

**The BCL11B Cancer Gene Codes for an Accessory Factor
in Base Excision Repair**

Fanny Lo

Department of Biochemistry

McGill University, Montreal

August 2019

A thesis submitted to McGill University in partial fulfillment
of the requirement for the degree of Master of Science

© Fanny Lo, 2019

ABSTRACT

Increased DNA repair efficiency, in particular of the base excision repair (BER) pathway, is essential to the survival of many cancer cells. BER not only enable cancer cells to resist to genotoxic treatments, but also to proliferate in the presence of elevated reactive oxygen species (ROS) produced by cancer-associated metabolic changes. Previous work from the laboratory demonstrated that CUX1 is required for the survival of Ras-driven cancer cells because it functions as an auxiliary factor in DNA repair, stimulating the enzymatic activities of core BER enzymes, OGG1, APE1, and DNA Pol β . OGG1 is a bifunctional DNA glycosylase that initiates the repair of oxidized purines by removing the faulty base; APE1 is an endonuclease that generates single-stranded DNA breaks at apurinic/apyrimidinic (AP) sites; and DNA Pol β is responsible for synthesizing the removed base. Our laboratory had launched a project to identify the full repertoire of accessory factors for other BER enzymes, with my project focusing on NTHL1, a DNA glycosylase that initiates the repair of oxidized pyrimidines. Using the BioID approach, we curated a list of potential candidates by identifying proteins that come into close contact with NTHL1 in the cell. After a thorough literature review, we focused our investigation on one protein, B-cell lymphoma/leukemia 11B (BCL11B). The purified BCL11B protein was found *in vitro* to stimulate the enzymatic activities of NTHL1. Structure/function analysis identified distinct BCL11B regions responsible for DNA binding and the stimulation of NTHL1. Consistent with the results of *in vitro* BER assay, BCL11B knockdown caused a decrease in DNA repair and an increase in DNA damage. BCL11B has previously been characterized as a haplo-insufficient tumor suppressor gene. Paradoxically, BCL11B is overexpressed in many cancer cells and its knockdown is synthetic lethal in T cell lymphomas. BCL11B's transcriptional activities fail to explain these contradictory observations. I propose that the DNA repair

functions of BCL11B explains why inactivation of one BCL11B allele promotes tumor development, why BCL11B knockdown results in apoptosis of malignant T cells but not normal mature T cells, and why resistance to genotoxic treatments is increased by BCL11B overexpression.

RÉSUMÉ

La survie de nombreuses cellules cancéreuses dépend de mécanismes efficaces de réparation de l'ADN, en particulier la voie de réparation des bases (VRB). Cette voie de réparation permet non seulement aux cellules cancéreuses de résister aux traitements génotoxiques, mais encore de proliférer en présence d'un excès de radicaux libres qui résultent d'un métabolisme perturbé. Des études antérieures de notre laboratoire ont montré que la protéine CUX1 est nécessaire à la survie des cellules cancéreuses dont la voie de signalisation RAS est activée, parce que CUX1 fonctionne comme facteur auxiliaire qui stimule l'activité enzymatique d'enzymes VRB, notamment OGG1, une glycosylase qui enlève les bases purines oxydées dans l'ADN génomique. Notre laboratoire a dès lors démarré un projet pour identifier le répertoire complet des facteurs auxiliaires de la voie VRB. Mon projet était d'identifier les facteurs auxiliaires de NTHL1, une glycosylase qui enlève les bases pyrimidiques dans l'ADN. La technique de BioID nous a permis d'identifier un grand nombre de protéines qui viennent en contact avec l'enzyme NTHL1 dans la cellule. Après une revue exhaustive de la littérature, nous avons décidé de porter notre attention sur la protéine BCL11B (B-cell lymphoma/leukemia 11B). Des tests de réparation de l'ADN *in vitro* ont montré que la protéine BCL11B purifiée est capable de stimuler l'activité enzymatique de NTHL1. L'analyse structure/fonction a permis d'identifier deux régions distinctes de BCL11B responsables respectivement de la liaison à

l'ADN et de la stimulation de NTHL1. En accord avec les résultats obtenus *in vitro*, l'inhibition de l'expression de BCL11B dans les cellules Jurkat cause une diminution dans leur capacité à réparer l'ADN et une augmentation dans la quantité de dommages oxydatifs dans leur ADN génomique. Le gène *BCL11B* a été précédemment caractérisé génétiquement comme étant un gène suppresseur de tumeur haplo-insuffisant. De façon paradoxale, *BCL11B* est surexprimé dans plusieurs cancers et l'inhibition de son expression cause une létalité synthétique dans les cellules de lymphomes T. Ces observations contradictoires ne peuvent être expliquées par les fonctions de *BCL11B* dans la transcription. J'émetts l'hypothèse que les fonctions de *BCL11B* dans la réparation de l'ADN permettent d'expliquer pourquoi l'inactivation d'un seul allèle de *BCL11B* augmente le risque de développer un cancer, pourquoi l'inhibition de son expression cause la létalité synthétique dans les lymphomes T mais pas dans les cellules T normales, et pourquoi la surexpression de BCL11B augmente la résistance des cellules cancéreuses aux traitements génotoxiques.

ACKNOWLEDGMENTS

I would first like to thank my supervisor, Dr. Alain Nepveu, for his intellectual support and guidance throughout my master education. I have learned valuable lessons under his supervision. I am very thankful for the time he invested in my training, assisting me with studentship applications, RAC meeting presentation, the revision of this thesis, and translating the abstract included in this thesis.

I would like to thank my Research Advisory Committee members, Dr. Xiang-Jiao Yang and Dr. William Pastor, for giving me valuable suggestions on my project during my RAC meeting.

I would like to acknowledge the financial support I received from McGill Department of Biochemistry and Goodman Cancer Center.

I would like to thank the past and present members of the Nepveu laboratory who supported me during this journey: Lam Leduy, Zubaidah Ramdzan, Daniela Gilardi, Camila Faraco, Elise Vickridge, Sara Wang, Simran Kaur, and Li Li. I am extremely grateful to Lam Leduy and Zubaidah Ramdzan. They gave me immense support and patience since day one, and provided technical and intellectual support whenever needed. Without their help, the work presented in this thesis would not be possible. Their remarkable work ethic and dedication is especially inspiring to me. The findings on CUX1 published by Dr. Zubaidah Ramdzan and former members in the laboratory set the ground work for my project. Other members in the laboratory also provided assistance to my project by giving valuable suggestions during weekly lab meetings, and lending a hand whenever I needed it.

Finally, I would like to thank my parents and my sister for their unconditional love and constant support. My parents inspire me to become better versions of myself, and I am the person

I am because of them.

ACKNOWLEDGMENT OF WORK

Ramdzan, Z. M:

Establishing cell lines for BioID (Fig. 1), DNA Pol β strand-displacement assay (Fig. 8), Single cell gel electrophoresis assay (Fig 15)

Leduy, L:

Creating BirA* construct for BioID (Fig. 1), Schiff base trapping assay of NTHL1 (Fig. 5), BCL11B on and off rate experiment (Fig. 6), Calculation of K_D for BCL11B (Fig. 7), Establishing cell lines for laser micro-irradiation (Fig. 14)

Tehrani, P: (Laboratory of Dr. Anne-Claude Gingras)

Mass spectrometry analysis of BioID (Fig. 1)

Djerir, B: (Laboratory of Alexandre Maréchal)

Laser micro-irradiation and corresponding microscopy works (Fig. 14)

Table of Contents

1. Introduction	1
1.1 DNA Damage and Repair	1
1.1.1 Types of DNA damage	1
1.1.2 Base Excision Repair	5
1.1.3 Mismatch Repair	6
1.1.4 Nucleotide Excision Repair	7
1.1.5 Double Strand Break Repair	8
1.2 Oncogenes and Tumor Suppressors	12
1.2.1 Oncogene Addiction	13
1.2.2 Non-oncogene Addiction and Synthetic Lethality.....	14
1.3 CUX1	15
1.3.1 CUX1 as a Tumour Suppressor	17
1.3.2 CUX1 as an Oncogene.....	17
1.3.3 CUX1 in Base Excision Repair.....	18
1.3.4 Duality of CUX1 in Cancer	19
1.4 Project Rationale	20
2. Materials and Methods	21
Cell Culture and Viral Production	21
Bacterial Protein Expression	22
Immunoblotting	22
BioID	23
Immunoprecipitation	23
<i>In Vitro</i> Fluorescent Cleavage Assay	23
<i>In Vitro</i> Radioactive Cleavage Assay	24
Sodium Borohydride Trapping of NTHL1	24
Electrophoretic Mobility Shift Assay (EMSA)	25
On/Off Rate of DNA binding	25
Strand-Displacement Assay of DNA PolB	26
Oligonucleotide Sequences	26
GST-pull down assay	27
Single Cell Gel Electrophoresis	27
3. Results	28
BCL11B Interacts with NTHL1	28
BCL11B Stimulates Enzymatic activities of NTHL1	31
BCL11B Increases the Formation of the Schiff Base Intermediate	32
BCL11B Binds DNA with Rapid On and Off Rate	33
BCL11B Preferentially Binds to DNA with Thymine Glycol Damage	34
BCL11B stimulates Polymerase Activity of DNA Polβ	34
Residue 213-420 of BCL11B Stimulates NTHL1 Enzymatic Activity	35
C-terminal Zinc Fingers of BCL11B are Responsible for DNA Binding	37
BCL11B Recruitment to Site of DNA Damage	38
BCL11B Knockdown in Jurkat Cells Reduces DNA Repair Efficiency	38
4. Figures	40
Fig. 1: Biotin Dependent Proximity Identification Experiment of NTHL1	40

Fig. 2: Co-Immunoprecipitation of NTHL1 and BCL11B.....	42
Fig. 3: BCL11B Stimulates Enzymatic Activities of NTHL1 in Fluorescent Cleavage Assay.	44
Fig. 4: BCL11B Stimulates Enzymatic Activities of NTHL1 in Radioactive Cleavage Assay.	46
Fig. 5: BCL11B Promotes Schiff Base Formation by NTHL1.....	48
Fig. 6. BCL11B Binds DNA with Rapid On and Off Rate.	50
Fig. 7: Calculations of Apparent Equilibrium Dissociation Constant (K_{Dapp}) of BCL11B for Binding to DNA with a Normal or an Oxidized Thymine.	52
Fig. 8: BCL11B Stimulates the Polymerase and Strand-Displacement Activities of DNA Pol β.....	54
Fig. 9: Amino Acids 213-420 of BCL11B Are Sufficient to Stimulate the Enzymatic Activities of NTHL1.....	56
Fig. 10: BCL11B Recombinant Protein Fragments Stimulates Schiff Base Formation by NTHL1.	58
Fig. 11: BCL11B Recombinant Protein Fragments Interacts with NTHL1 in Pull Down Assay.	60
Fig. 12: C-terminal Zinc Fingers of BCL11B are Responsible for DNA Binding.....	62
Fig. 13: C-terminal Zinc Fingers of BCL11B Binds DNA with Rapid On and Off Rate.	64
Fig. 14: BCL11B is Rapidly Recruited to Laser Micro-Irradiation-Induced DNA Lesions.	66
Fig. 15: BCL11B Knockdown in Jurkat cells Decreases DNA Repair Efficiency.	68
5. Tables	70
Table 1: NTHL1 Preys Identified in Unirradiated and Irradiated Cells.	70
6. Discussion	79
Mechanism of BCL11B Stimulation of NTHL1.....	79
Impact of BCL11B on DNA Repair in Cells.....	82
Potential for Novel Classification of Cancer Gene.....	84
7. Conclusion	85
8. Reference	85

1. Introduction

1.1 DNA Damage and Repair

Mammals have developed efficient response to DNA insults from both external and internal origins. Exogenous sources of DNA damage include ionizing radiation (IR), ultraviolet (UV) radiation, and chemical agents^{1, 2}. The vast majority of DNA damage is from endogenous sources, including hydrolysis, deamination, alkylation, mis-incorporation of DNA by polymerase, and exposure to reactive oxygen species (ROS) and other reactive metabolites produced from cellular processes^{3, 4}.

It has been estimated that roughly 70,000 lesions occur in each human cell per day⁵. Given the large amount of daily DNA damage, it is evident that effective DNA repair pathways are essential for the maintenance of genome integrity. When damaged DNA is not repaired correctly and efficiently, it can result in genomic instability, apoptosis, and senescence. Cells have evolved an intricate network of DNA repair mechanisms to deal with the different types of DNA damage. A highly coordinated process known as the DNA damage response (DDR) is responsible to sense the DNA damage, signal its presence, and mediate its repair⁶.

DNA damage comes in different forms, such as base modifications or loss, strand breaks, cross-linkage, and mismatches⁶. Each type of DNA damage has its own dedicated DNA repair pathway. The major DNA repair pathways are base excision repair (BER), mismatch repair (MMR), nucleotide excision repair (NER), and double-strand break repair⁶.

1.1.1 Types of DNA damage

Spontaneous hydrolysis of DNA is one of the simplest form of endogenous damage⁷. The N-glycosidic bond between the deoxyribose and nucleobase is particularly susceptible to

hydrolysis, and hydrolysis leaves behind an apurinic/aprimidinic site (AP site), also known as an abasic site⁷. Abasic sites are also intermediates produced in base excision repair, when DNA glycosylases remove the damage bases⁸. In addition to abasic sites, hydrolysis of DNA can cause deamination of nitrogenous bases carrying exocyclic amino groups, such as cytosine, adenine and guanine, creating uracil, hypoxanthine and xanthine, respectively^{7, 9, 10}.

Reactive molecules produced from cellular metabolism are also sources of DNA damage. Some examples of reactive molecules are reactive carbonyl species, created from lipid peroxidation and glycation; reactive nitrogen species, important cell signaling molecules; and reactive oxygen species, produced in the mitochondria as a result of oxidative respiration^{4, 11, 12}. Among the different reactive molecules, DNA damage by reactive oxygen species occur most frequently⁴. ROS can be generated by NADPH oxidases (NOX) and dual oxidases (DUOX), as well as by mitochondria and peroxisome as a result of aerobic metabolic processes such as oxidative respiration^{13, 14}. They function as mediators for signal transduction related to growth, angiogenesis, and apoptosis¹⁴. These reactive oxygen species include superoxide ($O_2^{\cdot-}$), hydrogen peroxide (H_2O_2), and hydroxyl radicals (OH^{\cdot})¹³. Some highly reactive ROS, such as hydroxyl radicals, react non-specifically and extremely rapidly with biomolecules, causing molecular damage such as DNA mutation, lipid peroxidation and protein oxidations¹⁴. Therefore, ROS production is elevated in phagocytes, mainly neutrophils and macrophages, by phagocyte NADPH oxidases (Phox) as a part of microbicidal mechanism, making ROS an important component of host defence and innate immune response. However, presence of ROS in cells also makes proteins and DNA of the host organism susceptible to damage. ROS in cells can be converted to less reactive molecules, such as water, by antioxidant systems including catalase, peroxidase, and superoxide dismutase¹⁵. ROS can generate over a hundred different oxidative

DNA adducts, such as base modification, deoxyribose oxidation, and single- or double-strand breaks¹⁶. Aberrant levels of intracellular ROS have been attributed to diseases including cancer, neurodegenerative disorders, cardiovascular diseases, and endocrine disorders¹⁴. For example, elevated ROS level caused by overexpression of ROS producing enzymes, especially NOX1, and altered levels of antioxidant enzymes is essential for Ras-driven oncogenic transformation¹⁷⁻¹⁹. Alkylation is another type of DNA damage caused by endogenous reactive molecules. Molecules generated by lipid peroxidation including S-adenosylmethionine (SAM) and methyl radicals are methyl donors, and are capable of alkylating the O- and N- atoms of nucleobases^{4, 20}.

Error in DNA synthesis reactions is another source of endogenous damage of DNA. For example, mis-incorporation of bases during DNA replication by DNA polymerases introduces mismatches, as well as insertions and deletions of bases²¹. In addition, accidental incorporation of chemically altered nucleotide precursors, such as dUTP, can also lead to replication-related DNA damage²². Importantly, DNA polymerases involved in DNA repair and damage tolerance exhibit higher error rates, in part because they do not carry proofreading exonuclease activity, result in an increase of fidelity up to 100 fold^{23, 24}. Moreover, DNA polymerases that perform translesion synthesis have a larger catalytic pocket to enable them to accommodate various types of DNA structures²⁵. The larger catalytic pocket has lower specificity requirement of correct base-pairing and results in lower fidelity rate compare to DNA polymerases involved in DNA replication. Since the fidelity of replicating DNA polymerases is conferred by the tightness of their catalytic pocket, which can accommodate only the correct base-pairs (C-G, G-C, A-T and T-A), the larger pockets of DNA polymerases involved in repair and damage tolerance increases error rates by 100 to 1000 fold, depending on the specific enzyme^{21, 25}. As a result, mutations that

arise following DNA damage are the products or errors made by DNA polymerases involved in DNA repair, the very enzymes whose role is to maintain genomic integrity.

Topoisomerases are additional enzymes participating in DNA synthesis that can cause DNA damage. The function of topoisomerase is to relax DNA supercoils by generating transient single- or double- strand breaks, depending on the topoisomerase²⁶. Furthermore, abortive topoisomerase activity can sometimes create an irregular lesion where a covalent linkage is formed between termini of DNA strand breaks and the enzyme^{26, 27}.

In addition to endogenous sources of damage, DNA is also susceptible to damage by exogenous agents, both physical and chemical. Physical stresses are caused by exposure to electromagnetic radiations. Ultraviolet (UV) exposure from the sun can cause the formation of abnormal covalent bond between adjacent pyrimidine bases²⁸. Ionizing radiation, originated from cosmic radiation or from medical sources (X-rays and radiotherapy), directly induces double strand breaks²⁹. In aqueous systems, ionizing radiation can ionize water molecules and produce reactive oxygen species such as hydroxyl radicals³⁰. These free radicals is capable of base modifications, generation of abasic sites, and strand breaks³¹. Chemical sources of exogenous damage include alcohol, tobacco, and chemotherapeutics. Chronic exposure to ethanol in the body results in elevated production of hydrogen peroxide, and when reacted with metal ions (i.e. iron), can produce reactive oxygen species^{32, 33}. It's well-established that smoking is a risk factor for developing lung cancer, because tobacco contains many carcinogens that cause genome instability³⁴. Clinical drugs, especially those developed specifically to combat cancer, often induce DNA damage. These include mono- and bi-functional alkylators (i.e. temozolomide and cisplatin), topoisomerase inhibitors, and replication inhibitors³⁵.

1.1.2 Base Excision Repair

Base excision repair is the predominant mechanism for the repair of damaged DNA bases, and does not distort the overall DNA helix structure during the course of the repair. BER repairs most base lesions including deaminated, alkylated, and oxidized bases, as well as apurinic/aprimidinic (AP) sites⁸. The BER pathway is initiated by the excision of damaged bases by a monofunctional or bifunctional DNA glycosylase³⁶. Uracil and alkylation lesions in DNA are targeted by monofunctional glycosylases, uracil DNA glycosylase (UDG) and N-methylpurine glycosylase (MPG)³⁷. Monofunctional glycosylases hydrolyze the N-glycosylic linkage between the damaged base and its corresponding deoxyribose, a process that produces an AP site³⁶. Following the base removal, AP endonuclease 1 (APE1) will incise the DNA backbone 5' to the AP site, resulting in a single-strand break with a 3' hydroxyl group and 5' deoxyribose phosphate group (dRP)³⁶⁻³⁸. The dRP moiety at the 5'-end is subsequently removed by the dRP-lyase activity of DNA polymerase β to produce a 5'-phosphate (5'-P) that is suitable for ligation with a 3'-hydroxyl group (3'-OH)³⁹. On the other hand, repair of oxidative DNA lesions is initiated by bifunctional glycosylases that are also capable of AP/lyase activity, which generates a single-strand nick 3' to the AP site via either β - (NTHL1, OGG1) or β,δ -elimination (NEIL1, NEIL2)⁴⁰. These glycosylases replace the damaged base by nucleophilic substitution, and covalently bond to the deoxyribose of the AP site, forming the Schiff base intermediate³⁶. In general, OGG1 is the primary glycosylase for oxidized purines, whereas oxidative pyrimidine lesions are removed primarily by NTHL1, NEIL1, or NEIL2⁴⁰. The resulting single-stranded nicked DNA undergoes end-processing by APE1 and polynucleotide kinase 3'-phosphatase (PNKP) depending on the specific nature of the terminus^{41, 42}. End-processing allows DNA to become suitable substrate for DNA pol β to synthesize the removed base in the long or short

patch pathway^{43, 44}. In short-patch repair, DNA pol β adds a single base and removes the 5'-dRP to allow ligation⁴³. In long patch repair, 2 to 13 bases are synthesized by pol β or δ/ϵ , thereby generating a displaced strand that is cleaved by FEN1 prior to ligation^{43, 45}. Base excision repair is completed by DNA ligase III to seal the nick in the DNA backbone^{36, 44}.

1.1.3 Mismatch Repair

Mismatch repair (MMR) is especially important in detecting replication errors and targets mismatches due to mis-incorporation, and insertion and deletion loops (IDLs) resulting from polymerase slippage. Mismatch repair is very crucial because it prevents a mistake from becoming permanent and passed on to daughter cells. This system reduces the overall error rate by approximately 1000 folds⁴⁶. There are three major steps in mismatch repair, recognition of mismatches, excision of the mismatch on the error-containing strand, and repair the damage resulted from the excision⁴⁷.

MMR pathway is highly conserved between human and *E. coli*⁴⁷. In *E. coli*, the homodimer MutS recognizes and binds to base mismatches and insertion/deletions⁴⁷. Although the two MutS dimers are sequentially identical, they act as virtual heterodimers because they are structurally and functionally different⁴⁸. The human MutS homologues are formed using true heterodimers: hMSH2 dimerizes with hMSH6 or hMSH3 to form hMutS α or hMutS β . hMutS α recognizes mismatches of 1 to 2 nucleotides, and hMutS β recognizes larger mismatches⁴⁷. In *E. coli*, MutL interacts with MutS to recruit and activates MutH. MutH, in the presence of ATP, will nick the DNA at the unmethylated strand⁴⁷. DNA in *E. coli* is methylated at GATC sites, and MMR enzymes use hemi-methylated sites to distinguish newly synthesized DNA from the

template DNA. In human, three MutL heterodimers have been identified⁴⁹. MLH1 and PMS2 forms the heterodimer MutL α , and is responsible for majority of DNA incision⁴⁷. In *E. coli*, depending on the position of the mismatch to the site of the nick, different exonucleases will remove DNA up to slightly past the mismatch site. The resulting single-stranded gap is repaired by DNA polymerase III and DNA ligase⁵⁰. In human, after incision by MutL α , exonuclease 1 (Exo1) degrades DNA via 5' to 3' direction. Then MMR is completed by DNA polymerase δ and DNA ligase I⁴⁷.

1.1.4 Nucleotide Excision Repair

Nucleotide excision repair recognizes and removes a wide range of bulky, helix-distorting lesions, such as UV-induced photoproducts, cyclobutane pyrimidine dimers (CPD) and pyrimidine 6-4 pyrimidone photoproducts, and cisplatin-DNA intra-strand crosslinks⁵¹. NER defects threaten genome integrity, and are the underlying causes of several human genetic disorders, such as xeroderma pigmentosum and Cockayne syndrome. NER deficiencies are also implicated in cancer, immunological defects and neurodegeneration⁵².

There are two subsets of NER, global genomic nucleotide excision repair (GG-NER) and transcription-coupled nucleotide excision repair (TC-NER), with the two subsets only differ in the mechanism by which DNA lesions are recognized⁵³. GG-NER eliminates DNA lesions throughout the genome, and TC-NER, due to its method of lesion recognition, preferentially targets the coding strand of genes actively being transcribed⁵¹. In GG-NER, the XPC/HR23B/CEN2 complex detects and binds to DNA lesions, which results in further distortion of the double helix⁵¹. HR23B and CEN2 are accessory factors that increase the affinity

and specificity of XPC binding to damaged DNA. The affinity of XPC to DNA is positively correlated with the amount of DNA distortion. Therefore, in the case of CPD damage that results in minor distortions, additional auxiliary factors, such as UV-damaged DNA binding complex (UV-DDB) are needed to increase helix distortion and facilitate XPC binding⁵⁴. Contrastingly, damage recognition in TC-NER is initiated by the stalling of RNA polymerase II (RNAPII) upon encountering a damage. Next, Cockayne syndrome A and B (CSA/CSB) displace RNAPII to allow the binding of other NER proteins⁵⁵.

Following damage recognition, both GG-NER and TC-NER accomplish the rest of NER in the same manner. XPC, from GG-NER, and CSA/CSB, from TC-NER, recruit the transcription factor complex TF11H. Two TF11H-associated helicases, XPB and XPD, then flank the damage site at opposite directions, creating a bubble of unwound DNA (~30 nt) at the damaged region⁵³. Following the unwinding of DNA, XPA and RPA (replication protein A) bind to the damaged DNA. XPA acts as a secondary damage recognition and RPA stabilizes the pre-incision complex. Subsequently, two endonucleases, XPG and XPF/ERCC1, cleave DNA 3' and 5' to the damage, resulting in a 30 nucleotide-long of excised DNA strand containing the damage. Lastly, DNA polymerase δ and ϵ re-synthesize the gap, and the nicks are sealed by DNA ligases.

1.1.5 Double Strand Break Repair

Double-strand breaks (DSBs) are the most detrimental type of DNA damage. When left unrepaired, DSBs often result in cell death. In addition, inaccurate repair can lead to permanent chromosomal alterations⁵⁶. Therefore, once a DSB occur, rapid detection and repair is key to

maintain genome integrity. In mammalian cells, there are two main mechanisms to repair DSBs, homologous recombination (HR) and non-homologous end-joining (NHEJ). The two pathways differ in their dependency on homologous templates and the fidelity of the repair⁵⁷. Homologous recombination-directed repair utilizes the undamaged sister chromatid as a template for repair, and it's a generally error-free repair. Since it requires a sister chromatid template, it only occurs during late-S and G2 phases. On the other hand, non-homologous end-joining repairs DSBs by direct ligation. While it's more error-prone than HR, it is the predominant pathway for DSB repair in cells.

