DOCTOR OF PHILOSOPHY

A ‘bump-&-hole’ approach for engineering allele-selective inhibition of the BET bromodomains

Runcie, Andrew C.

Award date:
2019

A ‘bump-&-hole’ approach for engineering allele-selective inhibition of the BET bromodomains

Andrew Colin Runcie

Supervisor:
Professor Alessio Ciulli

Thesis submitted for the degree of
Doctor of Philosophy
School of Life Sciences
University of Dundee
2019

To my wife Ketsuda

Acknowledgements

I would first like to thank my supervisor Prof. Alessio Ciulli for giving me this opportunity. Thank you for your advice, encouragement and fruitful discussions. Thank you for providing me with the resources to pursue my own ideas and fuel my growth as a researcher.

Many thanks to the AC lab members, past and present, for their help and guidance. Dr Kwok-Ho Chan for help with protein purification, X-ray crystallography and tissue culture. Dr Michael Zengerle and Andrea Test for help with chemistry. Dr Morgan Gadd for guidance with crystallography.

Large parts of this work relied heavily on interdisciplinary nature of the School of Life Sciences. Thank you to Lorna Campbell and Sandra O'Neill, from the DDU's David Gray lab, for their assistance. Also from the DDU, Dr Ola Epemolu and Dr Lucy Ellis provided DMPK data invaluable to evaluation of our compounds. Thank you to Dr Sam Swift and the Light Microscope Facility for their help with fluorescent microscopy. In addition, reagents crucial to this research were provided by the Division of Signal Transduction Therapy and the Structural Genomics Consortium.

Thank you to the BBSRC and to the EastBio group for funding and training, and specifically to Maria Filippakopoulou, Dr Caroline Pope and Caroline Proctor. As a part of the EastBio DTP programme I carried out a 3-month internship at Collagen Solutions Plc in Glasgow, and I would like to thank my supervisors Dr Grahame Busby of Collagen Solutions and Alex Muhlholzl of Mikota for their support and guidance. Thank you as well to my colleagues Leigh Adamson and Steven Lee for their help with training me and making me feel at home in a new lab.

This work also relied on the work of many School of Life Science staff, and apologies to those not specifically named. Thank you to the BCDD lab managers Shona McInroy and Tracey Norris, and PhD administrative staff Lesley Coats, Gail Guild and Nikki Wilson. Many thanks to MRC tissue culture staff, managed by Dr Laura Finn.

Finally, special thanks to my wife Ketsuda, my family and my friends for their love and support throughout these 4 years. It would not have been possible without you.

Declaration

I declare that the work described in this thesis was carried out by me under the supervision of Professor Alessio Ciulli and, unless otherwise stated, all references cited have been consulted by me. To my knowledge, the work of which this thesis is a record of is original and has not been submitted for any other degree at this or another university.

Andrew Colin Runcie

August 2018

Publications

The work described in this thesis led to the following publications:

1. Runcie AC, Chan KH, Zengerle M & Ciulli A. *Chemical genetics approaches for selective intervention in epigenetics*. Current Opinion in Chemical Biology **2016**, 33, 186-194.

2. Runcie AC, Zengerle M, Chan KH, Testa A, van Beurden L, Baud MGJ, Epemolu O, Ellis LCJ, Read KD, Coulthard V, Brien A & Ciulli A. *Optimization of a "bump-and-hole" approach to allele-selective BET bromodomain inhibition*. Chemical Science **2018,** 9, 2452-2468

3. Testa A, Lucas X, Castro GV, Chan KH, Wright JE, Runcie AC, Gadd MS, Harrison WTA, Ko EJ, Fletcher D & Ciulli A. *3-Fluoro-4-hydroxyprolines: Synthesis, conformational analysis and stereoselective recognition by the VHL E3 ubiquitin ligase for targeted protein degradation.* Journal of the American Chemical Society **2018**, 140, 9299-9313

Abstract

Chemical tools can be used to investigate the function of proteins and their importance to biological processes and disease. This approach has many advantages over classical genetic tools, but requires compounds with high selectivity for their targets. When single-target selectivity cannot be achieved through normal means it can be engineered through the development of orthogonal protein:ligand pairs, wherein target proteins are mutated and an existing ligand is modified to become mutant-selective.

An ideal model for this engineered, allele-specific chemical genetic work is the BET family of bromodomains – 8 closely related, and structurally conserved, epigenetic acetyl-lysine reader domains that so far has defied single-bromodomain inhibition. The Alessio Ciulli lab has previously designed a proof-of-concept 'bump-&-hole' system to selectively target these bromodomains by mutating a conserved leucine to an alanine (L/A mutation) and adding an ethyl bump to an existing inhibitor scaffold (ET compound). This system was promising, but required optimisation of the mutation and modified compound before it could be used to answer biological questions.

Here the L/A mutation is replaced with a more conservative leucine/valine (L/V) mutation, which is shown to be significantly more functional with regards to acetylated chromatin binding and downstream gene expression. Simultaneously, a large number of new modifications were incorporated into a library of bumped compounds, which were then screened for their selectivity for L/V BET bromodomains (over WT). This lead to the selection of '9-ME-1' as the preferred L/V-selective compound.

This optimised BET bump-&-hole system was validated through a number of cellular assays. By repurposing some of these cellular assays the bump-&-hole system was then used to answer biological questions – specifically the relative importance of different BET bromodomains to BET protein function. These results, in connection to recent advances in the literature, provides evidence for a new model of how the BET bromodomains work together to drive downstream gene expression.

Finally, the need to use CRISPR/Cas9 gene-editing technology to generate L/V-mutant cell-lines was identified, and an attempt was made to knock-in the L/V mutation to the first bromodomain of BRD4 in a cancer cell-line.

Abbreviations

A_3: adenosine A3 receptor, A.K.A. ADORA3

AA: amino-acid

Ac-CoA: acetyl-coenzyme A

Alpha: amplified luminescent proximity homogenous assay

AML: acute myeloid leukaemia

ApoA1: apolipoprotein A1

AU: absorption units

BCA assay: bicinchoninic acid assay

BD: bromodomain

BD1: N-terminal BET bromodomain

BD2: C-terminal BET bromodomain

BET: bromodomain & extra-terminal domain

BLI: bio-layer interferometry

BPPM: bio-orthogonal profiling of protein methylation

BRD4: bromodomain containing protein 4

BRD3: bromodomain containing protein 3

BRD2: bromodomain containing protein 2

BRDT: bromodomain testis-specific protein

BRET: bioluminescence resonance energy transfer

BSA: bovine serum albumin

Cas: CRISPR-associated system

CETSA: cellular thermal shift assay

CHAPS: 3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate

CHD: chromodomain

ChIP-Seq: chromatin immunoprecipitation, with massively parallel DNA sequencing

CliEn-seq: clickable chromatin enrichment with parallel DNA sequencing

CL_{int}: intrinsic clearance

CLogP: predicted (via group contribution method) log partition-coefficient of a compound between octanol and H_2O

CRBN: cereblon

CRISPR: clustered regularly interspaced short palindromic repeats

CRPC: Castration resistant prostate cancer

CTCL: cutaneous T-cell lymphoma

CTD: C-terminal domain

CYP: cytochrome P450 proteins

DAPI: 4',6-diamidino-2-phenylindole, dihydrochloride

DLBCL: diffuse large B-cell lymphoma

DMEM: Dulbecco's modified eagle medium

DMPK: drug metabolism and pharmacokinetics

DNMT: DNA methyltransferase enzyme

DSF: differential scanning fluorimetry

E. coli: Escherichia coli

eGFP: enhanced GFP, with S65T and F64L mutations and human codon optimisation

ET: extra-terminal domain

EtOH: ethanol

EMA: European medicines agency

FACS: fluorescence-activated cell sorting

FBS: foetal bovine serum

FDA: food and drug administration

FEN: flap endonuclease

FP: fluorescence polarisation

FRAP: fluorescence recovery after photobleaching

FRET: Förster resonance energy transfer

ΔG: Gibbs free energy

GATA1: GATA-binding factor 1. A.K.A. erythroid transcription factor

gDNA: genomic DNA

GFP: green fluorescent protein, from *Aequorea victoria*

GPCR: G-protein-coupled receptor

ΔH: system change in enthalpy

HAT: histone acetyltransferase

HDAC: histone deacetylase

HDL: high-density lipoprotein

HDR: homology-directed repair

HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

HKMT: histone lysine methyltransferase

Interleukin 6: IL-6

IPTG: isopropyl-β-D-thiogalactopyranoside

ITC: isothermal titration calorimetry

Kac: acetyl-lysine

KAT: protein lysine acetyltransferase

k_{cat}: enzyme catalytic rate (number of substrate molecules converted to product over time)

K_d: dissociation constant

KI: knock-in

K_m: concentration of enzyme substrate allowing for half maximal reaction rate

KO: knock-out

LAC: lung adenocarcinoma

LB: lysogeny broth

LC-MS: liquid chromatography-mass spectrometry

MBD: methyl-CpG-binding domain

MBT domains: malignant brain tumor domains

MeOH: methanol

MM: multiple myeloma

MMEJ: microhomology-mediated end joining

MWCO: molecular weight cut-off

MT_1: melatonin receptor type 1A

NADPH: nicotinamide adenine dinucleotide phosphate

NF-κB: nuclear factor kappa-light-chain-enhancer of activated B cells

NF-κB-RE: NF-κB response element

NHEJ: non-homologous end joining

NHP: non-histone protein

NK_2: substance-K receptor 2

NMC: NUT-midline carcinoma

NUT: nuclear protein in testes

OD_{600}: optical density at 600 nm

PAMPA: parallel artificial membrane permeability assay

PBS: phosphate-buffered saline

PCR: polymerase chain reaction

PDB: protein data bank

Pe: effective permeability coefficient

PEG: polyethylene glycol

PEST: proline, glutamic acid, serine & threonine rich sequence

PHD: plant homeodomain

PITCH: precise integration into target chromosome

PKMT: protein lysine methyltransferase

POI: protein of interest

PPI: protein-protein interaction

PRC2: polycomb repressive complex

PROTAC: proteolysis targeting chimera

PTCL: peripheral T-cell lymphoma

P-TEFb: positive transcription elongation factor

PTM: post-translational modification

RelA: NF-κB p-65 subunit

RFU: relative fluorescent units

RMSD: root mean square deviation

RNAi: RNA interference

RNP: ribonucleoprotein

rpm: rotations per minute

RPMI: media developed at Roswell Park Memorial Institute

ΔS: system change in entropy

SAH: S-adenosyl-L-homocysteine

SAHA: suberanilohydroxamic acid. A.K.A. vorinostat

SAM: S-adenosyl-l-methionine

SDS: sodium dodecyl sulphate

SDS-PAGE: sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC: size exclusion chromatography

SFC: supercritical fluid chromatography

SGC: structural genomics consortium

shRNA: short hairpin RNA

siRNA: silencing RNA

SNPs: single-nucleotide polymorphisms

SPR: surface plasmon resonance

SUMO: small ubiquitin-like modifier

Sy.x: standard deviation of residuals. Calculated as square root of (sum of squared residuals, divided by n-K), where K is number of parameters fit by regression

t½: half-life (RE: stability) or recovery-time (RE: FRAP)

TBS: tris-buffered saline

TBST: TBS with 0.1% Tween-20

TEV: tobacco etch virus

T_m: melting temperature

TNF-α: tumour necrosis factor alpha

TR-FRET: time-resolved FRET

UPLC-MS/MS: ultra-performance liquid chromatography-tandem mass spectrometry

UPS: ubiquitin proteasome system

VHL: von Hippel–Lindau protein

WPF shelf: W81, P82, F83 residues (BRD4 BD1 numbering)

WT: wild type

Z': Z-factor

Contents

Chapter 5: Advanced Biology

Chapter 6: Discussion

Chapter 7: Conclusions

Chapter 8: Materials & Methods

List of Figures

List of Tables

List of Schemes

Chapter 1
Introduction

1.1 – Chemical Genetics:

Chemical genetics is the use of biologically active 'chemical probe' compounds (typically small-molecule inhibitors) to investigate the role, and elucidate the function, of genes and their encoded proteins [1-3]. Thus, chemical genetics parallels conceptually classical genetics in which genetic tools such as RNA interference (RNAi), gene mutations and knock-outs (KOs) are used to study genes and gene products. While classical genetics typically interrupts the gene – protein – phenotype system at the DNA or RNA level chemical genetics acts through disruption of the protein stage (figure 1.1). There is a great deal of overlap between chemical probes and therapeutic drugs, with many compounds being used as both. However, the focus of therapeutic drugs is to be efficacious in select indications, which means they may not necessarily have to be overly selective or even have a fully characterised mode of action. In contrast to therapeutic drugs, chemical probes must be thoroughly characterised and selective, but there is less need for high quality drug metabolism and pharmacokinetics (DMPK) characteristics and safety profiles as they are typically used *in vitro* [4].

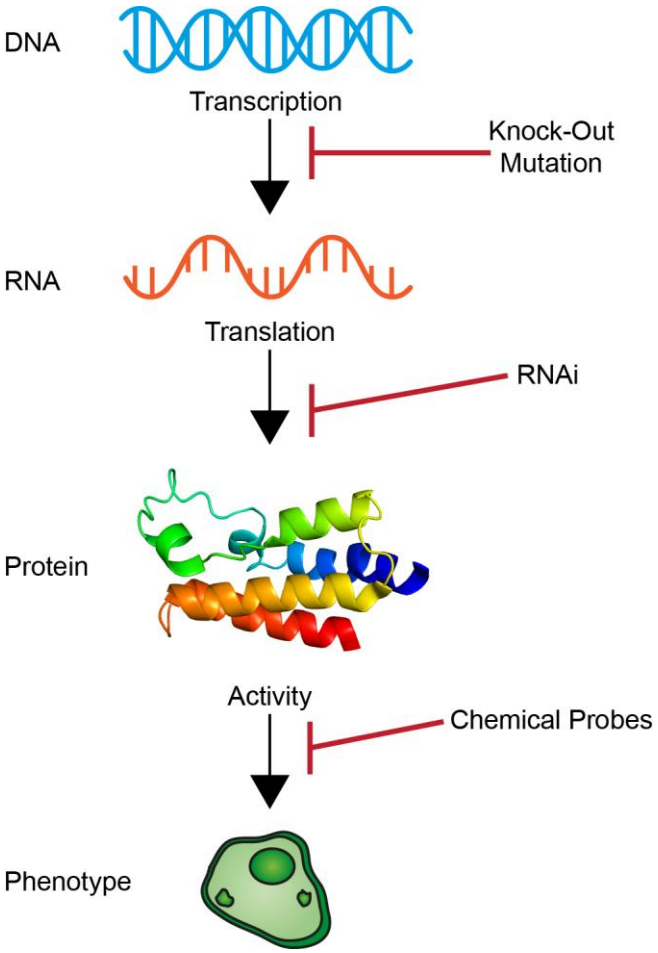


Figure 1.1. 'The Central Dogma' and genetic/chemical tools.

Protein is crystal structure of BRD4 BD1 (PDB: 3UVW).

Chemical and classical genetics can both be performed in a ‘forward’ or ‘reverse’ manner [3, 5] (figure 1.2). In forward genetics a library of small molecules (or random RNAi/mutations) is screened against cell-line(s) of interest. When a molecule or mutation is found to elicit a phenotype of interest it undergoes further studies to identify the molecule’s target protein, or the gene affected by the mutation. Such processes are typically referred to as ‘phenotypic screens’. In reverse genetics a specific protein or gene of interest is already identified, and small-molecules or mutations are deliberately designed to target it. These chemical or genetic tools are then used in a variety of biological contexts to investigate the phenotype induced by their disruption of the gene/protein of interest. In practice the forward/reverse dichotomy is not always so distinct, and often important gene or protein families will be identified first through forward genetics, then for additional tools to be developed in a reverse genetics manner.

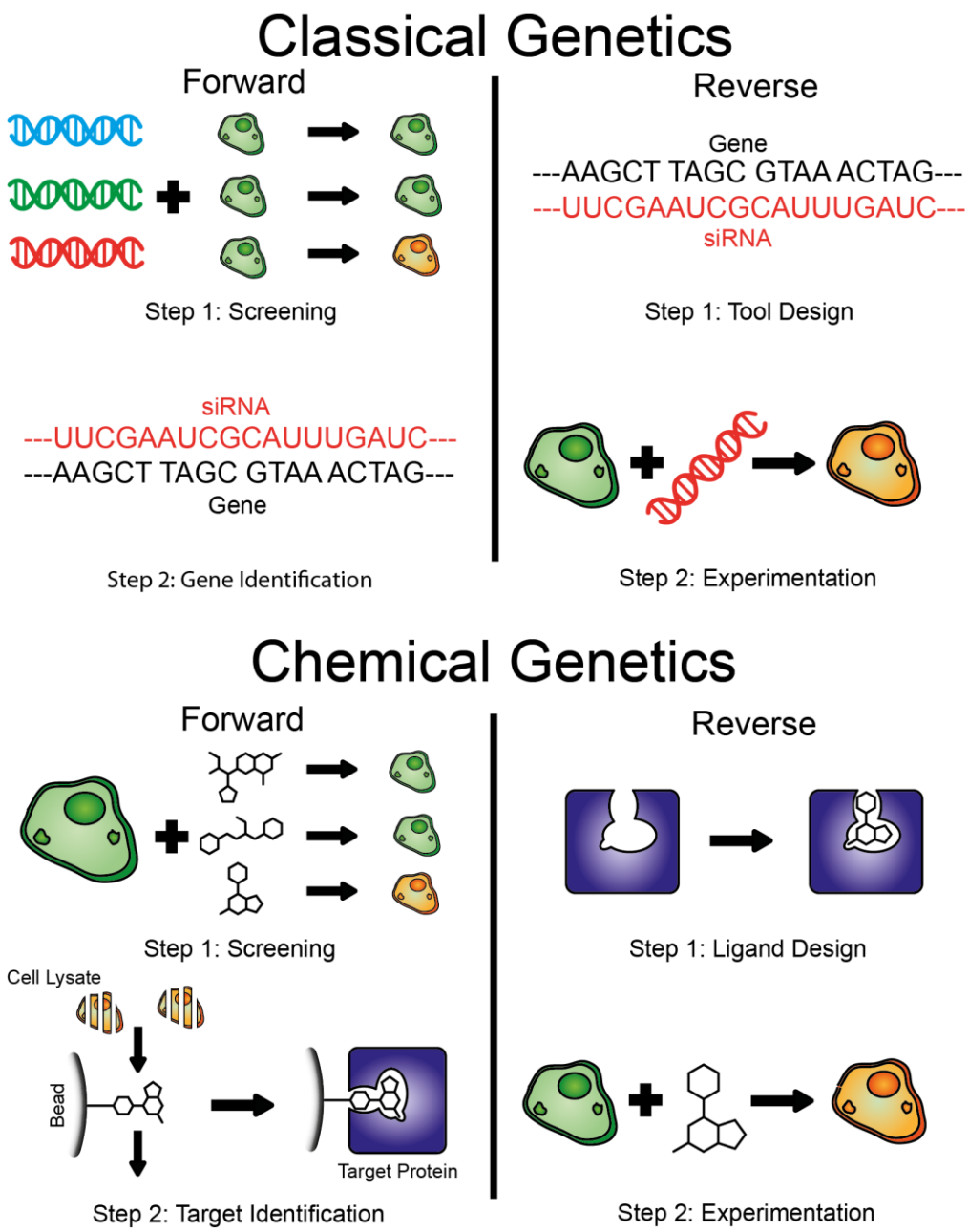


Figure 1.2. Schematic and comparison between classical & chemical genetics.

In forward genetics random compounds, mutations or RNAi tools are screened, and the target gene/protein of any hits are identified. In reverse genetics a target gene/protein is known, and genetic or chemical tools custom designed towards it. Modified from Runcie et al, COCB, 2016 [1].

Both classical and chemical genetics have distinct advantages and disadvantages, with different techniques being preferable in different contexts [5]. Classical tools such as mutations and gene knock-outs can be very specific, only affecting the gene within which said alterations are made. Unfortunately such modifications are not necessarily tuneable, with 100% of the gene product affected (potentially 50% in diploid heterozygous systems), and hence one cannot observe dose-response effects and may face complications due to cell death or, in the case of

in vivo systems, potentially embryonic lethality following gene deletion. Additionally, such effects are uniform, and hence cannot be controlled temporally or spatially. Finally, introducing mutations into endogenous genes can be a very complicated affair, despite constant advances in gene editing technology. More complicated systems, such as the use of inducible systems such as tetracycline-controlled genes [6], can allow for some degree of 'dosage' and temporal control, although the compounds used to regulate expression may alter cell phenotypes themselves by affecting other proteins. RNA interference is considered a genetic tool, although faces a very different set of challenges [7]. RNAi tools do not require gene editing, and can be used in various doses and timings, but suffer from mixed efficacy (not completely removing gene products) and tend to cause off-target effects due to low selectivity. The CRISPR/Cas9 gene-editing system is an increasingly valuable genetic tool, which will be discussed fully in chapter 5 of this thesis.

Chemical genetics can offer a much greater degree of control than classical genetic tools, as doses can be applied at desired time points, and once the probe is washed out one may see its effect be reversed. When working with animal models one can try to limit the effect of the probe to specific tissues for a more nuanced experiment. Dose-response relationships can be observed, and the probe dose tuned to allow for experimentation without killing the cells or animals tested. While many genetic tools (gene KO, RNAi) work by removing the entire protein of interest, chemical tools will only inhibit specific functions or block individual binding sites. This can be very significant when working with multi-domain proteins where one may wish only to investigate the role of a specific domain's function, while not affecting other active domains or scaffolding roles. Chemical genetics is generally favourable to classical genetics when performing target validation, to determine if a specific protein can be inhibited for therapeutic effects. As therapeutics typically take the form of small-molecule inhibitors, chemical genetics is a much more accurate representation of their potential biological impact. The main drawback of chemical genetics is the need for well characterised and highly-selective inhibitors [1, 4]. While mutations will only affect the genes they are integrated into, small-molecules will often inhibit multiple structurally-related proteins. This can be very evident when working with strongly conserved protein families, which exhibit very high sequence identity and

conservation of the targeted binding site. Moreover, it is also not uncommon for small-molecules to affect multiple proteins with little genetic similarity and very different functions. It is also important to ensure that the chemical tools used are sufficiently potent to disturb the function of the protein of interest, and possesses the necessary DMPK characteristics to remain active in cellular or physiological contexts.

1.2 – Epigenetics:

Epigenetics generally lacks a universal definition, but here will refer to reversible and partially heritable chemical modifications made to DNA (that don't alter its sequence) and its associated chromatin components that serve to regulate genetic processes. Epigenetics is involved in the regulation of DNA repair, replication and recombination and, most importantly for our purposes, gene expression through transcription [8, 9]. These chemical modifications are referred to as 'epigenetic marks'. Epigenetics involves the interplay of three types of protein (figure 1.3): 1) writer enzymes (e.g. histone acetyltransferases HATs) add epigenetic marks to specific DNA or histone regions; 2) eraser enzymes (e.g. histone deacetylases HDACs) remove said marks; and 3) reader proteins (e.g. bromodomains BDs) recognise these marks, and in turn recruit additional proteins that induce or inhibit various important processes regarding DNA [9].

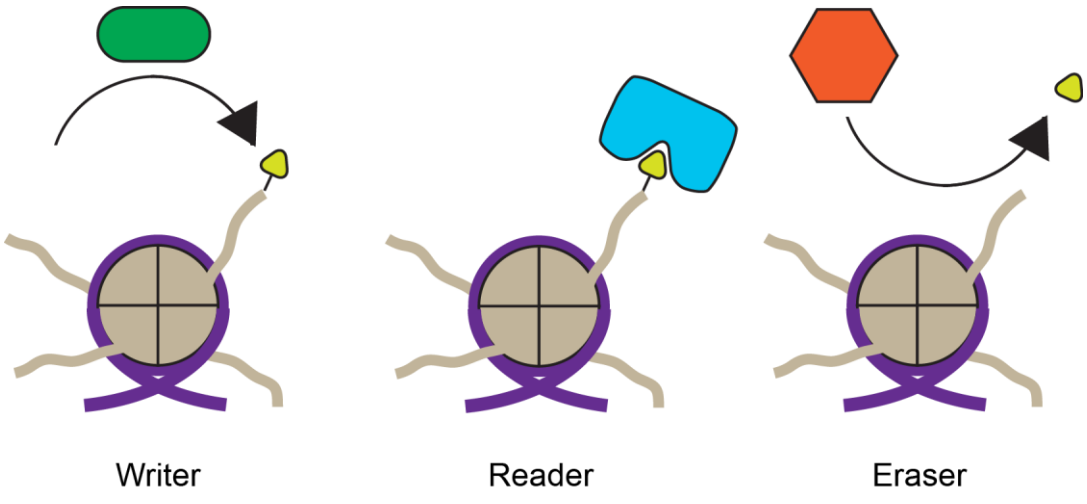


Figure 1.3. Epigenetic proteins.

Writer, reader and eraser proteins add, recognise and remove (respectively) epigenetic marks (yellow) from nucleosome (purple DNA & grey histone protein).

The main epigenetic modification of DNA itself is cytosine methylation. Conversion of cytosine to 5-methylcytosine by DNA methyltransferase (DNMT) enzymes is associated with a reduction in gene expression [10, 11]. This gene repression is believed to be caused in part by altering the physical properties of the DNA and hindering its binding to transcription complexes, and mainly by the recruitment of methyl-CpG binding domain (MBD) proteins, which in return recruit histone-modifying epigenetic proteins [10, 11]. In mammals C methylation mainly occurs at CpG sites, although can occur at CpApC sites in some progenitor cells. Aberrant hypo- or hyper-methylation at CpG sites has been associated with some cancers [10, 11].

The majority of epigenetic marks occur at the histone level. Inside the cell nucleus DNA is packaged into chromatin. Chromatin is, fundamentally, constructed of a series of 'nucleosomes' - ~146bp of DNA wrapped in a double super-helix around 8 histone proteins (2 copies each of H2A, H2B, H3 and H4) [12] (figure 1.4A). Each histone protein possesses an unstructured N-terminal 'tail', of 40–60 amino acids (AA) in length, that are accessible for a wide range of covalent and reversible post-translational modifications (PTMs) [12] (figure 1.4B). The first structure of the nucleosome was published in 1997 [13], while the exact manner in which nucleosomes and additional scaffolding proteins come together to form higher order structures is not fully understood. Areas of the genome packaged into isolated nucleosomes ('beads on a string') is referred to as euchromatin and is typically more transcriptionally active. The recruitment of the H1 histone protein, together with certain epigenetic marks, causes nucleosomes to coil into the transcriptionally silent '30 nm fibre', or heterochromatin, which again can further condense into chromosomes for cell division [14].

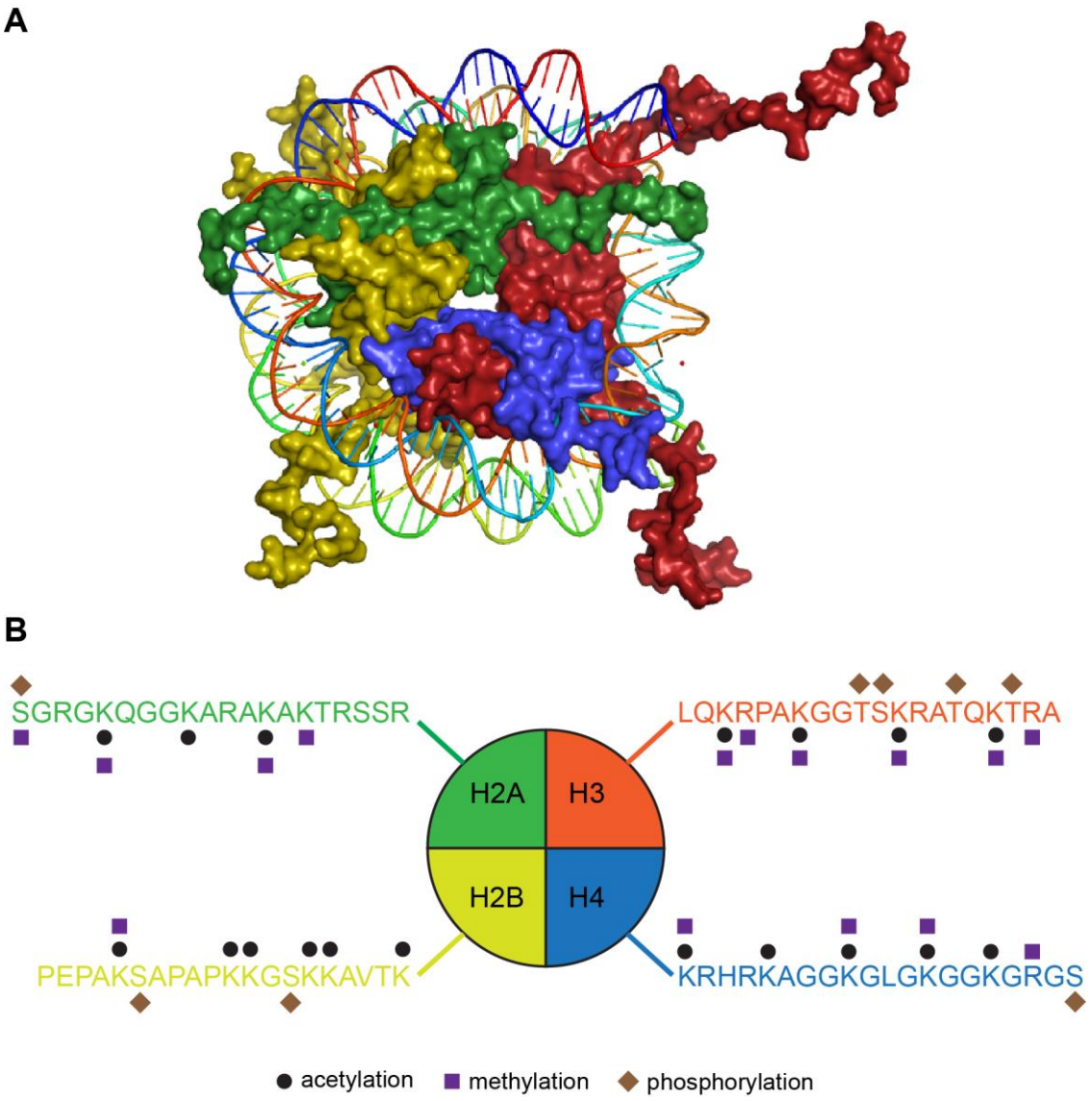


Figure 1.4. Nucleosome structure and histone tail modifications

A) X-ray crystal structure of nucleosome (PDB: 1KX5). B) Amino acid sequence of histone N-terminal tails (first 20 residues shown), with common epigenetic marks. DNA and second copy of each histone protein omitted for clarity. Colour scheme: green – H2A, yellow – H2B, red – H3, blue – H4.

The most common, and best understood, histone modifications are acetylation and methylation, although phosphorylation, ubiquitination, citrullination, SUMOylation and ADP-ribosylation are also known [15]. Histone modifications can directly regulate cellular processes, without recognition by reader proteins, by altering the structure of chromatin. Histone acetylation and phosphorylation can be seen as neutralising the positive charge of histone proteins, reducing their affinity for DNA and producing less compact euchromatin. There is evidence that this occurs following acetylation both of histone tail residues (such as H4 K16 [16]) and of globular residues (H3 K122 [17]).

Ubiquitination and SUMOylation may also induce large conformational changes that alter chromatin structure, due to the large size of the added modification [15].

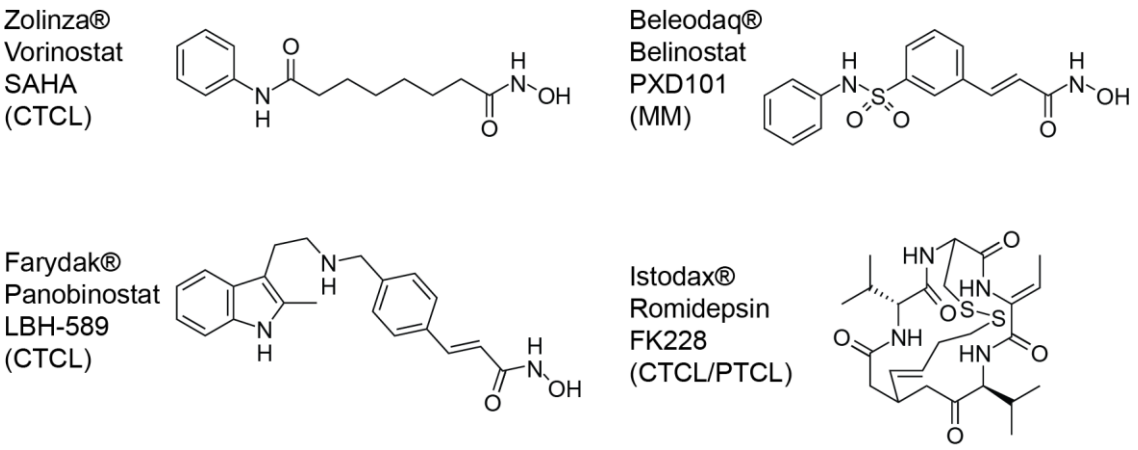
Lysine acetylation is a major epigenetic mark, installed by HAT enzymes, removed by HDACs and read by bromodomains. It is strongly associated with increased gene transcription and more open euchromatin, both for the structural reasons mentioned above but also due to the recruitment/activation of transcription machinery and chromatin remodelling complexes by bromodomain-containing proteins [15]. Lysine methylation is another important epigenetic mark. Lysine residues can be mono-, di- or tri-methylated, with each mark (as well as unmodified lysine) being recognised by different readers [18]. Broadly speaking, di- and tri-methylation are associated with reduced gene expression, while mono- and un-methylated lysine residues are associated with active expression, but even these trends are not absolute. Tri-methylated H3 K4, common at transcription start sites, can be recognised by 3 different types of reader – PHD (plant homeodomain), CHD (chromodomain) and Tudor – which generally leads to the recruitment of pro-transcription writer and eraser enzymes. Tri-methylated H3 K9 is strongly associated with heterochromatin and permanently silenced genes, and is thought to be recognised by CHD-proteins to recruit additional HKMT (histone lysine methyltransferase) writers to maintain heterochromatin status during DNA replication. Mono-methylated H3 K4 is found in abundance at active transcriptional enhancers, and is strongly associated with active gene expression [15, 18].

As many epigenetic marks have been clearly shown to influence gene expression and other processes it has been proposed that there exists a ‘histone-code’, wherein the regulation of a given gene may be determined by the exact combination of epigenetic marks around it and its promoter [12]. If such a precise code does exist it is likely to be highly complex, with each histone tail possessing dozens of potential epigenetic marks and an incredible array of conceivable combinations of marks. Furthermore, the effects of some marks are heavily context-dependent and influenced by inter-modification cross-talk [15]. One study identified ~4000 distinct combinations of histone modifications at CD+ T cell promoters. While some patterns could be observed in this study the majority of combinations (~3000) occurred at only a single promoter [19].

1.3 – Chemical Genetics for Epigenetics:

The development of chemical probes for epigenetic proteins has had a major impact on our understanding of their function and role in disease, although unfortunately the majority of epigenetic targets are still essentially unexplored [20]. The development of chemical probes has likely suffered somewhat from the reliance of epigenetic proteins on PPIs (protein:protein interactions), which are typically considered 'undruggable' due to their large, shallow and hydrophobic nature [21]. This has not been a total block on epigenetic chemical probe development, as the PPIs that epigenetic proteins are reliant upon typically involve well-defined and often deep binding pockets that have evolved to recognise very specific chemical modifications (e.g. acetylated lysine) or enzyme cofactors (e.g. SAM S-adenosyl-l-methionine). Additionally, due to the regulatory nature of epigenetics processes the modulation of any one gene will likely have many severe downstream impacts, complicating further analysis. This may make the need for selectivity even more stringent when working in epigenetics. The development of epigenetic chemical probes, either out of unbiased phenotypic screens or following genetic target validation, has had the effect of greatly increasing the level of research into the relevant target [20].

The most established example of epigenetic inhibitor design is that of HDAC eraser enzymes. Following a series of hits in phenotypic screens many HDAC inhibitors have now been developed, with 4 approved for use in the clinic for treatment of cancers such as lymphoma and multiple myeloma (MM) (scheme 1.1) [22-24]. The use of HDAC inhibitors in chemical genetics studies have shown that HDACs affect many biological responses. HDAC inhibition has been shown to alter the expression of 2-20% of genes, as well as alter chromatin structure, likely through fairly immediate effects on histone acetylation levels. Downstream functions include, but are not limited to, DNA repair, cell-cycle control, angiogenesis and metastasis [22-24]. The use of HDAC inhibitors has been somewhat hindered by the difficulty in developing inhibitors for all 18 HDACs, and also to ensure selective inhibition of specific enzymes [25].



Scheme 1.1. HDAC inhibitors in clinical use.

Structures of FDA-approved HDAC inhibitors, alongside commercial and chemical names. FDA approved indication in brackets: CTCL – cutaneous T-cell lymphoma, MM – multiple myeloma, PTCL – peripheral T-cell lymphoma.

The development of inhibitors for epigenetic readers has generally lagged behind that of writers and erasers, but recent work on BET bromodomains (discussed later) has shown that readers are viable therapeutic targets and has helped lead to the development of many reader domain inhibitors. This reader inhibitor development has been most successful with bromodomains, but there has also been a wealth of methyl-lysine reader inhibitor design [26, 27]. The Stephen Frye lab has developed several inhibitors of malignant brain tumor (MBT) domains, specifically targeting L3MBTL1 and L3MBTL3, with the most potent being UNC1215, which has an affinity of 120 nM [28]. In addition, several compounds with micromolar affinity for CHD and Tudor domains have been published, and several fragments have been identified with affinity for Pygo and BAZ PHD domains [29, 30].

1.4 – Bromodomains:

As mentioned previously, bromodomains are the main reader domain for lysine acetylation (Kac marks). Originally identified in the *Brahma*/*brm* gene of Drosophila in 1992 [31], 59 human bromodomains have now been identified, located within 43 proteins (11 proteins contain 2 BDs, 1 protein contains 6 BDs) (figure 1.5A). All bromodomains possess a structurally conserved fold comprising of 4 alpha helices (αZ, αA, αB, αC) linked by loops named ZA and BC [32] (figure 1.5B). Despite being a PPI (typically involving shallow and featureless areas of protein) bromodomains bind Kac-modified histone tails within a deep hydrophobic pocket, which typically contains a conserved asparagine residue that forms important hydrogen bonds to the Kac mark and water molecules [32]. This deep hydrophobic pocket has the effect of making some bromodomains unusually 'druggable' by the standards of PPIs, although there is variation within the family [33]. In addition to their Kac binding pocket bromodomains possess positively or negatively charged patches on their surface, which may act to help recognise more basic/acidic histone tails, or to interact directly with DNA [32, 34].

Following recognition of acetylated chromatin, bromodomain-containing proteins then bring about a number of potential biological activities. These activities can be mediated directly by other domains within the proteins [32], with BD-containing proteins also possessing helicase [35], HAT [36] and HKMT [37] domains, or through the recruitment of additional proteins or protein complexes. By recruiting or modulating transcription factors and other transcription machinery components BDs can, relatively directly, modulate the expression of select genes. Additionally, through the recruitment of chromatin remodelling complexes BDs can drive changes to the epigenetic labelling of histones and through that the packaging and accessibility of chromatin. By switching gene loci between heterochromatin and euchromatin states, through histone tail modification and nucleosome addition/removal, these complexes can further regulate gene expression as well as other processes such as DNA repair and replication. In several proteins BDs are paired with other epigenetic reader domains, such as the methyl-lysine recognising PHD domain [38], which may help to provide additional specificity with regards to recognition of specific regions of the genome and subsequent regulation of specific downstream genes.

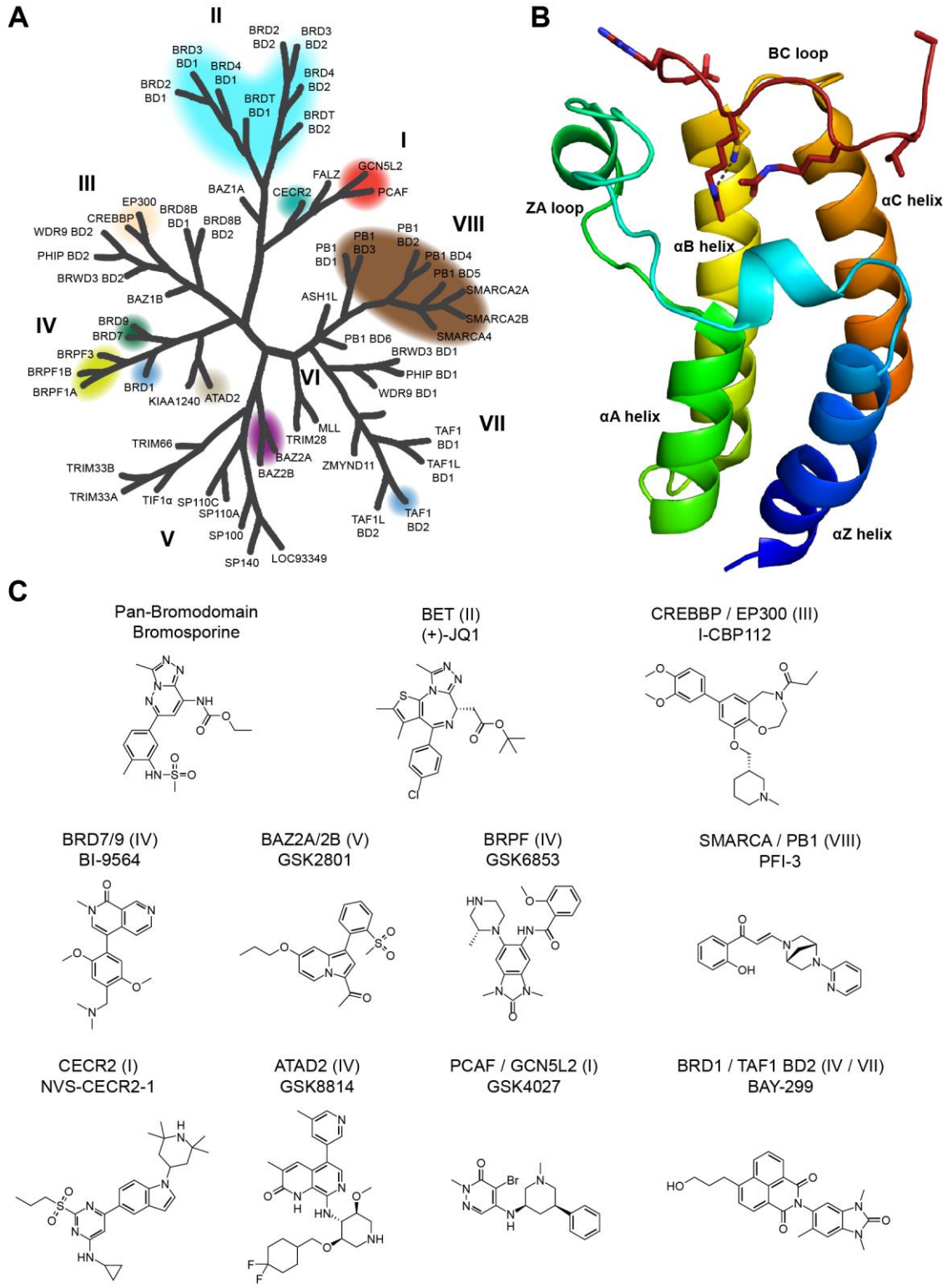


Figure 1.5. Human bromodomains, their structure & their inhibitors

A) Phylogenetic tree of human bromodomains, modified from Filippakopoulos et al, Cell, 2012 [32]. BDs targeted by chemical probe from [39] highlighted. BET subfamily (II) is highlighted in light blue. B) X-ray co-crystal structure of BRD4 BD1 bound to $H4_{(1\text{-}11)}$ K(5,8)ac (red). PDB: 3UVW. Critical hydrogen bond between Kac5 and N140 are shown with black dash. C) Chemical structures of a representative inhibitor for each group of targeted BDs in A. Subfamily of each group of BDs in brackets.

A number of bromodomain inhibitors have been developed (figure 1.5C) [40, 41], so far primarily being used as chemical probes to study protein function. To date, chemical probes exist for ~31 human bromodomains from all but one main branch of the bromodomain family, with selectivity varying from a single bromodomain to entire subfamilies [39] (figure 1.5A). In addition, a promiscuous bromodomain inhibitor, bromosporine, has been developed with binding observed against >20 human bromodomains [42].

1.5 – BET Proteins:

Some of the best studied bromodomains are those of the bromodomain and extra-terminal domain (BET) subfamily. There are 4 human BET proteins (BRD4, BRD3, BRD2 & BRDT), which each possess two tandem bromodomains (BD1 & BD2) (figure 1.6). The BET BDs are generally understood to recognise acetylated lysine residues in histone tails [32]. No single Kac pattern is associated with BET BDs, although H4 di- and tetra-acetylation is commonly used for model substrates (histone tail peptides and whole nucleosomes). The possibility of the BD1s and BD2s recognising different Kac marks has been raised, with the BD1s being suggested to prefer H4 histone tails while the BD2s preferring H3 tails [32] or being promiscuous. BET bromodomain prefer to bind pairs of nearby Kac marks, with the H4 K(5,8)ac pair (separated by two glycines) seemingly being optimal [32]. The BET proteins are known to possess homologues in mice [43], drosophila [44] and fungi [45].

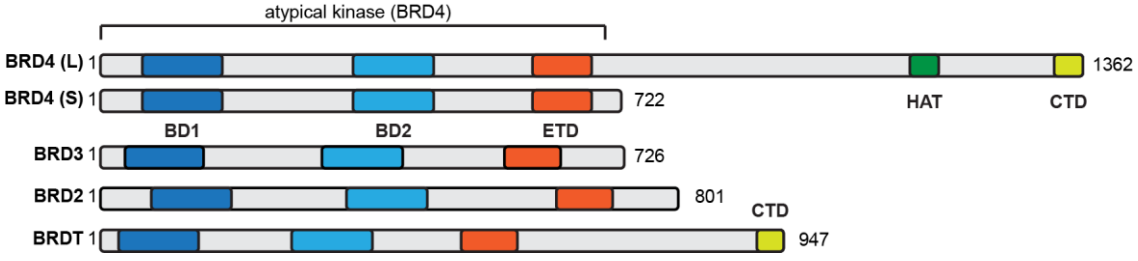


Figure 1.6. Domain organisation of human BET proteins.

BRD4 long and short isoforms shown. Based on Wu & Chiang, JBC, 2007 [46] and Devaiah et al, J. Leukoc. Biol., 2016 [47].

In addition to the Kac-recognising BDs all four BET proteins possess an extra-terminal domain (ETD), a helical domain that has not yet been fully characterised. ETDs are believed to form PPIs with epigenetic/transcriptional proteins/complexes that the

BET proteins direct to certain loci (figure 1.7A). Known ETD binding partners include NSD3 [48, 49], JMDJ6, GLTSCR1, CHD4 and ATAD5 [50], as well as a number of viral proteins [51]. The recruitment of NSD3, JMDJ6 and GLTSCR1 is believed to contribute to up-regulating gene expression, potentially through histone demethylation (NSD3) and transcript splicing (JMDJ6). CHD4, meanwhile, is a component of the NuRD complex associated with heterochromatin formation and gene silencing, while ATAD5 may play a role in DNA repair [50]. Viral integrases may hijack the BET ETDs in order to ensure that the viral genome is integrated into transcriptionally active gene loci [51]. The ETD hence appears to work with the BET BDs in a variety of cellular processes, from directing chromatin remodelling and DNA repair complexes to Kac-rich gene loci and ensuring splicing of BET-regulated transcripts.

BRD4 has also been shown to be an atypical kinase, capable of both autophospohorylation and the phosphorylation of the S2 residue of the RNA Pol II C-terminal tail [52]. This RNA Pol II phosphorylation has been show to occur *in vivo*, and under conditions where other RNA Pol II kinases are inactive. Because RNA Pol II phosphorylation is necessary for the elongation phase of transcription and for mRNA capping and splicing, this is likely one mechanism by which BRD4 binding to acetylated chromatin at the promoters of select genes can then drive their transcription. BRD4 has also been shown to phosphorylate the CDK9 subunit of the positive transcription elongation factor (P-TEFb), and the TFIID component of TAF7, two other kinases key to RNA Pol II activity [47] (figure 1.7B). Although not investigated as thoroughly BRD2 has also been identified as a potential atypical kinase [53], and BRD3 and BRDT may yet be shown to possess this activity. BRD4, furthermore, possesses a 1362-AA ‘long’ isoform containing two additional domains. One is a HAT domain which is believed to acetylate lysine residues in H3/H4 histone tails and the K122 residue of H3’s globular domain. This activity decompacts chromatin around downstream genes, enhancing transcription [54] (figure 1.7D). Finally, the CTD domain (also found in BRDT) interacts with P-TEFb, which is essential for transitioning from the initiation phase to the elongation phase of transcription [55]. The CTD directly interacts with both the CDK9 and the cyclin T1 subunits of P-TEFb, and stimulates the phosphorylation of the RNA Pol II C-terminal tail at BET-regulated genes [56, 57] (figure 1.7C). The short isoform of BRD4 likely functions to maintain BD, ETD and atypical kinase-based functions, while

potentially acting as a competitive inhibitor of the long isoform with regards to HAT and CTD-mediated activities [54]. The HAT and atypical kinase activity of BRD4 is prevented by BET BD inhibitors [52, 54], highlighting the dominant nature of the bromodomains to BET protein function.

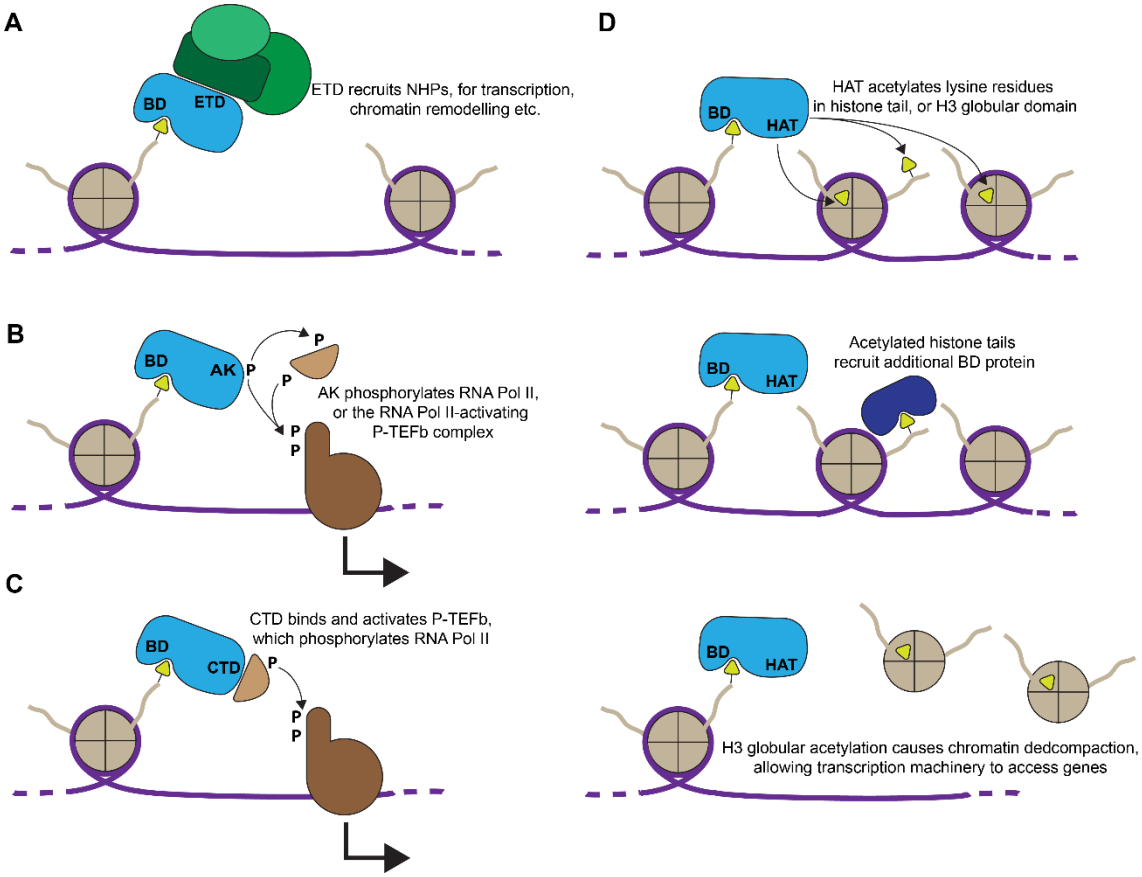


Figure 1.7. BET protein activities.

BD1 and BD2 bind acetylated lysine residues in histone tails (for clarity only 1 BD shown). A) ETD recruits NHPs (non-histone proteins), for various functions. B) Atypical kinase (AK) activity phosphorylates and activates RNA Pol II's C-terminal tail directly, or through P-TEFb. C) CTD binds and stimulates P-TEFb activity. D) HAT acetylates histone tails, or H3's globular K122 residue. Acetylated histone tails recruit additional BD proteins. H3 K122 acetylation causes chromatin decompaction.

In addition to inducing transcription of protein-coding mRNAs BRD4 has also been shown to bind acetylated histone at enhancer sites and, through P-TEFb/RNA Pol II interactions, stimulate the transcription of non-coding enhancer RNA (eRNA) [58]. Enhancer RNAs are not yet fully understood but are believed to help regulate transcription [59], and in a later study eRNAs were shown to interact with BRD4 BDs and stabilise interactions with acetylated histone tails [60].

The clinical significance of the BET proteins was first realised through their role in NUT-midline carcinoma (NMC), a squamous cell carcinoma. NMC's highly-aggressive nature, under-diagnosis and a lack of effective treatments gives it a median survival time of 6.7 months after diagnosis [61, 62], and as of 2006 only one 'long-term survivor' was known of [62]. NMC has been shown to be caused by a chromosomal re-arrangement fusing the nuclear protein in testes (NUT) gene to another gene, which in ~80% cases is BRD4 or BRD3, leading to the expression of a fusion NUT-BET oncoprotein [63, 64]. NUT's ability to recruit the HAT p300, and BRD4/3's ability to bind acetylated chromatin and increase gene expression are believed to drive NMC's proliferation and inhibit cell differentiation in two ways. First BRD4/3 binds acetylated chromatin around MYC promoters, and up-regulates its expression. Secondly, a positive feedback loop is created as NUT recruits p300 which acetylates the nearby chromatin, increasing NUT-BRD4/3 recruitment and thus depriving other pro-differentiation genes of HAT activity [62].

BRD4, BRD3 and BRD2 are expressed in a wide range of tissues and cell-type and hence likely involved in a large number of biological functions. BRDT, in contrast, is expressed solely in the testes where it is necessary for spermatogenesis [65, 66]. BRDT is expressed during meiosis and is believed to contribute to spermatogenesis both by directly increasing transcription of downstream genes and by contributing to chromatin compaction. Knock-out or mutation of BRDT has been linked to low sperm count and male infertility [67, 68]. BRDT has also been seen to be ectopically expressed in some lung cancers [69]. Since BRDT is unlikely to be involved in any clinically important biological processes outside of spermatogenesis, it is not studied as thoroughly as the other BET proteins. Consequently at times both in this thesis and in the literature certain experiments may be performed on the other BET proteins but not BRDT.

1.6 – BET Probes / Inhibitors:

Within the last decade there has been a surge in BET bromodomain inhibitor development. In addition to an array of 'academic' inhibitors used only in research (most famously (+)-JQ1) 17 BET inhibitors have been submitted for a total of 45 registered clinical trials (as of June 2018) [70] (scheme 1.1) (figure 1.8). BET inhibitor development has been discussed thoroughly in the literature [41, 71-73].

BET inhibitors were first published on by Mitsubishi Tanabe in a series two patents in 2008 [74] and 2009 [75], outlining the use of thienotriazolodiazepine compounds for anti-tumour and anti-inflammation uses. Despite this work BET inhibitors did not become widely used until 2010 with the key, concomitant publication of (+)-JQ1 [76] and I-BET762 [77]. (+)-JQ1, the premier academic BET inhibitor, was published by the laboratory of James Bradner in collaboration with the Structural Genomics Consortium and others, in a paper focusing on NMC. The compound was selected following optimisation of the previously-mentioned compounds discovered by Mitsubishi Tanabe in an anti-cancer screen [75]. I-BET762, meanwhile, was originally discovered in a GSK screen for ApoA1 up-regulators [78]. I-BET762 was first published in the context of inflammation suppression, but has since entered clinical trials for anti-cancer indications. Both compounds possess ~100 nM K_d values towards BET BDs and are based on the 'privileged' benzodiazepine scaffold [79]. This scaffold has since been used for the development of the clinical inhibitors OTX015 [80] and CPI-0610 [81], as well as several academic inhibitors [82]. A variety of other scaffolds have since been utilised to produce BET inhibitors, with slight changes in potency and DMPK characteristics [41, 71, 72]. The only feature shared by all BET inhibitors is the presence of a Kac mimic moiety (scheme 1.2) (figure 1.8). In order to boost overall potency bivalent BET inhibitors have also been developed. AstraZeneca produced the asymmetric compound AZD5153 which has now entered clinical trials [83, 84], while the Bradner lab developed MT1 by conjugation of two (+)-JQ1 molecules via a polytheylene glycol (PEG) linker [85].

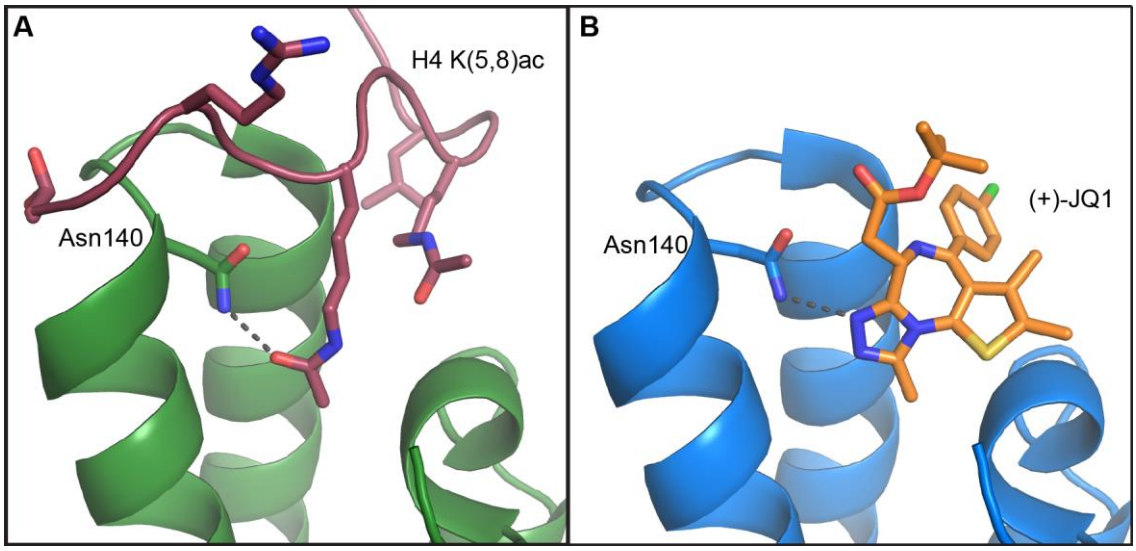
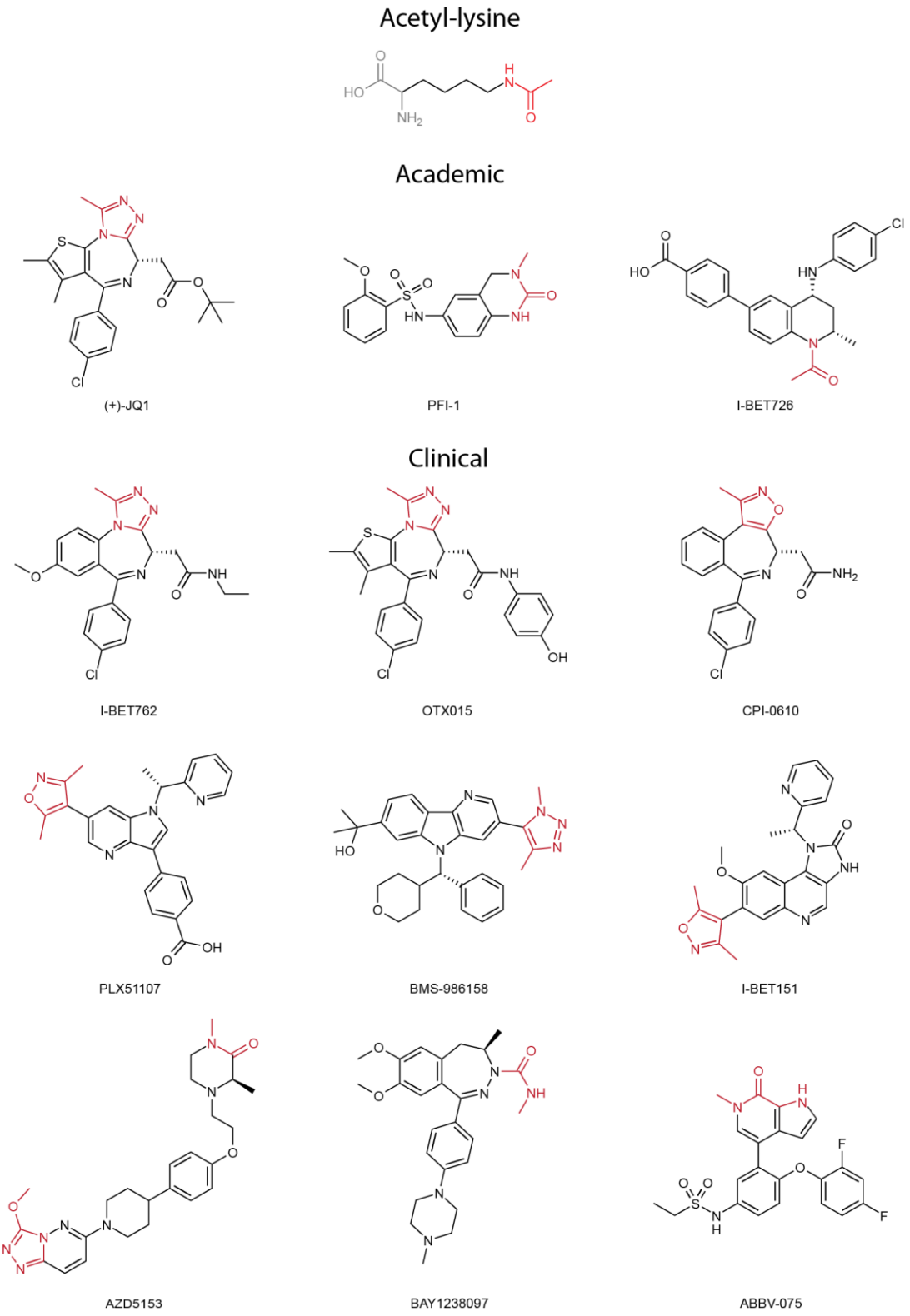


Figure 1.8. BET inhibitors mimic the Kac mark

X-ray co-crystal structures of BRD4 BD1 in complex with an H4 K(5,8)ac peptide (A) (3UVW) and (+)-JQ1 (B) (3MXF). Critical N140 residue highlighted. Grey dash represents critical hydrogen bond.

In addition to 'standard' inhibitors a few molecules have been developed that have been functionalised for specific purposes. Bio-JQ1 was developed by linking (+)-JQ1 to a biotin unit through a triple-PEG linker, and was used to show the localisation of (+)-JQ1 and its targets throughout the genome [86]. A photo-reactive (+)-JQ1 cross-linker has also been developed, able to label bound proteins with azides/biotin (for on-target and potential off-target pull-down) and fluorophores (for fluorescence microscopy) [87]. Later cross-linkers, based on GW841819X, were used in chemoproteomic experiments to identify a novel off-target [88]. (+)-JQ1 compounds with click-chemistry compatible labels have also been developed for protein labelling and pulldown [89].



Scheme 1.2. BET bromodomain inhibitors developed as probes and clinical candidates

Kac mimics in red.

BET bromodomain inhibitors have gone on to be used to study the BET proteins in a wide range of fields, thanks in part due to the availability of compounds like (+)-JQ1. BET inhibitors have been shown to reduce the proliferation of a number of cancers [72], such as acute myeloid leukaemia (AML) [90]. This is understood to be primarily due to a reduction in BET-regulated expression of the oncogene C-MYC [91]. The anti-MYC activity of BET inhibitors has led to their use against a number of other MYC-dependent cancers [92, 93]. Additionally, BET inhibitors have shown to be active against a number of other cancer types, such as lung adenocarcinoma (LAC), in a MYC-independent fashion. This activity is due to the role of BET proteins in regulating oncogenic signalling pathways such as FOSL1 and NF-κB [94-96].

In addition to cell proliferation and differentiation the BET bromodomains have also been shown to regulate inflammation [77, 97, 98]. BET proteins are necessary for the production of inflammatory cytokines by macrophages, and BET inhibition has been shown to alleviate sepsis by reducing interleukin-6 (IL-6) and tumour necrosis factor alpha (TNF-α levels). The importance of BET proteins to the activity of immune cells, such as TH17 helper cells, is likely connected to their significance in the previously mentioned haematological malignancies.

The BET proteins also regulate the apolipoprotein A1 (ApoA1) protein, part of the high-density lipoprotein (HDL) complex and a key indicator of cardiovascular health. One effort to identify and develop ApoA1 up-regulators was led by GSK, for which they generated a HepG2 cell-line transfected with an ApoA1 luciferase reporter gene, and this led to the discovery of benzodiazepine (I-BET762) and tetrahydroquinoline (I-BET726) BET bromodomain inhibitors [78, 99]. A similar effort by Resverlogix led to the discovery of RVX-208 [100]. These represent prominent examples of successful forward chemical genetics.

As BET inhibitors are used in more and more scientific fields and biological contexts more potential BET protein functions are revealed. These include such disparate roles as male fertility [101], HIV infection [102] and neuronal development [103, 104].

As is clear from the above chemical genetics studies, BET bromodomain inhibitors are believed to possess great therapeutic potential [72], leading to the

previously mentioned 45 federal drug administration (FDA)-registered clinical trials with 17 compounds from 15 companies [70]. About 75% (32) of these clinical trials are targeted towards anti-cancer indications (figure 1.9). Some of these anti-cancer trials (3) relate to the BET proteins key (and dysfunctional) role in NMC, but most (29) aim to take advantage of the BET protein's control of key proliferative or anti-differentiation signalling pathways such as C-MYC. Outside of NMC the cancers targeted by BET bromodomain inhibitors are 'advanced solid tumours', 'neoplasms' and hematologic malignancies; specifically AML, castration-resistant prostate cancer (CRPC), MM, diffuse large B-cell lymphoma (DLBCL) and lung cancers. These anti-cancer trials are overwhelmingly at an early stage, with 24 in phase 1, 8 in 1/2, 3 in phase 2 and none so far in phase 3 [70] (figure 1.9). The majority of these trials utilise BET inhibitors as a monotherapy, or with standard background care (24), although in 16 trials they are used in combination with treatments such as androgen/estrogen receptor antagonism or monoclonal antibodies. The overall efficacy of BET inhibition in these trials, where complete, is difficult to assess. In general, BET inhibitors appear to be generally efficacious, but limited by a feedback looping triggering BET protein up-regulation. A major complication appears to be dose-dependent toxicity. The most common adverse effect observed is thrombocytopenia (platelet depletion), alongside more common toxicities such as fatigue and diarrhoea [105-108]. As these are observed in multiple trials, using differing scaffolds they are likely the result of undesired on-target activity. There is also concern that BRD4/2 inhibition may reactivate latent HIV (or other) infections [55, 102, 109, 110].

All non-cancer-targeting clinical trials so far utilise Resverlogix's RVX-208, which will be discussed in greater detail later. These trials possess indications focusing on cardiovascular health (atherosclerosis, coronary artery disease and acute coronary syndrome), diabetes, kidney disease and the metabolic disorder Fabry disease. Due to its earlier development RVX-208 has reached the later stages of clinical trials, with five phase 2 and one phase 3 clinical trials (figure 1.9). The phase 3 clinical trial, BETonMACE, aims to determine if RVX-208 treatment in type-2 diabetes can reduce the onset of major cardiovascular events [111].

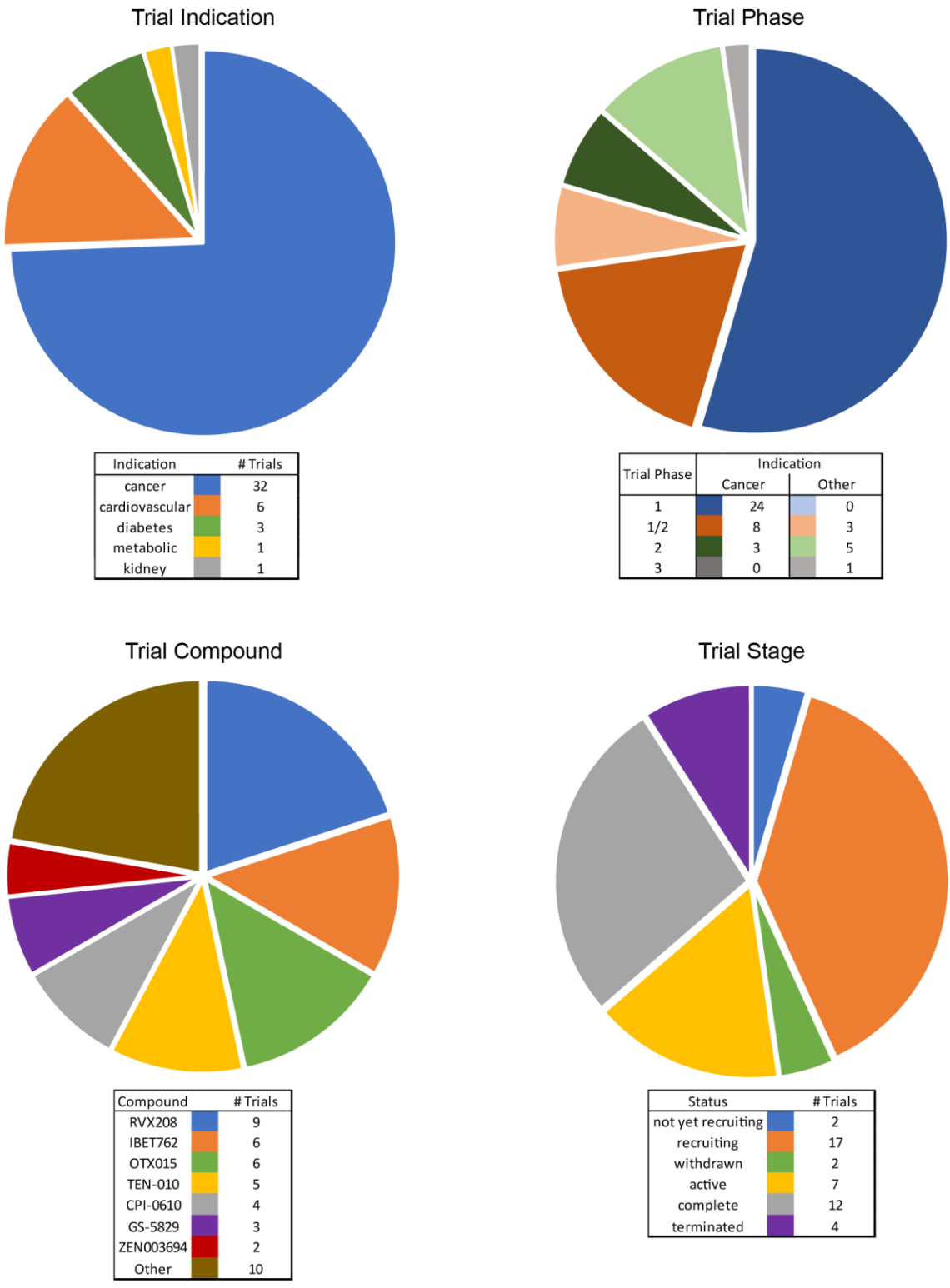


Figure 1.9. BET bromodomain clinical trials.

Data from clinicialtrial.gov, as of June 2018. In chart 3 'other' refers to the following 10 compounds used in only one clinical trial each: AZD5153, ABBV-075, ABBV-744, BAY1238097, BI894999, BMS-986158, FT-1101, INCB054329, PLX51107 and I-BET151.

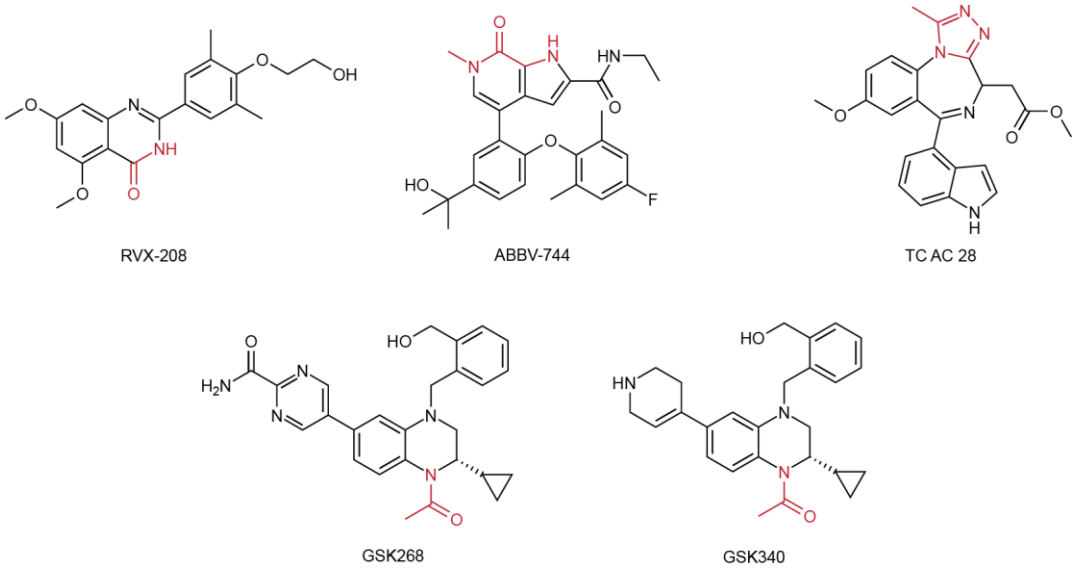
1.7 – Selective BET Probes:

Unfortunately, the majority of BET bromodomains inhibitors developed so far have been pan-selective, meaning that though they are selective for the BET family over other bromodomains they are not selective *within* the family [41]. As these compounds bind to all BET bromodomains and proteins with roughly equal potency it has proved difficult to illustrate the significance of specific bromodomains/proteins or to link specific phenotypes and cellular functions with specific BET bromodomains/proteins. By combining chemical inhibitors with genetic tools and molecular biology some advances have been made on the protein-specific level [43, 112-114], but this has failed to keep pace with the wealth of research illustrating the importance of the BET bromodomains on a family-wide scale. Thankfully, a small number of compounds have been released showing inter-BET selectivity, taking advantage of the small differences in sequence and conformation that exist between different bromodomains (figure 1.10). These efforts have been most successful in generating selectivity for the BET BD2s over BD1s (scheme 1.3).

For some years the only reasonably high-quality, selective BET inhibitor was Resverlogix's RVX-208. Developed from a screen for ApoA1 up-regulators [100] RVX-208 entered clinical trials before later being identified as a BET inhibitor with ~20-fold selectivity and ~200 nM potency for the BD2s [115, 116]. Structural biology and molecular dynamics simulations have been used to rationalise this selectivity. RVX-208 is shown to pack against the H437 residue of BD2, an interactions not as favoured in the case of the BD1 equivalent residue (D144) [115, 116]. MD simulations, meanwhile, showed that RVX-208 can selectively re-orient the binding pocket of BRD2 BD2 in a more favourable fashion than in BRD2 BD1, improving the BD:RVX-208 affinity. This re-orientation upon RVX-208 binding is most beneficial to binding with regards to the BD2s [117]. RVX-208 was confirmed to still be active in cells and could displace BRD3 from chromatin, though to a lesser extent than the pan-selective PFI-1 [116]. RVX-208 treatment was also shown to alter the regulation of genes in HepG2 cells, though again in a reduced fashion compared to pan-selective (+)-JQ1. RVX-208 affected ~10% of the genes affected by (+)-JQ1 [116]. Unlike the bulk of non-selective BET inhibitors, which are focused on cancer indications, RVX-208 has mainly been targeted against heart diseases and metabolic disorders. RVX-208 is, interestingly, the only BET inhibitor to

reach phase 3 clinical trials, although it is hard to say if this is due to its BD2 selectivity, its differing indications or pure chance.

RVX-208 has, in the last year been followed by Abbvie's ABBV-744 [118-120] and GSK'S GSK268 and GSK340 BD2-selective inhibitors [121]. Finally, in 2016 the Alessio Ciulli lab developed a BD2-selective indole-derivative of the benzodiazepine scaffold [122].



Scheme 1.3. BD2-selective BET bromodomain inhibitors.

Potential Kac mimics in red.

GSK268 and GSK340 were both developed from the very potent but pan-selective tetrahydroquinoline I-BET726 [99] and display very promising qualities. Both compounds display, in a time-resolved-FRET (TR-FRET) assay, ~100 nM potency against BET BD2s, selectivity against BD1s from ~7-fold (BRDT) to ~45-fold (BRD4), anti-BET activity in cells at low concentrations and a wealth of good to moderate quality DMPK properties [121]. GSK340 was also tested in the DiscoveRX BROMOscan panel, wherein it possessed *Kd* values against BET BD2s of 100 to 10 nM and 40-230-fold selectivity against BD1s, in addition to negligible binding to non-BET BDs [121]. To achieve this remarkable selectivity GSK exploited two differences in protein sequence between BET BD1s and BD2s. The first variation is the identity of the 'gatekeeper' residue, which controls access to the WPF-shelf, corresponding to I146 in BD1s and V439 in BD2s (BRD4 numbering) – allowing for a differential steric clash with the ligand. The second difference was on the BC loop, wherein the BD1's D144 and D145 contrast with H437 and E438 in the BD2. The presence of E438 on the BD2 allows for the formation of a

BD2-selective water-mediated hydrogen bond, while H437 alters the conformation of an aryl moiety, improving interactions with the WPF shelf in BD2s [121]. Despite the strong potency and selectivity, cellular activity and strong DMPK characteristics neither GSK compound has been submitted for clinical trials, as their high intrinsic clearance (CL_{int}) in rats prevents pre-clinical testing.

Abbvie's BD-selective ABBV-744, disclosed close to GSK's compounds, also displays very promising properties. ABBV-744 was developed from the pan-selective pyrrolopyridone ABBV-075, which in a TR-FRET assay displays 1-15 nM potency for all BET BDs [123]. Through exploiting the I146/V439 and D144/H437 differences discussed above ABBV-744 achieved BD2 potency of 2-5 nM, and selectivity against BET BD1s of 250-640-fold [118, 120]. In the cellular NanoBRET assay ABBV-744 possessed 28 nM potency against BRD4 BD2, and 750-fold selectivity against the BD1. ABBV-744 appears to retain ABBV-075's strong DMPK properties, with similar oral bioavailability in mice, and nM potency against some BET-dependent cancer cell-lines. Interestingly, ABBV-744 was shown to lack cytotoxicity for a number of (non-AML) cell-lines vulnerable to the pan-selective ABBV-075 [119], suggesting BD2 inhibition may have different therapeutic uses. Excitingly, ABBV-744 has already been submitted for a phase 1 clinical trial against CRPC and AML [124], suggesting BD2-selective inhibitors may have therapeutic use beyond RVX-208's cardiovascular indications.

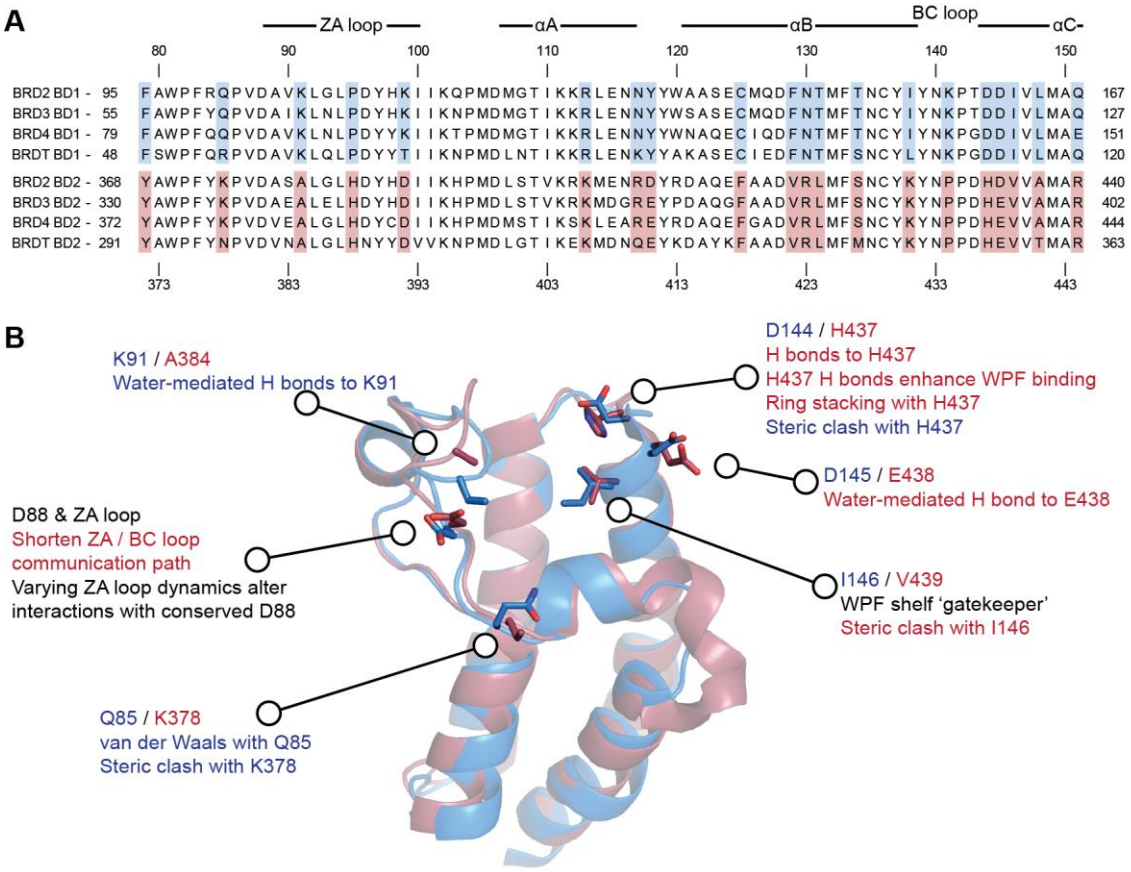


Figure 1.10. Structural basis of BD1 / BD2 selectivity.

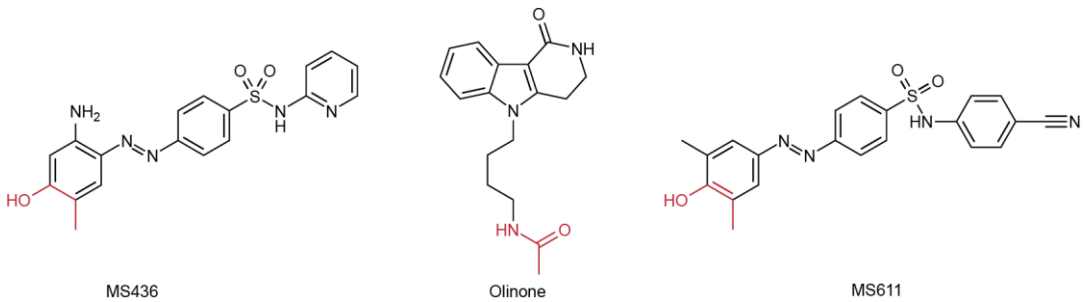
A) Alignment of BET BD protein sequence, with notable differences between BD1s and BD2s highlighted. Numbering refers to BRD4 BD1 (top) and BD2 (bottom). B) Crystal structures of apo forms of BRD4 BD1 (blue) (2OSS) and BD2 (red) (2OUO), with varying residues highlighted. Text explains how BD1/BD2 differences are exploited by selective compounds, text colour refers to which BD group said interactions enhanced selectivity for (blue – BD1, red – BD2).

The Ciulli lab's benzodiazepine-indole derivative was synthesised during attempts to develop allele-selective benzodiazepine analogues targeting W97F/H and W370F/H (BRD2 numbering) BET bromodomain mutants (discussed later). The indole compound was later seen to bind BRD4 and BRD2 BD2s with ~50 nM K_d values, and possess 10-15-fold selectivity against the relevant BD1s [122]. A co-crystal structure of the indole with BRD2 BD2 W370F showed it to possess the same overall binding mode to non-selective analogues, but with the BD2-specific H433 moved from an 'open' conformation pointing away from ligands to a 'closed' conformation wherein it formed a π-stack with the indole ring [122]. The synthesis of this indole compound was later optimised, and it was named TC AC 28 [125].

In addition to the above published work, the pharmaceutical company Zenith Epigenetics is believed to have developed a small number of BD2-selective inhibitors

(ZEN-222, 297 & 2135) with <100 nM potency, ~50-fold selectivity and some degree of activity in cells [41]. No structures have been published nor have any such compounds progressed to clinical trials alongside their pan-selective ZEN-3694 [126], and they are likely intended for use as chemical tools.

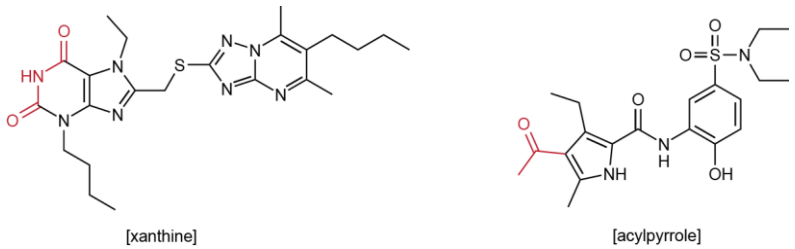
The development of BD1-selective inhibitors has, by comparison, been less successful, with no such compounds developed by pharmaceutical companies or submitted for clinical trials. All published BD1-selective inhibitors (scheme 1.4) were developed by the Ming-Ming Zhou lab, starting with the diazobenzene MS436 compound published in 2013. This compound is highly potent (30-50 nM K_i) but has limited selectivity, with only a 10-fold preference for the BD1, and only in the context of BRD4 [127]. Structural studies suggest this selectivity was due to water-mediated hydrogen bonding to the K91 residue of BRD4 BD1, which corresponds to A384 in BRD4 BD2. In a later 2014 paper MS611 and Olinone were discussed. Olinone displayed a much weaker affinity (3.4 µM K_d) but enhanced selectivity for the BD1 over the BD2 of all somatic BET proteins (no quantifiable BD2 binding) [128]. This enhanced selectivity was shown to be due to favourable interactions with D144 in the BRD4 BD1 binding pocket, contrasted to steric clashes with its BD2 equivalent, H437. Olinone, alongside additional BD1/BD2-selective inhibitors was used to probe the role of BET proteins in oligodendrocyte differentiation. Interestingly, this work showed that while pan- and BD2-selective inhibitors inhibited differentiation; BD1-selective inhibitors instead up-regulated this process – hinting at differing roles for the different bromodomains [128]. Finally, MS611, another diazobenzene compound, possesses a K_i of 400 nM against BRD4 BD1, and 100-fold selectivity against BRD4 BD2 [128]. Unfortunately this very impressive potency and selectivity is not replicated with other BET proteins, with 1-6 µM K_i values against both BRD2 and BRD3 BDs.



Scheme 1.4. BD1-selective BET bromodomain inhibitors.

Potential Kac mimics in red.

In addition to BD1 selectivity, there has been some recent progress of inhibitors specific for a single bromodomain (scheme 1.5). BD1 and BD2-selective inhibitors are of use, seem able to generate differing phenotypes in cells and may be of differing therapeutic quality, but do not allow for interrogation of specific BET proteins or single bromodomains. A xanthine-based Kac mimic, published in 2016, showed moderate selectivity for BRD4 BD1, but with an overall potency of 1-5 µM [129]. In a TR-FRET assay this compound showed ~10-fold selectivity for BRD4 BD1 over BRD3 BD1 and BRD2 BD1, >20-fold selectivity against BRDT BD1 and no quantifiable binding to any BET BD2. Despite its low potency this xanthine did appear capable, at high enough concentrations, to supress the expression of C-MYC in Jurkat cells, and to impair the viability of said cells [129]. Through the use of co-crystal structures, and the synthesis and testing of less selective analogues, it was theorised that this selectivity is not, as one might expect, caused by ligand:protein interactions via residues unique to BRD4 BD1. While the xanthine core allows for pan-selective BET inhibition a triazolo moiety forms an important bond with the D88 residue on the ZA loop. This D88 residue is conserved within the BET family, but the differing dynamics of the ZA loop between BDs limits the ability of the compound to make this hydrogen bond with other BDs – generating selectivity. In addition to this conformation-specific D88 interaction, selectivity against BRDT BD1 and the BD2s was enhanced through exploitation of sequence variation. While the xanthine compound forms van der Waals interactions with the side-chain of Q85 in BRD4 BD1 it clashes sterically with the equivalent arginine (BRDT BD1) and lysine (BD2s) residues [129].



Scheme 1.5. BRD4 BD1-selective BET bromodomain inhibitors.

Potential Kac mimics in red.

Another paper, also published in 2016, revealed a 4-acyl pyrrole derivative with moderate selectivity for BRD4 BD1 [130]. The selectivity of this compound was not as thoroughly characterised as the xanthine, although as ITC was used instead of TR-FRET the measured K_d values are expected to be more reliable. The acyl pyrrole displayed ~12-fold selectivity for BRD4 BD1 over BRD3 BD1 (similar to the xanthine compound) but only 10-fold selectivity against BRD3 BD2 [130]. No other BET bromodomains were tested. In contrast to the xanthine compound, however, the acyl pyrrole was a potent binder, with a K_d for BRD4 BD1 of 240 nM [130]. A number of features/interactions are presented as possible caused for said selectivity, mainly the occupation of the 'WL trap' formed by W81 and L92. In addition a hydrogen bond between the pyrrole unit and BRD4 BD1's P82 residue, as well as a perpendicular T-shaped π–stacking interaction with W81, were suggested as means to generate selectivity. Through the use of QSAR it was shown that selectivity benefited from exploiting an acceptor feature within BRD4 BD1's ZA channel [130]. Despite its superior potency to the xanthine compound the acyl pyrrole was not tested in cells. To date there is no known high-quality (<500 nM potency, >30x selectivity, acceptable DMPK) single-BD-selective inhibitors.

1.8 – Selective BET Degradation

Interestingly, the parallel development of proteolysis-targeting chimera (PROTAC) compounds has led to the observation of inter-BET protein selectivity. PROTACs are heterobifunctional ligands, in which a ligand that binds to an E3 ubiquitin ligase is linked to a ligand that binds one's protein of interest. Such a compound will then bring together the target protein and the E3 ligase to form a ternary structure that triggers ubiquitination of the target, leading to its proteasomal degradation [131] (figure 1.11A).

