OPTIMIZING NECEEM-BASED APTAMER SELECTION USING EMULSION PCR TO OBTAIN POTENTIAL DNA INHIBITORS FOR THE HUMAN DEALKYLATING ENZYME ABH2

ROMAN YUFA

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE

STUDIES IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

GRADUATE PROGRAM IN BIOLOGY

YORK UNIVERSITY

TORONTO, ONTARIO

MAY 2015

Abstract

The oldest chemotherapeutic drugs still in use today are the alkylating agents. In many cases, their effectiveness is hindered by the ability of cancerous cells to develop resistance using the cell's innate DNA repair enzymes, such as hABH2. Targeted inhibition of such enzymes can potentially greatly improve chemotherapy. Aptamers are more efficient, versatile and theoretically easier to discover than conventional small molecule inhibitors. However, Aptamer selection is regularly unsuccessful and screening inhibitors is a lengthy process. Here we optimize the aptamer selection process using Emulsion PCR to improve the amplification efficiency post NECEEM separation and rapidly select high affinity aptamers to untagged hABH2 in 4 rounds. We then show the efficiency and robustness of the recently introduced direct CE-based approach to rapidly measure the demethylation activity of hABH2 and hABH3 towards ssDNA 3-meC substrates and discuss its potential as method to identify aptamers able to selectively inhibit hABH2.

Acknowledgments

First and foremost, I would to thank my supervisor, Professor Sergey Krylov, for his support, guidance and inspiration throughout the duration of my research. He has taught me how to think as a scientist and provided a model to strive towards. I would also like to extend my gratitude to Professor Svetlana Krylova for guiding my researching, for her constructive criticism and helping me to strive to become the best researcher I can be. Their combined support is truly invaluable and I will carry their lessons and example throughout the rest of my career.

I would also like to thank my committee member, Dr. Mark Bayfield, who has always been available for helping me with my research and has provided invaluable ideas during my research evaluations.

I would like to thank my lab members: Ruchi Liyanage, Fletcher Agostino, Dave Wegman, Jiayin Bao, Natasha Obrecht, Victor Galievsky, Stephanie De Jong and Vasili Koshkin for never hesitating to help me with any issues I encountered or discussing any idea I may have had. I would like to especially thank Mirzo Kanoatov for his patience over the years, for taking the time to answer my questions, regardless of how simple or foolish they might be and for never hesitating to lend me his expertise. I have learned invaluable skills and made lifelong friends in the lab for which I am forever grateful.

Finally, I would like to thank my family for their lifelong support, their unconditional love, and for making me who I am. This work would have not been possible without them.

Table of contents

ABSTRACT ..... **II**

ACKNOWLEDGMENTS ..... **III**

TABLE OF CONTENTS ..... **IV**

LIST OF TABLES ..... **X**

LIST OF FIGURES ..... **XI**

LIST OF ABBREVIATIONS ..... **XVII**

CHAPTER 1: INTRODUCTION **1**

1.1 DNA Alkylation **1**

1.2 Alkylating Antineoplastic Agents **1**

1.2.1 Nitrogen Mustards 2

1.2.2 Aziridines 4

1.2.3 Alkyl Sulfonates 6

1.2.4 Nitrosoureas 6

1.2.5 Platinum-Based Drugs 7

1.3 Alkylation DNA Repair **9**

1.4 Oxoglutarate and Fe^{2+} Dependent Dioxygenases **10**

1.5 ABH2 ..... **12**

1.5.1 Discovery 12

1.5.2 Substrate Specificity 14

1.5.3 Kinetic Parameters 20

1.5.4 Structure 21

1.5.5 Expression and Localization 26

1.5.6 Pathological Role 27

1.5.7 Inhibition 29

1.6 ABH3 30

1.6.1 Discovery 30

1.6.2 Substrate Specificity 31

1.6.3 Kinetic Constants 33

1.6.4 Structure 33

1.6.5 Expression and Localization 36

1.6.6 Physiological Role 36

1.6.7 Inhibition 37

1.7 AlkB and the other Human Homologues 38

1.7.1 AlkB 38

1.7.2 ABH1 41

1.7.3 ABH4-7 42

1.7.4 ABH8 42

1.7.5 FTO 43

1.7.6 ABH Localization 43

1.8 AlkB and ABH Assays 44

1.8.1 Radioactivity 46

1.8.2 HPLC 48

1.8.3 Mass Spectrometry 50

1.8.4 Continuous Coupled Assay 51

1.8.5 Restriction Endonuclease 52

1.8.6 Direct Approach using CE 53

1.9 Aptamers 54

1.9.1 AlkB Aptamer Inhibitors 55

1.10 Research Objective 56

CHAPTER 2: APTAMER SELECTION OPTIMIZATION 57

2.1 Introduction 57

2.1.1 SELEX 57

2.1.2 Homogeneous Partitioning 58

2.1.3 Kinetic Capillary Electrophoresis 59

2.1.4 NECEEM 59

2.1.4.1 Issues 60

2.1.4.2 Solutions 61

2.1.5 Emulsion PCR 62

2.1.6 ABH2 as a Target 63

2.2 Results 64

2.2.1 Unsuccessful aptamers selection for hABH2 by NECEEM with conventional PCR 64

2.2.2 Options for improving efficiency of PCR 65

2.2.3 Emulsion versus Conventional PCR 66

2.2.4 Aptamer selection for hABH2 based on NECEEM-ePCR 68

2.2.5 Aptamer Synthesis and Binding 71

2.2.6 Cross Reactivity 74

2.3 Conclusion 74

2.4 Materials and Methods 76

2.4.1 Materials 76

2.4.2 Instrumentation 76

2.4.3 Determination of the aptamer-collection window 77

2.4.4 Aptamer selection 77

2.4.5 Propagation past uncooled region 78

2.4.6 Emulsion PCR 78

2.4.7 Regeneration of enriched pool 78

2.4.8 Affinity analysis 79

2.4.9 Cloning 79

CHAPTER 3: DIRECT MEASUREMENT OF THE DEMETHYLATION ACTIVITY OF HABH2 AND HABH3 81

3.1 Introduction 81

3.1.1 DNA Alkylation and Repair 81

3.1.2 Chemotherapeutics 81

3.1.3 Alkylation Activity Studies 82

3.1.4 Objective 83

3.2 Results and Discussion 83

3.3 Conclusion 86

3.4 Materials and Methods 86

3.4.1 Materials 86

3.4.2 Instrumentation 86

3.4.3 Measuring Enzyme Kinetics 87

LIMITATIONS 88

CONCLUDING REMARKS 89

FUTURE DIRECTION 90

REFERENCES.....92

List of Tables

Table 1: Known substrate specificity for ABH2 and ABH3 14

Table 2: Steady-state kinetic constants for removal of 1-meA and 3-meC by hABH2 and hABH3. Adapted from [14]. 21

Table 3: Binding constants and sequences of synthesized hABH2 aptamers. 73

Table 4: Kinetic constants for AlkB and AlkB mammalian homologues for demethylating 3-meC on ssDNA substrates. 85

List of Figures

Figure 1: Alkylation mechanism of nitrogen mustards. **A)** Metabolism of cyclophosphamide. **B)** Nitrogen mustard induced DNA alkylation and crosslinking between adenine moieties. Generated by ChemSketch. Adapted from [5]. 4

Figure 2: Busulphan mediated interstrand crosslinking between guanine moieties. Generated by ChemSketch. Adapted from [5]. 5

Figure 3: Metabolism of Carmustine which causes DNA crosslinking. Generated by ChemSketch. Adapted from [5]. 7

Figure 4: A) Conversion of cisplatin to positively charged reactive species via reversible aquation reactions. B) Cisplatin induced DNA adducts. Generated by ChemSketch. Adapted from [5]. 9

Figure 5: Proposed ABH2 mechanism. Generated using ChemSketch. Adapted from [1-3]. 11

Figure 6: Repair of 1-meA and 3-meC lesions by ABH2 and hABH3. 1 nM of protein was incubated with [^{14}C]MeI-treated poly(dA) or poly(dC) for 30 min under standard assay conditions (section 1.8). The radioactively labeled methylated bases remaining in the substrates were analyzed by HPLC and scintillation counting. Triangle – no enzyme; Square – with enzyme. Adapted from [13]. 15

Figure 7: **A**) Comparison of the activity of AlkB, hABH2 and hABH2 towards ssDNA (filled circles) and dsDNA (open circles) substrates. **B**) Comparison of the activity of activity of AlkB, hABH2 and hABH3 towards DNA and RNA substrates as well as M13 ssDNA (filled squares). Enzyme was incubated with radioactive [H^3]methylated substrate. The DNA was precipitated by ethanol and the released radioactivity in the supernatant was measured by scintillation counting. Experiments were carried out at least three times with similar results. Adapted from [4]. 16

Figure 8: The effect of Magnesium on the preference of hABH2 for dsDNA. [^{3}H]methylated ACGT-oligo (single or double stranded) was incubated with varying concentrations of hABH2 in the absence (circles) or presence (triangles) of $MgCl_2$. The ethanol soluble radioactivity was measured by scintillation counting. Adapted from [11]. 17

Figure 9: Accumulation of 1-meA in aging repair-deficient mice. **A.** Accumulation of 1-meA in genomic DNA from liver of 1-12 month-old wild-type, mABH2-and mABH3-targeted mice. The steady-state level of 1-meA was determined by the HPLC-MS/MS. **B**. Numerical values of 1-meA in genomic DNA from liver of 8- and 12–month-old mABH2-null mice. **C**. HPLC-MS/MS chromatogram of 1-meA in genomic DNA from liver of 8-month-old wild-type (upper panel) and mABH2-targeted mice (middle panel), compared with 1me(dA) standard (lower panel). Adapted from [18]. 18

Figure 10: Fe-2OG-dependent dioxygenase sequence alignment. **A**) Similarity of the C-terminus core region of AlkB, hABH1, hABH2 and hABH3 and other known Fe–2OG-dependent dioxygenase: isopenicillin N synthase (IPNS), deacetoxycephalosporin C synthase (DAOCS), and anthocyanidin synthase (ANS). **B**) The predicted secondary structures of AlkB, hABH2 and hABH3 core region aligned with that of known Fe–2OG-dependent dioxygenase. Green - known secondary structures. Red - predicted secondary structures. Background colors: green-conserved hydrophobics; red-conserved Fe^{2+} coordinate; blue-conserved substrate binding pocket; violet- conserved substrate binding pocket of known structures only; magenta - conserved binding pocket for predicted structures; yellow - substrate binding pocket not conserved. Adapted from [13]. 22

Figure 11: **A)** AlkB–DNA complex. Green - protein (the cross-linked Cys 135 is labeled); Orange – Mg^{2+}; blue - 2OG; yellow- DNA; magenta - flipped base C8* and the two bases flanking C8*; red - disulphide bond. **B)** ABH2–DNA complex. Same colour coding as in b. DNA-binding loop containing the RKK sequence is labeled. Adapted from [6]. 23

Figure 12:hABH2 base flipping and activity site structure. **a**) hABH2–DNA complex with 1-meA in the active site. Green - protein (the cross-linked Cys 135 is labeled); Orange –

Mg^{2+}; blue - 2OG; yellow- DNA; magenta - flipped base C8* and the two bases flanking C8*; red - disulphide bond. **b**) Active site of hABH2 with Mg^{2+} (orange), 2OG (blue) and 1-meA (light magenta). Adapted from [6]. 24

Figure 13: structures of AlkB (magenta), ABH2 (protein in green and DNA backbone in orange) and hABH3 (blue). The crystal structure of hABH3 was obtained without a substrate. Adapted from [6]. 25

Figure 14: hABH3 structure and active site. The β-strand jelly roll coloured orange and blue. The β4–β5 hairpin is coloured green. α-Helices are coloured yellow, and key residues as well as the co-substrate 2OG are shown as balls and sticks. The iron (violet) and the iron bound water (red) are shown as spheres. Adapted from [17]. 34

Figure 15: Crystal and domain structures of the Fe-2OG-dependent oxygenase AlkB and its mammalian homologues. **(a)** Crystal structures of AlkB and its mammalian homologues. The conserved double-stranded ß-helix (DSBH, purple) is seen in all of the dioxygenases. Enzyme specific domains (brown or green) help with substrate recognition (FTO) or additional catalysis (ABH8). 2OG (yellow) is seen in the active site with the metal ion (green). **(b)** Domain structures of the Fe-2OG dependent oxygenases seein in (a). The conserved DSBH (purple) and additional domains (brown or green) are indicated. The HxD..H motif is a part of the DSBH and coordinates iron binding. **(c)** AlkB inhibitors (i) structures of 2OG, its analogue NOG (*N*-oxalyl glycine, and MD316. (ii) The AlkB active site co-crystallized with the MD316 inhibitor. Adapted from [19]..... 40

Figure 16: Simulated product formation, [P], of an enzyme over time. Curves for different substrate concentrations ([S]) are shown. **B)** Michaelis-Menten curve of initial rate, $v_{0,}$ of the enzyme in (A). The initial rate is plotted over substrate concentration, [S], and the solid line represents the nonlinear regression fit using the Michaelis-Menten formula. K_m can be obtained from this fit (shown) as well as v_{max} and k_{cat} (not shown). Adapted from [8]. 45

Figure 17: Radioactivity released from [^{3}H]methyl DNA by AlkB. A) AlkB activity on different substrates: poly(dA) (filled circles), single-stranded, A-rich oligonucleotide AAAGCAAGAAACGAAAAAGCGAAA) (open circles), double-stranded oligonucleotide (same sequence as the previous substrate) (filled triangles), M13 ssDNA (open triangles) and poly(dA)•poly(dT) (filled squares). Adapted from [9] 47

Figure 18: HPLC (Reverse phase) elution profiles of the radioactivity released from methylated ssDNA labelled with [^{3}H] by AlkB with and without acid treatment. **A)** HPLC elution profiles of the intact ethanol soluble supernatant incubations with (open circles) or without (filled circles) AlkB. **B)** HPLC elution profiles of methylated purines post the acid hydrolysis treatment of the ethanol precipitated DNA incubations with (open circles) or without (filled circles) AlkB. The elution time of 1-meA, 3-meA, 7-meA and 7-meG is indicated by the arrows. Adapted from [3,9]. 48

Figure 19: HPLC detection of AlkB induced demethylation. **a, b)** HPLC separation followed by scintillation count of [^{14}C]Methyl iodide-treated **a)** poly(dA) and **b)** poly(dC). HPLC analysis was performed on the remaining unmethylated substrate after incubation of 2.5 pM AlkB at 37°C for 30 minutes. **c,d**) Analysis of non-radioactive methylated adenine with (d) and without (c) AlkB. 920 pmol AlkB was incubated with the methylated substrate at 37 °C for 30 min. Adapted from [16]. 49

Figure 20: MS analysis of 1,N^{6}-ethenoadenine (εA) lesion repair by AlkB *in vitro*. **A)** MALDI-TOF MS of εA reacting with AlkB for 90 min. The -1 charged species are shown. **B)** ESI-TOF MS **C)** GC-MS. Adapted from [12]. 51

Figure 21: Oxidative demethylation of 1-meA or 3-meC by AlkB coupled with FDH reduction reaction. The AlkB demethylation reaction produces a single molecule of formaldehyde which can be used as a substrate for FDH. FDH oxidizes formaldehyde to formate, converting coenzyme NAD+ to a UV detectable NADH. Adapted from [7]. 52

Figure 22: Capillary Electorphoresis with LIF separation of a fluorescently labeled ssDNA substrate (S) from synthetic demethylated DNA product (P) differing by a single methyl. Product is formed from the methylated substrate after 0, 30, and 120 s of the AlkB-catalyzed reaction. Adapted from [15]..... 54

Figure 23: The concept of aptamer selection: major steps (A) and (B) major approaches for partitioning and PCR amplification..... 58

Figure 24: A NECEEM electropherogram, where DNA is fluorescently tagged, shows three regions corresponding to the DNA peak (red), complex peak (blue) and DNA dissociation region (green). Kinetic constants can be derived by integrating these peaks. Adapted from [10]. .. 60

Figure 25: Emulsion PCR involves vigorously mixing a conventional PCR mix (with the addition of BSA) with emulsion oil. The resulting emulsion is composed of micelles which contain (ideally) a single template and all necessary PCR components, thus creating a homogeneous-like amplification conditions. 62

Figure 26: Comparison of product loss due to byproduct formation in conventional PCR and emulsion PCR..... 67

Figure 27: Determination of aptamer-collection window in NECEEM for ABB2 as a target. In NECEEM, an equilibrium mixture containing the protein and DNA aptamer library is injected (t_1) into the capillary and an electric field is applied to separate the complex from the protein and ligand (t_2). The electropherogram contains peaks of the complex and ligand due to the fluorescent tag on the DNA. A window is chosen to capture the intact complex and a part of the dissociation region. 69

Figure 28: Progress of library enrichment with number of SELEX rounds. NECEEM-based aptamer selection combined with ePCR was used to select aptamers for ABH2 from an unbiased ssDNA library. 71

Figure 29: ABH2 Aptamer binding improves with the addition of OG. 100nM of synthetic aptamer was incubated with 400nM ABH2. Low binding was observed when the selecting conditions were replicated. Addition of 2OG restored the binding. 72

Figure 30: Aptamer binding studied by NECEEM. Two hundred nM aptamer was mixed with 200 nM of ABH2 in the presence of 160 μM 2OG. 73

Figure 31: Cross reactivity of the C2 and C10 aptamers with related and unrelated proteins. The synthetic ABH2 C2 and C10 aptamers do not significantly bind to any of the proteins tested. Dissociation is observed with MutS and to a lesser extent with the rest of the proteins. 100nM aptamer was incubated with 400nM protein and incubated for 15 minutes prior to NECEEM analysis. 74

Figure 32: NECEEM-based Aptamer Selection with Emulsion PCR. In this version of the protocol Emulsion PCR is used to amplify an aptamer pool selected using NECEEM. Once amplified, single stranded DNA is generated using asymmetric PCR and purified using streptavidin magnetic beads. The purified pool is then incubated with the protein again and the process is repeated. 75

Figure 33: Michaelis Menten plots for demethylation of the TTCm substrate catalyzed by 50 nM hABH2 **(left)** and 50 nM hABH3 **(right)**. Initial reactions rates were determined though CE separation of the demethylated product from the substrate. Curve fitting was performed with GraphPad Prism 5 software. Kinetic constants (K_m and V_{max}) were determined for the best fitting. 84

List of Abbreviations

1-etA: 1-ethyladenine

1-meA: 1-methyladenine

1-meG: 1-methylguanine

2OG: 2-Oxoglutarate

3-etC: 3-ethylcytidine

3-meA: 3-methyladenine

3-meC: 3-methylcytosine

3-meU: 3-methyluracil

3-meT: 3-methylthymine

6-meA: 6-methyladenine

7-meG: 7-methylguanine

AA: Ascorbic Acid

ABH: AlkB homologue

Ada response: Adaptive response

AMD: Age-related Macular Degeneration

ANPG: Alkyl-N-adenine-DNA glycosylases

AP: Apurinic/apyrimidinic

ASCC: Activating Signal Cointegrator Complex

BSA: Bovine Serum Albumin

CE: Capillary Electrophoresis

CENU: 2-chloroethylnitrosoureas

Cisplatin: Cis-diamminedichloroplatinum(II)

DNA: Deoxyribonucleic acid

dsDNA: Double-Stranded DNA

ssDNA: Single-Stranded DNA

DSBH: Double-Stranded β-helix

DTT: Dithiothreitol

DMOG: Dimethyl Oxalyglycine

EC_{50}: Half Maximal Effective Concentration

E. coli: *Escherichia coli*

EDTA: Ethylenediaminetetraacetic acid

EOF: Electroosmotic Flow

εA: $1,N^6$-Ethenoadenine

εC: $3,N^4$-ethenocytosine

EMT: Epithelial to Mesenchymal Transition

ePCR: Emulsion PCR

ESI-TOF MS: Electrospray Time of Flight MS

EST: Expressed Sequence Tags

F488: Fluorescein (Absorbs at 488nm)

FDH: Formaldehyde Dehydrogenase

FTO: Fat Mass and Obesity-associated protein

GC-MS: Gas Chromatography MS

GBM: Glioblastoma Multiforme

HIF-1α: Hypoxia Inducible Factor 1- α

HPLC: High-Performance Liquid Chromatography

KCE: Kinetic Capillary Electrophoresis

K_d: Dissociation constant

k_{off}: Dissociation rate constant

k_{on}: Association rate constant

K_m: Michaelis Constant

K_i: Inhibition Constant

K_{cat}: Catalytic Constant

LC-MS: Liquid chromatography–mass spectrometry

LIF: Laser-Induced fluorescence

MALDI-TOF MS: Matrix Assisted Laser Desorption Ionisation with Time of Flight MS

MMS: Methyl Methanesulfonate

MS: Mass Spectrometry

NECEEM: Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures

NOG: N-oxalyglycine

NSCLC: Non-small cell lung cancers

Nt: Nucleotide

P4H: Prolyl 4-hydroxylase

PBT: Pediatric Brain Tumors

PCA: Prostate Cancer Antigen

PCNA: Proliferating-Cell Nuclear Antigen

PCR: Polymerase Chain Reaction

aPCR: Asymmetric PCR

ePCR: Emulsion PCR

sPCR: Symmetric PCR

PDT: Photodynamic therapy

Ph-PDT photofrin-mediated PDT

qRT-PCR: Quantitative Real Time Polymerase Chain Reaction

RNA: Ribonucleic acid

RNAi: RNA interference

ROS: Reactive Oxygen Species

SELEX: Systematic Evolution of Ligands by Exponential Enrichment

Thio-tepa: Triethylenethiophosphoramide

UV: ultraviolet

V_{max}: Maximum Velocity

Chapter 1: Introduction

1.1 DNA Alkylation

DNA is naturally susceptible to cytotoxic, mutagenic and clastogenic modifications. These include oxidation, deamination, and alkylation of the nucleobases. In humans, DNA damage may occur at a daily rate of 20,000 bases per cell.[23] If left unrepaired, this damage could cause widespread mutation and apoptosis.[24] The relative contribution of each aforementioned DNA damage mechanisms is not well established. However, the ubiquitous expression and conservation of several types of alkylation repair enzymes indicates that alkylation damage is a significant one.[25-27]

DNA alkylation can be classified into two broad categories: epigenetic and mutagenic. Epigenetic methylation, such as 5-methylcytosine and 6-methyladenine, is a key mechanism of gene regulation which is tightly controlled by methyltransferases. Such modifications can alter chromatin structure and control the access of transcription factors to gene promoters.[28] In contrast, mutagenic alkylation, such as 1-methyladenine (1-meA) and 3-methylcytosine (3-meC), is an uncontrolled process occurring due to the nucleophilic nature of oxygen and nitrogen atoms in nucleobases. Such lesions were found to block cell replication and cause mutation and apoptosis when left unrepaired.[29,30] For example, 3-meA has been found to cause sister chromatid exchange, DNA breaks, S-phase arrest, accumulation of p53 and apoptosis.[31] While unwelcome in somatic cells, these cytotoxic effects make alkylation an attractive mechanism for chemotherapeutics.[32] Alkylating drugs aim to disrupt DNA replication in cancer cells, leading to cell cycle arrest and cell death.[33]

1.2 Alkylating Antineoplastic Agents

The aforementioned toxic nature of alkylation is utilized by a large class of chemotherapeutic drugs termed the alkylating antineoplastic agents.[34] These drugs are the oldest cancer chemotherapeutics, dating back to the realization that mustard gas used during the First World War can be used to treat a variety of cancers.[35,36] There are currently five major structural classes of alkylating agents: nitrogen mustards,[37] aziridines,[38] alkyl sulfonates,[39] nitrosoureas,[40] and platinum-containing drugs.[41] These drugs are both monofunctional and bifunctional agents,

forming a single or double alkyl lesion, respectively. Both monofunctional and bifunctional lesions can block DNA replication and cause apoptosis. In general, bifunctional lesions are considered more cytotoxic than monofunctional ones.[42] Monofunctional agents can alkylate base-pairing nitrogen or oxygen atoms in DNA. This prevents the alkylated strands from hydrogen bonding with their compliments or may block DNA polymerase from replicating the DNA.[30,43,44] Bifunctional agents are able to form interstrand or intrastand crosslinks. Interstrand crosslinks join two complementary strands, preventing them from uncoiling and replicating, whereas intrastrand crosslinks join two adjacent bases, blocking DNA polymerase from properly replicating the DNA.[45,46]

Although alkylating agents can be divided into five distinct structural classes, their mechanisms of actions can be similar. It is important to understand the structural implications as well as the mechanisms of action of each class as they help determine the scope of clinical application. Detailed below are the specific mechanisms of alkylation and clinical uses of each class of alkylating agent.

1.2.1 Nitrogen Mustards

Nitrogen mustards were the first agents tested for their chemotherapeutic potential. After the end of the Second World War, nitrogen mustard was found to be able to kill tumor cells. Since then, hundreds of nitrogen mustard-based alkylating agents were tested for their chemotherapeutic potential and a small portion was found to be successful in treating cancer. These include nitrogen mustard (mechlorethamine), and its derivatives chlorambucil, melphalan, cyclophosphamide and ifosfamide. These drugs share a common bischloroethyl group, $Cl(CH_2)_2O(CH_2)_2Cl$, responsible for the alkylation reaction.[47] This group mainly alkylates N^7 on guanines as well as N^3 and N^7 on adenine.[47] Mechlorethamine spontaneously loses a chloride atom in aqueous solutions and forms a reactive ethylenimmonium ion. Due to this rapid activation, the drug is known to cause major side-effects and is generally avoided. However, it is still used in combination with other drugs to treat Hodgkin's disease, lymphosarcoma, leukemia and bronchogenic carcinoma.

Alternatively, cyclophosphamide is not reactive in aqueous solutions and requires activation by cytochrome P450-mediated microsomal oxidation in the liver.[48] When metabolized, cyclophosphamide forms 4-hydroxycyclophosphamide, a non-polar compound which can enter cells readily. In the cell, 4-hydroxycyclophosphamide spontaneously degrades into a reactive bischloroethyl containing phosphoramide mustard or nornitrogen mustard (**Figure 1A**). In order to turn into an active alkylating agent, the bischloroethyl group of nitrogen mustards cyclizes into an imonium ion, a highly reactive alkylating moiety, which can react with DNA and form an alkyl monoadduct. This monoadduct can then form another imonium ion, which can react with a second base, forming a crosslink (**Figure 1B**).[49] This more controlled release of the drug and the cell-permeation mechanism makes cyclophosphamide an effective chemotherapeutic. As such, this alkylating agent is one of the most commonly used chemotherapeutics, able to treat a wide variety of cancer. Unfortunately, a byproduct of the nitrogen mustards cyclization reaction is acrolein, an agent responsible for hemorrhagic cystitis.[50] Of the nitrogen mustards, cyclophosphamide was found to be more potent and produce less acrolein than other mustard-based agents.[47]

The effectiveness of mustard agents is reduced, however, due to DNA repair enzymes in cancer cells. These agents repair the DNA lesions induced by mustard agents, thus reversing their cytotoxic effects. The overexpression of the base excision repair enzyme 3-methyladenine-DNA glycosylase as well as the nucleotide excision repair enzyme ERCC-1 has been attributed to nitrogen mustard resistance in Chronic Lymphocytic Leukemia cells.[51] Additionally, the

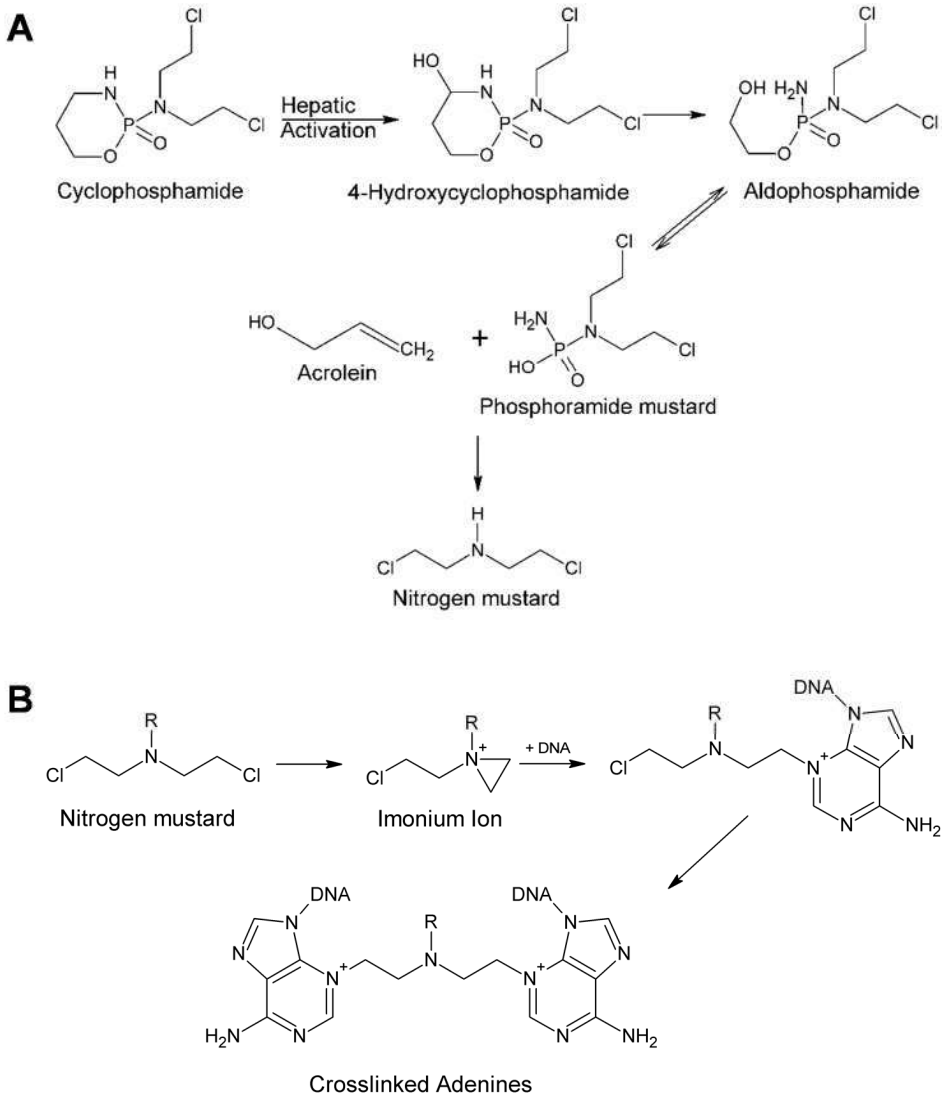


Figure 1: Alkylation mechanism of nitrogen mustards. **A)** Metabolism of cyclophosphamide. **B)** Nitrogen mustard induced DNA alkylation and crosslinking between adenine moieties. Generated by ChemSketch. Adapted from [5].

Rad51-related homologous recombination repair enzyme was found to contribute to removal of DNA crosslinks induced by nitrogen mustards.[52]

1.2.2 Aziridines

The aziridines are characterized by a three membered aziridine ring structure similar to the reactive imonium ion ring formed by nitrogen mustards. The aziridine ring does not carry a positive charge like the imonium ion and is therefore less reactive than nitrogen mustards.

Members of this family include triethylenemelamine, triethylenethiophosphoramide (thio-tepa) and mitomycin C. Thio-tepa has been used since the 1950's and is still one of the most popular aziridines. It is most effective in treating ovarian and breast cancer as well as superficial papillary carcinoma of the urinary bladder.[53] Thio-tepa forms alkylation lesions on N^1 of thymine, O^2 of cytosine, N^1, N^6 and N^7 of adenine, as well as N^1, N^7 and O^6 of guanine.[38] Similar to nitrogen mustards, thio-tepa is a bifunctional alkylating agent able to form guanine-guanine and adenine-guanine crosslinks.[54] The drug is also able to induce monofunctional adducts that lead to DNA strand breaks and apoptosis.[38] Uniquely, the aziridine mitomycin C requires enzymatic reduction to activate the aziridine ring, forming a positively charged imodium-like ring.[55] Once reduced, the ring can interact with DNA, forming cytotoxic crosslinks. Mitomycin C is used to treat bladder, breast, cervix, stomach and pancreatic cancers.[56] The drug is not recommended for prolonged use, however, due to permanent bone-marrow damage, lung fibrosis and renal damage. As with nitrogen mustard, increase in DNA crosslink repair has been attributed to cancer resistance to aziridines chemotherapeutic agents, such as mitomycin C.[57] Specifically, inhibition of the Brca2- Rad51 homologous recombination repair has been found to significantly increase the sensitivity of cancer cells to mitomycin C.[58]

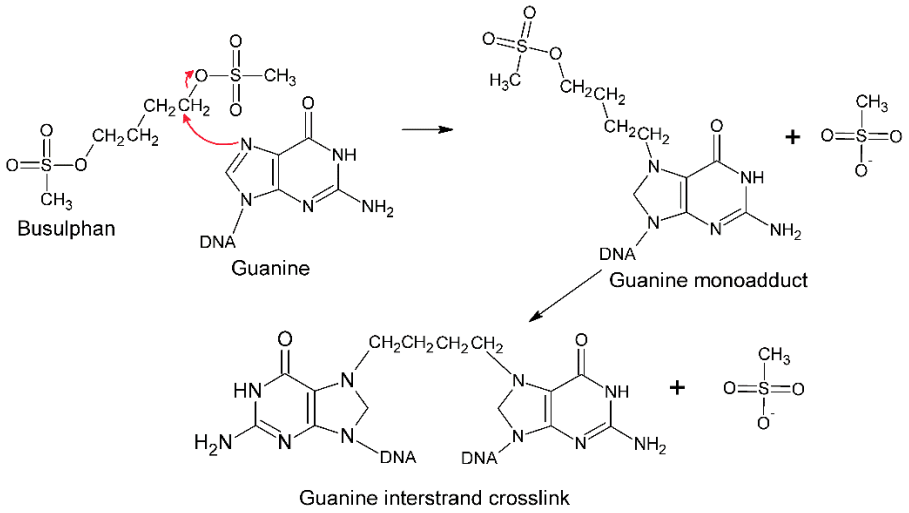


Figure 2: Busulphan mediated interstrand crosslinking between guanine moieties. Generated by ChemSketch. Adapted from [5].

1.2.3 Alkyl Sulfonates

Alkyl sulfonates are sulfonic acid esters with the general formula R-SO_2-O-R'. Busulfan is the only member of this group with clinical applications. The drug forms DNA crosslinks in a mechanism distinct from that of nitrogen mustards and aziridines. Busulfan is able to alkylate N^7 on guanine without the generation of intermediate reactive species (**Figure 2**).[39,59] Busulfan has been used as the drug of choice for treatment of chronic myeloid leukemia for a number of years. Recently, however, the drug was replaced by the more effective imatinib, a tyrosine kinase inhibitor.[60] Busulfan is still in moderate use due to its low cost. As with the nitrogen mustards and aziridins, DNA crosslink repair has been attributed to cancer cell resistance to Busulfan.[61] However, no specific enzyme has yet been identified as a mediator for this resistance.

