Investigating Bromodomain Proteins as Targets for Antileishmanial Drug Discovery

Catherine Nicola Russell

Doctor of Philosophy

University of York

Chemistry

December 2022

Abstract

Leishmania is a kinetoplastid parasite and the causative agent of the neglected tropical disease, leishmaniasis. Despite the high disease burden, current treatments for leishmaniasis show toxic side effects and growing resistance, so new antileishmanials are urgently needed. Bromodomain inhibitors have shown huge potential for the treatment of other diseases, and may also present an attractive drug target in *Leishmania*. Five bromodomain factor (BDF) proteins are found in *Leishmania* which are essential for promastigote survival. Bromodomains are small, epigenetic reader domains which recognise acetylated lysine residues on histone tails and regulate gene expression. The aim of this work was to validate *Leishmania* bromodomain proteins as targets for the development of new antileishmanials. The binding of ligands to recombinant *Leishmania donovani* BDF bromodomains was investigated using biophysical methods including thermal shift assays, fluorescence polarisation assays and NMR. Exploring the peptide recognition profile of the tandem bromodomain protein, *Leishmania donovani* BDF5 (*Ld*BDF5) revealed binding sites in acetylated peptides derived from the tails of histones H2B and H4. Screening a library of 15 commercially available human bromodomain inhibitors against *Ld*BDF proteins identified three binding compounds, including SGC-CBP30 which bound to *Ld*BDF5 in biophysical assays and also displayed significant cytotoxicity against *Leishmania* promastigotes in cell viability assays. Subsequent work sought to identify novel inhibitors of BDF5. Collaborative high-throughput screens and SAR-optimisation efforts yielded a series of hit compounds. Using orthogonal biophysical assays, binding interactions were characterised and a smaller cohort of promising lead compounds selected. Gratifyingly, these compounds were phenotypically active against promastigotes, with EC_{50} values comparable to miltefosine, a leading antileishmanial drug. These findings validate *Leishmania* BDF5 as a viable drug target and describe the first steps taken towards the generation of novel antileishmanial compounds to tackle the global burden of this neglected tropical disease.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to thank my supervisors Professor Tony Wilkinson and Professor Jeremy Mottram for giving me the opportunity to undertake this PhD project, and for their invaluable support and guidance throughout. I would also like to thank my thesis advisory panel member Dr Martin Fascione for his continual advice and encouragement.

I am also very grateful to Dr Nathanial Jones for his invaluable input throughout the project, for many helpful discussions and for teaching me the promastigote cell viability assay. Thank you also to Dr Jim Brannigan for support with lab work, and for always being so helpful with advice and trouble-shooting.

I would also like to thank Juliet Borgia for her many contributions and for helping me get started at the beginning of the project. I would additionally like to extend my deepest thanks to the full YSBL technical support staff, including Simon Grist and Louise Haigh for making sure lab work always ran smoothly.

I am very grateful to all members of the York Structural Biology Laboratory and the Mottram group for always being supportive and willing to help both in and out of the lab. Thank you also to project students Camila Torres, Catrina Kenyon and Yitao Zhang for their assistance at various points in the project.

Many thanks to Sam Hart and Dr Johan Turkenburg for their assistance in X-ray data collection and analysis, and Dr Alex Heyam for his help with NMR. Within the University of York Technology Facility, I would particularly like to thank Dr Jared Cartwright and Rebecca Preece for protein production, and Dr Andrew Leech for carrying out SEC-MALLS and for assistance with the fluorescence polarisation assay.

I am especially grateful to my collaborators Professor Stuart Conway and Jennifer Carter at the University of Oxford, for synthesising and gifting compounds, and for their invaluable input and advice. Thank you also to Dr Corentine Laurin for teaching me the fluorescence polarisation assay technique. I would also like to extend my gratitude to collaborators at GSK, particularly to Jacob Bush, Arron Aatkar and Emma Grant for performing high-throughput screens, for exchange of materials

and for many helpful discussions. Thank you also to Raquel Gabarro-Carrion and Felix Calderon from GSK for kindly gifting compounds, as well as Dr Raymond Hui and the Structural Genomics Consortium for gifting compounds and plasmids.

I am very grateful to the Biotechnology and Biological Sciences Research Council for providing me with the funding to carry out this PhD project.

Thank you also to my wonderful family and friends for being a continual source of support and encouragement. Thank you mum and dad, for every science article you cut out of the newspaper for me, for learning how to pronounce leishmaniasis perfectly and for phoning me excitedly every time you heard someone talking about it on the radio. Thank you Daniel, for your endless support during the whole PhD and for your confidence in me which has always reassured me I'm on the right path. And thank you to my wonderful grandmothers, for inspiring me with your strength, and for your calls which always brightened my day during the long hours of thesis writing.

Finally, I wish to thank Steve Kendall and Rodney Kendall who brought the magic of science into my life. From analysing pond samples with an old microscope, to writing computer code, thank you for sharing your love of science with me and for being my inspirations.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. All experimental work carried out in this thesis was completed by the author with the exception of collaborative work outlined below. The writing of the thesis is all the authors own work. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Work carried out or overseen by others:

- SEC-MALLS experiments were performed and analysed by Dr Andrew Leech in the York Technology Facility

- *Ld*BDF5.1 and *Ld*BDF5.2 recombinant proteins were expressed and purified both by Juliet Borgia in the York Structural Biology Laboratory and by Rebecca Preece in the University of York Technology Facility

- Dr Nathaniel Jones at the University of York performed site-directed mutagenesis to produce *Ld*BDF2, 5.1 and 5.2 mutant plasmids

- Project student Catrina Kenyon assisted with primer design and experimental work for cloning *Ld*BDF1, 2, 3 and 4, and Camila Torres assisted with expression and purification of *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F), both under the direct supervision of the author

- X-ray diffraction data was collected in-house and at the Diamond light source (Harwell Science & Innovation Campus, Didcot) by Dr Johan Turkenburg and Sam Hart of York Structural Biology Laboratory

- High-throughput screens were performed by Jacob Bush, Arron Aatkar and Emma Grant at GSK, and acetyllysine mimic compounds were synthesised by Jennifer Carter and Professor Stuart Conway at the University of Oxford

More detailed acknowledgements are also given in the relevant sections of the text.

1. Chapter 1: Introduction

1.1. Leishmaniasis

Leishmaniasis is a parasitic disease classed by the World Health Organisation as one of 20 neglected tropical diseases, due to its high prevalence in tropical areas associated with poverty (WHO, 2022b). The disease is a huge burden in developing countries, with up to a million new cases predicted annually (WHO, 2022a). Leishmaniasis is caused by *Leishmania* parasites, which belong to a group of flagellated protozoa called the kinetoplastids (Stuart *et al.*, 2008).

1.1.1. *Leishmania* life cycle and genetics

Leishmania parasites are transmitted via a sandfly insect vector, and exhibit a complex lifecycle (Figure 1.1) involving human and insect stages, as well as differentiation through two main forms of the parasite; amastigotes and promastigotes. The amastigote form is found in the mammalian host and at this developmental stage, the parasite exhibits a smaller, spherical form lacking a functionally motile flagellum (Sunter and Gull, 2017). Promastigotes are the motile form found in the insect vector which have an elongated body and a flagellum important for motility, as well as conveying possible attachment and sensory functions (Gluenz, Ginger and McKean, 2010). *Leishmania* promastigotes can be further divided into four main forms, classified by the size of the cell body and flagellum. These forms are procyclic, nectomonad, leptomonad and metacyclic, where all four morphologies occur in the sand fly vector and metacyclic promastigotes are the mammalian infective form (Gossage, Rogers and Bates, 2003; Sunter and Gull, 2017).

Leishmania are first transmitted into a human host when a female sandfly vector takes a blood meal and metacyclic promastigotes are introduced into the skin. Within the host, the parasites infect immune cells such as macrophages. In the phagolysosome, metacyclic promastigotes differentiate into intracellular amastigotes, multiply, are released, and go on to infect other cells and tissues. The cycle passes back into the insect stage when another sandfly takes a blood meal from an infected individual and ingests cells infected with amastigotes. In the

sandfly midgut, amastigotes are released as the cells are ingested, and differentiate into procyclic promastigotes and subsequently nectomonad promastigotes which migrate to the proboscis where they again differentiate into leptomonad promastigotes. The cycle is then repeated when the parasites differentiate back to metacyclic promastigotes and are transmitted to another mammalian host when the infected sandfly takes another blood meal (Gossage, Rogers and Bates, 2003; Jain and Jain, 2013).

In addition to the human host, other mammals (both domestic and wild) can act as reservoirs for the parasite, such as dogs and rodents, with increasingly challenging disease control pertaining to these zoonotic leishmaniases (WHO, 2010).

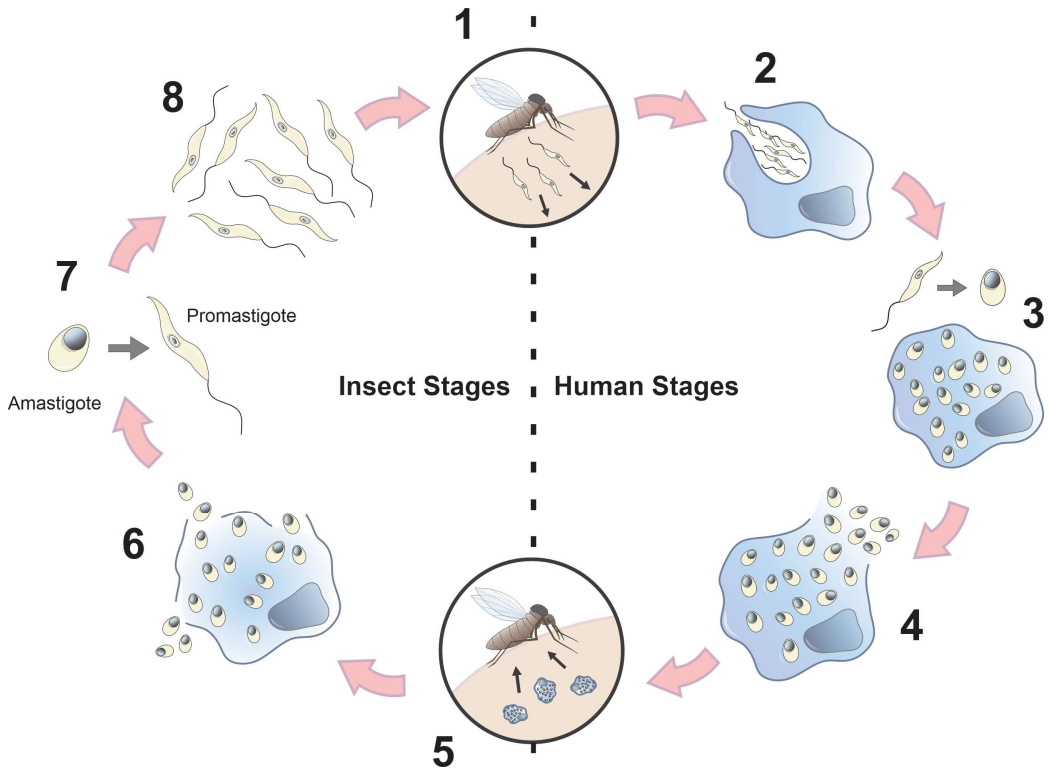


Figure 1.1. *Leishmania* parasite life cycle. 1 The sandfly takes a bloodmeal, injecting *Leishmania* promastigotes into a human host. 2 Promastigotes are phagocytosed. 3 promastigotes differentiate into amastigotes and multiply. 4 Amastigotes are released from the cell and infection spreads. 5 A sandfly ingests infected cells. 6 Amastigotes are released into the sandfly. 7 Amastigotes differentiate back into promastigotes. 8 promastigotes differentiate through procyclic, nectomonad, leptomonad, metacyclic forms and replicate.

Navigating through this complex life cycle, it is crucial that *Leishmania* can enact precise gene regulation mechanisms to alternate between these structurally and phenotypically distinct differentiation stages. *Leishmania* fall into a category of flagellated protozoans called the kinetoplastids, characterised by presence a DNA-containing structure known as the kinetoplast. Other kinetoplastids include the *Trypanosoma* species, including *Trypanosoma brucei* (*T. brucei*) and *Trypanosoma cruzi* (*T. cruzi*) which are the causative agents of human African trypanosomiasis (sleeping sickness) and Chagas disease, respectively (Stuart *et al.*, 2008). Kinetoplastids exhibit unusual genome architecture distinct from the mammalian host. In contrast to other eukaryotes, genes are not organised into single intron-containing transcription units with an associated promoter and terminator corresponding to expression of a single protein. Instead, the genome is arranged into clusters of non-functionally related genes called polycistronic transcription units (PTUs). These intron-lacking PTUs are associated with transcriptional start regions (TSR) and transcriptional termination sites (TTS).

1.1.2. Leishmaniasis disease forms

There are over 20 different species of *Leishmania* responsible for causing leishmaniasis and different forms of the disease are caused by different species in different geographical locations. These different disease forms can be characterised as Old World or New World, with prevalence in the Eastern and Western hemispheres, respectively. Causative agents of Old World leishmaniasis include the *Leishmania infantum* (*L. infantum*)*, Leishmania donovani* (*L. donovani*)*, Leishmania major* (*L. major*)*, Leishmania tropica* (*L. tropica*) and *Leishmania aethiopica* (*L. aethiopica*) species. New World Leishmania species include *Leishmania amazonensis* (*L. amazonensis*)*, Leishmania mexicana* (*L. mexicana*), *L. infantum,* and the *Viannia* subgenus (WHO, 2010; Burza, Croft and Boelaert, 2018).

The three most prevalent forms of leishmaniasis are cutaneous, mucocutaneous, and visceral or 'kala-azar'. Cutaneous leishmaniasis is not typically life-threatening, with clinical manifestations such as skin lesions and ulcers which often self-heal or require only local treatment. Mucocutaneous leishmaniasis is a more severe

progression of cutaneous leishmaniasis, caused exclusively by New World species including *L. braziliensis* and *L. panamensis*. Here, infection spreads to mucosal tissue in areas such as the mouth and upper respiratory tract and can cause destructive lesions and permanent disfigurement. Finally, the most severe form of the disease is visceral leishmaniasis, mostly caused by *L. donovani* and *L. infantum* species. In 2012 it was reported that over 90% of worldwide visceral leishmaniasis cases occurred in just six countries (Bangladesh, Brazil, Ethiopia, India, South Sudan and Sudan) (Alvar *et al*., 2012), and children under 10 years old are at the highest risk of developing this form of the disease. Onset of visceral leishmaniasis is gradual with an incubation period of anywhere from 10 days to over a year. The parasite infects organs including the spleen and liver, with symptoms such as fever, weight loss and acute malnutrition. These are often accompanied by clinical manifestations such as damage to mucous membranes and enlargement of the spleen. Mortality is usually a result of severe anaemia or intercurrent infection, often pneumonia. Coinfection with HIV is also being increasingly reported. The disease forms described here are summarised from more comprehensive reviews of leishmaniasis epidemiology and pathology (WHO, 2010; Burza, Croft and Boelaert, 2018; Scarpini *et al.*, 2022).

1.1.3. Current treatments

There are various diagnostic tools used to identify leishmaniasis infections, with molecular techniques such as PCR offering one of the preferred methods. Treatment of leishmaniasis varies depending upon factors such as geographical location, parasite species and disease form; for example cutaneous leishmaniasis is often self-healing whereas, without treatment, visceral leishmaniasis fatality rates are estimated to be over 90% (Georgiadou, Makaritsis and Dalekos, 2015).

There are no approved vaccines currently available to combat leishmaniasis, though clinical trials have begun for a vaccine recently developed by researchers at the University of York which has shown promising results in phase 1 and 2a trials (Musa *et al.*, 2022). Treatment of leishmaniasis with existing chemotherapies is limited by high costs, long treatment time, need for a temperature-controlled supply chain, toxicity and growing resistance (Jain and Jain, 2018). Drugs used for

the treatment of leishmaniasis include pentavalent antimonials, amphotericin B and miltefosine (Figure 1.2), as well as paromomycin and pentamidine, but all have drawbacks.

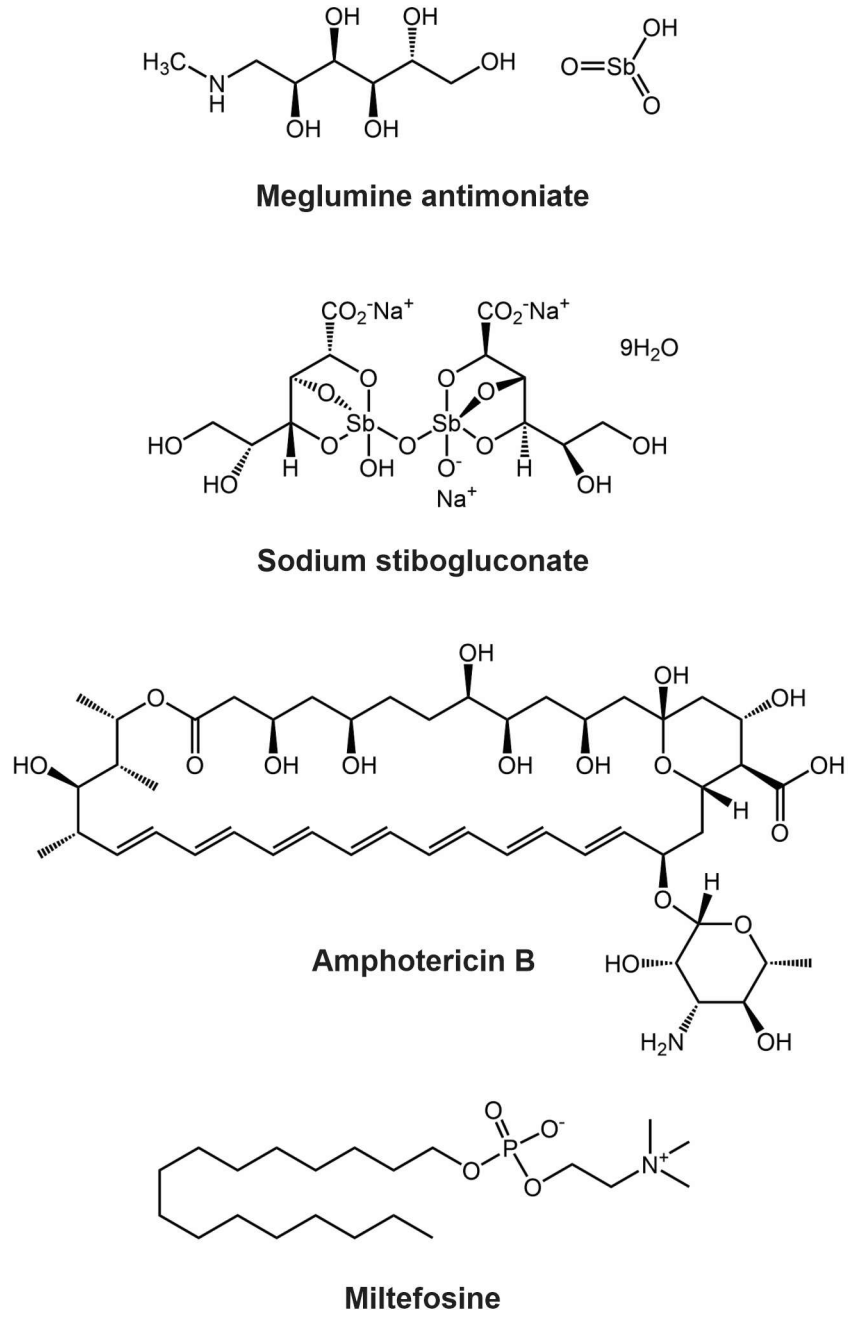


Figure 1.2. Current antileishmanial drugs. Structures of pentavalent antimonials (meglumine antimoniate and sodium stibogluconate), amphotericin B and miltefosine which currently used to treat leishmaniasis.

Pentavalent antimonials were one of the first classes of antileishmanials to be used clinically and remain one of the first-line treatment options. Two pentavalent antimonials commonly used to treat leishmaniasis are meglumine antimoniate and sodium stibogluconate (Figure 1.2). Pentavalent antimonials become biologically active after being reduced to a trivalent form which is able to disrupt the thiol redox potential and trypanothione metabolism in the parasite. Two mechanisms have been described through which this may be achieved. The first is by inducing efflux of trypanothione and glutathione from the cell, causing a reduction in the thiol-buffering capacity, and the second is inhibition of the trypanothione reductase enzyme (Wyllie, Cunningham and Fairlamb, 2004). *Leishmania* are especially reliant on a functioning trypanothione system to overcome oxidative stress within macrophages which may explain why pentavalent antimonials are such an effective treatment. Despite being a front-line drug for leishmaniasis and showing high cure rates, pentavalent antimonials require parenteral administration and are associated with toxic side effects ranging from nausea and lethargy to cardiotoxicity and pancreatitis. Additionally, resistance is a major problem in areas such as India and Nepal (Frézard *et al.*, 2009; Georgiadou, Makaritsis and Dalekos, 2015).

Amphotericin B (Figure 1.2) is another widely used treatment for visceral leishmaniasis, commonly used as a second-line drug in regions where there is resistance to pentavalent antimonials (WHO, 2010). Amphotericin B has a high affinity for ergosterol and upon binding to this sterol in the parasite, causes pore formation in phospholipid bilayers accompanied by the leaking of small molecules out of the cell. The activity of this compound may also be attributed to oxidative damage via formation of reactive oxygen species (Kamiński, 2014). Like pentavalent antimonials, amphotericin B requires parenteral administration, and despite high cure rates, this drug is also associated with adverse side effects including myocarditis, hypokalemia and nephrotoxicity (Sundar and Chakravarty, 2015).

Miltefosine (Figure 1.2) is an alkyl-lysophospholipid that was originally developed as an anti-cancer drug and is the only oral treatment of visceral leishmaniasis. Miltefosine exerts its antileishmanial effect through several complex mechanisms. These include inhibition of lipid metabolism (Rakotomanga *et al.*, 2007), disruption of calcium homeostasis (Pinto-Martinez *et al.*, 2018), and disruption of

mitochondrial function via inhibition of cytochrome c oxidase (Luque-Ortega and Rivas, 2007). Additionally, miltefosine has a host cell-dependent mechanism whereby the compound induces nitric oxide synthetase 2 in infected macrophages, stimulating the production of toxic nitric oxide which kills the parasite (Wadhone *et al*., 2009). However, miltefosine too has side effects, including diarrhoea, vomiting and dehydration. Additionally, miltefosine cannot be used during pregnancy and resistance has been observed (WHO 2010; Scarpini *et al*., 2022).

Combination therapy is another approach used for leishmaniasis treatment and offers advantages in the reduction of both treatment time and toxic side effects due to lower doses of each drug being required (Georgiadou, Makaritsis and Dalekos, 2015). Nevertheless, the inadequacy of current treatment options is indisputable and there is a pressing international need for the development of new antileishmanials.

In response to these needs, a not-for-profit research and development organisation, the Drugs for Neglected Diseases Initiative (DNDi), was founded in 2003 with the aim to deliver new treatments for neglected tropical diseases including visceral leishmaniasis. To address this aim, DNDi have established a research and development portfolio covering the entire drug development pipeline. The first stage in the pipeline is drug discovery, which involves initial research, compound screens, lead selection and subsequent optimisation. Translational research then progresses promising drug candidates to pre-clinical and clinical trials, with successful treatments finally becoming available to patients (Chatelain and Ioset, 2011). Through collaborative efforts, research and clinical trials, DNDi have already delivered four new visceral leishmaniasis treatments (DNDi, 2023) including an improved, combination therapy treatment option for patients co-infected with visceral leishmaniasis and HIV (Burza *et al*., 2022). Furthermore, there are also several promising compounds progressing through the pipeline, including DNDI-0690, DNDI-6148, DNDI-6174, GSK3186899/ DD853651, GSK3494245/ DDD1305143, and LXE408 (DNDi, 2023). Whilst these examples show that there are new antileishmanial compounds in development, there are still several major hurdles to be overcome before they can reach patients. Therefore, it is essential that research into new antileishmanials continues. With the lack of well-

defined drug targets in *Leishmania*, new avenues need to be explored. In light of the vastly different gene expression system of the parasite compared with its human host, selective disruption of *Leishmania* gene regulation may offer an enticing new strategy for the development of new antileishmanials.

1.2. Epigenetic histone modification

Epigenetics is defined as 'heritable changes in gene function that cannot be explained by changes in DNA sequence' (Riggs, Martienssen and Russo, 1996). In other words, phenotypes that do not arise from changes to the DNA sequence, but rather where alterations to gene expression are brought about by the presence or absence of chemical marks on histones or DNA.

1.2.1. Histone modifications

Within chromatin, DNA is packaged into structures called nucleosomes (Figure 1.3A) which consist of around 150 bp DNA wrapped around an octamer of histone proteins in a left handed superhelix. The histone octamer contains two copies of each of the four core histone proteins, H2A, H2B, H3 and H4. These small proteins are typically around 11 – 15 kDa and highly conserved across eukaryotes (Cutter and Hayes, 2015). There are also variant forms of histones which share a degree of homology with the canonical proteins and have distinct functions, such as H2A.Z, which has specialised roles in gene regulation (Giaimo *et al.*, 2019).

Histone proteins (Figure 1.3B) have a largely α-helical core, and unstructured N- and C-terminal tails which can undergo post-translational modifications (PTMs) such as acetylation, methylation and phosphorylation (Figure 1.4A). These covalent modifications form what is known as the histone code; specific arrangements of PTMs on histones which can influence chromatin organisation and gene expression (Strahl and Allis, 2000). Histone modifications are one of the two main mechanisms of epigenetic regulation, the other being DNA methylation.

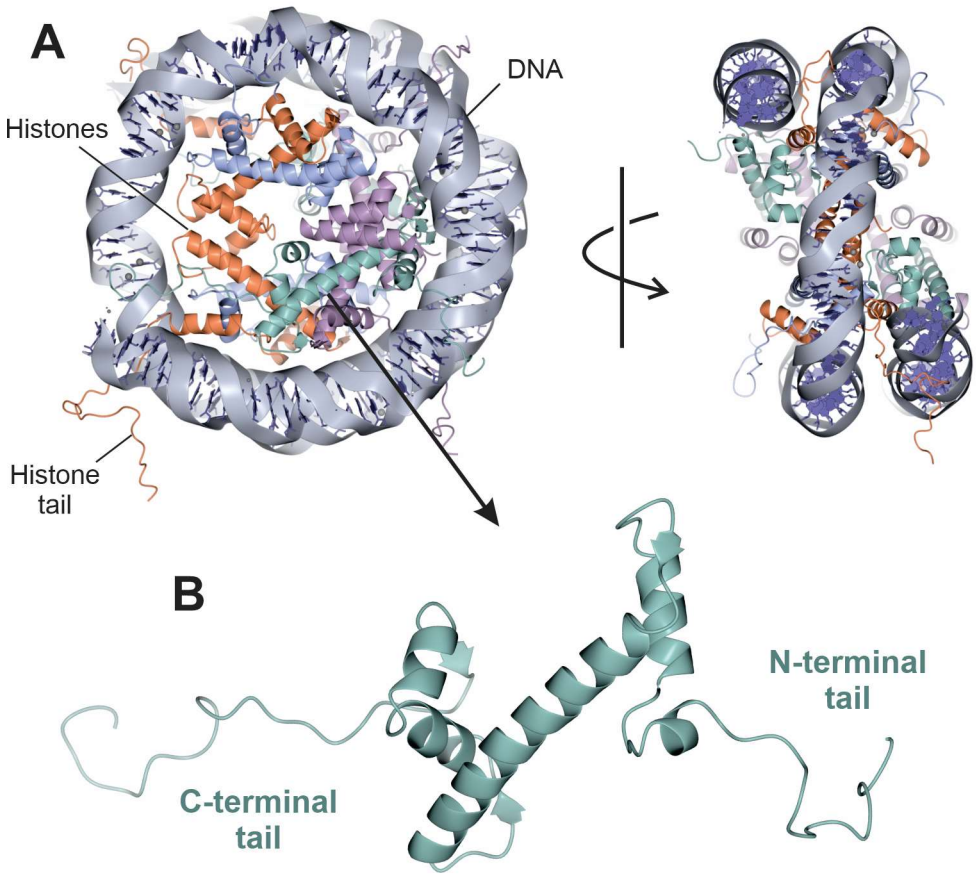


Figure 1.3. Structure of the nucleosome (A) Structure of a nucleosome core particle with DNA wrapped around an octamer of H2A, H2B, H3 and H4 histones, with each histone protein shown in a different colour. **(B)** Structure of a histone protein from the nucleosome, with the N- and C- terminal tails indicated. PDB code 1EQZ (Harp *et al.*, 2000). Image produced using CCP4mg software (McNicholas *et al.*, 2011).

1.2.2. Lysine acetylation

Lysine acetylation (Figure 1.4) is a common epigenetic histone modification, usually associated with open chromatin (euchromatin) and active gene transcription. This is because at physiological pH the amino group of lysine is protonated (pK_a = 10.5), but acetylation neutralises this positive charge, disrupting electrostatic interactions between histones and DNA, leading to unravelling of chromatin allowing access to transcriptional machinery (Ali *et al.*, 2018; Ringel, Tucker and Haigis, 2018).

Epigenetic histone modifications are dynamic; they can be added and removed by catalytic 'writer' and 'eraser' proteins and recognised by 'reader' proteins. In this case, lysine acetylation is catalysed by histone acetyltransferases (HATs) and deacetylation by histone deacetylases (HDACs) (Figure 1.4B). The reader proteins of acetylated lysine, or 'acetyllysine', recognise this mark using a domain known as the bromodomain (Figure 1.4C). Bromodomains are the best characterised readers of histone acetyllysine marks, though other domains that can bind this PTM include PHD fingers, YEATS domains, and PH domains (Lange *et al.*, 2008; Zeng *et al.*, 2010; Ali *et al.*, 2012; Su *et al.*, 2012; Li *et al.*, 2014).

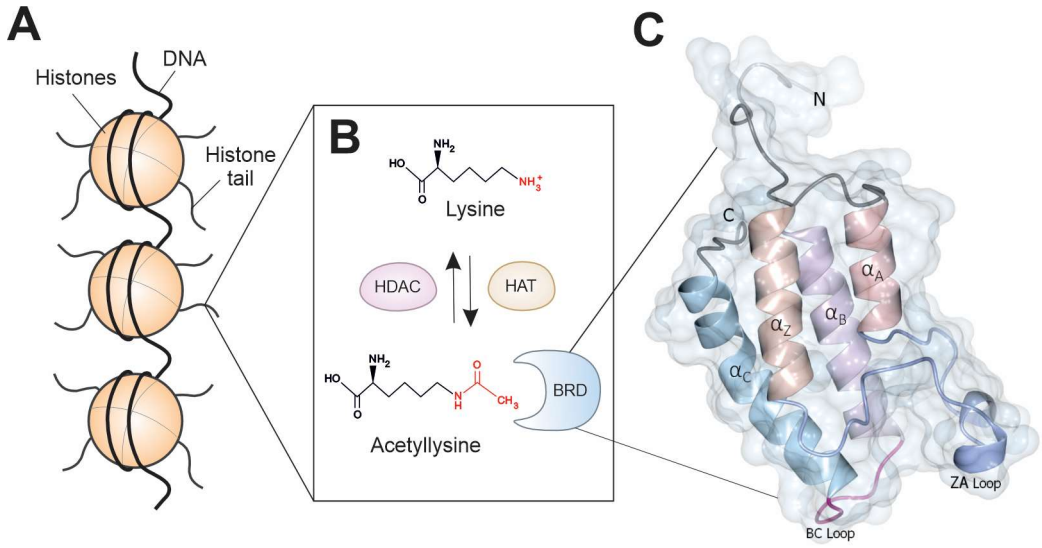


Figure 1.4. Acetylation of histone lysine residues as a post-translational epigenetic modification. (A) Structure of chromatin showing nucleosomes formed by DNA packaged around histone proteins. **(B)** Acetylation of lysine residues on histone tails added by histone acetyltransferase (HAT) and removed by histone deacetylase (HDAC) proteins, with recognition by bromodomain proteins (BRD). **(C)** Conserved bromodomain structure with four anti-parallel alpha helices linked by two prominent loops which form the binding pocket (demonstrated by *L. donovani* BDF2 bromodomain; PDB code 5C4Q, image produced using CCP4mg software).

1.3. Bromodomains

1.3.1. Bromodomain structure and acetyllysine binding

Bromodomains are small protein domains of around 110 amino acid residues which were first reported in a nuclear remodelling protein encoded by the *brahma* gene in *Drosophila melanogaster* (Tamkun *et al.*, 1992). Structurally, they consist of a left-handed bundle of four antiparallel alpha helices, termed αZ, αA, αB and αC, connected by loops called the ZA loop and a BC loop (Figure 1.4C). These loops form a hydrophobic pocket which acts as the binding site of the protein (Filippakopoulos and Knapp, 2012). Acetyllysine residues are the best characterised binding partners of bromodomains, though they have also been shown to bind larger lysine modifications such as butyryllysine and crotonyllysine (Flynn *et al.*, 2015).

Neutralisation of the positive charge on lysine allows it to be accommodated in the bromodomain binding pocket, where it is recognised by two conserved residues; asparagine and tyrosine (Figure 1.5). Acetyllysine is anchored via a hydrogen bond with the side chain amide NH_2 group of the conserved asparagine residue. Another conserved feature of bromodomains is a network of water molecules in the binding cavity, and acetyllysine interacts with the phenol group of the conserved tyrosine residue through a water-mediated hydrogen bond involving one of these structured water molecules (Conway, 2012; Filippakopoulos and Knapp, 2012).