Several BET PROTACs have been developed by a number of labs, with most remaining pan-selective, however the Alessio Ciulli lab has developed a number of PROTACs showing varying degrees of preference for BRD4 degradation [132]. The selectivity of the PROTAC MZ1 (figure 1.11B) does not appear to be due to any differences in bromodomain binary binding affinity. Instead, MZ1 has been shown to induce novel PPIs between the BET bromodomains and their E3 ligase (the von Hippel-Lindau protein VHL), leading to the formation of cooperative ternary complexes. In this context cooperativity refers to the binding of the PROTAC to one of its protein binding partners enhancing its affinity for the remaining protein. At this 'cooperativity' level MZ1 shows greater cooperative binding and hence selectivity for BRD4 BD2 over other BET BDs (figure 1.11C). X-ray crystallography-derived 'ternary' structures of BRD4 BD2-PROTAC-VHL complexes revealed PROTAC-induced interactions between BRD4 and VHL residues, as well as interactions between the BET-targeting ligand and VHL, and again between the VHL-targeting ligand and BRD4 BD2 [133]. Said structures also informed the development of newer PROTACs, like AT1, with varying selectivity for BRD4 [133]. This selectivity is ultimately based upon the ability of VHL and BRD4 BD2 to interact with each other and does not lend itself to rational design of PROTACs selective for other proteins [131].

BRD4-selective PROTACs have also been developed by a second lab, instead utilising the cereblon (CRBN) E3 ubiquitin ligase and not exploiting cooperativity [134]. These PROTACs take advantage of the plastic nature of the induced BRD4:CRBN interactions and a variety of potential conformations the two proteins can form.

PROTACs induce the degradation of their target proteins in their entirety. As such BET-targeting PROTACs cannot be used to interrogate the differing roles of BD1s and BD2s, but only those of different BET proteins. Additionally the phenotypes they generate in cells could be caused by the removal of non-bromodomain domains within BRD4, such as its ET domain, which would be unaffected by standard small-molecule inhibitors. As PROTACs act catalytically, that is each individual PROTAC molecule will over time bind and induce the degradation of many protein molecules, they do not require large concentrations to be effective, and MZ1 doses of 10-100 nM have been shown to completely degrade BRD4 within 24 hours [132]. This suggests PROTACs may be very effective treatments for BET-sensitive cancers, such as AML, although this will likely be impeded by the DMPK characteristics such large compounds will possess and a potential increase in on-target toxicity.

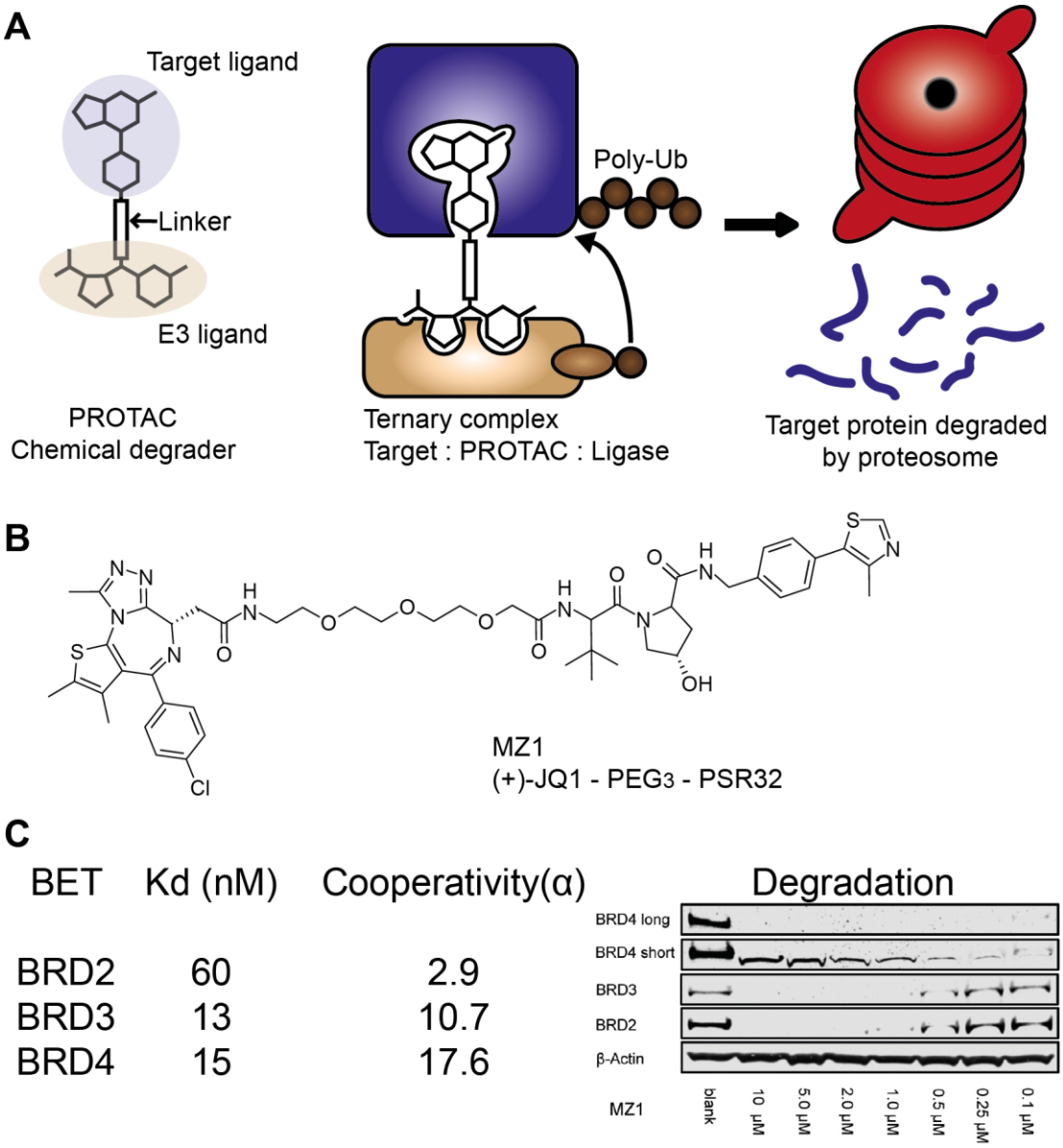


Figure 1.11. Inter-BET selectivity with PROTACs

A) Mode-of-action of PROTAC compounds, modified from Runcie et al, COCB, 2016 [1]. B) Chemical structure of BRD4-selective PROTAC MZ1. C) ITC-measured affinities of MZ1 for BET bromodomains (BD2), and associated cooperativity (X-fold increase in affinity for VHL, when MZ1 pre-incubated with BD2). Western-blot observed degradation of BET proteins in HeLa cells, following 24 h treatment. Affinity and cooperativity data from Gadd et al, Nat. Chem. Biol., 2017 [133], western-blot from Zengerle et al, ACS chem. biol., 2015 [132].

1.9 – Bump-&-Hole Theory:

One means by which selective inhibitors can be acquired is the design of an orthogonal protein/ligand pair, by engineering fully functional mutant proteins targeted by an allele-selective analogue of an existing probe [1, 135, 136]. By modifying a single existing pan-selective probe one can achieve a single allele-selective compound that can be used to inhibit multiple target proteins bearing the mutation. By doing this potentially years of compound design can be bypassed. This method is dependent on the orthogonal protein mutants retaining sufficient functionality in cells and *in vivo* to allow generation of viable mutant cell-lines or animal models. The resulting compounds will not have any immediate therapeutic use, as they will not be effective against wild-type (WT) proteins involved in disease.

There are several methods available to generate orthogonal protein/ligand pairs [135] (figure 1.12). Protein/ligand interactions that rely on charge-charge interactions can be made allele-selective through reversing the charge complementation: i.e. taking a positively-charged probe binding a negatively-charged binding site, and developing a negatively-charged probe that binds a positively-charged mutant. A similar approach is to replace hydrophobic protein:ligand interactions with polar interactions, or vice-versa. Allele-selectivity can also be engineered through the introduction of a novel covalent interaction. Introducing a cysteine residue into a suitable position can allow for selective inhibition by probe bearing a suitable electrophile group (e.g. thiol or acrylamide) [137]. This method is best applied when structure-activity relationship (SAR) studies have been unable to produce high-affinity probes, as the covalent interaction should strongly boost potency. The reactive probe can suffer from low selectivity if it is able to form covalent bonds with cysteine residues on other, undesired proteins.

Perhaps the most successful strategy to engineer selective protein:ligand binding is the exploitation of steric complementation – the ‘bump-&-hole’ approach (figure 1.13) [135, 136]. Replacing a large or medium-sized residue (phenylalanine, tyrosine, tryptophan, leucine, isoleucine) with a smaller residue (valine, alanine, glycine) can create a void (‘hole’) in the target protein’s binding site. Meanwhile, introducing a complementary steric ‘bump’ to an existing probe in a suitable position can confer upon it allele-selective binding. This is because the bump will weaken or

prevent binding to WT proteins through the introduction of a new steric clash, while increasing its affinity to mutant proteins by the introduction of a novel hydrophobic interaction if correctly complemented by the mutation 'hole'. By not altering the intrinsic chemical nature of any residues (unlike charge or polar/hydrophobic reversal) this method is also the least likely to alter the functionality of the protein, a typically desirable characteristic of an allele-selective system.

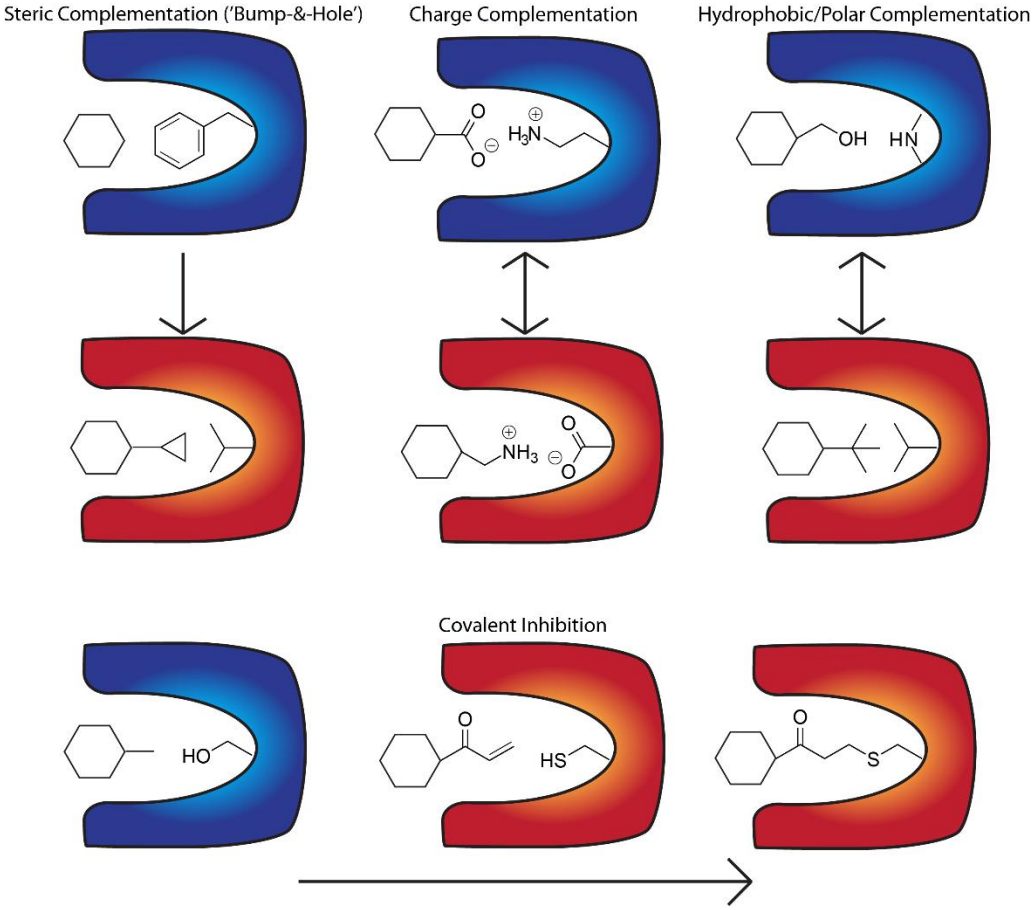


Figure 1.12. Strategies for design of orthogonal protein:ligand pairs.

Example ligand, with and without orthogonal modification, paired with WT (blue) and mutant (red) proteins, with mutated residue shown.

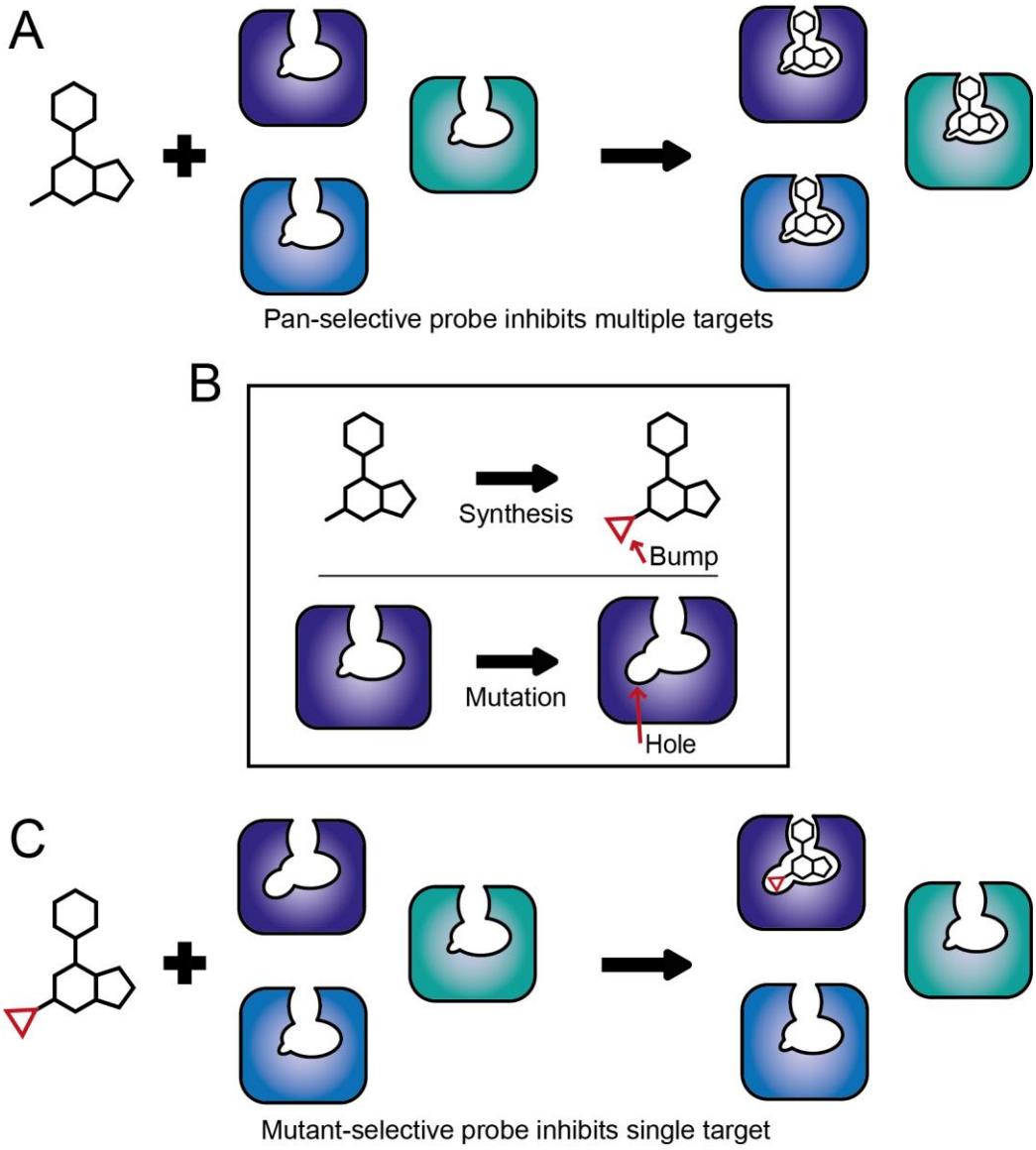


Figure 1.13. 'Bump-&-Hole' strategy

Modified from Runcie et al, COCB, 2016 [1].

1.10 – Bump-&-Hole Examples (non-epigenetics):

Perhaps the best known application of the bump-&-hole method of engineered selectivity was the Kevan Shokat lab's development of mutant selective kinase cofactors and inhibitors [138]. Kinase-dependent protein phosphorylation is a vital part of biological signalling and is involved in almost all cellular processes, but single-target inhibition of all ~500 human kinases is far-fetched and, even if attainable, would require the deployment of significant efforts and resources. This motivated the development of allele-selective strategies.

The bump-&-hole approach was first performed on the viral tyrosine kinase vSRC, which was known to phosphorylate over 50 distinct proteins and to be responsible for the transformation of fibroblasts by Rous sarcoma virus [139]. The specific substrates of vSRC could not be sufficiently distinguished from those of other kinases, nor was any selective, potent and cell-permeable inhibitor in use. In its first iteration this system paired a V323A / I338A mutant kinase with an N6-cyclopentyl – bumped ATP analogue, which was unreactive with WT vSRC but could be successfully catalysed by the mutant [139]. Although the mutant kinase retained some kinase activity, its catalytic efficiency (k_{cat}/K_m) with N6-cyclopentyl-ATP was 50-fold lower than that of WT vSRC for unmodified ATP (table1.1). This significant loss of activity made this bump-&-hole system unlikely to be able to compete with WT kinases in cells and allow for successful substrate identification. Addition structural biology and SAR work generated an optimised system, using an I338G mutation and N6-benzyl – bumped ATP, which showed improved catalytic efficiency [140] (table 1.1) (scheme 1.5). To emphasise generality and potential wide applicability of the bump-&-hole approach a related Src kinase, Fyn, was also engineered to produce a catalytically functional T339G mutant [140].

Enzyme	Substrate	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}S^{-1}$)
vSRC WT	ATP	2	12	2.8×10^3
	N6-cyclopentyl-ATP		-	
	N6-benzyl-ATP		-	
vSRC V323A / I338A	ATP	0.8	150	8.9×10^1
	N6-cyclopentyl-ATP	0.05	15	$5.6 \times x10^1$
vSRC I338G	ATP	0.8	80	1.7×10^2
	N6-cyclopentyl-ATP	0.1	15	1.1×10^2
	N6-benzyl-ATP	0.2	5	6.7×10^2
Fyn WT	ATP	2.5	70	6.0×10^2
	N6-benzyl-ATP		-	
Fyn T339A	ATP	0.8	90	1.5×10^2
	N6-benzyl-ATP	0.5	25	3.3×10^2
Fyn T339G	ATP	0.5	100	1.0×10^2
	N6-benzyl-ATP	0.3	7	7.1×10^2

Table 1.1. Kinetics of kinase mutants and cofactor analogues

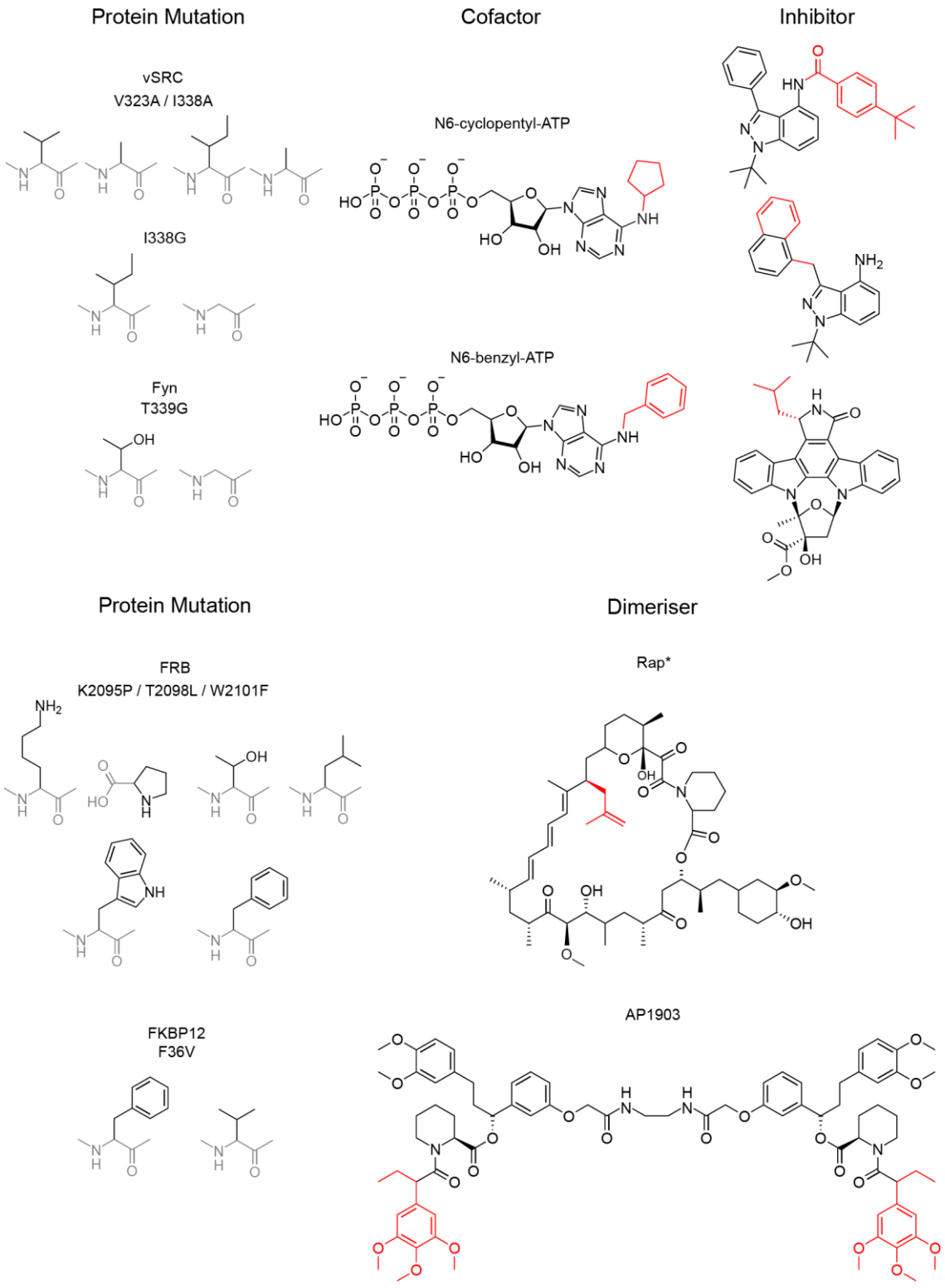
Symbol '-' denotes no observed activity. Data from [139, 140]

In addition to a mutant-selective cofactor a mutant-selective small-molecule inhibitor was developed. Based on the Src pan-selective inhibitor PP1 this bumped inhibitor possessed ~500 nM potency and >600-fold selectivity for vSrc and Fyn mutants over their WT counterparts [141]. This selectivity was then confirmed in cells, where treatment of mutant (but not WT) cells with the bumped inhibitor blocked the phosphorylation of protein and the vSRC-dependent formation of fibroblasts [141].

The use of mutant-selective inhibitors was later expanded to the rest of the kinome [142]. Kinases from five distinct families were engineered to include mutations comparable to the vSRC I338G mutations and tested against bumped modifications of two inhibitor scaffolds: the Src inhibitor PP1 used previously and the pan-kinase K252a inhibitor. This proved to be successful, with each tested kinase being targetable by at least one selective inhibitor, and the bumped analogues of the originally Src pan-selective PP1 being potent and selective for mutants of other kinase families [138, 142]. This approach was then applied to the fungal kinase Fus3, where the mutation had little impact on cell viability, and the bumped inhibitor showed high potency and a stark selectivity for the mutant [138, 142]. The same was done with the fungal CDK2 homologue Cdc28, and the resulting phenotypes differed from that displayed by the

previously used method of engineering temperature-sensitive mutants [142]. An alternative approach to engineering orthogonal protein:ligand pairs, the introduction of a covalent interaction, has also been trialled, with a c-SRC T338C mutant retaining activity and being selectively and covalently bound by thiol-modified kinase inhibitors [143].

Another significant application of the bump-&-hole approach was the development of orthogonal FKBP dimerisers (scheme 1.6). A family of prolyl isomerases, the FKBP proteins are involved in protein folding and trafficking. Bivalent analogues of the FKBP-inhibiting immunosuppressant FK506 are able to induce dimerization of FKBP12, and this process has been used as a tool to engineer and manipulate PPIs to study and control protein localisation and function [144]. The bump-&-hole approach was first utilised in 1997. At this point in time rapamycin was used to induce dimerization of FKBP12 and the FRB domain of RAFT, but this approach suffered from rapamycin's innate off-target activity. Using previously solved binary and ternary crystal structures a number of inactive, bumped rapamycin analogues were developed, and then screened against a library of FRB mutants. This approach eventually lead to the pairing of a methallyl bump and a triply-mutated FRB (T2098L, W2101F and K2095P), which was shown to be able to alter the cellular localisation of proteins in cells, by recruiting a kinase domain to the plasma membrane [145]. The methallyl-bumped compound displayed ~15 nM potency and ~50-fold selectivity for the FRB triple-mutant over WT [145].



Scheme 1.6. Applications of bump-&-hole system.

Chemical structures of WT and mutated residues of target protein, allele-specific enzyme co-factors, inhibitors and chemical dimerisers (modifications in red).

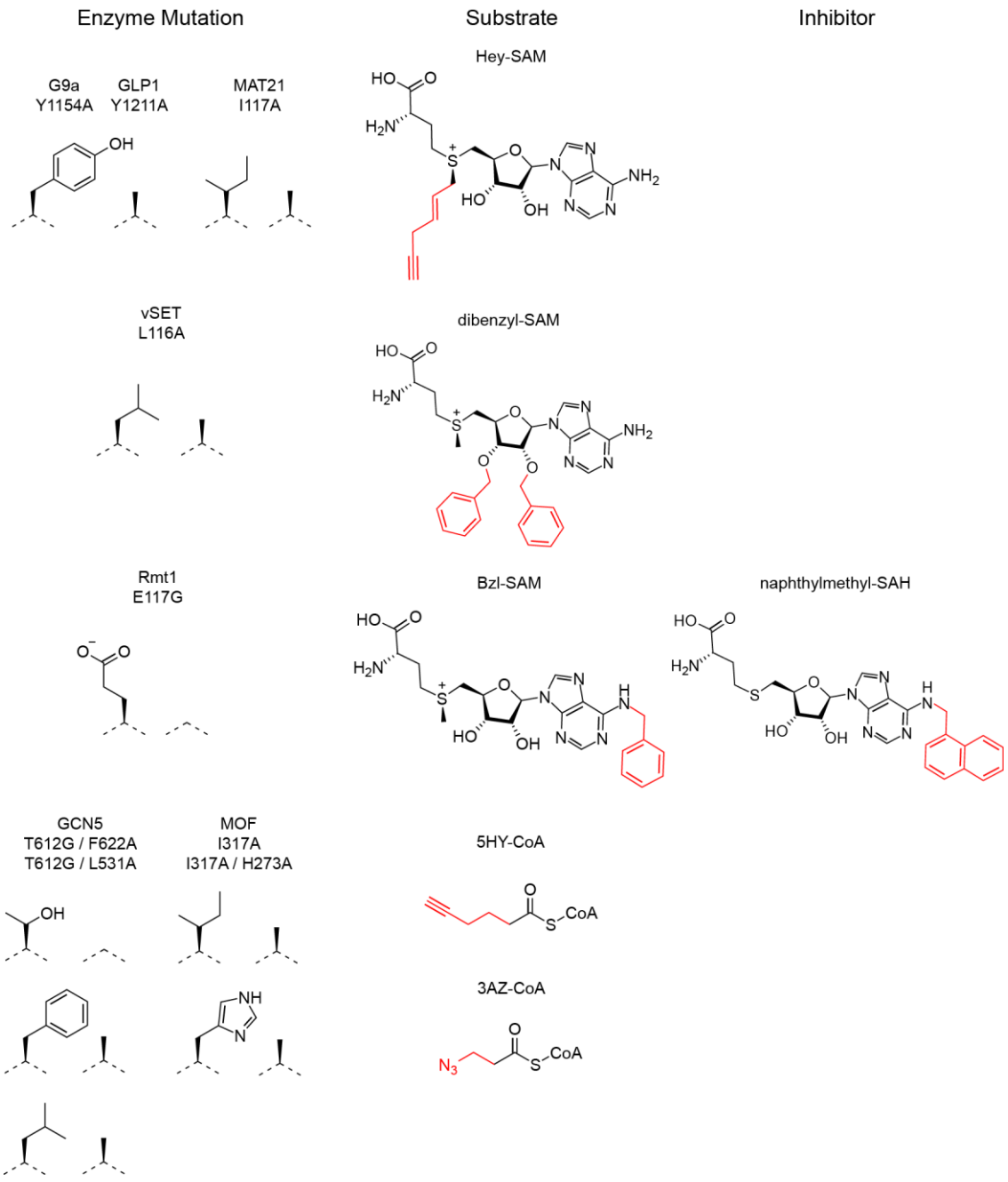
In a parallel study an orthogonal FKBP12 mutant and ligand pair was developed [146]. Mutations on the F36 residue were shown to retain binding to the FK506 ligand, and these mutants were then screened against a library of bumped FK506 compounds. The compound that emerged as best, bore a trimethoxyphenol bump, possessed high potency (IC_{50} = 2 nM) against a functional F36V mutant, and showed weak potency (IC_{50} = 3 μM) against the WT FKBP – resulting in outstanding selectivity of >3000-fold [146]. This compound (later termed Shield-1) was then coupled together to form a homobifunctional ligand (AP1903), able to recruit and dimerise two molecules of mutant FKBP protein. This system was then shown to activate Fas signalling in cells and in a mouse model [146]. The later inclusion of an additional L106P mutation created an unstable FKBP12 protein, which when fused to a protein of interest would trigger its rapid degradation unless stabilised by the bumped Shield-1 compound [147-149]. A fluorescent analogue of Shield-1 coupled with genetic tagging of the F36V FKBP12 mutant to a target gene allowed for fluorescent microscopy study of the tagged target protein in cells [150].

This system was later developed further, to develop the dTAG system for targeted protein degradation [151]. The dTAG system is premised upon the idea of creating fusion proteins of a target of interest and FKBP12, and targeting this protein with a FKBP12-targeting PROTAC molecule developed previously. As the FKBP12 component of the fusion protein is polyubiquitinated and degraded so too will the target protein be degraded, removing the need to develop a PROTAC molecule for each protein of interest. To avoid off-target effects emerging from endogenous FKBP12 degradation, a mutant F36V FKBP12 tag was utilised and a PROTAC using the trimethoxyphenol-bumped ligand [151]. The bumped PROTAC was shown to selectively degrade the F36V FKBP12 mutant in cells, and the dTAG system was then used to degrade other target proteins such as BRD4, MYC and KRAS G12V [151]. For a final validation the dTAG system was used to degrade a luciferase enzyme *in vivo*, by expressing the tagged luciferase in MV4-11 cells grafted into the bone marrow of mice [151].

1.11 – Bump-&-Hole Examples (epigenetics):

Within the context of epigenetics engineered orthogonal selectivity has been used, almost exclusively on the enzyme/cofactor level, to better understand the activity of epigenetic writer proteins (scheme 1.7) (table 1.2).

Over a series of publications a bump-&-hole system was developed for the HKMT enzymes G9a and GLP1, which are known to introduce epigenetic marks such as H3 K9 methylation, as well as to methylate NHPs such as the tumour suppressor p53 [152-154]. A number of bulky binding site residues (tyrosines, phenylalanines and a tryptophan) were replaced with alanine residues and tested for their ability to methylate H3 K9 peptides, with a Y/A mutation proving to be the most promising for both enzymes. The G9a Y1154A / GLP1 Y1211A mutants were tested against a range of SAM analogues wherein the donor CH_3 was replaced with long carbon chains with terminal alkyne groups (for click-chemistry). The Y/A mutants were able to process an (E)-hex-2-en-5-ynyl-SAM analogue (Hey-SAM) with only a small (<20%) drop in efficiency [153, 154] and very similar kinetics [152]. This system originally relied on treatment with exogenous Hey-SAM, until later work developed a mutant enzyme capable of synthesising Hey-SAM in cells. The SAM synthase MAT2A, which catalyses the formation of SAM from methionine and ATP, underwent a mutagenesis screen following structural analysis of its binding site. An I117L mutation was shown to be capable of producing Hey-SAM from ATP and exogenous Hey-Met, with no observable drop in efficiency [153]. This optimised system was used in a ‘clickable chromatin enrichment with parallel DNA sequencing’ (CliEn-seq) approach through which the genome-wide chromatin methylation activities of G9a and GLP1 were studied and substrate loci were identified to a previously unobtainable enzyme-specific level [153]. In addition, bioorthogonal profiling of protein methylation (BPPM), was used to identify non-histone substrates, by pulling down the substrates and carrying out LC-MS/MS analysis. These experiments identified a number of previously unknown G9a/GLP1 substrates, most of which were surprisingly revealed to be cytosolic, membrane bound or mitochondrial [152].



Scheme 1.7. Applications of the bump-&-hole system in epigenetics

Chemical structures of WT and mutated residues of epigenetic writer proteins (and SAM synthase MAT2A), allele-specific enzyme co-factors and inhibitors (modifications in red).

Another study attempted a similar strategy with the vSET HKMT from *Paramecium bursaria* chlorella virus, which specifically methylates H3 K27 and can be used as a more practical model for the SET-domain possessing EZH2 component of the PRC2 complex [155]. Rather than introducing a functional group (i.e. alkyne) this research aimed simply to generate a mutant vSET with no activity with the regular SAM co-factor, paired with a mutant-selective SAM analogue. Potential mutations were identified using a vSET/SAH/K3 K(27)me1 crystal structure, while bulky groups

were added to the 2' and 3'-hydroxyl groups of the SAM ribose unit. Mass spectrometry and kinetic analysis identified the ideal combination as that of the L116A mutation and a dibenzyl-modified SAM [155]. Although the SAM analogue is highly selective for the mutant vSET, the mutant enzyme was still more active with unmodified SAM. The lack of cofactor selectivity from the mutant enzyme limits the usability of this system in cells [155].

Enzyme	Substrate	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m ($M^{-1}S^{-1}$)
G9a WT	SAM	2.6	4	1.1×10^4
	Hey-SAM		-	
G9a Y1154A	SAM	0.1	39	4.3×10^1
	Hey-SAM	3.6	12	5.0×10^3
GLP1 WT	SAM	2.0	3	1.1×10^4
	Hey-SAM		-	
GPL1 Y1211A	SAM	0.1	36	4.6×10^1
	Hey-SAM	3.8	11	1.8×10^2
vSET WT	SAM	3.9	126	5.2×10^2
	dibenzyl-SAM	0.4	173	3.9×10^1
vSET L116A	SAM	7.4	60	2.1×10^3
	dibenzyl-SAM	1.3	72	3.0×10^2
RMT1 WT	SAM	1.1	6.7	2.7×10^3
	Bzl-SAM	3.2	4.9	1.1×10^4
RMT1 E117G	SAM		-	
	Bzl-SAM	18	0.2	1.5×10^6

Table 1.2. Kinetics of epigenetic enzyme mutants and cofactor analogues

Symbol '-' denotes no observed activity. Data from [152, 155, 156]

One other HKMT study focused on the yeast enzyme RMT1, and like the vSET study aimed to produce an orthogonal enzyme:cofactor pair, without the addition of functional chemical groups. An E117G mutation generated an enzyme with a ~4-fold drop in activity with SAM compared to WT [156]. Interestingly, given the accessibility of yeast genetics, the mutant RMT1 was incorporated into yeast cells, and displayed only a very minor drop in growth rates. A benzyl-modified SAM analogue was shown to be inactive with WT RMT1, but unfortunately also showed a ~100-fold drop in activity with the E117G mutant, limiting its utility [156]. This study did not just generate an orthogonal enzyme cofactor, but also produced two SAH analogues that acted as

inhibitors with >20-fold selectivity for the mutant enzyme [156]. This selectivity was shown both kinetically, and through the selective inhibition of downstream methylation in WT and mutant cell extracts. One selective inhibitor was then used, with the mutant yeast cells, to monitor changes in gene expression, identifying several genes involved in cellular organisation and metabolism as regulated by RMT1 [156]. As the selective inhibitors possessed low potency (K_i ~5 μM) these experiments required high doses (e.g. 500 μM) and the need for improved inhibitors was identified [156].

The final example of a bump-&-hole system for an epigenetic enzyme is a 2013 study on a number of lysine acetyltransferase (KAT) enzymes. The main enzymes tested were GCN5 and MOF, from 2 separate KAT families, important to biological acetylation [157]. Like the studies into G9a and GLP1 this work aimed not just to engineer selectivity but also to introduce reactive groups (alkynes and azides) onto substrate proteins for further analysis. A number of mutation/cofactor pairs were identified with varying properties: 5HY-CoA was effective with GCN5 T612G/F622A and TG12G/L531A mutants and MOF I317A and I317A/H273A mutants, while 3AZ-CoA was effective with MOF I137A [157]. The ability to react with, and identify substrate proteins incorporating the modified analogues was validated through fluorescent labelling and SDS-PAGE. This bump-&-hole system was not actually used to study the full ‘acetylomes’ of these enzymes, perhaps due to difficulties in introducing modified acetyl-CoA analogues into cells and the lack of comprehensive kinetics data [157]. As such this research would greatly benefit from the development of a mutant enzyme for the biosynthesis of modified acetyl-CoA cofactors, as was done in [153].

1.12 – A 'Bump-&-Hole' System for BET Bromodomains

Moving beyond writer enzymes into reader domains, the BET bromodomains provide, in principle, an ideal opportunity and a compelling case for the application of the bump-&-hole approach for developing an allele-selective, orthogonal protein:ligand pair. Many high quality, but pan-selective, chemical probes already exist and a wealth of chemical, SAR and cellular data is available in the literature, which should aid the development of allele-selective analogues without the need to develop new probes *de novo*. Additionally, the study of the BET proteins/bromodomains (both regarding basic biology but also for therapeutic use) would greatly benefit from the ability to selectively inhibit specific BET proteins and bromodomains. Pan-BET inhibition is already associated with several phenotypes, some of which are therapeutically advantageous or deleterious. In this context the BET proteins are often viewed as a monolith, with a list of biological roles. In reality this is much more complicated (figure 1.14) with each BET protein (and BD) generating different phenotypes when inhibited. In some cases different BET protein/BDs have been shown to have directly opposing roles [128, 158]. BET protein and BD-specific inhibition should reveal and dissect these differing roles, and may identify BET BDs whose inhibition generates therapeutically beneficial phenotypes, and others with more deleterious roles such as the observed on-target toxicity. This information would then go on to inform the development of selective BET inhibitors that may show more promise in the clinic. Selectively interrogating each BET bromodomain traditionally, through developing eight bromodomain-specific probes, might be unattainable given the high conservation of the binding pockets described, and even if achievable it would take a great deal of time and money, and would be hard to justify without the very bromodomain-specific phenotypic data we wish to obtain. By being able to (relatively) quickly engineer selective inhibition of all eight BET BDs we would have a new tool at our disposal that can help to better understand these proteins on an individual level and generate the data needed for the proper targeting of selective BET inhibitors.

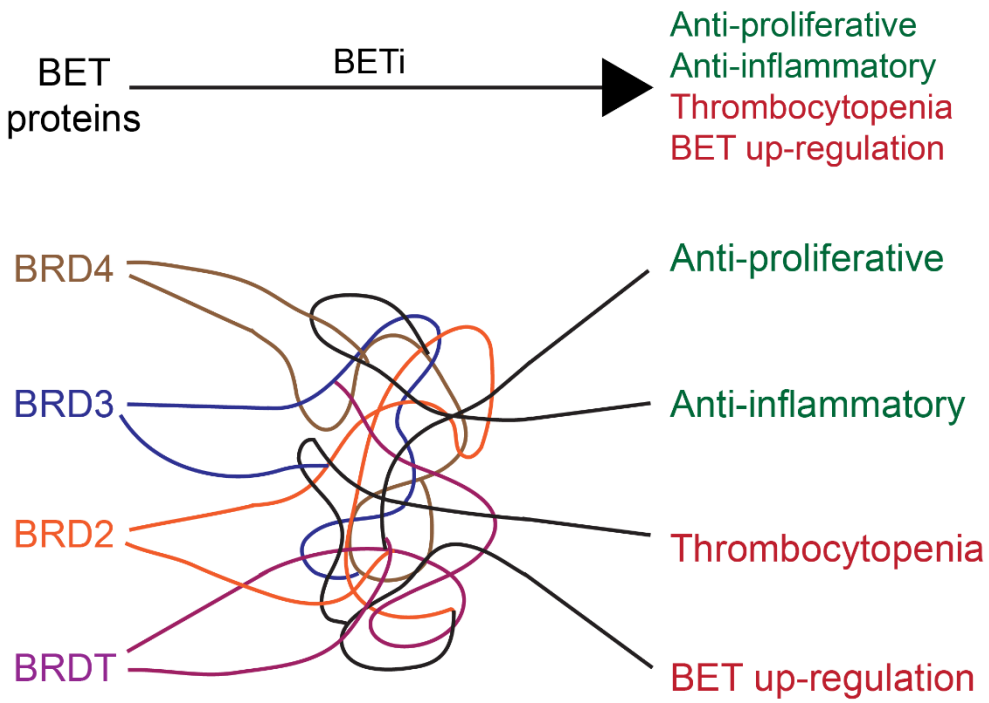


Figure 1.14. Differing BET protein functions obscured by pan-selective inhibition

Inhibiting all BET bromodomains with pan-selective BET inhibitors reveals a number of positive (green) and negative (red) biological effects. These phenotypes are actually generated by the inhibition of 8 BET bromodomains, with varying (if not directly opposed) biological functions. The exact roles of different BDs, and how they interact to generate observed phenotypes, remains poorly understood.

This work also provides an opportunity to broaden the general applicability of the bump-&-hole approach, as this system is rarely applied to small-molecule inhibitors, and instead typically focuses on enzymes and co-factors. In particular, the bump-&-hole approach had not previously been applied in the context of a protein-protein interaction, and there was some scepticism that it would be possible to engineer sufficient selectivity within a PPI binding site.

The Ciulli laboratory decided to take on the challenge of developing allele-selective PPI inhibitors in 2010, using the BET bromodomain as the system of choice. Since then development of the BET bump-&-hole system has required a combined effort using a number of disciplines such as computational biology, structural biology, chemistry and cell biology. Starting with initial analysis of BET bromodomain sequences and structures, followed by computational simulations of the impact of mutations, three binding site residues were identified as potential ways to engineer selectivity. These residues were (BRD2 BD1 numbering): W81, V103, and L110 [122] (figure 1.15).

The valine/alanine mutation (V/A) was quickly discarded, as initial characterisation revealed it to be highly unstable and showed the greatest loss in affinity for a tetra-acetylated H4 peptide. Additionally, thermal shift experiments suggested that the ligands designed as V/A-targeting were not selective for the mutants over WT [122].

The tryptophan mutations: W/F and W/H were more successful, but ultimately did not lead to a usable system. The W/F mutation had only a minor impact on the affinity for acetylated peptides, which was deemed a positive feature of these mutant proteins. Unfortunately, again the corresponding ligands were not sufficiently selective [122]. Nonetheless this research did generate some useful outputs. To design W/F-targeting compounds, chemical means of substituting the chlorophenyl ring of I-BET762 were developed, which opened up possibilities for developing new, non-orthogonal, BET inhibitors. Additionally, as discussed previously one of the indole compounds designed to target the W/F mutations showed unexpected BD2-selectivIty [122].

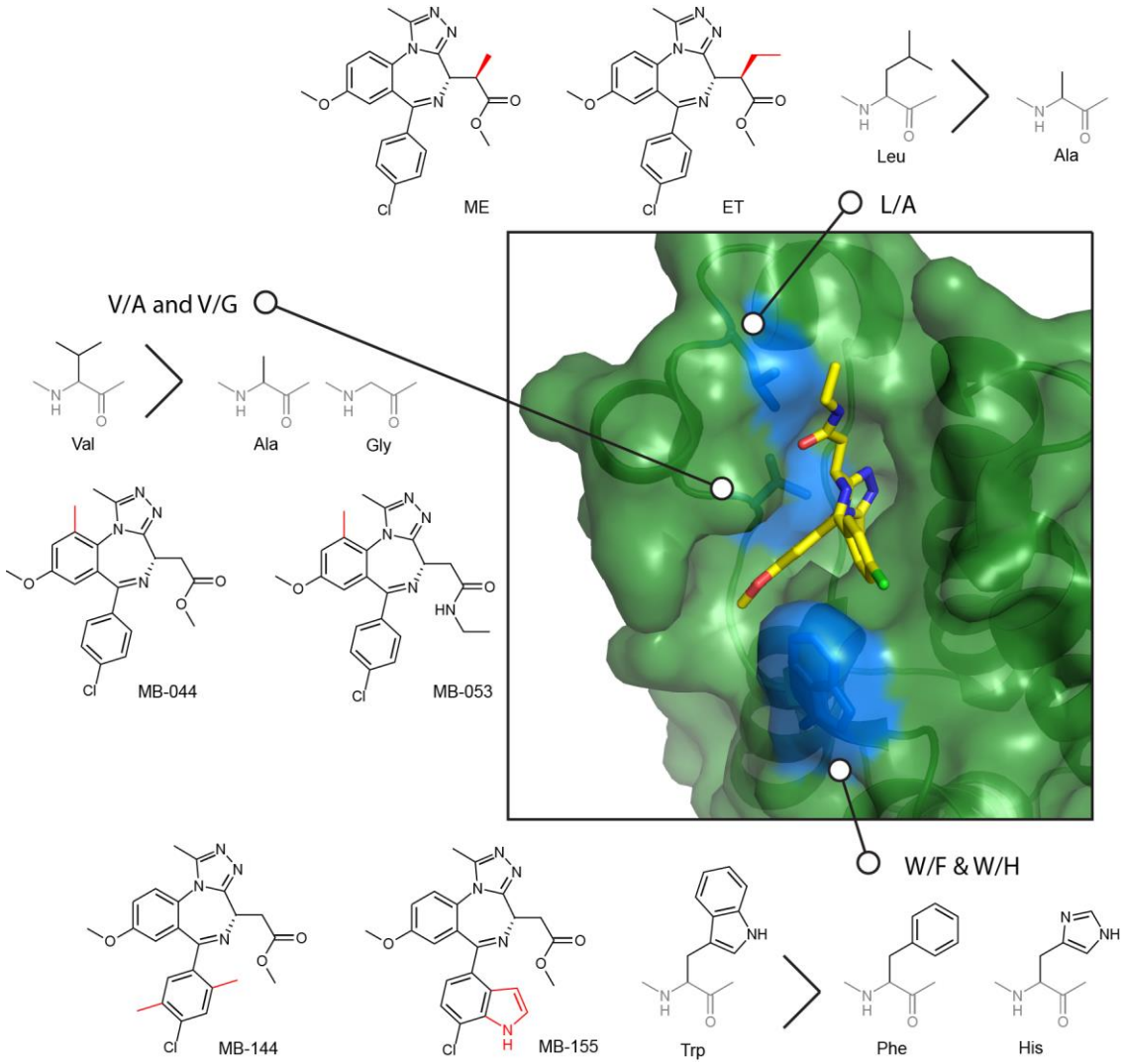


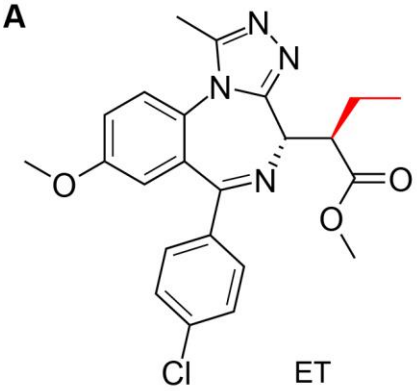
Figure 1.15. Alternative BET bump-&-hole mutations & ligands.

Co-crystal structure of BRD2 BD1 in complex with I-BET762 (2YEK), with mutated residues highlighted. Chemical structures of WT and mutant amino acids, and mutant-targeting, I-BET762-derived ligands (bumps in red). MB-155 is earlier referred to as TC AC 28.

The leucine mutants, specifically the leucine/alanine (L/A) was by far the most successful strategy. Its success led to the first publication from the group on the BET bump-&-hole approach and is the foundation for the research in this thesis [159]. This initial, proof-of-concept, publication showed that the L/A BET bromodomain mutants retained the stability and overall structure of the WT domain and the ability to bind acetylated histone peptides [159]. It also discussed the synthesis of its allele-selective inhibitors. These ligands utilised the benzodiazepine scaffold, identical to I-BET762 but with a methyl-ester replacing the di-ethyl-amide moiety. Alkyl bumps on the 2nd carbon (adjacent to the ester Cα) were shown to generate potent and highly-selective inhibitors ME (with methyl bump) and ET (with ethyl bump). ET was shown to be

potent and selective across all BET bromodomains, with an average L/A K_d value of ~100 nM and ~160-fold selectivity [159] (figure 1.16A-C). The system was validated through ITC experiments with tandem-BD constructs and using a cellular FRAP assay.

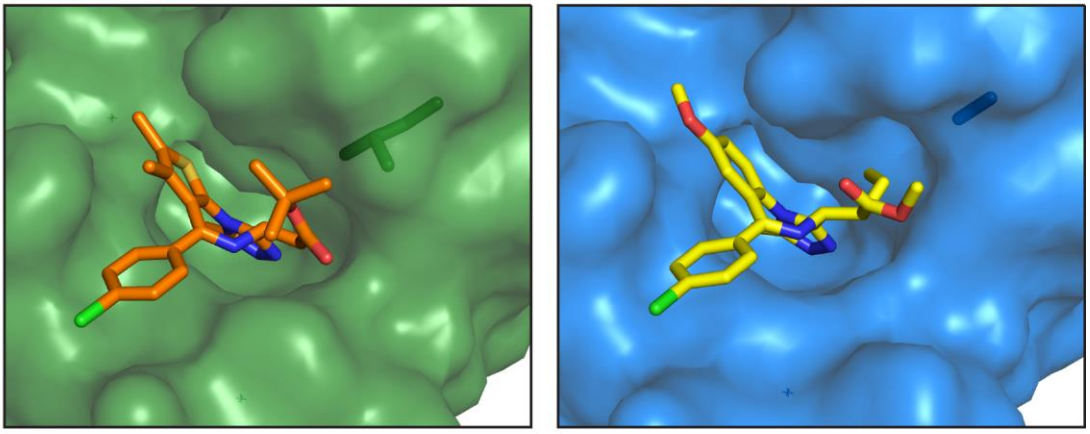
A



B

ET	L/A	2.2	2.2	3.1	3.3	4.1	4.2
WT	Kd (nM)	60	50	40	80	40	200
2.1	9300	150	180	230	120	230	50
2.2	17200	280	340	430	220	430	90
3.1	6100	100	120	150	80	150	30
3.2	6500	100	130	160	80	160	30
4.1	8600	140	170	210	110	210	40
4.2	9200	150	180	230	120	230	50

C



D

H4 tetra	BRD2 BD1		BRD4 BD1	
Kac	WT	L/A	WT	L/A
K_d (µM)	27	121	10	94
SE (µM)	1	5	1	3
LV/WT	4		9	

E

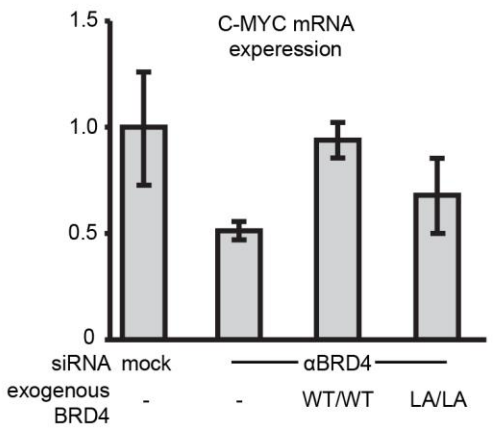


Figure 1.16. BET bump-&-hole proof-of-concept

A) Chemical structure of bumped ET compound, with ethyl bump highlighted. B) Selectivity box of ET for L/A vs WT BET bromodomains. C) Co-crystal structures of (+)-JQ1 bound to BRD2 BD2 WT (3ONI) and ET bound to BRD2 BD2 L383A (4QEW), with leucine/alanine residue highlighted. D) ITC-measured affinity of WT and L/A BET BDs to tetra-acetylated H4 peptide. E) Expression of C-MYC mRNA in U2OS cells, treated with mock or BRD4-targeting siRNA. U2OS cells stably transfected with exogenous WT or L/A BRD4, induced by treatment with 0.1 µg/ml tetracycline. Selectivity, peptide-binding and C-MYC expression data from Baud et al, Science, 2014 [159].

As successful as the initial research was, there was a clear need for further optimisation before biological questions could be properly addressed. The L/A mutation allowed for potent and selective inhibition while retaining stability and basic functionality but further experiments, however, showed that L/A BET proteins would be significantly less functional in cells. This could result in mutant cell-lines or animal models being non-viable (a severe waste of resources) or, if less severe, reducing the value of any results obtained using the system. The L/A BET bromodomains were able to bind acetylated histone peptides, but with up to 10-fold reductions in affinity for the BD1s. In a cellular rescue assay, L/A BRD4 was not able to induce the expression of C-MYC as well as the WT form [159] (figure 1.16D-E). As a result, alternative leucine mutations were designed and trialled to identify a mutation that, while still allowing for selective inhibition, would likely have a more conservative impact on protein function. On the chemical side of the project, it was hoped that examination of additional modifications would allow for increased selectivity, potency and DMPK characteristics. This optimised bump-&-hole system could then be validated in cells and, once introduced into appropriate cell-lines, be applied to answer biological questions in a BET-family wide fashion.

1.13 – Aims & Objectives:

The goal of this PhD project was to optimise the bump-&-hole approach for selective inhibition of BET bromodomains and utilise it to investigate the functions of the different BET proteins, hopefully gaining important biological insights that will also have therapeutic implications.

The first aim of the project was to optimise the 'biology' aspect of the bump-&-hole system – the mutation. Alternative leucine mutants will be characterised, looking at stability, structure and their 'functionality' (ability to bind acetylated histone peptides / acetylated chromatin). This work will hopefully identify a mutation with little-to-no impact on bromodomain function while allowing for selective and potent inhibition.

The second aim was to optimise the 'chemistry' aspect of the system – the bumped compounds. The impacts of new chemical modifications, and the quality of resulting compounds, will be evaluated with regard to selectivity, potency and DMPK characteristics. Bromodomain:ligand co-crystal structures will be solved to rationalise observed SAR trends. The best compound(s) will be selected based on desirable criteria: the best compounds should possess >100-fold selectivity for the mutants, high potency (ideally 100 nM) and strong DMPK qualities (stability and cell permeability). Another way to improve upon the chemistry of the initial system is to develop a means to separate the active and inactive enantiomers of the best compounds, instead of using the racemic mixture as in the previous publication [159]. Finally, a selection of protein panels will be used to screen for off-targets. This will focus on non-BET bromodomains but will also check for unexpected, unrelated off-targets across other druggable protein families e.g. G-protein-coupled receptors (GPCRs) that would impact the usability of our compounds.

Following the system's optimisation, the next aim will be to validate it in cells. Cellular assays will be developed and optimised to confirm the ability of bumped compounds to inhibit mutant BET proteins without effecting WT BET proteins or, through off-target activity, altering the phenotype of WT cell-lines.

As a final aim, the optimised bump-&-hole system will then be utilised to answer biological questions concerning the BET proteins/bromodomains. Following on

from previous experiments [159], the optimised system will be used in cellular assays such as FRAP to compare the relative importance to chromatin binding for each BET proteins BD1 and BD2. More advanced biological investigations will likely require the development of genome editing approaches to generate mutant cell-lines. These cell-lines could be engineered to express mutant BET proteins at endogenous levels in place of the WT, without the need for transient overexpression of exogenous mutant proteins and concomitant knock-down of WT endogenous proteins.

The aims & objectives may be summarised as:

1 – Optimise biology
- Characterise alternate mutation
  - Mutant stability
  - Mutant structure
  - Mutant functionality

2 – Optimise chemistry
- Evaluate new modifications/compounds, based on selectivity, potency & DMPK
- Triage and select best compound(s)
- Separate and characterise enantiomers
- Solve co-crystal structures with new mutant bound to bumped compounds
- Screen for off-targets

3 – Validate system
- Assess mutant functionality in cells
- Measure compound efficacy/selectivity in cells
- Check bumped compounds for impact on WT cells

4 – Use optimised system to answer biological questions
- Compare importance of BD1s and BD2s to each BET protein's function
- Develop cell-lines incorporating new mutation
- Use bumped compounds with mutant cell-lines for more advanced biological questions

Chapter 2
L/V Mutant Characterisation

2.1 – Introduction:

The first iteration of the BET bump-&-hole approach utilised a leucine/alanine mutation, which allowed for strong selectivity by bumped compounds while retaining stability and the ability to bind acetylated histone peptides [159]. A deeper examination, however, raised concerns that introduction of an L/A mutation on BET proteins might impact on their cellular functions. This could compromise the viability of mutant cell-lines and render any obtained results less relevant to WT biology. In response alternate leucine mutations were trialled, with the aim of finding a mutation with minimal impact on protein function while allowing for selective inhibition by bumped compounds.

Two other leucine mutants have been trialled in the bump-&-hole system, but ultimately were not investigated to the same extent as the L/A mutation (figure S2.1). In an attempt to boost selectivity through generating a larger 'hole' a leucine/glycine mutation was trialled early on. Although the bromodomain was not destabilised its functionality was greatly affected, and in isothermal titration calorimetry (ITC) experiments it showed no sign of binding an acetylated H4 peptide. A leucine/isoleucine mutation met with more success, with early peptide ITC experiments suggesting very strong functionality. Despite having the same mass as leucine, the isoleucine residue did seem to create some sort of 'hole' allowing for selective inhibition. Unfortunately, obtaining high selectivity proved difficult, with ET showing only ~10-fold selectivity.

With the L/A mutation proving too disruptive and L/I hindering efforts to achieve selectivity a leucine/valine mutation was selected as the obvious next step (figure 2.1). Valine remains chemically similar to the original leucine residue: hydrophobic, non-polar, uncharged and lacking carbon cycles. Valine retains the branch at the end of its side-chain and is situated between leucine and alanine in terms of size. As such L/V mutations were introduced into BET constructs which were then purified for a thorough characterisation to investigate the impact of the mutation on structure, stability and function.

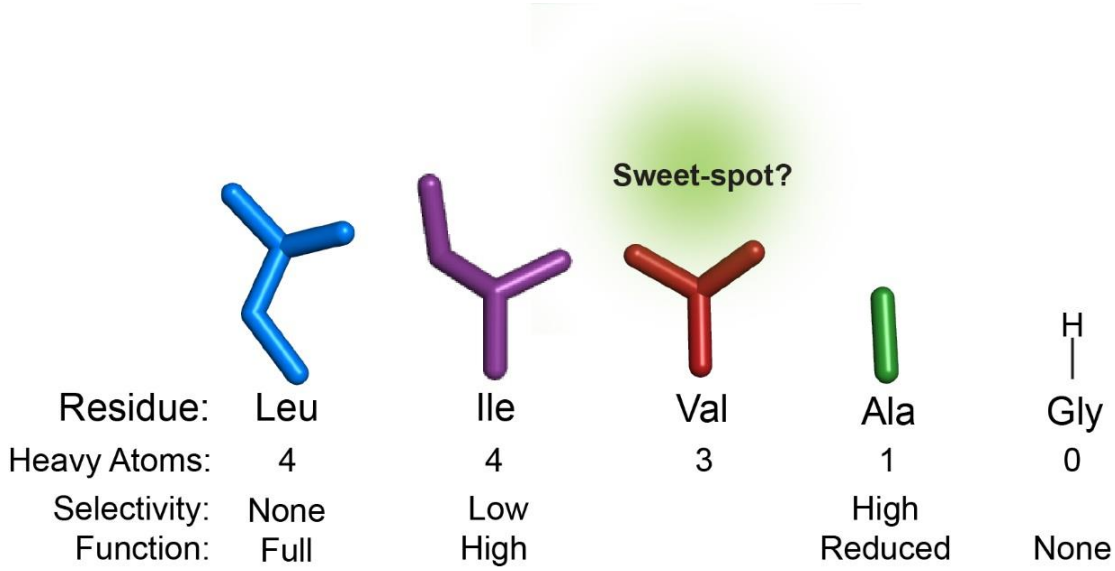


Figure 2.1. Amino-acid side-chain comparison.

Comparison of WT leucine and alternate residues. Comparison of side-chain structure and mass (excluding Cα), selectivity of bumped compounds for mutant and ability of mutant bromodomains to bind acetylated histone peptides. Residue side chains shown as sticks, and prepared in PyMol.

2.2 – Bromodomain Mutation & Purification:

L/V mutant constructs were created by site-directed mutagenesis of previously cloned WT-encoding plasmids (residue numbering in figure S2.2). pNIC28 plasmids, for single-bromodomain purification, and pcDNA5 plasmids, encoding GFP-BRD4 for cellular experiments were mutated by Dr Kwok-Ho Chan. pcDNA5 GFP-BRD3 plasmids were mutated by Lars van Beurden. pcDNA5 GFP-BRD2 and pcDNA6 GFP-BRDT plasmids were mutated by myself. Mutagenesis was straightforward, although LV/LV double-mutant GFP-BET plasmids could not be created through a single, double-mutation protocol, and instead through 2 rounds of mutagenesis.

Single bromodomain constructs, used for the bulk of *in vitro* work, were expressed in *E. coli* BL21 (DE3) cells and purified by Ni^{+} column and then size exclusion chromatography. This protocol was developed in the lab, as described in Baud et al. [159] and was utilised for this project by several researchers. Final protein samples were consistently of high purity (figure S2.3) and gave final yields from ~2 mg/ml (BRDT BD2 and BRD3 BD1) to ~20 mg/ml (BRD2 BD2 and BRD4 BD2). Helpfully, this purification could be carried out at room temperature with no issues, an early indicator that the L/V mutation did not destabilise the bromodomains.

2.3 – Stability of Mutant BET Bromodomains:

To confirm that the L/V mutation did not destabilise the BET bromodomains differential scanning fluorimetry (DSF) was performed. DSF is used to measure the 'melting temperature' (T_m) of proteins through the use of the SYPRO orange dye. As proteins are heated and begin to denature this dye interacts with newly exposed hydrophobic regions, where it is protected from the quenching effects of water and its quantum yield greatly increases. By incubating proteins with SYPRO orange and heating them step-by-step while measuring fluorescence a sigmoid curve is generated and its inflection point used to calculate the T_m. These experiments showed that the L/V mutation did not greatly affect the stability of any BET bromodomains (figure 2.2), with melting temperatures dropping by only 1-3°C (broadly consistent with the impact of the previous L/A mutation).

A

Bromodomain	Tm (°C) ±SE			ΔTm (°C) ±SE	
	WT	L/V	L/A	L/V	L/A
BRD2 BD1	45 ±0.0	43 ±0.6	45 ±0.5	-2 ±0.6	-1 ±0.6
BRD2 BD2	49 ±0.1	48 ±0.1	48 ±0.3	-1 ±0.1	-1 ±0.3
BRD3 BD1	46 ±0.1	43 ±0.2	44 ±0.1	-3 ±0.2	-2 ±0.1
BRD3 BD2	44 ±0.1	45 ±0.9	44 ±0.0	1 ±0.9	0 ±0.1
BRD4 BD1	43 ±0.2	41 ±0.2	44 ±0.2	-2 ±0.3	0 ±0.3
BRD4 BD2	46 ±0.1	45 ±0.3	43 ±0.1	-1 ±0.3	-4 ±0.1
BRDT BD1	45 ±0.2	45 ±0.1	43 ±0.2	1 ±0.2	-1 ±0.3
BRDT BD2	44 ±0.1	42 ±1.6	47 ±0.1	-2 ±1.6	3 ±0.2

B

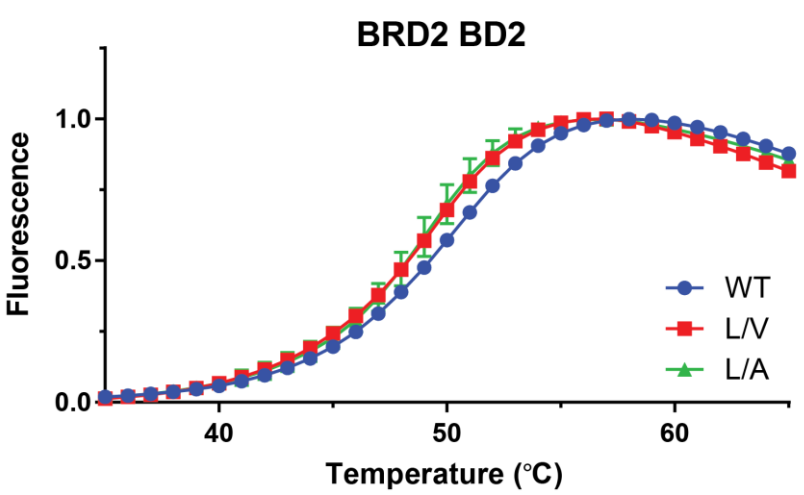


Figure 2.2. Thermal stability of BET bromodomain mutants.

A) Differential Scanning Fluorimetry (DSF)-derived melting points of WT, L/V and L/A BET single-bromodomain constructs. Mean and standard error of three technical replicates. B) Normalised fluorescence of SYPRO orange against temperature, when mixed with BRD2 BD2 bromodomain constructs.

2.4 – Effect of L/V Mutation on Histone Peptide Binding (ITC):

Moving on from structural stability, the 'functionality' of the BET bromodomain mutants was assessed. To this end we investigated the impact of the L/V mutation on the recognition and binding of acetylated histone peptides. These acetylated histone peptides act as binding epitope models for acetylated chromatin, which the mutant BET proteins must be able to recognise and bind to in order to drive downstream gene expression. Bromodomain binding to purified, acetylated nucleosomes would represent a biophysical model more closely representative of the endogenous biological system, and is described in the literature [34]. The use of acetylated nucleosomes would require either the purification of nucleosomes from human cells, which is a material-intensive process, and can create difficulties in obtaining nucleosomes with specific PTMs [160]. Alternatively, nucleosomes can be produced from the assembly of modified histone monomers, but this is also very time and material-consuming and has a low throughput [161]. Due to these practical limitations nucleosome-based assays were deprioritised in place of easily obtained synthetic acetylated histone peptides.

ITC [162] was used to quantitatively assess the effect of the L/V mutation on the affinity of BET bromodomains for di- and tetra-acetylated H4 histone peptides: H4 K(5,8)ac and H4 K(5,8,12,16)ac (commonly used in the literature as BET bromodomain substrates [32]). Titrations of peptides against WT and L/V bromodomains showed a ~2-fold decrease in affinity (largest increase 2.4-fold, smallest 1.2-fold) (table 2.1). This effect is small and close to experimental errors but was also fairly consistent. This marks a clear improvement over the L/A mutants, which previously showed up to 10-fold decreases in affinity [159]. To confirm this improved functionality, H4 K(5,8,12,16)ac was titrated against BRD2 BD1 and BRD4 BD1 L/A, and showed much weaker binding and K_d values ~10-fold higher than against WT (figure 2.3A). The impact of the L/V mutation on binding affinity did not seem to change significantly between peptides (p = 0.3) or between BD1s and BD2s (p = 0.7). Our acetylated peptides interacted more strongly with BD1s (17 µM average K_d) than with BD2s (93 µM average).

H4 K5,8ac	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2		BRD4 BD1		BRD4 BD2		BRDT BD1	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V
Kd (μM)	31	42	91	138	18	22	105	154	7.6	9.6	83	198	23	40
SE	2.0	3.0	19.0	16.0	1.0	1.0	12.0	16.0	1.0	1.0	2.0	22.0	4.0	8.0
ΔH (kcal/mol)	-8.8	-9.8	-1.7	-1.4	-9.4	-10.7	-6.7	-11.6	-10.5	-10.9	-4.1	-5.6	-14.4	-16.9
SE	0.2	0.5	0.2	0.2	0.1	0.2	1.1	5.7	0.1	0.1	0.1	0.9	1.4	3.0
N	2.1	1.8	2.3	2.0	2.0	1.7	1.2	0.5	1.7	1.7	1.4	1.6	1.3	1.1
SE	0.04	0.06	1.14	0.17	0.02	0.02	0.18	0.22	0.02	0.01	0.02	0.21	0.10	0.16

H4 K5,8, 12,16ac	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2		BRD4 BD1		BRD4 BD2		BRDT BD1	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V
Kd (μM)	16	34	133	169	20	36	63	128	9	13	83	138	15	33
SE	1.0	1.0	7.0	12.0	1.0	3.0	6.0	7.0	1.0	1.0	3.0	10.0	1.0	4.0
ΔH (kcal/mol)	-11.9	-15.7	-6.3	-3.8	-9.1	-10.5	-9.3	-7.0	-9.7	-9.8	-8.3	-7.3	-14.8	-14.9
SE	0.1	0.2	0.4	0.3	0.1	0.3	0.9	0.2	0.1	0.1	0.2	0.4	0.4	1.0
N	1.1	1.2	1.1	1.5	2.3	2.0	0.9	1.0	1.8	1.6	0.9	1.1	1.3	1.1
SE	0.01	0.01	0.05	0.08	0.01	0.03	0.08	(FIX)	0.02	0.02	0.02	0.04	0.02	0.06

Table 2.1. Effect of L/V mutation on BET bromodomain binding to acetylated H4 peptides.
Results of ITC titrations of di- and tetra-acetylated H4 peptides against WT and L/V BET bromodomains constructs. For some weak interactions N was fixed to 1.0.

In addition to measure binding affinity, ITC also allows for analysis of the thermodynamic parameters of an interaction – enthalpy (ΔH), entropy (TΔS) and Gibbs free-energy (ΔG). The L/V mutation did not substantially alter these values for any bromodomain:peptide interactions, suggesting that the binding mode of these interactions has not been changed (i.e. from principally hydrogen bond-mediated to hydrophobically driven). Additionally the titrations of peptides against L/A bromodomains shows much greater changes in the thermodynamic parameters, consistent with this mutation having a much more deleterious effect on binding (figure 2.3B).

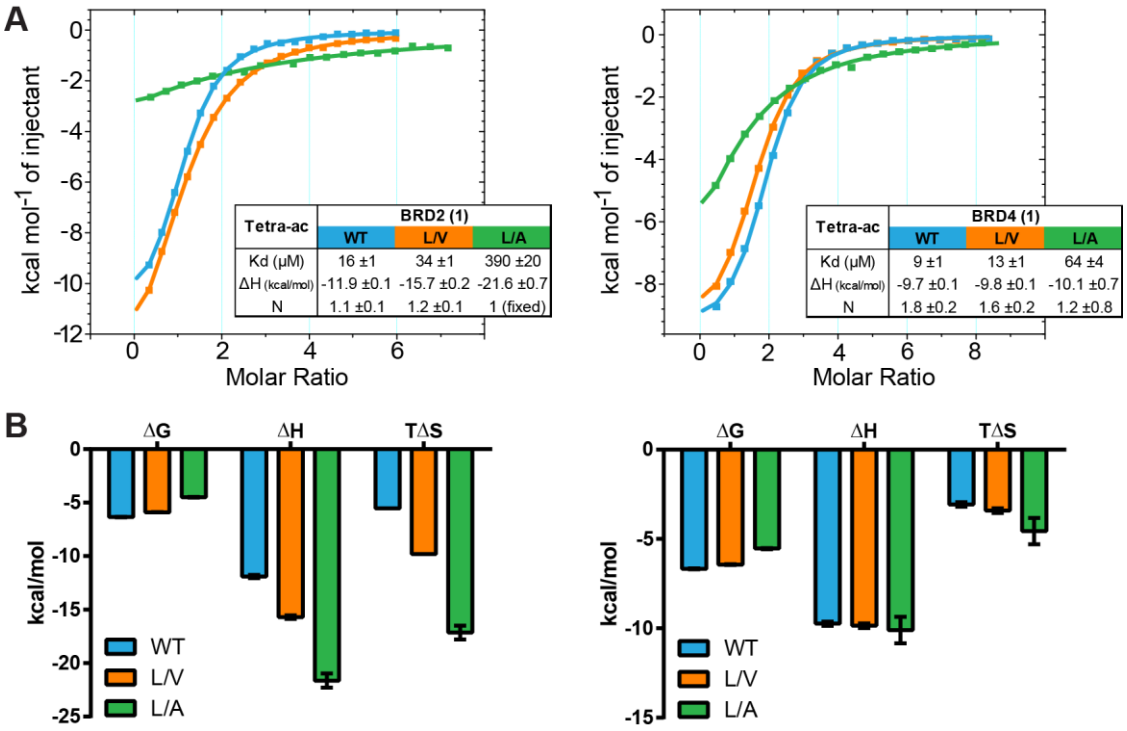


Figure 2.3. The impact of L/V and L/A mutations on BET bromodomain binding to acetylated H4 peptides.

Results of ITC titration of H4 K(5,8,12,16)ac peptide against BRD2 BD1 and BRD4 BD1 bromodomain constructs.

2.5. – Effect of L/V Mutation on Histone Peptide Binding (BLI):

Our peptide ITC titrations gave powerful insights into how the L/V mutation affected a few select BE:peptide interactions. The ITC work, however, did not cover many different acetylation marks or combinations thereof, and ITC has too low a throughput to study a large library of acetylated histone peptides. Without studying a larger number of Kac-combinations it cannot be determined if the L/V mutation has altered the specificity of the BET bromodomains (i.e. which Kac marks/combination they do and do not recognise). To determine this, a library of histone peptides, bearing many combinations of Kac modifications, was next purchased and screened against WT and L/V bromodomains through bio-layer interferometry (BLI) [163]. This gave a single-dose response for each peptide against each bromodomain, and normalising these responses for each construct gave a number of binding profiles to compare (figure 2.4). All WT bromodomains, and their L/V mutants, show a marked preference for a selection of poly-acetylated H4 peptides such as H4 K(5,8)ac, H4 K(5,8,12,16)ac and H4 K(5,8,12,16)ac, consistent with previous investigations showing the di-acetylation of lysines 5 and 8 to be optimal for BET bromodomain binding [32]. A visual comparison shows that the L/V mutation has had little impact on the binding profiles of our bromodomains, with several of the BD1s showing almost identical binding profiles with WT and L/V constructs.

The L/V mutation is shown to be less disruptive than the L/A mutation when compared to WT. The L/A mutation in the binding profiles from [159] shows a much more severe impact, often resulting in more promiscuous binding to di- and tri-acetylated H4 peptides. For a direct comparison BRD4 BD1 L/A was re-tested alongside its WT and L/V counterparts, confirming the supremacy of the L/V mutation (figure 2.5). Non H4 peptides showed significantly weaker responses, and again little difference between WT and L/V constructs (figure S2.4). Some example curves for the BRD4 BD1 constructs are presented in figure S2.5.

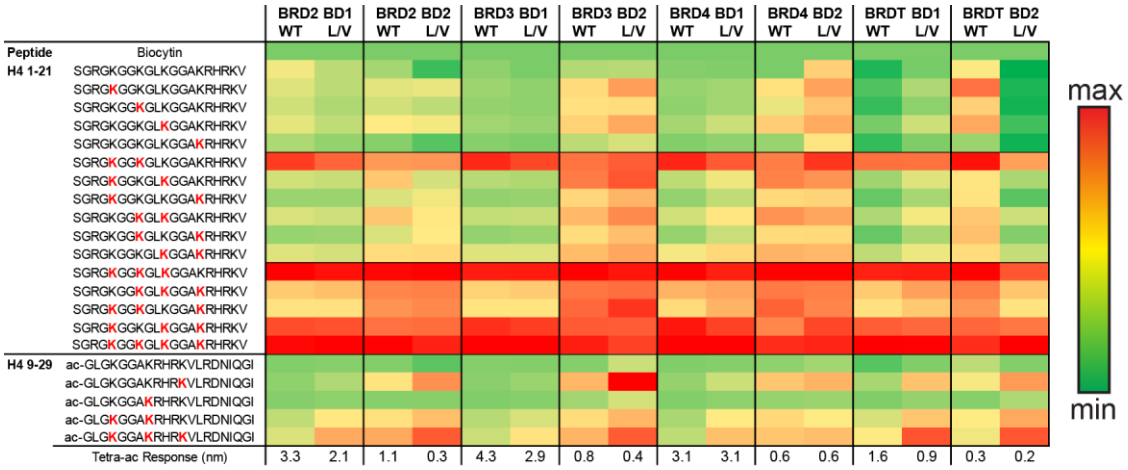


Figure 2.4. The effect of the L/V mutation on binding profiles of BET bromodomains.

Binding profiles of WT and L/V BET bromodomains for acetylated H4 peptides, derived from BLI screen. Responses normalised to strongest response of each construct, and colour-coded. Red 'K' in sequence denotes lysine acetylation.

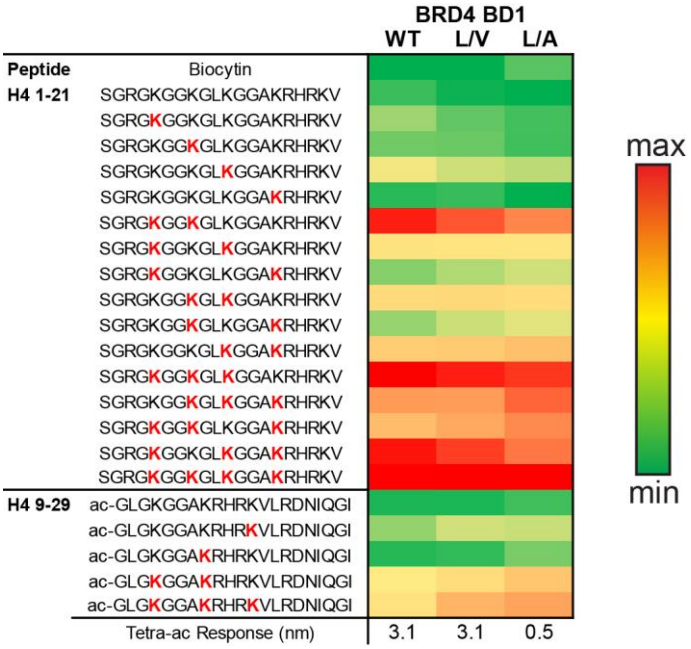


Figure 2.5. BRD4 BD1 histone peptide binding profiles.

Binding profiles of WT, L/V and L/A BRD4 BD1 constructs for acetylated H4 peptides, derived from BLI screen. Responses normalised to strongest response of each construct, and colour-coded. Red 'K' in sequence denotes lysine acetylation.

2.6 – Structural Impacts of BET Bromodomain Mutation:

In an effort to rationalise the differing impacts of the L/V and L/A mutation on histone binding the X-ray crystal structure of BRD2 BD2 L/V *apo* was solved by Dr Kwok-Ho Chan, and compared to the previously solved WT *apo* structure (figure 2.6). Both mutants retain the broad, overall structure, and when aligned (in PyMol) with the WT structure the L/V mutant generated RMSD values of 0.15 Å and 0.73 Å (with and without outlier rejection), while the L/A mutant gives 0.31 Å and 1.43 Å. Close analysis of the binding pocket shows the L/V mutant to perfectly mirror the WT structure, while the L/A mutant shows a noticeable re-orientation of the ZA loop upon which the mutated alanine resides. The L/V apo structure was also compared to previously solved structures of WT BRD2 BD2 bound to two peptides bearing acetylated lysine marks. As expected, the L/V structure aligns extremely well with the WT bromodomain bound to a histone peptide, and similarly well to the WT bromodomain bound to a potential non-histone binding partner (Stat3 K(87)ac [112]). Although these promising alignments cannot tell us that the L/V mutant will show high functionality they do suggest that there are no major concerns.

Comparing the apo crystal structures may also shed light on the reduced functionality of the L/A mutant. The ZA-loop re-orientation to a 'closed' state seen in the L/A mutant may be an attempt to 'plug' the void created by the mutation (figure 2.6), and when bound to ligands this reverts to the WT-style 'open' conformation (figure S2.6). This suggested a need to re-orient the ZA loop during binding which might contribute to the loss in affinity between L/A bromodomains and histone peptides. It must be noted that the L/A mutant crystal structure presents an alternate space group to those of the WT and L/V bromodomains (C1,2,1 vs P21,21,2) and changes to crystal packing around the ZA loop (figure S2.7). With this data it cannot be determined whether the alternate conformation presented by the L/A mutant is caused by these crystallisation changes or if the unusual conformation has altered how the protein crystallises.

This 'ZA-loop reorganisation' hypothesis could be better tested by crystallising each BRD2 BD2 construct under similar crystallisation conditions and obtaining identical space groups (if possible). A better response may be to use NMR to solve the structure of these constructs in solution, without crystallisation artefacts.

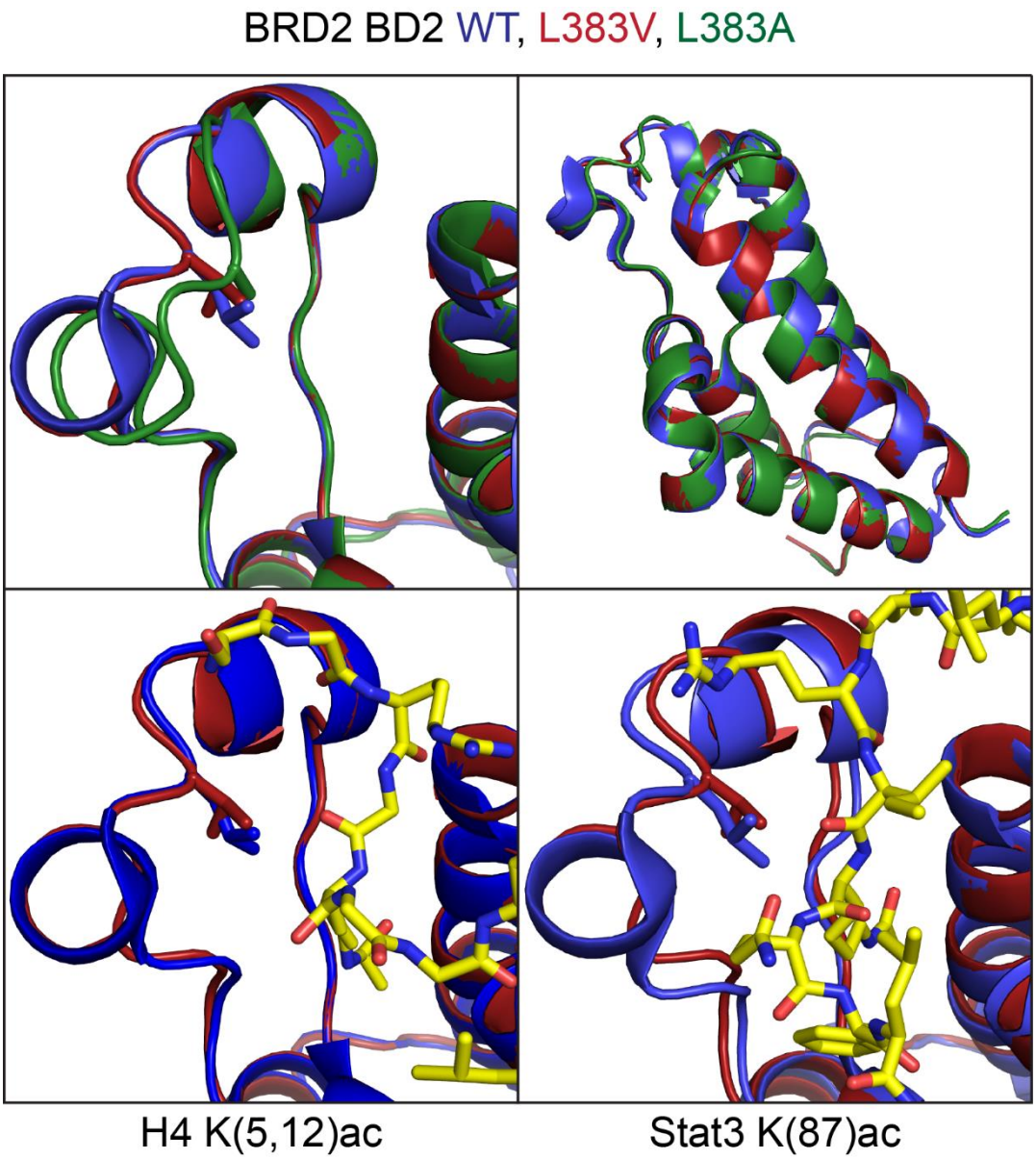


Figure 2.6. Structural comparison of BET bromodomain mutants.

Alignment of crystal structures of BRD2 BD2 WT (blue) (2DVV), L/V (red) (5O38) and L/A (green) (4QEU). BRD2 BD2 L/V also aligned with crystal structure of BRD2 BD2 WT bound to 15-mer H4 K(5,12)ac peptide (2E3K) and NMR structure of BRD2 BD2 WT bound to 12-mer Stat3 K(87)ac peptide (5U5S). Key leucine/valine/alanine residue highlighted.

2.7 – Assessing Cellular Function of Mutant BET Proteins with FRAP:

Thus far our characterisation of the L/V mutants is restricted to single-bromodomain constructs and histone peptides, which cannot fully replicate the interactions between full-length BET proteins and chromatin in cells. To monitor the functionality of mutant BET proteins within living cells we utilised a fluorescence recovery after photobleaching (FRAP) assay, which monitors the interaction of full-length, nuclear-localised BET proteins (GFP-tagged) with chromatin [164]. In a FRAP assay U2OS cells are transfected to transiently overexpress GFP-BET proteins, creating brightly fluorescent nuclei. An area of a fluorescent nucleus is then bleached by a laser, and the random movement of GFP-BET proteins (in and out of the bleached spot) leads to a recovery of the fluorescence (figure S2.8) [165]. If the GFP-BET proteins are bound to (the largely immobile) chromatin the rate of exchange (of bleached protein out of the spot and unbleached protein into the spot) will be reduced and the recovery will be rather slow. The presence of mutations, small-molecules that inhibit chromatin binding and displace the GFP-BET proteins will increase the rate of exchange and hasten this recovery, resulting in a reduction of the measured recovery time (t½) (figure S2.9) [164, 165]. This technique, and later assays using GFP-BET constructs, use the long isoform of BRD4.

In our assay WT BRD4, BRD3 and BRD2 gave recovery times of 3.5 – 4s, with BRDT presenting recovery times of ~1.5s (figure 2.7). To show the recovery rate for a protein completely unable to bind chromatin, and that its recovery is limited only by diffusion, GFP alone was tested in the assay – giving an average recovery time of 0.4s. All double LV/LV mutants showed statistically significant reductions in t½, suggesting weaker chromatin binding. However these decreases were fairly minor, in contrast to the LA/LA mutant of BRD4 which showed a much more severe reduction in recovery time (figure 2.8). Single L/V mutants (LV/WT or WT/LV) showed less noticeable changes in recovery time, with most not being statistically significant. Thus, with regard to BET protein:chromatin interactions the L/V mutation has a minor impact on binding, relative to WT, and exhibits much improved performance on the previous L/A.

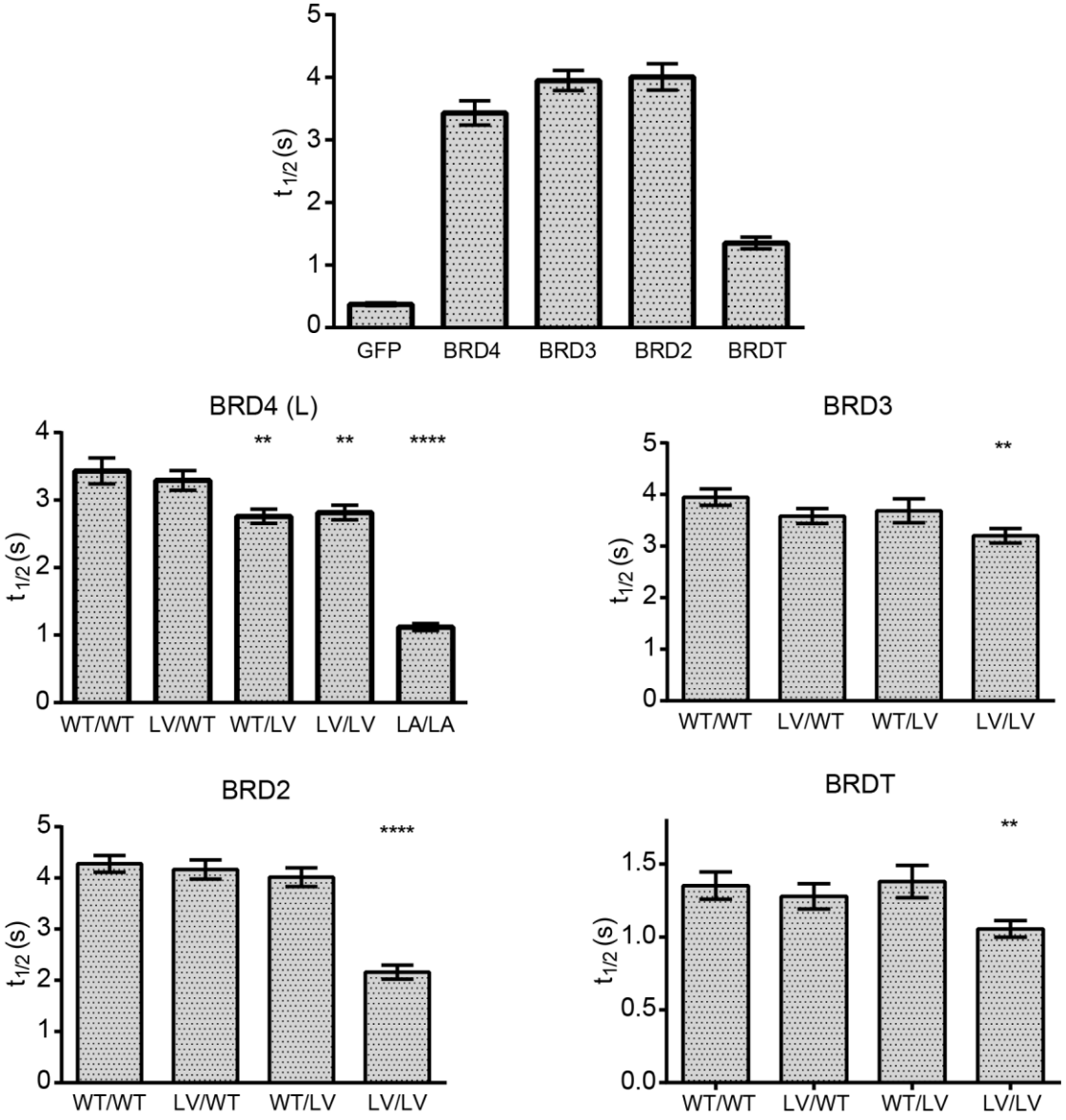


Figure 2.7. BET proteins & chromatin binding.

Recovery times of GFP-labelled BET protein constructs following 0.5s laser bleach event, at 2 µM SAHA and 0.03% DMSO. Each bar is mean ± SE of ~40 U2OS cells tested over two separate experiments. Statistical significance determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$.