1.2.4 Nitrosoureas

Nitrosoureas are a group of alkylating chemotherapeutic agents composed of a nitroso (R-NO) as well as a urea group. Commonly, these drugs contain a 2-chloroethylnitrosoureas (CENUs) active group (**Figure 3**). The most common nitrosourea chemotherapeutics are carmustine, lomustine, semustine and chlorozotocin. Due to their lipophilic nature, nitrosoureas are able to cross the blood–brain barrier and are therefore used in treating brain tumors such as glioblastoma multiforme.[62] CENUs are highly unstable and spontaneously form a number of products of which 2-chloroethyldiazene hydroxide is the most important. 2-chloroethyldiazene hydroxide rapidly converts to the alkylating 2-chloroethylcarbonium ion, which produces DNA crosslinks (**Figure 3**).[63] CENUs are among the most active chemotherapeutics, commonly alkylating N^7 and O^6 on guanine.[64] As with other alkylating agents, the effectiveness of nitrosoureas is reduced though the removal of alkyl lesions by DNA repair enzymes. The O^6-

alkylguanine–DNA alkyltransferase repair enzyme has been implicated in inducing nitrosoureas drug resistance.[65]

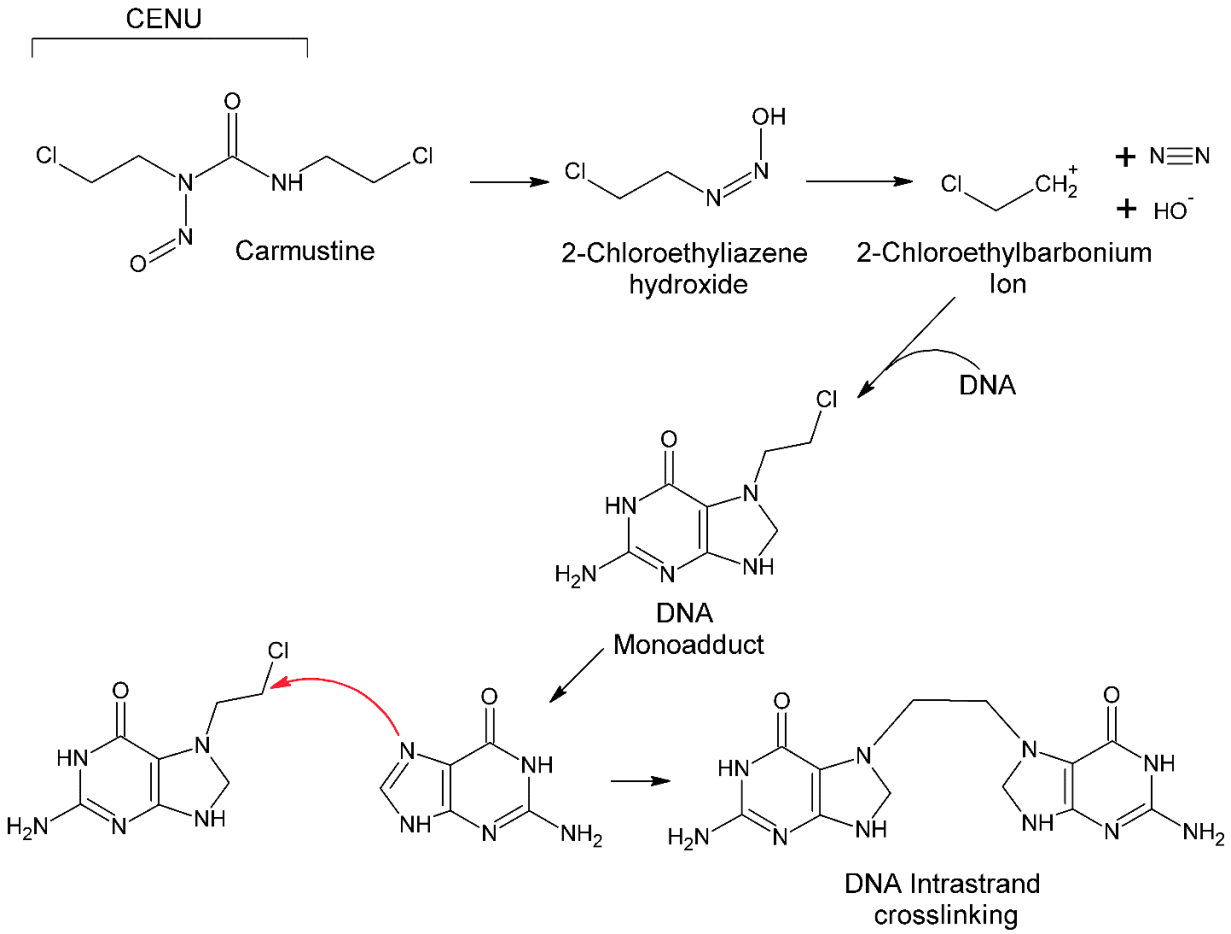


Figure 3: Metabolism of Carmustine which causes DNA crosslinking. Generated by ChemSketch. Adapted from [5].

1.2.5 Platinum-Based Drugs

Platinum-based drugs are arguably the most effective, potent and well-studied chemotherapeutics currently in clinical use. Cis-diamminedichloroplatinum(II) (Cisplatin), the most commonly used drug of this class, is often termed the "Penicillin of Cancer" because it has been highly effective in treating a wide variety of malignancies.[41] The platinum-based drugs can exist in bivalent or tetravalent states. Tetravalent drugs, such as ormaplatin, become biologically active by converting into the bivalent state. Once in a bivalent state, these drugs can exist in cis- or trans-configurations, of which only the former is effective as an alkylating agent. For

unknown reasons, other metals used in place of platinum such as gold, ruthenium, rhodium and palladium, have not been shown to be cytotoxic to cancer cells.[5]

Cisplatin exists in a bivalent state and is composed of two chloro and two amine groups in a cis-configuration. This drug is non-enzymatically activated through a displacement reaction in solution. The chloro groups are exchanged in a stepwise manner with water molecules, resulting in an equilibrium between a variety of species.[66] Of the highly reactive charged species, chloro-monoaquo is the most reactive with DNA at physiological pH (**Figure 4A**). These reactive species interact with DNA, forming monoadducts and crosslinks (**Figure 4B**). Carboplatin, an analogue of cisplatin, has a similar mechanism of action but is less nephrotoxic, due to a slower aquation rate which allows for more controlled dosage. Similar to the previous classes of alkylating antineoplastic agents, platinum-based drugs are rendered ineffective by DNA repair enzymes in cancer cells. Inhibition of DNA alkylation repair enzymes, specifically hABH2, has been shown to increase the effectiveness of cisplatin in treating lung cancer.[67]

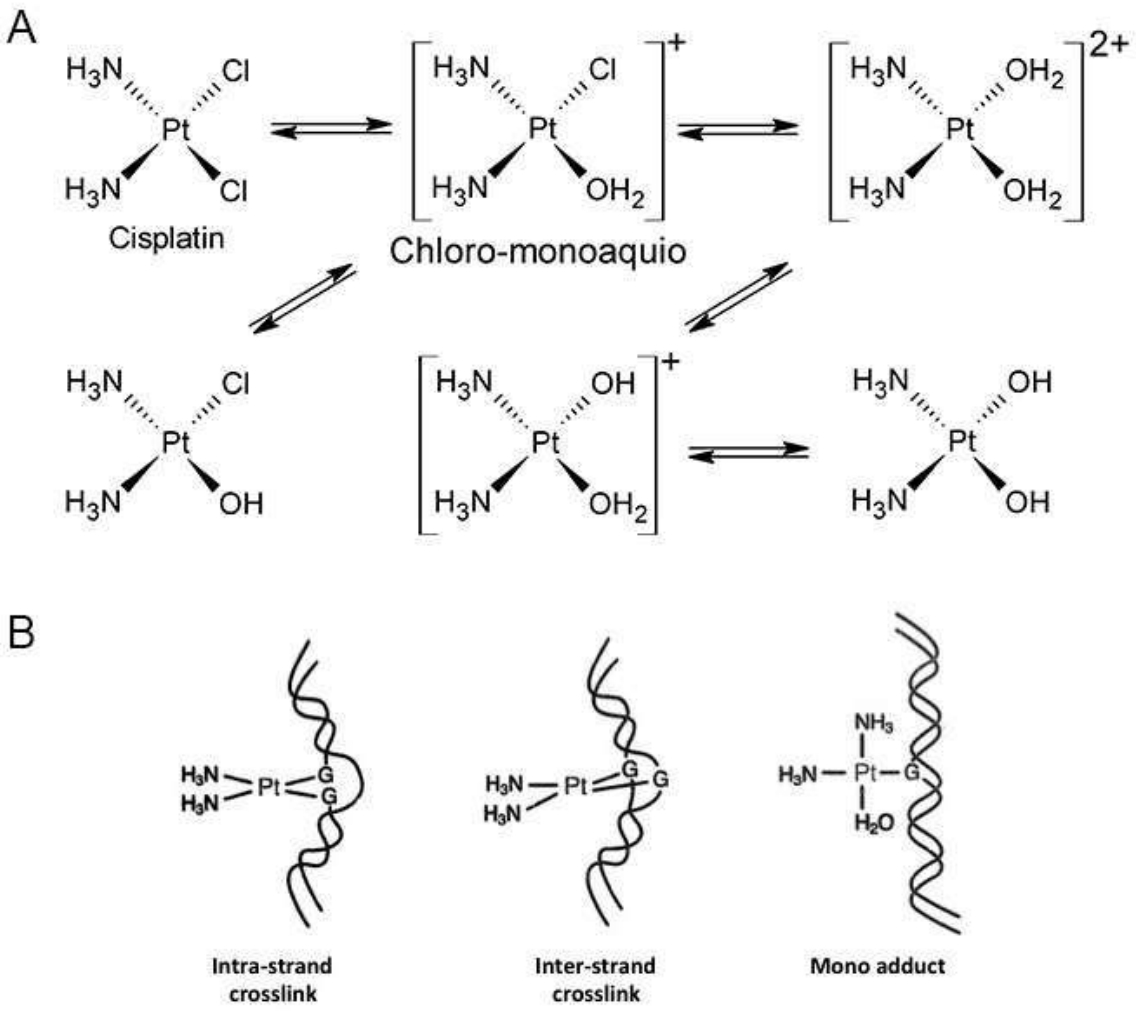


Figure 4: A) Conversion of cisplatin to positively charged reactive species via reversible aquation reactions. B) Cisplatin induced DNA adducts. Generated by ChemSketch. Adapted from [5].

1.3 Alkylation DNA Repair

Alkylating agents have been used to treat a variety of cancers, but their efficacy is limited by the development of drug resistance induced by DNA repair enzymes. [68,69] Such enzymes are able to rapidly reverse the cytotoxic DNA lesions induced by the chemotherapeutic drugs.[35,70] As drug resistance develops, an increased dose of the chemotherapeutic is required to have an adequate effect. Most chemotherapeutics are toxic to cells; as the required dose of drug approaches the limit of the tolerable dose of the patient, the chemotherapeutic can no longer be used.[71] In order to develop more effective chemotherapeutics, it is important to identify and characterize the repair enzymes involved in chemotherapeutic drug resistance.

Both prokaryotic and eukaryotic cells have evolved conserved mechanisms to dealkylate cytotoxic DNA lesions.[72] Alkylation repair mechanisms include base excision repair, nucleotide excision repair, recombination repair, and direct damage reversal.[4,72] Of all these techniques, the direct damage reversal mechanism is the most efficient as it prevents apurinic/apyrimidinic (AP) site intermediates and does not required additional enzymes to maintain DNA integrity.[73]

The direct damage reversal mechanism in *Escherichia coli* (*E. coli)* has been well-studied and resulted in the discovery of similar mechanisms in eukaryotes.[74] When an alkylation lesion occurs in the *E.coli* genome, the adaptive response (Ada response) is activated.[75] The Ada response is initiated when the Ada methyltransferases transfers a methyl group from DNA onto itself. This transfer activates the Ada protein, enabling it to upregulate the expression of AlkA, AlkB and AidB DNA repair enzymes. Ada methyltransferase is upregulated in response to cytotoxic alkylating agents and significantly increases cell survival in response to such agents.[75]

The mechanism of the AlkB family of alkylation repair enzymes remained elusive for nearly two decades. Unlike the other enzymes controlled by the Ada response, AlkB does not have apparent DNA methyltransferase, DNA glycosylase, nuclease, or DNA-dependent ATPase activity.[76] In 2001, bioinformatic protein fold-recognition studies and sequence analysis of the enzyme finally revealed that AlkB, as well as its human homologues, belong to the Fe^{2+} and 2-Oxoglutarate dependent dioxygenases superfamily of enzymes (Fe-2OG-dependent dioxygenase).[74] This has allowed for the development of enzymatic assays, uncovering the specific function of these proteins.

1.4 Oxoglutarate and Fe^{2+} Dependent Dioxygenases

Fe-2OG-dependent dioxygenases were first identified in prolyl and lysyl hydroxylase reactions involved in collagen biosynthesis.[77] As the name suggests, these enzymes catalyze the addition of oxygen from a dioxygen source to a target molecule, requiring Fe^{2+} and 2OG as cofactors. The oxidation of the alkyl groups results in a C–N bond cleavage to produce an unmodified base and a formaldehyde molecule in the case of demethylation. Fe-2OG-dependent dioxygenases all share a conserved double stranded β-helix (DSBH) core with flanking α-helices which coordinate iron binding.[3] **Figure 5** illustrates the proposed mechanism by which the

human AlkB Homologue 2 (ABH2) dealkylates a methylated adenine (1-meA). Briefly, the DSBH core coordinates iron binding through a conserved His-Asp-His motif and three water molecules. Upon 2OG binding, two water molecules are displaced, allowing for the bidentate binding of 2OG (C-1 carboxyl, C-2 keto). The binding of 2OG weakens the bond to the third water molecule, exposing the iron binding site to molecular oxygen (O_2), forming a super-oxo radical anion (O^{2-}). This oxygen radical then forms a peroxo bridge intermediate produced by the attack of the peroxide to the α-keto of 2OG, forming a high-valent iron(IV)-oxo intermediate. The iron-oxo intermediate then hydroxylates the methyl lesion. This hydroxymethyl group is unstable and spontaneously detaches as formaldehyde, resulting in a demethylated base.[1-3]

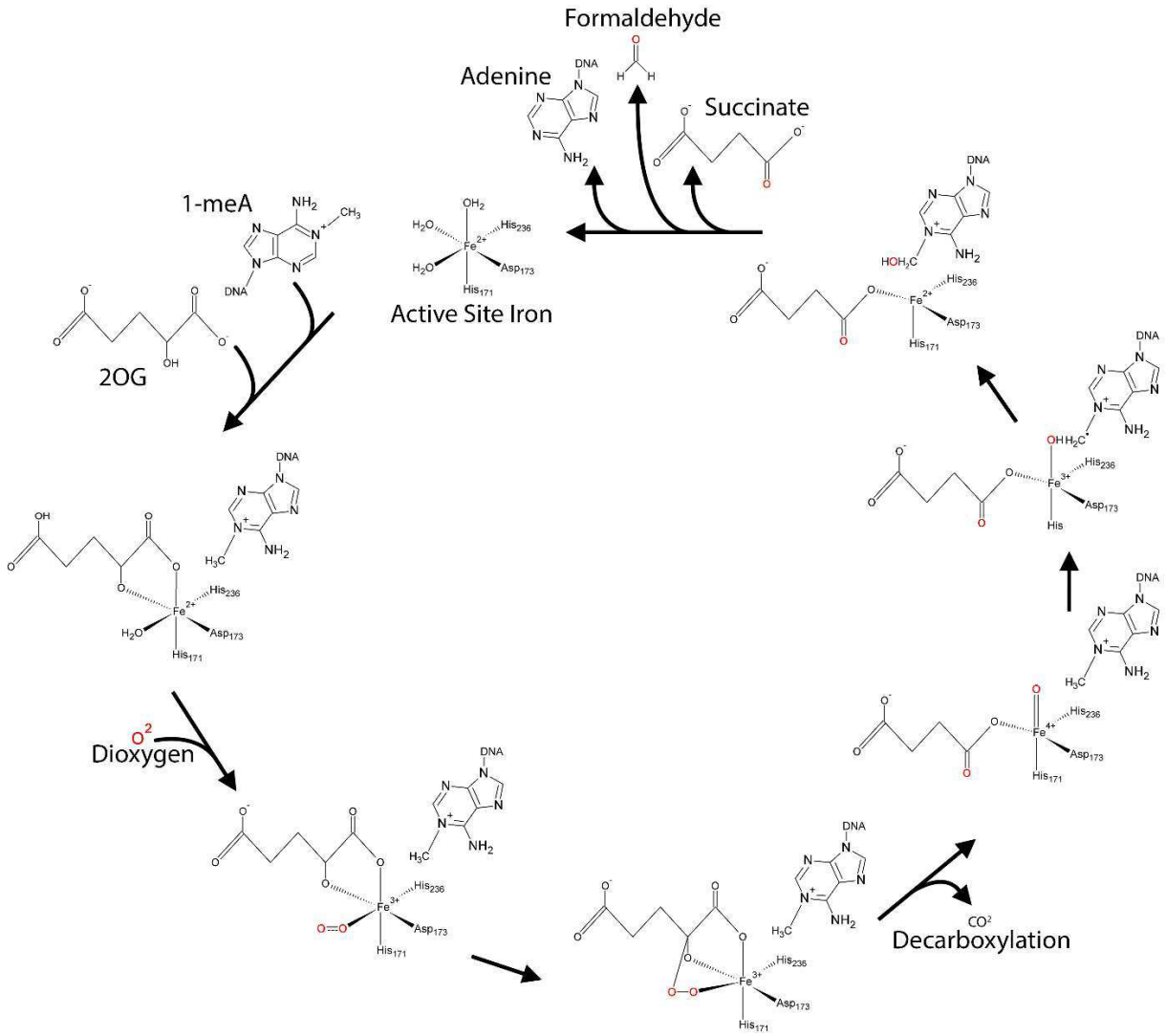


Figure 5: Proposed ABH2 mechanism. Generated using ChemSketch. Adapted from [1-3].

Although this direct repair processes is only known to repair a relatively small number of DNA lesions, the simplicity and practically error-free nature makes it an attractive means for DNA repair in a cell.[73]

Enzymes of the Fe-2OG-dependent dioxygenase family have a wide variety of functions in both eukaryotes and prokaryotes. The human genome has been found to contain over 60 different Fe-2OG-dependent dioxygenase genes.[78] In addition to DNA repair, many enzymes in this family were found to play a role in other physiologically important process, such as chromatin modification,[79] oxygen sensing,[80] and fatty acid metabolism,[81] making them an attractive target for therapeutics.[82] This family of proteins are ubiquitous throughout prokaryotic as well as the eukaryotic domains and is the largest oxidizing family of enzymes which does not have a heme group catalyzing the oxidation of a wide variety of reactions.[3,83]

The identification of AlkB as a member of the Fe-2OG-dependent family of dioxygenases has allowed for the development of AlkB enzymatic assays. AlkB has been found to repair the most common and stable forms of cytotoxic alkylation, 7-methylguanine (7-meG), 1-meA and 3-meC by a direct damage reversal mechanism in double-stranded DNA (dsDNA), single-stranded DNA (ssDNA) as well as RNA.[43]

Humans were found to have nine AlkB homologues termed ABH 1 through ABH8, as well as the Fat Mass and Obesity-associated protein (FTO). Of these, only hABH2 and hABH3 were found to be functional homologues of AlkB. These two proteins have similar substrate specificities to AlkB and were able to repair the most common alkylated targets: nitrogen atoms of purines.[84] The significance of DNA repair in chemotherapeutic resistance has raised the potential of targeting hABH2 and hABH3 to increase the effectiveness of alkylating neoplastic agents. To this end, using the assays developed for AlkB, the function and medical significance of these proteins has been slowly uncovered over the years.

1.5 ABH2

1.5.1 Discovery

Cancer drug resistance is the main cause the failure of many chemotherapeutics.[85] In the case of alkylating neoplastic agents, the involvement of DNA repair enzymes has been well

established.[35,70,86] As the importance of AlkB as an alkylation repair enzyme has emerged in *E.coli*, the discovery of human homologues was motivated by the potential of revealing novel repair mechanisms, possibly involved in the aforementioned chemotherapeutic resistance. AlkB expression in humans cells has been shown to increase cell survival in response to the cytotoxic methylating agent Methyl methanesulfonate (MMS), suggesting that humans may have a similar repair mechanism.[87] To find human AlkB homologs (hABH), Wie *et al*. compared the amino acid sequence of AlkB with a human cDNA database containing ~500 000 human expressed sequence tags, or ESTs (fragments of cDNA library sequences). Using the blastn and tblastn algorithms, a sequence match was found in an EST derived from a human synovial sarcoma cDNA library. The sequence has a 34% identity and 59% similarity at the amino acid level to a region within the C-terminus of AlkB. This region is conserved in all Fe-2OG-dependent dioxygenases.[12] The protein encoded by the full length cDNA sequence contained 299 amino acid residues encoding a 34 kDa polypeptide. As with AlkB, expressing the protein in *E. coli* was able to rescue AlkB-mutants from MMS-induced cell death, suggesting that the protein is a function homologue of AlkB.[88]

In attempt to further characterize the homologue, Duncan *et al.* was not able to reproduce the results obtained by Wie *et al*. Using a BLAST search of the NCBI human DNA sequence databases, the team found two additional potential AlkB homologues. The original AlkB homologue was named ABH1 and the newly discovered homologues were named, appropriately, hABH2 and hABH3. hABH2 has a 30.8% core region identity to AlkB whereas hABH3 has a 23.1% identity to AlkB.[12]

Likely unaware of Duncan *et al.*'s work, two months after the publication of hABH2 and hABH3, Aas *et al.* published the same AlkB human homologue sequences which he discovered while characterizing a uracil-DNA glycosylase gene.[4,89] Due to the established existence of hABH1, they named the homologue hABH2. Follow-up BLAST search resulted in the identification of a second homologue which they named hABH3.[88] Aas *et al.*, like Duncan *et al.*, were not able to replicate Wei *et al.*'s hABH1 results, but did show that hABH2 and hABH3 have activity analogous to AlkB, providing further evidence that hABH2 and hABH3, but not hABH1, are functional human homologues of AlkB. To further characterize the function of

hABH2 in the cell, it is necessary to identify its specific substrates. The similarity of hABH2 to AlkB allowed for the use of existing assays to reveal its function.

1.5.2 Substrate Specificity

The potential of hABH2 to be a true functional homologue of AlkB was extensively tested. Radiometric assays combined with High-Performance Liquid Chromatography (HPLC) analysis (**section 1.8.2**) were carried out to determine the substrate specificity of the enzyme.[12,16] Similar to AlkB, hABH2 was shown to remove both 1-meA and 3-meC, but not 3-methyladenine (3-meA) lesions in ssDNA (**Figure 6**, **Table 1**). Additionally, using a semi-quantitative radioactive formaldehyde release assay (**Section 1.8.4**), hABH2 was found to be 4 times more active on 1-meA than 3-meC. These assays also revealed that hABH2 is completely dependent on 2OG, stimulated by AA and inhibited by EDTA through Fe^{2+} chelating, providing further confirmation that hABH2 is a member of the Fe-2OG-dependent superfamily.[12] Similar to AlkB, using a radioactive cytosine (as opposed to a radioactive methyl group as in previous assays),

Table 1: Known substrate specificity for ABH2 and ABH3

	ABH2	ABH3
Good	1-meA (dsDNA)[4] 3-meC (dsDNA)[4]	1-meA (ssDNA)[12] 3-meC (ssDNA)[12] 1-meA (RNA)[4] 3-meC (RNA)[4]
Moderate	3-meT (dsDNA)[20] 1-meA (ssDNA)[12] 3-meC (ssDNA)[12] εA (dsDNA) [22] εA (ssDNA) [22]	3-meT (ssDNA)[20]
Weak	3-meT (ssDNA)[20] εC (dsDNA) [22] 1-etA (dsDNA)[12]	1-meA (dsDNA)[12] 3-meC (dsDNA)[12] 3-meT (dsDNA)[20] 1-etA (dsDNA)[12]

hABH2 was shown to be able to directly remove 3-meC lesions, producing an unmethylated base without an AP intermediate.[12,16] Using a radiometric assay, similar activity was observed by hABH2 as it successfully repaired 1-ethyladenine (1-etA), however at approximately 0.5% of the rate of 1-meA.[12]

The hABH2 discovered by Aas e*t al.* alongside Duncan *et al.*, show similar demethylation activity, confirming them to be the same proteins discovered in parallel. Using HPLC analysis of radioactive methylated ([^{3}H]methylated) DNA (**Section 1.8.2**) as well as radioactive formaldehyde release assay, hABH2 was able to demethylate 1-meA and 3-meC. The enzymes were also shown to be completely dependent on both Fe^{2+} and 2OG and the direct demethlyation mechanism of a 1-meA and 1-meC was confirmed by LC-MS/MS (**section 1.8.3**).

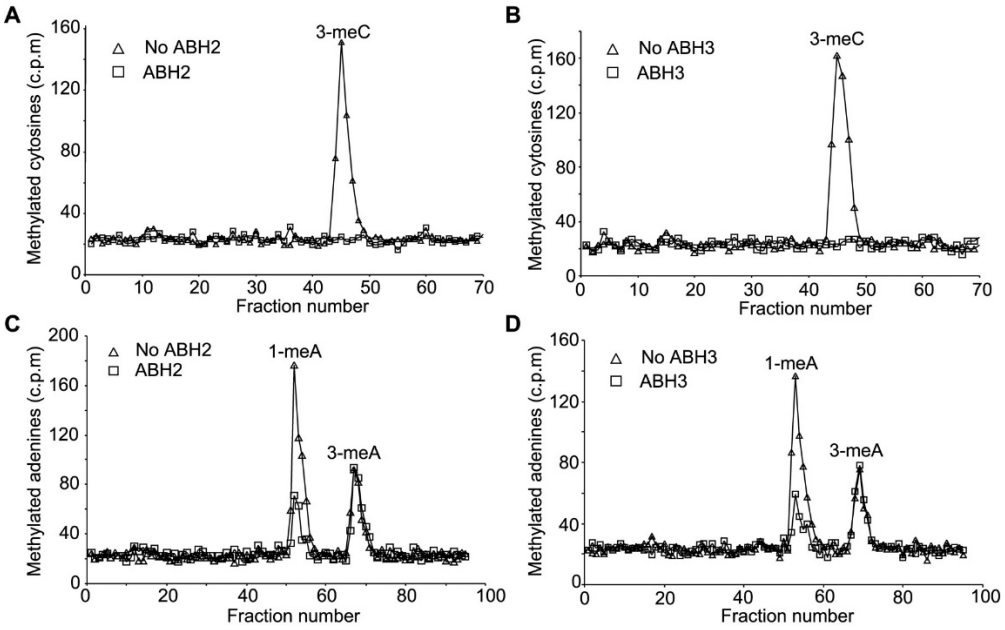


Figure 6: Repair of 1-meA and 3-meC lesions by ABH2 and hABH3. 1 nM of protein was incubated with [^{14}C]MeI-treated poly(dA) or poly(dC) for 30 min under standard assay conditions (section 1.8). The radioactively labeled methylated bases remaining in the substrates were analyzed by HPLC and scintillation counting. Triangle – no enzyme; Square – with enzyme. Adapted from [12].

hABH2 has an apparent preference for dsDNA substrates (**Figure 7A**). However, further analysis revealed that hABH2 has the highest activity towards dealkylation of M13 bacteriophage (ssDNA genome), followed by 1-meA and 3-meC lesions in dsDNA and almost no activity towards similar lesions in ssDNA (**Figure 7B).** Lesions on these bases disrupt base-pair binding and cause local unwinding of dsDNA, thus disrupting replication and triggering cellular checkpoints.[90,91] hABH2, but not hABH3, was able to fully recover M13 as well as λ bacteriophages (dsDNA genome) inactivated by MMS, similar to AlkB. However, hABH2, unlike AlkB and hABH3, was not able to recover ssRNA bacteriophage MS2, suggesting that hABH2 acts mainly on dsDNA.[4]

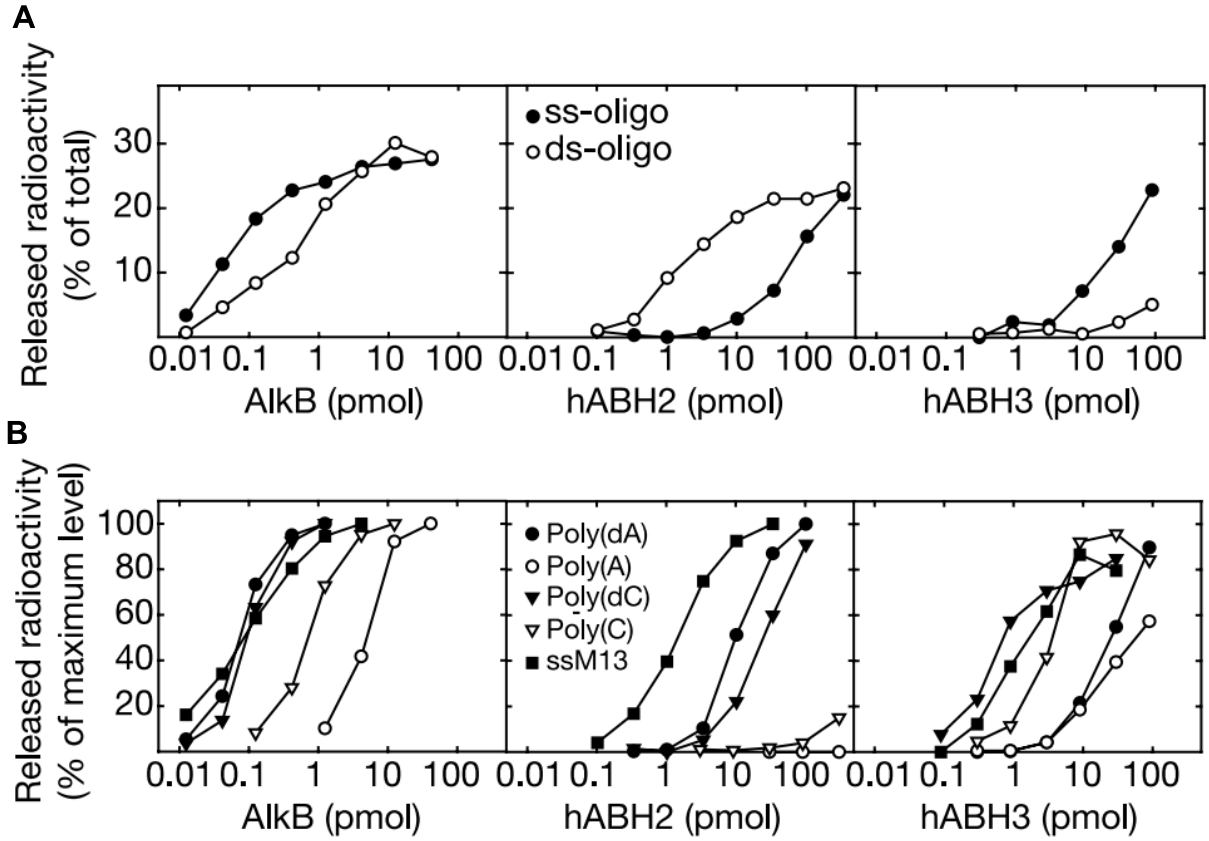


Figure 7: **A**) Comparison of the activity of AlkB, hABH2 and hABH2 towards ssDNA (filled circles) and dsDNA (open circles) substrates. **B**) Comparison of the activity of activity of AlkB, hABH2 and hABH3 towards DNA and RNA substrates as well as M13 ssDNA (filled squares). Enzyme was incubated with radioactive [H^3]methylated substrate. The DNA was precipitated by ethanol and the released radioactivity in the supernatant was measured by scintillation counting. Experiments were carried out at least three times with similar results. Adapted from [4].

Despite Aas *et al.*'s findings that hABH2 has clear preference towards dsDNA, Duncan *et al.* did not observe any such differences. To resolve this discrepancy, Falnes *et al*. tested hABH2 ssDNA and dsDNA substrate specificities using different oligonucleotide sequences under different conditions.[11] hABH2 showed a moderate preference towards dsDNA substrates with no significant difference between substrates with different melting points.[11] The presence of magnesium (Mg^{2+}) was found to affect the apparent substrate specificity of hABH2 towards dsDNA (**Figure 8**). This specificity was observed exclusively when Mg^{2+} was present in the reaction. This is likely caused by the known influence of Mg^{2+} on the structure of DNA as well as its established effect on the activity and specificity of DNA repair enzymes.[92] While Mg^{2+} containing buffers were used in some studies, it was not present in others. The results showed that Mg^{2+} appeared to stimulate the activity of hABH2 towards dsDNA but inhibited it towards ssDNA substrates. It is, therefore, possible that the discrepancy seen by Duncan *et al.* can be attributed to the lack of Mg^{2+} in the reaction. Unlike hABH3, hABH2 is not able to demethylate : 3-methylthymine (3-meT) lesions in ssDNA efficiently; however it is able to repair 3-meT lesions is dsDNA.[20]

To investigate the importance of hABH2 *in-vivo*, mABH2-null mice were generated.[18] hABH2 was knocked out by mutating the Fe^{2+}-binding motif, a highly conserved region essential for enzymatic activity.[14] mABH2-null mice were found to be completely viable. Homozygous mutant mice were phenotypically indistinguishable from heterozygous or wild-type mice well

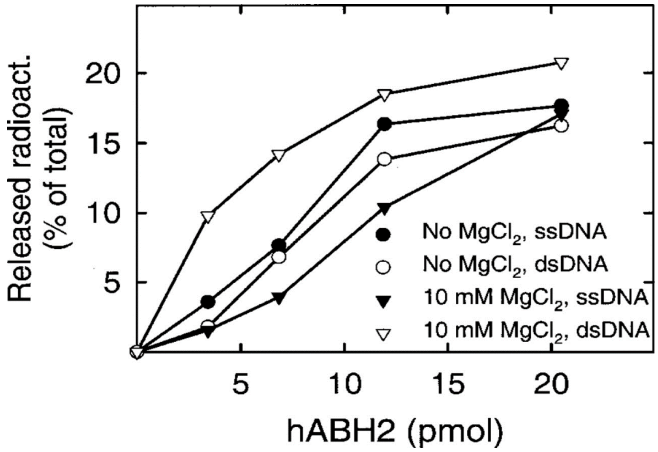
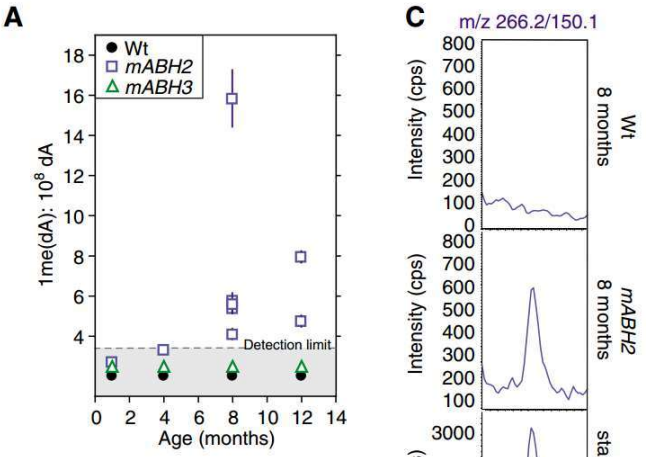


Figure 8: The effect of Magnesium on the preference of hABH2 for dsDNA. [^{3}H]methylated ACGT-oligo (single or double stranded) was incubated with varying concentrations of hABH2 in the absence (circles) or presence (triangles) of $MgCl_2$. The ethanol soluble radioactivity was measured by scintillation counting. Adapted from [11].

into adulthood. Additionally, no abnormalities were observed in the tissues of homozygous or heterozygous mice during histopathological studies. To determine the effect of mABH2 knockout, LC-MS/MS (**section 1.8.3**) was utilized to determine the degree of methylation of DNA extracted from the liver of 1, 4, 8 and 12 month old mice (**Figure 9C**). Whereas no 1-meA was detected in mABH3-null mice, a clear accumulation of 1-meA lesions was observed in mABH2-null mice over the 12 months (**Figure 9A, B**).[18] This finding supports the claim that spontaneous methylation lesions arise in the genome and ABH2 acts as a housekeeping gene, repairing such lesions in the genome as they arise.

Utilizing a restriction enzyme based activity assay (**section 1.8.5**), Ringvoll *et al.* tested the ability of cell extract from wild type as well as mABH2-null mice liver, testis and kidney tissue to repair 1-meA and 3-meC lesions in dsDNA. In all tissues, mABH2 appeared to be the



B

Individual	Age (months)	1me(dA) : 10^8 dA Sample a	Sample b
1	8	5.5	5.2
2	8	4.1	4.0
3	8	6.0	5.5
4	8	5.3	5.8
5	8	14.8	16.8
6	12	7.7	8.1
7	12	4.5	4.9

Figure 9: Accumulation of 1-meA in aging repair-deficient mice. **A.** Accumulation of 1-meA in genomic DNA from liver of 1-12 month-old wild-type, mABH2-and mABH3-targeted mice. The steady-state level of 1-meA was determined by the HPLC-MS/MS. **B**. Numerical values of 1-meA in genomic DNA from liver of 8- and 12–month-old mABH2-null mice. **C**. HPLC-MS/MS chromatogram of 1-meA in genomic DNA from liver of 8-month-old wild-type (upper panel) and mABH2-targeted mice (middle panel), compared with 1me(dA) standard (lower panel). Adapted from [18].

primary enzyme responsible for dsDNA 1-meA and 3-meC lesions repair. mABH3-null extracts were similar to wild type and mABH2/mABH3-null extracts were similar to that of mABH2-null cells, confirming hABH2's importance. Cellular extract from the kidneys, in which mABH2 expression is low, had a significantly reduced methylation repair activity as compared with the liver, where mABH2 expression was high.[18]

In addition to methylated and ethylated DNA, hABH2 has been shown to repair etheno lesions.[22] Such adducts are formed by oxidative stress through lipid peroxidation as well as by vinyl chloride, an organochloride used in PVC production. Etheno lesions were found ubiquitously in genomic DNA in both humans and mice.[93] Two major types of etheno lesions are 1,N^6-Ethenoadenine (εA) and 3,N^4-ethenocytosine (εC). εA adducts are highly mutagenic in mammals.[94] Reduced εA and εC repair has been associated with lung cancer.[95] Alkyl-N-adenine-DNA glycosylases (ANPG) is currently the major DNA glycosylases able to remove εA, leaving behind an AP site.[96] However, this finding has since been disputed as ANPG knockout mice did not show reduced εA lesions upon vinyl carbanate treatment.[97] This has inspired the search to find the enzymes responsible for etheno lesion repair. Recently, AlkB has been demonstrated to reverse etheno lesions, raising the possibility that hABH2, and possibly hABH3, possess a similar activity.[13] Using a restriction enzyme assay (**section 1.8.5**), hABH2 has been shown to successfully remove εA lesions in both ssDNA and dsDNA, with a marginally higher activity towards dsDNA substrates.[22] mABH2 mediated εA repair was also observed in mice cellular extracts. Additionally, hABH2 expression in *E. coli* has rescued etheno induced cell death mediated by ethenolated M13 phage infected *E. coli*, although not as effectively as AlkB. Notably, whereas the above *in-vitro* assays were performed on dsDNA substrates, M13 phage has ssDNA genome. AlkB's preference of single stranded substrates may explain its increased effectiveness over hABH2. Additionally, whereas εA lesions were used for the *in-vitro* assays, chloroacetaldehyde introduces εC more frequently.[98] No significant difference in viability was observed between ANPG-knockout and mABH2-knockout cells treated with chloroacetaldehyde. A time course LS MS/MS analysis of mice with $hABH2^{-/-}$, $hABH3^{-/-}$ and $ANPG^{-/-}$ knockouts over 24 months showed no accumulation of εA in wild-type, $mABH2^{-/-}$, or $mABH3^{-/-}$ null mice up to 12 months of age. However, significant εA accumulation was observed in $ANPG^{-/-}$ as

early as 9 months. This suggests that hABH2 alone is not sufficient for clearing all of the endogenously generated εA lesions.[22]

Once the substrates have been identified, it is important to determine the kinetic parameters, such as K_m and k_{cat}, to accurately measure the enzyme's efficiency and preference towards different substrates. Determining such constants are also important as they help determine the inhibition efficiency of potential chemotherapeutic enhancers.