Binding specificity is conveyed in part by sequence variations in the flexible ZA and BC loops, which vary in length and structure between different bromodomains. On the other side of this interaction, bromodomains affinity for a single acetyllysine residue is low, and the flanking residues are needed to provide specificity and elicit a higher affinity interaction (Mujtaba, Zeng and Zhou, 2007).

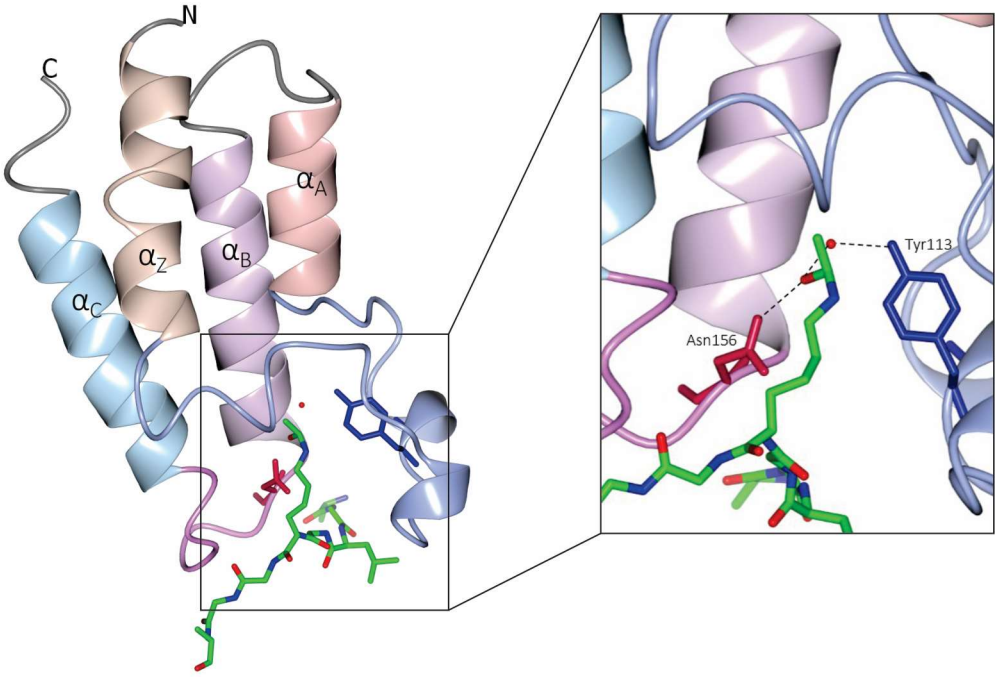


Figure 1.5. Bromodomain binding acetyllysine. Binding mechanism of bromodomains illustrated by human BRD2 complexed with acetylated histone H4 peptide (PDB code 2DVQ; (Umehara *et al.*, 2010) where the carbonyl group in acetyllysine forms a hydrogen bond with the conserved asparagine residue (pink) and a water-mediated hydrogen bond with the conserved tyrosine (dark blue), with the water molecule shown in red. Image produced using CCP4mg software.

1.3.2. Human BET bromodomains

The human genome encodes 46 bromodomain-containing proteins and 61 distinct bromodomains. These bromodomains have been comprehensively studied both in regards to their structure and histone substrates, and have been categorised into eight defined families (I – VIII) based on sequence and structure homology (Filippakopoulos and Knapp, 2012; Filippakopoulos *et al.*, 2012). Some families have been much better characterised than others, with arguably the most well studied being family II, containing the bromodomain and extra-terminal (BET) bromodomains. This has led to the common classification of human bromodomains as either BET or non-BET. BET proteins are characterised by their conserved protein domain architecture containing two tandem N-terminal bromodomains and an extra-terminal (ET) protein-protein interaction domain. This family contains four

members; BRD2, BRD3, BRD4 which are expressed ubiquitously, and BRDT which has testis-specific expression (Chung *et al.*, 2011). BET proteins regulate gene expression with well explored roles in various processes including cell cycle progression, erythropoiesis and spermatogenesis (Maruyama *et al.*, 2002; Berkovits and Wolgemuth, 2013; Stonestrom *et al.*, 2015). The relative wealth of research pertaining to these proteins forms the basis for much of our understanding and predictions about other bromodomains outside the human BET family.

1.3.3. Tandem bromodomain proteins

Bromodomain-containing proteins typically have one or two bromodomains, though they can have as many as six, as is the case for polybromo-1 protein (Thompson, 2009). Many bromodomain proteins, including those of the aforementioned BET family, have two tandem bromodomains (commonly referred to as BD1 and BD2). There has been enormous interest in understanding how tandem bromodomains function in relation to one another, and numerous studies have sought to answer this question. In spite of this, their architecture and binding mechanisms remain largely elusive.

To date, structures of only two tandem bromodomain-containing proteins are available in the protein data bank (PDB). The first is human TAF1 (TAFII250), the largest subunit of the TFIID multiprotein complex which is involved in transcription initiation. The structure of this protein has been solved by X-ray crystallography (PDB code 1EQF) and shows tandem bromodomains arranged side-by-side with both binding sites facing the same direction, separated by around 25 Å. The bromodomains in this protein are connected by an additional 'linking helix' between αC of BD1 and αZ of BD2. Based on the arrangement of the bromodomains in this structure, and isothermal titration calorimetry (ITC) data showing a weaker affinity for a mono-acetylated H4 peptide compared with di-acetylated peptides, it was suggested that BD1 and BD2 may act in a cooperative fashion to engage di-acetylated histone peptides (Jacobson *et al.*, 2000). However, a more recent study proposed an alternative model whereby the tandem bromodomains of TAF1 independently recognise their substrates, based on ITC and photo-cross-linking

data (Yadav *et al.*, 2022). Subsequent unliganded and inhibitor-bound structures solved of the TAF1 tandem bromodomains also showed different orientations of the bromodomains, indicating the protein may exhibit distinct conformational states which can be stabilised by inhibitors (Filippakopoulos *et al.*, 2012; Karim *et al.*, 2022).

The other tandem bromodomain structure that has been solved is that of Rsc4, a subunit of the yeast chromatin structure remodelling complex (RSC). Like TAF1, the tandem bromodomains were similarly oriented side-to-side with binding pockets facing the same way. However, this structure was more compact and there was no helix in the linker region between BD1 and BD2. There was, however, a short extra helix in BD2 between αZ and αA, hypothesised to facilitate histone recognition. Furthermore, a crystal structure of the protein in complex with an acetylated H3 peptide revealed that BD2 is responsible for binding the histone substrate. In an attempt to better visualise this interaction, the peptide was fused to the N-terminus of the Rsc4 tandem bromodomain protein and, following expression, acetylated by the HAT protein Gcn5. Interestingly, mass spectrometry revealed that not only was the histone peptide acetylated, but so too was a lysine residue (K25) in the N-terminus the Rsc4 protein itself. Even more surprisingly, an X-ray structure of this fusion protein showed that the acetylated H3 peptide did not bind BD2, and instead the acetylated K25 of Rsc4 was bound by BD1 which prevented interaction of BD2 with the histone peptide (Figure 1.6). Thus, a novel tandem bromodomain auto-regulation mechanism was proposed in which both the bromodomain protein and its histone substrate are both acetylated by the same HAT protein, and acetylation of the bromodomain protein causes binding of BD1 to another region of the protein itself, which antagonises binding of BD2 to the natural histone substrate (VanDemark *et al.*, 2007).

It is evident from these pieces of research that there is still much to be learned about tandem bromodomains and their mechanisms of action, however these proteins clearly have complex functions.

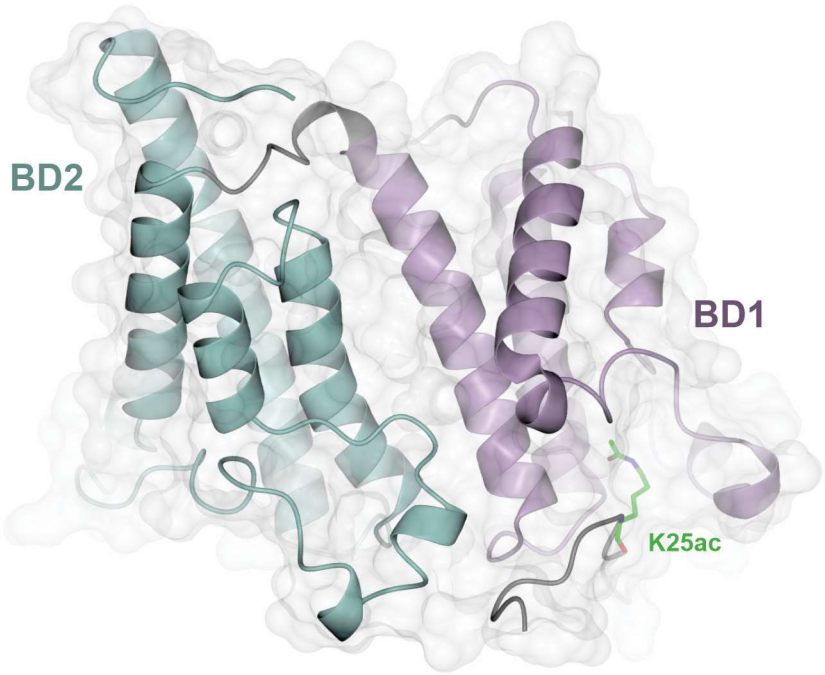


Figure 1.6. Structure of Rsc4 tandem bromodomains exhibiting auto-regulatory binding. Structure of the tandem bromodomains BD1 (purple) and BD2 (teal) with acetylated K25 in the N-terminal tail adjacent to BD1 bound within the BD1 binding pocket (PDB code: 2R10) (VanDemark *et al.,* 2007). Image produced using CCP4mg software.

1.3.4. Bromodomain roles in homeostasis and disease

Bromodomain proteins can regulate gene expression through various mechanisms. One example is that bromodomain-containing proteins can act as scaffolds for the recruitment of other proteins to DNA. The BET proteins are an example of this, and can use their bromodomain to recognise acetylated chromatin, whilst the ET domain recruits transcription machinery (Fujisawa and Filippakopoulos, 2017). Bromodomains are commonly found in larger multidomain proteins alongside functional domains such as PHD, HAT and methyltransferase domains (Filippakopoulos *et al.*, 2012). Bromodomains with these additional catalytic domains can have histone modification roles, for example the protein MLL contains both a bromodomain and methyltransferase and its methyltransferase activity is stimulated with histone H3 peptide acetylation (Milne *et al.*, 2002). Bromodomain proteins themselves can also be part of larger protein complexes, such as the

SWI/SNF complex which contains a central bromodomain-containing ATPase and has roles in chromatin remodelling (Singh *et al.*, 2007). Another role of bromodomain-containing proteins is acting as transcription co-modulators, such as TRIM28 which associates with HP1 family proteins and acts as a co-repressor of KRAB-domain-containing zinc finger genes involved in embryogenesis (Cammas *et al.*, 2002).

However, aberrant histone modification patterns and dysregulation of reader, writer and eraser proteins can lead to the development and exacerbation of disease. Bromodomain proteins in particular have been implicated in many different diseases. These range from haploinsufficiency of BAZ1B contributing to neurological disorders, to histone acetylation by CREBBP being linked to lung inflammation (Ito, Charron and Adcock, 2007; Lalli *et al.*, 2016). Bromodomain proteins also have involvement in pathogenic diseases, including HIV (Boehm *et al.*, 2013; Alamer *et al.*, 2021). The most well studied diseases in which bromodomain proteins have been implicated are undoubtedly cancers, including NUT midline carcinoma, leukaemia and breast cancer, to name just a few (Muller, Filippakopoulos and Knapp, 2011).

1.4. Bromodomain inhibitors

Consistent with the observation that bromodomain proteins are involved in a range of different diseases, copious research is now focusing on the development of bromodomain inhibitors. These efforts have been particularly rewarding with regard to cancer research, and many bromodomain inhibitors have now progressed through to clinical trials (Theodoulou, Tomkinson, *et al.*, 2016; Shorstova, Foulkes and Witcher, 2021).

A large proportion of bromodomain inhibitors have been optimised to target BET proteins due to the relative wealth of information about this bromodomain family, and increasing evidence of their involvement in disease. One of the first bromodomain inhibitors developed was JQ1 (Figure 1.7A) which is an example of a BET inhibitor. This compound showed selectivity for the BET proteins with

particularly high affinity binding for BRD4. A crystal structure of this compound in complex with BRD4 BD1 showed that the triazole ring formed a hydrogen bond with the conserved asparagine in the bromodomain binding pocket. Furthermore, this binding effectively displaced BRD4 from chromatin and induced an antiproliferative effect in NUT midline carcinoma cells, highlighting its therapeutic potential (Filippakopoulos *et al.*, 2010). Non-BET inhibitors have also been developed, such as I-BRD9 (Figure 1.7B), a compound developed based on a thienopyridone scaffold which binds to BRD9 with 700-fold selectivity for this protein over BET bromodomains (Theodoulou, Bamborough, *et al.*, 2016). Whilst many bromodomain inhibitors have shown potential for clinical applications, compounds have also been developed specifically as tools to probe cellular function of bromodomains. One notable example is bromosporine (Figure 1.7C), a broad-spectrum bromodomain inhibitor derived from a triazolopyridazine dicyclic core which binds to a large, diverse range of bromodomains, with up to nanomolar affinity (Picaud *et al.*, 2016).

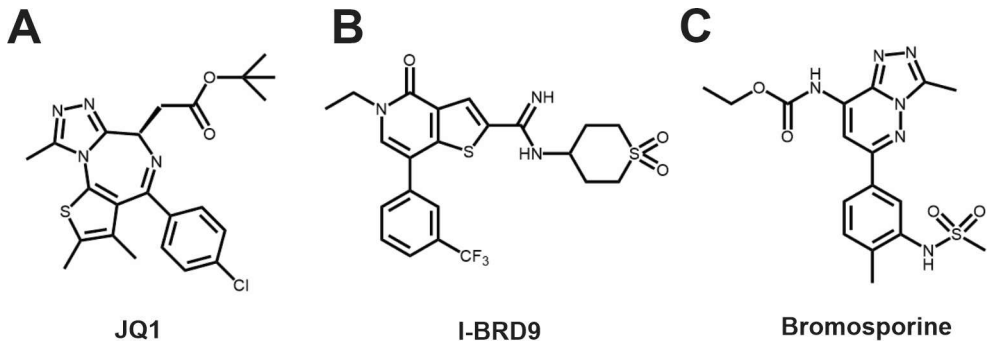


Figure 1.7. Structures of three bromodomain inhibitors. Structures of three human bromodomain inhibitors; **(A)** BET inhibitor, JQ1, **(B)** non-BET inhibitor, I-BRD9 and **(C)** pan-bromodomain inhibitor, bromosporine.

As mentioned, many bromodomain proteins contain two tandem bromodomains, and details are beginning to emerge as to how these two domains may be targeted, either individually or dually. Compounds which selectively bind one bromodomain in a tandem bromodomain pair have been highly sought after. One such compound

is RVX-208 which showed 23-fold selectivity for BRD2 BD2 over BD1, with binding associated with a conformational change which may explain the selectivity for this bromodomain (Picaud *et al.*, 2013). RVX-208 was also one of the first bromodomain inhibitors to enter clinical trials for non-oncological applications, for the treatment of high-density lipoprotein-linked diseases (Nicholls *et al.*, 2011). Another recent example is ABBV-744 which showed potent and selective binding to BET BD2 and holds promise for further applications in cancer therapy (Faivre *et al.*, 2020). Dual-bromodomain inhibitors have also been developed such as MT1 which binds the tandem bromodomains of BRD4 in a bivalent fashion. A co-crystal structure of this interaction showed the compound bound within a large hydrophobic pocket formed at the interface between the two bromodomains, simultaneously interacting with both BD1 and BD2 (Tanaka *et al.*, 2016).

Another exciting new avenue for bromodomain inhibitor research is the use of proteolysis targeting chimeras (PROTACS). Tethering of the BET inhibitor JQ1 (Figure 1.7A) to an E3 ubiquitin ligase VHL ligand generated a PROTAC compound named MZ1 which showed selectivity for BRD4, despite JQ1 alone binding BET proteins BD2, 3 and 4 with similar affinities. Possible explanations for this selectivity include distinct polyubiquitination of lysine residues on the surfaces of the three proteins, or a preferential interaction of the E3 ubiquitin ligase VHL with BRD4 compared with BRD2 or BRD3. Furthermore, the PROTAC was able to elicit intracellular degradation of BRD4, presenting another exciting new method for the therapeutic targeting of bromodomains (Zengerle, Chan and Ciulli, 2015).

These successes seen in bromodomain inhibitor research have served to validate these proteins as chemically tractable drug targets for the treatment of many diseases. In light of their essential roles in gene expression regulation, and their proven druggability, bromodomains may also present an attractive target for the development of inhibitors in organisms aside from humans.

1.5. Epigenetics in kinetoplastid parasites

Kinetoplastid parasites have complex life cycles with several differentiation stages, so precise regulation of gene expression is crucial. Epigenetics may offer one way in which parasites achieve this tight control over transcription regulation, and epigenetic proteins, including bromodomain proteins, have been identified in kinetoplastids.

1.5.1. *T. cruzi* and *T. brucei* bromodomain proteins

Leishmania, along with *T. brucei* and *T. cruzi* comprise a group of mono-flagellated protozoan parasites known as the 'TriTryps'. Five orthologous bromodomain factor (BDF) proteins have been identified across the TriTryps (BDF1 – 5) alongside an additional three non-canonical bromodomain proteins (BDF6 – 8) (Jones *et al.*, 2022). Efforts have been made to characterise the bromodomain factor (BDF) proteins in all three species.

In *T. brucei*, for example, BDF3 was found to co-localise with histones H4K10ac, H2AZ and H2BV at RNA polymerase II transcription start sites, suggesting a role in transcription initiation (Siegel *et al.*, 2009). It has also been suggested that *T. brucei* BDF5 is part of a transcription initiation complex on account of its association with HAT2 and transcriptional start regions (TSRs) (Staneva *et al.*, 2021). One study has proposed that the bromodomain proteins in *T. brucei* contribute to maintenance of the bloodstream form of the parasite, using both genetic manipulation and targeted inhibition using the bromodomain inhibitor I-BET151. In this work, a co-crystal structure was also solved, showing binding of *T. brucei* BDF2 to this bromodomain inhibitor (Schulz *et al.*, 2015). Another paper also showed inhibition of *T. brucei* BDF2 using the inhibitor GSK2801, which prevented binding to the proteins acetylated H2AZ substrate and disrupted parasite growth (Yang *et al.*, 2017).

Previous research has also uncovered properties of the *T. cruzi* BDF proteins and investigated their inhibition. Histone binding interactions of *T. cruzi* BDF2 have been uncovered, with the protein exhibiting binding to acetylated H4 (Villanova *et al.*, 2009). Whilst bromodomains are typically nuclear, *T. cruzi* proteins BDF1 and 3

both showed non-nuclear localisation in Western blot and immunofluorescence analyses. BDF1 localises to the glycosome, and its overexpression was disruptive in multiple parasite life cycle stages (Ritagliati *et al.*, 2016). Contrastingly, *T. cruzi* BDF3 was found in the flagellum where it associates with α-tubulin, thus a novel, non-nuclear bromodomain role was proposed in cytoskeleton dynamics (Alonso *et al.*, 2014). Recently, a fragment-based approach has been used to identify novel ligands of *T. cruzi* BDF3, with development of a fragment hit compound yielding tool compounds for application in future biophysical assays (Laurin *et al.*, 2021). With regards to *T. cruzi* bromodomain inhibition, research has shown that human bromodomain inhibitors, JQ1 and I-BET151 bind to *T. cruzi* BDF3 and disrupt epimastigote growth as was evidenced by a reduction in the number of viable epimastigotes counted 72 hrs after treatment with these compounds (Alonso *et al.*, 2016).

This work has not only begun to shed light on the roles of BDFs in kinetoplastids, but crucially, also provides validation that the BDF proteins are druggable and their inhibition has disruptive effects on the parasite.

1.5.2. *Leishmania* bromodomain proteins and histone acetylation

The H2A, H2B, H3 and H4 histone proteins show relatively high homology among different *Leishmania* species, with *L. donovani* H3 and H4 displaying around 90% homology with their orthologues in *L. major* and *L. infantum* (Baharia *et al.*, 2014). Details have begun to emerge regarding acetylation of histones by HAT proteins in *Leishmania*, and in *L. donovani* acetylation sites of HAT1, 2, 3 and 4 have been identified in histones H2A.Z and H4 (Kumar *et al.*, 2012; Maity and Saha, 2012; Kumar and Saha, 2015; Chandra *et al.*, 2017; Jha *et al.*, 2017). These findings are covered in more detail in Chapter 4, and provide an initial indication that transcription regulation in *Leishmania* may be mediated by histone acetylation.

Five bromodomain proteins have been identified in *Leishmania* (BDF1 – 5) (Figure 1.8) which are essential for promastigote survival, as determined by Cas9-targeted gene deletion (Jones *et al.*, 2022). The conserved asparagine and tyrosine residues are present in all five proteins, and all contain a single bromodomain with the exception of BDF5 which contains two tandem bromodomains (termed BDF5.1 and

BDF5.2). BDF5 also appears to contain an additional C-terminal domain, predicted to be an MORF4-related gene (MRG) domain. These MRG domains are thought to have roles in chromatin remodelling and transcription regulation (Bowman *et al.*, 2006). BDF4 is predicted to contain an additional, zinc finger domain. A further three non-canonical/ pseudo-bromodomain proteins have also been identified (BDF6, 7 & 8) which lack some of the typical conserved features of bromodomains. The bromodomain of BDF6 is notably longer than the other BDFs, whilst BDF7 and BDF8 appear to lack the conserved asparagine and tyrosine residues (Jones *et al.*, 2022).

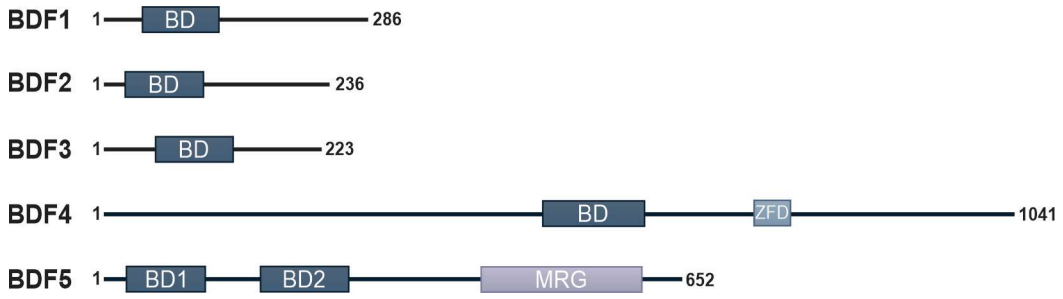


Figure 1.8. *Leishmania* bromodomain proteins BDF1-5. Protein domain architecture of the five essential *Leishmania* bromodomain factor proteins, showing bromodomains (BD), zinc finger domains (ZFD) and MRG domains (MRG), with length given in terms of amino acids.

Structures have been deposited to the protein data bank (PDB) for four *L. donovani* BDF (*Ld*BDF) bromodomains in complex with different human bromodomain inhibitors, solved by X-ray crystallography at resolutions of 1.93 – 2.60 Å. *Ld*BDF2, *Ld*BDF3 and *Ld*BDF5.2 were co-crystallised with bromosporine (PDB codes 5C4Q, 5FEA and 5TCK, respectively), whilst *Ld*BDF5.1 was co-crystallised with two different human bromodomain inhibitors; BI 2536 and SGC-CBP30 (PDB codes 5TCM and 6BYA, respectively) (Figure 1.9).

These structures demonstrate that the *L. donovani* BDF proteins exhibit a typical bromodomain structure comprising four anti-parallel alpha helices linked by two loops forming a hydrophobic binding pocket. These structures also provide

evidence that existing human bromodomain inhibitors are able to target bromodomains in *L. donovani,* and that these bind as predicted within the hydrophobic pockets of the proteins.

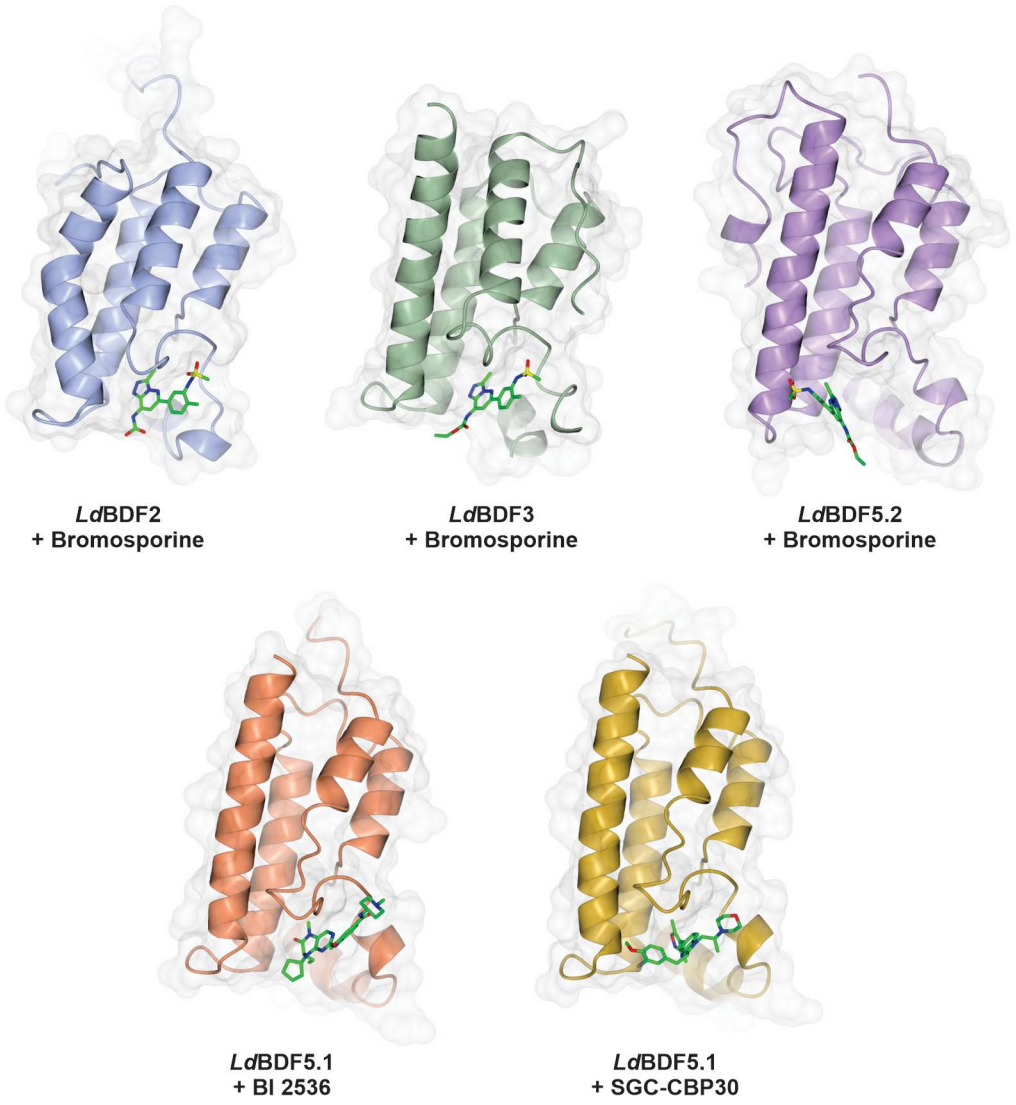


Figure 1.9. Structures of *L. donovani* BDF bromodomains in complex with human bromodomain inhibitors. Structrues of *Ld*BDF2 (PDB code 5C4Q), *Ld*BDF3 (PDB code 5FEA), *Ld*BDF5.2 (PDB code 5TCK) and *Ld*BDF5.1 (PDB codes 5TCM & 6BYA) in complex with the human bromodomain inhibitors bromosporine, BI 2536 and SGC-CBP30. Note the *Ld*BDF3 structure is domain swapped. Images produced using CCP4mg software.

1.5.3. *Leishmania* BDF5

BDF5 is the only *Leishmania* bromodomain protein known to have two bromodomains. Sequence and structure comparisons generated using the Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) and the Dali server (Holm and Sander, 1993) showed that the two individual bromodomains of *Ld*BDF5 exhibit high sequence and structural similarity with one other compared to bromodomains in other proteins. Recent work by Jones et al. (2022) has characterised the cellular interactions and roles of BDF5 in *L. mexicana*. BDF5 was found to be expressed in both lifecycle stages, and localised to the nucleus. In addition to being essential for promastigote survival, inducible knockout of BDF5 using a dimerisable split Cre recombinase (DiCre) system indicated that the protein is essential for amastigote survival and infection of the mammalian host. ChIP-seq data also revealed that BDF5 is enriched at TSRs and furthermore, RNA analysis revealed a global decrease in gene transcription following BDF5 deletion, indicating an essential role of the protein in RNA polymerase II-mediated transcription within the parasite.

Using an *in situ* proximity labelling technique (XL-BioID), the BDF5-proximal proteome was also characterised. This identified 156 proteins, 22 of which were further analysed by reciprocal co-immunoprecipitation which revealed 19 that co-precipitated with BDF5. Common functions of the 156 proximal proteins included roles in DNA replication, DNA repair, transcription and epigenetic reader & writer functions. Four histone proteins were identified; H2A.Z, H2B.V, H3 and H3.V, as well as the acetyltransferase protein, HAT2, which was among the most highly enriched and also one of the proteins verified by co-immunoprecipitation. Two of the other most highly enriched proteins were BDF3 and BDF4, with BDF6 and 8 also found in the proximal proteome. BDF3, BDF6 and BDF8 were all validated by co-immunoprecipitation. Based upon these findings, and alongside related data from immunoprecipitation experiments in *T. brucei* (Staneva *et al.*, 2021), it was proposed that BDF5 forms part of a complex with proteins including BDF3, BDF8 and HAT2, named the Conserved Regulators of Kinetoplastid Transcription (CRKT) complex, which associates with TSRs and influences transcription (Jones *et al.*, 2022).

There is currently no published work concerning inhibition of *Leishmania* BDF proteins. However, the genetic target validation of *Leishmania* BDF5 alongside the aforementioned research surrounding *T. brucei* and *T. cruzi* BDFs provides a rationale for the selective targeting of bromodomains in *Leishmania*, and highlights the potential of manipulating parasite epigenetics as an approach for the development of antiparasitic treatments.

1.6. Project aims

New treatments for leishmaniasis are urgently needed, and in light of evidence suggesting an important role for epigenetics in differential gene expression in parasites, this may offer an attractive new avenue for antileishmanial drug discovery. This project sought to investigate bromodomain proteins in *Leishmania* as targets for the development of new antileishmanials.

The three main aims of this project were:

1. Characterise the biological functions of bromodomain proteins in *Leishmania*

2. Validate *Leishmania* bromodomains as chemically tractable drug targets

3. Identify novel, high affinity inhibitors of the *Leishmania* bromodomain proteins

To realise these aims, it was necessary to develop a supply of recombinant *Leishmania* bromodomain proteins and establish assays to probe the interaction of these proteins with different ligands. Combined with existing genetic evidence of *Leishmania* BDF protein essentiality, achieving the aims set out here will strengthen our understanding of parasite epigenetics and firmly establish bromodomains as viable targets for antileishmanial drug discovery.

2. Chapter 2: Materials and Methods

2.1. General molecular biology and biochemical techniques

2.1.1. Reagents, media and buffers

Antibiotic selection was used during the cloning and protein expression process with final concentrations of 100 µg/ml ampicillin, 35 µg/ml kanamycin and 35 µg/ml chloramphenicol. Antibiotic stocks at 1000x concentration were prepared and filter sterilised using 0.22 µm filters (Sarstedt) and stored in aliquots at -20 °C.

Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Bio Basic) and dithiothreitol (DTT) (Melford) were prepared as 1 M stock solutions in distilled water, filter sterilised using 0.22 µm filters and stored as 1 ml aliquots at -20 °C. Glucose (Scientific Laboratory Supplies) was produced as a 20% stock in distilled water, sterilised using 0.22 µm filters and stored at 4 °C. The NMR standard, sodium trimethylsilylpropanesulfonate (DSS) (Sigma-Aldrich) was prepared as a 275 mM stock in distilled water and stored at room temperature.

Media used for the culture of *Escherichia coli* (*E. coli*) was Luria-Bertani broth (LB), prepared by dissolving one 25 g LB broth capsule (Melford) per 1 L distilled water. Once dissolved, LB media was sterilised by autoclave at 121 °C for 20 minutes. All buffers are detailed in Appendix D. pH measurements were made using a bench pH meter (Hanna) following prior calibration, with the electrode stored in 3 M potassium chloride solution between measurements.

2.1.2. Bacterial strains

NEB® 5-alpha competent *E. coli* cells (NEB #C2987) (New England Biolabs) were used for transformation with assembled constructs in cloning. *E. coli* Stellar™ competent cells were used for plasmid maintenance and amplification. *E. coli* Rosetta™ (DE3) pLysS competent cells were used for protein expression, and this strain carries chloramphenicol resistance. Genotypes of these bacterial cell lines are given in Table 2.1.