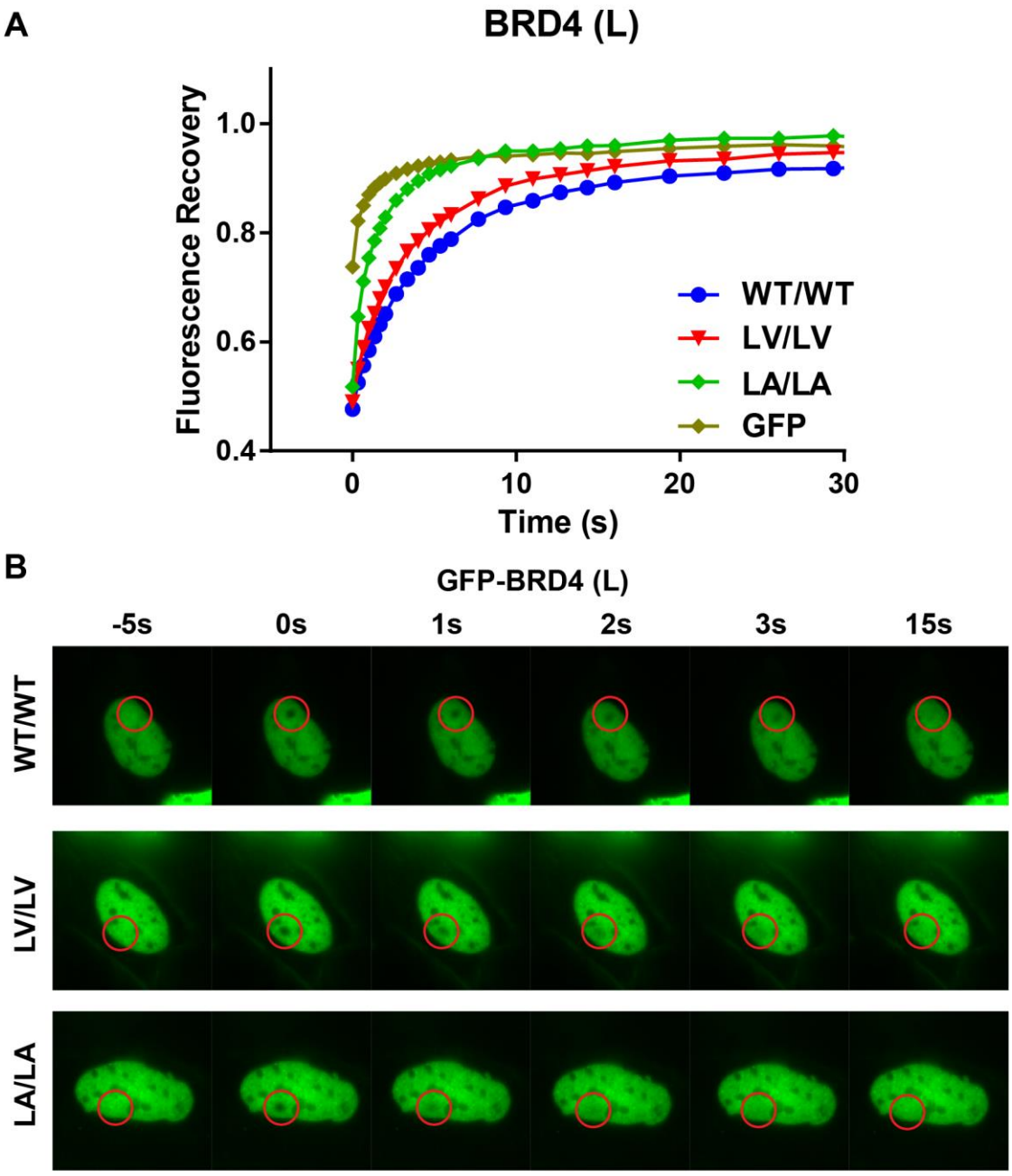


Figure 2.8. FRAP recovery of BRD4 constructs.

A) Normalised fluorescence recovery of GFP-BRD4 (L) constructs (or GFP) transiently overexpressed in U2OS cells, with 0.03% DMSO and 2 µM SAHA, following 0.5s laser bleach event. Each data point is mean ± SEM of ~40 cells from two independent experiments. Data points after 30s omitted for clarity. B) Nuclei of cells from A. Cells shown are representative of their population. Bleach spot highlighted with red circle.

2.8 – Assessing Cellular Function of Mutant BET Proteins with a Luciferase Assay:

BET proteins do not exist solely to bind chromatin, but instead to regulate the expression of downstream genes. To assess this function a luciferase reporter assay was developed, inspired by [95], measuring the ability of over-expressed GFP-tagged BET proteins to induce the expression of a luciferase enzyme regulated by a NF-κB response element (NF-κB-RE). This NF-κB-RE luciferase can be used as a reporter of BET protein functionality due to the observation [95] that BRD4 interacts with NF-κB to up-regulate its target genes. In this assay HEK293T cells are transfected with both a plasmid for overexpressing GFP-BET constructs (or GFP as a control) and the reporter plasmid dependent on NF-κB signalling (figure 2.9A). The impact of NF-κB activity on the luciferase is maximised by the combination of a low-activity promoter with 5 NF-κB response elements. Additionally, the Kozak sequence (nucleotides at the translation start site) of the gene is designed to maximise the translation of any transcribed mRNA and a PEST sequence reduces the half-life of the resulting luciferase, lowering overall signal but increasing how responsive the reporter is to varying NF-κB levels.

In this assay, WT BET proteins showed a clear 2.5 to 6-fold increase in luciferase expression compared to GFP alone and, as expected, this increase is maintained by the L/V mutants (again with some small but statically significant differences) (figure 2.9B). The LA/LA mutant of BRD4 was not able to maintain this increased transcription, giving luminescence values barely above the GFP control. This shows that, as the L/V bromodomains retain their ability to bind acetylated histone tails and chromatin, the L/V BET proteins can still be used to recruit complexes to specific downstream genes and regulate their transcription.

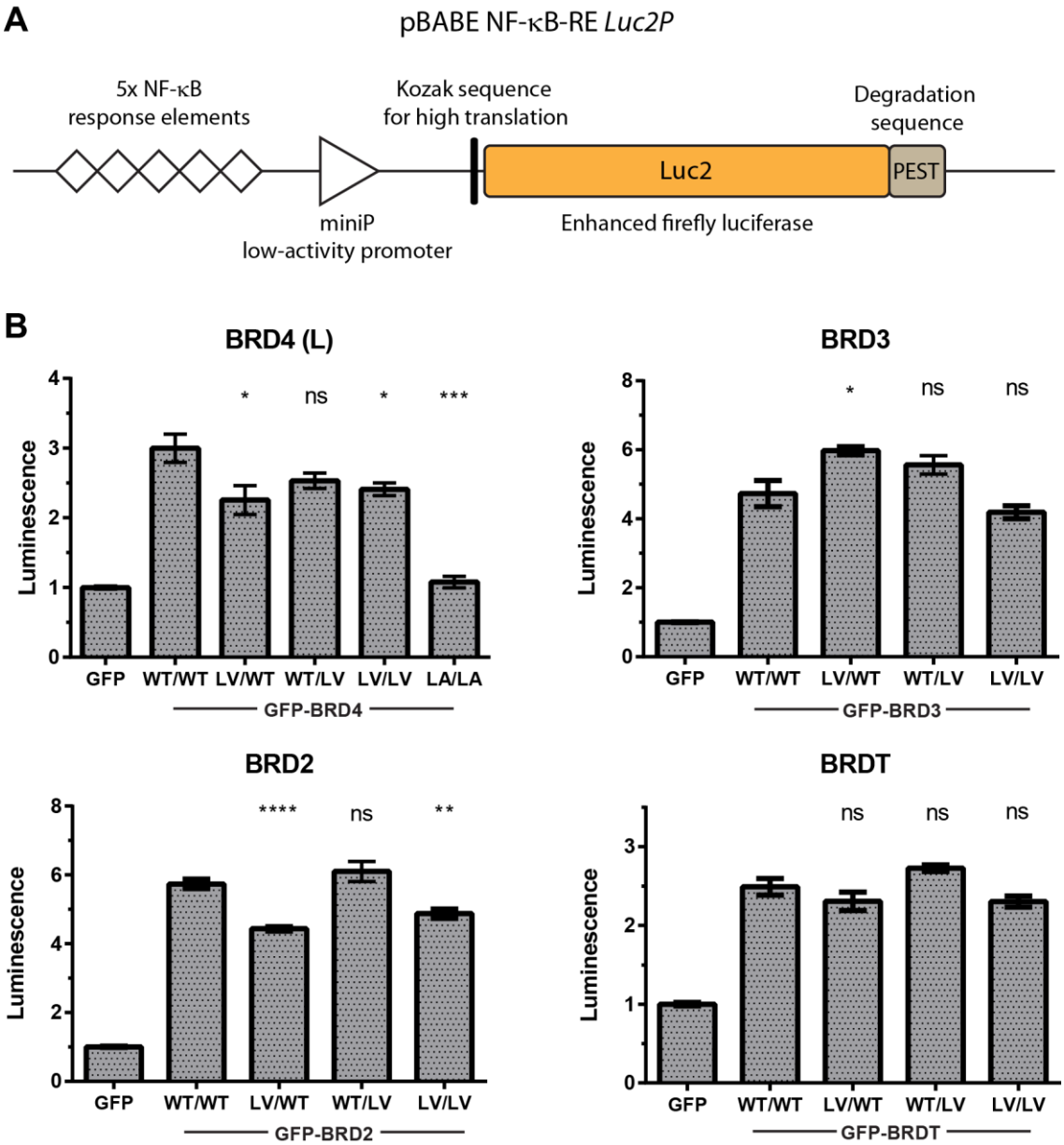


Figure 2.9. BET proteins and NF-κB-mediated luciferase expression.

A) Diagram of key components of NF-κB-RE / Luciferase plasmid. B) Luminescence of HEK293T cells transfected with GFP-labelled BET proteins and a NF-κB-RE/*luc2P* reporter plasmid. Luminescence normalised to GFP control. Statistical significance (related to WT/WT) determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$.

2.9 – Discussion:

For the cellular FRAP assay, GFP alone was chosen as a non-binding control to give context to the reductions in t½ caused by mutations, and later, the addition of inhibitors. Even a protein that lacks any interaction with chromatin will present a low but non-zero t½, as its movement will be limited by diffusion and hence its fluorescence recovery will not be instant. Alternatively, BD1/2-deletion BET proteins or inactive N/F mutants could have been used as controls. Their t½ values may be a better representation of BET proteins completely displaced by mutations or compounds, as they would be roughly the same size and would contain other domains (i.e. ETD) that may contribute recovery time. On the other hand, this would have required an investment of time and resources into mutagenesis and would require a different control be used for each BET protein tested. GFP does not contain a nuclear localisation signal, but due to its small size GFP monomers and dimers are often present in the nucleus [166], which for many of our cells could be visually identified (figure S2.10) and bleached.

Despite giving similar K_d values in the peptide ITC experiments BRDT showed much weaker chromatin binding in the FRAP assay (t½ of 1.5 s compared to ~4 s). Without further data this should likely not be taken as evidence that BRDT is in fact any worse at chromatin binding than the other BET proteins. As the different BET proteins likely prefer different epigenetic marks it may be that BRDT, a testes-specific protein [167], does not bind strongly to the epigenetic marks presented in the SAHA-treated U2OS cells (epithelial cells from an osteosarcoma) used in FRAP.

In our FRAP assay BRD2 showed a much more severe response to the L/V mutation, with the LV/LV construct giving a t½ of 2s. Both the ITC and BLI data with histone peptides do not show BRD2 to be any more affected by the L/V mutation. This could be an indicator that ITC assays with peptides are not a reliable indicator of interactions between full-length proteins and chromatin. BET protein:chromatin interactions are complicated by numerous factors that do not contribute to BD:peptide binding. Nucleosomes, with the presence of 8 globular domains and 8 histone tails, as well as associated DNA, may generate a large number of steric clashes that alter binding. Again, the presence of multiple histone tails, and the presence of 2 BDs in a BET protein, creates opportunity for multivalent interactions. Finally, there are the

previously mentioned possibilities that BET proteins may interact with DNA [34] or even interact with chromatin indirectly through ETD-recruited NHPs. The recovery time of BRD2 might not be dependent on interactions with the H4 histone tail, but instead other histones or chromatin-bound proteins which we have not tested in ITC and hence could be impacted more by the L/V mutation.

The unusually severe impact of the L/V mutation on BRD2 in the FRAP assay is most likely an artefact of said FRAP assay. In the cellular luciferase assay the LV/LV BRD2 construct did not show such a severe impact, and was much more consistent with the BLI and ITC data. Additionally given the conserved nature of the BET BDs it would be surprising if one protein were to be affected so much more severely than the others. Finally, subsequent FRAP work (discussed later) generated unusual results with GFP-BRD2. Specifically, when both the BD1 and BD2 of GFP-BRD2 were inhibited the construct was shown to aggregate (or undergo a significant change in localisation) which complicated the acquisition of t½ values. The low t½ values of uninhibited LV/LV GFP-BRD2 are likely connected to this phenomenon, and are not an accurate indicator of the construct's ability to bind chromatin.

2.10 – Conclusion:

Through extensive characterisation of the L/V mutation: covering stability, *in vitro* peptide binding, crystal structures and cellular function we have consistently seen the L/V mutation to be very conservative and a clear improvement over the L/A mutation, rendering a functionally silent mutant behaving as closely as reasonably required to wild-type. L/V BET bromodomains are stable, bind acetylated histone peptides both with high affinity and similar binding profiles (to WT) and show no marked conformational changes; while full-length L/V BET proteins recognise and bind chromatin almost as well as WT and can translate this into increased expression of downstream genes. The L/V mutation is not perfect, and produces a small but consistent impact on functionality. However, in all cases (except interestingly stability) the L/V mutation is a clear improvement on the previously used L/A mutation, which our cellular experiments show would not function properly in cell or animal models. Once implemented into cells we expect the L/V mutation to be viable and allow for successful interrogation of the BET proteins by our optimised bumped compounds.

Chapter 3
Evaluation of Bumped Compounds

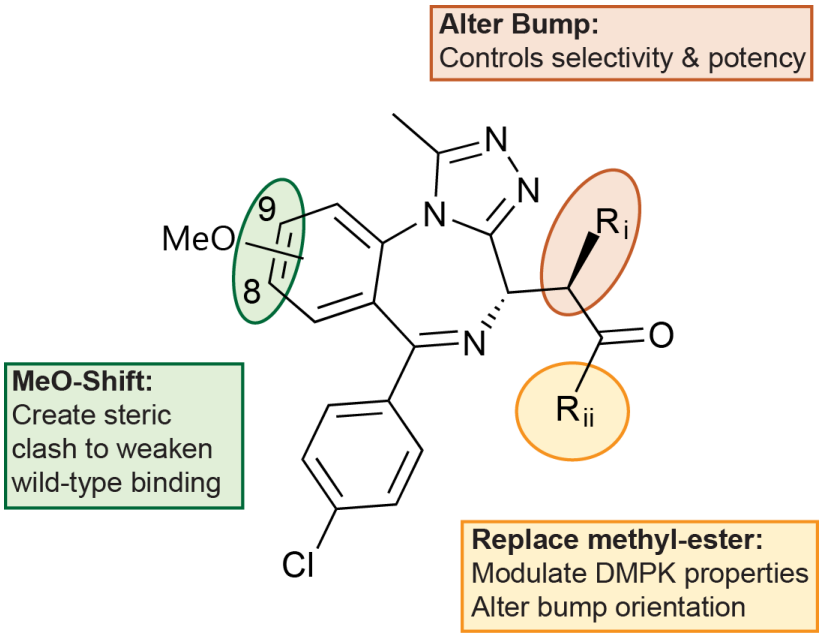
3.1 – Introduction

With the new and more functional L/V mutation characterised and validated we decided to shift attention to the chemical aspect of the bump-&-hole system. Specifically, the 'bump' and other modifications used, as well as the selectivity, potency and DMPK qualities of the final compound had scope for optimisation at this stage of the project. By moving to a more conservative mutation existing bumped compounds will display different SAR, with a smaller mutation potentially hindering selectivity. By trialling a greater range of chemical modifications it was hoped that, despite the 'more difficult' mutation, bumped compounds could be developed with selectivity greater than that observed with the ET compound in the old L/A mutant bump-&-hole system.

3.2 – Bumped Compound Rational & Synthesis

To optimise the chemical aspect of the bump-&-hole system we trialled a range of modifications to three parts of our scaffold (scheme 3.1 & 3.2) (scheme S3.1). While the original bump-&-hole research tested only the methyl and ethyl bumps the longer propyl bump was trialled too. Additionally an allyl bump was tested, to observe the importance of the bump's degree of rotational freedom. The wealth of BET inhibitor SAR in the literature contains several simple modifications that reduce the WT potency of the I-BET762 scaffold, and it was thought that these might be ways of generating selectivity. Moving the methoxy-group from the 8' to the 9' position is one such modification [78], which creates a steric clash that we hypothesise could be better accommodated by the mutated binding site. This modification was also expected to not alter our compounds' DMPK properties. Our original compounds feature a methyl-ester 'side-group' adjacent to the alkyl bump, which we decided to replace for two reasons. The residual WT binding displayed by early compounds is likely the result of the flexible methyl-ester side-group rotating and re-orienting in ways that relieve the key steric clash. This hypothesis is supported by co-crystal structures of bumped compounds bound to L/A bromodomains and I-BET762 bound to WT, which showed significant movement of the side-group [159]. To 'lock' the side-group in place we introduced a polar ethyl-amide side-group, as in I-BET762, which we expected to form

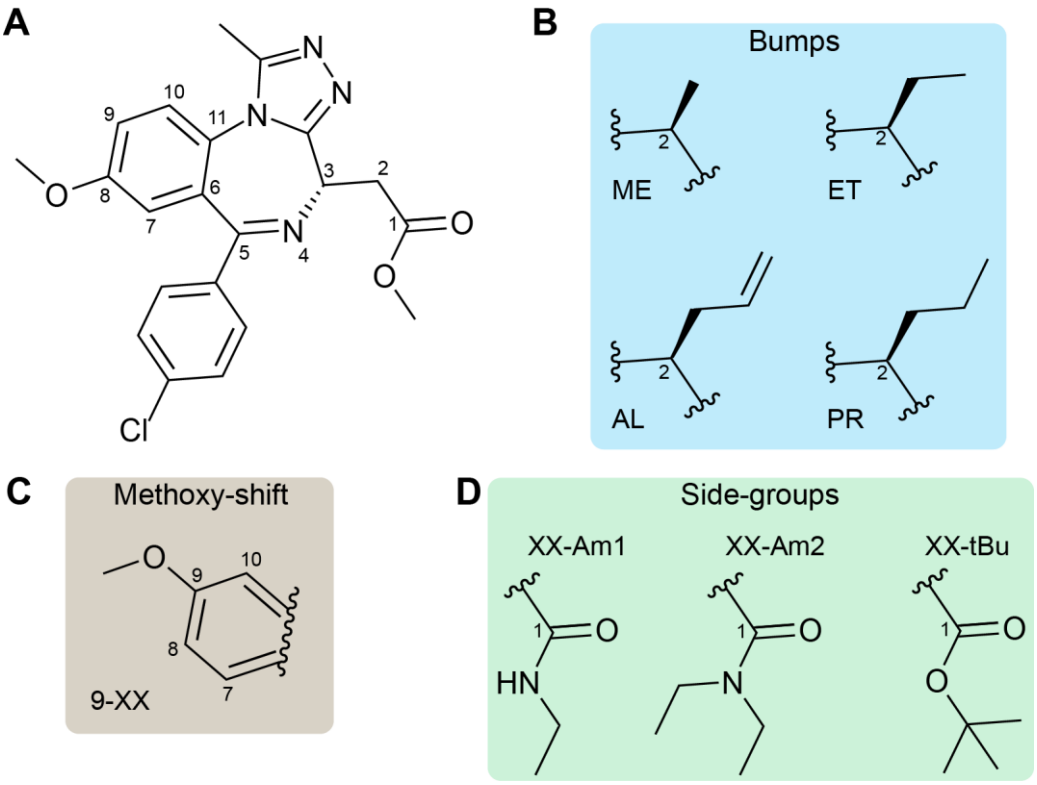
a hydrogen bond with the side-chain of N140. We also tried a *tert*-butyl-ester side-group, based on (+)-JQ1, which may form favourable hydrophobic interactions and may display limited flexibility due to its bulky nature. These alternate side-groups were also expected to fine-tune the DMPK properties of our compounds.



Scheme 3.1. Rational for bumped compound optimisation.

Chemical structure of scaffold, with modification sites highlighted.

Compounds were synthesised by Dr Michael Zengerle and Dr Matthias Baud, the details of which are published [122, 159, 168]. As carbons 2 (when bumped) and 3 form chiral centres this synthesis produced a mixture of 4 isomers (scheme S3.2). Previous analysis of co-crystal structures showed that only the 2R,3S enantiomer can bind BET bromodomains [159]. The inactive *2S,3S* and *2R,3R* diastereomers were removed through reverse-phase HPLC, but the active *2R,3S* and inactive *2S,3R* enantiomers were not separated at this stage, and unless stated otherwise they were tested as a racemic mixture. As such, ligand concentrations, IC_{50} values and related binding and activity measurements refer to only the active enantiomer.



Scheme 3.2. Bumped compound modifications.

Chemical structure of scaffold compound (A), alkyl bumps (B), methoxy-shift modification (C) and alternate side-groups (D).

3.3 – Development of AlphaLISA Assay

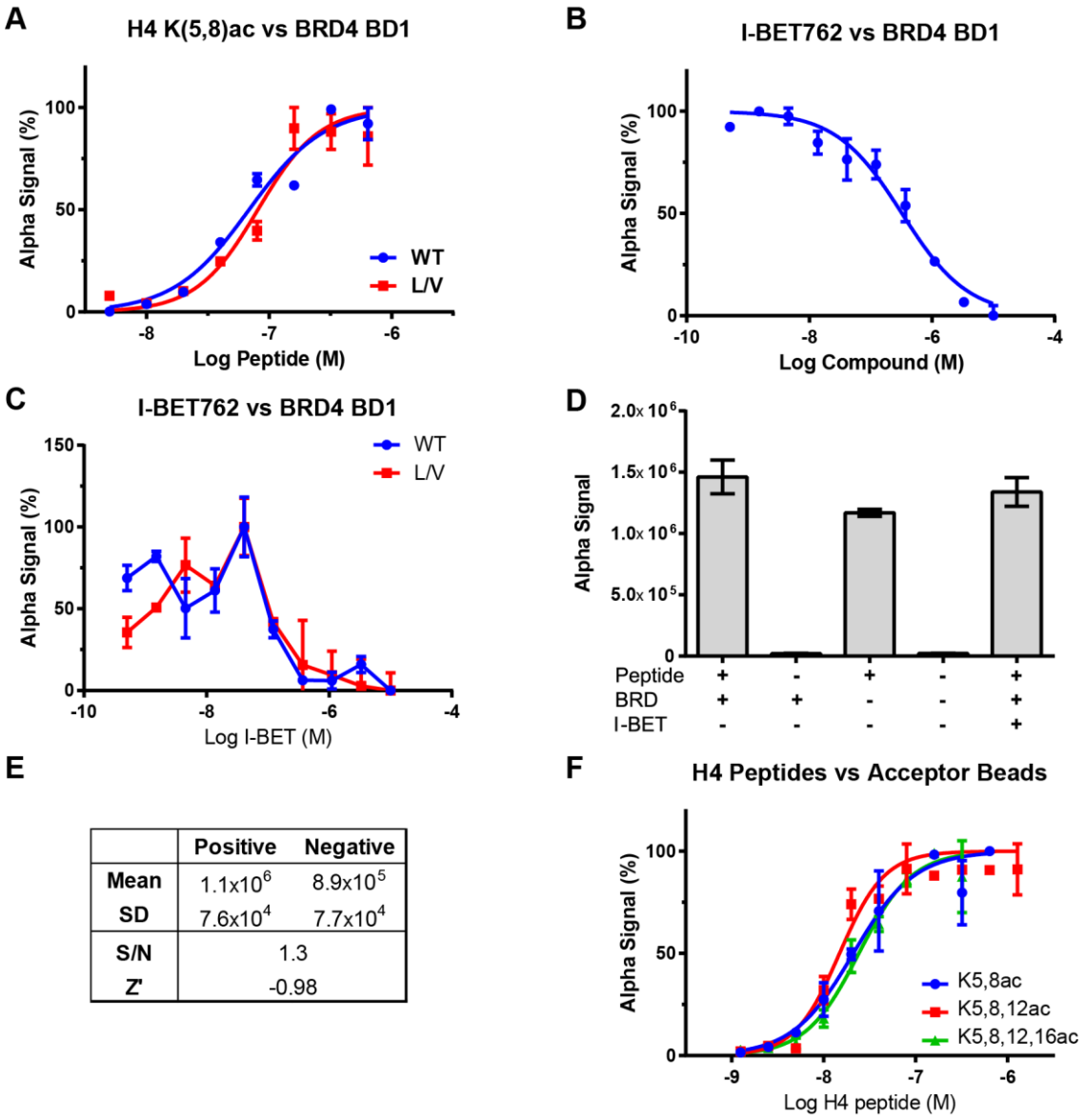
In the early stages of the bump-&-hole project compounds were tested through ITC and the thermal shift assay [169]. ITC is very powerful and reliable, but is not suitable for testing a large number of interactions, while the project necessitated testing >20 compounds against 16 proteins (8 BET bromodomains, WT and L/V). The thermal shift assay does possess reasonably high throughput, but is not very reliable nor is quantitative. The thermal shift assay measures the change in protein stability upon ligand binding (ΔT_m values). While the size of the observed ΔT_m is often seen to be related to binding strength, it cannot be directly correlated to K_d or IC_{50} values. Hence it was decided to develop a high-throughput, quantitative assay in a 384-well format that could be used to titrate all bumped compounds against all WT and L/V bromodomains.

Following a review of the literature, it was decided to develop an AlphaLISA assay, wherein a biotinylated histone peptide is displaced from a his-tagged bromodomain. The assay signal is generated by a pair of AlphaLISA beads, with donor beads binding to the peptide through a streptavidin group and acceptor beads binds the bromodomain through an anti-His6 antibody. During the assay the donor beads, when excited by light at 680 nm, produce singlet oxygen which triggers emission of light at 615 nm from the acceptor beads, provided they are within 200 nm. This condition is satisfied when the peptide is interacting with the bromodomain [170].

Titrations of biotinylated, di-acetylated H4 K(5,8)ac peptides against WT and L/V BET bromodomains showed that concentrations of 200 nM bromodomains and 200 nM peptide would generate the strongest signal (figure 3.1A). These experiments also showed similar affinity for WT and L/V bromodomains, suggesting WT and L/V IC_{50} values could be reliably compared (table S3.1). Despite early successes (figure 3.1B), when the screening began the results were very poor and generated unusable dose-response curves (figure 3.1C). This suggested that the assay was not as robust as believed, and may only generate usable results when performed in a low-throughput capacity with the use of very exact concentrations, timings or handling procedures. Further investigations revealed that, although removing the peptide probe reduced the signal ~70-fold, removing the bromodomain had little impact on signal (figure 3.1D-E). This explained why even large concentrations of potent inhibitors had little

impact on assay signal, and suggested that the biotinylated peptides were generating Alpha signal independent of any bromodomain binding. Titrations of biotinylated peptides against AlphaLISA beads in the absence of bromodomain generated sigmoid curves with EC_{50} values of ~20 nM (figure 3.1F), showing that the biotinylated peptides were somehow interacting directly with the acceptor beads or their anti-his6 antibodies. Increasing bovine serum albumin (BSA) concentrations, and the introduction of Tween-20 and Triton-x100 reduced these non-specific interactions but failed to improve assay quality (figure 3.2A-B). Alternate acceptor beads that bind his-tags through a Ni-chelator also generated this non-specific signal (figure 3.2C-D). Based on these results it was decided that, for an AlphaLISA assay to work, it was not possible to use a labelled peptide probe to monitor displacement and labelled small-molecules would have to be used instead.

Following the failure to 'fix' the peptide-based AlphaLISA assay it was decided to measure instead the displacement of a biotinylated derivative of (+)-JQ1. One biotinylated (+)-JQ1, using an aminohexanoic linker, (MZ-70) (figure 3.3A) failed to generate any assay signal, presumably as its linker was not long enough to accommodate both the bromodomain and streptavidin (figure 3.3C). Later a 3-PEG linker was tried, based on the development of Bio-JQ1 by Anders et al. [86], and succeeded in generating Alpha signal with bromodomains (but crucially not beads) (figure 3.3B-C). Bio-JQ1 was titrated against all BET bromodomains and all showed Alpha signal generation. Additionally, Bio-JQ1 bound all WT and L/V BET bromodomains with similar affinity, with a ~2-fold decrease in affinity for L/V bromodomains compared to their WT equivalents (figure 3.3D-E) (table S3.2). Concentrations of 100 nM bromodomain and 5 nM Bio-JQ1 provided high-quality dose-response curves with IC_{50} values close to known K_d values (figure 3.3F). Tests with inhibitors showed a ~100-fold assay window and a Z-factor of 0.77 (indicating a high-quality assay) (figure 3.3G) [171]. Test titrations with ET against WT and L/V BDs gave high-quality data in-line with expectations (figure 3.3H) and in additional experiments the assay was able to accurately rank BET inhibitors and even detect fragment binding (figure S3.1). Following these tests it was decided to proceed with the full compound screen with this new optimised JQ1-based assay.



	Positive	Negative
Mean	$1.1x10^6$	$8.9x10^5$
SD	$7.6x10^4$	$7.7x10^4$
S/N	1.3	
Z'	-0.98	

Figure 3.1. Peptide-based AlphaLISA assay.

A) Titration of biotinylated H4 K(5,8)ac against BRD4 BD1 constructs in AlphaLISA assay. B) Successful titration of I-BET762 against BRD4 BD1 in peptide AlphaLISA assay C) Failed titration of I-BET762 against BRD4 BD1. D) Signal of AlphaLISA assay in absence/presence of assay components. E) Signal/noise and Z-factor of peptide AlphaLISA assay. F) Titration of biotinylated, acetylated H4 peptides against AlphaLISA beads in absence of bromodomain.

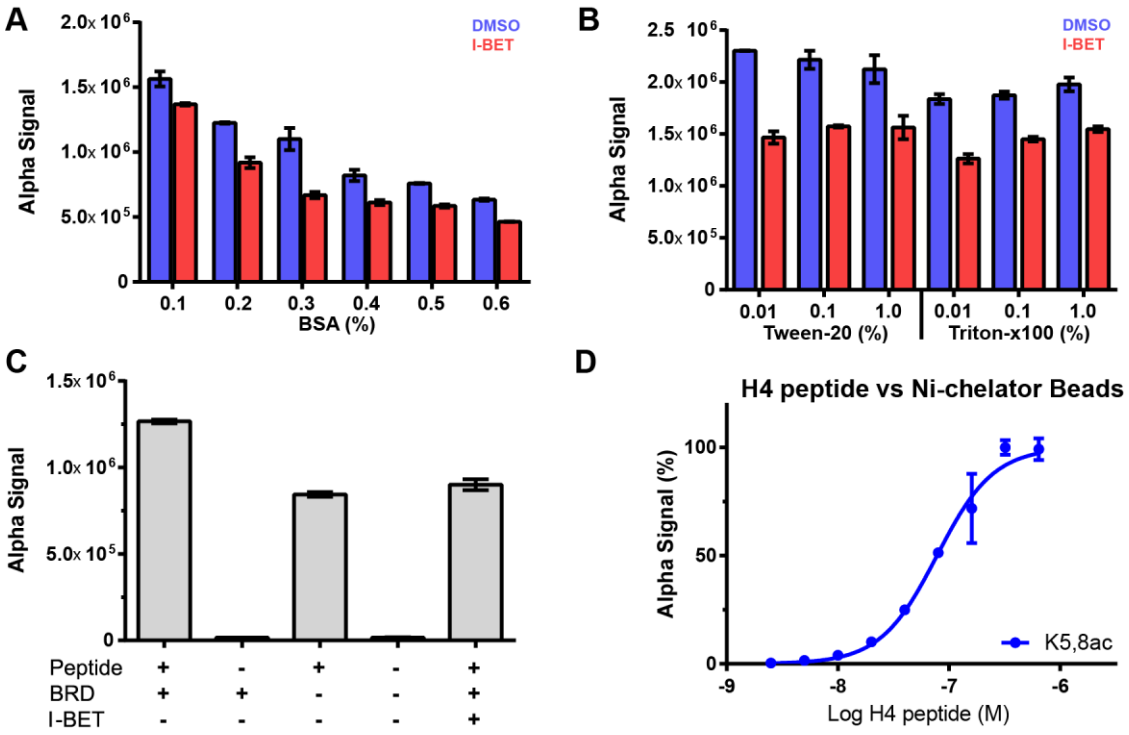


Figure 3.2. Attempts to fix peptide-based AlphaLISA assay.

A) Signal of AlphaLISA assay in presence of BSA. B) Signal of AlphaLISA assay in presence of Tween-20 and Triton-100. C) Signal of AlphaLISA assay, with Ni-chelator acceptor beads, in absence/presence of assay components. D) Titration of biotinylated, acetylated H4 peptide against AlphaLISA beads (with Ni-chelator acceptor beads) in absence of bromodomain.

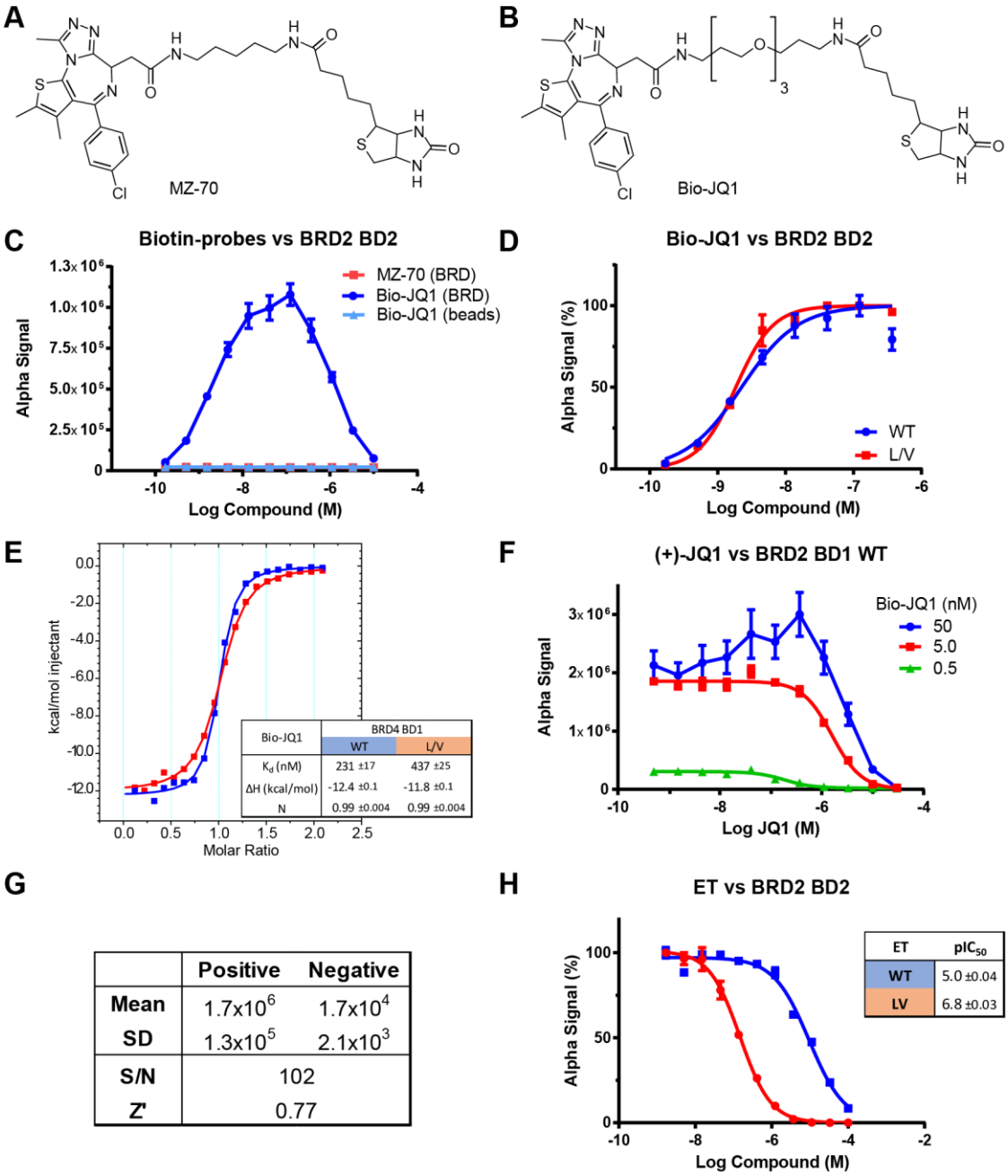


Figure 3.3. Development of Bio-JQ1-based AlphaLISA assay.

Chemical structures of MZ-70 (A) and Bio-JQ1 (B). C) AlphaLISA titration of MZ-70 and Bio-JQ1 against BRD2 BD2 WT and Bio-JQ1 against beads alone. D) AlphaLISA titrations of Bio-JQ1 against BRD2 BD2 WT and L/V. E) ITC titrations of Bio-JQ1 against BRD4 BD1 WT and L/V. F) AlphaLISA titrations of (+)-JQ1 against 100 nM BRD2 BD1 WT and varying concentrations of Bio-JQ1. G) Robustness of final Bio-JQ1/AlphaLISA assay. Positive control = no inhibitor, negative control = 30 µM (+)-JQ1. H) AlphaLISA titrations of ET against BRD2 BD2 WT and L/V.

3.4 –Primary AlphaLISA Screen of Bumped Compounds

The AlphaLISA screen produced a large amount of SAR data (table 3.1) (figure 3.4) (table S3.3), showing several clear trends associated with various modifications. One immediately noticeable trend is that, overall, our compounds are very successful as L/V bromodomain probes. Almost all compounds have L/V potencies below 1 μM and selectivity against WT bromodomains of at least 10-fold. Many bumped compounds are as potent (or more) against L/V bromodomains as our scaffold and established BET inhibitors (+)-JQ1 and I-BET762 are against WT. Meanwhile we are able to greatly weaken WT binding, in some cases comparable to the inactive (-)-JQ1 molecule. This screen showed that the methyl, ethyl and allyl bumps, the methoxy-shift modification and the methyl-ester and ethyl-amide side-groups were all tolerated. Finally, the propyl bump, the di-ethyl-amide and *tert*-butyl-ester side-groups were shown to not be beneficial.

The alkyl bumps are the main determinant of potency and selectivity. Smaller bumps show stronger binding to both WT and L/V bromodomains, with these potencies decreasing with larger bumps. Against L/V bromdodomains, methyl-ester compounds with the ethyl and allyl bumps are about as potent as the scaffold, suggesting these bumps are not introducing steric clashes. The methyl bump meanwhile, shows a clear increase in potency suggesting this small bump essentially forms a novel hydrophobic interaction with the valine 'hole'. Regarding overall selectivity, the ethyl bump is optimal, while the long propyl bump is the least effective, with <10-fold selectivity and a clear drop in potency. Interestingly, the low potency and selectivity of the propyl bump is likely not the result of a steric clash, but instead due to its high degree of rotational freedom which must be constrained upon binding. The similarly sized allyl bump is much more potent, and its SAR is much more similar to the ethyl bump with which it shares the same degree of rotational freedom.

Shifting the methoxy group from the 8' position to the 9' has a clear impact on the potency of the scaffold, without altering selectivity. Meanwhile when the methoxy group is shifted on bumped compounds it generally maintains potency and selectivity. When paired with the large allyl bump or other modifications such as the amide side-groups the methoxy shift noticeably reduces potency, and it only produces a clear increase in L/V selectivity when paired with the methyl bump. Hence the methoxy shift

is best paired with more subtle modifications, likely as it is in these conditions where the methoxy is best accommodated in the mutant binding site.

The ethyl-amide side-group, introduced to replace the methyl-ester moiety and hopefully 'lock' the bump in place, had a clear impact on binding. WT binding was greatly weakened, by roughly ~0.7 log units, which is likely due to the desired hydrogen bond to N140 preventing bump re-orientation and accommodation within the WT binding site. This hydrogen bond did not strengthen L/V binding, as hoped, but instead also weakened binding to L/V bromodomains by again ~0.7 log units, resulting in broadly similar overall selectivity. Combining the ethyl-amide side-group, ethyl bump and 9' methoxy shift generates 9-ET-Am1, which showed the weakest WT binding of all bumped compounds. Meanwhile, the di-ethyl-amide side-group has very similar SAR to the ethyl-amide, but it does show slightly reduced selectivity (~1.2 vs ~1.5 ΔpIC_{50}).

The *tert*-butyl ester side-group is perhaps the most deleterious modification introduced, with all compounds possessing it showing less than 10-fold selectivity for the L/V bromodomains. Combining this with the methoxy-shift modification creates an even more severe drop in selectivity. This modification, in most cases, both reduces potency against the L/V bromodomains while increasing off-target WT binding. As the highly-potent (+)JQ1 features an identical *tert*-butyl ester this modification is likely not an issue on its own, but is instead highly sensitive to the addition of alkyl bumps.

Compound	R_i	R_{ii}	MeO-Position	AlphaLISA pIC_{50} ±SD		
				WT	L/V	Δ
(+) JQ1	H	O^{t}Bu	-	6.6 ±0.1	6.5 ±0.1	-0.1 ±0.1
(-) JQ1	H	O^{t}Bu	-	4.8 ±0.2	4.6 ±0.1	-0.2 ±0.2
I-BET762	H	NHEt	8	6.5 ±0.2	6.5 ±0.3	0.0 ±0.3
Scaffold	H	OMe	8	6.6 ±0.2	6.6 ±0.3	0.0 ±0.2
9-Scaffold	H	OMe	9	5.8 ±0.2	5.9 ±0.2	0.1 ±0.1
Carboxylic Acid	H		8	4.7 ±0.4	5.6 ±0.3	0.9 ±0.2
ME	methyl	OMe	8	6.2 ±0.3	7.4 ±0.1	1.2 ±0.2
ET	ethyl	OMe	8	5.0 ±0.1	6.8 ±0.2	1.7 ±0.2
AL	allyl	OMe	8	5.1 ±0.1	6.6 ±0.3	1.5 ±0.3
PR	propyl	OMe	8	4.5 ±0.2	5.0 ±0.1	0.6 ±0.2
9-ME	methyl	OMe	9	5.4 ±0.2	6.8 ±0.2	1.4 ±0.2
9-ET	ethyl	OMe	9	5.5 ±0.2	6.8 ±0.2	1.4 ±0.2
9-AL	allyl	OMe	9	4.4 ±0.1	5.7 ±0.2	1.3 ±0.2
ME-Am1	methyl	NHEt	8	5.3 ±0.2	6.7 ±0.2	1.4 ±0.2
ET-Am1	ethyl	NHEt	8	4.4 ±0.1	5.9 ±0.3	1.5 ±0.2
AL-Am1	allyl	NHEt	8	4.3 ±0.1	5.9 ±0.2	1.5 ±0.1
ME-Am2	methyl	N[Et]$_2$	8	4.9 ±0.2	6.1 ±0.2	1.2 ±0.2
ET-Am2	ethyl	N[Et]$_2$	8	4.6 ±0.1	5.8 ±0.2	1.2 ±0.2
9-ME-Am1	methyl	NHEt	9	4.7 ±0.3	6.0 ±0.2	1.3 ±0.2
9-ET-Am1	ethyl	NHEt	9	<4.0	5.2 ±0.2	>1.2
ET-tBu	ethyl	O^{t}Bu	8	5.3 ±0.1	6.0 ±0.2	0.7 ±0.1
AL-tBu	allyl	O^{t}Bu	8	5.3 ±0.1	6.2 ±0.2	0.9 ±0.2
9-ET-tBu	ethyl	O^{t}Bu	9	4.6 ±0.1	5.0 ±0.2	0.4 ±0.2
9-AL-tBu	allyl	O^{t}Bu	9	5.2 ±0.2	5.8 ±0.2	0.5 ±0.1

Table 3.1. AlphaLISA primary bumped-compound screen.

Mean pIC_{50} values of bumped compounds for all somatic bromodomains. JQ1 possesses a thiophene ring and hence has no methoxy group. Refer to scheme 3.1 for positions of R_i and R_{ii}. pIC_{50} values for individual bromodomains can be found appendix table S3.3.

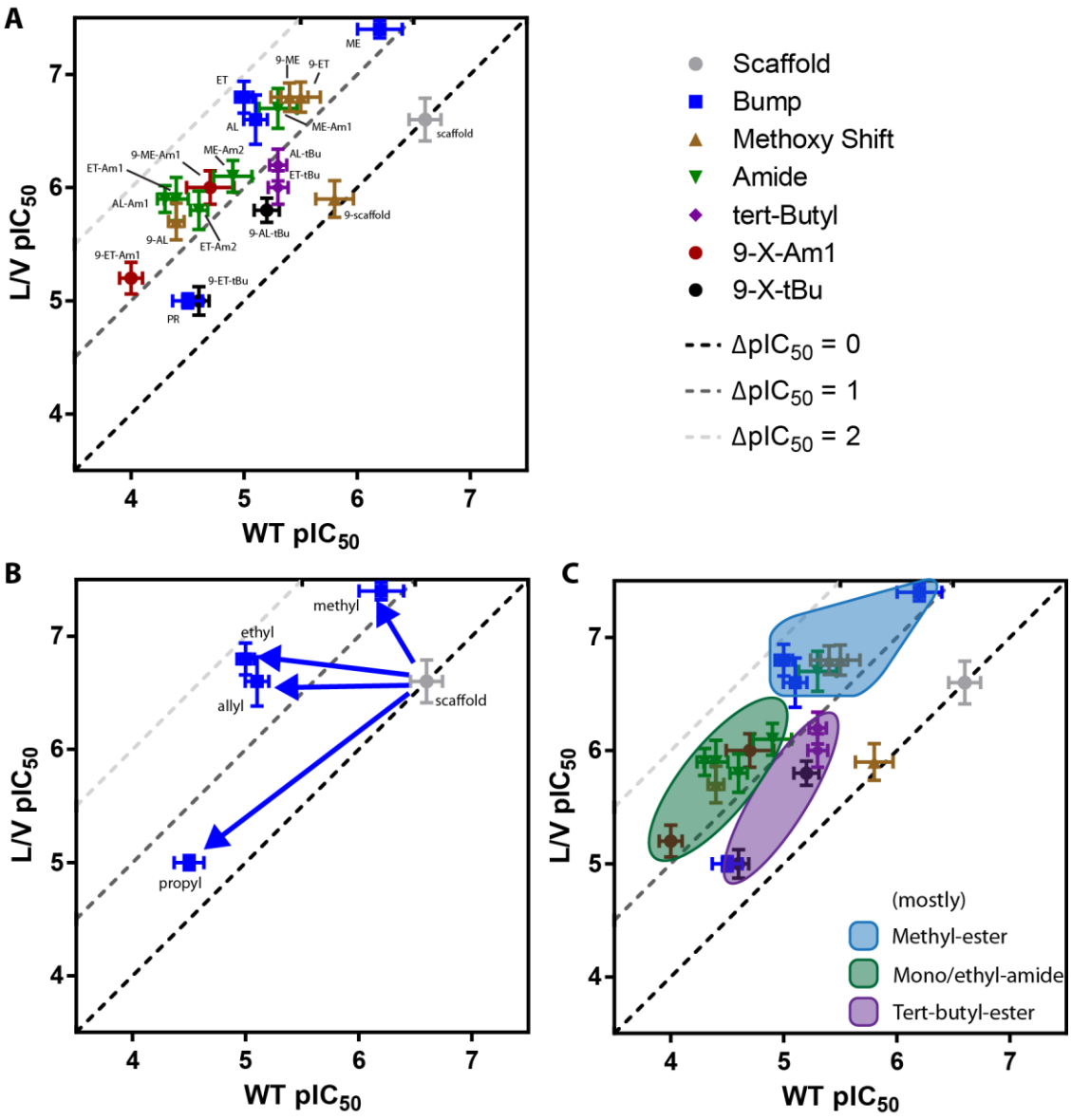


Figure 3.4. Structure-activity relationships of bumped compounds in AlphaLISA screen.

A) Mean pIC_{50} values of compounds for all somatic WT and L/V BET bromodomains. B) Effect of different alkyl bumps on scaffold. C) Different side-groups create 'clusters' of SAR. Each cluster has 1 structural outlier: ME-Am1 in the methyl-ester cluster, 9-AL in the amide cluster and PR in the *tert*-butyl-ester cluster.

3.5 – Investigation of AL-tBu Binding

Although generating a large amount of useful SAR data, the AlphaLISA screen did show one significant problem. Previous to this work the bumped compounds had been assessed by Dr Michael Zengerle through ITC (table S3.4). This had shown AL-tBu to be the best bumped compound – showing no detectable WT binding and an average L/V K_d of 320 nM. The AlphaLISA screen instead showed Al-tBu to be one of the worst compounds and in fact all four *tert*-butyl-ester compounds show less than 10-fold selectivity for L/V (table 3.1). This discrepancy seems much too severe to blame on inter-assay variation so the assessment of bumped compounds was put on hold until this issue could be addressed and either dataset could be determined reliable.

The divergence between ITC and AlphaLISA results was greatest with regards to WT binding (figure 3.5A), suggesting that one possible cause of the disagreement was that AL-tBu did show moderate WT binding, but this was obscured in its ITC titrations. Further reading revealed that the tBu compounds were too insoluble for the default ITC titration and had to undergo a reverse titration of protein injected into compound (as opposed to compound into protein). Any change in protocol for a binding assay like this will change the threshold at which point weak binding will not be detected, and in general such reverse ITC titrations are believed to be less sensitive. To confirm that AL-tBu's WT binding was obscured by the change in ITC protocol, compounds that possessed similar AlphaLISA WT potency, and measurable WT binding in the default ITC protocol, were re-tested in the reverse ITC protocol (figure 3.5B-D). Using the reverse ITC protocol this previously-observed moderate WT potency disappeared, leaving no detectable binding. This confirmed that the high selectivity displayed by AL-tBu in the early ITC work was due to poor interpretation of 'no binding' results, and validated the AlphaLISA dataset as a more accurate representation of AL-tBu compound.

A

BRD2 BD1	AlphaLISA pIC_{50} WT	AlphaLISA pIC_{50} L/V	ITC pK_D WT	ITC pK_D L/V
ET	4.9	6.5	4.7	6.6
ME-Am1	4.9	6.4	5.6	6.9
AL-tBu	5.3	5.9	<4.3	6.3

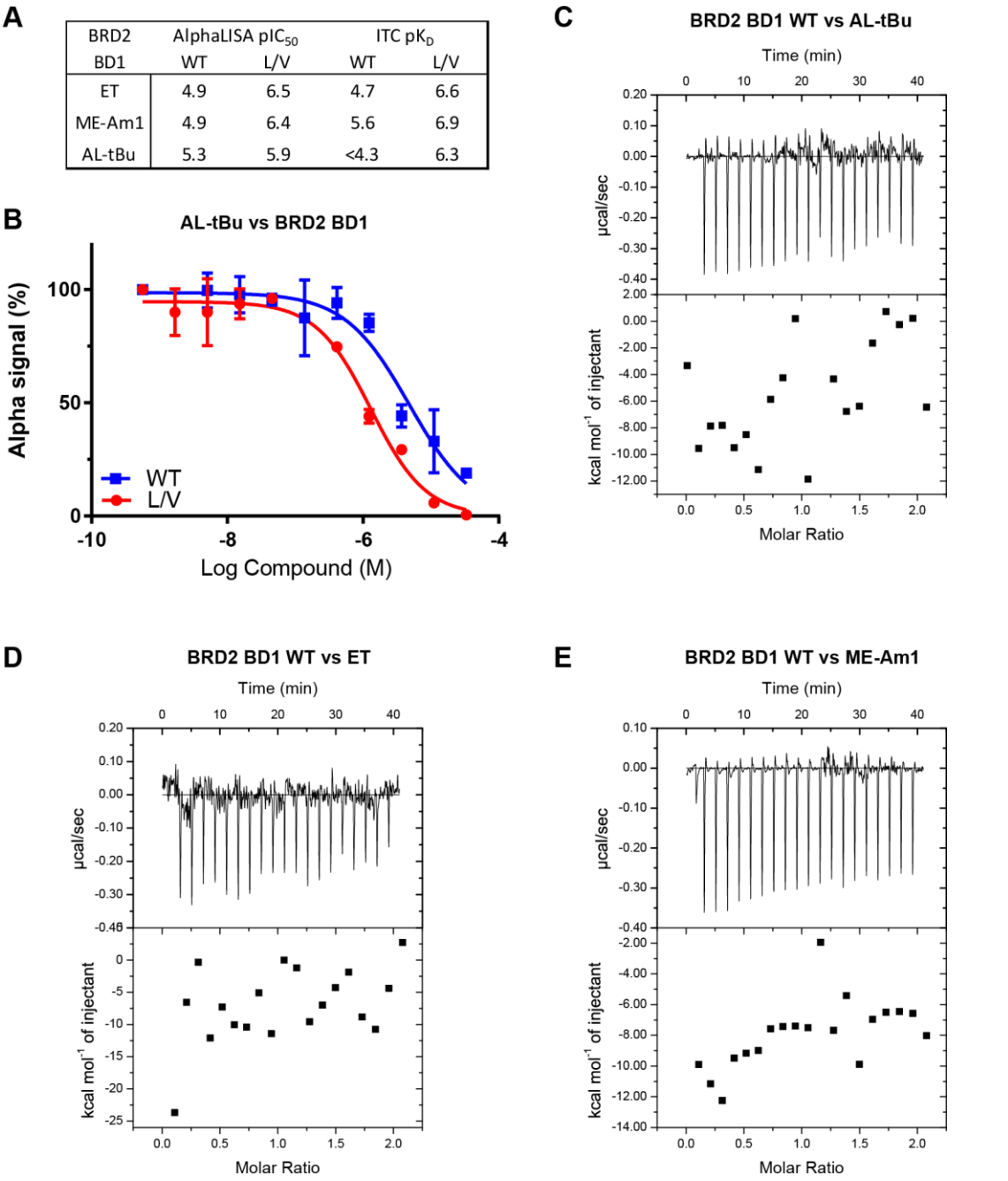


Figure 3.5. Investigating AL-tBu binding

A) AlphaLISA and ITC-derived affinity values for bumped compounds and BRD2 BD1 constructs. ET and ME-Am1 results from standard compound-into-protein titrations, AL-tBu results from reverse protein-into-compound titration. B) AlphaLISA titration of AL-tBu against BRD2 BD1 WT and L/V. C-E) ITC reverse titration of BRD2 BD1 WT against AL-tBu (C), ET (D) and ME-Am1 (E).

3.6 – Secondary DMPK triage of Bumped Compounds

Determining the best bumped compounds clearly required intensive ITC screening of select compounds against all WT and L/V BET bromodomains. To narrow down the list of potential compounds, and ensure our best compound would be usable *in vivo*, a secondary triage was carried out focusing on DMPK characteristics. These assays were carried out by Dr Ola Epemolu and Dr Lucy Ellis of the Drug Development Unit (DDU) (Kevin Read lab) and looked at the stability of compounds, in both plasma and liver microsomes, and their apparent intrinsic permeability as measured by the *in vitro* parallel artificial membrane permeability assay (PAMPA) (table 3.2).

The stability of our compounds in plasma (recorded as half-life t½) is a strong indicator of vulnerability to esterase activity [172]. This has previously been identified as a potential threat to compound cellular activity, as cleavage of the methyl-ester side-chain to a carboxylic acid leads to a marked reduction in potency and selectivity (table 3.1). Both our scaffold compounds were shown to be unstable in plasma, with half-lives of ~1 hour, while (+)-JQ1 and I-BET762 which instead possess a *tert*-butyl ester and ethyl-amide side-groups, respectively, show no significant degradation. Interestingly, these alternate side-groups were shown not to be necessary to prevent breakdown of our bumped compounds, as none of them showed any degradation in plasma. Fortunately, the presence of our alkyl bump on the ester-bonds α-carbon is enough to prevent any significant esterase activity. The esterase reaction relies on interactions between enzyme residues and the ester bond's two oxygen atoms and carboxylic carbon atom [173]. As the α-carbon is not involved in the reaction mechanism it is likely that the alkyl bumps protect the ester bond through sterically impairing enzyme:compound binding and not through effects on the reaction chemistry itself.

Compound	R_i	R_{ii}	MeO-Position	CLogP	Plasma t1/2 (min)	CL int (ml/min/g)	Pe (nm/s)
(+) JQ1	H	O^{t}Bu	-	4.8	>180	7.5	
I-BET762	H	NHEt	8	2.8	>180	1.3	25
Scaffold	H	OMe	8	3.3	54	<0.5	149
9-Scaffold	H	OMe	9	3.3	67	1.7	
ME	methyl	OMe	8	3.5	>180	0.7	185
ET	ethyl	OMe	8	3.9	>180	1.5	153
AL	allyl	OMe	8	3.9	>180	6.7	127
PR	propyl	OMe	8	4.3	>180	5.2	
9-ME	methyl	OMe	9	3.5	>180	<0.5	158
9-ET	ethyl	OMe	9	3.9	>180	4.1	155
9-AL	allyl	OMe	9	3.9	>180	9.4	136
ME-Am1	methyl	NHEt	8	3.1	>180	1.2	26
ET-Am1	ethyl	NHEt	8	3.5	>180	1.4	45
AL-Am1	allyl	NHEt	8	3.5	>180	2.6	52
ME-Am2	methyl	N[Et]$_2$	8	3.8	>180	>50	
ET-Am2	ethyl	N[Et]$_2$	8	4.1	>180	>50	
9-ME-Am1	methyl	NHEt	9	3.1	>180	1.5	40
9-ET-Am1	ethyl	NHEt	9	3.5	>180	2.1	59
ET-tBu	ethyl	O^{t}Bu	8	4.8	>180	30.7	
AL-tBu	allyl	O^{t}Bu	8	4.9	>180	39.1	
9-ET-tBu	ethyl	O^{t}Bu	9	4.8	>180	31.0	
9-AL-tBu	allyl	O^{t}Bu	9	4.9	>180	39.4	
Proacaine					13		
Verrapamil						30.0	
Atenolol							0
Propanolol							82

Table 3.2. Bumped compound secondary DMPK triage

(+)-JQ1 possessed a thiophene ring and hence has no methoxy group. Refer to scheme 3.1 for positions of R_i and R_{ii}. CLogP measured in Stardrop. Controls: proacaine (low plasma half-life), verapamil (high intrinsic clearance), atenolol (low permeability) and propranolol (high permeability).

In addition to plasma stability we also looked at stability in the presence of liver microsomes, a way of monitoring CYP-mediated metabolism key to *in vivo* stability [172]. In this assay both our scaffolds showed low CL_{int} values which were largely unaffected by alkyl bumps, the shifted methoxy group or the ethyl-amide side group. As CYP enzyme binding is primarily driven by hydrophobic interactions [174] it was not surprising to see increased clearance for compounds with the *tert*-butyl ester (and to a lesser extent the propyl bump). Compounds with the di-ethyl-amide side-group also showed excessively high intrinsic clearance, although it is likely that these compound are simply being converted to the mono-ethyl-amide side-group [175]. Finally, the allyl bump showed some small increases in clearance that did not necessarily disqualify any compounds but was recognised as not ideal moving forwards. This effect is probably

the result of epoxide formation [176], a known issue limiting the utilisation of allyl groups on drug development. This assay proved very useful in reducing the list of candidate compounds for ITC screening, by disqualifying the propyl bump, di-ethyl-amide side-group and the *tert*-butyl ester side-group, while highlighting the methyl and ethyl bumps, methoxy-shift and methyl-ester/ethyl-amide side-groups.

As several compounds had already been disqualified at this stage, the PAMPA assay was next carried out on a more select range of compounds. By measuring the movement of compounds across an artificial membrane this assay makes for an easy and reliable (although not perfect) indicator of passive cell membrane permeability [172]. The PAMPA assay's main limit is that it does not cover active transport into and out of cells. The key outcome of this assay was to show that, although all tested compounds could be considered to possess high-permeability (>20 nm/s) the presence of the original methyl-ester side-group generated very high permeability (127-185 nm/s) compared to the ethyl-amide side-group (25-59 nm/s), and highlighted these compounds as the favourites going into ITC screening. The nature of the bumps and position of the methoxy group show little impact on permeability.

3.7 – Tertiary ITC Screen of Bumped Compounds

Taking together the results of the primary AlphaLISA screen and the secondary DMPK triage the following modifications were chosen as the most promising: the methyl, ethyl and allyl bumps (with allyl as the least promising), the methoxy group at the 8' or 9' position, and the methyl-ester and ethyl-amide side-groups (with the methyl-ester most promising). This resulted in the following compounds undergoing further characterisation: ET, AL, 9-ME, 9-ET, ME-Am1, ET-Am1 and 9-ME-Am1.

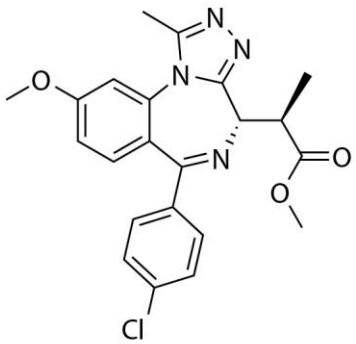
As expected, based on their success in previous SAR tests, all compounds excelled in this thorough ITC screening, with L/V affinities close to or exceeding that of our un-bumped scaffold (180 nM) and L/V selectivity of >30-fold (table 3.3). Of these compounds 9-ME and 9-ET were the clear outliers and their ITC titrations were replicated until three consistent K_d values were recorded for each construct (table S3.5). This additional testing showed these compounds to possess mean selectivity values >100-fold while maintaining potency (figure 3.6). Of the two compounds, 9-ET

showed greater overall selectivity, but at the cost of reduced L/V potency. Both compounds can be used as highly selective chemical probes, but for the bulk of later experiments 9-ME was preferred as it is the more potent compound while showing >100-fold selectivity, a slightly better DMPK profile and a lower molecular weight.

Compound	R_i	R_{ii}	MeO-Position	ITC pKd ±SD WT	L/V	Δ
Scaffold	H	OMe	8	6.8	6.5	-0.3
ET	ethyl	OMe	8	5.1 ±0.5	6.9 ±0.3	1.8 ±0.4
AL	allyl	OMe	8	5.1 ±0.4	6.6 ±0.2	1.5 ±0.5
9-ME	methyl	OMe	9	4.5 ±0.2	6.6 ±0.3	2.0 ±0.4
9-ET	ethyl	OMe	9	<4.2	6.4 ±0.2	>2.2 ±0.2
ME-Am1	methyl	NHEt	8	5.6 ±0.3	6.9 ±0.2	1.3 ±0.3
ET-Am1	methyl	NHEt	8	<4.2	6.0 ±0.2	>1.8 ±0.2
9-ME-Am1	methyl	NHEt	9	4.5 ±0.4	6.3 ±0.2	1.8 ±0.5

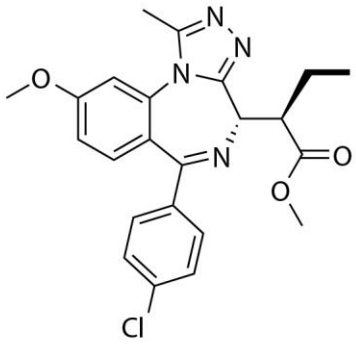
Table 3.3. Bumped compound tertiary ITC screen

Results of ITC titrations of bumped compounds against all somatic WT & L/V BET bromodomains. Refer to scheme 3.1 for positions of R_i and R_{ii}. Scaffold tested only against BRD4 BD1. 9-ME & 9-ET titrations replicated until 3 consistent K_d values observed for each construct.



9-ME

	L/V	2.1	2.2	3.1	3.2	4.1	4.2
WT	Kd (nM)	310	330	140	230	140	980
2.1	32000	100	100	230	140	230	30
2.2	>60000	190	180	430	260	430	60
3.1	12000	40	40	90	50	90	10
3.2	34000	110	100	240	150	240	40
4.1	28000	90	90	200	120	200	30
4.2	33000	110	100	240	140	240	30



9-ET

	L/V	2.1	2.2	3.1	3.2	4.1	4.2
WT	Kd (nM)	340	300	270	410	530	740
2.1	>60000	180	200	230	150	110	80
2.2	>60000	180	200	230	150	100	80
3.1	29000	90	100	110	70	60	40
3.2	>60000	180	200	230	150	110	80
4.1	>60000	180	200	230	150	100	80
4.2	>60000	180	200	230	150	100	80

Figure 3.6. Selectivity of 9-ME & 9-ET.

Chemical structures and selectivity grids of 9-ME & 9-ET. Selectivity grids based on ITC titrations, replicated until 3 consistent K_d values were observed for each construct.

3.8 – Crystallographic Structural Analysis of Bumped Compounds

Throughout the compound evaluation process a number of compounds were co-crystallised with BRD2 BD2 L/V and their structures solved to better inform observed SAR (figure 3.7). All of these compounds were shown to bind in the same basic orientation, with none of the modifications drastically altering binding. Some modifications did lead to slight re-organisations of the ligand or nearby protein. These structures also confirmed that only the 2R,3S enantiomer is an active inhibitor of the BET bromodomains (figure S3.2). Ligand Fo-Fc maps are shown in figure S3.3.

While the AlphaLISA SAR emphasises the importance of each bump's rotational freedom to L/V potency analysis of co-crystal structures does show that the bump's size alone also plays a role. Moving from smaller to larger bumps the alkyl bump and valine 'hole' can be seen to interact less intimately, with the measured distance between the valine Cα and the ligand's C2 (the base of the bump) increasing from 6.8 Å for ME to 7.2 Å for ET and 7.5 Å for AL (figure 3.7A). This is not surprising, with the L/V mutation only 'removing' one heavy atom from the binding site and clearly not making sufficient room to fully 'bury' 2 or 3 carbon chains.

The co-crystal structures of 9-ME (figure 3.7C) shows that shifting the methoxy group to the 9' position has not visibly altered how the ligand binds the L/V bromodomains. While the methoxy group has now been moved towards the binding site surface there are no residues it can be seen to clash with. The co-crystal structure of 9-ET however, still does not show any clear steric clash, but seems to have slightly re-oriented the ligand and the valine-bearing ZA loop. It is possible that the methoxy-shift only creates a serious steric clash with regard to WT bromodomains, where one can see how it would be difficult for a ligand to orient itself such that the alkyl bump does not clash with the leucine side-chain, while the 9' methoxy does not clash with any residues in the WPF shelf or ZA loop region.

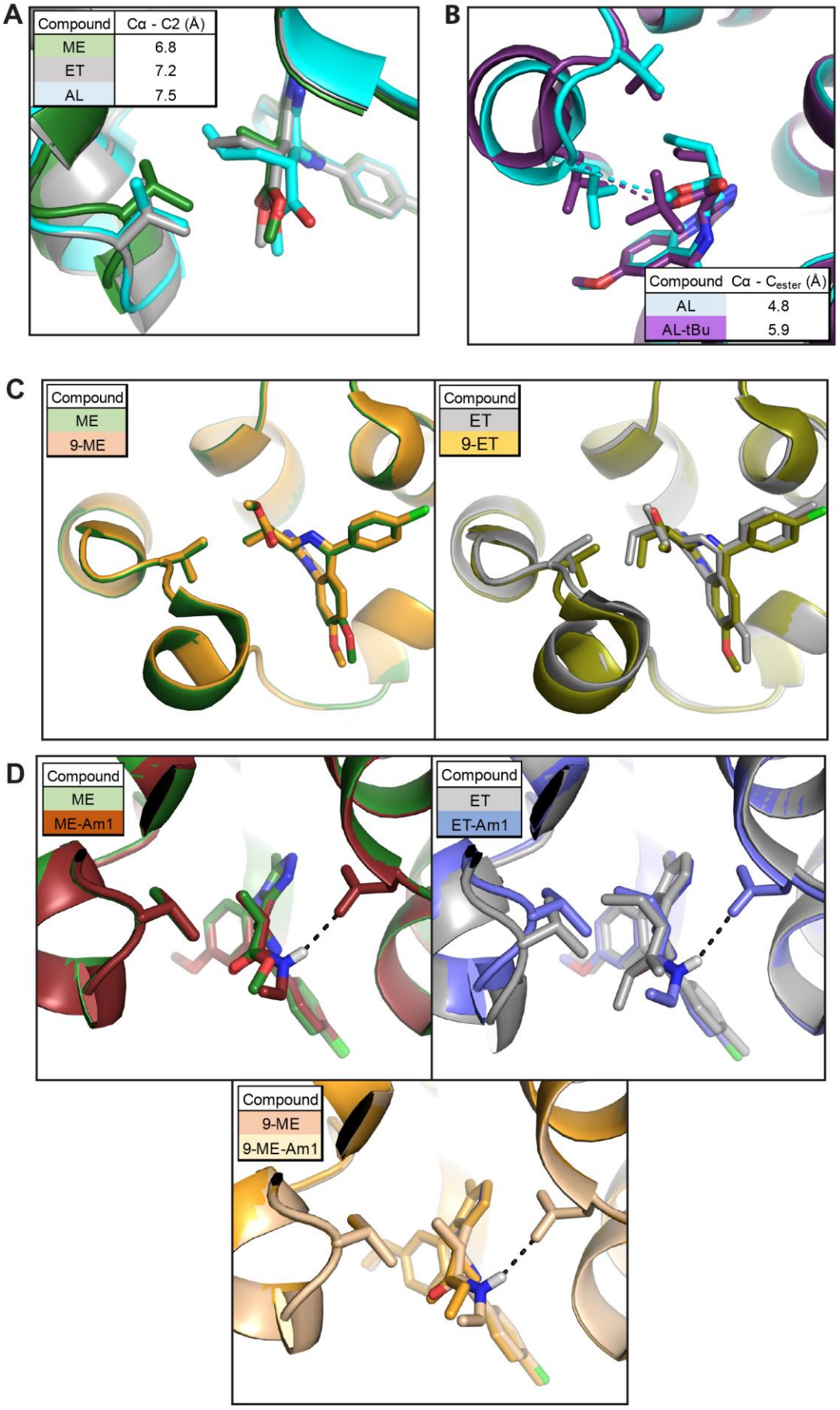

A
Compound
Cα - C2 (Å)
ME
6.8
ET
7.2
AL
7.5
B
Compound
Cα - Cester (Å)
AL
4.8
AL-tBu
5.9
C
Compound
ME
9-ME
Compound
ET
9-ET
D
Compound
ME
ME-Am1
Compound
ET
ET-Am1
Compound
9-ME
9-ME-Am1

Figure 3.7. Analysis of bumped-compound co-crystal structures.

Comparison of aligned BRD2 BD2 L/V : bumped compound co-crystal structures. BRD2 BD2 L/V in complex with (left-to-right, top-to-bottom): ME (5O39), ET (5O3A) and AL (5O3B); AL and AL-tBu (5O3I); ME and 9-ME (5O3C); ET and 9-ET (5O3D); ME and ME-Am1 (5O3E); ET and ET-Am1 (5O3F); 9-ME and 9-ME-Am1 (5O3H). Black dashes denote amide:asparagine hydrogen bonds. Coloured dashes represent inter-atom distances of interest. Cα – alpha atom of valine residue; C2 – carbon #2 in ligand; C_{ester} – carbon after ester bond.

The ethyl-amide and di-ethyl-amide side-group modifications were introduced with the aim of restricting the movement of the alkyl bump through hydrogen bonding to a conserved asparagine. The co-crystal structures of ME-Am1, ET-Am1 and 9-ME-Am1 all show this hydrogen bond (to N429 in this case) to be present (figure 3.7D). For ME-Am1 and 9-ME-Am1 this hydrogen bond has not altered the orientation of the alkyl bump or surrounding protein. ET-Am1, however, shows some disruption to the rest of the compound and the ZA loop around the mutated valine, which may be connected to its reduced L/V potency. The large drop in WT potency observed by all amide compounds can be seen to be the product of these hydrogen bonds, preventing the bump and flexible methyl-ester from moving into more WT-favourable orientations. Unfortunately our SAR also shows similar reductions in L/V potency (for some compounds) – leaving similar overall selectivity – which may be due to the entropic penalty associated with ligand:protein hydrogen bonding. This impact is most severe with ET-Am1, where our structures also show the amide group has disturbed much of the binding site.

Comparing the co-crystal structures of AL and AL-tBu gives insights into the poor performance of the *tert*-butyl-ester compounds (figure 3.7B). The *tert*-butyl group appears to be clashing sterically with L381, with the distance between the L381 Cα and the C_{ester} (carbon atom following the ester bond) changing from 4.8 Å with AL to 5.9 Å with AL-tBu. This clash is likely the main cause of the drop in potency observed, although the high-potency of (+)-JQ1 suggests this occurs only when paired with our alkyl bumps.

3.9 – Separation and Evaluation of Enantiomers

As discussed previously, our compounds are synthesised as a mixture of four isomers with the carbons in positions 2 and 3 forming chiral centres. The inactive *2R,3R* / *2S,3S* diastereoisomer was removed by reverse-phase HPLC, leaving the active *2R,3S* and inactive *2S,3R* enantiomers. Due to practical issues with separating these enantiomers, compounds were tested as racemic mixtures, through the described screening, using just the concentration of the active form. Following discussions with Reach Separations Ltd (a company specialising in analytical and preparative supercritical fluid chromatography SFC) it was decided to send in samples of 9-ME and 9-ET for separation. Separated enantiomers would then be assayed to confirm our active/inactive hypothesis, inform to what extent using the racemic mixture was still advisable, and test if using the purified active enantiomer would bring any benefits.

9-ME and 9-ET were re-synthesised in bulk by Dr Andrea Testa and sent to Reach Separations Ltd. Racemate powder was solubilised in ethanol at 20 mg/ml, and then purified by HPLC on a Lux A1 column (21.2 mm x 250 mm, 5 µm). Samples were injected at a volume of 1 ml with 4:6 HEPT:EtOH (0.1% v/v NH_3) and ran at a flow rate of 21 ml/min. This protocol gave recovery rates of ~50% (25% for each enantiomer (Table 3.4.). Compounds could be separated into two clear peaks of high purity (figure S3.4.), for testing by AlphaLISA and ITC. For the purposes of these assays ligand concentrations refer to total compound and not just the active enantiomer, for example 100 nM racemate means 50 nM each of the 2 enantiomers.

Compound	Sample (mg)	Sample purity (%)	Peak 1 yield (mg)	Peak 1 enantiomeric excess	Peak 2 yield (mg)	Peak 2 enantiomeric excess
9-ME	42	96	11	98	9	97
9-ET	15	95	5	100	4	97

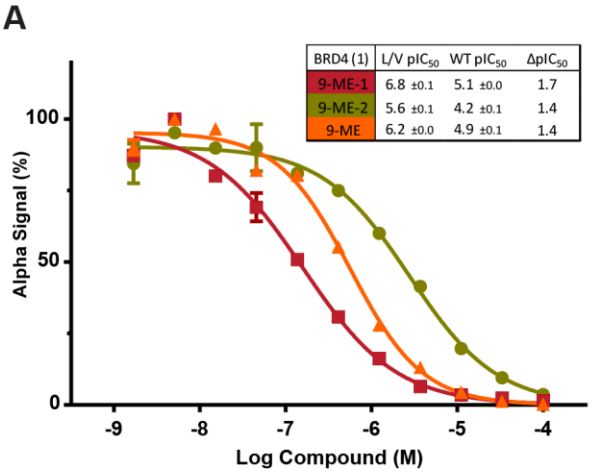
Table 3.4. Separation of bumped-compound enantiomers

Data from Reach Separations Ltd.

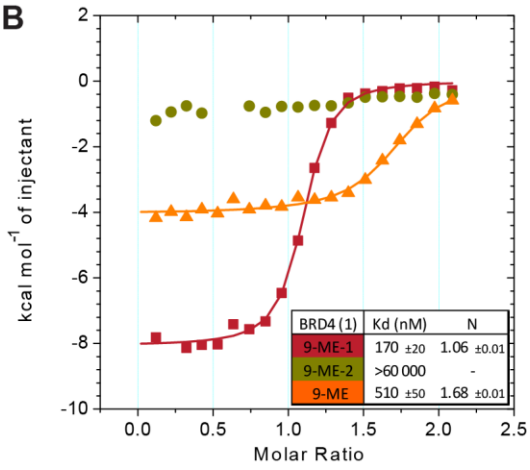
Separated 9-ME and 9-ET enantiomers were tested first in the AlphaLISA assay, against BRD4 BD1 WT and L/V (figure 3.8A,C). For both compounds the first peak (9-ME-1 and 9-ET-1) showed as both more potent and selective than the racemic mixture, while the second peak (9-ME-2 and 9-ET-2) was the least potent and L/V-selective. This

strongly suggested that the first peak was the active *2R,3S* enantiomer and the second peak the inactive *2S,3R* enantiomer, but it was decided to use ITC to fully confirm this. ITC titrations of the 9-ME and 9-ET enantiomers (again against BRD4 BD1 WT and L/V) confirm that the first peak is the active enantiomer, with high-affinity interactions with N values close to 1, and that the second peak is inactive, with no measurable binding even to the L/V bromodomain (figure 3.8B,D). The racemic mixture then shows less potent binding and with elevated N values – suggesting only half of the mixture is active.

A



B



C

BRD4	9-ME (mix)		9-ME-1		9-ME-2		9-ET (mix)		9-ET-1		9-ET-2	
BD1	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V
pIC_{50}	4.9	6.2	5.1	6.8	4.2	5.6	4.3	5.7	4.1	5.9	4.6	5.0
SE	0.07	0.04	0.03	0.07	0.07	0.05	0.08	0.06	0.11	0.06	0.18	0.12
ΔpIC_{50}	1.4		1.7		1.4		1.4		1.8		0.4	

D

BRD4	9-ME (mix)	9-ME-1		9-ME-2	
BD1	L/V	WT	L/V	WT	L/V
Kd (nM)	510 ±50	57500 ±4000	170 ±20	>60000	>60000
ΔH	-4.2 ±0.0	-13.3 ±0.5	-7.9 ±0.1		
N	1.7 ±0.0	1.0 (FIX)	1.1 ±0.1		

BRD4	9-ET (mix)	9-ET-1		9-ET-2	
BD1	L/V	WT	L/V	WT	L/V
Kd (nM)	840 ±170	>60000	440 ±50	>60000	>60000
ΔH	-3.5 ±0.1		-6.5 ±0.1		
N	1.4 ±0.0		0.85 ±0.0		

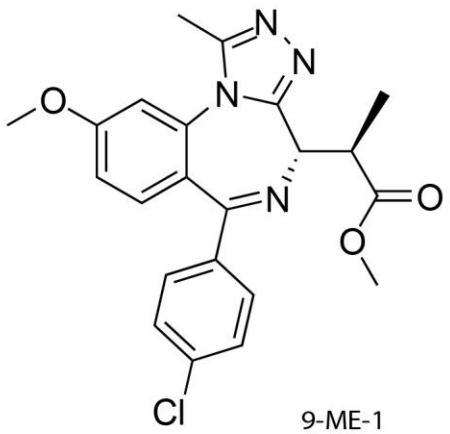
Figure 3.8. Identification of active/inactive enantiomers

Purified 9-ME enantiomers, and racemic mixture thereof, titrated against BRD4 BD1 L/V in AlphaLISA assay (A) and ITC (B). C) Expanded AlphaLISA data. D) Expanded ITC data.

Following the identification of the active *2R,3S* 9-ME enantiomer (hereafter named 9-ME-1) this compound was titrated against all WT and L/V BET bromodomains in ITC, to determine the new potency and selectivity values (figure S3.5). This testing showed 9-ME-1 to be, on average, slightly more potent and selective than 9-ME. This improvement was not universal however, with a slight increase in potency for WT BRD3 BD1 and BD2 and a corresponding decrease in selectivity against these bromodomains.

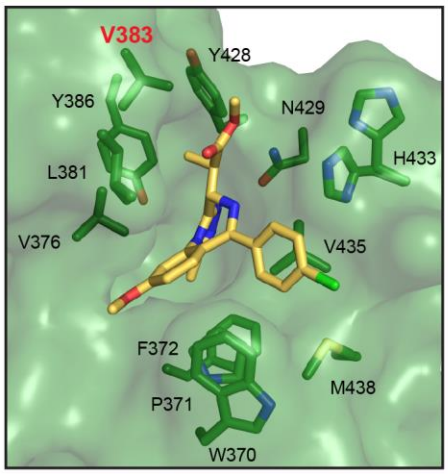
Based on this ITC data, alongside our AlphaLISA SAR and DMPK investigations, 9-ME-1 was selected as the preferred bumped compound for selective inhibition of L/V bromodomains. Starting with our scaffold compound, with no selectivity, ~150 nM affinity and excellent DMPK properties (with the exception of esterase vulnerability) we have progressed to a compound with ~200 nM affinity, ~200-fold selectivity and even better DMPK characteristics (figure 3.9).

A



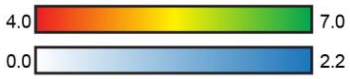
9-ME-1

B



C

		Scaffold	9-ME	9-ME-1
AlphaLISA pIC_{50}	WT	6.6 ± 0.2	5.4 ± 0.2	5.1
	L/V	6.6 ± 0.3	6.8 ± 0.2	6.8
	Δ	0.0 ± 0.2	1.4 ± 0.2	1.7
Plasma t½	(mins)	54	>180	
CL_{int}	(ml/min/g)	<0.5	<0.5	
P_e	(nm/s)	149	158	
ITC pKd	WT	6.8	4.5 ± 0.2	4.5 ± 0.3
	L/V	6.5	6.6 ± 0.3	6.7 ± 0.2
	Δ	-0.3	2.0 ± 0.4	2.2 ± 0.4



D

	L/V	2.1	2.2	3.1	3.2	4.1	4.2
WT	Kd (nM)	210	200	90	310	170	300
2.1	43000	210	220	480	140	250	140
2.2	>60000	>290	>300	>670	>190	>350	>200
3.1	7500	40	40	80	20	40	30
3.2	24000	110	120	270	80	140	80
4.1	58000	280	290	640	190	340	190
4.2	29000	140	150	320	90	170	100

Figure 3.9. 9-ME-1 as an L/V-selective bumped BET inhibitor

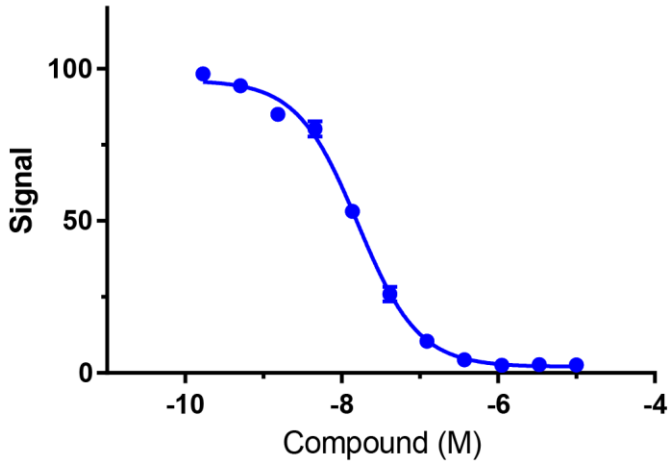
A) Chemical structure of 9-ME-1. B) Co-crystal structure of 9-ME-1 bound to BRD2 BD2 L/V, with key residues highlighted. C) SAR of scaffold, 9-ME and 9-ME-1. 9-ME-1 AlphaLISA data and scaffold data from BRD4 BD1 only, other data is mean ±SE of titrations against all somatic BET bromodomains. D) Selectivity grid of 9-ME-1.

3.10 – Screening for Off-Targets

Throughout our compound evaluation we focused on the selectivity of our compounds for L/V BET bromodomains over their WT counterparts. This approach did not take into consideration the possibility of off-target binding to other bromodomains, or completely unrelated proteins. Off-target screening of other related BET-inhibitor scaffolds [76, 99, 115] suggest our scaffold should show little binding to non-BET bromodomains, and presumably the presence of our bumps should minimize this even more. Despite this, it was decided to confirm these assumptions thoroughly, and to check for binding to completely unrelated proteins, which has been shown by other BET inhibitors [76]. To do this 9-ME-1 was screened against three commercial protein panels: 1) DiscoverRX's BROMOscan panel of 32 (of 56) human bromodomains, 2) the MRC ICKP's express screen of 50 kinases and 3) CEREP's ExpresS Profile of 55 receptors, ion channels and transporters.

As a positive control for the BROMOscan assay we first had our scaffold titrated against BRD4 BD1 WT (figure 3.10A), which produced a K_d value of 15 nM – 10-fold lower than the equivalent values from ITC (150 nM) and AlphaLISA (140 nM) (figure 3.10B). This increased sensitivity in BROMOscan assays has been observed previously and theorised to be due to differences in construct design and assay format [99, 121]. 9-ME-1 was then screened against their WT bromodomain panel at a concentration of 1 μM, this dose was chosen as it was generally realistic, yet high enough to allow detection of even weak off-target activity. Against the panel 9-ME-1 showed 50-80% displacement of WT BET bromodomains, which is surprisingly high given our ITC data, however our scaffold titration data suggests the assay is ~10-fold more sensitive than other assays, so this activity was not discouraging. More importantly, 9-ME-1 displaced other, non-BET bromodomains to a lesser extent than the WT BET bromodomains (figure 3.10C). Based on this data, it can be concluded that if 9-ME-1 has ~200x selectivity for L/V BET BDs versus WT BET BDs, it should possess >200x selectivity for L/V BET BDs over other bromodomains.

A



B

Scaffold vs BRD4 BD1 WT	
ITC Kd	150 nM
AlphaLISA IC_{50}	140 nM
BROMOscan Kd	15 nM

C

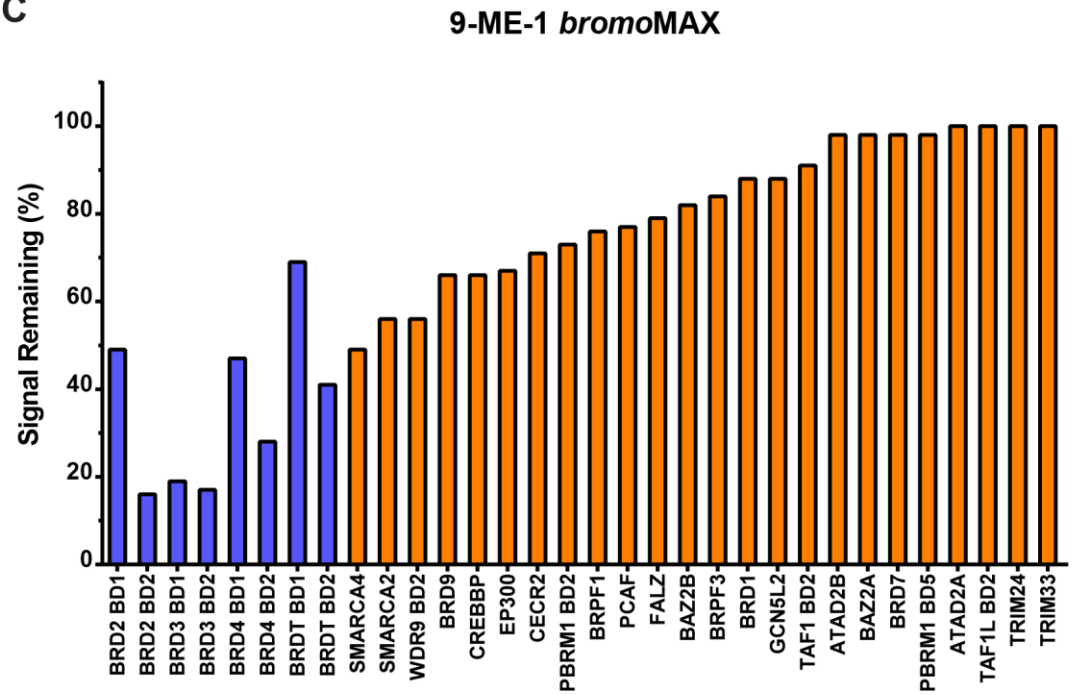


Figure 3.10. 9-ME-1 BROMOscan results

A) Titration of scaffold compound against WT BRD4 BD1 in *bromo*kdELECT assay. B) Affinity of scaffold compound for WT BRD4 BD1, measured by different assays. C) Retention of WT bromodomains, following incubation with 1 µM 9-ME-1, with BET bromodomains in blue and non-BET bromodomains in orange. DiscoveRX BROMOscan *bromo*MAX panel (https://www.discoverx.com/services/drug-discovery-development-services/epigenetic-profiling/bromoscan-epigenetic-profiling/bromomax).

Outside of the bromodomain family, 9-ME-1 showed no off-target binding to any kinase tested in the MRC express screen, with no more than 20% inhibition at a concentration of 1 μM (figure 3.11). In the CEREP ExpresS Profile of receptors, ion channels and transporters 9-ME-1 at 1 μM also showed less than 20% inhibition against all targets, with the sole exception of the MT1 melatonin receptor (figure 3.12). The relevance and implications of this identified off-target are elaborated upon in the Discussion (section 3.11) .Pleasingly, despite being based on a benzodiazepine scaffold, 9-ME-1 showed no binding to the central benzodiazepine or GABA receptors, likely due to the presence of the bump and side-group [76].

Following the off-target screening in three diverse panels, containing 137 proteins, 9-ME-1 was shown to be highly specific for the L/V BET bromodomains, with little binding to WT non-BET bromodomains and only a single strong off-target. As such, we can confidently conclude that phenotypes arising from its use in cells can be reliably attributed to L/V BET bromodomain inhibition and not to the inhibition of unrelated proteins.

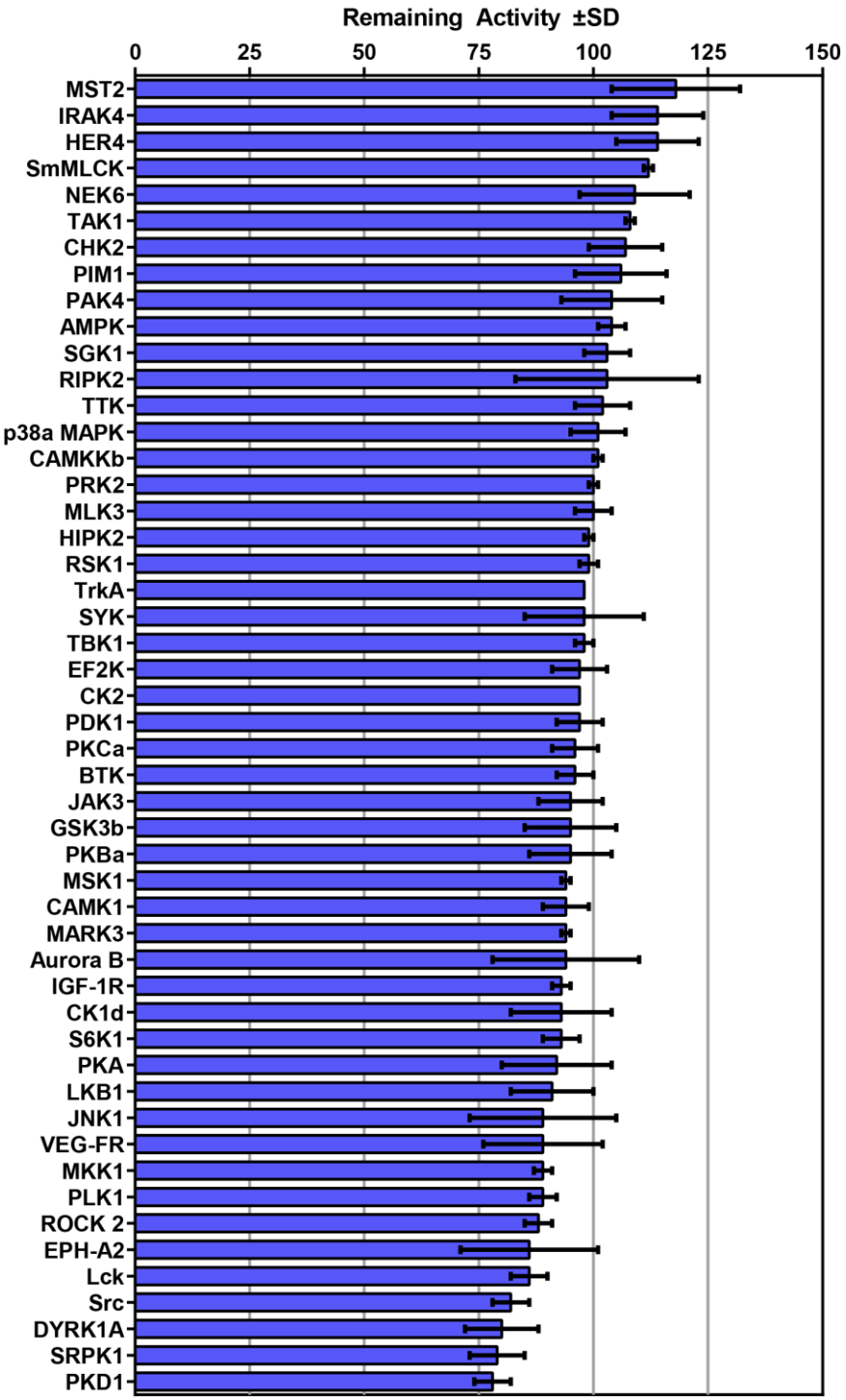


Figure 3.11. 9-ME-1 kinase screen.

