
Length Range bp	70-200		
Advance Product Parameters			
Cross Dimer Max Δ G	9	9 (3'End)	
TaqMan - agTcgCtgCctAcctgga			
Type of Probe	Sense	Rating	85.9
Position	741	Length bp	19
Tm $^{\circ}$ C	76.6	GC%	57.9
Hairpin Δ G	0		
Self Dimer Δ G	0		
Probe Search Parameters			
Length Range	16-19	Target TM Range $^{\circ}$ C	77.5 (+/-5.2)
Hairpin(Internal) Δ G			6
Self Dimer(Internal) Δ G	10		
Report Comments:			

Beacon Designer, PREMIER Biosoft International. 6/22/2016

Figure 7. Taq man probe designed using Beacon designer.

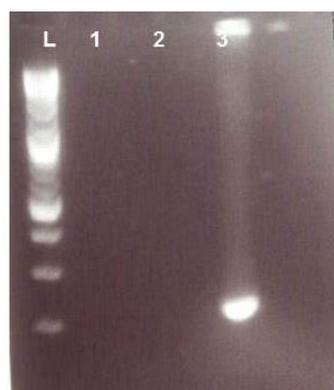


Figure 8. Amplification of the cells expressed renilla luciferase.

Agarose gel (2%) showing the amplification of renilla luciferase by PCR in HEK293T cells, and confirming the primers specificity. (1) No DNA template as negative control, (2) Cells not expressed renilla luciferase, (3) cells expressed renilla luciferase. HEK293T cells transfected with pRL-5box B, after 24 hours RNA isolated. Then, perparign the cDNA using QuantiTect® Reverse Transcription kit. (L)DNA 2-log ladder(marker) were used.