Table 2.1. *E. coli* strain genotypes. Genotypes given for the three strains used throughout DNA manipulation and recombinant protein production.

Strain	Genotype
NEB® 5-alpha	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*
Stellar™	F-, *endA1*, *supE44, thi-1, recA1, relA1, gyrA96, phoA*, *Φ80d lacZΔ M15, Δ (lacZYA - argF)* U169, *Δ (mrr - hsdRMS - mcrBC), ΔmcrA, λ–*
Rosetta™ (DE3) pLysS	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (DE3) pLysSRARE (Cam^R)

2.1.3. Polymerase chain reaction (PCR)

PCR experiments were performed based on the Q5® High-Fidelity DNA Polymerase PCR protocol for a 50 µl reaction (New England Biolabs). PCR mixtures were prepared on ice, containing 10 µl 5x Q5 reaction buffer (New England Biolabs), 1 µl dNTPs (10 mM) (Thermo Scientific), 2.5 µl forward primer (10 µM), 2.5 µl reverse primer (10 µM), 0.5 µl Q5 High-Fidelity DNA Polymerase (New England Biolabs), 1 – 10 µl template DNA and the remaining volume made up by nuclease-free water. Template DNA was added at an appropriate volume to give 0.001 – 2 ng. Primer sequences are listed in Appendix B and were supplied by Eurofins. The PCR thermal cycling conditions are detailed in Table 2.2 and performed using a LifeECO Thermal Cycler (Bioer). The annealing temperature for different primers was calculated using the online NEB T_m calculator tool and extension time was calculated as 20 – 30 seconds per kb.

Table 2.2. Cloning PCR thermal cycling conditions. Annealing temperature and extension time were dependent upon the primers and template DNA being used. Denaturation, annealing and extension steps were repeated for 35 cycles total.

Step	Temperature	Time	
Initial denaturation	98 °C	30 seconds	
Denaturation	98 °C	10 seconds	35 cycles
Annealing	Variable	30 seconds	
Extension	72 °C	Variable	
Final extension	72 °C	2 minutes	
Hold	4 °C	∞	

2.1.4. Agarose gel electrophoresis

DNA samples were analysed by agarose gel electrophoresis. 1% agarose gels were prepared by heating to dissolve 0.5 g agarose powder (Melford) in 50 ml 1x Tris-acetate-EDTA (TAE) buffer, prepared as a 1/50 dilution of 50x TAE buffer (Table D.1) in distilled water, with the addition of 0.5 µl SYBR safe dye (Thermo Scientific). The molten agarose was set in a casting tray with insertion of a 10-well comb. Gels were run by transferring into a horizontal gel electrophoresis system (Bio-Rad) filled with 1x TAE buffer. DNA samples were prepared by addition of 6X blue/ orange loading dye (Promega) to give a final 1X concentration. DNA samples were loaded onto the gel in volumes no greater than 25 µl, alongside a Quick-load® purple 1 kb plus ladder (New England Biolabs), prepared as 2 µl ladder, 2 µl 6X blue/ orange loading dye (Promega) and 8 µl distilled water. The gel was run at 110 V for 40 minutes then imaged under UV using a transilluminator imaging system (Syngene).

2.1.5. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE gels were prepared using 20% resolving gels for single bromodomain recombinant proteins and 12% resolving gels for the larger *Ld*BDF5T tandem recombinant protein. Resolving gels and stacking gels were formulated using the quantities given in Table D.1, and cast within gel casting apparatus (Bio-Rad) using

1.0 mm plates. Resolving gels were cast first with a layer of butanol added on top to produce a level surface on the top of the gel. Once resolving gels had set, butanol was removed and stacking gels cast with the addition of a 10-well comb. Once set, gels were transferred into the tank of a vertical gel electrophoresis system (Bio-Rad) filled with SDS-PAGE running buffer (Table D.1) and the comb removed.

Protein samples were prepared by mixing 5 µl 4x SDS-PAGE sample buffer (Table D.1) with 15 µl protein and boiling at 95 °C for 5 minutes. The 20 µl denatured protein samples were then loaded onto the gels alongside 10 µl SDS-PAGE standard broad range molecular weight marker (Bio-Rad). Gels were run at 200 V for between 1 and 1.5 hours.

To stain the gels, they were first boiled in distilled water and placed on a rocker for 5 minutes, following which the water was replaced with Magic stain dye (60 mg/L Coomassie brilliant blue G-250 dye, 40 mM HCl). Gels were boiled again then placed on the rocker for a further 5 minutes. De-staining was achieved by replacing the dye with distilled water and incubating on the rocker overnight. Gels were subsequently imaged using a gel imaging system (Syngene).

2.1.6. Size exclusion chromatography multi-angle laser light scattering (SEC-MALLS)

Recombinant proteins were analysed by SEC-MALLS, with experiments performed by Dr Andrew Leech in the University of York Technology Facility. Purified recombinant proteins in HEPES buffer (20 mM HEPES, 500 mM NaCl, pH 7.5) were injected in 100 µl volumes onto a Superdex S75 size exclusion column (GE Healthcare), run at a flow rate of 0.5 ml/min using an HPLC system (Shimadzu). Data were collected over 59 minutes using the HPLC system SPD-20A UV detector (Shimadzu) and a HELEOS-II multi-angle light scattering detector and rEX refractive index detector (both Wyatt). Data analysis was performed using Astra V software and protein molecular weight (MW) estimated by the Zimm fit method with degree 1. Graphs were plotted using GraphPad Prism® software.

2.2. DNA cloning

Constructs for the production of the recombinant bromodomains of *Ld*BDF1, 2, 3 and 4 were generated in the plasmid vector pET-YSBLLIC3C (Fogg and Wilkinson, 2008) (further described in Chapter 3; Figure 3.1) which carries a kanamycin resistance gene. For *Ld*BDF1, three truncated versions of the gene were cloned with different length N-termini, named *Ld*BDF1a, 1b and 1c. Full details of the native genes, plasmids and truncations are given in Appendices A and B. Aseptic technique was followed throughout.

2.2.1. Plasmid linearisation

The pET-YSBLLIC3C plasmid was linearised by PCR using primers pET-YSBLLIC3C F and pET-YSBLLIC3C R (Table B.1), with an annealing temperature of 65 °C and an extension time of 2 minutes 40 seconds. The reaction products were treated with DpnI to degrade the original template DNA which is methylated. A reaction mixture containing 50 µl linearised plasmid PCR product, 1 µl DpnI (10 U/µl) and 6 µl 10x Tango buffer (both Thermo Scientific) was incubated at 37 °C for 1 hour then at 80 °C for 20 minutes to deactivate the enzyme. The DNA products were resolved by agarose gel electrophoresis (Section 2.1.4) and fragments of the desired size were extracted using a QIAquick gel extraction kit (Qiagen), according to manufacturer's instructions. This yielded 50 µl DNA at 20 ng/µl, measured using a NanoPhotometer® NP80 (Implen).

2.2.2. Insert preparation

Vectors containing the bromodomain regions of *Ld*BDF1, 2, 3 or 4 in pET-15, pEX-A128 or pUC-SP plasmid backbones were used as template DNA to produce the inserts for cloning with overhangs complementary to the plasmid cloning sites (insert points). *Ld*BDF2 and 3 pET-15 plasmids were gifted by Dr Raymond Hui at the Structural Genomics Consortium, Toronto. Template DNA for the *Ld*BDF1 and *Ld*BDF4 bromodomains in pEX-A128 and pUC-SP backbone plasmids, respectively, were designed by Dr Nathaniel Jones and synthesised externally (Eurofins & DC Biosciences). Full details of the plasmids and cloned regions are given in Appendix B. The template DNA plasmids were diluted in Milli-Q water to concentrations of 0.6 – 1 ng/µl and PCR performed as outlined in Section 2.1.3, using primers listed

in Table B.1 An annealing temperature of 72 °C and extension time of 20 seconds were used. PCR products were analysed by agarose gel electrophoresis, excised from the gel and extracted from gel fragments as previously described, eluting in 40 µl volumes. The concentration of eluted products were measured using the NanoPhotometer® NP80 (Implen).

2.2.3. Assembly and transformation

Prepared bromodomain insert DNA was joined to the linearised pET-YSBLLIC3C plasmid, utilizing an assembly method in which an exonuclease enzyme digests 5 prime ends of the DNA, leaving single-stranded 3 prime overhangs in the plasmid and insert. After annealing, DNA polymerase and dNTPs are used to fill in the single-stranded gaps in the DNA and DNA ligase is used to seal the nicks to produce the fully assembled plasmid. Assembly was performed using an NEBuilder® HiFi DNA assembly cloning kit (New England Biolabs) following the manufacturer's instructions, with reaction mixtures containing 0.046 pmol insert DNA, 0.023 pmol vector (80.7 ng), 10 µl NEBuilder HiFi DNA assembly master mix and Milli-Q water in total sample volumes of 20 µl, incubated in a thermocycler at 50 °C for 15 minutes. The reaction products were then used directly to transform NEB® 5-alpha competent *E. coli* cells (NEB #C2987) (New England Biolabs) (Section 2.1.2) by adding 2 µl of the reaction mix to 50 µl cells and incubating on ice for 30 minutes before heat shock at 42 °C for 30 seconds and a further 2 minute incubation on ice. 950 µl of room temperature super optimal broth with catabolite repression (SOC) media (New England Biolabs) was then added, and the cells incubated at 37 °C with shaking at 250 rpm for 1 hour. 200 µl of cells were then plated onto agar plates containing kanamycin and incubated at 37 °C overnight.

2.2.4. Colony PCR screen

To screen for *E. coli* colonies that contained the desired plasmid, eight colonies were selected from each agar plate and the cells lysed to extract DNA by mixing with 55 µl Milli-Q water and heating in a thermocycler at 95 °C for 5 minutes. Each sample was then centrifuged at 16,000 x g, room temperature for 2 minutes in a benchtop microcentrifuge (Progen) and PCR performed using 10 µl the supernatant with primers complementary to the T7 promoter and terminator regions of the

plasmid, which flank the DNA insert site (T7 F and R primer sequences given in Table B.1). PCR used an annealing temperature of 58 °C and an extension time of 25 seconds, and PCR products were analysed by agarose gel electrophoresis alongside a control sample of the insert DNA used for assembly. Colonies that gave rise to PCR products of the expected sizes were selected for further analysis.

2.2.5. Plasmid preparation, final PCR and sequencing

For each *Ld*BDF clone, three colonies were used to inoculate 5 ml LB with 5 µl kanamycin, incubated at 37 °C, 180 rpm overnight. The plasmids were then extracted and purified using a Monarch® Plasmid Miniprep Kit (New England Bioscience) according to manufacturer's instructions. A second PCR using the T7 primers was then performed on the three isolated plasmids for each clone and PCR products analysed by agarose gel electrophoresis. One plasmid was selected for each bromodomain clone for subsequent use in expression tests, and was sequenced by Eurofins Genomics using the T7 F and T7 R (Eurofins) primers (Table B.1) to confirm there were no mutations in the bromodomain sequences and they were in-frame.

2.3. Recombinant protein production

2.3.1. Preparation of glycerol stocks

*Ld*BDF plasmids were transformed into *E. coli* Stellar™ competent cells and *E. coli* Rosetta™ (DE3) pLysS competent cells (Section 2.1.2) by adding 5 – 100 ng plasmid DNA to 50 µl of cells and incubating on ice for 30 minutes. Cells were then subjected to heat shock at 42 °C for 1 minute and returned to ice for 2 minutes. 250 µl LB media was then added to the cells followed by incubation at 37 °C for 1 hour. 200 µl of cells were plated onto agar plates containing the appropriate antibiotic(s) and incubated at 37 °C overnight. A single bacterial colony was used to inoculate 10 ml LB media containing 10 µl of the appropriate antibiotic(s), incubated at 37 °C with shaking at 180 rpm overnight. 1 ml of the overnight culture was mixed with 1 ml 80% glycerol to produce glycerol stocks, stored at -70 °C.

2.3.2. Small-scale expression tests

Small-scale expression tests were performed in order to establish optimal conditions for the overproduction of recombinant *Ld*BDF proteins in pET-15 and pET-YSBLLIC3C plasmids which carry ampicillin and kanamycin resistance genes, respectively. All proteins contained a cleavable N-terminal polyhistidine-tag (His-tag), and a full list of protein sequences is given in Appendix A.

Overnight starter cultures were prepared by inoculating 10 ml LB media and the appropriate antibiotics with a scraping of the glycerol stock of *E. coli* Rosetta™ cells harbouring the relevant plasmid. Cultures were grown at 37 °C, 180 rpm overnight. Small-scale expression tests were then performed following a standardised procedure, with specific parameters varied to screen for optimal expression conditions. Optical density measurements were made at 600 nm (OD_{600}) using a BioPhotometer spectrophotometer (Eppendorf). In 50 ml conical tubes, 10 ml LB media supplemented with the appropriate antibiotic(s) and 0.2% glucose was inoculated with 200 – 300 µl starter culture and incubated at either 30 or 37 °C until OD_{600} reached 0.6 – 0.8, or at 37 °C until OD_{600} reached 0.4 then at 18 °C until OD_{600} reached 0.6 – 0.8. At this point, the inducer IPTG (0.1 or 1 mM) was added to the cultures and incubation continued at 18, 30 or 37 °C with shaking at 180 – 200 rpm.

1 ml culture samples were collected before induction and at various time-points following induction. Cells were harvested by centrifuging at 16,000 x g, room temperature for 3 minutes and the pellets resuspended in 200 – 300 µl buffer A (Table D.1). Cells were then lysed by three cycles of sonication for 10 seconds using a Soniprep 150 Plus sonicator (MSE) followed by 1 minute on ice. Cell debris was pelleted by centrifugation at 16,000 x g, room temperature for 3 minutes. Samples of 'total' and 'soluble' protein were taken before and after centrifugation, respectively, and analysed alongside the samples taken before induction by SDS-PAGE.

2.3.3. Large-scale protein expression

The most promising expression conditions were scaled up for large-scale production of *Ld*BDF2, 5T, 2(N87F) and 5.1(N90F) proteins in *E. coli* Rosetta™ (DE3)

pLysS competent cells. In some cases, further optimisation of large-scale expression conditions was conducted. Overnight starter cultures were prepared as described above, in either 10, 50 or 60 ml volumes. 10 – 17 ml starter culture was used to inoculate 0.5 – 1 L LB, supplemented with the appropriate antibiotics and 0.2% glucose. Cultures were incubated at 37 °C until the OD_{600} reached 0.6. Production of recombinant protein was induced with the addition of 1 mM IPTG followed by incubation at 37 °C for 2.5 hours for *Ld*BDF2, 30 °C for 3.5 hours or overnight for *Ld*BDF5T, or 30 °C for 3.5 hours for *Ld*BDF2(N87F) and 5.1(N90F). Cells were harvested by centrifugation at 5,000 x g, 4 °C for 30 minutes using a Sorvall Lynx 6000 Superspeed Centrifuge with an F9-6 x 1000 fixed angle rotor (Thermo Scientific) and stored at -70 °C pending protein purification.

2.3.4. Protein purification

Cell pellets were resuspended at 5 ml per 1 g cell pellet in protein purification buffer A (Table D.1), supplemented with 1 cOmplete™ mini EDTA-free protease inhibitor cocktail tablet (Roche), $MgCl_2$ (5 mM) and DNase I (5 µg/ml). Lysis was carried out by sonication using 30 second bursts followed by 30 second pauses on ice for 15 cycles, and the cell debris pelleted by centrifugation at 18,000 x g, 4 °C for 40 minutes using a Sorvall Lynx 6000 Superspeed Centrifuge with a SS-34 fixed angle rotor (Thermo Scientific).

The soluble lysate was then resolved by immobilised metal affinity chromatography (IMAC) and size-exclusion chromatography (SEC), using an ÄKTA pure protein purification system (GE Healthcare), monitoring absorbance of the eluate at 280 nm (A_{280}). Protein purification buffers A, B and C are listed in Table D.1, where DTT was added following column equilibration.

The supernatant was applied to a 5 ml HisTrap™ FF column (GE Healthcare) pre-equilibrated in buffer A, subsequently developed with a linear gradient of buffer A to buffer B over 10 column volumes (CV) at a flow rate of 2 ml/min. Fractions were analysed by SDS-PAGE and those containing protein of the expected molecular weight were pooled and concentrated using a Vivapsin 3K MWCO spin column (Sartorius) according to manufacturer's instructions.

His-tags were cleaved by addition of recombinant Tobacco etch virus (TEV) protease at 1 mg per 50 mg protein. The sample was loaded into 3.5 kDa MWCO dialysis tubing (Spectra/Por) and dialysed into imidazole-free protein purification buffer C at 4 °C overnight. The proteins were then concentrated as before and applied a second pre-equilibrated HisTrap™ column using a 5 ml loop, washing with 15 ml buffer A. The column was then washed with 5 CV buffer A, developed with a linear gradient of buffer A to buffer B over 20 CV at 2 ml/min. Fractions were again analysed by SDS-PAGE, pooled and concentrated.

Proteins were further purified using a HiLoad 16/600 Superdex 75 pg size exclusion chromatography column (GE Healthcare), pre-equilibrated with buffer C. The concentrated protein sample was applied to the column in volumes no greater than 2.5 ml using a 5 ml loop, emptied with 15 ml buffer C, and the column washed with 1 CV buffer C at a flow rate of 1 ml/min to elute the protein. Fractions were analysed by SDS-PAGE, and those containing the highest levels of protein pooled and concentrated as before. Protein concentrations were calculated by measuring A_{280} using a NanoPhotometer® NP80 (Implen), with relevant the molar extinction coefficient at 280 nm (ε_{280}) determined by the Protparam online tool within the Expasy Swiss bioinformatics resource portal.

2.4. Histone peptide design and synthesis

Peptides were designed based on the *L. donovani* histone H2A.Z, H2B and H4 sequences (Appendix C) for use in biophysical assays to investigate the interactions of the *Ld*BDF bromodomains with the histones. A variety of peptides were designed, including acetylated versions and unmodified control peptides to confirm binding specificity, as well as fluorescently-labelled peptides to be used in a fluorescence polarisation assay (full list of peptides provided in Chapter 4, Figure 4.1). The fluorophore used to label peptides was 5-carboxyfluorescein (5-FAM) which has excitation and emission of around 492 and 518 nm, respectively. Peptides were synthesised by Cambridge Research Biochemicals at >90% purity with analysis by HPLC and mass determination by MALDI coupled to time-of-flight (MALDI-TOF) mass spectrometry.

2.5. Ligand-binding assays

2.5.1. Ligand stocks

Peptides were supplied as 5 mg or 10 mg quantities (Cambridge Research Biochemicals) and dissolved to 10, 50 or 100 mM in dimethyl sulfoxide (DMSO). Human bromodomain inhibitor compounds bromosporine, I-BET151, SGC-CBP30, BI 2536 and JQ1 were purchased commercially. GSK8814 was supplied by the Structural Genomics Consortium under an Open Science Trust Agreement. All other human bromodomain inhibitors were gifted by Raquel Gabarro-Carrion and Felix Calderon from GSK. Compound stock solutions were prepared to concentrations of 10 – 100 mM in DMSO, or deuterated DMSO to allow them to also be used in the NMR assay. Acetyllysine mimic fragments were synthesised and gifted by Jennifer Carter from the Conway group (University of Oxford), supplied as 10 – 100 mM stocks in DMSO or deuterated DMSO. Hit compounds from high-throughput screens were gifted by GSK as 10 – 100 mM stocks in DMSO, or purchased commercially in the case of certain sulfonyl fluoride compounds. All peptide and compound stocks were stored at -70 °C.

2.5.2. Thermal shift assay (TSA)

TSA (also known as differential scanning fluorimetry) (Figure 2.1) was used to investigate the interactions of recombinant *Ld*BDF2 and 5 bromodomain proteins with different compounds and peptides. Initial optimisation experiments were performed to explore different protein (1 – 5 μM) and dye (1x, 2x and 3x) concentrations. The chosen optimised conditions (Table E.1) were used in subsequent TSA experiments unless stated otherwise. SYPRO orange (Merck) was used as the fluorescent dye in the TSA assay, and HEPES buffer (20 mM HEPES, 500 mM NaCl, pH 7.5) was used for dilutions. Control experiments were performed to confirm that compounds and peptides did not give fluorescent signals themselves or interfere with the assay.

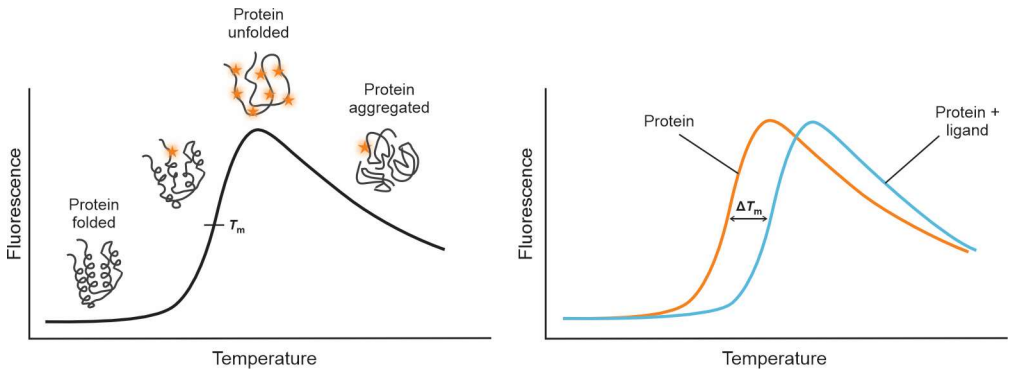


Figure 2.1. TSA principle. Left: Increasing temperature causes the protein to unfold so that a fluorescent dye (orange stars) binds non-specifically to the exposed hydrophobic surfaces, causing a fluorescent signal from which the protein melting temperature (T_m) can be determined. **Right**: ligand binding stabilises the protein, causing a positive shift in the thermal stability from which the thermal shift (ΔT_m) can be calculated.

Sample volumes of 25 µl were prepared containing the protein and dye plus the ligand at a final concentration in the range 10 – 400 µM. Reference samples contained DMSO/ deuterated DMSO at the same final concentration as in the ligand samples. Control samples were also included of the protein alone and dye alone. Samples were prepared in 96-well qPCR plates (Agilent) using six replicates of each test sample and three replicates of the control samples. Plates were sealed and centrifuged at 500 x g, 4 °C for 1 minute using a Sigma 3-16KL benchtop centrifuge. TSA experiments were then carried out using a Stratagene Mx3005P real-time PCR instrument (Agilent) which increased the temperature from 25 – 95 °C at 30 seconds per degree, taking fluorescence readings after each 30 second increment. Data were imported into an online JTSA tool for analysis, with a five parameter sigmoid equation curve fitting applied to the data and melting points calculated as the midpoints (Schulz, Landström and Hubbard, 2013; Bond, 2017). Any anomalous results within the six replicates were excluded from the analysis. Data were also analysed in Microsoft Excel and GraphPad Prism® and ΔT_m calculated as the difference between the T_m of the protein-ligand samples and the reference DMSO/ deuterated DMSO samples. Where appropriate, statistical analysis was performed using an unpaired t-test, with P values displayed on graphs as * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

2.5.3. Fluorescence polarisation (FP) assay

The FP assay utilises a fluorescent probe which gives a low FP signal when free in solution but a high FP signal when bound to a protein. Such a probe can be used in competition assays with unlabelled competitor compounds; where a drop in FP is observed with increasing competitor concentration, this indicates binding is occurring and from the resulting data binding affinity can be calculated (Figure 2.2).

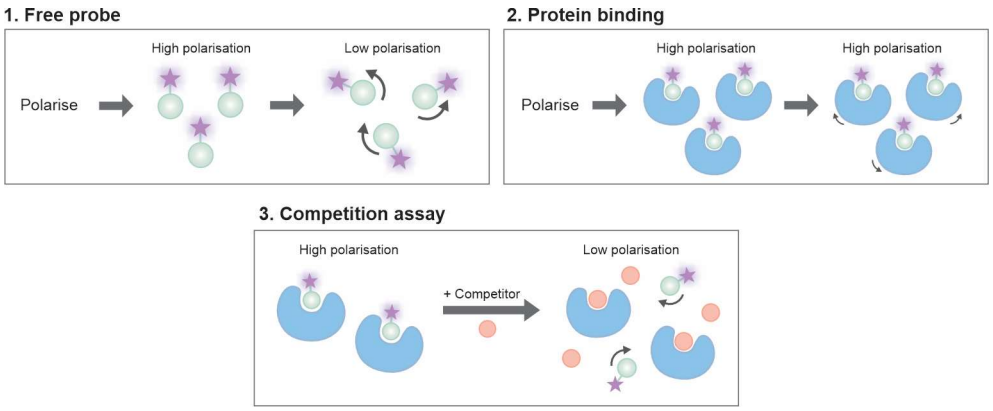


Figure 2.2. FP principle. 1 Unbound probes rotate rapidly in solution so lose polarisation in the time between excitation and emission, resulting in a low polarisation signal. **2** The probe binds to a larger protein molecule (blue) which rotates less so the probe retains its high polarisation signal. **3** Competition with an unlabelled ligand (orange) causes displacement of the probe from the protein, once again resulting in low polarisation.

The FP assay was performed to measure binding of different 5-FAM-labelled histone peptide probes (Figure 4.1) to the recombinant *Ld*BDF2 and *Ld*BDF5 bromodomain proteins. All FP experiments were set up in 384-well black flat bottom plates (Corning) using sample volumes of 20 µl, diluting in FP buffer (20 mM HEPES, 500 mM NaCl, 1 mg/ml BSA added fresh, pH 7.5). BSA was included in the buffer to block non-specific probe binding. All fluorescence intensity and polarisation readings were taken using a BMG LABTECH CLARIOstar® microplate reader using appropriate excitation and emission settings for the 5-FAM fluorophore (excitation 482-16 nm; emission 530-40 nm) following gain and focus

adjustment. Data were processed using BMG LABTECH Mars software with subsequent analysis in Microsoft Excel and GraphPad Prism®.

Firstly, probe optimisation experiments were performed by preparing triplicate samples of 0 – 1000 nM probes alongside triplicate control samples of buffer alone for blank correction. Fluorescence intensity and polarisation readings were taken as described above. Measurements were blank-corrected and FP calculated using the equation:

$$FP = \frac{I_{II} - I_{\perp}}{I_{II} + I_{\perp}}$$

In this equation, I_{II} is intensity of emitted light polarised parallel to the excitation light, and $I_{\perp}$ is intensity of emitted light polarised perpendicular to the excitation light. Average, blank-corrected fluorescence intensity and polarisation values were plotted against probe concentration to determine an appropriate concentration to use in subsequent experiments (Figure E.1, Figure 4.5).

Secondly, binding of the probes to the recombinant *Ld*BDF bromodomain proteins was assayed. Samples containing the probe (800 nM) and proteins covering a range of 0 – 300 µM were prepared in triplicate, alongside control samples of probe alone and buffer alone also in triplicate. Additionally, samples with and without BSA were included to confirm there was no interference from the probe binding to BSA. Samples were mixed in the wells, and the plate incubated in the dark for 30 minutes at room temperature before taking FP readings. Average, blank corrected FP values were transformed by subtracting the mP of the free probe, with ΔmP plotted against protein concentration and for single bromodomain proteins (*Ld*BDF2, 5.1 and 5.2), fitted with a one site-specific binding non-linear regression model:

$$Y = \frac{B_{max} \times X}{(K_D + X)}$$

In this equation, X and Y are the protein concentration and FP, respectively, B_{max} is the maximum specific binding. For the dual bromodomain protein, the two sites-specific binding non-linear regression model was instead used:

$$Site1 = \frac{B_{max}Hi \times X}{(K_D Hi + X)}$$

$$Site2 = \frac{B_{max}Lo \times X}{(K_D Lo + X)}$$

$$Y = Site1 + Site2$$

Here, X and Y are the protein concentration and FP, respectively, $\mathrm{B_{max}Hi}$ and $\mathrm{B_{max}Lo}$ are maximum specific bindings to the two sites, and $\mathrm{K_DHi}$ and $\mathrm{K_DLo}$ are the equilibrium dissociation constants.

The H2B-7mer-Ac-F peptide (Figure 4.1) showed the strongest binding to *Ld*BDF5.1 and 5.2 and was therefore used as a probe in subsequent competition assays to screen inhibitor compounds. 2.5 µM and 10 µM were chosen as the protein concentrations of *Ld*BDF5.1 and 5.2, respectively, to be used in the competition assays, alongside the probe at 800 nM. Fixed concentrations of the protein and probe were titrated against increasing concentrations of the competitor ligand up to 100 µM, obtained by serial dilution in DMSO/ deuterated DMSO, with a final concentration of 1% DMSO/ deuterated DMSO. Protein and compounds were incubated together for 30 minutes at room temperature before addition of the probe, then FP readings were taken as before. For covalent inhibitors, a pre-incubation period of 1 hour was used before addition of the probe, and readings were taken at specific time-points up to 20 hours. Controls included buffer alone, probe alone, protein alone and ligands alone. Experiments were also performed in which probes were competed against non-labelled peptides, which followed a similar methodology with 30 minutes pre-incubation, 800 nM probe, but with different protein concentrations of 2.5 – 150 µM, and titrating peptides at various concentrations up to a maximum of 500 µM.

In all competition assay experiments, average, blank-corrected FP values were plotted against compound concentration and IC_{50} values calculated by fitting an [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression model:

$$Y = Bottom + \frac{(Top - Bottom)}{(1 + \left(\frac{IC50}{X}\right)^{Hillslope})}$$

X and Y are the ligand concentration and FP, respectively, Top and Bottom are plateaus (mP), $\mathrm{IC_{50}}$ is the concentration that gives a response half way between the top and bottom of the curve (µM) and $\mathrm{HillSlope}$ is the steepness of the curve.

2.5.4. Nuclear magnetic resonance (NMR)

Three different ligand-observed proton NMR experiments were used to investigate the ligand-binding interactions of *Ld*BDF5.1 and 5.2; water ligand-observed via gradient spectroscopy (waterLOGSY), saturation transfer difference (STD), and Carr-Purcell-Meiboom-Gill (CPMG). Buffer exchange was first performed to transfer proteins from HEPES buffer to a sodium phosphate (NaPi) NMR buffer (20 mM NaPi, 100 mM NaCl, pH 7.5) using Zeba 0.5 ml 7K MWCO columns (Thermo Scientific) following the manufacturer's instructions. DSS was used as the NMR reference standard to provide the standard peak, set to chemical shift 0 ppm.

NMR samples (550 µl) were prepared in 5 mm NMR tubes (Wilmad) including the appropriate controls, using concentrations of 2.5 – 20 µM recombinant protein and 0.2 – 0.6 mM compound (prepared from 10 – 100 mM stock solutions in deuterated DMSO). Each sample also contained 13 – 17.5% D_2O (Sigma-Aldrich), 80 µM DSS (Sigma-Aldrich), 20 mM NaPi and 100 mM NaCl. Spectra were recorded using a 700 MHz Bruker Avance Neo spectrometer, equipped with a cryoprobe at 298 K, with 16 – 24 scans. In cases where the cryoprobe was not equipped, the number of scans was increased 4-fold. 1D proton spectra were recorded with water suppression, and additional acquisition parameters are given in Appendix E. Spectra were analysed using TopSpin NMR data analysis software (Bruker) and proton chemical shifts predicted using the NMRDB NMR spectral predictor tool (Banfi and Patiny, 2008; Castillo, Patiny and Wist, 2011).

2.6. Protein crystallography

Bromodomain proteins were concentrated using Vivaspin® 3 kDa MWCO centrifugal concentrators (Sartorius) according to manufacturer's instructions. DTT (1 mM) was added to the protein along with the relevant compound for co-crystallisation, all solutions prepared in HEPES buffer (20 mM HEPES, 500 mM NaCl, pH 7.5). The solution was incubated for up to 24 hours at 4 °C, shaking gently. Samples were then centrifuged at 16,000 x g, room temperature for 2 – 10 minutes and crystal screens set up using commercial screen solutions pre-dispensed into reservoirs of 96-well MRC crystallisation trays, using the sitting-drop vapour diffusion method with protein and reservoir solution drops dispensed by a Mosquito robot (TTPlabtech). Plates were sealed and either stored at 19 °C and checked at regular intervals or loaded into a Minstrel™ crystal monitoring and imaging robot (Rigaku) with images taken at regular intervals, including under UV, visualised using XtalTrak software (Rigaku). Optimisation of successful crystallisation conditions was performed in 48-well plates also using the sitting-drop method, with relevant buffer solutions prepared and plates set up by hand using 1 – 4 µl drops and 100 – 200 µl reservoir solution. Optimisation plates were stored at 19 °C and checked regularly for crystal growth. Appendix F outlines all the crystallisation conditions trialled for both screens and optimisation. When setting up co-crystallisation screens for *Ld*BDF5.2 with the covalent inhibitor, SO2FC, the protein (1.78 mg/ml) and ligand (0.2 mM) were incubated together with DTT (1 mM) for 24 hours at 4 °C, shaking gently. The protein was then concentrated to 20 or 22 mg/ml using Vivaspin® 3 kDa MWCO centrifugal concentrators (Sartorius) according to manufacturer's instructions. The sample was then centrifuged and crystal screens (Table F.2) set up as described above.

Where crystals were large enough to harvest, fishing was carried out using a nylon loop to transfer crystals from the tray into vials of liquid nitrogen. Where necessary, crystals were transferred into a cryoprotectant (25% ethylene glycol, 75% mother liquor) between the drop and storage vial. Isolated crystals were tested by X-ray diffraction either in-house or at the Diamond light source (Harwell Science & Innovation Campus, Didcot) by Dr Johan Turkenburg and Sam Hart of the York Structural Biology Laboratory.