There are four general steps in HR, DNA resection, strand exchange, branch migration and DNA synthesis, and resolution. The purpose of DNA resection is to prepare the site of DSB to become a suitable substrate for subsequent repair. DSB ends are processed through 5' to 3' end resection to generate a single-strand overhang with a 3' end⁵⁸. The resection is facilitated by the MRN complex (Mre11-Rad50-Nbs1), CtIp, BLM helicase, and Exo1 exonuclease^{59, 60}. Then Rad51 recombinase, Rad51 paralogs, RPA, together with other mediator proteins, including BRCA1, BRCA2, and Rad52, bind and coat the 3' overhanging DNA, this is known as the Rad51 nucleoprotein filament⁶¹. The Rad51 nucleoprotein filament facilitates homology search and strand invasion of the template DNA in the sister chromatid. DNA polymerase synthesizes DNA based on the template, extending in the 3' direction of the invading strand to allow branch migration. This results in a four-way DNA junction known as the Holliday junction⁶². DNA ligase 1 seals the nicks, and the Holliday junction is resolved to complete homologous recombination repair. The Holliday junction can be resolved in one of three ways, disengagement mediated by the BLM-TopIIIa complex, symmetric cleavage by GEN1/YEN1 or

Slx1/Slx4, or asymmetric cleavage by Mus81/Eme1⁶³. Depending on the resolution method, a non-crossover or crossover product will be produced.

The first step in non-homologous end-joining is the recognition and binding of the Ku70/Ku80 heterodimer to the exposed termini of DSB. Upon binding, the Ku-DNA complex recruits the catalytic subunit of DNA-dependent protein kinase (DNA-PK), which promotes tethering of the two DNA termini⁶⁴. Depending on the nature of the double-strand break, the ends may require modification prior to ligation. In order for termini to be ligatable, there should be no overhangs, and the appropriate 5' phosphate and 3' hydroxyl groups need to be present. This processes, while not fully understood, may involve the exonuclease activity of FEN1, WRN and Artemis⁶⁵. DNA polymerases then perform any necessary DNA repair synthesis. Lastly LIG4–XRCC4–XLF complex is recruited to perform the ligation and complete NHEJ. Consequently, after ligation of the two DNA ends, there may be a few bases that were added or removed in the process. This is the reason why NHEJ is considered error-prone. Within coding sequences, addition or loss of bases will cause a frameshift two third of the time, potentially leading to production of truncated proteins with different c-terminal regions. This process is exploited in the recently developed method of CRISPR/Cas gene inactivation.

1.2 Paradoxical Role of DNA repair in Cancer

Living organisms are constantly experiencing DNA damage, and require effective repair pathway to fix the damage. When the repair pathway is overpowered by the presence of excess carcinogens, or when repair pathways are defective, it leaves the genome vulnerable to mutations. Somatic mutations in DNA repair pathways lead to accumulation of mutations and genome instability, and germline aberrations in these pathways leads to cancer predisposition⁶⁶.

⁶⁷. Although DNA repair defects contribute to cancer initiation and progression, DNA repair pathways are also needed by cancer cells for their survival³⁵. Cancer cells typically proliferate more rapidly, and many are burdened with elevated reactive oxygen species (ROS) produced by cancer-associated metabolic changes⁶⁸. In addition, classical cancer treatments such as radiotherapy and chemotherapy employ DNA-damaging agents to induce DNA lesions or interfere with DNA synthesis, and cause cell cycle arrests, apoptosis and senescence³⁵. Increased DNA repair efficiency, in particular of the base excision repair (BER) pathway, is essential for many cancer cells, not only to resist genotoxic treatments but also to proliferate in the presence of elevated ROS⁶⁸. For example, cancer cells in which the RAS pathway is activated produce elevated ROS levels¹⁷. Although ROS contribute to RAS transformation of cells to tumorigenesis, there is a fine balance of ROS level that cancer cells need to maintain so that they benefit from the mutational advantage without suffering from excessive DNA damage. Cancer cells can adapt by increasing expression of ROS scavenging proteins. This is often accomplished through inactivation of the *KEAP1* tumor suppressor gene. KEAP1 is a cytoplasmic protein that interacts with the transcription factor NRF2, resulting in NRF2 retention in the cytoplasm. When cells sense signal for oxidative stress, NRF2 is activated and upregulates transcription of antioxidant proteins⁶⁹. Mutations that prevent KEAP1 expression or reduce its affinity for NRF2 cause an accumulation of NRF2 in the nucleus and the anti-oxidant genes under its regulation⁷⁰.⁷¹. Inactivation mutations within the *KEAP1* gene are found in 15 to 30% of cancers depending on the tissue of origin⁷¹. As an alternative mechanism of adaptation to high ROS level, cancer cells can enhance their capacity to repair oxidative DNA damage⁸. This is accomplished mostly by increasing the expression and efficiency of proteins involved in the base excision repair (BER) pathway. It's important to note that increased BER efficiency enables RAS-transformed

cells to avoid senescence (or apoptosis, depending on cell-type) and proliferate, however, there is a cost associated with this⁷². As described earlier, enzymes involved in DNA repair are more error-prone than the replicative DNA polymerases. The BER DNA polymerase, DNA pol β , exhibits an error rate close to 10^{-3} ⁷³. The increased error rate is compounded by the fact that Ras-cancer cells suffer more DNA damage from ROS than normal cells. As a result, Ras-transformed cells are expected to acquire mutations faster than normal cells, a phenomenon that may contribute to their genetic heterogeneity.

1.2 Oncogenes and Tumor Suppressors

It's well recognized that cancer is a genetic disease that arises as a result of mutations and abnormal expression of specific cancer genes. Genetic alterations in these two types of genes are responsible for tumorigenesis: oncogenes, and tumor suppressors. Proto-oncogenes, when mutated, become oncogenes. Oncogenes have genetic alterations that are activating in nature, such as gain-of-function mutation, amplification in copy number, or overexpression. These mutations generally lead to an increase of activity of the mutated oncogenes, promoting tumorigenesis. Some classical functions of oncogenes are growth factors and their receptors, cell cycle regulators, and apoptosis inhibitors. Contrastingly, tumor suppressors often have alterations that's deactivating in nature, with loss-of-function mutation, deletion, and epigenetic silencing⁷⁴. These genes often regulate essential cellular processes, such as cell proliferation, differentiation, and development. There are generally two types of tumor suppressors, termed the gatekeepers and the caretakers. Gatekeeper genes directly regulate cell proliferation and their mutations are rate-limiting steps for tumorigenesis. Inactivation or down-regulation of gatekeepers generally results in unregulated, rapid cell proliferation, a hallmark of cancer⁷⁵.

Caretaker genes do not directly partake in proliferation regulation, but contribute to prevent genomic instability. A famous example of caretakers is the *p53* tumor suppressor gene. Inactivation of *p53* is found in over 50% of human tumours⁷⁶. It acts as a cell cycle “checkpoint” prior to DNA replication. If a cell is suffering from DNA damage, *p53* either stalls DNA replication until the damage is repaired, or in the case of high level DNA damage, signal the cell to apoptosis⁷⁶. Cells with *p53* inactivation will enter S phase and replicate DNA regardless of the damaged, producing daughter cells containing the same DNA damage.

A single mutational event in proto-oncogene is often enough to activate it and contribute to abnormal cell growth. However, this is not the case for tumor suppressors. The inactivation of one tumor suppressor allele leaves behind one intact and functional allele that is usually sufficient enough to perform the tumor suppression task. As a result, both alleles need to be inactivated to disable tumor suppression. This is known as the Knudson two-hit hypothesis⁷⁷. Haplo-insufficient tumor suppressors are the exceptions, where the loss of function on one allele is enough to impart a phenotype. Due to the recessive nature of most inactivated tumor suppressor, the mutation can be inherited through germline, since a carrier would show no physical consequences, other than being more at risk for cancers. Indeed, familial cancers such as familial breast and ovarian cancers often involve the inheritance of mutated BRCA1/2 protein, which are caretaker tumour suppressors.

1.2.1 Oncogene Addiction

Oncogene addiction describes the phenomenon when some cancers are so heavily reliant on the survival benefits provided by an oncogene, such that the inhibition of said oncogene can

lead to serious deleterious effects on cancer cells. In this case, we will say that cancer cells have become addicted to an oncogene. Some oncogenes are so coveted that their inhibition can lead to tumor cell death, arrest or senescence. These oncogenes are often of great clinical interest as targets for cancer therapy. A good example of such is the usage of imatinib, an inhibitor of BCR-ABL to treat chronic myeloid leukemia (CML). A large number of CML cases are caused by the oncogene *BCR-ABL*, which is a fusion gene coding for an active tyrosine kinase signaling protein that's always active⁷⁸.

1.2.2 Non-oncogene Addiction and Synthetic Lethality

Since the original six hallmarks of cancer proposed in 2000, additional hallmarks have been proposed as our understanding on cancer initiation and progression expands⁷⁹. Some of these new hallmarks are stress phenotypes of cancer, such as oxidative stress and DNA damage stress⁷⁴. Some of these stress phenotypes are not uniquely observed in cancer cells, and can be found in other conditions such as chronic inflammation. These stress phenotypes are likely oncogenesis-associated cellular stresses. The concept of non-oncogene addiction depicts the dependence of cancer cells on genes involved in normal cellular functions that are not classical oncogenes, but can assist cancer cells on overcoming the stress phenotype⁸⁰. Due to this dependence, a new therapeutic strategy is to target the function of these “non-oncogenes” to cause synthetic lethality in cancer cells. PARP1 inhibition in BRCA defective cells is one such success story. Poly(ADP-ribose) polymerase (PARP1) facilitates repair of single-stranded breaks⁸¹. Inhibition of PARP1 in cells will result in accumulation of SSBs, and if these breaks are present in close proximity, double-strand breaks are created. DSBs in cells are repaired by

either homologous recombination-mediated DNA repair and non-homologous end joining. Therefore, PARP1 inhibition is well-tolerated in normal cells. BRCA1/2 are mediators of HR repair, and are tumour suppressors frequently inactivated in cancers, especially in breast and ovarian cancers⁸². In cancer cells with inactivated BRCA1/2, PARP1 inhibition is synthetic lethal, because the generation of double strand breaks cannot be effectively repaired due to BRCA1/2 mutations^{83, 84}. Synthetic lethality refers to the idea that a combination of two or more individually non-lethal mutations, when present together, result in unviability⁸⁵. In this case, PARP1 itself is not an oncogene, however, BRCA inactivated cancer cells rely on the function of PARP1 to survive.

This concept of non-oncogene addiction and synthetic lethality can be exploited for novel therapies by inhibition of these non-oncogenes. Cancer cells are under constant oxidative stress and DNA damage stress⁷⁴. In order to manage these stress, they rely on the activity of proteins involved in DNA repair to combat these stresses. As a result, some cancers cells have up-regulation or overexpression of proteins involved in DNA repair, such an example is the CUX1 protein⁸⁰.

1.3 CUX1

The *cut* gene was initially identified and characterized in *Drosophila melanogaster*, where it plays a role in cell-type specification in the peripheral nervous system and the wing margin^{86, 87}. The mammalian orthologue of the *cut* gene in *Drosophila* is *CUX1* (CUT-like homeobox 1). The human CUX1 protein was initially identified as the CCAAT-displacement protein (CDP). Following the cloning of its cDNA, the sequence conservation with the *cut* gene revealed that these genes are orthologs⁸⁸. Since then, CUX1 has been named

successively as CDP, CDP/Cut, Cut-like 1 (CUTL1). Upon consultation with a world expert on homeobox genes, Dr. William J. Muller, the gene has finally been called *CUX1*⁸⁹. Three CUT domains, a Cut homeodomain, and a coiled-coil region are evolutionarily conserved between drosophila and human.

CUX1 encodes two main isoforms, the full-length protein p200, and a smaller isoform, p110⁸⁹. The p200 protein contains all three Cut repeats (CR1, CR2 and CR3), and a Cut homeodomain (HD); while the p110 protein contains only two Cut repeats and the homeodomain. The p110 isoform is produced by proteolytic processing of p200 during late-G1 phase^{88, 90}. The p110 isoform is much less abundant than p200 in cells. It exhibits stable interaction with DNA, and functions as a transcription regulator, the promoter context dictating whether it behaves as a transcription activator or repressor^{91, 92}. It's been reported as a transcription factor that stimulates cell cycle progression and the expression of DNA replication genes^{93, 94}. In addition, p110 CUX1 is also involved in the strengthening of the spindle assembly checkpoint and the establishment of a transcriptional program that ensures that cells are prepared to launch a rapid DNA damage response^{95, 96}. It should be noted that the DNA damage response is triggered and is amplified by a series of post-translational modifications of several types⁹⁶. The transcriptional program governed by p110 CUX1 ensures that the necessary proteins are in place ahead of time, such that they can rapidly be activated by post-translational modifications in response to DNA damage^{97, 98}.

In contrast, p200 CUX1 is more abundant, and binds DNA with extremely rapid “on” and “off” rate, rare for a classical transcription factor⁹⁹. This property precludes p200 CUX1 from functioning as a transcriptional activator. However, p200 CUX1 was reported to function as a transcription repressor by competing for binding site occupancy¹⁰⁰⁻¹⁰². This was the function

originally described for the CCAAT-displacement protein¹⁰³.

1.3.1 CUX1 as a Tumour Suppressor

CUX1 has a very complex role in cancer, in that although it is characterized genetically as a haplo-insufficient tumor suppressor, the gene has characteristics that defy that classification. *CUX1* is located at chromosome 7q22.1, a region where frequent loss of heterozygosity is found in various cancer, namely in 14% of uterine leiomyomas, 18% of breast cancers, and 15–25% of acute myeloid leukemia (AMLs) and myeloproliferative disorder (MPD)¹⁰⁴⁻¹⁰⁸. There are also RNAseq and RT-PCR data showing drastic reduction of *CUX1* mRNA expression in leukemia patient samples¹⁰⁹. Among the cancer cells with both *CUX1* allele present, 1-5% of them have inactivating point mutation in one allele¹¹⁰. In human blood progenitors cells, knockdown of *CUX1* led to approximately 40% increase in engraftment on transplantation into immunodeficient mice¹⁰⁹. In addition, subcutaneous injection of T cell acute lymphoblastic leukemia (T-ALL) with *CUX1* knockdown into mice showed an increase in tumour¹¹⁰.

1.3.2 CUX1 as an Oncogene

Paradoxically, copy number variation (CNV) analysis indicated gains of copy numbers of *CUX1* gene, are more frequent than losses in cancer of different tissues. For example, in the case of cancers of the large intestine, 2.9% of the cancer has copy number reduction while 37.2% has a gain of copy number. Interestingly, approximately one-third (25 of 77) of tumors and cancer cell lines with *CUX1* LOH show amplification of the remaining allele. A plausible

explanation for this observation is that inactivation of one *CUX1* allele facilitates tumor initiation. In tumor cells, the remaining allele is then amplified to increase *CUX1* expression to support cancer cell survival and tumorigenic progression.

Transgenic mice expressing p200 CUX1 protein under a cytomegalovirus promoter had striking multi-organ hyperplasia and organomegaly¹¹¹. Expression of p200 CUX1 protein under the control of the mouse mammary tumor virus long terminal repeat (MMTV-LTR) resulted in the development of mammary tumors with long latency and a penetrance of 21%. Spontaneous activating mutation of Kras (G12V or Q61L) were observed in 45% of mammary tumors. Lung tumor formation was observed in 20% of transgenic mice as well¹¹². Transgenic mice expressing the p75 CUX1 or p110 CUX1 isoform also developed mammary tumors after a long latency period. Metastasis to the lung was observed in small portion of p75 CUX1 transgenic mice¹¹³. In addition to mammary tumors, sarcomas were also observed in the uterus and liver of p110 CUX1 and p75 CUX1 transgenic mice¹¹⁴. These transgenic mice models demonstrated that the expression of CUX1 isoforms increase tumor incidence in mouse genetic background.

1.3.3 CUX1 in Base Excision Repair

Oncogenic Ras cannot transform primary culture cells on its own, and induces senescence in primary cells by elevating ROS level. However, co-expression of p200 CUX1 and activating RAS mutation was found to prevent cellular senescence in human primary fibroblasts¹¹². In addition, mouse embryonic fibroblasts (MEF) isolated from *CUX1* null mice showed a longer G1 phase and slower proliferation compare to wildtype MEFs when cultured

in 3% oxygen^{89, 115}. However, when these MEFs were cultured in 20% oxygen, striking proliferation stress was observed in *Cux1*^{-/-} MEFs. *Cux1*^{-/-} MEFs also showed a reduced efficiency to repair DNA damage induced by hydrogen peroxide compared to wildtype MEFs¹¹⁵. These observations from our laboratory led to the discovery of CUX1 as an accessory factor of base excision repair. BER enzymes in mammalian cells, unlike their bacterial homologs, can be stimulated by accessory factors¹¹⁵⁻¹²⁰. Consequently, BER efficiency can be modulated by controlling the expression of accessory factors.

p200 CUX1 functions as an accessory factor in base excision repair through its three CUT repeats^{112, 115}. p200 CUX1 is able to both accelerate the binding of OGG1 to damaged DNA, and stimulate the enzymatic activity of OGG1^{112, 115, 117, 121}. As a result, p200 CUX1 accelerates the removal of oxidized base, 8-oxoguanine. In addition, CUX1 also stimulates the APE1 endonuclease activity⁴⁵. Preliminary work from the laboratory has generated evidence suggesting that Cut repeats of CUX1 can stimulate the polymerase and dRP-lyase activities of DNA pol β (Ramdzan *et al*, submitted).

1.3.4 Duality of CUX1 in Cancer

CUX1 is heavily involved in maintaining genome stability. The p110 CUX1 protein transcriptionally activates genes in DNA replication, DNA damage response and spindle assembly checkpoint, together ensuring proper DNA replication and cell cycle progression. In addition, p200 CUX1 directly promotes genome stability by its involvement in the base excision repair pathway⁹³⁻⁹⁶. Such functions are atypical for oncogenes, but overexpression of p200 or p110 CUX1 contributes to tumorigenicity in both cell lines and transgenic mice¹¹².

This phenomenon suggests that cancer cells exhibit non-oncogene addiction to the roles CUX1 plays in maintaining genome stability.

In Ras-transformed cells, high level of reactive oxygen species is produced from prolonged activation of the Ras pathway. This causes oxidative DNA damage and sustained DNA damage would typically lead to cellular senescence. However, high level of CUX1 present in cells tolerates the ROS level, as base excision repair is the main pathway for the repair of oxidative DNA damage. Indeed, CUX1 knockdown is synthetic lethal with activation of the Ras pathway, including activating mutations in KRas, HRas, NRas, BRAF or EGFR^{112, 122, 123}. Lastly, CUX1 was shown to increase resistance of glioblastoma cells treated with temozolimide⁴⁵. The paradoxical role of CUX1 in cancer is quite different from that of classical tumor suppressors. Thus, it represents the potential for a new class of cancer gene.

1.4 Project Rationale

CUX1 doesn't stimulate the activity of DNA glycosylases other than OGG1 in base excision repair. Therefore, we hypothesized that additional accessory factors must exist to stimulate other enzymes in the pathway. Studies have identified some additional accessory factors that stimulate BER enzymes: for example, YB-1 stimulates NTHL1, NEIL1, and NEIL2^{124, 125}. However, in order to fully elucidate the complex roles BER accessory factors play in cancer, we will need a complete repertoire of accessory factors. Therefore, our laboratory decided to perform systematic identification of accessory factors of BER enzymes.

My project is one of the first initiatives of this ambitious goal, which we attempt to identify accessory factors for NTHL1. NTHL1, is a bifunctional glycosylase in BER and it is responsible for removing oxidative pyrimidine bases, specifically oxidative variants of thymine,

such as thymine glycol (Tg) or 5,6-dihydrothymine (DHT)¹²⁶.

This is achieved by using the powerful technique of BioID. Proximity identification by biotinylation (BioID) is a powerful method for the study of protein interactions in living cells. This technique fuses a promiscuous biotin ligase mutant derived from *Escherichia coli*, coined BirA*, to a protein of interest¹²⁷. This biotin ligase can efficiently modify biotin to biotinyl-AMP, an activated form for biotinylation, yet has poor affinity for the molecule¹²⁷. As a result, biotinyl-AMP diffuses away from the fusion protein, reacts with nearby lysine residues, and biotinylates neighboring proteins within a 10 nm radius^{127, 128}. The greatest advantage of using BioID is its sensitivity, it can label binding partners with weak or transient interactions. Moreover, since biotin is linked through a covalent bond to target proteins, this made it possible to use harsh purification methods to prepare protein extracts.

2. Materials and Methods

Cell Culture and Viral Production

HEK293FT and U2OS were cultured in Dulbecco's modified eagle medium (DMEM, Wisent), and Jurkat was cultured in Roswell Park Memorial Institute medium (RPMI, Wisent), both media were supplemented with 10% fetal bovine serum (FBS, Gibco) and 1% penicillin-streptomycin (Invitrogen). All cells were maintained at 37°C, 5% CO₂ and atmospheric O₂. Lentiviruses were produced by co-transfecting 293-FT cells with plasmids encoding BCL11B, BCL11B EGFP, short hairpin RNA against BCL11B (Misson shRNA pLKO.1 panel, Sigma) with packaging plasmid psPAX2 and envelop plasmid pMD2G. The medium of the transfected cells were collected for 3 days, 48 h post-transfection.

Bacterial Protein Expression

Plasmids expressing histidine-tagged BCL11B fragments (F1-4) were prepared by inserting gBlocks gene fragments (Integrated DNA Technologies) into pET-30a vector using restriction enzymes. C-His-BCL11B plasmid was purchased from GeneCopoeia in vector system pReceiver-B31. All proteins were expressed in the BL21 strain of *Escherichia coli* and were induced with isopropyl- β -D-thiogalactopyranoside. His-tagged fusion proteins were purified using nickel beads (Qiagen) and were eluted with 300 mM imidazole. Imidazole was eliminated by buffer exchanges with PBS molecular weight cut-off dialysis membrane (Spectra/Pro Dialysis tubing, Spectrum Laboratories).

Immunoblotting

Protein extracts were re-suspended in Laemmli buffer, boiled for 5 min, resolved by SDS-PAGE. The resolved gel was electrophoretically transferred to a PVDF membrane in transfer buffer (20% (v/v) methanol, 25 mM Tris, 19.2 mM glycine). Membranes were blocked in TBS-T (10 mM Tris, pH 8, 150 mM NaCl, 0.1% Tween X-100) with 5% milk and 2% bovine serum albumin for 2 hours. Following blocking, membranes were incubated with indicated primary antibodies diluted in TBS-T, washed with TBS-T, and incubated with corresponding secondary antibodies conjugated to horseradish peroxidase for 1 hr. Proteins were then visualized using the ECL system according to the manufacturer's instructions (BioRad). The following antibodies were used: anti- γ -tubulin (1:10,000, Sigma), BCL11B (1:2000, Bethyl), NTHL1 (1:1000, proteintech), anti-Flag (1:1000, Sigma), anti-streptavidin (1:2000, BioLegend), anti-GST (1:1000, Abcam), anti-His (1:3000, sigma).

BioID

Human NTHL1 was cloned into pcDNA5 FRT/ FLAGBirA* vector (from Dr. Anne-Claude Gingras' laboratory). Using the Flp-InTM T-RExTM system (Thermofisher), we generated 293 T-REx Flp-In cells stably expressing NTHL1-BirA*-FLAG, eGFP-NLS-BirA*-FLAG, and eGFP-BirA*-FLAG under a tetracyclin inducible system. Cells were treated with 1µg/mL of tetracyclin for 24 hr to induce protein expression. Six hours before cell collection, 50µM of biotin was added and immediately followed by irradiation treatment. Cells were pelleted and delivered to our collaborator Dr. Anne Claude Gingras' laboratory to be lysed and analyzed with mass spectrometry¹²⁷.

Immunoprecipitation

293T cells were co-transfected with 3xHA-BCL11B and c-avi NTHL1 (plasmids purchased by GeneCopoeia) using Lipofectamine3000 (Invitrogen) according to the manufacturer's instructions. 20µM of biotin was added into the media 48hr after transfection, and cells were collected after 24 hr. Cells were lysed using lysis buffer (20 mM Tris (pH 8.0), 150 mM NaCl, 1% NP40) supplemented with a protease inhibitor cocktail (Sigma) and centrifuged at 13000g for 20 min. c-avi NTHL1 was immunoprecipitated with magnetic streptavidin beads (GE Healthcare). The samples were separated by SDS-PAGE followed by immunoblotting with anti-NTHL1 antibody and anti-BCL11B antibody.

***In Vitro* Fluorescent Cleavage Assay**

The fluorescent cleavage assay was performed using a 43-mer oligonucleotide (Midland) with DHT modification at its sixth position, modified from Svilar *et al*¹²⁹. The 5' end of the DNA was

conjugated to a FAM fluorophore, and 3' end conjugated to a Dabcyl quencher. Cleavage reactions were conducted in 25 μ L reaction containing using 50 nM of oligonucleotide, 10nM of enzyme and proteins, in 20 mM Tris (pH 8), 1 mM EDTA (pH 8.0), 1 mM DTT.

***In Vitro* Radioactive Cleavage Assay**

Double-stranded 32-mer oligonucleotides containing a DHT modification at the 18th position (Midland) were labeled with ³²P-gamma ATP at the 5' end of the top strand (*) using poly nucleotide kinase. Cleavage reactions were conducted using indicated concentration of bacterially purified proteins and enzyme in 20 mM Tris (pH 8), 1 mM EDTA (pH 8.0), 1 mM DTT and 1 pmol of labelled probe. Reactions were performed for 30 min in 37°C and terminated by formamide DNA loading buffer (90% formamide with 0.05% bromphenol blue and 0.05% xylene cyanol). The DNA was loaded on a pre-warmed 20% polyacrylamide-urea gel (19:1) and separated by electrophoresis in Tris-borate and EDTA (TBE; pH 8.0) at 20 mA. The radio-labeled DNA fragments were visualized by storage phosphor screen (GE Healthcare).

Sodium Borohydride Trapping of NTHL1

5'-end-labeled 32-mer duplex containing DHT modification was incubated with purified NTHL1, and BCL11B proteins or BSA at the indicated concentrations. After incubation at 37°C for 30 mins, 50 mM sodium borohydride was added and the reactions were incubated for 15 minutes at 37 °C. The trapped complexes were separated from free substrate by 10% SDS-PAGE gel.