Remaining activity of panel of human kinases following incubation with 1 µM 9-ME-1. Dundee MRC-PPU Express Screen (http://www.kinase-screen.mrc.ac.uk/services/express-screen).

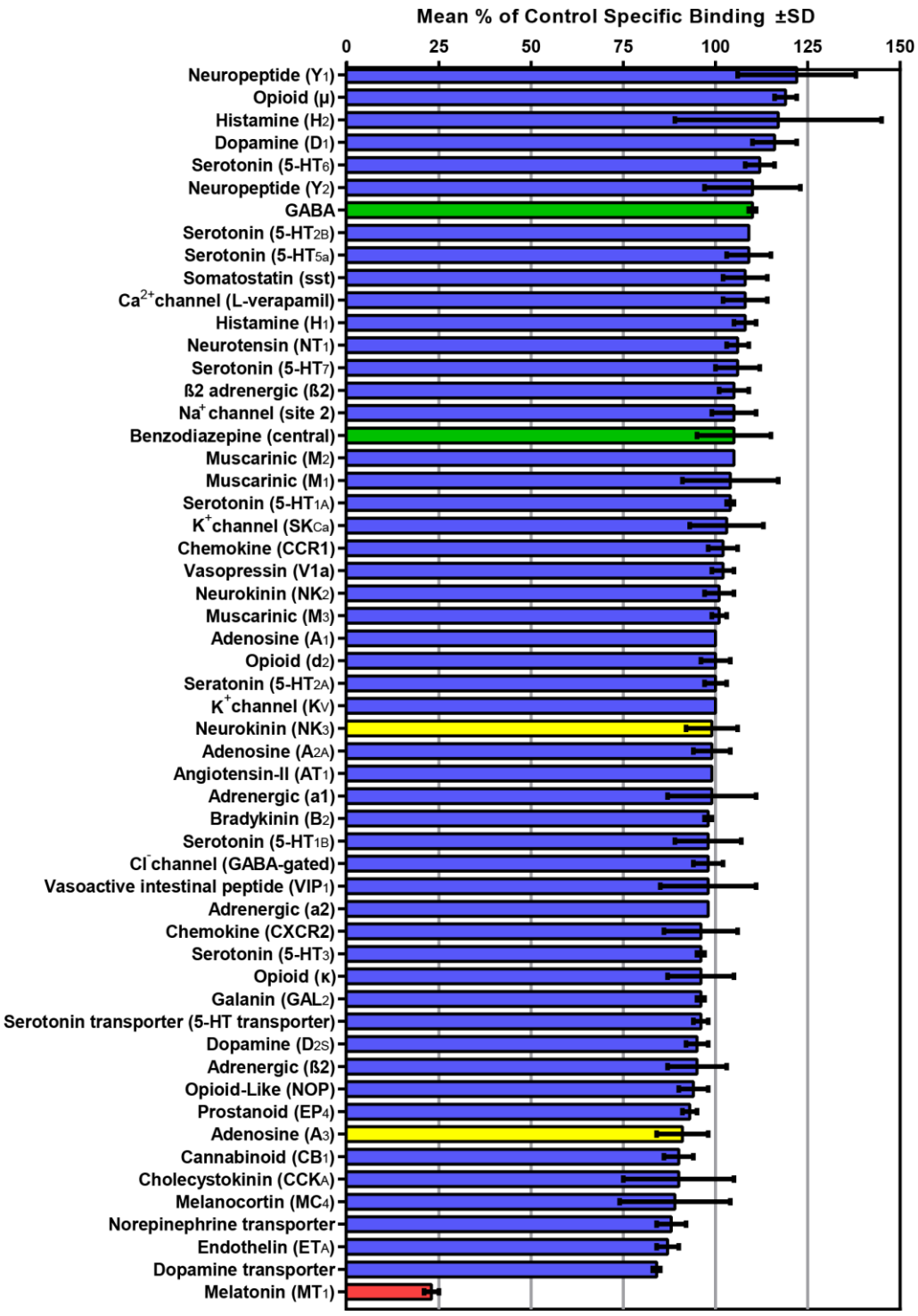


Figure 3.12. 9-ME-1 receptor, ion-channel & transporter screen.

Activity of panel of human receptors, ion-channels and transporters following incubation with 1 μM 9-ME-1. Green denotes central benzodiazepine and GABA receptors, yellow denotes known (+)-JQ1 off-targets and red denotes 9-ME-1 off-target (MT1). CEREP ExpresS Profile (http://www.cerep.fr/cerep/utilisateurs/pages/downloads/Documents/Marketing/Pharmacology%20&%20ADME/Standard%20profiles/ExpresSProfile_2014v2LD.pdf)

3.11 – Discussion

For the high-throughput bromodomain-inhibition assay the newer and more expensive AlphaLISA format was chosen over the older AlphaScreen technology. AlphaLISA uses europium-chelate based acceptor beads, instead of Alphascreen's anthracene and rubrene dyes (both beads also utilise thioxene). This results in AlphaLISA beads emitting light over a much narrower range of wavelengths – 607-623 mm compared to 520-620 nm – reducing its susceptibility to absorbance by interfering compounds or protein/bead precipitation [170]. Additionally, the AlphaLISA acceptor beads are recruited to His-tagged proteins through an anti-His6 antibody instead of the Ni chelator used by Alphascreen, which is also more vulnerable to interference. Previous projects within the group faced difficulties in developing reliable Alphascreen assays, but saw a clear improvement with the AlphaLISA beads. The AlphaLISA format is, theoretically, still vulnerable to interference from compounds that quench singlet oxygen or compete with Bio-JQ1 for the streptavidin conjugated to the donor beads. Despite this, given the limited number of compounds tested, and their high structural similarity, such interference was not expected to be an issue. During this project a similar AlphaLISA-based Bio-JQ1 / BET bromodomain assay was developed and validated by another lab [177]. The proximity-based nature of the AlphaLISA assay lead to its re-purposing for another project in the lab: PROTAC development. In this assay instead of measuring the ability of a biotinylated probe to bring together a BRD and a streptavidin-labelled bead, PROTAC compounds were assessed by their ability to bring together BRDs and biotinylated E3 ubiquitin ligases [133, 178, 179]. This assay was able not just to confirm whether specific PROTACs could recruit both their BRD and E3 targets, but was also able to indicate which PROTACs may display stronger cooperativity of ternary complex formation, a parameter now known to be very important to PROTAC design [131].

If our compounds did interfere with, and prevent the use of, the AlphaLISA assay format then some alternate technologies could be used. TR-FRET assays are very similar to AlphaLISA, monitoring the proximity of a fluorophore-labelled probe compound and a Europium chelate recruited to His-tagged proteins, but relies on the FRET phenomenon rather than singlet oxygen diffusion to trigger a signal. This technology has previously been used to study BET inhibitors (especially by scientists at

GSK) with much success [99, 121, 180]. The assays would likely be more practical to run, as there are fewer components to add and no need to protect the assay plate from light, although the shorter range of the FRET phenomenon may require careful construct and probe design. One final plate-based technique is fluorescence polarisation (FP), in which a fluorophore-labelled probe is observed. The light emitted by the excited fluorophore is polarised and the difference between the parallel and perpendicular intensity is quantified. As interacting with the BRD will, in effect, increase the MW of the fluorophore it will slow its rotation and increase the polarisation of the emitted signal [181]. This technology allows for an even simpler protocol than TR-FRET, and is also seen in the literature [99], but is limited by the inability to detect binding stronger than that possessed by the labelled probe. The biophysical techniques of BLI and SPR [163] could also be developed for evaluating BET bromodomain inhibitors, but this would prevent the use of high-throughput 384-well plates, and such techniques will detect binding to any part of the bromodomain construct while the competitive AlphaLISA / TR-FRET / FP assays require compounds to actually bind the Kac binding site of the bromodomain.

One rational behind the decision to trial the methoxy-shift modification was that, as no chemical groups are removed or added, it was not expected to greatly alter the compound's DMPK characteristics compared to the non-shifted compound. This was largely observed to be true (table 3.5.) with no changes to CLogP or plasma t½. The only large changes were in the CL_{int} of ET and AL compounds, and the P_e of ME-Am1 and ET-Am1. Additionally, the hypothesis that changes to the side-groups (originally a methyl-ester) would be an effective means of tuning DMPK properties was also confirmed (figure 3.13). Alternate side-groups did not have any visible impact on plasma t½, as it was revealed that the mere presence of the alkyl bump was enough to prevent esterase metabolism. The *tert*-butyl ester side-group had a massive effect on CL_{int}, due to its increased hydrophobicity, and the main determinant of Pe was the presence of a methyl-ester or ethyl-amide side-group.

Compound	R_i	R_{ii}	MeO-Positio	CLogP	Plasma t1/2	CL int (ml/min/g)	Pe (nm/s)
Scaffold	H	OMe	8	3.3	54	<0.5	149
9-Scaffold	H	OMe	9	3.3	67	1.7	
ME	methyl	OMe	8	3.5	>180	0.7	185
9-ME	methyl	OMe	9	3.5	>180	<0.5	158
ET	ethyl	OMe	8	3.9	>180	1.5	153
9-ET	ethyl	OMe	9	3.9	>180	4.1	155
AL	allyl	OMe	8	3.9	>180	6.7	127
9-AL	allyl	OMe	9	3.9	>180	9.4	136
ME-Am1	methyl	NHEt	8	3.1	>180	1.2	26
9-ME-Am1	methyl	NHEt	9	3.1	>180	1.5	40
ET-Am1	ethyl	NHEt	8	3.5	>180	1.4	45
9-ET-Am1	ethyl	NHEt	9	3.5	>180	2.1	59
ET-tBu	ethyl	O^{t}Bu	8	4.8	>180	30.7	
9-ET-tBu	ethyl	O^{t}Bu	9	4.8	>180	31.0	
AL-tBu	allyl	O^{t}Bu	8	4.9	>180	39.1	
9-AL-tBu	allyl	O^{t}Bu	9	4.9	>180	39.4	

Table 3.5. Effect of methoxy-shift on DMPK properties.

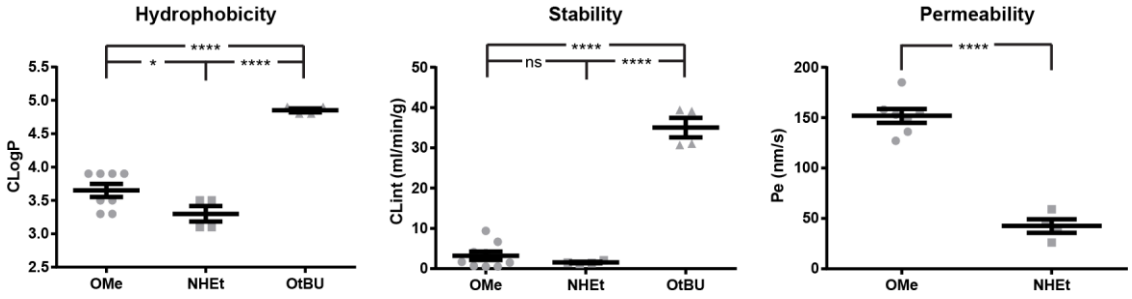


Figure 3.13. DMPK properties of different side-groups.

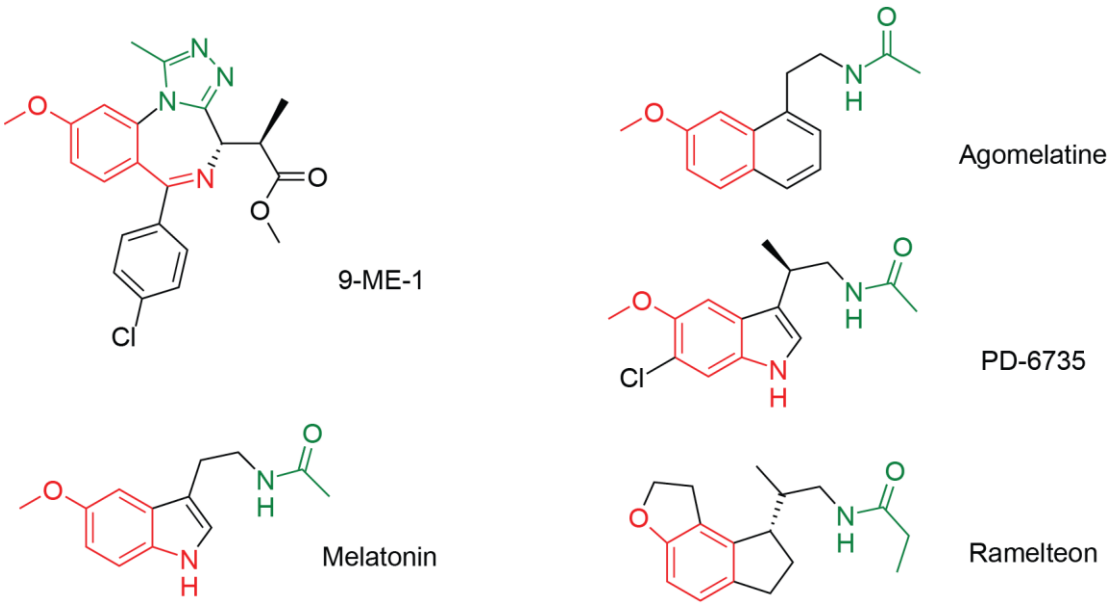
Bars represent the mean ± SE of each compound with each listed side-group. Statistical significance determined with two-tailed t tests: ns P>0.05; * P≤0.05; ** P ≤0.01; *** P≤0.001; **** P<0.0001.

9-ME-1 was chosen as the 'best' bumped compound, as out of the two compounds showing >100-fold selectivity it possessed greater potency and stability. Several of our other compounds could, nonetheless, still prove useful, as different projects or experiments may prioritise different parameters. 9-ET-1, while not as potent, possessed even greater overall selectivity; while ET does not show consistently >100-fold selectivity but features potency of almost 100 nM. Furthermore, ME-Am1 was the most potent compound to reach the final ITC screen, while ET-Am1 showed no detectable binding to any WT bromodomain.

The bump-&-hole concept need not be utilised solely in the context of small-molecule inhibitors. Bumped compounds could be attached to functional groups, in a manner similar to Bio-JQ1 [86], photo-crosslinkers [87, 88] and click-chemistry enabled probes [89], for protein pulldown, fluorescence microscopy and other techniques. Instead of functional groups bumped compounds could be linked to other small-molecules, either to form bivalent BET inhibitors like MT1 [83, 85], dual-inhibitors [182], or PROTAC compounds. Hopefully such efforts will be aided both by our extensive bumped-compound SAR and the extensive characterisation of benzodiazepine-scaffold BET inhibitors in the literature.

Our off-target screening efforts were, ultimately, encouraging but unfortunately did show significant off-target activity against the melatonin receptor MT1, against which we estimate 9-ME-1 has a K_d of roughly ~500 nM. If 9-ME-1 binds MT1 with 1:1 stoichiometry there should be 1 log unit of 9-ME-1 concentration between 50% and 90% inhibition, hence ~77% inhibition at a dose of 1 µM suggests an IC_{50}/K_d value well between 100 nM and 1 µM. Although this off-target screen is not perfect it can still be considered very encouraging, as the widely used (+)-JQ1 also bound MT1 (more weakly) and in addition presented strong binding to the NK2 neurokinin and A3 adenosine receptors [76, 85]. As MT1 is primarily expressed in the CNS and regulates processes such as the circadian rhythm off-target MT1 inhibition is unlikely to complicate the sort of experiments we aim to use 9-ME-1 for [183].

Structural analysis of MT1, and computational docking experiments with 9-ME-1 were not performed as there are currently no crystal or NMR structures of MT1 and there is some disagreement between homology models [184]. Nevertheless, comparing the chemical structures of 9-ME-1 and a selection of MT1 agonists [185] showed some structural similarities (scheme 3.3). MT1 binding may be enhanced by the 9’ methoxy ring, and altering this group may weaken MT1 binding, at the cost of L/V selectivity. The use of further modifications may weaken MT1 binding, but is unlikely to remove this issue. Despite lacking features like the 9’ methoxy ring (+)-JQ1 also binds the MT1 receptor. Additionally, melatonin and other MT1 agonists contain a Kac mimic, suggesting the triazolo moiety of 9-ME-1 and (+)-JQ1 is key to MT1 binding; but this cannot be removed from 9-ME-1 without compromising BD-binding.



Scheme 3.3. 9-ME-1 & MT1 agonists.

Green denotes Kac mimics. Red denotes other perceived structural similarities between 9-ME-1 and MT1 agonists.

3.12 – Conclusions

A small library of novel bumped compounds were designed and synthesised, combining a number of modifications to three distinct sites of the scaffold. Through a 3-step compound evaluation workflow, alongside extensive structural studies, a wealth of SAR data was collected illustrating the workings of the bump-&-hole system. The chemistry of the bump-&-hole system was then further improved through the separation of the active *2R,3S* and inactive *2S,3R* enantiomers, which further boosted selectivity for L/V bromodomains. From this work, 9-ME-1 was highlighted as the ideal bumped compound, with >200-fold selectivity, high potency and excellent DMPK characteristics. The BET bump-&-hole system, in addition to possessing a new and more functional mutation, has now also been bolstered by the development of more selective and better-characterised bumped compounds.

Chapter 4
Validation & Implementation

4.1 – Introduction:

With the development of the more conservative and functional L/V mutation (chapter 2), and the development of the highly L/V-selective bumped inhibitor 9-ME-1 (chapter 3), it was next attempted to validate the bump-&-hole system in cells and confirm its ability to provide BD-specific inhibition. It was also necessary to confirm that our bumped compounds would not affect the phenotype of cells not possessing any L/V-mutant BET proteins. Once the system was validated it could then be used to investigate the BET proteins and answer biological questions regarding their cellular function.

4.2 – Cellular Validation of Bumped Compounds

The main technique by which we verified that our bumped compounds were active & selective in cells is the FRAP assay, previously used to monitor the impact of the L/V mutation on chromatin binding. In this assay as our bumped compounds bind to L/V bromodomains the mutant BET proteins will be displaced from chromatin, dramatically reducing the measured t½. Following on from experiments in [159] some FRAP experiments were carried out by Lars van Beurden, showing that several compounds were active and selective for BRD4 L/V constructs (figure S4.1). The assays were hampered however, by the low t½ values recorded for uninhibited BET proteins (~1s) leading to a ~2-fold assay window. This small window will obscure small effects, like the impact of L/V mutations or weak off-target WT binding from bumped compounds. This small assay window is presumably principally caused by the fact that our GFP-BET proteins are overexpressed, meaning that the majority of proteins at any one point are not actually bound to chromatin. Although several ideas were suggested this issue was tackled by altering the acetylation state of the chromatin. By treating with an HDAC inhibitor, SAHA (suberoylanilide hydroxamic acid, Vorinostat) (scheme 4.1), the acetylation state of the chromatin was increased, reducing the proportion of GFP-BET protein that was not bound to chromatin [164]. Dose-response experiments showed an assay window of ~6-fold at SAHA concentrations of 1 µM or higher, with 2 µM taken as a concentration that would reliably generate a large-enough window (figure 4.1).

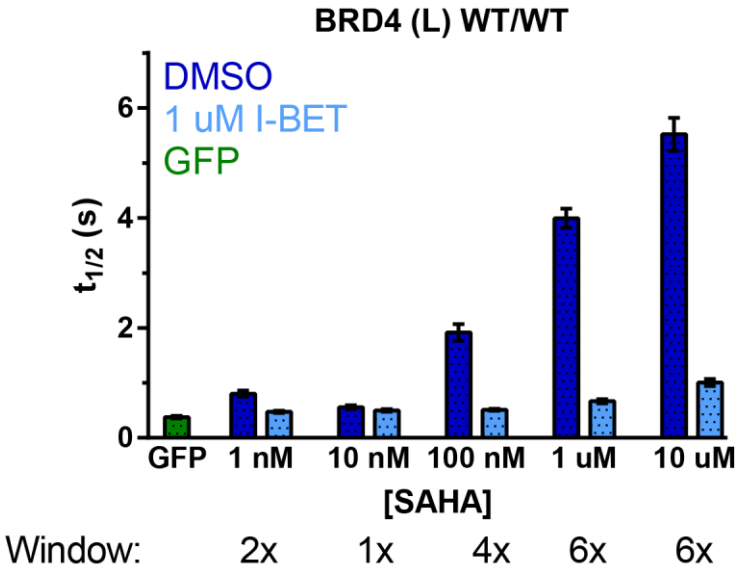


Figure 4.1. FRAP assay optimisation.

Recovery times of GFP-labelled WT BRD4 (L) or GFP alone (green) following 0.5 s laser bleach event, at varying concentrations of SAHA, within nuclei of U2OS cells treated with DMSO (dark blue) or 1 µM I-BET762 (light blue). Each bar is mean ± SE of ~20 cells.

With the FRAP assay now optimised a selection of compounds were tested against WT and LV/LV BRD4. This experiment showed that our bumped compounds were able to enter our U2OS cells, cross into the nucleus, bind BET bromodomains in an L/V-selective fashion and displace L/V proteins from chromatin (figure 4.2A). Several compounds were able to displace L/V BRD4 to a similar extent as (+)JQ1, while others had no impact on WT BRD4, in line with what is observed with the non-binding control (-)JQ1. Additionally, 9-ME-1 (our preferred compound) was tested at a range of concentrations against WT & L/V BRD4, giving a cellular IC_{50} of ~50 nM (figure 4.2B-C). This showed that, with some tweaking of dosage, it is possible to almost completely inhibit L/V BET proteins, while leaving their WT counterparts relatively untouched. As a result of this experiment 9-ME-1 was used at 200 nM for further FRAP experiments to ensure >50% of the target was displaced from chromatin while keeping a minimal impact on the WT protein. Finally, the inactivity of the *2S,3R* enantiomer was again shown, now in a cellular context (figure 4.2D).

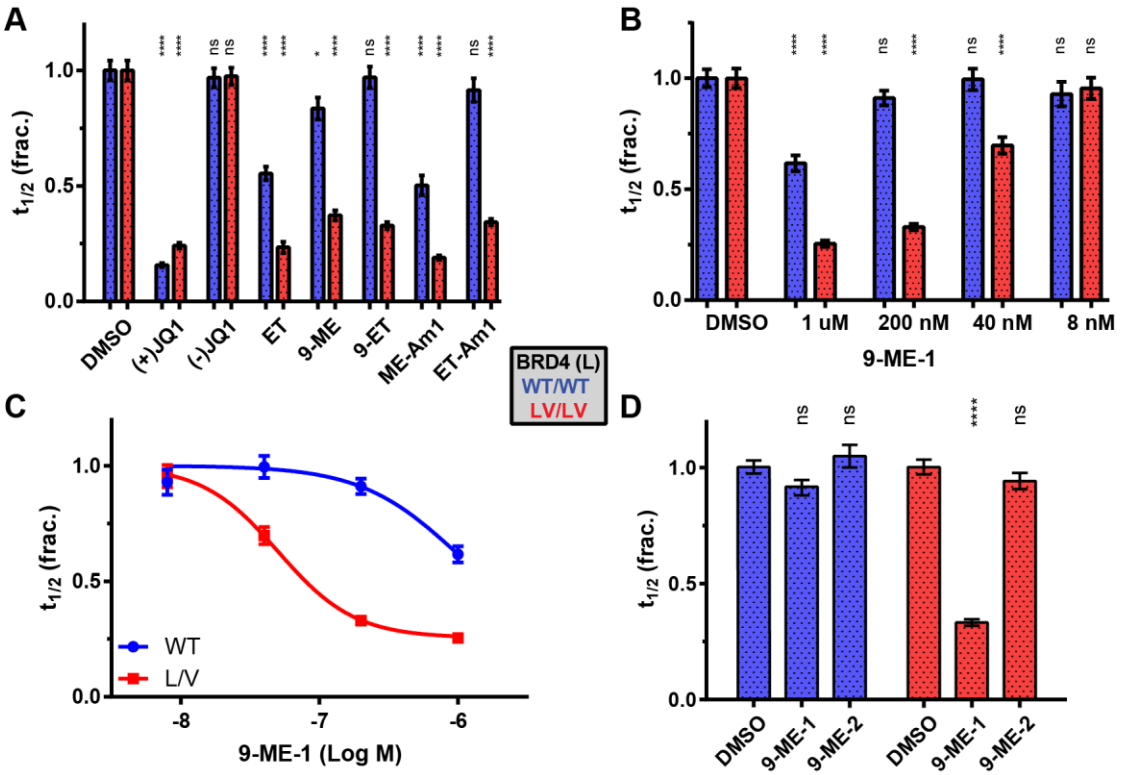


Figure 4.2. Validation of bumped compounds with FRAP.

A) Recovery times of GFP-labelled BRD4 (L) constructs following 0.5 s laser bleach event, at 2μM SAHA and 0.03% DMSO, treated with 1μM bumped compound. Each bar is mean ± SE of ~40 U2OS cells tested over two separate experiments. Statistical significance (compared to each construct's DMSO treatment) determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$. B) As for A, but cells treated with range of 9-ME-1 doses. C) Data from B, with dose-response curve fitted. D) As for A, but final compound concentrations of 200 nM.

4.3 – Biological Question – Chromatin Binding

With the bump-&-hole system validated in cells we were now able to utilise it to answer a biological question. By testing the impact of 9-ME-1 on each individual BET BD their importance to chromatin binding can be determined. BET BDs necessary for chromatin recognition will cause large drops in t½ values when mutated and inhibited by 9-ME-1, while BDs less important to chromatin binding will produce smaller changes in recovery rates. To this end 9-ME-1 was tested, at 200 nM, against each WT and L/V GFP-BET construct. Mean raw and normalised t½ values are listed in table S4.1.

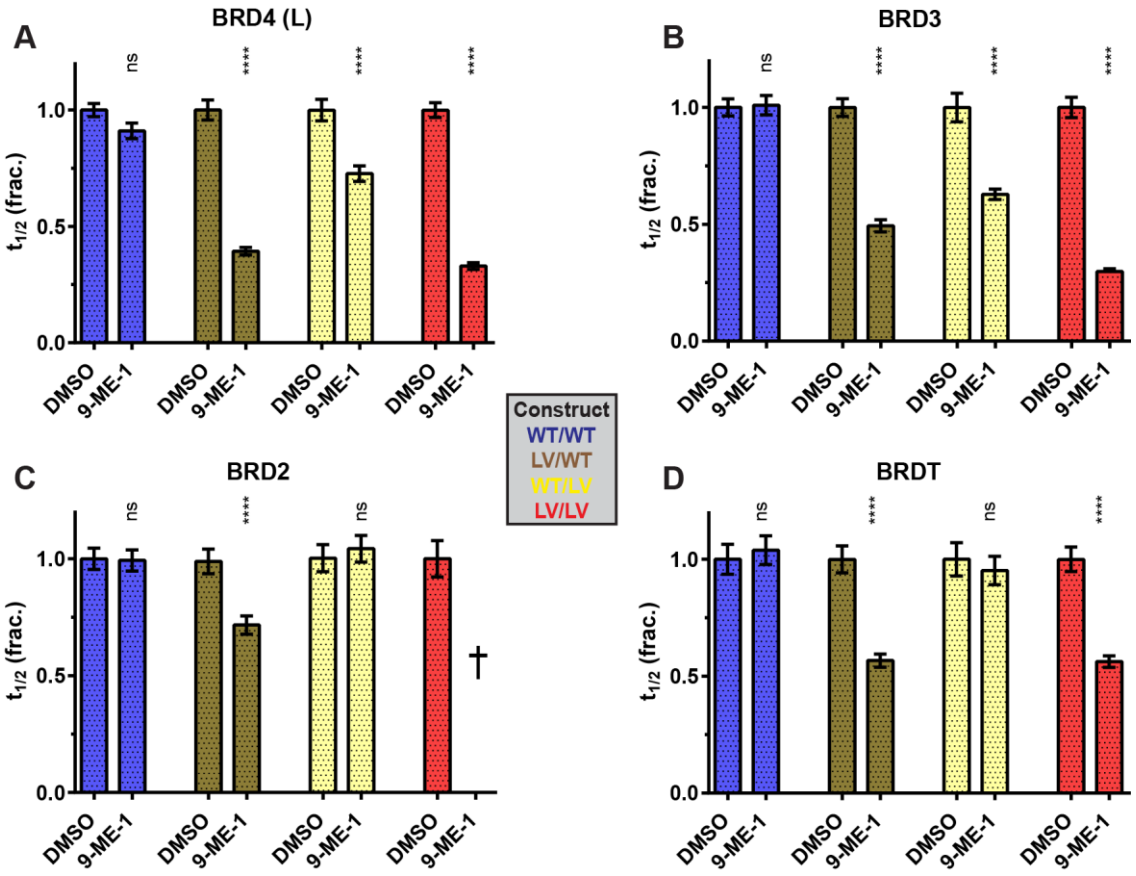


Figure 4.3. Importance of individual BET BDs to chromatin binding.

Recovery times of GFP-labelled BET protein constructs following 0.5 s laser bleach event, at 2μM SAHA and 0.03% DMSO, treated with 200 nM 9-ME-1. Each bar is mean ± SE of ~40 U2OS cells tested over two separate experiments. Statistical significance (compared to each construct's DMSO treatment) determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$. † denotes aggregation of construct.

One immediately visible trend between BET proteins is that targeting (mutating and inhibiting) the BD1 has a larger impact than targeting the BD2 (figure 4.3), showing that the BD1s are more important to chromatin binding. Additionally, the degree to which BD1 inhibition has a larger impact varies between BET proteins. BRDT shows no response to BD2 blockade, while targeting the BD1 alone has the same impact as

targeting the LV/LV double-mutant (figure 4.3D). BRD4 shows some response to BD2 inhibition, but it is significantly smaller than the BD1, and BD1 inhibition gives a response almost as strong as the double-mutant (figure 4.3A). BRD3, meanwhile, shows the strongest BD2 response, which is almost as strong as the BD1, targeting both BDs generates further displacement (figure 4.3B). BRD2 was an interesting case, with no BD2 response and a small BD1 response, meanwhile treating the double-mutant with 9-ME-1 caused aggregation of the GFP-BRD2, preventing the measurement of recovery times (figure 4.3C).

The aggregation of BRD2 was unexpected, and not seen in any other BET proteins even with higher compound doses. It is also unusual as targeting the BD2 had no impact on recovery times (figure 4.3C), which would imply that inhibition of LV/WT and LV/LV would generate the same results. This aggregation could be quantified by taking a line across a cell, plotting the brightness along this line, plotting a linear line-of-best-fit through said plot and then recording the quality of said fit (figure 4.4). Further experiments showed that this aggregation is caused by strong inhibition of both BRD2 BDs, as 200 nM (+)-JQ1 causes aggregation of WT GFP-BRD2 (which also shows this aggregation is not the result of the L/V mutation), while 9-ME-1 is only able to induce this effect when targeting the LV/LV double-mutant. Reducing the 9-ME-1 dose to 40 nM produced less severe aggregation, and t½ could again be measured. The FRAP assay was also repeated in HEK293T cells, to investigate whether the observed aggregation effect might be a cell-line-specific phenomenon. Unfortunately GFP-BRD2 also aggregated in these cells when inhibited, although to a slightly reduced extent (figure 4.5). Additionally, the assay window was reduced and the semi-adherent nature of HEK293T cells complicated the practical elements of the assay. The faster growth of HEK293T (compared to U2OS), combined with their vulnerability to being accidentally dislodged from the microscope dish, resulted in extreme variance in transfected cell-density. Many areas of the microscope dish would show no transfected cells, while transfected cells were often found so densely packed together than FRAP experiments were not possible.

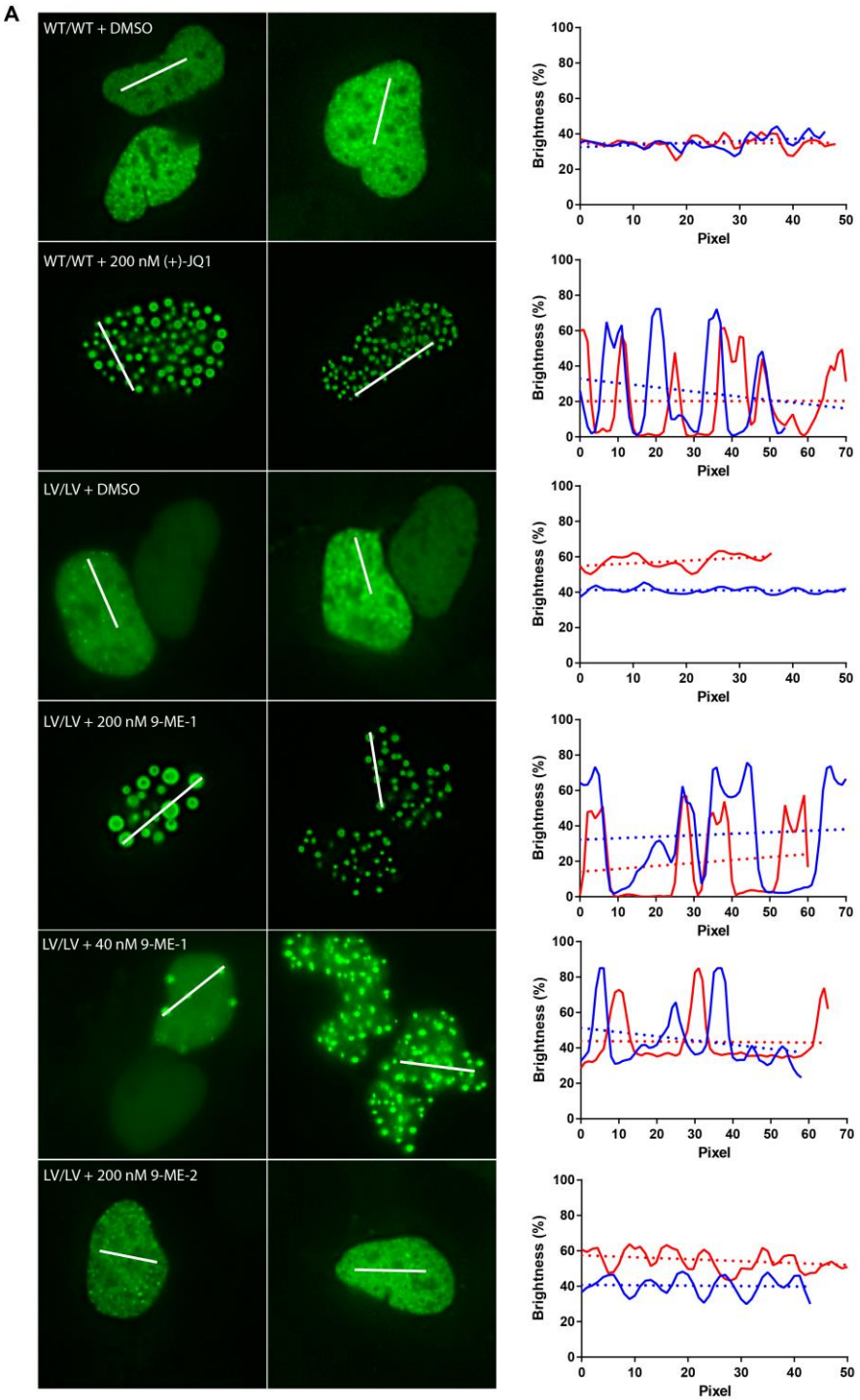

A
WT/WT + DMSO
WT/WT + 200 nM (+)-JQ1
LV/LV + DMSO
LV/LV + 200 nM 9-ME-1
LV/LV + 40 nM 9-ME-1
LV/LV + 200 nM 9-ME-2
Brightness (%)
Pixel
B
Linear Sy.x
DMSO
(+)-JQ1
9-ME-1
9-ME-2
Compound (nM):
GFP-BRD2: WT/WT LV/WT WT/LV LV/LV

Figure 4.4. Inhibition-induced aggregation of GFP-BRD2.

A) Nuclei of U2OS cells transfected with GFP-BRD2 constructs and treated with 2 µM SAHA, 0.03% DMSO and test compound. Line drawn cross cells and brightness profile plotted, and linear line of best fit added (dotted lines). B) Quality of fit of linear line recorded as SD of residuals (Sy.x) and used to quantify aggregation, with larger values representing greater aggregation. Each bar is mean and SE of >25 U2OS cells from two separate experiments. Statistical significance (compared to each construct's DMSO treatment) determined with two-tailed t tests: ns P>0.05; * P≤0.05; ** P ≤0.01; *** P≤0.001; **** P<0.0001.

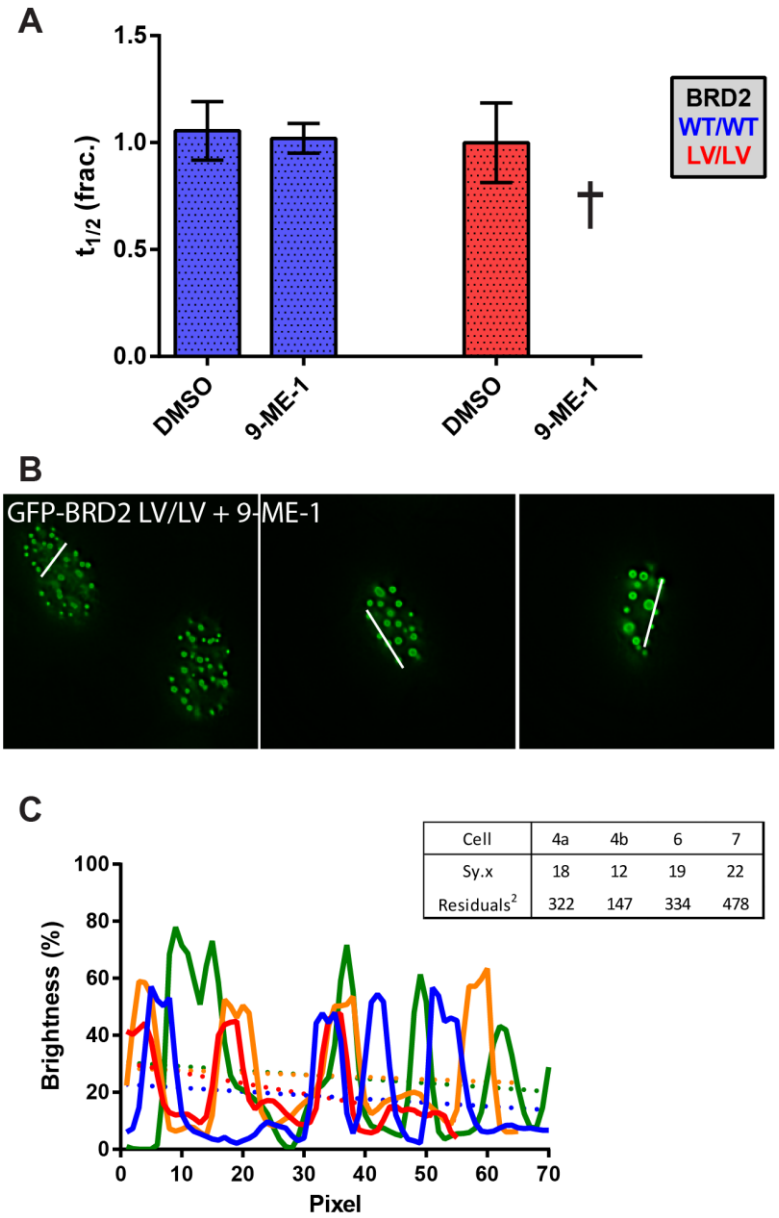


Figure 4.5. GFP-BRD2 aggregation in HEK293T cells.

A) Recovery times of GFP-labelled BRD2 constructs in HEK293T cell FRAP assay, following 0.5 s laser bleach event, at 2 µM SAHA and 0.03% DMSO. Each bar is mean and SE of ~30 HEK293T cells. † denotes aggregation of construct. B) Nuclei of transfected HEK293T cells, treated with 2 µM SAHA, 0.03% DMSO and 200 nM 9-ME-1. C) Brightness profile of nuclei plotted. Linear line of best fit (dotted line) plotted, and quality of fit tabulated.

BRD2 was re-tested in the U2OS FRAP assay, but with a reduced dose of 40 nM 9-ME-1, with the hope that double-mutant t½s could be measured and compared to those of the BD1 and BD2 mutants. This experiment was successful, and produced a set of results similar to that displayed by BRDT – no response from BD2, BD1 inhibition matching that of the double-mutant (figure 4.6).

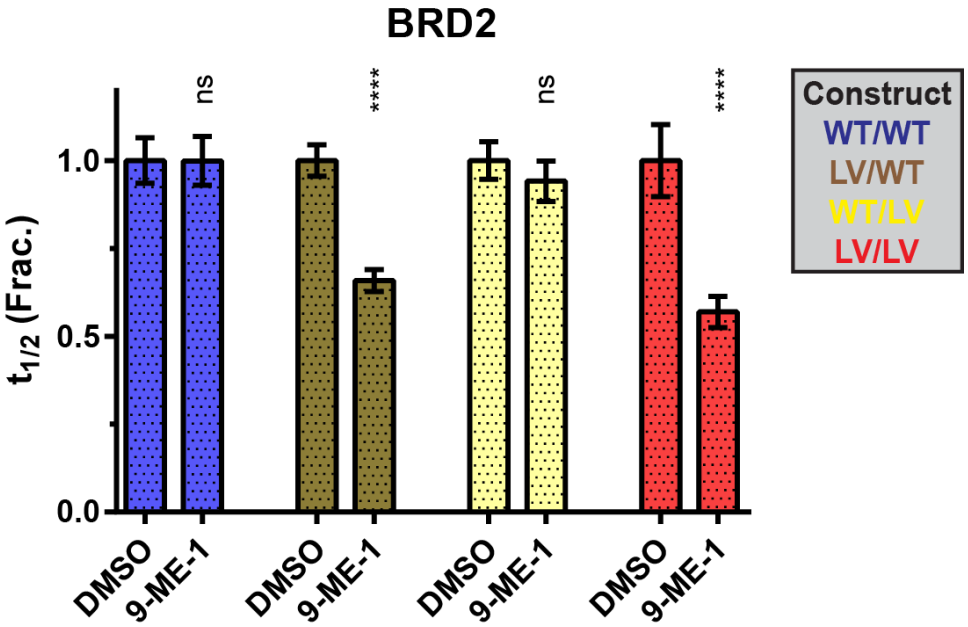


Figure 4.6. Importance of individual BRD2 BDs to chromatin binding.

Recovery times of GFP-labelled BRD2 constructs following 0.5s laser bleach event, at 2μM SAHA and 0.03% DMSO, treated with 40 nM 9-ME-1. Each bar is mean ± SE of ~40 U2OS cells tested over two separate experiments. Statistical significance (compared to each construct's DMSO treatment) determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$.

All of these results taken together show the BD1 to be the dominant domain for acetylated chromatin binding (table 4.1). Both BRD2 and BRDT appear to bind chromatin solely through their BD1s, and their BD2s have no observable significance. BRD4 BD2 does play a small role in chromatin binding, but BRD4 BD1 is both necessary and sufficient for chromatin binding as of itself. Chromatin binding of BRD3 meanwhile is almost balanced, with the BD2 playing only a slightly smaller role than the BD1. Our results for BRD4 and BRDT match those previously obtained in the literature using alternate techniques. Pulldown of BRD4 constructs mixed with custom nucleosomes bearing specific Kac combinations showed BRD4 BD2 to play a minor but non-zero role in chromatin binding [186], while ITC titrations and NMR experiments showed BRDT BD2 to be incapable of binding acetylated nucleosomes [34].

	9-ME-1 inhibition (%)			
L/V BD	WT	BD1	BD2	Both
BRD4	13	90	41	100
BRD3	-1	72	53	100
BRD2	0	79	14	100
BRDT	-9	99	11	100

Table 4.1. Importance of each BET BD to chromatin binding

Percentage inhibition of chromatin binding by 9-ME-1 treatment (200 nM for BRD4, BRD3 and BRDT; 40 nM for BRD2), calculated with regard to the impact of 9-ME-1 on each BET proteins LV/LV construct. Note: LV/LV construct may retain some chromatin binding even after 9-ME-1 treatment. Colour scale: 0% = white, 100% = dark green. All inhibition values below 15% were not statistically significant (p>0.05).

4.4 – Biological Question - Transcription

In addition to measuring BD1/BD2 importance to chromatin binding through FRAP, the previously discussed NF-κB/luciferase assay (chapter 2.14) was applied to measure the importance of each BD to inducing the transcription of NF-κB target genes (figure 4.7). In the NF-κB/luciferase assay HEK293T cells are transfected with a plasmid encoding a luciferase reporter enzyme controlled by NF-κB signalling through 5 NF-κB response elements. Co-transfection and transient overexpression of GF-BET constructs increases the expression of said luciferase.

For BRD4, transcription could be severely inhibited by targeting either the BD1 or the BD2 (figure 4.7A). A similar trend was shown by BRD3 (figure 4.7B). BRD2 showed a significantly greater impact from BD1 inhibition, although targeting the BD2 did reduce the expression of the luciferase (figure 4.7C). Finally, BRDT showed a strong response to BD1 and double-mutant inhibition, but again presented a smaller albeit measurable response with the BD2 (figure 4.7D). In some cases targeting a single BD has a greater impact on transcription than targeting both with the LV/LV double-mutant. This could indicate that inhibiting either the BD1 or BD2 is sufficient for reducing transcription, while mutating both BDs may reduce the effective inhibitor dose (by halving the compound : L/V BD ratio).

In contrast with the FRAP assay on chromatin binding, these results show the BD1 and BD2 to be equally important to transcription for BRD4 and BRD3. In contrast, in the case of BRD2 and BRDT, the BD1 is still pre-eminent but the BD2 does now play an observable role (table 4.2).

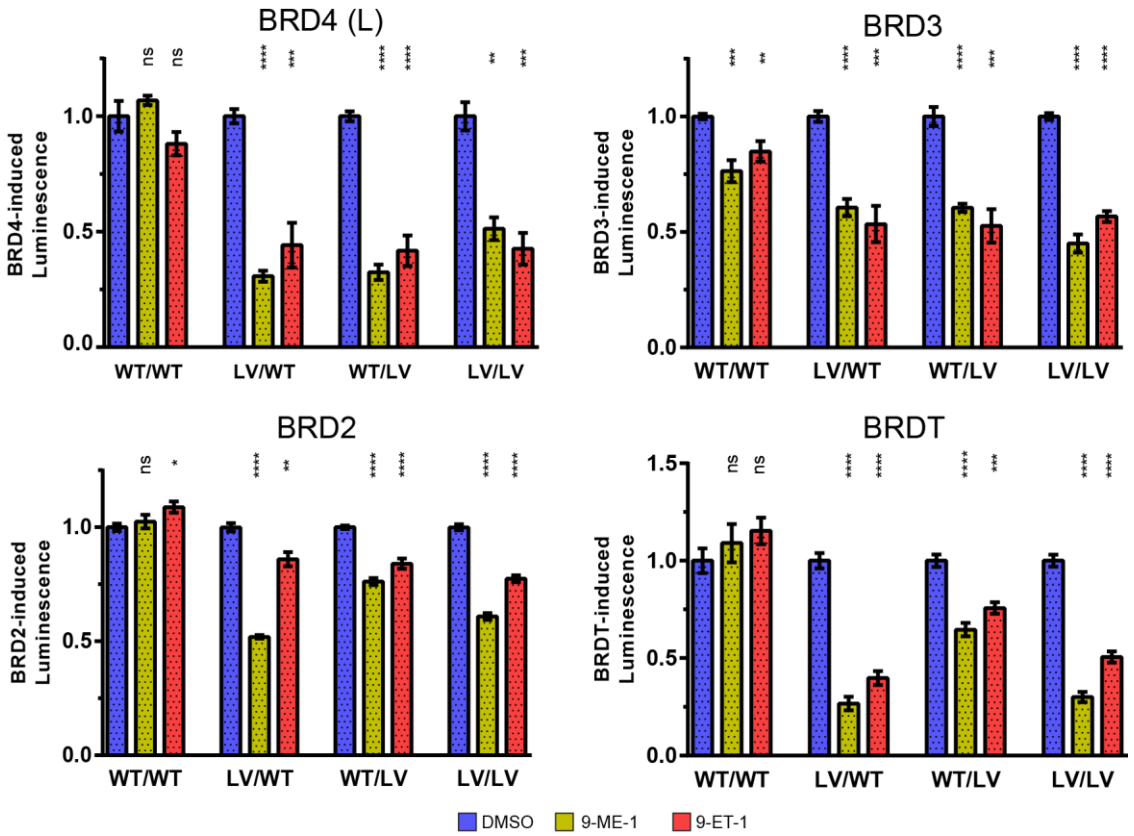


Figure 4.7. Importance of individual BET BDs to the transcription of NF-κB signalling genes.

Normalised luminescence of lysate from HEK293T cells transfected with GFP-labelled BET constructs and a NF-κB-RE/*luc2P* reporter plasmid, and treated with 1 µM compound and/or 0.01% DMSO. Basal luciferase expression (with GFP plasmid) subtracted as background. Signal normalised to each construct's DMSO control. Results are mean and standard error of six technical replicates spread over two independent experiments. Statistical significance determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$.

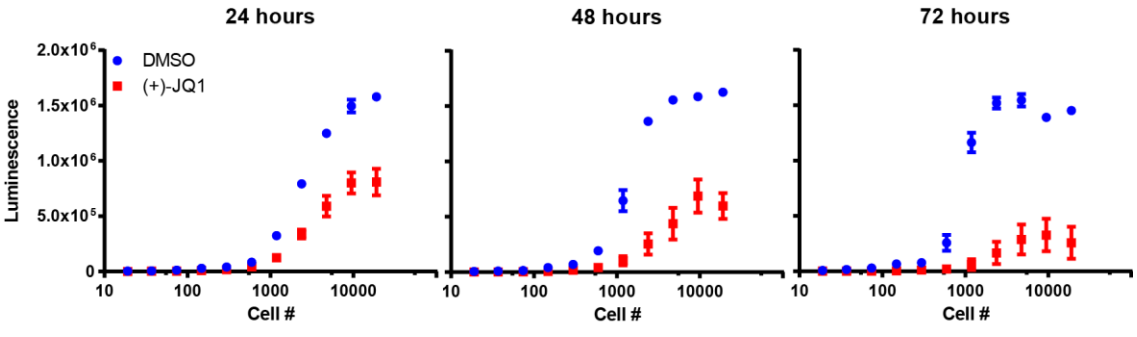
	9-ME-1 inhibition (%)			
L/V BD	WT	BD1	BD2	Both
BRD4	-14	142	139	100
BRD3	43	72	72	100
BRD2	-6	123	61	100
BRDT	-13	105	51	100

Table 4.2. Importance of each BET BD to NF-κB-mediated transcription

Percentage inhibition of NF-κB-mediated transcription binding by 9-ME-1 treatment (200 nM for BRD4, BRD3 and BRDT; 40 nM for BRD2), calculated with regard to the impact of 9-ME-1 on each BET proteins LV/LV construct. Note: LV/LV construct may retain some activity even after 9-ME-1 treatment. Colour scale: 0% = white, 100% = dark green. Note: all inhibition values below 15% were not statistically significant ($p>0.05$). 9-ET-1 results omitted for clarity.

4.5 – Effects of 9-ME-1 on WT Cell-viability:

We have shown through both biochemical/biophysical and cellular assays that our bumped compounds have little activity against WT BET bromodomains, and our off-target screening has only uncovered one notable off-target (MT1). As such, we hypothesised that our bumped compounds should show little to no impact on the phenotype of WT cells, as these cells will lack any L/V-mutated BET proteins. To test this hypothesis, the viability of BET-sensitive cell-lines was measured at increasing concentrations of 9-ME-1, using Promega's ATP-measuring CellTiter-Glo assay [187]. This assay had already been optimised for suspension cell-lines, specifically BET-sensitive AML-derived MV4-11 and HL-60 [91], by Dr Kwok-Ho Chan [188]. In order to test adherent cell-lines, for example BET-sensitive, LAC-derived A549 and H23 [94], this assay had to be re-optimised to work with much lower cell densities.



A549	24 hours		48 hours		72 hours	
Cell #	Assay Window	Z'	Assay Window	Z'	Assay Window	Z'
20000	2.0 ±0.1	0.56	2.7 ±0.1	0.66	5.7 ±0.3	0.64
10000	1.9 ±0.1	0.46	2.3 ±0.1	0.50	4.2 ±0.2	0.64
5000	2.1 ±0.1	0.60	3.5 ±0.2	0.61	5.4 ±0.2	0.62
2500	2.3 ±0.1	0.66	5.4 ±0.2	0.73	9.2 ±0.3	0.73
1250	2.6 ±0.1	0.46	6.3 ±2.2	0.38	19.4 ±0.4	0.69
625	2.0 ±0.1	-0.23	4.8 ±0.2	0.51	12.3 ±0.3	0.12
313	2.4 ±0.1	0.00	4.0 ±1.1	0.64	7.6 ±0.3	0.21
156	3.2 ±0.1	0.28	6.5 ±3.1	0.56	17.3 ±0.2	0.49
78	4.1 ±0.1	0.61	4.8 ±0.2	0.14	17.2 ±0.2	0.42
39	3.5 ±0.1	0.20	7.1 ±0.1	0.22	15.0 ±0.2	-0.21
20	3.2 ±0.1	0.32	6.8 ±0.2	0.25	9.4 ±0.2	0.04
10	1.1 ±0.1	-2.93	1.3 ±0.1	-0.21	1.3 ±0.1	-0.42

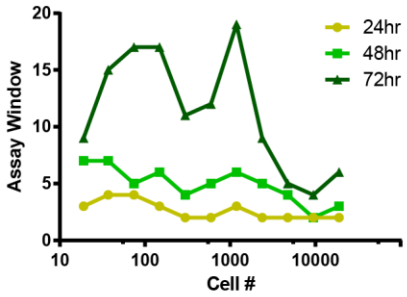


Figure 4.8. Cell viability assay LAC optimisation.

Luminescence of A549 cells mixed with Promega CellTiter-Glo reagent, after being seeded over a range of densities, treated with 0.05% DMSO and/or 50 µM (+)-JQ1 and incubated for 24, 48 or 72 hours. Assay quality (assay window and Z') calculated for each condition.

To discover the optimal assay conditions for adherent cell-lines, A549 cells were seeded in 384-well assay plates over a range of cell-densities and treated with either DMSO or 50 µM (+)-JQ1 for 24, 48 or 72 hours. Increasing the cell number showed a clear increase in signal, saturating at 1000 – 10,000 cells depending on incubation time. Longer incubation times also in general produced larger assay windows and Z' values,

with slightly lower optimal cell numbers (figure 4.8). These experiments revealed the optimal conditions for this assay to be ~2000 cells with a 72 hour incubation time, giving a 10-20-fold assay window and Z' values of ~0.7 [171].

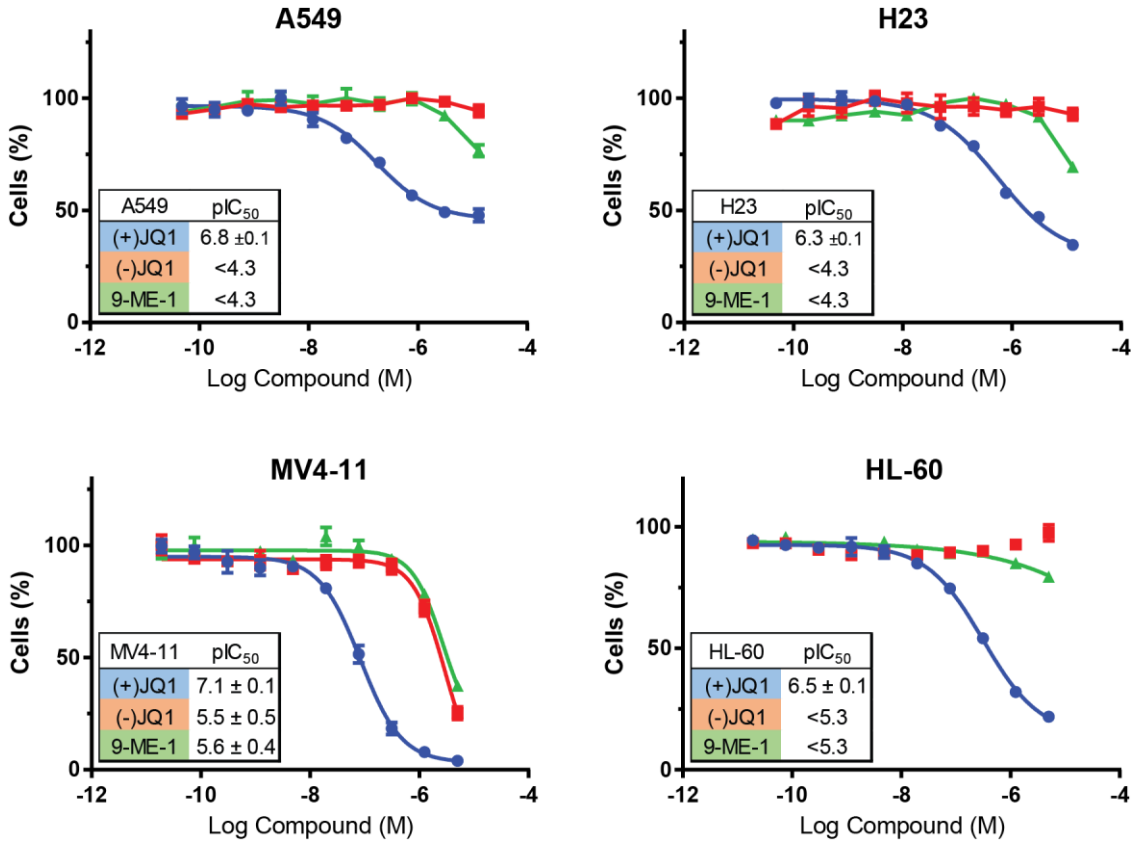


Figure 4.9. Impact of 9-ME-1 on WT cell-viability

Viability of BET-sensitive cells: adherent, LAC-derived A549 & H23; suspension, AML-derived MV4-11 & HL-60. Cell viability determined via Promega CellTiter-Glo measurement of ATP-levels.

Titrating 9-ME-1 against BET-sensitive cells showed it to be inactive, when compared to active and inactive JQ1 compounds (figure 4.9). While (+)-JQ1 shows a strong reduction in cell viability and IC_{50} values of 600 to 100 nM 9-ME-1 acts more like the inactive (-)-JQ1, and shows no impact below 1 µM. This clearly shows 9-ME-1 to be A) relatively inactive against WT BET proteins in cells, and B) to not have any biologically significant off-targets that might impact the viability of these cells.

4.6 – Effects of 9-ME-1 on WT Cell Colony formation:

As an alternative to the cell-viability assay discussed previously a colony-formation assay was developed. These assays involve seeding small numbers of cells, which are treated with compound for long timescales (~1 week) before being fixed and stained. The readout of this assay is the number of colonies left after staining, and the degree to which small-molecules reduce this. These such assays require adherent cell-lines, and was hence carried out with A549 cells. Initial testing, wherein a range of cell-densities was incubated for 7 days and fixed/stained with 0.01% crystal-violet in 25% MeOH for 1 hour did leave visible cells, but colony counting was not possible and washing the cells with H_2O displaced large amounts of cells (figure 4.10A). This fixing problem was solved by separating the fixing and staining processes, and fixing the cells with 100% MeOH at 4°C, which showed no displacement despite multiple H_2O washes (figure 4.10B).

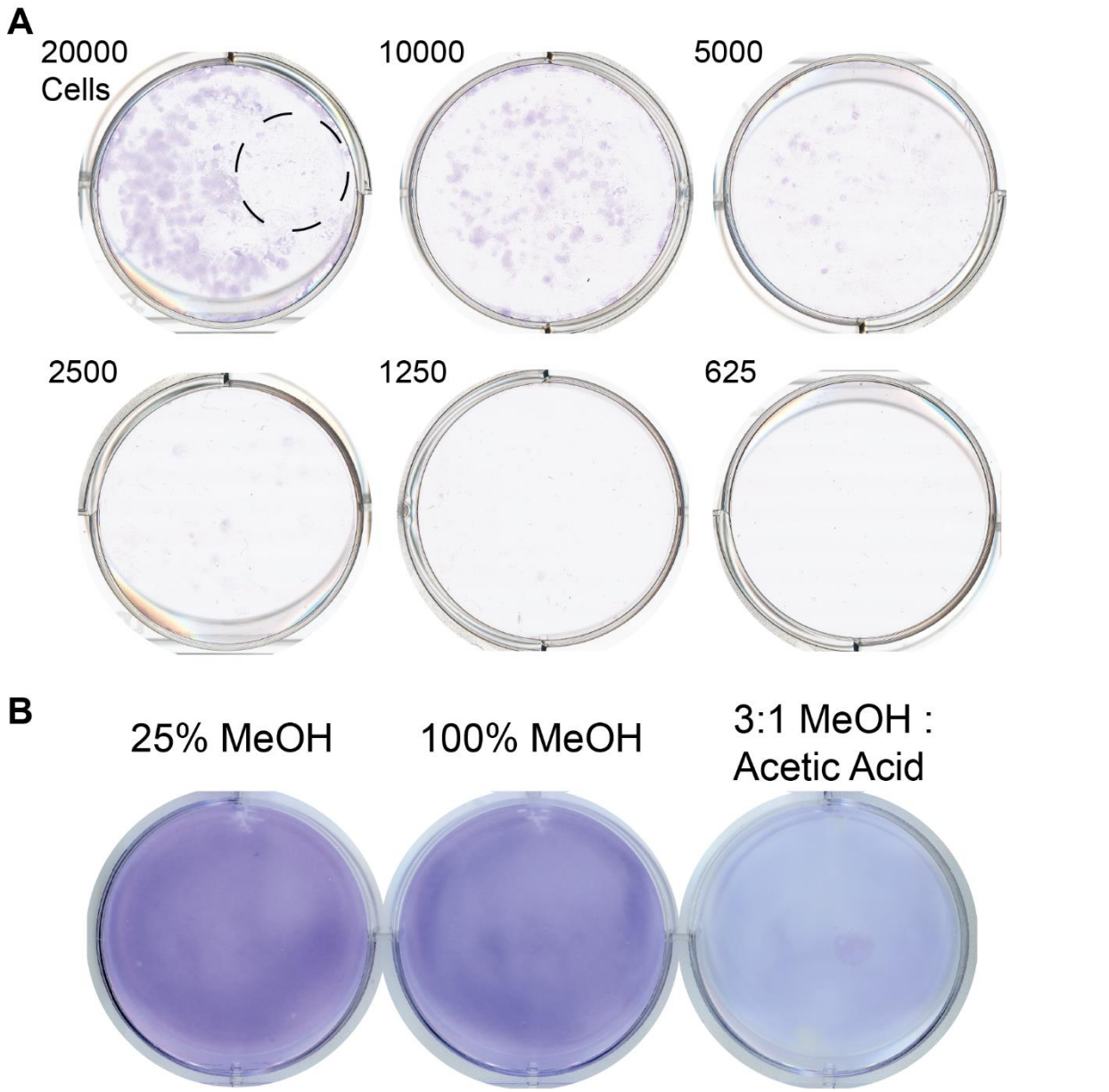


Figure 4.10. A549 colony-formation assay optimisation.

A) 6-well plates seeded with A549 cells (number indicated) and grown for 1 week. Cells then fixed at room temperature with 25% MeOH and stained with 0.01% crystal-violet. Dotted circle shows area where cells removed during attempted second wash with water. B) 6-well plates seeded with A549 cells and grown to 100% confluence before being fixed at 4°C with indicated cell-fixing solution, and stained with 0.1% crystal violet.

Following this optimisation, the effect of 9-ME-1 on WT A549 cells was measured (figure 4.11) (figure S4.2). While the presence of 1 µM (+)-JQ1 killed essentially all cells, 1 µM 9-ME-1 reduced colony numbers by only ~10% on average – showing similar results to the cell-viability experiments and confirming that 9-ME-1 does not inhibit WT BET bromodomains, or other unknown off-targets, at the doses recommended for use.

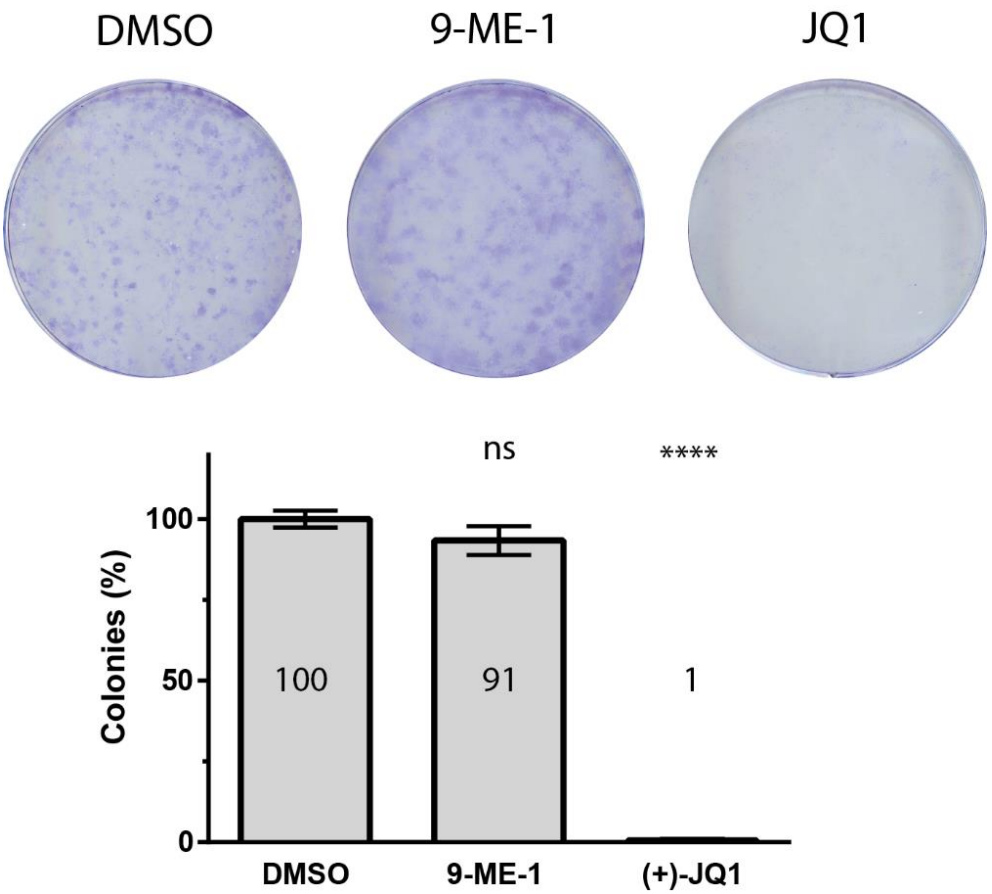


Figure 4.11. Effects of 9-ME-1 on WT A549 colony formation.

6-well plates seeded with 2000 A549 cells, treated with 1 µM compound and/or 0.01% DMSO and incubated for 8 days before being fixed with 100% MeOH and stained with 0.1% crystal-violet. Assay carried out with two biological replicates, each with two technical replicates. Statistical significance (compared to DMSO treatment) determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$.

A dose-response experiment was then performed, to get an idea of what sort of 9-ME-1 dose would have a strong impact on any L/V-mutated A549 cells while avoiding any significant WT effects (figure S4.3). (+)-JQ1 was shown to be fully-lethal at 1 µM, partially lethal at 100 nM and have no impact at lower doses. 9-ME-1, meanwhile, was shown to have no impact until dosed at 10 µM. Although it is possible that (+)-JQ1 might be more potent against WT cells than 9-ME-1 would be against L/V cells, together this data suggests that a dose between 100 nM and 1 µM would likely give full L/V lethality alongside little impact on WT.

In order to better perform colony-formation dose-response experiments the decision was taken to move into a 96-well format and to bypass the counting of colonies. A549 were seeded in 96-well plates, treated with a gradient of compound doses and grown until DMSO-treated wells were fully confluent before fixing and

staining. The potent inhibitor and PROTAC, (+)-JQ1 and MZ1, here used as positive controls, showed inhibition of colony-formation from 400 nM to 1 µM, while (-)JQ1 showed high confluence growth even at 10 µM. As expected, 9-ME-1 showed no inhibition at 3 µM, but with a noticeable reduction in signal (though not visible confluence) at 10 µM (figure 4.13).

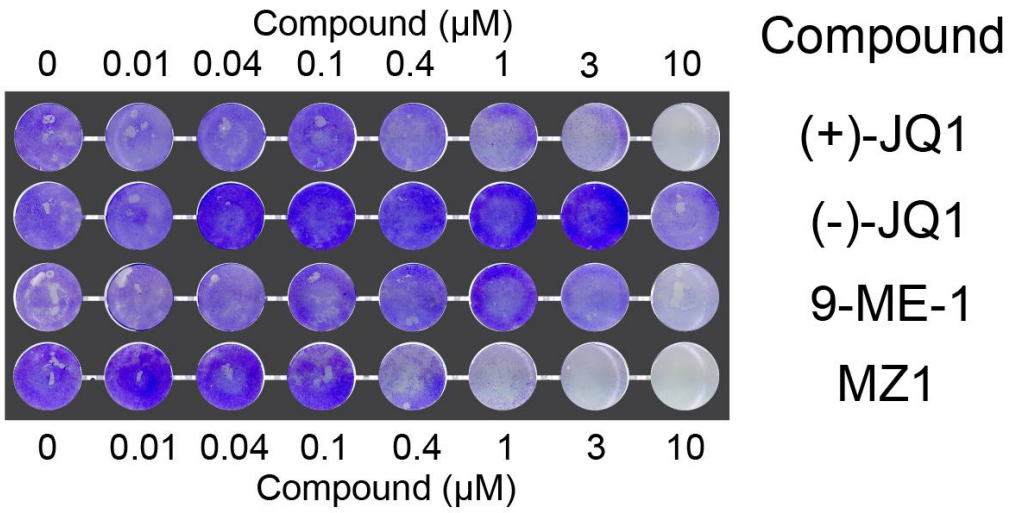


Figure 4.13. 96-well A549 colony-formation assay

96-well plate seeded with A549 cells, treated with compound and/or 1% DMSO and incubated for 1 week before being fixed with 100% MeOH and stained with 0.1% crystal-violet.

The dose-response, 96-well colony-formation assay developed above is a valuable complement to the CellTiter-Glo assay, for assessing cell viability and the cytotoxicity of compounds. The colony-formation assay is not easily quantifiable, requires adherent cells and may be difficult to transfer to a 384-well plate. Additionally, the quality of any specific assay run will be strongly affected by the growth-rate of the cells and exactly how long they are incubated with compounds. On the other hand, this assay is easier to run for extended periods of time, allowing one to study cell-viability over short time periods (2-3 days) with CellTiter-Glo and longer periods (1+ week) with the colony-formation assay. Additionally, as the colony-formation assay works by direct visualisation of cells it is less vulnerable to assay interference and false negatives/positives than the ATP/luminescence measuring CellTiter-Glo assay. This assay has since been used in other projects, such as the assessment of the cellular growth inhibition activity of PROTAC compounds (figure S4.4) [189].

4.7 – Attempts to Answer Other Biological Questions:

The previously described FRAP and luciferase assays were very successful in allowing us to probe BET bromodomain function with the bump-&-hole system. While performing these experiments, however, several similar experiments were tried and failed. These assays attempted to look into the importance of BET BD1s and BD2s to gene expression and overall cell-viability, relying on GFP-BET overexpression and, in some cases, siRNA of endogenous proteins. These experiments were inconclusive and thus highlight the need to, ultimately, produce model cell-lines with the L/V mutation incorporated into endogenous BET genes that will be expressed and regulated appropriately, while eliminating the need to knock-out/down WT genes. Although this work failed to answer biological questions a great deal was learnt from these experiments that should inform future experiments, hence their description within this thesis.

The BRD4 / NF-κB luciferase assay, as originally developed by Dr Kwok-Ho Chan, was quite different from the final version. It was performed in A549 cells (more biologically relevant, and in line with literature [95]) instead of HEK293T, used siRNA to suppress endogenous WT BRD4 expression and featured TNF-α induction, looking at the differing increase in luciferase expression between different constructs (figure 4.14).

The transient overexpression of BRD4 was generally unable to rescue the luciferase signal following siRNA knock-down of the endogenous protein, but an increase in signal was observed by Dr Chan in one experiment where the siRNA was also shown to have had no impact on the endogenous protein (for unknown reasons) (figure 4.14A). This suggested that, going forward, siRNA knock-down of endogenous protein was not necessary nor recommended.

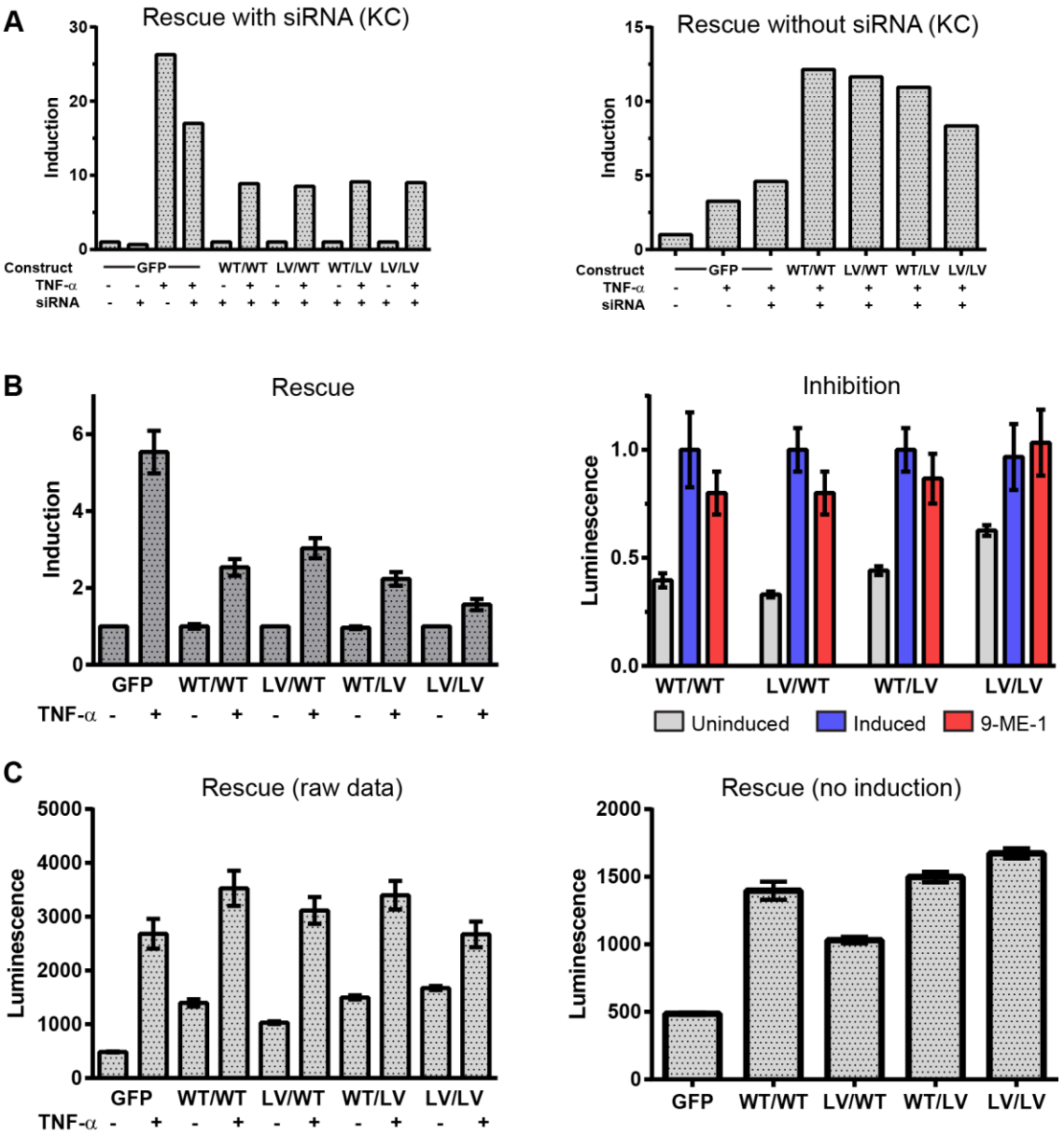


Figure 4.14. Development of luciferase assay.

A) Luciferase experiments run by Dr Kwok-Ho Chan, looking for an increase in luciferase signal upon GFP-BRD4 (L) overexpression. Experiment run with siRNA (targeting 5' UTR and hence not effecting exogenous protein) that was mildly effective shows a decrease in luciferase induction with BRD4 overexpression. Experiment run with 3' UTR targeting siRNA, that was not effective, shows a noticeable increase in induction with BRD4. B) Replicate experiment, where BRD4 overexpression again reduced luciferase induction, and where 9-ME-1 had a very weak and non-selective impact. C) Data from B, without induction normalisation, showing that without TNF-α induction BRD4 overexpression strongly increases luciferase expression.

Even following the removal of the siRNA step no clear 'rescue' was observed with BRD4, and no meaningful impact was seen with 9-ME-1 treatment (figure 4.14B). Overexpression of GFP alone consistently led to a much higher increase in luciferase expression upon TNF-α induction (figure 4.14b). Analysis of the uninduced samples revealed that this was not because the signal with TNF-α induction was higher for GFP compared to GFP-BRD4, but instead a result of the uninduced signal, being substantially lower for the GFP control than with the GFP-BRD4 constructs (figure 4.14C). This led to the decision to remove the TNF-α induction step, which allowed for the observation of strong increases in luciferase expression from GFP-BRD4 WT and L/V mutants, with the attendant L/V-selective decrease in signal from bumped compounds. As the removal of TNF-α led to a large decrease in signal it became difficult to work with A549 cells, which consistently produced lower luminescence than the HEK293T.

As we had previously utilised transient overexpression of mutant GFP-BET constructs to study chromatin binding and transcription it was hoped a similar assay could be developed to investigate the importance of BD1s/BD2s to the viability of BET-dependent cells. As such, BET-dependent A549 cells were transfected with GFP-BET constructs and their viability measured through the CellTiter-Glo assay (figure 4.15). The first experiment ran showed some promise: BRD4 overexpression did not increase the viability of the cells, but 9-ME-1 was shown to reduce their viability when paired with the LV/LV BRD4 double mutant, suggesting that both bromodomains need to be inhibited to effect this cell-line (figure 4.15A).

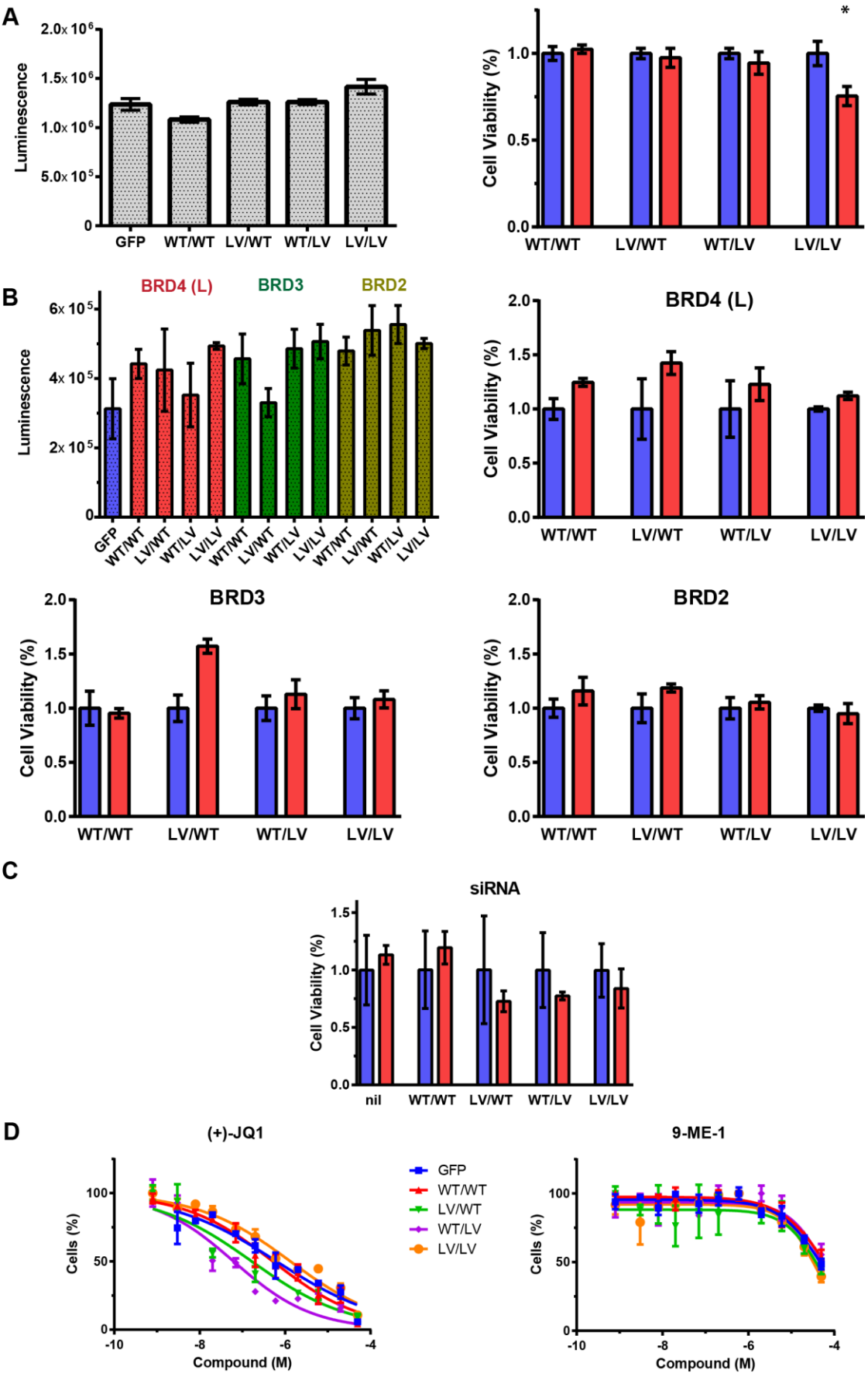

A
Luminescence
2.0x 10^6
1.5x 10^6
1.0x 10^6
5.0x 10^5
0
GFP
WT/WT
LV/WT
WT/LV
LV/LV
Cell Viability (%)
1.0
0.5
0.0
*
WT/WT
LV/WT
WT/LV
LV/LV
B
BRD4 (L)
BRD3
BRD2
Luminescence
6x 10^5
4x 10^5
2x 10^5
0
GFP
WT/WT
LV/WT
WT/LV
LV/LV
BRD4 (L)
Cell Viability (%)
2.0
1.5
1.0
0.5
0.0
BRD3
BRD2
C
siRNA
Cell Viability (%)
nil
D
(+)-JQ1
9-ME-1
Cells (%)
100
50
0
-10
-8
-6
-4
Compound (M)
GFP
WT/WT
LV/WT
WT/LV
LV/LV

Figure 4.15. Attempts to study cell-viability through BET overexpression.
A) CellTiter-Glo signal produced by A549 cells following GFP or GFP-BRD4 (L) overexpression and treatment with DMSO or 9-ME-1. B) Repeat of A, but including overexpression of GFP-BRD3 and GFP-BRD2 constructs. C) BRD4 (L) assay repeated again, but with addition of BRD4 UTR targeting siRNA (which does not affect exogenous protein). D) BRD4 (L) experiment repeated but with gradient of (+)-JQ1 and 9-ME-1 doses.

Unfortunately, this result could not be recreated in later experiments with BRD4 and other BET proteins, nor was any strong or consistent increase in cell-viability recorded as a result of BET overexpression (figure 4.15B). Attempts to knock-down the endogenous BRD4 through siRNA, using a protocol developed by Dr Kwok-Ho Chan, merely increased the experimental variance (figure 4.15C). Finally, a dose-response experiment was carried out, with cells transfected in a 6-well plate then re-seeded in 384-well assay plates, and no BRD4 construct was shown to alter the dose-response curve of 9-ME-1 or (+)-JQ1 (figure 4.15D).

4.8 – Discussion:

The FRAP assay was chosen to confirm L/V-selective bromodomain inhibition in cells as it was a technique that could be used to confirm mutant functionality (chapter 2), show target engagement (validate the system) and measure the importance of each BD to chromatin binding (utilise the system for a biological question). The FRAP assay is a powerful way to observe the displacement of full-length BET proteins from chromatin in cells, but it does suffer from low throughput and is not optimised for dose-response experiments. FRAP also requires the expression of exogenous GFP-labelled protein, though this is less of a drawback in this context, as testing L/V mutant BET proteins would require such exogenous expression anyway. One other issue is the need to treat cells with an HDAC inhibitor to ensure a usable assay window. This issue, however, is not FRAP-specific and would likely apply to any assay seeking to monitor displacement of overexpressed BET proteins from chromatin. The assay-window problem had previously been solved in the course of assay optimisation by previous members of the group while obtaining data for [159]. This was achieved by bleaching an area much larger than the area in which fluorescence was measured. Unfortunately changes in the hardware available prevented this option from being reused. Lowering

the expression of the GFP-BET protein may have increased the assay window, but may leave the transfected nuclei too dim for the FRAP assay to function. Finally, former lab member Lars van Beurden observed that running the assay at lower temperatures increased the assay window, but also greatly increased variance. Replacing the GFP tag with a brighter fluorophore (such as Clover and its derivatives [190]) may allow for further optimisation of the FRAP assay. A brighter tag could allow for FRAP assays to be carried out in cells expressing a much lower (and closer to endogenous) level of tagged BET protein. Reducing the levels of tagged protein will both improve the signal to noise by re-balancing the BET:ac-chromatin ratio to near-WT, and generate more realistic dose-response relationships.

Chromatin immunoprecipitation, with massively parallel DNA sequencing (ChIP-Seq) is another means of monitoring the interaction of BET proteins to chromatin, that could be employed in the future to complement the FRAP assay [191, 192]. An advantage of ChIP-Seq is the ability to monitor the recruitment of BET proteins to, and their displacement from, multiple specific gene loci or epigenetic marks. This additional data is not, however, required for this project as the focus was on the displacement of BET proteins from chromatin in general. Additional ChIP-Seq experiments could be worthwhile in the case of BRD4 and BRD3, as it would be interesting to dissect to what extent BD1 versus BD2 blockade might displace BET proteins from different gene loci / epigenetic marks.

One highly unexpected outcome of the FRAP work was the observation of GFP-BRD2 aggregation in response to inhibition. Such distributions of nuclear proteins are often observed in response to DNA damage and repair, however a recent study suggested that BRD2 is not involved in any such process [193]. Such aggregation could be due to the addition of the GFP-tag, issues with the specific BRD2 construct used or the high cellular concentration of the overexpressed protein. As mentioned in chapter 2, this aggregation of GFP-BRD2 may be connected to the unusually severe impact the L/V mutation has on GFP-BRD2 LV/LV's ability to bind chromatin. Aggregation of GFP-labelled bromodomains in response to inhibition has been observed previously, and developed into an assay to measure inhibitor potency [194]. Gacias et al., in their publication of the BD2-selective inhibitor Olinone, also observed this aggregation (or "punctate distribution") of GFP-BRD2 [128]. As this redistribution was observed with a

pan-selective BET inhibitor, but not the BD2-selective Olinone they put forth that the BD2 of BRD2 is important for its subcellular localisation. With the bump-&-hole system we are able to interrogate this phenomenon much more selectively, and have shown that GFP-BRD2 aggregation requires inhibition of both BDs – showing how the bump-&-hole system can be a much more effective means for investigating biology than current BET inhibitors. Further investigation is warranted to determine if this phenomenon is simply an artefact specific to GFP-BRD2 constructs, or actually represents how endogenous BRD2 is localised and functions within the cell nucleus.

There is currently constant progress in the development of novel in-cell target engagement assays, which could be used to confirm L/V-selective BET inhibition of the bumped compounds (but would not be useful for studying chromatin binding). The cellular thermal shift assay (CETSA) is a commonly used means of measuring target engagement in cells, and the ability of ligands to thermally stabilise proteins, without the need for much specialised reagents or equipment. Attempts were made to implement this assay for the project, utilising cell-lines developed by Dr Kwok-Ho Chan that used the Flip-in system to stably transfect U2OS cells with tetracycline-controlled, GFP-labelled BRD4 constructs. As the endogenous WT BRD4 and the exogenous, GFP-Labelled proteins could be monitored simultaneously by western blot this assay had the potential to show the selective engagement of L/V BRD4 in preference to WT BRD4 in the same system (as opposed to testing the two separately and comparing the results). The tetracycline-induced expression of GFP-BRD4 proteins was confirmed (figure S4.5A), and protocols for western-blotting of BRD4 have already been extensively developed in-house [132, 188]. However, no clear thermostabilisation could be observed, both in an unselective fashion with (+)-JQ1 and in a selective fashion with 9-ME-1 (figure S4.5B-C). A dose-response CETSA assay has been utilised on the BET proteins [85], but no such assay was successfully developed for this project.

Finally, some plate-based cellular target engagement assays have been developed that aim to translate the plate-based biochemical assays discussed earlier (chapter 3) from a purified, buffer-based context to a lysate or cellular environment. The most successful so far is the NanoBRET assay [195], in which a fluorescent probe (or covalently-labelled histone peptide) excites a NanoLuc luciferase recruited to the target protein, unless displaced by test compounds. As this assay requires the

transfection of a NanoLuc-tagged target protein it should be possible to alter the expression vector to express L/V mutants of the target BET proteins.

4.9 – Conclusions:

Using the cellular FRAP assay our bumped compounds were shown to be able to selectively displace L/V BET proteins from chromatin, and through varying the exact dose it was possible to find the concentration range in which this was optimal. The bump-&-hole system was further validated, as 9-ME-1 was shown not to effect the cell viability or colony-formation of WT BET-dependent cell-lines expressing no L/V mutants.

The bump-&-hole system was then utilised to answer biological questions regarding how the BET proteins, and their tandem BDs, work, on a molecular level, to bind acetylated chromatin and induce transcription of downstream genes. Through the FRAP assay it was shown that, for all BET proteins, the BD1 was most important to chromatin binding. Excitingly though, the importance of the BD1 varied between BET proteins, from being the sole driver of chromatin binding in BRDT to almost parity with BD2 in BRD3. Utilising a luciferase reporter-gene assay, the bump-&-hole system was then applied to investigate transcription of NF-κB-RE controlled genes. Although it is difficult to identify how exactly a given BD contributes to transcription (and not just chromatin binding), several BD2s appeared to play a greater role in transcription than could be explained by their importance to chromatin binding.

Further experiments, using overexpressed GFP-BET proteins to investigate cell-viability were attempted without success, outlining the limits of what could be achieved without more sophisticated genetic implementation of the L/V mutation.

Chapter 5
Advanced Biology

5.1 – Introduction:

The ultimate goal of the BET bump-&-hole system is to chemically interrogate the function and importance of the BET proteins/BDs, both in healthy biological systems and during disease. Although we have been able to investigate BET protein function in cells by transient expression of L/V BET constructs such experiments are not sophisticated enough to properly investigate the BET proteins. Further experiments require mutant BET proteins to be expressed at a level similar to their WT counterparts, and for their expression to be regulated and vary appropriately.