2.7. *Leishmania* promastigote cell viability assays

2.7.1. Promastigote cell culture

L. mexicana (MNYC/BZ/62/M379) and *L. donovani* LV9 (MHOM/ET/67/HU3) promastigotes were grown in hemoflagellate-modified minimum essential medium (HOMEM) (Gibco) supplemented with 10% (v/v) heat-inactivated foetal calf serum (hi-FCS) (Gibco) and 1% (v/v) penicillin/ streptomycin solution (Sigma-Aldrich) at 25 °C. Cell cultures were passaged weekly into fresh medium using a dilution of between 1/40 and 1/100 culture in fresh medium.

2.7.2. Cell density determination

To determine the parasite cell density, cells were fixed by making a 1/10 dilution in 2% (v/v) formaldehyde and manually counted using a Neubauer haemocytometer.

2.7.3. Cell viability single point-screens

Compounds were screened for activity against *L. mexicana* and *L. donovani* promastigotes at 10 µM. 20 µM compound solutions were prepared by dissolving 10 mM stocks in medium (HOMEM with 10% hi-FCS and 1% penicillin/ streptomycin solution). Parasite cultures were grown to mid-log (exponential) phase (cells counted as described above) and dilutions prepared to 5 x 10^3 cells/ml in medium. 100 µl cells were then seeded into triplicate wells of 96-well plates at 500 cells per well, along with 100 µl test compounds to give a final concentration of 10 µM. Also included were a miltefosine positive control, and controls of DMSO, parasites-only and media-only. Empty wells were filled with 200 µl phosphate-buffered saline (PBS). Parasites were exposed to the compounds for 5 days at 25 °C, after which time 40 µl resazurin (Sigma-Aldrich) was added to each well (final concentration 80 µM) and plates incubated for a further 8 hours at 25 °C. Resazurin is reduced by cells to resorufin which emits fluorescence at 590 nm, therefore fluorescence can be used to measure cell viability and cytotoxicity of the compounds towards the parasite. Fluorescent readings were taken using a BMG LABTECH CLARIOstar® microplate reader (excitation 545-20 nm; emission 600-40 nm) and data processed using BMG LABTECH Mars software with subsequent analysis in GraphPad Prism®. Average, blank-corrected fluorescence measurements were normalised to give values as % growth of DMSO control. Compounds were tested in two independent

biological replicate experiments for each parasite species, using the same experimental setup. Control experiments were also performed to confirm the compounds did not interact with the resazurin to give a fluorescent signal in the absence of parasites.

2.7.4. Cell viability dose-response assays

Potent compounds were tested against *L. mexicana* and *L. donovani* promastigotes in dose-response assays. 60 µM compound solutions were prepared in medium (HOMEM with 10% hi-FCS and 1% penicillin/ streptomycin solution) and parasite cell dilutions were prepared as described above. 100 µl cells were seeded into triplicate wells of 96-well plates to give 500 cells per well, along with 100 µl test compounds serially diluted in medium to obtain final concentrations of 0 – 30 µM. A miltefosine positive control was again included, along with DMSO, parasites-only and media-only controls. Empty wells were filled with 200 µl PBS and plates incubated for 5 days at 25 °C, then 40 µl resazurin (Sigma-Aldrich) added to each well (final concentration 80 µM) and plates incubated for 8 hours at 25 °C, after which time fluorescence was measured as before. Data were again processed using BMG LABTECH Mars software and analysed in GraphPad Prism®. Average, blank-corrected fluorescence measurements (0.23 – 30 µM compound) were normalised to give values as % cell viability, and data fitted to an [Inhibitor] vs. normalized response -- variable slope dose-response curve to calculate EC_{50} using the equation:

$$Y = \frac{100}{1 + (\frac{EC_{50}}{X})^{HillSlope}}$$

Here, X and Y are compound concentration and fluorescence intensity, respectively, EC_{50} is the concentration that gives a response half way between the top and bottom of the curve, and $\mathrm{HillSlope}$ is the steepness of the curve. Compounds were tested in two independent biological replicates for each parasite species, using the same experimental setup.

3. Chapter 3: Production of Recombinant *Ld*BDF Proteins

3.1. Introduction

Leishmania have five conserved canonical bromodomain proteins; BDF1, 2, 3, 4 and 5, which are essential for promastigote survival (Jones *et al.*, 2022). Bromodomains have been extensively investigated and have shown great potential as drug targets for a diverse range of diseases, with several bromodomain inhibitors progressing through to clinical trials as cancer treatments (Theodoulou, Tomkinson, *et al.*, 2016; Cochran, Conery and Sims, 2019; Zaware and Zhou, 2019; Kougnassoukou Tchara, Filippakopoulos and Lambert, 2020). Bromodomains may also present a propitious new target for the development of antileishmanials, which are greatly needed in light of the ever growing limitations of current treatments (Burza, Croft and Boelaert, 2018).

The five essential *Leishmania* BDF proteins each contain a single bromodomain, with the exception of BDF5 which has two, termed BDF5.1 and BDF5.2. BDF5 is essential for both promastigotes and amastigotes, and its deletion causes a downregulation of RNA polymerase II transcription. Additionally, the protein is enriched at transcriptional start regions, with these observations together indicating that BDF5 may have an key role in gene transcription in *Leishmania* (Jones *et al.*, 2022). Structures have been solved for *Ld*BDF2, 3, 5.1 and 5.2 in complex with different human bromodomain inhibitors (PDB codes 5C4Q, 5FEA, 5TCM, 6BYA, 5TCK), although further biophysical characterisation and quantification of these binding interactions have not been reported.

To investigate the *Leishmania* bromodomain proteins, it was first necessary to produce recombinant proteins that could be used in biophysical assays such as thermal shift assays, fluorescence polarisation assays and NMR. These protein constructs contained the bromodomain regions of the *L. donovani* BDFs as well as a tandem bromodomain-containing protein *Ld*BDF5T, and mutant versions of *Ld*BDF2, 5.1 and 5.2 (full details are listed in Appendix A). The *Ld*BDF5T construct comprises the region of *Ld*BDF5 containing the two tandem bromodomains, so comparing this to the single bromodomains (*Ld*BDF5.1 and 5.2) in biophysical assays could reveal how the bromodomains may act both independently and in

relation to one another. *Ld*BDF2(N87F), *Ld*BDF5.1(N90F) and *Ld*BDF5.2(N257F) are mutant versions of these bromodomains in which the conserved asparagine has been mutated to a phenylalanine. The presence of this bulkier side chain within the binding pocket should disrupt ligand binding in this site both through steric occlusion and the loss of any interactions with the polar amide of asparagine. This mutation can therefore be used to inactivate the bromodomain and characterise bromodomain ligands, as has been used to this effect previously (Philpott *et al.*, 2014; Hammitzsch *et al.*, 2015; Picaud *et al.*, 2015).

This chapter details the cloning, expression and purification of these recombinant *Ld*BDF proteins, with the aim of producing soluble protein that could subsequently be used in biophysical assays to investigate their binding interactions with histone peptides and inhibitor compounds.

3.2. Cloning *Ld*BDF1, 2, 3 and 4 bromodomains

To achieve a comprehensive understanding the bromodomains in *Leishmania*, one aim was to generate recombinant protein for the bromodomains of all five essential *Ld*BDF proteins. Expression vectors for *Ld*BDF2 and 5 bromodomains were already available and had been used in prior work to express soluble protein. However, this was not the case for *Ld*BDF1, 3 and 4 so truncated genes encoding these bromodomains were subcloned into a new vector, alongside *Ld*BDF2 to investigate whether expression conditions may be improved in a new plasmid.

The chosen plasmid vector for cloning was pET-YSBLLIC3C (Fogg and Wilkinson, 2008) (Figure 3.1). This plasmid allows for IPTG-induced protein expression of an inserted gene under the control of a T7 promoter. DE3 *E. coli* strains carry a chromosomal copy of a T7 RNA polymerase gene under control of a *lac* promoter, and addition of IPTG removes the LacI repressor from the *lac* operator in the chromosome, allowing T7 RNA polymerase expression, with simultaneous removal of LacI from the *lac* operator in the plasmid allowing transcription of the inserted gene by the T7 RNA polymerase. The pET-YSBLLIC3C plasmid also harbours a hexahistidine sequence which allows a cleavable His-tag to be attached to the N-

terminus of the protein. In total, six clones were to be produced, one for each bromodomain, with the exception of *Ld*BDF1 for which three different cassettes were cloned with three different extensions of the N-terminus (*Ld*BDF1a, 1b, 1c).

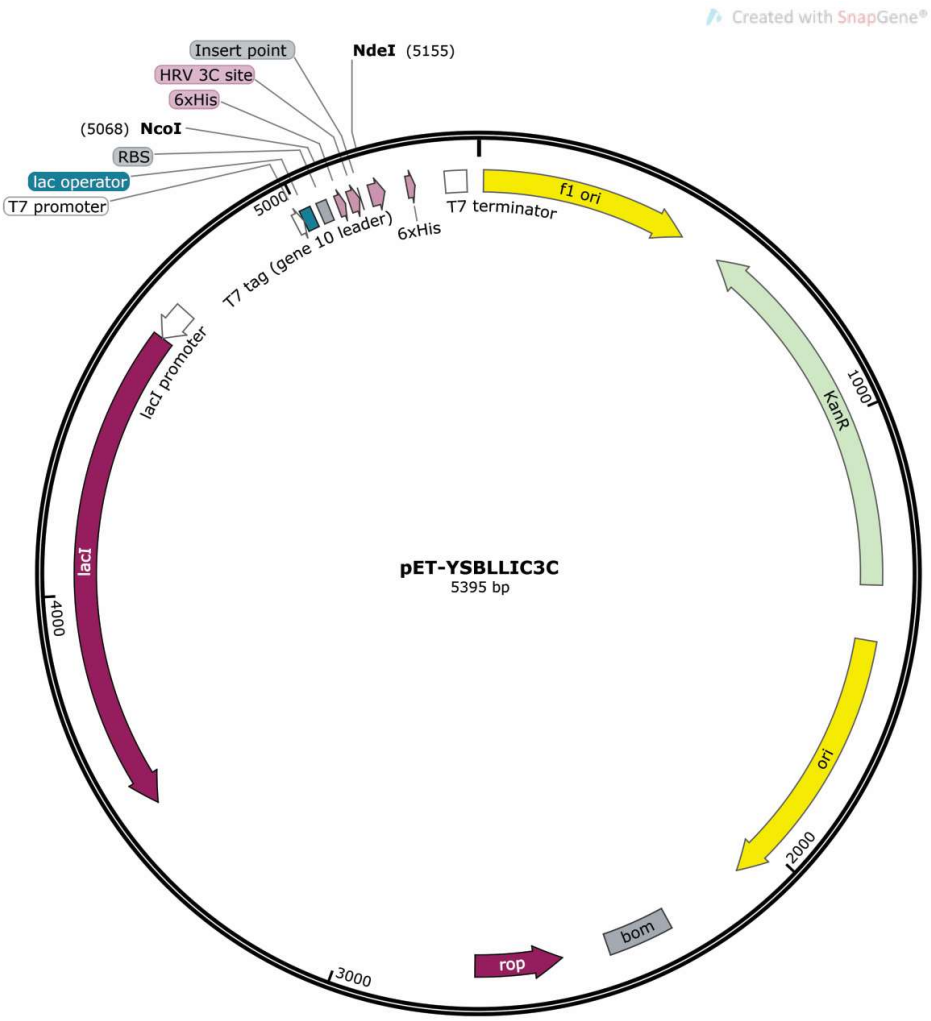


Figure 3.1. pET-YSBLLIC3C vector plasmid map. Notable features for cloning and protein expression include the origin of replication site (ori), kanamycin resistance gene (KanR), T7 promoter and T7 terminator, lac repressor (lacI), lacI promoter, lac operator, rop protein (rop), ribosome binding site (RBS), His-tag (6xHis), HRV 3C protease cleavage site (LEVLFQ/GP), and the insert point of the bromodomain genes. Map was created and annotated using SnapGene® software.

Template DNA for *Ld*BDF1 and *Ld*BDF4 was provided in pEX-A128 and pUC-SP backbone plasmids, respectively, designed by Dr Nathaniel Jones and synthesised externally (Eurofins & DC Biosciences). *Ld*BDF2 and *Ld*BDF3 pET-15 template plasmids were gifted by Dr Raymond Hui at the Structural Genomics Consortium, Toronto (full details in Table B.2). Using these *Ld*BDF bromodomain-encoding plasmids as templates, the new bromodomain gene fragments were amplified by PCR using specific primers and assembled into the linearised pET-YSBLLIC3C vector. *E. coli* cells were then transformed using the assembled constructs and grown on agar plates containing kanamycin to detect colonies that had successfully taken up the plasmid. Kanamycin-resistant transformants, indicating promising candidates, were used for isolation of plasmid DNA which was analysed by PCR using primers complementary to the vector DNA flanking the insert site (T7 F & T7 R, Table B.1) to amplify the insert regions.

Gel electrophoresis confirmed presence of insert DNA of the expected length for the bromodomain fragments. For each bromodomain, three colonies containing plasmids supporting expression of the bromodomain gene were grown up and the plasmid DNA extracted and purified by miniprep. A final PCR was performed using the same T7 F & T7 R primers to re-confirm presence of bromodomain insert DNA (Figure 3.2).

For each *Ld*BDF clone, one verified plasmid construct was sequenced to establish whether the bromodomain genes were in-frame and identify any mutations in the sequences. *Ld*BDF2, 3 and 4 sequences were all as expected, however for *Ld*BDF1, sequencing was inconclusive for *Ld*BDF1b, and showed a frameshift mutation in the bromodomain gene for *Ld*BDF1c. In comparison, the *Ld*BDF1a clone sequence was as predicted, so this construct was taken forward for expression tests alongside the *Ld*BDF2, 3 and 4 clones.

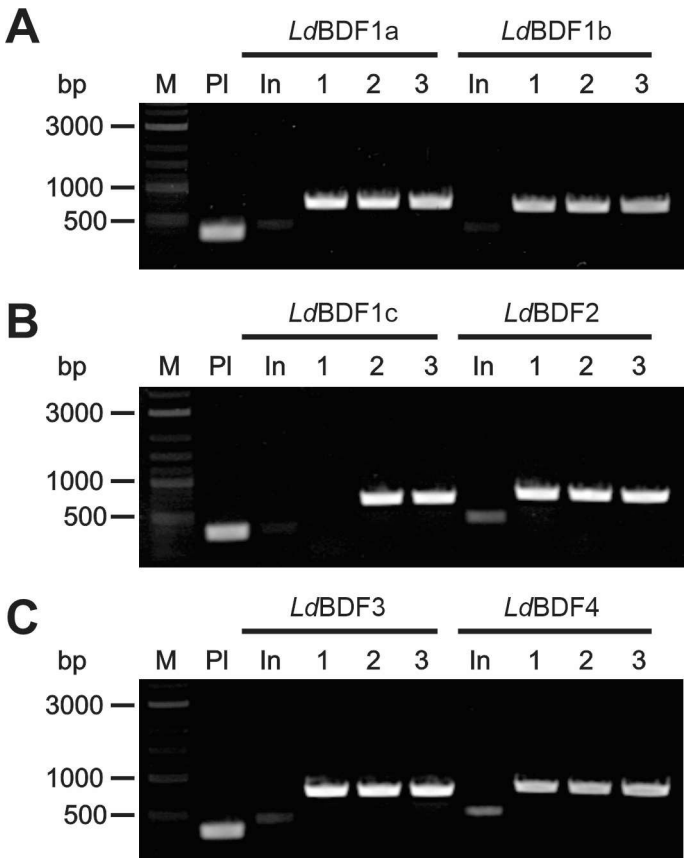


Figure 3.2. *Ld*BDF1, 2, 3 and 4 bromodomain DNA in newly cloned plasmids. Insert sites amplified by PCR for three miniprepped plasmids (**1**, **2**, **3**) for each bromodomain and resolved using 1% agarose gels alongside the original pET-YSBLLIC3C plasmid (**Pl**), and bromodomain inserts (**In**) used for assembly. Expected sizes of the PCR products are 718, 688, 658, 721, 754 and 817 bp for *Ld*BDF1a, 1b, 1c, 2, 3 and 4, respectively. **M** = DNA ladder. One *Ld*BDF1c construct was unsuccessful likely due to an error during plasmid preparation.

3.3. Establishing conditions for the expression of recombinant *Ld*BDF proteins

Of the five bromodomain proteins, BDF5 has been the best characterised (Jones *et al.*, 2022), and optimal expression conditions had previously been established for *Ld*BDF5.1 and *Ld*BDF5.2 using existing plasmids (pGL2776 pET-15 and pGL2775 pET-15) gifted by Dr Raymond Hui at the Structural Genomics Consortium, Toronto. These two recombinant proteins were produced by Juliet Borgia in the York Structural Biology Laboratory and Rebecca Preece in the University of York Technology Facility by expression in *E. coli* and purification by IMAC and SEC.

New small-scale expression trials were performed using the newly cloned bromodomain constructs for *Ld*BDF1, 2, 3 and 4 alongside pET-15 plasmids vectors encoding the bromodomains of *Ld*BDF2, *Ld*BDF5T, *Ld*BDF2(N87F), *Ld*BDF5.1(N90F) and *Ld*BDF5.2(N257F). *Ld*BDF2 and *Ld*BDF5T bromodomains in pET-15 vectors were gifted by the Dr Raymond Hui at the Structural Genomics Consortium, Toronto and *Ld*BDF2(N87F), *Ld*BDF5.1(N90F) and *Ld*BDF5.2(N257F) mutants were produced using site-directed mutagenesis by Dr Nathaniel Jones at the University of York using the pET-15 template plasmids. Expression tests were conducted for all nine plasmids using 10 ml LB and the relevant antibiotics, inoculated using glycerol stocks of transformed *E. coli* Rosetta™ (DE3) pLysS competent cells harbouring the relevant plasmid. Expression conditions such as IPTG concentration, incubation temperature and incubation time were varied and samples were taken before IPTG induction and of both total (insoluble and soluble) and soluble protein, analysed by SDS-PAGE to identify conditions yielding soluble protein.

For the *Ld*BDF2 bromodomain, expression tests were first carried out using a pET-15 plasmid as the backbone of the construct which had been used for protein expression in prior work. Initially, expression was performed following this established methodology, however, *Ld*BDF2 did not express well under these conditions once scaled up, so new expression tests were conducted. Incubation was carried out at a continuous 37 °C, with induction once OD_{600} reached 0.6 using 1 mM IPTG and expression over 2.5 hours. These new conditions generated high levels of soluble protein, and were therefore used to scale up expression of *Ld*BDF2.

For the new *Ld*BDF1, 2, 3 and 4 clones, the transformed *E. coli* Rosetta™ cultures were incubated at 37 °C until OD_{600} reached 0.6 then the inducer IPTG (0.1 or 1 mM) added and further incubation at 30 °C. Expression was successful for *Ld*BDF2 and *Ld*BDF3, with soluble protein produced in almost all expression conditions and the highest levels correlating with 1 mM IPTG and 3.5 hours incubation (Figure 3.3). No soluble protein was observed for *Ld*BDF1 and 4, though bands were seen for *Ld*BDF1 insoluble fractions particularly using 3.5 hour incubation, so it remains possible that further optimisation could be explored to generate soluble protein in the future. For example, a HEPES lysis buffer was used here but others such as Tris or phosphate could also be tested.

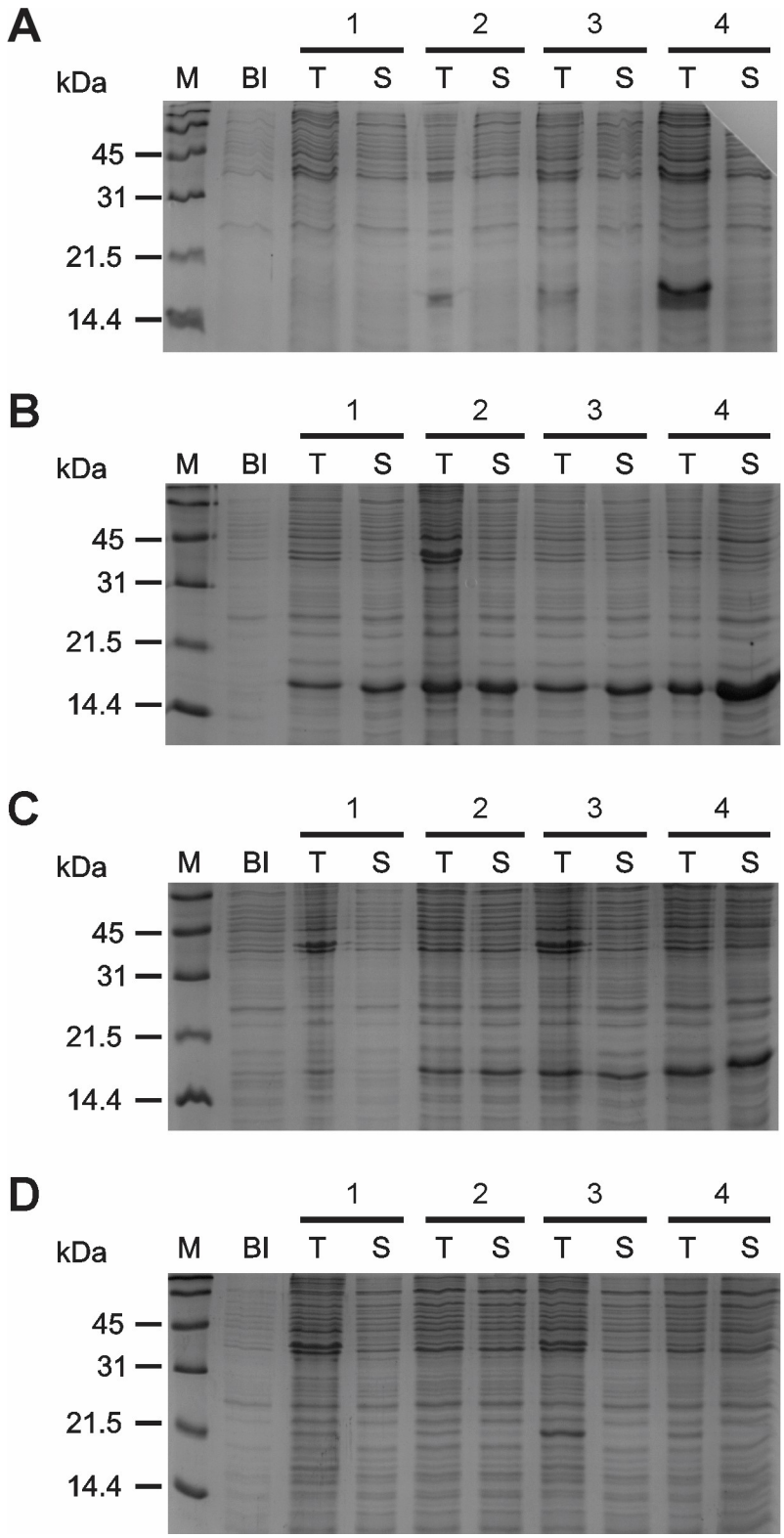


Figure 3.3. Expression tests for *Ld*BDF1, 2, 3, 4 clones in pET-YSBLLIC3C. (A) *Ld*BDF1, **(B)** *Ld*BDF2, **(C)** *Ld*BDF3 and **(D)** *Ld*BDF4 expression tests analysed by 20% SDS-PAGE where **1** = 0.1 mM IPTG, 1 hr 45 min expression; **2** = 0.1 mM IPTG, 3 hr 30 min expression; **3** = 1 mM IPTG, 1 hr 45 min expression; **4** = 1 mM IPTG, 3 hr 30 min expression. **BI** = before induction, **T** = total protein, **S** = soluble protein. Expected protein sizes are 17.3, 16.8, 18.1 and 19.7 kDa for *Ld*BDF1, 2, 3 and 4, respectively. M = protein marker.

For the *Ld*BDF5T tandem bromodomain construct, good levels of expression and solubility were observed in small-scale tests (Figure 3.4A), however when scaling up protein expression, analysis found soluble protein levels to be lower than expected (Figure 3.4B). Therefore, additional optimisation was performed using the large-scale setup, comparing different starter culture inoculation volumes, temperatures, incubation times and inducing at different OD_{600} values. The highest levels of soluble protein were obtained by inoculating 500 ml LB with 10 ml starter culture, incubating at 37 °C until OD_{600} exceeded 0.7 then adding 1 mM IPTG and incubating at 30 °C for 3.5 hours or overnight (Figure 3.4C). Pellets from these large-scale expressions were harvested for subsequent purification.

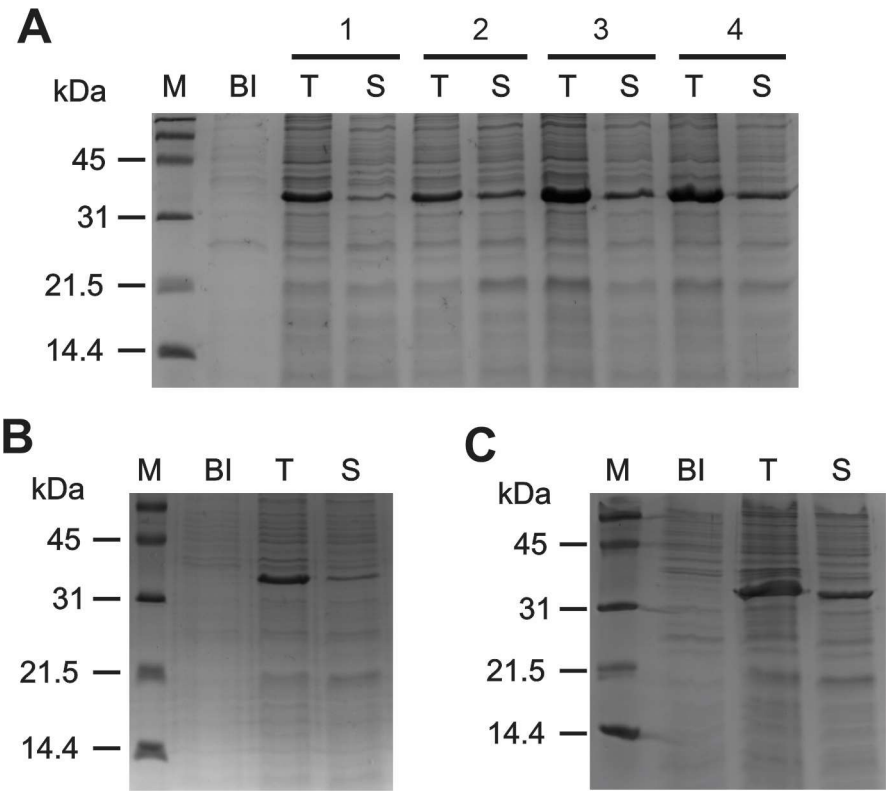


Figure 3.4. Expression tests for *Ld*BDF5T. **(A)** Small-scale expression test where **1** = 0.1 mM IPTG, 2 hrs expression; **2** = 0.1 mM IPTG, 3.5 hrs expression; **3** = 1 mM IPTG, 2 hrs expression; **4** = 1 mM IPTG, 3.5 hrs expression. **(B)** Large-scale expression using incubation temperature of 37 °C, inducing with 1 mM IPTG once OD_{600} reached 0.6, then incubating for 3.5 hours. **(C)** Optimised large-scale expression, all analysed by 20% SDS-PAGE. Expected molecular weight of *Ld*BDF5T is 35.2 kDa. **BI** = before induction, **T** = total protein, **S** = soluble protein, **M** = protein marker.

Finally, expression tests for the mutant proteins, *Ld*BDF2(N87F), *Ld*BDF5.1(N90F) and *Ld*BDF5.2(N257F) followed a similar methodology, varying IPTG concentration and incubation time. Figure 3.5 shows that *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F) expressed well under all conditions, whereas no expression of *Ld*BDF5.2(N257F) was detected and cell growth actually declined throughout incubation, particularly following addition of 1 mM IPTG (Figure 3.5C). For the other two proteins, conditions were scaled up using 37 °C incubation, with 1 mM IPTG added once OD_{600} reached 0.6 and further incubation at 30 °C for 3.5 hours.

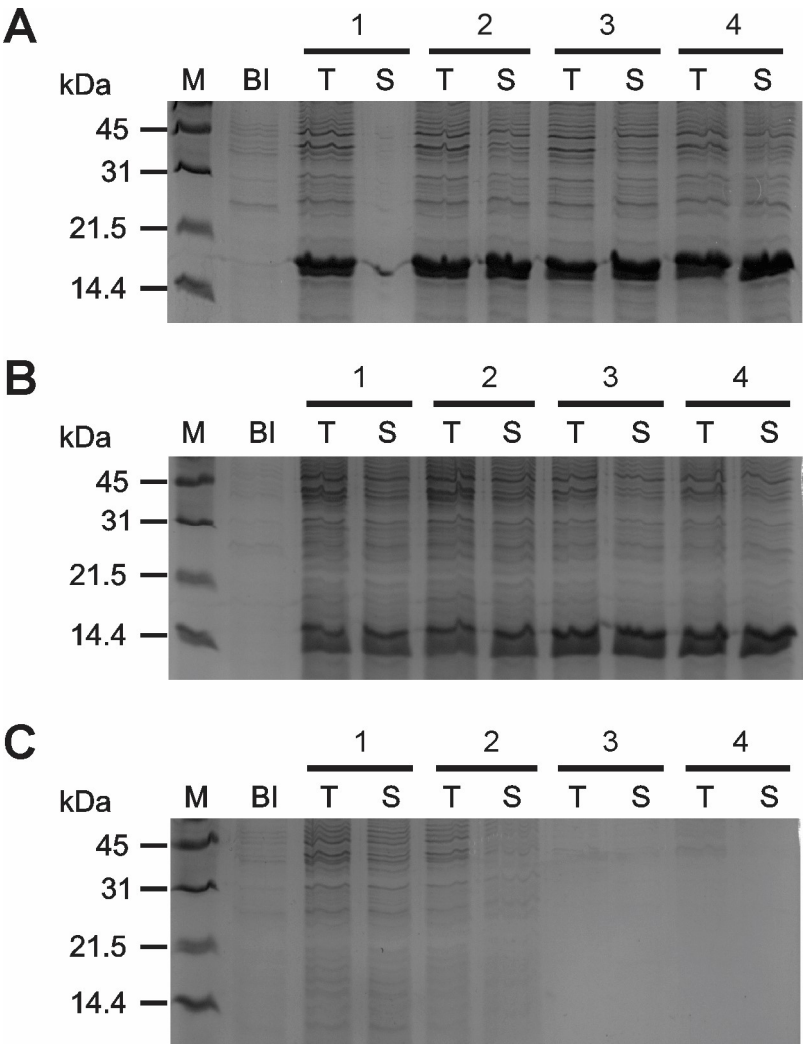


Figure 3.5. Expression tests for mutant *Ld*BDFs. 12% SDS-PAGE analysis of expression tests for **(A)** *Ld*BDF2(N87F), **(B)** *Ld*BDF5.1(N90F) and **(C)** *Ld*BDF5.2(N257F) where expected MWs are 16.6, 15.2 and 18.0 kDa, respectively. **1** = 0.1 mM IPTG, 2.5 hr expression; **2** = 0.1 mM IPTG, 3.5 hr expression; **3** = 1 mM IPTG, 2.5 hr expression; **4** = 1 mM IPTG, 3.5 hr expression. **BI** = before induction, **T** = total protein, **S** = soluble protein, **M** = protein marker. Note there was an error with the loading of sample 1, S in lane 4 of (A).

3.4. Large-scale expression and purification of *Ld*BDF2, *Ld*BDF5T, *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F)

After identifying optimal expression conditions, large-scale overexpression and purification was carried out for four recombinant proteins; *Ld*BDF2 *Ld*BDF5T, *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F) (all pET-15 plasmids, full details in Appendix A). Glycerol stocks of *E. coli* Rosetta^{TM} (DE3) pLysS competent cells harbouring the relevant plasmid vectors were used to inoculate 500 ml or 1 L LB containing the relevant antibiotics. Expression was then performed using the conditions described above, and SDS-PAGE analysis showed good levels of expression and solubility for all proteins. Protein purification was subsequently achieved through IMAC and SEC chromatography methods. The purification process is outlined below with typical example chromatograms and SDS-PAGE analysis for *Ld*BDF5T purification.