Electrophoretic Mobility Shift Assay (EMSA)

EMSAs were performed using indicated amount of purified proteins and incubated at room temperature for 15 minutes in a final volume of 30 μ L of 25mM HEPES (pH 7.5), 50mM KCl, 5mM MgCl₂, 10 μ M ZnCl₂, 5% glycerol, 0,1M DTT, with 0.1 μ g poly dIdC and 1 μ g BSA as nonspecific competitors of DNA and protein. End-labeled double-stranded oligonucleotides (1 pmol) were added and further incubated for 15 min at room temperature. Samples were loaded on 5% native polyacrylamide gels and separated by electrophoresis at 100V. Gels were dried and visualized by autoradiography.

On/Off Rate of DNA binding

EMSAs were performed using indicated amount of purified proteins and incubated at room temperature for 15 minutes in a final volume of 30 μ L of 25mM HEPES (pH 7.5), 50mM KCl, 5mM MgCl₂, 10 μ M ZnCl₂, 5% glycerol, 0,1M DTT, with 0.1 μ g poly dIdC and 1 μ g BSA as nonspecific competitors of DNA and protein.

On rate: End-labeled double-stranded oligonucleotides (0.1 pmol) were added to incubate for the indicated amount of time.

Off rate: End-labeled double-stranded oligonucleotides (0.1 pmol) were added to incubate for 15 minutes to establish stable binding. 100 pmol of unlabeled oligonucleotides of the identical sequence were added for the indicated amount of time.

Samples were loaded on 5% native polyacrylamide gels and separated by electrophoresis at 100V. Gels were dried and visualized by autoradiography.

Strand-Displacement Assay of DNA PolB

Oligonucleotides used in this experiment were purchased from Midland Oligos. The short complementary oligonucleotide primer (5'-TCACCCTCGTACGACTC) and reporter labeled (5' TTTTTTTGC - FAM 3') were annealed to oligonucleotide template (3'-AGTGGGAGCATGCTGAGAAAAAACG -DABCYL-5') to create quenched double-stranded DNA substrates. These were annealed in 50mM Tris-HCl, pH 8.0 and 100mM NaCl by heating at 95°C for 5 min and allowed to cool gradually to room temperature. Polymerase and strand-displacement activity was conducted by incubating 25 nM of PolB and the indicated concentration of BCL11B or BSA in the presence of 50 mM Tris-HCl, pH 8.0, 10 mM KCl, 1 mM MgCl₂, 2 mM dithiothreitol, 0.01% Tween-20, and 500 ng BSA.

Oligonucleotide Sequences

Fluorescent cleavage

(X= dihydrothymine)

5'(FAM) CCACTXTTGAATTGACACGCCATGTCGATCAATTCAACAGTGG (3'-Dabcyl)3'

Radioactive cleavage/ Schiff base trapping assay

(X= thymine glycol, sequences modified from Paz *et al.*¹³⁰)

5' CCGGTGCATGACACTGTXACCTATCCTCAGCG 3'

3' GGCCACGTACTGTGACAATGGATAGGAGTCGC 5'

EMSA/On/off rate/ K_D calculation

(bolded sequences are conserved binding motif of BCL11B)¹³¹

5' CCGGTA**ACCACATGATGCTTGCCTAGTGCTATCCTCA** 3'

3' GGCCATT**GGTGTACTACGAACGGATCACGATAGGAGTCGC** 5'

Strand-displacement assay

5' TCACCCTCGTACGACTC 3' 5' TTTTTTTGC (FAM) 3'

3' AGTGGGAGCATGCTGAGAAAAAACG (Dabcyl) 5'

GST-Pull Down Assay

GST-NTHL1 and GST are bound to glutathione beads (GE healthcare) in 50mM Tris (pH 8), 150mM NaCl, 1% NP40, 50mM NaF, 0.5% BSA and protease inhibitor in room temperature for 4 hours. Then 1µg of purified his-tagged BCL11B fragments was incubated with 1µg of GST/GST-NTHL1 bound to glutathione beads. Add NP40 lysis buffer (50mM Tris (pH 8), 150mM NaCl, 1% NP40, 50mM NaF) until the total volume is 700µL, then incubate at 4 °C overnight. Wash the beads three times with 1mL of NP40 lysis buffer the next day. Run the samples in 12% SDS-PAGE and proteins are visualized by immunoblotting against GST and His.

Single Cell Gel Electrophoresis

Single cell electrophoresis (comet assay) was used to measure DNA damage. 50µM of H₂O₂ was used to treat cells on ice for 20 min to induce DNA damage. Immediately after treatment, cells were washed with PBS to eliminate H₂O₂ residue and allowed to recover at 37°C in fresh medium for the indicated amount of time. Then cell pallets were collected and re-suspended in 1% agarose. The mixture was loaded on pre-coated slides and allowed to solidify (Trevigen). The slides were lysed in alkaline condition and subjected to electrophoresis to resolve damaged DNA from intact DNA. The slides were stained with propidium iodide and visualize with Axiovert 200M microscope with an LSM 510 laser module (Zeiss). Comet tail moments were

measured on a minimum of 50 cells using the Comet Score software (TriTeck Corp).

Methodology of single cell gel electrophoresis was modified from Olive and Banath (2006)¹³².

3. Results

BCL11B Interacts with NTHL1

The first aim of this project was to create a list of potentially NTHL1-interacting proteins, from which we can identify an accessory factor that is capable of stimulating the activity of NTHL1. We achieved this by performing a BioID experiment where we expressed a fusion protein containing the human NTHL1 enzyme, BirA* and a FLAG tag (Fig. 1B). The BirA* moiety is a mutant of a biotin ligase from *E. coli*¹²⁷. Two negative controls were used in this experiment, eGFP-BirA*-Flag, a cytoplasmic control, and eGFP-NLS-BirA*-Flag, a nuclear control. It is important that negative controls were included, because any proteins identified from these controls are likely caused by nonspecific interactions, and should be eliminated from our curated list of NTHL1 interacting protein. The expression of the fusion proteins was controlled under a tetracycline inducible system and biotin was added to the media to allow for the biotinylation of proteins in close proximity to NTHL1. Immunoblotting for NTHL1 and the FLAG tag were performed to confirm expression of the intended fusion proteins (Fig. 1D); and immunoblotting for streptavidin, which has a strong binding affinity for biotin and therefore recognizes biotinylated proteins, was done to confirm successful biotinylation (Fig. 1E, lanes 2, 6 and 8). Immediately after biotin addition, cells were submitted to ionizing irradiation to induce DNA damage and collected 6 hours later. Cell pellets were lysed, and biotinylated proteins were isolated using streptavidin beads and identified using mass spectrometry (schematics of BioID see Fig. 1A and C). All mass spectrometry work was performed by Dr. Payman S. Tehrani in the

laboratory of Dr. Anne-Claude Gingras at the University of Toronto. All protein candidates identified in mass spectrometry were given a Bayesian false discovery rate (BFDR), defined as the expected proportion of false positives. After eliminating all candidates with a BFDR greater than 0.2, 223 candidates remained: 41 unique to unirradiated cells, 36 unique to irradiated cells, and 146 are common (Fig. 1F). Table 1 provides a list of all these proteins together with their BFDR in unirradiated and irradiated cells. The number of amino acids of each protein is also indicated. This criterion enabled us to eliminate proteins that would be too large to purify from bacteria for eventual enzymatic assays with the NTHL1 enzyme. After intensive literature review of some of the more interesting candidates, we decided to investigate further with the BCL11B protein.

BCL11B, B-cell leukemia/lymphoma 11B, belongs to the Kruppel-like C₂H₂ zinc finger transcription factor family, and contains 6 zinc finger domains for DNA binding^{133, 134}. The *BCL11B* gene consists of 4 exons, and two alternatively spliced variants, with the smaller isoform, V2, lacking exon 3^{134, 135}. No known function has been attributed to the smaller BCL11B variant. BCL11B is primarily expressed in T cells, thymocytes and brain tissue¹³⁶. It is a bi-functional transcription regulator, and can function both as a transcriptional repressor or activator depending on posttranslational modifications and promoter context¹³⁷⁻¹⁴². Loss-of-function studies performed in mouse demonstrates that BCL11B plays a significant role in T-cell development and lineage commitment¹⁴³.

BCL11B has also been implicated in cancer in multiple ways. BCL11B is identified as a haplo-insufficient tumour suppressor¹⁴⁴. It was initially named radiation-induced tumor suppressor gene 1 (*Rit1*), because it contained homozygous deletions and point mutations in gamma-ray induced mouse thymic lymphomas^{143, 145-147}. Mice with heterozygous BCL11B are

more susceptible to thymic lymphomas after exposure to γ -radiation^{135, 148, 149}. The absence of a BCL11B allele also contribute to vulnerability to DNA replication stress and damage¹⁴⁸. Heterozygous BCL11B mutations and deletions were also found in major molecular subtypes of T-cell acute lymphoblastic leukemia (T-ALL)¹⁵⁰. Subsequent analysis in human T-cell acute lymphoblastic leukemia (T-ALL) revealed monoallelic *BCL11B* inactivating point mutations or deletions in 9 to 16% of cases¹⁵¹⁻¹⁵³.

All these findings are evidence that BCL11B is a haplo-insufficient tumor suppressor, yet there are contradicting observations suggesting otherwise. BCL11B is found overexpressed in T-ALL, cutaneous T-cell lymphomas (CTCL), squamous cell carcinomas, Ewing sarcomas and glioblastomas^{138, 154-157}. The Sanger/COSMIC site reports BCL11B copy number gain or loss in 35 and 15 cancer patient samples, respectively; while overexpression and underexpression are observed in 394 and 2 cases, respectively.

[<http://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=BCL11B>]. Down-regulation of BCL11B led to growth inhibition and synthetic lethality in T-ALL and glioblastoma cells, but not in normal mature T cells¹⁵⁸⁻¹⁶⁰. *BCL11B* overexpression also increases resistance to radiomimetic drugs^{161, 162}.

The status of *BCL11B* as a haplo-insufficient tumor suppressor gene, with the fact that it is overexpressed in several cancers, and its knockdown is synthetic lethal in some T-cell lymphomas, are reminiscent of the situation with the *CUX1* gene.

After we decided to investigate further with BCL11B, we performed a co-immunoprecipitation in 293FT cells to validate the protein-protein interaction between NTHL1 and BCL11B. Two vectors, one coding for c-3xHa-BCL11B, another for c-avi-NTHL1 and a biotin ligase (vectors purchased from GeneCopoeia), were introduced into 293FT by

transfection. The avi tag is a 15 amino acid peptide that can be biotinylated by biotin ligase in the presence of biotin. The biotinylated avi tag allowed us to pull-down NTHL1 using magnetic streptavidin beads, and immunoblot analysis showed successful pull-down of NTHL1, as well as its binding partner, BCL11B (Fig. 2A). This result, in combination of the identification of BCL11B from BioID of NTHL1, are strong evidences of the protein-protein interaction between BCL11B and NTHL1.

BCL11B Stimulates Enzymatic Activities of NTHL1

We next investigated whether BCL11B is capable of stimulating the glycosylase and AP/lyase activities of NTHL1. First, we used a fluorescence-based assay to evaluate the cleavage efficiency of NTHL1. This assay utilised a 43 base oligonucleotide, with an oxidized thymine base (dihydrothymine, DHT) at its sixth position, designed specifically to form a hairpin loop when hybridized (Fig 3A). A green FAM fluorophore was conjugated to the 5' end of the oligonucleotide, positioned adjacent to a Dabycl quencher when the probe is in its hairpin form. When NTHL1 was in contact with this fluorescent probe, the enzyme excised the damaged DHT base, cleaved the resulting AP site through its AP-lyase activity, thereby releasing the short oligonucleotide containing the fluorophore. Hence, fluorescence emission was observed. Therefore, by measuring the fluorescent intensity, we were indirectly measuring the enzymatic activities of NTHL1. When purified BCL11B was added into the reaction mixture, we observed a significant increase in fluorescent intensity, compared to reactions in the presence of non-stimulating proteins such as BSA, CUX1 and HOXB3 (Fig. 3B). Next, we confirmed this observation using a radioactive cleavage assay, with a 32 base pair double-stranded oligonucleotide that was radioactively labelled at the 5' end, containing another variant of the

oxidized thymine base, thymine glycol (Tg) (Fig. 4A). When NTHL1 was in contact with the DNA probe, it excised the damaged base and cleaved the AP site, resulted in a 17 nucleotide labelled product that can be resolved from the 32 nucleotide substrate by electrophoresis and visualized on a radiograph. When BCL11B was added in the reaction, there was significantly more product formation compared to enzyme alone (Fig. 4B, compare lane 6 with lanes 7 to 10).

BCL11B Increases the Formation of the Schiff Base Intermediate

A key intermediate is produced in between the glycosylase and AP/lyase activity of NTHL1, where the catalytically active lysine residue of NTHL1 is covalently bonded to the deoxyribose of the damaged base to form a Schiff base intermediate. Typically, the subsequent chemical reaction would release NTHL1 from the Schiff base, allowing the AP/lyase reaction to take place. In the Schiff base trapping assay, sodium borohydride (NaBH_4), a reducing agent is added into the reaction mixture. As a result, the double bond is reduced, and NTHL1 is permanently bonded to DNA (Fig. 5A). This trapped complex can be resolved from the unbound DNA using 10% SDS-PAGE. In this experiment, we allowed varying concentrations of BCL11B and BSA, a control protein, to incubate with NTHL1 and radioactively labelled oligonucleotide (identical to the one used in the cleavage assay), and the resulting Schiff base was trapped using NaBH_4 . Lanes with the addition of BCL11B showed significantly more trapped complex compare to lanes with BSA, suggesting that BCL11B increased the formation of Schiff base intermediate (Fig. 5B, compare lane 6-8 to lane 3-5). As the concentration of BCL11B increased, we surprisingly saw diminishing amount of the trapped complex. A possible explanation is that the high concentration of BCL11B increased the reaction rate such that much of the Schiff base intermediate was already resolved and couldn't be trapped by NaBH_4 .

BCL11B Binds DNA with Rapid On and Off Rate

BCL11B is a transcription factor with six C₂H₂ zinc finger domains for DNA binding. Therefore, we wanted to investigate the DNA binding behaviour of BCL11B. A study from 2011 performed a ChIP-seq analysis of BCL11B in striatal cells and found consensus DNA binding motifs¹³¹. Based on the three identified DNA binding motifs from the study (ACCACA, TGCTTGC, AGTGCT), we designed a 39-mer double stranded oligonucleotide and performed the on/off rate experiments (Fig. 6A). To examine the on rate, the rate at which BCL11B binds to DNA, we allowed the protein to incubate with the radioactively labelled oligonucleotide for various periods of time. When protein is bound to DNA, the resulting complex, termed retarded complex, can be resolved from the free unbound DNA probe using gel electrophoresis. In as few as 30 seconds, BCL11B already formed a retarded complex with DNA (Fig. 6B, lane 9). In order to examine the off rate, the rate at which BCL11B dissociates from a bound DNA, we first incubated the protein with radioactively labelled DNA for 15 minutes to reach an equilibrium, as seen in lane 4. Then, 1000-fold excess of unlabelled DNA of the same sequence was added into the reaction mixture and allowed to incubate for various amount of time. When BCL11B dissociates from the labelled probe, due to the excess amount of unlabelled probe present in the environment, it will likely bind to an unlabelled probe, and therefore no visible retarded complex should be observed on a radiograph. From the off rate portion of the experiment, BCL11B dissociated from the labelled DNA in as few as 30 seconds (Fig. 6B, lane 11). In conclusion, BCL11B binds DNA with a very rapid on and off rate. As will be discussed later, this DNA binding behavior may be important in the cellular context where large regions of DNA must be patrolled by the NTHL1 enzyme.

BCL11B Preferentially Binds to DNA with Thymine Glycol Damage

Next, we wanted to explore whether BCL11B binds preferably to DNA with a damaged base, and potentially assists NTHL1 in the recognition of damaged bases. Using the same oligonucleotide sequence from the on/off rate experiment, we modified the base at position 24 from a wildtype thymine to an oxidative variant, thymine glycol, and evaluated the DNA binding affinity of BCL11B to the two different oligonucleotides (Fig. 7A). The binding affinity is assessed in the form of equilibrium dissociation constant, K_D , which is calculated based on when 50% of the DNA is bound by the protein of interest¹⁶³. By definition, a smaller K_D value indicates that a lower concentration of protein is required to bind 50% of DNA. Therefore, equilibrium dissociation constant and binding affinity are reciprocal in relationship, in that smaller K_D calculation suggests higher binding affinity. We allowed BCL11B in a wide range of concentrations to incubate with 10 pmol of DNA, both wildtype and with thymine glycol (Tg) modification, and the protein-bound DNA was separated from the unbound DNA using electrophoretic mobility shift assay (Fig. 7B). The radioactive signals were quantified using the software ImageQuant, and the percentage of free DNA in each lane was plotted against protein concentration (Fig. 7C). The final calculations of the apparent equilibrium dissociation constant, K_{Dapp} , is 5.0×10^{-10} nM and 1.6×10^{-9} nM for DNA containing thymine glycol damage and wildtype DNA, respectively (Fig. 7D). The smaller K_{Dapp} value for the Tg probe suggests that BCL11B has a slightly higher preference for DNA with thymine glycol base.

BCL11B Stimulates Polymerase Activity of DNA Pol β

DNA Pol β plays an important role in base excision repair, as it synthesizes the base removed by glycosylases and fills the gap via either the short patch or long patch repair pathway.

In addition, Pol β is involved in end processing where it removes the 5'-deoxyribose phosphate (5'-dRP), produced when APE1 cleaves the DNA backbone during an earlier step of BER. Completion of base excision repair by Pol β and ligase III is crucial because single-strands that remain untended during DNA replication can generate a special type of double strand breaks that are conducive to DNA rearrangements. Therefore, stimulation of BER enzymes upstream of Pol β will not provide survival advantage to cells if more single-strand breaks were produced than can be processed by Pol β , as the bottleneck can lead to accumulation of single-strand breaks and eventually double-strand breaks. Consequently, we wanted to examine if BCL11B is capable of stimulating the polymerase activity of Pol β . This was assessed using a fluorescence-based strand displacement assay. The DNA probe used in this experiment was produced by annealing three oligonucleotides to mimic a nicked DNA substrate for Pol β (Fig. 8A). Pol β , in the presence of dTTP, will polymerize and elongate the DNA probe at the site of the nick, resulting in the displacement of the oligonucleotide containing the FAM fluorophore. The displacement separates the fluorophore from its quencher, and fluorescent emission can be detected in real time using a qPCR machine (Fig. 8B). The result of the displacement assay showed that the addition of BCL11B significantly increased the fluorescent intensity, when compared to the addition of the control protein, BSA (Fig. 8C). This observation suggests that BCL11B stimulates the polymerase and strand displacement activities of Pol β in the long patch repair pathway.

Residue 213-420 of BCL11B Stimulates NTHL1 Enzymatic Activity

Next, we aimed to dissect BCL11B into fragments and identify the domain(s) or region(s) of BCL11B responsible for the stimulation of NTHL1 activity. We performed a

function/structure analysis where we expressed four shorter fragments containing key domains: 1) fragment 1, residue 1-213, contains the region encoded by the first 3 exons of *BCL11B*; 2) fragment 2, residue 213-420, contains the first zinc finger domain, and a proline rich region; 3) fragment 3, residue 411-550, contains two zinc finger domains; 4) fragment 4, residue 711-894, contains the three C-terminal zinc fingers (Fig. 9A). These four fragments were expressed as His-tagged proteins, purified using immobilized metal affinity chromatography, and tested in the previously mentioned fluorescent cleavage assay with NTHL1. The result of the cleavage assay suggests that at 10 nM of proteins and 10 nM of NTHL1, fragment 2 is the only fragment capable of stimulation to the extent similar to that of the full length BCL11B protein (Fig. 9C). Next, a Schiff base trapping assay was performed using the purified fragments (Fig. 10B). The full length BCL11B used in this experiment was from a different purification than the previously shown Schiff base experiment, and likely contained a lower amount of active protein, as we observed less trapped Schiff base in this experiment. Nonetheless, there is noticeable increase in trapped Schiff base upon addition of full length BCL11B when compared to lanes with the control protein, BSA (Fig. 10B, lane 3 and 4). Among the four BCL11B fragments tested in Schiff base assay, it is clear that fragment 2 significantly increases Schiff base formation, a result that is consistent with our fluorescent cleavage data (Fig. 10B, lane 6). In addition, we also performed a pull down assay using the four BCL11B fragments and NTHL1 to identify the region(s) on BCL11B responsible for the interaction. We purified His-tagged BCL11B fragments and GST-tagged NTHL1. GST-NTHL1 and GST, a negative control for non-specific binding to GST tags, were bound to glutathione beads and then incubated with the His-tagged fragments of BCL11B. We visualized the proteins using immunoblot against the His and the GST tags (Fig. 11B). GST-NTHL1 is approximately 62 kDa (GST is 25 kDa and NTHL1 is 37 kDa). The two

GST western blots confirmed successful binding of GST proteins to glutathione beads. Of the four BCL11B fragments tested, GST-NTHL1 was able to pull-down fragment 2 and 3 of BCL11B (Fig. 11 B, histidine blot, lane 4 and 5). However, the binding of fragment 3 appears to be non-specific, as it is also pull-downed by GST (Fig. 11B, histidine blot, lane 9). Although fragment 2 is also pull-downed by GST alone, it is present at a much lesser amount, suggesting its binding to GST is much weaker compared to GST-NTHL1. The data presented on this pull down experiment is only preliminary, and will be optimized.

C-terminal Zinc Fingers of BCL11B are Responsible for DNA Binding

Our next course of action was to determine which fragment(s) can bind to DNA, and characterize the binding behaviour of such fragment(s). We performed an EMSA experiment with the four fragments and the full length protein, where we allowed the proteins to incubate with radioactively labelled oligonucleotides for 15 minutes, and the retarded complex was resolved from the unbound DNA. The full length protein used in this experiment is from the same purification as the Schiff base assay from Fig.10, and therefore, although still present, we observed less retarded complex formation than in the previous EMSAs shown in Fig. 6 and 7. Of the four fragments tested, fragment 4, which contains the three C-terminal zinc finger domains, is the only fragment capable of DNA binding (Fig. 12B, lane 9-10). Next we examined the DNA binding behaviour of fragment 4 by evaluating its on and off rate. Much like the full length protein, fragment 4 exhibits rapid on and off rate of binding to the oligonucleotides: it is able to bind to DNA in less than 12 seconds, and dissociates from previously bound DNA within 12 seconds (Fig. 13B, lane 9 and 12, respectively).

BCL11B Recruitment to Site of DNA Damage

After investigating BCL11B's interaction with NTHL1 *in vitro*, we decided to investigate the recruitment of BCL11B to the site of DNA damage in cells using micro-irradiation. The micro-irradiation and microscopy work was done in collaboration with the laboratory of Dr. Alexandre Marechal at the University of Sherbrooke. We created a U2OS cell line stably expressing a GFP-BCL11B fusion protein (Fig. 14A). DNA damage was induced using micro-irradiation, and cells were fixed and stained by immunofluorescence for GFP and γ -H2AX, a marker for double strand breaks. There is a higher concentration of GFP signal at the site of double strand breaks, inferring BCL11B recruitment to the site of DNA damage (Fig. 14B). Life-imaging was also performed on these cells once per minute post-irradiation, and BCL11B was recruited to the site of damage in less than one minute and lingers until at least four minutes post irradiation.

BCL11B Knockdown in Jurkat Cells Reduces DNA Repair Efficiency

Following the laser micro-irradiation, we investigated BCL11B's impact on DNA repair efficiency using single cell gel electrophoresis. We generated Jurkat cells stably expressing BCL11B shRNAs from a lentiviral vector. The decision to use Jurkat cells was based on the fact that BCL11B is overexpressed in Jurkat cells and BCL11B knockdown in these cells was previously reported to cause synthetic lethality^{158, 159}. A panel of 5 shRNAs were tested, and two out of five, shBCL11B #3 and #4, showed significant reduction of BCL11B expression and were used in the subsequent experiment (Fig. 15B). We chose single cell gel electrophoresis, also known as comet assay, to evaluate the efficiency of DNA repair. In this assay, cells were treated with hydrogen peroxide, H₂O₂, to induce DNA damage, and allowed to recover for 0, 60 or 120

minutes. Then, cells were embedded in agarose droplets and lysed under alkaline condition. The alkaline lysis buffer enables hydrolysis of AP sites and alkali-labile bases, thereby producing single strand breaks. Using gel electrophoresis, we can separate the denatured DNA fragments from the intact genome. Damaged DNA migrates slower than intact DNA, resulted in the formation of a “comet tail”. These agarose beads were stained with propidium iodide, and the comet tails were visualized under a microscope and quantified into comet tail moments (schematics of single cell gel electrophoresis shown in Fig. 15A). The scored comet tail moment of the three Jurkat cell lines, one cell line with empty vector acting as negative control, and two cell lines each expressing one of the two selected shRNAs, were plotted (Fig. 15C). Reduced BCL11B expression was associated with an increase in DNA damage at baseline level in untreated samples. In addition, at both 60 or 120 minutes of recovery time, cell lines with BCL11B knockdown showed significantly higher DNA damage compared to control, suggesting BCL11B knockdown reduces DNA repair efficiency, resulting in higher amount of unrepaired damage (Fig. 15C).

4. Figures

Fig. 1: Biotin Dependent Proximity Identification Experiment of NTHL1.

A) Schematics of biotin dependent proximity identification (BioID).

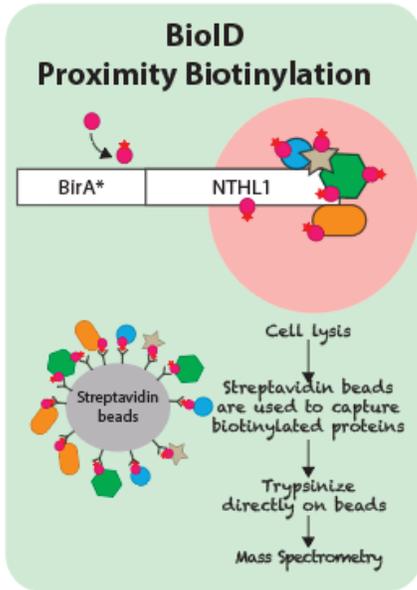
B) Flp-InTM T-RExTM 293 cell lines stably expressing NTHL, eGFP, or eGFP-NLS fusion proteins conjugated to a modified biotin ligase, BirA^{*}, and a flag tag was generated using transfection.