The first attempt to develop such a system, performed by Dr Kwok-Ho Chan, used the Flp-in and T-Rex systems to stably transfect U2OS cells with tetracycline-dependent GFP-BET constructs, while knocking-down endogenous genes with siRNA (figure S5.1A-B) [6, 159]. The tetracycline dose was optimised to equalise GFP-BRD4 (long) expression with that of the endogenous gene, but unfortunately the exogenous gene (even when WT) could not restore the expression of BRD4-regulated genes following siRNA treatment (figure S5.1C). This is due to the GFP-BRD4 expression being broadly comparable to the endogenous gene, but not actually regulated in the same manner. In addition this Flp-in system only allows for a single isoform to be expressed, but BRD4 has multiple splice variants with differing functions [46, 47]. In addition the αBRD4 siRNA was 'leaky' and allowed for some endogenous expression. Pleasingly, ET was still shown to be active and L/V-selective in altering the expression of BRD4-regulated genes (figure S5.1D). Given these results it was clear that the ideal model cell-line would contain the L/V mutations knocked-into endogenous BET genes, both ensuring proper regulation and splicing, while removing the endogenous WT genes. Given the different gene-editing tools currently in use the best way to resolve these issues is to use the CRISPR/Cas9 gene-editing system.

The CRISPR/Cas9 system, now used for gene editing, evolved as an anti-viral/phage component of the adaptive immune systems in bacteria and archaea [196]. In infected prokaryotes, viral DNA is recognised and inserted into CRISPR arrays within the host genome. This vDNA is modified and transcribed into gRNA strands, which bind the Cas9 nuclease and direct it to cleave and inactivate matching vDNA. The CRISPR system appears to exist in ~45% of bacteria and ~85% of archaea species, and there is some evidence that it has been adapted for use in DNA repair and gene regulation.

Some viruses and phages have evolved to counteract the CRISPR system, either through random mutation of PAM sites or the expression of CRISPR-inhibiting proteins [196].

From 2013 onwards, the CRISPR/Cas9 system has been repurposed for deliberate gene-editing [197-199], gradually replacing the older and less practical zinc-finger nucleases and TALENS systems [200-203]. In this context, target cells are transfected to express a nuclease (typically Cas9), which will recognise PAM sites within DNA ("NGG" for Cas9) and gRNA strands [204-206]. gRNAs consist of 2 RNA strands: crRNA possesses homology for the target DNA sequence (which must be within ~20bp of a PAM site), and tracrRNA which forms a hairpin loop with the crRNA to form an active complex and direct Cas9 to the target DNA. The Cas9 nuclease will then generate a double-strand break at the target sequence, which is generally repaired through either non-homologous end joining (NHEJ) or homology-directed repair (HDR) (figure 5.1.) [204-206]. In NHEJ the broken strands are directly reconnected, but with the generation of random indel mutations. These random indels can often lead to the knock-out of all or part of the target gene, for which the CRISPR/Cas9 system is frequently utilised [204-206]. HDR instead repairs the DNA break using a homologous donor DNA template (typically a sister chromatid). When a DNA template is also introduced into the target cell a sequence of desired DNA, containing insertions, deletions or mutations, can be intentionally knocked-into the target gene. These knock-ins can be used to introduce novel genes or regulatory units, create reporter genes, introduce tags onto endogenous proteins or introduce small mutations [204-206]. The need for an additional component (DNA template), and the overall dominance of NHEJ-mediated DNA repair, results in knock-outs being far more common and easier to generate outcome of CRISPR/Cas9 transfections than knock-ins.

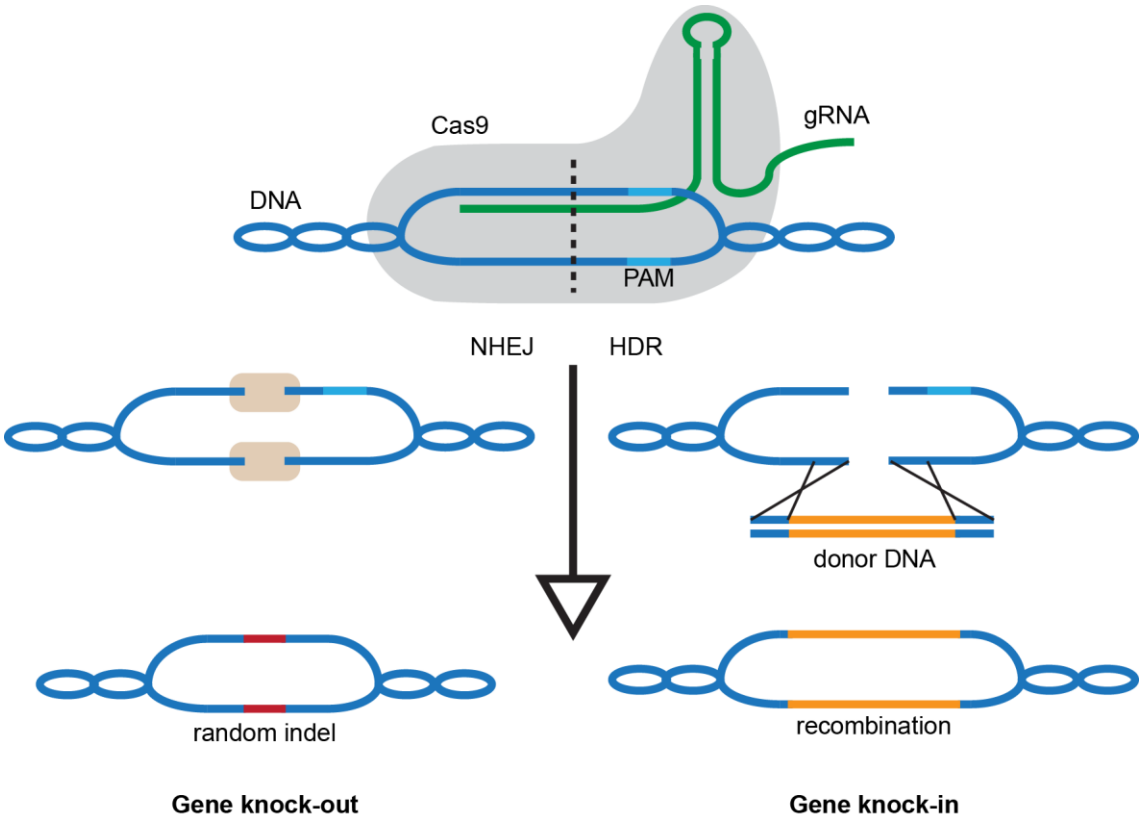


Figure 5.1. CRISPR/Cas9 mediated gene editing.

The Cas9 nuclease binds to a gRNA with homology to target DNA, and to a PAM site in the target DNA. Following cleavage the double-strand break can be repaired via NHEJ, which can create random indel mutations, or by HDR using a donor DNA template. The random mutations from NHEJ can create knock-out mutants, while the right donor DNA template used in HDR can be used to create knock-in mutants.

In CRISPR/Cas9 experiments the Cas9 nuclease, gRNA and (if performing a knock-in) DNA template must be introduced into the target cell. The gRNA strand may consist of 2 separate strands (tracrRNA and crRNA) that later interact, or a single strand combining the 2. Transfection of cells with plasmids can be used to introduce all 3 components, with the transcription of gRNA strands, transcription and translation of a Cas9 gene and containing the DNA template. Instead of a plasmid containing a Cas9 gene some protocols rely on transfection of a Cas9 mRNA which can then be immediately translated [207]. Several cell-lines have also been developed that stably express a permanently integrated Cas9 gene, and hence require transfection of only gRNA and DNA template [208, 209]. The DNA template can also be introduced as a single or double-stranded DNA molecule instead of being contained within a larger plasmid. The exact transfection protocol varies with target cell-line and nature of CRISPR components, with cation-lipid transfection reagents, electroporation, microinjection and viral transfection used in the literature [205].

5.2 –Design of CRISPR Knock-In System:

The first CRISPR knock-ins designed for the BET bump-&-hole system were designed by Dr Kwok-Ho Chan and aimed to knock-in the L/V mutation onto the BD1 and BD2 of BRD4 in A549 cells (figure 5.2A-B). The decision of which cell-line to focus on is not a simple one. Creating L/V mutants of otherwise normal human cell-lines would better allow for an accurate investigation into the function of the BET proteins, however focusing on cell-lines derived from cancers in which the BET proteins are validated targets will help understand the role of BET proteins in disease and their potential as therapeutic targets. Here A549 cancer cells were selected so as to sooner generate clinically relevant data.

A549 is a hypotriploid cell-line, meaning it may have elevated copy numbers for the BET genes [210]. Increased copy numbers will both complicate the generation of homozygous knock-ins, and potentially complicate the analysis of later experiments into knock-in/knock-out detection and genetic characterisation. One published A549 karyotype shows 2 copies of chromosomes 6 (BRD2), 9 (BRD3) and 19 (BRD4) and 1 copy of chromosome 1 (BRDT), but 13 abnormal chromosomes generated by translocation of chromosome regions. The abundance of abnormal chromosomes complicates estimations of gene copy numbers. The CanSAR database [211] gives elevated copy numbers for all BET-encoding chromosomes except chromosome 6. However the same database shows that only BRD3 is triploid, while the other BET genes are diploid. The Harmonizome database shows no BET gene to have an increased or decreased copy number [212]. Finally, the CellMiner NCI-60 Analysis Tool suggests BRD4, BRD3 and BRDT are diploid while giving no data on BRD2 [213]. Based on the above sources (specifics in table S5.1) it was decided to assume that A549 cells possess 2 copies of the BRD4 gene. It must be noted, however, that A549 cells have displayed cytogenic instability [210] and the BRD4 copy number of the specific A549 cell-line in our possession was not experimentally confirmed.

The full DNA sequence for the BRD4 gene (including introns) was obtained from the NCBI (NC_000019.10 reference GRCh38.p12 primary assembly), and the exons that form the long BRD4 isoform were identified (reference sequences NM_058243.2 / NP_490597.1 [214]). The codons encoding for L94 and L387 were identified and the requisite DNA changes for the L/V mutation were determined. To aid with later

analysis of potential KI clones a novel restriction site was introduced into each BD, using silent mutations when this fell into an exon region. Based on DNA sequence, commercial availability and activity in relevant buffers restriction sites for enzymes BamHI and ApaL1 were introduced into BD1 and BD2, respectively. Next, two PAM sites (sense and antisense) were identified, to allow for Cas9 binding. To minimise off-target double-strand breakage it was decided to utilise the *S. pyogenes* Cas9 D10A nickase mutant, which uses a PAM site with sequence "NGG". By reducing off-target activity this nickase mutant should reduce toxicity and increase the reliability of results obtained using any knocked-in cells [215, 216]. As the nickase will only cleave one strand of DNA a pair (sense/antisense) of PAM sites and gRNAs are required. The 20 nucleotides up-sequence of each PAM site was used to form the target-binding crRNA component of the gRNAs. To prevent repeated cleavage of knocked-in DNA the L/V knock-in was also engineered to remove the PAM sites through the introduction of silent mutations in the donor DNA sequence. In some cases a single mutation could be used for two KI elements (i.e. the L/V mutation also destroying a PAM site). To this donor DNA 500bp stretches of sequence before and after the L/V knock-in were added as homology arms for HDR. Finally, a set of primers were synthesised for later PCR-based assays and sequencing (figure 5.2C) (table S5.2). Both Fwd and Rev primers were generated, with 1 pair located at the outer edges of the homology arms/insert (inner) and 1 pair just outside the homology arms/insert (outer). The last primer was developed to be selective for knocked-in DNA sequences, exploiting the sequence changes introduced for the L/V mutation, restriction site introduction or PAM site deletion.

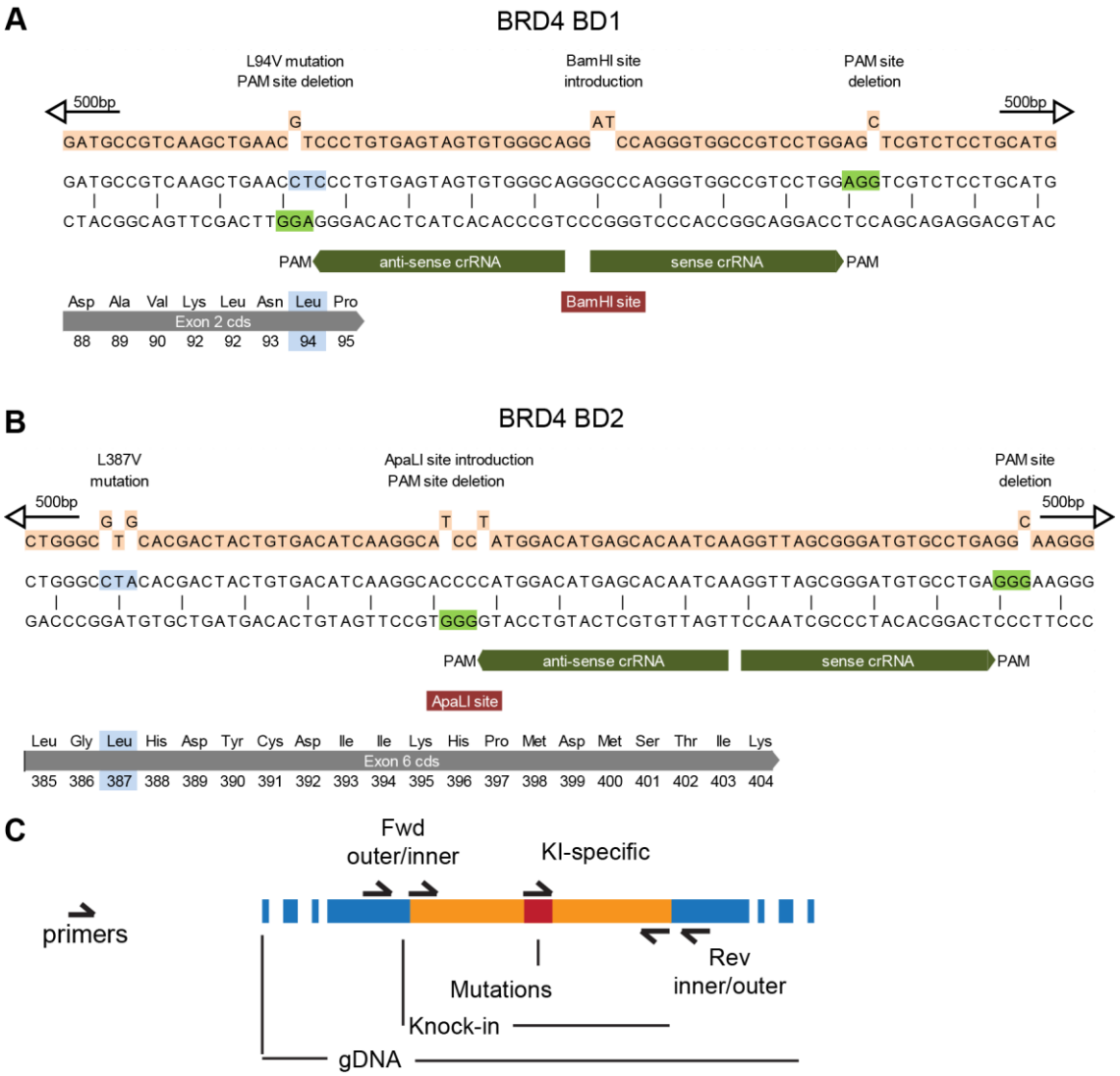


Figure 5.2. BRD4 L/V knock-in design

Design of BRD4 L/V knock-in reagents targeting BRD4 BD1 (A) and BD2 (B). Sense and anti-sense reference DNA sequence shown, alongside donor DNA template (orange) and protein sequence. L94 codon highlighted in blue. PAM sites and gRNA regions highlighted in green. Introduced restrictions sites shown in red. ~500bp homology arms on DNA template omitted for clarity. C) Design of primers for KI testing and sequencing.

The original knock-in protocol, developed by Dr Kwok-Ho Chan, utilised transfection of multiple plasmid vectors. A pX335 plasmid encoded a Cas9 D10A nickase gene and the gRNA, while a pBABE vector had the knock-in DNA templates encoded in its sequence. Plasmids were transfected into cells using Lipofectamine LTX (Thermo), and transfected cells were later selected for using treatment.

The design of CRISPR reagents for BRD2, BRD3 and BRDT knock-ins was begun by myself. These designs were not completed as it was anticipated that, following

BRD4 knock-ins, better ways of designing reagents may be realised. DNA templates and gRNA sequences were based on NCBI chromosome sequences (table S5.3) (figure 5.3). The codon responsible for the critical leucine residue was mutated to encode a valine. To develop novel restriction sites a semi-computational approach was utilised. The website tool re site finder [217] can be given a DNA sequence, and will check several commercial catalogues for restriction sites either already present in the given sequence or that can be introduced through silent mutations. Re site finder was given the DNA sequence adjacent to the L/V mutation and generated many potential restrictions sites that required 3 or less mutations. Priority was given to restriction enzymes that were either already used in the AC lab, or readily available. Enzymes with specific restrictions sites were preferred over those with variable sites (i.e. enzyme AleI with site 'GACNNNNNGTC'). Finally to simplify later identification of knock-ins, enzymes with medium to high activity in the buffers of high-fidelity DNA polymerases Phusion and Q5 were preferred (table 5.1). Various PAM sites were identified around the L/V and restriction site mutations. The ideal PAM sites would be those that can be deleted through a silent mutation and are 40-70 bp apart and close to the introduced mutations. In addition, PAM sites could only be exploited if their relevant crRNA (20 bp up-sequence) did not possess significant potential off-targets in the genome, as off-target cleavage is a serious concern with the CRISPR/Cas9 system [218, 219], which can be predicted through several computational tools [220-222].

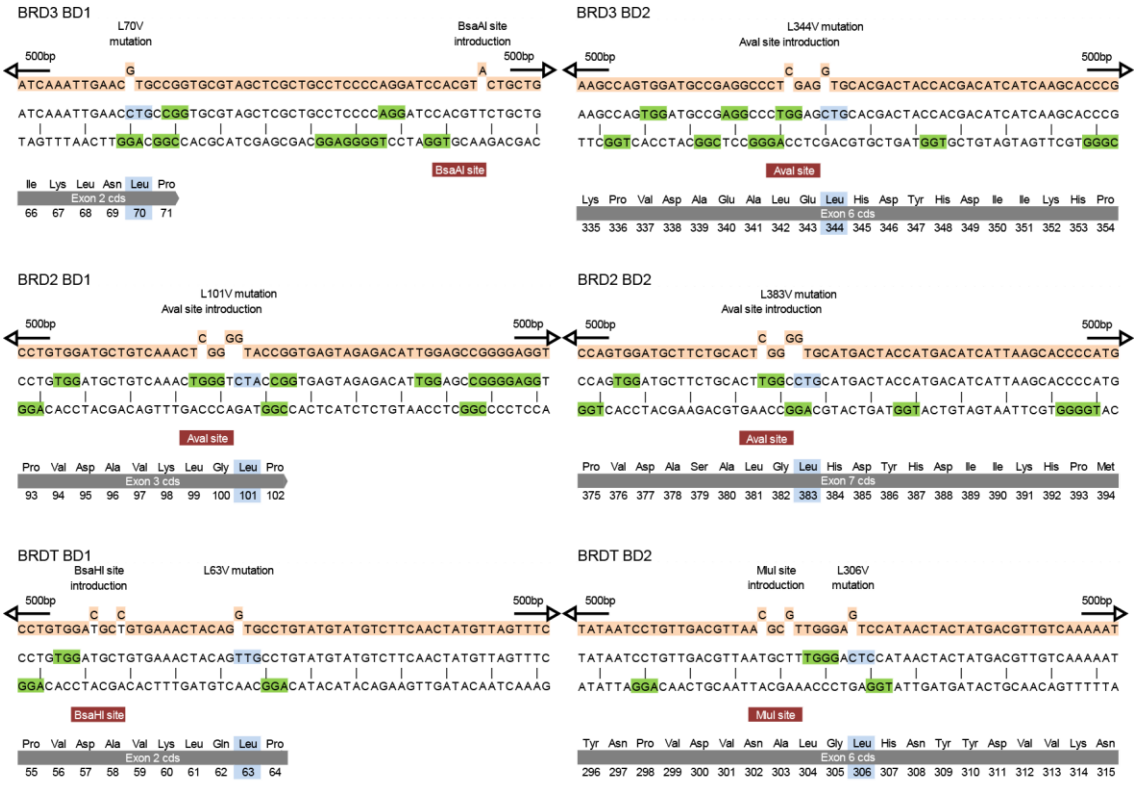


Figure 5.3. Non-BRD4 CRISPR designs

Design of L/V knock-in DNA templates for all non-BRD4 BDs (orange), based on reference gDNA sequences. Encoded protein sequence included, and target leucine codon/residue highlighted in blue. Potential PAM sites (NGG) shown in green. Actual ideal PAM sites to be exploited may be outside of depicted region.

BET protein	BRD3		BRD2		BRDT		BRD4	
BD	BD1	BD2	BD1	BD2	BD1	BD2	BD1	BD2
Leu residue #	70	344	101	383	63	307	94	387
Leu codon	CTG	CTG	CTA	CTG	TTG	CTG	CTG	CTA
Val mutation	GTG	GTG	GTA	GTG	GTG	GTG	GTG	GTG
Enzyme	BsaAI	AvaI	AvaI	AvaI	BsaHI	MluI	BamHI	ApaLI
WT sequence	CAGGTG	CTGGAG	CTGGGT	CTTGGC	GATGCT	ATGCTT	GGGCCC	CTACAC
Restriction site	CAGGTA	CTCGAG	CTCGGG	CTCGGG	GACGCC	ACGCGT	GGATCC	GTGCAC
Phusion activity	+++	+++	+++	+++	+++	++	+++	<+
Q5 activity	++	<+	<+	<+	+	++	<+	<+
Modifications	2	2	3	3	3	3	3	2

Table 5.1. CRISPR knock-in L/V mutations and novel restriction sites.

DNA sequence for WT leucine codons, and intended valine mutants. Restriction enzyme for which novel restriction site is introduced. WT DNA sequence, and mutations creating novel restriction site. Values of enzyme activity in PCR buffers is from New England Biolabs [223]. +++ - 100%, ++ - 50%, + - 25%. Total sequence changes required for L/V mutation and novel restriction site numbered.

5.3 – Optimisation of RNP Transfections

While an abundance of CRISPR work has been carried out with the kind of plasmid-based systems described above that majority of contemporary work is now performed using so-called 'RNP transfections', wherein purified Cas9 proteins are pre-incubated with gRNAs (and in the case of knock-ins template DNA) and these ribonucleoprotein (RNP) complexes are directly transfected into cells (figure 5.4) [224-226]. This new system has various advantages over the traditional plasmid-based transfection. Target cells are no longer required to produce the Cas9 nuclease or gRNA strands; the chance of cells not being transfected with all necessary plasmids (or not expressing all components at the same time) are greatly reduced; the Cas9 nuclease is rapidly cleared from the cells which greatly limits off-target mutagenesis and cytotoxicity [226]. One notable disadvantage of the system is that, without the presence of a plasmid expressing a selection marker (i.e. puromycin resistance) it can be more difficult to select for transfected cells and hence identify successful knock-ins.

A number of short experiments were planned, attempting to knock-in the BRD4 L/V mutations into A549 cells, with the aim of confirming the suitability of RNP-based knock-ins, determine the optimal protocol and compare the results to the plasmid-based system. It was hoped that the knock-in efficiency of different protocols could be compared using gDNA extracted from non-sorted transfected cells (a mix of untransfected cells and hetero/homozygous knock-outs/ins) using two PCR-based assays. Fully evaluating the efficacy of each protocol would require single-cell sorting of the resulting populations and one-by-one screening for the presence of the knock-in and was hence not practical at this stage.

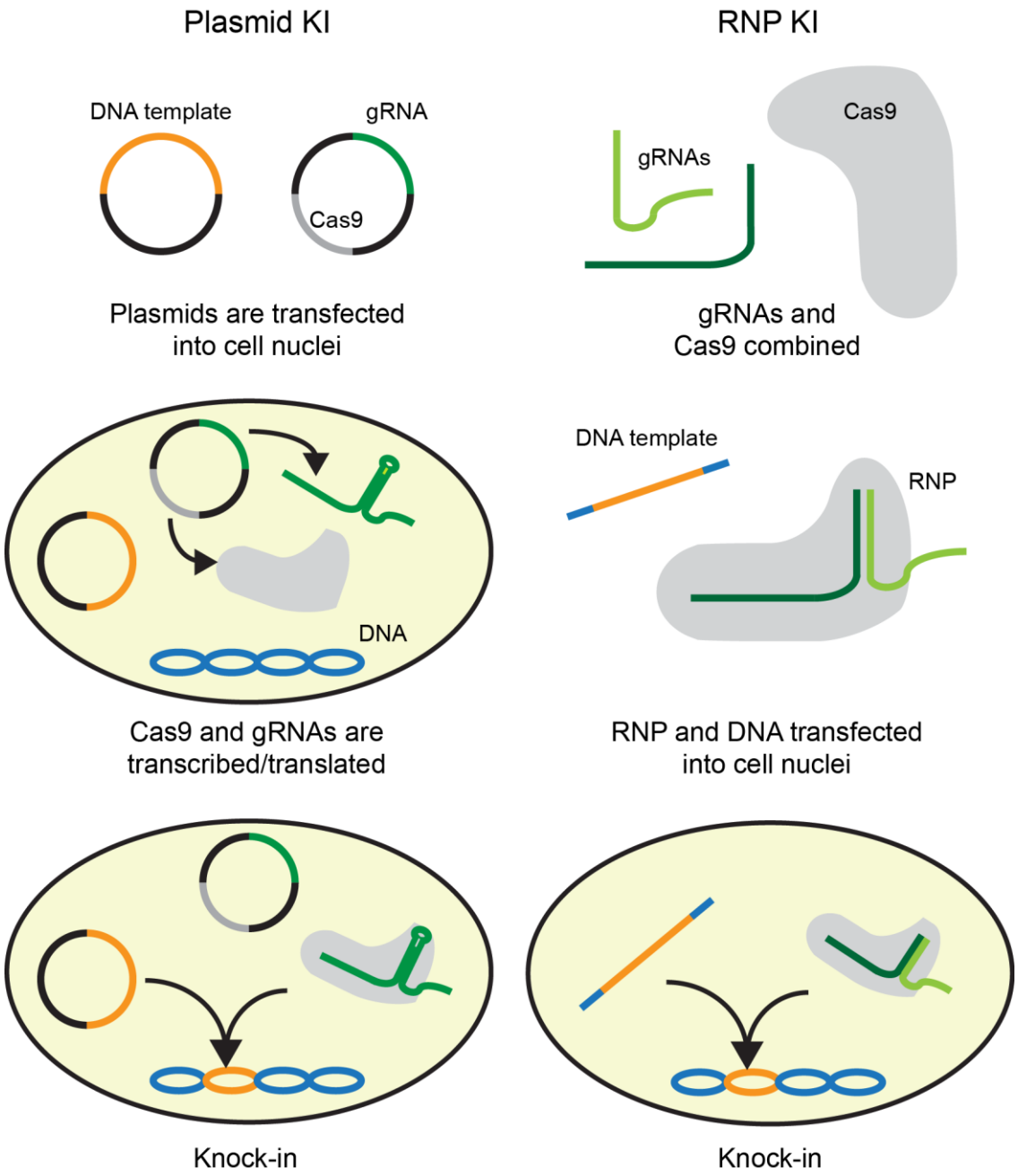


Figure 5.4. Comparison of plasmid and RNP-based knock-ins.

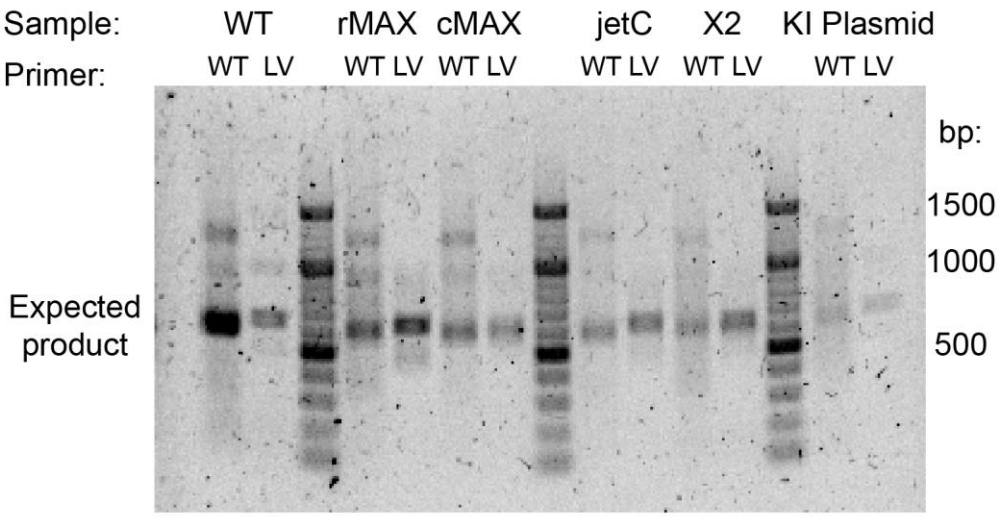
Plasmids method: plasmid containing template DNA and plasmid encoding single gRNA and Cas9 gene transfected into nucleus. gRNA and Cas9 are expressed, combine and use template DNA for knock-in. RNP method: separate crRNA and tracrRNA are combined, then combined with purified Cas9 protein. RNP complex and ssDNA template are transfected into nucleus, where they work to perform knock-in.

The two knock-in detection assays, developed by Dr Kwok-Ho Chan rely on PCR amplification of the BRD4 BD1/BD2 regions of extracted gDNA (figure S5.2). In the original 'cleavage' assay an ~1100 bp region of the genome, encompassing the knock-in insert, is amplified before undergoing an overnight digestion with a restriction enzyme. As the L/V knock-ins contain engineered restriction sites (BamHI for BRD4 BD1, ApaLI for BD2) cleavage of the PCR product will only be observed with successful

knock-ins. This assay is quantitative, and theoretically allows for estimations of the knock-in efficiency of different protocols (by analysing the gDNA of mixed cell populations) and can discriminate between hetero and homozygous knock-ins (when analysing FACS-sorted cells). Preliminary experiments validated the assay (figure S5.3A), although the 'knock-in' signal from a population of unsorted cells was almost undetectable. An alternate 'amplification' assay foregoes the restriction step and instead aims to amplify the knock-in region through a knock-in specific primer. By utilising the various knock-in mutations it is hoped that a primer can be specific enough to only amplify DNA when the knock-in is present. This assay should be able to generate a strong signal even when the knock-in rate is very low, but this assay will not be quantitative. Again, preliminary experiments validated the assay (figure S5.3B), but did raise the possibility of false-positives arising from interactions with leftover template DNA in the cells. Another serious concern was just how selective the primers could be, and how low a false-positive rate was acceptable given the anticipated low knock-in efficiency.

BRD4 BD1 L/V knock-ins were attempted with the RNP protocol and a range of transfection reagents: RNAiMAX (Thermo), CRISPRMAX (Thermo), jetCRISPR (Polyplus) and TransIT-X2 (Mirus), with cells transfected by Dr Kwok-Ho Chan with the plasmid-based protocol used for comparison. Knock-in efficiency was first analysed using the amplification assay, with mixed results (figure 5.5). The expectation was that the WT primer would generate a strong ~500 bp band for all samples, while the L/V primer would produce a fainter but visible band for any sample with a successful knock-in (i.e. the plasmids transfection, hopefully the RNP transfections but not the WT cells). Instead the results were far less clear as the L/V primer produced a band with WT cells and the amount of DNA loaded onto the gel varied greatly between samples. It was necessary to look not at the intensity of the 'L/V band' in isolation, but to compare it to the intensity of the bands produced by WT-specific primers (calculating the LV/WT ratio) in order to control for varying levels of gDNA. This, pleasingly, saw that all RNP transfections were more effective than the plasmid protocol, with TransIT-X2 being clearly the most effective. The reliability of the data is questionable however, as the L/V-specific primer still generated a strong band with WT gDNA (although this sample did give the lowest LV/WT ratio).

A



B

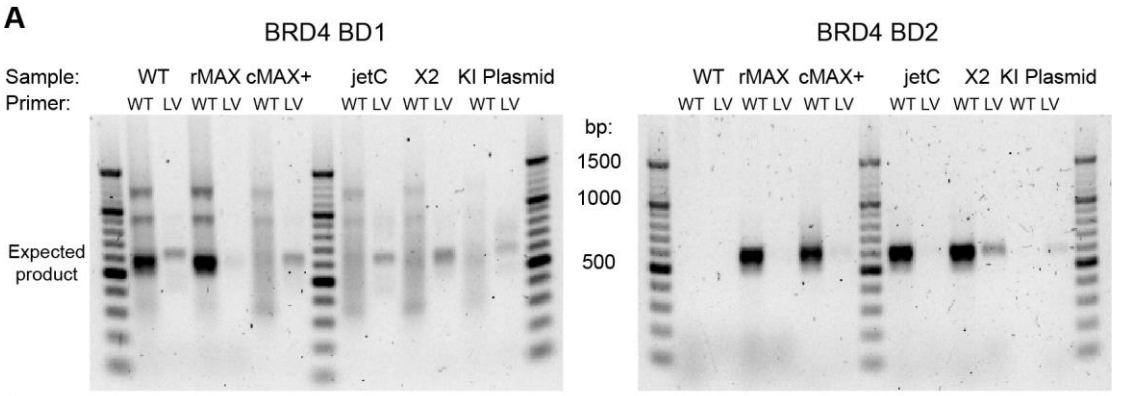
BRD4 BD1	WT	rMAX	cMAX	jetC	X2	KI Plasmid
L/V Intensity	14400	23000	13100	12600	16100	3290
LV/WT ratio	0.4	1.1	1.0	1.1	1.4	0.7

Figure 5.5. CRISPR RNP Test 1

A549 cells were transfected with RNP complexes for BRD4 BD1 L/V knock-in, using transfection reagents RNAiMAX (rMAX), CRISPRMAX (cMAX), jetCRISPR (jetC) TransIT-X2 (X2), or by Dr Kwok-Ho Chan using a plasmid-based protocol (KI Plasmid). Genomic DNA was extracted, and the BRD4 BD1 region amplified with either WT or L/V-specific primers. A) PCR products were ran on a 1% agarose gel alongside Generuler 1kbp PLUS size ladder (Thermo) and visualised with SYBR safe. B) Intensities of ~700 bp bands were quantified in Image Studio Lite (Licor). LV/WT ratio = intensity of expected product band with L/V primer / intensity with WT primer.

The transfections were repeated for both BRD4 BD1 and BD2 knock-ins, and gDNA of higher quality was extracted. Additionally, following the manufacturer's protocol, the PLUS reagent was used alongside the CRISPRMAX transfection reagent. Both the BD1 and BD2 datasets indicated that TransIT-X2 was again the superior transfection reagent, and both showed CRISPRMAX (with PLUS reagent) to also be effective (figure 5.6). Both datasets also show RNAiMAX (not designed specifically for RNP transfection) as the worst transfection reagent. Strangely, the BD1 assay again showed knock-in signal for the WT cells (again, the lowest of all samples) and suggested that the plasmid-protocol was actually superior to RNP transfection. The BD2 assay meanwhile again indicated the superiority of TransIT-X2 and CRISPRMAX,

but failed to generate any signal with the WT primer with the WT and plasmid-transfected samples – again calling the reliability of the assay into question.



B

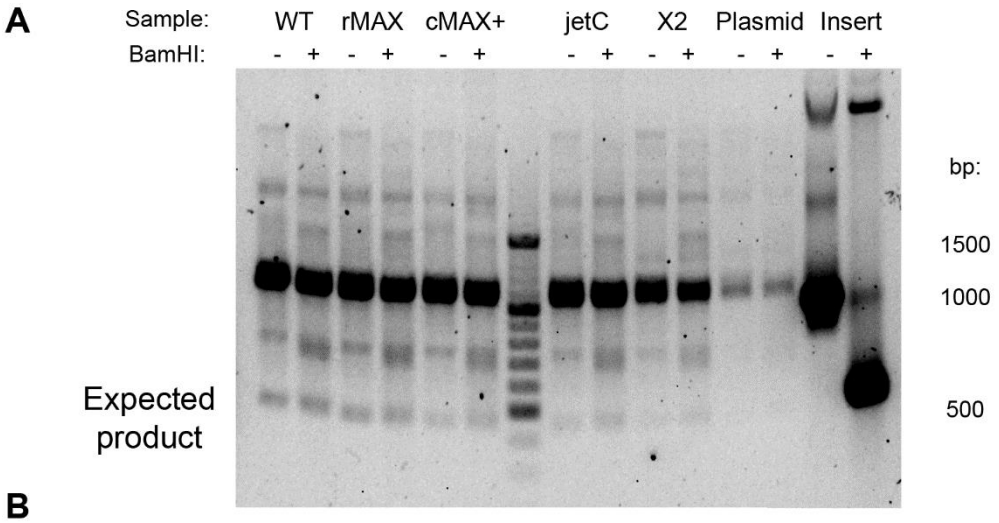
BRD4 BD1	WT	rMAX	cMAX+	jetC	X2	KI Plasmid
L/V Intensity	6720	1920	6360	6660	10000	3210
LV/WT ratio	14	17	323	31	350	783

BRD4 BD2	WT	rMAX	cMAX+	jetC	X2	KI Plasmid
L/V Intensity	550	2050	3600	2830	11700	3450
LV/WT ratio		19	183	89	409	

Figure 5.6. CRISPR RNP Test 2.

A549 cells were transfected with RNP complexes for BRD4 BD1 and BD2 L/V knock-in, using transfection reagents RNAiMAX (rMAX), CRISPRMAX with PLUS reagent (cMAX+), jetCRISPR (jetC) TransIT-X2 (X2), or by Dr Kwok-Ho Chan using a plasmid-based protocol (KI Plasmid). Genomic DNA was extracted, and the BRD4 BD1 region amplified with either WT or L/V-specific primers. A) PCR products were ran on a 1% agarose gel alongside Generuler 1kbp PLUS size ladder (Thermo) and visualised with SYBR safe. B) Intensities of ~700 bp bands were quantified in Image Studio Lite (Licor). LV/WT ratio = intensity of expected product band with L/V primer / intensity with WT primer.

The cleavage assay was then run on the BD1 samples, in the hopes it would prove more reliable and quantitative. Here a strong ~1100 bp 'uncut' band was expected from all samples, while successful knock-ins would produce a very faint ~500 bp 'cut' band. The best KI method would show a greater proportion of 'cut' PCR product. Again the results were highly mixed, with all samples showing a range of bands in addition to the expected ~1100 bp PCR product and the ~500 bp cleaved fragment (figure 5.7). This assay also, like the amplification assay, indicated the presence of knock-ins with WT cells, and actually suggested that the WT cells possessed a greater proportion of knock-ins than the RNAiMAX or TransIT-X2 transfected cells.



B

BRD4 BD1	WT		rMAX		cMAX		jetC		X2		KI Plasmid	
	uncut	cut	uncut	cut	uncut	cut	uncut	cut	uncut	cut	uncut	cut
+/- Ratio	1.3		1.0		1.5		1.4		1.0		3.2	
Cleave %	16	20	14	16	10	14	8	11	6	7	3	1

Figure 5.7. CRISPR RNP Test 3.

A549 cells were transfected with RNP complexes for BRD4 BD1 and BD2 L/V knock-in, using transfection reagents RNAiMAX (rMAX), CRISPRMAX (cMAX), jetCRISPR (jetC) TransIT-X2 (X2), or by Dr Kwok-Ho Chan using a plasmid-based protocol (KI Plasmid). Genomic DNA was extracted, and the BRD4 BD1 region amplified, then digested overnight with BamHI. A) PCR products (and knock-in template) were ran on a 1% agarose gel alongside Generuler 1kbp PLUS size ladder (Thermo) and visualised with SYBR safe. B) Intensities of 'cut' ~500 bp bands and 'uncut' ~1100 bp bands quantified in Image Studio Lite (Licor). +/- ratio = intensity of 500 bp band with BamHI / intensity without BamHI. Cleave % = intensity of 'cut' bands as a percentage of the 'uncut' ~1100 bp bands.

The RNP transfections could have been replicated, and modifications to the protocols for each assay could have been made to increase their reliability. Higher-fidelity DNA polymerase enzymes could have reduced the false negative/positive rate, while PCR clean-up kits could remove many of the unwanted bands. The main optimisation step would likely be the tweaking of the PCR protocol, to both remove unwanted bands and to improve the accuracy of the WT and L/V-specific primers. This may have required the use of different PCR cycles for WT versus L/V primers (and BD1 vs BD2 primers), which would make the assays much less practical.

Repetition and optimisation of the RNP transfections and knock-in assays was not carried out, and the decision was taken to progress to a genuine knock-in attempt. Identifying the best transfection protocol, and maximising the knock-in efficiency would greatly aid in designing a protocol for obtaining knocked-in cells, but it was not

clear if that could actually be done. Even with improvements to PCR cycles it was not clear if any reliable and meaningful data would be obtained while analysing mixed, unsorted populations of cells following a transfection, and the cleavage and amplification assays may only be reliable or sensitive enough when utilised on sorted cells wherein each sample would contain only WT cells or only knocked-in cells. Optimisation of the knock-in strategy would have to be done in an incremental fashion, with multiple approaches done in tandem, with the protocol that generated the most (or any) knocked-in cells used for knocking-in the L/V mutation on the next bromodomain.

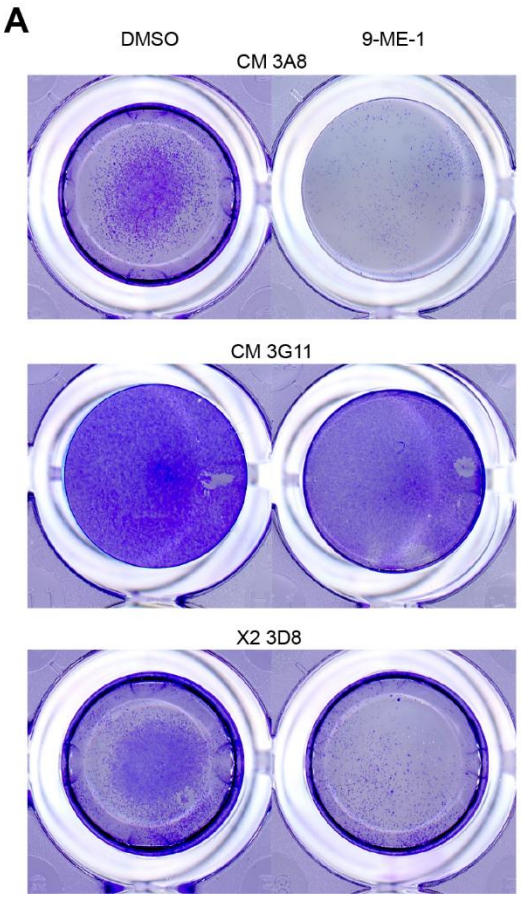
5.4 – Knock-In Attempt – BRD4 BD1 L/V:

The first proper attempt to knock-in an L/V mutation was carried out on A549 cells and aimed to introduce the L94V mutation into the BD1 of BRD4. It was planned to later attempt the BD2 knock-in on both WT and BD1 L/V cells, so as to acquire all three BRD4 L/V cell-lines (LV/WT, WT/LV and LV/LV). Low-passage number A549 cells were revived from an aliquot in cryostorage and grown-up in a 6-well plate. Once at 60% confluency the CRISPR transfection was performed. Due to the additional length of the plasmid-based system and the slow growth of the revived cells only RNP-based knock-ins were attempted. Specifically the knock-in was attempted using the TransIT-X2 and the CRISPRMAX (with PLUS reagent) transfection reagents, given their slight superiority in the (albeit unreliable) optimisation experiments.

Following transfection the cells were then given 3 days to grow in the incubator before being sorted. This allowed for new cells to replace those killed by the transfection (or resulting CRISPR activity) and for any RNP complexes to be cleared from the cells. The cells were not left to grow any longer, in case any L/V mutant cells that were generated were then out-competed by WT cells and made harder to identify later. The two populations of cells were then sorted by the School of Life Science Flow Cytometry and Cell Sorting Facility. From each population three 96-well plates were seeded, each well containing a single clone, giving potentially 576 distinct clones to analyse. Clones were hereafter referred to by their location post-sort, for example CM 2D7 refers to the CRISPRMAX-transfected clone sorted into well D7 of plate 2.

In the hope of reducing the number of clones requiring testing through PCR-based assays a triage was attempted using the 96-well colony-formation assay described in chapter 4. As cells started to reach confluence each of the six 96-well plates was passaged and 3 new plates seeded. Of these 3 plates 1 was handled normally, and its cells used for later experiments. The other 2 plates were treated with either DMSO or 2 μM 9-ME-1 for 1 week, and then fixed and stained to look at cell growth and cytotoxicity. It was hoped that, if BRD4 BD1 activity is essential for A549 cells, only successfully knocked-in cells would show a response to 9-ME-1. The cells grown in DMSO could also be used to confirm that the cells were still viable. Of course, given the lack of the BD1 mutant cell-lines the assumptions this assay rested on could not be confirmed in advance, nor was it known whether or not it would identify heterozygous knock-ins. As such it was felt that this experiment may not work, but if it was successful it would greatly reduce the amount of PCR work required later.

The assay was carried out as planned, revealing the difficulty on performing such an experiment on so many 96-well plates without the use of specialised hardware. The multiple pipetting steps were complicated by the need to replace tips between wells, to prevent cross-contamination; and efficient trypsinisation and passage of A549 cells proved difficult and likely inconsistent. The results of the assay were somewhat mixed. Checking each well with the light microscope suggests about 50% of the clones survived the cell sorting process and went on to grow and divide, while the colony-formation assay shows far fewer clones growing in the DMSO-treated plates. Pleasingly, there were several clones that seemed to grow quite well in DMSO and were significantly perturbed by 9-ME-1 treatment (some examples in figure 5.8). Additionally, after scanning the stained plates 1% SDS was used to extract the dye from the cells, allowing for measurement of absorbance at 595 nm and a more quantitative analysis. This dye extraction also highlighted some clones that visually appeared unaffected by 9-ME-1 but accumulated significantly less crystal violet. Due to doubts about the assay's reliability it was decided to take a larger number of clones forward for later analysis. By collecting any clone to show any degree of 9-ME-1 sensitivity (either visually or by A595 measurement) two 24-well plates were filled with 48 clones. Meanwhile the 96-well plates were frozen and stored at -80°C, with the hope that some clones could be revived for further analysis at a later date.



C

A595	CM 3A8	CM 3G11	X2 3D8	X2 3B5
DMSO	0.42	0.99	0.49	0.64
9-ME-1	0.22	0.66	0.39	0.69
Δ	0.20	0.34	0.10	-0.05

Figure 5.8. Colony-Formation Assay on CRISPR Clones.

CRISPR-transfected and FACS-sorted A549 cells seeded in 96-well plates, treated with 0.02% DMSO or 2 µM 9-ME-1 and incubated for 7 days, fixed with 100% MeOH and stained with 0.1% crystal-violet. A) Clones showing apparent 9-ME-1 sensitivity, identified by visual analysis of stained cells or A595 measurement of wells following dye extraction with 1% SDS. B) A clone showing no such 9-ME-sensitivity. C) A595 values for the above clones.

The 48 clones were grown up in 24-well and later 12-well plates, and their gDNA extracted using the Qiagen DNeasy blood & tissue kit. As the L/V-specific PCR primer amplification assay was seen as unreliable the knock-in could be directly confirmed through PCR amplification of the BD1 region and either BamHI digestion of the knocked-in restriction site, or DNA sequencing of the PCR product directly. As DNA sequencing is obviously highly dependent on the minimisation of PCR-associated errors in DNA sequence it was decided to switch from Taq polymerase to a high fidelity enzyme. The QF polymerase would be ideal, but BamHI is not active in its reaction buffer, which would complicate the restriction assay. The Phusion polymerase also possesses high fidelity while allowing for BamHI activity. A trial PCR amplification and

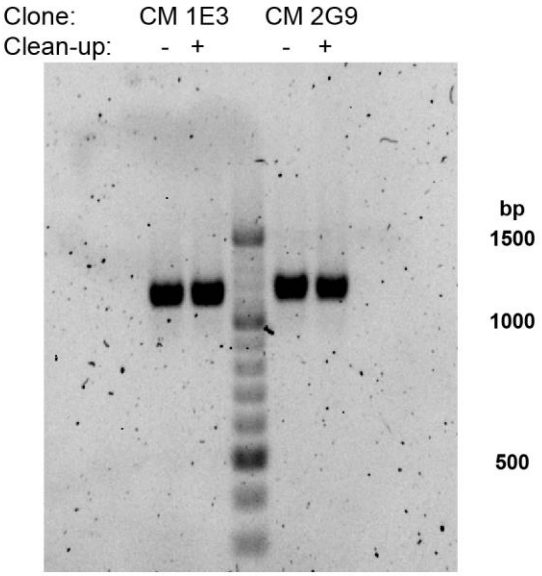
clean-up (NEB Monarch kit) was performed on two samples, to ensure later samples would be of high enough quality for DNA sequencing (figure 5.9A). For both samples the major PCR product was of the correct length, and little PCR product was lost in the clean-up (figure 5.9B). In the gel electrophoresis there were no visible impurities in the raw PCR product, and hence no improvement in the 'clean' samples. Meanwhile UV nanodrop analysis of the samples showed that the cleaned samples were of much higher quality. The A260/A280 ratio was much improved by the clean-up protocol. The removal of proteins and other impurities has greatly lowered the apparent DNA concentration (~135 to ~10 ng/µl) while the gel electrophoresis shows no actual reduction in DNA content.

The two purified PCR products were then submitted for DNA sequencing, using the 'Fwd Inner' primer (~80 bp from start of PCR product, ~450 bp from knock-in mutations). The sequencing reactions were successful, and the provided sample was both abundant and pure enough to generate reads with reliable coverage of the knock-in region. Neither clone tested was shown to be successfully knocked-in (hetero or homozygous), but the results did show that PCR-amplification of the clone gDNA, and using DNA sequencing to test for the knock-in, was a practical way to test the remaining 45 clones (figure 5.9C-D). For discussing subsequent sequencing results it was decided to refer to individual nucleotides/basepairs not by their absolute position within the reference gDNA sequence, but by their position relative to the critical L/V mutation (C59726G for BRD4 BD1). Both samples did show an unintended, off-target mutation (G to T, at position +49) which was originally thought to be a PCR error but later was seen to be a single-nucleotide polymorphism (SNP).

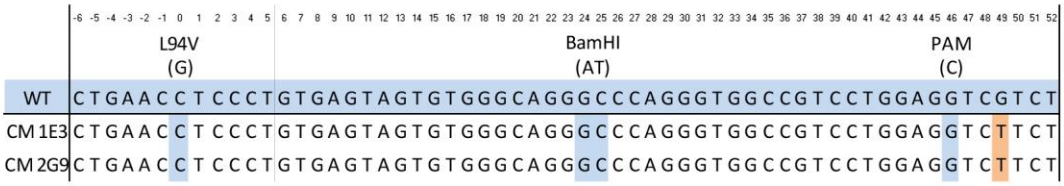
A

	CM 1E3		CM 2G9	
	Raw	Clean	Raw	Clean
ng/µl	139	8.9	134	13.8
A260	0.28	0.18	0.27	0.28
A280	0.21	0.10	0.22	0.16
260/280	1.3	1.7	1.2	1.7

B



C



D

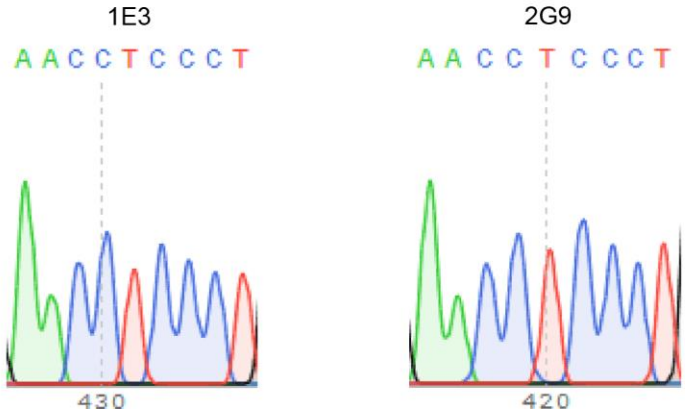


Figure 5.9. Validation of BRD4 BD1 PCR, Clean-Up and DNA Sequencing.

BRD4 BD1 region amplified from gDNA by PCR (Phusion polymerase), and resulting reaction mixture purified. A) Nanodrop analysis of raw PCR mixture and purified product. B) 10 µl of raw PCR product and purified product run on a 1% agarose gel alongside Generuler 1kbp PLUS size ladder and visualised with SYBR safe. C) DNA sequence of original WT gDNA and purified PCR products. Blue = WT, orange = non knock-in mutations. Location and nature of desired knock-in mutations are given. D) Portion of sequencing chromatograms, centred on the target Leu codon, displayed in SnapGene Viewer.

The PCR and sequencing protocol was applied to all 44 remaining clones (2 of the 48 selected clones failed to grow) (figure 5.10). Of the 46 clones that underwent sequencing 6 failed to generate any reads (figure 5.11A) (discussed later), 2 appeared to contain heterozygous knock-outs and the remaining 36 showed no sign of any knock-in. The 35 clones showing no knock-in were disposed of, and the 8 clones with potential knock-outs were frozen and stored at -80°C for later investigation. The G/T mutation at position +49 was seen in all PCR products, showing it to be a SNP and not an off-target mutation or PCR error.

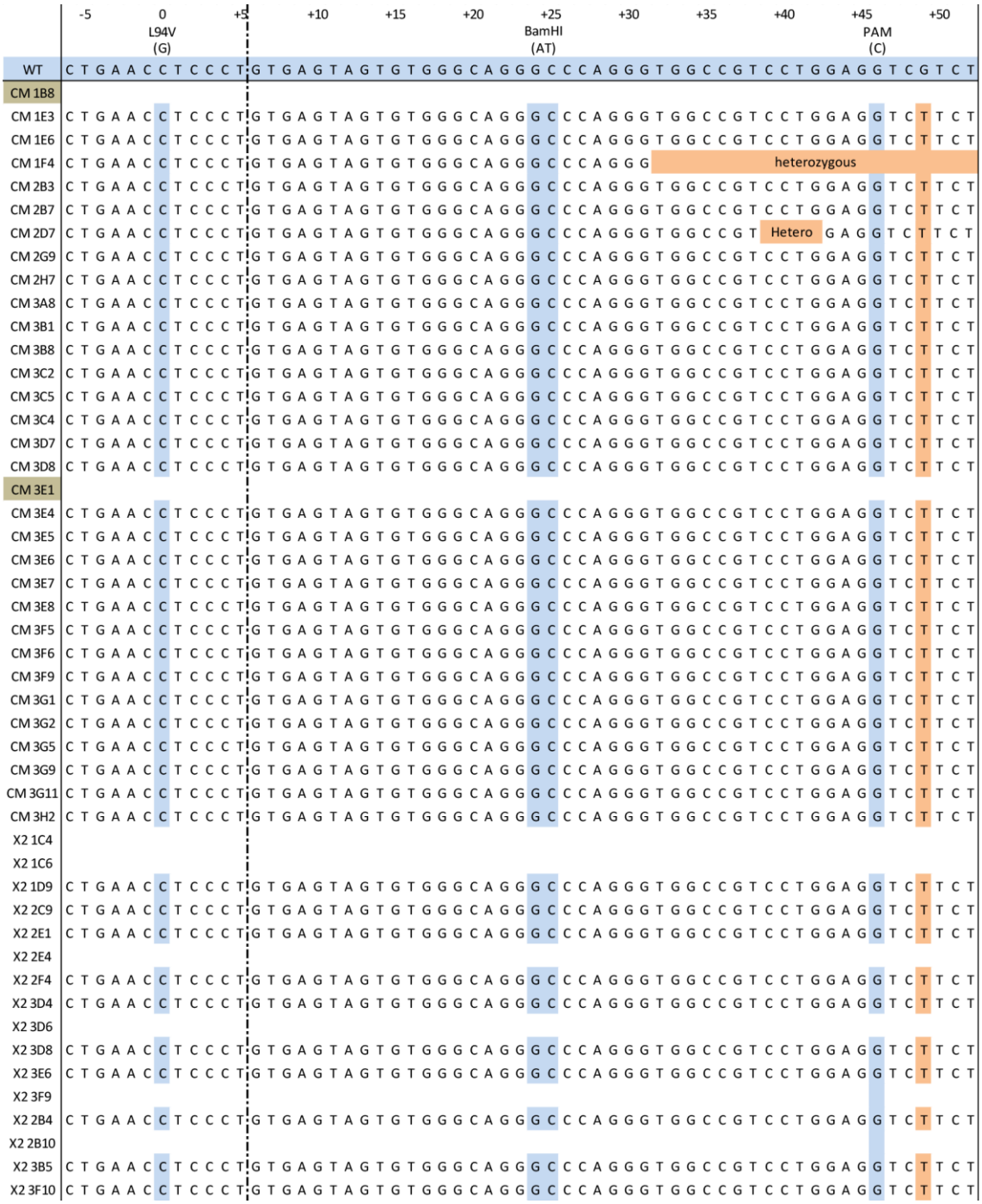


Figure 5.10. CRISPR Sequencing Screen.

DNA sequence of PCR products derived from A549 clones targeted for BRD4 BD1 L/V knock-in, compared to WT reference sequence (NC_000019.10) in blue. Nucleotides are numbered relative to the key L/V C59726G mutation. Location and nature of knock-in mutations given. Key knock-in nucleotides that were not mutated are highlighted in blue. Divergences from WT sequence are highlighted in orange. Clones CM 1B8 and 3E1 failed to grow after being transferred to 48-well plate, prior to gDNA extraction and PCR. Clones X2 1C4, 1C6, 2E4, 3D6, 3F9 and 2B10 did not produce any PCR product with the BRD4 exon2 Fwd and Rev outer primers. Sequencing performed with BRD4 exon2 Fwd inner primer.

5.5 – Knock-Out Characterisation:

Potential heterozygous BRD4 BD1 knock-outs were identified in clones CM 2D7 and CM 1F4. They were first identified by a sudden breakdown in the coherence of the reads at positions +39 and +32, respectively, and overlapping signals at many positions (figure 5.11C). A look at the chromatogram for CM 2D7 showed that some stretches of the WT sequence were still identifiable, like the 2 Cs at positions +39 and +40 and 6 Cs from positions +83 to +88. The PCR products of these clones was sequenced using the Rev Inner primer, which generated reads suggesting the potential knock-outs CM 2D7 and 1F4 terminated at positions +42 (2D7) +66 (1F4), giving knock-outs with length 4 and 35 bp, respectively (figure 5.11D). The PCR products were ran on a 1% agar gel, and no clear difference in size is visible (figure 5.11B). Very faint ~1000 bp bands are visible for clone 1F4, but these are likely from an adjacent size-marker. The knock-outs are contained entirely within the intron sequences.

If the BRD4 gene is triploid in A549 cells then, strictly, the term ‘heterozygous’ cannot be used as it refers to diploid systems. As the triploid nature of BRD4 is not certain, and lacking any better term, the term ‘heterozygous’ will still be used. But, it must be noted that ‘heterozgous’ knock-out clones may in fact only have 1 (or 2) of 3 BRD4 gene copies modified instead of the traditional 1 of 2.

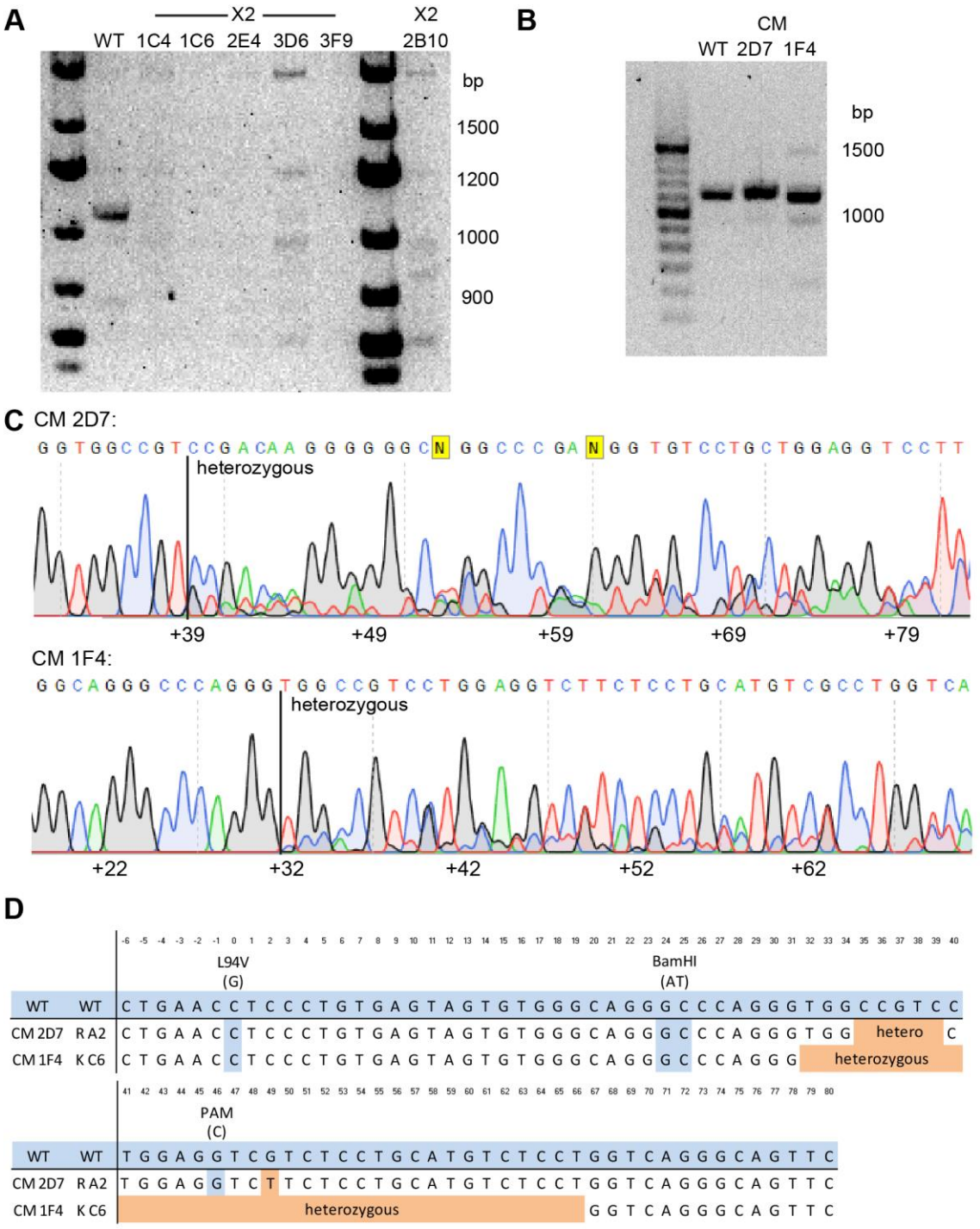


Figure 5.11. Genetic characterisation of potential knock-outs.

A) Gel electrophoresis of BRD4 BD1 PCR products from WT A549 cells and potential KO clones ran alongside GeneRuler 100bp Plus DNA Ladder (Thermo) and visualised with SYBR safe. B) Gel electrophoresis of BRD4 BD1 PCR products from WT A549 cells and potential heterozygous KO clones run alongside Generuler 1kbp Plus DNA ladder (Thermo) and visualised with SYBR safe. C) Portion of sequencing chromatograms, centred on beginning of suspected heterozygous knock-out, displayed in SnapGene Viewer. D) DNA sequence of PCR products derived from potential KO clones, compared to WT reference sequence (NC_000019.10) in blue. Nucleotides are numbered relative to the key L/V C59726G mutation. Location and nature of knock-in mutations given. Key knock-in nucleotides that were not

mutated are highlighted in blue. Divergences from WT sequence are highlighted in orange. Sequencing performed with BRD4 exon2 Fwd inner primer.

To investigate the significance (if any) of these potential knock-outs the viability of the clones was assessed using the CellTiter-glo and 96-well colony-formation assays (figure S5.4). The colony-formation assay, ran on all potential KOs, showed all clones to have reduced viability (figure 5.12A). Clones X2 2E4, X2 3F9 and CM 2D7 showed some visible cell growth during the assay, while the remaining clones showed no visible colonies. Meanwhile the CellTiter-glo assay, ran only on the CM 2D7 clone, showed no drop in cell viability for the potential KO (figure 5.12B). The pronounced difference in the growth of clone CM 2D7 between the two assays could be due to a technical error when seeding the colony-formation assay, or suggest a loss in cell adherence.

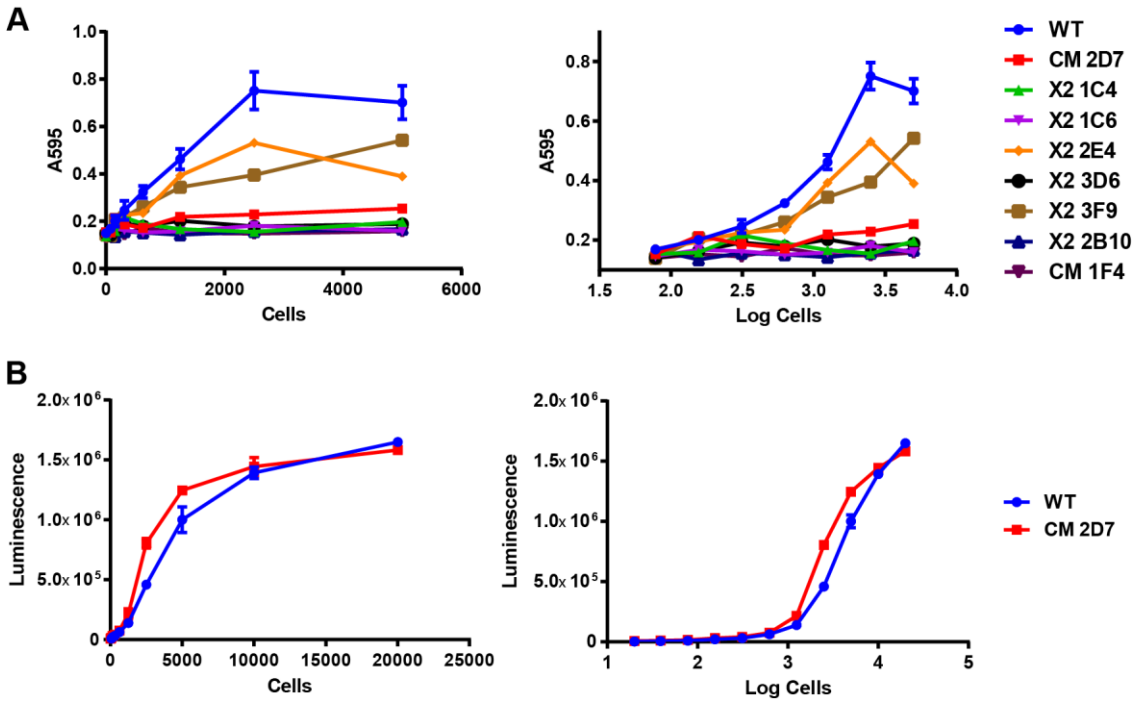


Figure 5.12. Viability of Potential BRD4 BD1 Knock-Outs

A) Absorbance at 595 nm of A549 cells stained with crystal violet dye (7 days' growth), and dye extraction in 1% SDS. B) Luminescence of A549 cells measured via Celltiter-Glo assay (3 days' growth.)

The remaining 6 clones (X2 1C4, 1C6, 2E4, 3D6, 3F9 and 2B10) failed to generate any reads at all, and were later shown to have generated no PCR product, due to either serious problems with the gDNA extraction or hinting at sequence changes so widespread that the annealing sites of one or both PCR primers were deleted (such as an extensive and homozygous knock-out). Unfortunately there was

insufficient time to further investigate these clones and determine the exact cause of the lack of sequence reads.

5.6 – Discussion:

One early step that could have aided this knock-in attempt would be ensuring that the starting population of A549 cells was clonal. As A549 cells often display cytogenic instability the cells utilised in CRISPR transfection was likely polyclonal. This would add layer of variability to later experiments, which may account for some odd results. Different populations within our A549 cell-line may grow at different rates, complicating the passaging of the many clones that were sorted post-transfection. When the colony-formation assay was repurposed to screen for 9-ME-1 sensitivity (a potential marker of L/V knock-in) differences in the copy number or expression of BRD4 (or other proteins) make it unlikely that (excluding knock-ins) all A549 cells are equally sensitive to the 9-ME-1 dose. This could also explain why, when characterising the viability of potential BRD4 knock-outs, different clones showed different growth-rates. Before future knock-in attempts a monoclonal population should be established. This can be done through cell-sorting, as was done post CRISPR transfection, or even by extensive dilution of a polyclonal population and seeding into a multi-well plate.

Both the optimisation of the knock-in transfection and the detection of knocked-in clones is limited by the assays involved in knock-in detection. The amplification assay, using WT and L/V-specific PCR primers, did not give the clear and reliable binary response hoped for, and the occasional lack of bands generated by WT primers casts doubt on its results. It is difficult to tell how ‘selective’ a given primer is, and given that our knock-in templates were not designed for such an assay we may not be able to generate very selective primers from the small changes in DNA sequence introduced. New knock-in templates, with additional silent mutations may allow for much more selective primers. Another potential optimisation is to alter the annealing temperature used, which should reduce ‘off-target’ amplification. This optimisation could be quite complex, especially if the WT and L/V-selective primers possess differing optimal annealing temperatures. The cleavage assay, utilising the introduced restriction site, should be more reliable and quantitative but does rely on a sufficiently

high knock-in efficiency. Using the cleavage assay on unsorted cells may not be realistic, although some techniques discussed below may help with this. If new KI templates are designed then it might be worthwhile to remove an endogenous restriction site, rather than introducing a novel one. One could then run a 3-part assay: non-selective PCR amplification, cleavage, and a second round of PCR using primers selective for uncleaved PCR products. The second round of amplification should then boost the sensitivity of the assay.

Instead of our in-house developed PCR-based assays for KI detection the SURVEYOR nuclease assay, from the literature [227, 228], could be used instead. In this assay the target locus (in this case BRD4 BD1/BD2) is amplified from gDNA using PCR. This PCR product will (hopefully) contain a mixture of homozygous 'WT' and 'KI' DNA duplexes, but through denaturation and re-annealing any KI duplexes can be converted into partially mismatched heterozygous WT/KI duplexes. This mismatched DNA will then be selectively cleaved by nucleases such as CEL II and T7 endonuclease I, producing additional, smaller bands following gel electrophoresis [227-229]. Quantification of gel bands can then be used to calculate an estimated of the proportion of gDNA containing a modification. The SURVEYOR assay has been used often in the literature when discussing CRISPR-mediated gene editing [197, 199, 207]. The SURVEYOR assay in many ways acts like our existing cleavage assay, which has provided mixed results, but it does remove the need to introduce a new restriction site during the knock-in. Its main advantage is that, as the WT and KI DNA duplexes are re-annealed the final percentage of DNA that is cleaved by the nuclease should theoretically be double the percentage that was originally knocked-in, doubling the sensitivity of the assay [227]. The SURVEYOR assay is hence very effective for measuring knock-out efficiency, but it may be far less suited for knock-ins. Unlike our cleavage assay the SURVEYOR assay measures all changes to DNA sequence – including KOs and off-target mutations [227], and hence the majority of 'gene-modification' signal we observe may not be due to knock-ins. Comparing the strengths of different assays (table 5.2) the SURVEYOR assay may not be as accurate as the cleavage assay when determining if individual clones possess knock-ins, but it can be used to determine if gRNA pairs are working, and for optimising transfection/knock-in protocols.

Assay	Sensitivity	Quantitative?	KI-specific?	Throughput	Source
Amplification	High	No	In theory	High	In-house
Cleavage	Low	Yes	Yes	Mid	In-house
SURVEYOR	High	Yes	No	Low	Literature

Table 5.2. Comparison of knock-in detection assays.

Knock-in detection could also be carried out by progressing from PCR and gel-electrophoresis to Southern blots [230]. In such experiments PCR products from BD sequences would be transferred to into membranes where they can be tested with hybridisation probes. This could be used alongside the cleavage assay, where the use of probes could increase sensitivity and clarify results otherwise complicated by unwanted DNA bands (as in figure 5.7). Protocols for assessing CRISPR knock-out and knock-in efficiencies using next generation sequencing are also being developed. One such protocol [231] can be performed on an unsorted population of cells and is highly sensitive.

Outside of CRISPR-based gene-editing, the literature features many examples of assays developed to detect specific single-nucleotide polymorphisms (SNPs) (too many to describe here) which could be repurposed to detect our desired knock-ins [232]. In the Invader assay [233] two oligonucleotides with partially overlapping homology for the SNP form a three-part structure with the sample DNA strand, which can be cleaved by the Flap endonuclease (FEN). One of the overlapping oligonucleotides possesses both a fluorophore and complementary quencher label, hence cleavage by FEN will generate a fluorescence signal. If the SNP is not present in the sample DNA the three-part structure is not formed properly, is not cleaved by FEN and the fluorophore is unable to generate any signal. Another assay, TaqMan [234, 235], works on a similar basis with a SNP-commentary oligonucleotide probe possessing a fluorophore/quencher pair on opposite ends of the probe. In the absence of the SNP this probe forms a hairpin structure, keeping the fluorophore/quencher pair in proximity and preventing signal. When this probe binds to a SNP-containing DNA sample the fluorophore and quencher will be brought apart, generating signal. These SNP assays may be more complex and require more expensive reagents than our PCR-based assays, and may still face issues with selectivity and false-positives. Additionally

the oligonucleotide probes will need to be carefully designed and assay conditions optimised. The relatively high-throughput nature of these assays, and their apparent success in detecting very small changes in DNA sequence does mean they may be ideal for future knock-in detection.

In order to maximise the chance of obtaining a heterozygous or homozygous BRD4 BD1 L/V knock-in almost 600 A549 cells were sorted into 6 96-well plates, to be grown and analysed. This proved to be too many cells to properly care for and study. A549 cells appear to be somewhat resistant to trypsinisation, often requiring ~7-10 minutes instead of ~5 minutes for the other adherent cell-lines worked on in this project. This, combined with the 96-well format, made passaging the clones very difficult and inconsistent. For many reasons (polyclonal starting population, variation in media volume, edge-effect in plate) different clones would grow at different rates, needing to be passaged on different days and split to a greater extent. Additionally, the mere act of checking the confluence of 600 clones was impractical, and clones could not really be passaged at different rates. Finally, running the PCR-based KI-detection assays on hundreds of clones would have been very difficult, time-consuming and would require large amounts of expensive enzymes and gDNA extraction kits. For this reason the colony-formation assay was used to look for 9-ME-1 sensitivity, and narrowed the population down to ~45 clones. This assay, however, was likely not very reliable due to the above mentioned differences in clone confluence and growth, and the difficulty in running said assay in such a small format and such large numbers. A better way to utilise 9-ME-1 to detect L/V knock-ins might be to instead develop a 9-ME-1-dye/fluorophore conjugate and see if it accumulates in KI versus WT cells

CRISPR work by Dr Kwok-Ho Chan in a separate project suggests that obtaining heterozygous knock-ins can be reasonably easy, and may only require 100-200 clones. Hence it may help with later knock-in attempts to run much smaller, easier to manage, experiments with the aim of obtaining heterozygous knock-ins (figure 5.13). These heterozygous knock-ins can then undergo a second round of transfection and sorting to obtain homozygous knock-ins. One downside to this could be the risk of doubling the likelihood of deleterious off-target double-strand breakage during the knock-in process, which could be especially severe when developing double-mutants (i.e. BRD4

BD1 L/V, BD L/V cell-lines) which would undergo four rounds of transfection. As this approach would double the amount of 'work' done on said knock-in cells they accumulate other, not immediately noticeable, changes in the genome. This increased divergence from WT may hinder the usability of the knock-ins, as results obtained from them may not be of any use in understanding WT cells. Heterozygous knock-in cell-lines may also be of use while, although bumped compounds will only be able to inhibit ~50% of the expressed target BET protein. Whether or not this would be sufficient to investigate BET protein/BD function will likely depend on the exact context one is studying. If the target BET gene is triploid then performing the L/V knock-in in only 1 of 3 genes would further limit the impact of bumped compounds, while generating homozygous (3 of 3 genes) in a stepwise fashion would require an additional round of transfection.

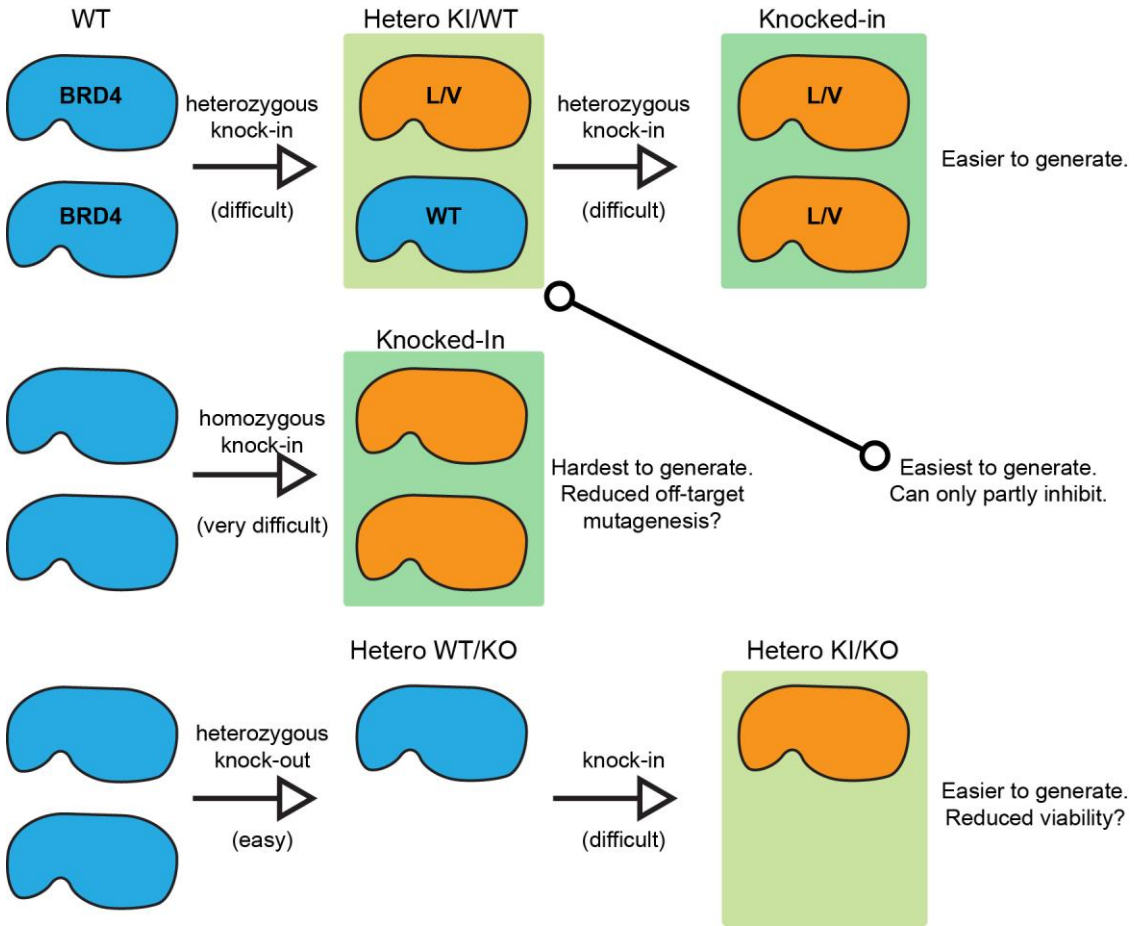


Figure 5.13. Different methods to generate knock-in clones.

Four different, potentially usable, clones can be generated through CRISPR, following one or two rounds of transfection. Target genes (e.g. BRD4) can be heterozygously or homozygously knocked-out or knocked-in, with heterozygous being easier than homozygous and knock-outs being easier than heterozygous.

Despite not detecting any knocked-in clones several knock-out clones may have been generated. Two clones are believed to possess short heterozygous knock-outs within an intron sequence in the BD1 region of BRD4, while 6 clones may have much more widespread, homozygous knock-outs. Western blot analysis of BRD4 levels could identify whether these potential knock-outs have prevented BRD4 expression, or potentially lead to the expression of truncated isoforms lacking the BD1. Due to the pre-eminence of NHEJ-mediated DNA repair it is expected to generate substantially more KO than KI clones following a CRISPR experiment, and the appearance of these potential KO clones can be taken as evidence that the gRNAs used and RNP transfections are functional.

One potential strategy, moving forwards, is to create heterozygous knock-out/knock-in cell-lines, instead of the much more difficult homozygous knock-in (figure 5.13). If knock-outs are generated that abolish BD function, but do not affect cell viability (will be dependent on specific BD and cell-line) then the other, still WT, allele could be targeted for a L/V knock-in. If cells are able to survive with only one functional allele of a the BET protein in question then such a system should be a viable way to study said protein and its associated phenotypes with bumped compounds like 9-ME-1. Some knock-outs will remove or frameshift the entire BET gene, while others may only remove small stretches of the gene sequence and make relatively minor changes to the final protein sequence. In such cases site-directed mutagenesis could be used to recreate the effects of the knock-out on the protein level and produce protein constructs for characterisation. The effects of a knock-out on BD function could then be assessed by peptide ITC titrations on single-BD constructs, or FRAP experiments with full-length, GFP-tagged constructs.

The use of RNP transfections, as opposed to transfecting cells with plasmids expressing Cas9 and containing KI template DNA, is believed to have many advantages but does have one serious downside. Plasmid-based transfections allows for the use of FACS to select for successfully transfected cells, as one can use plasmids that also express GFP proteins. In order to allow for FACS-enrichment of transfected cells IDT has recently developed fluorophore-labelled gRNAs. By attaching the ATTO-550 dye to the long, invariable tracrRNA FACS can be used to select for cells that have been successfully transfected with RNP complexes [236]. This approach may be somewhat

limited if a sizable proportion of the fluorescent RNA is not packaged into full RNP complexes, but is still transfected into cells. These cells would then still be selected for but would not have any chance of being knocked-in. The ATTO-550 dye has been confirmed to be compatible with the FACS equipment used in the School of Life Sciences. By enriching for transfected cells it may become possible to use the PCR cleavage assay to quantify the overall knock-in efficiency of a given transfection protocol, while without said enrichment the efficiency is too low for detection without single-cell sorting. If said overall efficiency can be measured, without single-cell sorting, then optimisation of transfection protocols will be much easier. Fluorescent gRNAs can also be used to measure simple transfection efficiency, perhaps by transfecting cells in a microscope dish and later imaging them for ATTO-550 and a DAPI-style nuclear stain [236].

One other change to our knock-in strategy that would allow for fluorescent selection of knocked-in cells would be to switch from a (relatively) subtle L/V knock-in onto endogenous BET proteins and instead knock-in a full-length, L/V-mutated GFP-BET construct (like those used in FRAP assays), either on top of endogenous BET genes or into other regions of the genome. This would allow for selection of knocked-in cells, as opposed to transfected cells with the fluorescent gRNA strategy. The reason this strategy has not been employed is the presence of multiple, potentially conflicting, BRD4 isoforms. Yet BRD2 and BRD3 do not possess such divergent isoforms and hence this strategy could be employed to more easily produce BRD2/3 mutant cell-lines. Depending on where these constructs are knocked-in there could still be issues with proper regulation of expression and silencing of endogenous genes; while depending on the promoters used one may only need to knock-in a single copy of the construct, further simplifying the process.

Several approaches have been trialled in the literature that seek to boost HDR-mediated DNA repair and knock-in efficiencies, which could be adapted for future bump-&-hole knock-ins. Attempts have been made to boost knock-in efficiencies by inhibiting NHEJ processes with small-molecules, thus increasing the utilisation of HDR by targeted cells. An inhibitor of DNA ligase IV, Src7, increased the efficiency of HDR-mediated gene editing at 4 gene loci in several cell-lines (including A549) from 2 to 19-fold [237]. In a later paper cytochalasin, an inhibitor of actin polymerisation, aided the

generation of knock-in mice embryos [238]. Similar results have also been seen with the B3-adronergic receptor agonist L755507 and the natural product resveratrol [239, 240]. Use of these inhibitors have not become particularly widespread due to the often small improvements in knock-in efficiency and concerns over toxicity [239, 241].

Small-molecules, and other tools, can also improve KI-efficiency by interfering with the cell-cycle (figure 5.14). NHEJ dominates DNA repair from the G1 to the G2 phase, while HDR-mediated NDA repair occurs only in the S and G2 phases (alongside NHEJ), as only at this stage can sister chromatids be used as repair templates [242]. As such perturbation of the cell-cycle, or temporally control of the CRISPR-process may increase knock-in efficiency. A combination of small-molecule cell-cycle synchronisation and well-timed CRISPR RNP nucleofection increased the efficiency of HDR-mediated DNA repair from ~10% to ~33% [243]. By fusing Cas9 and the N-terminal region of Geminin one lab created a 'Cas9-hGem(1/110)' construct that would be ubiquitinated and degraded during the G1 phase of the cell-cycle, and limit gene-editing activity to the less NHEJ-dominated S, G2 and M phases. This process almost doubled the efficiency of HDR in HEK293T cells, albeit from an unusually low baseline of ~0.5% [244]. Finally, a recent paper has shown that HDR efficiencies can be increased 2-10-fold in HEK293 cells through 'cold shocking' cells at 32°C for 24-48 h after transfection [245]. An accumulation of cells at G2 and M phases was an early explanation for this effect, but this has not yet been confirmed, while the thermodynamic stabilisation of DNA recombination intermediaries is another likely factor. Crucially, several of these studies [239, 243, 245] that showed increases in HDR efficiencies also observed a plateau at ~33%, suggesting this may be an absolute limit (without more extensive perturbation of biology). Another point to consider is that the above referenced works typically report on HDR efficiency, or the percentage of DNA repair mediated by HDR, and not a parameter more directly relevant to this work (i.e. percentage of transfected cells shown to possess a desired KI). The use of A549 cells stably transfected to express the nickase enzyme would simplify subsequent knock-ins, but would require a significant up-front investment of resources, would not help with knock-ins in other cell-lines and may perturb natural cellular functions. Said stably transfected nickase gene could potentially be placed under the control of certain

promoters and regulatory elements that maximise Cas9 levels during the S phase of the cell-cycle, when HDR occurs.

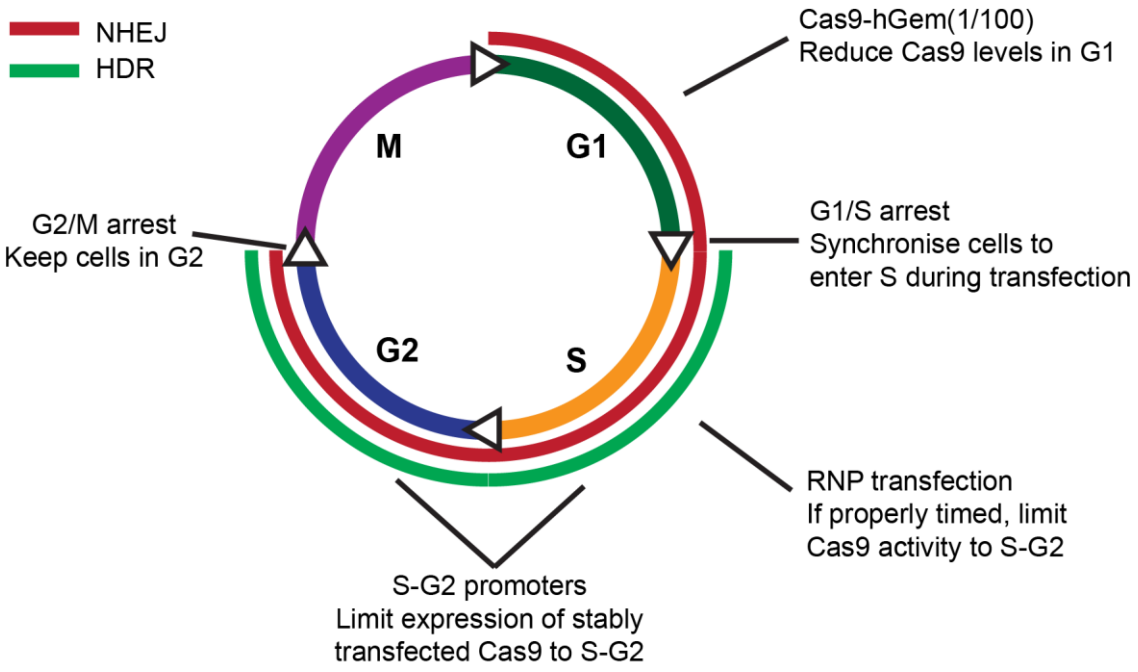


Figure 5.14. Exploitation and perturbation of the cell-cycle to boost knock-in efficiency.

G1 – growth. S – DNA replication. G2 – preparation for division. M – mitosis.

In addition to the above CRISPR knock-on optimisations there have also been advances in the overall methodology used to generate KIs, combining multiple changes in protocol and reagents. By switching from short ssODN to long ssDNA knock-in templates the *Easi*-CRISPR system increased knock-in efficiency from ~10% to ~50% and simplified template design [241]. The *Easi*-CRISPR system is limited with regards to insert size (~1.5 kbp limit), but this should still be suitable for our purposes. In the Tild-CRISPR system PCR amplification or enzyme cutting is used to generate dsDNA templates with ~800 bp homology arms, which are delivered into cells alongside gRNA and Cas9 mRNA [246]. Compared to a standard homologous recombination DNA template Tild-CRISPR increased overall knock-in efficiency 10-20-fold, while avoiding the insert size and financial cost of *Easi*-CRISPR templates. Finally, the PITCH system attempts to optimise gene editing by utilising MMEJ instead of HDR [247]. Double-strand DNA breaks can be repaired with MMEJ when short (~20 bp) microhomologies are present up-and down-stream of the break site, and this occurs in the M and S

phases of the cell-cycle. This system generated knock-in efficiencies of 60-85% in HEK293T cells and zebrafish embryos, and also allows for easy template design.

Our sequencing reads, derived from the gDNA extracted from transfected A549 cells, consistently showed several single base-pairs differing from the database-derived gDNA sequences we used for designing primers and templates. The most notable example was a T at position +49 instead of a G, and there were several more. A few isolated examples could be explained as errors introduced during PCR, and may be removed by the use of an even higher fidelity enzyme. However, the consistency at which these were observed suggests they are not errors, but simply examples of the A549 genome differing from those used to make database sequences. As far as can be discerned, these SNPs should not affect the BRD4 BD1 L/V knock-in, as they do not alter the PAM sites or regions targeted by gRNAs, and the large homology arms used in CRISPR can tolerate several mismatches. SNPs could be an issue when working with other BDs or in other cell-lines, where they may alter PAM sites and gRNA sites. Additionally, if these SNPs impact on protein sequence, splicing or other regulatory systems it would be ideal to not alter them in the course of the L/V knock-in, as we want to minimise sequence changes. In the future CRISPR knock-ins should be designed based on DNA sequences derived from A549 cells (or other intended targets), with sequencing primers designed using database gDNA sequences.