Following cell lysis, the soluble fraction was resolved first by IMAC where the His-tagged protein binds to the column via interactions with nickel. Following a column wash to remove unbound proteins, the column was developed with an imidazole gradient. In the case of *Ld*BDF5T, the chromatogram displayed one large broad peak and a smaller preceding one detected during the elution (Figure 3.6A) which were both analysed by SDS-PAGE (Figure 3.6B). Based upon this, fractions from peak B were pooled and concentrated. The cleavable N-terminal His-tag was then removed by treatment with TEV protease for *Ld*BDF2, *Ld*BDF5T, and *Ld*BDF2(N87F), though this was not possible for *Ld*BDF5.1(N90F), likely due to inaccessibility of the TEV cleavage site. Following His-tag cleavage, and a second IMAC step was performed, this time with the cleaved His-tag and TEV binding on the column whilst the protein of interest was eluted in the flow-through (Figure 3.6C). Peaks were analysed again by SDS-PAGE and those containing the highest levels of soluble protein combined. For *Ld*BDF5T this was peaks A and B (Figure 3.6D). Finally, the protein was concentrated and applied to a SEC column which separates molecules based on their size. The chromatogram for *Ld*BDF5T showed overlapping peaks which SDS-PAGE analysis revealed corresponded with a first peak (A) containing the protein of interest and a second (B) with lower molecular weight species (Figure 3.7). The fractions containing highest levels of protein (peak A for *Ld*BDF5T) were combined and concentrated to give the final protein sample, and the full

purification process visualised by a final SDS-PAGE gel (Figure 3.8). In this case, for purification of *Ld*BDF5T, some faint bands remained in the final sample, in particular a species at around 20 kDa which may have been a proteolytic fragment, likely a single bromodomain (BDF5.1 or 5.2). However, these bands were significantly fainter than the *Ld*BDF5T protein at 33 kDa which made up the majority of the sample, and gave a final yield of 12 mg purified protein from 2 L culture. Comparable yields were obtained for the other three proteins, and with higher purity of the final protein sample compared with *Ld*BDF5T (Figure 3.9).

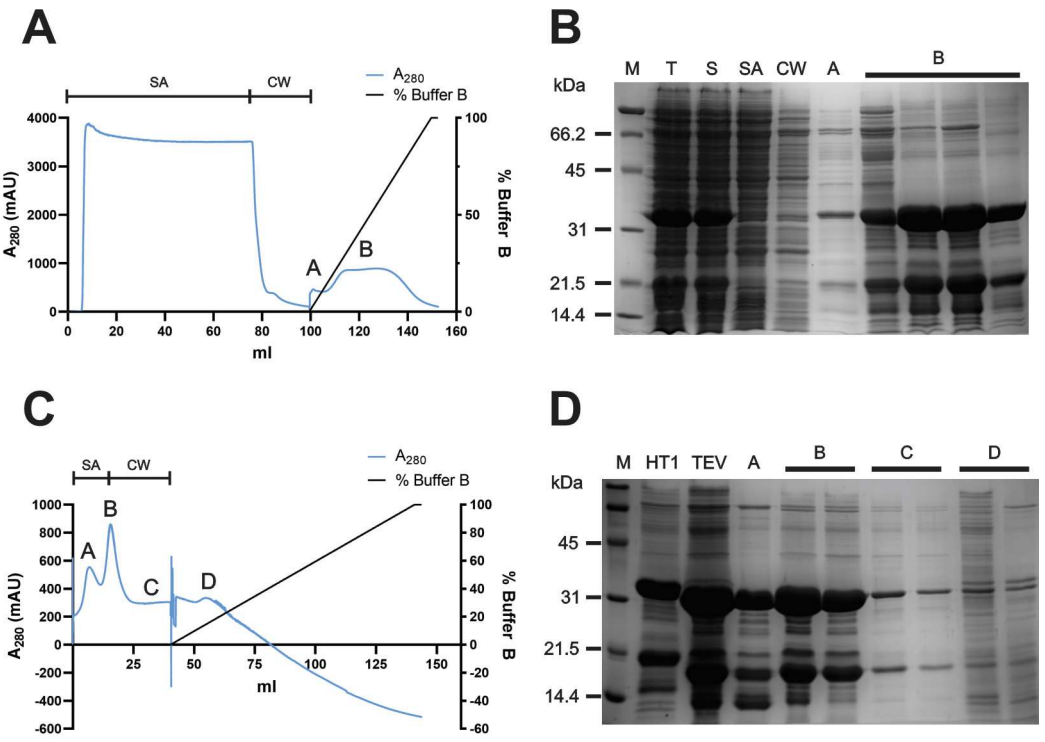


Figure 3.6. Protein purification of *Ld*BDF5T by IMAC. Top panel shows first IMAC before His-tag cleavage and bottom panel shows the second IMAC following His-tag cleavage. **(A)** & **(C)** Chromatograms showing A_{280} of the flow-through during sample application (SA) to the column, column wash (CW) with buffer A and over a linear imidazole gradient of 0% to 100% buffer B (Table D.1), with protein peaks labelled. **(B)** & **(D)** 12% SDS-PAGE analysis of IMAC purification where **M** = protein marker, **T** = total protein, **S** = soluble protein, **SA** = sample application, **CW** = column wash, **HT1** = pooled protein from the first IMAC HisTrap, **TEV** = TEV-cleaved protein, and **A-D** = chromatogram protein peaks. The expected protein MW is 35.2 and 33.0 kDa with and without His-tag, respectively.

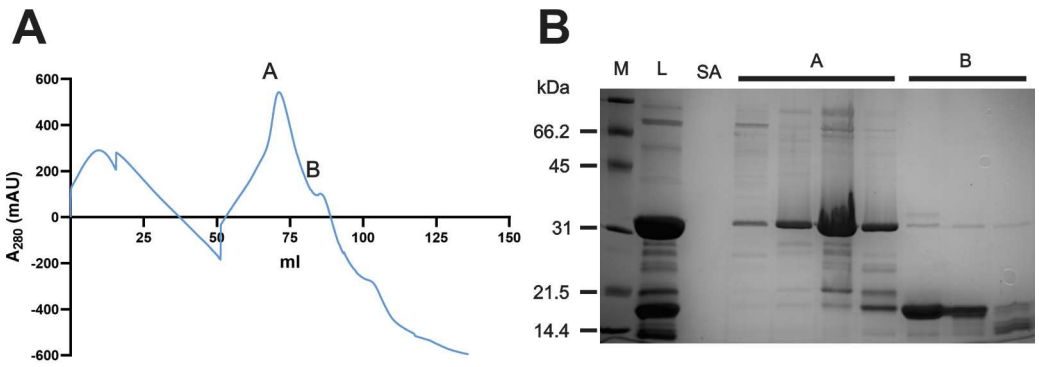


Figure 3.7. Protein purification of *Ld*BDF5T by SEC. (A) Chromatogram showing A_{280} of the protein elution during SEC with two overlapping protein peaks, **A** and **B**. **(B)** 12% SDS-PAGE analysis of SEC purification where SA is elution during sample application and wash of the sample loop. The expected His-tag-cleaved protein MW is 33.0 kDa.

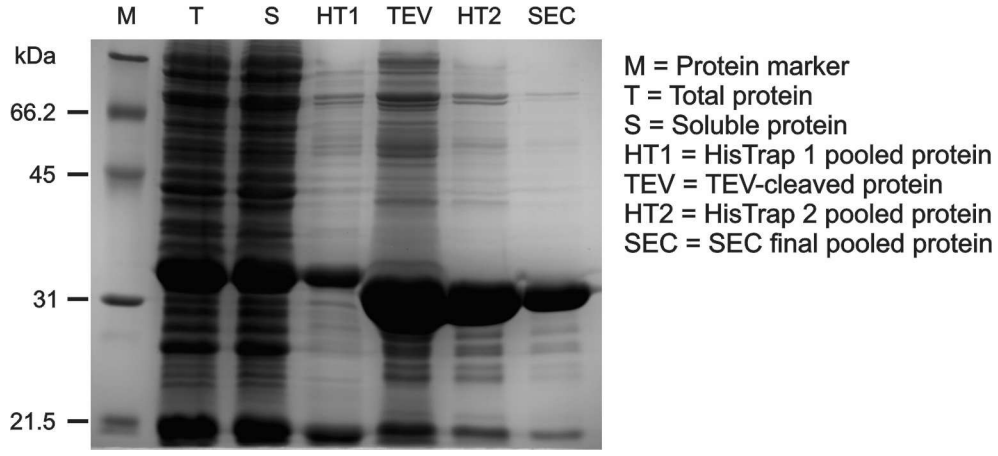


Figure 3.8. Purification of recombinant *Ld*BDF5T. Samples taken at different stages of purification process were resolved by 12% SDS-PAGE. Total and soluble protein pertain to those fractions of the lysed cell pellet; HisTrap 1 and 2 pooled protein are the combined fractions from the two IMAC purifications; TEV-cleaved protein is the sample following cleavage of the His-tag; SEC is the final pooled protein following SEC. The expected *Ld*BDF5T protein MW is 35.2 and 33.0 kDa with and without His-tag, respectively.

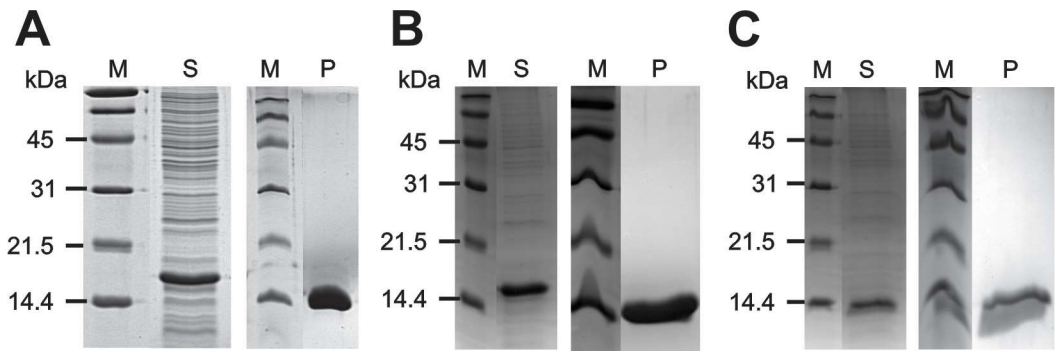


Figure 3.9. Purification of *Ld*BDF2, *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F) proteins. 20% SDS-PAGE gels showing **(A)** *Ld*BDF2, **(B)** *Ld*BDF2(N87F) and **(C)** *Ld*BDF5.1(N90F) purification where **M** = protein marker, **S** = soluble protein before purification and **P** = final purified protein. Note His-tags were cleaved for *Ld*BDF2 and *Ld*BDF2(N87F) but not *Ld*BDF5.1(N90F). Expected protein MWs for both *Ld*BDF2 proteins before and after His-tag cleavage are 16.6 and 14.5 kDa, respectively. Expected MW of *Ld*BDF5.1(N90F) is 15.2 kDa. Note gels are cropped to show only the samples of interest but the scale and alignment has been maintained.

3.5. SEC-MALLS analysis of the recombinant *Ld*BDF proteins

The four recombinant proteins; *Ld*BDF2, *Ld*BDF5T, *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F) were analysed by SEC-MALLS alongside the *Ld*BDF5.1 and *Ld*BDF5.2 bromodomain proteins which were produced separately. Note SEC-MALLS for *Ld*BDF2 were not performed using the protein purified in Section 3.4 but rather using recombinant protein produced previously using the same expression construct but without cleavage of the His-tag. This protein was also produced by Rebecca Preece in the University of York Technology Facility. SEC-MALLS experiments were performed by Dr Andrew Leech in the University of York Technology Facility. All proteins were shown to be monomeric and MALLS analysis generated molecular weight estimates close to the theoretical values (in all cases within 10%) (Figure 3.10). The oligomeric properties of the full-length *Ld*BDF proteins are currently unknown, though PTMs such as phosphorylation may influence oligomerisation and subsequent functionality as is the case for bromodomains such as human BRD4 (Malvezzi *et al.*, 2021).

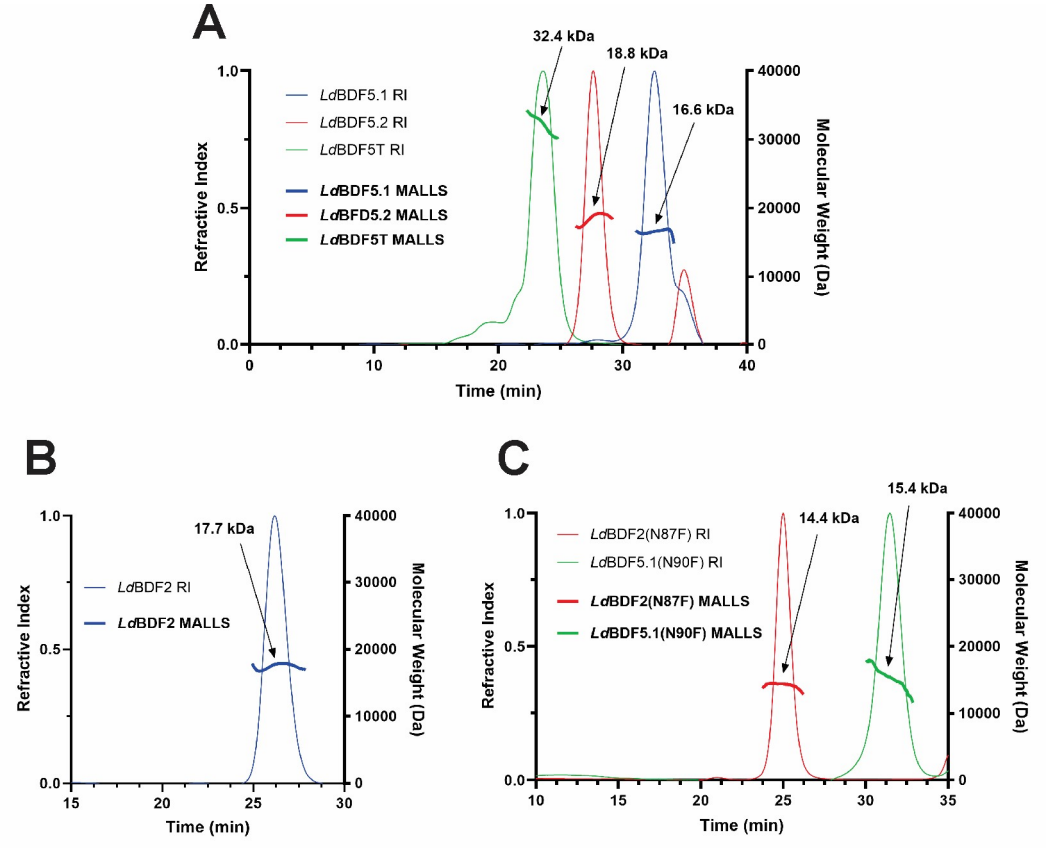


Figure 3.10. SEC-MALLS analysis of recombinant *Ld*BDF2 and 5 proteins. SEC-MALLS analysis of **(A)** *Ld*BDF5 protein variants: *Ld*BDF5.1 (blue), *Ld*BDF5.2 (red) and *Ld*BDF5T (green), **(B)** *Ld*BDF2, and **(C)** *Ld*BDF2(N87F) (red) and *Ld*BDF5(N90F) (green). Arrows pointing to MALLS curves are labelled with the associated estimated molecular weights which were in all cases within 10% of the expected values. Note for *Ld*BDF5.2 in **(A)** an artifact peak seen at 35 min elutes after the main protein peak at 27.5 min.

3.6. Discussion

Through cloning and expression trials, conditions have been established for the production of soluble recombinant proteins for several *Ld*BDF bromodomain constructs and mutant variants. Recombinant *Ld*BDF2, *Ld*BDF5T, *Ld*BDF2(N87F) and *Ld*BDF5.2(N90F) bromodomains were produced on a large-scale and purified to generate protein for use in subsequent biophysical assays. Soluble recombinant protein could not be produced for bromodomain proteins *Ld*BDF1, *Ld*BDF4 and *Ld*BDF5.2(N257F), which is consistent with past attempts to express these proteins. Future work could explore new expression conditions or look to design alternative

constructs using different length sequences for these bromodomain proteins which for now remain elusive.

Recombinant protein production was a significant achievement and formed an essential foundation for this research, with subsequent biophysical assays only possible due to the availability of these proteins. Hereinafter, *Ld*BDF2 and *Ld*BDF5 became the main focus of investigation due in part to the pre-existing genetic and structural information available for these two proteins (Jones *et al.*, 2022) (PDB codes 5C4W, 5TCM, 5TCK, 6BYA) and also in accordance with these bromodomains being the focus of high-throughput screens and novel inhibitor design, as described in Chapter 6.

4. Chapter 4: Biochemical Analysis of *Ld*BDF Histone Binding Interactions

4.1. Introduction

Epigenetic regulation in the context of histone modification is a dynamic process involving writer, eraser and reader proteins. In the case of lysine acetylation, histone acetyltransferase (HAT), histones deacetylase (HDAC) and bromodomain proteins are the respective writer, eraser and readers of this PTM. The interplay between these proteins and their histone targets has been comprehensively researched in humans, and the histone interactions of bromodomains well documented (Filippakopoulos *et al.*, 2012). In *Leishmania*, however, significantly less is known about histone acetylation and the associated protein interactions.

The proximal proteome of *L. mexicana* BDF5 has previously been defined using *in situ* proximity labelling, and includes four histone proteins; H2A.Z, H2B.V, H3 and H3.V, as well as the acetyltransferase protein, HAT2 (Jones *et al.*, 2022). Additionally, a histone peptide microarray experiment was performed by Dr Nathaniel Jones at the University of York in which *Ld*BDF2, 5.1 and 5.2 proteins were screened against a panel around 1000 unmodified, mono-acetylated and pan-acetylated histone peptides, with binding interactions identified using a fluorescent antibody binding the protein His-tags. This identified a 15 amino acid residue sequence in the *L. donovani* H2B histone which appeared to show binding to *Ld*BDF5.1 and 5.2 (unpublished data). This $H2B_{9-23}$ sequence contained four acetyllysines; H2BK9, K15, K19 and K21. In *T. brucei*, H2B exhibits minor acetylation at K12 and K16 which correspond to K15 and K19 in *L. donovani* H2B (Mandava *et al.*, 2007).

Aside from this, our understanding of lysine acetylation in *Leishmania* is limited and largely pertains to the HAT proteins. This includes data from *in vitro* assays as well as cellular evidence of HAT protein activity. In *L. donovani*, HAT1 and HAT2 have been found to acetylate H4K10 (Maity and Saha, 2012; Chandra *et al.*, 2017), whilst H4K4 was shown to be an acetylation site of both HAT2 and HAT3 (Kumar and Saha,

2015; Jha *et al.*, 2017). Additionally, H4K14 was identified as a major acetylation site and H4K2 as a potential minor acetylation site of HAT4 (Kumar *et al.*, 2012).

Previous research also points towards histone H2A.Z acetylation being important for transcriptional regulation in kinetoplastids. In addition to being identified in the proximal proteome of BDF5 in *L. mexicana* (Jones *et al.*, 2022), studies in *T. brucei* have also found that hyperacetylation of the H2A.Z N-terminus is associated with binding of *Tb*BDF2, the *T. brucei* homologue of *Ld*BDF2 (Yang *et al.*, 2017).

Whilst the *Ld*BDF proteins are predicted to bind acetyllysine residues on histone tails, their specific targets in *Leishmania* have not been identified. Therefore, the aim of the work described in this chapter was to characterise the interactions of *Leishmania* bromodomains with histones. Specifically, based upon the above research, H2A.Z, H2B and H4 were investigated using specifically designed peptides, applied in biophysical assays including a thermal shift assay (TSA) and fluorescence polarisation (FP) assay along with recombinant *Ld*BDF proteins (described in Chapter 3). These findings will contribute to our understanding of epigenetic regulation mechanisms in *Leishmania*.

4.2. Histone peptide design

Peptides were designed based on amino acid sequences of the N-terminal tails of *L. donovani* H2A.Z, H2B and H4 histones (full histone sequences in Appendix C). For each histone, the peptides included acetylated, unmodified, and fluorescently-labelled variants (Figure 4.1). The unmodified peptides were to be used as controls to establish the importance of acetylation of the lysine residues for binding to the bromodomains. To enable peptide binding to be monitored by FP, fluorophores were attached at the peptide termini. This section explains the strategic considerations behind designing probes for the FP assay and the rationale behind design of the peptides based on the H2A.Z, H2B and H4 protein sequences.

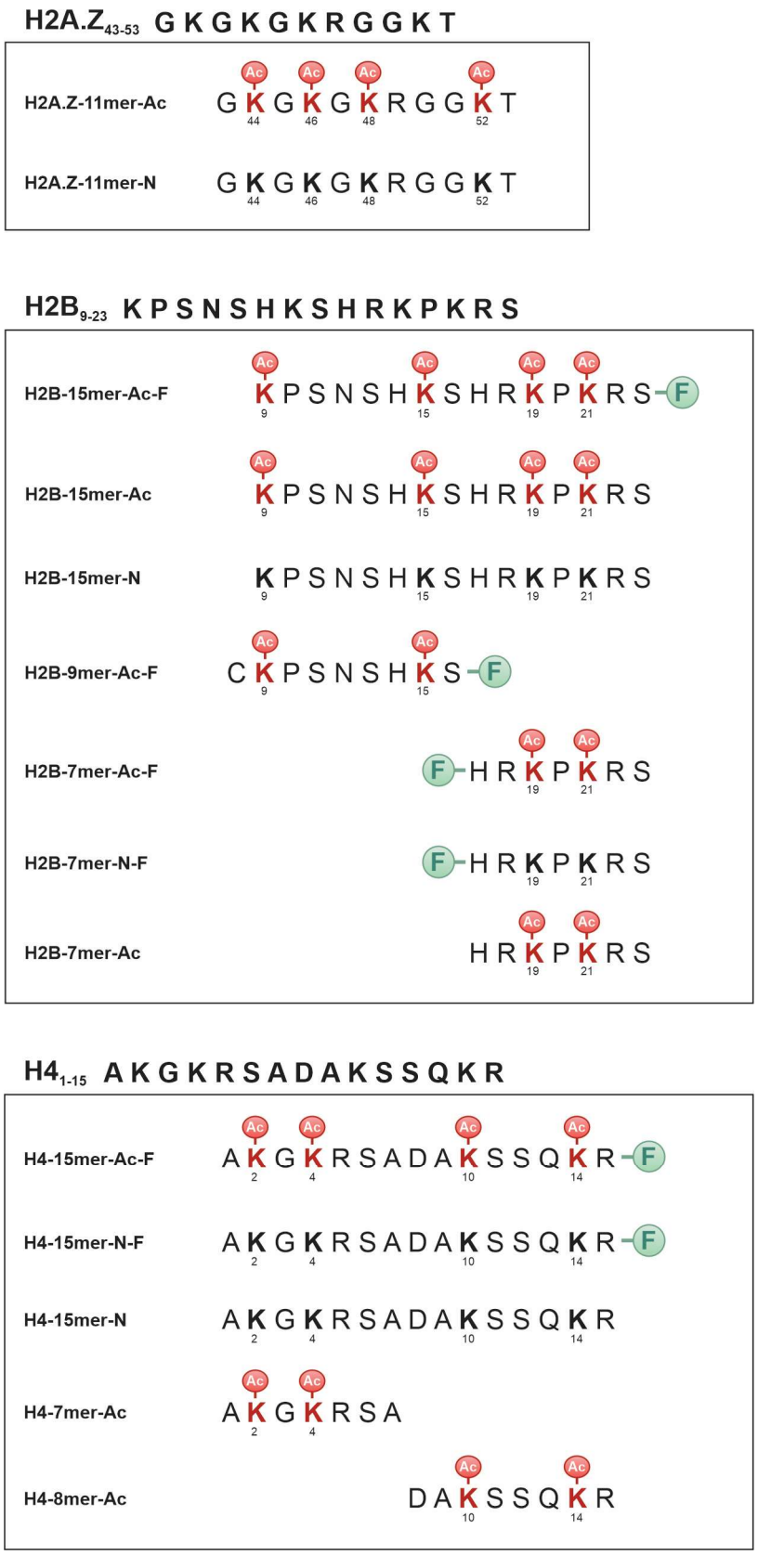


Figure 4.1. Histone peptides based on *L. donovani* H2A.Z, H2B and H4 sequences. Naming convention includes the **histone**, followed by peptide **length**, followed by **Ac** or **N** to indicate wither the peptide is acetylated or not, and finally **F** to denote peptides with a conjugated 5-FAM fluorophore.

4.2.1. Peptides mimicking histones H2A.Z, H2B and H4

Peptides were designed based on amino acid sequences of H2B, H2A.Z and H4 sequences, as labelled or unlabelled and pan-acetylated or unmodified variants (Figure 4.1). Whilst in the parasite, selective acetylation of individual lysine residues in histones is more likely than pan-acetylation, and previous work has often utilised mono-acetylated peptides (Filippakopoulos *et al.*, 2012), here pan-acetylated peptides were used due to limited resources. H2B peptide sequences were based on the 15mer sequence identified in the peptide microarray experiment designed above. Initially, H2B peptides were designed as labelled and acetylated (H2B-15mer-Ac-F), unlabelled and acetylated (H2B-15mer-Ac) and unmodified (H2B-15mer-N) forms. Following TSA and FP analysis of bromodomain binding, shorter peptides designed on similar principles were used in follow up experiments (Figure 4.1). H2A.Z 11mer peptides were designed based on residues 43 – 53 which contains four of the hyper-acetylated lysine residues seen to interact with BDF2 in *T. brucei* (H2A.ZK44, K46, K48 and K52). Unlabelled acetylated and unmodified versions of this peptide were produced for analysis.

As discussed above (Section 4.1), lysine residues K2, K4, K10 and K14 in histone H4 have been identified as potential acetylation sites of HAT1, 2, 3 in *L. donovani* (Kumar *et al.*, 2012; Maity and Saha, 2012; Kumar and Saha, 2015; Chandra *et al.*, 2017; Jha *et al.*, 2017). 15mer H4 peptides were designed spanning the region containing all four of these lysine residues. These peptides included labelled and acetylated (H4-15mer-Ac-F), labelled and unmodified (H4-15mer-N-F) and unmodified (H4-15mer-N) variants. As with H2B, additional truncated peptides were also subsequently designed to further probe the binding interaction.

4.2.2. Designing peptide probes for fluorescence polarisation

When designing peptides for use as FP probes, a number of factors had to be considered regarding both the peptide sequence and fluorescent labelling. The conjugated fluorophore in this case was 5-FAM, a fluorescein-based fluorophore commonly used to label peptides. The fluorophore was covalently attached directly to the N-terminus or via an ethylene diamine linker to the C- terminus (Figure 4.2).

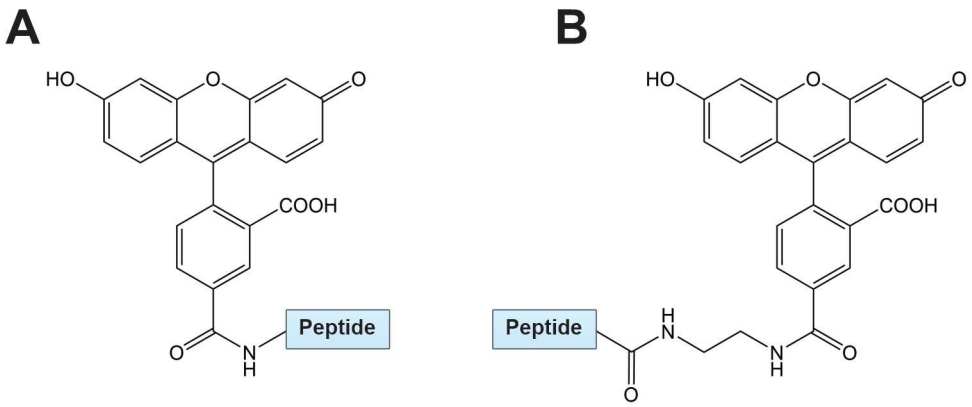


Figure 4.2. 5-carboxyfluoroscein peptide labelling. Conjugation of the 5-FAM fluorophore to peptides at either the **(A)** N-terminus or **(B)** C-terminus via an ethylene diamine linker.

Fluorophore positioning was one key consideration. To avoid interference with acetyllysine binding, the fluorophore was never positioned directly adjacent to an acetyllysine residue. On the other hand, the distance between the acetyllysine and fluorophore could not be so long as to cause the 'propeller effect' whereby the fluorophore remains highly mobile even when the probe is bound (Moerke, 2009). In parallel with this, peptide length was also an important consideration. The greater the size difference between the probe and the protein binding partner, the better the FP signal. As the recombinant *Ld*BDF proteins have low molecular weights, it was therefore particularly important to minimise the length of the peptides to optimise the FP signal. However, peptides still needed to be long enough to establish the determinants of bromodomain binding. Therefore, the peptide length was limited to a maximum of 15 amino acid residues. Terminal acetyllysines were also avoided to minimise the likelihood of non-specific binding, which is why peptide **H2B-9mer-Ac-F** contains an additional cysteine residue at the N-terminus.

4.3. Identifying *Ld*BDF-peptide interactions using TSA

The first biophysical assay used to investigate interactions between *Ld*BDF bromodomains and the histone peptides was a thermal shift assay. Using a SYPRO orange fluorescent dye which binds non-specifically to hydrophobic surfaces, the

thermal denaturation of folded proteins was measured by an increase in fluorescence, and the T_m quantified. Ligand-binding can then be inferred as this generally stabilises proteins and causes a positive thermal shift (ΔT_m). TSA is a rapid technique which is not protein-intensive and does not require displacement of a specifically designed probe, and was therefore utilised for preliminary screening of bromodomain interactions. TSA was used to assay the binding of *Ld*BDF5 to a range of the H2B and H2A.Z peptides. Binding to H4 peptides was not assessed in this assay. The **H2B-15mer-Ac** peptide was screened against *Ld*BDF5.1, *Ld*BDF5.2 and *Ld*BDF5T at concentrations of 75 and 200 µM, using optimised protein and dye concentrations (Table E.1). The recombinant *Ld*BDF5.1 protein was not successful in this TSA assay, and melting curves could not be clearly identified. For *Ld*BDF5.2 and 5T, very small positive thermal shifts were observed (Figure 4.3).

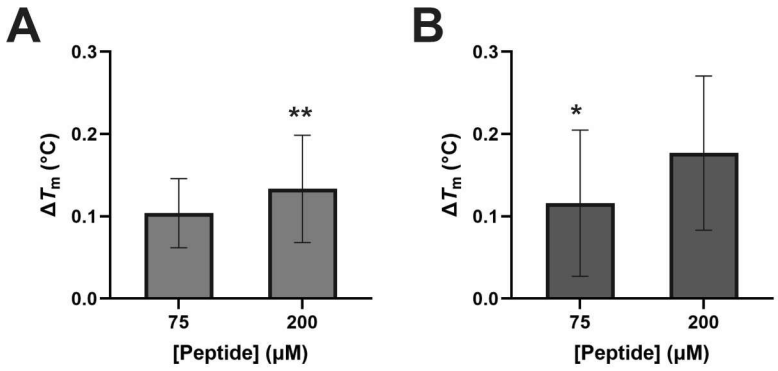


Figure 4.3. Thermal shifts for *Ld*BDF5.2 and 5T proteins with the tetra-acetylated H2B 15mer peptide at 75 and 200 µM. ΔT_m calculated as the difference between average T_m of test samples and DMSO controls for **(A)** *Ld*BDF5.2 and **(B)** *Ld*BDF5T with peptide **H2B-15mer-Ac**, showing small positive shifts. Error bars represent standard deviation (SD) of the replicate samples (n ≥ 5). P values were calculated using an unpaired t-test, * P < 0.05, ** P < 0.01, *** P < 0.001 vs respective protein + DMSO sample as indicated.

The assay was repeated using a range of higher peptide concentrations (250 – 400 µM). In these experiment, **H2B-15mer-N** was included, in which the lysine residues are not acetylated. The acetylated peptide produced greater ΔT_m values compared with the unmodified control for both proteins at all concentrations tested (Figure

4.4A,B). The greatest thermal shifts were seen for *Ld*BDF5T with 400 µM **H2B-15mer-Ac** (ΔT_m = 0.60 °C) (Figure 4.4C). The melting curves for this interaction and the corresponding unmodified peptide (Figure 4.4C,D) show that whilst small, these thermal shifts were visually distinguishable from the DMSO control, providing the first indication that *Ld*BDF5 is specifically interacting with the acetylated H2B sequence.

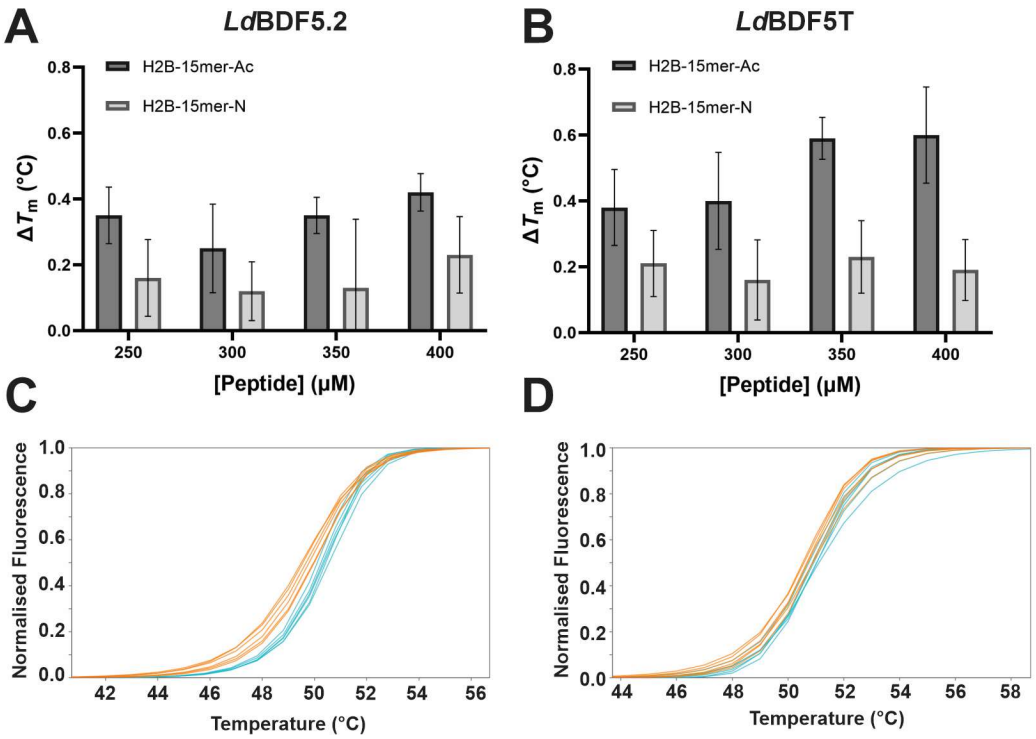


Figure 4.4. Thermal shifts for *Ld*BDF5.2 and 5T with tetra-acetylated H2B 15mer acetylated and unmodified peptides. (A) ΔT_m calculated as the difference between average T_m of test samples and DMSO controls for *Ld*BDF5.2 and **(B)** *Ld*BDF5T with peptide **H2B-15mer-Ac** compared with **H2B-15mer-N** at 250 – 400 µM. Error bars represent standard deviation (SD) of the replicate samples (n = 6). **(C)** Melting curves for *Ld*BDF5T with **(A) H2B-15mer-Ac** acetylated peptide and **(B) H2B-15mer-N** unmodified peptide at 400 µM, with normalised curves for five parameter sigmoid equation model fitting for each of the six replicate samples are shown for the protein with peptides (blue) alongside DMSO control samples (orange). Graphs produced using the online JTSA tool.