1.5.3 Kinetic Parameters

To date, no kinetic parameters for the human ABH2 were determined. However, in 2005, Lee *et al*. has determined the kinetics for the mouse homologues of ABH2 and ABH3. mABH2 shares 75.1% identity with the human homologue with the differences in the sequences located mostly in the N-terminus of the proteins, away from the active core region. Such similarity suggests that the activity between the proteins is conserved. This was confirmed as, similar to hABH2, mABH2 was able to rescue AlkB-deficient *E. coli* strains from MMS induced cell death with similar efficiency. Additionally, both enzymes showed comparable demethylation activity towards a variety of substrates.[14]

To obtain kinetic parameters, the team used restriction enzyme assay (**Section 1.8.5**) to measure mABH2 demethylation activity of both 1-meA and 3-meC lesions with both ssDNA and dsDNA substrates (**Table 2**). mABH2 appears to remove dsDNA 1-meA and 3-meC lesions more efficiently than any other substrate-enzyme pair. hABH2 repairs dsDNA lesions three times more efficiently than the most efficient hABH3 repair.[14]

To gain further insight into the protein's specific function, structural and comparative studies were performed. Such studies revealed a great deal about the protein dependence on 2OG and Fe as well as its mechanism of function. Such structural analyses show potential inhibition targets for targeted therapeutics.

Table 2: Steady-state kinetic constants for removal of 1-meA and 3-meC by hABH2 and hABH3. Adapted from [14].

	k_{cat} (s^{-1})	Rel. k_{cat}	K_m (nM)	Rel. $1/K_m$	k_{cat}/K_m (M^{-1} s^{-1})	Rel. k_{cat}/K_m
			1-meA			
ss-ABH2	3.30 ± 0.27	87.3	183 ± 23	1.75	1.80×10^7	4.00
ds-ABH2	13.72 ± 2.0	363	320 ± 73	1.00	4.29×10^7	9.51
ss-ABH3	2.98 ± 0.74	78.8	182 ± 140	1.76	1.63×10^7	3.63
ds-ABH3	1.83 ± 0.038	48.5	263 ± 110	1.22	6.97×10^6	1.55
			3-meC			
ss-ABH2	1.06 ± 0.12	28.0	82.2 ± 22	3.89	1.29×10^7	2.86
ds-ABH2	8.84 ± 0.88	234	167 ± 27	1.91	5.29×10^7	11.7
ss-ABH3	2.06 ± 0.32	54.4	162 ± 48	1.97	1.27×10^7	2.81
ds-ABH3	0.0378 ± 0.0077	1.00	8.40 ± 16	38.1	4.51×10^6	1.00

Relative values (Rel.) are calculated by dividing the lowest value of k_{cat} or K_m by individual values. All values were determined by performing the reactions in triplicates.

1.5.4 Structure

Protein domain analysis using DomainFishing[99] revealed that the C-terminus domain of the hABH2 is highly similar to that of other Fe-2OG-dependent dioxygenases (**Figure 10A**).[99] Similar to other Fe-2OG-dependent dioxygenases hABH2 contained a Fe^{2+}-binding motif ($HXDX_NH$), as well as a conserved arginine involved in binding C^5 carboxylate of 2OG (**Figure 10B**).[12] Mutational studies showed that a single mutation of the aspartic acid (D) in the iron binding motif ($HX\mathbf{D}X_NH$) abolishes the demethylation activity of the protein, likely by preventing iron binding.[14] Mutation of the last histidine in the ($HXDX_N\mathbf{H}$), however, reduced but did not eliminate the function of the proteins, suggesting that this mutation weakens but does not prevent Fe^{2+} binding. Furthermore, the secondary structure prediction of hABH2 revealed a clear alignment between the proteins and the overall crystal structures of known Fe-2OG-dependent dioxygenases (**Figure 10B**). Specifically, the proteins share several β-strands which form a structure known as 'jellyroll' fold. The crystal structure of hABH2 bound to dsDNA was solved in 2008 by Yang *et al.*[6] The structure revealed the aforementioned predicted jellyroll fold. This

fold consists of a C-terminus core double-stranded β-helix (DSBH) in which eight β-strands form a 'jellyroll' structure. This structure is shared by all Fe-2OG-dependent dioxygenases.[3]

The solved crystal structure of hABH2 also revealed differences in the way that hABH2 and AlkB interact with DNA. Whereas hABH2 contacts both complementary strands, AlkB only contacts the lesion containing stand (**Figure 11A and B**). hABH2 appears to use a short

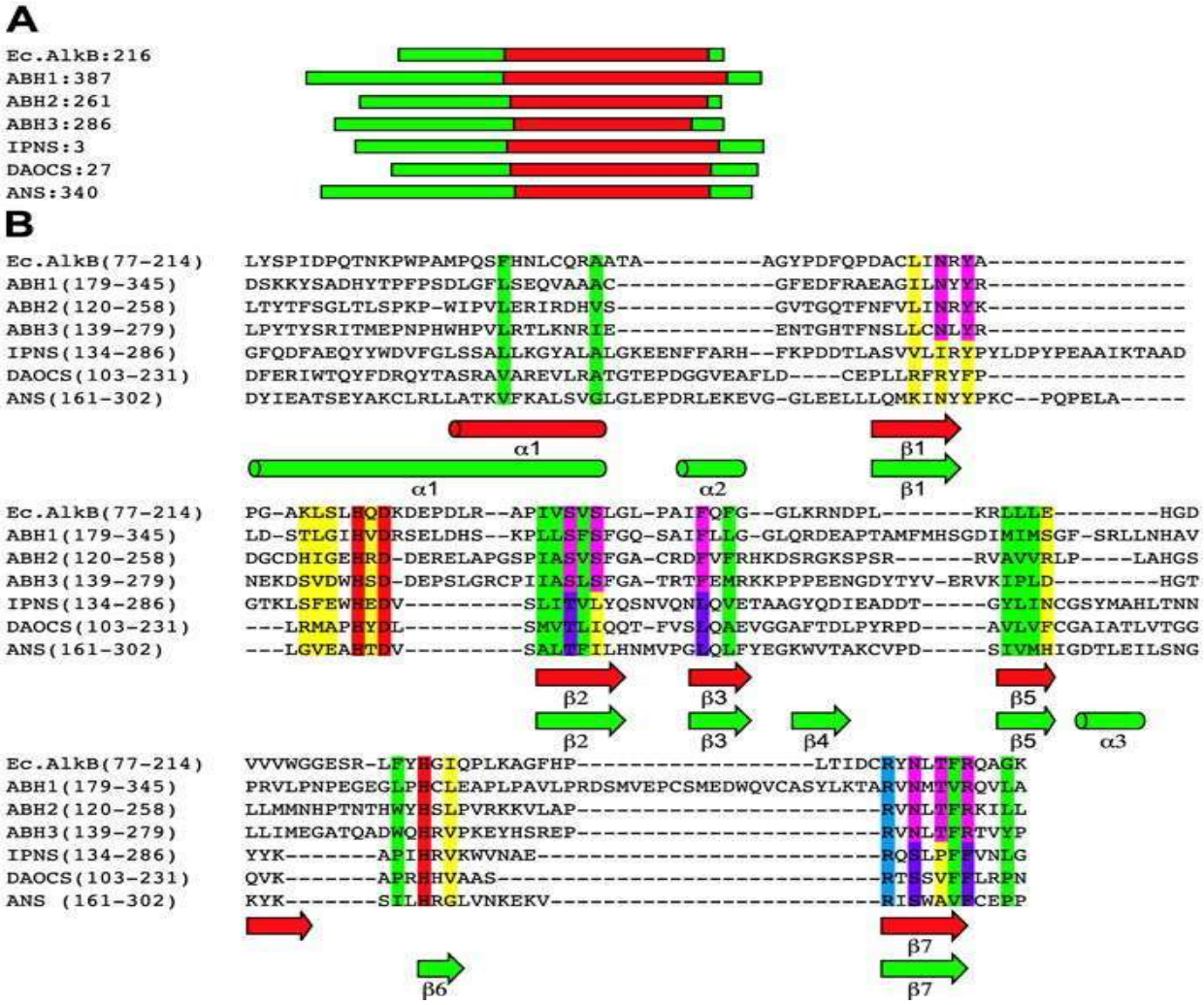


Figure 10: Fe-2OG-dependent dioxygenase sequence alignment. **A**) Similarity of the C-terminus core region of AlkB, hABH1, hABH2 and hABH3 and other known Fe–2OG-dependent dioxygenase: isopenicillin N synthase (IPNS), deacetoxycephalosporin C synthase (DAOCS), and anthocyanidin synthase (ANS). **B**) The predicted secondary structures of AlkB, hABH2 and hABH3 core region aligned with that of known Fe–2OG-dependent dioxygenase. Green - known secondary structures. Red - predicted secondary structures. Background colors: green-conserved hydrophobics; red-conserved Fe^{2+} coordinate; blue-conserved substrate binding pocket; violet-conserved substrate binding pocket of known structures only; magenta - conserved binding pocket for predicted structures; yellow - substrate binding pocket not conserved. Adapted from [12].

positively charged loop (RKK motif, Arg 241–Lys 243) and a flexible, longer loop which binds the complementary DNA strand using Arg 198, Gly 204 and Lys 205 residues **(Figure 11B).** Additionally, the proteins use different base flipping mechanisms to position the methylated base in their active site. hABH2 utilizes an aromatic finger residue (Phe 102), similar to DNA glycosylases, to intercalate into the duplex thereby flipping the base (**Figure 11B,** and **Figure 12A**). AlkB lacks such a finger residue and manipulates the DNA backbone around the base as to stack the flanking bases over one another, resulting in base flipping (**Figure 11A**).

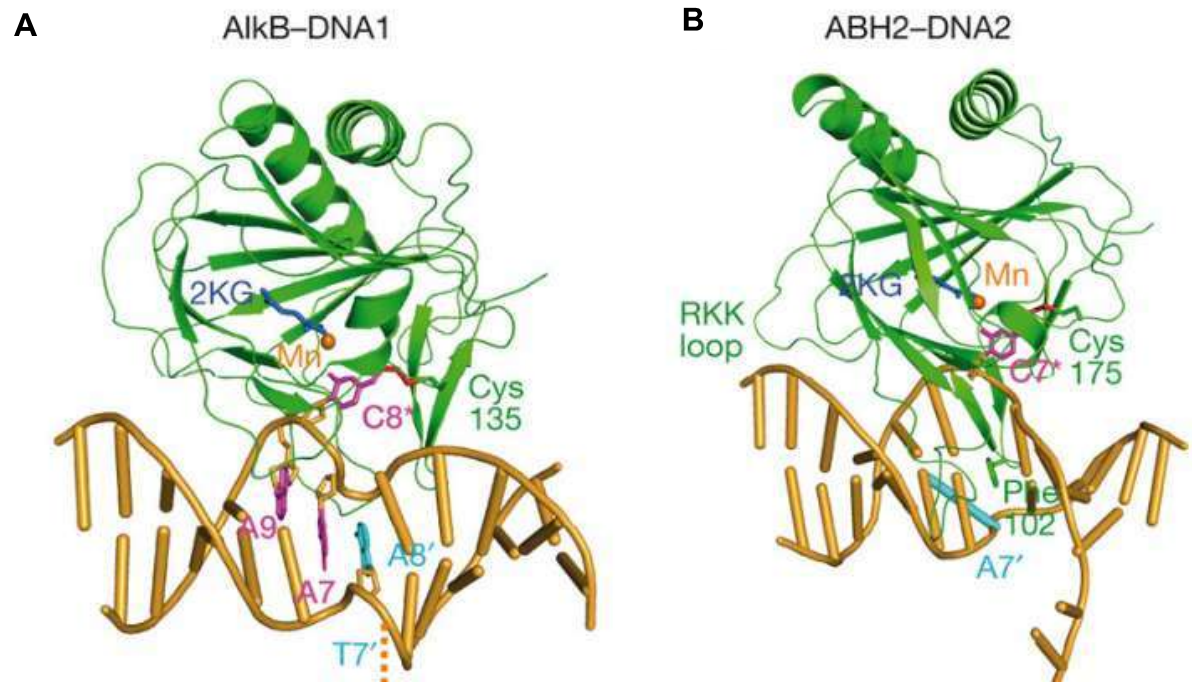


Figure 11: **A)** AlkB–DNA complex. Green - protein (the cross-linked Cys 135 is labeled); Orange – Mg^{2+}; blue - 2OG; yellow- DNA; magenta - flipped base C8* and the two bases flanking C8*; red - disulphide bond. **B)** ABH2–DNA complex. Same colour coding as in b. DNA-binding loop containing the RKK sequence is labeled. Adapted from [6].

The protein structure containing cofactors was determined by soaking the crystals in a solution containing 2OG and Mg^{2+} which occupies Fe^{2+} binding site. The cofactor binding site was found to resemble other Fe-2OG-dependent dioxygenases.[3,6,100] An octahedral structure is formed around the metal ion by His 171, Asp 173, His 236, 2OG and H_2O (**Figure 5**, **Figure 12B**). Residues Phe 124 and His 171 stack on both sides of the 1-meA lesion and position it in the active site (purple, **Figure 12b).**[6] A similar stacking was observed with AlkB.[6] In hABH2, residues Tyr 122, Glu 175, Asp 174 as well as H_2O form a hydrogen-bonding network that interacts with N^6 and N^7 of 1-meA (**Figure 12b)**. Such an interaction was not observed in AlkB where only Asp 135 and Glu 136 together with a H_2O contact N^6 of the methylated base. The negatively charged glutamic acid residue (Glu 175) likely interacts with the positively charged 1-meA base and stacks with the π system of Phe 124. A similar system is observed in AlkB.[6]

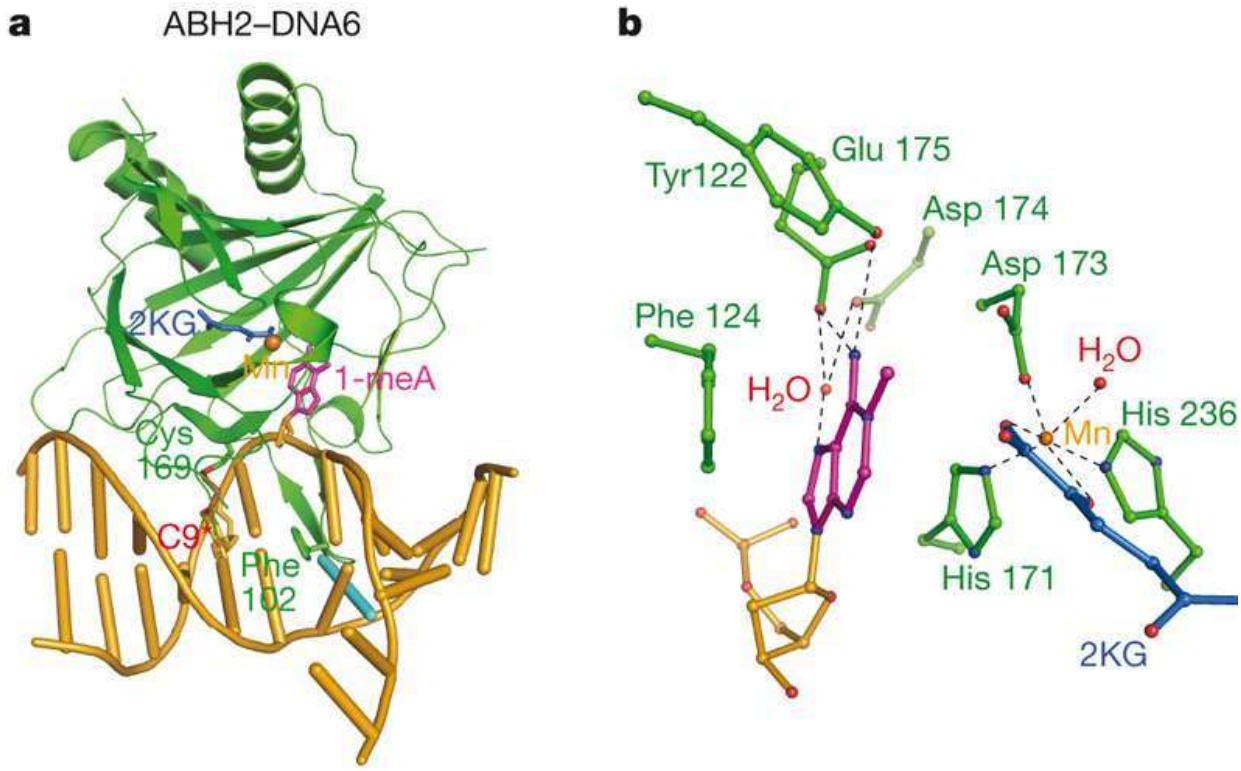


Figure 12:hABH2 base flipping and activity site structure. **a**) hABH2–DNA complex with 1-meA in the active site. Green - protein (the cross-linked Cys 135 is labeled); Orange – Mg^{2+}; blue - 2OG; yellow- DNA; magenta - flipped base C8* and the two bases flanking C8*; red - disulphide bond. **b**) Active site of hABH2 with Mg^{2+} (orange), 2OG (blue) and 1-meA (light magenta). Adapted from [6].

When comparing the structures of AlkB, hABH2 and hABH3, it appears that hABH3, which has a similar substrate specificity as that of AlkB, shares substrate binding motifs with hABH2 more so than with AlkB. The long loop utilized by hABH2 to bind the complementary DNA strand is partially present in hABH3 (**Figure 13**). Further similarity between hABH2 and hABH3 is seen in the finger residue, however, the aromatic ring is absent in hABH3. hABH3 also does not contain the positively charged RKK motif which hABH2 uses to contact the complementary strand of dsDNA (**Figure 13**).[6] In 2010, Monsen *et al.* attempted to identify the source of the divergent function between hABH2 and hABH3 using motif swapping and site-directed mutagenesis.[101] The team focused on the DNA binding finger motif, a feature present in both hABH2 and hABH3 but not AlkB (**Figure 13**). In hABH2, the finger motif is hydrophobic and composed of 20 residues. In contrast, the finger motif in hABH3 is hydrophilic and composed of 17 residues. Deletion of this motif abolishes the demethylation function of both ssDNA and dsDNA substrates in both proteins. Swapping the finger motif between hABH2 and hABH3 resulted in a loss of hABH2 dsDNA function and a greatly reduced ssDNA activity. However, when the hABH2 finger motif is inserted into hABH3, an increase in 3-meC repair activity towards ssDNA, comparable to that of hABH2, and a significant increase in dsDNA

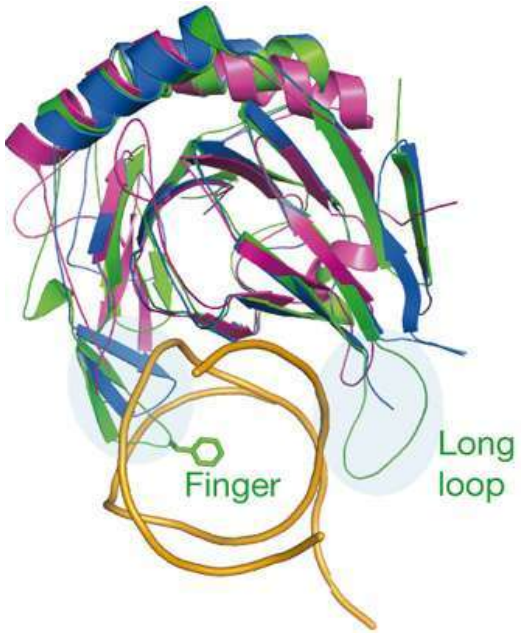


Figure 13: structures of AlkB (magenta), ABH2 (protein in green and DNA backbone in orange) and hABH3 (blue). The crystal structure of hABH3 was obtained without a substrate. Adapted from [6].

demethylation activity is observed. Therefore, the finger motif seems to be essential for the hABH2 dsDNA methylation repair activity.

Characterizing hABH2 would not be complete without expression and localization studies. The extent of the expression of the protein in somatic and pathologic cells could reveal the significance of the protein in cell maintenance as well as in pathology, information that may not be available through structural and substrate specificity studies alone.

1.5.5 Expression and Localization

hABH2 mRNAs expression was found in a variety of human tissues and cell lines. Relatively high levels of hABH2 mRNA were detected in liver and bladder tissues as well as in HeLa (cervical cancer) cells. A later study by Lee *et al.* detected the highest level of hABH2 mRNA in the testis, skeletal muscle, liver, prostate, ovary, and heart as well as lower amounts in the peripheral blood leukocytes, colon, thymus, spleen, pancreas, and kidney. No mRNA was observed in the lung, placenta, or brain tissue.[14]

The hABH2 protein was localized to cell nucleus, where it was present diffusely throughout the nucleoplasm and accumulated in nucleoli.[4] hABH2 exhibits bright, dot-like fluorescence structures in the nuclei but not in the nucleoli. This dot-like pattern might indicate the position of methylated DNA.[25] Outside S phase, hABH2 was homogeneously distributed in the nucleoplasm but showed some accumulation in nucleoli. During synthesis (S phase), hABH2 co-localized with proliferating-cell nuclear antigen (PCNA), a protein known to accumulate in the replication foci, suggesting a similar localization for hABH2.[4] PCNA is a sliding clamp protein important in DNA replication, DNA repair, epigenetic regulation, chromatin remodeling, cell cycle control and sister chromatid cohesion.[102] Many proteins which localize to the replication fork commonly interact with PCNA through a conserved PCNA-interacting peptide (PIP) box. hABH2 has no known PIP box.[4] Follow-up mutagenic studies revealed a novel PCNA-interacting motif termed AlkB homologue 2 PCNA-interacting motif (APIM).[103] A database search has found over 200 proteins containing this motif, many of which are involved in cell maintenance.

The ubiquitous expression of hABH2 as well as its accumulation in the nucleus suggests that the protein is important for genomic maintenance. Additionally, the high relative expression of hABH2 in HeLa cells implies a possible contribution to cancer development. As discussed previously, it is possible that hABH2 maintains cancer cell DNA, allowing it to replicate indefinitely and evade DNA-damaging chemotherapeutics.

1.5.6 Pathological Role

A large portion of chemotherapeutic agents function by damaging the DNA of tumor cells, thereby aiming to trigger cell cycle arrest and apoptosis. Increase in the DNA repair capability of cancer cells may, therefore, significantly reduce the effectiveness of such therapies. DNA repair enzymes are in fact often found to be over-expressed in a cancer cells.[35,104] Therefore, inhibiting such enzymes could increasing the effectiveness of modern chemotherapeutics.[70] Alkylating agents are among the most commonly used antineoplastic agents.[84,104] Recently, Fe-2OG-dependent dioxygenases emerged as important DNA alkylation repair enzymes and have, therefore, been suggested to contribute to chemotherapeutic resistance in various cancers.[27] Over the past decade, increasing evidence emerged to support this claim.

Non-small cell lung cancers (NSCLC) is the most common and deadly forms of lung cancer. Platinum-based alkylating drugs, such as cisplatin, are the main chemotherapeutic used to treat NSCLC.[105] Lung cancer tumors were shown to be able to develop a resistance to these drugs by repairing cisplatin induced lesions, eventually making the chemotherapeutic ineffective entirely.[106] Inhibition of hABH2 using a lentivirus-mediated RNA interference (RNAi) in non-small cell lung cancers cell line was found to increase pro-apoptotic Bax levels (member of the Bcl-2 family) and decrease cancer cell survival in response to Cisplatin treatment.[67] Notably, since Cisplatin is not strictly an alkylating agent as it does not add alkylating lesions (**Figure 4B**), it is likely that hABH2 has the ability to remove lesions beyond the established repertoire described previously (**Table 1**).

Photodynamic therapy (PDT) has emerged as a promising chemotherapeutic alternative to radiotherapy and surgery.[107] Photodynamic therapy excites photosensitizers, such as photofrin, in cancers cells which then induce cytotoxic reactive oxygen species (ROS) production. Clinical

trials using photofrin-mediated PDT (Ph-PDT) has shown success in treating gliomas, a classically difficult to treat brain tumor.[108] However, Ph-PDT resistance in glioma, as well as other cancers, has emerged.[109-111] Ph-PDT has been reported to function though DNA oxidative damage.[112] Although many mechanisms have been proposed to causes Ph-PDT resistance, DNA repair has been often overlooked. Both hABH2 mRNA and protein have been found to be overexpressed in glioma cells that have developed Ph-PDT resistance. hABH2 knockout intensified the cytotoxic effects of Ph-PDT, indicating a strong correlation between hABH2 and Ph-PDT chemotherapeutic resistance. The p53 tumor suppressing transcription factor has been found to bind the hABH2 gene and was upregulated alongside hABH2, suggesting that p53 promotes hABH2 overexpression during PDT resistance.[113,114] Although p53 is a known promoter of DNA repair,[115] the correlation between p53, and enhanced DNA repair leading to chemotherapeutic resistance requires further work.

ABH2 overexpression has also been linked to chemotherapeutic resistance in glioblastoma multiforme (GBM), the most common primary brain tumor in adults, as well as in pediatric brain tumors (PBT).[68,116] Drug resistance has been attributed as the main cause of chemotherapeutics failure in GBM. Temozolomide, a novel alkylating agent not a part of the aforementioned five families of alkylating antineoplastic agents, has shown to be successful in increasing the lifespan of GBM patients by up to 5 years.[117] However, during the course of the therapy, drug sensitivity commonly decreases and tumor growth reappears. In a recent study, hABH2 mRNA levels were found to be expressed up to 4 fold higher in both human GBM tumor tissue and GBM cell lines compared to normal human brain cells. hABH2 levels correlated with the degree of temozolomide treatment, increasing 2-fold 24 hours post temozolomide treatment. Additionally, induced overexpression of hABH2 in GBM cell lines have shown to increase the survival of the cells in response to temozolomide. Follow-up knockdown of hABH2 using siRNA restored the GBM cell line sensitivity to temozolomide.[68] This provides strong evidence that hABH2 has a central role in contributing to GBM drug resistance.

As with Cisplatin treatment of NSCLC, and Ph-PDT treatment of gliomas, hABH2 has been shown to correlate with the apoptotic mechanism in temozolomide treatment of GBM. Upregulation of the p53 pathway was found to downregulate hABH2 expression at both mRNA and protein levels in GBM cell lines.[68] This contradicts the aforementioned finding that p53

promotes hABH2 overexpression during PDT resistance.[113] Possibly, regulation of hABH2 by p53 may differ depending on cellular context, likely by the type and degree of DNA damage. A better understanding of this relationship requires further analysis.

In addition to GBM, hABH2 overexpression has been detected in bladder cancer cells.[118] Urinary bladder cancer is known to have a high degree of recurrence and invasiveness.[119] hABH2 overexpression was found to promote the expression of MUC1, a membrane-bound oncoprotein that is often expressed at the membranes of secretory epithelial cells.[120] MUC1 is overexpressed in a number of cancers where it was found to inhibit apoptosis, enhance invasion and metastasis, and promotes epithelial to mesenchymal transition (EMT).[113,118,121] Knockdown of hABH2 in urothelial carcinoma cell line was shown to increase E-cadherin mRNA expression, a protein involved in cell adhesion. The knockdown of hABH2 was also shown to reduce tumor volume in mice. Overall, these results suggest that hABH2 contributes to urothelial carcinoma progression through the regulation of MUC1-dependent epithial to mesenchymal transition as well as tumor cell growth.[118]

hABH2 has a clear association with cancer proliferation and chemotherapeutic drug resistance. Inhibiting the enzyme is an attractive approach for improving modern chemotherapeutics. Although several Fe-2OG-dependent dioxygenase inhibitors exist, a high affinity, hABH2-specific inhibitors are still sought after.

1.5.7 Inhibition

Apart from siRNAs used to knockout hABH2 in cell culture, no hABH2 specific inhibitions currently exist. As a therapeutic, siRNA suffers from a few setbacks preventing it from becoming a successful pharmaceutical. Since siRNA inhibits at the genetic level, it does not inhibit proteins which can be long lasting and can therefore take hours or days to take effect. Additionally, siRNA cannot penetrate the cell membrane and degrades rapidly in the body. As a result, despite many attempts, no siRNA drug has passed clinical trials.

Inhibitors have been previously developed for other Fe-2OG-dependent dioxygenases, some of which are likely to inhibit hABH2 as well. The first inhibitors were developed to target collagen Prolyl 4-hydroxylases (P4Hs), aiming to prevent excessive collagen synthesis,

potentially treating fibrotic diseases.[122] Classically, the main approach to inhibit P4H was through chelating iron.[123] Additionally, bivalent cations, such as Zn^{2+}, Co^{2+}, Cd^{2+}, and Ni^{2+}, are able to inhibit P4H by outcompeting iron.[122] Similarly, nitric oxide is able to inhibit P4H activity by outcompeting oxygen.[122] Several synthetically generated 2OG analogues were generated to outcompete 2OG, thereby inhibiting the enzyme. Two popular inhibitors are the 2OG analogues, N-oxalyglycine (NOG) and its cell penetrating derivative dimethyl oxalyglycine (DMOG).[124] Originally discovered to inhibit C-P4H, NOG has since been shown to inhibit many other Fe-2OG-dependent enzymes, including AlkB.[125-127] Finally, succinate, a product of Fe-2OG dioxygenase activity, generated from the reduction of 2OG during the enzymatic activity, can inhibit the enzyme at high concentrations.[128]

The use of generic inhibitors as therapeutics is undesirable. As over 60 different physiologically important Fe-2OG-dependent dioxygenases exist, such inhibitors will likely cause widespread side-effects.[78] Although no P4H inhibitors are currently in clinical use, some are in clinical trials.[122,129] For example, a drug targeting the Fe-2OG-dependent dioxygenase γ-Butyrobetaine hydroxylase is in clinical trials as a cardiac protectant.[130] In agriculture, Fe-2OG-dependent dioxygenases are commonly targeted by a class of compounds termed “onium”, used as plant growth retardants.[131]

To find specific, high affinity inhibitors, a method to rapidly discover such inhibitors should be developed. Finding inhibitors is relatively time consuming and labor intensive process. Recently, aptamers have come out as a potential solution to this issue (Section 1.9). High affinity aptamers can be potentially selected rapidly and, in combination with new alkylation repair assays, rapid screening of inhibitors is possible. Finding methods to select such inhibitors to enhance modern chemotherapeutics has motivated this research.

1.6 ABH3

1.6.1 Discovery

hABH3, a close structural and functional homologue of hABH2, was discovered alongside hABH2 using a BLAST search of the NCBI human DNA databases by Duncan *et al.* (**section 1.5**). It was found that hABH3 is identical to the previously known Prostate Cancer

Antigen-1 (PCA-1) protein, a highly expressed protein in prostate cancer found to contribute to apoptosis resistance, and has a 23.1% core region identity to AlkB.[12,132,133] Likely unaware of Duncan *et al.*'s work, hABH3 was rediscovered the following year by Aas *et al.*, using a similar BLAST search.[4]

As with hABH2, the substrate specificity of hABH3 was determined. Although the two homologues are able to dealkylate similar substrates, they repair different subsets of lesions, suggesting complementary rather than redundant functions.

1.6.2 Substrate Specificity

hABH3 is able to dealkylate similar lesions as hABH2, allowing for both proteins to be analyzed together using the same methods and substrates. Using radiometric assays and HPLC analyses (**Section 1.8.2**), hABH3 was shown to repair both 1-meA and 3-meC, but not 3-meA lesions, similar to AlkB and hABH2 (**Figure 6, Table 1**). hABH3, however, is 2 times more active on 3-meC than 1-meA substrates. Furthermore, hABH3 is able to remove 1-etA lesions from ssDNA at ≈0.5% of the rate of 1-meA. Similar to hABH2 and AlkB, hABH3 demethylation activity was found to be highly dependent on 2OG, enhanced by AA and inhibited by EDTA, likely by chelating Fe^{2+}.[12] hABH3 exhibits higher activity towards ssDNA substrates (**Figure 7A**), and preferentially demethylates 3-meC, whereas hABH2 exhibits higher activity towards dsDNA and demethylates 1-meA. (**Figure 7B**).[4] Similar to AlkB, hABH3 is able to efficiently demethylate 1-meA and 3-meC lesions in RNA. The conservation of such mechanism across multiple proteins suggests that RNA demethylation repair is important for genetic stability.[4] RNA modifications can affect both gene expression, transcription and translation. In addition to influencing RNA-protein interactions, RNA alkylation may affect the structural stability and folding of functional RNAs such as tRNA and rRNA.[134] 1-meA lesion in rRNA were found to interfere with codon-anticodon binding,[135] as well as resulted in miscoding by reverse transcriptase.[4,136]

To resolve discrepancies discussed in section 1.5.2, in which some studies, but not others, showed hABH3 preference towards single stranded substrates, Falnes *et al.* compared the ssDNA vs dsDNA activity of hABH3 under different conditions.[11] The presence of Mg^{2+} ions, known to

influence the structure of DNA and DNA repair enzymes,[92] was tested. Unlike hABH2, the preference of hABH3 (as well as AlkB) towards ssDNA was unaffected by the presence or absence of Mg^{2+}.[11] As hABH3 was shown to demethylate RNA substrates, its activity towards different form of RNA was analyzed.[4] hABH3 has significantly higher activity towards ssDNA and ssRNA than DNA:RNA or dsRNA hybrid.

In order to investigate the activity of the protein *in-vivo*, mABH3-null mice were generated.[18] As with mABH2, mABH3-null mice were fully viable and showed no apparent phenotypical differences as compared to wild-type mice. Furthermore, histopathological studies failed to reveal any abnormalities on the tissue level. Using LC-MS/MS (**section 1.8.3**), the accumulation of 1-meA lesions was investigated in the DNA from the liver of 1, 4, 8, and 12 month old mice. As opposed to mABH2, No accumulation of 1-meA was detected in hABH3-null mice (**Figure 9A**).[18] Mouse liver, testis and kidney tissue cell extracts were tested for their ability to repair 1-meA and 3-meC lesions in dsDNA. In all tissues, repair by mABH3-null extracts were indistinguishable from the wild-type extracts, and mABH2/mABH3 null extracts were similar to mABH2-null extracts.[18] This finding supports the role of ABH2 as the main alkylation repair protein of 1-meA and 3-meC lesions in dsDNA which suggests that ABH3 may have a separate role in the cell, likely repairing lesions in ssDNA and RNA.

A restriction enzyme assay (**section 1.8.5**) was performed to determine whether hABH3, like hABH2 and AlkB, is able to repair etheno lesions in ssDNA and dsDNA. Unlike its homologues, hABH3 did not appear to repair ethano εA lesions in either ssDNA or dsDNA substrates.[22] Additionally, unlike hABH2 and AlkB, the expression of hABH3 in *E. coli* cells was not effective at rescuing etheno induced cell death. A time course LS MS/MS analysis of mice with $hABH2^{-/-}$, $hABH3^{-/-}$, and $ANPG^{-/-}$ knockouts over 24 months showed no accumulation of εA in wild-type, $mABH2^{-/-}$, and $mABH3^{-/-}$ mice up to 12 months of age. However, significant εA accumulation was observed in $ANPG^{-/-}$ as early as 9 months. This suggests that hABH2 alone is not sufficient at clearing all of the endogenously generated εA lesions.[22]

The kinetic constants for mABH3 towards 1-meA and 3-meC substrates were also determined. Although the values were determined for the mouse homologue of the enzyme, they

provide adequate primary data for a more accurate comparison of the protein's affinity towards different substrates as well as comparing the enzyme to others in its family.

1.6.3 Kinetic Constants

The first kinetic parameters for the ABH3 demethylation reaction were performed on the mouse homologue of the protein by Lee *et al* (**Table 2**). mABH3 showed a 85.7% identity with hABH3, suggesting that the function is conserved between the enzymes. However, as opposed to hABH2, the bases that are different between mABH3 and hABH3 are spread throughout the proteins. Both mABH3 and hABH3 were found to demethylate 1-meA and 3-meC lesions from ssDNA substrates at a faster rate than dsDNA substrate.[14] To obtain kinetic constants, the team used a quantitative restriction enzyme assay (**Section 1.8.5).** Lee's team obtained parameters for both 1-meA and 3-meC lesions for both ssDNA and dsDNA substrates (**Table 2**).[14]

Further insight into the function of ABH3 in the cell can be obtained by determining the structure of the protein. The function of the protein can be inferred by comparing the substrate binding motif as well as the active site to its homologues as well as to other related proteins.