5.7 – Conclusions:

The development of cell-lines bearing the L/V mutation on endogenous, properly expressed BET proteins is key to the future use of the BET bump-&-hole system. This is likely best done through the CRISPR/Cas9 system, which we plan to utilise in an RNP-transfection method with the Cas9 D10A mutant nickase enzyme. Preliminary results suggest that both the plasmid and RNP-based transfections do generate knock-ins, and some optimisation may be possible but this is severely limited by our inability to reliably detect knock-ins within unsorted populations of cells. An attempt was made to knock-in the L/V mutation onto BRD4 BD1 in A549 cells, which appears to have been unsuccessful. This work has however taught many lessons on the practical aspects of CRISPR-mediated knock-ins, which in addition to advances in the literature and commercial CRISPR products should improve future attempts. In addition several clones were identified as potential hetero-/homo-zygous BRD4 or BRD4 BD1 knock-outs. This both suggests that our current CRISPR knock-in system is fundamentally sound, and may be useful for further studies.

Chapter 6
Discussion

6.1 – Implications for Bump-&-Hole Systems

After several years of developing, optimising and implementing the BET bump-&-hole system we believe some of our observations can be applied to bump-&-hole research in general (not just to BET bromodomains). These observations may aid in the design of similar systems with regards to other families of structurally conservative and therapeutically significant proteins.

The most fundamental component of any bump-&-hole system is the mutation utilised for selective inhibition of the preferred target. As any reliable investigation into the biology of the protein(s) of interest requires a functional mutant it is imperative to properly characterise any mutants developed, and to focus only on those that seem likely to function in cells or *in vivo*. Fortunately, we see a reasonable correlation in the ranking of mutants between low-level functionality assays with purified constructs (i.e. BLI / ITC) and higher-level cellular experiments (FRAP/luciferase), suggesting that functional mutants should be identifiable at early stages. Unfortunately, it is very difficult to know exactly how 'functional' a mutant needs to be to ensure viable mutant cell-lines or animal-models, and this will likely depend on the exact role of any target protein. Additionally, we do see some variance in the impact of mutations across different experiments, as the L/A mutants show a ~10-fold drop in affinity for peptides (ITC), a ~3-fold decrease in chromatin binding (FRAP) but an almost complete collapse of downstream gene expression induction (NF-κB luciferase assay). Researchers ought to focus on mutations that are as subtle and structurally conservative as possible (figure 6.1). The L/A mutation can be argued to not be too severe from a structural standpoint, but has since been shown to be suboptimal, and our preferred L/V mutation represents the loss of only a single heavy atom. Even the leucine/isoleucine L/I mutation could have been used, if the L/V mutation proved unusable, as without the loss of any heavy atoms it did allow for selectivity, albeit limited. Once an acceptable functional mutant has been characterised selective bumped compounds can then be assessed, and any limits on potential selectivity as a result of the necessary mutation must simply be accepted.

Other bump-&-hole projects outside of our work, like those in kinases or PKMTs, do utilise much more disruptive mutations (i.e. I/A, Y/A). This is likely the result of different proteins being potentially much more or much less tolerant of mutations,

in an essentially random fashion. Alternatively, this could be an indicator that PPIs are quite sensitive to mutations, while enzymatic catalytic sites are fairly tolerant of any mutations that don't directly affect catalytic residues.

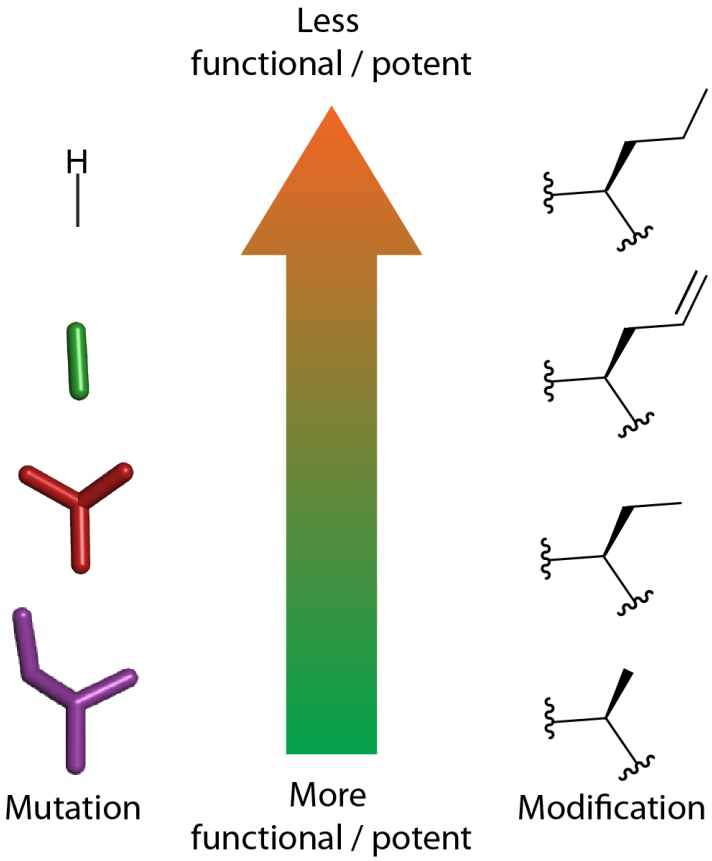


Figure 6.1. The supremacy of more conservative mutations and smaller bumps.
More severe mutations lead to a loss in mutant functionality, while larger and more flexible bumps tend towards a loss in compound potency.

Just as a subtle mutation is key, it seems that appropriately minor chemical modifications are likely the best way to generate the best mutant-selective inhibitors (figure 6.1). The BET bromodomain bump-&-hole system has progressed from a L/A mutant and a 2-carbon, ethyl bumped ligand to a L/V mutation paired with a 1-carbon, methyl bumped ligand (with additional methoxy-shift modification). Although larger bumps might be considered to offer enhanced selectivity they possess three severe drawbacks. The most obvious is the potential loss of affinity to the mutant, if the bump is too large or does not fit optimally into the created cavity. The second drawback, which has shown to be more severe in our SAR work, is that longer carbon chains will possess greater rotational flexibility, which is likely to seriously impair affinity/selectivity for mutants. Our 4-carbon, propyl-bumped compound was one of

the least potent and selective compounds trialled. The 4-carbon alkyl-bump, bearing a rigid double-bond, was much more effective, though notably still worse than the 2-carbon ethyl bump. The utilisation of double C=C bonds (as just discussed) or even the incorporation of 3-carbon rings into bumps could be beneficial, if they are possible synthetically. Thirdly, larger bumps can significantly alter the DMPK properties (LogP, membrane permeability and CL_{int}) of an inhibitor scaffold, which is not desirable for cellular investigations. Larger bumps (especially those possessing rings) will increase the hydrophobicity of the compound (as seen with the propyl bump) and potentially alter cell-membrane permeability and metabolic degradation by microsomes.

Through our SAR we have also revealed insights into other ways of modifying scaffold compounds to generate superior bumped inhibitors. Moving the methoxy group from the 8' to the 9' position on the scaffold (relatively distant from the bump and mutation) was quite beneficial and was present in two of the best compounds (9-ME and 9-ET), showing the value in trialling a range of modifications seemingly irrelevant to the target mutation. The strong impacts of the different alternate side-group modifications show that the flexibility of the part of the ligand the steric bump is derivatized from can be very important to affinity and selectivity. The residual WT binding of compounds like ET is best explained by the rotation of the methyl-ester side-group into a conformation that reduces the WT steric clash, and attempts to prevent this through amide side-group : protein hydrogen bonds had a strong impact on WT binding, producing several of the least WT-potent compounds tested. Rotational freedom of the bump site may also explain why the compounds targeting the W/F and W/H mutants (with modified chlorophenol rings) displayed low selectivity. Ideally, compounds would feature the bump coming off a part of the scaffold that is both rigid, yet correctly oriented to ensure any bump added is able to fit into the mutant cavity out of a suitably fixed vector.

For the bulk of this project the focus has been on the mutation and compound aspects of the bump-&-hole system, though the means by which the mutation is implemented into a model biological system is also key to its use in answering biological questions. Transient overexpression of GFP-BRD4 constructs was successful for investigating the importance of the BD1s and BD2s to chromatin binding and transcription, but attempts to utilise this to study the importance of BET BDs to cell

viability failed. These attempts showed that for more advanced experiments the mutant BET proteins must be stably expressed by model cell-lines, and the endogenous WT proteins must be suppressed or removed. Originally this was intended to be done by using the Flip-in system to stably integrate a tetracycline-dependent GFP-BRD4 gene into U2OS cells, while using siRNA to knock-down the WT gene [159]. Some results obtained with this system were promising, but the inability to properly regulate or splice the L/V BRD4 gene left it unable to fully rescue the expression of BRD4-dependent genes following knock-down of the WT gene. In order to both ensure that mutant BET genes are perfectly regulated and spliced, and completely remove endogenous WT expression, it was clear that the CRISPR/Cas9 system had to be used to knock-in the L/V mutation onto BET genes. This work has not yet been successful, but is likely the only way to use the bump-&-hole system to answer more detailed biological questions, like what BET proteins are needed for the proliferation of different cancers, or what downstream genes are regulated by which BET proteins.

6.2 – BET Protein Function Model:

Using the FRAP assay the importance of each BET proteins' BD1 and BD2 to chromatin binding was quantified. This work showed the BD1 to be the most important to chromatin binding, and that the BD2 had no role in chromatin binding for BRD2 and BRDT and little role in BRD4. In recent years several papers have been published showing evidence for interactions between BET bromodomains and acetylated non-histone peptides (figure 6.2). BRD4 BD2 has been shown to bind Twist and the RelA subunit of NF-κB (both when acetylated) [90, 95], BRD2 BD2 is believed to interact with acetylated STAT3 [112] and BRD3 BD1 has been shown to bind di-acetylated GATA1 [248]. The lack of chromatin binding displayed by these bromodomains was surprising, and it was decided to investigate their importance to transcription – a key cellular function of BET proteins. Work with a NF-κB luciferase assay showed that, with regards to NF-κB signalling, the BD2 of BRD4 is equally essential to transcription as its BD1. Additionally, the BD2 of BRD2 and BRDT were shown to play a role in transcription despite no observable chromatin binding. These findings collectively show that, although they do not bind chromatin, these BD2s and the non-histone proteins (NHPs) they interact with are involved in transcription.

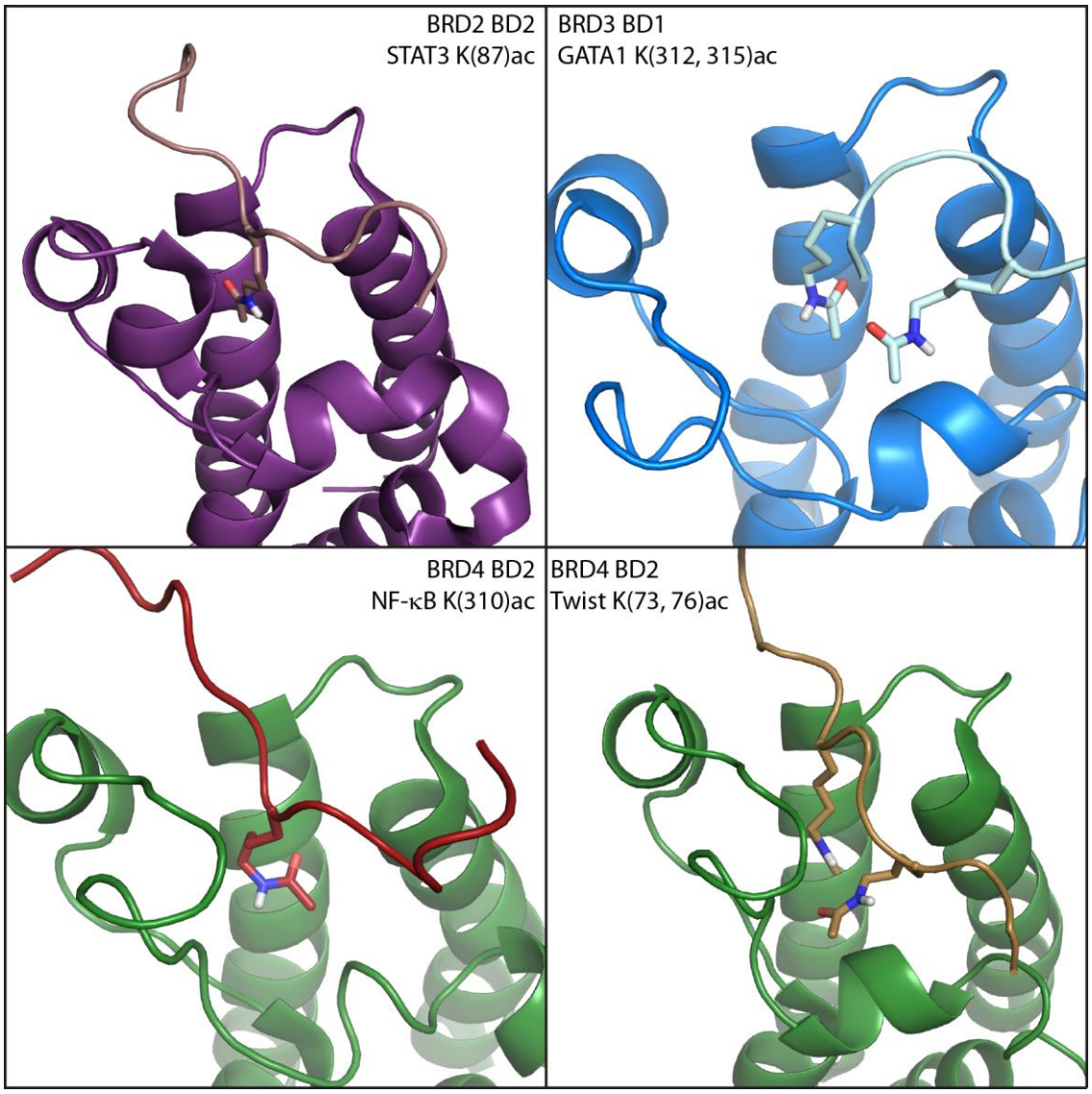


Figure 6.2. BET bromodomain : non-histone protein interactions.

Structures of BET bromodomains in complex with acetylated non-histone protein peptides, with acetylated lysine residues highlighted. PDB codes: BRD2 BD2 / STAT3 (5U5S), BRD3 BD1 / GATA1 (2L5E), BRD4 BD2 NF-κB (2LSP), BRD4 BD2 / Twist (2MJV).

Taking these observations and the literature we discussed we believe we have provided strong evidence for a new model for BET protein function. Older models often feature BET proteins binding acetylated chromatin through both the BD1 and the BD2 (sometimes the same pattern of Kac marks, sometimes with the BD1/BD2 binding different marks or different histone tails), while additional domains (e.g. ETD, CTD, atypical kinase) work to activate RNA POL II and initiate the elongation phase of transcription and up-regulate the expression of downstream genes (figure 6.3A) [47, 249]. One issue with these models is that, given the abundance of acetylated chromatin and gene promoters, these BET proteins would be very non-selective and

would affect a large number of downstream genes. Thus, they would represent a very 'blunt' tool for regulating cellular processes. Selectivity could be acquired through the BD1/BD2 recognising different Kac combinations, but evidence for this at the peptide level is mixed, and the BD2s we now know play little role in chromatin recognition. In this new model (best explained by [90]) the BD2 binds acetylated NHPs, like NF-κB and Twist, which then direct the complex to specific enhancers/promoters and hence to more specialised sets of downstream genes (figure 6.3B). Hence the expression of downstream genes can be regulated through: chromatin acetylation, BET expression, NHP expression and NHP acetylation. This BD2:NHP interaction may only matter for a narrow subset of BET-regulated genes, but given our data it is likely to be important in at least some circumstances. Interestingly, although BRD3 is also shown to bind an NHP it does so through its BD1 [248], so it may act with the functions of its BD1/BD2 occasionally reversed.

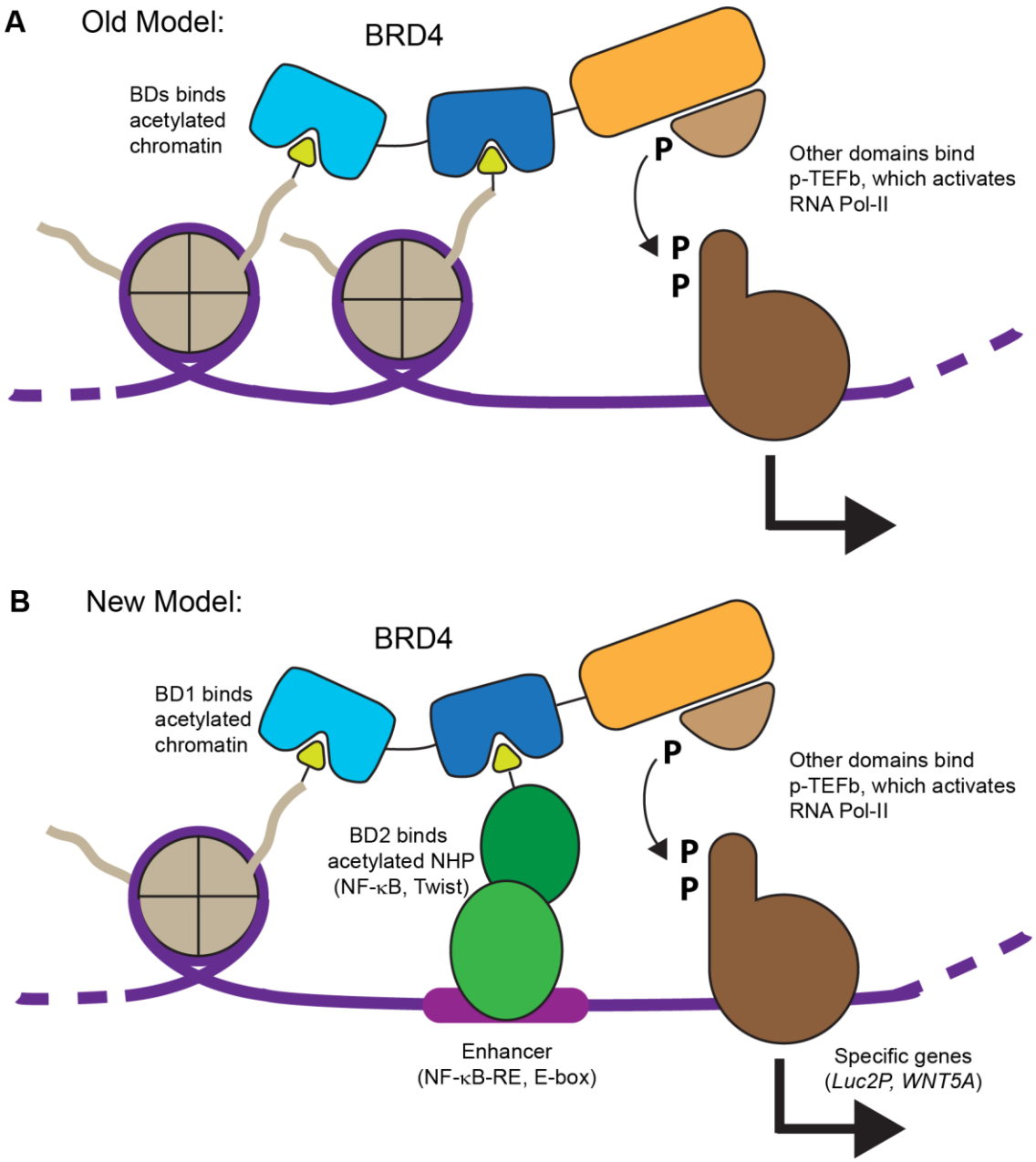


Figure 6.3. A revised model for BET protein function

Schematic representation of BET protein mode of action, using BRD4 as an example. In older BET protein models (A) both the BD1 and BD2 bind acetylated histone tails, directing additional domains to recruit and activate additional protein complexes, such as P-TEFb, which induce transcription. In the new model (B) BRD4 BD1 binds to acetylated histone tails. BD2 binds (in some cases) to acetylated non-histone proteins such as NF-κB, which themselves recognise and bind specific regulatory DNA sequences such as NF-κB-REs. Additional domains still induce transcription, but now of more specific sets of genes.

This model has several therapeutic implications. Firstly, regarding the displacement of BET proteins from chromatin, inhibition of BD1s alone is sufficient for the displacement of all BET proteins save BRD3. Conversely, BD2 inhibitors will displace only BRD3 (and perhaps BRD4 less so), which could explain differing cellular and (potentially) therapeutic impacts of pan-selective, BD1-selective and BD2-selective BET inhibitors. Full BRD3 displacement, meanwhile, requires inhibition of both BDs.

Taking a broader view, one can theorise that pan-BET inhibition, or BD1-selective inhibition will, generally, displace a given BET protein from chromatin and thus disturb the expression of all downstream genes. This mass perturbation will lead to a number of cellular and therapeutic effects, both positive and negative. BD2-selective inhibition meanwhile, will disrupt only BD2:ac-NHP interactions and hence only alter the expression of a narrower range of downstream genes (figure 6.4). This suggests that BD2-selective inhibition may be more successful in a clinical setting, as there is a chance that positive impacts (e.g. reduced cancer proliferation) may be maintained while negative on-target toxicity may be avoided (although the opposite is just as possible). Further experiments with said inhibitors will be needed to confirm this, as it is not clear what effects are dependent on BD2:ac-NHP interactions. This theory may explain the relative clinical success of BD2-selective RVX-208 (the only BET inhibitor to reach phase three clinical trials to date) [111], which has also been shown to alter the expression of far fewer genes than pan-selective (+)-JQ1 [116]. A growing awareness of this possibility may explain the recent development of BD2-selective inhibitors by AbbVie (ABBV-744) [120] and GSK (GSK268 and GSK340) [121]. ABBV7-744 has been shown to be less effective against a number of non-AML cell-lines [119], which this theory would explain as due to BET-regulated oncogenes relevant to these cancers not being reliant on BD2:ac-NHP interactions.

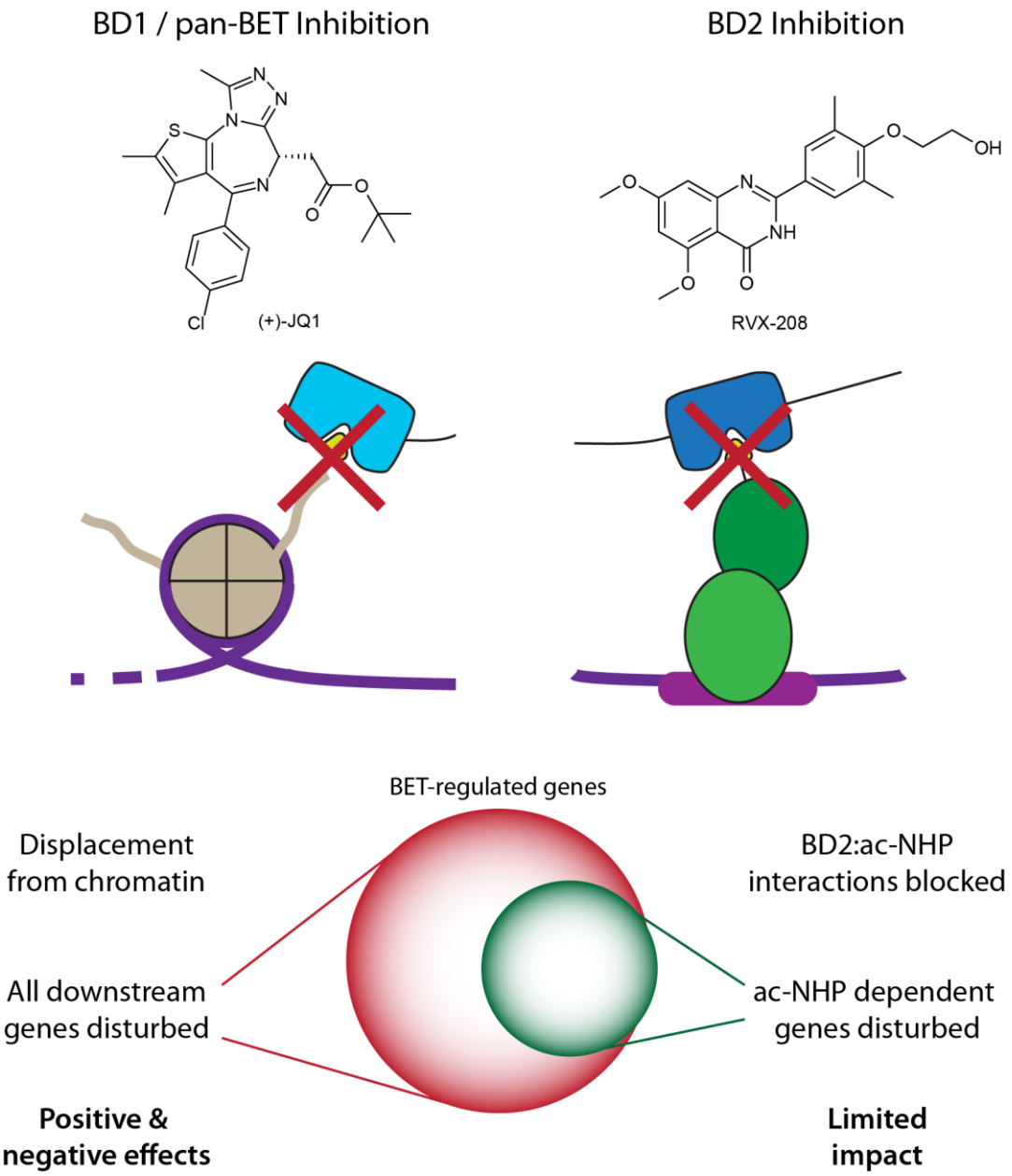


Figure 6.4. Possible therapeutic implications of pan-BET/BD1 inhibition vs BD2-selective inhibition. BD1 or pan-selective BET inhibitors (such as (+)-JQ1) will displace BET proteins from chromatin, disturbing all downstream genes and producing many positive and negative therapeutic effects. In contrast, BD2-selective inhibitors (such as RVX-208) will, mostly, not displace BET proteins from chromatin, and only disturb genes dependent on BD2:ac-NHP interactions. Hence BD2-selective inhibitors will generate a more limited therapeutic impact but potentially bear improved safety profiles.

6.3 – Future Development & Applications of BET Bump-&-Hole System:

The optimised BET bromodomain bump-&-hole system has been used in 2 cellular assays (FRAP and NF-κB luciferase) to study the BET proteins. Both of these assays have utilised transient overexpression of WT and L/V BET protein constructs,

and hence have been fairly easy to run. Transient construct overexpression has its limits, and could not be used to study the importance of BET BDs to cell viability, but there is likely several more experiments that could be run in such a fashion. Suitable overexpression assays will require one of two properties. The FRAP assays works very well as what is measured (GFP fluorescence and recovery) is directly generated by the overexpressed constructs. The luciferase assay meanwhile, uses a measurement that is affected by endogenous BET (and other) proteins. This measurement, however, is clearly increased by BET overexpression and hence its inhibition by bumped compounds can be assessed. As future bump-&-hole experiments relying on BET overexpression will need these properties options are limited. Luciferase gene reporter assays, looking at genes/signalling other than NF-κB, should be simple enough to develop. One problem with these assays is that comparisons are difficult to make between BET proteins, and are best used to compare the BD1s and BD2s.

Other, comparably straightforward, uses of the bump-&-hole system could be experiments to compare the importance of the BD1 and BD2 to the HAT and PK activity of BRD4. BRD4's HAT activity was shown with purified recombinant protein and transient overexpression [54], and it should be possible to replicate this work with L/V constructs. The atypical kinase activity of BRD4 was also assessed with purified protein and transient expression, and these experiments have already been performed with (deletion) mutant constructs [52]. As NMC is fundamentally caused by BRD4/3-NUT fusion proteins aspects of this disease can be studied through transient expression of BRD4/3-NUT constructs [64, 76, 250, 251]. These transient expression experiments can be repeated with L/V mutant constructs and 9-ME-1 treatments to better understand the role of different BET BDs in this disease. For example, such experiments with L/V-mutant BRD4/3-NUT constructs could determine the viability of treating NMC with BD1 or BD2-selecive BET inhibitors. As treatment of NMC requires displacement of BRD4/3-NUT from chromatin our FRAP experiments suggest that BRD4-NUT complexes could be treated with BD1-selective inhibitors, while BRD3-NUT would require both BD1 and BD2 blockade for effective displacement.

Many biological questions will not be answerable through the use of transient overexpression of mutant BET proteins, and will require the use of the CRISPR/Cas9 technology to create cell-lines expressing properly-regulated and spliced mutant BET

proteins from the same endogenous locus as the parental gene, with no remaining WT protein. Performing such knock-ins are possible with current technology, but as seen in chapter 5 it is a difficult and time-consuming process. The ideal mutant cell-line will have simply a small knock-in creating a few small changes in DNA sequence, but these knock-ins will be very difficult to detect. Alternatively one could knock-in a gene cassette expressing a full-length mutant BET protein, alongside a GFP or other easily identifiable tag; but this mutant gene will not be regulated or spliced appropriately and the endogenous WT gene will need to be dealt with. Hopefully, advances in CRISPR/Cas9 technology and optimisation of knock-in methodology will allow for generation of cell-lines containing L/V mutants of different BET BDs. These cell-lines, and the 9-ME-1 probe, can then be used by the Alessio Ciulli lab, sent to collaborators or distributed via third parties (such as Tocris [252], chemicalprobes.org [253] and cell-line databases [254] or other commercial vendors).

Once successfully implemented into a variety of cells the bump-&-hole system can be used to answer a number of pressing biological questions. The most immediate efforts should likely focus on, essentially, replicating observations previously made with BET inhibitors in a BD-specific fashion. Combining the mutant cell-lines with cytotoxicity/cell-viability assays like CellTiter-glo can be used to determine the BET BDs that are the best targets for different cancers. For example, cell-lines with BRD4 mutated may be more vulnerable to 9-ME-1 treatment, while experiments with ABBV-744 suggest several cancer types will not be affected by 9-ME-1 if only BD2s are mutated [119]. Such experiments can also be used to address the suspicion that the anti-cancer effects of BET inhibitors is reliant on the inhibition of multiple BET proteins.

Other advantageous effects of BET inhibition can be tested in a similar fashion. Implementing the L/V mutation into PBMC cells would allow for recreating GSK's cytokine release assay in a BD-specific fashion, and identify the BET proteins/BDs that are suitable targets for BET-based inflammation suppression. This would not be applicable to the similar whole-blood assay used in conjunction. In a similar fashion, mutant HepG2 cells can be used to determine which BET proteins/BDs are necessary for the ApoA1 up-regulation observed by GSK [99] and Resverlogix [100]. This would aid the development of BET inhibitors for cardiovascular indications.

This general approach could also be utilised for investigating and dissecting the unwanted effects of BET inhibition. BET inhibition has been shown to result in the up-regulation of BET proteins, limiting the efficacy of BET inhibitors against cancers. The exact fashion in which this occurs has not been determined, i.e. does inhibition of any one protein up-regulate all BET proteins, or just that specific protein; or is all up-regulation caused only be inhibition of one or two specific BET proteins? Recent data suggests that induced BET protein degradation and pan-inhibition leads to feedback upregulation of BRD2 specifically [255]. By tracking the expression of BET proteins in mutant cell-lines treated with 9-ME-1 this feedback system could be properly characterised. Perhaps the major on-target toxicity of BET inhibitors is thrombocytopenia, the depletion of platelets. The inhibition of which BET proteins/BDs contributes to thrombocytopenia could be addressed using one recent method, however this would require the insertion of the L/V mutation into an entire animal (rat) and is likely impractical [107]. Addressing this question would ideally rely on a cell-line based assay for measuring, or predicting, platelet depletion, meaning only one cell-line would need to be mutated. Additionally, most assays for measuring thrombocytopenia are diagnostic [256-258], and hence use patient sera, and focus on heparin-induced thrombocytopenia. It is possible that a single cell-line based thrombocytopenia assay could be developed using megakaryocyte cells [259], and the measurement of platelet production or certain biomarkers. Finally, it may be that one or more pharmaceutical companies has developed such an assay, for use in developing BET inhibitors, and has not yet disclosed it.

Following these experiments more advanced techniques could be paired with the bump-&-hole system. Over the years many papers have paired proteomic and transcriptomic techniques to measure changes in gene expression brought about by BET inhibitors [77, 116, 260, 261]. These have identified many downstream genes, and suggested many broader biological processes regulated by the BET proteins. Performing these experiments with mutant cell-lines and 9-ME-1 will allow identification of downstream genes on a protein and BD-specific level. This will help show to what extent differing BET proteins are involved in different biological processes, while also identifying genes regulated by multiple BET proteins.

Chapter 7
Conclusions

Conclusions

The aim of this work was to advance an existing, promising but unoptimised, bump-&-hole system for engineering allele-selective inhibition of specific BET bromodomains, and take it up to a stage where it can be applied to investigate the roles of individual BET proteins and their individual bromodomains.

The original L/A mutation, which allowed for selective inhibition but had reduced function, was replaced with an L/V mutation. The L/V mutation was thoroughly characterised biophysically, structurally and functionally in cells. *In vitro*, the introduced mutation was shown to have little to no impact on the stability, structure and affinity of recombinant BET BD proteins for acetylated histone peptides. Full-length L/V BET proteins were then tested in cells and shown to be able to bind chromatin and drive the transcription of downstream genes, to a much better extent than L/A BET proteins (figure 7.1).

While the original bump-&-hole system utilised an ethyl-bumped compound, ET, the switch to a new mutation required a fresh round of compound evaluation. By testing an expanded range of bumps and other chemical modifications, and purifying the active *2R,3S* enantiomer, improved mutant-selective inhibitors were developed. The bumped compound that emerged as best-in-class, 9-ME-1 showed increased overall selectivity and strong DMPK properties, albeit not being quite as potent against L/V BDs as ET was against L/A BDs (figure 7.1).

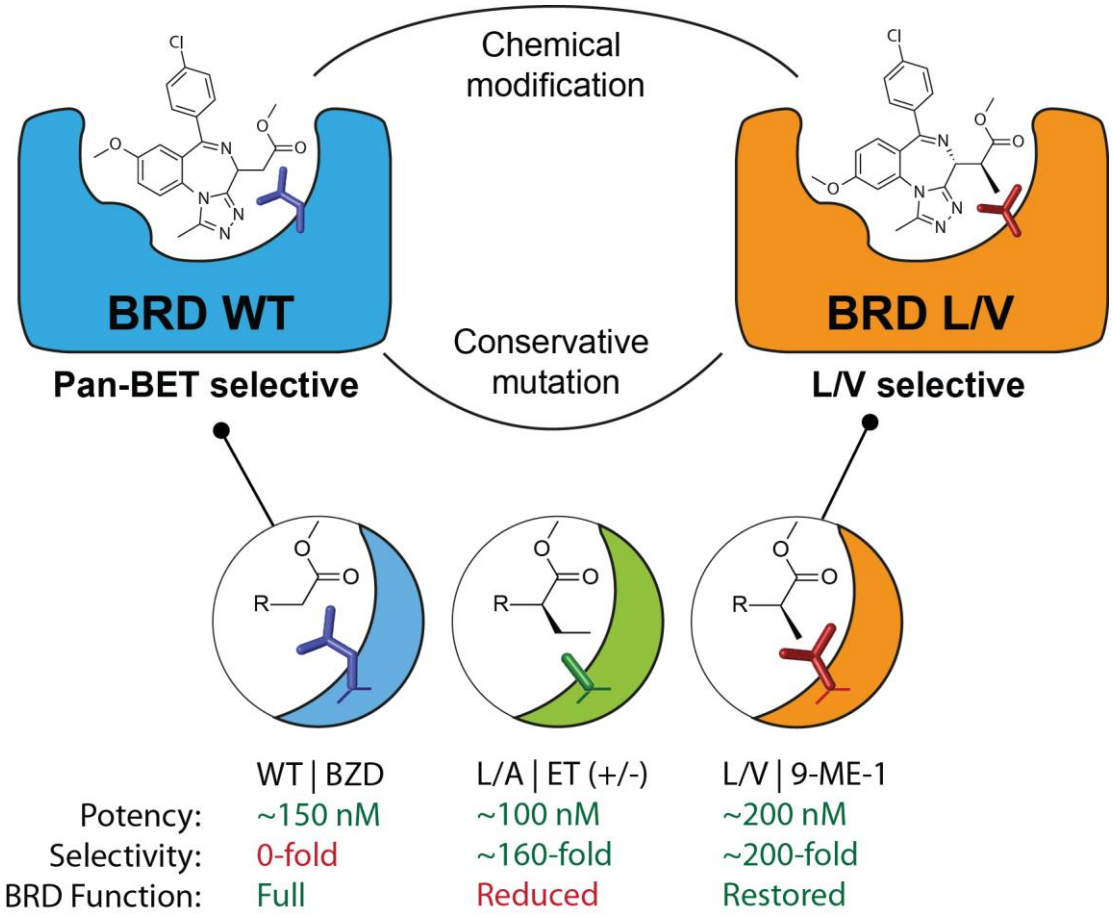


Figure 7.1. Optimisation of BET bromodomain 'bump-&-hole' system.

The optimised L/V | 9-ME-1 bump-&-hole system was validated in cells and used to investigate the importance of individual BET BDs to the BET proteins' ability to bind chromatin and drive the transcription of NF-κB-regulated genes. These experiments showed that chromatin recognition and binding is primarily driven by the BD1s of the BET proteins, with the BD2s played a less prominent role to varying degrees depending on the family member. Specifically, in the case of BRD2 and BRDT the BD2 was found to play no observable role in chromatin binding, and BRD4 BD2 played a very small role too. In contrast BRD3 was seen to rely on both BDs to achieve full chromatin recognition. Experiments on transcription, meanwhile, showed that for at least some downstream genes the BD2s still play a role in transcription. These results, combined with recent observations from the literature, were used as evidence for a new model of BET protein function. In this model the BET BD2s do not necessarily bind acetylated chromatin and, for certain downstream genes, they act through binding acetylated NHPs, which direct the BET protein and associated transcription machinery to more specific sets of genes.

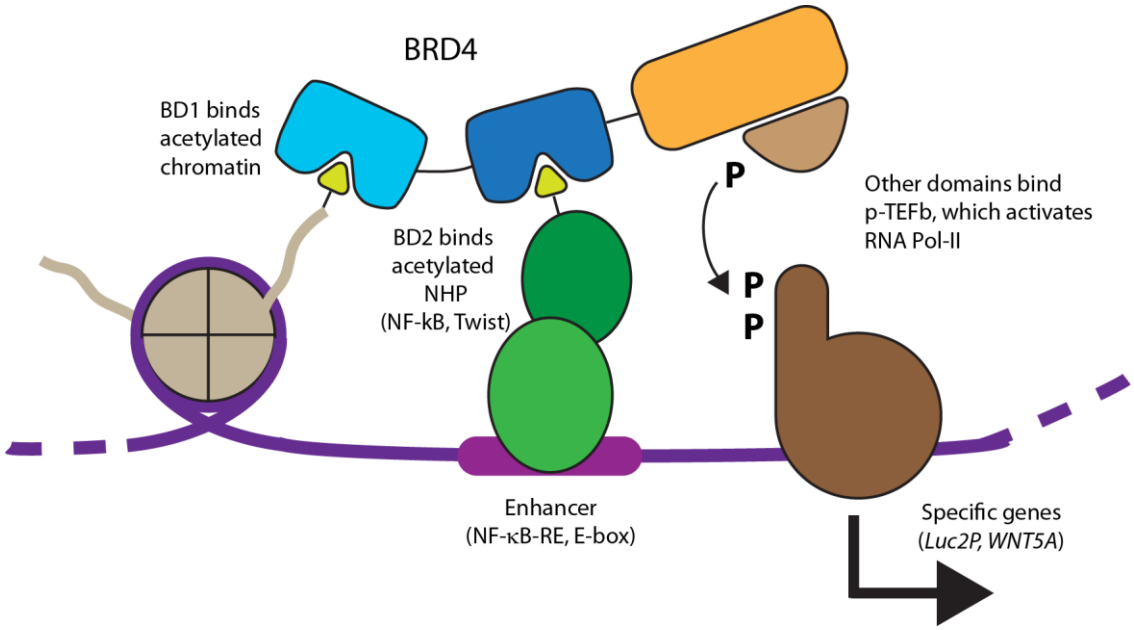


Figure 7.2. A model for BET protein function

Schematic representation of BET protein mode of action, using BRD4 as an example. BRD4 BD1 binds to acetylated histone tails. BD2 binds (in some cases) to acetylated non-histone proteins such as NF-κB, which themselves recognise and bind specific regulatory DNA sequences such as NF-κB-REs. Additional BRD4 domains bind, recruit and activate additional protein complexes, such as p-TEFb, which induce transcription.

Finally, it was recognised that to enable more sophisticated investigations into BET protein biology would require the introduction of the L/V mutation into endogenous BET genes in WT cell-lines. A method for inserting said mutation through the CRISPR/Cas9 gene editing system was developed, and an attempt was made to knock-in the L/V mutation onto BRD4 BD1 in A549 cells. This attempt has so far not been shown to be successful, but many lessons have been learnt that will aid in future, optimised knock-in attempts.

Future Directions

The highest priority, regarding future development and use of the BET bump-&-hole system, is the generation of L/V-mutant cell-lines as was attempted in chapter 5. The generation of apparent BRD4 KO cells suggests that existing gRNAs are suitable for targeting BRD4 BD1. Hence the focus should be on maximising the percentage of cells (following transfection) that have been correctly knocked-in, and optimising their detection. The use of fluorophore-labelled gRNA strands will allow for selection of transfected cells through FACS, which should greatly reduce the number of clones that must be screened to find successful knock-ins. If improved assays for detecting knock-ins can be developed, perhaps taking inspiration from research into SNP genotyping, then it will also become easier to optimise transfection conditions, to improve both the transfection efficiency and the knock-in efficiency. The decision may also be taken, following validation and characterisation of the apparent BRD4 KOs, on the value of generating heterozygous KO/KI cell-lines. With so much research being conducted on CRISPR/Cas9 gene editing there is also a reasonable likelihood that in the near future advances will be made in how CRISPR-mediated knock-ins are conducted, that may aid our own efforts in this direction.

Following generation of BRD4 L/V KI cell-lines (A549 LV/WT, WT/LV and LV/LV) the first step (following genetic validation of said knock-in) should be assessing the viability of said cell-lines and determining whether the effects of the L/V mutation on BRD4 has affected the cells in a more general fashion. This could be quite easily down through the use of the CellTiter-Glo and crystal-violet colony-formation assays, as was done for the potential KO cell-lines in chapter 5. These same assays could then be used to measure the vulnerability of these cell-lines to 9-ME-1 (or other bumped compound) treatment. Such experiments would show which (if any) BRD4 BDs are important to A549 cell-viability, for example is BRD4 BD1 inhibition sufficient to kill A549, or are inhibiting both BRD4 BDs necessary, or do multiple BET proteins need to be targeted? The next step would be to utilise proteomic and transcriptomic techniques, such as RNAseq and TBT-labelled LC-MS/MS to determine which downstream genes are up or down-regulated by BRD4 inhibition versus (+)-JQ1 pan-BET inhibition. Comparisons could also be made to datasets utilising BRD4 knock-out or RNAi, to compare the importance of BRD4's BDs and its other domains.

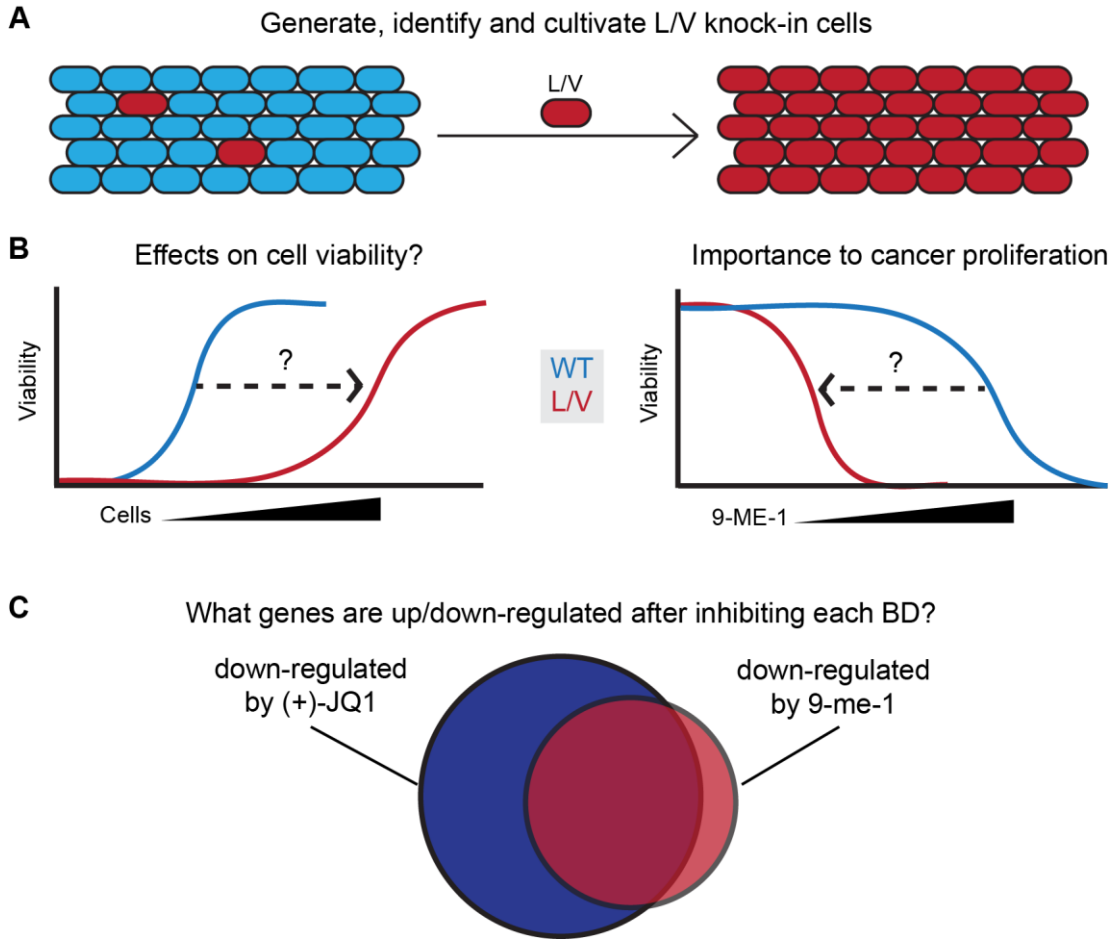


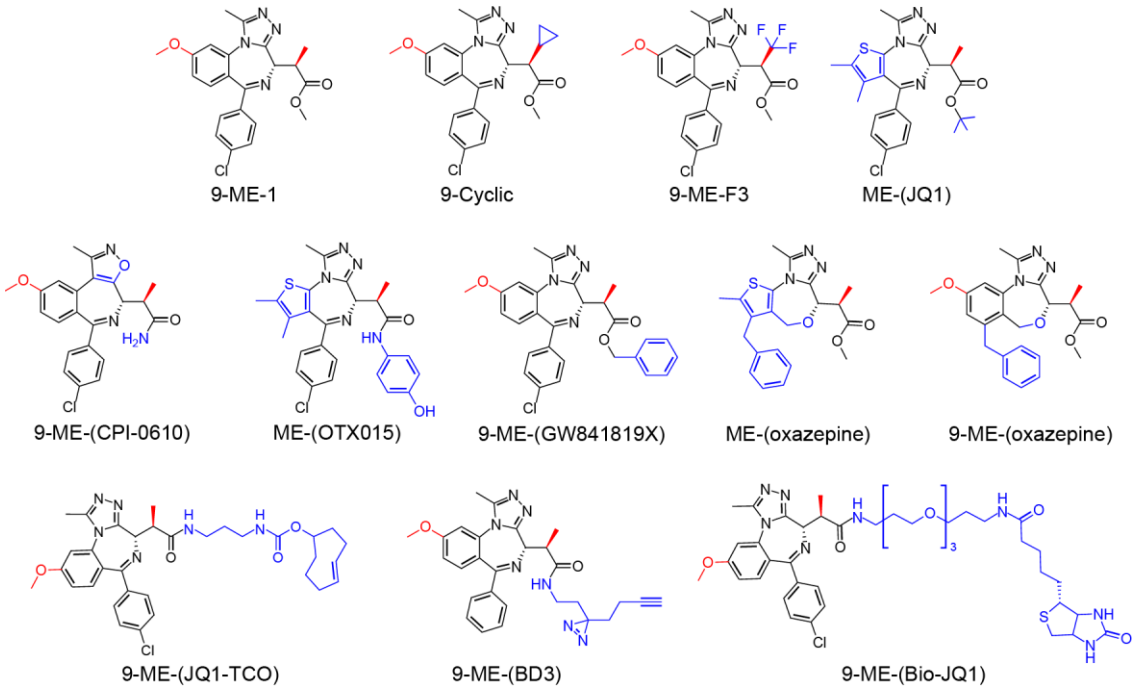
Figure 7.3. Future experiments with L/V knocked-in cells.

A) The L/V mutation (red) should be knocked-into WT cells (blue), and successful knock-ins identified and cultivated. B) Assays such as CellTiter-Glo and colony-formation measurement can be used to (left) determine if the L/V mutation reduces the viability of knock-in cells, and (right) determine which BDs are necessary for cell proliferation. C) Transcriptomic and proteomics can determine which genes are regulated by specific BET proteins and/or BDs.

Through our intensive SAR work we have settled on 9-ME-1 as the 'best' bumped compound, but additional chemistry could be done to produce an even better L/V-selective inhibitor. The introduction of small carbon cycles, such as cyclopropane, into bumps may boost L/V affinity by limiting the freedom of the bump and the corresponding entropic penalty to binding. Furthermore, it may reduce WT binding by removing the ability of the bump to twist into an orientation that relieves the steric clash. The use of double C=C bonds has been trialled with the AL bump, which can be used to modulate rotational freedom, but also reduces the length of the bump. Further use of C=C double bonds may help produce improved bumped compounds. One final bump modification to test is the use bioisosteres. Replacing the methyl bump

of 9-ME with a CF_3 group may be a way of fine-tuning the size of the bump, as the resulting bump would occupy a slightly larger space than a simple methyl group. This modification will also effect the electronic properties of the compound, which could have a differential impact on binding affinity between WT and L/V BDs. Such modifications could also alter DMPK properties and van der Waals interactions with the L/V cavity.

Many BET inhibitors are based on the basic benzodiazepine core but use structural variations such as amide side-groups, thiophene instead of methoxy-phenol rings and even alternate Kac mimics. Bumped analogues of these alternate scaffolds could be trialled as L/V-selective inhibitors. Recently a super-potent PROTAC compound was developed, using an oxazepine inhibitor that acted as a (+)-JQ1 mimic [262]. These compounds could also be modified to contain our L/V-selective bumps, and their enhanced affinity and hydrophobicity may be lead to bumped compounds with improved selectivity or cellular potency.



Scheme 7.1. Potential Future Bumped Compounds.

Chemical structure of 9-ME and potential future bumped compounds utilising new cyclic or fluoride-containing bumps, alternate scaffolds or linked to functional groups. L/V-selective modifications within 9-ME highlighted in red. Potential future modifications highlighted in blue.

9-ME-1 could also be modified to produce a set of functional derivatives, for use in protein pulldowns or labelling. 9-ME could be modified to develop L/V-selective analogues of the previously discussed Bio-JQ1 [86], photo-cross-linkers [87, 88] and click-reactive probes [89]. These compounds could then be used to pulldown or modify specific BET proteins, in advanced chemical genetics studies to investigate their protein binding partners or genome localisation.

In this thesis a system has been developed allowing for highly-selective chemical inhibition of the BET bromodomains – a protein family of great interest to epigenetics and a range of clinical indications. This system has already been used to better understand how these proteins function on an immediate, molecular level and with further work and the development of mutant cell-lines we will be in a position to answer a great number of questions on the BET proteins. This knowledge could then better inform future development of BET-targeting therapeutics.

Chapter 8
Materials & Methods

8.1 – Plasmids

Single BET bromodomain constructs (WT, L/V and L/A) were encoded in pNIC28-Bsa4 (KanR) plasmids. Constructs are as follows: BRD2 BD1 (K71-K176), BRD2 BD2 (G344-D455), BRD3 BD1 (P24-E144), BRD3 BD2 (G306- P416), BRD4 BD1 (N44-E168), BRD4 BD2 (K333-E460), BRDT BD1 (N21-E137) and BRDT BD2 (S257-M361). Uniprot accession numbers: BRD2 – P25440; BRD3 – Q15059; BRD4 – O60885; BRDT - Q58F21. WT DNA and protein sequences can be found in [32] and details of the site-directed-mutation can be found in table 8.1. Single BET BD constructs possess an N-terminal His6-tag and TEV protease cleavage site. WT plasmids were provided by the Oxford Structural Genomics Consortium. L/A mutants created by Terrence Kwan and Enrique Lin-Shiao, L/V mutants created by Dr Kwok-Ho Chan and Teresa Cardote. These plasmids were used for purification of single-bromodomain constructs from transformed *E. coli* cells, for use in assays and X-ray crystallography.

Full-length BET proteins (WT, L/V and L/A) with N-terminal eGFP-tag were encoded in pcDNA5 FRT/TO (AmpR) plasmids (pcDNA6.2 for BRDT). Constructs are as follows: BRD2 (M1-G801), BRD3 (M1-E726), BRD4 (M1-F1362) and BRDT (M1-D951). The GFP-BRDT construct also contained a C-terminal V5 tag. Details of site-directed-mutation can be found in table 8.1. pcDNA5 FRT/TO-GFP vector, and BRD2 (NM_001113182.2 [263]) and BRD3 cDNA (NM_007371.3 [264]) clones were provided by Dr Mark Peggie (DSTT, University of Dundee). BRD4 cDNA (NM_058243.2 [214]) was provided by the Oxford Structural Genomics Consortium, contained within a pEGFP-C1 plasmid. BET proteins were then cloned into pcDNA5 FRT/TO-GFP plasmid. WT GFP-BRDT plasmids purchased from Addgene (65381), originally produced by the Kyle Miller lab [193] using BRDT cDNA (NM_001242806.2 [265]). Cloning and site-directed-mutagenesis performed by Annica Pschibul (BRD4 L/A), Teresa Cardote (BRD4 L/V), Lars van Beurden (BRD3 L/V) and Andrew Runcie (BRD2 and BRDT L/V). These plasmids were used for expression in mammalian cells, primarily for FRAP & luciferase assays, but could also be used for CETSA or western blots (using αBET & αGFP antibodies).

The NF-κB-RE luciferase reporter was encoded in a pBABE plasmids, following cloning from a pGL4.32 vector. This reporter consists of one *Luc2P* gene, with C-terminal PEST tag, controlled by a miniP promoter and 5 copies of an NF-κB response element. This plasmids was acquired from Dr Mark Peggie (DSTT, University of Dundee). This plasmid

was used for expression in mammalian cells for the NF-κB luciferase assay. BRD4 (and to a lesser extent other BET proteins) required for NF-κB signalling. Expression can be induced with TNF-α treatment, but this is not always necessary/advisable (i.e. HEK293T cells).

Single point mutations were introduced using QuickChange II Site directed Mutagenesis Kit (Agilent). Primers were designed following the recommendations from the manufacturer. The polymerase chain reaction (PCR) was performed on a 2720 Thermal Cycler (Applied Biosystems®). Upon digestion of the parental DNA strands by DpnI restriction enzyme, the PCR product was transformed into competent E. coli DH5α cells and grown on lysogeny broth (LB) agar plates containing corresponding selection antibiotics at 37°C for 16h. Single colonies were then picked from the agar plates and grown for 12h in 10 mL of LB medium containing selection antibiotics. Extracted and purified plasmid DNA were then sequenced to confirm the presence of the desired mutation.

Bromodomain	WT Leucine codon	L/A codon	L/V codon
BRD2 BD1	CTG	GCG	GTG
BRD2 BD2	CTG	GCG	GTG
BRD3 BD1	CTG	GCG	GTG
BRD3 BD2	CTG	CTG	GTG
BRD4 BD1	CTC	GCC	GTC
BRD4 BD2	CTA	GCA	GTA
BRDT BD1	TTG	GCG	GTG
BRDT BD2	CTC	GCC	GTC

Table 8.1 Leucine mutant codons.

8.2 – Protein Expression & Purification

BL21 (DE3) *E. coli* cells were transformed with bromodomain-expressing pNIC plasmids using heat shock. Cell cultures were grown at 37°C and 200 rpm in LB media with 50 µg/ml kanamycin. A starter culture was incubated until saturation, then diluted 100-fold in fresh media and grown until reaching an optical density of 2.0 (OD_{600}). The temperature of the culture was decreased to 18°C, and protein production was induced overnight with 0.4 mM IPTG. Cells were harvested the next day by centrifugation (4500 g for 20 minutes at 4°C, swing-out JS-4.2 rotor on a Beckman J6-MC centrifuge) and stored as pellets at -20°C.

Cell pellets were resuspended in lysis buffer (50 mM HEPES, 500 mM NaCl, 10 mM imidazole & 2 mM 2-mercaptoethanol pH7.5) and treated with one complete protease inhibitor (Roche) tablet. Cells were lysed using a Stanstead pressure cell homogenizer and the lysate centrifuged at 50 000 g for 1 hour at 4°C (fixed-angle JA-25.50 rotor in Beckman Avanti J-25 centrifuge) and the supernatant transferred and passed through a 0.45 µm filter.

The lysate was purified using metal ion affinity chromatography, with a His Trap 5 ml Ni sepharose column (GE Healthcare) on an AKTApure™ system (GE Healthcare). The column was washed with 30 ml of lysis buffer, and the protein was eluted using 30 ml of elution buffer (50 mM HEPES, 500 mM NaCl, 250 mM imidazole and 2 mM 2-mercaptoethanol at pH 7.5). The Ni column elution was concentrated to 2 ml using a Vivaspin 20, 10 000 MWCO (Sartorius) before further purification using size exclusion chromatography. Concentrated Ni column elution was passed through a Superdex 75 16/60 Hiload column (GE healthcare) on an AKTApure™ system, using gel filtration buffer (20 mM HEPES, 150 mM NaCl and 1 mM DTT at pH 7.5). Desired fractions were pooled, concentrated and aliquoted before being flash frozen and stored at -80°C.

8.3 – Differential Scanning Fluorimetry (DSF / Thermal Shift)

DSF experiments were performed on a BioRad CFXconnect machine, using clear 96-well plates. Protein constructs were tested at 6 µM, with 2.5X SYPRO orange (Invitrogen), in 40 µl of buffer (10 mM HEPES, 100 mM NaCl, pH7.5). Samples underwent a heat cycle from 20°C to 95°C, heated at a rate of 1°C every minute. Plates

were read at 1 minute intervals. Each sample was tested in triplicate. Data was analysed as recommended by Niesen et al. [169], using GraphPad Prism 6 and the SGC's DSF Analysis 3.02 spreadsheet.

DSF can be used to screen inhibitors. This follows the same protocol but with the addition of ~1 µM compound and associated DMSO.

8.4 – BioLayer Interferometry (BLI)

All BLI experiments were carried out on a OctetRed 384 instrument (ForteBio), at 25°C and in 20 mM HEPES, 150 mM NaCl pH7.5 buffer. A histone peptide library, possessing both single and multiple lysine acetylation marks, was obtained from Alta Bioscience Ltd. Peptides were 20 residues long and were biotinylated on the C-terminal (with aminohexanoic linker). Streptavidin-coated BLI tips (ForteBio) were loaded in 100 µl 5 µM peptide (or 10 µg/ml biocytin) over 10 minutes. For the assay bromodomain constructs were kept at 20 µM in 100 µl buffer in black 384-well plates, agitated at 1000 rpm. Each assay involved exposing peptide-loaded BLI tips to buffer for 120 s to determine baseline signal, 240 s in protein to measure association and finally 120 s in buffer to measure dissociation. Raw data was then analysed using the ForteBio software, to account for background signal and non-specific binding.

8.5 – Isothermal Titration Calorimetry (ITC)

ITC titrations were performed on an ITC200 instrument (MicroCal™, GE Healthcare). Proteins, peptides and compounds were all dissolved in ITC buffer (20 mM HEPES, 100 mM NaCl, pH7.5), with protein samples buffer-exchanged through dialysis using D-tubes (Millipore). Each ITC titrations consisted of 20 titrations: 1 initial injection of 0.4 µl over 0.8 s, followed by 19 injections of 2 µl over 4 s, at 2 minute intervals. Data was analysed using the Microcal LLC ITC200 Origin data analysis software, using a single binding site model, to determine binding values such as K_d and enthalpy of binding.

Peptide titrations were carried out at 15°C, with 2 mM of acetylated histone peptide titrated against 50-150 µM bromodomain. Ligand titrations were performed at 30°C, with 250 µM active ligand titrated against 25 µM bromodomains, with a final DMSO

concentration of 2.5%. Reverse ligand titrations were also performed at 30°C, with 150 µM protein titrated against 15 µM ligand.

8.6 – X-ray Crystallography

Purified BRD2 BD2 L/V (8 mg/ml) was combined with 2-4 mM compound to form a protein:ligand complex. This complex was mixed with precipitant solution at a 2:1 ratio in a sitting-drop vapour diffusion format. A variety of crystallisation conditions were used, with the majority using 0.1 M Tris, pH 7-9, 45-60% pentaerythiol propoxylate (5/4 PO/OH) with or without 0.2 M imidazole. Crystals appeared within hours and were fully grown after 2-3 days.

Diffraction data were collected either in-house with a Rigaku M007H5 X-ray generator, with Varimax Cu-VH5 optics and a Saturn 944HG+ CCD using a wavelength of 1.5418 Å; or Diamond Light Source beamline I04-1 using a Pilatus 6M-F detector at a wavelength of 0.92819 Å and at a temperature of 100K.

Indexing and integration of reflections was performed either with XDS (with the XDSGUI interface) [266]or MOSFLM [267], and scaling and merging with AIMLESS [268] in CCP4i [269]. To solve the phase problem the molecular replacement method was used with the programs MOLREP [270] and Phaser [271], using search models derived from the coordinates of BRD2 BD2 WT (PDB 2DVV). The initial model was refined iteratively using PHENIX [272, 273] and COOT [274]. Ligand structures and restraints were generated using eLBOW [275], REEL [276] and PRODRG [277]. The MOLPROBITY [278] server was used to validate the geometry and steric clashes in the structures. The structure models have been deposited in the protein data bank (PDB) and data collection and refinement statistics are presented below. Figures were generated using PyMOL [279].

I generated crystals, processed data and refined structures for AL, ET-Am1, ME and 9-ME co-crystal structures. Dr Kwok-Ho Chan did the same for the *apo* structure and the ET, 9-ET, ME-Am1, 9-ME-Am1 and AL-tBu co-crystal structures. All in-house data collection, as well as PDB submissions, were performed by Dr Kwok-Ho Chan.

Structure	**Apo**	**ME**	**ET**	**AL**	**9-ME**	**9-ET**
PDB code	**5O38**	**5O39**	**5O3A**	**5O3B**	**503C**	**5O3D**
Data Collection						
Space group	P 21 21 2	P 21 21 2	I 41 2 2	I 222	P 2 21 21	P 21 21 2
Cell dimensions						
a, *b*, *c* (Å)	52.420, 71.409, 31.969	52.406 72.129 31.990	126.799, 126.799, 43.194	88.075 101.893 123.534	31.852 52.082 71.789	52.407 71.83 32.003
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	19.50 - 1.20 (1.243 - 1.20)	21.77-1.74 (1.802-1.74)	40.10 - 2.40 (2.486 - 2.40)	44.04 - 1.95 (2.02 - 1.95)	31.85 - 1.60 (1.657 - 1.60)	32.00 - 1.60 (1.657 - 1.60)
Rmerge	0.05918 (0.06509)	0.04883 (0.05565)	0.09961 (0.3312)	0.2827 (0.9697)	0.1111 (0.5627)	0.05236 (0.07519)
I/sigma(*I*)	17.14 (14.31)	43.16 (37.50)	17.11 (5.14)	10.53 (2.26)	7.67 (2.04)	32.70 (18.62)
$CC_{1/2}$	0.997 (0.998)	1 (0.999)	0.999 (0.984)	0.984 (0.535)	0.992 (0.755)	0.999 (0.994)
Completeness (%)	98.26 (93.37)	97.12 (79.28)	99.18 (100.00)	99.92 (99.98)	99.51 (99.56)	98.83 (94.76)
Redundancy	5.7 (5.7)	17.1 (15.3)	15.3 (14.6)	8.6 (7.8)	4.3 (4.2)	10.3 (7.2)
Refinement						
Resolution (Å)	19.50 - 1.20 (1.243 - 1.20)	21.77-1.74 (1.802-1.74)	40.10 - 2.40 (2.486 - 2.40)	44.04 - 1.95 (2.02 - 1.95)	31.85 - 1.60 (1.657 - 1.60)	32 .00 - 1.60 (1.657 - 1.60)
No. reflections	122513 (5072)	123746 (4452)	109065 (10146)	349867 (31395)	69717 (6561)	168336 (11018)
Unique no. of reflections	37657 (894)	12633 (291)	7114 (696)	40791 (4007)	16285 (1575)	16364 (1536)
Rwork	0.1178 (0.1068)	0.1443 (0.1466)	0.2055 (0.3278)	0.1972 (0.2710)	0.1615 (0.2197)	0.1483 (0.1324)
Rfree	0.1413 (0.1365)	0.1897 (0.2237)	0.2547 (0.3512)	0.2442 (0.3180)	0.1945 (0.2670)	0.1711 (0.1608)
No. atoms						
Protein	991	922	893	3613	918	948
Ligand/ion	31	40	31	128	40	41
Water	237	203	57	559	212	219
B factors						
Average	13.3	12.9	43.2	19.3	16.7	11.8
Protein	10.1	10.4	43.3	18.4	14.3	8.9
Ligand/ion	16.9	11.0	37.9	15.4	15.8	15.4
Water	26.0	24.9	44.4	25.9	27.6	24.0
R.m.s. deviations						
Bond lengths (Å)	0.005	0.007	0.011	0.009	0.007	0.009
Bond angles (°)	1.03	1.05	1.19	1.04	1.08	1.06

Table 8.2 X-ray crystallography data collection & refinement statistics 1

Structure	**ME-Am1**	**ET-Am1**	**AL-Am1**	**9-ME-Am1**	**AL-tBu**
PDB Code	**5O3E**	**5O3F**	**5O3G**	**5O3H**	**5O3I**
Data Collection					
Space group	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2	P 21 21 2
Cell dimensions					
a, *b*, *c* (Å)	52.177, 71.95, 31.917	51.965 71.548 31.704	73.74, 52.66, 63.12	52.143, 71.837, 31.97	53.12, 72.829, 32.359
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	35.98 - 1.40 (1.45 - 1.40)	42.05 - 1.75 (1.813 - 1.75)	25.41 - 1.85 (1.916 - 1.85)	35.92 - 1.40 (1.45 - 1.40)	14.31 - 1.20 (1.243 - 1.20)
Rmerge	0.1005 (0.2228)	0.1116 (0.1436)	0.07581 (0.3878)	0.06188 (0.1025)	0.1049 (0.5239)
I/sigma(*I*)	15.07 (4.57)	7.72 (6.60)	9.64 (3.00)	21.35 (7.92)	9.10 (3.25)
$CC_{1/2}$	0.998 (0.955)	0.992 (0.99)	0.997 (0.776)	0.999 (0.982)	0.994 (0.781)
Completeness (%)	99.77 (97.92)	99.66 (99.92)	95.19 (95.67)	98.72 (93.14)	99.19 (99.29)
Redundancy	9.6 (4.8)	3.9 (4.0)	3.9 (3.9)	7.5 (3.7)	5.8 (5.8)
Refinement					
Resolution (Å)	35.98 - 1.40 (1.45 - 1.40)	42.05 - 1.75 (1.813 - 1.75)	25.41 - 1.85 (1.916 - 1.85)	35.92 - 1.40 (1.45 - 1.40)	14.31 - 1.20 (1.243 - 1.20)
No. reflections	232876 (11299)	27589 (1075)	79692 (7845)	180745 (8305)	229769 (22505)
Unique no. of reflections	24337 (2352)	12439 (269)	20578 (2034)	24057 (2228)	39710 (3911)
Rwork	0.1315 (0.1401)	0.1721 (0.2604)	0.1888 (0.2460)	0.1477 (0.1678)	0.1557 (0.2044)
Rfree	0.1653 (0.1794)	0.2018 (0.2902)	0.2356 (0.2788)	0.1636 (0.1633)	0.1784 (0.2326)
No. atoms					
Protein	931	909	1818	932	939
Ligand/ion	41	42	66	41	45
Water	218	185	196	229	218
B factors					
Average	12.7	14.3	26.5	14.3	14.2
Protein	9.7	12.2	25.9	11.1	11.2
Ligand/ion	11.5	14.8	26.3	13.6	18.1
Water	26.0	24.7	32.0	27.5	26.2
R.m.s. deviations					
Bond lengths (Å)	0.006	0.006	0.014	0.011	0.019
Bond angles (°)	1.68	1.01	1.52	1.15	1.75

Table 8.3 X-ray crystallography data collection & refinement statistics 2

8.7 – AlphaLISA Assays

AlphaLISA assays were carried out in 384-well, light-grey Alphaplates (PerkinElmer), with an assay volume of 25 µl. To prepare each assay each component was diluted in AlphaLISA buffer (50 mM HEPES, 100 mM NaCl, 0.1% BSA, 0.02% CHAPS, pH7.5) to 5x its intended final concentration, and added at a volume of 5 µl, using a 12-channel Matrix P125 multi-channel pipette (Thermo). His-tagged protein, biotinylated probe/peptide and test compounds were added to the well, and incubated for 1 hour at room temperature. Then, in low-light conditions, acceptor beads were added and incubated for 1 hour, then donor beads were added and incubated for 1 hour. Alpha signal was measured in a Pherastar FS, using the AlphaLISA module. Dose-response curves were generated in Prism 6 (Graphpad), using the 'log(inhibitor) vs response – variable slope equation'.

Unless stated otherwise, the acceptor beads used were anti-His6 antibody-conjugated AlphaLISA acceptor beads (PerkinElmer) and streptavidin-coated Alphascreen donor beads (PerkinElmer), at a final concentration of 10 µg/ml. For the original, failed biotinylated-peptide displacement assay the final concentrations used were 200 nM BRD, 200 nM biotinylated H4 K(5,8)ac and 0.1% DMSO. For the successful, Bio-JQ1 assay final concentrations were 100 nM BRD, 5 nM Bio-JQ1 and 1% DMSO. For the SAR screen bromodomains and Bio-JQ1 were pre-mixed, aliquoted and stored at -80°C.

8.8 – DMPK

Compounds were prepared in-house and sent to the Kevin Read lab (DDU) for measurement of plasma half-life, intrinsic clearance and PAMPA permeability. Material requirements for these assays (combined) was 100µl of 10 mM compound in DMSO.

Plasma Half-life:

50 µM test compound incubated in mouse BALB/c plasma (pre-warmed to 37°C and buffered to pH7.4 in ratio of 70:30 plasma:buffer). At 0, 30, 60, 120 and 180 minutes a 50 µl aliquot of incubation mixture were removed and mixed with 200 µl acetonitrile, containing 50 ng/ml Donepezil as the internal standard, to stop the reaction. The samples were centrifuged to sediment any precipitated protein and the microplates sealed prior to UPLC-MS/MS analysis using a Quattro Premier XE (Waters

Corporation). XLfit (IDBS) was used to calculate the exponential decay and hence the rate constant (K) using the ratio of the peak areas of the test compound to the internal standard at each time point. The half-life was calculated for each test compound using the formula: t1/2 = 0.693/K.

Intrinsic Clearance:

0.5 µM test compound was incubated with BALB/c female CD1 mouse liver microsomes (Xenotech LLC TM; 0.5 mg/ml 50 mM potassium phosphate buffer, pH7.4) and the reaction started with addition of excess NADPH (8 mg/ml 50 mM potassium phosphate buffer, pH7.4). At 0, 3, 6, 9, 15 and 30 minutes a 50 µl aliquot of the incubation mixture was removed and mixed with 100 µl acetonitrile to stop the reaction. Internal standard was added to all samples, which were centrifuged to sediment precipitated protein and the microplates were then sealed prior to UPLC MS/MS analysis using a Quattro Premier XE. XLfit was used to calculate the exponential decay and hence the rate constant (K), based on the ratio of the peak areas of test compound to internal standard at each time point. The rate of intrinsic clearance (CL_{int}) was then calculated using the following formula:

C_{Lint} (ml/min/g liver) = K * V * microsomal protein yield

Where V (ml/mg protein) is the incubation volume/mg protein added, and microsomal protein yield is taken as 52.5 mg protein/g liver. Verapamil (0.5 µM) was used as a positive control to confirm assay performance.

PAMPA:

PAMPA assays were performed using a 96-well pre-coated BD Gentest™ PAMPA plate (BD Biosciences). Each well was divided into two chambers; donor and acceptor, separated by a lipid-oil-lipid tri-layer constructed in a porous filter. The effective permeability, Pe, of the compound was measured at pH 7.4. Stock solutions (5 mM) of the compound were prepared in DMSO. The compound was then further diluted to 10 µM in phosphate buffered saline at pH 7.4. The final DMSO concentration did not exceed 5% v/v. The compound dissolved in phosphate buffered saline was then added to the donor side of the membrane and phosphate buffered saline without compound was added to the acceptor side. The PAMPA plate was left at room temperature for 5 h. After which time, an aliquot (100 µl) was then removed from

both acceptor and donor compartments and mixed with acetonitrile (80 µl) containing an internal standard. The samples were centrifuged (10 min, 5^{o}C, 3270 g) to sediment precipitated protein and sealed prior to UPLC-MS/MS analysis using a Quattro Premier XE. Pe was calculated as shown in the below equation:

Pe (nm/sec) = 107 x - In [1 - CA (t) / Cequi]) / (A * [1/VD + 1/VA] * t)

Where:

CA(t) = peak area of compound present in acceptor well at time t = 18000 sec

Cequi = [CD(t) * VD + CA(t) * VA] / (VD + VA)

A = filter area

VD = donor well volume

VA = acceptor well volume

t = incubation time

CD(t) = peak area of compound present in donor well at time t = 18000 sec

Recovery of compound from donor and acceptor wells was calculated and data was only accepted when recovery exceeded 70 %.

8.9 – Off-Target Screening (companies)

Off-target screening was carried out by a number of organisations, who were provided with compounds in the form of 100% DMSO solutions.

BROMOscan experiments were performed by DiscoveRX (San Diego, California, USA), requiring 60 µl of 10 mM MZ-20 (for KdELECT assay) and 70 µl of 1 mM 9-ME-1 (for bromoMAX screen).

The kinase screen was carried out within the School of Life Sciences, at the MRC PPU International Centre for Kinase Profiling. 9-ME-1 was submitted for the Express Screen panel, which required 200 µl of 1 mM compound.

A broad-spectrum screen of transporters, ion channels and receptors was performed at Eurofins CEREP (Bois-l'Évêque, France). 9-ME-1 was tested against the Express Profile panel, requiring 140 µl of 10 mM compound.

8.10 – Tissue Culture

Human U2OS, HEK293T, HL-60 A549 and H23 cell lines were obtained from ATCC and MV4;11 cell line was obtained from DSMZ. U2OS, HEK293T, A549 & H23 were kept in DMEM medium (Gibco) supplemented with 10% FBS (Gibco), 1% L-glutamate (Gibco) and 100 U/ml penicillin and streptomycin. MV4-11 and HL-60 were kept in RPMI medium (Gibco) supplemented with 10% FBS, 1% L-glutamate (Gibco) and 100 U/ml penicillin and streptomycin. Cells were kept in an incubator at 37°C, 5% CO_2. All cell lines were tested for mycoplasma contamination every month using MycoAlert Mycoplasma detection kit (Lonza).

Cell-line	**U2OS**	**HEK293T**	**MV4-11**	**HL-60**	**A549**	**H23**
Source	osteosarcoma	Embryonic kidneys	AML	AML	LAC	LAC
Adherent?	Adherent	Semi-adherent	Suspension	Suspension	Adherent	Adherent
BET-dependent?	No	No	Yes	Yes	Yes	Yes
Assays	FRAP, CETSA	Luciferase, FRAP	Cell-viability	Cell-viability	Cell-viability, colony-formation, luciferase	Cell-viability

Table 8.4 Cancer cell-Line characteristics.

8.11 – Fluorescence Recovery after Photobleaching (FRAP)

FRAP experiments were performed in human osteosarcoma U2OS cells transfected with pcDNA5 FRT/TO plasmids encoding wild-type and mutant GFP chimeras of BET proteins (pcDNA6 for GFP-BRDT). Cells were seeded into glass-bottom, 35x10 mm dishes (WillCo) at ~200 000 cells in 2 ml media per dish ~66 h before FRAP. Cells were transfected using 3 µl FuGENE HD (Promega) and 1 µg of plasmid DNA ~42 h before FRAP. Eighteen hours before the FRAP assay cells were treated with test compounds and SAHA (Sigma-Aldrich) to a final concentration of 2 µM SAHA 0.03% DMSO.

FRAP experiments were carried out on a Deltavision Elite imaging system (GE Healthcare) running Resolve 3D (SoftWoRx) kept at 37°C and using a FITC filter set (488 nm excitation, 525 nm emission). Cells were imaged using a 60X objective lens (Olympus), with an exposure time of 0.05s and using 2x2 binning to give a 512x512

pixel image. Cells were excluded for FRAP experiments if they displayed aberrant morphology and signs of cell-death. Cells providing a signal of <1000 fluorescence units were excluded as they were too close to the background signal for photobleaching (requiring reduction to ~50% fluorescence) to be observed. Cells fluorescing over 3500 units were also excluded due to the risk of detector saturation.

Cells were photobleached using a quantifiable laser module (QLM), set to a wavelength of 488 nm and 100% power, for 0.5s, covering an area with a 0.5 µm radius. For each cell 5 images were captured pre-bleach (over 5 s) and 32 post-bleach, using a CoolSNAP HQ camera (Photometrics). Post-bleach imaging was usually spread over 60s and distributed to best measure a t½ of ~2 s, although this was altered in some cases. FRAP experiments were analysed in SoftWorx (GE Healthcare) using the PK analysis function, set to analyse a bleach event with 0.5 μm^2 radius. Calculated t½ values were extracted and analysed in Prism 6 (GraphPad). Each experiment was run twice on separate days, with ~20 cells tested on each day.

8.12 – Cell Viability Assays

Cell-viability, and the impact of our compounds, was measured using the CellTiter-Glo assay (Promega). Slightly different protocols were utilised for AML-derived suspension cell-lines (MV4-11 & HL-60) and LAC-derived adherent cell-lines (A549 & H23).

For suspension cells, test compounds were serially diluted in a sterile, white, clear-bottomed 384-well cell-culture microplate (Greiner Bio-one), at 2X concentration and a volume of 25µl. Cell suspension was then added, with 25 µl at 2X concentration. Both cells and compounds were diluted in RPMI medium. After 48 h incubation 25 µl of CellTiter-Glo reagent was added to each well. Following 15 minute incubation the luminescence signal was read on a Pherastar FS. The final concentration of assay components are as follows: $3x10^5$ cells/ml, 0.05% DMSO, 5 µM and below compound.

Adherent cell-lines were first seeded into the microplate, at 3000 cells in 25 µl. The next day 25 µl of 2X concentration, serially diluted test compounds were added. Both cells and compounds were diluted in DMEM medium. Following a 72 h incubation 25 µl of CellTiter-Glo reagent was added to each well. Following a 30 minute incubation the

luminescence signal was read on a Pherastar FS. The final concentration of assay components are as follows: 0.05% DMSO, 5 µM and below compound.

Data was processed and dose-response curves generated using Prism 6 (Graphpad).

For the failed GFP-BET overexpression experiments A549 cells were seeded in 96-well plates with 40000 cells in 100 µl. Cells were transfected the next day, using a scaled-down version of the transfection mixture used in luciferase assays. Compounds were again added at 2X concentration in 100 µl. For luminescence measurement 100 µl of reagent was added per well.

8.13 – Colony Formation Assays

For colony formation assays A549 cells were seeded in 6-well plates at a volume of 2 ml. For qualitative assays $2x10^4$ cells were seeded, while quantitative assays utilised $2x10^3$ cells. The next day the media was replaced with DMEM containing 1 µM compound and/or 0.1% DMSO (unless stated otherwise). Plates were then incubated for 7-14 days, until the desired confluence (in general or within colonies) was obtained. Qualitative assays took 7-9 days, while quantitative assays required 10-14 days.

Following incubation, media was removed and cells washed with 1 ml H_2O. Cells were fixed with 1 ml 100%, ice-cold MeOH, incubated for 30 minutes at 4°C. Cells were then stained with 500 µl 0.1% crystal-violet dye (in MeOH), with a 30 minute room-temperature incubation. Cells were then washed three times with H_2O and left to dry overnight.

For 96-well dose-response colony formation assays, 1000 A549 cells were seeded per well in 90 µl DMEM. The next day cells were treated with 10 µl 10x test compound in DMEM. Following 7-10 days incubation cells were fixed and stained as with 6-well plates, but with volumes of 50 µl.

Dried plates were scanned in an Epson Perfection V800 scanner, with a resolution of 1200 dpi and in 16-bit grayscale and 48-bit colour form. Colonies were counted manually from computer images.

For some assays, crystal-violet dye was later extracted by 1% SDS over 3 hours, and the absorbance at 595 nm measured in a Pherastar FS.

8.14 – Luciferase Assays

Luciferase experiments were performed in human HEK293T cells seeded in a 24-well plate at $1x10^5$ cells in 500 µl media, and transfected with pcDNA5 FRT/TO plasmids encoding WT and mutant GFP-BRD4 chimeras, and pBABE NF-κB-RE/*luc2P* reporter plasmid. Cells were transfected using Lipofectamine™ LTX (Invitrogen) with PLUS™ reagent, with each well receiving 0.5 µg of each plasmid, 2.5 µl LTX and 0.5 µl PLUS reagent in 100 µl OPTIMEM. Cells were treated overnight with 1 µM test compound and 0.02% DMSO. After treatment cells were washed with PBS, lysed with passive lysis buffer (Promega) and stored at -20°C.

Luciferase activity was measured suing the Promega luciferase assay. In a black low-volume 384-well microplate (Corning) 3 µl of clarified lysate was mixed with 15 µl luciferase assay reagent and after 15 minutes luminescence was measured in a Pherastar FS. Each experiment was run twice, on separate days, with each experiment containing 3 technical replicates.

8.15 – Cell Extract Thermal Shift Assay (CETSA)

U2OS FLP-IN cells, containing tetracycline-inducible GFP-BRD4 WT/WT and LV/LV, were cultured in 10 cm dishes. Unless stated otherwise, cells were treated with 0.1 µg/ml tetracycline, giving protein levels equivalent to the endogenous WT BRD4 (long). Cells were trypsinised, washed with PBS and lysed through repeated (4x) liquid-nitrogen freeze/thaw cycles. Before the fourth freeze the protein content of each lysate was quantified through the BCA assay (Thermo) and aliquoted, before being stored at -80°C.

For the compound assay, lysate was diluted in PBS to 3 mg/ml protein, and incubated for 1 hour with 1 µM compound / 0.01% DMSO. Aliquots of treated and untreated lysates were then heated for 5 minutes from 35 to 60°C, on a CFXconnect (BioRad). Heated lysates were then clarified by centrifugation at 17 000g for 10 minutes at 4°C,

before undergoing SDS-PAGE on a NuPAGE 3-7% tris-acetate gel (30 mg protein per well).

Bands were transferred from the gel to PVDF membrane in a 4°C, overnight transfer (BioRad Mini Trans-Blot) using TBST buffer (TBS buffer with 0.1% Tween-20) with 10% MeOH, and blocked for 1 hour at room temperature with 3.5% BSA in TBST. Blotting with the primary antibody was done either overnight at 4°C, or for 1 hour at room temperature, and secondary antibody blotting was done at 4°C for 1 hour. BRD4 was blotted for using an Abcam αBRD4 128874 rabbit antibody, at 1:1000, then a Licor αrabbit cw800 antibody at 1:10000. GFP was blotted for with a Chromatek αGFP 3H9 rat antibody at 1:1000, followed by a Licor 800cw αrat antibody at 1:10000. Bands were visualised using a Licor Odyssey Sa imaging system, and images were processed and band intensity measured using Licor Image Studio v5.2.5. Band intensities were normalised relative to those heated to the lowest temperature.

8.16 – CRISPR/Cas9 BRD4 BD1 L/V Knock-In

RNP Transfection:

The Cas9 nickase (Alt-R® S.p. Cas9 D10A Nickase V3) and the large, invariable, nickase-binding gRNA strand (Alt-R® CRISPR-Cas9 tracrRNA) were ordered from Integrated DNA Technologies (IDT). The short, variable, DNA-binding gRNA strands (Alt-R® CRISPR-Cas9 crRNA) and the knock-in template DNA were custom-ordered from IDT.

Cells were transfected at ~50% confluence. The following protocol is for a 10 cm dish, and can be scaled down for 6-well plates. First, 4 µl of 100 µM long, invariable gRNA, and 20 µl of 10 µM of each short, custom gRNA was diluted in 320 µl duplex buffer, heated to 95°C and slowly cooled to room temperature. The nickase (61x stock) was diluted to 1x in optimem, with 6.6 µl 61x stock in 400 µl media. Next 330 µl of the RNA mix was mixed with 330 µl of 1X nickase, and incubated at rt for 10 minutes. Template DNA was diluted with 10.5 µl 10 µM stock to 300 nM in duplex buffer. Next 535 µl of RNP and 270 µl of DNA was diluted in 1200 µl media, and incubated for 10 minutes. Meanwhile, 250 µl of diluted transfection reagent was prepared, with 25 µl of reagent (25 µl of PLUS reagent), and incubated for 10 minutes. The final transfection mixture was formed from 750 µl of the RNP/DNA solution combined with 250 µl of dilute

transfection reagent, incubated at rt for 15 minutes. Cells were washed with PBS and given fresh media, before being treated drop-by-drop with 1ml of the final transfection solution.

Cell Sorting:

4 days after transfection the cells underwent single-cell sorting. Cells were washed, trypsinised, centrifuged and re-suspended in 3 ml DMEM with 1% FBS, at 2-5x10^6 cells/ml. Cells were sorted onto 96-well plates, which the day before had been coated with 1% gelatin and filled with 150 µl of conditioned media (50% DMEM with 10% FBS, 40% used media (filtered), 10% FBS) before being incubated overnight at 37°C. After the addition of single cells the plates were centrifuged at 50xg for 1 minute, before being incubated at 37°C for 2 weeks.

Colony-Formation Assay:

The colony-formation assay was begun once a large number of clones had formed large, and densely confluent colonies. Cells were drained and detached with 50 µl trypsin (inactivated with 150 µl DMEM). For each 96-well plate containing sorted cells, 3 new 96-well plates were seeded. Two plates were seeded with 70 µl dissociated cells, and one was seeded with 50 µl cells. The plates with 50 µl of cell solution were used for continued growth and passaging of cells, and had 150 µl DMEM added. The plates with 70 µl cells were used to test for 9-ME-1 sensitivity, and the next day were treated with 140 µl of media containing 0.04% DMSO or 3 µM 9-ME-1. After 1 week DMSO-treated cells began to reach confluence, and so the cells were fixed with 100% MeOH for 1 hour at 4°C and stained with 0.1% crystal-violet for 1 hour at room temperature. Cells were then washed with H_2O and dried overnight, before being scanned in an Epson Perfection V800 scanner, with a resolution of 1200 dpi and in 16-bit grayscale and 48-bit colour form. Later, crystal-violet dye was extracted by 1% SDS over 3 hours, and the absorbance at 595 nm measured in a Pherastar FS.

Chapter 9
References

1. Runcie AC, et al., *Chemical genetics approaches for selective intervention in epigenetics.* Curr Opin Chem Biol, 2016. **33**: p. 186-94.
2. Schreiber SL, *Chemical genetics resulting from a passion for synthetic organic chemistry.* Bioorg Med Chem, 1998. **6**(8): p. 1127-52.
3. Walsh DP and YT Chang, *Chemical genetics.* Chem Rev, 2006. **106**(6): p. 2476-530.
4. Arrowsmith CH, et al., *The promise and peril of chemical probes.* Nat Chem Biol, 2015. **11**(8): p. 536-41.
5. Spring DR, *Chemical genetics to chemical genomics: small molecules offer big insights.* Chem Soc Rev, 2005. **34**(6): p. 472-82.
6. Ward RJ, E Alvarez-Curto, and G Milligan, *Using the Flp-In (TM) T-Rex (TM) System to Regulate GPCR Expression.* Receptor Signal Transduction Protocols, Third Edition, 2011. **746**: p. 21-37.
7. Ambesajir A, et al., *RNA interference: A futuristic tool and its therapeutic applications.* Saudi J Biol Sci, 2012. **19**(4): p. 395-403.
8. Bird A, *Perceptions of epigenetics.* Nature, 2007. **447**(7143): p. 396-8.
9. Tarakhovsky A, *Tools and landscapes of epigenetics.* Nat Immunol, 2010. **11**(7): p. 565-8.
10. Jin B, Y Li, and KD Robertson, *DNA methylation: superior or subordinate in the epigenetic hierarchy?* Genes Cancer, 2011. **2**(6): p. 607-17.
11. Moore LD, T Le, and G Fan, *DNA methylation and its basic function.* Neuropsychopharmacology, 2013. **38**(1): p. 23-38.
12. Jenuwein T and CD Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.
13. Luger K, et al., *Crystal structure of the nucleosome core particle at 2.8 angstrom resolution.* Nature, 1997. **389**(6648): p. 251-260.
14. Travers A, *The 30-nm Fiber Redux.* Science, 2014. **344**(6182): p. 370-372.
15. Bannister AJ and T Kouzarides, *Regulation of chromatin by histone modifications.* Cell Res, 2011. **21**(3): p. 381-95.
16. Shogren-Knaak M, et al., *Histone H4-K16 acetylation controls chromatin structure and protein interactions.* Science, 2006. **311**(5762): p. 844-847.
17. Tropberger P, et al., *Regulation of Transcription through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer.* Cell, 2013. **152**(4): p. 859-872.
18. Martin C and Y Zhang, *The diverse functions of histone lysine methylation.* Nat Rev Mol Cell Biol, 2005. **6**(11): p. 838-49.
19. Wang Z, et al., *Combinatorial patterns of histone acetylations and methylations in the human genome.* Nat Genet, 2008. **40**(7): p. 897-903.
20. Huston A, et al., *Probing the epigenome.* Nat Chem Biol, 2015. **11**(8): p. 542-5.
21. Arkin MR, YY Tang, and JA Wells, *Small-Molecule Inhibitors of Protein-Protein Interactions: Progressing toward the Reality.* Chemistry & Biology, 2014. **21**(9): p. 1102-1114.
22. Marks PA, *Histone deacetylase inhibitors: A chemical genetics approach to understanding cellular functions.* Biochimica Et Biophysica Acta-Gene Regulatory Mechanisms, 2010. **1799**(10-12): p. 717-725.
23. Yoon S and GH Eom, *HDAC and HDAC Inhibitor: From Cancer to Cardiovascular Diseases.* Chonnam Med J, 2016. **52**(1): p. 1-11.
24. Eckschlager T, et al., *Histone Deacetylase Inhibitors as Anticancer Drugs.* International Journal of Molecular Sciences, 2017. **18**(7).
25. Micelli C and G Rastelli, *Histone deacetylases: structural determinants of inhibitor selectivity.* Drug Discov Today, 2015. **20**(6): p. 718-35.
26. Milosevich N and F Hof, *Chemical Inhibitors of Epigenetic Methyllysine Reader Proteins.* Biochemistry, 2016. **55**(11): p. 1570-83.
27. Teske KA and MK Hadden, *Methyllysine binding domains: Structural insight and small molecule probe development.* Eur J Med Chem, 2017. **136**: p. 14-35.