TSA was next used to examine whether the *Ld*BDF5 bromodomains interacted with H2A.Z_{43-53}, using the peptides H2A.Z-11mer-Ac and H2A.Z-11mer-N. The peptides were tested at 200 and 400 µM using 3 µM protein and 3x dye. No positive thermal shifts were seen for either *Ld*BDF5.2 or *Ld*BDF5T with the acetylated peptide. This suggests that *Ld*BDF5 either does not interact with this region of the histone, or binding is too weak to be detected by TSA. It is possible that *Ld*BDF5 interacts with a different region of the hyper-acetylated H2A.Z N-terminal tail, which could be explored using peptides spanning these other regions of the histone sequence.

Overall, whilst largely empirical, the TSA corroborated peptide microarray data supporting an interaction of *Ld*BDF5 with H2B. The assay did present limitations including discrepancies between replicates and variations in T_m values recorded in different runs, and small shifts which could not be correlated with binding affinity, yet are indicative of weak binding interactions as is the nature of this low affinity system. Nevertheless, TSA provided a useful rapid first screening method to investigate *Ld*BDF histone binding interactions.

4.4. Development of a fluorescence polarisation assay to investigate *Ld*BDF interactions

Fluorescence polarisation was selected as a second assay to test histone binding. This technique is not highly protein-intensive, can be adapted to a high-throughput format, and has been successfully applied within previous bromodomain research including the development of a pan-BET bromodomain FP probe, and using mutant proteins to divulge the interactions of the inhibitor JQ1 with human BRD4 (Jung *et al.*, 2014; Paulson *et al.*, 2018). Interactions of the recombinant *Ld*BDF proteins with histone peptide probes was assayed in parallel with establishing a peptide probe for subsequent FP competition assays. Six fluorescently-labelled peptides were used in FP, based on H2B and H4 sequences (Figure 4.1).

4.4.1. Fluorescence polarisation probe optimisation

The fluorescence intensity and polarisation of the six histone peptide probes was first measured in order to determine an appropriate probe concentration for use

in subsequent FP binding experiments with the recombinant *Ld*BDF proteins. Serial dilutions of the probes (0 – 1000 nM) were prepared in triplicate in a 384-well plate, and fluorescence intensity and polarisation readings taken using a microplate reader. Average, blank-corrected values were plotted against probe concentration. The data showed the expected trends, as shown for H2B-15mer-Ac-F (Figure 4.5). Similar plots were obtained for the other probes (Appendix E, Figure E.1). Fluorescence intensity increased linearly with probe concentration, signifying that the fluorophore was emitting fluorescence as expected when conjugated to the peptides. FP values showed variability at low concentrations before stabilising above around 600 nM for all probes. Therefore, 800 nM was chosen as an appropriate probe concentration to use in the FP assay.

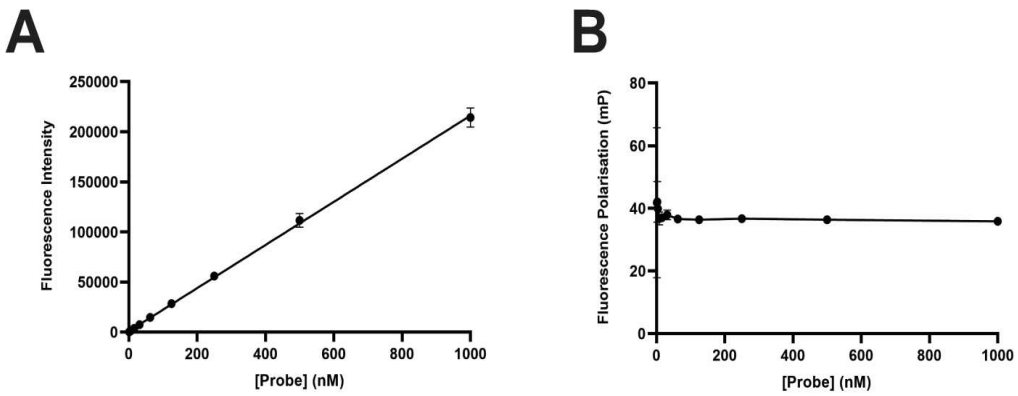


Figure 4.5. Optimisation of the H2B-15mer-Ac-F probe for FP. (A) Fluorescence intensity and (B) fluorescence polarisation in millipolarisation (mP) units for the histone peptide probe H2B-15mer-Ac-F. Average, blank-corrected values are plotted against probe concentration and fluorescence intensity data is fitted with linear regression. Error bars represent SD (n = 3).

4.4.2. Binding of *Ld*BDF bromodomains to peptide probes

The six probes were then screened for binding to the *Ld*BDF bromodomains by titrating a fixed concentration of probe (800 nM) against increasing concentrations of the proteins up to 300 µM in triplicate samples with FP measured using a microplate reader. The first probe to be examined was H2B-15mer-Ac-F, the tetra-acetylated 15mer H2B sequence which previously showed indication of binding to *Ld*BDF5. FP assays were performed to assess binding to *Ld*BDF2, 5.1, 5.2 and 5T

bromodomain proteins. FP data were analysed by transforming the average blank-corrected FP values (Figure 4.6A) to ΔmP by subtracting the FP reading from the probe alone control sample (Figure 4.6B) before fitting the data to one- or two-site specific non-linear regression models (Figure 4.6C) for the single or dual bromodomain proteins, respectively, to calculate K_D. Standard deviation (SD) was also calculated for the triplicate replicate samples, shown as error bars, though in some cases SD values were small and not clearly visible on these and subsequent graphs. H2B-15mer-Ac-F showed no clear indication of binding to *Ld*BDF2. The probe exhibited weak binding to *Ld*BDF5.2, stronger binding to *Ld*BDF5T and the strongest interaction with *Ld*BDF5.1 (K_D = 117 ± 2.3 µM). These data were consistent with both the peptide microarray and TSA indicating binding of *Ld*BDF5 to acetylated H2B, and additionally these findings suggest that the first bromodomain is more important for the binding interaction than the second.

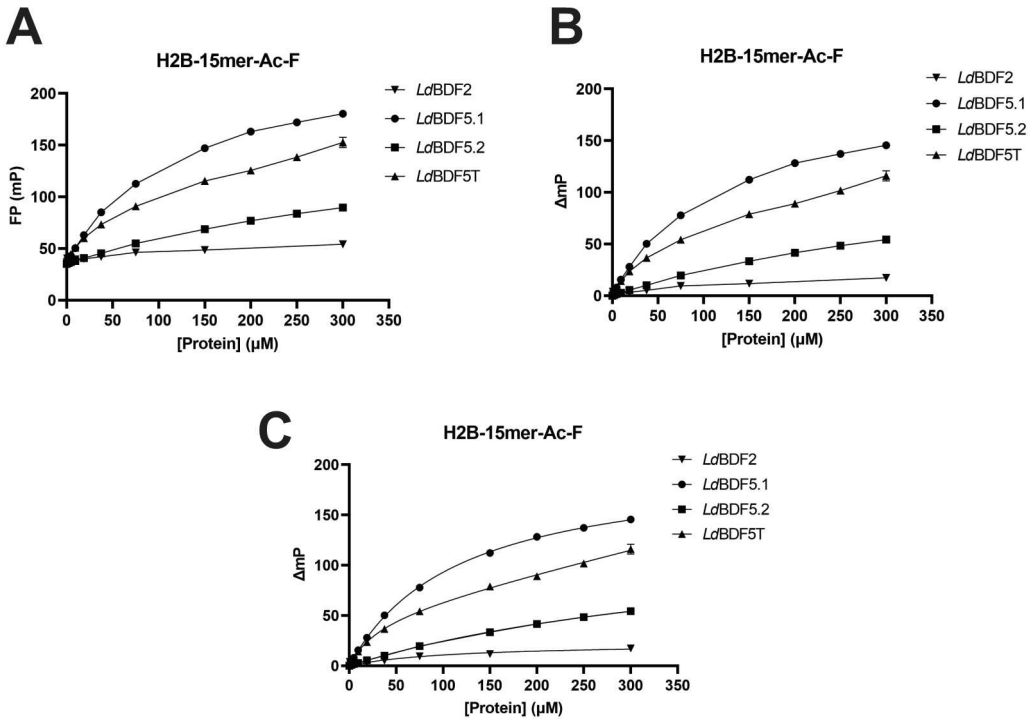


Figure 4.6. FP binding curves for H2B-15mer-Ac-F with *Ld*BDF2, 5.1, 5.2 and 5T. Data analysis showing protein concentration plotted against **(A)** raw average blank-corrected FP values, **(B)** transformed ΔmP values and **(C)** model-fitted ΔmP values, with one- or two-sites specific binding non-linear regression models fitted for K_D determination. Error bars represent SD (n = 3), though are largely not visible on account of low SD values.

The interaction of *Ld*BDF5 and the tetra-acetylated H2B histone peptide is complex, involving two distinct bromodomains and four acetyllysine residues (K9, K15, K19 & K21). Therefore, subsequent FP experiments were conducted to explore the *Ld*BDF5-H2B interaction in more detail. New peptides were designed, splitting the 15mer H2B sequence into two peptides each containing two acetyllysines; H2B-9mer-Ac-F and H2B-7mer-Ac-F (Figure 4.1). These were assayed for binding to *Ld*BDF5.1 and 5.2, with proteins titrated against the fixed concentration of probe and binding measured using FP readings. Data were analysed as described above and plotted with one-site specific models fitted (Figure 4.7). The H2B-7mer-Ac-F peptide containing acetylated K19 and K21 residues showed the stronger binding to both proteins, with the strongest overall binding between *Ld*BDF5.1 and this 7mer peptide (K_D of 43.3 ± 1.3 µM). These results suggest that the binding of *Ld*BDF5 to H2B primarily involves an interaction between the first bromodomain of *Ld*BDF5 and one or both of the acetylated K19 and K21 residues. Binding may be augmented by an additional weaker involvement of the second bromodomain.

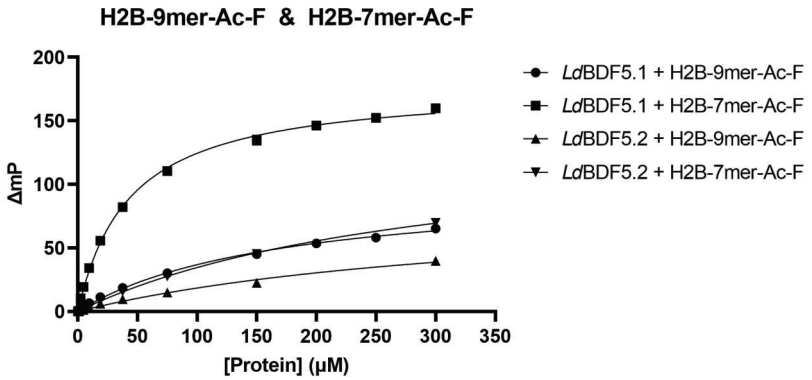


Figure 4.7. FP binding curves for H2B-9mer-Ac-F and H2B-7mer-Ac-F with *Ld*BDF5.1 and 5.2. Average blank-corrected ΔmP values plotted against protein concentration with one-site specific binding non-linear regression models fitted for K_D determination. Error bars represent SD (n = 3) though are not visible on account of low SD values.

The next histone to be analysed by FP was H4. In *L. donovani*, H4K4, K10 and K14 are known to be acetylated by HAT1, 2, 3 or 4 proteins, with H4K2 also being a potential minor acetylation site of HAT4 (Kumar *et al.*, 2012; Maity and Saha, 2012; Kumar and Saha, 2015; Chandra *et al.*, 2017; Jha *et al.*, 2017). The H4-15mer-Ac-F

15mer peptide used in FP contained these four acetylated lysine residues. The peptide exhibited a clear binding interaction with *Ld*BDF5.1 (K_D = 139 ± 7.9 µM), comparable to **H2B-7mer-Ac-F** which was included as a positive control (Figure 4.8). In contrast, the H4 probe did not show any clear indication of binding to *Ld*BDF2, and only a weak FP response with *Ld*BDF5.2. These findings suggest that *Ld*BDF5 also associates with the H4 histone, likely involving an interaction between the first bromodomain and acetylated H4K2, K4, K10 or K14.

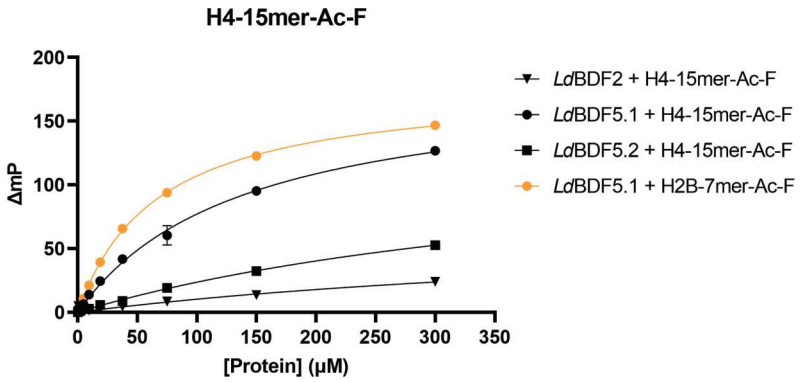


Figure 4.8. FP binding curves for H4-15mer-Ac-F with *Ld*BDF2, 5.1 and 5.2. Binding of recombinant bromodomain proteins to the H4 15mer peptide alongside the positive control **H2B-9mer-Ac-F** with *Ld*BDF5.1. Average blank-corrected ΔmP values are plotted against protein concentration with one-site specific binding non-linear regression models fitted for K_D determination. Error bars represent SD (n = 3) though are largely not visible on account of low SD values.

As illustrated by the binding experiments (Figure 4.6, 4.7, 4.8 & 4.10) and associated K_D values (Table 4.1), *Ld*BDF5.1 and 5.2 bromodomains interacted with both H2B and H4 histone peptides in the FP assay. *Ld*BDF5.1 showed the strongest binding to all peptides with K_D values within a relatively narrow range of 77.5 – 170 µM. *Ld*BDF5.2 showed notably weaker interactions, and strongest binding to **H2B-7mer-Ac-F** (K_D = 160 ± 10 µM). This was also the peptide which bound most strongly to *Ld*BDF5.1 (K_D = 77.5 ± 0.95 µM). These K_D values are comparable with the reported binding affinities for human bromodomains and fall well within the range of affinities for mono-acetylated histone peptides measured in a much wider analysis (Filippakopoulos *et al.*, 2012).

Table 4.1. K_D values for binding of H2B and H4 peptide probes to *Ld*BDF5.1 and *Ld*BDF5.2 determined by FP. Binding affinity calculated by plotting average blank-corrected ΔmP values against protein concentration and fitting one-site specific binding non-linear regression models. Heat map display shows 'hot' colours corresponding to lower K_D values. K_D values are reported ± standard error. Measurements for **H2B-15mer-Ac-F** correspond to binding curves in Figure 4.6; **H2B-9mer-Ac-F** correspond to Figure 4.7; and **H2B-7mer-Ac-F** and **H4-15mer-Ac-correspond to** Figure 4.10.

Probe	*Ld*BDF5.1 K_D (μM)	*Ld*BDF5.2 K_D (μM)
H2B-15mer-Ac-F 	117 ± 2.3 (R^2 = 1.00)	463 ± 45 (R^2 = 0.997)
H2B-9mer-Ac-F 	170 ± 10 (R^2 = 0.997)	366 ± 60 (R^2 = 0.984)
H2B-7mer-Ac-F 	77.5 ± 0.95 (R^2 = 1.00)	160 ± 10 (R^2 = 0.995)
H4-15mer-Ac-F 	137 ± 4.4 (R^2 = 0.999)	353 ± 49 (R^2 = 0.989)

Overall, these findings led to the identification of two histone target sequences of *Ld*BDF5 within histones H2B and H4. Additionally, results signify that the first bromodomain of *Ld*BDF5 is more important for these binding interactions, with the second bromodomain showing some interactions but significantly weaker binding in comparison to *Ld*BDF5.1.

4.4.3. FP competition assay using unlabelled histone peptides

Once a successful probe has been established for the FP assay, competition experiments can be performed in which displacement of the probe by unlabelled competitor compounds is evidenced by a reduction in FP. Here, several attempts were made to compete unlabelled peptides against the H2B-15mer-Ac-F, H2B-7mer-Ac-F and H4-15mer-Ac-F probes for binding to *Ld*BDF5.1, 5.2 or 5T using various conditions. In all cases, 800 nM probe was used, with protein concentrations ranging from 2.5 – 150 µM, titrating competitor peptides at various concentrations up to a maximum of 500 µM. As expected, H2B-15mer-N and H4-15mer-N caused no drop in FP. Surprisingly, the acetylated peptides H2B-15mer-Ac, H2B-7mer-Ac, H4-7mer-Ac and H4-8mer-Ac also caused no FP change. This is illustrated by data for titration of peptides H2B-15mer-N and H2B-7mer-Ac (0 – 100 µM) against *Ld*BDF5.1 (2.5 µM) and the H2B-7mer-Ac-F probe (800 nM) (Figure 4.9). Here, it can clearly be seen that neither competitor peptide causes a drop in FP, whereas a positive control compound, SGC-CBP30, elicited a clear FP response consistent with binding to the protein and displacement of the probe. These observations indicate that the unlabelled peptides were not able to displace the peptide probes under the experimental conditions tested.

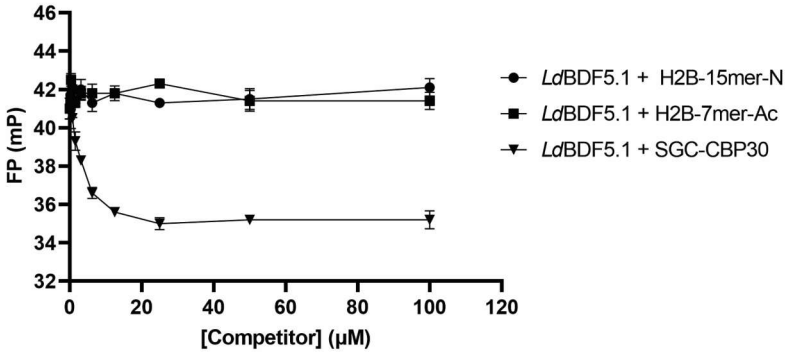


Figure 4.9. FP displacement curves for competition of peptides and positive control compound SGC-CBP30 with H2B-7mer-Ac-F probe for binding to *Ld*BDF5.1. Average, blank-corrected FP values plotted against concentration of the competitor peptide/ compound where a drop in FP indicates displacement of the probe (800 nM) and binding to *Ld*BDF5.1 (2.5 µM). Error bars represent SD (n = 3).

4.4.4. Confirmation of histone peptide binding specificity

The competitor peptides which failed to displace the histone peptide probes were identical to these probes other than lacking the 5-FAM fluorophore and in the case of H2B-15mer-Ac-F the ethylene diamine linker. Therefore, one possible explanation for the lack of competition would be that the proteins are binding non-specifically to the fluorophore or ethylene diamine linker rather than the histone sequence. To establish whether this was the case, an FP experiment was performed in which *Ld*BDF5.1 and 5.2 were titrated against both acetylated and unmodified versions of the H2B 7mer (H2B-7mer-Ac-F & H2B-7mer-N-F) and H4 15mer (H4-15mer-Ac-F & H4-15mer-N-F) probes. The FP titrations produced similar binding curves as observed previously for the acetylated peptides, whereas the unmodified versions failed to elicit an FP response (Figure 4.10). This indicated that *Ld*BDF5 bromodomains do indeed interact only with H2B and H4 peptides when the lysine residues are acetylated and thus the interaction is modification-specific.

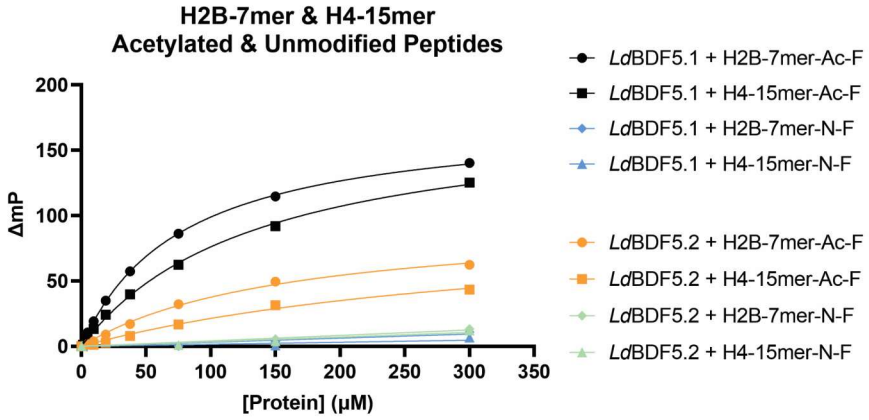


Figure 4.10. FP binding curves for H2B-7mer and H4-15mer acetylated & unmodified peptides with *Ld*BDF5.1 and 5.2. Average blank-corrected ΔmP values plotted against protein concentration with one-site specific binding non-linear regression models fitted for K_D determination. Error bars represent SD (n = 3) though are largely not visible on account of low SD values.

*Ld*BDF5.1 and 5.2 also showed no evidence of binding to the 5-FAM fluorophore alone in an accompanying negative control FP experiment in which 800 nM 5-FAM was titrated against 0 – 100 µM protein with FP measured as before (Figure 4.11A). Furthermore, the bromodomain inhibitor SGC-CBP30 did cause displacement of the probes as seen in the **H2B-7mer-Ac-F** competition experiment (Figure 4.9) and corresponding point concentration competition assays for **H4-15mer-Ac-F** (data not shown). SGC-CBP30 is known to bind within the binding pocket of *Ld*BDF5.1 as a co-crystal structure has been solved of this complex (PDB code 6BYA) (Figure 4.11B) so this provides direct evidence that the probe is also selectively binding within the bromodomain binding pocket as it is displaced by SGC-CBP30.

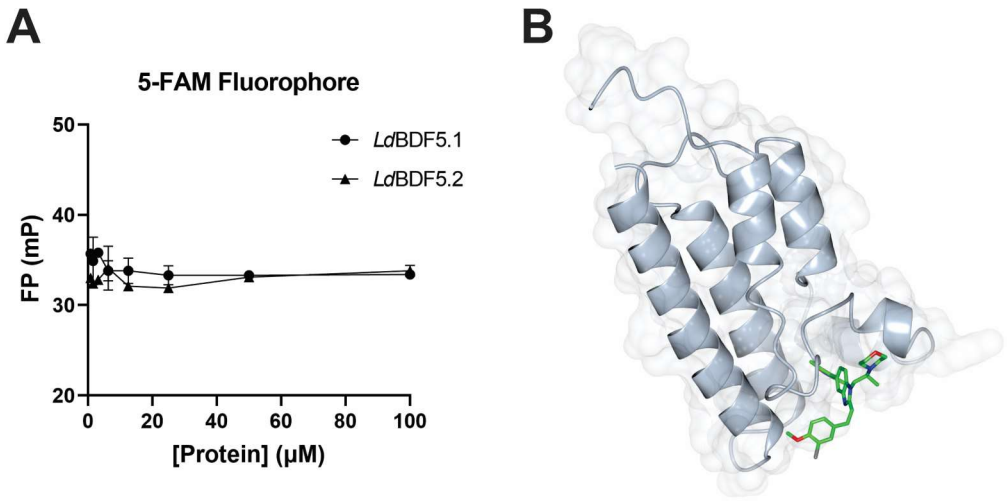


Figure 4.11. Evidence for specificity of histone peptide binding. (A) FP titration of 5-FAM against *Ld*BDF5.1 and 5.2 with Average blank-corrected FP values plotted against protein concentration and error bars showing SD (n = 3). **(B)** X-ray structure of *Ld*BDF5.1 in complex with SGC-CBP30 (PDB code 6BYA); image created using CCP4MG software.

Together, this evidence confirms that despite not seeing displacement of the **H2B-7mer-Ac-F** and **H4-15mer-Ac-F** probes by corresponding unlabelled peptides, the interaction is specific to the acetylated histone sequence and the probes bind within the binding pocket of the protein as expected.

4.5. Co-crystallisation screens for *Ld*BDF5.1 with an H2B peptide

After TSA and FP assays confirmed binding of *Ld*BDF5 to acetylated H2B, attempts were made to co-crystallise the protein bound to an H2B peptide and determine the three-dimensional structure of this complex. FP data indicated the strongest interaction between *Ld*BDF5.1 and the di-acetylated 7mer, so the unlabelled version of this peptide (H2B-7mer-Ac) was used in co-crystallisation screens with the purified recombinant *Ld*BDF5.1 protein. Crystallisation trials used a 96-well plate sitting drop format with JCSG+ and PACT crystallisation screens, with 19 mg/ml protein and either 1.2 or 1.5x excess peptide (1.9 or 1.5 mM) using 150:150 drops. Unfortunately, these screens failed to yield any crystals.

4.6. Discussion

Using the recombinant *Ld*BDF bromodomain proteins described in Chapter 3, alongside specifically designed histone peptide probes, TSA and FP assays were used to determine histone binding partners of *Ld*BDF5. Hitherto, the only direct evidence for histone binding by *Ld*BDF5 came from the peptide microarray described above (Section 4.1). Histone H2B binding was confirmed and further dissected using the FP assay, with the $H2B_{17-23}$ region containing acetylated residues K19 and K21 showing the strongest binding to both *Ld*BDF5 bromodomains (K_D = 77.5 ± 0.95 µM and 160 ± 10 µM for *Ld*BDF5.1 and *Ld*BDF5.2, respectively). Additionally, $H4_{1-15}$ (containing acetylated residues K2, K4, K10 and K14) was identified as a ligand of *Ld*BDF5.1 (K_D = 137 ± 4.4 µM), also displaying a weak interaction with *Ld*DBF5.2 (K_D = 353 ± 49 µM) in the FP assay. The interaction of *Ld*BDF5 with one or more of the H4 lysine residues K2, 4, 10 or 14 correlates well with existing knowledge that these residues are acetylated by HAT proteins in the parasite (Kumar *et al.*, 2012; Maity and Saha, 2012; Kumar and Saha, 2015; Chandra *et al.*, 2017; Jha *et al.*, 2017). In addition, these findings highlight the previously unknown importance of H2B acetylation in *Leishmania* epigenetics, though the acetylation status of H2BK19 and K21 in living parasites has yet to be determined.

4.6.1. *Ld*BDF5-histone binding mechanism

Regarding potential binding mechanisms of *Ld*BDF5 with $H2B_{17-23}$ and $H4_{1-15}$, the presence of multiple acetyllysines in these sequences could permit simultaneous engagement of both bromodomains. However, FP experiments indicate that the first bromodomain of *Ld*BDF5 binds both histone sequences with significantly higher affinity than the second bromodomain. Therefore, it may be postulated that the second bromodomain acts as a low affinity landing site to facilitate the interaction. In accordance with this, it may have been expected that the highest binding affinity would be observed with *Ld*BDF5T, in which both bromodomains are present, however in the case of H2B-15mer-Ac-F this was not the case. Instead, the peptide actually produced a greater FP response with *Ld*BDF5.1 than with the tandem bromodomain protein *Ld*BDF5T (Figure 4.6). There could be a mechanistic explanation for this observation, and indeed the complexity with which tandem bromodomains can bind their histone ligands is evident from previous research which has suggested an array of different mechanisms, including co-operative binding, independent binding and even auto-regulation (Jacobson *et al.*, 2000; VanDemark *et al.*, 2007; Yadav *et al.*, 2022).

In yeast, the tandem bromodomain protein Rsc4, was found to be acetylated itself at the N-terminal K25 residue by the HAT protein Gcn5 *in vivo*. This permitted binding of the first bromodomain to the protein itself via this acetyllysine, and this binding actually inhibited binding of the second bromodomain to its H3 histone substrate (Figure 1.6) (VanDemark *et al.*, 2007). An auto-regulation mechanism could also provide rationale as to why *Ld*BDF5.1 was able to bind to the 15mer H2B peptide with a higher affinity than the tandem bromodomain protein *Ld*BDF5T. Indeed, there are several lysine residues in *Ld*BDF5 both preceding BDF5.1 (K6), following BDF5.1 (K203) and in the linker between the two bromodomains (K147, K148 and K158). These lysine residues could be explored to further investigate whether *Ld*BDF5 may exhibit auto-regulation. Future work to better understand the interaction of *Ld*BDF5 with both H2B and H4 could utilise peptides carrying mono-acetylation modifications as well as continuing attempts to co-crystallise the *Ld*BDF proteins with histone peptides.

4.6.2. Application of the FP assay to screen inhibitor compounds against *Ld*BDF5.1 and *Ld*BDF5.2

The FP experiments described above highlighted H2B-7mer-Ac-F as the strongest peptide to bind *Ld*BDF5.1 and 5.2, with K_D values of 77.5 ± 0.95 µM and 160 ± 10 µM, respectively. Therefore, this fluorescently-labelled peptide was chosen as the probe for subsequent FP competition experiments to test the binding of human bromodomain inhibitors and novel compounds to the *Ld*BDF5.1 and 5.2 bromodomains. As detailed in Section 4.4.1, 800 nM was chosen as an appropriate probe concentration for use in these assays. It is recommended that a protein concentration be selected which gives rise to a 50 – 80 % increase in FP between free and bound probe (Moerke, 2009). However, the interaction between the histone peptides and *Ld*BDF5 is relatively weak in relation to other FP probe-protein binding affinities, therefore a very high concentration of protein would be required to produce this FP difference. Initial trial experiments used 150 µM protein but found that this did not provide a sufficient excess of the competitor compound to yield binding data representative of the actual binding interactions. Therefore, lower concentrations of 2.5 µM *Ld*BDF5.1 and 10 µM *Ld*BDF5.2 were chosen. Although these concentrations yielded only small changes in FP, results were found to be largely reproducible and allow calculation of more accurate IC_{50} values. Using these conditions, the FP competition was used as a tool to screen compounds for binding to *Ld*BDF5.1 and 5.2 bromodomains, as described in the following chapters.

5. Chapter 5: Binding of Human Bromodomain Inhibitors to the *Leishmania* Bromodomains

5.1. Introduction

The overarching aim of this project was to investigate the inhibition of the *Leishmania* BDFs in order to validate these proteins as viable drug targets and provide rationale for the development of new inhibitors. Whilst there are no existing bromodomain inhibitors optimised against *Leishmania*, there are many human bromodomain inhibitors which have been designed both for use as chemical probes, and as treatments for diseases arising from bromodomain dysregulation (Cochran, Conery and Sims, 2019; Wu *et al.*, 2019).

A small library of 15 commercially available human bromodomain inhibitors was assembled and screened against the recombinant *Ld*BDF bromodomain proteins (Figure 5.1). The library included compounds targeting bromodomains in all eight human bromodomain families with varying selectivity. Some of these compounds have been successfully used as tool compounds to investigate bromodomain function, whilst others show promising tractability as therapeutics. One such compound is OXT015, a thienotriazolodiazepine, which has progressed to clinical trials as a cancer treatment though it has displayed dose-limiting toxicities (Noel *et al.*, 2013; Amorim *et al.*, 2016). The compounds in the library have up to nanomolar binding affinities for their bromodomain targets and range in molecular weight from 347 Da (PFI-1) to 527 Da (GSK8814).

Bromosporine, I-BET151, SGC-CBP30, BI 2536 and JQ1 were purchased commercially. GSK8814 was supplied by the Structural Genomics Consortium under an Open Science Trust Agreement. All other compounds were gifted by Raquel Gabarro-Carrion and Felix Calderon from GSK.