1.6.4 Structure

The crystal structure of hABH3 was solved in 2006 by Sundheim *et al.* The structure revealed a "jelly roll" fold (composed of the eight of β-strands - β7–β14) common to all Fe-2OG dependent dioxygenases (**Figure 14**).[17] Similar to both hABH2 and AlkB, the active site iron is coordinated by HDH residues (His191, Asp193, His257), as well as 2OG (**Figure 14**).[6] Of these, His191 and Asp 193 were found to be essential for the function of hABH3.[17] Additionally, the crystal structure of hABH3 revealed a hairpin loop which creates a lid-like structure over the active site (β4 and β5, **Figure 14**). A similar structure is seen in AlkB and hABH2 (long loop, **Figure 13**).

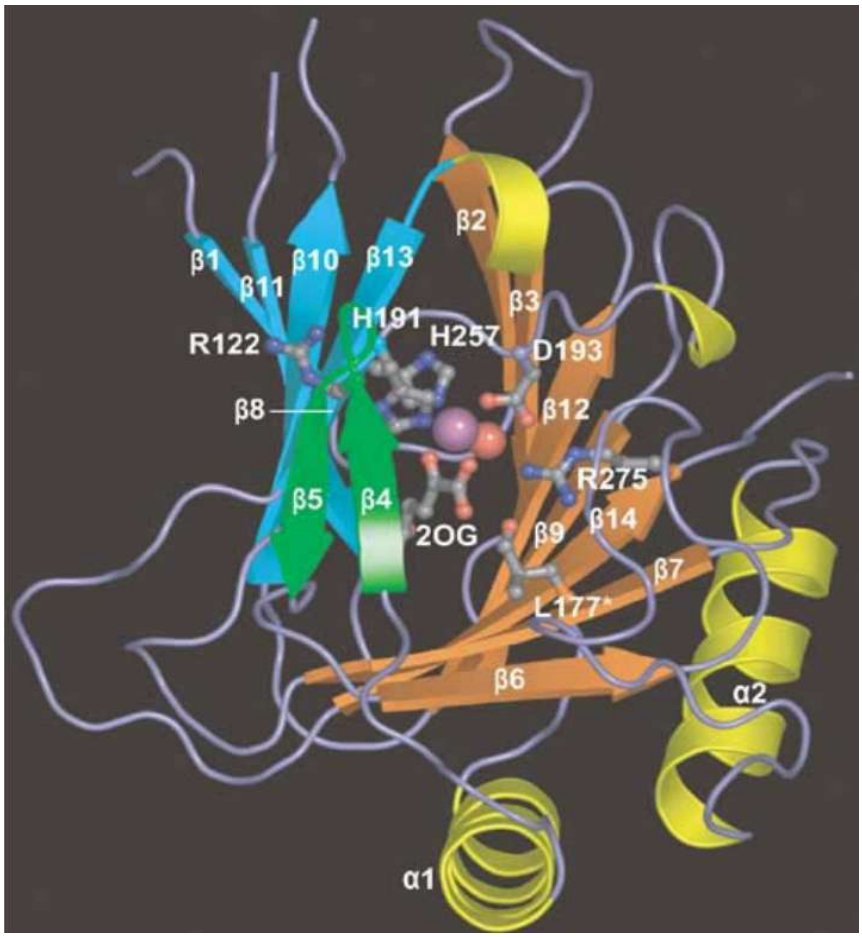


Figure 14: hABH3 structure and active site. The β-strand jelly roll coloured orange and blue. The β4–β5 hairpin is coloured green. α-Helices are coloured yellow, and key residues as well as the co-substrate 2OG are shown as balls and sticks. The iron (violet) and the iron bound water (red) are shown as spheres. Adapted from [17].

Similar to AlkB, The C-1 carboxyl of 2OG interacts with His257 and the C2 keto is interacting with Asp193 in the active site of hABH3.[100] Additionally, the C-5 carboxyl forms a salt bridge with Arg269 and a hydrogen bond to Tyr181, an interaction common to all Fe-2OG dependent dioxygenases. Mutational analysis revealed that these residues were in fact essential to the function of hABH3.[17] Despite having substrate specificity similar to AlkB, hABH3 has substrate binding motif more similar to hABH2 (**Figure 13**). Where AlkB has a hydrophobic substrate binding pocket which sandwiches the substrate (1-meA) between residues His131 and Trp69,[6,100] hABH3 has a more polar binding pocket and lacks a stacking base in the AlkB Trp69 position. Because hABH3 has been crystallized without a substrate, this interaction is unclear. The hABH3 structure suggests that Tyr143 and Arg131 hydrogen bond to N^6 and N^3 of 1-meA. This is partially conserved with hABH2 where residues Tyr 122, Glu 175, Asp 174 interacts with

N^6 and N^7 of 1-meA (**Figure 12**).[6] Additionally, the long loop hABH2 uses to interact with the complementary strand of a dsDNA substrate is partially present in hABH3 (**Figure 13**). hABH3 lacks the positively charged RKK motif which hABH2 uses to contact the complementary strand.[6] However, hABH3 possess antiparallel β4-β5 residues that form a hairpin ideal for intercalating DNA (**Figure 13** and **Figure 14**). Mutation of a conserved Arg122 residue in this hairpin motif, abolishes hABH3's activity towards dsDNA substrates.

The finger motif has emerged as an important feature for the function of the protein (**Figure 13**). The motif is present in both hABH2 and hABH3, however with some differences. Where hABH2 has a hydrophobic, 20 residue finger motif, hABH3 has a hydrophilic 17 residue motif. Deletion of the motif abolishes hABH3 function towards both ssDNA and dsDNA substrates. Swapping the finger motif from hABH2 to hABH3 allows the protein to demethylate 3-meC in ssDNA with efficiency similar to hABH2. This swap also allows hABH3 to repair 3-meC in dsDNA, similar to hABH2.

A unique feature of hABH3 is three negatively charged amino acid residues (Glu123, Asp189 and Asp194) located at the entrance of the substrate binding site. Mutation of either Glu123 or Asp189 to a neutral amino acid increases the activity of hABH3 to ssDNA by 1.5 fold, likely by increasing the accessibility of the substrate to the active site. The mutation of Glu123 to alanine results in a 5 fold increase in activity towards dsDNA, suggesting that the position of that amino acid is important to discriminate against dsDNA substrate. hABH2 has a valine residue in that position, likely allowing for the relatively increased activity towards dsDNA substrates. Additionally, where AlkB has a Val59 residue which fits the 5-methyl group of 3-meT, hABH3 has an Arg131 residue positioned in such a way as to occupy that region. This may explain the higher activity observed by AlkB towards 3-meT compared to hABH3.[20]

Structural analysis of ABH3 suggests that the protein has evolved specific functions different than hABH2. Its ability to demethylate RNA substrates may suggest that the enzyme may function outside the nucleus, repairing messenger as well as catalytic RNA molecules, such as rRNA and tRNA. In fact, ABH8, a close homologue of ABH3, has been shown to be involved in tRNA modifications.[137] To determine whether this is the case, it is important to ascertain the localization of the protein in the cell as well as its expression patterns in different tissues.

1.6.5 Expression and Localization

As discussed in **section 1.5.5**, hABH3 mRNAs was detected at different levels in all the tissues and cell lines. Higher levels of hABH3 were detected in the spleen, prostate, bladder, and colon tissues as well as carcinomas cells lines.[12] Lee *et al.* detected the highest levels in the testis, skeletal muscle, prostate, heart, pancreas, and ovary. Unlike hABH2, hABH3 was mostly absent from the nucleoli and nucleoplasm. The little staining observed in the nucleus did not colocalize with PCNA like hABH2. Diffuse hABH3 staining was observed in the cytoplasm of the cell.[4,25]

Although the substrate specificity of hABH3 suggests that it primarily repairs alkylated RNA,[11] recent finds showed that hABH3 forms a complex with the activating signal cointegrator complex 3 (ASCC3), and relying on it to function.[138] ASCC3 has a large DNA helicase subunit able to unwind DNA, likely providing a ssDNA substrate upon which hABH3 functions.[138] A knockdown of either hABH3 or ASCC3 results in the build-up of 3-meC lesions in a prostate cancer cells, but not in normal cells.

The localization of hABH3, its association with the ASCC3 helicase and its substrate specificity suggests that the enzyme may have roles in repairing both RNA and DNA in and outside of the nucleus. Further studies exploring this could shed light on hABH3 specific function.

1.6.6 Physiological Role

ABH3 overexpression was detected in various cancers, specifically in prostate,[118] bladder[139] and NSCLC.[138,140] The overexpression of hABH3 seems to be cancer specific, and is not present in nonmalignant prostate epithelial cell line, colon adenocarcinoma, osteosarcoma or urothelial carcinoma.[138] The alkylation resistant property of hABH3 appears to be specific to the cancer line as well. Whereas hABH3 knockout induced MMS hypersensitivity in prostate and lung cancer cell lines, it had no significant effect in a non-malignant prostate epithelial cancer cell lines.[138]

NSCLC is one of the deadliest cancers. 70% of NSCLC are adenocarcinomas (glandular cancers of the epithelial tissue). Currently, the number of NSCLC cases are increasing, partially attributed to drug resistance.[141] Recently, immunohistological studies showed that hABH3 is overexpressed in 50% of adenocarcinomas and 56.5% of squamous cell tumors. Interestingly, the inhibition of hABH3 using siRNA resulted in cell cycle arrest in a human lung adenocarcinoma cell line without chemotherapeutic treatment. In mice models, siRNA inhibition of hABH3 resulted in a significant decrease of tumors.[140] The mechanism for this is yet to be determined; however it is likely that this is caused due to decreased DNA repair.[141]

Prostate cancer is the most common form of cancer in men.[142] A new gene, termed PCA-1, later discovered to be hABH3, was found to be greatly overexpressed in 90% of tested prostate cancer tissues. The protein was found to be localized to the nucleus of the cancer cell. The expression of hABH3, as expected, resulted in a resistance to alkylation induced cell death, suggesting that it may induce alkylating chemotherapeutic resistance. PCA-1 has since been suggested to contribute to apoptosis resistance.[132,133] Inhibition of PCA-1/hABH3 using siRNA induced apoptosis, suppressed cell proliferation as well as down-regulated VEGF expression, thereby inhibiting angiogenesis. In contrast, overexpression of hABH3 increased anchorage-dependent growth, invasiveness and apoptosis resistance of the prostate cancer line.[133] Due to the prevalent activity of hABH3 in two of the most aggressive human cancers, it would be a logical target for new chemotherapeutic drugs.

1.6.7 Inhibition

As discussed in section 1.6.7, many generic Fe-2OG-dependent dioxygenase inhibitors exist, which are able to inhibit hABH2 and hABH3 alike. Nickel (Ni) was recently proposed to be such an inhibitor. Ni has been correlated to lung cancer, cardiovascular disease, as well as involved in chromosomal aberrations, DNA strand breaks, excessive ROS production, weakened DNA repair, hypoxia-mimic stress, aberrant epigenetic changes, and signaling cascade activation.[143] Until recently, no clear molecular target responsible for Ni toxicity has been identified. Recently, Fe-2Og-dioxygenases has been proposed to be potential targets.[144] Initial evidence showed that Ni, specifically Ni^{2+}, can induce hypoxia inducible factor 1-alpha (HIF-1α) expression, an important transcription factor involved in low oxygen adaptation. The

transcription factor is regulated though degradation and is stabilized upon hypoxic conditions. The Fe-2OG dioxygenase prolyl hydroxylase domain proteins 1–3 (PHD1-3) is involved in inducing HIF-1α degradation. Ni has been found to inhibit HIF-prolyl hydroxylases by replacing Fe^{2+}, thereby stabilizing HIF1α.[145] Follow-up investigations to determine whether Ni can inhibit other Fe-2OG dioxygenases revealed that it can indeed inhibit hABH3 in a dose dependent manner.[144] Ni^{2+} has an estimated three order of magnitude higher affinity to Fe-2OG dioxygenases than Fe^{2+}.[145] Although still in its early stages, this work shows that the inhibition of Fe-2OG dioxygenases may be the cause of the broad nickel toxicity. As with hABH2, no specific inhibitors for hABH3 exist. Such inhibitors could potentially greatly improve current chemotherapeutics. The development of such inhibitors motivated this work.

1.7 AlkB and the other Human Homologues

Since the discovery of hABH1-3, the advent of genome sequencing allowed for the identification of five additional AlkB human homologues, hABH4-8 as well as the previously known FTO.[146]

1.7.1 AlkB

The AlkB protein, discovered in *E. coli* over 30 years ago, is a Fe-2OG-dependent oxidative demethylase protein.[147] The function of AlkB became apparent during early *E. coli* knockout studies which showed that lack of AlkB increased the sensitivity of *E. coli* cells to MMS.[147] AlkB, as well as Ada and AlkA, were later found to be controlled by the Ada operon, as discussed previously.[76] In contrast to Ada and AlkA, AlkB directly counteracts the cytotoxicity of alkylating agents, and its expression has been found to be induced by such agents.[9,75,147]

AlkB repairs 1-alkylpurines and 3-alkylpyrimidines lesions. Specifically, AlkB has a preference for 1-meA and 3-meC demethylation, however, it can also repair 1-methylguanine (1-meG), 3-meT, 3-ethylcytidine (3-etC) and 1-etA lesions in ssDNA.[9,12,16,148] The preference of AlkB towards ssDNA substrates is consistent with the finding that AlkB binds dsDNA weakly.[149] All of the aforementioned lesions were found to block DNA replication.[29] Cells expressing AlkB were able to restore replication blocked by 1-meA, 3-meC, 3-etC and to a lesser

extent 1-meG, and scarcely restored replication blocked by 3-meT.[29,87] AlkB specificity towards ssDNA as well as the observation that AlkB mutants are more susceptible to MMS while in the growth stage suggests that the enzyme acts at the replication forks or during transcription.[150] Uniquely, AlkB and hABH3 have been found to repair RNA.[4,137,151] Specifically, AlkB was found to repair alkylation lesions on different types of RNA *in-vitro*, such as mRNA, rRNA, tRNA, and viral RNA.[151] AlkB has also showed low activity in repairing alkyl groups larger than methyl and ethyl, such as hydroxyethyl, propyl, and hydroxypropyl.[12,148]

In addition to alkylation repair, AlkB has been observed to repair DNA damaged induced by oxidative stress. The enzyme was found to repair endogenous εA and εC lesions generated as adenine and cytosine react with the breakdown product of oxidatively damaged unsaturated lipids.[13,152,153] Such lesions are toxic to *E. coli* lacking AlkB. Although etheno lesions were found to be repaired in mammal cells *in-vivo* by base excision mechanisms using glycosylases, it is likely that AlkB homologues play a similar role as well.[154] When repairing etheno groups, glyoxal rather than formaldehyde is released as a by-product and can be measured in activity assays.

Recent X-ray crystal structure of AlkB in an enzyme–substrate complex revealed three well-defined regions: a catalytic core (**HxD...H, Figure 15A, B**) in the carboxy-terminal domain, a unique nucleotide-recognition lid and an N-terminal extension.[100] The catalytic core is conserved among all Fe-2OG-dependent dioxygenases. The core consists of two β-sheets forming a jelly-roll structure (**DSBH, Figure 15A, B**) along with a conserved motif that coordinates Fe^{2+} binding **(Figure 15C)**.[3] Additionally, the catalytic core has two conserved arginine residues flanking the last β-sheet that act to stabilize salt bridges with 2OG.[74] AlkB, as well as hABH2, have been found to flip the alkylated base into a deep active site pocket in the catalytic region. Although AlkB and hABH2 use different mechanisms to flip the damaged base, a similar distortion in the DNA strand was observed.[6] AlkB lacks a finger residue present in hABH2, and facilitates base flipping by squeezing together two flanking bases so they overlap

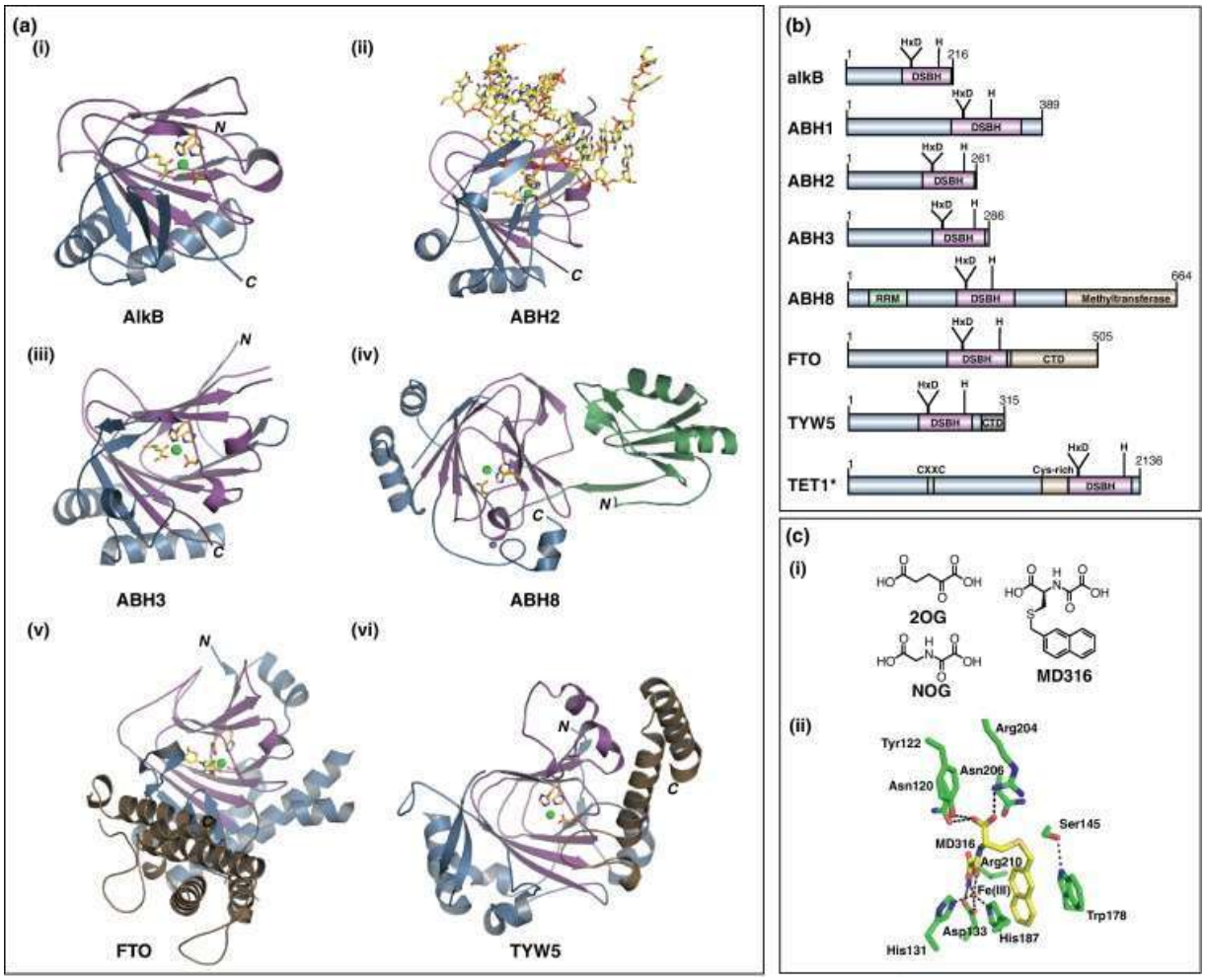


Figure 15: Crystal and domain structures of the Fe-2OG-dependent oxygenase AlkB and its mammalian homologues. **(a)** Crystal structures of AlkB and its mammalian homologues. The conserved double-stranded ß-helix (DSBH, purple) is seen in all of the dioxygenases. Enzyme specific domains (brown or green) help with substrate recognition (FTO) or additional catalysis (ABH8). 2OG (yellow) is seen in the active site with the metal ion (green). **(b)** Domain structures of the Fe-2OG dependent oxygenases seein in (a). The conserved DSBH (purple) and additional domains (brown or green) are indicated. The HxD..H motif is a part of the DSBH and coordinates iron binding. **(c)** AlkB inhibitors (i) structures of 2OG, its analogue NOG (*N*-oxalyl glycine, and MD316. (ii) The AlkB active site co-crystallized with the MD316 inhibitor. Adapted from [19].

one another. Additionally, whereas AlkB makes contact with the damaged strand alone in dsDNA substrates, hABH2 makes contact with both complementary strands, likely contributing to AlkB preference for ssDNA.[6] The nucleotide-recognition lid of AlkB is composed of a flexible β-sheet surrounding the substrate, thus locking the alkylated base over the catalytic-site.[12,148] A fraction of the substrate-binding residues are conserved in all of the eight human AlkB homologs, suggesting similar (but not identical) substrate specificity, however dealkylation

activity has not been observed for ABH1 or ABH4-8.[155] Almost all of the AlkB amino acids in contact with the substrate are in the nucleotide-recognition lid. Residues, W69 and Y76, are conserved between eukaryotes and prokaryotes.[100,146] Additionally, residue D135 in the catalytic core was found to form stabilizing H-bond with the external N6 of 1-meA and 3-meC and not with the N4 groups of 1-meG and 3-meT in DNA, explaining the increased activity towards those lesions.[100]

Detailed genetic mapping of AlkB homologues in the human genome has resulted in the discovery of nine human AlkB homologues, ABH1-8 and FTO as described in the following section.[88,146] AlkB expression in human cell lines showed the same alkylating-resistance properties as in *E. coli* and, equally, the human homologues provided alkylation resistance to *E. coli* AlkB mutants.[82,94] Additionally, AlkB and its homologues function readily *in-vitro*.[7,9] The ability to function in both prokaryotic and eukaryotic cells as well as without the presence of other proteins suggests that AlkB and its homologues function independently, not forming multiprotein complexes.[87,88]

1.7.2 ABH1

Early RNA blot analysis of ABH1 revealed that of the 16 normal adult human tissues analyzed, ABH1 mRNA was found ubiquitously in all of the tissues, including brain, kidney, small intestine, testis, pancreas, prostate, heart, liver, lung, thymus, spleen, placenta, colon, ovary, leukocyte and muscle. This finding suggests that hABH likely plays a fundamental housekeeping role in human tissues.[88] Unlike AlkB and the O^6-meG MTase and 3-meA glycosylase demethylation enzymes, ABH1 expression in human skin fibroblast cells was found to be consistent regardless of treatment with an alkylating agents (MMS).[88]

Early experiments have failed to show any AlkB-like function for the ABH1 protein. This was surprising since the highly conserved core residues found in AlkB, hABH2, and hABH3 (known functioning dealkylating proteins) were also found in ABH1 (**Figure 15A, B**).[12] It is, therefore, likely that ABH1 has a similar activity but for different substrates. Recently, ABH1 has been found to be localized to the mitochondria and low level 3-meC demethylase activity was finally detected.[156] Unlike AlkB and hABH2, however, it does not demethylate 1-meA.

Like AlkB, ABH1 has been found to demethylate ssDNA and RNA exclusively. In addition to its demethylating activity, ABH1 is able to cleave DNA at AP sites (abasic sites) on both single and dsDNA independently of 2OG and Fe^{2+}.[157] The specific biological role of the protein remains unclear.

1.7.3 ABH4-7

The specific function of ABH4-7 is still largely unknown as they do not show similar demethylating activity as their ABH counterparts.[14] As such, they are hypothesized to have novel, yet undetermined DNA or RNA repair activities. Phylogenomic studies have found that ABHs 4–8 are more closely related to hABH2 and hABH3 than AlkB and ABH1. This suggests that ABH 2–8 have a recent common origin and likely arose from duplication events.[146] An exception is the ABH5 gene that has been found to be upregulated under hypoxia and has HIF binding site in its promoter. Although chromatin immunoprecipitation has shown that HIF-1α upregulates ABH5 expression, siRNA inhibition of ABH5 did not affect survival after hypoxia or reoxygenation.[158] No substrate specificity has been identified for ABH5 as of yet.

1.7.4 ABH8

Although not showing similar activity as AlkB, ABH1, 2 or 3, recent studies have begun to shed light on ABH8. ABH8 possesses not only the conserved Fe-2OG-dependent dioxygenase DSBH domain , but also both an RNA-recognition motif and a C-terminus methyltransferase domain (**Figure 15A, B**).[25] The ABH8 sequences is similar to that of AdoMet binding motifs (a common co-substrate in methyl transfers) and to the C-terminal region of tRNA methyltransferase 9 (trm9), suggesting a role as a tRNA methylase.[137] To our knowledge, the exact physiological role of ABH8 in tRNA methylation is still unclear. ABH8 knockout cell lines showed increased sensitivity to MMS and to the bleomycin nuclease (a chemotherapeutic used to treat lymphoma as well as other cancers), indicating a role in alkylation repair.[137] ABH8 expression in HeLa cells was localized exclusively to the cytoplasm, suggesting that the target molecules are likely RNA. ABH8 may in fact possess both demethylation and methylation activities depending on the substrate. Additionally, inhibition of ABH8 using siRNA has been found to reduce bladder cancer tumor growth, making it an attractive target for

chemotherapeutics.[139]ABH8 has been found to be an upstream target for NOX-1, a ROS generating enzyme implicated in inflammation, cell growth, and angiogenesis in various human cancers, specifically colon cancer.[139,159] ABH8 was found to play a role in regulating human urothelial carcinoma cells survival and progression in both in-vitro and *in-vivo*. hABH8 silencing was found to upregulate the JNK, p38, and γH2AX enzymes, known inducer of cell death and was found to strongly suppress the malignancy of bladder carcinoma cells, likely through apoptosis.[139] This finding raises the potential of ABH8 inhibition as a treatment for bladder cancer.

1.7.5 FTO

FTO has been found to demethylate 3-meT and 3-methyluracil (3-meU) in ssDNA and RNA.[160] FTO has the highest activity towards 3-meU in RNA, suggesting that RNA is its preferred substrate. FTO showed no activity on dsDNA.[160] Recently, FTO was found to demethylate 6-methyladenine (6-meA) in both ssDNA and RNA, albeit weaker than 3-meU.[161] FTO dioxygenase has been associated with fat mass and obesity.[162-164] In addition to population genetic studies, FTO mRNA levels were found to be regulated by feeding and fasting in the brain.[81] Although FTO knockout mice displayed no changes in calorie intake relative to normal mice, mice expressing increased levels of FTO showed increased calorie intake. Whether this function relates to its dealkylation activity remains unclear. In addition to the conserved DSBH, the FTO crystal structure revealed a C-terminus α-helical domain not present in the other Fe-2OG-dependent dioxygenases (**Figure 15A, B**).[165]

1.7.6 ABH Localization

mRNA expression of the ABH human homologues has been found in all normal human tissues examined, suggesting a fundamental role in these cells.[12,14,18,25] ABH1 was found to localize to the mitochondria, hABH2 has been found to be present diffusely in the nucleoplasm, and accumulate in the nucleoli, hABH3-7 were detected in both the cytoplasm and nucleus of the cell, and ABH8 was found exclusively in the cytoplasm.[4,12] FTO mRNA was found to be abundant in the brain, particularly in hypothalamic nuclei, likely effecting energy balance.[81]

Uniquely, whereas hABH2 expression appears to form a dot-like pattern in the nucleus, ABH5 showed a dot-like expression in the cytoplasm.[25]

The increased understanding of the Fe-2OG-dependent dioxygenase family of dealkylating enzymes and the existence of established assays has opened the door to finding selective inhibitors to the enzymes.

1.8 AlkB and ABH Assays

As a direct result of the classification of AlkB and its homologues as members of the Fe-2OG-dependent dioxygenase family of enzymes, the specific substrates of AlkB, 1-meA and 3-meC, as well as its preference for ssDNA, was determined.[9,74,150] A number of activity assays have since been developed, uncovering an increasing number of substrates for both AlkB and many of its homologues, as well as their respective functions in the cell.[7,14]

The enzymatic assay conditions for Fe-2OG-dependent dioxygenases are now well established. In general, the reaction mixture contains a Fe^{2+} source, 2OG, ascorbic acid (AA), bovine serum albumin (BSA), an alkylated substrate, and a buffer at pH 7.0-8.0 depending on the enzyme. Although AA is not essential for the reaction, it was found to increase apparent enzymatic activity, likely by regenerating Fe^{2+} from Fe^{3+}.[16] Additionally, catalase may be added to protect the enzyme from hydrogen peroxide.[166] The reaction is stopped either through iron chelation with EDTA, destroying the protein with heat or both. The duration of the assay depends on the concentration and activity of the enzyme as well as the temperature used. The reaction is stopped once the initial linear portion of the product *vs* time curve is surpassed and the curve approaches a plateau (**Figure 16A**).

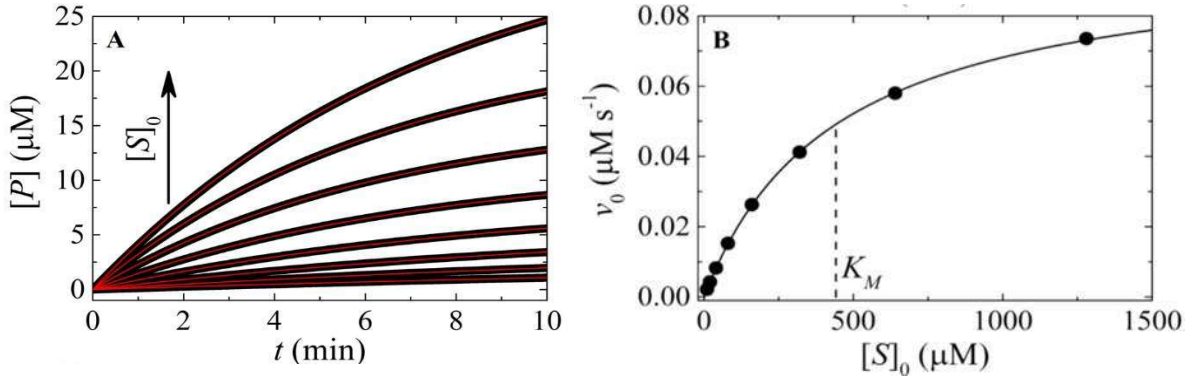


Figure 16: Simulated product formation, [P], of an enzyme over time. Curves for different substrate concentrations ([S]) are shown. **B)** Michaelis-Menten curve of initial rate, v_0, of the enzyme in (A). The initial rate is plotted over substrate concentration, [S], and the solid line represents the nonlinear regression fit using the Michaelis-Menten formula. K_m can be obtained from this fit (shown) as well as v_{max} and k_{cat} (not shown). Adapted from [8].

In **Figure 16A**, the slope of the linear portion is the initial rate (v_0) which is used for the calculation of kinetic constants using the Michaelis Menten kinetics.[167] The initial rate increases as the concentration of the substrate ($[S]_0$) increases (**Figure 16A**). The substrate concentration needs to be significantly higher than that of the enzyme to satisfy the steady-state kinetics assumption used in Michaelis-Menten analysis. Once the initial rates have been obtained for different concentrations of the substrate, a Michaelis-Menten graph can be plotted (**Figure 16B**). By performing nonlinear regression analysis on the plot using the Michaelis-Menten equation (**Eq. 1**), the maximum velocity of the enzyme, V_{max}, the Michaelis-Menten constant, K_m ($½V_{max}$), and the catalytic constant or turnover, k_{cat} ($V_{max} \times [E]$) can be determined.

$$v = \frac{d[P]}{dt} = \frac{V_{max}[S]}{K_m + [S]} \quad (1)$$

Depending on the type of assay used, the reaction mixture is analyzed either in a non-continuous (time-fixed) or a continuous (time-dependent) manner. In a non-continuous approach, a reaction sample is obtained in intervals and analyzed after the enzymatic reaction is completed. Less commonly, the enzymatic reaction in analyzed in a continuous manner, in which the reaction is monitored in real time while the assay is progressing. Although a continuous approach is preferable as initial rates can be determined in real-time, such assays are difficult to achieve as measuring the product without partitioning is seldom possible and often requires an elaborate apparatus. As a result, continuous assays are rare and often indirect, measuring co-

products as opposed to the product. In most cases, in which separating the substrate from the product is required for measurement, a discontinuous approach is utilized. Partitioning often requires a discontinuous approach as, in most cases, it disturbs the equilibrium and alters reaction kinetics. The following section will describe the methods most commonly used to measure Fe-2OG-dependent dioxygenase dealkylation activity.

1.8.1 Radioactivity

One of the simplest and most commonly used assays for dealkylating enzyme activity utilizes scintillation counting of a radioactively labelled compound released during the reaction. The methyl group is commonly labelled with ^{14}C or ^{3}H prior to incorporation into the polynucleotide substrate.[12] Once the enzymatic reaction is stopped, the DNA sample is ethanol precipitated and the radioactive label is quantified either in the supernatant or the DNA pellet.[11,14,168] As aldehydes (a product of the alkyl group after hydroxylation (**Figure 5**)) are soluble in ethanol, the radioactivity of the supernatant corresponds to the demethylated DNA product. Alternatively, the DNA pellet can be analyzed, the radioactivity reflecting the unreacted methylated DNA substrate. As an example, such an assay was performed in 2002 by two separate groups to first show that AlkB had a higher apparent activity towards single stranded methylated nucleotides (**Figure 17**).[9,16] In addition, this assay revealed 1-meA and 3-meC as the specific substrates of AlkB.[9]

An alternative approach to this assay was performed by Welford *et al.* in 2003 to measure the activity and inhibition of AlkB,[169] and by Aas *et al.* in the same year to characterize the substrate specificity of both AlkB and its homologues.[4] In this assay, a conventional activity reaction was set up with the exception of the inclusion of a radioactive 2OG (containing ^{14}C) as opposed to a radioactive methyl as discussed above.[170] After the reaction was completed, $^{14}CO_2$ gas (generated as a byproduct of the oxidative decarboxylation of 2OG (**Figure 5**)), was collected and quantified by scintillation.[169] Using this assay, the efficiency of AlkB in

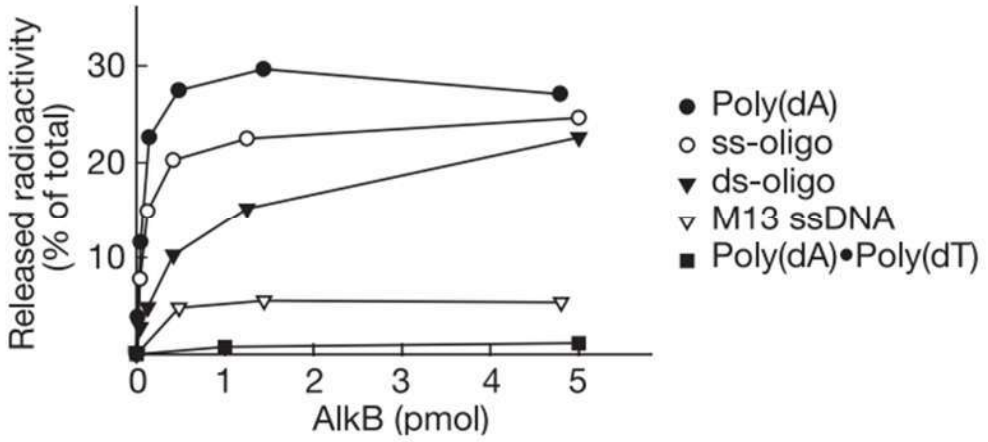


Figure 17: Radioactivity released from [^{3}H]methyl DNA by AlkB. A) AlkB activity on different substrates: poly(dA) (filled circles), single-stranded, A-rich oligonucleotide AAAGCAAGAAACGAAAAAGCGAAA) (open circles), double-stranded oligonucleotide (same sequence as the previous substrate) (filled triangles), M13 ssDNA (open triangles) and poly(dA)•poly(dT) (filled squares). Adapted from [9].

methylating specific substrates, as well as the effect of a variety of inhibitions, were indirectly quantified using 2OG turnover.

Radiometric assays are sensitive and specific as they allow for the labelling of a single atom. They can be designed to provide quick results with minimal equipment. Radioactive assays are also flexible, allowing for the labelling and detection of different components of the reaction. However, the use of radioactive substances is often undesirable. Additionally, ethanol precipitation, used to separate the product from the substrate, is rarely 100% efficient and is sensitive to assay conditions, such as temperature and duration of incubation, thus reducing the accuracy of the assay.[171] To gain additional insight into the enzymatic reaction, radiometric assays are often coupled to a variety of separation techniques such as HPLC as discussed below.

1.8.2 HPLC

Reverse phase HPLC separation (hydrophobic stationary phase) is commonly coupled with radioactivity assays for a more accurate measure of Fe-2OG-dependent dioxygenase activity. In contrast to radioactivity assays, the DNA pellet is subjected to DNA glycosylases or acid hydrolysis to release the methylated nitrogenous bases from the substrate prior to analysis.[9,168] This step is required as HPLC analysis of the unhydrolyzed substrate was found to result in poor separation efficiency of the methylated bases (**Figure 18A).**[9] Once the bases were released, HPLC was able to discriminate between different bases as well as different methylated positions on the same bases (**Figure 18B**).[16,148] This allows, in a single reaction, to identify the substrate specificity of a dealkylating enzyme when a multi-methylated substrate is used. Additionally, this approach can be used quantitatively to measure rate constants.[16,148] After acid treatment, it is clear in that whereas 3-meA and 7-meG are unaffected by AlkB, 1-meA is completely removed, indicating a clear specificity (**Figure 18B)**.