28. James LI, et al., *Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain.* Nature Chemical Biology, 2013. **9**(3): p. 184-191.
29. Amato A, et al., *Targeting Ligandable Pockets on Plant Homeodomain (PHD) Zinc Finger Domains by a Fragment-Based Approach.* Acs Chemical Biology, 2018. **13**(4): p. 915-921.
30. Miller TC, et al., *Competitive binding of a benzimidazole to the histone-binding pocket of the Pygo PHD finger.* ACS Chem Biol, 2014. **9**(12): p. 2864-74.
31. Tamkun JW, et al., *brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2.* Cell, 1992. **68**(3): p. 561-72.
32. Filippakopoulos P, et al., *Histone recognition and large-scale structural analysis of the human bromodomain family.* Cell, 2012. **149**(1): p. 214-31.
33. Vidler LR, et al., *Druggability analysis and structural classification of bromodomain acetyl-lysine binding sites.* J Med Chem, 2012. **55**(17): p. 7346-59.
34. Miller TC, et al., *A bromodomain-DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT.* Nat Commun, 2016. **7**: p. 13855.
35. *UniProt - SMARCA2*. 18/07/2018; Available from: https://www.uniprot.org/uniprot/P51531.
36. *UniProt - CREBBP*. 18/07/2018; Available from: https://www.uniprot.org/uniprot/Q92793.
37. *UniProt - ASH1L*. 18/07/2018; Available from: https://www.uniprot.org/uniprot/Q9NR48.
38. *UniProt - BAZ2A*. 18/07/2018; Available from: https://www.uniprot.org/uniprot/Q9UIF9.
39. *SGC - Chemical Probes*. July 2018; Available from: https://www.thesgc.org/chemical-probes.
40. Filippakopoulos P and S Knapp, *Targeting bromodomains: epigenetic readers of lysine acetylation.* Nat Rev Drug Discov, 2014. **13**(5): p. 337-56.
41. Galdeano C and A Ciulli, *Selectivity on-target of bromodomain chemical probes by structure-guided medicinal chemistry and chemical biology.* Future Med Chem, 2016. **8**(13): p. 1655-80.
42. Picaud S, et al., *Promiscuous targeting of bromodomains by bromosporine identifies BET proteins as master regulators of primary transcription response in leukemia.* Science Advances, 2016. **2**(10).
43. Korb E, et al., *BET protein Brd4 activates transcription in neurons and BET inhibitor Jq1 blocks memory in mice.* Nature Neuroscience, 2015. **18**(10): p. 1464-+.
44. Chang YL, et al., *A double-bromo domain protein, FSH-S, activates the homeotic gene Ultrabithorax through a critical promoter-proximal region.* Molecular and Cellular Biology, 2007. **27**(15): p. 5486-5498.
45. Mietton F, et al., *Selective BET bromodomain inhibition as an antifungal therapeutic strategy.* Nature Communications, 2017. **8**.
46. Wu SY and CM Chiang, *The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation.* Journal of Biological Chemistry, 2007. **282**(18): p. 13141-13145.
47. Devaiah BN, A Gegonne, and DS Singer, *Bromodomain 4: a cellular Swiss army knife.* Journal of Leukocyte Biology, 2016. **100**(4): p. 679-686.
48. Shen C, et al., *NSD3-short is an adaptor protein that couples BRD4 to the CHD8 chromatin remodeler.* Cancer Research, 2016. **76**.
49. Zhang Q, et al., *Structural Mechanism of Transcriptional Regulator NSD3 Recognition by the ET Domain of BRD4.* Structure, 2016. **24**(7): p. 1201-1208.
50. Rahman S, et al., *The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3.* Mol Cell Biol, 2011. **31**(13): p. 2641-52.

51. Crowe BL, et al., *Structure of the Brd4 ET domain bound to a C-terminal motif from gamma-retroviral integrases reveals a conserved mechanism of interaction.* Proc Natl Acad Sci U S A, 2016. **113**(8): p. 2086-91.
52. Devaiah BN, et al., *BRD4 is an atypical kinase that phosphorylates Serine2 of the RNA Polymerase II carboxy-terminal domain.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(18): p. 6927-6932.
53. Denis GV and MR Green, *A novel, mitogen-activated nuclear kinase is related to a Drosophila developmental regulator.* Genes & Development, 1996. **10**(3): p. 261-271.
54. Devaiah BN, et al., *BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin.* Nature Structural & Molecular Biology, 2016. **23**(6): p. 540-548.
55. Bisgrove DA, et al., *Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription.* Proc Natl Acad Sci U S A, 2007. **104**(34): p. 13690-5.
56. Itzen F, et al., *Brd4 activates P-TEFb for RNA polymerase II CTD phosphorylation.* Nucleic Acids Research, 2014. **42**(12): p. 7577-7587.
57. Jang MK, et al., *The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription.* Molecular Cell, 2005. **19**(4): p. 523-534.
58. Kanno T, et al., *BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones.* Nature Structural & Molecular Biology, 2014. **21**(12): p. 1047-1057.
59. Natoli G and JC Andrau, *Noncoding Transcription at Enhancers: General Principles and Functional Models.* Annual Review of Genetics, Vol 46, 2012. **46**: p. 1-19.
60. Rahnamoun H, et al., *RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation.* Nat Struct Mol Biol, 2018. **25**(8): p. 687-697.
61. Bauer DE, et al., *Clinicopathologic features and long-term outcomes of NUT midline carcinoma.* Clin Cancer Res, 2012. **18**(20): p. 5773-9.
62. French C, *NUT midline carcinoma.* Nat Rev Cancer, 2014. **14**(3): p. 149-50.
63. French CA, et al., *BRD4 bromodomain gene rearrangement in aggressive carcinoma with translocation t(15;19).* Am J Pathol, 2001. **159**(6): p. 1987-92.
64. French CA, et al., *BRD-NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells.* Oncogene, 2008. **27**(15): p. 2237-42.
65. Fagerberg L, et al., *Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics.* Mol Cell Proteomics, 2014. **13**(2): p. 397-406.
66. Paillisson A, et al., *Bromodomain testis-specific protein is expressed in mouse oocyte and evolves faster than its ubiquitously expressed paralogs BRD2, -3, and -4.* Genomics, 2007. **89**(2): p. 215-223.
67. Barda S, et al., *New insights into the role of the Brdt protein in the regulation of development and spermatogenesis in the mouse.* Gene Expression Patterns, 2016. **20**(2): p. 130-137.
68. Berkovits BD and DJ Wolgemuth, *The Role of the Double Bromodomain-Containing BET Genes During Mammalian Spermatogenesis.* Gametogenesis, 2013. **102**: p. 293-326.
69. Bourova-Flin E, et al., *The Role of Bromodomain Testis-Specific Factor, BRDT, in Cancer: A Biomarker and A Possible Therapeutic Target.* Cell J, 2017. **19**(Suppl 1): p. 1-8.
70. *US National Library of Medicine - Clinical Trials Database*. [cited 2018 June]; Available from: https://clinicaltrials.gov/.
71. Liu Z, et al., *Drug Discovery Targeting Bromodomain-Containing Protein 4.* J Med Chem, 2017. **60**(11): p. 4533-4558.
72. Perez-Salvia M and M Esteller, *Bromodomain inhibitors and cancer therapy: From structures to applications.* Epigenetics, 2017. **12**(5): p. 323-339.

73. Hewings DS, et al., *Progress in the development and application of small molecule inhibitors of bromodomain-acetyl-lysine interactions.* J Med Chem, 2012. **55**(22): p. 9393-413.
74. Adachi K, et al., *THIENOTRIAZOLODIAZEPINE COMPOUND AND A MEDICINAL USE THEREOF*. 2008.
75. Miyoshi S, et al., *ANTITUMOR AGENT*. 2009.
76. Filippakopoulos P, et al., *Selective inhibition of BET bromodomains.* Nature, 2010. **468**(7327): p. 1067-73.
77. Nicodeme E, et al., *Suppression of inflammation by a synthetic histone mimic.* Nature, 2010. **468**(7327): p. 1119-23.
78. Mirguet O, et al., *Discovery of epigenetic regulator I-BET762: lead optimization to afford a clinical candidate inhibitor of the BET bromodomains.* J Med Chem, 2013. **56**(19): p. 7501-15.
79. Welsch ME, SA Snyder, and BR Stockwell, *Privileged scaffolds for library design and drug discovery.* Curr Opin Chem Biol, 2010. **14**(3): p. 347-61.
80. Noel JK, et al., *Abstract C244: Development of the BET bromodomain inhibitor OTX015.* Molecular Cancer Therapeutics, 2013. **12**: p. C244-C244.
81. Albrecht BK, et al., *Identification of a Benzoisoxazoloazepine Inhibitor (CPI-0610) of the Bromodomain and Extra-Terminal (BET) Family as a Candidate for Human Clinical Trials.* Journal of Medicinal Chemistry, 2016. **59**(4): p. 1330-1339.
82. Filippakopoulos P, et al., *Benzodiazepines and benzotriazepines as protein interaction inhibitors targeting bromodomains of the BET family.* Bioorganic & Medicinal Chemistry, 2012. **20**(6): p. 1878-1886.
83. Waring MJ, et al., *Potent and selective bivalent inhibitors of BET bromodomains.* Nat Chem Biol, 2016. **12**(12): p. 1097-1104.
84. Rhyasen GW, et al., *AZD5153: A Novel Bivalent BET Bromodomain Inhibitor Highly Active against Hematologic Malignancies.* Mol Cancer Ther, 2016. **15**(11): p. 2563-2574.
85. Tanaka M, et al., *Design and characterization of bivalent BET inhibitors.* Nat Chem Biol, 2016. **12**(12): p. 1089-1096.
86. Anders L, et al., *Genome-wide localization of small molecules.* Nat Biotechnol, 2014. **32**(1): p. 92-6.
87. Li Z, et al., *"Minimalist" cyclopropene-containing photo-cross-linkers suitable for live-cell imaging and affinity-based protein labeling.* J Am Chem Soc, 2014. **136**(28): p. 9990-8.
88. Pan S, et al., *A Suite of "Minimalist" Photo-Crosslinkers for Live-Cell Imaging and Chemical Proteomics: Case Study with BRD4 Inhibitors.* Angew Chem Int Ed Engl, 2017. **56**(39): p. 11816-11821.
89. Tyler DS, et al., *Click chemistry enables preclinical evaluation of targeted epigenetic therapies.* Science, 2017. **356**(6345): p. 1397-1401.
90. Shi J and CR Vakoc, *The mechanisms behind the therapeutic activity of BET bromodomain inhibition.* Mol Cell, 2014. **54**(5): p. 728-36.
91. Zuber J, et al., *RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia.* Nature, 2011. **478**(7370): p. 524-8.
92. Delmore JE, et al., *BET Bromodomain Inhibition as a Therapeutic Strategy to Target c-Myc.* Cell, 2011. **146**(6): p. 903-916.
93. Aird F, et al., *Replication Study: BET bromodomain inhibition as a therapeutic strategy to target c-Myc.* Elife, 2017. **6**.
94. Lockwood WW, et al., *Sensitivity of human lung adenocarcinoma cell lines to targeted inhibition of BET epigenetic signaling proteins.* Proc Natl Acad Sci U S A, 2012. **109**(47): p. 19408-13.
95. Zou Z, et al., *Brd4 maintains constitutively active NF-kappaB in cancer cells by binding to acetylated RelA.* Oncogene, 2014. **33**(18): p. 2395-404.

96. Xu Y and CR Vakoc, *Targeting Cancer Cells with BET Bromodomain Inhibitors.* Cold Spring Harb Perspect Med, 2017. **7**(7).
97. Mele DA, et al., *BET bromodomain inhibition suppresses TH17-mediated pathology.* J Exp Med, 2013. **210**(11): p. 2181-90.
98. Belkina AC, BS Nikolajczyk, and GV Denis, *BET Protein Function Is Required for Inflammation: Brd2 Genetic Disruption and BET Inhibitor JQ1 Impair Mouse Macrophage Inflammatory Responses.* Journal of Immunology, 2013. **190**(7): p. 3670-3678.
99. Gosmini R, et al., *The discovery of I-BET726 (GSK1324726A), a potent tetrahydroquinoline ApoA1 up-regulator and selective BET bromodomain inhibitor.* J Med Chem, 2014. **57**(19): p. 8111-31.
100. Bailey D, et al., *RVX-208 A Small Molecule That Increases Apolipoprotein A-I and High-Density Lipoprotein Cholesterol In Vitro and In Vivo.* Journal of the American College of Cardiology, 2010. **55**(23): p. 2580-2589.
101. Matzuk MM, et al., *Small-Molecule Inhibition of BRDT for Male Contraception.* Cell, 2012. **150**(4): p. 673-684.
102. Banerjee C, et al., *BET bromodomain inhibition as a novel strategy for reactivation of HIV-1.* Journal of Leukocyte Biology, 2012. **92**(6): p. 1147-1154.
103. Benito E, et al., *The BET/BRD inhibitor JQ1 improves brain plasticity in WT and APP mice.* Transl Psychiatry, 2017. **7**(9): p. e1239.
104. Li JJ, et al., *BET bromodomain inhibition promotes neurogenesis while inhibiting gliogenesis in neural progenitor cells.* Stem Cell Research, 2016. **17**(2): p. 212-221.
105. Amorim S, et al., *Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study.* Lancet Haematol, 2016. **3**(4): p. e196-204.
106. Berthon C, et al., *Bromodomain inhibitor OTX015 in patients with acute leukaemia: a dose-escalation, phase 1 study.* Lancet Haematol, 2016. **3**(4): p. e186-95.
107. Collins TA, et al., *Translational Modeling of Drug-Induced Myelosuppression and Effect of Pretreatment Myelosuppression for AZD5153, a Selective BRD4 Inhibitor.* CPT Pharmacometrics Syst Pharmacol, 2017. **6**(6): p. 357-364.
108. Mantena SR, et al., *Preclinical evidence for management of thrombocytopenia associated with bromodomain extra-terminal (BET) inhibition therapy.* Journal of Clinical Oncology, 2018. **36**(15_suppl): p. e24207-e24207.
109. Andrieu G, AC Belkina, and GV Denis, *Clinical trials for BET inhibitors run ahead of the science.* Drug Discov Today Technol, 2016. **19**: p. 45-50.
110. Lu P, et al., *BET inhibitors RVX-208 and PFI-1 reactivate HIV-1 from latency.* Sci Rep, 2017. **7**(1): p. 16646.
111. *Clinical Trial NCT02586155*. 28/09/2017; Available from: https://clinicaltrials.gov/ct2/show/NCT02586155.
112. Cheung KL, et al., *Distinct Roles of Brd2 and Brd4 in Potentiating the Transcriptional Program for Th17 Cell Differentiation.* Mol Cell, 2017. **65**(6): p. 1068-1080 e5.
113. Khan YM, et al., *Brd4 Is Essential for IL-1 beta-Induced Inflammation in Human Airway Epithelial Cells.* Plos One, 2014. **9**(4).
114. Michaeloudes C, et al., *Bromodomain and Extraterminal Proteins Suppress NF-E2-Related Factor 2-Mediated Antioxidant Gene Expression.* Journal of Immunology, 2014. **192**(10): p. 4913-4920.
115. McLure KG, et al., *RVX-208, an inducer of ApoA-I in humans, is a BET bromodomain antagonist.* PLoS One, 2013. **8**(12): p. e83190.
116. Picaud S, et al., *RVX-208, an inhibitor of BET transcriptional regulators with selectivity for the second bromodomain.* Proc Natl Acad Sci U S A, 2013. **110**(49): p. 19754-9.
117. Wang Q, et al., *Selective inhibition mechanism of RVX-208 to the second bromodomain of bromo and extraterminal proteins: insight from microsecond molecular dynamics simulations.* Sci Rep, 2017. **7**(1): p. 8857.

118. Kati W, *Abstract DDT01-05: ABBV-744: A first-in-class highly BDII-selective BET bromodomain inhibitor.* Cancer Research, 2018. **78**: p. DDT01-05-DDT01-05.
119. Lin X, et al., *Abstract 800: ABBV-744, a first-in-class and highly selective inhibitor of the second bromodomain of BET family proteins, displays robust activities in preclinical models of acute myelogenous leukemia.* Cancer Research, 2018. **78**: p. 800-800.
120. Sheppard GS, et al., *Abstract 931: Discovery of ABBV-744, a first-in-class highly BDII-selective BET bromodomain inhibitor.* Cancer Research, 2018. **78**: p. 931-931.
121. Law RP, et al., *Discovery of Tetrahydroquinoxalines as Bromodomain and Extra-Terminal Domain (BET) Inhibitors with Selectivity for the Second Bromodomain.* J Med Chem, 2018. **61**(10): p. 4317-4334.
122. Baud MG, et al., *New Synthetic Routes to Triazolo-benzodiazepine Analogues: Expanding the Scope of the Bump-and-Hole Approach for Selective Bromo and Extra-Terminal (BET) Bromodomain Inhibition.* J Med Chem, 2016. **59**(4): p. 1492-500.
123. McDaniel KF, et al., *Discovery of N-(4-(2,4-Difluorophenoxy)-3-(6-methyl-7-oxo-6,7-dihydro-1H-pyrrolo[2,3-c]pyridin -4-yl)phenyl)ethanesulfonamide (ABBV-075/Mivebresib), a Potent and Orally Available Bromodomain and Extraterminal Domain (BET) Family Bromodomain Inhibitor.* J Med Chem, 2017. **60**(20): p. 8369-8384.
124. *Clinical Trial NCT03360006*. 02/08/2018; Available from: https://clinicaltrials.gov/ct2/show/NCT03360006.
125. Khan R, et al., *Gram-Scale Laboratory Synthesis of TC AC 28, a High-Affinity BET Bromodomain Ligand.* Acs Omega, 2017. **2**(8): p. 4328-4332.
126. *Clinical Trial NCT02705469*. 20/11/2017; Available from: https://clinicaltrials.gov/ct2/show/NCT02705469.
127. Zhang G, et al., *Structure-guided design of potent diazobenzene inhibitors for the BET bromodomains.* J Med Chem, 2013. **56**(22): p. 9251-64.
128. Gacias M, et al., *Selective chemical modulation of gene transcription favors oligodendrocyte lineage progression.* Chem Biol, 2014. **21**(7): p. 841-854.
129. Raux B, et al., *Exploring Selective Inhibition of the First Bromodomain of the Human Bromodomain and Extra-terminal Domain (BET) Proteins.* J Med Chem, 2016. **59**(4): p. 1634-41.
130. Hugle M, et al., *4-Acyl Pyrrole Derivatives Yield Novel Vectors for Designing Inhibitors of the Acetyl-Lysine Recognition Site of BRD4(1).* J Med Chem, 2016. **59**(4): p. 1518-30.
131. Hughes SJ and A Ciulli, *Molecular recognition of ternary complexes: a new dimension in the structure-guided design of chemical degraders.* Essays Biochem, 2017. **61**(5): p. 505-516.
132. Zengerle M, KH Chan, and A Ciulli, *Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4.* ACS Chem Biol, 2015. **10**(8): p. 1770-7.
133. Gadd MS, et al., *Structural basis of PROTAC cooperative recognition for selective protein degradation.* Nat Chem Biol, 2017. **13**(5): p. 514-521.
134. Nowak RP, et al., *Plasticity in binding confers selectivity in ligand-induced protein degradation.* Nat Chem Biol, 2018.
135. Islam K, *Allele-specific chemical genetics: concept, strategies, and applications.* ACS Chem Biol, 2015. **10**(2): p. 343-63.
136. Islam K, *The Bump-and-Hole Tactic: Expanding the Scope of Chemical Genetics.* Cell Chem Biol, 2018.
137. Erlanson DA, et al., *Site-directed ligand discovery.* Proc Natl Acad Sci U S A, 2000. **97**(17): p. 9367-72.
138. Bishop AC, O Buzko, and KM Shokat, *Magic bullets for protein kinases.* Trends Cell Biol, 2001. **11**(4): p. 167-72.
139. Shah K, et al., *Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3565-70.

140. Liu Y, et al., *Engineering Src family protein kinases with unnatural nucleotide specificity.* Chem Biol, 1998. **5**(2): p. 91-101.
141. Bishop AC, et al., *Design of allele-specific inhibitors to probe protein kinase signaling.* Curr Biol, 1998. **8**(5): p. 257-66.
142. Bishop AC, et al., *A chemical switch for inhibitor-sensitive alleles of any protein kinase.* Nature, 2000. **407**(6802): p. 395-401.
143. Garske AL, et al., *Chemical genetic strategy for targeting protein kinases based on covalent complementarity.* Proc Natl Acad Sci U S A, 2011. **108**(37): p. 15046-52.
144. Spencer DM, et al., *Controlling signal transduction with synthetic ligands.* Science, 1993. **262**(5136): p. 1019-24.
145. Liberles SD, et al., *Inducible gene expression and protein translocation using nontoxic ligands identified by a mammalian three-hybrid screen.* Proc Natl Acad Sci U S A, 1997. **94**(15): p. 7825-30.
146. Clackson T, et al., *Redesigning an FKBP-ligand interface to generate chemical dimerizers with novel specificity.* Proc Natl Acad Sci U S A, 1998. **95**(18): p. 10437-42.
147. Banaszynski LA, et al., *A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules.* Cell, 2006. **126**(5): p. 995-1004.
148. An W, et al., *Engineering FKBP-Based Destabilizing Domains to Build Sophisticated Protein Regulation Systems.* PLoS One, 2015. **10**(12): p. e0145783.
149. Egeler EL, et al., *Ligand-switchable substrates for a ubiquitin-proteasome system.* J Biol Chem, 2011. **286**(36): p. 31328-36.
150. Robers M, et al., *Fluorescent Labeling of Proteins in Living Cells Using the FKBP12(F36V) Tag.* Cytometry Part A, 2009. **75A**(3): p. 207-224.
151. Nabet B, et al., *The dTAG system for immediate and target-specific protein degradation.* Nat Chem Biol, 2018. **14**(5): p. 431-441.
152. Islam K, et al., *Defining efficient enzyme-cofactor pairs for bioorthogonal profiling of protein methylation.* Proc Natl Acad Sci U S A, 2013. **110**(42): p. 16778-83.
153. Wang R, et al., *Profiling genome-wide chromatin methylation with engineered posttranslation apparatus within living cells.* J Am Chem Soc, 2013. **135**(3): p. 1048-56.
154. Islam K, et al., *Expanding cofactor repertoire of protein lysine methyltransferase for substrate labeling.* ACS Chem Biol, 2011. **6**(7): p. 679-84.
155. Li J, H Wei, and MM Zhou, *Structure-guided design of a methyl donor cofactor that controls a viral histone H3 lysine 27 methyltransferase activity.* J Med Chem, 2011. **54**(21): p. 7734-8.
156. Lin Q, et al., *Design of allele-specific protein methyltransferase inhibitors.* J Am Chem Soc, 2001. **123**(47): p. 11608-13.
157. Yang C, et al., *Labeling lysine acetyltransferase substrates with engineered enzymes and functionalized cofactor surrogates.* J Am Chem Soc, 2013. **135**(21): p. 7791-4.
158. Andrieu GP and GV Denis, *BET Proteins Exhibit Transcriptional and Functional Opposition in the Epithelial-to-Mesenchymal Transition.* Mol Cancer Res, 2018. **16**(4): p. 580-586.
159. Baud MGJ, et al., *Chemical biology. A bump-and-hole approach to engineer controlled selectivity of BET bromodomain chemical probes.* Science, 2014. **346**(6209): p. 638-641.
160. Schnitzler GR, *Isolation of histones and nucleosome cores from mammalian cells.* Curr Protoc Mol Biol, 2001. **Chapter 21**: p. Unit 21 5.
161. Fierz B and TW Muir, *Chromatin as an expansive canvas for chemical biology.* Nature Chemical Biology, 2012. **8**(5): p. 417-427.
162. Pierce MM, CS Raman, and BT Nall, *Isothermal titration calorimetry of protein-protein interactions.* Methods, 1999. **19**(2): p. 213-21.
163. Cooper M and A., *Current biosensor technologies in drug discovery.* Drug Discovery World, 2006: p. 68-82.
164. Philpott M, et al., *Assessing cellular efficacy of bromodomain inhibitors using fluorescence recovery after photobleaching.* Epigenetics Chromatin, 2014. **7**: p. 14.

165. Ishikawa-Ankerhold HC, R Ankerhold, and GPC Drummen, *Advanced Fluorescence Microscopy Techniques-FRAP, FLIP, FLAP, FRET and FLIM.* Molecules, 2012. **17**(4): p. 4047-4132.
166. Seibel NM, et al., *Nuclear localization of enhanced green fluorescent protein homomultimers.* Anal Biochem, 2007. **368**(1): p. 95-9.
167. Shang E, et al., *Identification of unique, differentiation stage-specific patterns of expression of the bromodomain-containing genes Brd2, Brd3, Brd4, and Brdt in the mouse testis.* Gene Expr Patterns, 2004. **4**(5): p. 513-9.
168. Runcie AC, et al., *Optimization of a "bump-and-hole" approach to allele-selective BET bromodomain inhibition.* Chem Sci, 2018. **9**(9): p. 2452-2468.
169. Niesen FH, H Berglund, and M Vedadi, *The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability.* Nat Protoc, 2007. **2**(9): p. 2212-21.
170. Eglen RM, et al., *The use of AlphaScreen technology in HTS: current status.* Curr Chem Genomics, 2008. **1**: p. 2-10.
171. Zhang JH, TD Chung, and KR Oldenburg, *A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays.* J Biomol Screen, 1999. **4**(2): p. 67-73.
172. Chung TDY, DB Terry, and LH Smith, *In Vitro and In Vivo Assessment of ADME and PK Properties During Lead Selection and Lead Optimization - Guidelines, Benchmarks and Rules of Thumb*, in *Assay Guidance Manual*, GS Sittampalam, et al., Editors. 2004: Bethesda (MD).
173. Steiner K and H Schwab, *Recent advances in rational approaches for enzyme engineering.* Comput Struct Biotechnol J, 2012. **2**: p. e201209010.
174. Lewis DF, MN Jacobs, and M Dickins, *Compound lipophilicity for substrate binding to human P450s in drug metabolism.* Drug Discov Today, 2004. **9**(12): p. 530-7.
175. Cai J and J Henion, *Elucidation of LSD in vitro metabolism by liquid chromatography and capillary electrophoresis coupled with tandem mass spectrometry.* J Anal Toxicol, 1996. **20**(1): p. 27-37.
176. Patel JM, JC Wood, and KC Leibman, *The biotransformation of allyl alcohol and acrolein in rat liver and lung preparations.* Drug Metab Dispos, 1980. **8**(5): p. 305-8.
177. Roberts JM and JE Bradner, *A Bead-Based Proximity Assay for BRD4 Ligand Discovery.* Curr Protoc Chem Biol, 2015. **7**(4): p. 263-78.
178. Maniaci C, et al., *Homo-PROTACs: bivalent small-molecule dimerizers of the VHL E3 ubiquitin ligase to induce self-degradation.* Nat Commun, 2017. **8**(1): p. 830.
179. Whitworth C, et al., *PO-449 Optimisation of an alphalisa assay for the characterisation of protac-induced ternary complexes within cell lysates.* ESMO Open, 2018. **3**(Suppl 2): p. A198-A198.
180. Chung CW, et al., *Discovery and characterization of small molecule inhibitors of the BET family bromodomains.* J Med Chem, 2011. **54**(11): p. 3827-38.
181. Lea WA and A Simeonov, *Fluorescence polarization assays in small molecule screening.* Expert Opin Drug Discov, 2011. **6**(1): p. 17-32.
182. Ciceri P, et al., *Dual kinase-bromodomain inhibitors for rationally designed polypharmacology.* Nat Chem Biol, 2014. **10**(4): p. 305-12.
183. Dubocovich ML and M Markowska, *Functional MT1 and MT2 melatonin receptors in mammals.* Endocrine, 2005. **27**(2): p. 101-10.
184. Zlotos DP, et al., *MT1 and MT2 melatonin receptors: ligands, models, oligomers, and therapeutic potential.* J Med Chem, 2014. **57**(8): p. 3161-85.
185. Pala D, et al., *Homology models of melatonin receptors: challenges and recent advances.* Int J Mol Sci, 2013. **14**(4): p. 8093-121.
186. Nguyen UTT, et al., *Accelerated chromatin biochemistry using DNA-barcoded nucteosome libraries.* Nature Methods, 2014. **11**(8): p. 834-840.

187. Riss TL, et al., *Cell Viability Assays*, in *Assay Guidance Manual*, GS Sittampalam, et al., Editors. 2004: Bethesda (MD).
188. Chan KH, et al., *Impact of Target Warhead and Linkage Vector on Inducing Protein Degradation: Comparison of Bromodomain and Extra-Terminal (BET) Degraders Derived from Triazolodiazepine (JQ1) and Tetrahydroquinoline (I-BET726) BET Inhibitor Scaffolds.* J Med Chem, 2018. **61**(2): p. 504-513.
189. Testa A, et al., *3-Fluoro-4-hydroxyprolines: Synthesis, Conformational Analysis, and Stereoselective Recognition by the VHL E3 Ubiquitin Ligase for Targeted Protein Degradation.* J Am Chem Soc, 2018. **140**(29): p. 9299-9313.
190. Bajar BT, et al., *Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting.* Sci Rep, 2016. **6**: p. 20889.
191. Furey TS, *ChIP-seq and beyond: new and improved methodologies to detect and characterize protein-DNA interactions.* Nat Rev Genet, 2012. **13**(12): p. 840-52.
192. Arrigoni L, et al., *Standardizing chromatin research: a simple and universal method for ChIP-seq.* Nucleic Acids Res, 2016. **44**(7): p. e67.
193. Gong F, et al., *Screen identifies bromodomain protein ZMYND8 in chromatin recognition of transcription-associated DNA damage that promotes homologous recombination.* Genes Dev, 2015. **29**(2): p. 197-211.
194. Conery AR, et al., *Bromodomain inhibition of the transcriptional coactivators CBP/EP300 as a therapeutic strategy to target the IRF4 network in multiple myeloma.* Elife, 2016. **5**.
195. Machleidt T, et al., *NanoBRET--A Novel BRET Platform for the Analysis of Protein-Protein Interactions.* ACS Chem Biol, 2015. **10**(8): p. 1797-804.
196. Rath D, et al., *The CRISPR-Cas immune system: biology, mechanisms and applications.* Biochimie, 2015. **117**: p. 119-28.
197. Cho SW, et al., *Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease.* Nat Biotechnol, 2013. **31**(3): p. 230-2.
198. Mali P, et al., *CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering.* Nat Biotechnol, 2013. **31**(9): p. 833-8.
199. Cong L, et al., *Multiplex Genome Engineering Using CRISPR/Cas Systems.* Science, 2013. **339**(6121): p. 819-823.
200. Christian M, et al., *Targeting DNA double-strand breaks with TAL effector nucleases.* Genetics, 2010. **186**(2): p. 757-61.
201. Joung JK and JD Sander, *TALENs: a widely applicable technology for targeted genome editing.* Nat Rev Mol Cell Biol, 2013. **14**(1): p. 49-55.
202. Bibikova M, et al., *Stimulation of homologous recombination through targeted cleavage by chimeric nucleases.* Molecular and Cellular Biology, 2001. **21**(1): p. 289-297.
203. Carroll D, *Genome engineering with zinc-finger nucleases.* Genetics, 2011. **188**(4): p. 773-82.
204. Adli M, *The CRISPR tool kit for genome editing and beyond.* Nat Commun, 2018. **9**(1): p. 1911.
205. Lino CA, et al., *Delivering CRISPR: a review of the challenges and approaches.* Drug Deliv, 2018. **25**(1): p. 1234-1257.
206. Wang H, M La Russa, and LS Qi, *CRISPR/Cas9 in Genome Editing and Beyond.* Annu Rev Biochem, 2016. **85**: p. 227-64.
207. Wang H, et al., *One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering.* Cell, 2013. **153**(4): p. 910-8.
208. *Cas9 Stable Cell Lines*. Available from: https://www.creative-biogene.com/product/cas9-cell-line.html.
209. Munoz IM, et al., *Improved Genome Editing in Human Cell Lines Using the CRISPR Method.* Plos One, 2014. **9**(10).

210. ATCC. *A549 Cell-Line (CCL-185)*. [cited 2018 September]; Available from: https://www.lgcstandards-atcc.org/products/all/CCL-185.
211. Coker EA, et al., *canSAR: update to the cancer translational research and drug discovery knowledgebase.* Nucleic Acids Res, 2018.
212. Rouillard AD, et al., *The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins.* Database (Oxford), 2016. **2016**.
213. Varma S, et al., *High resolution copy number variation data in the NCI-60 cancer cell lines from whole genome microarrays accessible through CellMiner.* PLoS One, 2014. **9**(3): p. e92047.
214. GenBank. *NM_058243.2. Homo sapiens bromodomain containing 4 (BRD4), transcript variant long, mRNA.* [cited 2018 December]; Available from: https://www.ncbi.nlm.nih.gov/nuccore/NM_058243.2.
215. Shen B, et al., *Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects.* Nat Methods, 2014. **11**(4): p. 399-402.
216. Ran FA, et al., *Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity.* Cell, 2013. **154**(6): p. 1380-9.
217. Song K and H Gudjonson. *re site finer*. 2013; Available from: http://resitefinder.appspot.com/.
218. Anderson KR, et al., *CRISPR off-target analysis in genetically engineered rats and mice.* Nat Methods, 2018. **15**(7): p. 512-514.
219. Zhang XH, et al., *Off-target Effects in CRISPR/Cas9-mediated Genome Engineering.* Mol Ther Nucleic Acids, 2015. **4**: p. e264.
220. Aach J, P Mali, and GM Church, *CasFinder: Flexible algorithm for identifying specific Cas9 targets in genomes.* bioRxiv, 2014.
221. Stemmer M, et al., *CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool.* PLoS One, 2015. **10**(4): p. e0124633.
222. Zhao C, et al., *CRISPR-offinder: a CRISPR guide RNA design and off-target searching tool for user-defined protospacer adjacent motif.* Int J Biol Sci, 2017. **13**(12): p. 1470-1478.
223. Biolabs NE. *Activity of Restriction Enzymes in PCR Buffers*. Available from: https://international.neb.com/tools-and-resources/usage-guidelines/activity-of-restriction-enzymes-in-pcr-buffers.
224. Vakulskas CA, et al., *A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells.* Nat Med, 2018. **24**(8): p. 1216-1224.
225. DeWitt MA, JE Corn, and D Carroll, *Genome editing via delivery of Cas9 ribonucleoprotein.* Methods, 2017. **121-122**: p. 9-15.
226. Kim S, et al., *Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins.* Genome Res, 2014. **24**(6): p. 1012-9.
227. Guschin DY, et al., *A rapid and general assay for monitoring endogenous gene modification.* Methods Mol Biol, 2010. **649**: p. 247-56.
228. Qiu P, et al., *Mutation detection using Surveyor (TM) nuclease.* Biotechniques, 2004. **36**(4): p. 702-+.
229. Yeung AT, et al., *Enzymatic mutation detection technologies.* Biotechniques, 2005. **38**(5): p. 749-58.
230. Southern EM, *Detection of Specific Sequences among DNA Fragments Separated by Gel-Electrophoresis.* Journal of Molecular Biology, 1975. **98**(3): p. 503-+.
231. Guell M, L Yang, and GM Church, *Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA).* Bioinformatics, 2014. **30**(20): p. 2968-70.
232. Syvanen AC, *Accessing genetic variation: genotyping single nucleotide polymorphisms.* Nat Rev Genet, 2001. **2**(12): p. 930-42.
233. Olivier M, *The Invader assay for SNP genotyping.* Mutat Res, 2005. **573**(1-2): p. 103-10.
234. McGuigan FEA and SH Ralston, *Single nucleotide polymorphism detection: allelic discrimination using TaqMan.* Psychiatric Genetics, 2002. **12**(3): p. 133-136.

235. Abravaya K, et al., *Molecular beacons as diagnostic tools: Technology and applications.* Clinical Chemistry and Laboratory Medicine, 2003. **41**(4): p. 468-474.
236. Schubert M, et al. *Fluorescently labeled tracrRNA provides efficient genome editing while allowing cellular microscopy and FACS analysis*. 2017. Integrated DNA Technologies.
237. Maruyama T, et al., *Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining (vol 33, pg 538, 2015).* Nature Biotechnology, 2016. **34**(2): p. 210-210.
238. Nakao H, et al., *A possible aid in targeted insertion of large DNA elements by CRISPR/Cas in mouse zygotes.* Genesis, 2016. **54**(2): p. 65-77.
239. Li G, et al., *Small molecules enhance CRISPR/Cas9-mediated homology-directed genome editing in primary cells.* Sci Rep, 2017. **7**(1): p. 8943.
240. Yu C, et al., *Small molecules enhance CRISPR genome editing in pluripotent stem cells.* Cell Stem Cell, 2015. **16**(2): p. 142-7.
241. Miura H, et al., *Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors.* Nature Protocols, 2018. **13**(1): p. 195-215.
242. Hustedt N and D Durocher, *The control of DNA repair by the cell cycle.* Nat Cell Biol, 2016. **19**(1): p. 1-9.
243. Lin S, et al., *Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery.* Elife, 2014. **3**: p. e04766.
244. Gutschner T, et al., *Post-translational Regulation of Cas9 during G1 Enhances Homology-Directed Repair.* Cell Rep, 2016. **14**(6): p. 1555-1566.
245. Guo Q, et al., *'Cold shock' increases the frequency of homology directed repair gene editing in induced pluripotent stem cells.* Sci Rep, 2018. **8**(1): p. 2080.
246. Yao X, et al., *Tild-CRISPR Allows for Efficient and Precise Gene Knockin in Mouse and Human Cells.* Developmental Cell, 2018. **45**(4): p. 526-+.
247. Sakuma T, et al., *MMEJ-assisted gene knock-in using TALENs and CRISPR-Cas9 with the PITCh systems.* Nature Protocols, 2016. **11**(1): p. 118-133.
248. Gamsjaeger R, et al., *Structural basis and specificity of acetylated transcription factor GATA1 recognition by BET family bromodomain protein Brd3.* Mol Cell Biol, 2011. **31**(13): p. 2632-40.
249. Taverna SD and PA Cole, *Drug discovery: Reader's block.* Nature, 2010. **468**(7327): p. 1050-1.
250. Alekseyenko AA, et al., *Ectopic protein interactions within BRD4-chromatin complexes drive oncogenic megadomain formation in NUT midline carcinoma.* Proc Natl Acad Sci U S A, 2017. **114**(21): p. E4184-E4192.
251. Yan J, et al., *Perturbation of BRD4 protein function by BRD4-NUT protein abrogates cellular differentiation in NUT midline carcinoma.* J Biol Chem, 2011. **286**(31): p. 27663-75.
252. *Tocris*. Available from: https://www.tocris.com/.
253. *Chemical Probes*. Available from: http://www.chemicalprobes.org/.
254. *ATCC - Human Cell-Lines*. Available from: https://www.lgcstandards-atcc.org/Products/Cells_and_Microorganisms/Cell_Lines/Human/Alphanumeric.aspx.
255. Riching KM, et al., *Quantitative live-cell kinetic degradation and mechanistic profiling of PROTAC mode of action.* ACS Chem Biol, 2018.
256. Cuker A, et al., *Novel diagnostic assays for heparin-induced thrombocytopenia.* Blood, 2013. **121**(18): p. 3727-32.
257. Mullier F, et al., *Platelet microparticle generation assay: A valuable test for immune heparin-induced thrombocytopenia diagnosis.* Thrombosis Research, 2014. **133**(6): p. 1068-1073.
258. Ivetic N, et al., *Performance Characteristics of a Novel Platelet Viability Assay for Heparin-Induced Thrombocytopenia.* Blood, 2017. **130**.

259. Nakamura S, et al., *Expandable megakaryocyte cell lines enable clinically applicable generation of platelets from human induced pluripotent stem cells.* Cell Stem Cell, 2014. **14**(4): p. 535-48.
260. Savitski MM, et al., *Multiplexed Proteome Dynamics Profiling Reveals Mechanisms Controlling Protein Homeostasis.* Cell, 2018. **173**(1): p. 260-274 e25.
261. Jung KH, et al., *RNA sequencing reveals distinct mechanisms underlying BET inhibitor JQ1-mediated modulation of the LPS-induced activation of BV-2 microglial cells.* J Neuroinflammation, 2015. **12**: p. 36.
262. Qin C, et al., *Discovery of QCA570 as an Exceptionally Potent and Efficacious Proteolysis Targeting Chimera (PROTAC) Degrader of the Bromodomain and Extra-Terminal (BET) Proteins Capable of Inducing Complete and Durable Tumor Regression.* Journal of Medicinal Chemistry, 2018.
263. GenBank. *NM_001113182.2. Homo sapiens bromodomain containing 2 (BRD2), transcript variant 2, mRNA*. [cited 2018 December]; Available from: https://www.ncbi.nlm.nih.gov/nuccore/NM_001113182.2.
264. GenBank. *NM_007371.3. Homo sapiens bromodomain containing 3 (BRD3), mRNA.* [cited 2018 December]; Available from: https://www.ncbi.nlm.nih.gov/nuccore/NM_007371.3.
265. GenBank. *NM_001242806.2. Homo sapiens bromodomain testis associated (BRDT), transcript variant 4, mRNA.* [cited 2018 December]; Available from: https://www.ncbi.nlm.nih.gov/nuccore/NM_001242806.2.
266. Kabsch W, *Xds.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 125-32.
267. Battye TG, et al., *iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 271-81.
268. Evans PR and GN Murshudov, *How good are my data and what is the resolution?* Acta Crystallogr D Biol Crystallogr, 2013. **69**(Pt 7): p. 1204-14.
269. Winn MD, et al., *Overview of the CCP4 suite and current developments.* Acta Crystallogr D Biol Crystallogr, 2011. **67**(Pt 4): p. 235-42.
270. Vagin A and A Teplyakov, *Molecular replacement with MOLREP.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 22-5.
271. McCoy AJ, et al., *Phaser crystallographic software.* J Appl Crystallogr, 2007. **40**(Pt 4): p. 658-674.
272. Afonine PV, et al., *Towards automated crystallographic structure refinement with phenix.refine.* Acta Crystallogr D Biol Crystallogr, 2012. **68**(Pt 4): p. 352-67.
273. Adams PD, et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 2): p. 213-21.
274. Emsley P, et al., *Features and development of Coot.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
275. Moriarty NW, RW Grosse-Kunstleve, and PD Adams, *electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation.* Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 10): p. 1074-80.
276. Moriarty NW, EJ Draizen, and PD Adams, *An editor for the generation and customization of geometry restraints.* Acta Crystallogr D Struct Biol, 2017. **73**(Pt 2): p. 123-130.
277. Schuttelkopf AW and DM van Aalten, *PRODRG: a tool for high-throughput crystallography of protein-ligand complexes.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 8): p. 1355-63.
278. Chen VB, et al., *MolProbity: all-atom structure validation for macromolecular crystallography.* Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 1): p. 12-21.
279. Schrodinger, LLC, *The PyMOL Molecular Graphics System, Version 1.8*. 2015.

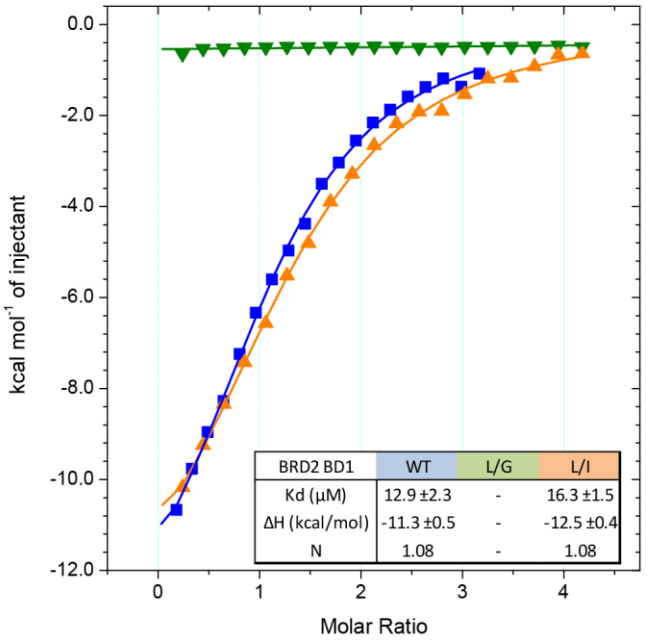
Appendix 1
Supplemental Data

Chapter 2 Supplementary Figures

A

BRD2 BD1	T_m (°C)	ΔT_m (°C)
WT	46.2 ±0.2	-
L/G	43.8 ±0.1	-2.4
L/I	43.8 ±0.0	-2.4

B



C

ET	BRD2 BD1		BRD2 BD2	
	WT	L/I	WT	L/I
Kd (nM)	1471 ±300	257 ±114	3130 ±2390	160 ±11
Selectivity	6x		20x	

Figure S2.1. Characterisation of L/G and L/I Mutant Bromodomains.

Unpublished data from Dr Michael Zengerle and Enrique Lin-Shiao. A) DSF-measured melting temperature of WT, L/I and L/G BRD2 BD1 bromodomain constructs. B) ITC titrations of tetra-acetylated H4 K(5,8,12,16)ac peptide into BRD2 BD1 bromodomain constructs, at 15°C. C) Results of ITC titrations of ET compound into WT and L/I BET bromodomain constructs.

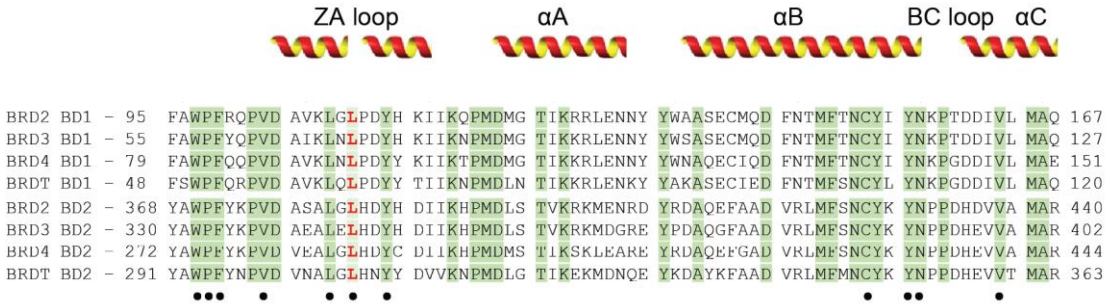


Figure S2.2. BET Bromodomain Sequence alignment

Sequence alignment of the eight human BET bromodomains, with positions of α–helices and associated loops. Conserved residues highlighted in green, with mutated leucine in red. Conserved residues making direct contact with I-BET shown with black dot. Ruler numbering based on BRD4 BD1 sequence.

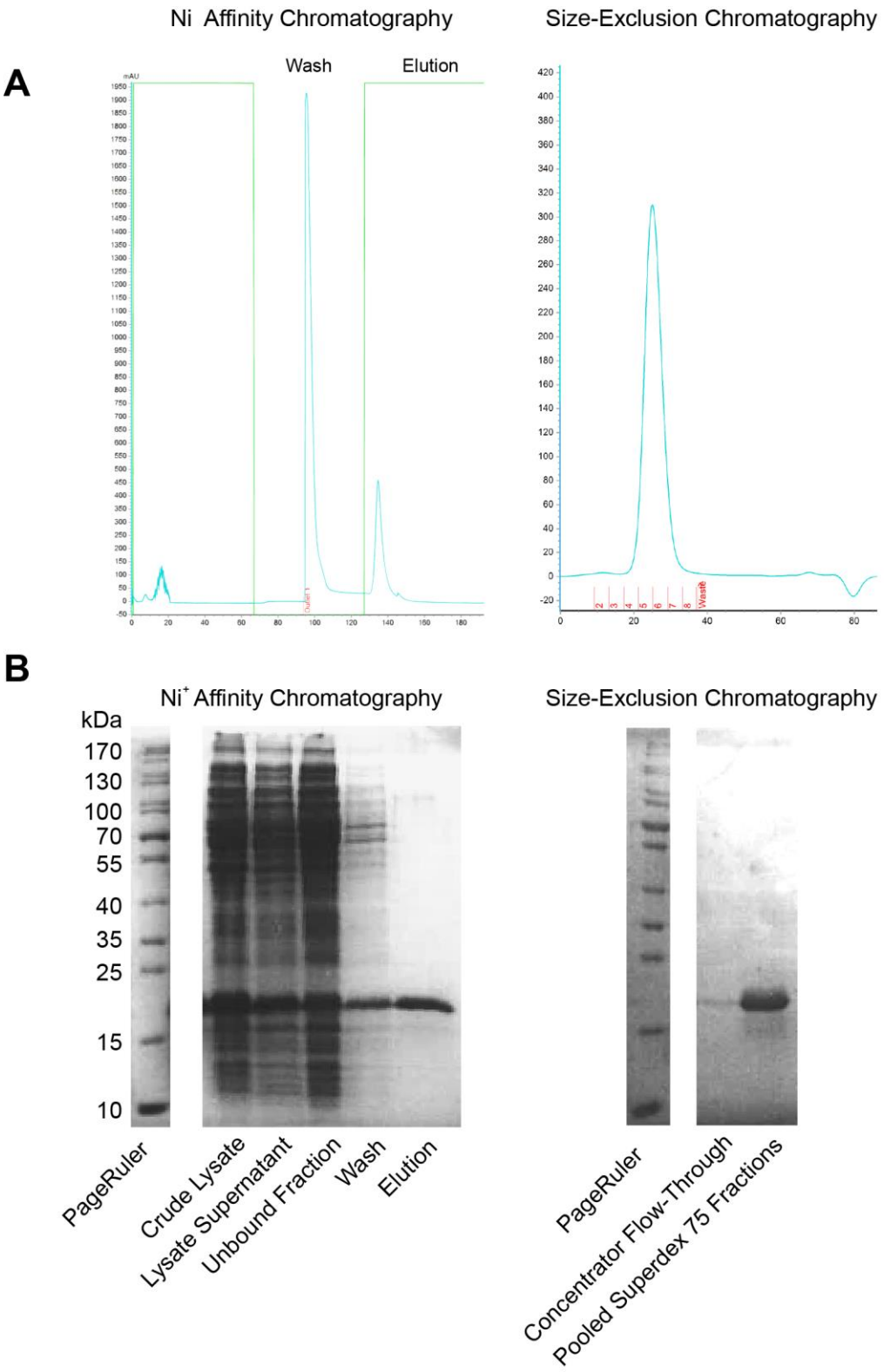


Figure S2.3. BET bromodomain construct purification.

A) UV Chromatograms from AKTApure purification. B) SDS-PAGE of purification samples.

Peptide	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2		BRD4 BD1		BRD4 BD2		BRDT BD1		BRDT BD2	
	WT	LV	WT	LV	WT	LV	WT	LV	WT	LV	WT	LV	WT	LV	WT	LV
H4 1-21	0.62	0.21	0.07	-0.03	0.14	-0.01	0.08	0.04	0.05	0.03	0.00	0.19	-0.23	0.00	0.06	-0.04
H4 1-21 K5ac	0.53	0.21	0.18	0.06	0.23	0.10	0.23	0.21	0.14	0.18	0.16	0.29	-0.11	0.07	0.19	-0.03
H4 1-21 K8ac	0.44	0.17	0.15	0.03	0.17	0.08	0.21	0.12	0.10	0.19	0.11	0.22	-0.17	0.03	0.09	-0.04
H4 1-21 K12ac	0.56	0.23	0.23	0.06	0.26	0.14	0.26	0.20	0.21	0.37	0.20	0.27	-0.01	0.12	0.13	-0.02
H4 1-21 K16ac	0.26	0.08	0.02	-0.02	0.09	0.02	0.09	0.06	0.04	0.10	0.03	0.15	-0.17	0.01	0.02	-0.04
H4 1-21 K5acK8ac	2.70	1.40	0.54	0.16	3.77	2.18	0.54	0.31	2.74	2.12	0.38	0.53	1.00	0.59	0.29	0.11
H4 1-21 K5K8ac	0.46	0.25	0.37	0.04	0.40	0.23	0.51	0.32	0.33	0.55	0.37	0.32	0.09	0.15	0.07	0.02
H4 1-21 K5K16ac	0.16	0.11	0.16	0.06	0.16	0.12	0.24	0.18	0.12	0.32	0.17	0.18	-0.05	0.07	0.07	-0.01
H4 1-21 K8K12ac	0.49	0.28	0.37	0.07	0.49	0.34	0.33	0.24	0.42	0.68	0.33	0.28	0.13	0.18	0.10	0.03
H4 1-21 K8K16ac	0.15	0.12	0.16	0.07	0.17	0.15	0.23	0.16	0.13	0.37	0.17	0.17	-0.05	0.07	0.11	0.00
H4 1-21 K12K16ac	0.52	0.29	0.29	0.09	0.66	0.43	0.32	0.20	0.59	0.82	0.26	0.24	0.17	0.17	0.08	0.03
H4 1-21 K5K8K12ac	3.38	2.01	1.10	0.33	3.93	2.60	0.89	0.42	3.12	2.72	0.65	0.64	1.43	0.86	0.31	0.17
H4 1-21 K8K12K16ac	1.06	0.75	0.62	0.18	1.26	0.90	0.53	0.29	1.22	1.33	0.38	0.37	0.51	0.43	0.18	0.08
H4 1-21 K5K12K16ac	0.79	0.51	0.56	0.18	0.99	0.66	0.57	0.38	0.81	1.20	0.43	0.35	0.38	0.30	0.15	0.05
H4 1-21 K5K8K16ac	2.51	1.54	0.69	0.21	3.65	2.28	0.62	0.31	2.88	2.37	0.35	0.49	1.10	0.62	0.20	0.14
H4 1-21 K5K8K12K16ac	3.31	2.12	1.12	0.29	4.34	2.86	0.80	0.36	3.11	3.05	0.55	0.56	1.61	0.95	0.26	0.23
H4 9-29	0.04	0.04	0.01	-0.02	0.07	0.04	0.01	0.06	0.03	0.05	0.02	0.04	-0.09	0.01	0.03	0.00
H4 9-29 K20ac	0.11	0.18	0.26	0.17	0.10	0.14	0.34	0.46	0.13	0.38	0.22	0.25	0.11	0.33	0.06	0.11
H4 9-29 K16ac	0.05	0.06	0.02	0.01	0.04	0.05	0.05	0.06	0.03	0.09	0.03	0.03	0.01	0.04	0.01	0.02
H4 9-29 K12K16ac	0.36	0.47	0.27	0.12	0.40	0.41	0.23	0.18	0.23	0.63	0.18	0.13	0.23	0.32	0.06	0.10
H4 9-29 K12K16K20ac	0.50	0.90	0.48	0.23	0.53	0.66	0.35	0.32	0.33	1.07	0.22	0.23	0.33	0.68	0.06	0.17
H2a 1-21	0.45	0.09	0.04	-0.02	0.09	0.03	0.01	0.08	0.03	0.07	0.02	0.21	-0.14	0.01	0.04	-0.03
H2a 1-21 K5ac	0.54	0.23	0.17	0.04	0.22	0.04	0.18	0.17	0.14	0.17	0.16	0.32	-0.12	0.06	0.11	-0.02
H2a 1-21 K9ac	0.50	0.20	0.20	0.03	0.22	0.03	0.16	0.12	0.11	0.13	0.13	0.25	-0.16	0.04	0.10	-0.03
H2a 1-21 K13ac	0.36	0.13	0.07	0.01	0.16	0.05	0.03	0.08	0.07	0.08	0.04	0.20	-0.12	0.04	0.09	-0.02
H2a 1-21 K15ac	0.33	0.10	0.07	0.00	0.15	0.04	0.04	0.07	0.06	0.07	0.04	0.17	-0.15	0.03	0.08	-0.03
H2a 1-21 K5K9ac	0.24	0.12	0.30	0.10	0.22	0.13	0.34	0.24	0.16	0.28	0.28	0.31	-0.07	0.08	0.12	0.01
H2a 1-21 K5K13ac	0.10	0.07	0.14	0.03	0.17	0.09	0.20	0.17	0.11	0.18	0.15	0.19	-0.11	0.05	0.08	0.00
H2a 1-21 K5K15ac	0.12	0.07	0.07	0.09	0.14	0.07	0.20	0.17	0.09	0.19	0.07	0.21	-0.10	0.05	0.07	-0.01
H2a 1-21 K9K13ac	0.18	0.09	0.17	0.04	0.22	0.11	0.21	0.07	0.13	0.24	0.18	0.17	-0.09	0.06	0.09	0.00
H2a 1-21 K9K15ac	0.16	0.12	0.29	0.09	0.22	0.14	0.29	0.18	0.16	0.29	0.27	0.21	-0.03	0.06	0.11	0.02
H2a 1-21 K13K15ac	0.10	0.04	0.04	-0.01	0.13	0.06	0.05	0.07	0.05	0.10	0.06	0.09	-0.12	0.03	0.04	-0.02
H2a 1-21 K5K9K13ac	0.02	0.05	0.00	0.03	0.01	0.06	0.10	0.04	0.01	0.05	0.05	0.03	0.07	0.05	0.03	0.07
H2a 1-21 K5K13K15ac	0.12	0.08	0.17	0.06	0.15	0.11	0.20	0.17	0.13	0.23	0.17	0.15	0.02	0.05	0.03	0.00
H2a 1-21 K9K13K15ac	0.20	0.13	0.30	0.08	0.30	0.11	0.19	0.12	0.18	0.31	0.21	0.13	-0.03	0.06	0.12	-0.01
H2a 1-21 K5K9K15ac	0.19	0.13	0.18	0.15	0.32	0.13	0.40	0.20	0.22	0.35	0.35	0.26	-0.02	0.08	0.11	-0.02
H2a 1-21 K5K9K13K15ac	0.30	0.23	0.51	0.15	0.35	0.28	0.40	0.23	0.29	0.49	0.35	0.23	0.14	0.12	0.12	0.05
H2b 1-21	0.02	0.00	0.00	0.00	0.08	0.02	-0.02	0.04	0.01	0.01	0.02	0.04	-0.14	0.02	0.04	-0.02
H2b 1-21 K5ac	0.01	0.02	0.06	0.07	0.06	0.05	-0.02	0.06	0.02	0.06	0.02	0.02	-0.03	0.03	0.06	0.01
H2b 1-21 K11ac	0.00	0.00	0.00	0.01	0.00	0.02	0.00	0.04	0.01	0.01	0.01	0.02	-0.04	0.02	0.02	-0.01
H2b 1-21 K12ac	0.01	0.02	0.00	0.00	0.02	0.04	0.03	0.06	0.02	0.05	0.03	0.02	-0.01	0.02	0.04	0.01
H2b 1-21 K15ac	0.01	0.04	0.00	0.01	0.03	0.06	0.05	0.06	0.02	0.08	0.02	0.03	0.00	0.04	0.04	0.01
H2b 1-21 K5K11K16ac	0.04	0.04	0.02	0.03	0.04	0.08	0.04	0.04	0.03	0.10	0.03	0.03	0.04	0.04	0.04	0.02
H2b 1-21 k5k12AC	0.02	0.01	0.00	0.02	0.00	0.03	0.02	0.03	0.01	0.02	0.02	0.02	0.00	0.03	0.01	0.00
H2b 1-21 K5K15ac	0.01	0.03	0.01	0.01	0.02	0.06	0.06	0.04	0.02	0.08	0.02	0.03	0.01	0.04	0.02	0.01
H2b 1-21 K5K16ac	0.02	0.03	0.01	0.02	0.04	0.05	0.03	0.03	0.02	0.09	0.02	0.02	0.03	0.03	0.00	0.01
H2b 1-21 K11K12ac	0.03	-0.02	0.06	0.01	0.03	-0.02	0.01	0.00	0.02	0.00	0.01	0.00	-0.02	0.00	0.00	-0.01
H2b 1-21 K11K15ac	0.04	0.03	0.11	0.05	0.07	0.02	0.04	0.01	0.03	0.07	0.02	0.01	0.00	0.02	0.01	0.00
H2b 1-21 K11K16ac	0.04	0.05	0.03	0.04	0.00	0.05	0.04	0.03	0.03	0.08	0.02	0.02	0.02	0.03	0.03	0.01
H2b 1-21 K5K11K12ac	0.03	0.01	0.05	0.03	0.03	0.04	0.04	0.03	0.02	0.05	0.02	0.02	0.01	0.01	0.03	0.01
H2b 1-21 K5K15K16ac	0.03	0.06	0.06	0.03	0.04	0.06	0.03	0.03	0.04	0.09	0.03	0.03	0.04	0.04	0.02	0.02
H2b 1-21 K5K15K16ac	0.26	0.12	0.17	0.03	0.50	0.20	0.06	0.03	0.17	0.17	0.03	0.03	0.13	0.06	0.04	0.03
H2b 1-21 K11K15K16ac	0.02	0.04	0.02	0.00	0.05	0.08	0.04	0.03	0.03	0.08	0.02	0.02	0.04	0.05	0.02	0.02
H2b 1-21 K12K15K16ac	0.47	0.17	0.20	0.04	1.02	0.30	0.07	0.02	0.45	0.34	0.05	0.03	0.30	0.10	0.03	0.02
H2b 1-21 K11K12K15ac	0.17	0.10	0.13	0.04	0.30	0.16	0.06	0.03	0.11	0.15	0.03	0.03	0.11	0.06	0.04	0.03
H2b 1-21 K11K12K16ac	0.04	0.04	0.01	0.02	0.06	0.07	0.05	0.04	0.04	0.07	0.02	0.03	0.05	0.04	0.02	0.02
H2b 1-21 K5K11K12K15K16ac	0.07	0.12	0.01	0.06	0.01	0.05	0.09	0.04	0.03	0.03	0.03	0.01	0.07	0.09	0.01	0.05
H2b 13-33	0.73	0.14	0.09	-0.02	0.18	0.06	0.02	0.11	0.05	0.08	0.02	0.25	-0.25	0.02	0.08	-0.04
H2b 13-33 K20ac	1.09	0.39	0.18	0.01	0.19	0.04	0.04	0.10	0.09	0.09	0.03	0.29	-0.28	0.03	0.12	-0.02
H2b 13-33 K23ac	0.88	0.35	0.14	-0.01	0.20	0.02	0.07	0.09	0.06	0.05	0.02	0.26	-0.27	0.02	0.10	-0.05
H2b 13-33 K24ac	0.83	0.26	0.11	0.00	0.14	0.06	0.02	0.10	0.07	0.10	0.02	0.26	-0.24	0.03	0.11	-0.03
H2b 13-33 K20K23K24ac	0.43	0.18	0.23	0.02	0.23	0.09	0.08	0.10	0.09	0.12	0.07	0.19	-0.21	0.05	0.10	-0.05
H3 1-21	0.09	0.02	0.01	0.01	0.08	0.03	-0.01	0.06	0.03	0.02	0.01	0.10	-0.10	0.02	0.04	-0.02
H3 1-21 K4ac	0.02	0.02	0.01	0.00	0.10	0.05	0.03	0.07	0.03	0.07	0.04	0.06	-0.13	0.03	0.04	-0.01
H3 1-21 K9ac	0.02	0.01	-0.01	0.00	0.07	0.05	0.03	0.06	0.02	0.04	0.02	0.05	-0.11	0.02	0.04	0.00
H3 1-21 K14ac	0.07	0.03	0.06	0.03	0.12	0.06	0.10	0.10	0.04	0.09	0.10	0.10	-0.09	0.03	0.05	0.00
H3 1-21 K4K9ac	0.04	0.05	0.05	0.02	0.07	0.09	0.06	0.09	0.06	0.13	0.06	0.06	0.00	0.05	0.06	0.02
H3 1-21 K4K14ac	0.07	0.04	0.08	0.03	0.10	0.07	0.10	0.12	0.05	0.12	0.11	0.08	-0.02	0.04	0.03	0.00
H3 1-21 K9K14ac	0.05	0.04	0.04	0.02	0.04	0.08	0.05	0.05	0.03	0.10	0.03	0.03	0.05	0.03	0.03	0.02
H3 1-21 K4K9K14ac	0.11	0.07	0.10	0.04	0.09	0.09	0.15	0.10	0.07	0.17	0.10	0.07	0.06	0.05	0.05	0.02
H3 11-31	0.01	-0.03	0.05	0.01	0.11	-0.01	-0.03	0.05	0.02	0.02	0.00	0.04	-0.17	-0.01	0.04	-0.03
H3 11-31 K14ac	0.03	-0.02	0.04	0.00	0.09	0.01	0.00	0.05	0.01	0.03	0.04	0.03	-0.07	-0.01	0.02	-0.03
H3 11-31 K18ac	0.04	0.03	0.11	0.06	0.07	0.08	0.15	0.14	0.06	0.16	0.19	0.10	-0.01	0.04	0.13	0.04
H3 11-31 K23ac	0.11	0.08	0.03	0.00	0.14	0.04	0.01	0.09	0.05	0.06	0.02	0.12	-0.22	0.01	0.08	-0.03
H3 11-31 K27ac	0.01	0.01	0.01	0.00	0.04	0.04	0.00	0.07	0.02	0.06	0.03	0.03	-0.02	0.01	0.04	0.00
H3 11-31 K14K18ac	0.06	0.04	0.13	0.06	0.08	0.10	0.18	0.12	0.05	0.15	0.15	0.09	0.03	0.05	0.11	0.05
H3 11-31 K14K23ac	0.04	0.02	0.03	0.02	0.03	0.06	0.07	0.05	0.04	0.07	0.04	0.04	0.03	0.02	0.03	0.02
H3 11-31 K14K27ac	0.05	0.02	0.08	0.07	0.03	0.06	0.07	0.05	0.02	0.09	0.05	0.04	0.03	0.03	0.03	0.02
H3 11-31 K18K23ac	0.06	0.05	0.10	0.05	0.05	0.09	0.13	0.07	0.05	0.15	0.13	0.07	0.05	0.05	0.06	0.03
H3 11-31 K18K27ac	0.05	0.04	0.13	0.06	0.08	0.12	0.19	0.12	0.07	0.18	0.18	0.11	0.06	0.07	0.10	0.06
H3 11-31 K23K27ac	0.03	0.04	0.01	0.02	0.02	0.06	0.04	0.04	0.02	0.08	0.02	0.03	0.05	0.03	0.02	0.02
H3 11-31 K14K23K27ac	0.10	0.06	0.07	0.02	0.07	0.10	0.08	0.05	0.05	0.11	0.05	0.03	0.08	0.05	0.03	0.03
H3 11-31 K18K23K27ac	0.16	0.13	0.30	0.11	0.14	0.10	0.24	0.09	0.12	0.22	0.23	0.10	0.08	0.11	0.11	0.06
H3 11-31 K14K18K27ac	0.11	0.06	0.14	0.05	0.12	0.06	0.16	0.05	0.07	0.15	0.15	0.06	0.03	0.03	0.08	0.03
H3 11-31 K14K18K23K27ac	0.16	0.13	0.29	0.10	0.14	0.13	0.18	0.09	0.10	0.20	0.18	0.08	0.10	0.09	0.11	0.06
H3 23-43	0.10	0.08	0.05	0.02	0.13	0.07	0.02	0.09	0.05	0.05	0.02	0.12	-0.22	-0.01	0.06	-0.04
H3 23-43 K36ac	0.04	0.06	0.11	0.03	0.11	0.07	0.19	0.14	0.06	0.08	0.11	0.09	-0.08	0.04	0.07	-0.01
H3 23-43 K37ac	0.02	0.02	0.11	0.00	0.11	0.09	0.05	0.09	0.04	0.06	0.09	0.06	-0.08	-0.01	0.08	-0.02
H3 23-43 K36K37ac	0.05	0.04	0.23	0.04	0.05	0.07	0.18	0.11	0.05	0.08	0.16	0.07	-0.01	0.04	0.10	0.01
H3 23-43 K27ac	0.02	0.02	0.01	0.00	0.08	0.08	0.01	0.08	0.03	0.07	0.04	0.05	-0.09	0.00	0.03	-0.02
H3 23-43 K27K36ac	0.06	0.05	0.17	0.06	0.07	0.10	0.23	0.14	0.07	0.14	0.11	0.08	0.00	0.06	0.08	0.02
H3 23-43 K27K37ac	0.05	0.03	0.19	0.03	0.07	0.09	0.16	0.10	0.06	0.10	0.14	0.09	-0.01	0.00	0.08	0.01
H3 23-43 K27K36K37ac	0.09	0.07	0.31	0.08	0.06	0.09	0.22	0.07	0.07	0.12	0.17	0.08	0.05	0.05	0.11	0.03

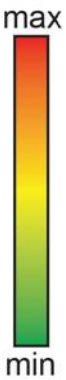


Figure S2.4. Full BET bromodomain histone peptide binding profiles.

Binding profiles of WT and L/V BET bromodomain constructs for acetylated histone peptides, derived from the BLI screen. Average response of two negative controls (biocytin-loaded tips) subtracted, to account for any non-specific binding. Responses color-coded to maximum and minimum response of each construct.

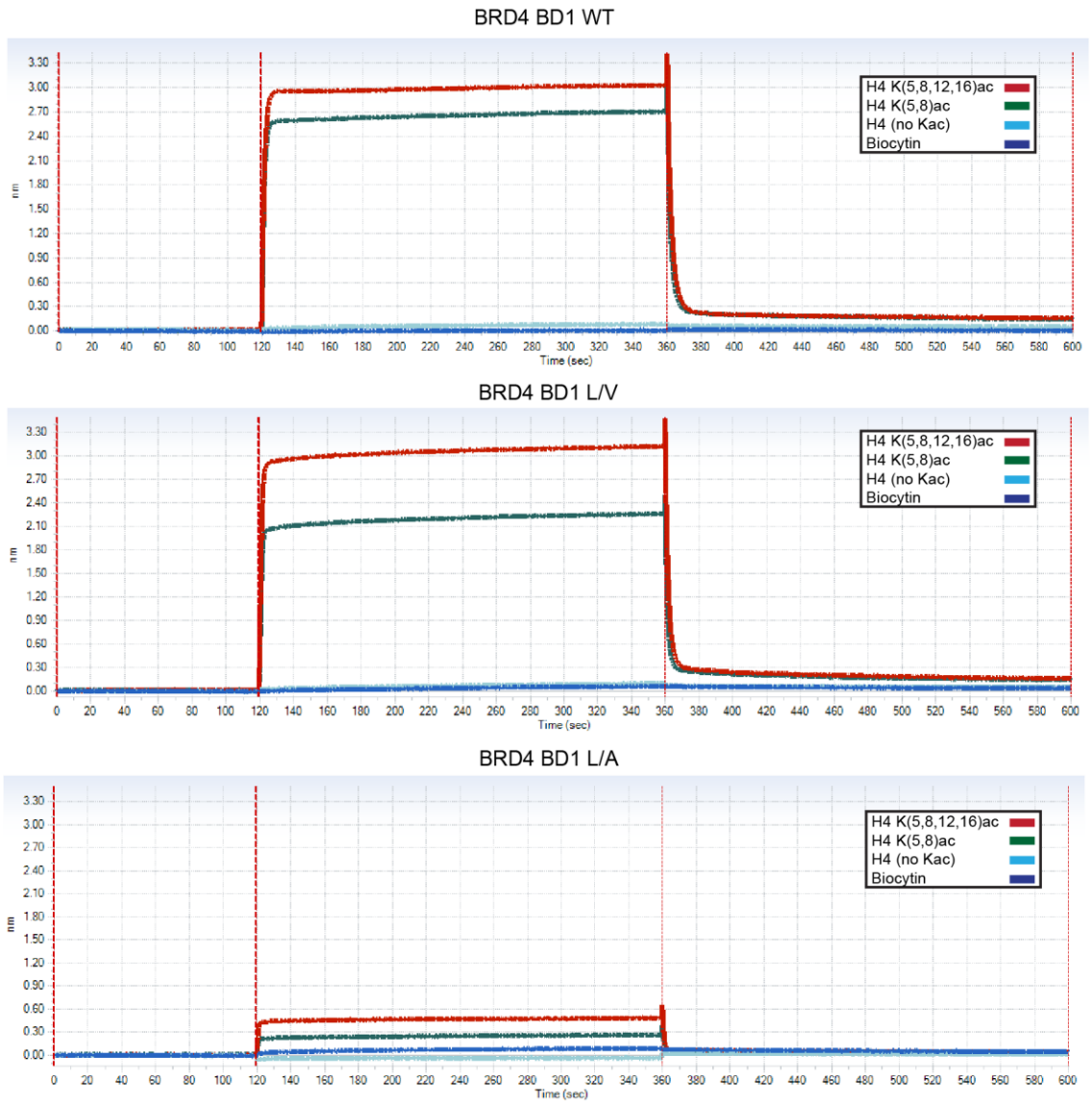


Figure S2.5. Sample BLI curves.

Streptavidin-labelled BLI tips were loaded with biocytin or biotinylated histone peptides, and then submerged in buffer for 120s, 20 µM bromodomain solutions for 120s and then buffer again for 120s.

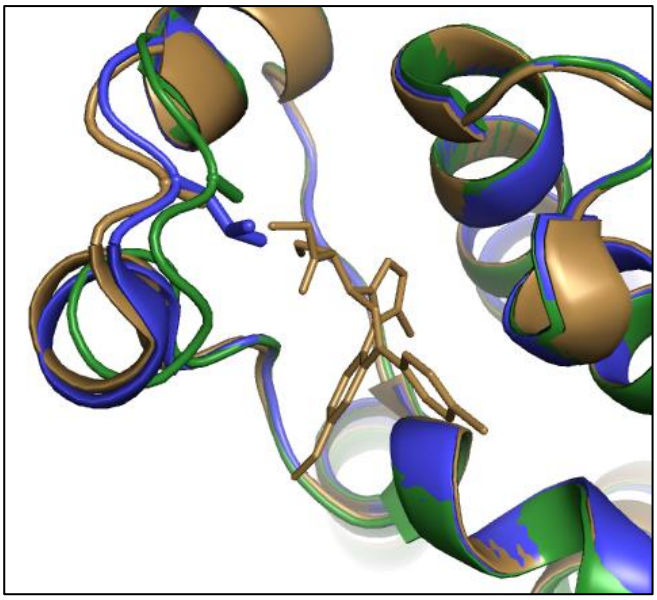


Figure S2.6. Crystallographic analysis of ZA loop during binding

Aligned X-ray crystal structures of BRD2 BD2 L/A *apo* (green) (4QEU) and bound to ET (brown) (4QEW), as well as BRD2 BD2 WT bound to an acetylated H4 peptide (blue) (2E3K). Bound peptide not shown. Mutated leucine/alanine side-chain highlighted.

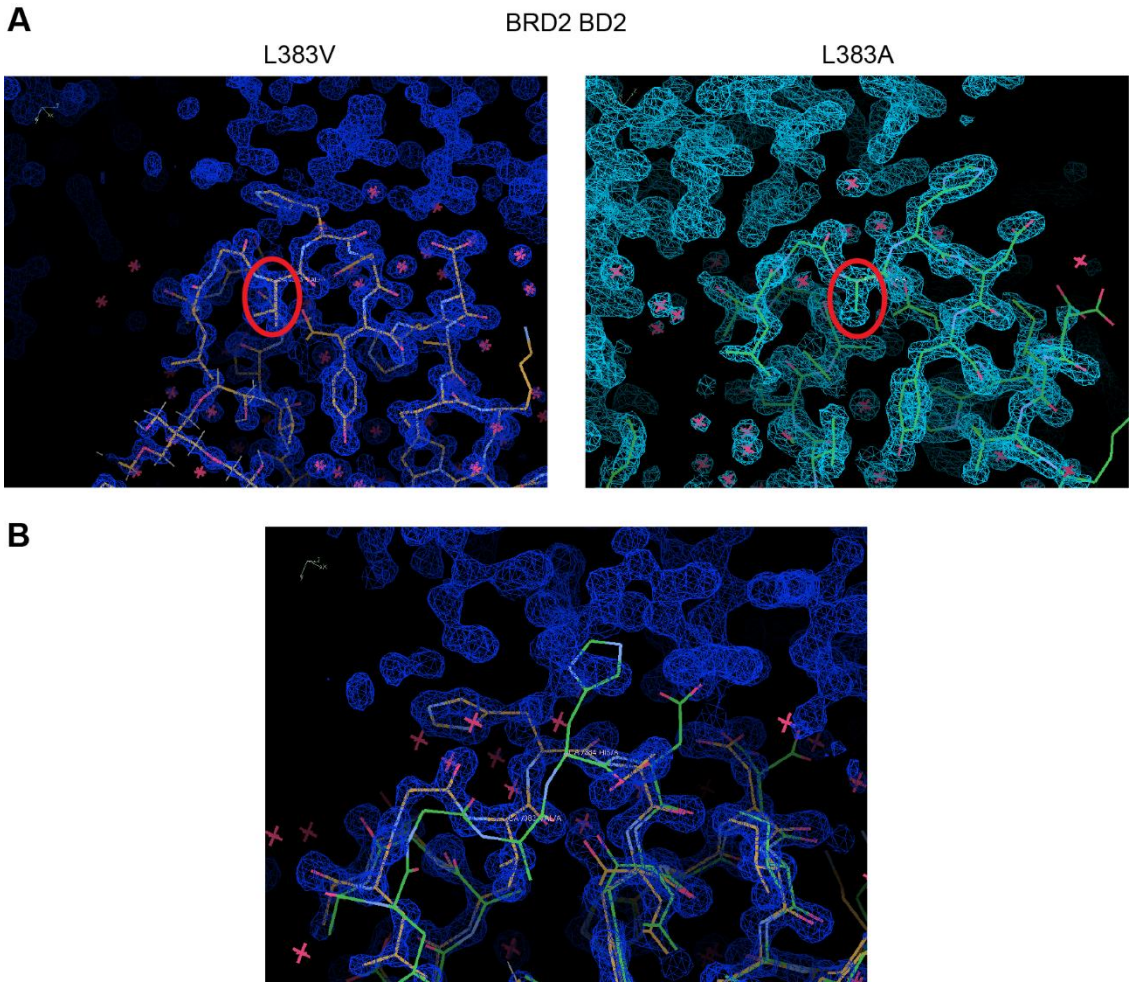


Figure S2.7. BRD2 BD2 mutant alternate crystal packing.

A) Coot screenshot of a BRD2 BD2 *apo* structure, and MTZ map, for L/V (left) and L/A (right) mutants. Valine/alanine residues are circled in red. Crystal contacts (electron density map lacking fitted atoms) are observed interacting with the ZA loop from clearly different orientations. B) BRD BD2 L/V *apo* structure and MTZ map, with L/A structure superimposed. A potential steric clash, between the L/A BD's H384 and Y385 side-chains and the L/V BD's crystal contacts, can be observed.

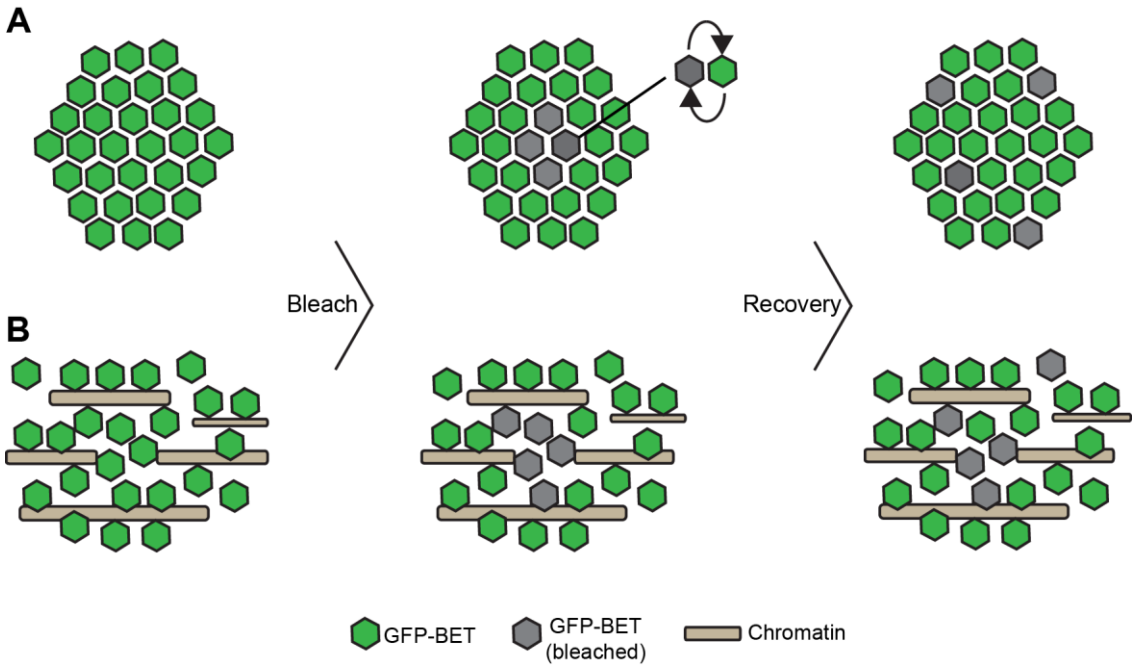
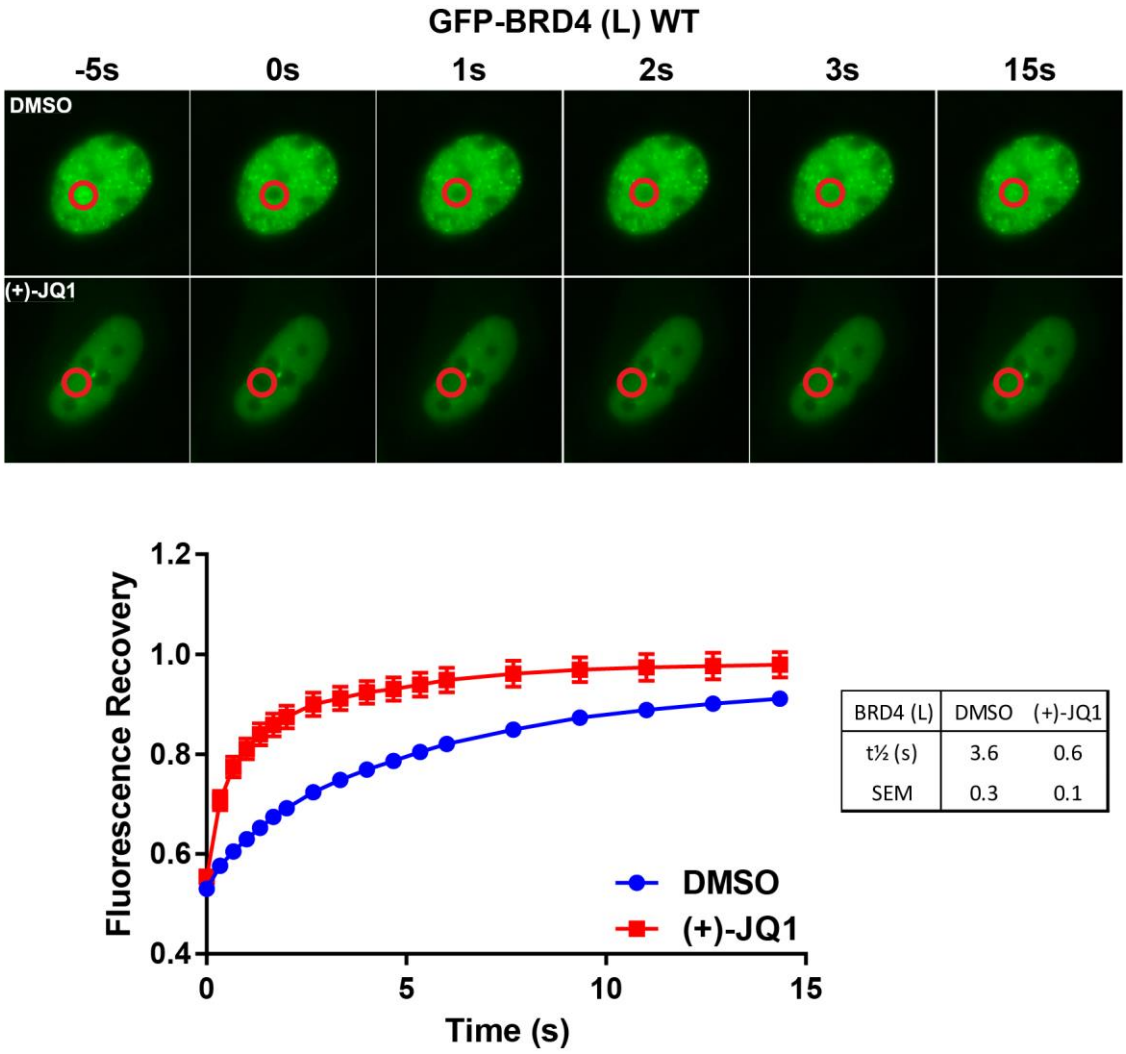


Figure S2.8. FRAP theory and impact of chromatin binding.

A) In group of free GFP-BET, bleached and unbleached protein can move freely in and out of bleach spot, leading to rapid recovery. B) When GFP-BET is mainly bound to immobile chromatin protein cannot move freely in or out of bleach spot and recovery is slow.



BRD4 (L)	DMSO	(+)-JQ1
t½ (s)	3.6	0.6
SEM	0.3	0.1

Figure S2.9. Example FRAP datasets.

Nuclei of U2OS cells transiently overexpressing GFP-BRD4 (L) and treated with 2 µM SAHA, 0.03% DMSO and 1 µM (+)-JQ1. Nuclei are bleached with laser at 0s (red circle) and imaged while the fluorescence recovers. Cells shown are representative of population from B. B) Fluorescence recovery plots for ~20 U2OS cells per treatment, with mean t½. Data points after 15s omitted for clarity.

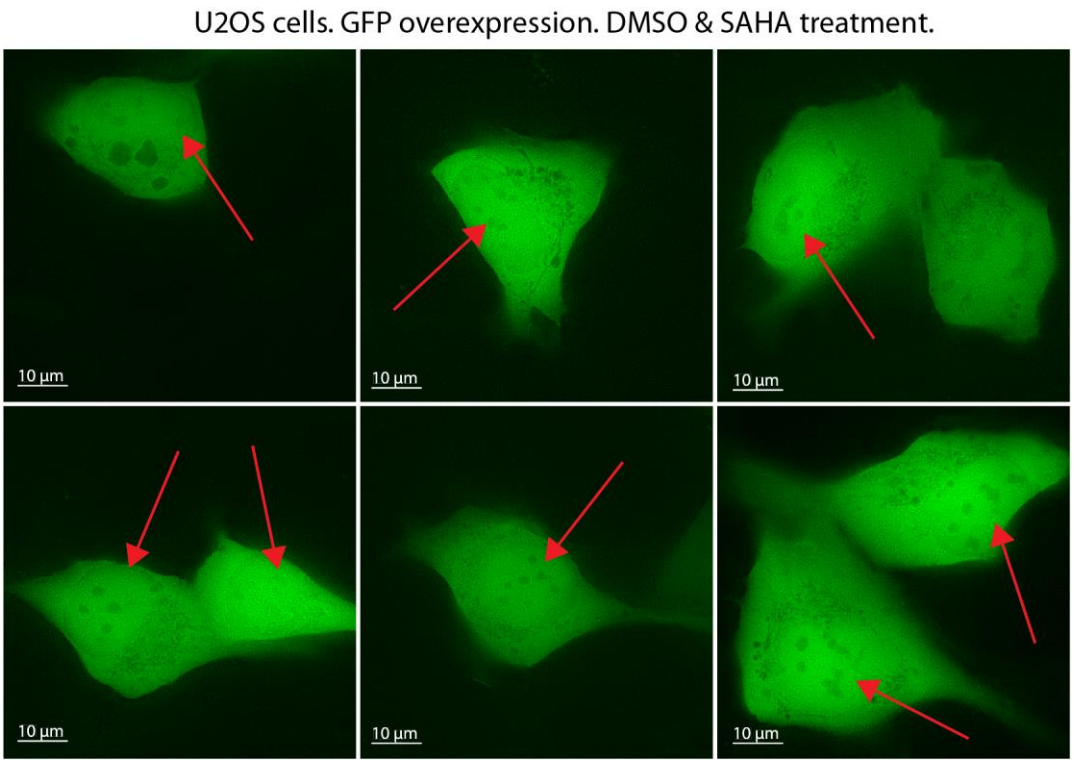
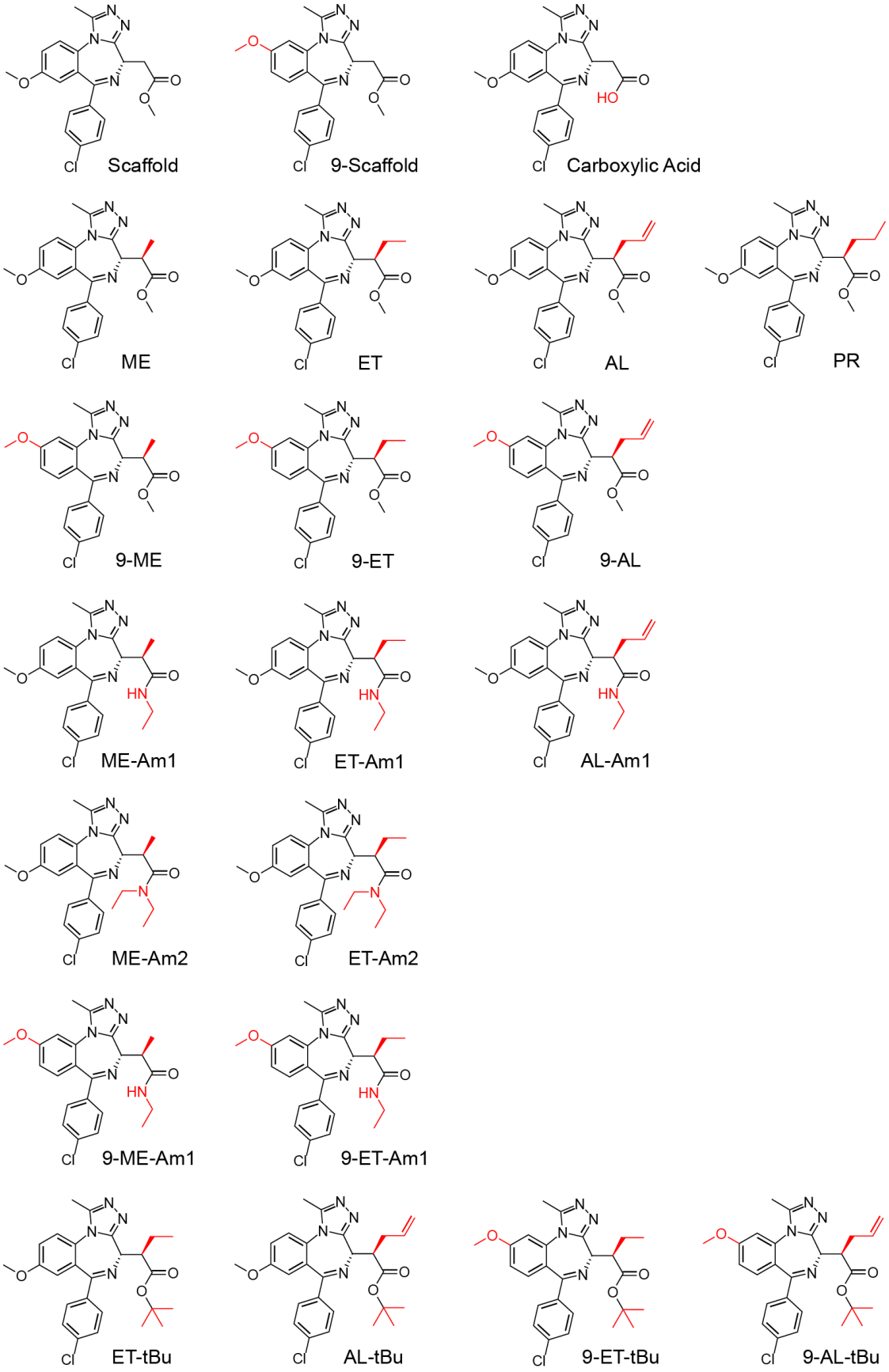


Figure S2.10. GFP cellular localisation.