C) Flow chart of events took place during the BioID experiment for NTHL1.

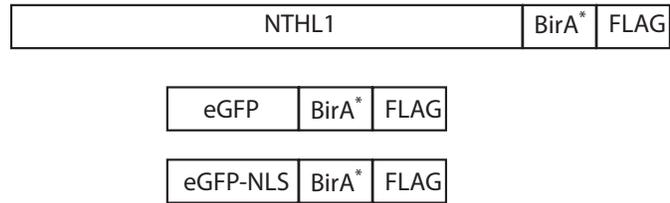
D,E) Immunoblots confirming the expression of BirA^{*} fusion proteins and successful biotinylation by BirA^{*}.

F) Venn diagram showing the number of NTHL1 target proteins identified from mass spectrometry after removing all candidates with BFDR above 0.2, and proteins identified in cells expressing eGFP-BirA^{*}-FLAG and eGFP-NLS-BirA^{*}-FLAG.

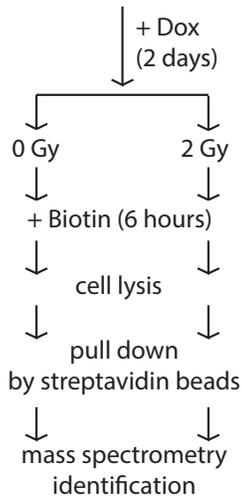
A)



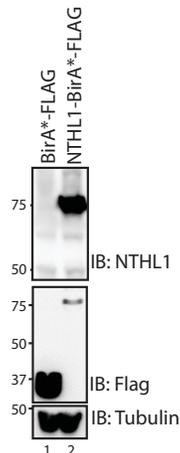
B)



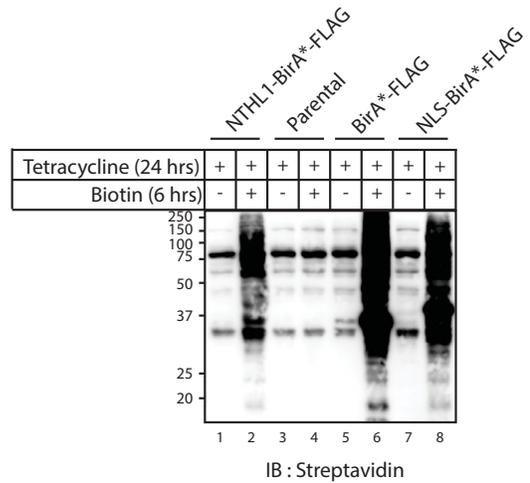
C) NTHL1-BirA*-FLAG
or eGFP-BirA*-FLAG
or eGFP-NLS-BirA*-FLAG



D)



E)



F)

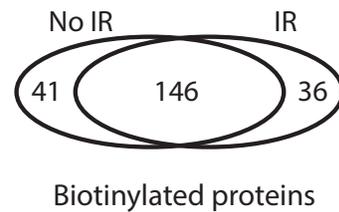


Fig. 2: Co-Immunoprecipitation of NTHL1 and BCL11B.

A) HEK293FT cells were transfected with two vectors, expressing c-3xHA BCL11B and c-avi NTHL1. Immunoprecipitation was performed using magnetic streptavidin beads and analyzed by immunoblotting with BCL11B antibody. Input (1%) was loaded as a protein expression control.

A)

1% input IP: streptavidin

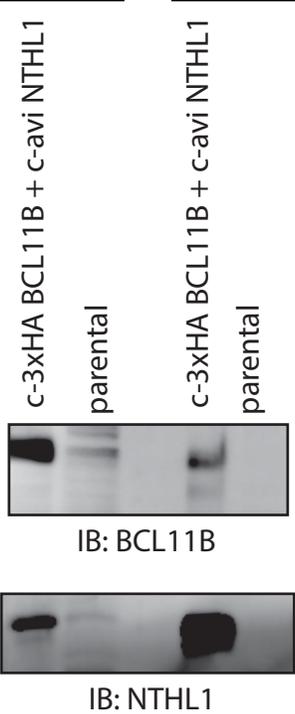


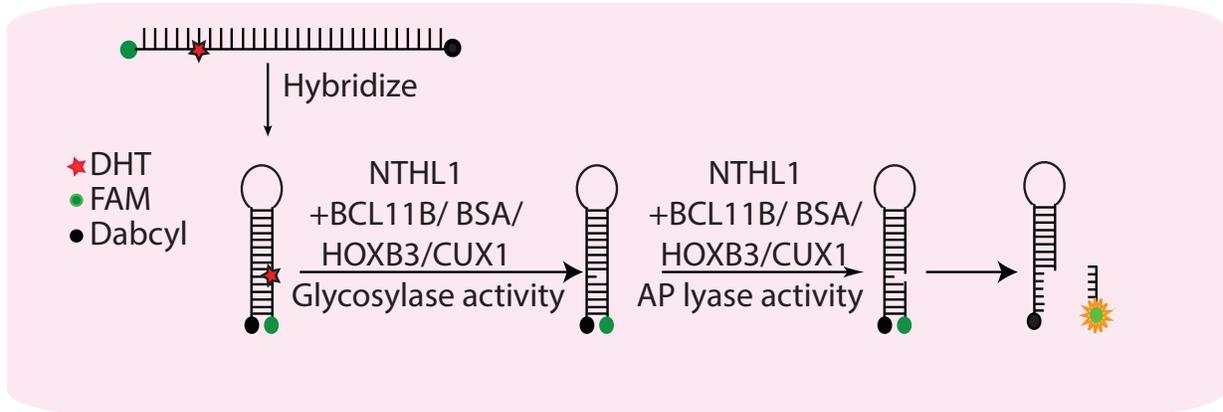
Fig. 3: BCL11B Stimulates Enzymatic Activities of NTHL1 in Fluorescent Cleavage Assay.

A) Schematic representation of the NTHL1 cleavage assay using a fluorophore-based probe that contains a green FAM fluorophore, a dihydrothymine (DHT) base at position 6 and the Dabcyl quencher.

B) NTHL1 cleavage assay was performed using the fluorophore-based probe and histidine-tagged affinity purified NTHL1 (10 nM) in the presence of another protein (5 nM), either BSA, BCL11B, HOXB3 or CUX1 CUT domains 1 and 2. The experiment was performed in triplicates. Error bars represent standard deviation. Curves were plotted using Prism (6.0) Graphpad.

A)

5'-(FAM)-CCACTXTTGAATTGACACGCCATGTCGATCAATTCAATAGTGG-(DabcyI)-3'
 where X is an oxidized thymine, dihydrothymine (DHT)



B)

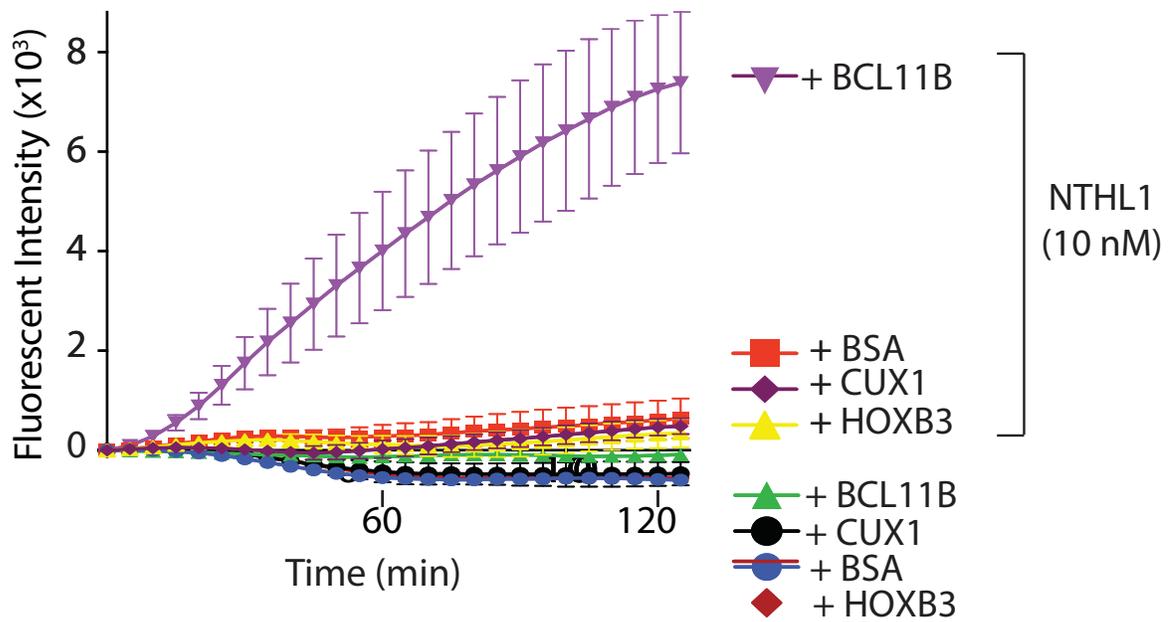


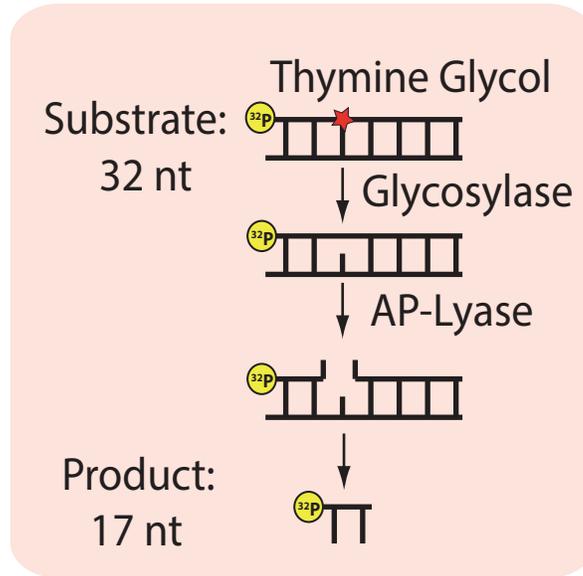
Fig. 4: BCL11B Stimulates Enzymatic Activities of NTHL1 in Radioactive Cleavage Assay.

(A) Double-stranded oligonucleotides with a thymine glycol base (red star) were labeled with ^{32}P - γ ATP at the 5' end of the top strand using polynucleotide kinase.

(B) His-tagged NTHL1 and varying concentrations of BCL11B proteins were incubated at 37°C for 30 min.

A)

Substrate: CCGGTGCATGACACTGTXACCTATCCTCAGCG
 X=Tg GGCCACGTACTGTGACAATGGATAGGAGTCGC



B)

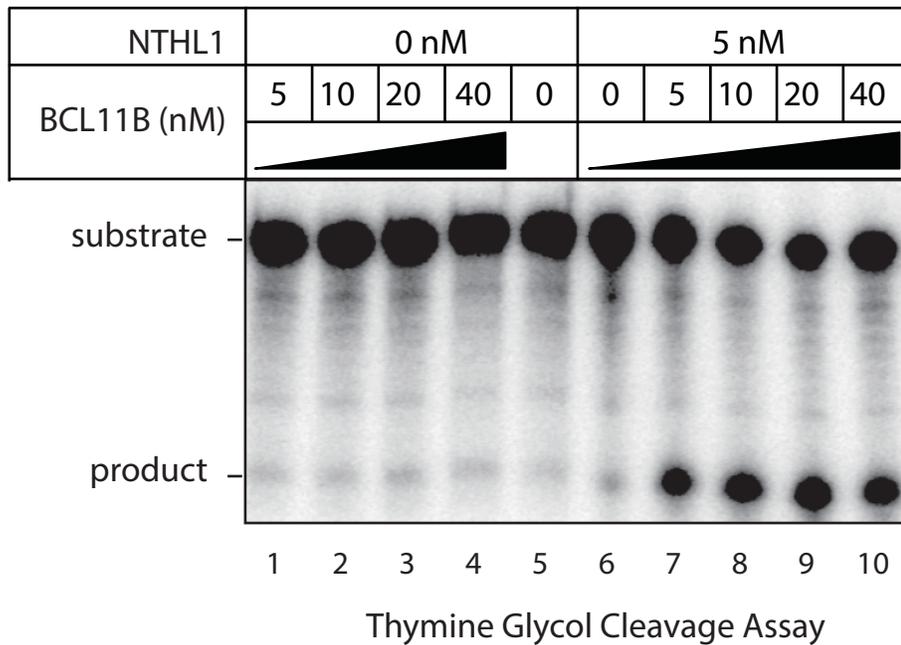


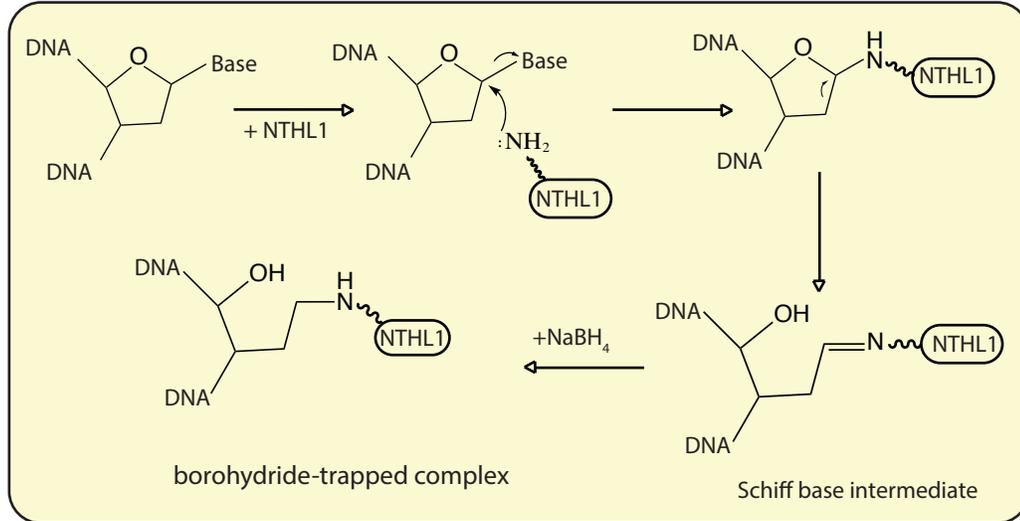
Fig. 5: BCL11B Promotes Schiff Base Formation by NTHL1.

A) Schematics of the Schiff base assay.

B) Radioactively end-labeled double-stranded oligonucleotides containing a thymine glycol (Tg) base were incubated with his-NTHL1, in the presence of His-BCL11B or BSA, and the Schiff base complex was trapped using 50 mM sodium borohydride. The trapped complexes were separated from free substrate using 10% SDS-PAGE.

A)

Substrate: CCGGTGCATGACACTGTXACCTATCCTCAGCG
 X=Tg or T
 GGCCACGTACTGTGACAATGGATAGGAGTCGC



B)

Nth1 (nM)	20	0	2.5	2.5	2.5	2.5	2.5	2.5	0	2.5	2.5
BSA (nM)	10		2.5	5	10					5	
BCL11B (nM)						2.5	5	10			5
wt probe									+	+	+
Tg probe	+	+	+	+	+	+	+	+			
Lane #	1	2	3	4	5	6	7	8	9	10	11

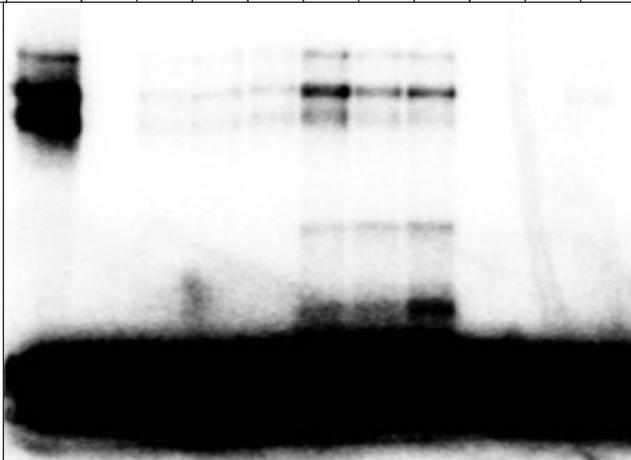


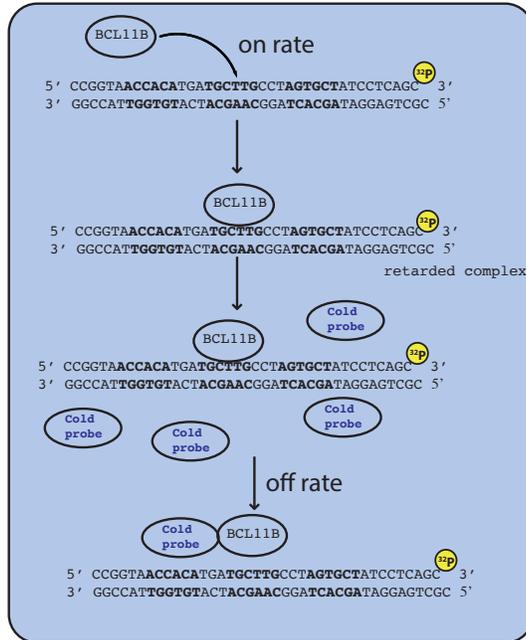
Fig. 6. BCL11B Binds DNA with Rapid On and Off Rate.

A) Schematics of the on/off rate electrophoretic mobility shift assay. The bolded sequences are BCL11B binding motifs based on ChIP-seq analysis¹³¹.

B) Left: 9 ng of purified BCL11B protein was incubated with radiolabeled oligonucleotides at room temperature for the indicated time points to determine the on rate for DNA binding.

Right: The off rate experiment was performed by allowing purified BCL11B protein to bind to radiolabeled oligonucleotides for 15 min to reach equilibrium at room temperature (see lane 4). Then 1000-fold molar excess of unlabeled oligonucleotides was added and allowed to incubate for the indicated amount of time. The reaction mixtures were resolved on a non-denaturing polyacrylamide gel and analyzed for EMSA.

A)



B)

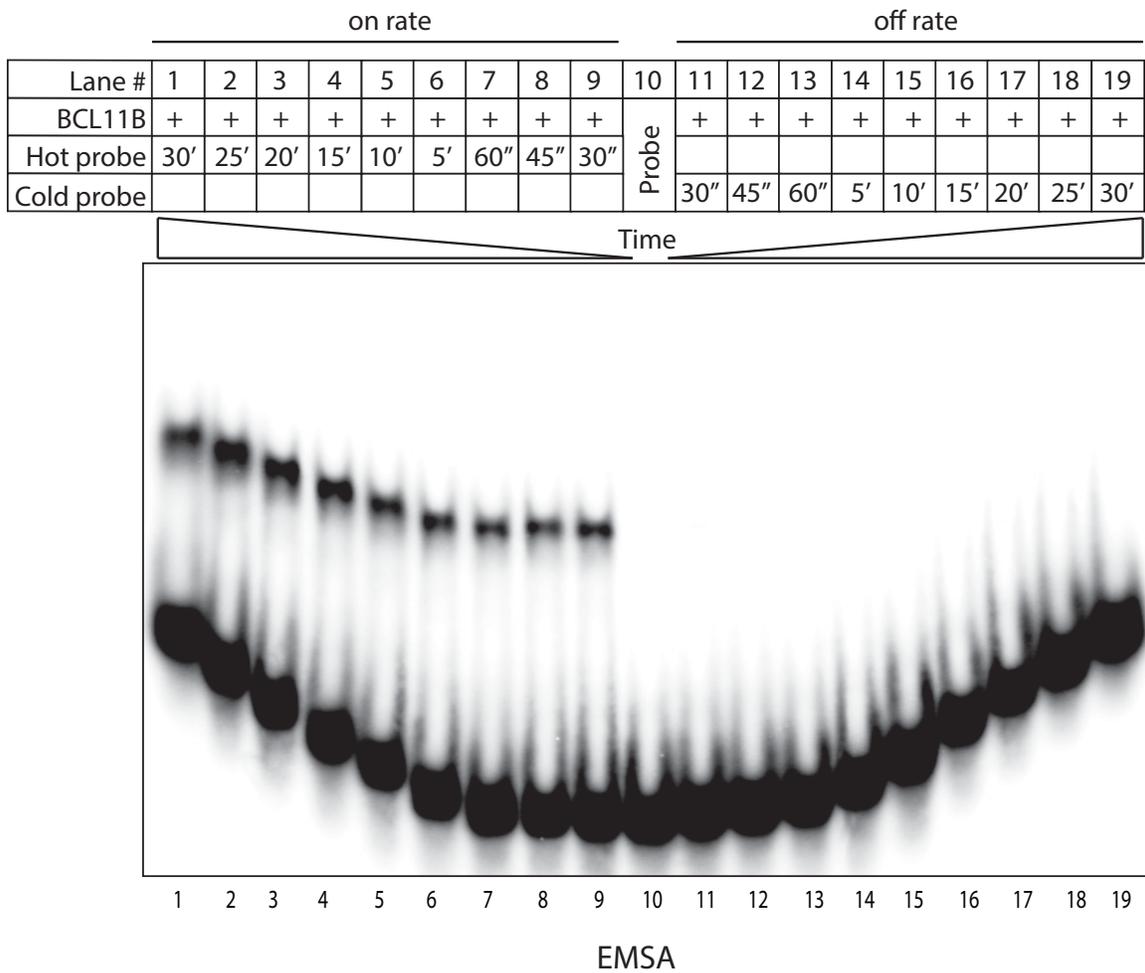


Fig. 7: Calculations of Apparent Equilibrium Dissociation Constant (K_{Dapp}) of BCL11B for Binding to DNA with a Normal or an Oxidized Thymine.

A) Sequence of the radiolabeled oligonucleotides used in the EMSA experiment. The underlined base is thymine in WT probe (wild type) and thymine glycol in Tg probe.

B) 10 pM of radiolabeled oligonucleotides, WT or Tg probes, containing conserved BCL11B binding sites were incubated with varying concentrations of BCL11B fusion proteins and analyzed in EMSA.

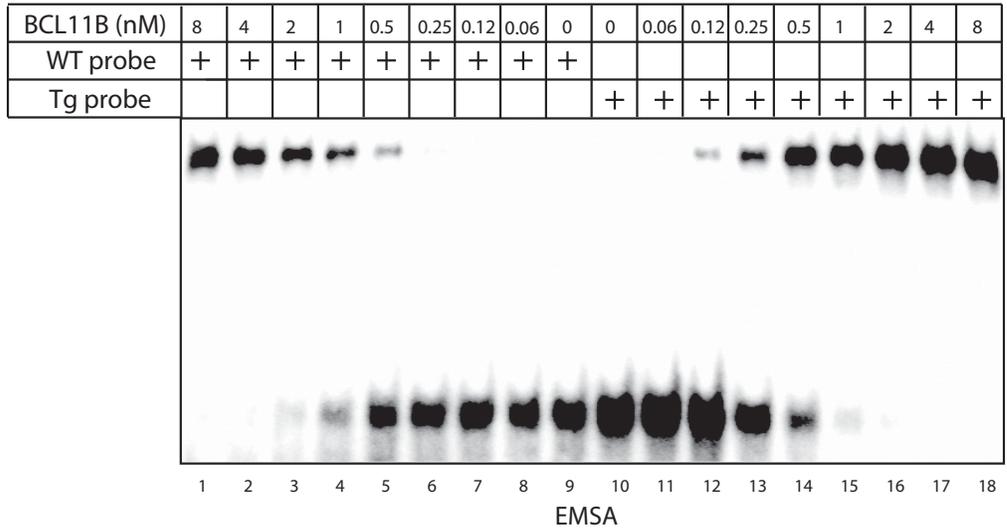
C) The percentage of free DNA in each lane was plotted against protein concentration.

D) Calculation of $K_{D(app)}$ to DNA containing oxidized and wildtype thymine.

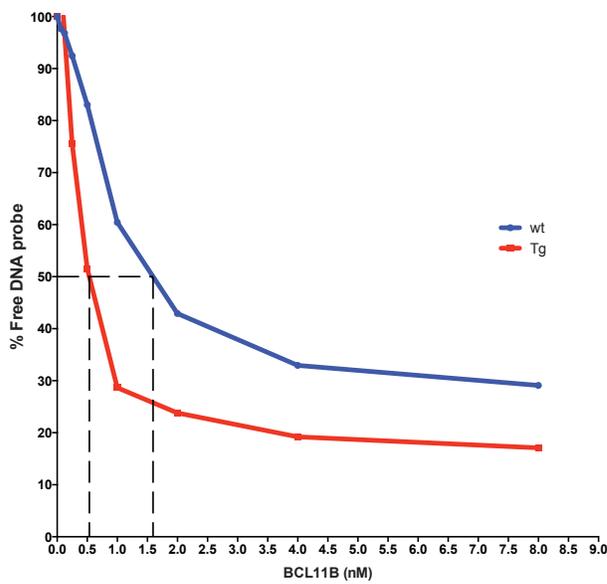
A)



B)



C)



D)

$$K = \frac{[\text{Free DNA}] [P]}{[\text{Bound DNA}]}$$

$$\therefore K = [P], \text{ when } [\text{Free DNA}] = [\text{Bound DNA}]$$

% Free DNA Probe Calculation:
 Tg probe $K_{d_{app}} = 5.0 \times 10^{-10} \text{ M}$

wt probe $K_{d_{app}} = 1.6 \times 10^{-9} \text{ M}$

Fig. 8: BCL11B Stimulates the Polymerase and Strand-Displacement Activities of DNA Pol β .

A) The nicked probe was designed by annealing a 17-mer and a 15-mer forward oligonucleotide with a 32-mer reverse oligonucleotide.

B) Schematics of the fluorescent strand-displacement assay. A FAM fluorophore conjugated with the forward oligonucleotide is placed adjacent to a quencher. Upon incubation with Pol β and dTTPs, in the presence or BSA or purified BCL11B protein, the polymerase and strand-displacement activities of Pol β release the short oligo with the fluorophore to generate fluorescence.