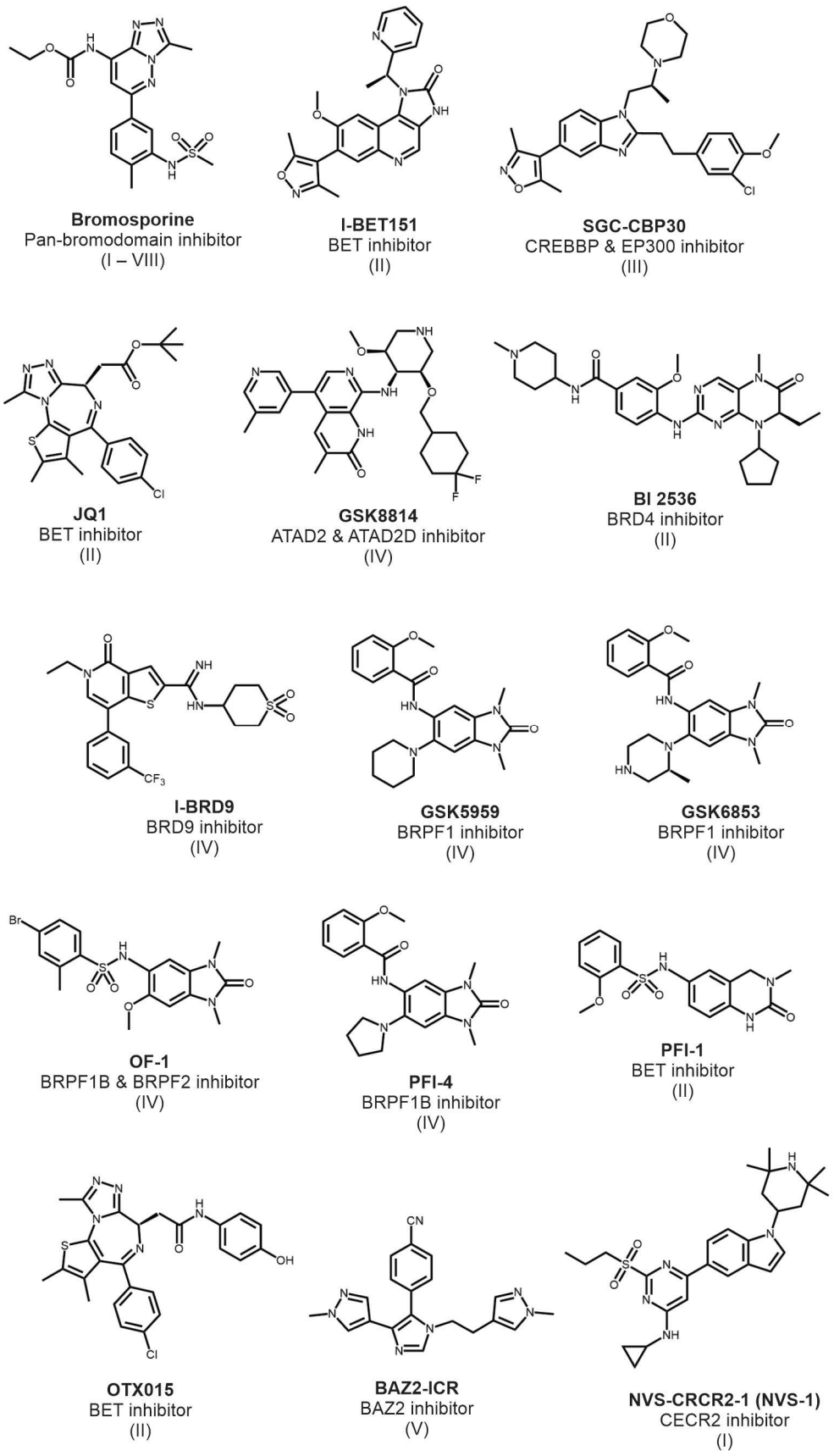


Figure 5.1. Human bromodomain inhibitor library. Structures of the 15 compounds screened against *Ld*BDF proteins, alongside their human protein target(s) and the corresponding human bromodomain families of these target(s) in brackets.

Three of these compound have previously been co-crystallised with *Ld*BDF bromodomains; bromosporine, SGC-CBP30 and BI 2536. Bromosporine is a widely used tool compound with broad-range activity against human bromodomains, and up to nanomolar affinity (Picaud *et al.*, 2016). Bromosporine has been co-crystallised in complex with *Ld*BDF2, *Ld*BDF3 and *Ld*BDF5.2 (PDB codes 5C4Q, 5FEA, 5TCK). SGC-CBP30 binds bromodomains of the HAT proteins CREBBP and EP300 with nanomolar binding affinities, with this proving to be a successful strategy for targeting multiple myeloma cell lines (Hay *et al.*, 2014; Conery *et al.*, 2016). A crystal structure has been solved for SGC-CBP30 in complex with *Ld*BDF5.1 (PDB code 6BYA). BI 2536 was originally developed as a kinase inhibitor (Steegmaier *et al.*, 2007), but has also shown potent binding to the first bromodomain of the BET protein BRD4 (K_D = 25 nM) (Ember *et al.*, 2014). This compound has also been co-crystallised with *Ld*BDF5.1 (PDB code 5TCM). Certain chemical motifs are common in bromodomain inhibitors, such as 3,5-dimethylisoxazole (Figure 5.2) which appears in both I-BET151 and SGC-CBP30. This moiety acts as an acetyllysine bioisostere and has been employed to develop bromodomain inhibitors which successfully prevent histone binding (Hewings *et al.*, 2011). Similarly, compounds OF1, PFI-4, GSK5959 and GSK6853 all share a benzimidazolone core and all target bromodomain and PHD finger-containing (BPRF) proteins.

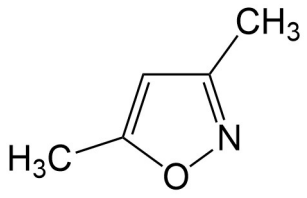


Figure 5.2. Structure of 3,5-dimethylisoxazole. An acetyllysine mimic which appears in several human bromodomain inhibitors including SGC-CBP30 and I-BET151.

This chapter describes screening of this compound library against *Ld*BDF bromodomains proteins using thermal shift assays, with further characterisation of hits by fluorescence polarisation, NMR, co-crystallography and finally cell viability assays with *Leishmania* promastigotes. The aim of this work was to provide chemical validation of *Ld*BDFs as targets for drug-discovery and explore how existing compounds interact, to inform the design of new compounds specifically targeting the *Leishmania* bromodomains.

5.2. TSA screen of bromodomain inhibitors

5.2.1. Screening inhibitors against *Ld*BDF2, 5.1 and 5.2

The 15 human bromodomain inhibitor compounds were screened against *Ld*BDF2, 5.1, 5.2 and 5T bromodomains in the thermal shift assay using optimised concentrations of protein and ligand (Table E.1) and various inhibitor concentrations between 10 and 200 µM (Figure 5.3).

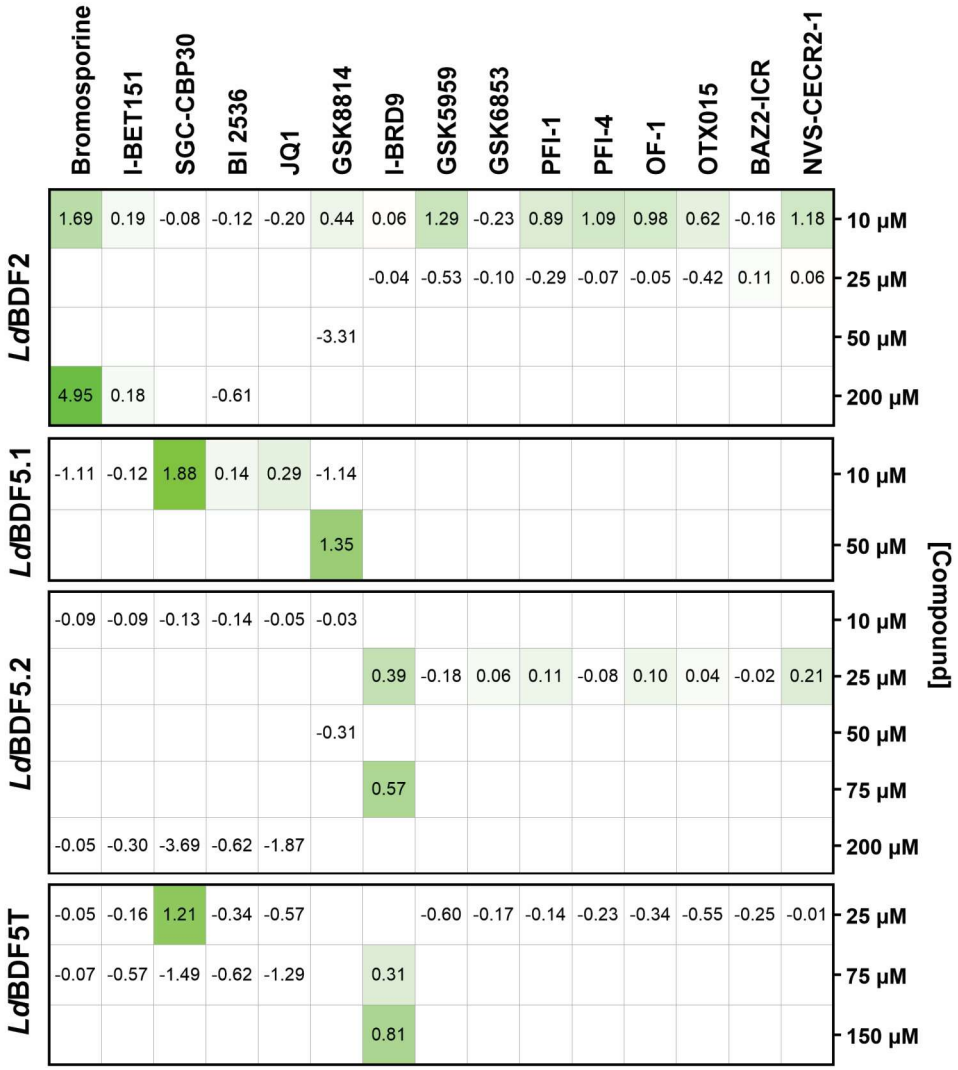


Figure 5.3. Thermal shift assays for human bromodomain inhibitors screened against *Ld*BDF2, 5.1, 5.2 and 5T. ΔT_m for compounds screened against the recombinant proteins at concentrations between 10 and 200 µM. Average T_m values were calculated for replicate samples ($n \geq 4$) and ΔT_m calculated as the difference in T_m between inhibitor samples and DMSO control samples. ΔT_m values are shown in green to indicate positive shifts, where colour intensity correlates with the degree of thermal shift.

As seen previously, *Ld*BDF5.1 showed inconsistent results in this assay and thus the data for this protein are limited to six compounds. This is also reflected by a large range of T_m values seen between replicates (n ≥ 4). By comparison, *Ld*BDF5.2 and 5T performed well and showed high consistency between replicates, providing more reliable data despite smaller ΔT_m values.

Thermal shifts were expected for *Ld*BDF2 and 5.2 with bromosporine and *Ld*BDF5.1 with SGC-CBP30 & BI 2536 as crystal structures of these complexes have previously been reported. Indeed, bromosporine showed binding to *Ld*BDF2 at both 10 and 200 µM (ΔT_m = 1.69 ± 0.80 °C and 4.95 ± 0.52 °C), but conversely gave no positive shifts for *Ld*BDF5.2 or *Ld*BDF5T. SGC-CBP30 showed binding to *Ld*BDF5.1 at 10 µM (ΔT_m = 1.88 ± 0.24 °C) but only negative shifts for *Ld*BDF2 and 5.2, indicating some degree of specificity. SGC-CBP30 also showed binding to *Ld*BDF5T at 25 µM (ΔT_m = 1.21 ± 0.06 °C), indicating that binding of this compound to the first bromodomain is not limited to the truncated protein containing the first bromodomain alone, but also occurs in the extended protein in which both bromodomains are present. One contradictory observation was that SGC-CBP30 produced a negative shift for *Ld*BDF5T when tested at 75 µM. This was also observed for other compounds, such as those screened at 200 µM against *Ld*BDF5.2. Recording negative ΔT_m values for validated ligands is not uncommon in the literature, particularly at high concentrations, and one proposed explanation is that ligands may bind to an unfolded state of the protein (Cimmperman *et al.*, 2008). A small positive thermal shift was observed for BI 2536 at 10 µM with *Ld*BDF5.1 (ΔT_m = 0.14 ± 0.23 °C), however this falls within the range of the standard deviation, and was not reproducible for the *Ld*BDF5T protein at 25 or 75 µM BI 2536. The lack of thermal shifts elicited by bromosporine with *Ld*BDF5.2, and BI 2536 with *Ld*BDF5.1 could indicate a weak binding interaction or one that does not stabilise the protein to such an extent that a thermal shift can be detected.

As well as confirming the binding interactions of bromosporine and SGC-CBP30, the TSA screen also highlighted a new hit compound for *Ld*BDF5.2. This was I-BRD9 which caused thermal shifts for *Ld*BDF5.2 when assayed at 25 and 75 µM (ΔT_m = 0.39 ± 0.07 °C and 0.57 ± 0.05 °C). The compound also showed binding to *Ld*BDF5T at 75 and 150 µM (ΔT_m = 0.31 ± 0.07 °C and 0.81 ± 0.05 °C). Whilst small, these

thermal shifts were in all four cases statistically significant (unpaired t-test, $P < 0.01$, $n = 6$). This was the first new compound to be identified as a ligand of *Ld*BDF5, binding to both the second bromodomain alone and the tandem bromodomain protein.

GSK8814 produced a relatively high thermal shift of 1.35 °C when screened at 50 µM against *Ld*BDF5.1, however the range of T_m values for the DMSO replicates far exceeded this (4.58 °C), and no thermal shift was observed at 10 µM suggesting this is a false positive. Therefore, GSK8814 was not determined to be a hit based on these results. Small positive shifts were also observed for other compounds, particularly those screened at 10 µM against *Ld*BDF2, however these were typically not reproducible and were not followed up.

Based on these distinctions, the hit compounds identified by the TSA screen were SGC-CBP30, bromosporine and I-BRD9, which were taken forward for further investigation.

5.2.2. Simultaneous binding of SGC-CBP30 and I-BRD9 to *Ld*BDF5

Building on the finding that SGC-CBP30 and I-BRD9 are ligands of *Ld*BDF 5.1 and 5.2, respectively, TSA was used to investigate the effect of incubating the *Ld*BDF5T tandem bromodomain protein with both compounds simultaneously. In this assay, SGC-CBP30 (25 µM) and I-BRD9 (150 µM) were assayed alone and together against *Ld*BDF5T using the same optimised conditions. The compounds alone produced thermal shifts of 1.18 ± 0.06 °C for SGC-CBP30 and 0.81 ± 0.12 °C for I-BRD9 (Figure 5.4) which correlated well with those seen previously at the same concentrations (SGC-CBP30 ΔT_m = 1.21 ± 0.06 °C & I-BRD9 ΔT_m = 0.81 ± 0.05 °C). The thermal shift observed for the compounds together (ΔT_m = 1.21 ± 0.05 °C) was not a significant increase compared with SGC-CBP30 alone, suggesting that either the compounds cannot bind simultaneously, or in doing so they do not stabilise the protein significantly more than binding of SGC-CBP30 alone.

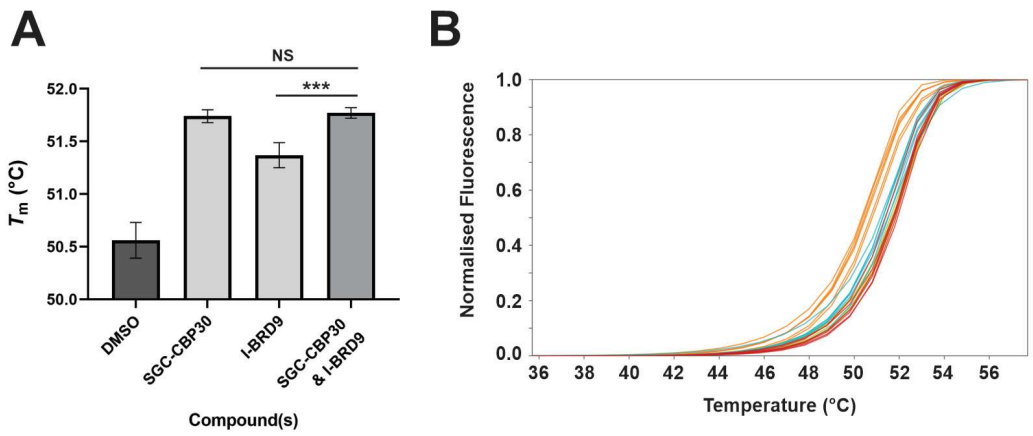


Figure 5.4. Simultaneous binding of SGC-CBP30 and I-BRD9 to *Ld*BDF5T. (A) T_m values for the compounds alone or together compared with the DMSO control for TSA screening against *Ld*BDF5T, where error bars represent SD (n = 6). **(B)** Corresponding melting curves for the DMSO control (orange), SGC-CBP30 (green), I-BRD9 (light blue), and SGC-CBP30 + I-BRD9 (red) showing normalised curves for five parameter sigmoid equation model fitting for each of the six replicates; produced using the online JTSA tool. P values were calculated for the compounds together vs alone using an unpaired t-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, NS not significant, as indicated.

5.2.3. Binding of inhibitors to mutant *Ld*BDF2 and *Ld*BDF5.1

To explore specificity and investigate binding sites of bromosporine and SGC-CBP30 with *Ld*BDF2 and 5.1, respectively, TSA was repeated using mutant proteins in which the conserved asparagine residues in the binding sites are mutated to phenylalanine (described in Chapter 3). Bromosporine (100 µM) and SGC-CBP30 (50 µM) were screened against *Ld*BDF2(N87F) and *Ld*BDF5.1(N90F), respectively using the optimised protein and dye concentrations. Comparing the melting curves to those recorded with the native proteins at the nearest concentrations, mutation of the *Ld*BDF2 bromodomain appeared to abolish binding of bromosporine (Figure 5.5A & B), suggesting strongly that the compound binds specifically via the binding pocket of the protein. Unexpectedly, the *Ld*BDF5.1 mutant exhibited a very significant positive thermal shift with SGC-CBP30 ($\Delta T_m = 4.32 \pm 0.23$°C), suggesting that the compound was still able to bind despite the presence of the bulky phenylalanine residue (Figure 5.5C & D). In the co-crystal structure of SGC-CBP30 and *Ld*BDF5.1 (Figure 5.6), the compound does bind in the binding pocket, however it is plausible that it may be able to accommodate the phenylalanine residue.

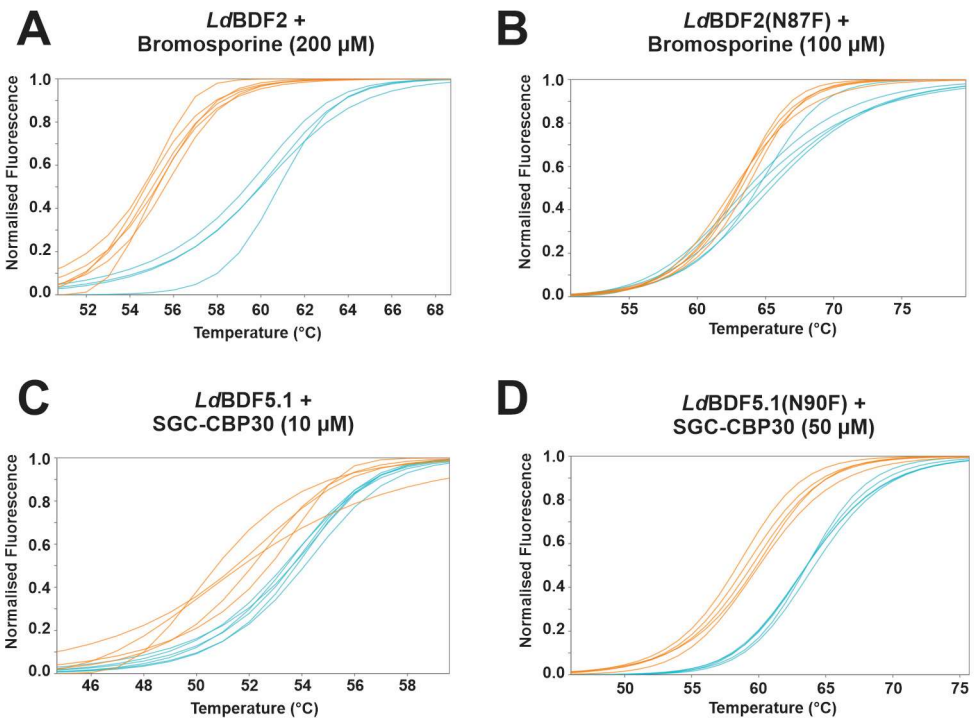


Figure 5.5. Binding of bromodomain inhibitors to native and mutant *Ld*BDF bromodomains. Melting curves for the protein DMSO control (orange) and with inhibitors (blue) for **(A)** & **(B)** *Ld*BDF2 and **(C)** & **(D)** *Ld*BDF5.1 native and mutant bromodomains, showing normalised curves for five parameter sigmoid equation model fitting for each of the replicates; produced using the online JTSA tool.

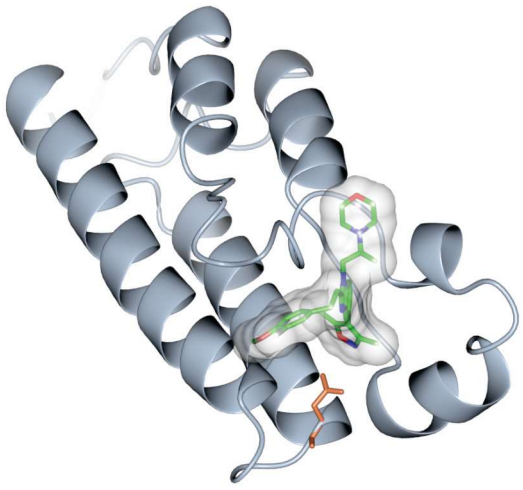


Figure 5.6. Position of Asn90 in *Ld*BDF5.1 complexed with SGC-CBP30. Crystal structure of *Ld*BDF5.1 bound to SGC-CBP30 (PDB code 6BYA) with the conserved Asn90 shown in orange to indicate position in relation to the bound inhibitor. Image created using CCP4mg software.

5.3. Biophysical characterisation of hit compounds

5.3.1. Fluorescence polarisation assay

Following the optimisation of the fluorescence polarisation assay using a histone peptide probe based on H2B (H2B-7mer-Ac-F) described in Chapter 4, this technique was implemented to validate and further characterise the interactions of the hit compounds identified by thermal shift assay; bromosporine, SGC-CB30 and I-BRD9. The displacement assay was performed with SGC-CBP30 for *Ld*BDF5.1, and bromosporine and I-BRD9 for *Ld*BDF5.2.

The compounds (0 – 100 µM) were titrated against fixed concentrations of *Ld*BDF5.1 or 5.2 (2.5 and 10 µM, respectively) and the H2B-7mer-Ac-F probe (800 nM) in triplicate samples, with FP measured using a microplate reader. Also included as a negative control in these assays was an unmodified peptide, H2B-15mer-N. Average, blank corrected FP readings were plotted against ligand concentration and data fitted with an [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression model to calculate IC_{50} values (Figure 5.7, Table 5.1).

All three compounds caused displacement of the H2B7-mer-Ac-F probe, with IC_{50} values ranging from 3 – 25 µM. The two *Ld*BDF5.2 ligands gave similar IC_{50} values though it should be noted that the displacement curves did not appear to fully plateau at the maximum compound concentration of 100 µM tested and this is reflected in the lower R^2 values for model fitting. Whilst the compounds all exhibited relatively low potency, this was expected from non-optimised inhibitors and the FP assay served its main objective of confirming the hits identified in the TSA screen.

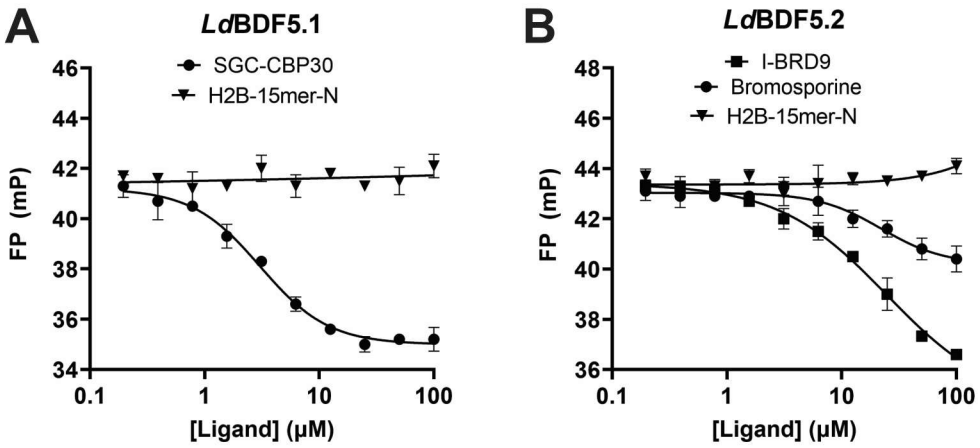


Figure 5.7. FP competition assays for *Ld*BDF5.1 & 2 with displacement of the H2B-7mer-Ac-F probe by bromodomain inhibitors. FP response for the binding of **(A)** SGC-CBP30 to *Ld*BDF5.1 and **(B)** bromosporine and I-BRD9 to *Ld*BDF5.2, compared with negative control competitor peptide H2B-15mer-N. Average, blank-corrected FP values plotted against ligand concentration, and data fitted with [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; error bars representing SD (n = 3).

Table 5.1. IC_{50} values for the binding of bromodomain inhibitors to *Ld*BDF5.1 and 5.2 as determined by FP. Values calculated by plotting average blank-corrected mP values against inhibitor concentration and fitting [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; reported as IC_{50} ± standard error (n = 3) with R^2 for model fitting.

Protein	Inhibitor	IC_{50} (μM)	R^2
*Ld*BDF5.1	SGC-CBP30	3.06 ± 0.30	0.977
*Ld*BDF5.2	Bromosporine	22.1 ± 6.5	0.875
*Ld*BDF5.2	I-BRD9	24.7 ± 6.2	0.985

5.3.2. Ligand-observed NMR

Another technique used to investigate binding of the three human bromodomain inhibitors to *Ld*BDF5.1 and 5.2 was ligand-observed proton NMR. Ligand-observed NMR screening requires fast ligand-protein exchange and is a good technique for

low-to-moderate affinity ligands but can also be prone to false negatives when screening tight binders (nM affinities). To mitigate this, and identify false negatives, three distinct experiments were performed to probe the protein-ligand interactions; waterLOGSY (Dalvit *et al.*, 2000, 2001), saturation transfer difference (STD) (Mayer and Meyer, 1999) and CPMG (Carr and Purcell, 1954; Meiboom and Gill, 1958). All three are based on the principle that nuclei in ligands give different signals when they are free or bound to a protein. WaterLOGSY uses magnetisation of the water molecules in a sample, where the transfer of magnetisation to a ligand varies depending on whether it's bound to protein or not. In saturation transfer difference, the protein is selectively saturated and the magnetisation is transferred to any bound ligands. STD can also be used to characterise the binding epitope, as protons in groups in direct contact with the protein give rise to stronger STD signals. Finally, CPMG utilises the principle that nuclei in free ligands relax slowly, so signal intensity remains consistent when using short or long relaxation times, whereas ligands that bind to a target protein behave as larger species and exhibit faster relaxation times which can be observed by a drop in intensity between the short and long relaxation times. Whilst the NMR experiments were not applied to quantify binding affinities of the compounds, they provide an orthogonal approach to complement the TSA and FP assays.

SGC-CBP30, bromosporine and I-BRD9 (0.6 mM) were screened against the relevant *Ld*BDF5 bromodomain protein (19 µM) using the three NMR experiments, recorded alongside a water-suppressed proton NMR. Illustrative spectra for SGC-CBP30 with *Ld*BDF5.1 are shown as examples of positive binding results, with insets for representative peaks at 2.30 and 2.44 ppm which were assigned to the methyl groups of the 3,5-dimethylisoxazole group (Figure 5.8, 5.10, 5.11). In waterLOGSY experiments, bound compounds are differentiated from non-binding compounds based on opposite signal peaks. Typically, bound ligands produce positive peaks whilst non-binders show negative peaks. As anticipated, the non-binding calibration standard compound DSS gave a negative signal at 0 ppm in all protein-ligand waterLOGSY spectra. Positive peaks were seen for SGC-CBP30 in the presence of *Ld*BDF5.1, indicating a binding interaction (Figure 5.8). I-BRD9 also showed a binding interaction with *Ld*BDF5.2, with most compound peaks giving positive

signals in the presence of *Ld*BDF5.2, albeit with lower peak intensities. By comparison, waterLOGSY was not able to definitively confirm binding of bromosporine to *Ld*BDF5.2, as only two peaks gave weak positive signals in the presence of the protein.

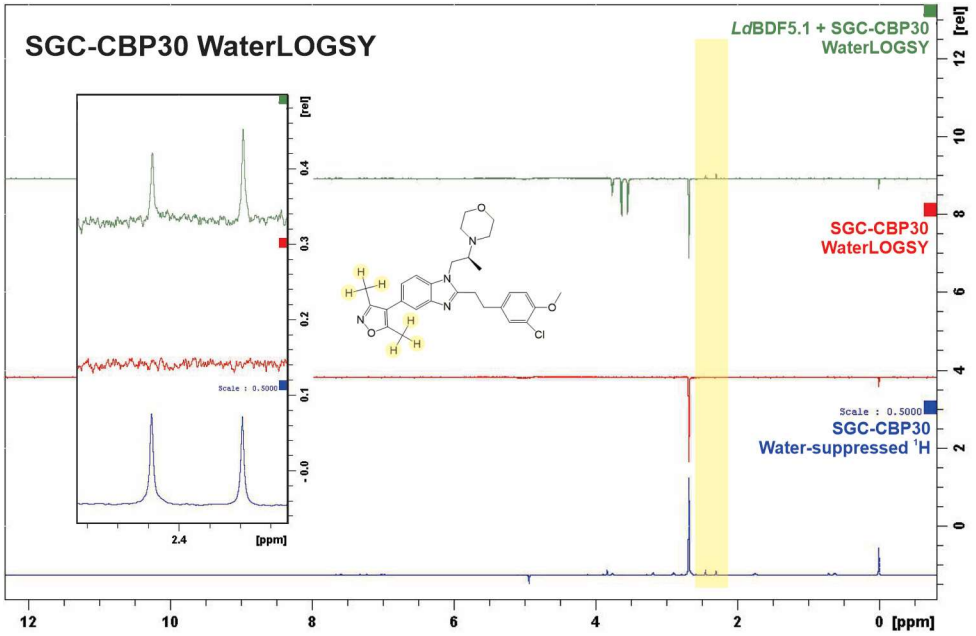


Figure 5.8. WaterLOGSY NMR spectra showing binding of SGC-CBP30 to *Ld*BDF5.1. Spectra recorded for SGC-CBP30 alone (red) and with *Ld*BDF5.1 (green) alongside the reference water-suppressed ^{1}H spectrum for the compound alone (blue). Representative ligand peaks are highlighted and shown as an inset, where these correspond to the methyl groups highlighted in the SGC-CBP30 compound structure.

In the STD experiment, saturation is transferred from the protein to any bound compounds, observed as peaks in the 'difference' spectra. Furthermore, comparing the intensities of different ligand peaks can reveal the binding epitope of the compound. STD difference peaks were present for SGC-CBP30 with *Ld*BDF5.1, and those arising from the methyl groups of the 3,5-dimethylisoxazole group showed markedly higher intensities than other peaks (Figure 5.9). This correlates with the co-crystal structure of SGC-CBP30 in complex with *Ld*BDF5.1 (Figure 4.11B; PDB code 6BYA) in which the 3,5-dimethylisoxazole group is positioned well within the hydrophobic binding pocket. Results for I-BRD9 were ambiguous; most compound peaks were either absent or very weak, with the exception of those at higher chemical shifts, particularly singlets around 8.0 and 8.3 ppm which were predicted

to correspond to two hydrogen atoms in the thienopyridone core (Figure 5.10). Bromosporine did show small, broad compound peaks above the noise, however these were comparable to peaks seen in the STD spectrum for the compound alone were not a definitive confirmation of binding.

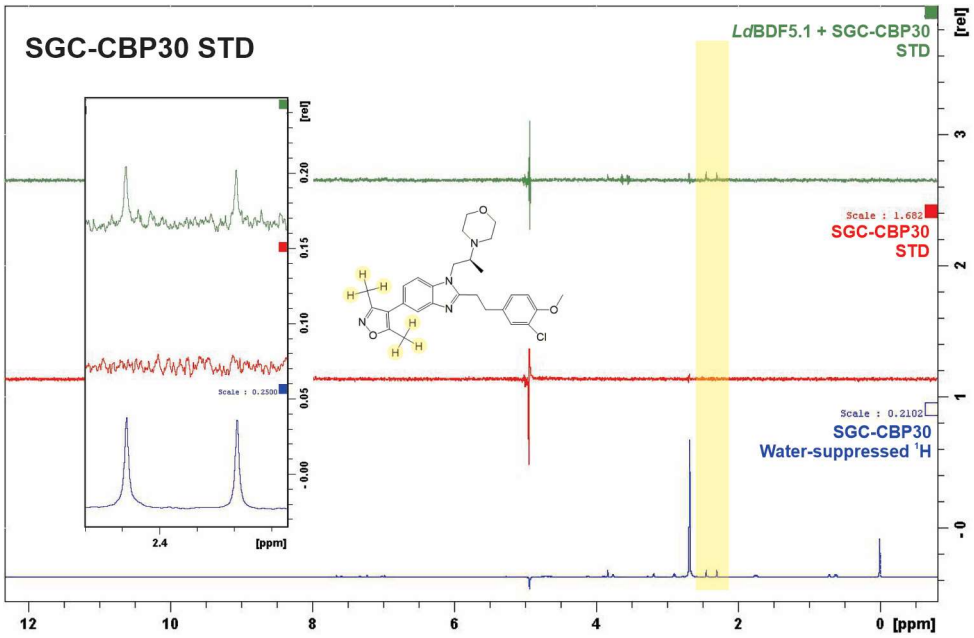


Figure 5.9. STD NMR spectra showing binding of SGC-CBP30 to *Ld*BDF5.1. Spectra recorded for SGC-CBP30 alone (red) and with *Ld*BDF5.1 (green) alongside the reference water-suppressed ^{1}H spectrum for the compound alone (blue). Representative ligand peaks are highlighted and shown as an inset, where these correspond to the methyl groups highlighted in the SGC-CBP30 structure.