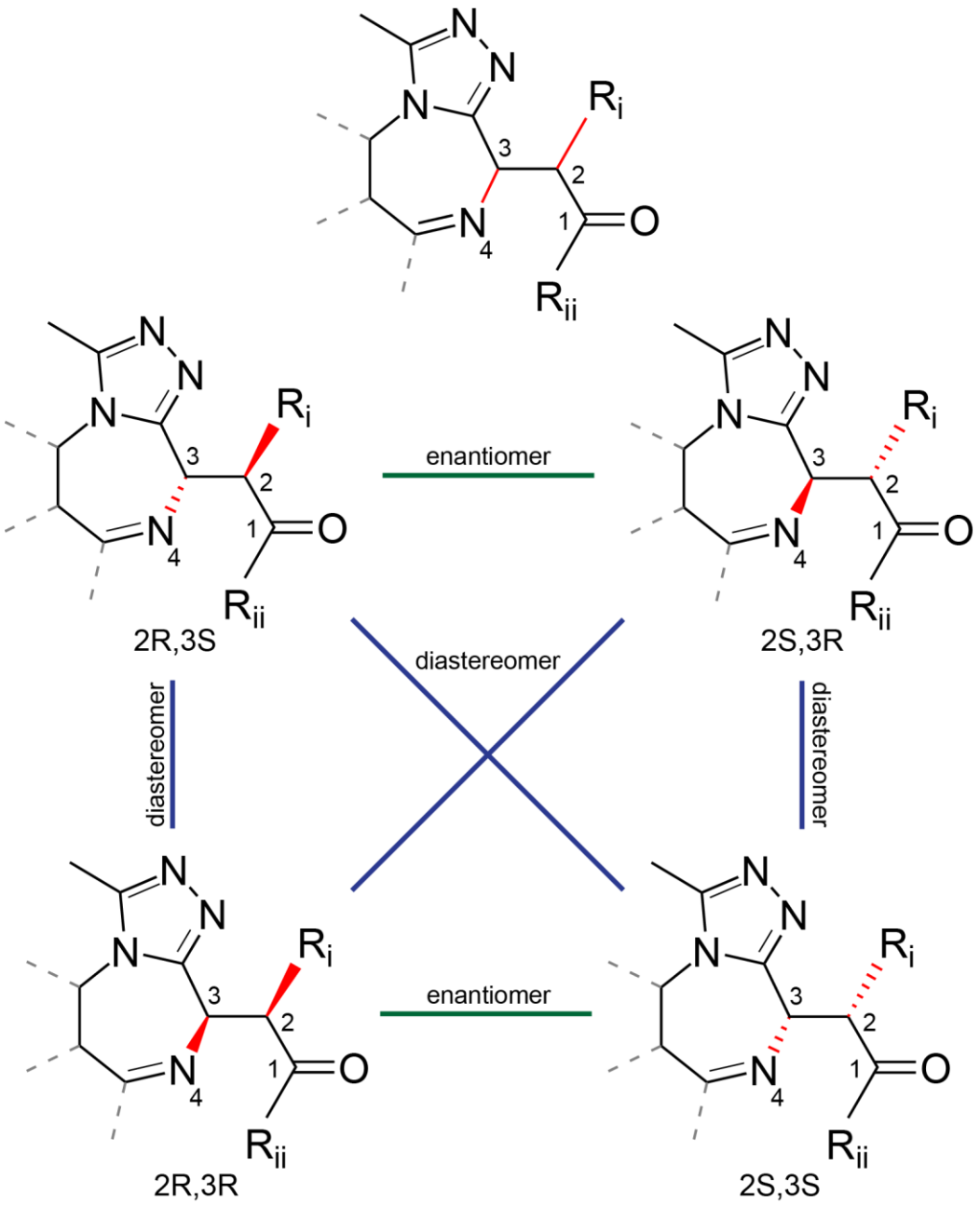
U2OS cells transfected with monomeric GFP and treated with 2 µM SAHA and 0.03% DMSO. Visually identifiable nuclei highlighted with red arrows.

Chapter 3 Supplementary Figures



Scheme S3.1 Bumped compounds.

Modifications over scaffold compounds are shown in red.



Scheme S3.2 Bumped compound isomers structures and relationships

H4 K(5,8)ac	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V
EC_{50} (nM)	58	77	61	92	77	78	84	27.0
95% C.I.	40-80	50-120	50-80	50-120	70-90	50-110	70-100	20-40
LV/WT	1.3		1.5		1.0		0.3	

H4 K(5,8)ac	BRD4 BD1		BRD4 BD2		BRDT BD1		BRDT BD2	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V
EC_{50} (nM)	73.0	81.0	122	73	47	69.0	39	64
95% C.I.	60-90	60-110	90-160	60-90	40-60	50-100	30-50	50-80
LV/WT	1.1		0.6		1.5		1.6	

Table S3.1. AlphaLISA biotinylated-peptide titrations

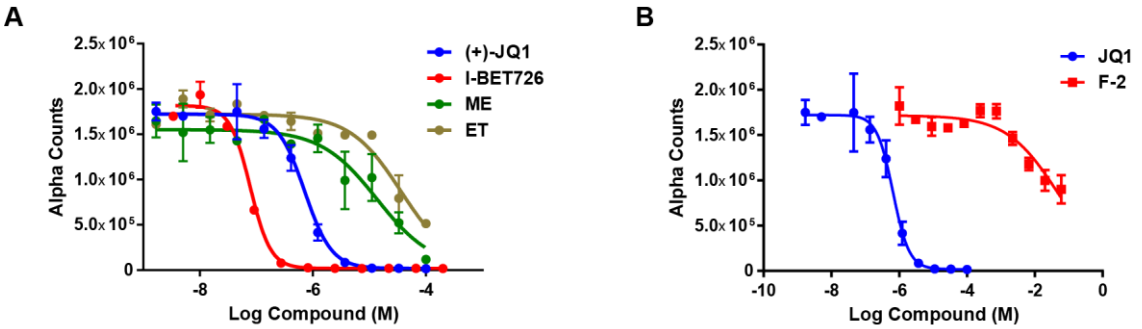
EC_{50} values of Bio-JQ1 against each WT and L/V BET bromodomain, in AlphaLISA assay.

Bio-JQ1	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V
EC_{50} (nM)	1.4	3.2	2.3	1.8	0.9	1.2	1.9	4.0
95% C.I.	1.0-2.6	1.2-2.1	1.6-3.2	1.6-2.2	0.6-1.3	1.1-1.4	1.7-2.3	3.7-4.4
LV/WT	2.3		0.8		1.3		2.1	

Bio-JQ1	BRD4 BD1		BRD4 BD2		BRDT BD1		BRDT BD2	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V
EC_{50} (nM)	1.0	2.0	1.3	2.9	3.2	4.0	3.1	4.9
95% C.I.	0.7-1.5	1.6-2.6	0.8-2.2	2.7-3.8	2.7-3.7	3.7-4.4	2.4-3.9	3.9-6.1
LV/WT	2.0		2.2		1.3		1.6	

Table S3.2. AlphaLISA Bio-JQ1 titrations.

EC_{50} values of Bio-JQ1 against each WT and L/V BET bromodomain, in AlphaLISA assay.



C

BRD2 BD2	(+)-JQ1	I-BET726	ME	ET	F2
pIC_{50}	6.2	7.1	4.9	4.4	~1.4
SE	0.1	0.0	0.1	0.1	

D

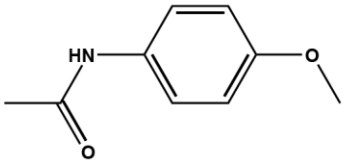


Figure S3.1. Bio-JQ1 AlphaLISA assay accurately ranks inhibitors and can detect fragment binding.

A) AlphaLISA titrations of BET inhibitors against Bio-JQ1 and BRD2 BD2 WT. B) AlphaLISA titrations of (+)-JQ1 and fragment F-2 (acetaminophen) against Bio-JQ1 and BRD2 BD2 WT. C) Results of previous titrations. D) Chemical structure of fragment F-2 (acetaminophen).

Compound	AlphaLISA pIC_{50}															
	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2		BRD4 BD1		BRD4 BD2		BRDT BD1		BRDT BD2	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V
(+)JQ1	6.4	6.5	6.5	6.5	6.8	6.6	6.5	6.4	6.6	6.5	6.5	6.4	6.6	6.0	6.3	6.1
(-)JQ1	4.8	4.8	4.8	4.4	5.0	5.0	4.6	4.4	4.9	4.6	4.7	4.5	4.7	4.6	4.6	4.6
I-BET762	6.2	6.1	6.3	6.7	6.7	6.9	6.6	6.3	6.6	6.3	6.5	6.5	6.3	6.0	6.3	6.3
Scaffold	6.5	6.6	6.5	6.5	6.9	7.2	6.6	6.3	6.9	6.6	6.4	6.6	6.3	6.2	6.6	6.0
9-Scaffold	5.7	5.8	5.6	5.7	6.2	6.3	5.6	6.0	6.0	6.0	5.6	5.6	5.7	5.8	5.4	5.5
Carboxylic Acid	4.0	5.0	4.4	5.6	4.9	5.8	5.1	5.8	4.6	5.4	5.0	5.7	3.9	4.9	4.6	5.4
ME	5.7	7.2	6.0	7.3	6.6	7.5	6.3	7.5	6.2	7.4	6.1	7.3	5.7	7.3	5.9	7.1
ET	4.9	6.5	5.0	6.8	5.2	7.0	5.0	7.0	5.1	6.6	4.9	6.6	4.8	6.3	4.8	6.2
AL	5.0	6.0	5.0	6.8	5.4	6.9	5.0	6.9	5.1	6.5	5.0	6.6	4.7	5.7	4.9	6.2
PR	4.5	5.0	<4	5.0	4.7	5.2	4.7	5.0	4.6	5.1	4.2	4.9	4.5	4.6	4.9	6.2
9-ME	5.5	6.7	5.1	6.5	5.8	7.0	5.3	7.0	5.5	6.7	5.2	6.7	5.3	6.6	5.1	6.2
9-ET	5.0	6.5	5.4	6.7	5.6	6.9	5.4	7.1	5.8	6.9	5.5	6.7	4.8	6.2	5.4	6.4
9-AL	4.3	5.4	4.4	5.9	4.5	5.8	4.3	6.0	4.4	5.5	4.2	5.5	4.2	5.2	4.3	5.3
ME-Am1	4.9	6.4	5.3	6.6	5.7	7.0	5.4	7.1	5.4	6.6	5.3	6.6	4.9	6.2	5.1	6.4
ET-Am1	4.2	5.5	4.6	5.9	4.5	6.3	4.4	6.1	4.3	5.7	4.2	5.7	4.2	4.9	4.1	5.3
AL-Am1	4.3	5.6	4.3	6.0	4.5	6.1	4.3	5.9	4.4	5.9	4.2	5.7	4.1	5.3	4.1	5.3
ME-Am2	4.9	6.0	4.6	6.0	5.3	6.4	4.9	6.3	4.9	5.9	4.6	5.9	4.4	5.4	4.6	5.5
ET-Am2	4.5	5.7	4.8	5.8	4.7	6.3	4.6	5.8	4.6	5.9	4.5	5.5	4.4	5.2	4.4	5.2
9-ME-Am1	4.4	5.7	4.8	6.0	5.3	6.3	4.8	6.3	4.6	6.0	4.5	5.9	4.3	5.6	4.6	5.8
9-ET-Am1	<4	5.2	<4	5.1	4.2	5.5	<4	5.4	4.2	5.1	<4	4.9	<4	4.7	<4	4.7
ET-tBu	5.3	5.8	5.3	6.0	5.6	6.4	5.3	6.1	5.3	6.1	5.2	5.8	5.2	5.6	5.2	5.6
AL-tBu	5.3	5.9	5.1	6.3	5.4	6.5	5.2	6.1	5.3	6.2	5.4	6.0	5.1	5.7	5.0	5.8
9-ET-tBu	4.4	5.0	4.7	4.9	4.7	5.3	4.6	5.0	4.8	5.1	4.6	4.7	4.6	4.9	4.4	4.9
9-AL-tBu	5.1	5.6	5.2	5.8	5.5	5.9	5.3	5.9	5.2	5.8	5.0	5.5	5.0	5.4	5.0	5.6

Table S3.3. AlphaLISA bumped compound screen results.

pIC_{50} values of all compounds in AlphaLISA assay, for each WT and L/V BET bromodomain.

Compound	ITC pKd											
	BRD2 BD1		BRD2 BD2		BRD3 BD1		BRD3 BD2		BRD4 BD1		BRD4 BD2	
	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V	WT	L/V
ME			5.7	7.1								
ET			5.5	7.5								
AL			5.2	7.3								
PR												
9-ME			-	5.9								
9-ET			-	6.5								
9-AL			-	6.5								
ME-Am1			4.9	6.8								
ET-Am1			-	6.5								
AL-Am1			-	6.6								
ME-Am2			-	6.0								
ET-Am2			-	6.0								
9-ME-Am1	-	6.0	-	6.6	5.1	6.4	5.0	6.1	4.4	6.4	-	6.5
9-ET-Am1			-	5.3								
ET-tBu			-	6.3								
AL-tBu	-	6.3	-	6.6	-	6.8	-	6.5	-	6.4	-	6.6
9-ET-tBu												
9-AL-tBu												

Table S3.4. Previously-obtained ITC data

pK_d values for bumped compounds against select WT and L/V BET bromodomains. "-"denotes no visible binding. Data by Dr Michael Zengerle.

	BRD2 BD1			BRD2 BD2			BRD3 BD1			BRD3 BD2			BRD4 BD1			BRD4 BD2		
	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ
ME-Am1	5.6	6.9	1.3	5.6	6.7	1.1	5.8	7.2	1.4	5.8	6.9	1.1	5.0	6.8	1.8	5.8	7.1	1.3
ET	4.7	6.6	2.0	5.5	7.3	1.8	5.5	7.1	1.6	4.8	7.0	2.2	5.6	6.8	1.2	4.5	6.5	2.0
ET-Am1	4.2	6.0	1.8	4.2	5.9	1.7	4.2	6.1	1.9	4.2	6.2	2.0	4.2	5.8	1.6	4.2	6.0	1.8
9-ME-Am1	4.2	6.0	1.8	4.2	6.6	2.4	5.1	6.4	1.3	5.0	6.1	1.1	4.4	6.4	2.0	4.2	6.5	2.3
AL	5.1	6.3	1.3	5.2	6.6	1.4	5.8	6.6	0.7	4.6	6.5	1.9	5.0	6.7	1.7	4.8	6.8	2.0

	BRD2 BD1			BRD2 BD2			BRD3 BD1			BRD3 BD2			BRD4 BD1			BRD4 BD2		
	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ	WT	L/V	Δ
9-ME (1)	4.6	6.3		4.2	6.5		4.9	7.0		4.7	6.7		4.9	7.0		4.5	5.8	
9-ME (2)	4.5	6.4		4.2	6.5		4.9	6.9		4.3	6.9		4.4	6.8		4.5	6.3	
9-ME (3)	4.4	6.6		4.2	6.4		4.9	6.7		4.4	6.4		4.6	6.8		4.6	6.1	
9-ME (4)		6.5						6.9									5.9	
9-ME *	4.5	6.5	2.0	4.2	6.5	2.3	4.9	6.9	2.0	4.5	6.7	2.2	4.6	6.9	2.3	4.5	6.0	1.5
9-ET (1)	4.2	6.4		4.2	6.6		4.6	6.6		4.2	6.4		4.2	6.2		4.2	6.2	
9-ET (2)	4.2	6.5		4.2	6.6		4.5	6.6		4.2	6.4		4.2	6.4		4.2	6.2	
9-ET (3)	4.2	6.5		4.2	6.4		4.5	6.6		4.2	6.4		4.2	6.2		4.2	6.0	
9-ET (4)																	6.2	
9-ET *	4.2	6.5	2.3	4.2	6.5	2.3	4.2	6.6	2.4	4.2	6.4	2.2	4.2	6.3	2.1	4.2	6.1	1.9

Table S3.5. ITC bumped-compound screen results.

A) K_d and pK_d values from ITC titrations of bumped compounds into WT and L/V BET bromodomain constructs.

B) Replicated ITC titrations of 9-ME and 9-ET into WT and L/V BET bromodomain constructs. * denotes mean WT and L/V K_d values, and resulting selectivity.

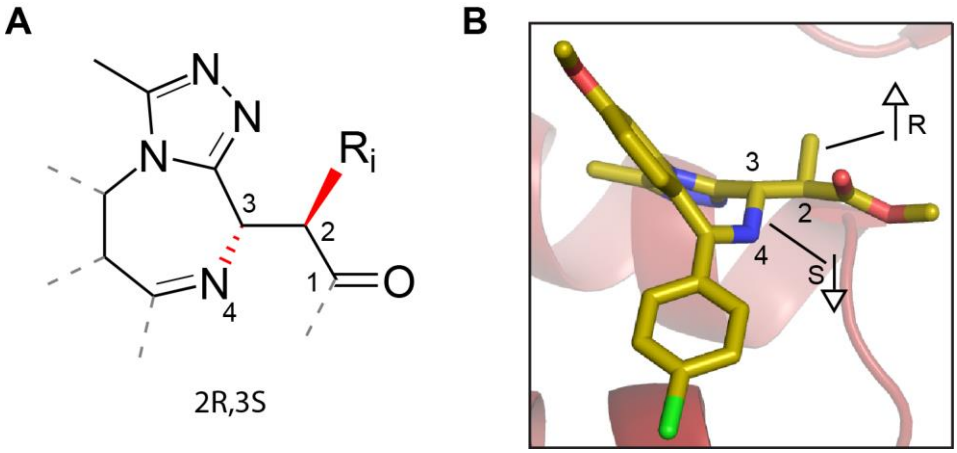


Figure S3.2 Only the 2R,3S enantiomer binds BET bromodomains.

A) Part of the chemical structure of the 2R,3S enantiomer of 9-ME, with relevant atoms and bonds numbered and highlighted. B) X-ray co-crystal structure of BRD2 BD2 L/V in complex with 9-ME, with relevant atoms and bonds again numbered and highlighted.

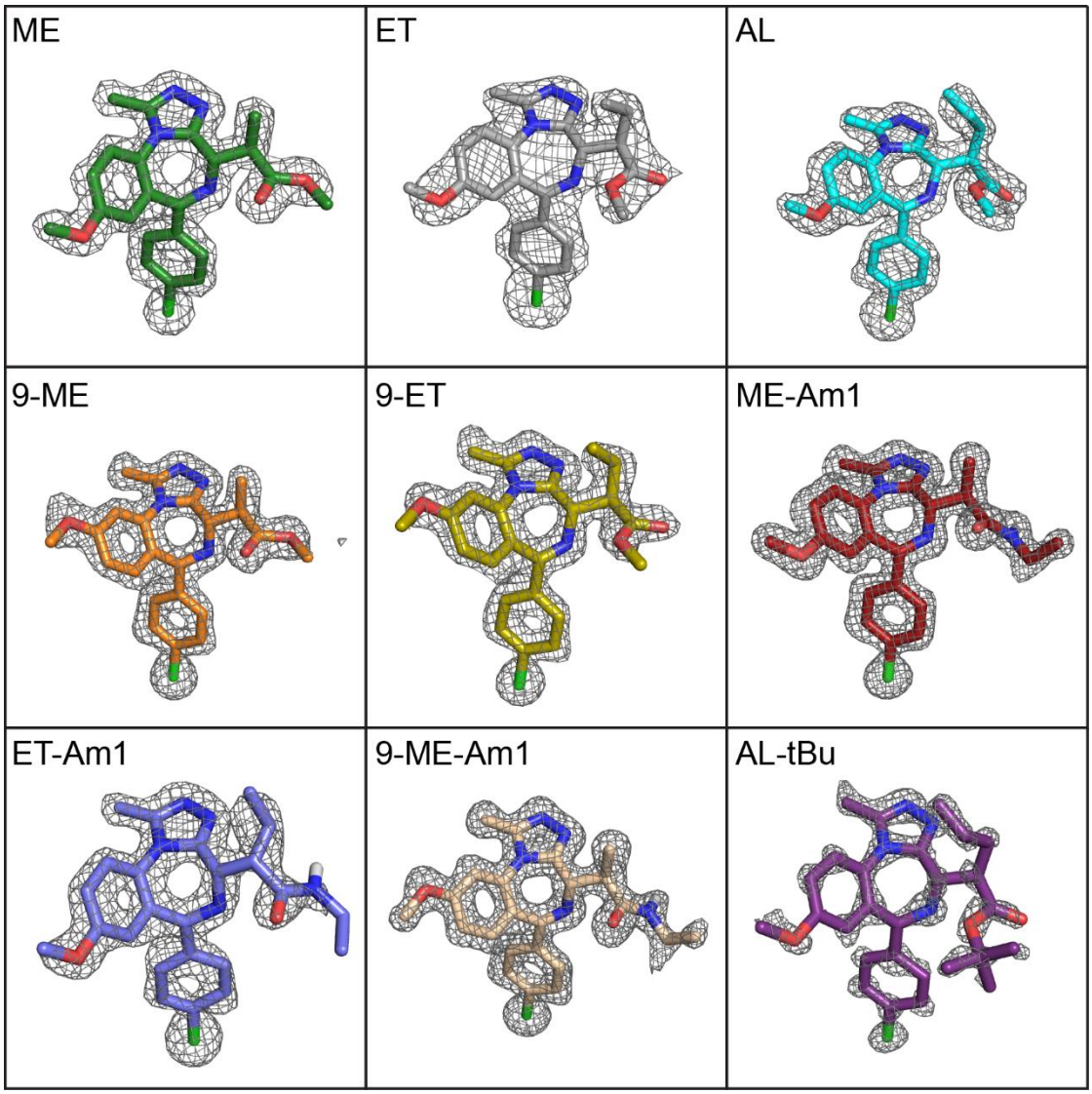
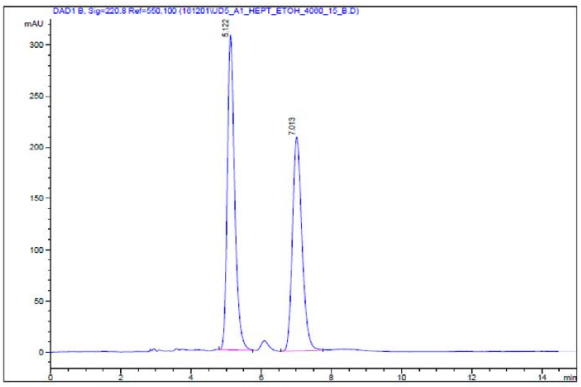


Figure S3.3. Ligand omit maps from bumped compound co-crystal structures.

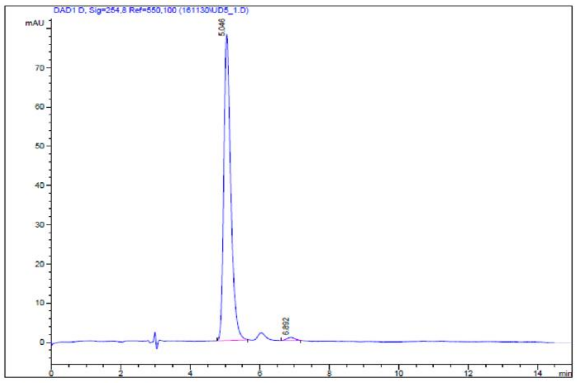
Electron density maps observed in X-ray crystal structures of BRD2 BD2 L/V in complex with bumped ligands (bromodomain omitted for clarity). The unbiased Fo-Fc omit map (grey mesh) is shown superimposed around the ligand and contoured at 3 sigma level.

9-ME (*R,*S mixture)



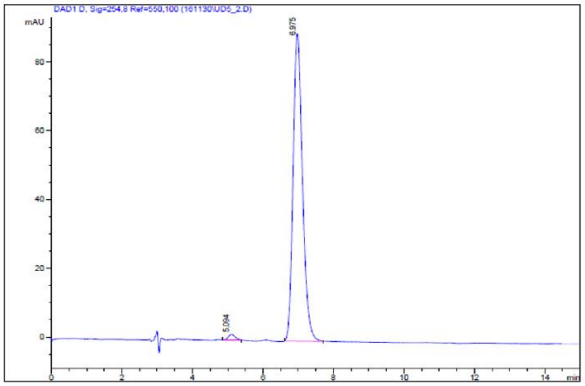
#	Meas. R	Height	Width	Area	Area %	Symmetr
1	5.122	307.488	0.243	4.484e3	53.076	0.730
2	7.013	208.429	0.294	3.965e3	46.924	0.789

9-ME-1 *(2R,3S)*



#	Meas. R	Height	Width	Area	Area %	Symmetr
1	5.046	78.080	0.234	1.094e3	98.786	0.692
2	6.892	0.828	0.271	13.447	1.214	0.986

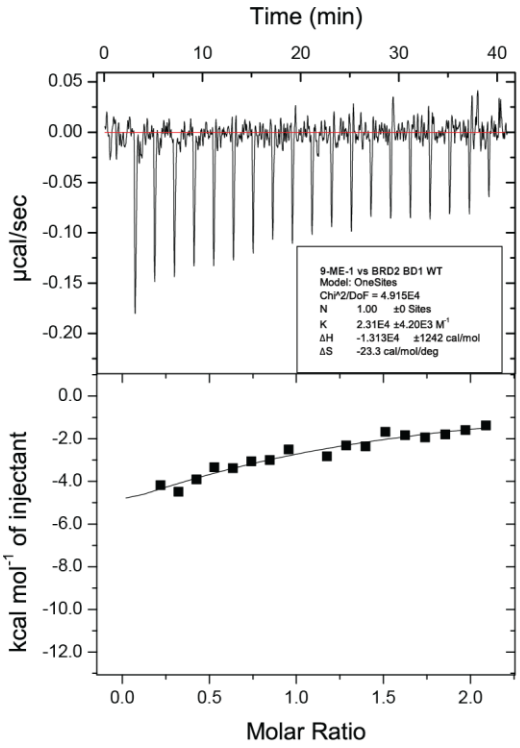
9-ME-2 *(2S,3R)*

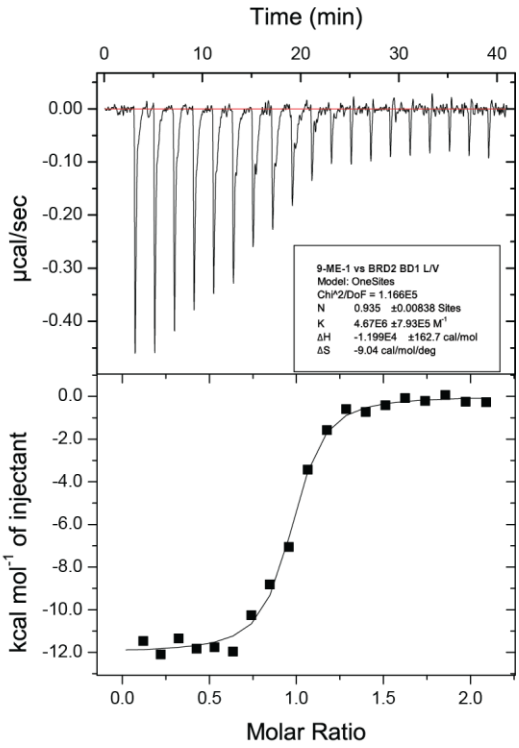


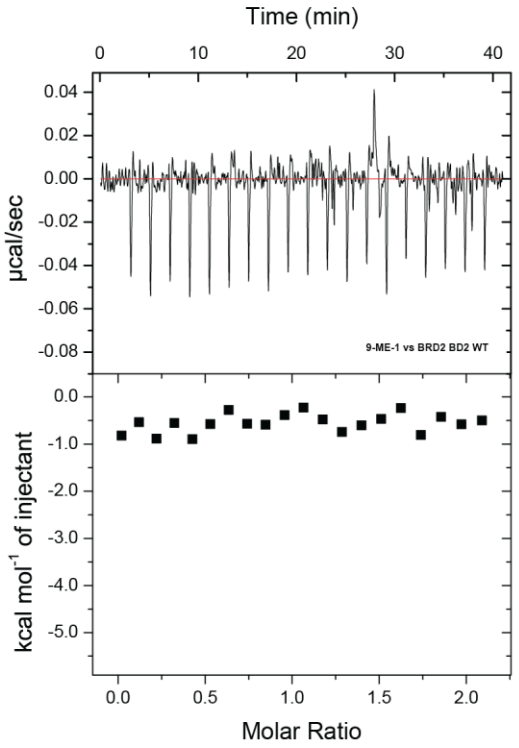
#	Meas. R	Height	Width	Area	Area %	Symmetr
1	5.094	1.621	0.227	22.051	1.314	0.650
2	6.975	88.925	0.297	1.657e3	98.686	0.793

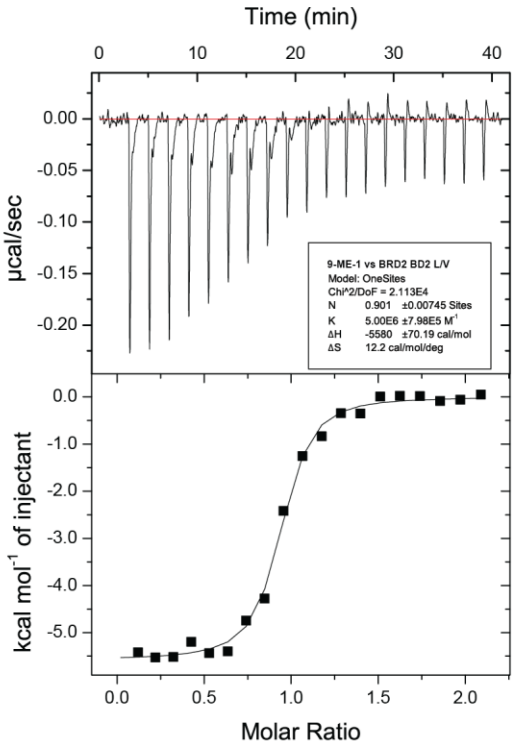
Figure S3.4. Separation of 9-ME enantiomers.

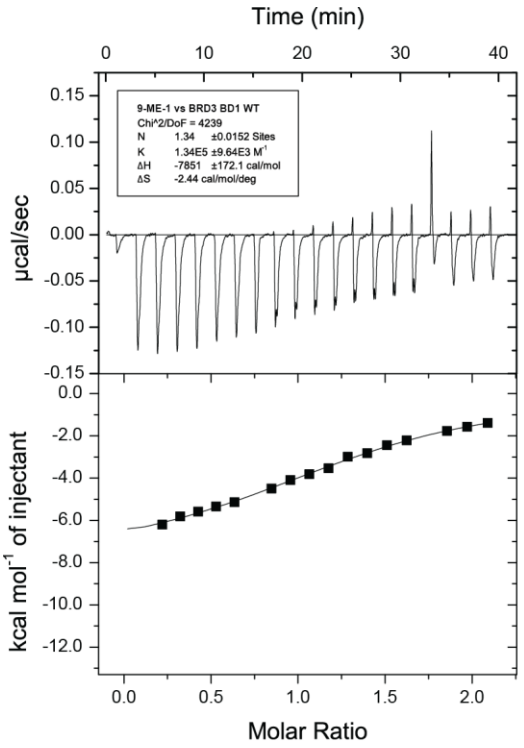
HPLC traces of 9-ME mixture and separated enantiomers. Data from Reach Separations Ltd.

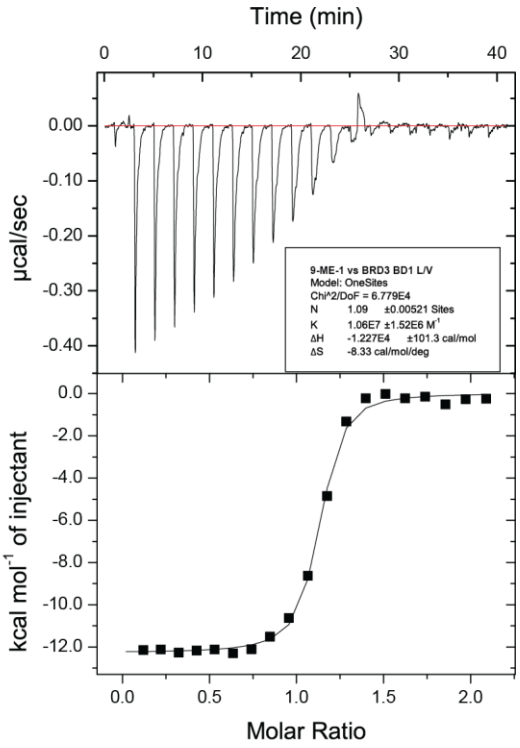

Time (min)
µcal/sec
9-ME-1 vs BRD2 BD1 WT
Model: OneSites
Chi^2/DoF = 4.915E4
N 1.00 ±0 Sites
K 2.31E4 ±4.20E3 M-1
ΔH -1.313E4 ±1242 cal/mol
ΔS -23.3 cal/mol/deg
kcal mol-1 of injectant
Molar Ratio

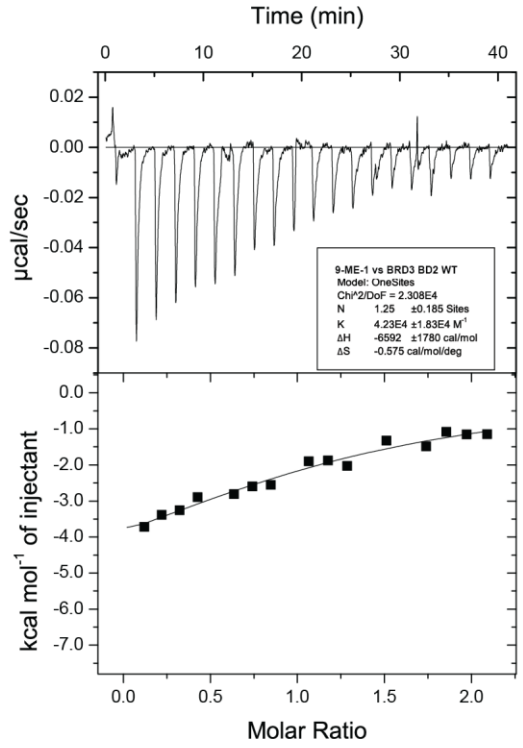

Time (min)
µcal/sec
9-ME-1 vs BRD2 BD1 L/V
Model: OneSites
Chi^2/DoF = 1.166E5
N 0.935 ±0.00838 Sites
K 4.67E6 ±7.93E5 M-1
ΔH -1.199E4 ±162.7 cal/mol
ΔS -9.04 cal/mol/deg
kcal mol-1 of injectant
Molar Ratio

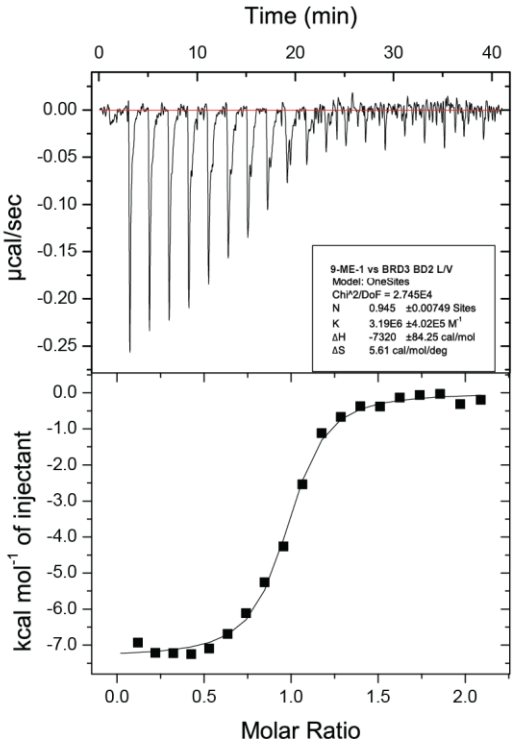

Time (min)
µcal/sec
9-ME-1 vs BRD2 BD2 WT
kcal mol-1 of injectant
Molar Ratio

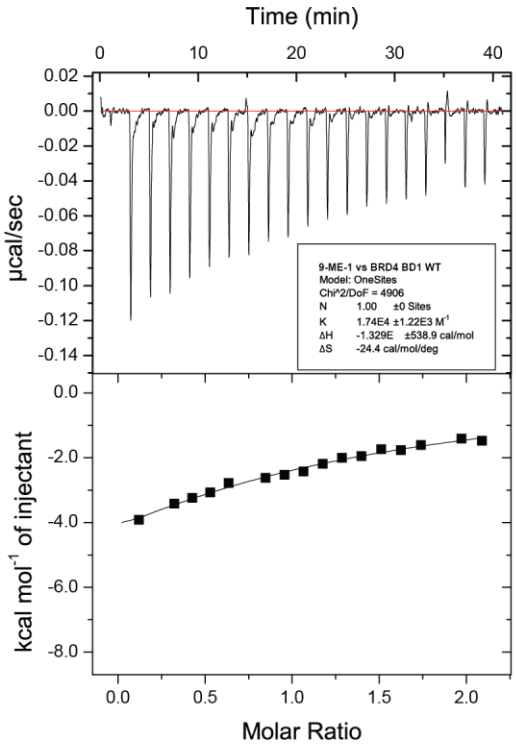

Time (min)
µcal/sec
9-ME-1 vs BRD2 BD2 L/V
Model: OneSites
Chi^2/DoF = 2.113E4
N 0.901 ±0.00745 Sites
K 5.00E6 ±7.98E5 M-1
ΔH -5580 ±70.19 cal/mol
ΔS 12.2 cal/mol/deg
kcal mol-1 of injectant
Molar Ratio

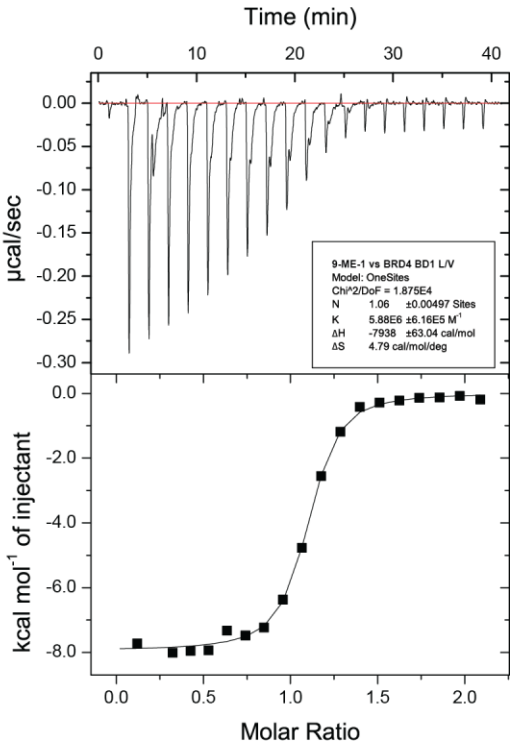

Time (min)
9-ME-1 vs BRD3 BD1 WT
Chi^2/DoF = 4239
N 1.34 ±0.0152 Sites
K 1.34E5 ±9.64E3 M-1
ΔH -7851 ±172.1 cal/mol
ΔS -2.44 cal/mol/deg
µcal/sec
kcal mol-1 of injectant
Molar Ratio

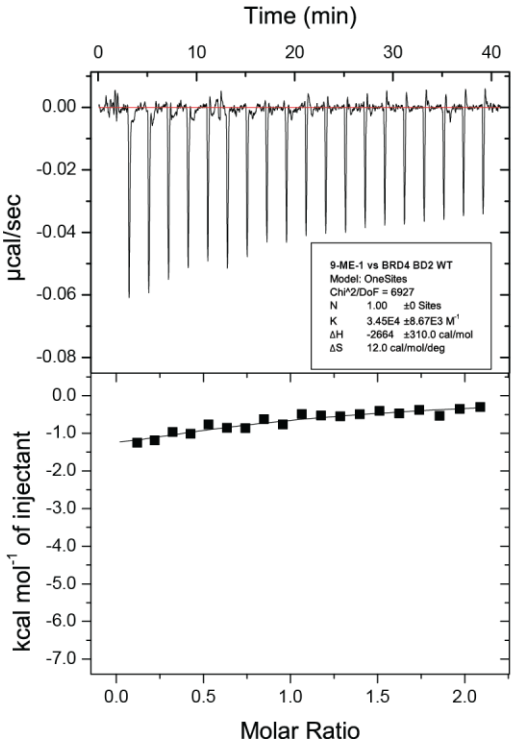

Time (min)
9-ME-1 vs BRD3 BD1 L/V
Model: OneSites
Chi^2/DoF = 6.779E4
N 1.09 ±0.00521 Sites
K 1.06E7 ±1.52E6 M-1
ΔH -1.227E4 ±101.3 cal/mol
ΔS -8.33 cal/mol/deg
µcal/sec
kcal mol-1 of injectant
Molar Ratio

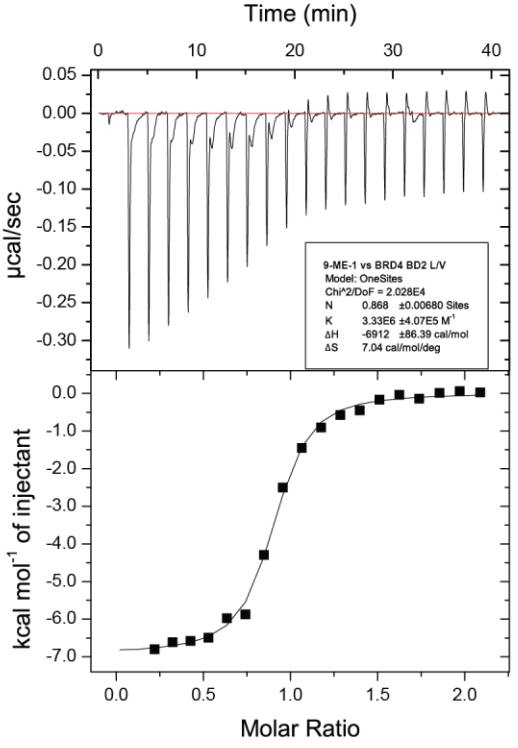

Time (min)
9-ME-1 vs BRD3 BD2 WT
Model: OneSites
Chi^2/DoF = 2.308E4
N 1.25 ±0.185 Sites
K 4.23E4 ±1.83E4 M-1
ΔH -6592 ±1780 cal/mol
ΔS -0.575 cal/mol/deg
µcal/sec
kcal mol-1 of injectant
Molar Ratio


Time (min)
9-ME-1 vs BRD3 BD2 L/V
Model: OneSites
Chi^2/DoF = 2.745E4
N 0.945 ±0.00749 Sites
K 3.19E6 ±4.02E5 M-1
ΔH -7320 ±84.25 cal/mol
ΔS 5.61 cal/mol/deg
µcal/sec
kcal mol-1 of injectant
Molar Ratio


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
9-ME-1 vs BRD4 BD1 WT
Model: OneSites
Chi^2/DoF = 4906
N 1.00 ±0 Sites
K 1.74E4 ±1.22E3 M-1
ΔH -1.329E ±538.9 cal/mol
ΔS -24.4 cal/mol/deg


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
9-ME-1 vs BRD4 BD1 L/V
Model: OneSites
Chi^2/DoF = 1.875E4
N 1.06 ±0.00497 Sites
K 5.88E6 ±6.16E5 M-1
ΔH -7938 ±63.04 cal/mol
ΔS 4.79 cal/mol/deg


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
9-ME-1 vs BRD4 BD2 WT
Model: OneSites
Chi^2/DoF = 6927
N 1.00 ±0 Sites
K 3.45E4 ±8.67E3 M-1
ΔH -2664 ±310.0 cal/mol
ΔS 12.0 cal/mol/deg


Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio
9-ME-1 vs BRD4 BD2 L/V
Model: OneSites
Chi^2/DoF = 2.028E4
N 0.868 ±0.00680 Sites
K 3.33E6 ±4.07E5 M-1
ΔH -6912 ±86.39 cal/mol
ΔS 7.04 cal/mol/deg

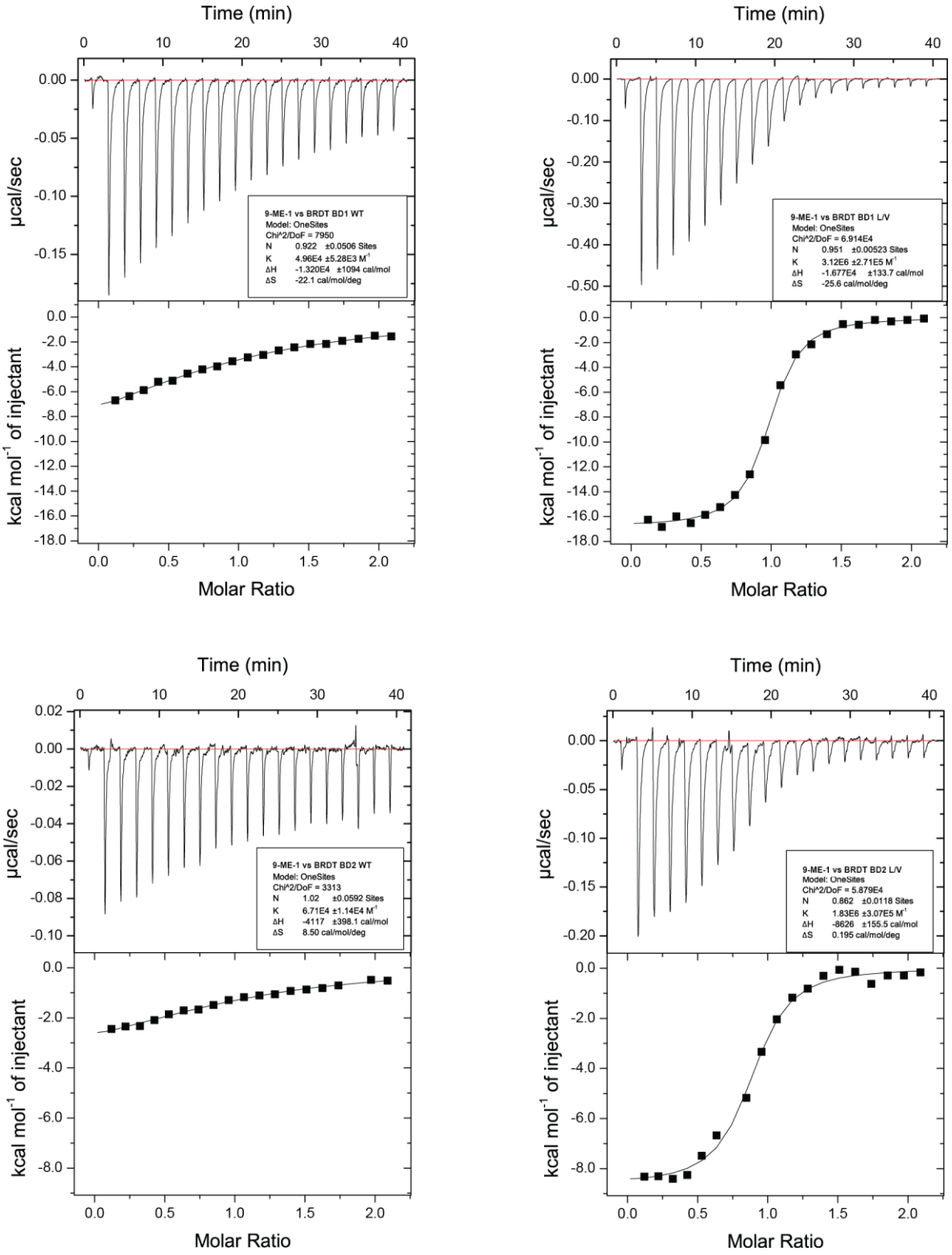


Figure S3.5. 9-ME-1 ITC titrations.

ITC titrations of 250 µM 9-ME-1 into 25 µM WT or L/V BET bromodomain construct, at 30°C and 2.5% DMSO.

Chapter 4 Supplementary Figures

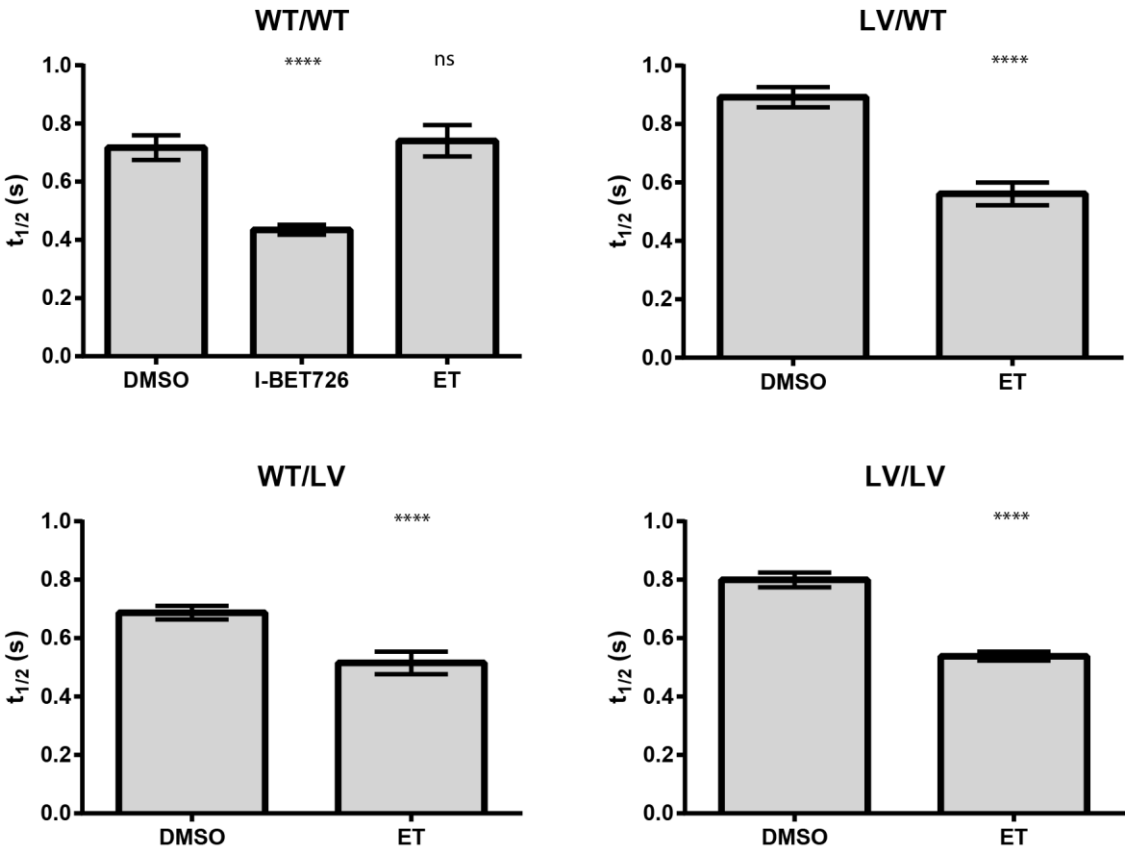
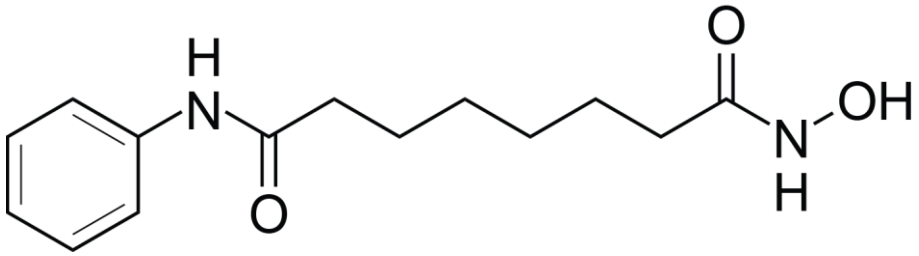


Figure S4.1. Preliminary FRAP data.

Recovery times of GFP-labelled BRD4 (L) constructs following 0.5 s laser bleach event, at 0.01% DMSO, treated with 1 µM I-BET762 or ET. Each bar is mean ± SE of ~50 U2OS cells tested over two separate experiments. Statistical significance (compared to each construct's DMSO treatment) determined with two-tailed t tests: ns $P>0.05$; * $P\leq0.05$; ** $P\leq0.01$; *** $P\leq0.001$; **** $P<0.0001$. Experiment performed by Lars van Beurden.



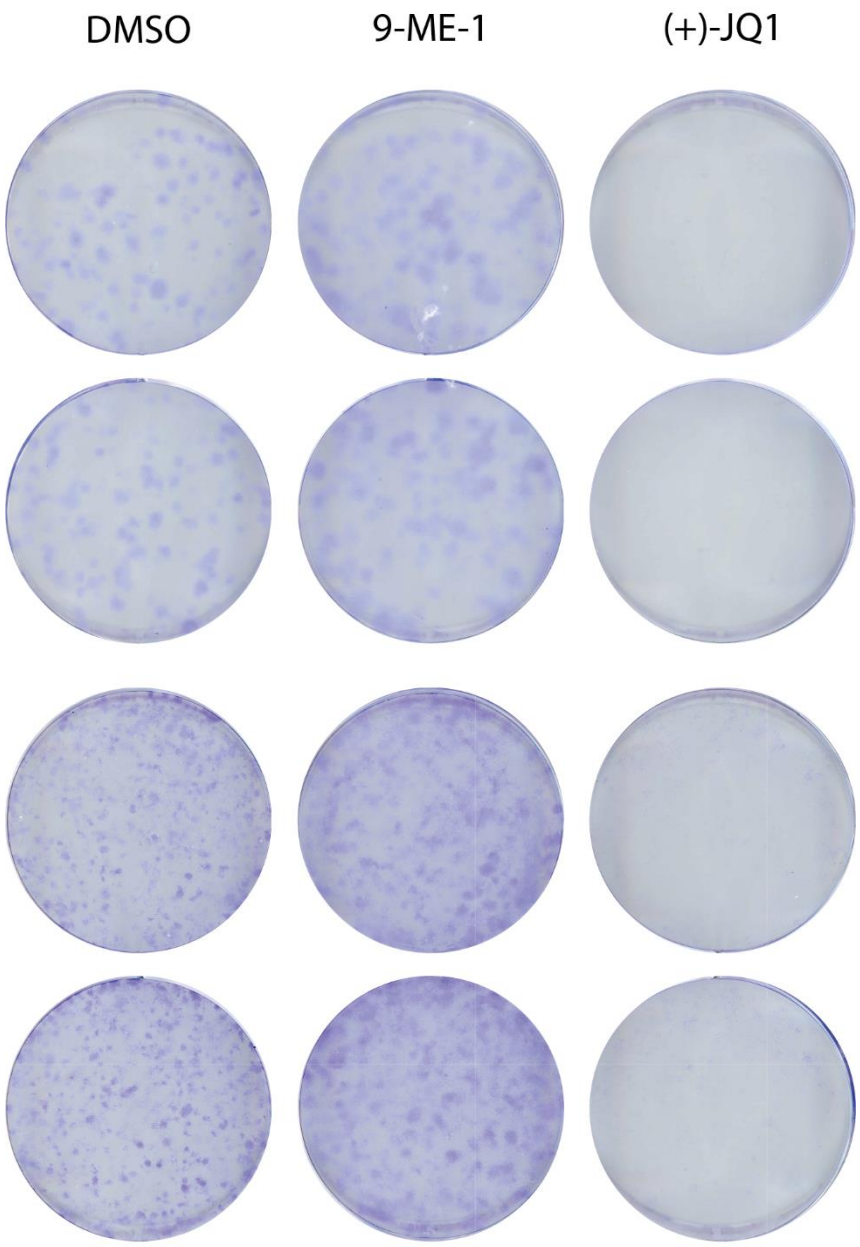
Scheme S4.1. SAHA

Chemical structure of SAHA, a.k.a Suberoylanilide Hydroxamic Acid, Vorinostat

BRD4 (L)	WT/WT		LV/WT		WT/LV		LV/LV	
Treatment	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1
t½ (s)	3.5	3.0	3.2	1.3	3.0	2.2	2.3	0.8
SEM	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.0
t½ (frac.)	1.00	0.91	1.00	0.39	1.00	0.73	1.00	0.33
SEM	0.04	0.03	0.04	0.02	0.05	0.03	0.04	0.01
P	0.09		<0.0001		<0.0001		<0.0001	
BRD3	WT/WT		LV/WT		WT/LV		LV/LV	
Treatment	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1
t½ (s)	4.0	3.9	3.6	1.8	3.7	2.3	3.2	1.0
SEM	0.2	0.2	0.1	0.1	0.2	0.1	0.1	0.0
t½ (frac.)	1.00	1.01	1.00	0.49	1.00	0.63	1.00	0.30
SEM	0.04	0.04	0.04	0.03	0.06	0.02	0.04	0.01
P	0.9		<0.0001		<0.0001		<0.0001	
BRD2	WT/WT		LV/WT		WT/LV		LV/LV	
Treatment	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1
t½ (s)	4.0	4.0	4.0	2.9	4.0	4.2	2.3	N/A
SEM	0.2	0.2	0.2	0.2	0.3	0.3	0.2	
t½ (frac.)	1.00	0.99	0.99	0.72	1.00	1.04	0.56	
SEM	0.05	0.05	0.05	0.04	0.06	0.06	0.04	
P	0.9		<0.0001		0.6			
BRDT	WT/WT		LV/WT		WT/LV		LV/LV	
Treatment	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1	DMSO	9-ME-1
t½ (s)	1.4	1.4	1.3	0.7	1.4	1.3	1.1	0.6
SEM	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.0
t½ (frac.)	1.00	1.04	1.00	0.57	1.00	0.95	1.00	0.56
SEM	0.06	0.06	0.06	0.03	0.07	0.06	0.05	0.02
P	0.7		<0.0001		0.6		<0.0001	
BRD2	WT/WT		LV/WT		WT/LV		LV/LV	
Treatment	DMSO	9-ME-1 *	DMSO	9-ME-1 *	DMSO	9-ME-1 *	DMSO	9-ME-1 *
t½ (s)	4.2	4.2	3.8	2.5	3.6	3.3	2.6	1.5
SEM	0.3	0.3	0.2	0.1	0.2	0.2	0.3	0.1
t½ (frac.)	1.00	1.00	1.00	0.66	1.00	0.94	1.00	0.57
SEM	0.06	0.07	0.04	0.03	0.05	0.06	0.10	0.04
P	0.99		<0.0001		0.5		<0.0001	

Table S4.1. Detailed FRAP results.

Average t½ values of GFP-BET constructs transiently overexpressed in U2OS cells, following 0.5s laser bleach, at 0.03% DMSO and/or 9-ME-1. 9-ME-1 dose is 200 nM or 40 nM (denoted by '9-ME-1 *'). Values are mean of ~40 cells, over 2 separate experiments. GFP-BRD2 LV/LV aggregated when treated with 200 nM 9-ME-1 and hence no t½ was recorded.



	A549 Colonies			
	Replicate #1		Replicate #2	
	#	%	#	%
DMSO	76	101	102	106
	74	99	90	94
9-ME-1	65	87	78	81
	67	89	102	106
(+)-JQ1	0	0	0	0
	0	0	2	2

Figure S4.2. 9-ME-1 assessment in A549 colony-formation assay.

6-well plates seeded with 2000 A549 cells, treated with 1 µM compound and/or 0.01% DMSO and incubated for 8-12 days before being fixed with 100% MeOH and stained with 0.1% crystal-violet.

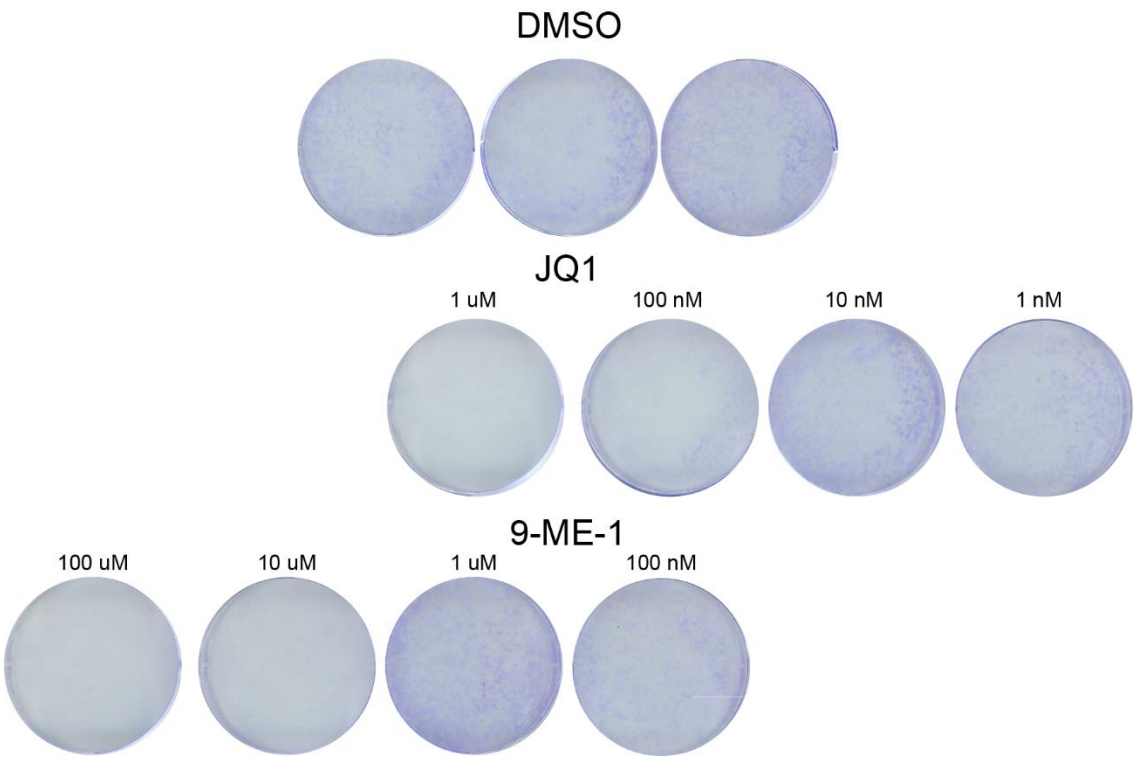


Figure S4.3. Dose-response A549 colony-formation assay.

6-well plates seeded with 20000 A549 cells, treated with compound and/or 0.1% DMSO and incubated for 8 days before being fixed with 100% MeOH and stained with 0.1% crystal-violet.

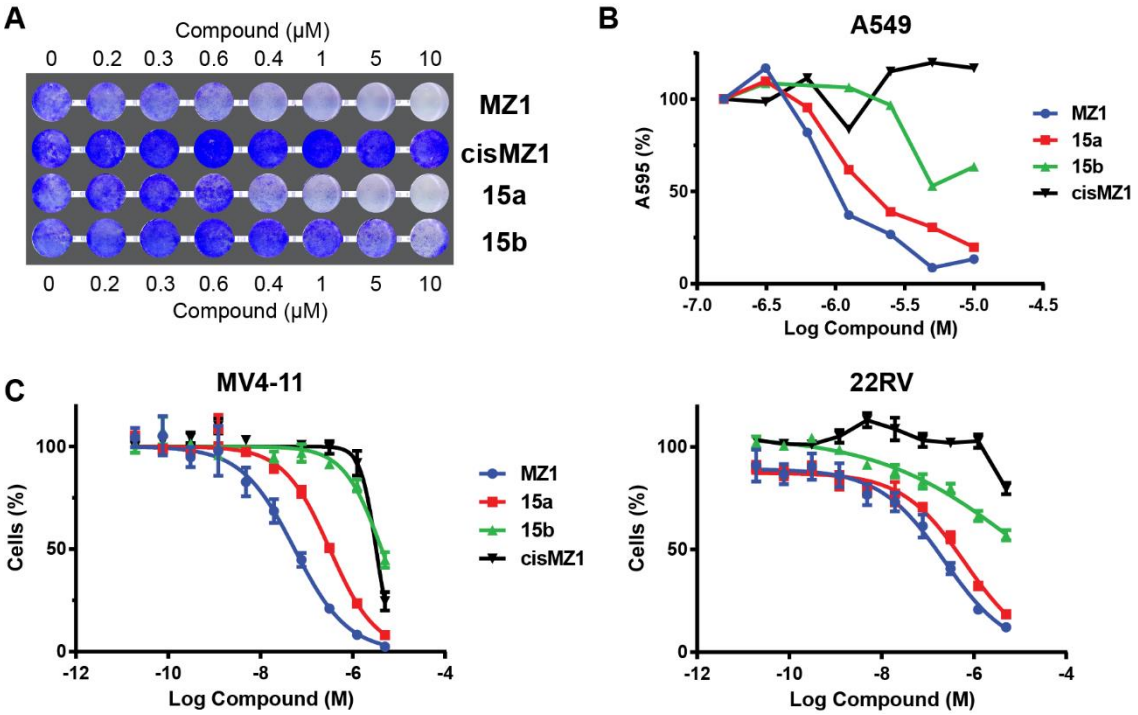


Figure S4.4. PROTAC assessment with colony-formation assay.

From [189]. A) 96-well plate seeded with A549 cells, treated with compound and/or 1% DMSO and incubated for 1 week before being fixed with 100% MeOH and stained with 0.1% crystal-violet. B) Crystal-violet dye extracted with 1% SDS, and absorbance measured at 595 nm. C) Viability of BET-sensitive MV4-11 and 22RV cells, determined via Promega CellTiter-Glo, following 72 h incubation with compounds. CellTiter-Glo data by Dr Jane Wright.

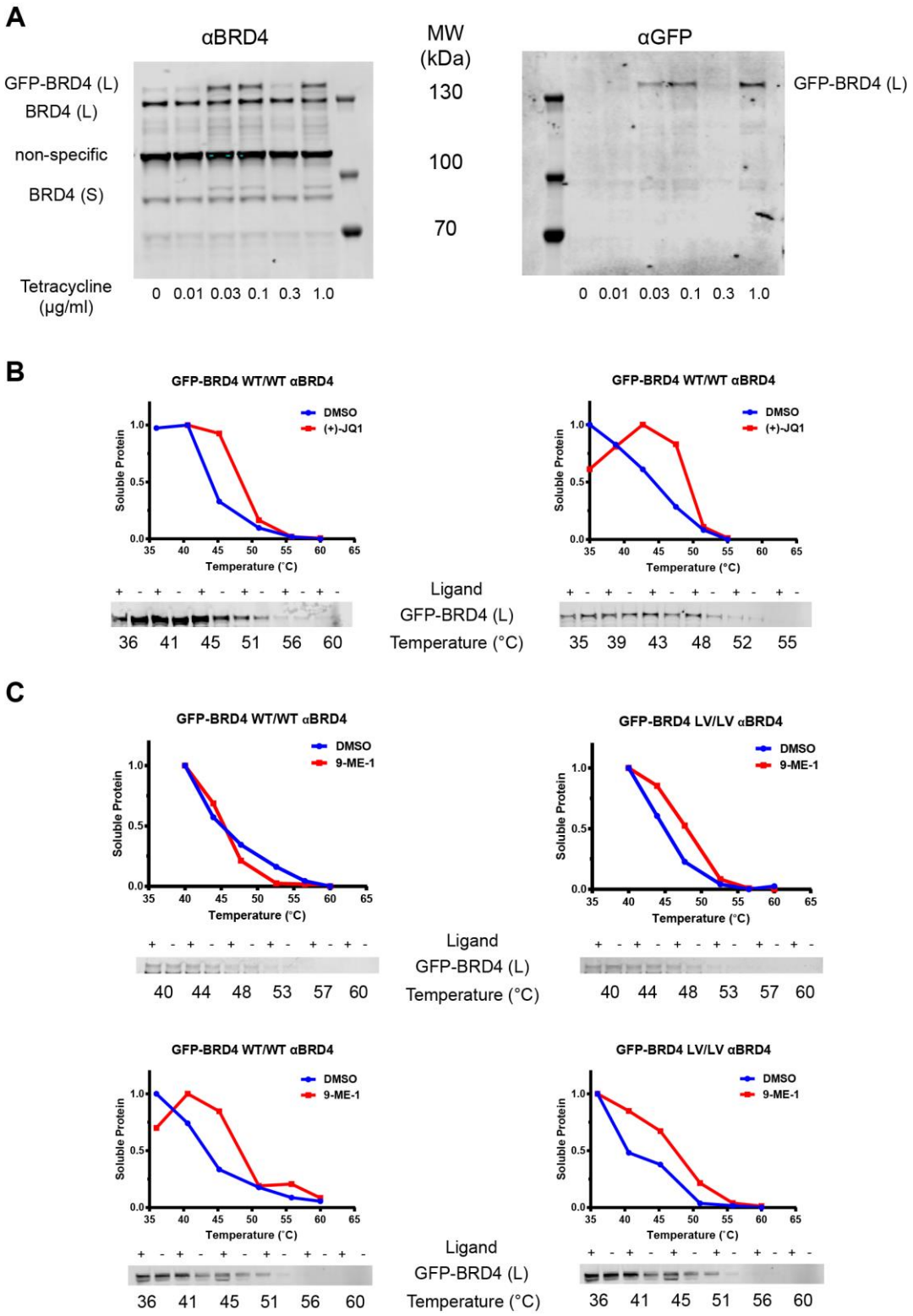


Figure S4.5. CETSA assay for assessing cellular target engagement of BET inhibitors.

A) U20S Flp-In cells, stably expressing GFP-BRD4 (L) WT in response to tetracycline treatment, were treated with a gradient of tetracycline doses, and the expression of the GFP-BRD4 (L) was observed through αBRD4 and αGFP western-blotting. Cells treated with 0.3 µg/ml tetracycline failed to express any protein, likely due to experimental error with the treatment. B) Thermal-gradient CETSA curve of GFP-BRD4 (L) WT, expressed from cells in A with 0.1 µg/ml tetracycline, treated with 0.01% DMSO and 1 µM (+)-JQ1. C) Thermal-gradient CETSA curve of GFP-BRD4 (L) WT/WT and LV/LV, expressed from cells in A with 0.1 µg/ml tetracycline, treated with 0.01% DMSO and 1 µM 9-ME-1.

Chapter 5 Supplementary Figures

Source	Measurement	BRD4	BRD3	BRD2	BRDT
CellMiner	Gene No.	1.7	2.0	N/A	1.6
CanSAR	Gene No.	2.0	3.0	2.0	2.0
	Chromosome	19	9	6	1
	Ch. No.	3.5	3.0	2.0	2.5

Table S5.1 Reported BET Copy Number in A549.
CellMiner reported no cell-number for BRD2 in A549.

BRD4 BD	Primer	Sequence
BD1	Fwd Outer	CCCTGCTGGAGGGTAAGTC
	Fwd Inner	CCTCTGGCCAACTTGGCTAG
	Rev Outer	TGCTGTAGCCAGAGGGAG
	Rev Inner	TGTACTTGCACTCCCTGCGA
	KI-specific	TGAGTAGTGTGGGCAGGAT
	WT	TGAGTAGTGTGGGCAGGGC
BD2	Fwd Outer	GCTTCCAAGTGAGCACACAG
	Fwd Inner	GAGCAGAGTGGTGGGACAC
	Rev Outer	AGGAGCTGGTCAAAGGGAC
	Rev Inner	GCACTGTTCCCTGCTCTGC
	KI-specific	GTGGAGGCACTGGGCGTG
	WT	GTGGAGGCACTGGGCCTA

Table S5.2 Primers for analysing BRD4 knock-ins.

Red nucleotides in KI-specific primers are complementary to KI sequence but not WT.

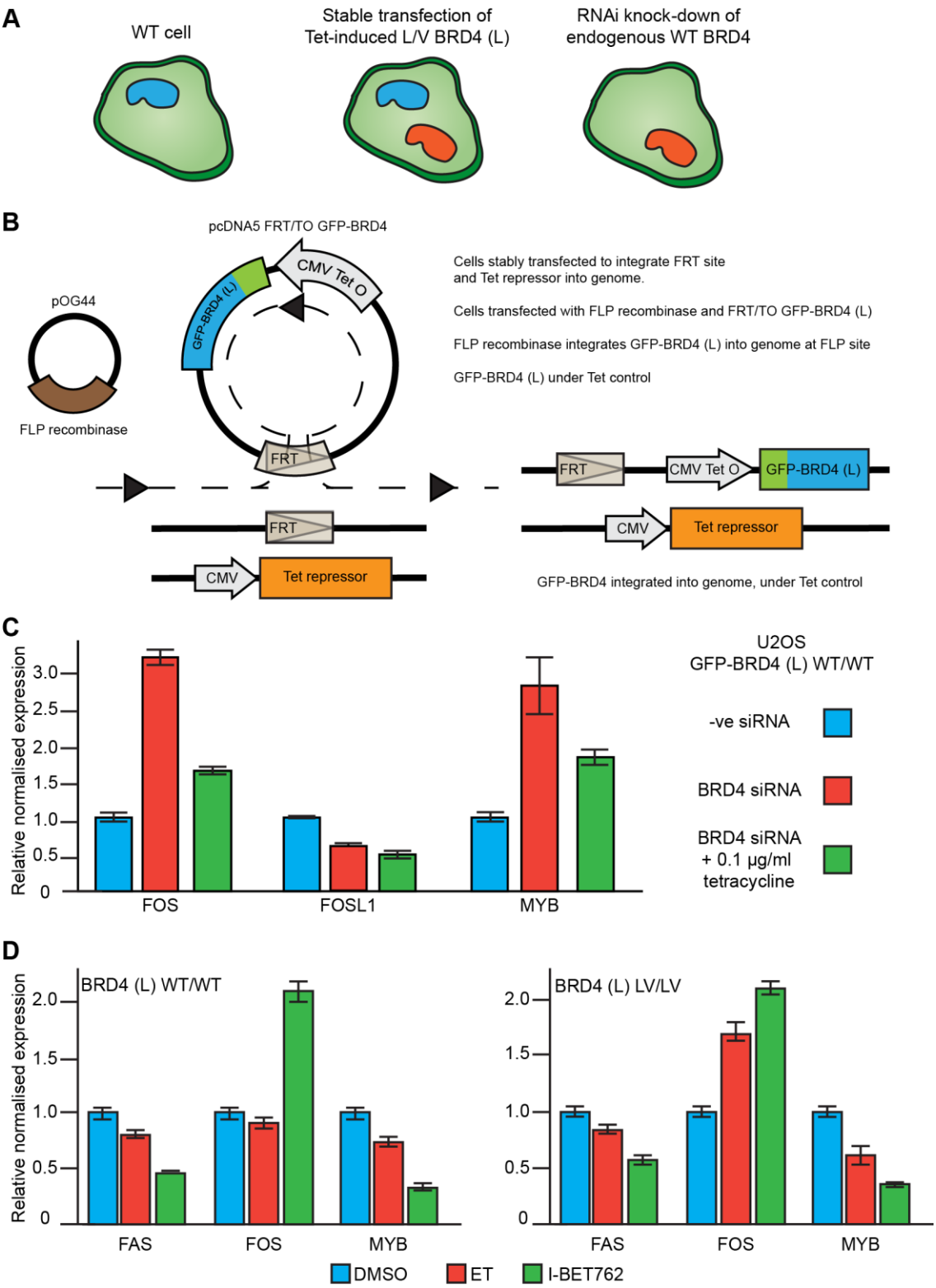


Figure S5.1. Implementation of bump-&-hole system through Flp-in system

A) Original plan for implementing L/V mutation into cell lines. Mutant BET proteins introduced through stable transfection of Tet-induced GFP-BET constructs, while endogenous WT gene knocked-down by RNAi (targeting non-coding region not found in GFP-BET construct).B) Integration of GFP-BET constructs using the Flp-in T-Rex system. U2OS cells stably transfected with pFRT/*lacZeo* and pCDNA6 T/R vectors, randomly integrating FRT site and Tet repressor into genome. Transfection with FRT-recombinase encoding pOG44 vector and GFP-BET encoding pcDNA5 FRT/TO vector leads to FRT-targeting

recombination event, integrating GFP-BET construct under CMV/Tet-operator control into genome. Tet repressor protein will inhibit GFP-BET expression, unless inhibited by Tetracycline. C) qPCR-measured expression of BRD4-regulated genes in Flp-in U2OS cells stably transfected with tetracycline-dependent GFP-BRD4 WT/WT, following treatment with negative siRNA, αBRD4 siRNA and a combination of αBRD4 siRNA and 0.1 μg/ml tetracycline. D) qPCR-measured expression of BRD4-regulated genes in Flp-in U2OS cells stably transfected with tetracycline-dependent GFP-BRD4 WT/WT and LV/LV, following treatment with αBRD4 siRNA and test compounds (1 μM). All work by Dr Kwok-Ho Chan.

Reference sequences	BRD4	BRD3	BRD2	BRDT
gDNA	NC_000019.10	NC_000009.12	NC_000006.12	NC_000001.11
mRNA	NM_058243.2	NM_007371.3	NM_001113182.1	NM_001242806.2
Protein	NP_490597.1	NP_031397.1	NP_001106653.1	NP_001229735.2

Table S5.3. Reference sequences for BET knock-in design.

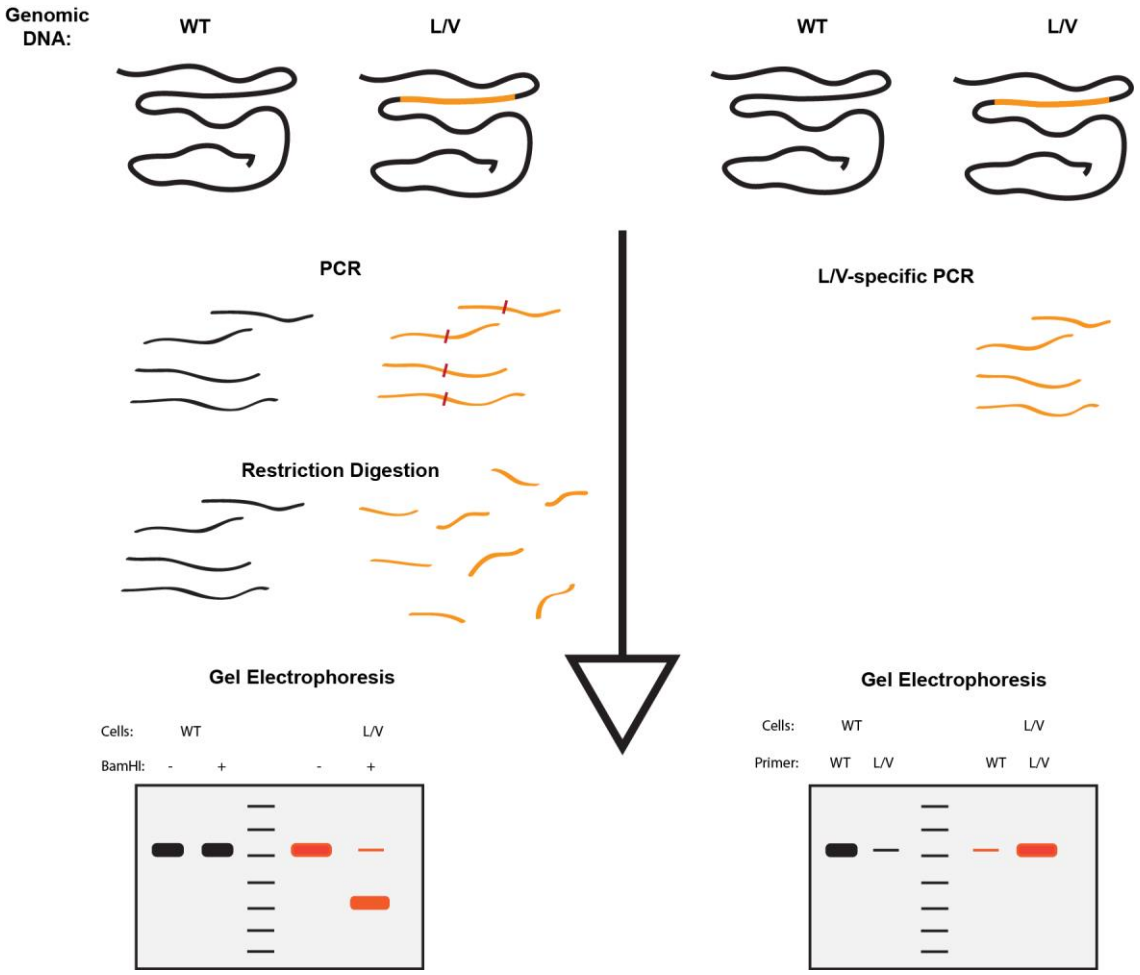
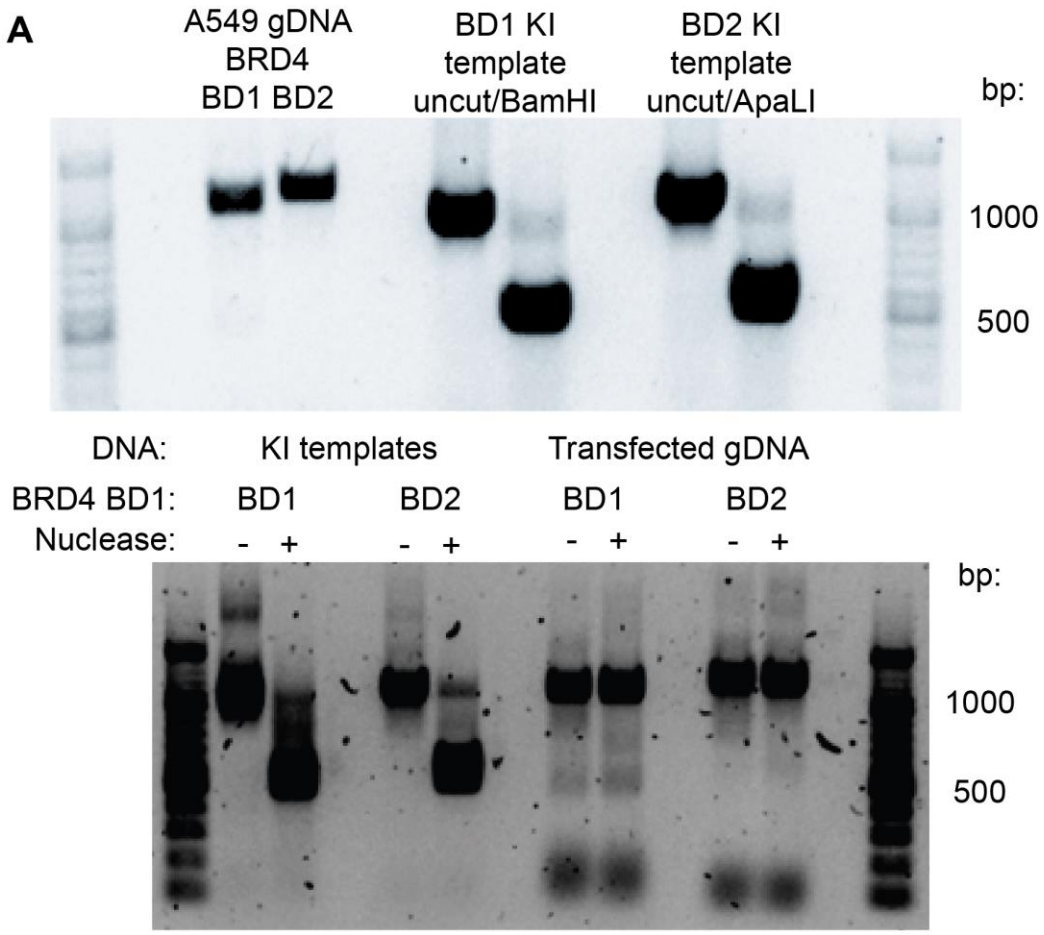


Figure S5.3. PCR Assays for Knock-In Detection.

Starting with extracted gDNA from cells that are either WT or contain L/V knock-in (orange), bromodomain region amplified via PCR reaction. L/V-agnostic PCR will amplify region from all cells, but restriction enzyme will only digest this DNA if it contains the knocked-in restriction site (red line). Alternatively, a primer specific for the knocked-in sequence (utilising L/V mutation and/or restriction site and/or PAM-site modifications) will only amplify DNA from knocked-in cells.


A
A549 gDNA
BRD4
BD1 BD2
BD1 KI
template
uncut/BamHI
BD2 KI
template
uncut/ApaLI
bp:
1000
500
DNA:
KI templates
Transfected gDNA
BRD4 BD1:
BD1
BD2
BD1
BD2
Nuclease:
- + - + - + - +
bp:
1000
500

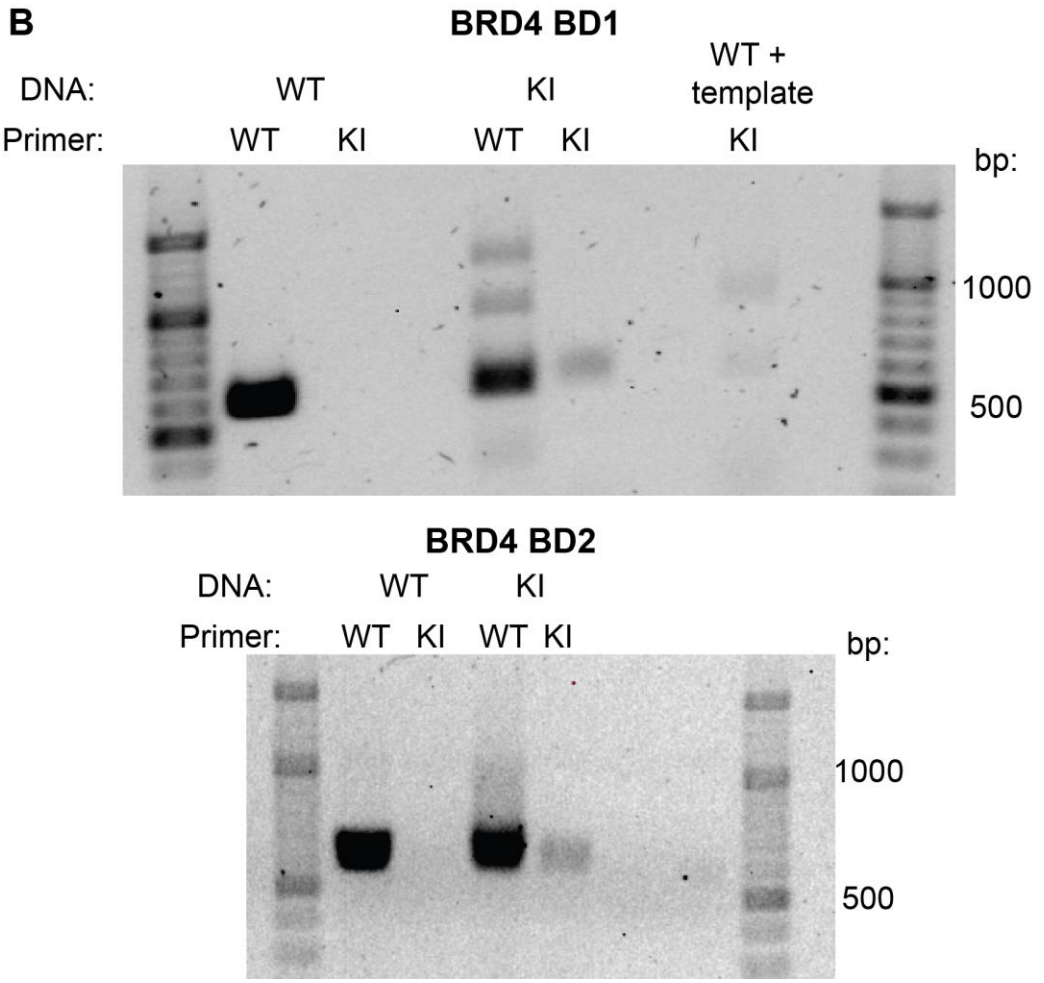

B
BRD4 BD1
DNA:
WT
KI
WT +
template
Primer:
WT KI WT KI KI
bp:
1000
500
BRD4 BD2
DNA:
WT
KI
Primer:
WT KI WT KI
bp:
1000
500

Figure S5.3. Preliminary testing of knock-in detection assays.

A) PCR was used to amplify BRD4 BD1 and BD2 regions from WT A549 gDNA, BRD4 BD1 and BD2 KI templates and gDNA extracted from A549 cells transfected with plasmids targeting BRD4 BDs for knock-in. To test for presence of restriction site PCR products were then treated with restriction enzymes BamHI (for BD1) and ApaLI (for BD2). Final products ran on 1% agar gel, alongside Generuler 1kb PLUS size marker (Thermo). All work by Dr Kwok-Ho Chan. B) A549 cells were transfected with plasmids to generate BRD4 BD1 and BD2 L/V knock-ins. Genomic DNA was extracted from these and WT cells, and the BRD4 BD1 and BD2 regions amplified with either WT or specific primers. As a control WT gDNA was also incubated with knock-in DNA template before PCR. PCR products ran on 1% agar gel, alongside Generuler 1kb PLUS size marker (Thermo). All work by Dr Kwok-Ho Chan.

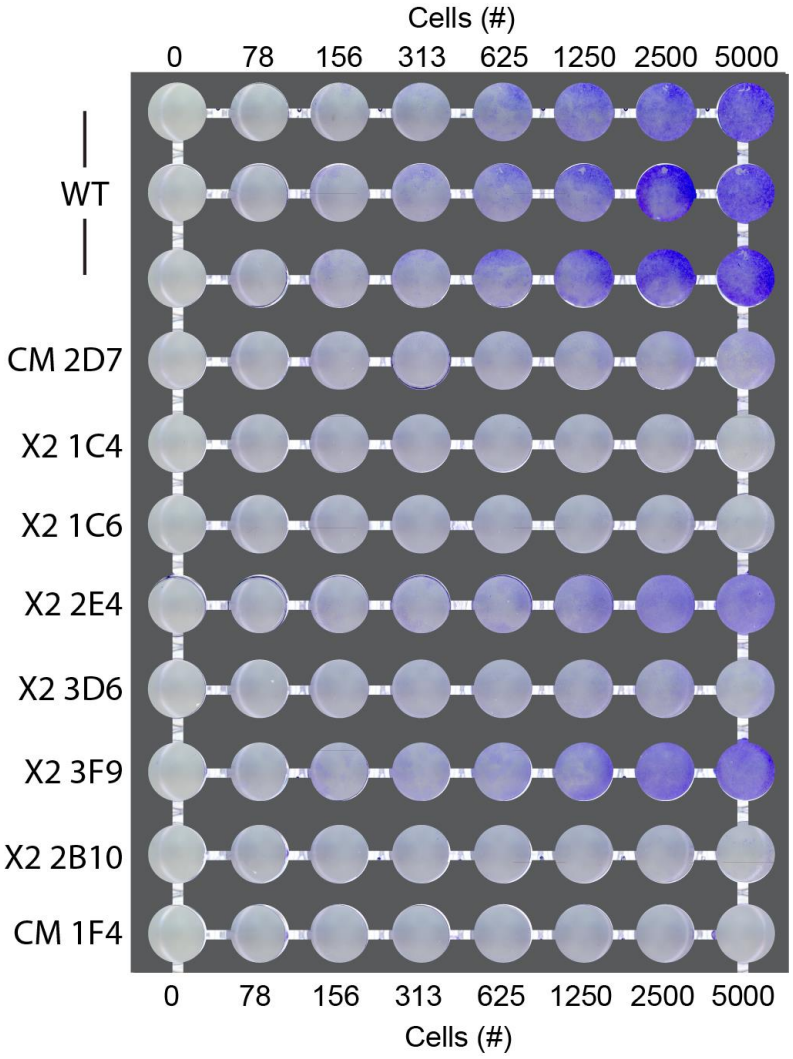


Figure S5.4. Effects of potential knock-out colony-formation.

WT A549 cells, and potential BRD4 BD1 knock-outs, seeded over a range of cell densities and incubated for 7 days, fixed with 100% MeOH and stained with 0.1% crystal-violet.