C) The nicked fluorophore-based probe was incubated with Pol β (25 nM) and 5 or 10 nM of BSA or recombinant BCL11B protein, in the presence or absence of dTTP.

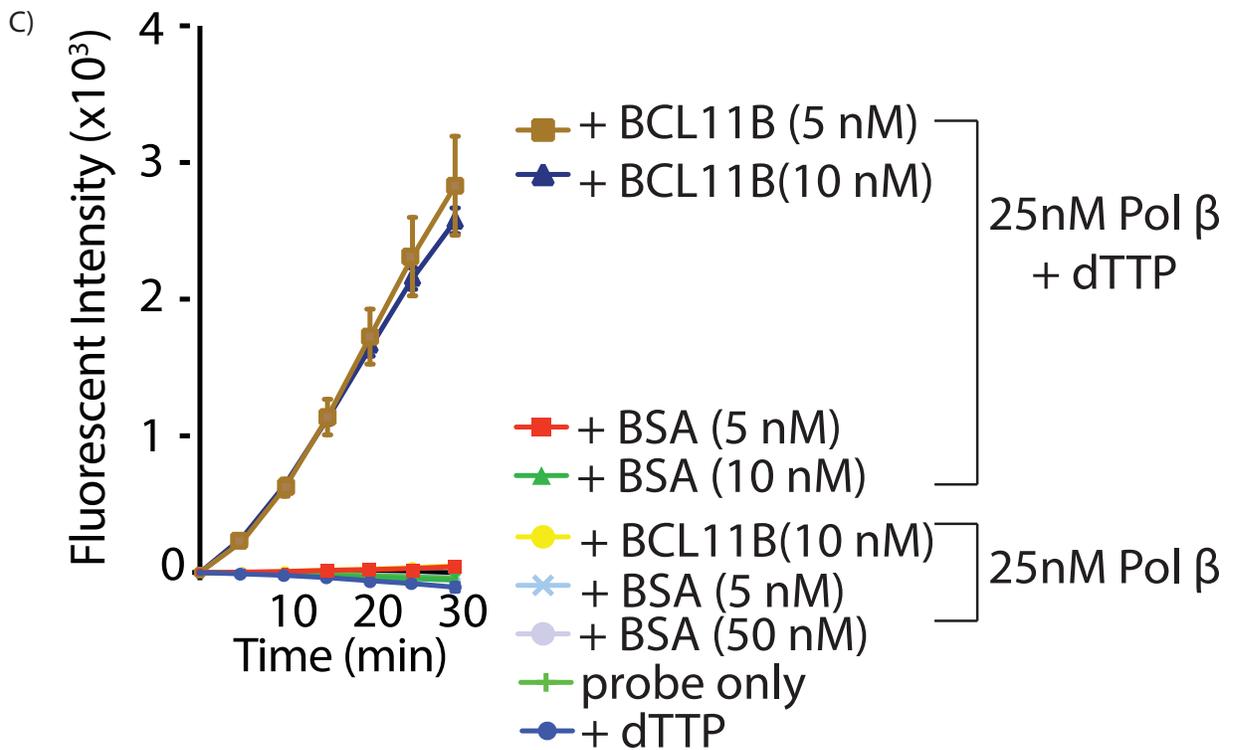
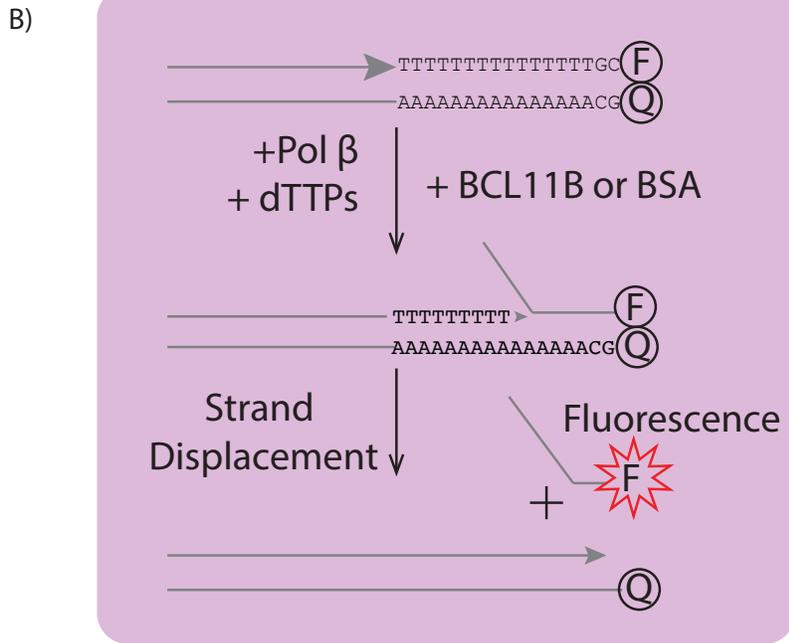
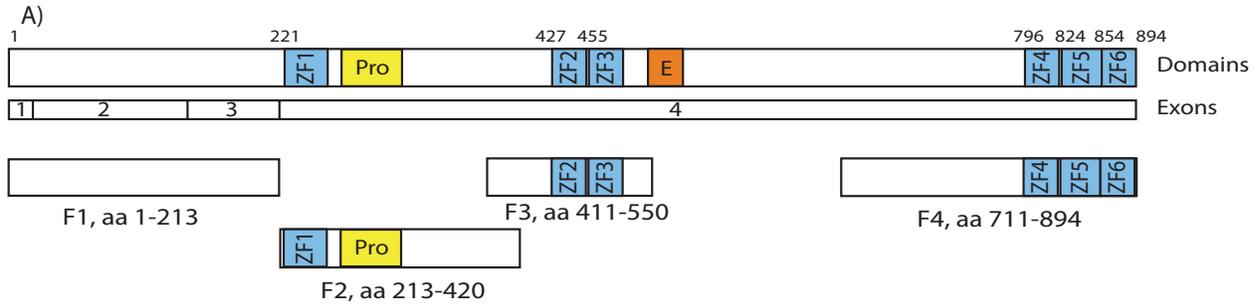


Fig. 9: Amino Acids 213-420 of BCL11B Are Sufficient to Stimulate the Enzymatic Activities of NTHL1.

A) Diagrammatic representation of the four BCL11B protein fragments. ZF represents zinc fingers, Pro represents a proline rich region, and E is a glutamic acid rich region. The four BCL11B fragments were expressed with a His-tag and purified using immobilized metal affinity chromatography.

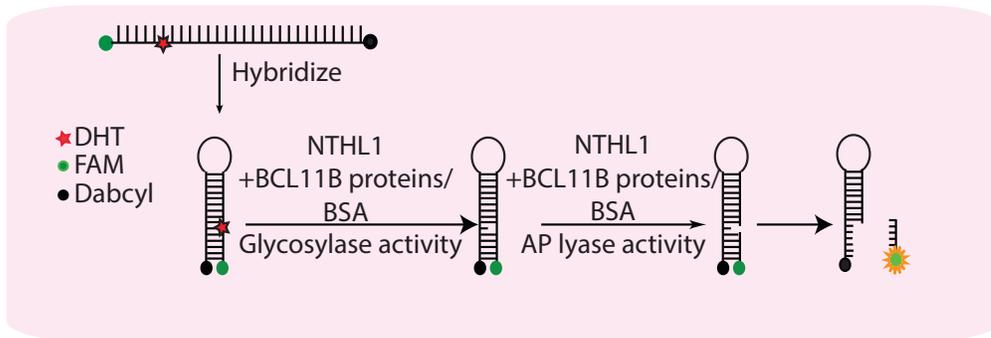
B) Schematics of the fluorescent cleavage assay.

C) NTHL1 cleavage assay was performed in triplicates using 10nM of NTHL1 enzyme and 10 nM of other proteins, as indicated. Error bars represent standard deviation. Curves were plotted using Prism (6.0) Graphpad.



5'-(FAM)-CCACTXTTGAATTGACAGCCATGTCGATCAATTCAATAGTGG-(Dabcyl)-3'
 where X is an oxidized thymine, dihydrothymine (DHT)

B)



C)

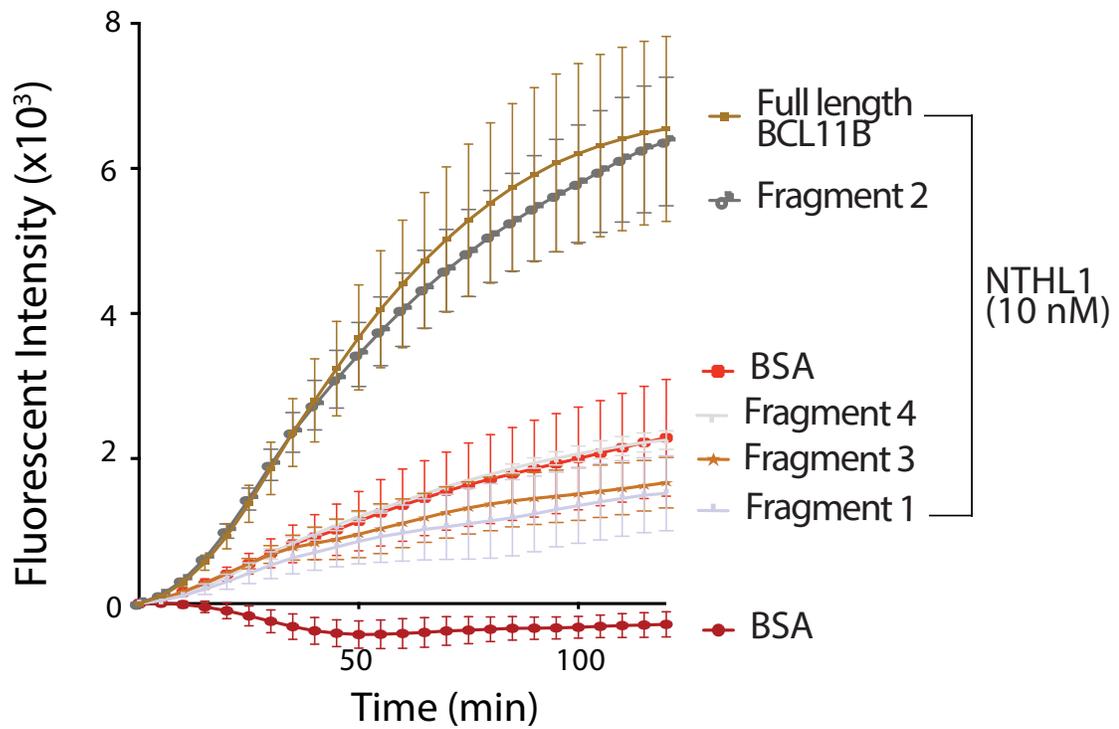
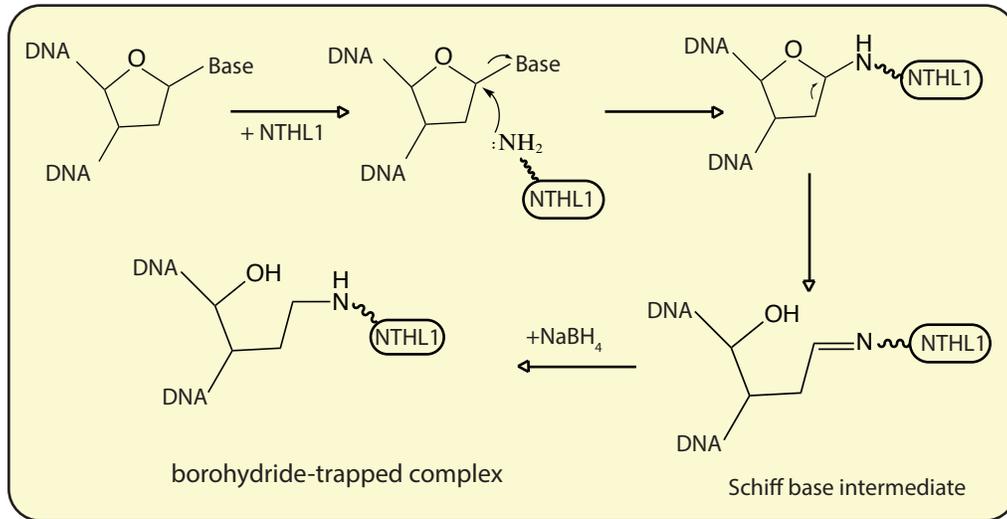


Fig. 10: BCL11B Recombinant Protein Fragments Stimulates Schiff Base Formation by NTHL1.

A) Schematics of the Schiff base assay.

B) Radioactively end-labeled double-stranded oligonucleotides containing a thymine glycol (Tg) base were incubated with his-NTHL1 protein, in the presence of his-tagged BCL11B protein fragments or BSA. FL is the BCL11B full length protein, F1-4 are the four BCL11B protein fragments. The Schiff base complex was trapped using 50 mM sodium borohydride, and the trapped complex was separated from free substrate using 10% SDS-PAGE.

A)



B)

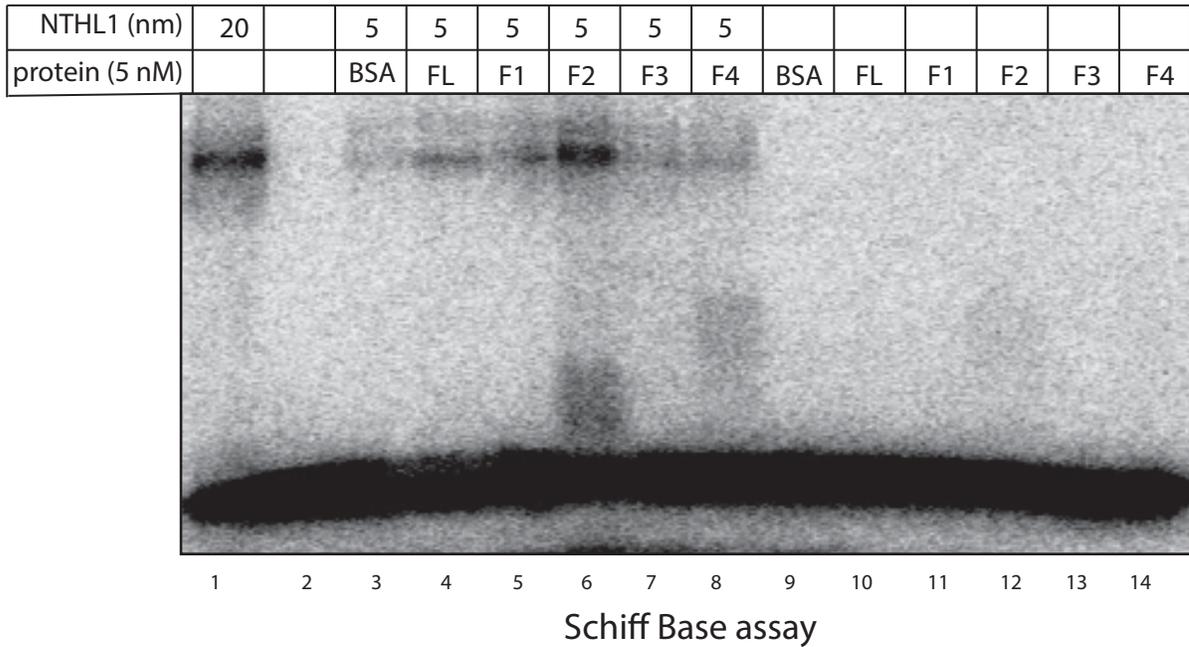
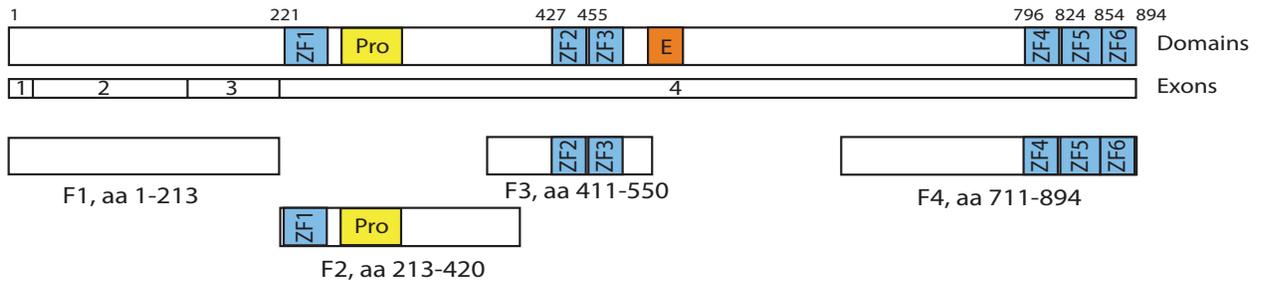


Fig. 11: BCL11B Recombinant Protein Fragments Interacts with NTHL1 in Pull Down Assay.

A) Diagrammatic representation of the four BCL11B protein fragments.

B) GST-pull down assay was performed using purified GST-NTHL1 and His-tagged BCL11B fragments. GST-tagged proteins, GST-NTHL1 and GST, were immobilized using glutathione beads and incubated with four his-tagged BCL11B protein fragments. GST-tagged and His-tagged proteins were visualized by immunoblotting.

A)



B)

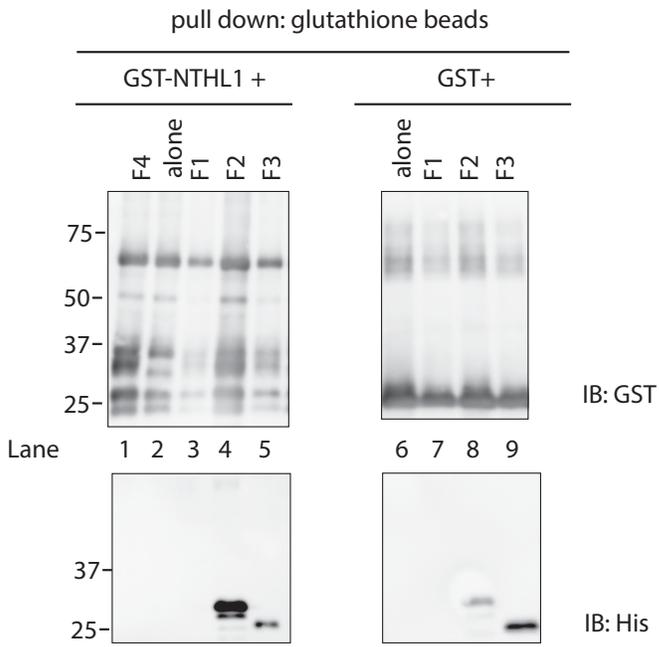


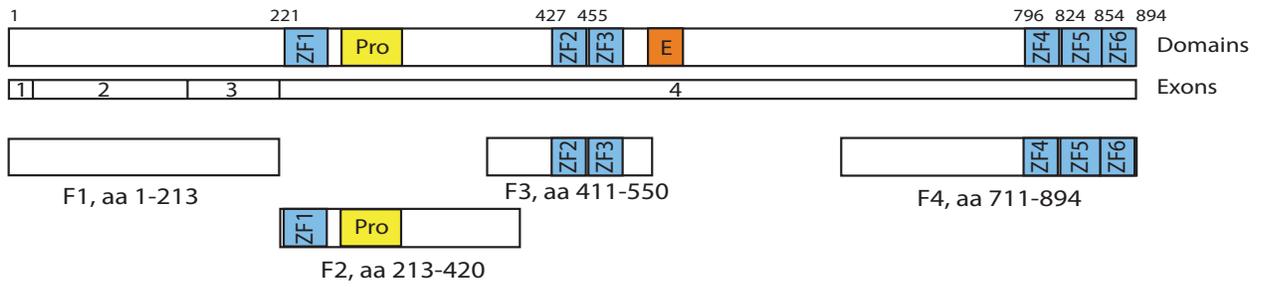
Fig. 12: C-terminal Zinc Fingers of BCL11B are Responsible for DNA Binding.

A) Diagrammatic representation of the four BCL11B protein fragments.

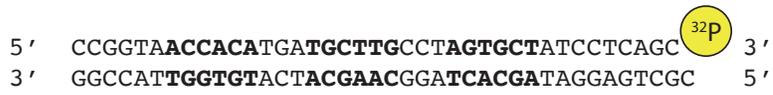
B) Sequence of radiolabeled probe used in EMSA, the bolded regions are the putative BCL11B binding sites.

C) EMSA experiment performed by incubating 3 or 9 ng of various BCL11B proteins with radiolabeled probe for 15 minutes.

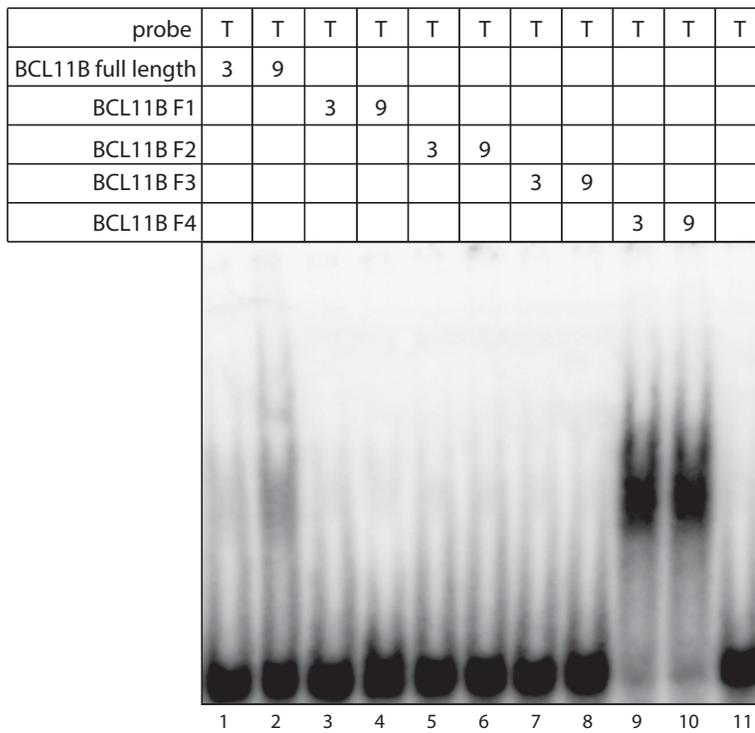
A)



B)



C)



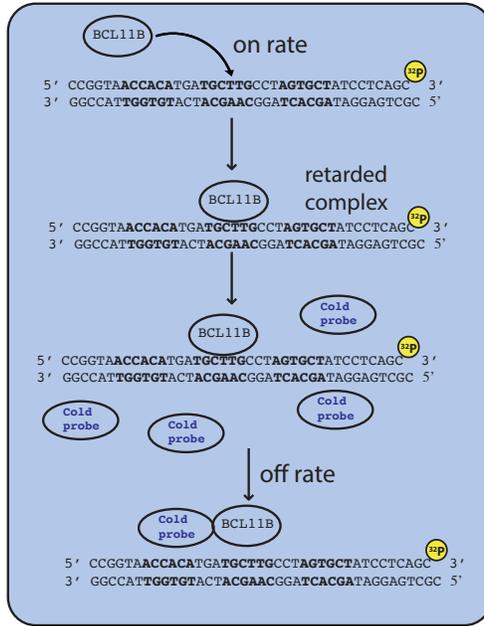
EMSA

Fig. 13: C-terminal Zinc Fingers of BCL11B Binds DNA with Rapid On and Off Rate.

A) Schematics of the on and off rate experiment.

B) Purified BCL11B fragment 4, containing three C-terminal zinc fingers, was incubated with radiolabeled oligonucleotides at room temperature for the indicated time points to determine the on rate for DNA binding. The off rate experiment was performed by allowing purified BCL11B fragment proteins to bind to radiolabeled oligonucleotides for 15 min to reach equilibrium at room temperature. Then 1000-fold molar excess of unlabeled oligonucleotides was added and allowed to incubate for the indicated amount of time. The reaction mixtures were resolved on a non-denaturing polyacrylamide gel.

A)



B)

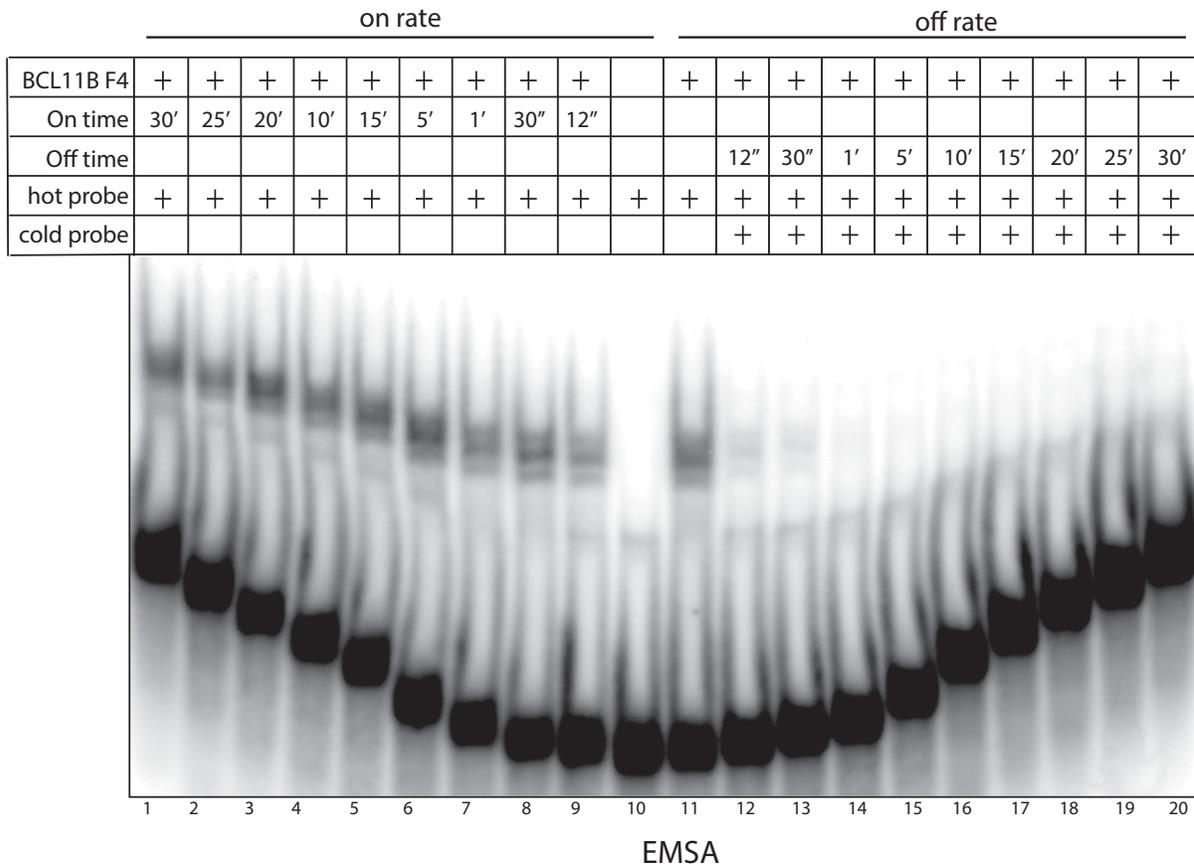


Fig. 14: BCL11B is Rapidly Recruited to Laser Micro-Irradiation-Induced DNA Lesions.