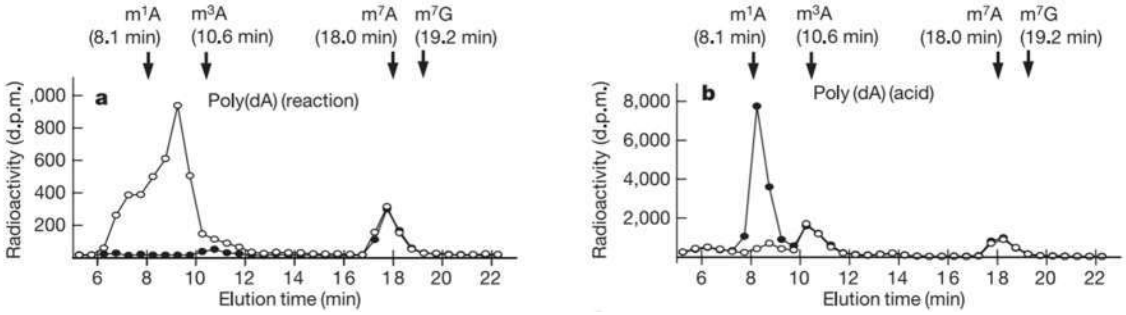


Figure 18: HPLC (Reverse phase) elution profiles of the radioactivity released from methylated ssDNA labelled with [^{3}H] by AlkB with and without acid treatment. **A)** HPLC elution profiles of the intact ethanol soluble supernatant incubations with (open circles) or without (filled circles) AlkB. **B)** HPLC elution profiles of methylated purines post the acid hydrolysis treatment of the ethanol precipitated DNA incubations with (open circles) or without (filled circles) AlkB. The elution time of 1-meA, 3-meA, 7-meA and 7-meG is indicated by the arrows. Adapted from [3,9].

HPLC detection commonly uses radioactive substrate as describe above. However, unlabeled substrates can be used and detected by ultraviolet (UV) absorbance (A_{260}).[9,16,148] When absorbance is utilized, ethanol precipitation is not required, but instead acid hydrolysis of the DNA is performed. Absorbance detection allows for the visualization of both the methylated as well as the unmethylated bases (**Figure 19**). Such an assay was performed to confirm AlkB as a

Fe-2OG-dependent dioxygenase,[4,9,16] to identify the substrate specificity of hABH2 and hABH3 soon after their discovery,[4,12] and to measure various rate constants, such as K_m and V_{max}.[21,148]

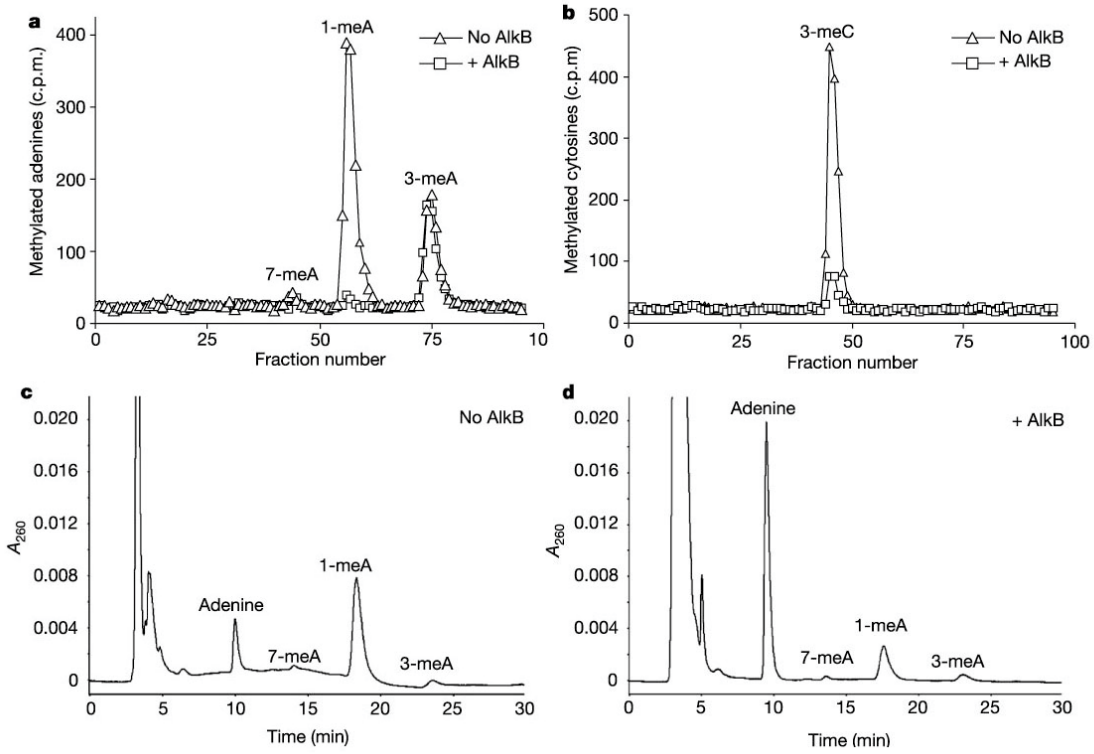


Figure 19: HPLC detection of AlkB induced demethylation. **a, b)** HPLC separation followed by scintillation count of [^{14}C]Methyl iodide-treated **a)** poly(dA) and **b)** poly(dC). HPLC analysis was performed on the remaining unmethylated substrate after incubation of 2.5 pM AlkB at 37°C for 30 minutes. **c,d**) Analysis of non-radioactive methylated adenine with (d) and without (c) AlkB. 920 pmol AlkB was incubated with the methylated substrate at 37 °C for 30 min. Adapted from [16].

The ability of HPLC to discriminate between different methylated bases is beneficial, allowing for the determination of enzyme specificity in a single reaction. HPLC can be used quantitatively, measuring rate constants in a discontinuous manner. However, when used with radioactive tags, the assay suffers from the same setbacks discussed for the radioactive assays previously (**section 1.8.1**). Although the use of UV absorbance (A_{260}) provides the benefit of using untagged, nonradioactive substrates, it comes at the price of a lower detection limit. Additionally, the need to release the nucleobase from the sugar phosphate backbone prior to

analysis makes the assay less direct and may introduce systematic error, reducing the accuracy of the assay.

1.8.3 Mass Spectrometry

Many Mass Spectrometry (MS) techniques have been utilized to futher characterize the Fe-2OG-dependent dioxygenase demethylation reactions. These include Liquid chromatography–mass spectrometry (LC-MS), Gas Chromatography MS (GC-MS) (**Figure 20A)**, Matrix Assisted Laser Desorption Ionisation with Time of Flight MS (MALDI-TOF MS) (**Figure 20B**) as well as Electrospray Time of Flight MS (ESI-TOF MS) (**Figure 20C**). In early experiments, LC–MS/MS has been utilized to estimate AlkB turnover values.[4] Here, the substrate was subjected to nuclease and phosphatase digestion prior to detection.[13] GC-MS, MALDI-TOF MS and ESI-TOF have all been used to characterize and detect the ability of AlkB to repair etheno lesions.[13,29] Although MALDI-TOF MS is able to differentiate between adenine (A), εA, and εA-epocide, an intermediate of the repair reaction, ESI-TOF MS is able to provide a much better separation between the species (**Figure 20A, B**). This separation allows for a time-course analysis of the AlkB etheno repair reaction and shows a clear conversion of εA to A as well as the epoxide and glycol (reaction intermediates) over the 30 minutes (**Figure 20B**). This method has also been shown to work for 1-meA and 6-meA, in which a complete repair of 1-meA but not 6-meA was observed, confirming previous studies.[13] GC-MS was utilized to analyze AlkB repair of εA as well as 1-meA lesions (**Figure 20C**). Here, the demethylation activity is visualized by the accumulation of the formaldehyde co-product (**Figure 5**), whereas the ethanol repair activity is visualized by the accumulation of glyoxal, a co-product of the reaction (**Figure 20C**).

Mass spectrometry is a very powerful tool which is able to detect and identify many of the reaction components. This is not possible with other techniques, in which the studied products need to be predetermined. MS can also be used quantitatively for a variety of reactions. However, optimizing MS analysis conditions is not trivial. Finding ionization parameters may become a labour intensive process. Additionally, MS cannot tolerate salts as they interfere with sample ionization. Salt is often an essential component of biological reactions, especially when

DNA is involved. This requires post reaction modifications, such as desalting, making the assay more complex and less direct.

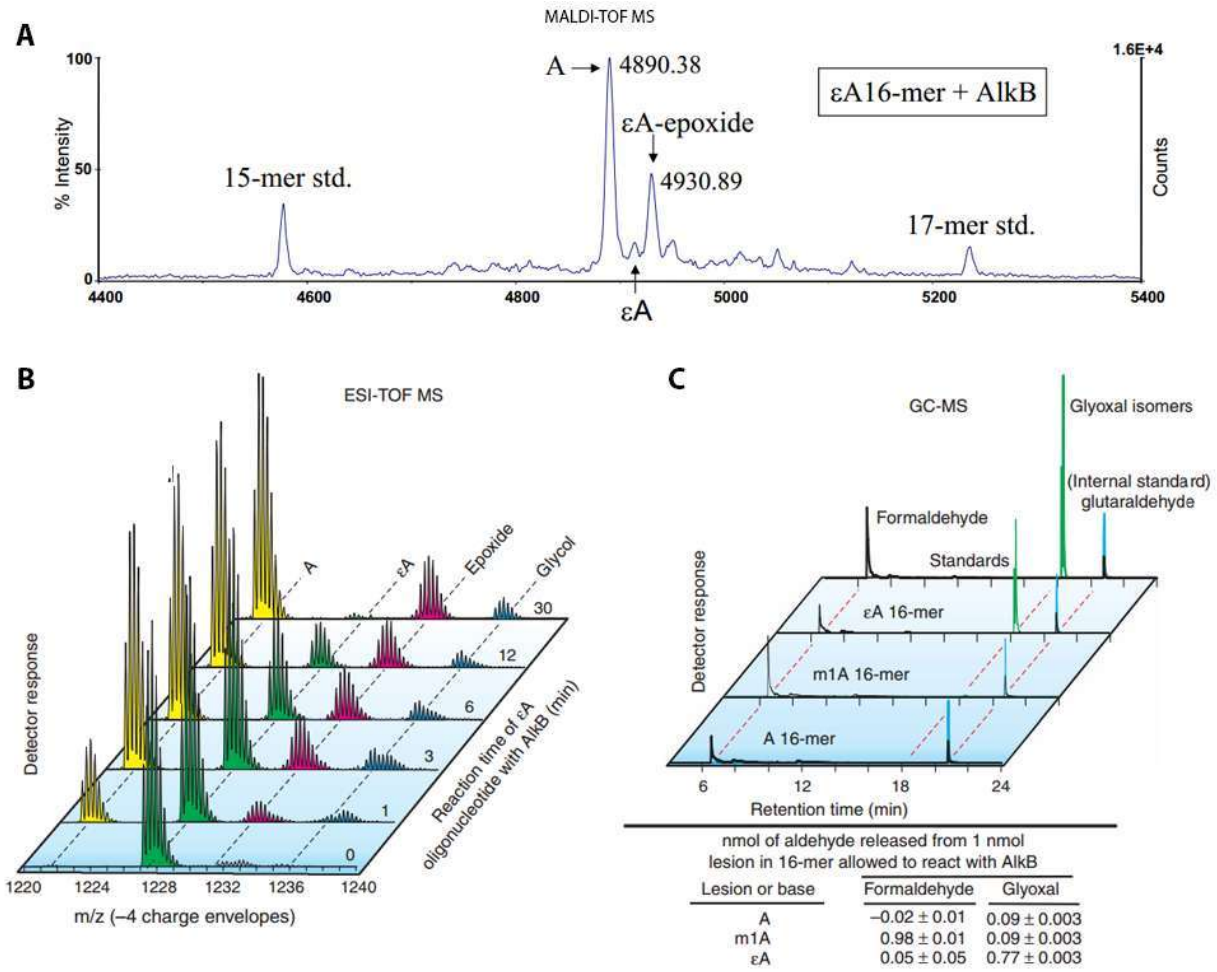


Figure 20: MS analysis of 1,N^6-ethenoadenine (εA) lesion repair by AlkB *in vitro*. **A)** MALDI-TOF MS of εA reacting with AlkB for 90 min. The -1 charged species are shown. **B)** ESI-TOF MS **C)** GC-MS. Adapted from [13].

1.8.4 Continuous Coupled Assay

In 2006, a continuous enzymatic assay was introduced. In this assay, formaldehyde dehydrogenase (FDH) was used to oxidize formaldehyde to formate, reducing coenzyme NAD+ to NADH in the process (**Figure 21**).[7] NADH was then monitored using UV light absorption at 340 nm in real time. As each demethylation reaction releases a single formaldehyde molecule, the absorbance output can be related to product formation using a standard curve. Such an assay

was again used in 2007 by Roy and Bhagwat to obtain rate constants of AlkB to a variety of substrates.[7]

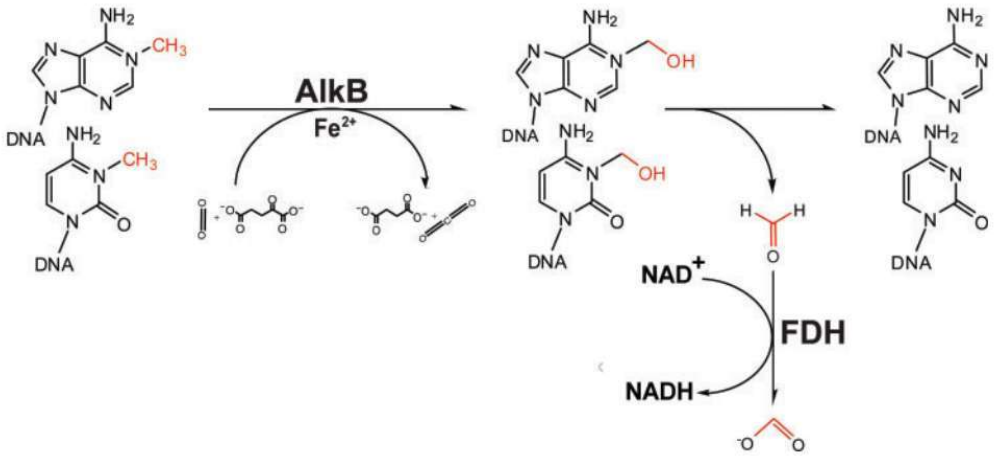


Figure 21: Oxidative demethylation of 1-meA or 3-meC by AlkB coupled with FDH reduction reaction. The AlkB demethylation reaction produces a single molecule of formaldehyde which can be used as a substrate for FDH. FDH oxidizes formaldehyde to formate, converting coenzyme NAD+ to a UV detectable NADH. Adapted from [7].

To my knowledge, this assay is the only continuous quantitative assay utilized for Fe-2OG-dependent dioxygenase dealkylation reactions. The assay is highly beneficial as it allows for the rapid determination of initial rates in a single reaction and in real time. The method allowed the authors to monitor four reactions, and obtain rate constant, at the same time.[7] However, the indirect nature of the assay is a great disadvantage. The assay relies on the activity of a second enzyme (FDH) to convert a coproduct of the reaction to a detectable compound. This introduces a great deal of error to the method. Additionally, A_{340} absorbance has a lower limit of detection than both radiometric and fluorescence detection based assays.

1.8.5 Restriction Endonuclease

Restriction endonucleases were utilized to measure enzymatic activity based on the finding that methylation blocks the enzyme's activity when added to the restriction site. Such an assay was used in 2005 by Lee *et al*. to quantitatively measure the activity of hABH2 and hABH3,[14] as well as investigate the structural features responsible for ssDNA vs dsDNA preference in 2010.[172] In this assay, the oligos themselves were labeled with ^{32}P at the 5' end. A methylated base was selected on the restriction enzyme cleavage site as to interfere with the

nuclease activity. The methylated substrate was then allowed to react with the Fe-2OG-dependent dioxygenase after which the DNA substrate and product were phenol/chloroform extracted and ethanol precipitated. The DNA was then subjected to restriction enzyme cleavage. Since methylation blocks restriction enzyme activity, cleavage only occurred at those sites where the DNA was demethylated, corresponding to product formation. The cleaved DNA product is then partitioned from the uncleaved substrate using a denaturing PAGE gel and quantified by scintillation counting.[14]

Although used quantitatively, this assay is indirect and error prone. The number of compatible combinations of the methylated bases and restriction enzymes is very limited. For example, whereas 3-meC was shown to block HindIII activity, 1-meA had no effect; However, 1-meA was able to block EcoR1 activity. The enzymes KpnI and Hhal only showed partial cleavage.[14] As such, great care needs to be taken when choosing an appropriate restriction site and enzyme. Additionally, the use of a second enzymatic processing suffers from the same setbacks discussed for the continuous coupled assay (**section 1.8.4).** Additionally, as discussed in **section 1.8.1**, ethanol precipitation is seldom 100% efficient. As such, a direct approach is highly desirable. Such a method eliminates many sources of error associated with indirect methods, likely resulting in more accurate quantitative assays.

1.8.6 Direct Approach using CE

In 2009, the first direct method to quantitatively study demethylating Fe-2OG-dependent dioxygenases has emerged.[15] This method uses a commercial capillary electrophoresis (CE) instrument to directly quantify both ssDNA substrate and product differing by a single methyl group (**Figure 22**). The method does not require ethanol precipitation, additional enzymatic modifications and does not rely on detecting co-products. This method uses laser induced florescence (LIF) detection and, therefore, does not require the use of radioactive substances. Additionally, the use of CE has the advantage of high speed and resolution as well as automation, and low sample consumption. Originally, the method has been used to accurately quantitate the activity of AlkB and measure the inhibition of a number of small molecule

inhibitors.[15] The method was later used to probe the inhibition potential of high-affinity aptamers capable of binding AlkB protein. The results proved that aptamers are potent AlkB inhibitions.[127]

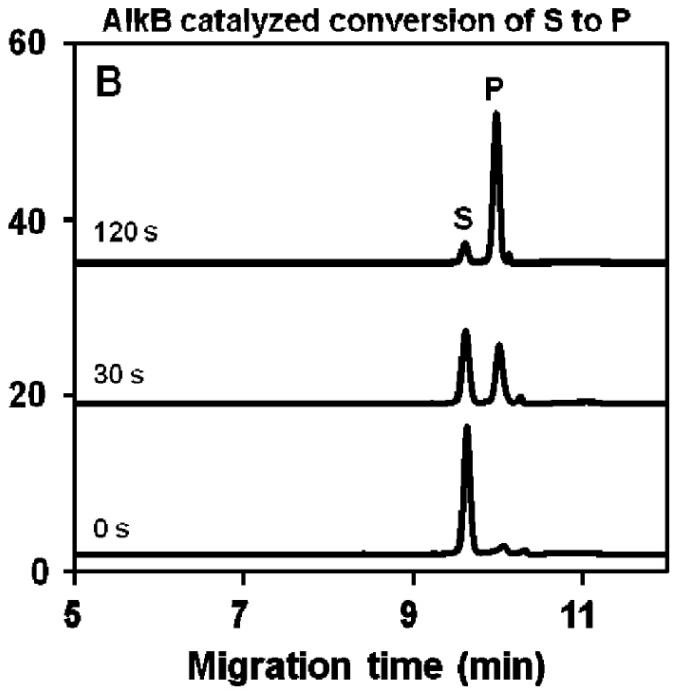


Figure 22: Capillary Electorphoresis with LIF separation of a fluorescently labeled ssDNA substrate (S) from synthetic demethylated DNA product (P) differing by a single methyl. Product is formed from the methylated substrate after 0, 30, and 120 s of the AlkB-catalyzed reaction. Adapted from [15].

The direct quantification of both the substrate and product of Fe-2OG dioxygenase demethylation activity using CE is the first of its kind. The method eliminates many of the disadvantages of indirect methods discussed previously. The use of LIF detection is sensitive and does not require the use of radioactive substances. As opposed to HPLC, the method does not require the hydroxylation of nucleobases and is sensitive enough to separate DNA strands differing by a single methyl group.

1.9 Aptamers

Aptamers are short sequences of ssDNA or RNA strands, usually between 20-100 nucleotides in length, able to form a wide variety of secondary structures. These secondary structures allow aptamers to bind with high specificity and affinity to a variety of targets, including polysaccharides, small molecules, whole cells, and most notably in this case,

proteins.[173] Unlike antibodies, aptamers are synthesized *in-vitro* with ease, high purity and minimal cost. In addition, due to their nucleic acid nature, aptamers can be chemically modified to be tailored for various applications. For example, aptamers can be fluorescently tagged,[174] modified to resist nuclease activity,[175] or altered to increase their binding affinity.[176] The equilibrium dissociation constants (K_d) of aptamers were found to be very similar to that of antibodies.[177,178] In contrast to antibodies, aptamers are produced *in-vitro* and can be selected towards protein targets that do not elicit an immune response. As such, the repertoire of protein targets available to aptamers is larger than that of antibodies. Aptamers are stable in a wide range of temperatures and can be stored for long periods of time without compromising their function.[179] Due to their robustness and versatility, the possible applications of aptamers are almost limitless. Among other applications, aptamers have been used as biosensors in conjunction with electrochemical,[180,181] optical,[182] and mass-sensitive detection modes.[183-185] Aptamers which are selected to internalizable cell surface receptors can be attached to pharmaceuticals and used to improve the efficiency of intercellular drug delivery. This has been previously achieved using a dual aptamer probe to deliver the anticancer drug doxorubicin to prostate cancer cells,[186] and has been suggested as a delivery vehicle for siRNA inhibitors.[187]

Most importantly, aptamers can be used as new class of drugs. This takes advantage of the fact that aptamers can bind to a protein and inhibit its activity, either by causing a change in conformation or by sterically blocking the protein's active site.[188,189] The first therapeutic aptamer, Pegaptanib, discovered in 2000 and approved by the FDA in 2004, is currently being used to treat age-related macular degeneration (AMD), a debilitating disorder which often leads to partial blindness.[190] This aptamer works by specifically targeting and inhibiting the $VEGF_{165}$ isoform, a protein which initiates angiogenesis and increases permeability of blood vessels in the retina which leads to AMD.[191] Currently, there are a number of other aptamers undergoing clinical trials and many more in development.[192] Recently, inhibitory aptamers were selected to AlkB, uncovering their potential as modulators of alkylating chemotherapeutic resistance.[127]

1.9.1 AlkB Aptamer Inhibitors

Due to their potential as inhibitors, aptamers were selected to the AlkB enzyme in 2011.[193] The aptamers were selected using a Systematic Evolution of Ligands by EXponential

enrichment (SELEX) method which utilizes repetitive rounds of Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures (NECEEM) separation, followed by symmetric and asymmetric amplification. NECEEM is a homogeneous separation method that has the great advantage of using unmodified proteins in solution as the target. Using this method, aptamers with K_d values ranging from 20 – 160 nM were obtained. Aptamers selected to unmodified proteins typically have an increased chance of acting as inhibitors *in-vivo*. When aptamers are selected using heterogeneous partitioning techniques, the protein is immobilized on a surface, a process which may alter the protein structure. Aptamers selected to such a protein may therefore be unable to bind the protein in its native form. These aptamers were indeed found to be able to inhibit the AlkB enzyme with unmatched efficiency.[127] This finding has motivated this research to find similar inhibitors to hABH2 and hABH3.

1.10 Research Objective

Selecting inhibitors towards human dealkylating enzymes is an essential step towards improving the efficiency of chemotherapeutics. The goal of this research, therefore, is to select an aptamer which could potentially inhibit hABH2, an AlkB human homologue. Firstly, to accomplish this goal, the current aptamer selection protocol was optimized for efficient selection towards complex targets, such as sensitive, charged proteins which may pose a challenge for NECEEM. After optimization, the new protocol was used to select a panel of high affinity hABH2 aptamers. Lastly, a method to determine the DNA repair activity of demethylating enzymes was developed and tested for hABH2 and hABH3 and proved to be a viable approach to test the inhibition ability of the aptamers.

Chapter 2: Aptamer Selection Optimization

This material is based on a published article: **Yufa, R.** Krylova, S.M., Bruce, C. Bagg, E.A. Schofield, C.J. Krylov, S.N. Emulsion PCR significantly improves NECEEM-based aptamer selection, allowing for efficient and rapid selection of aptamer to unmodified ABH2 protein. Analytical Chemistry 2015, 87, 1411-1419.

Contributions: performed all selection and binding experiments, analyzed all data and prepared figures, participated in majority of manuscript writing.

2.1 Introduction

As discussed above, inhibiting DNA repair enzymes has the potential to greatly improve chemotherapeutics. Recently, aptamers have emerged as promising affinity probes able to inhibit enzymes better than any other known small molecule inhibitor.[127] Although aptamers are theoretically easily selected, recent attempts to select aptamers for the hABH2 repair protein have failed. As such, exploring the possible factors attributing to this failure may potentially result in more successful selection of aptamers to a wide variety of protein targets.

2.1.1 SELEX

Aptamers are typically obtained by a theoretically straightforward process termed SELEX. SELEX involves repetitive rounds of two alternating processes: 1) partitioning of protein-bound DNA from free DNA, and 2) PCR amplification of the collected target-bound DNA (**Figure 23A**). Despite the theoretical simplicity of SELEX and promise of aptamers, useful aptamers have been obtained for a much smaller number of protein targets than antibodies, in spite of the significant efforts over twenty five years.[194] Here we explore a possible reason and apply a solution to this issue.

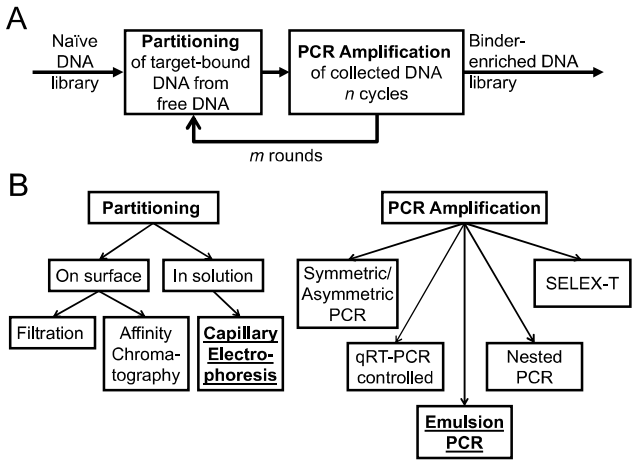


Figure 23: The concept of aptamer selection: major steps (A) and (B) major approaches for partitioning and PCR amplification.

One possible cause for the failure of conventional SELEX may arise from the classical partitioning methods used (**Figure 23B**). Partitioning has been typically done by surface based heterogeneous methods: either filter-based separation, in which filters retain protein-DNA complexes better than free DNA, or affinity chromatography, in which the protein is immobilized on a surface and retains protein-bound DNA.[195] These surface-based techniques suffer from non-specific adsorption of free DNA to the surface of the filters or stationary phase. The non-specific binding reduces the efficiency of partitioning and leads to the necessity of multiple rounds of partitioning/amplification as well as negative selection steps, typically requiring between 10 and 30 rounds of SELEX.[196-198] A homogeneous environment, in contrast, is performed in solution without the need of target immobilization.

2.1.2 Homogeneous Partitioning

Homogeneous partitioning methods are performed in solution, thus avoiding non-specific binding and the adverse effects associated with it. Ideally, SELEX partitioning is done under such conditions. Homogeneous methods are highly desirable for the selection of aptamers to protein targets in cases where the aptamers are to be used *in-vivo*, such as for therapeutic applications. The immobilization process used in heterogeneous partitioning techniques can affect protein structure; potentially reducing biological relevance. Aptamers selected to a protein modified by immobilization may be unable to bind the protein in its native form. Therefore, additional confirmation of their affinity to the native structure of the protein using solution-based

approaches is required. In contrast, aptamers selected to unmodified proteins *in-vitro* using homogeneous techniques are more likely to maintain their function and binding affinity *in-vivo* and do not require extra confirmation steps.

2.1.3 Kinetic Capillary Electrophoresis

Methods employing Kinetic Capillary Electrophoresis (KCE) have been suggested as an alternative to the surface-based methods.[199] KCE methods employ CE separation of species which interact during the electrophoresis process. These methods are homogeneous and, therefore, do not suffer from non-specific binding in the way surface-based methods do. KCE methods can drastically increase the efficiency of partitioning and facilitate selection of aptamers with desirable binding parameters.[200] Additionally, KCE methods do not assume equilibrium and can therefore facilitate measurements of equilibrium and rate constants of aptamer-protein binding, allowing for easy assessment of selection progress from round to round.[10]

2.1.4 NECEEM

NECEEM is the most widely used KCE method for the highly-efficient selection of aptamers to unmodified protein targets.[10,201-206] NECEEM partitioning in SELEX involves preparing an equilibrium mixture of a ssDNA library and a protein target, and injecting a small plug of this mixture into a capillary followed by separation of the target, ligand and complex by applying a current.[10] In most cases, the DNA is fluorescently tagged and LIF detection is utilized to visualize the species. The resulting electropherogram produces three regions: A DNA peak (**A1,Figure 24**), a DNA-protein complex peak (**A2, Figure 24)** and DNA dissociated from the complex (**A3, Figure 24**). By integrating these peaks, kinetic constants can be derived (**Figure 24**). The efficiency of partitioning by NECEEM was found to be at least two orders of magnitude higher than those of conventional surface-based selection techniques.[201] As a result, library enrichment can be completed in as few as 2-4 rounds of NECEEM-based SELEX, in contrast to the 10-30 rounds required in SELEX based on heterogeneous partitioning.[196-198,201]

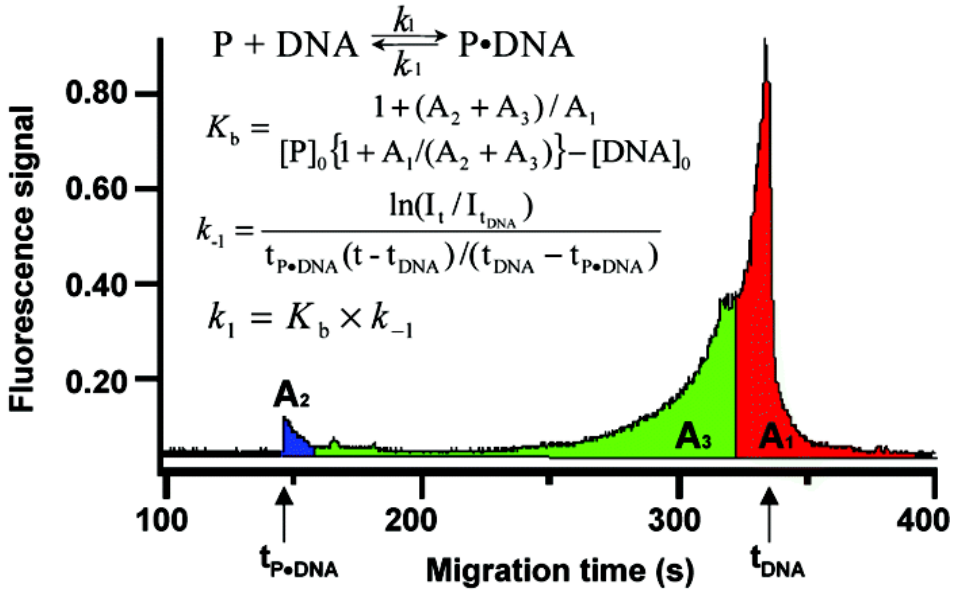


Figure 24: A NECEEM electropherogram, where DNA is fluorescently tagged, shows three regions corresponding to the DNA peak (red), complex peak (blue) and DNA dissociation region (green). Kinetic constants can be derived by integrating these peaks. Adapted from [10].

2.1.4.1 Issues

As PCR follows partitioning (**Figure 23A**), inefficient amplification can drastically affect the overall performance of SELEX. Despite this, relatively little attention has been paid to improving PCR in SELEX. Commonly, a classical PCR amplification protocol is used during SELEX, one that is optimized for homogeneous DNA templates. Such conventional PCR amplification, when used with heterogeneous DNA pools, such as the one used in SELEX, was found to quickly degrade the products, rapidly form byproducts through the formation of product heterodimers, and is susceptible to amplification bias.[207,208] Additionally, DNA amplification by conventional PCR is susceptible to nonspecific primer hybridization, primer dimers, and potential aptamer loss due to their adhesion to walls of PCR tubes.[209]

PCR bias when amplifying aptamers is a potentially significant issue that is often overlooked. This bias arises from the tight secondary structures of aptamers as compared to the non-binding oligonucleotides or weakly binding aptamers. Tight secondary structure has been

previously shown to reduce amplification efficiency, either by interfering with primer binding or preventing polymerase extension.[210-214]

PCR issues are particularly detrimental to NECEEM-based aptamer selection. An important intrinsic property of NECEEM is that the number of collected aptamers is very small (due to the small injection volume and stringency) which creates a significant challenge for the amplification step. To successfully amplify a small number of aptamers, the PCR step must be highly-efficient and specific, *i.e.* a high yield of a single amplicon product has be generated per corresponding intended template sequence, as opposed to byproducts generated due to non-specific primer binding and amplification of template heterodimers or preferential amplification of non-binders. These issues are most significant during the initial cycles of SELEX where aptamers may appear in single numbers. Inefficient amplification can therefore counteract all the benefits of NECEEM and is detrimental to any partitioning method, homogeneous or otherwise. For successful SELEX, these PCR issues must be addressed.

2.1.4.2 Solutions

A number of approaches have been attempted over the years to mend the aforementioned inefficiencies of PCR amplification in SELEX. To this end, quantitative real-time PCR (qRT-PCR) was used to determine the optimal number of PCR cycles by gauging the cycle at which 50% of maximal yield is reached (usually 11-13 cycles), thus reducing byproduct formation by preventing over-amplification. Once the cycle number was determined, symmetrical PCR (same concentrations of forward and reverse primers) was combined with asymmetric PCR (excess of forward primer) to decrease byproduct formation and increase product yield.[215] Although effective in its goal to reduce byproduct formation, this method does not address amplification bias, byproduct formation due to primer-dimers, nonspecific primer binding in early amplification cycles, or DNA loss due to adsorption. More recently, Nested-PCR, a method which uses a series of two or more sets of flanking primers to amplify an increasingly shorter sequence in a master template, was introduced to increase amplification specificity.[216] This approach, however, would exacerbate the bias and product-heterodimer-induced byproducts due to the increased size of the intermediate amplicon and the increased number of required PCR reactions. For instance, a 64-bp sequence was required to amplify a 16-bp random aptamer

sequence using six separate PCR reactions per selection round.[216] The bias issue was tackled specifically in SELEX-T.[217,218] This method first generates cDNA from RNA aptamers, followed by T7 RNA polymerase transcription to linearly amplify the RNA aptamers. Using excessive primers and low number of PCR cycles, bias due to heteroduplex formation is prevented by amplifying the cDNA template solely and avoiding the RNA product. This approach, however, only applies to RNA aptamers and introduces additional purification and bias-prone enzymatic reaction to the selection process.

2.1.5 Emulsion PCR

Here we suggest a possible solution to the PCR issue by combining NECEEM partitioning with a highly-efficient single molecule amplification procedure termed Emulsion PCR (ePCR). ePCR is a method to compartmentalize and miniaturize the PCR reaction *in-vitro*.[219,220] Micro-reactors enclosing PCR components and (ideally) a single template are created by vigorously mixing mineral oil and surfactants with a conventional PCR mix (**Figure 25**).[219] In effect, a local homogeneous-like PCR amplification environment is created.[221] These conditions prevent the formation of the highly detrimental product-product hybrids (the main source of byproducts).[207] Additionally, an increased effective local concentration of the template DNA is created in each micelle thus reducing primer-dimer formation and non-specific primer hybridization (the major cause of PCR inefficiencies in early amplification cycles).[219] Importantly, the presence of a single template in each micelle reduces amplification bias, a theoretically important factor which hinders the success of aptamer selection.[219,220] Furthermore, the hydrophobic properties of the oil does not allow for the DNA template to adsorb to the reactor tube wall,[209] a significant issue in alternative single molecule amplification methods.[219]

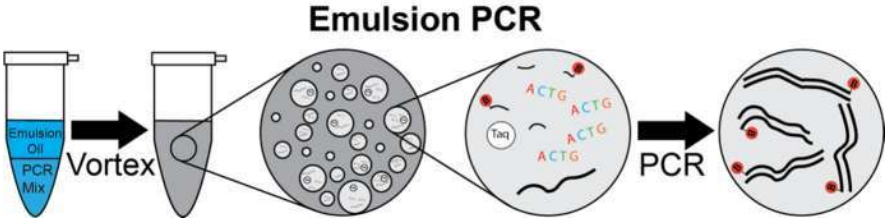


Figure 25: Emulsion PCR involves vigorously mixing a conventional PCR mix (with the addition of BSA) with emulsion oil. The resulting emulsion is composed of micelles which contain (ideally) a single template and all necessary PCR components, thus creating a homogeneous-like amplification conditions.

The efficient amplification of aptamers through their isolation in micelles, as provided by ePCR, could potentially significantly reduce aptamer loss due to stresses introduces by NECEEM, allowing it to function as an efficient homogenous partitioning method in the context of aptamer selection.

The utility of ePCR was previously suggested for the amplification of aptamers.[220] A recent semi-automated aptamer selection method, dubbed "just in time-selection", uses magnetic beads and emulsion PCR to amplify aptamers.[222] The method, however, relies on a heterogeneous separation approach and requires elaborate washing and negative selection steps. This results in the need for 15 rounds of SELEX. In addition, this approach requires a separate method (e.g. surface plasmon resonance) for affinity measurements after aptamer synthesis. To our knowledge, ePCR has never been used in conjunction with homogeneous separation techniques to select aptamers to difficult protein targets.