A) U2OS cells stably expressing a GFP-BCL11B fusion protein from a lentiviral vector were submitted to 405 nm laser micro-irradiation.

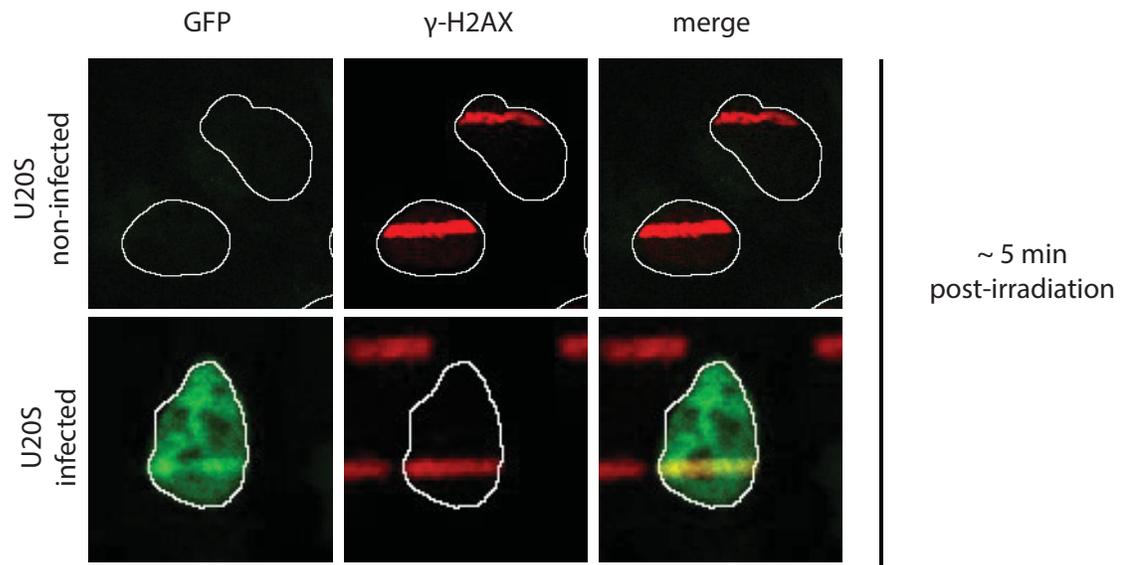
B) U2OS cells were fixed and stained for γ -H2AX by immunofluorescence at 5 min post irradiation. The GFP signal was obtained and the two images were merged.

C) Live-imaging of U2OS cells imaged once per minute post-irradiation. The 00:00 time point was taken before micro-irradiation. Scale bars = 10 μ m.

A)



B)



C)

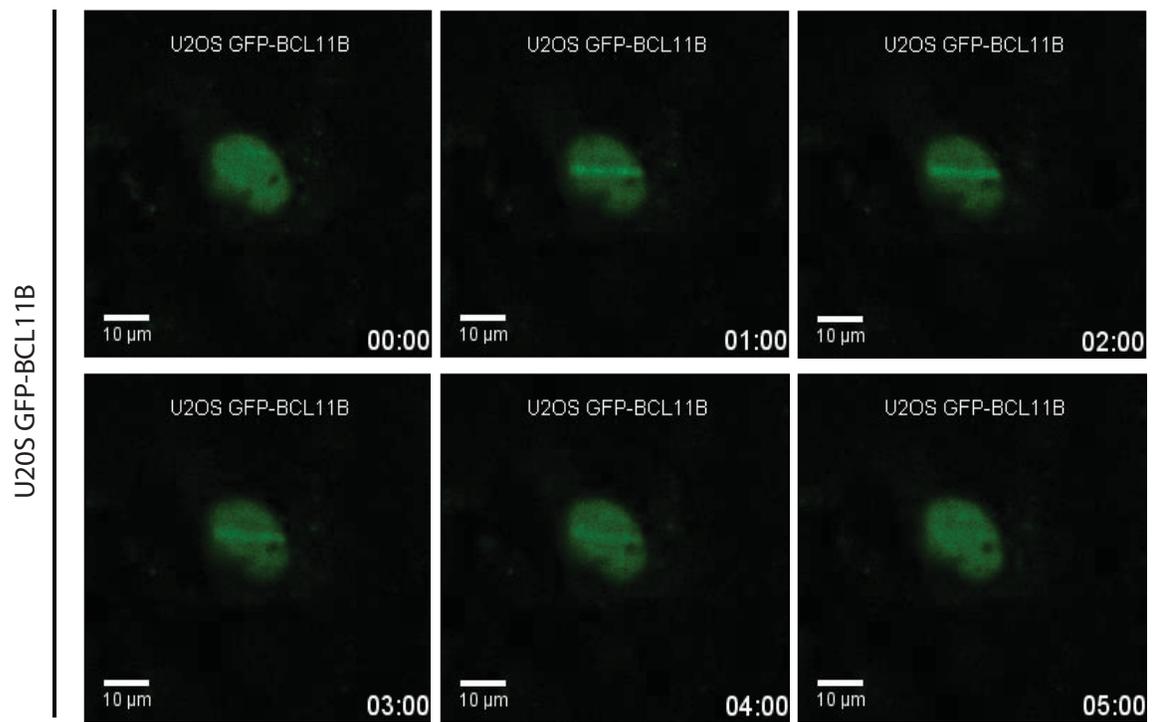


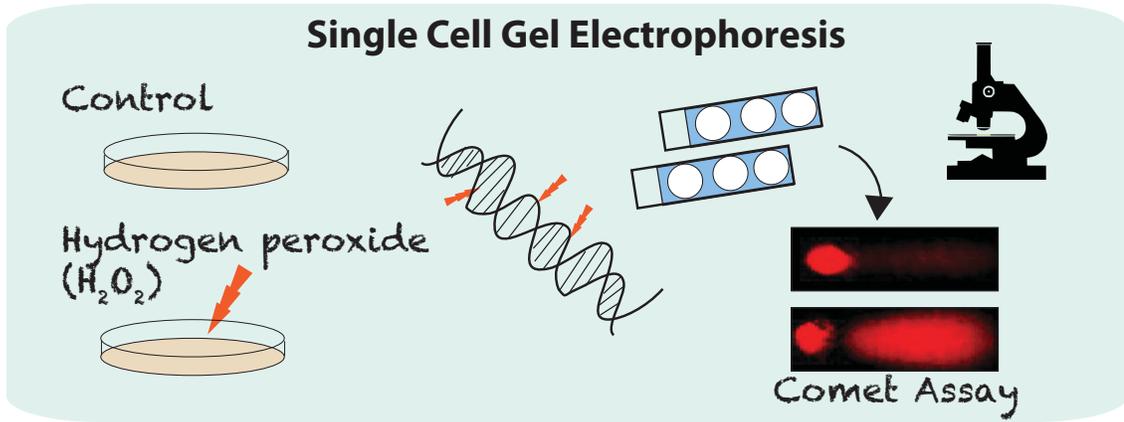
Fig. 15: BCL11B Knockdown in Jurkat cells Decreases DNA Repair Efficiency.

A) Schematics of single cell gel electrophoresis.

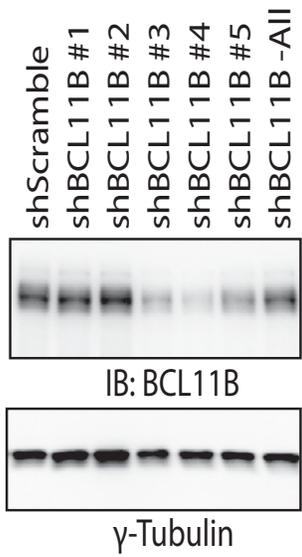
B) Immunoblot of BCL11B expression in Jurkat cells infected with a set of five pLKO.1 vectors each expressing a specific BCL11B shRNA.

C) Cells were exposed to 50 μ M of H₂O₂ for 20 min and allowed to recover for the indicated time before carrying out single cell gel electrophoresis at pH > 13. Comet tail moments were scored for at least 50 cells per condition. Error bars represent standard error. *p<0.05, **p<0.01, ***p<0.001; Student's t-test.

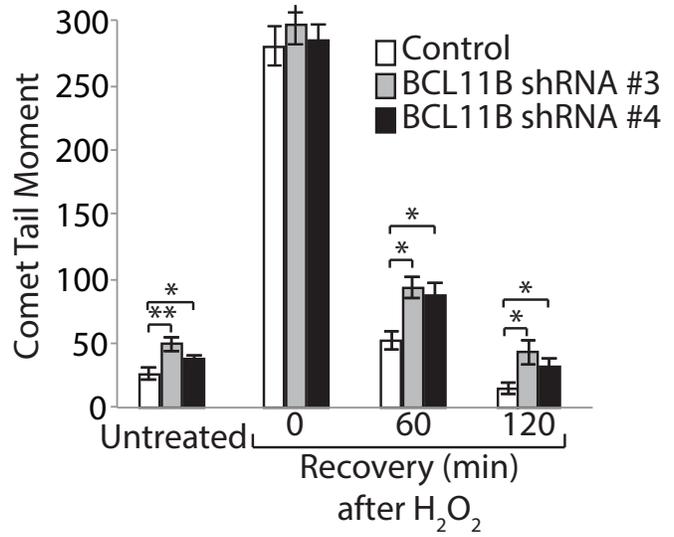
A)



B)



C)



5. Tables

Table 1: NTHL1 Preys Identified in Unirradiated and Irradiated Cells.

The table lists all proteins (preys) identified from mass spectrometry with a BFDR below 0.2. Preys in the table is sorted in the order of increasing BFDR value in unirradiated cells (0 Gy), starting from “0” to 0.2. Some preys are identified from irradiated cells (2 Gy) but not in unirradiated cells, they were not given a BFDR value, and are located towards the bottom of the table. Cases left blank in this column of BFDR 2 Gy indicate that the prey was not identified in irradiated cells.

Prey Gene	BFD R 0 Gy	BFDR 2 Gy	Prey Sequence Length (aa)	Notes
ABCE1	0	0.16	599	ATP-binding cassette
ACOT1	0	0	421	Acyl-coenzyme thioesterase 1
AFG3L2	0	0	797	AFG-like 3
ARID5B	0	0	1188	AT-rich interaction domain
ATP5B	0	0	529	Mitochondrial ATP synthase
ATPAF1	0	0	283	
BAZ2B	0	0	2168	Acetyl-transferase, bromodomain + Zn finger
BCL11A	0	0	773	B-Cell CLL/Lymphoma 11A, not a tumor suppressor
CBX2	0	0	532	Chromobox 2, Polycomb complex
CBX8	0	0	389	Chromobox 8, Polycomb complex
CCDC71L	0	0	235	Coiled-coil domain
CIZ1	0	0	842	Zn finger CDKN1A, p21 interactor
CLPX	0	0	633	Mitochondrial protease
DBT	0	0	482	Mitochondrial acetyltransferase
DDX39A	0	0	427	DEAD box RNA helicase
DDX56	0	0.01	507	DEAD box RNA helicase
DHX30	0	0	1155	DEAD box RNA helicase
ETFA	0	0	333	Electron transfer flavoprotein
FLYWCH1	0	0	703	FLYWCH-type Zn finger
GNL3L	0	0	582	Guanine nucleotide binding protein-like 3-like protein
GUF1	0	0	669	Translation Factor GUF1, mitochondrial
H2AFX	0	0.14	143	H2AX
HSD17B10	0	0	252	3-hydroxyacyl-COA dehydrogenase type-2
IARS2	0	0	1012	isoleucine—tRNA ligase, mitochondrial
KDM5A	0	0	1690	lysine-specific demethylase 5a
LETM1	0	0	739	letm1 and ef-hand domain-containing, mitochondrial
LRPPRC	0	0	1394	leucine-rich ppr motif-, mitochondrial
MAZ	0	0	493	ZNF801, myc- associated Zn
MBD4	0	0	574	Methyl-CpG binding protein
MRPL37	0	0	423	Mitochondrial

MRPL42	0	0.16	142	
MRPS24	0	0	167	
MRPS31	0	0	395	
MRPS9	0	0	396	
MYBL2	0	0.02	700	MYB-like 2
NDUFAF4	0	0.01	169	mitochondrial
NDUFV3	0	0	175	
NF1	0	0	2818	Neurofibromin 1, negative regulator of RAS
NFIA	0	0	509	
PHB2	0	0	299	Prohibitin 2
PHF2	0	0	1096	Lysine demethylase
PITX2	0	0	324	HD protein, bicoid type
POLRMT	0	0	1230	mitochondrial
PREPL	0	0.01	661	Serine protease
RBAK	0	0	714	RB Associated KRAB Zinc Finger
SHMT2	0	0	483	methyltransferase
SUV39H1	0	0	412	Suppressor of variegation
TACO1	0	0	297	
TIMM44	0	0	452	
TRAP1	0	0	651	
TRIP12	0	0	1992	
TRMT10C	0	0	403	
TSR1	0	0	804	
VAR52	0	0	923	
WDR76	0	0	562	DDR responder
ZMAT2	0	0.14	199	
ZNF12	0	0	659	
ZNF518A	0	0	1483	
ZNF91	0	0	1191	
ARL2	0.01	0.01	157	ADP-rybosylation like
BUD31	0.01	0	144	
FASTKD2	0.01	0	710	FAST kinase domain
FBL	0.01	0.03	321	Fibrillar nucleolar
GLS	0.01	0.01	598	mitochondrial
HOXD13	0.01	0.01	343	
LRIF1	0.01	0	769	Nuclear factor interacting
MIS18BP1	0.01	0	1132	MIS18 binding protein
REV3L	0.01	0.01	3130	

RLF	0.01	0.01	1914	Insulin-like 3
RPS7	0.01			ribosomal
RTEL1	0.01	0.15	1219	Telomere helicase
ZNF512B	0.01	0	892	amyotrophic lateral sclerosis
ZNF644	0.01	0	1327	Replication associated DNA damage
BAZ1A	0.02	0	1556	Zn finger
CDKAL1	0.02	0.01	579	
ELAC2	0.02	0.03	786	
ESCO2	0.02	0	601	
KIF22	0.02	0.01	597	
PBRM1	0.02	0.1	1582	
PHB	0.02	0	272	Prohibitin Tumor suppressor, protection against oxidative stress
POU2F1	0.02	0.02	755	POU domain
TFAM	0.02			Transcription factor A, Mitochondrial Methyl transferase
XRCC1	0.02	0.04	633	
ZKSCAN4	0.02	0.01	545	Zn finger KRAB domain
ZNF107	0.02			Zn finger KRAB domain
ZNF780A	0.02	0.04	641	Zn finger
ARHGAP11A	0.03	0.09	1023	Rho GTPase activating protein 11A
RALY	0.03	0.08	290	mRNA processing factor
RNF169	0.03	0.01	708	E3 ubiquitin-protein ligase RNF169
SALL1	0.03	0.01	1227	Spalt Like Transcription Factor 1 Zn finger
UNG	0.03	0.03	304	U DNA glycosylase
ZNF507	0.03	0.02	953	Zn finger
KDM2A	0.04	0.01	723	LYSINE-SPECIFIC DEMETHYLASE 2A
L3MBTL3	0.04	0.01	755	Lethal (3) malignant brain tumor-like protein 3, methyl-lysine binding repressor, deletion in medulloblastoma
SMCHD1	0.04	0.03	2005	Structural maintenance of chromosomes flexible hinge domain-containing protein 1
TARDBP	0.04	0.02	246	TAR DNA-binding protein 43
ZNF136	0.04	0.03	540	Zn finger
ZNF711	0.04	0.06	761	Zn finger
ZNF8	0.04	0.03	575	Zn finger
BEND3	0.05	0.08	828	BEN domain, repressor with NoRC, nucleolar remodeling complex

GTF3C3	0.05			General transcription factor 3C polypeptide 3
HLTF	0.05	0.04	1009	Helicase-like transcription factor, SWI/SNF member
UHRF1	0.05	0.05	793	E3 ubiquitin-protein ligase UHRF1
ZSCAN21	0.05	0.09	473	Zinc finger and SCAN domain containing protein 21
BRIX1	0.06			Biogenesis Of Ribosomes
CDCA8	0.06	0.05	280	Borealin, chromosomal passenger complex
KHDRBS1	0.06	0.2	404	K homology domain-containing, RNA-binding
PDS5B	0.06	0.11	1447	Cohesin Associated Factor B
PIAS1	0.06	0.06	651	protein inhibitor of activated STAT (PIAS) SUMO E3 ligase
ZNF195	0.06	0.06	606	Gemcitabine association
ZNF92	0.06			Zn finger
DDX24	0.07			
INCENP	0.07	0.04	918	Inner centromere protein
JADE3	0.07			
PRDM2	0.07			
RB1	0.07			
XPC	0.07	0.03	903	
ZNF174	0.07	0.07	407	
BCORL1	0.08	0.03	1711	BCL-6 corepressor-like protein 1
NOL10	0.08	0.09	662	Nucleolar Protein 10
NUSAP1	0.09	0.02	439	Nucleolar and spindle associated protein 1
ZBTB10	0.09			
ZNF516	0.09			
ATP5A1	0.1	0.06	503	
DGCR14	0.1	0.1	476	
PHF8	0.1	0.05	1060	PHD Zn finger lysine demethylase
ZNF770	0.1	0.03	691	Zn finger
CDC23	0.11			
CWC27	0.11	0.11	472	CWC27 Spliceosome Associated Protein Homolog
HNRNPAB	0.11			
LAS1L	0.11			
MCM3	0.11			
NUP107	0.11			
SUGT1	0.11	0.08	365	MIS12 Kinetochores Complex Assembly Cochaperone

ZNF280D	0.11			
CASZ1	0.12	0.09	1759	Castor Zn finger 1, tumor suppressor overexpressed in some cancer
LCORL	0.12			
PRPF31	0.12	0.16	499	Pre-mRNA Processing Factor 31
SMARCA1	0.12	0.12	1054	SWI/SNF Related, Matrix Associated, Actin-Dependent Regulator Of Chromatin
ATP5J2-PTCD1	0.13	0.01	749	PENTATRICOPEPTIDE REPEAT-CONTAINING PROTEIN 1, MITOCHONDRIAL-RELATED
BCL11B	0.13	0	823	B-Cell CLL/Lymphoma 11B, Zn finger, Haplo-insufficient tumor suppressor Ctip2 (chicken ovalbumin upstream promoter transcription factor (COUP-TF) interacting protein 2 Rit1 zinc finger protein hRit1 alpha
C17orf80	0.13	0	573	Uncharacterized protein
GRSF1	0.13	0	318	G-rich sequence factor 1, ribosomal
SSBP1	0.13	0.01	148	Single-stranded DNA-binding protein, mitochondrial
PIAS4	0.14			
STRBP	0.14			
ZNF462	0.14	0	2506	Zn finger
C1QBP	0.15	0	282	Complement component 1 Q subcomponent binding protein, mitochondrial
DNMT3A	0.15	0.15	912	DNA Methyltransferase 3 Alpha
FASTKD5	0.15			
HKR1	0.15	0	659	transmembrane
MRPL45	0.15	0.16	306	Mitochondrial Ribosomal Protein L45
MRPS22	0.15	0.01	360	Mitochondrial Ribosomal Protein S22, interacts with p53
NSD1	0.15	0.02	2696	Histone-lysine N-methyl transfer
NUDT19	0.15			
PTCD3	0.15	0.01	689	Pentatricopeptide repeat domain containing protein 3, mitochondrial
QRSL1	0.15			
RBM15	0.15			
ATAD2B	0.16	0.05	1390	ATPase Family, AAA Domain Containing 2B
BRD1	0.16	0.14	1058	Bromodomain Containing 1

CBX4	0.16			
DCLRE1A	0.16			
EHMT2	0.16	0.16	1210	Euchromatic Histone Lysine Methyltransferase 2
ENSEMBL:ENSBTAP00000038253	0.16			
ERAL1	0.16	0.01	437	GTPase Era, mitochondrial
ERICH1	0.16			
FKBP4	0.16			
H2AFZ	0.16	0	128	H2A Histone Family Member Z
HMGB3	0.16			
IBA57	0.16			
MORC2	0.16	0.01	970	MORC Family CW-Type Zinc Finger 2 Zn finger
MORC3	0.16			MORC Family CW-Type Zinc Finger 2 Zn finger
MRPL10	0.16	0.01	261	Mitochondrial Ribosomal Protein L10
MRPS28	0.16			
NFS1	0.16			
NOM1	0.16	0.15	860	Nucleolar Protein With MIF4G Domain 1
PDE12	0.16			
RBM19		0.02	960	Probable RNA binding protein 19
RECQL5	0.16			
RUNX1	0.16			
USP22	0.16	0	525	Ubiquitin carboxyl terminal hydrolase 22
ZNF160	0.16	0	818	Zn finger
CHD9	0.2			
CLSPN	0.2			
FOXC1	0.2			
LIG3	0.2			
TOP1	0.2	0.06	765	Topoisomerase (DNA) I
XRN2	0.2			
BRD7		0.09	652	Bromodomain Containing 7
C1orf131		0.16	293	Chromosome 1 Open Reading Frame 131
C9orf78		0.15	289	Chromosome 9 Open Reading Frame 78
CDCA5		0.15	252	Cell Division Cycle Associated 5, cohesin association
CENPC		0.04	943	Centromere protein C

CENPF		0.1	3114	Centromere protein F
CENPT		0.16	561	Centromere protein T
CHD6		0.1	2715	Chromodomain Helicase DNA Binding Protein 6
DNAJA2		0	412	DnaJ homolog subfamily A member 2, heat shock
FOXK1		0.05	733	Forkhead box protein K1
FOXM1		0	748	Forkhead box protein M1 Activates DDR genes, increases resistance Not a tumor suppressor except in the context of urethane-induced lung tumorigenesis Gagoski2016
GLTSCR2		0.05	478	Glioma tumor suppressor candidate region gene 2; DDR
HDGF		0.08	240	Hepatoma-Derived Growth Factor, DNA binding High Mobility Group Protein 1-Like 2 Not a tumor suppressor
HIST1H2AC		0.14	130	Histone Cluster 1 H2A Family Member C
HOXB9		0.16	250	Homeobox B9
HSPE1		0.01	102	Heat shock protein E1, mitochondrial
KIAA0391		0.16	567	Mitochondrial RNase P Subunit 3
KIAA1143		0.07	154	Not known
KIF18A		0.06	898	Kinesin Family Member 18A
KIFC1		0.08	673	Kinesin Family Member C1
NDUFV3		0	473	NADH dehydrogenase [ubiquinone] flavoprotein 3, mitochondrial
PRC1		0.01	525	Protein regulator of cytokinesis 1
PSIP1		0.03	530	PC4 and SFRS1-interacting protein
RBM22		0.01	420	RNA Binding Motif Protein 22, Zn finger
RBMX		0.08	196	RNA Binding Motif Protein, X-Linked
RECQL5		0.01	991	RecQ Like Helicase 5
SATB2		0.16	733	
SGO2		0.2	1260	Shugoshin 2, cohesin association
UIMC1		0.1	719	Ubiquitin Interaction Motif Containing 1
UTP3		0.2	479	UTP3, Small Subunit Processome Component Homolog
WIZ		0.03	794	Widely Interspaced Zinc Finger Motifs
ZBED4		0.01	1171	Zinc Finger BED-Type Containing 4
ZBTB4		0.01	1013	Zinc Finger And BTB Domain Containing 4

ZMYM3		0.07	1358	Zinc Finger MYM-Type Containing 3
ZNF106		0.01	1883	
ZNF226		0.03	803	
ZNF486		0.16	463	
ZNF850		0.16	1090	

6. Discussion

Mechanism of BCL11B Stimulation of NTHL1

Our BioID identification of BCL11B (Fig. 1), as well as the co-immunoprecipitation result (Fig. 2), both suggest interaction between BCL11B and NTHL1. Both the fluorescent cleavage assay (Fig. 3) and radioactive cleavage assay (Fig. 4) indicate that BCL11B stimulates the enzymatic activity of NTHL1. Since the two cleavage assays used different oxidative DNA modifications (DHT for fluorescent cleavage and Tg for radioactive cleavage sequence), DNA structure (hairpin for fluorescent cleavage, and double-stranded DNA for radioactive cleavage), DNA sequence, length of the sequence, and nature of the reporter, I am confident that our results were not artifacts impacted by these factors. Both experiments measured the end product, therefore we could not deduct which, if not both, activity of NTHL1, glycosylase or AP/lyase, is stimulated by BCL11B from the cleavage assays alone.

The Schiff base assay using the full length BCL11B protein (Fig. 5) shed some light on that aspect. The Schiff base assay captured the amount of Schiff base intermediates present at the time they were trapped by sodium borohydride. The Schiff base intermediate is produced before the AP/lyase activity occur, but after the removal of the damaged base has occurred. Our observations from the Schiff base assay shows that BCL11B increases the amount of Schiff base intermediates produced by NTHL1. Therefore, BCL11B must be able to stimulate at least one step of action for NTHL1 prior to Schiff base formation. Two possible steps are that BCL11B stimulates the glycosylase activity of NTHL1, or, it enhances the binding or recognition of NTHL1 to damaged DNA. Our work of measuring the binding affinity of wildtype DNA versus one with a thymine glycol modification showed preference of binding to damaged DNA (Fig. 7). It is possible that BCL11B enhances NTHL1 activity by assisting in the recognition and binding

of NTHL1 to damaged DNA. Unfortunately, from the experiments we performed, we were unable to determine whether this preference is specific to thymine glycol damage, or also applied to other modified bases NTHL1 repairs, for example, DHT. It is possible that BCL11B preferentially binds to damaged DNA in cells, and through interaction with NTHL1, BCL11B brings NTHL1 closer to the damaged base its bound to, thus increasing the efficiency of DNA repair. Our findings of the DNA binding dynamics of BCL11B also supports this idea. BCL11B binds DNA with extremely rapid on and off rate (Fig. 6), which is atypical for classical transcription factors. Transcription factors normally binds DNA stably over a period of time to regulate the expression of genes. A “jumpy” transcription factor seemingly defeats that purpose. I speculate that BCL11B actually scans the genome for damaged DNA, which explains the rapid association and dissociation of DNA. Whenever BCL11B senses DNA damage, due to its preference for binding, will linger in that region, and increases the chance of interacting with NTHL1 at the site of damage, thus repairing the damage.

However, from our structure/function analysis of BCL11B, the findings do not support a direct association between NTHL1 stimulation and DNA binding properties of BCL11B. In the four protein fragments of BCL11B we expressed, fragment 2 (residue 213-420) was the only fragment capable of NTHL1 stimulation in fluorescent cleavage assay (Fig. 9), as well as the only fragment stimulating Schiff base formation by NTHL1 (Fig. 10). In addition, fragment 2 also exhibit specific binding to purified GST-NTHL1 (Fig. 11). However, the EMSA experiment showed that fragment 4 (residue 711-849) was the only fragment capable of DNA binding, at least with the DNA probe we used (Fig. 12). In addition, similar to the full length protein, fragment 4 also exhibits rapid on and off rate for DNA binding (Fig. 13). Since fragment 2 doesn't possess DNA binding capability, our previous hypothesis of attributing NTHL1

stimulation to DNA binding alone is not correct. Based on our observations from the GST-pulldown assay, I speculate it is the protein-protein interaction between BCL11B and NTHL1 that is responsible for the stimulation of NTHL1 activities. Indeed, fragment 2 of BCL11B contains a proline rich region, which is reported to be responsible for protein binding and interactions in various families of proteins, including transcription factors¹⁶⁴. Unfortunately, there are currently no published crystal structure on BCL11B and human NTHL1 for us to postulate the nature of how this interaction can lead to stimulation of NTHL1 activities. It is possible that BCL11B binds to NTHL1 at specific regulatory sites or the active site, which lead to a change in activity. The crystal structure of prokaryotic orthologues of NTHL1 has been solved^{165, 166}. Based on these structures, the group of Robey-Bond *et al.* used site-directed mutagenesis and identified the residue Gln287 near the active site is important for catalysis¹⁶⁷. For future work, we can create smaller fragments of NTHL1 containing the catalytically active residues and verify their interactions with BCL11B.

Although fragment 4 doesn't stimulate NTHL1 in *in vitro* cleavage assay, we cannot deny the possibility of its involvement in NTHL1 stimulation. It is entirely likely the dual functions of fragment 2 and 4 of BCL11B work cohesively in stimulating the enzymatic activity of NTHL1. Published crystallographic, kinetics, and single molecules studies suggested DNA glycosylases locate and identify DNA damage in the genome by one dimensional diffusion rotationally along DNA helix¹⁶⁸⁻¹⁷⁰. This is a daunting task for NTHL1 given the size of the human genome, and it is possible that BCL11B assists NTHL1 in damage recognition. I propose that both fragment 2 and 4 of BCL11B play important role in NTHL1 stimulation: fragment 4 with DNA binding ability scans and binds to damaged DNA, and fragment 2 is responsible for interacting with NTHL1, and assisting in recognition of damaged base, as well as direct

stimulation of activity by interaction. It is important to note that the *in vitro* cleavage assays were done in small volume with high concentration of proteins and DNA probe, and NTHL1 can find its substrate with relative ease; whereas in the cells, NTHL1 needs to navigate through a network of proteins and complex chromosomal structures.

Impact of BCL11B on DNA Repair in Cells

While our *in vitro* work showed interaction and stimulation of NTHL1 by BCL11B, our cell line works cemented the importance of BCL11B in DNA repair. In our micro-irradiation experiment, we saw clear evidence of BCL11B migration to sites of DNA damage. This migration was rapid and directional, triggered by the laser-induced DNA damage, as recruitment was observed within 60s (Fig. 14). Given that NTHL1 participates in the initiation of base excision repair, this observation was not surprising. The group Lan *et al.* published a study in 2004 where they induced DNA damage in mammalian cells using 365-nm UVA irradiation and examined the recruitment of various enzymes involved in repair of oxidative damage¹⁷¹. They reported the recruitment of NTHL1 to UVA-induced damage in as little as thirty seconds. Both NTHL1 and BCL11B are recruited to site of damage in similar time frames post-damage, which agrees with our hypothesis of BCL11B assisting NTHL1 during base excision repair.

Next, in our single cell gel electrophoresis experiment, we observed a difference in DNA damage level (represented by comet score), between our control cell line and a cell line with BCL11B knockdown in Jurkat cells (Fig. 15). Moreover, following treatment with H₂O₂, an agent that causes oxidative DNA damage, cells in which BCL11B expression were reduced exhibited a delay in DNA repair (Fig. 15). Jurkat is a T-ALL cell line with BCL11B overexpression. A previous study showed that BCL11B knockdown is synthetic lethal in Jurkat

cells, while it does not affect normal T cells¹⁷². The authors of this study, Grabarezyk *et al.* showed 50% of Jurkat cells undergo apoptosis 72 hr after transfection of BCL11B siRNA. In addition, using global gene expression profiling, they found upregulation of TRAIL (the tumor necrosis factor-related apoptosis-inducing ligand) mRNA and downregulation of the anti-apoptotic BCLxL mRNA in cells with reduced BCL11B expression. From these findings, the authors concluded BCL11B is involved in regulation of apoptosis and is anti-apoptotic in nature. As a result, its suppression in cells induces apoptosis, which explains the synthetic lethality of its knockdown in Jurkat cells, as well as the paradoxical role of BCL11B in cancer.

However, the findings presented in this thesis proposed an alternative explanation to the observations by Grabarezyk *et al.* BCL11B is an accessory factor in base excision repair. Therefore, reduction of BCL11B expression decreases the efficiency of DNA repair. Cancer cells are under constant oxidative stress as a result of cancer-associated metabolic changes. As a result, cancer cells become reliant on enhanced DNA repair to combat the oxidative DNA damage. We propose cancer cells with BCL11B overexpression, such as Jurkat cells, exhibit non-oncogene addiction to BCL11B, and become dependant on BCL11B for efficient DNA repair. When BCL11B expression is reduced in Jurkat cells, they lost their DNA repair efficiency, and cells undergo apoptosis as a result of excessive DNA damage. Therefore, at baseline level of DNA damage, without H₂O₂ treatment, reduction of BCL11B expression already had an impact on DNA repair efficiency (Fig. 15). This impact was also observed after we induced DNA damage with hydrogen peroxide and allowed the cells to recover over time. Hydrogen peroxide is responsible for a large amount of oxidative damage and generation of single-strand breaks, both of which are repaired by base excision repair. However, we would not

be surprised if BCL11B is also involved in other pathways of DNA repair, since many DNA repair proteins are found in more than one repair pathway.

Potential for Novel Classification of Cancer Genes

While reviewing the literatures on the various BioID candidates for NTHL1, we decided to investigate further with BCL11B for its many similarities to CUX1. Both CUX1 and BCL11B are transcription factors and classified as haplo-insufficient tumour suppressors. Furthermore, both tumour suppressors have increased expression or copy number in cancer cells that's paradoxical to the role of traditional tumor suppressors. In addition, both proteins are found to participate and enhance DNA repair. This leads to the interesting and exciting possibility that they belong to a special class of cancer genes that's beyond the traditional classification of oncogenes and tumour suppressors. During early stages of tumorigenesis, these genes function as tumor suppressors, and are inactivated to decrease DNA repair efficiency, thus contributing to tumor initiation and progression. However, as the DNA damage burden increases in cancer cells due to their abnormal metabolism, they become dependent on DNA repair to maintain the integrity of the genome. They exhibit non-oncogene addiction to these accessory factors, which explains the phenomenon that some cancer cell lines are synthetic lethal to BCL11B knockdown.

In addition, the discovery of these unique role of accessory factors in cancer opens doors for novel therapeutic options. Inhibitors of these accessory factors can reduce the efficiency of base excision repair. Therefore, these inhibitors could be used in combination therapy to sensitize currently available therapeutic options. Indeed, studies have shown overexpression of CUX1 and BCL11B in cells contribute to resistance of chemo- and radio-therapy, and reduction of their expression by shRNA sensitizes cells to these therapy^{45, 162}.

7. Conclusion

In conclusion, utilizing the powerful technique of BioID, we identified BCL11B as an accessory factor of NTHL1. We demonstrated protein-protein interaction between BCL11B and NTHL1. We also have convincing evidence of BCL11B stimulating the enzymatic activities of NTHL1 in two different *in vitro* assays. We studied the DNA binding kinetics of BCL11B, and showed that BCL11B binds DNA with extremely rapid association and dissociation rate. In addition, we showed evidence of BCL11B preferentially binds DNA sequence with thymine glycol modification. We also performed structure/function analysis of BCL11B, and identified regions of BCL11B responsible for stimulating NTHL1 and DNA binding. Based on our *in vitro* findings, we proposed plausible mechanisms of how BCL11B stimulates NTHL1 activities. Our cell line works demonstrated recruitment of BCL11B to sites of laser-induced DNA damage. Using single cell gel electrophoresis, we showed evidence that BCL11B knockdown in Jurkat cells reduces the DNA repair efficiency. Lastly, we proposed the possibility of a new class of cancer genes with paradoxical roles during progression of cancer, based on work presented on this thesis, as well as past work on CUX1.

8. Reference

- [1] Hansen, W. K., and Kelley, M. R. (2000) Review of Mammalian DNA Repair and Translational Implications, *Journal of Pharmacology and Experimental Therapeutics* 295, 1.
- [2] Stokes, M., and Comb, M. (2008) A wide-ranging cellular response to UV damage of DNA, *Cell Cycle* 7, 2097-2099.
- [3] Jackson, A. L., and Loeb, L. A. (2001) The contribution of endogenous sources of DNA damage to the multiple mutations in cancer, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 477, 7-21.
- [4] De Bont, R., and Van Larebeke, N. (2004) Endogenous DNA damage in humans: a review of quantitative data, *Mutagenesis* 19, 169-185.
- [5] Lindahl, T., and Barnes, D. (2000) Repair of endogenous DNA damage, In *Cold Spring Harbor symposia on quantitative biology*, pp 127-134, Cold Spring Harbor Laboratory Press.

- [6] Dexheimer, T. S. (2013) DNA Repair Pathways and Mechanisms, In *DNA Repair of Cancer Stem Cells* (Mathews, L. A., Cabarcas, S. M., and Hurt, E. M., Eds.), pp 19-32, Springer Netherlands, Dordrecht.
- [7] Lindahl, T. (1993) Instability and decay of the primary structure of DNA, *nature* 362, 709-715.
- [8] David, S. S., O'Shea, V. L., and Kundu, S. (2007) Base-excision repair of oxidative DNA damage, *Nature* 447, 941.
- [9] Yonekura, S.-I., Nakamura, N., Yonei, S., and Zhang-Akiyama, Q.-M. (2008) Generation, biological consequences and repair mechanisms of cytosine deamination in DNA, *Journal of radiation research*, 0811040073-0811040073.
- [10] Kow, Y. W. (2002) Repair of deaminated bases in DNA, *Free Radical Biology and Medicine* 33, 886-893.
- [11] Patel, R. P., McAndrew, J., Sellak, H., White, C. R., Jo, H., Freeman, B. A., and Darley-Usmar, V. M. (1999) Biological aspects of reactive nitrogen species, *Biochimica et Biophysica Acta (BBA)-Bioenergetics* 1411, 385-400.
- [12] Murphy, Michael P. (2009) How mitochondria produce reactive oxygen species, *Biochemical Journal* 417, 1-13.
- [13] Apel, K., and Hirt, H. (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction, *Annu. Rev. Plant Biol.* 55, 373-399.
- [14] Lambeth, J. D. (2004) NOX enzymes and the biology of reactive oxygen, *Nature Reviews Immunology* 4, 181-189.
- [15] Sun, Y. (1990) Free radicals, antioxidant enzymes, and carcinogenesis, *Free Radical Biology and Medicine* 8, 583-599.
- [16] Cadet, J., Berger, M., Douki, T., and Ravanat, J.-L. (1997) Oxidative damage to DNA: formation, measurement, and biological significance, In *Reviews of Physiology Biochemistry and Pharmacology, Volume 131*, pp 1-87, Springer.
- [17] Yang, J. Q., Li, S., Huang, Y., Zhang, H. J., Domann, F. E., Buettner, G. R., and Oberley, L. W. (2001) V-Ha-Ras overexpression induces superoxide production and alters levels of primary antioxidant enzymes, *Antioxidants & redox signaling* 3, 697-709.
- [18] Maciag, A., Sithanandam, G., and Anderson, L. M. (2004) Mutant K-ras V12 increases COX-2, peroxides and DNA damage in lung cells, *Carcinogenesis* 25, 2231-2237.
- [19] Mitsushita, J., Lambeth, J. D., and Kamata, T. (2004) The superoxide-generating oxidase Nox1 is functionally required for Ras oncogene transformation, *Cancer research* 64, 3580-3585.
- [20] Major, G. N., and Collier, J. D. (1998) Repair of DNA lesion O 6-methylguanine in hepatocellular carcinogenesis, *Journal of hepato-biliary-pancreatic surgery* 5, 355-366.
- [21] McCulloch, S. D., and Kunkel, T. A. (2008) The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases, *Cell research* 18, 148.
- [22] Shimizu, M., Gruz, P., Kamiya, H., Kim, S. R., Pisani, F. M., Masutani, C., Kanke, Y., Harashima, H., Hanaoka, F., and Nohmi, T. (2003) Erroneous incorporation of oxidized DNA precursors by Y-family DNA polymerases, *EMBO reports* 4, 269-273.
- [23] Goodman, M. F. (2002) Error-prone repair DNA polymerases in prokaryotes and eukaryotes, *Annual review of biochemistry* 71, 17-50.
- [24] Rattray, A. J., and Strathern, J. N. (2003) Error-prone DNA polymerases: when making a mistake is the only way to get ahead, *Annual review of genetics* 37, 31-66.

- [25] Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication, *Cell* 107, 91-102.
- [26] Pourquier, P., and Pommier, Y. (2001) Topoisomerase I-mediated DNA damage.
- [27] McClendon, A. K., and Osheroff, N. (2007) DNA topoisomerase II, genotoxicity, and cancer, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 623, 83-97.
- [28] Ravanat, J.-L., Douki, T., and Cadet, J. (2001) Direct and indirect effects of UV radiation on DNA and its components, *Journal of Photochemistry and Photobiology B: Biology* 63, 88-102.
- [29] Borrego-Soto, G., Ortiz-López, R., and Rojas-Martínez, A. (2015) Ionizing radiation-induced DNA injury and damage detection in patients with breast cancer, *Genet Mol Biol* 38, 420-432.
- [30] Ward, J. F. (1988) DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability, In *Progress in nucleic acid research and molecular biology*, pp 95-125, Elsevier.
- [31] Téoule, R. (1987) Radiation-induced DNA Damage and Its Repair, *International Journal of Radiation Biology and Related Studies in Physics, Chemistry and Medicine* 51, 573-589.
- [32] Kukielka, E., and Cederbaum, A. (1992) The effect of chronic ethanol consumption on NADH- and NADPH-dependent generation of reactive oxygen intermediates by isolated rat liver nuclei, *Alcohol and alcoholism (Oxford, Oxfordshire)* 27, 233-239.
- [33] Halliwell, B., and Gutteridge, J. M. (2015) *Free radicals in biology and medicine*, Oxford University Press, USA.
- [34] Irigaray, P., and Belpomme, D. (2009) Basic properties and molecular mechanisms of exogenous chemical carcinogens, *Carcinogenesis* 31, 135-148.
- [35] Helleday, T., Petermann, E., Lundin, C., Hodgson, B., and Sharma, R. A. (2008) DNA repair pathways as targets for cancer therapy, *Nature Reviews Cancer* 8, 193.
- [36] Fromme, J. C., and Verdine, G. L. (2004) Base Excision Repair, In *Advances in Protein Chemistry*, pp 1-41, Academic Press.
- [37] Jeppesen, D. K., Bohr, V. A., and Stevnsner, T. (2011) DNA repair deficiency in neurodegeneration, *Progress in Neurobiology* 94, 166-200.
- [38] David, S. S., and Williams, S. D. (1998) Chemistry of Glycosylases and Endonucleases Involved in Base-Excision Repair, *Chemical Reviews* 98, 1221-1262.
- [39] Allinson, S. L., Dianova, I. I., and Dianov, G. L. (2001) DNA polymerase β is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts, *The EMBO journal* 20, 6919-6926.
- [40] Hegde, M. L., Hazra, T. K., and Mitra, S. (2008) Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells, *Cell research* 18, 27-47.
- [41] Weinfeld, M., Mani, R. S., Abdou, I., Aceytuno, R. D., and Glover, J. M. (2011) Tidying up loose ends: the role of polynucleotide kinase/phosphatase in DNA strand break repair, *Trends in biochemical sciences* 36, 262-271.
- [42] Demple, B., and Sung, J.-S. (2005) Molecular and biological roles of Ape1 protein in mammalian base excision repair, *DNA repair* 4, 1442-1449.

- [43] Dianov, G. L., Sleeth, K. M., Dianova, I. I., and Allinson, S. L. (2003) Repair of abasic sites in DNA, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 531, 157-163.
- [44] Lindahl, T., and Wood, R. D. (1999) Quality control by DNA repair, *Science* 286, 1897-1905.
- [45] Kaur, S., Ramdhan, Z. M., Guiot, M.-C., Li, L., Leduy, L., Ramotar, D., Sabri, S., Abdulkarim, B., and Nepveu, A. (2017) CUX1 stimulates APE1 enzymatic activity and increases the resistance of glioblastoma cells to the mono-alkylating agent temozolomide, *Neuro-oncology* 20, 484-493.
- [46] Kunkel, T. A. (2003) Considering the cancer consequences of altered DNA polymerase function, *Cancer cell* 3, 105-110.
- [47] Li, G.-M. (2008) Mechanisms and functions of DNA mismatch repair, *Cell research* 18, 85.
- [48] Obmolova, G., Ban, C., Hsieh, P., and Yang, W. (2000) Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA, *nature* 407, 703.
- [49] Barras, F., and Marinus, M. G. (1989) The great GATC: DNA methylation in *E. coli*, *Trends in Genetics* 5, 139-143.
- [50] Modrich, P., and Lahue, R. (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology, *Annual review of biochemistry* 65, 101-133.
- [51] Costa, R. M., Chiganças, V., da Silva Galhardo, R., Carvalho, H., and Menck, C. F. (2003) The eukaryotic nucleotide excision repair pathway, *Biochimie* 85, 1083-1099.
- [52] Cleaver, J. E., Lam, E. T., and Revet, I. (2009) Disorders of nucleotide excision repair: the genetic and molecular basis of heterogeneity, *Nature Reviews Genetics* 10, 756.
- [53] Shuck, S. C., Short, E. A., and Turchi, J. J. (2008) Eukaryotic nucleotide excision repair: from understanding mechanisms to influencing biology, *Cell research* 18, 64.
- [54] Sugawara, K. (2010) Regulation of damage recognition in mammalian global genomic nucleotide excision repair, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 685, 29-37.
- [55] Hanawalt, P. C., and Spivak, G. (2008) Transcription-coupled DNA repair: two decades of progress and surprises, *Nature reviews Molecular cell biology* 9, 958.
- [56] van Gent, D. C., Hoeijmakers, J. H., and Kanaar, R. (2001) Chromosomal stability and the DNA double-stranded break connection, *Nature Reviews Genetics* 2, 196.
- [57] Li, X., and Heyer, W.-D. (2008) Homologous recombination in DNA repair and DNA damage tolerance, *Cell research* 18, 99.
- [58] Karran, P. (2000) DNA double strand break repair in mammalian cells, *Current opinion in genetics & development* 10, 144-150.
- [59] Nimonkar, A. V., Özsoy, A. Z., Genschel, J., Modrich, P., and Kowalczykowski, S. C. (2008) Human exonuclease 1 and BLM helicase interact to resect DNA and initiate DNA repair, *Proceedings of the National Academy of Sciences* 105, 16906-16911.
- [60] Sartori, A. A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S. P. (2007) Human CtIP promotes DNA end resection, *Nature* 450, 509.
- [61] Wyman, C., Ristic, D., and Kanaar, R. (2004) Homologous recombination-mediated double-strand break repair, *DNA Repair* 3, 827-833.
- [62] McIlwraith, M. J., Vaisman, A., Liu, Y., Fanning, E., Woodgate, R., and West, S. C. (2005) Human DNA polymerase η promotes DNA synthesis from strand invasion intermediates of homologous recombination, *Molecular cell* 20, 783-792.

- [63] Ip, S. C., Rass, U., Blanco, M. G., Flynn, H. R., Skehel, J. M., and West, S. C. (2008) Identification of Holliday junction resolvases from humans and yeast, *Nature* 456, 357.
- [64] DeFazio, L. G., Stansel, R. M., Griffith, J. D., and Chu, G. (2002) Synapsis of DNA ends by DNA-dependent protein kinase, *The EMBO journal* 21, 3192-3200.
- [65] Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. (2003) Mechanism and regulation of human non-homologous DNA end-joining, *Nature Reviews Molecular Cell Biology* 4, 712-720.
- [66] Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L. A., de la Chapelle, A., Peltomäki, P., Mecklin, J. P., and Järvinen, H. J. (1999) Cancer risk in mutation carriers of DNA-mismatch-repair genes, *International journal of cancer* 81, 214-218.
- [67] Negrini, S., Gorgoulis, V. G., and Halazonetis, T. D. (2010) Genomic instability — an evolving hallmark of cancer, *Nature Reviews Molecular Cell Biology* 11, 220.
- [68] Klaunig, J. E., and Kamendulis, L. M. (2004) The role of oxidative stress in carcinogenesis, *Annu. Rev. Pharmacol. Toxicol.* 44, 239-267.
- [69] Kobayashi, A., Kang, M.-I., Okawa, H., Ohtsuji, M., Zenke, Y., Chiba, T., Igarashi, K., and Yamamoto, M. (2004) Oxidative Stress Sensor Keap1 Functions as an Adaptor for Cul3-Based E3 Ligase To Regulate Proteasomal Degradation of Nrf2, *Molecular and Cellular Biology* 24, 7130.
- [70] Hayes, J. D., and McMahon, M. (2009) NRF2 and KEAP1 mutations: permanent activation of an adaptive response in cancer, *Trends in biochemical sciences* 34, 176-188.
- [71] Kansanen, E., Kuosmanen, S. M., Leinonen, H., and Levonen, A.-L. (2013) The Keap1-Nrf2 pathway: Mechanisms of activation and dysregulation in cancer, *Redox Biology* 1, 45-49.
- [72] Lee, A. C., Fenster, B. E., Ito, H., Takeda, K., Bae, N. S., Hirai, T., Yu, Z.-X., Ferrans, V. J., Howard, B. H., and Finkel, T. (1999) Ras proteins induce senescence by altering the intracellular levels of reactive oxygen species, *Journal of Biological Chemistry* 274, 7936-7940.
- [73] Osheroff, W. P., Jung, H. K., Beard, W. A., Wilson, S. H., and Kunkel, T. A. (1999) The fidelity of DNA polymerase β during distributive and processive DNA synthesis, *Journal of Biological Chemistry* 274, 3642-3650.
- [74] Luo, J., Solimini, N. L., and Elledge, S. J. (2009) Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction, *Cell* 136, 823-837.
- [75] Hanahan, D., and Weinberg, Robert A. (2011) Hallmarks of Cancer: The Next Generation, *Cell* 144, 646-674.
- [76] Levine, A. J., Momand, J., and Finlay, C. A. (1991) The p53 tumour suppressor gene, *nature* 351, 453.
- [77] Knudson, A. G. (1971) Mutation and cancer: statistical study of retinoblastoma, *Proceedings of the National Academy of Sciences* 68, 820-823.
- [78] Druker, B. J., Tamura, S., Buchdunger, E., Ohno, S., Segal, G. M., Fanning, S., Zimmermann, J., and Lydon, N. B. (1996) Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr–Abl positive cells, *Nature medicine* 2, 561.
- [79] Hanahan, D., and Weinberg, R. A. (2000) The hallmarks of cancer, *cell* 100, 57-70.
- [80] Solimini, N. L., Luo, J., and Elledge, S. J. (2007) Non-Oncogene Addiction and the Stress Phenotype of Cancer Cells, *Cell* 130, 986-988.