2.1.6 ABH2 as a Target

To test our approach, we selected aptamers to hABH2, a DNA damage repair enzyme. This human protein is a challenging target for aptamer selection as it is relatively unstable and, therefore, requires very careful handling and sensitive partitioning conditions. Additionally, hABH2 is a difficult target for CE as the protein is positively charged at neutral pH and, thus, can potentially adhere to negatively charged capillary walls during separation. This adsorption can further reduce the amount of collected aptamers and, therefore, applies additional stress on PCR.

Normally, hABH2 acts as one of the main alkylation housekeeping enzymes, repairing alkylation that can result in the disruption of replication and transcription, triggering of the cell-cycle checkpoints, induction of apoptosis or, in some cases, cancerous growth.[27] Such cytotoxic alkylation has been taken advantage of by chemotherapeutics to disrupt the replication of cancer cells. Many cancers, however, develop a resistance to chemotherapeutic alkylating agents by taking advantage of the cell's innate natural DNA damage defenses, such as hABH2 and related enzymes. Using these DNA repair enzymes, cancer cells rapidly dealkylate the DNA, thus requiring higher doses of the drug to inhibit their growth and, as a result, increase the chance for

lethal side effects, or rendering the drug ineffective altogether.[223-225] hABH2 expression has been directly implicated in a variety of cancers as well as neurological and developmental disorders.[116,226] Aptamers able to specifically bind to hABH2 can potentially be used to detect tumors, deliver drugs to the therapeutic target (thus lowering the chance for potentially lethal side effect), or act as drugs through binding hABH2 and inhibiting its enzymatic activity.[1]

A number of attempts to select aptamers for the hABH2 protein by NECEEM partitioning with conventional PCR (combination of symmetric and asymmetric PCR) were unsuccessful; they showed no detectable library enrichment in 5 rounds of SELEX. The very first attempt of aptamer selection using the combination of NECEEM and ePCR was successful, thus suggesting that ePCR is a viable approach to improve NECEEM-based aptamer selection for challenging protein targets.

2.2 Results

2.2.1 Unsuccessful aptamers selection for hABH2 by NECEEM with conventional PCR

Aptamers able to inhibit hABH2 enzymatic activity can be potentially used to improve the effect of chemotherapies. Maintaining the hABH2 protein unmodified during the selection procedure should help in finding such inhibitors binding to the protein in its native structure *in-vivo*. Thus, NECEEM, a homogeneous partitioning approach, is desirable. In a previous study, aptamers for a bacterial homologue of hABH2, AlkB, were selected using a SELEX protocol based on NECEEM partitioning and amplification using a combination of symmetric and asymmetric PCR (sPCR and aPCR).[193] AlkB aptamers were able to inhibit DNA demethylation by AlkB 1000 times more efficiently than the best reported small-molecule inhibitors.[127] This success was the major motivation for the present attempt to select aptamers for human hABH2 using the same selection approach.

The hABH2 aptamer selection was performed using an 80-nt long ssDNA library containing 20-nt long primer regions and 40-nt random-sequence region (the library was identical to the one used in aptamer selection for AlkB). DNA has the advantage of being more stable and easier to processes than RNA as it does not require a transcription step. Additionally, DNA aptamers, in general, were found to have similar binding affinities to their RNA

counterparts.[227] Three attempts with five rounds of SELEX in each were conducted with monitoring library enrichment after every round. No detectable enrichment was observed indicating that SELEX was failing despite the known high efficiency of NECEEM partitioning and the previous success with AlkB. Analysis of hABH2 crystal structures reveals that it is a larger, more complex protein than its bacterial homologue, and thus, it may behave differently in selection.[228] hABH2 contains a short loop of positively charged arginine and lysine residues, and an additional relatively long, flexible dsDNA-binding loop containing arginine, glycine, and lysine residues.[6] The positive charge imposed by these residues likely increases the propensity of hABH2 to interact with the negatively charged capillary walls, thus, potentially retarding the aptamer-ABH2 complexes. Additionally, the activity of hABH2 has been found to decrease within hours of thawing. As such, the protein may be partially degraded by the time it is used for selection, thus further decreasing the amount of the aptamer-protein complex. All these factors could contribute to the potential reduction of the number of DNA templates bound to the protein target during NECEEM, thus challenging the efficiency of PCR and resulting in the observed failure of SELEX. This failure motivated us to study these issues in detail and search for a solution that could rescue the failing SELEX for hABH2 as well as other similar proteins.

2.2.2 Options for improving efficiency of PCR

NECEEM, although an advantageous partitioning technique, comes at a cost by imposing challenges on the amplification portion of the selection process. First, whereas conventional partitioning techniques use the entire aptamer pool, NECEEM uses nanoliter volumes for analysis and separation. Although this has the benefit of conserving reagents and costly protein targets, it also means that during the first round of selection only a small fraction of the available sequences in the initial library are used in partitioning. In subsequent rounds, if the desired aptamers are not present in sufficient quantity (e.g. due to biased amplification), they may not be present in the following partitioning rounds due to the small injection volume. Additionally, the unmatched stringency of NECEEM, although enabling isolation of high affinity aptamers, greatly reduces the number of aptamers which are present in the final fractioned pool. NECEEM therefore requires highly efficient amplification that will preserve, and sufficiently amplify, the relatively small number of aptamers between the partitioning rounds.

The major problem of PCR amplification is rapid conversion of products into byproducts and potential bias towards the amplification of non-binders. Many methods attempting to solve this issue only address one of the problems, while ignoring or, in some cases, aggravating the other. As mentioned previously, qRT-PCR-controlled amplification reduces over-amplification of byproducts, but does not address the primer dimerisation or bias issue.[215] Nested-PCR increases amplification specificity,[216] but worsens the bias due to the increased size of the intermediate amplicon. Other methods, such as SELEX-T,[217] reduce bias and byproduct formation but only apply to RNA aptamers. Emulsion PCR, which isolates PCR reactions in microdroplets can help alleviate both product conversion to byproducts and PCR bias. For these reasons, emulsion PCR is commonly used in next-generation sequencing, including the 454, the Polonator, and SOLiD platforms.[229]

2.2.3 Emulsion versus Conventional PCR

As compared to ePCR, conventional symmetric amplification of a heterogeneous ssDNA library was found to produce a significant number of byproducts by the 10th cycle of amplification and, more importantly, the rate of product loss due to the conversion to byproducts surpassed product generation by as early as the 15th cycle (**Figure 26A, C-D**). By the 30th cycle, practically all the products have been converted to byproducts. These byproducts have been previously found to arise from formation of primer dimers and non-specific primer binding in early PCR cycles when primer concentration is high, and, more significantly, from the formation of product-product hybrids later in the PCR process when the template concentration is high.[207] In contrast, ePCR produced virtually no byproducts in comparison to sPCR and, as a result, no clear product loss was observed (**Figure 26B, C**). Conventional PCR product yield peaked at cycle 15 whereas an exponential increase in product was observed in ePCR well into the completion of the procedure. By the 40th cycle, over 10 times greater amount of products and two orders of magnitude lower amount of byproducts have been observed in ePCR as compared to conventional symmetric PCR (**Figure 26C**). This is consistent with a previous study that

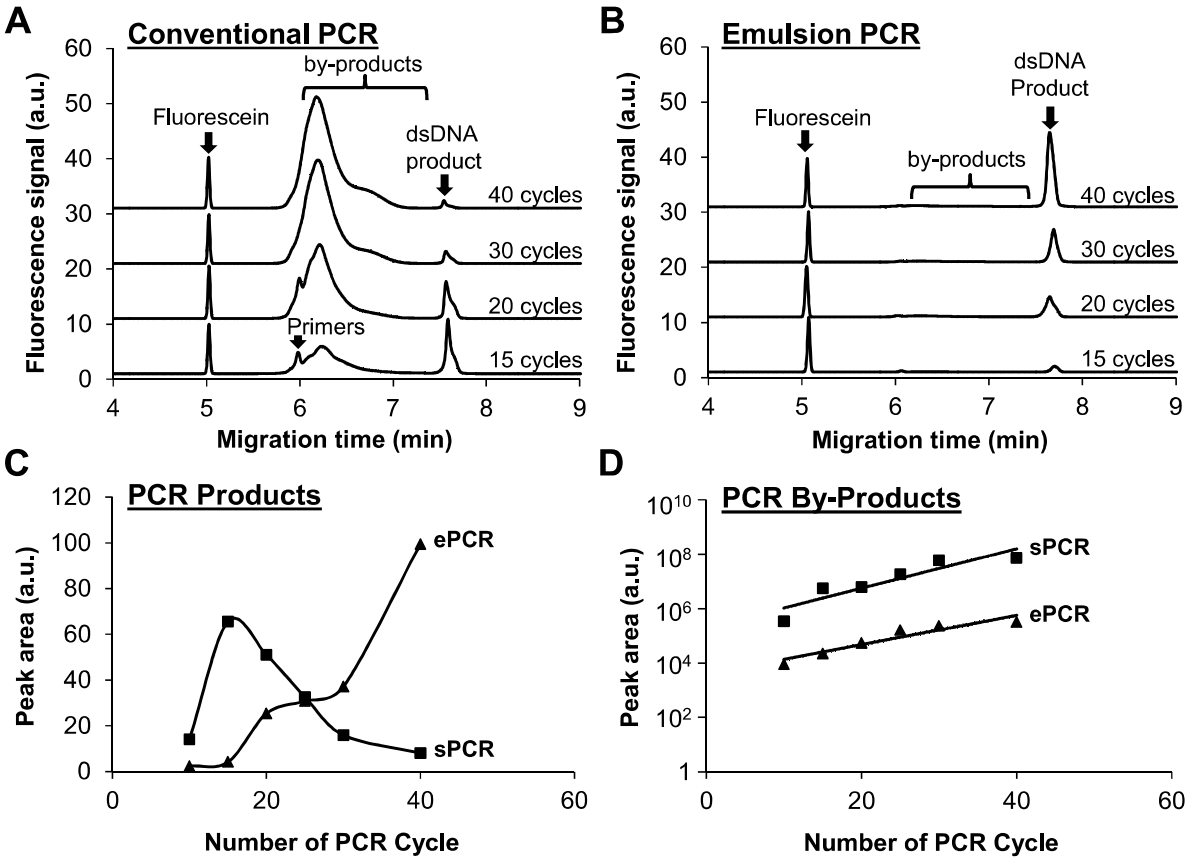


Figure 26: Comparison of product loss due to byproduct formation in conventional PCR and emulsion PCR.

found that conventional PCR produced byproducts as early as the 15th cycle and significant product loss by the 20th round.[207]

PCR product loss, such as that occurring due to their conversion to byproducts, is highly detrimental to the aptamer selection process, particularly when using NECEEM for partitioning. During the first round of selection, high affinity aptamers are present in single copies in the partitioned aptamer pool. The stringency and low volumes used by NECEEM reduce the quantity of aptamers further. Any loss of these sequences permanently removes them from the pool and may introduce a significant bottleneck to the aptamer selection process.

ePCR has been found to follow a Poisson distribution of templates with most micelles containing a single template. A small fraction of micro-reactors may have a larger number of templates or no templates at all.[219] In these cases, primer dimers, non-specific primer binding,

and product-product hybrids may form. Nevertheless, ePCR prevents any such byproducts, or other contaminations, from propagating throughout the PCR reaction thus limiting the amplification of byproducts to negligible amounts. In contrast, byproducts formed during conventional PCR are able to propagate throughout the entire process, increase in numbers exponentially, and, in addition to converting products to byproducts, limit the resources available for aptamer amplification. This allows the byproducts to outnumber the aptamers rapidly, significantly decreasing the chance of aptamers to be injected into the capillary in the following NECEEM partitioning round.

The remaining aptamers, which were not converted to byproducts in conventional PCR, are subjected to amplification bias. The variety of aptamer sequences present have different amplification efficiencies owing to factors including steric hindrance imposed by the secondary structure and polymerase sequence preference.[208,213,214,230] Non-binding DNA (possessing weak secondary structure) and aptamer sequences that are amplified more efficiently are overrepresented in the final pool as opposed to the desired high-affinity aptamers. This bias is potentially bypassed in ePCR as the polymerase is exposed to a single template and the template concentration is effectively increased in their respective micelles, thus eliminating competition. To summarize, our results showed that ePCR leads to a greater amount of desired products and 100-times lower amount of byproducts formed.

2.2.4 Aptamer selection for hABH2 based on NECEEM-ePCR

The experimental evidence showing significantly reduced product loss (see the previous section) prompted us to test the combination of NECEEM and ePCR as applied to hABH2 as a challenging target for aptamer selection. The selection process proceeded as follows: first, the library is allowed to equilibrate with the target hABH2 protein and the mixture is injected into a capillary (time t_1 in **Figure 27**). High affinity oligonucleotides (aptamers) bound to the target are partitioned from the unbound or weakly bound and dissociated sequences using an electric field of 400 V/cm (time t_2 in **Figure 27**). This causes the mixture components to separate into distinct regions, or zones, determined by charge to friction-coefficient ratios of the analytes.[231] Fluorescent tagging of the DNA enables visualization of the unbound DNA peak as well as the complex of DNA with hABH2 and the so-called dissociation region (**Figure 24**). A fraction-

collection window is determined to include the complex as well as a portion of the DNA pool that dissociated from the protein-DNA complex (**Figure 27**). Essentially, the closer the right boundary of the window to the ligand peak, the weaker ligands (higher rate constants of complex dissociation, k_{off}) will be collected.[199] The boundary is generally chosen towards the middle of the dissociation region with a bias towards the complex peak. This allows for sufficient distance as to not contaminate the high affinity aptamers with non-binders or low affinity aptamers and ensures that the entire complex peak is collected. This classic NECEEM experiment allows for the estimation of EC_{50} (analog of K_d used for a pool of different aptamers, EC_{50} is defined as the concentration of target at which 50% of aptamers in the pool are bound) and k_{off} constants during the fraction-collection stage for future reference and monitoring of the selection progression.[206]

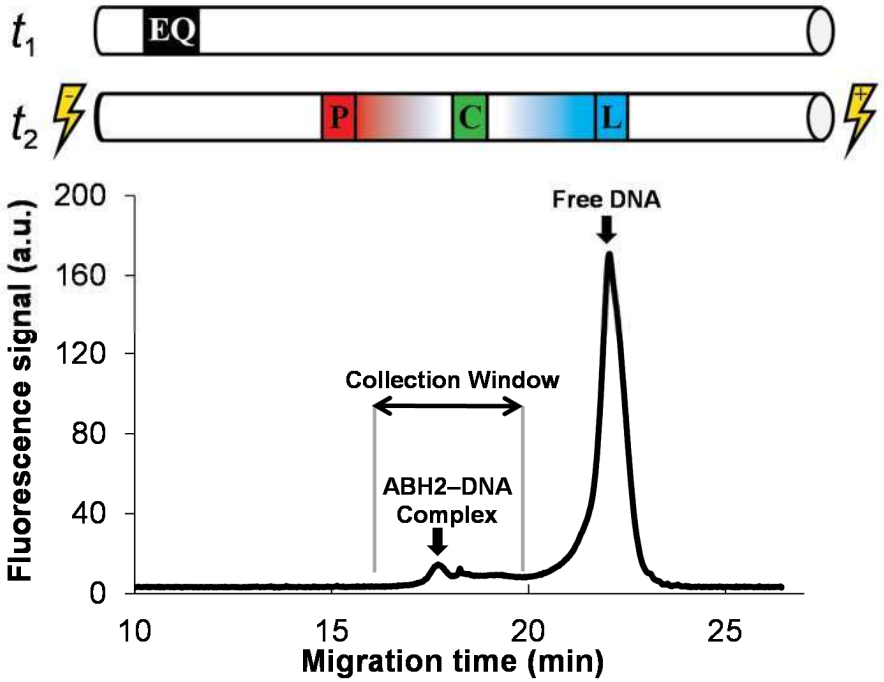


Figure 27: Determination of aptamer-collection window in NECEEM for ABB2 as a target. In NECEEM, an equilibrium mixture containing the protein and DNA aptamer library is injected (t_1) into the capillary and an electric field is applied to separate the complex from the protein and ligand (t_2). The electropherogram contains peaks of the complex and ligand due to the fluorescent tag on the DNA. A window is chosen to capture the intact complex and a part of the dissociation region.

Once obtained, the enriched DNA aptamer pool is amplified using ePCR followed by the generation of an enriched single stranded pool using asymmetric PCR,[232] completing the selection round. The next round is then initiated by incubating the enriched pool with hABH2 and subjecting the mixture to NECEEM partitioning. This classical NECEEM experiment allows

for the determination of the binding constants of the new enriched pool of DNA and for the selection of a new fraction at the same time. The rounds of alternating partitioning and amplification rounds are then repeated until no further enrichment is observed.

After amplification, the ePCR emulsion can be easily broken using organic solvents. Originally, the highly volatile and flammable organic solvent diethyl ether was used followed by a precipitation of the amplified DNA. Here, we follow Schütze's work to simplify the process and replace diethyl ether with isobutanol. We also replace the emulsification method with vortexing as opposed to the original contamination prone magnetic stirring method.[221] To regenerate the ssDNA library, aPCR was utilized at a 20:1 forward to biotinylated reverse primer ratio.[232] The residual dsDNA remaining after aPCR amplification was removed using streptavidin magnetic beads and biotinylated reverse primers. The resulting amplified DNA was purified from other components of the solution (primers, nucleotides, and enzymes) by using the MinElute purification method (QIAGEN, Mississauga, ON, Canada). This combination of ePCR and aPCR is a very simple, rapid, and efficient way to regenerate an enriched aptamer pool.

The affinity of the enriched pools, as assessed through EC_{50} values was measured using NECEEM immediately following aPCR purification and a new fractions was collected at the same time.[206] Obtaining binding constants prior to sequencing is of substantial benefit for aptamer selection. By following the progression of enrichment, one can halt selection as soon as no further enrichment is observed so as to avoid extra selection steps and labour intensive affinity measurements as required in other selection methods. An EC_{50} value of 600 nM was reached by the fourth round of SELEX (**Figure 28**). The fifth and sixth selection rounds showed no further improvement and the selection procedure was stopped. A common trend for EC_{50} enrichment, particularly in CE selection techniques, is for affinity to reach a maximum, followed by a worsening in EC_{50} following each successive round.[233,234] Although the reason for this phenomenon is not well understood, we believe that it occurs, at least partly, due to the limitations imposed by PCR. It is likely that at the point of loss of enrichment, the rate of PCR-related aptamer loss surpasses the rate of enrichment. Here, we did not observe this pattern. Although enrichment was not observed past the fourth round, no loss in enrichment was

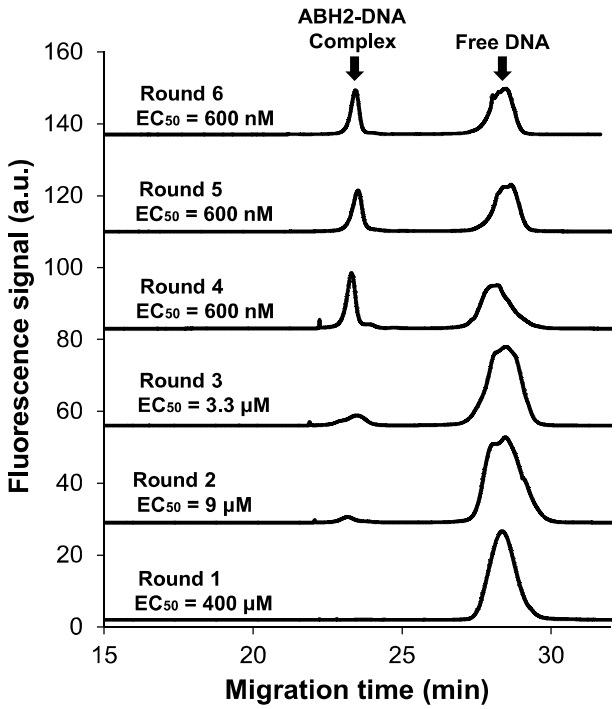


Figure 28: Progress of library enrichment with number of SELEX rounds. NECEEM-based aptamer selection combined with ePCR was used to select aptamers for ABH2 from an unbiased ssDNA library.

observed. It is likely that this occurs due to the improvements provided by ePCR which curbs PCR related aptamer loss.

2.2.5 Aptamer Synthesis and Binding

Pools displaying the lowest EC_{50} values were cloned and eight aptamers with the lowest K_d values were sequenced and synthesized. Although 2OG was not present during the selection process, binding measurements of the synthetic C2 aptamer without 2OG had a K_d of 1.8 µM whereas the addition of 2OG increased the strength of binding almost 10 folds, decreasing the K_d value to 190 nM (**Figure 29**).

This is explained by the ordered sequential mechanism of Fe-2OG-dependent dioxygenases, including hABH2, in which 2OG binding precedes that of substrate; 2OG coordinates to the active site Fe^{2+} and which induces substantial structural changes in hABH2 that reorganize the active site for substrate binding.[235,236] Selecting aptamers in the presence of co-factors will be the subject of future work. The synthetic aptamers obtained from the fourth

round of selection were found to be characterized by K_d values ranging from 38 nM to 1 μM (**Figure 30, Table 3**).

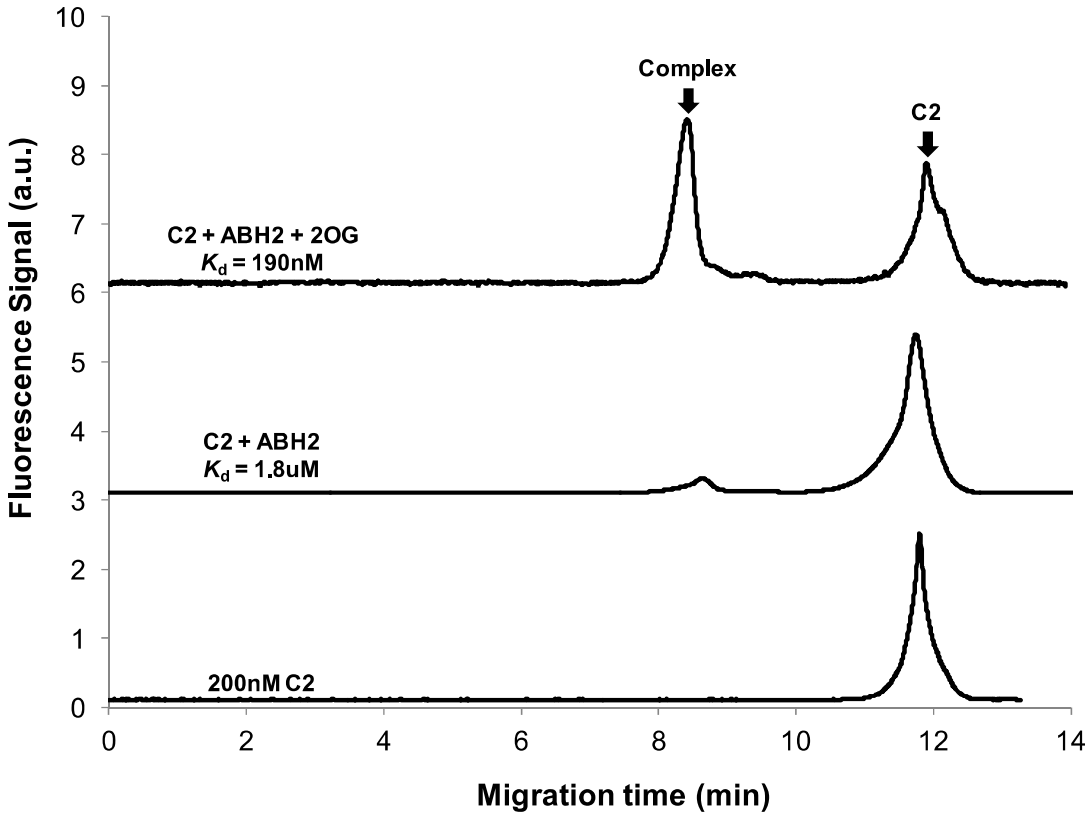


Figure 29: ABH2 Aptamer binding improves with the addition of OG. 100nM of synthetic aptamer was incubated with 400nM ABH2. Low binding was observed when the selecting conditions were replicated. Addition of 2OG restored the binding.

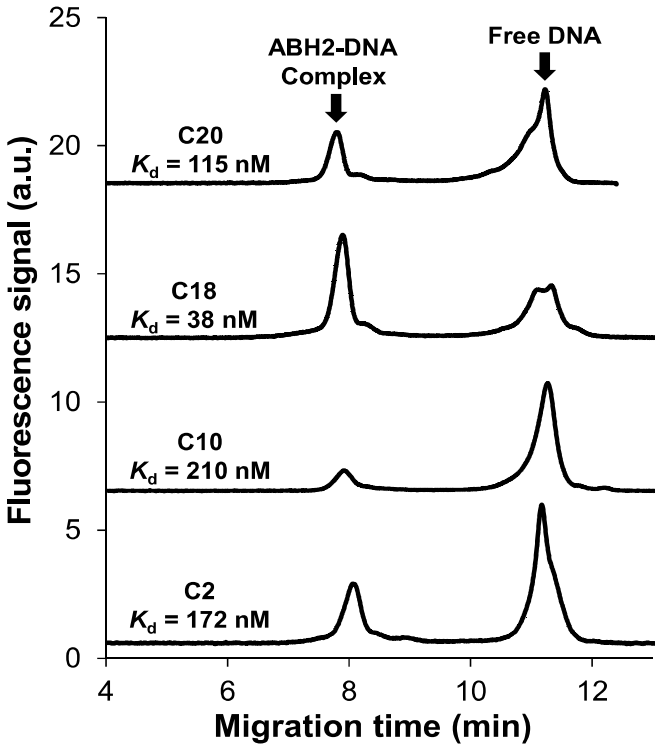


Figure 30: Aptamer binding studied by NECEEM. Two hundred nM aptamer was mixed with 200 nM of ABH2 in the presence of 160 μM 2OG.

Table 3: Binding constants and sequences of synthesized hABH2 aptamers.

ID	K_d (nM)*	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})	Core Sequence (5'→3')
1	215	1.0E+04	2.2E-03	AAGCTATTACCCGCGTTCATAGCTTTGCGTCAGAGCCTCA
2	154	3.3E+03	5.1E-04	CTGGAAGGGTTCAGAGGTAAGACTACGATCTTGGCGGGCC
10	178	1.3E+03	2.2E-04	GGTTCGGAGGAATCCATGTTGCGAACGACCGCTGCTAGCA
13	238	8.0E+03	1.9E-03	AAGTGATGGGGCCAGGTTCAGAATGAGGACATACGGGAGG
15	391	4.7E+03	1.8E-03	ATACCAGGGCTCAGAGGATCTGGGGGCCGATGCGTTGGAT
17	631	1.5E+03	9.3E-04	AGCGGGGTTGCGGAGTTCTGAGGTCGGCCACTCATGGTGC
18	35	7.0E+03	2.4E-04	GCATTACTGGTTCTGAGGAAAGTGCGACTCAGCCGTACGC
20	86	1.1E+04	9.3E-04	TGACCGGTTCAGATGGATGGGTCCTTCGCATAGACCTACA

Primers: forward, CTCCTCTGACTGTAACCACG; reverse: GCATAGGTAGTCCAGAAGCC

*Experiments were done in duplicates and the results of a single set of experiments are shown.

2.2.6 Cross Reactivity

The identified hABH2 aptamers were tested for cross-reactivity with the homologous Fe-2OG-dependent dioxygenases, AlkB and hABH3, as well as BSA and the MutS mismatch repair protein. The aptamers were found to successfully discriminate between the human and bacterial homologues as well as MutS and BSA (**Figure 31**). The aptamers' specificity to hABH2 suggest that they interact with the protein's unique DNA binding site, instead of other non-specific regions. The binding site of hABH2 is different from its homologues as it involves a finger residue and a positively charged loop involved in double stranded DNA binding that is not present in ABH3 or AlkB. To determine the specific binding region requires further mutagenic analysis.

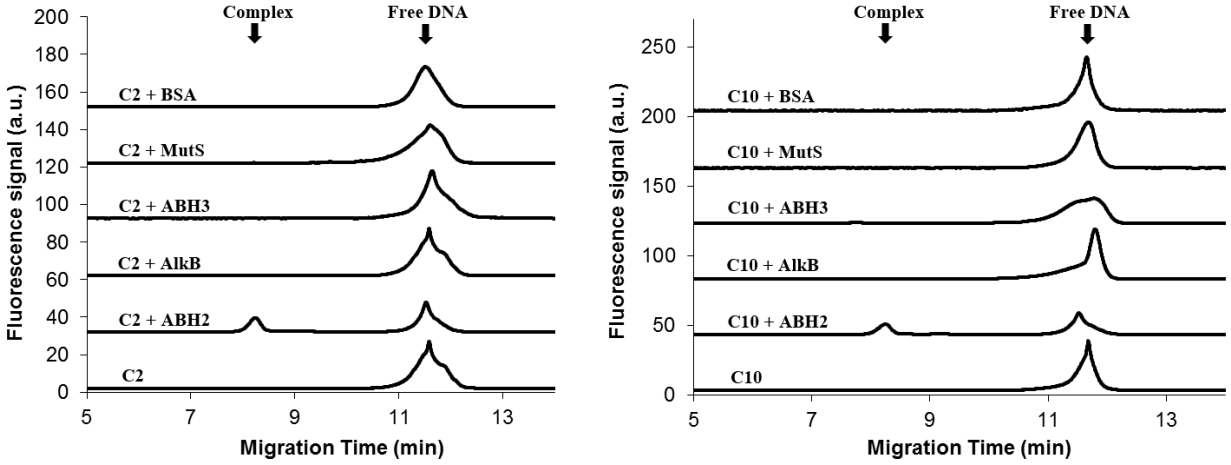


Figure 31: Cross reactivity of the C2 and C10 aptamers with related and unrelated proteins. The synthetic ABH2 C2 and C10 aptamers do not significantly bind to any of the proteins tested. Dissociation is observed with MutS and to a lesser extent with the rest of the proteins. 100nM aptamer was incubated with 400nM protein and incubated for 15 minutes prior to NECEEM analysis.

2.3 Conclusion

In conclusion, the procedure for efficient and rapid aptamer selection by NECEEM has been significantly improved by the addition of ePCR (**Figure 32**). By creating a homogeneous-like amplification environment using ePCR, aptamers with K_d values ranging from 38 nM to 1

μM were selected in as few as four rounds whereas the same method using conventional PCR failed to show any detectable enrichment.

The work leading to the selection of ABH2 aptamers reported here will serve as the basis in the development of efficient hABH2 and other DNA repair enzyme inhibitors for their potential use in combination with current alkylating drugs.[45,46] This approach could greatly reduce the chemotherapy resistance developed by many neoplastic cells. Such aptamers also have the potential to be used as delivery vehicles for cytotoxic drugs to cancer cells.[237] Due to their high affinity and relatively small size, aptamers exhibit durable target tissue retention as well as rapid blood clearance, essential features for their use in *in-vivo* imaging and drug delivery.[237] The newly developed method can be employed for the rapid discovery of aptamers with potential therapeutic use for other unstable proteins.

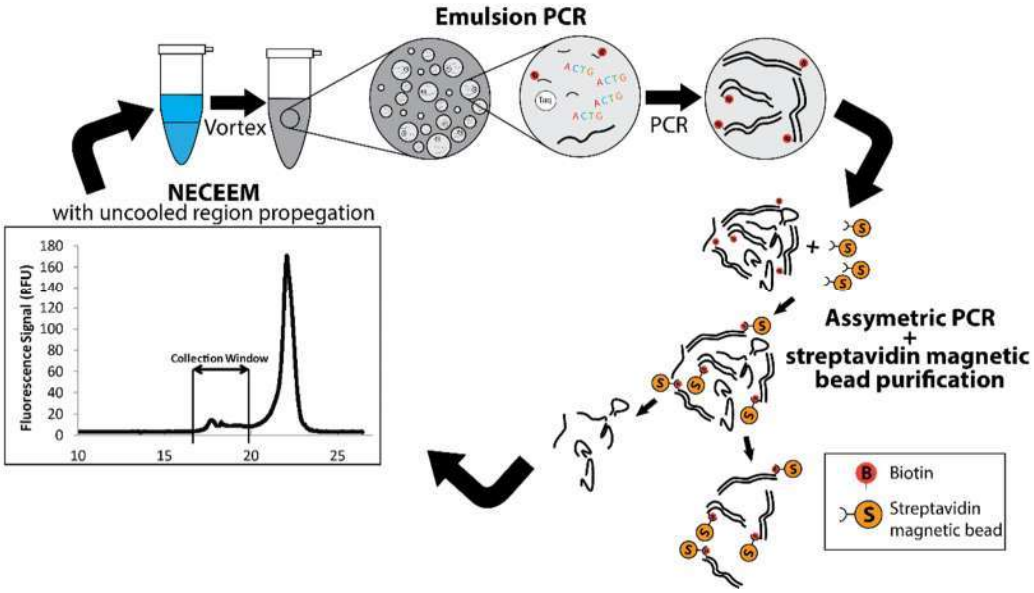


Figure 32: NECEEM-based Aptamer Selection with Emulsion PCR. In this version of the protocol Emulsion PCR is used to amplify an aptamer pool selected using NECEEM. Once amplified, single stranded DNA is generated using asymmetric PCR and purified using streptavidin magnetic beads. The purified pool is then incubated with the protein again and the process is repeated.

2.4 Materials and Methods

2.4.1 Materials

Fused silica capillary, 75 µm inner diameter, 365 µm outer diameter, was purchased from Polymicro Technologies (Phoenix, AZ). hABH2 was expressed and purified as described elsewhere,[169] and stored in 50 mM Tris-HCl, 500 mM NaCl, 1 mM DTT pH 7.5 buffer at −80°C. A ssDNA library, labeled at the 5' end with fluorescein (F488) for detection, 80 bp in length, including two 20-bp primer regions, was synthesized and HPLC-purified by Integrated DNA Technologies (IDT, Mississauga, ON, Canada) and stored at −20 °C. F488-labeled as well as unlabeled ssDNA forward primers, as well as 3' biotinylated and un-biotinylated reverse primers, 20 bp in length, were synthesized and purified with HPLC by IDT (Mississauga, ON, Canada) and stored at −20°C. Single stranded, 5' F488-labeled DNA aptamers (80 bp in length, containing 20-bp primer regions) were synthesized by IDT (Mississauga, ON, Canada). Streptavidin-labeled magnetic beads as well as all other chemical reagents and buffers were obtained from Sigma-Aldrich (Toronto, ON, Canada). Solutions were prepared using double distilled deionized water and filtered through a 0.22 µm filter (Millipore, Nepean, ON, Canada).

2.4.2 Instrumentation

CE experiments were performed using a P/ACE MDQ instrument from Beckman Coulter (Fullerton, CA, U.S.) utilizing LIF detection with excitation at 488 nm and emission at 520 nm. Uncoated fused silica capillaries were used with total lengths of 50 cm for K_d and EC_{50} measurements and 80 cm for aptamer partitioning; the distances to the detection window were 40 and 70 cm, respectively.

Brand new capillaries were preconditioned using a wash with 2 capillary volumes of methanol, followed by 5 capillary volumes of sequential washes with 100 mM HCl, 100 mM NaOH, deionized water, and, finally, equilibrated with the running buffer for 10 min. The sample (either the equilibrium mixture of aptamers and protein or aptamers alone) was introduced into the capillary by pressure at 1 psi for 28 s, and was propagated past the uncooled region of the capillary when the protein was present (see section 2.5). The capillary temperature was kept at

15°C. Electrophoresis was carried out at electric field of 400 V/cm. PCR amplification was performed with a Bio-Rad iCycler (Mississauga, ON, Canada).

2.4.3 Determination of the aptamer-collection window

To determine bulk affinity (EC_{50}) prior to the first selection round, a plug of 200 nM DNA library incubated with 1.6 μM hABH2 for 15 min was injected into a 80-cm-long capillary by a 10 s × 0.5 psi pressure pulse followed by buffer propagation by 0.3 psi pressure for 135 s. In the first selection round, a plug of 25 μM DNA library incubated with 500 nM hABH2 for 15 min was used for injection. In all subsequent rounds, 100 nM aptamer pool and 400 nM hABH2 were used. The components were separated by electrophoresis at 400 V/cm. Once the protein-aptamer complex was observed, an aptamer-collection window was determined to include the complex as well as a part of the region corresponding to aptamers dissociated from the complex during separation. Care was taken to maintain a 0.5-1 min gap between the collection window boundary and the beginning of the peak of unbound DNA. The following formula was used to calculate elution time ($t_{elution}$) of the selected fraction based on the observed time of travel to detector ($t_{detector}$), the total length of the capillary (L_{total}) and distance from the inlet to the detector ($L_{detector}$):

$$t_{\text{elution}} = t_{\text{detector}} \frac{L_{\text{total}}}{L_{\text{detecotor}}} \quad (2)$$

Aptamer collection was repeated 3 separate times per pool with slightly adjusted right boundaries. qRT-PCR was used to confirm a successful collection by comparing the collected fraction to a negative collection (without the presence of the protein target).

2.4.4 Aptamer selection

For the first round of selection, 25 μM ssDNA library was mixed with 500 nM hABH2 in 50 mM HEPES buffer at pH 7.5. The binding reaction was allowed for 15 min to approach equilibrium and the DNA-ABH2 complex was separated from the unbound DNA by using NECEEM in an electric field of 400 V/cm. 50 mM Tris-acetate pH 8.2 was used as a separation buffer. The complex was collected by using a method described in in section 2.3, and the amount

of DNA was quantified by qRT-PCR using iQ™ SYBR® Green Supermix Bio-Rad iCycler (Mississauga, ON, Canada).

2.4.5 Propagation past uncooled region

In order to move the sample past the uncooled region[238] of the capillary the pressure-injected plug was further pressure propagated 5 cm into the cooled region of the capillary. The required pressure pulses were 0.3 psi × 2.25 min for an 80-cm long capillary and 0.3 psi × 1 min for a 50-cm long capillary.

2.4.6 Emulsion PCR

Emulsion PCR oil was prepared by combining three emulsion oil components provided by the Micellula DNA Emulsion and Purification Kit (EURx, Poland) in a 73:7:20 ratio as recommended by the manufacturer. A conventional 50 µL PCR mixture with the addition of BSA was added to 300 µL of the oil phase and contained the following: 10× PCR buffer (New England BioLabs, Toronto, ON, Canada) , 400 nM forward and reverse primers, 200 nM dNTP mix (each, IDT, Mississauga, Canada), 5U/µL Taq (New England BioLabs, Toronto, ON, Canada), 0.75 mg/mL BSA (New England BioLabs, Toronto, ON, Canada) , nuclease free H_2O (IDT, Mississagua, Canada) as well as 10-20 pM of the collected aptamers. Once combined, an emulsion was formed by vortexing the mixture at a maximum speed for 5 min on ice. After amplification (15 cycles), the contents of PCR tubes were pooled and the emulsion was broken through the addition of 1 mL of isobutanol. The PCR mixture was then purified using the Micellula DNA Emulsion and Purification Kit (EURx, Poland) as per manufacturer's instructions. The purified amplified PCR reaction was eluted with 20 µL of provided elution buffer (EURx, Poland) incubated for 2 min followed by a 1-min centrifugation at 1.3×10^4 rpm into a collection tube.

2.4.7 Regeneration of enriched pool

The purified emulsion PCR product was amplified asymmetrically in triplicates with a 20:1 ratio of F488-labeled forward primer to biotinylated reverse primer to generate ssDNA product. Twenty µL of streptavidin magnetic beads (Sigma Aldrich, Oakville, ON, Canada) was

magnetized and re-suspended in 10 mM Tris-HCl buffer containing 10 mM NaCl and 1 mM EDTA at pH 8.0. Once amplified, the triplicate PCR reactions were combined and incubated with the streptavidin magnetic beads for 30 min at room temperature. The beads were magnetized, discarded and the PCR product was purified using a MinElute purification kit (QIAGEN, Mississauga, ON, Canada) as per manufacturer's recommendation. DNA was then eluted using 10 µL of 50 mM HEPES pH 7.5.

2.4.8 Affinity analysis

K_d and k_{off} values were determined from NECEEM electropherograms as described elsewhere;[10,239] k_{on} values were calculated as k_{off}/K_d. The NECEEM experiments were performed as follows: A 50-nL plug of the equilibrium mixture containing 100 nM aptamer and 400 nM protein in 50 mM Tris-HCl pH 7.5 was injected into a 50-cm-long capillary by a 0.5 psi × 10 s pressure pulse. An electric field of 400 V/cm was applied to the capillary with a positive electrode at the inlet. The electrophoresis run buffer was 50 mM Tris-acetate, pH 8.2. The sample was detected by LIF detection with excitation at 488 nm and emission at 520 nm.

2.4.9 Cloning

Aptamer pools were amplified using symmetric PCR and unlabeled primers. The formation of dsDNA was assessed using a 4% agarose gel and the appropriate band was extracted and purified using the QIAquick Gel Extraction Kit (QIAGEN, Mississauga, ON, Canada). The concentration of the DNA was quantified using qRT-PCR. Cloning was performed using the pT7 Blue-3 Perfectly Blunt cloning kit (VWR, Mississauga, ON, Canada) as follows. The eluted DNA was ensured to have blunt ends by using an End-Conversion Mix (provided with the cloning kit). The mix is then ligated into a Blunt vector using T4 DNA Ligase (VWR, Mississauga, ON, Canada). Finally, NovaBlue Singles Competent Cells (VWR, Mississauga, ON, Canada) were transfected with the ligation mixture prepared above. SOC medium (VWR, Mississauga, ON, Canada) was added to the cells and the mixture was plated on ampicillin agar plates. The cultures were allowed to grow in an incubator at 37^0C overnight. Twenty four colonies were picked and resuspended in 200 µL of H_2O. The resuspended colonies were then placed in a boiling water-bath for 5 min followed by centrifugation at 1.3×10^4 rpm for 10 min

to pellet the bacterial debris. The supernatant containing the DNA was then amplified in 30 cycles of symmetric PCR as described above using F488-labeled forward primers followed by 15 cycles of asymmetric PCR as described in section 2.7. The K_d value was measured for each clone prior to sequencing. Plasmids for sequencing were extracted using the GenElute Plasmid Miniprep Kit (Sigma-Aldrich, Toronto, ON, Canada). Once extracted, the plasmids were sent to TCAG DNA Sequencing facilities (The Hospital for Sick Children, Toronto, ON, Canada)

CHAPTER 3: Direct Measurement of the Demethylation Activity of hABH2 and hABH3

3.1 Introduction

3.1.1 DNA Alkylation and Repair

Alkylation is a ubiquitous form of DNA damage, resulting in cytotoxic, mutagenic, and carcinogenic effects.[12,27] The Fe-2OG-dependent dioxygenase family of repair enzymes catalyze the removal of N-alkyl groups from damaged DNA.[3,240] This cellular defense mechanism of dealkylation involves the hydroxylation of alkyl groups, followed by the release of succinate, formaldehyde and a dealkylated nucleic acid product.[1,241,242] Fe-2OG-dependent dioxygenase repair proteins were found to be highly conserved and ubiquitously expressed in normal human tissue, suggesting a critical role in DNA repair.[25] The enzymatic activity of human Fe-2OG-dependent dioxygenases has also been suggested to be linked to various types of cancers,[67,118,140,243] neurological diseases,[244] and obesity in humans.[81] 3-meC is among the most abundant and cytotoxic form of alkylation.[43,90] Such N-methylation adducts comprise more than 80% of methylated bases.[76] These lesions can block DNA replication and are commonly repaired via direct oxidative reversal, mediated by the Fe-2OG-dependent dioxygenases: AlkB in prokaryotes and ABH2 and 3 in humans.[4,12,14,18] hABH2 has been found to preferentially dealkylate dsDNA, whereas hABH3 prefers ssDNA as well as RNA.[4,12]

3.1.2 Chemotherapeutics

The cytotoxic nature of alkylation is taken advantage of by a specific class of anti-cancer drugs, termed alkylating antineoplastic agents.[34] These agents aim to damage DNA in tumors as to prevent the DNA strands from uncoiling and replicating, or trigger cellular checkpoint to induce cell cycle arrest, p53 response, and apoptosis, thus halting cancerous growth.[5] These alkylating drugs are susceptible to the cell's dealkylating repair systems, such as the aforementioned Fe-2OG-dependent dioxygenases hABH2 and hABH3.[76,245] Such enzymes were indeed found to reduce the success of alkylating chemotherapeutics.[35,70,246] Detection and inhibition of Fe-2OG-dependent dioxygenases, therefore, has a significant therapeutic potential.[122,130,131]

3.1.3 Alkylation Activity Studies

A detailed understanding of the dealkylation repair mechanism is essential for finding new drug candidates. Both the *in-vitro* study of the enzymatic activity of Fe-2OG-dependent dioxygenases and search for inhibitors require a simple and accurate method for a quantitative analysis of their enzymatic activity. Such a method should ultimately utilize biologically-relevant substrates, oligonucleotides rather than single nucleosides or other co-products. A classical approach for such an analysis is indirect; involving the detection of co-products formed, such as CO_2 and formaldehyde, or coupled reactions.[169] For example, NADH production is detected using absorption when the reaction is coupled with formaldehyde dehydrogenase. Such techniques are undesirable as they are prone to systematic errors.[7,14,148,247] A common assay for analyzing the substrate directly utilizes a methylated radioactively labeled substrate and HPLC analysis. After the enzymatic reaction, the DNA is precipitated and the radioactive methylated nitrogenous bases are released from the substrate using an enzymatic reaction or acid hydrolysis. The bases are then analyzed by HLPC and the amount of methyl groups remaining on the substrate is compared to the amount before the reaction.[12,168] Although superior to indirect methods, this approach requires error-prone post enzymatic modifications and radioactive labeling which introduces complexity to the reaction.

A direct enzymatic assay utilizing separation of a 15-nt single-site methylated ssDNA substrate from its demethylated product by capillary electrophoresis without the need of post reaction manipulations was recently introduced.[15] This method allows for quantitative analysis for both product formation as well as substrate consumption. It avoids the use of radioactive compounds, does not require a calibration curve or sample manipulation, and uses small (nanoliter) sample volumes. It is also rapid, cost-effective, and accurate. The proof of principle and demonstration of application to inhibitor screening was previously done with AlkB, a bacterial Fe-2OG-dependent dioxygenase, but was never applied to human proteins.[15,193] The method was used to demonstrate that DNA aptamers can inhibit the demethylation activity of AlkB more effectively than previously reported inhibitors.[127]

3.1.4 Objective

As discussed in chapter 2, aptamers specific to the human Fe-2OG-dependent dioxygenase hABH2 have been selected.[248] The goal of this work was to determine whether the direct CE enzyme activity assay could be utilized for the human Fe-2OG-dependent dioxygenases, hABH2 and hABH3. This work would allow for the future analysis of aptamers as inhibitors of human Fe-2OG-dependent dioxygenase DNA alkylation repair protein.

Here I demonstrate that the reported method is indeed directly transferrable to human enzymes, allowing for the fast and accurate measurement of hABH2 and hABH3 enzymatic activities. The technique can potentially be used as a universal approach for discovering and analyzing inhibitors for this family of enzymes, allowing for screening of large libraries of potential drug candidates.

3.2 Results and Discussion

To study the demethylation activity of hABH2 and hABH3, a synthetic substrate with a single methylated cytosine group labeled with fluorescein at the 5' end is used to directly monitor both substrate consumption and product formation. The enzyme assay is based on the CE separation of 3-meC in fluorescently labelled ssDNA substrate from the demethylated fluorescently labelled product. The speed, automation and sensitivity of CE allow for fast and accurate measurements of initial reaction rates at different substrate concentrations without any post enzymatic manipulation.

Michaelis-Menten plots were obtained for product formation in reactions catalyzed by 50 nM enzyme for varying concentrations of the substrate. The plots are shown in **Figure 33**. Using nonlinear regression, K_m values of 430 ± 27 nM and 1.0 ± 0.2 μM, and V_{max} values of 34 ± 1 and 103 ± 5 nM/min for hABH2 and hABH3, respectively, were determined (**Table 4**). These values imply that whereas hABH2 has a higher affinity for the substrate, hABH3 is able to achieve higher maximum velocities and, therefore, has a higher turnover (k_{cat}). This is consistent with previous published results in which hABH3 had a higher activity towards single stranded substrates.[4,12] A vast majority of previously published papers comparing the enzymatic activity of human hABH2 and hABH3 did not carry out kinetic analyses. For example, Duncan *et al.*

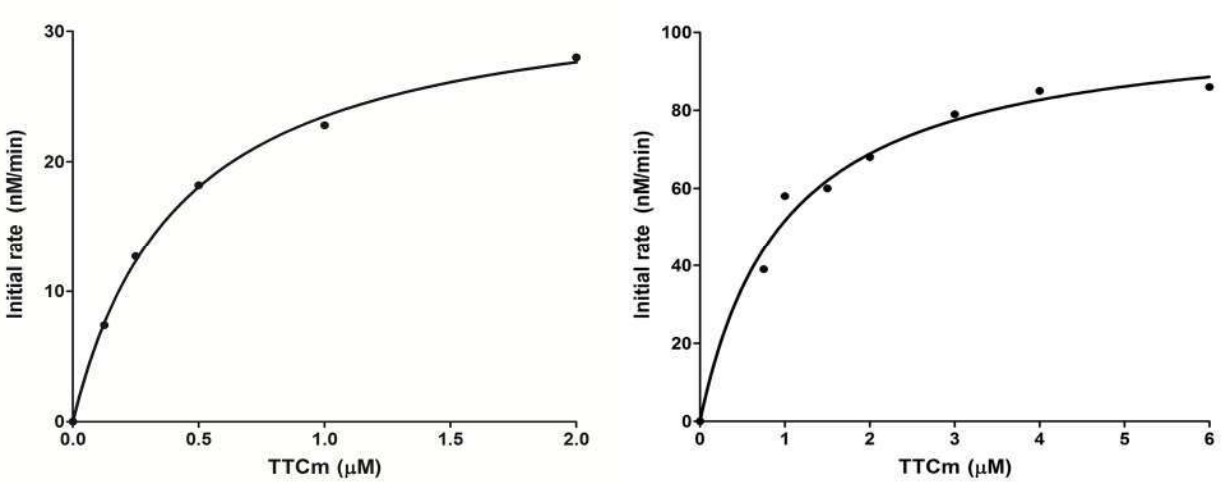


Figure 33: Michaelis Menten plots for demethylation of the TTCm substrate catalyzed by 50 nM hABH2 **(left)** and 50 nM hABH3 **(right)**. Initial reactions rates were determined though CE separation of the demethylated product from the substrate. Curve fitting was performed with GraphPad Prism 5 software. Kinetic constants (K_m and V_{max}) were determined for the best fitting.

mentioned that both GST- and His-tagged hABH2 and hABH3 are capable of demethylation; however, they did not carry out any kinetic measurements.[12] Lee *et al*., conversely, published K_m values for mouse hABH2 and hABH3 for ssDNA 3-meC substrate and report K_m values of 82 ± 22 nM and 162 ± 48 nM for mABH2 and mABH3, respectively (**Table 4**).[14] Although the K_m values reported by Lee *et al.* for the mouse enzymes were 5 fold lower than the values we obtained for the human enzymes, the difference in K_m between the enzymes is similar in both assays; hABH2 has a 2-fold lower K_m than hABH3 (**Table 4**). Similarly, Lee has found a comparable difference in the maximum velocity (V_{max}) and turnover (k_{cat}) between the enzymes. Whereas Lee found hABH3 to have a twofold higher V_{max} and k_{cat} than hABH2, we found a 3 fold difference (**Table 4**). The variance between the values becomes marginal when accounting for standard error. The difference between the results obtained here and the values reported by Lee *et al*. could be partially explained by the use of mouse protein as opposed to human as well as by differences in the assay conditions. Lee has used longer (24 nt), radioactively labelled substrates and performed the essay at 37°C for 30 min with lower concentrations of iron, AA, and BSA as well as higher pH and higher concentration of 2OG. In contract to our approach, Lee's method relied on phenol/chloroform extraction, ethanol precipitation and restriction enzyme cleavage of the product. Such post-processing steps, in addition to being time and resource consuming, could contribute to error and, therefore, cause the differences observed.

The results of the kinetic study of demethylation activity of human hABH2 and hABH3 enzymes were compared with constants obtained for the *E. coli* AlkB enzyme. As demonstrated in **Table 4,** AlkB has a similar turnover (k_{cat}) to hABH3 and a 4-fold higher turnover than hABH2. This is consistent with the known similarity in activity between AlkB and hABH3 towards single stranded substrates, and the preference of hABH2 towards dsDNA substrates.[4,12] AlkB also has a significantly higher efficiency (k_{cat}/K_M) than both hABH2 and hABH3 (30-50 folds) when a similar-size substrate was used.[15] This finding is consistent with the previously published results.[148] The significant variation between the constants obtained by different assays, as exemplified in **Table 4**, is indicative of the sensitivity of enzymatic assays to different assay conditions. Notably, different assays utilized substrates of different lengths, suggesting that not only do the enzymes have different activities towards different types of alkylation, but also towards different oligonucleotide lengths and sequence as previously shown.[21] Such variation could be a subject of future studies.

The consistency of this method with previous works confirms that this method can be applied to human demethylation enzymes. Ultimately, this approach can be used to determine the

Table 4: Kinetic constants for AlkB and AlkB mammalian homologues for demethylating 3-meC on ssDNA substrates.

Protein	K_m (nM)	V_{max} (nM/min)	k_{cat} (min^{-1})	k_{cat}/K_M ($nM^{-1}min^{-1}$)	Substrate	Protein (nM)	Ref
hABH2	430 ± 27	34 ± 1	$(6.7 \pm 0.1) \times 10^{-1}$	1.5×10^{-3}	15mer	50	This work
	82 ± 22	$(9 \pm 1) \times 10^{-4}$	$(1.8 \pm 0.2) \times 10^{-2}$	2.2×10^{-4}	24mer	0.05	14(a)
hABH3	1,000 ± 160	103 ± 5	2.1 ± 0.1	2.1×10^{-3}	15mer	50	This work
	162 ± 48	$(1.7 \pm 0.3) \times 10^{-3}$	$(3.4 \pm 0.5) \times 10^{-2}$	2.1×10^{-4}	24mer	0.05	14(a)
AlkB	35.3 ± 6.6	13 ± 1.5	2.6 ± 0.3	7.4×10^{-2}	15mer	5	15
	3,400 ± 600	30.8 ± 1.4	2.2 ± 0.1	6.4×10^{-4}	19mer	14	7
	24,000 ± 5,000	2,100 ± 400	21 ± 4	8.8×10^{-4}	3mer	100	21
	290 ± 30	2,300 ± 1,000	23 ± 10	7.9×10^{-2}	5mer	100	21

Initial rates were analyzed using nonlinear regression curve fitting with GraphPad Prism 5.
(a) Values were measured for the mouse variant of the protein.

inhibition of drug candidates, such as the aptamers selected in the previous chapter. This approach would allow for accurate measurement of inhibition constants such as K_i and IC_{50}.

3.3 Conclusion

In this work, the method to measure the activity of bacterial demethylation enzymes was adapted to human dealkylation repair enzymes involved in cancer chemotherapeutic resistance. Kinetic analysis of fluorescently labelled demethylated substrates is simple, fast, cost effective, and uses readily available, non-radioactive substances. Such a technique has great potential to become a universal approach to both measure the activity and potentially discovering and analyzing inhibitors for this family of enzymes. This method allows for the screening of large libraries of potential drug candidates in a search for potent modulators.

3.4 Materials and Methods

3.4.1 Materials

Synthetic fluorescently labelled DNA substrate (5'-TTCmTTTTTTTTTTT-3'-fluorescein) was synthesized by ATDbio (University of Southampton, UK). hABH2 and hABH3 were expressed and purified as described elsewhere,[169] and stored in 50 mM Tris-HCl, 500 mM NaCl, 1 mM DTT pH 7.5 buffer at −80°C. An uncoated fused-silica capillary was purchased from Polymicro (Phoenix, AZ, USA). DNA aptamers used in inhibition assay were synthesized and HPLC-purified by Integrated DNA Technologies (IDT, Mississauga, ON, Canada). All other chemicals were obtained from Sigma-Aldrich (Toronto, ON, Canada). All solutions were made using deionized water filtered through a 0.22 μm filter (Millipore, Nepean, ON).

3.4.2 Instrumentation

An uncoated fused silica capillary with a length of 80 cm (70 cm to the detection window) and inner and outer diameters of 75 and 365 μm, respectively, was used in all experiments. The capillary was mounted on a CE instrument (P/ACE MDQ, Beckman Coulter, Fullerton, CA) utilizing LIF detection with excitation at 488 nm and emission at 520 nm.

Brand new capillaries were preconditioned using a wash with 2 capillary volumes of methanol, followed by 5 capillary volumes of sequential washes with 100 mM HCl, 100 mM NaOH, deionized water, and, finally, equilibrated with the run buffer for 10 min.

A sample was introduced into the capillary by a pressure pulse of 1 psi (~7 kPa) for 28 s. The demethylated product and methylated substrate were separated by CE using a 20 mM Borax, 60 mM SDS run buffer at pH 8.6 and quantitated with LIF detection.

3.4.3 Measuring Enzyme Kinetics

The demethylation reaction was initiated by the addition of enzyme (50 nM hABH2 or hABH3) to a mixture containing 50 mM Tris-HCl at pH 7.5, 4 mM L-ascorbic acid, 160 μM 2-oxoglutarate, 80 μM $FeSO_4 \cdot 7H_2O$, 2740 units of catalase, 100 μg/mL BSA and 125 nM – 2 μM or 750 nM - 6 μM 5'-TTCmTTTTTTTTTTTT-3'-fluorescein substrate for hABH2 and hABH3 respectively, as final concentrations. The final volume of the reaction mixture was 100 μL.

Fifteen μL of the reaction mixture was taken at different time points, added to pre-chilled EDTA at a final concentration of 5 mM, followed by heating at 75°C for 10 min, to stop the reaction. The samples were then stored at −20°C until their analysis by CE. The initial reaction rate (nM/min) for Michaelis-Menten analysis was measured as a slope of the linear part of the "product versus reaction time" curve. K_m and V_{max} values were calculated from the Michaelis-Menten plot utilizing nonlinear regression curve fitting using GraphPad Prism 5. The value of k_{cat} was calculated as V_{max}/[E] where [E] is the enzyme concentration.

Limitations

Although NECEEM-based aptamer selection has unmatched partitioning efficiency, it has inherit limitations. A major limitation is the nature of the instrumentation. The silica capillary in which the separation occurs is composed of silanol groups (SiOH). Ionization of the silanol groups (SiO^-) at pH values above 3 is the fundamental property for the electroosmotic flow (EOF) formation. The negative charge poses problems, however, for the type of protein targets that can be used. Positively charged proteins can easily stick to the negatively charged capillary wall and hinder separation and therefore aptamer selection. This limits the possible protein target repertoire. Although a wide variety of coatings exist to reduce protein sticking by masking the charge, the obvious consequence is diminished EOF and therefore hindered separation ability.

Measuring enzyme kinetics and inhibition using our method, although effective, can be a lengthy and labour intensive process. If a substrate or exact assay conditions are unknown, testing different substrates and assay components, in addition to testing inhibitors, lengthens the process even further. Additionally, due to the speed of the reaction, measurement of initial rates is prone to human error. There is a need, therefore, for a more rapid approach, either a continuous assay or automated non-continuous method.

Concluding Remarks

The utility and potential of aptamers continues to grow year by year. However, in spite of their benefits over antibodies, aptamers have yet to gain widespread recognition and utilization. In fact, other than a select few, many researchers are not even aware of their existence and potential. This is likely due to two main issues: First, for laboratories, particularly clinical laboratories that have used antibodies for many years, have established protocols and training for handling antibodies are naturally resistant to change. Secondly, aptamers suffer from the "Thrombin problem". The thrombin aptamer is the most well-known and most widely used aptamer. Hundreds of labs study this aptamer rather than develop aptamers and assays for more clinically relevant targets. For aptamers to penetrate more labs and the clinical field, they have to not only have a substantial advantage over antibodies, but their utility has to be proven for a variety of targets and techniques.

To combat these issues, a universal approach to select aptamers to clinically relevant targets is essential. As such, SELEX development has to go in the direction of homogeneous partitioning and automation. Such techniques will pave the way for more *in-vivo* clinically relevant applications. Once established, such aptamers can start making a more significant penetration in the medical and research field.

Future Direction

In this work, we focused on improving the efficiency of PCR following NECEEM amplification by isolating the templates in separate micro-reactors. Although effective, there are many areas that could be optimized further. Among them, combining asymmetric PCR with emulsion could be attempted using the BEAMing (beads, emulsions, amplification and magnetics) method. Here, the reverse primer is attached to magnetic beads which allows for the isolation of the forward strands under denaturating conditions after amplification and purification of the strands by removing the magnetic beads. This will further automate the method and reduce possible contamination arising from sample manipulation. Additionally, the polymerase used for amplification can be optimized. A higher quality polymerase used in early rounds will increase efficiency (thus preserving the aptamers). Conversely, an error-prone polymerase used in later rounds could be introduced to induce mutations to already enriched aptamers in hopes of improving binding affinity.

An immediate next step to continue this work is to use the direct CE activity assay approach to measure the inhibition potential of the selected hABH2 aptamers. Measuring the activity in the presence of varying concentrations of aptamers will allow one to measure the inhibition constant of the aptamers. These aptamer inhibitors can then be used for *in-vivo* studies. An important factor to consider before doing so is nuclease resistance. Nuclease digestion is a major factor for the short half-life of aptamers *in-vivo*. For example, unmodified thrombin aptamer only lasts 108 seconds *in-vivo*.[249] If not addressed, aptamers are likely to be degraded prior to performing their intended function. The solution to this issue lies in modified nucleotides. Either a fluorine (F) or an amino group (NH2) can be used to substitute the 2'OH of the ribose moiety of pyrimidines. Such modification was found to increase the nuclease resistance by several orders of magnitude.[250] This modification has in fact been used with the Pegaptanib therapeutic aptamer for this purpose.[251] Although this modification can be done to existing aptamers[252] Such nucleotide modification were successfully used in CE-SELEX and would be a useful addition to the aptamer selection process.[203] An additional possible modification is the substitution of 2'-O-methyl (2'-OMe) at the 2'-OH position of purines.[253]

As mentioned previously, our activity and inhibition assay approach, although successful, is lengthy and labour intensive. To increase accuracy, reproducibility and efficiency, an automated or continuous approach should be integrated with our method. Possible approaches include mixing reagents in the capillary[254] or starting the reaction in a tube and repeatedly injecting reactions at specific time points.[255] Such improvements are essential for future research to characterize and find inhibitors for other enzymes of this family.

References

1 Yi, C., Yang, C.-G. & He, C. A non-heme iron-mediated chemical demethylation in DNA and RNA. *Accounts of chemical research* **42**, 519-529 (2009).

2 Que, L. One motif-many different reactions. *Nature structural biology* **7**, 182-183 (2000).

3 Schofield, C. J. & Zhang, Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Current opinion in structural biology* **9**, 722-731 (1999).

4 Aas, P. A. *et al.* Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **421**, 859-863, doi:Doi 10.1038/Nature01363 (2003).

5 Siddik, Z. H. Mechanisms of Action of Cancer Chemotherapeutic Agents: DNA-Interactive Alkylating Agents and Antitumour Platinum-Based Drugs. *The cancer handbook* (2002).

6 Yang, C.-G. *et al.* Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA. *Nature* **452**, 961-965 (2008).

7 Roy, T. W. & Bhagwat, A. Kinetic studies of Escherichia coli AlkB using a new fluorescence-based assay for DNA demethylation. *Nucleic acids research* **35**, e147-e147 (2007).

8 Cao, W. & Enrique, M. Quantitative full time course analysis of nonlinear enzyme cycling kinetics. *Scientific reports* **3** (2013).

9 Falnes, P. Ø., Johansen, R. F. & Seeberg, E. AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. *Nature* **419**, 178-182 (2002).

10 Berezovski, M. & Krylov, S. N. Nonequilibrium capillary electrophoresis of equilibrium mixtures-A single experiment reveals equilibrium and kinetic parameters of protein-DNA interactions. *Journal of the American Chemical Society* **124**, 13674-13675 (2002).

11 Falnes, P. Ø., Bjørås, M., Aas, P. A., Sundheim, O. & Seeberg, E. Substrate specificities of bacterial and human AlkB proteins. *Nucleic acids research* **32**, 3456-3461 (2004).

12 Duncan, T. *et al.* Reversal of DNA alkylation damage by two human dioxygenases. *Proceedings of the National Academy of Sciences* **99**, 16660-16665 (2002).

13 Delaney, J. C. *et al.* AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. *Nature structural & molecular biology* **12**, 855-860 (2005).

14 Lee, D.-H. *et al.* Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. *Journal of Biological Chemistry* **280**, 39448-39459 (2005).

15 Karkhanina, A. A. *et al.* Direct analysis of enzyme-catalyzed DNA demethylation. *Analytical chemistry* **81**, 5871-5875 (2009).

16 Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T. & Sedgwick, B. Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. *Nature* **419**, 174-178 (2002).

17 Sundheim, O. *et al.* Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage. *The EMBO journal* **25**, 3389-3397 (2006).

18 Ringvoll, J. *et al.* Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. *The EMBO journal* **25**, 2189-2198 (2006).

19 Thalhammer, A., Aik, W., Bagg, E. A. & Schofield, C. J. The potential of 2-oxoglutarate oxygenases acting on nucleic acids as therapeutic targets. *Drug Discovery Today: Therapeutic Strategies* **9**, e91-e100 (2012).

20 Falnes, P. Ø. Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. *Nucleic acids research* **32**, 6260-6267 (2004).

21 Yu, B. & Hunt, J. F. Enzymological and structural studies of the mechanism of promiscuous substrate recognition by the oxidative DNA repair enzyme AlkB. *Proceedings of the National Academy of Sciences* **106**, 14315-14320 (2009).

22 Ringvoll, J. *et al.* AlkB Homologue 2–Mediated Repair of Ethenoadenine Lesions in Mammalian DNA. *Cancer research* **68**, 4142-4149 (2008).

23 Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709-715 (1993).

24 Friedberg, E. C., Walker, G. C. & Siede, W. *DNA repair and mutagenesis*. Second edn, (American Society for Microbiology, 2006).

25 Tsujikawa, K. *et al.* Expression and sub-cellular localization of human ABH family molecules. *Journal of cellular and molecular medicine* **11**, 1105-1116 (2007).

26 Lombard, D. B. *et al.* DNA repair, genome stability, and aging. *Cell* **120**, 497-512 (2005).

27 Drabløs, F. *et al.* Alkylation damage in DNA and RNA—repair mechanisms and medical significance. *DNA repair* **3**, 1389-1407 (2004).

28 Ramchandani, S., Bhattacharya, S. K., Cervoni, N. & Szyf, M. DNA methylation is a reversible biological signal. *Proceedings of the National Academy of Sciences* **96**, 6107-6112 (1999).

29 Delaney, J. C. & Essigmann, J. M. Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14051-14056 (2004).

30 Safihill, R. Differences in the promutagenic nature of 3-methylcytosine as revealed by DNA and RNA polymerising enzymes. *Carcinogenesis* **5**, 691-693 (1984).

31 Engelward, B. P. *et al.* A chemical and genetic approach together define the biological consequences of 3-methyladenine lesions in the mammalian genome. *Journal of Biological Chemistry* **273**, 5412-5418 (1998).

32 Siddik, Z. H. *Mechanisms of Action of Cancer Chemotherapeutic Agents: DNA-Interactive Alkylating Agents and Antitumour Platinum-Based Drugs*. (2002).

33 Pieper, R. O. & Erickson, L. C. DNA adenine adducts induced by nitrogen mustards and their role in transcription termination in vitro. *Carcinogenesis* **11**, 1739-1746 (1990).

34 Hurley, L. H. DNA and its associated processes as targets for cancer therapy. *Nature Reviews Cancer* **2**, 188-200 (2002).

35 Chaney, S. G. & Sancar, A. DNA repair: enzymatic mechanisms and relevance to drug response. *Journal of the National Cancer Institute* **88**, 1346-1360 (1996).

36 Jacobson, L. O. *et al.* Nitrogen mustard therapy: Studies on the effect of methyl-bis (beta-chloroethyl) amine hydrochloride on neoplastic diseases and allied disorders of the hemopoietic system. *Journal of the American Medical Association* **132**, 263-271 (1946).

37 Karnofsky, D. A., Abelmann, W. H., Craver, L. F. & Burchenal, J. H. The use of the nitrogen mustards in the palliative treatment of carcinoma. With particular reference to bronchogenic carcinoma. *Cancer* **1**, 634-656 (1948).

38 Van Maanen, M., Smeets, C. & Beijnen, J. Chemistry, pharmacology and pharmacokinetics of N, N′, N″-triethylenethiophosphoramide (ThioTEPA). *Cancer treatment reviews* **26**, 257-268 (2000).

39 Tong, W. P. & Ludlum, D. B. Crosslinking of DNA by busulfan Formation of diguanyl derivatives. *Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis* **608**, 174-181 (1980).

40 Chen, F.-X., Zhang, Y., Church, K. M., Bodell, W. J. & Gold, B. DNA crosslinking, sister chromatid exchange and cytotoxicity of N-2-chloroethylnitrosoureas tethered to minor groove binding peptides. *Carcinogenesis* **14**, 935-940 (1993).

41 Prestayko, A., d'Aoust, J., Issell, B. & Crooke, S. Cisplatin (cis-diamminedichloroplatinum II). *Cancer treatment reviews* **6**, 17-39 (1979).

42 Roberts, J., Brent, T. & Crathorn, A. The mechanism of the cytotoxic action of alkylating agents on mammalian cells. *The interaction of drugs and subcellular components in animal cells* **5** (1968).

43 Bodell, W. & Singer, B. Influence of hydrogen bonding in DNA and polynucleotides on reaction of nitrogens and oxygens toward ethylnitrosourea. *Biochemistry* **18**, 2860-2863 (1979).

44 Boiteux, S. & Laval, J. Mutagenesis by alkylating agents: coding properties for DNA polymerase of poly (dC) template containing 3-methylcytosine. *Biochimie* **64**, 637-641 (1982).

45 Hartley, J., O'Hare, C. & Baumgart, J. DNA alkylation and interstrand cross-linking by treosulfan. *British journal of cancer* **79**, 264 (1999).

46 Rajski, S. R. & Williams, R. M. DNA cross-linking agents as antitumor drugs. *Chemical reviews* **98**, 2723-2796 (1998).

47 Colvin, M. E., Sasaki, J. C. & Tran, N. L. Chemical factors in the action of phosphoramidic mustard alkylating anticancer drugs: roles for computational chemistry. *Current pharmaceutical design* **5**, 645-664 (1999).

48 Emadi, A., Jones, R. J. & Brodsky, R. A. Cyclophosphamide and cancer: golden anniversary. *Nature reviews. Clinical oncology* **6**, 638-647, doi:10.1038/nrclinonc.2009.146 (2009).

49 Garcia, S. T., McQuillan, A. & Panasci, L. Correlation between the cytotoxicity of melphalan and DNA crosslinks as detected by the ethidium bromide fluorescence assay in the F 1 variant of B 16 melanoma cells. *Biochemical pharmacology* **37**, 3189-3192 (1988).

50 Cox, P. J. Cyclophosphamide cystitis—identification of acrolein as the causative agent. *Biochemical pharmacology* **28**, 2045-2049 (1979).

51 Geleziunas, R. *et al.* Increased DNA synthesis and repair-enzyme expression in lymphocytes from patients with chronic lymphocytic leukemia resistant to nitrogen mustards. *Journal of the National Cancer Institute* **83**, 557-564 (1991).

52 Panasci, L., Xu, Z.-Y., Bello, V. & Aloyz, R. The role of DNA repair in nitrogen mustard drug resistance. *Anti-cancer drugs* **13**, 211-220 (2002).

53 Ultmann, J. E., Hyman, G. A., Crandall, C., Naujoks, H. & Gellhorn, A. Triethylenethiophosphoramide (Thio-TEPA) in the treatment of neoplastic disease. *Cancer* **10**, 902-911 (1957).

54 Grigorii, V. A. *et al.* Direct observation of the alkylation products of deoxyguanosine and DNA by fast atom bombardment mass spectrometry. *Biological mass spectrometry* **20**, 665-668 (1991).

55 D'lncalci, M., Broggini, M. & Hartley, J. A. in *New Approaches in Cancer Pharmacology: Drug Design and Development* 5-11 (Springer, 1992).

56 Crooke, S. T. & Bradner, W. T. Mitomycin C: a review. *Cancer treatment reviews* **3**, 121-139 (1976).

57 Matsumoto, A., Vos, J.-M. H. & Hanawalt, P. C. Repair analysis of mitomycin C-induced DNA crosslinking in ribosomal RNA genes in lymphoblastoid cells from Fanconi's anemia patients. *Mutation Research/DNA Repair* **217**, 185-192 (1989).

58 Patel, K. J. *et al.* Involvement of Brca2 in DNA repair. *Molecular cell* **1**, 347-357 (1998).

59 Ponti, M., Souhami, R., Fox, B. & Hartley, J. DNA interstrand crosslinking and sequence selectivity of dimethanesulphonates. *British journal of cancer* **63**, 743 (1991).

60 Goldman, J. M. & Melo, J. V. Chronic myeloid leukemia—advances in biology and new approaches to treatment. *New England Journal of Medicine* **349**, 1451-1464 (2003).

61 Bedford, P. & Fox, B. W. Repair of DNA interstrand crosslinks after busulphan. A possible mode of resistance. *Cancer chemotherapy and pharmacology* **8**, 3-7 (1982).

62 Schabel Jr, F. Nitrosoureas: a review of experimental antitumor activity. *Cancer treatment reports* **60**, 665-698 (1976).

63 Ludlum, D. B. The chloroethylnitrosoureas: sensitivity and resistance to cancer chemotherapy at the molecular level. *Cancer investigation* **15**, 588-598 (1997).

64 Lemoine, A., Lucas, C. & Ings, R. Metabolism of the chloroethylnitrosoureas. *Xenobiotica* **21**, 775-791 (1991).

65 Walker, M., Masters, J. & Margison, G. O6-alkylguanine-DNA-alkyltransferase activity and nitrosourea sensitivity in human cancer cell lines. *British journal of cancer* **66**, 840 (1992).