- [81] Fisher, A. E., Hohegger, H., Takeda, S., and Caldecott, K. W. (2007) Poly (ADP-ribose) polymerase 1 accelerates single-strand break repair in concert with poly (ADP-ribose) glycohydrolase, *Molecular and cellular biology* 27, 5597-5605.
- [82] Welsh, P. L., and King, M.-C. (2001) BRCA1 and BRCA2 and the genetics of breast and ovarian cancer, *Human molecular genetics* 10, 705-713.
- [83] Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N. J., and Helleday, T. (2005) Specific killing of BRCA2-deficient tumours with inhibitors of poly (ADP-ribose) polymerase, *Nature* 434, 913.
- [84] Farmer, H., McCabe, N., Lord, C. J., Tutt, A. N., Johnson, D. A., Richardson, T. B., Santarosa, M., Dillon, K. J., Hickson, I., and Knights, C. (2005) Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy, *Nature* 434, 917.
- [85] Kaelin Jr, W. G. (2005) The concept of synthetic lethality in the context of anticancer therapy, *Nature reviews cancer* 5, 689.
- [86] Micchelli, C. A., Rulifson, E. J., and Blair, S. S. (1997) The function and regulation of cut expression on the wing margin of *Drosophila*: Notch, Wingless and a dominant negative role for Delta and Serrate, *Development* 124, 1485-1495.
- [87] Blochlinger, K., Bodmer, R., Jack, J., Jan, L. Y., and Jan, Y. N. (1988) Primary structure and expression of a product from cut, a locus involved in specifying sensory organ identity in *Drosophila*, *Nature* 333, 629.
- [88] Neufeld, E. J., Skalnik, D. G., Lievens, P. M., and Orkin, S. H. (1992) Human CCAAT displacement protein is homologous to the *Drosophila* homeoprotein, cut, *Nature genetics* 1, 50.
- [89] Sansregret, L., and Nepveu, A. (2008) The multiple roles of CUX1: insights from mouse models and cell-based assays, *Gene* 412, 84-94.
- [90] Hulea, L., and Nepveu, A. (2012) CUX1 transcription factors: from biochemical activities and cell-based assays to mouse models and human diseases, *Gene* 497, 18-26.
- [91] Truscott, M., Harada, R., Vadnais, C., Robert, F., and Nepveu, A. (2008) p110 CUX1 cooperates with E2F transcription factors in the transcriptional activation of cell cycle-regulated genes, *Molecular and cellular biology* 28, 3127-3138.
- [92] Vadnais, C., Awan, A. A., Harada, R., Clermont, P.-L., Leduy, L., Bérubé, G., and Nepveu, A. (2013) Long-range transcriptional regulation by the p110 CUX1 homeodomain protein on the ENCODE array, *BMC genomics* 14, 258.
- [93] Coqueret, O., Bérubé, G., and Nepveu, A. (1998) The mammalian Cut homeodomain protein functions as a cell-cycle-dependent transcriptional repressor which downmodulates p21WAF1/CIP1/SDI1 in S phase, *The EMBO Journal* 17, 4680-4694.
- [94] Harada, R., Vadnais, C., Sansregret, L., Leduy, L., Berube, G., Robert, F., and Nepveu, A. (2007) Genome-wide location analysis and expression studies reveal a role for p110 CUX1 in the activation of DNA replication genes, *Nucleic acids research* 36, 189-202.
- [95] Sansregret, L., Vadnais, C., Livingstone, J., Kwiatkowski, N., Awan, A., Cadieux, C., Leduy, L., Hallett, M. T., and Nepveu, A. (2011) Cut homeobox 1 causes chromosomal instability by promoting bipolar division after cytokinesis failure, *Proceedings of the National Academy of Sciences* 108, 1949-1954.
- [96] Vadnais, C., Davoudi, S., Afshin, M., Harada, R., Dudley, R., Clermont, P.-L., Drobetsky, E., and Nepveu, A. (2012) CUX1 transcription factor is required for optimal ATM/ATR-mediated responses to DNA damage, *Nucleic acids research* 40, 4483-4495.

- [97] Vadnais, C., Shooshtarizadeh, P., Rajadurai, C. V., Lesurf, R., Hulea, L., Davoudi, S., Cadieux, C., Hallett, M., Park, M., and Nepveu, A. (2014) Autocrine activation of the Wnt/ β -catenin pathway by CUX1 and GLIS1 in breast cancers, *Biology open* 3, 937-946.
- [98] Kedinger, V., Sansregret, L., Harada, R., Vadnais, C., Cadieux, C., Fathers, K., Park, M., and Nepveu, A. (2009) p110 CUX1 homeodomain protein stimulates cell migration and invasion in part through a regulatory cascade culminating in the repression of E-cadherin and occludin, *Journal of Biological Chemistry* 284, 27701-27711.
- [99] Moon, N. S., Bérubé, G., and Nepveu, A. (2000) CCAAT displacement activity involves Cut repeats 1 and 2, not the Cut homeodomain, *Journal of Biological Chemistry* 275, 31325-31334.
- [100] Skalnik, D., Strauss, E., and Orkin, S. (1991) CCAAT displacement protein as a repressor of the myelomonocytic-specific gp91-phox gene promoter, *Journal of Biological Chemistry* 266, 16736-16744.
- [101] Lievens, P. M., Donady, J. J., Tufarelli, C., and Neufeld, E. J. (1995) Repressor activity of CCAAT displacement protein in HL-60 myeloid leukemia cells, *Journal of Biological Chemistry* 270, 12745-12750.
- [102] Mailly, F., Berube, G., Harada, R., Mao, P., Phillips, S., and Nepveu, A. (1996) The human cut homeodomain protein can repress gene expression by two distinct mechanisms: active repression and competition for binding site occupancy, *Molecular and Cellular Biology* 16, 5346-5357.
- [103] Nepveu, A. (2001) Role of the multifunctional CDP/Cut/Cux homeodomain transcription factor in regulating differentiation, cell growth and development, *Gene* 270, 1-15.
- [104] Ozisik, Y. Y., Meloni, A. M., Surti, U., and Sandberg, A. A. (1993) Deletion 7q22 in uterine leiomyoma: a cytogenetic review, *Cancer genetics and cytogenetics* 71, 1-6.
- [105] Zeng, W. R., Watson, P., Lin, J., Jothy, S., Lidereau, R., Park, M., and Nepveu, A. (1999) Refined mapping of the region of loss of heterozygosity on the long arm of chromosome 7 in human breast cancer defines the location of a second tumor suppressor gene at 7q22 in the region of the CUTL1 gene, *Oncogene* 18, 2015.
- [106] Zeng, W. R., Scherer, S. W., Koutsilieris, M., Huizenga, J. J., Filteau, F., Tsui, L.-C., and Nepveu, A. (1997) Loss of heterozygosity and reduced expression of the CUTL1 gene in uterine leiomyomas, *Oncogene* 14, 2355.
- [107] Pedersen-Bjergaard, J., Andersen, M. T., and Andersen, M. K. (2007) Genetic pathways in the pathogenesis of therapy-related myelodysplasia and acute myeloid leukemia, *ASH Education Program Book 2007*, 392-397.
- [108] Zhang, Y., and Rowley, J. D. (2006) Chromatin structural elements and chromosomal translocations in leukemia, *DNA repair* 5, 1282-1297.
- [109] McNerney, M. E., Brown, C. D., Wang, X., Bartom, E. T., Karmakar, S., Bandlamudi, C., Yu, S., Ko, J., Sandall, B. P., and Stricker, T. (2013) CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia, *Blood* 121, 975-983.
- [110] Wong, C. C., Martincorena, I., Rust, A. G., Rashid, M., Alifrangis, C., Alexandrov, L. B., Tiffen, J. C., Kober, C., Green, A. R., and Massie, C. E. (2014) Inactivating CUX1 mutations promote tumorigenesis, *Nature genetics* 46, 33.
- [111] Ledford, A. W., Brantley, J. G., Kemeny, G., Foreman, T. L., Quaggin, S. E., Igarashi, P., Oberhaus, S. M., Rodova, M., Calvet, J. P., and Heuvel, G. B. V. (2002) Deregulated expression of the homeobox gene Cux-1 in transgenic mice results in downregulation of

- p27kip1 expression during nephrogenesis, glomerular abnormalities, and multiorgan hyperplasia, *Developmental biology* 245, 157-171.
- [112] Ramdzan, Z. M., Vadnais, C., Pal, R., Vandal, G., Cadieux, C., Leduy, L., Davoudi, S., Hulea, L., Yao, L., and Karnezis, A. N. (2014) RAS transformation requires CUX1-dependent repair of oxidative DNA damage, *PLoS biology* 12, e1001807.
- [113] Cadieux, C., Kedinger, V., Yao, L., Vadnais, C., Drossos, M., Paquet, M., and Nepveu, A. (2009) Mouse mammary tumor virus p75 and p110 CUX1 transgenic mice develop mammary tumors of various histologic types, *Cancer research* 69, 7188-7197.
- [114] Siam, R., Harada, R., Cadieux, C., Battat, R., Vadnais, C., and Nepveu, A. (2009) Transcriptional activation of the Lats1 tumor suppressor gene in tumors of CUX1 transgenic mice, *Molecular cancer* 8, 60.
- [115] Ramdzan, Z. M., Pal, R., Kaur, S., Leduy, L., Bérubé, G., Davoudi, S., Vadnais, C., and Nepveu, A. (2015) The function of CUX1 in oxidative DNA damage repair is needed to prevent premature senescence of mouse embryo fibroblasts, *Oncotarget* 6, 3613.
- [116] Kaur, S., Ramdzan, Z. M., Guiot, M. C., Li, L., Leduy, L., Ramotar, D., Sabri, S., Abdulkarim, B., and Nepveu, A. (2017) CUX1 Stimulates APE1 Enzymatic Activity and Increases the Resistance of Glioblastoma Cells to the Mono-Alkylating Agent, Temozolomide, *Neuro-oncology*.
- [117] Pal, R., Ramdzan, Z. M., Kaur, S., Duquette, P. M., Marcotte, R., Leduy, L., Davoudi, S., Lamarche-Vane, N., Iulianella, A., and Nepveu, A. (2015) CUX2 protein functions as an accessory factor in the repair of oxidative DNA damage, *Journal of Biological Chemistry* 290, 22520-22531.
- [118] Das, S., Chattopadhyay, R., Bhakat, K. K., Boldogh, I., Kohno, K., Prasad, R., Wilson, S. H., and Hazra, T. K. (2007) Stimulation of NEIL2-mediated oxidized base excision repair via YB-1 interaction during oxidative stress, *Journal of Biological Chemistry* 282, 28474-28484.
- [119] Marenstein, D. R., Ocampo, M. T., Chan, M. K., Altamirano, A., Basu, A. K., Boorstein, R. J., Cunningham, R. P., and Teebor, G. W. (2001) Stimulation of Human Endonuclease III by Y Box-binding Protein 1 (DNA-binding Protein B) INTERACTION BETWEEN A BASE EXCISION REPAIR ENZYME AND A TRANSCRIPTION FACTOR, *Journal of Biological Chemistry* 276, 21242-21249.
- [120] Prasad, R., Liu, Y., Deterding, L. J., Poltoratsky, V. P., Kedar, P. S., Horton, J. K., Kanno, S., Asagoshi, K., Hou, E. W., Khodyreva, S. N., Lavrik, O. I., Tomer, K. B., Yasui, A., and Wilson, S. H. (2007) HMGB1 is a cofactor in mammalian base excision repair, *Mol Cell* 27, 829-841.
- [121] Kaur, S., Coulombe, Y., Ramdzan, Z. M., Leduy, L., Masson, J.-Y., and Nepveu, A. (2016) Special AT-rich Sequence-binding Protein 1 (SATB1) Functions as an Accessory Factor in Base Excision Repair, *Journal of Biological Chemistry* 291, 22769-22780.
- [122] Luo, J., Emanuele, M. J., Li, D., Creighton, C. J., Schlabach, M. R., Westbrook, T. F., Wong, K.-K., and Elledge, S. J. (2009) A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene, *Cell* 137, 835-848.
- [123] Ramdzan, Z. M., Ginjala, V., Pinder, J. B., Chung, D., Donovan, C. M., Kaur, S., Leduy, L., Dellaire, G., Ganesan, S., and Nepveu, A. (2017) The DNA repair function of CUX1 contributes to radioresistance, *Oncotarget* 8, 19021.
- [124] Marenstein, D. R., Ocampo, M. T., Chan, M. K., Altamirano, A., Basu, A. K., Boorstein, R. J., Cunningham, R. P., and Teebor, G. W. (2001) Stimulation of human endonuclease

- III by Y box-binding protein 1 (DNA-binding protein B). Interaction between a base excision repair enzyme and a transcription factor, *J Biol Chem* 276, 21242-21249.
- [125] Pestryakov, P., Zharkov, D. O., Grin, I., Fomina, E. E., Kim, E. R., Hamon, L., Eliseeva, I. A., Petrusheva, I. O., Curmi, P. A., and Ovchinnikov, L. P. (2012) Effect of the multifunctional proteins RPA, YB-1, and XPC repair factor on AP site cleavage by DNA glycosylase NEIL1, *Journal of Molecular Recognition* 25, 224-233.
- [126] Krokan, H. E., and Bjørås, M. (2013) Base excision repair, *Cold Spring Harbor perspectives in biology* 5, a012583.
- [127] Coyaud, E., Mis, M., Laurent, E. M., Dunham, W. H., Couzens, A. L., Robitaille, M., Gingras, A.-C., Angers, S., and Raught, B. (2015) BioID-based identification of Skp Cullin F-box (SCF) β -TrCP1/2 E3 ligase substrates, *Molecular & Cellular Proteomics* 14, 1781-1795.
- [128] Varnaité, R., and MacNeill, S. A. (2016) Meet the neighbors: Mapping local protein interactomes by proximity-dependent labeling with BioID, *PROTEOMICS* 16, 2503-2518.
- [129] Svilar, D., Vens, C., and Sobol, R. W. (2012) Quantitative, real-time analysis of base excision repair activity in cell lysates utilizing lesion-specific molecular beacons, *Journal of visualized experiments: JoVE*.
- [130] Paz-Elizur, T., Elinger, D., Leitner-Dagan, Y., Blumenstein, S., Krupsky, M., Berrebi, A., Schechtman, E., and Livneh, Z. (2007) Development of an enzymatic DNA repair assay for molecular epidemiology studies: distribution of OGG activity in healthy individuals, *DNA repair* 6, 45-60.
- [131] Tang, B., Di Lena, P., Schaffer, L., Head, S. R., Baldi, P., and Thomas, E. A. (2011) Genome-wide identification of Bcl11b gene targets reveals role in brain-derived neurotrophic factor signaling, *PLoS One* 6, e23691.
- [132] Olive, P. L., and Banáth, J. P. (2006) The comet assay: a method to measure DNA damage in individual cells, *Nature Protocols* 1, 23-29.
- [133] Liu, P., Li, P., and Burke, S. (2010) Critical roles of Bcl11b in T-cell development and maintenance of T-cell identity, *Immunol Rev* 238, 138-149.
- [134] Satterwhite, E., Sonoki, T., Willis, T. G., Harder, L., Nowak, R., Arriola, E. L., Liu, H., Price, H. P., Gesk, S., Steinemann, D., Schlegelberger, B., Oscier, D. G., Siebert, R., Tucker, P. W., and Dyer, M. J. (2001) The BCL11 gene family: involvement of BCL11A in lymphoid malignancies, *Blood* 98, 3413-3420.
- [135] Kominami, R. (2012) Role of the transcription factor Bcl11b in development and lymphomagenesis, *Proc Jpn Acad Ser B Phys Biol Sci* 88, 72-87.
- [136] Li, P., Burke, S., Wang, J., Chen, X., Ortiz, M., Lee, S. C., Lu, D., Campos, L., Goulding, D., Ng, B. L., Dougan, G., Huntly, B., Gottgens, B., Jenkins, N. A., Copeland, N. G., Colucci, F., and Liu, P. (2010) Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion, *Science* 329, 85-89.
- [137] Zhang, L. J., Vogel, W. K., Liu, X., Topark-Ngarm, A., Arbogast, B. L., Maier, C. S., Filtz, T. M., and Leid, M. (2012) Coordinated regulation of transcription factor Bcl11b activity in thymocytes by the mitogen-activated protein kinase (MAPK) pathways and protein sumoylation, *J Biol Chem* 287, 26971-26988.
- [138] Wiles, E. T., Lui-Sargent, B., Bell, R., and Lessnick, S. L. (2013) BCL11B is up-regulated by EWS/FLI and contributes to the transformed phenotype in Ewing sarcoma, *PLoS One* 8, e59369.

- [139] Topark-Ngarm, A., Golonzhka, O., Peterson, V. J., Barrett, B., Jr., Martinez, B., Crofoot, K., Filtz, T. M., and Leid, M. (2006) CTIP2 associates with the NuRD complex on the promoter of p57KIP2, a newly identified CTIP2 target gene, *J Biol Chem* 281, 32272-32283.
- [140] Cismasiu, V. B., Paskaleva, E., Suman Daya, S., Canki, M., Duus, K., and Avram, D. (2008) BCL11B is a general transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through recruitment of the NuRD complex, *Virology* 380, 173-181.
- [141] Cismasiu, V. B., Ghanta, S., Duque, J., Albu, D. I., Chen, H. M., Kasturi, R., and Avram, D. (2006) BCL11B participates in the activation of IL2 gene expression in CD4+ T lymphocytes, *Blood* 108, 2695-2702.
- [142] Cismasiu, V. B., Adamo, K., Gecewicz, J., Duque, J., Lin, Q., and Avram, D. (2005) BCL11B functionally associates with the NuRD complex in T lymphocytes to repress targeted promoter, *Oncogene* 24, 6753-6764.
- [143] Wakabayashi, Y., Watanabe, H., Inoue, J., Takeda, N., Sakata, J., Mishima, Y., Hitomi, J., Yamamoto, T., Utsuyama, M., Niwa, O., Aizawa, S., and Kominami, R. (2003) Bcl11b is required for differentiation and survival of alphabeta T lymphocytes, *Nat Immunol* 4, 533-539.
- [144] Gingras, H., Cases, O., Krasilnikova, M., Bérubé, G., and Nepveu, A. (2005) Biochemical characterization of the mammalian Cux2 protein, *Gene* 344, 273-285.
- [145] Wakabayashi, Y., Inoue, J., Takahashi, Y., Matsuki, A., Kosugi-Okano, H., Shinbo, T., Mishima, Y., Niwa, O., and Kominami, R. (2003) Homozygous deletions and point mutations of the Rit1/Bcl11b gene in gamma-ray induced mouse thymic lymphomas, *Biochem Biophys Res Commun* 301, 598-603.
- [146] Sakata, J., Inoue, J., Ohi, H., Kosugi-Okano, H., Mishima, Y., Hatakeyama, K., Niwa, O., and Kominami, R. (2004) Involvement of V(D)J recombinase in the generation of intragenic deletions in the Rit1/Bcl11b tumor suppressor gene in gamma-ray-induced thymic lymphomas and in normal thymus of the mouse, *Carcinogenesis* 25, 1069-1075.
- [147] Matsumoto, Y., Kosugi, S., Shinbo, T., Chou, D., Ohashi, M., Wakabayashi, Y., Sakai, K., Okumoto, M., Mori, N., Aizawa, S., Niwa, O., and Kominami, R. (1998) Allelic loss analysis of gamma-ray-induced mouse thymic lymphomas: two candidate tumor suppressor gene loci on chromosomes 12 and 16, *Oncogene* 16, 2747-2754.
- [148] Kamimura, K., Mishima, Y., Obata, M., Endo, T., Aoyagi, Y., and Kominami, R. (2007) Lack of Bcl11b tumor suppressor results in vulnerability to DNA replication stress and damages, *Oncogene* 26, 5840-5850.
- [149] Ohi, H., Mishima, Y., Kamimura, K., Maruyama, M., Sasai, K., and Kominami, R. (2007) Multi-step lymphomagenesis deduced from DNA changes in thymic lymphomas and atrophic thymuses at various times after gamma-irradiation, *Oncogene* 26, 5280-5289.
- [150] Gutierrez, A., and Kentsis, A. (2018) Acute myeloid/T-lymphoblastic leukaemia (AMTL): a distinct category of acute leukaemias with common pathogenesis in need of improved therapy, *Br J Haematol* 180, 919-924.
- [151] Gutierrez, A., Kentsis, A., Sanda, T., Holmfeldt, L., Chen, S. C., Zhang, J., Protopopov, A., Chin, L., Dahlberg, S. E., Neuberg, D. S., Silverman, L. B., Winter, S. S., Hunger, S. P., Sallan, S. E., Zha, S., Alt, F. W., Downing, J. R., Mullighan, C. G., and Look, A. T. (2011) The BCL11B tumor suppressor is mutated across the major molecular subtypes of T-cell acute lymphoblastic leukemia, *Blood* 118, 4169-4173.

- [152] De Keersmaecker, K., Real, P. J., Gatta, G. D., Palomero, T., Sulis, M. L., Tosello, V., Van Vlierberghe, P., Barnes, K., Castillo, M., Sole, X., Hadler, M., Lenz, J., Aplan, P. D., Kelliher, M., Kee, B. L., Pandolfi, P. P., Kappes, D., Gounari, F., Petrie, H., Van der Meulen, J., Speleman, F., Paietta, E., Racevskis, J., Wiernik, P. H., Rowe, J. M., Soulier, J., Avran, D., Cave, H., Dastugue, N., Raimondi, S., Meijerink, J. P., Cordon-Cardo, C., Califano, A., and Ferrando, A. A. (2010) The TLX1 oncogene drives aneuploidy in T cell transformation, *Nat Med* 16, 1321-1327.
- [153] Przybylski, G. K., Dik, W. A., Wanzeck, J., Grabarczyk, P., Majunke, S., Martin-Subero, J. I., Siebert, R., Dolken, G., Ludwig, W. D., Verhaaf, B., van Dongen, J. J., Schmidt, C. A., and Langerak, A. W. (2005) Disruption of the BCL11B gene through inv(14)(q11.2q32.31) results in the expression of BCL11B-TRDC fusion transcripts and is associated with the absence of wild-type BCL11B transcripts in T-ALL, *Leukemia* 19, 201-208.
- [154] Gu, X., Wang, Y., Zhang, G., Li, W., and Tu, P. (2013) Aberrant expression of BCL11B in mycosis fungoides and its potential role in interferon-induced apoptosis, *J Dermatol* 40, 596-605.
- [155] Ganguli-Indra, G., Wasylyk, C., Liang, X., Millon, R., Leid, M., Wasylyk, B., Abecassis, J., and Indra, A. K. (2009) CTIP2 expression in human head and neck squamous cell carcinoma is linked to poorly differentiated tumor status, *PLoS One* 4, e5367.
- [156] Oshiro, A., Tagawa, H., Ohshima, K., Karube, K., Uike, N., Tashiro, Y., Utsunomiya, A., Masuda, M., Takasu, N., Nakamura, S., Morishima, Y., and Seto, M. (2006) Identification of subtype-specific genomic alterations in aggressive adult T-cell leukemia/lymphoma, *Blood* 107, 4500-4507.
- [157] Liao, C. K., Fang, K. M., Chai, K., Wu, C. H., Ho, C. H., Yang, C. S., and Tzeng, S. F. (2016) Depletion of B cell CLL/Lymphoma 11B Gene Expression Represses Glioma Cell Growth, *Mol Neurobiol* 53, 3528-3539.
- [158] Grabarczyk, P., Przybylski, G., Depke, M., Völker, U., Bahr, J., Assmus, K., Bröker, B., Walther, R., and Schmidt, C. (2007) Inhibition of BCL11B expression leads to apoptosis of malignant but not normal mature T cells, *Oncogene* 26, 3797-3810.
- [159] Huang, X., Chen, S., Shen, Q., Chen, S., Yang, L., Grabarczyk, P., Przybylski, G. K., Schmidt, C. A., and Li, Y. (2011) Down regulation of BCL11B expression inhibits proliferation and induces apoptosis in malignant T cells by BCL11B-935-siRNA, *Hematology* 16, 236-242.
- [160] Shen, Q., Huang, X., Chen, S., Yang, L., Chen, S., Li, B., Wu, X., Grabarczyk, P., Przybylski, G. K., Schmidt, C. A., and Li, Y. (2012) BCL11B suppression does not influence CD34(+) cell differentiation and proliferation, *Hematology* 17, 329-333.
- [161] Zweier-Renn, L. A., Riz, I., Hawley, T. S., and Hawley, R. G. (2013) The DN2 Myeloid-T (DN2mt) Progenitor is a Target Cell for Leukemic Transformation by the TLX1 Oncogene, *J Bone Marrow Res* 1.
- [162] Grabarczyk, P., Nahse, V., Delin, M., Przybylski, G., Depke, M., Hildebrandt, P., Volker, U., and Schmidt, C. A. (2010) Increased expression of bcl11b leads to chemoresistance accompanied by G1 accumulation, *PLoS One* 5.
- [163] Heffler, M. A., Walters, R. D., and Kugel, J. F. (2012) Using electrophoretic mobility shift assays to measure equilibrium dissociation constants: GAL4-p53 binding DNA as a model system, *Biochemistry and Molecular Biology Education* 40, 383-387.

- [164] Williamson, M. P. (1994) The structure and function of proline-rich regions in proteins, *The Biochemical journal* 297 (Pt 2), 249-260.
- [165] Fromme, J. C., and Verdine, G. L. (2003) Structure of a trapped endonuclease III–DNA covalent intermediate, *The EMBO journal* 22, 3461-3471.
- [166] Thayer, M., Ahern, H., Xing, D., Cunningham, R., and Tainer, J. (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure, *The EMBO journal* 14, 4108-4120.
- [167] Robey-Bond, S. M., Benson, M. A., Barrantes-Reynolds, R., Bond, J. P., and Wallace, S. S. (2017) Probing the activity of NTHL1 orthologs by targeting conserved amino acid residues, *DNA Repair (Amst)* 53, 43-51.
- [168] Banerjee, A., Santos, W. L., and Verdine, G. L. (2006) Structure of a DNA glycosylase searching for lesions, *Science* 311, 1153-1157.
- [169] Qi, Y., Nam, K., Spong, M. C., Banerjee, A., Sung, R.-J., Zhang, M., Karplus, M., and Verdine, G. L. (2012) Strandwise translocation of a DNA glycosylase on undamaged DNA, *Proceedings of the National Academy of Sciences* 109, 1086-1091.
- [170] Nelson, S. R., Dunn, A. R., Kathe, S. D., Warshaw, D. M., and Wallace, S. S. (2014) Two glycosylase families diffusively scan DNA using a wedge residue to probe for and identify oxidatively damaged bases, *Proceedings of the National Academy of Sciences* 111, E2091-E2099.
- [171] Lan, L., Nakajima, S., Oohata, Y., Takao, M., Okano, S., Masutani, M., Wilson, S. H., and Yasui, A. (2004) In situ analysis of repair processes for oxidative DNA damage in mammalian cells, *Proceedings of the National Academy of Sciences* 101, 13738-13743.
- [172] Grabarczyk, P., Przybylski, G. K., Depke, M., Volker, U., Bahr, J., Assmus, K., Broker, B. M., Walther, R., and Schmidt, C. A. (2007) Inhibition of BCL11B expression leads to apoptosis of malignant but not normal mature T cells, *Oncogene* 26, 3797-3810.