66 El-Khateeb, M. *et al.* Reactions of cisplatin hydrolytes with methionine, cysteine, and plasma ultrafiltrate studied by a combination of HPLC and NMR techniques. *Journal of inorganic biochemistry* **77**, 13-21 (1999).

67 Wu, S.-s. *et al.* Down-regulation of ALKBH2 increases cisplatin sensitivity in H1299 lung cancer cells. *Acta pharmacologica Sinica* **32**, 393-398 (2011).

68 Johannessen, T.-C. A. *et al.* The DNA repair protein ALKBH2 mediates temozolomide resistance in human glioblastoma cells. *Neuro-oncology* **15**, 269-278 (2012).

69 Marcu, L., Bezak, E., Olver, I. & Van Doorn, T. Tumour resistance to cisplatin: a modelling approach. *Physics in medicine and biology* **50**, 93 (2005).

70 Martin, N. DNA repair inhibition and cancer therapy. *Journal of Photochemistry and Photobiology B: Biology* **63**, 162-170 (2001).

71 Kiberstis, P. A. Cancer Therapy on Target. *Science* **292**, 399-401, doi:10.1126/science.292.5516.399e (2001).

72 Christmann, M., Tomicic, M. T., Roos, W. P. & Kaina, B. Mechanisms of human DNA repair: an update. *Toxicology* **193**, 3-34 (2003).

73 Yi, C. & He, C. DNA repair by reversal of DNA damage. *Cold Spring Harbor perspectives in biology* **5**, a012575 (2013).

74 Aravind, L. & Koonin, E. V. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate-and iron-dependent dioxygenases. *Genome Biol* **2**, 1-0007.0008 (2001).

75 Sedgwick, B. & Lindahl, T. Recent progress on the Ada response for inducible repair of DNA alkylation damage. *Oncogene* **21**, 8886 (2002).

76 Kondo, H. *et al.* Structure and expression of the alkB gene of Escherichia coli related to the repair of alkylated DNA. *Journal of Biological Chemistry* **261**, 15772-15777 (1986).

77 Hutton, J. J., Jr., Trappel, A. L. & Udenfriend, S. Requirements for alpha-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase. *Biochemical and biophysical research communications* **24**, 179-184 (1966).

78 Loenarz, C. & Schofield, C. J. Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases. *Trends in biochemical sciences* **36**, 7-18 (2011).

79 Trewick, S. C., McLaughlin, P. J. & Allshire, R. C. Methylation: lost in hydroxylation? *EMBO reports* **6**, 315-320 (2005).

80 Rose, N. R., McDonough, M. A., King, O. N., Kawamura, A. & Schofield, C. J. Inhibition of 2-oxoglutarate dependent oxygenases. *Chemical Society reviews* **40**, 4364-4397 (2011).

81 Gerken, T. *et al.* The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. *Science* **318**, 1469-1472 (2007).

82 Loenarz, C. & Schofield, C. J. Expanding chemical biology of 2-oxoglutarate oxygenases. *Nature chemical biology* **4**, 152-156 (2008).

83 Flashman, E. & Schofield, C. J. The most versatile of all reactive intermediates? *Nature chemical biology* **3**, 86-87 (2007).

84 Kondo, N., Takahashi, A., Ono, K. & Ohnishi, T. DNA damage induced by alkylating agents and repair pathways. *Journal of nucleic acids* **2010** (2010).

85 Persidis, A. Cancer multidrug resistance. *Nature biotechnology* **17**, 94-95 (1999).

86 Zamble, D. B. & Lippard, S. J. Cisplatin and DNA repair in cancer chemotherapy. *Trends in biochemical sciences* **20**, 435-439 (1995).

87 Chen, B. J., Carroll, P. & Samson, L. The Escherichia coli AlkB protein protects human cells against alkylation-induced toxicity. *Journal of bacteriology* **176**, 6255-6261 (1994).

88 Wei, Y.-F., Carter, K. C., Wang, R.-P. & Shell, B. K. Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair. *Nucleic acids research* **24**, 931-937 (1996).

89 Haug, T. *et al.* Human uracil-DNA glycosylase gene: sequence organization, methylation pattern, and mapping to chromosome 12q23–q24. 1. *Genomics* **36**, 408-416 (1996).

90 Larson, K., Sahm, J., Shenkar, R. & Strauss, B. Methylation-induced blocks to in vitro DNA replication. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **150**, 77-84 (1985).

91 Saffhill, R. Differences in the promutagenic nature of 3-methylcytosine as revealed by DNA and RNA polymerising enzymes. *Carcinogenesis* **5**, 691-693 (1984).

92 Hartwig, A. Role of magnesium in genomic stability. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **475**, 113-121 (2001).

93 Nair, J., Barbin, A., Guichard, Y. & Bartsch, H. 1,N6-Ethenodeoxyadenosine and 3,N4-ethenodeoxycytidine in liver DNA from humans and untreated rodents detected by immunoaffinity/32P-postlabelling. *Carcinogenesis* **16**, 613-617, doi:10.1093/carcin/16.3.613 (1995).

94 Levine, R. L. *et al.* Mutagenesis induced by a single 1, N6-ethenodeoxyadenosine adduct in human cells. *Cancer research* **60**, 4098-4104 (2000).

95 Speina, E. *et al.* Decreased repair activities of 1, N6-ethenoadenine and 3, N4-ethenocytosine in lung adenocarcinoma patients. *Cancer research* **63**, 4351-4357 (2003).

96 Hang, B., Singer, B., Margison, G. P. & Elder, R. H. Targeted deletion of alkylpurine-DNA-N-glycosylase in mice eliminates repair of 1, N6-ethenoadenine and hypoxanthine but not of 3, N4-ethenocytosine or 8-oxoguanine. *Proceedings of the National Academy of Sciences* **94**, 12869-12874 (1997).

97 Barbin, A., Wang, R., O'Connor, P. J. & Elder, R. H. Increased formation and persistence of 1, N6-Ethenoadenine in DNA is not associated with higher susceptibility to carcinogenesis in alkylpurine-DNA-N-Glycosylase knockout mice treated with vinyl carbamate. *Cancer research* **63**, 7699-7703 (2003).

98 Guichard, Y., El Ghissassi, F., Nair, J., Bartsch, H. & Barbin, A. ACCELERATED PAPER: Formation and accumulation of DNA ethenobases in adult Sprague-Dawley rats exposed to vinyl chloride. *Carcinogenesis* **17**, 1553-1559 (1996).

99 Contreras-Moreira, B. & Bates, P. A. Domain fishing: a first step in protein comparative modelling. *Bioinformatics* **18**, 1141-1142 (2002).

100 Yu, B. *et al.* Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. *Nature* **439**, 879-884 (2006).

101 Monsen, V. T. *et al.* Divergent β-hairpins determine double-strand versus single-strand substrate recognition of human AlkB-homologues 2 and 3. *Nucleic acids research*, gkq518 (2010).

102 Moldovan, G.-L., Pfander, B. & Jentsch, S. PCNA, the maestro of the replication fork. *Cell* **129**, 665-679 (2007).

103 Gilljam, K. M. *et al.* Identification of a novel, widespread, and functionally important PCNA-binding motif. *The Journal of cell biology* **186**, 645-654 (2009).

104 Scott, R. B. Cancer chemotherapy--the first twenty-five years. *British medical journal* **4**, 259-265 (1970).

105 Oliver, T. G. *et al.* Chronic cisplatin treatment promotes enhanced damage repair and tumor progression in a mouse model of lung cancer. *Genes Dev* **24**, 837-852, doi:10.1101/gad.1897010 (2010).

106 Marcu, L., Bezak, E., Olver, I. & van Doorn, T. Tumour resistance to cisplatin: a modelling approach. *Physics in medicine and biology* **50**, 93-102 (2005).

107 O'Connor, A. E., Gallagher, W. M. & Byrne, A. T. Porphyrin and nonporphyrin photosensitizers in oncology: preclinical and clinical advances in photodynamic therapy. *Photochem Photobiol* **85**, 1053-1074, doi:10.1111/j.1751-1097.2009.00585.x (2009).

108 Gerstner, E. R. & Fine, R. L. Increased permeability of the blood-brain barrier to chemotherapy in metastatic brain tumors: establishing a treatment paradigm. *J Clin Oncol* **25**, 2306-2312, doi:10.1200/jco.2006.10.0677 (2007).

109 Ferrario, A., Rucker, N., Wong, S., Luna, M. & Gomer, C. J. Survivin, a member of the inhibitor of apoptosis family, is induced by photodynamic therapy and is a target for improving treatment response. *Cancer Res* **67**, 4989-4995, doi:10.1158/0008-5472.can-06-4785 (2007).

110 Stylli, S. S., Howes, M., MacGregor, L., Rajendra, P. & Kaye, A. H. Photodynamic therapy of brain tumours: evaluation of porphyrin uptake versus clinical outcome. *J Clin Neurosci* **11**, 584-596, doi:10.1016/j.jocn.2004.02.001 (2004).

111 Rigual, N. R. *et al.* Photodynamic therapy for head and neck dysplasia and cancer. *Arch Otolaryngol Head Neck Surg* **135**, 784-788, doi:10.1001/archoto.2009.98 (2009).

112 Saczko, J. *et al.* Photooxidative action in cancer and normal cells induced by the use of photofrin in photodynamic therapy. *Folia Biol (Praha)* **54**, 24-29 (2008).

113 Roy, L. D. *et al.* MUC1 enhances invasiveness of pancreatic cancer cells by inducing epithelial to mesenchymal transition. *Oncogene* **30**, 1449-1459 (2010).

114 Lehmann, A. R. & Taylor, E. M. in *DNA Damage and Repair* 377-401 (Springer, 2001).

115 Smith, M. L. & Seo, Y. R. p53 regulation of DNA excision repair pathways. *Mutagenesis* **17**, 149-156 (2002).

116 Cetica, V. *et al.* Pediatric brain tumors: mutations of two dioxygenases (hABH2 and hABH3) that directly repair alkylation damage. *Journal of neuro-oncology* **94**, 195-201 (2009).

117 Stupp, R. *et al.* Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *New England Journal of Medicine* **352**, 987-996 (2005).

118 Fujii, T., Shimada, K., Anai, S., Fujimoto, K. & Konishi, N. ALKBH2, a novel AlkB homologue, contributes to human bladder cancer progression by regulating MUC1 expression. *Cancer science* **104**, 321-327 (2013).

119 Mitra, A. P., Datar, R. H. & Cote, R. J. Molecular pathways in invasive bladder cancer: new insights into mechanisms, progression, and target identification. *Journal of Clinical Oncology* **24**, 5552-5564 (2006).

120 Kawano, T. *et al.* MUC1 oncoprotein promotes growth and survival of human multiple myeloma cells. *Int J Oncol* **33**, 153-159 (2008).

121 Agata, N. *et al.* MUC1 oncoprotein blocks death receptor–mediated apoptosis by inhibiting recruitment of caspase-8. *Cancer research* **68**, 6136-6144 (2008).

122 Myllyharju, J. Prolyl 4-hydroxylases, key enzymes in the synthesis of collagens and regulation of the response to hypoxia, and their roles as treatment targets. *Annals of medicine* **40**, 402-417 (2008).

123 Myllyharju, J. HIF prolyl 4-hydroxylases and their potential as drug targets. *Current pharmaceutical design* **15**, 3878-3885 (2009).

124 Mole, D. R. *et al.* 2-oxoglutarate analogue inhibitors of HIF prolyl hydroxylase. *Bioorganic & medicinal chemistry letters* **13**, 2677-2680 (2003).

125 Stubbs, C. J. *et al.* Application of a proteolysis/mass spectrometry method for investigating the effects of inhibitors on hydroxylase structure. *Journal of medicinal chemistry* **52**, 2799-2805 (2009).

126 Cunliffe, C. J., Franklin, T. J., Hales, N. J. & Hill, G. B. Novel inhibitors of prolyl 4-hydroxylase. 3. Inhibition by the substrate analog N-oxaloglycine and its derivatives. *Journal of medicinal chemistry* **35**, 2652-2658 (1992).

127 Krylova, S. M., Koshkin, V., Bagg, E., Schofield, C. J. & Krylov, S. N. Mechanistic studies on the application of DNA aptamers as inhibitors of 2-oxoglutarate-dependent oxygenases. *Journal of medicinal chemistry* **55**, 3546-3552 (2012).

128 Selak, M. A. *et al.* Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. *Cancer cell* **7**, 77-85 (2005).

129 Kivirikko, K. I. & Myllyharju, J. Prolyl 4-hydroxylases and their protein disulfide isomerase subunit. *Matrix Biology* **16**, 357-368 (1998).

130 Dambrova, M., Liepinsh, E. & Kalvinsh, I. Mildronate: cardioprotective action through carnitine-lowering effect. *Trends in cardiovascular medicine* **12**, 275-279 (2002).

131 Rademacher, W. Growth retardants: effects on gibberellin biosynthesis and other metabolic pathways. *Annual review of plant biology* **51**, 501-531 (2000).

132 Konishi, N. *et al.* High expression of a new marker PCA-1 in human prostate carcinoma. *Clinical cancer research* **11**, 5090-5097 (2005).

133 Yamato, I. *et al.* PCA-1/ALKBH3 contributes to pancreatic cancer by supporting apoptotic resistance and angiogenesis. *Cancer research* **72**, 4829-4839 (2012).

134 Agris, P. F. The importance of being modified: roles of modified nucleosides and Mg2+ in RNA structure and function. *Prog Nucleic Acid Res Mol Biol* **53**, 79-129 (1996).

135 Yoshizawa, S., Fourmy, D. & Puglisi, J. D. Recognition of the codon-anticodon helix by ribosomal RNA. *Science* **285**, 1722-1725 (1999).

136 Matsugi, J. & Murao, K. Study on construction of a cDNA library corresponding to an amino acid-specific tRNA and influence of the modified nucleotide upon nucleotide misincorporations in reverse transcription. *Biochim Biophys Acta* **1521**, 81-88 (2001).

137 Fu, D. *et al.* Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. *Molecular and cellular biology* **30**, 2449-2459 (2010).

138 Dango, S. *et al.* DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation. *Molecular cell* **44**, 373-384 (2011).

139 Shimada, K. *et al.* A novel human AlkB homologue, ALKBH8, contributes to human bladder cancer progression. *Cancer research* **69**, 3157-3164 (2009).

140 Tasaki, M., Shimada, K., Kimura, H., Tsujikawa, K. & Konishi, N. ALKBH3, a human AlkB homologue, contributes to cell survival in human non-small-cell lung cancer. *British journal of cancer* **104**, 700-706 (2011).

141 Linardou, H., Dahabreh, I. J., Bafaloukos, D., Kosmidis, P. & Murray, S. Somatic EGFR mutations and efficacy of tyrosine kinase inhibitors in NSCLC. *Nature reviews Clinical oncology* **6**, 352-366 (2009).

142 Hanchette, C. L. & Schwartz, G. G. Geographic patterns of prostate cancer mortality. Evidence for a protective effect of ultraviolet radiation. *Cancer* **70**, 2861-2869 (1992).

143 Lu, H., Shi, X., Costa, M. & Huang, C. Carcinogenic effect of nickel compounds. *Molecular and cellular biochemistry* **279**, 45-67 (2005).

144 Chen, H. & Costa, M. Iron- and 2-oxoglutarate-dependent Dioxygenases: an emerging group of molecular targets for nickel toxicity and carcinogenicity. *Biometals* **22**, 191-196, doi:10.1007/s10534-008-9190-3 (2009).

145 Davidson, T. L., Chen, H., Di Toro, D. M., D'Angelo, G. & Costa, M. Soluble nickel inhibits HIF-prolyl-hydroxylases creating persistent hypoxic signaling in A549 cells. *Molecular carcinogenesis* **45**, 479-489 (2006).

146 Kurowski, M. A., Bhagwat, A. S., Papaj, G. & Bujnicki, J. M. Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. *BMC genomics* **4**, 48 (2003).

147 Kataoka, H., Yamamoto, Y. & Sekiguchi, M. A new gene (alkB) of Escherichia coli that controls sensitivity to methyl methane sulfonate. *Journal of bacteriology* **153**, 1301-1307 (1983).

148 Koivisto, P., Duncan, T., Lindahl, T. & Sedgwick, B. Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. *Journal of Biological Chemistry* **278**, 44348-44354 (2003).

149 Mishina, Y., Chen, L. X. & He, C. Preparation and characterization of the native iron (II)-containing DNA repair AlkB protein directly from Escherichia coli. *Journal of the American Chemical Society* **126**, 16930-16936 (2004).

150 Dinglay, S., Trewick, S. C., Lindahl, T. & Sedgwick, B. Defective processing of methylated single-stranded DNA by E. coli AlkB mutants. *Genes & development* **14**, 2097-2105 (2000).

151 Ougland, R. *et al.* AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation. *Molecular cell* **16**, 107-116 (2004).

152 El Ghissassi, F., Barbin, A., Nair, J. & Bartsch, H. Formation of 1, N6-ethenoadenine and 3, N4-ethenocytosine by lipid peroxidation products and nucleic acid bases. *Chemical research in toxicology* **8**, 278-283 (1995).

153 Nair, J., Barbin, A., Guichard, Y. & Bartsch, H. 1, N6-Ethenodeoxyadenosine and 3, N4-ethenodeoxycytidine in liver DNA from humans and untreated rodents detected by immunoaffinity/32P-postlabelling. *Carcinogenesis* **16**, 613-617 (1995).

154 Saparbaev, M. & Laval, J. 3, N4-ethenocytosine, a highly mutagenic adduct, is a primary substrate for Escherichia coli double-stranded uracil-DNA glycosylase and human mismatch-specific thymine-DNA glycosylase. *Proceedings of the National Academy of Sciences* **95**, 8508-8513 (1998).

155 Mishina, Y., Lee, C. H. J. & He, C. Interaction of human and bacterial AlkB proteins with DNA as probed through chemical cross-linking studies. *Nucleic acids research* **32**, 1548-1554 (2004).

156 Westbye, M. P. *et al.* Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. *Journal of Biological Chemistry* **283**, 25046-25056 (2008).

157 Muller, T. A., Meek, K. & Hausinger, R. P. Human AlkB homologue 1 (ABH1) exhibits DNA lyase activity at abasic sites. *DNA repair* **9**, 58-65 (2010).

158 Thalhammer, A. *et al.* Human AlkB homologue 5 is a nuclear 2-oxoglutarate dependent oxygenase and a direct target of hypoxia-inducible factor 1α (HIF-1α). *PloS one* **6**, e16210 (2011).

159 de Carvalho, D. D. *et al.* Nox1 downstream of 12-lipoxygenase controls cell proliferation but not cell spreading of colon cancer cells. *International journal of cancer* **122**, 1757-1764 (2008).

160 Jia, G. *et al.* Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. *FEBS letters* **582**, 3313-3319 (2008).

161 Jia, G. *et al.* N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nature chemical biology* **7**, 885-887 (2011).

162 Sanchez-Pulido, L. & Andrade-Navarro, M. A. The FTO (fat mass and obesity associated) gene codes for a novel member of the non-heme dioxygenase superfamily. *BMC biochemistry* **8**, 23 (2007).

163 Fischer, J. *et al.* Inactivation of the Fto gene protects from obesity. *Nature* **458**, 894-898 (2009).

164 Dina, C. *et al.* Variation in FTO contributes to childhood obesity and severe adult obesity. *Nature genetics* **39**, 724-726 (2007).

165 Han, Z. *et al.* Crystal structure of the FTO protein reveals basis for its substrate specificity. *Nature* **464**, 1205-1209 (2010).

166 Prescott, A. G. & John, P. Dioxygenases: molecular structure and role in plant metabolism. *Annual review of plant biology* **47**, 245-271 (1996).

167 Michaelis, L. & Menten, M. L. Die kinetik der invertinwirkung. *Biochem. z* **49**, 352 (1913).

168 Sedgwick, B., Robins, P. & Lindahl, T. Direct removal of alkylation damage from DNA by AlkB and related DNA dioxygenases. *Methods in enzymology* **408**, 108-120 (2006).

169 Welford, R. W., Schlemminger, I., McNeill, L. A., Hewitson, K. S. & Schofield, C. J. The selectivity and inhibition of AlkB. *Journal of Biological Chemistry* **278**, 10157-10161 (2003).

170 Sabourin, P. & Bieber, L. L. Purification and characterization of an alpha-ketoisocaproate oxygenase of rat liver. *Journal of Biological Chemistry* **257**, 7460-7467 (1982).

171 Crouse, J. & Amorese, D. Ethanol precipitation: ammonium acetate as an alternative to sodium acetate. *Focus* **9**, 3-5 (1987).

172 Chen, B., Liu, H., Sun, X. & Yang, C.-G. Mechanistic insight into the recognition of single-stranded and double-stranded DNA substrates by ABH2 and ABH3. *Molecular BioSystems* **6**, 2143-2149 (2010).

173 Han, K., Liang, Z. & Zhou, N. Design strategies for aptamer-based biosensors. *Sensors* **10**, 4541-4557 (2010).

174 Jhaveri, S., Rajendran, M. & Ellington, A. D. In vitro selection of signaling aptamers. *Nature biotechnology* **18**, 1293-1297 (2000).

175 Williams, K. P. *et al.* Bioactive and nuclease-resistant L-DNA ligand of vasopressin. *Proceedings of the National Academy of Sciences* **94**, 11285-11290 (1997).

176 Kusser, W. Chemically modified nucleic acid aptamers for in vitro selections: evolving evolution. *Reviews in Molecular Biotechnology* **74**, 27-38 (2000).

177 Stoltenburg, R., Reinemann, C. & Strehlitz, B. SELEX—a (r) evolutionary method to generate high-affinity nucleic acid ligands. *Biomolecular engineering* **24**, 381-403 (2007).

178 Hamula, C. L., Guthrie, J. W., Zhang, H., Li, X.-F. & Le, X. C. Selection and analytical applications of aptamers. *TrAC Trends in Analytical Chemistry* **25**, 681-691 (2006).

179 Song, K.-M., Lee, S. & Ban, C. Aptamers and their biological applications. *Sensors* **12**, 612-631 (2012).

180 Numnuam, A. *et al.* Aptamer-based potentiometric measurements of proteins using ion-selective microelectrodes. *Analytical chemistry* **80**, 707-712 (2008).

181 Xu, D. *et al.* Label-free electrochemical detection for aptamer-based array electrodes. *Analytical chemistry* **77**, 5107-5113 (2005).

182 Stojanovic, M. N., de Prada, P. & Landry, D. W. Fluorescent sensors based on aptamer self-assembly. *Journal of the American Chemical Society* **122**, 11547-11548 (2000).

183 Liss, M., Petersen, B., Wolf, H. & Prohaska, E. An aptamer-based quartz crystal protein biosensor. *Analytical chemistry* **74**, 4488-4495 (2002).

184 Wang, L. *et al.* Label-free, regenerative and sensitive surface plasmon resonance and electrochemical aptasensors based on graphene. *Chem. Commun.* **47**, 7794-7796 (2011).

185 Hwang, K. S. *et al.* Nanomechanical microcantilever operated in vibration modes with use of RNA aptamer as receptor molecules for label-free detection of HCV helicase. *Biosensors and Bioelectronics* **23**, 459-465 (2007).

186 Min, K. *et al.* Dual-aptamer-based delivery vehicle of doxorubicin to both PSMA (+) and PSMA (−) prostate cancers. *Biomaterials* **32**, 2124-2132 (2011).

187 Chu, T. C., Twu, K. Y., Ellington, A. D. & Levy, M. Aptamer mediated siRNA delivery. *Nucleic acids research* **34**, e73-e73 (2006).

188 Nimjee, S. M., Rusconi, C. P. & Sullenger, B. A. Aptamers: an emerging class of therapeutics. *Annu. Rev. Med.* **56**, 555-583 (2005).

189 Bock, L. C., Griffin, L. C., Latham, J. A., Vermaas, E. H. & Toole, J. J. Selection of single-stranded DNA molecules that bind and inhibit human thrombin. (1992).

190 Ng, E. W. *et al.* Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nature reviews drug discovery* **5**, 123-132 (2006).

191 Nowak, J. Z. Age-related macular degeneration (AMD): pathogenesis and therapy. *Pharmacol rep* **58**, 353-363 (2006).

192 Sundaram, P., Kurniawan, H., Byrne, M. E. & Wower, J. Therapeutic RNA aptamers in clinical trials. *European Journal of Pharmaceutical Sciences* **48**, 259-271 (2013).

193 Krylova, S. M. *et al.* DNA aptamers for as analytical tools for the quantitative analysis of DNA-dealkylating enzymes. *Analytical Biochemistry* **414**, 261-265 (2011).

194 Lee, J. F., Hesselberth, J. R., Meyers, L. A. & Ellington, A. D. Aptamer database. *Nucleic acids research* **32**, D95-D100 (2004).

195 Gopinath, S. C. B. Methods developed for SELEX. *Analytical and bioanalytical chemistry* **387**, 171-182 (2007).

196 Irvine, D., Tuerk, C. & Gold, L. SELEXION: Systematic evolution of ligands by exponential enrichment with integrated optimization by non-linear analysis. *Journal of molecular biology* **222**, 739-761 (1991).

197 Vant-Hull, B., Payano-Baez, A., Davis, R. H. & Gold, L. The mathematics of SELEX against complex targets. *Journal of molecular biology* **278**, 579-597 (1998).

198 Chen, H. & Gold, L. Selection of high-affinity RNA ligands to reverse transcriptase: inhibition of cDNA synthesis and RNase H activity. *Biochemistry* **33**, 8746-8756 (1994).

199 Krylov, S. N. Kinetic CE: Foundation for homogeneous kinetic affinity methods. *Electrophoresis* **28**, 69-88 (2007).

200 Drabovich, A. P., Berezovski, M., Okhonin, V. & Krylov, S. N. Selection of smart aptamers by methods of kinetic capillary electrophoresis. *Analytical chemistry* **78**, 3171-3178 (2006).

201 Berezovski, M. *et al.* Nonequilibrium capillary electrophoresis of equilibrium mixtures: a universal tool for development of aptamers. *Journal of the American Chemical Society* **127**, 3165-3171 (2005).

202 Krylov, S. N. Nonequilibrium capillary electrophoresis of equilibrium mixtures (NECEEM): A novel method for biomolecular screening. *Journal of biomolecular screening* **11**, 115-122 (2006).

203 Kasahara, Y., Irisawa, Y., Ozaki, H., Obika, S. & Kuwahara, M. 2', 4'-BNA/LNA aptamers: CE-SELEX using a DNA-based library of full-length 2'-O,4'-C-methylene-bridged/linked bicyclic ribonucleotides. *Bioorganic & medicinal chemistry letters* **23**, 1288-1292 (2013).

204 Yu, X. & Yu, Y. A Mathematical Analysis of the Selective Enrichment of NECEEM-Based Non-SELEX. *Applied biochemistry and biotechnology* **173**, 2019-2027 (2014).

205 Berezovski, M., Musheev, M., Drabovich, A. & Krylov, S. N. Non-SELEX selection of aptamers. *Journal of the American Chemical Society* **128**, 1410-1411 (2006).

206 Krylov, S. N. & Berezovski, M. Non-equilibrium capillary electrophoresis of equilibrium mixtures—appreciation of kinetics in capillary electrophoresis. *Analyst* **128**, 571-575 (2003).

207 Musheev, M. U. & Krylov, S. N. Selection of aptamers by systematic evolution of ligands by exponential enrichment: addressing the polymerase chain reaction issue. *Analytica chimica acta* **564**, 91-96 (2006).

208 Kanagawa, T. Bias and artifacts in multitemplate polymerase chain reactions (PCR). *Journal of Bioscience and Bioengineering* **96**, 317-323 (2003).

209 Kanoatov, M. & Krylov, S. N. DNA Adsorption to the Reservoir Walls Causing Irreproducibility in Studies of Protein-DNA Interactions by Methods of Kinetic Capillary Electrophoresis. *Analytical chemistry* **83**, 8041-8045 (2011).

210 Prince, J. A. *et al.* Robust and accurate single nucleotide polymorphism genotyping by dynamic allele-specific hybridization (DASH): design criteria and assay validation. *Genome research* **11**, 152-162 (2001).

211 Ogino, S. & Wilson, R. B. Quantification of PCR Bias Caused by a Single Nucleotide Polymorphism in< i> SMN</i> Gene Dosage Analysis. *The Journal of molecular diagnostics* **4**, 185-190 (2002).

212 Nguyen, H.-K. & Southern, E. M. Minimising the secondary structure of DNA targets by incorporation of a modified deoxynucleoside: implications for nucleic acid analysis by hybridisation. *Nucleic acids research* **28**, 3904-3909 (2000).

213 Gamper, H. B., Cimino, G. D. & Hearst, J. E. Solution hybridization of crosslinkable DNA oligonucleotides to bacteriophage M13 DNA: Effect of secondary structure on hybridization kinetics and equilibria. *Journal of molecular biology* **197**, 349-362 (1987).

214 Fredman, D., Jobs, M., Strömqvist, L. & Brookes, A. DFold: PCR design that minimizes secondary structure and optimizes downstream genotyping applications. *Human mutation* **24**, 1-8 (2004).

215 Savory, N. *et al.* Selection of DNA aptamers against uropathogenic *Escherichia coli* NSM59 by quantitative PCR controlled Cell-SELEX. *Journal of microbiological methods* **104**, 94-100 (2014).

216 Gu, G., Wang, T., Yang, Y., Xu, X. & Wang, J. An Improved SELEX-Seq Strategy for Characterizing DNA-Binding Specificity of Transcription Factor: NF-κB as an Example. *PloS one* **8**, e76109 (2013).

217 Schneider, J. *et al.* Systematic analysis of T7 RNA polymerase based in vitro linear RNA amplification for use in microarray experiments. *Bmc Genomics* **5**, 29 (2004).

218 Tsuji, S. *et al.* Effective isolation of RNA aptamer through suppression of PCR bias. *Biochemical and biophysical research communications* **386**, 223-226 (2009).

219 Nakano, M. *et al.* Single-molecule PCR using water-in-oil emulsion. *Journal of biotechnology* **102**, 117-124 (2003).

220 Shao, K. *et al.* Emulsion PCR: A high efficient way of PCR amplification of random DNA libraries in aptamer selection. *PloS one* **6**, e24910 (2011).

221 Schütze, T. *et al.* A streamlined protocol for emulsion polymerase chain reaction and subsequent purification. *Analytical biochemistry* **410**, 155-157 (2011).

222 Hünniger, T., Wessels, H., Fischer, C., Paschke-Kratzin, A. & Fischer, M. just in time-Selection: A rapid semi-automated SELEX of DNA aptamers using magnetic separation and BEAMing. *Analytical Chemistry* (2014).

223 Hegi, M. E. *et al.* Clinical trial substantiates the predictive value of O-6-methylguanine-DNA methyltransferase promoter methylation in glioblastoma patients treated with temozolomide. *Clinical Cancer Research* **10**, 1871-1874 (2004).

224 Hegi, M. E. *et al.* MGMT gene silencing and benefit from temozolomide in glioblastoma. *New England Journal of Medicine* **352**, 997-1003 (2005).

225 Sharma, S. *et al.* Role of MGMT in tumor development, progression, diagnosis, treatment and prognosis. *Anticancer research* **29**, 3759-3768 (2009).

226 Urdinguio, R. G., Sanchez-Mut, J. V. & Esteller, M. Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. *The Lancet Neurology* **8**, 1056-1072 (2009).

227 Jenison, R. D., Gill, S. C., Pardi, A. & Polisky, B. High-resolution molecular discrimination by RNA. *Science* **263**, 1425-1429 (1994).

228 Yang, P. *et al.* Multiplexed detection of protein-peptide interaction and inhibition using capillary electrophoresis. *Analytical chemistry* **79**, 1690-1695 (2007).

229 Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nature biotechnology* **26**, 1135-1145 (2008).

230 Dabney, J. & Meyer, M. Length and GC-biases during sequencing library amplification: a comparison of various polymerase-buffer systems with ancient and modern DNA sequencing libraries. *Biotechniques* **52**, 87-94 (2012).

231 Berezovski, M., Nutiu, R., Li, Y. & Krylov, S. N. Affinity analysis of a protein-aptamer complex using nonequilibrium capillary electrophoresis of equilibrium mixtures. *Analytical Chemistry* **75**, 1382-1386 (2003).

232 Marimuthu, C., Tang, T.-H., Tominaga, J., Tan, S.-C. & Gopinath, S. C. Single-stranded DNA (ssDNA) production in DNA aptamer generation. *Analyst* **137**, 1307-1315 (2012).

233 Mosing, R. K., Mendonsa, S. D. & Bowser, M. T. Capillary electrophoresis-SELEX selection of aptamers with affinity for HIV-1 reverse transcriptase. *Analytical Chemistry* **77**, 6107-6112 (2005).

234 Mendonsa, S. D. & Bowser, M. T. In vitro selection of aptamers with affinity for neuropeptide Y using capillary electrophoresis. *Journal of the American Chemical Society* **127**, 9382-9383 (2005).

235 Clifton, I. J. *et al.* Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. *Journal of inorganic biochemistry* **100**, 644-669 (2006).

236 McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. & Schofield, C. J. Structural studies on human 2-oxoglutarate dependent oxygenases. *Current opinion in structural biology* **20**, 659-672 (2010).

237 Hicke, B. J. & Stephens, A. W. Escort aptamers: a delivery service for diagnosis and therapy. *The Journal of clinical investigation* **106**, 923-928 (2000).

238 Musheev, M. U., Filiptsev, Y. & Krylov, S. N. Noncooled Capillary Inlet: A Source of Systematic Errors in Capillary-Electrophoresis-Based Affinity Analyses. *Analytical chemistry* **82**, 8637-8641 (2010).

239 Krylova, S. M., Dove, P. M., Kanoatov, M. & Krylov, S. N. Slow-Dissociation and Slow-Recombination Assumptions in Nonequilibrium Capillary Electrophoresis of Equilibrium Mixtures. *Analytical chemistry* **83**, 7582-7585 (2011).

240 Welford, R. W. *et al.* Incorporation of oxygen into the succinate co-product of iron (II) and 2-oxoglutarate dependent oxygenases from bacteria, plants and humans. *FEBS letters* **579**, 5170-5174 (2005).

241 Falnes, P., Klungland, A. & Alseth, I. Repair of methyl lesions in DNA and RNA by oxidative demethylation. *Neuroscience* **145**, 1222-1232 (2007).

242 Mishina, Y., Duguid, E. M. & He, C. Direct reversal of DNA alkylation damage. *Chemical reviews* **106**, 215-232 (2006).

243 Shimada, K. *et al.* ALKBH3 contributes to survival and angiogenesis of human urothelial carcinoma cells through NADPH oxidase and tweak/Fn14/VEGF signals. *Clinical Cancer Research* **18**, 5247-5255 (2012).

244 Shan, X., Tashiro, H. & Lin, C.-l. G. The identification and characterization of oxidized RNAs in Alzheimer's disease. *The Journal of neuroscience* **23**, 4913-4921 (2003).

245 Beranek, D. T. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **231**, 11-30 (1990).

246 Margison, G. P., Koref, M. F. S. & Povey, A. C. Mechanisms of carcinogenicity/chemotherapy by O6-methylguanine. *Mutagenesis* **17**, 483-487 (2002).

247 Dinovo, E., Miyada, D. & Nakamura, R. Evaluation of direct and indirect coupled enzyme assay systems for measurement of creatine kinase activity. *Clinical chemistry* **19**, 994-997 (1973).

248 Yufa, R. *et al.* Emulsion PCR Significantly Improves NECEEM-based Aptamer Selection, Allowing for Efficient and Rapid Selection of Aptamer to Unmodified ABH2 Protein. *Analytical chemistry* **87**, 1411-1419 (2015).

249 Griffin, L. C., Tidmarsh, G. F., Bock, L. C., Toole, J. J. & Leung, L. In vivo anticoagulant properties of a novel nucleotide-based thrombin inhibitor and demonstration of regional anticoagulation in extracorporeal circuits. *Blood* **81**, 3271-3276 (1993).

250 Pieken, W. A., Olsen, D. B., Benseler, F., Aurup, H. & Eckstein, F. Kinetic characterization of ribonuclease-resistant 2'-modified hammerhead ribozymes. *Science* **253**, 314-317 (1991).

251 Green, L. S. *et al.* Nuclease-resistant nucleic acid ligands to vascular permeability factor/vascular endothelial growth factor. *Chem Biol* **2**, 683-695 (1995).

252 Obika, S. *et al.* Stability and structural features of the duplexes containing nucleoside analogues with a fixed N-type conformation, 2′-O, 4′-C-methyleneribonucleosides. *Tetrahedron letters* **39**, 5401-5404 (1998).

253 Burmeister, P. E. *et al.* Direct in vitro selection of a 2'-O-methyl aptamer to VEGF. *Chem Biol* **12**, 25-33, doi:10.1016/j.chembiol.2004.10.017 (2005).

254 Van Dyck, S., Vissers, S., Van Schepdael, A. & Hoogmartens, J. Kinetic study of angiotensin converting enzyme activity by capillary electrophoresis after in-line reaction at the capillary inlet. *Journal of Chromatography A* **986**, 303-311 (2003).

255 Glatz, Z. Determination of enzymatic activity by capillary electrophoresis. *Journal of Chromatography B* **841**, 23-37 (2006).