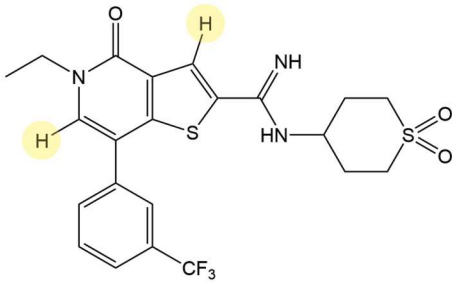


Figure 5.10. Hydrogens in I-BRD9 corresponding to peaks in STD spectra. Peaks around 8.0 and 8.3 ppm predicted to correspond to the highlighted hydrogens.

Finally, in CPMG, binding compounds are expected to show a reduction in peak intensity with a longer relaxation delay. Comparing spectra for SGC-CBP30 alone

(Figure 5.11A) and with *Ld*BDF5.1 (Figure 5.11B), a greater reduction in peak intensity is observed in the presence of the protein, indicating a binding interaction. I-BRD9 also showed peak reductions at all resonances in the presence of *Ld*BDF5.2 which were notably larger than the compound alone, confirming binding. For bromosporine, small peak reductions were observed, but these were not distinguishable from the compound alone and therefore did not constitute a positive result. Table 5.2 summarises the outcomes of the three NMR experiments.

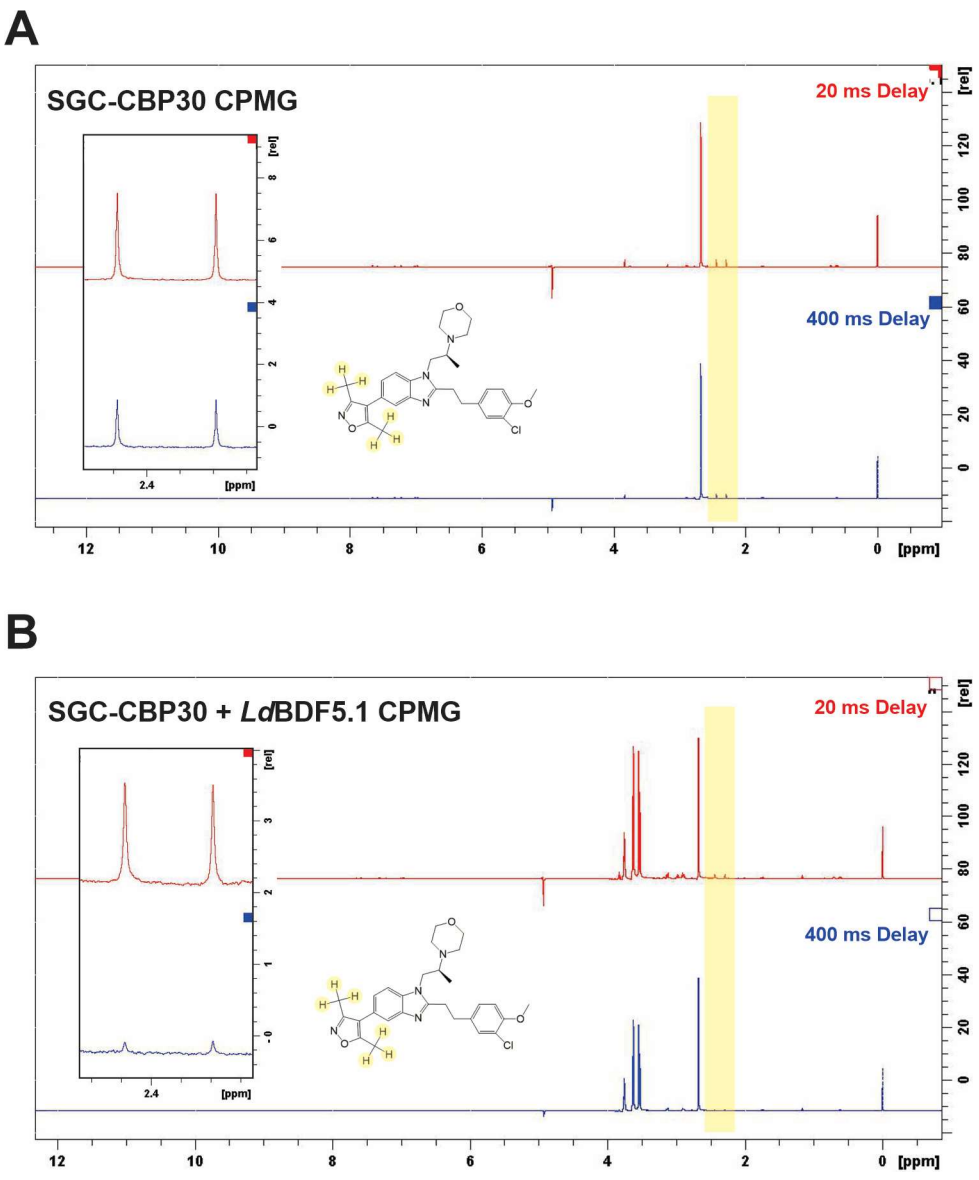


Figure 5.11. CPMG NMR spectra showing binding of SGC-CBP30 to *Ld*BDF5.1. Spectra recorded for **(A)** SGC-CBP30 alone and **(B)** SGC-CBP30 with *Ld*BDF5.1, comparing long 400 ms (blue) and short 20 ms (red) relaxation delays. Representative ligand peaks are highlighted and shown as an inset, which correspond to the methyl groups highlighted in the SGC-CBP30 structure.

Table 5.2. Outcomes of ligand-observed NMR experiments for hit human bromodomain inhibitors with *Ld*BDF5.1 and 5.2. Tick marks represent positive confirmation of binding, whilst question marks indicate inconclusive results and crosses indicate results of experiments which did not indicate a binding interaction.

	WaterLOGSY	STD	CPMG
*Ld*BDF5.1 + SGC-CBP30	✓	✓	✓
*Ld*BDF5.2 + bromosporine	?	X	X
*Ld*BDF5.2 + I-BRD9	✓	?	✓

5.3.3. Co-crystallisation screens with I-BRD9

Following the identification of I-BRD9 as a new ligand of *Ld*BDF5, attempts were made to co-crystallise I-BRD9 with *Ld*BDF5.2 and *Ld*BDF5T recombinant proteins for structure determination by X-ray diffraction (full crystallisation conditions listed in Appendix F).

Initially, screens were set up for *Ld*BDF5.2 using high concentrations of I-BRD9 as the compound showed a relatively weak binding affinity for *Ld*BDF5.2 in FP (24.7 ± 6.2 µM). However, solubility issues were encountered and these screens yielded no hits. In subsequent screens with *Ld*BDF5T, lower I-BRD9 concentrations were tested. A small number of crystals were observed in a Hampton and $AmSO_4$ screens with 8.5 mg/ml protein and 11 mM I-BRD9, using a 150:150 drop ratio. However, no fluorescence was observed when imaging these crystals under UV and those which were fished and tested at the synchrotron did not diffract, indicating these may have been salt crystals.

5.4. Activity of human bromodomain inhibitors against *Leishmania* promastigotes

Using biophysical assays, bromosporine, SGC-CBP30 and I-BRD9 were uncovered as ligands of *Ld*BDF5. However, whether this binding interaction also occurred in *Leishmania* and caused cytotoxicity towards the parasite was yet to be established. Using a cell viability assay, the compounds were screened against *L. mexicana* and *L. donovani* promastigotes by exposing the parasites to each compound (0 – 30 μM) for five days before measuring cell viability by adding resazurin. This chemical compound is metabolically reduced by cells to the fluorescent compound resorufin. Fluorescence was measured using a microplate reader, where a decrease in fluorescence intensity indicates fewer viable cells and therefore cytotoxicity of the compound. Average, blank-corrected fluorescence intensity measurements were normalised to give values as % cell viability, and data fitted to an [Inhibitor] vs. normalized response -- variable slope dose-response model for EC_{50} determination. Included as a positive control was miltefosine, an established antileishmanial (Sundar and Chakravarty, 2015).

Two biological replicate experiments including triplicate technical replicates were performed, with dose-response curves shown for the first biological replicate (Figure 5.12), alongside EC_{50} values for both experiments which showed good consistency (Table 5.3). The human bromodomain inhibitors showed similar activity against both species of *Leishmania*. As expected, miltefosine showed the highest activity but SGC-CBP30 also showed moderate cytotoxicity, with EC_{50} values ranging from 4.43 – 8.37 μM. In contrast, bromosporine showed considerably lower cytotoxicity, with EC_{50} values around 10-fold higher than SGC-CBP30, and I-BRD9 showed little to no reduction in cell viability. Future work could examine whether the SGC-CBP30 has activity in a shorter time frame (< 5 days) by repeating the assay with different periods of exposure to the compounds.

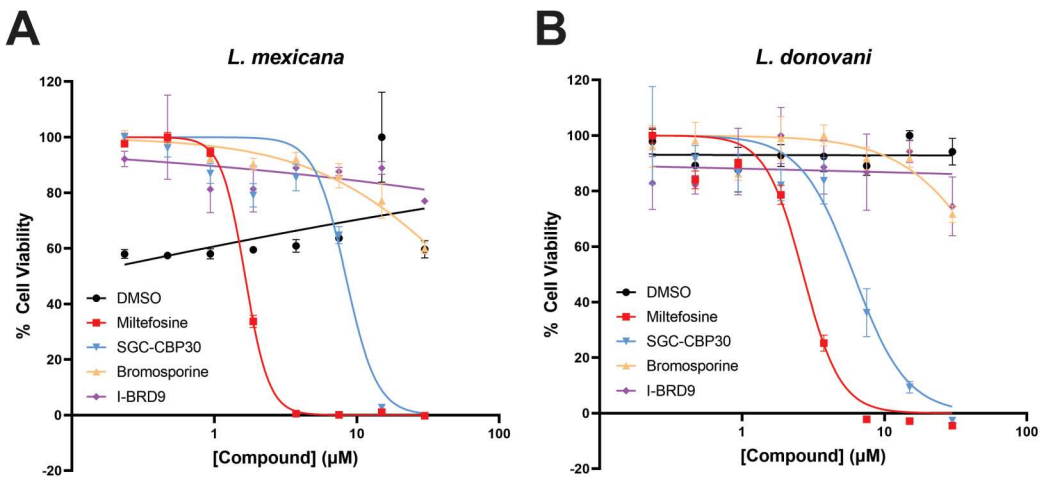


Figure 5.12. Dose-response curves for *Leishmania* promastigotes with bromodomain inhibitors. Cell viability assays for 5-day incubation of three human bromodomain inhibitors and miltefosine (0.23 – 30 μM) with **(A)** *L. mexicana* and **(B)** *L. donovani* promastigotes. Average, blank-corrected fluorescence intensity measurements were normalised and fitted to an [Inhibitor] vs. normalized response -- variable slope dose-response model. Error bars represent SEM (n = 3). Data shown for the first of two biological replicate experiments.

Table 5.3. EC_{50} values for bromodomain inhibitors with *Leishmania* promastigotes. EC_{50} values calculated from dose-response activity of three human bromodomain inhibitors and miltefosine against *L. mexicana* and *L. donovani* promastigotes, reported as EC_{50} ± standard error (n = 3). Heat map display shows 'hot' colours corresponding to lower EC_{50} values; #1 and #2 denote the two biological replicates; dashes are EC_{50} that could not be reliably determined.

Compound	*L. mexicana* #1 EC_{50} (μM)	*L. mexicana* #2 EC_{50} (μM)	*L. donovani* #1 EC_{50} (μM)	*L. donovani* #2 EC_{50} (μM)
Miltefosine	1.64 ± 0.019 (R^2 = 1.00)	0.959 ± 0.50 (R^2 = 0.605)	2.68 ± 0.24 (R^2 = 0.976)	1.67 ± 0.088 (R^2 = 0.993)
SGC-CBP30	8.37 ± 0.96 (R^2 = 0.932)	4.43 ± 2.5 (R^2 = 0.637)	6.07 ± 0.70 (R^2 = 0.960)	7.22 ± 0.67 (R^2 = 0.966)
Bromosporine	54.4 ± 11 (R^2 = 0.954)	-	60.9 ± 29 (R^2 = 0.617)	-
I-BRD9	-	-	-	-

5.5. Discussion

5.5.1. Human bromodomain inhibitors show binding to *Ld*BDFs

TSA screens identified three human bromodomain inhibitors, bromosporine, SGC-CBP30 and I-BRD9 which bound to *L. donovani* BDF2, 5.1 and 5.2 bromodomains, respectively. Identification of I-BRD9 as a ligand of *Ld*BDF5.2 was a novel finding as this interaction was not previously known, whereas co-crystal structures had been solved of the other two compounds in complex with the relevant bromodomains. Bromosporine was also understood to bind to *Ld*BDF5.2 as shown by an existing crystal structure.

Following the initial TSA screen, binding of the human bromodomain inhibitors to *Ld*BDF5.1 and 5.2 was validated and quantified by FP which yielded micromolar binding affinities ranging from 3 – 25 μM. Ligand-observed NMR experiments were able to further confirm the interaction of SGC-CBP30 with *Ld*BDF5.1 and analysis of the STD spectra also gave insight into the binding epitope, highlighting the 3,5-dimethylisoxazole group (Figure 5.2) as being prominent in the interaction. Two out of the three NMR experiments were also successful in confirming binding of I-BRD9 to *Ld*BDF5.2, however the interaction of bromosporine with this bromodomain could not be confirmed by NMR.

The three hit compounds were also applied as tools for chemical target-validation of *Ld*BDF5 within a *Leishmania* promastigote cell viability assay. Whilst bromosporine and I-BRD9 exhibited minimal cytotoxicity, SGC-CBP30 showed significant phenotypic activity against both *L. mexicana* and *L. donovani* promastigotes (EC_{50} values of 4.43 – 8.37 μM). This validated inhibition of the *Ld*BDF5 protein as a strategy to target the parasite, and proved a useful starting point for the investigation of novel inhibitors of *Ld*BDF5. The work described in this chapter also served to validate the biophysical assays as robust tools for the investigation of ligands of the *Ld*BDF proteins, which could then be implemented to screen novel compounds as described in Chapter 7.

5.5.2. Assessing the interaction of bromosporine with *Ld*BDF5.2

One notable observation was the inability to detect binding of bromosporine to *Ld*BDF5.2 using TSA or NMR despite the existence of a co-crystal structure showing this complex (PDB code 5TCK). In comparison, the FP assay did indicate a binding interaction between bromosporine and *Ld*BDF5.2 (K_D = 22.1 ± 6.5 µM). In analysing the co-crystal structure of this complex, it became apparent that bromosporine exhibits an atypical binding conformation, compared with bromosporine binding to *Ld*BDF2, *Ld*BDF3 and bromosporine human bromodomains (Figure 5.13). The triazolopyridazine core of bromosporine is oriented in the same way in the human bromodomains, *Ld*BDF2 and *Ld*BDF3, whereas in *Ld*BDF5.2 it is positioned almost perpendicular to this. It appears that this binding mode may mediate the dimerization of these *Ld*BDF5.2 bromodomain proteins seen in the co-crystal structure. This anomalous binding mode and the inconsistent results for this inhibitor binding *Ld*BDF5.2 in the TSA and NMR assays is an interesting observation and further work could look to explore this interaction further.

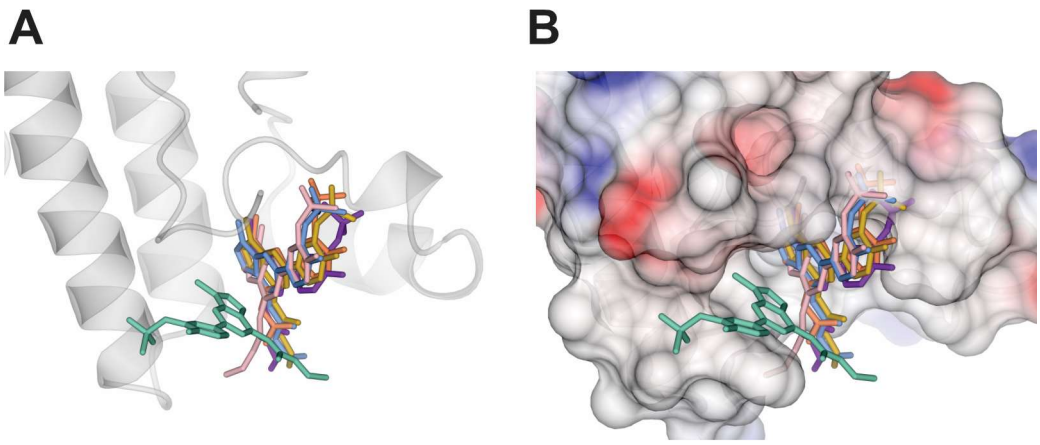


Figure 5.13. Bromosporine binding mode in different bromodomains. Overlay of co-crystal structures showing bromosporine binding to *Ld*BDF5.2 (PDB code 5TCK, green); *Ld*BDF2 (PDB code 5C4Q, orange); *Ld*BDF3 (PDB code 5FEA, pink), human BRD4 (PDB code 5IGK, purple); human BRD7 (PDB code 6V1H, blue); and human BRD9 (PDB code 5IGM, yellow). Protein structure of *Ld*BDF5.2 is shown as **(A)** ribbon representation and **(B)** electrostatic surface representation. Image created using CCP4mg software.

6. Chapter 6: Identification and Characterisation of Novel *Leishmania* Bromodomain Inhibitor Compounds

6.1. Introduction

The work described thus far has provided the foundation for the investigation of novel compounds that inhibit the *Leishmania* bromodomain proteins. Protein target-validation and optimisation of biophysical assays were essential steps within the overarching antileishmanial drug discovery aim, and the work described in this chapter covers the next stage in the pipeline. Two main approaches were employed to identify and develop hit compounds; high-throughput target-based screens and structure activity relationship (SAR) optimisation.

One frequently adopted strategy in drug discovery is the design of inhibitors that mimic the native substrate of the protein. Small molecules such as JQ1 and I-BET which contain acetyllysine mimics have been shown to successfully bind to BET bromodomains to disrupt their histone-binding abilities and related functions (Filippakopoulos *et al.*, 2010; Nicodeme *et al.*, 2010). The 3,5-dimethylisoxazole moiety (Figure 5.2) has been identified as an acetyllysine bioisostere, and previous optimisation of compounds containing this group and its derivatives has proved a useful strategy for the development of potent and selective BET ligands (Hewings *et al.*, 2011, 2013). Bromodomain inhibitors I-BET151 and SGC-CBP30 both contain this moiety, and the work described in Chapter 5 highlighted SGC-CBP30 as a ligand of *Ld*BDF5 with activity against *Leishmania* promastigotes. In a collaboration with Jennifer Carter and Professor Stuart Conway at the University of Oxford, this approach was adopted for the design of compounds targeting *Ld*BDF5.

Another common approach within discovery is the use of high-throughput screens of large compound libraries to identify ligands of a target protein. Affinity selection mass spectrometry (ASMS) is one such screening technique in which a protein of interest is incubated with one or more compounds from a library, followed by isolation of any protein-ligand complexes and identification of the binding ligands by mass spectrometry (Annis *et al.*, 2007). ASMS screening has proved a successful and rapid route for ligand discovery, for example in one study, this approach yielded

382 target-selective hits from a library of 2 million compounds screened against proteins associated with the NF-κB pathway (Kutilek *et al.*, 2016). This screening strategy was also employed for the discovery of *Ld*BDF2, 5.1 and 5.2 ligands through a collaboration with GSK.

Whilst the above ligand identification strategies concern reversibly binding compounds, ligands which bind covalently to target proteins also provide useful tools for studying protein function and can also be used as therapeutics. Covalent compounds offer stronger, irreversible binding and can be optimised to generate highly potent inhibitors. These compounds contain a reactive functional group or 'warhead' which reacts with the protein to form a covalent bond. Sulfur (VI) fluoride groups are emerging as promising warheads which can target several different nucleophilic amino acids (Lonsdale and Ward, 2018). The application of this electrophilic group in compounds for the covalent modification of proteins has been explored in previous research, highlighting the potential of their use within covalent ligand design and fragment-based screening platforms (Gilbert *et al.*, 2022). Covalent reactive fragments containing a sulfonyl fluoride (SO_2F) warhead were also investigated as potential ligands of the *Ld*BDF bromodomains in collaboration with GSK.

Herein is described the application of the previously defined biophysical assays using recombinant *Ld*BDF proteins, and phenotypic screens using *Leishmania* promastigotes, for the identification and characterisation of novel *Leishmania* bromodomain inhibitors as part of these drug-discovery routes. This work formed part of a collaborative effort with the University of Oxford and GSK to develop new inhibitors of the *Leishmania* bromodomains.

6.2. Acetyllysine mimic ligands of *Ld*BDF5.1

6.2.1. TSA screen of 31 acetyllysine mimic fragments

A library of 31 acetyllysine mimic fragments (133 – 326 Da) were synthesised by Jennifer Carter from the Conway group (University of Oxford). The structures of these compounds are given in the Appendix G. Several contain the previously discussed 3,5-dimethylisoxazole moiety also present in SGC-CBP30. These compounds were supplied and screened in two separate batches using the thermal shift assay using the recombinant *Ld*BDF bromodomain proteins to identify compounds that showed binding.

The first 11 compounds (Table G.1) were screened against *Ld*BDF2 and 5.2 at 25 and 150 µM using the optimised TSA conditions (Appendix Table E.1). At these concentrations, no positive thermal shifts were observed (data not shown). The second batch of 20 library compounds (Table G.2) was screened against *Ld*BDF5T by TSA at the higher concentration of 400 µM. Whilst high ligand concentrations in TSA can lead to non-specific binding, here the low molecular weight nature of the fragment compounds meant any binding interactions were likely to be low affinity and a higher concentration gave a better chance of detection. Note JC-A12 and JC-A15 were also included in this screen along with the 20 new compounds. Optimised protein and dye concentrations were used as before. Data for AS-A7 are omitted as this compound exhibited auto-fluorescence which interfered with the assay. Thermal shifts were negative for all other compounds with the exception of JC-B4, JC-B87 and JC-B93 (Figure 6.1). The largest and most significant thermal shift was observed for JC-B4 (ΔT_m = 0.60 ± 0.06 °C; unpaired t-test $P < 0.0001$, n = 6). Orthogonal TSA, isothermal titration calorimetry (ITC) and microscale thermophoresis (MST) screens performed by Jennifer Carter at the University of Oxford corroborated the identification of JC-B4 as a ligand of *Ld*BDF5, and indicated that it specifically binds the first bromodomain of the protein.

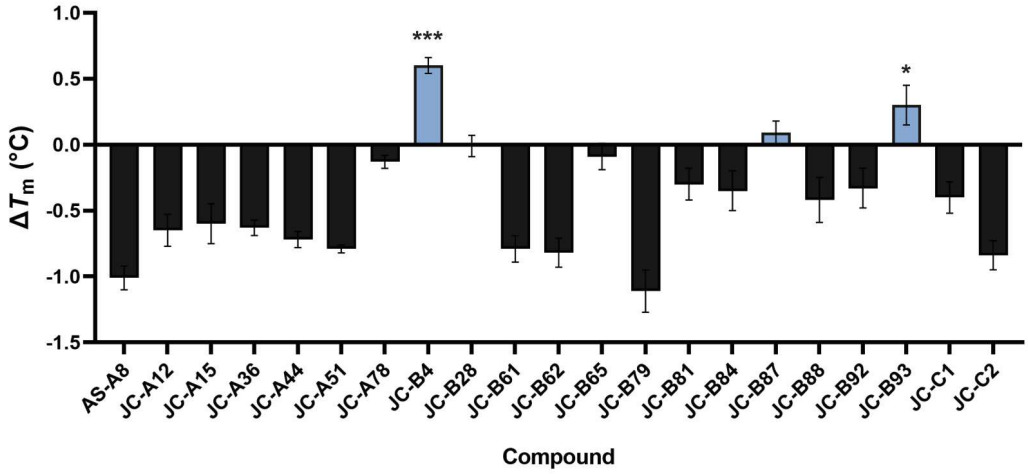


Figure 6.1. Thermal shifts of *Ld*BDF5T incubated with acetyllysine mimic fragments at 400 µM. ΔT_m calculated as the difference between the average melting temperature compared with deuterated DMSO reference samples, with positive shifts in blue. Error bars represent standard deviation (SD) of the replicate samples (n ≥ 4). P values were calculated for positive shifts using an unpaired t-test, * P < 0.05, ** P < 0.01, *** P < 0.001 vs protein + DMSO sample, as indicated.

6.2.2. Validation of JC-B4 hit by NMR

Ligand-observed NMR was used to validate the interaction of JC-B4 with *Ld*BDF5.1. As described in Chapter 5 (Section 5.3.2), three different NMR experiments were applied; waterLOGSY, STD and CPMG. Alongside JC-B4, the compound JC-A15 (Figure 6.2) was included as a negative control as this showed no indication of binding to *Ld*BDF5 in TSA.

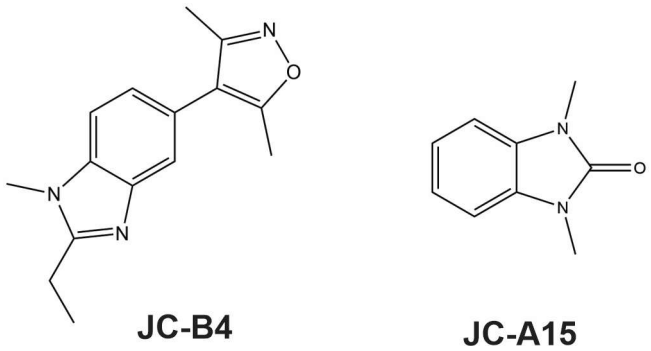


Figure 6.2. Structures of JC-B4 and JC-A15 compounds used in NMR. Hit compound JC-B4 and negative control JC-A15 screened for binding to *Ld*BDF5.1 in NMR experiments.

Spectra are shown for JC-B4 with representative signal peaks of the 3,5-dimethylisoxazole group at 2.27 and 2.42 ppm highlighted (Figure 6.3, Figure 6.4, Figure 6.5). In the waterLOGSY experiment, the non-binding reference compound DSS gave a negative signal peak at 0 ppm in both JC-B4 and JC-A15 samples and therefore positive compound peaks were required to signify a binding interaction. JC-B4 showed negative peaks alone but predominantly positive peaks in the presence of *Ld*BDF5.1, particularly clear for the 3,5-dimethylisoxazole resonance (Figure 6.3). In comparison, JC-A15 showed negative signal peaks both in the absence and presence of the protein (data for JC-A15 not shown).

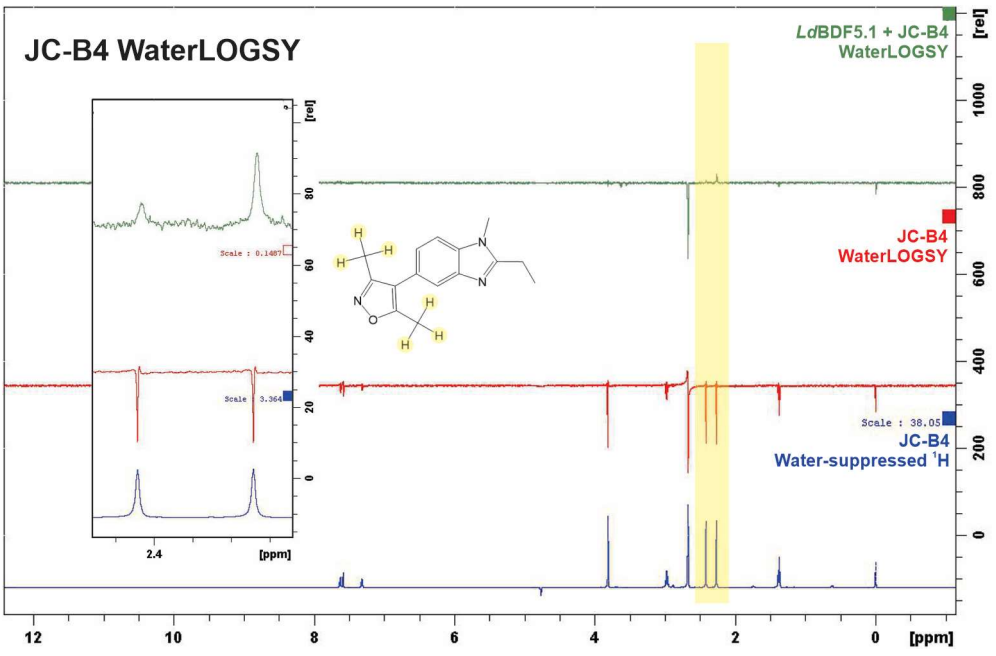


Figure 6.3. WaterLOGSY NMR spectra showing binding of JC-B4 to *Ld*BDF5.1. Spectra recorded for the compound alone (red) and with *Ld*BDF5.1 (green) alongside the reference water-suppressed 1H spectrum for JC-B4 alone (blue). Representative ligand peaks are highlighted and shown as an inset, where these correspond to the methyl groups highlighted in the JC-B4 compound structure.

In the STD experiment, JC-B4 compound peaks were present in the 'difference' spectrum for the sample with *Ld*BDF5.1, but not in the absence of the protein (Figure 6.4). As previously discussed, comparison of the STD signal peak intensities can give suggestions as to the binding epitope. Here, the resonances of the 3,5-dimethylisoxazole group were markedly stronger than other peaks which indicates this group is involved in the binding of JC-B4 to *Ld*BDF5.1. The JC-B4 compound shares the same core group as SGC-CBP30 which was also found to interact with *Ld*BDF5.1 via the 3,5-dimathylisoxazole acetyllysine mimic moiety. JC-A15 gave very weak signal peaks of similar intensity in both the samples with and without *Ld*BDF5.1, indicating no binding to the protein.

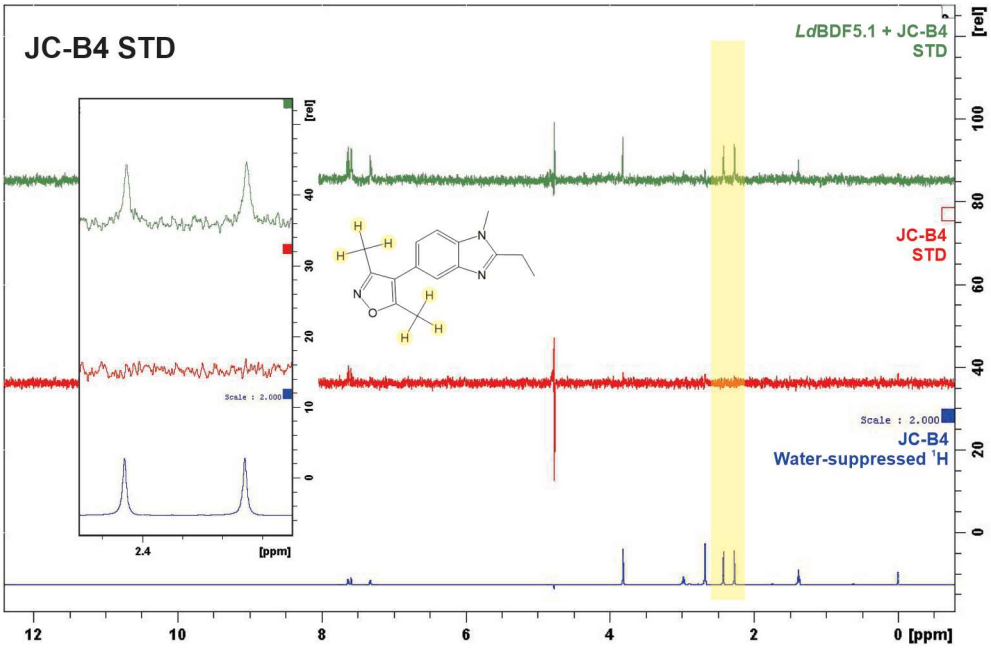


Figure 6.4. STD NMR spectra showing binding of JC-B4 to *Ld*BDF5.1. Spectra recorded for JC-B4 alone (red) and with *Ld*BDF5.1 (green) alongside the reference water-suppressed 1H spectrum for the compound alone (blue). Representative ligand peaks are highlighted and shown as an inset, where these correspond to the methyl groups highlighted in the JC-B4 structure.

The CMPG experiment uses different relaxation delays to exploit the different relaxation rates of nuclei in bound and un-bound ligands, where a drop in peak intensity with increased relaxation delay indicates binding. In agreement with the waterLOGSY and STD experiments, JC-B4 showed binding to *Ld*BDF5.1 as peaks displayed a greater drop in intensity in the sample with *Ld*BDF5.1 compared with the compound alone sample (Figure 6.5). The negative control JC-A15 again did not show binding to *Ld*BDF5.1, where peaks showed similar intensities using both relaxation delays, both in presence and absence of the protein.

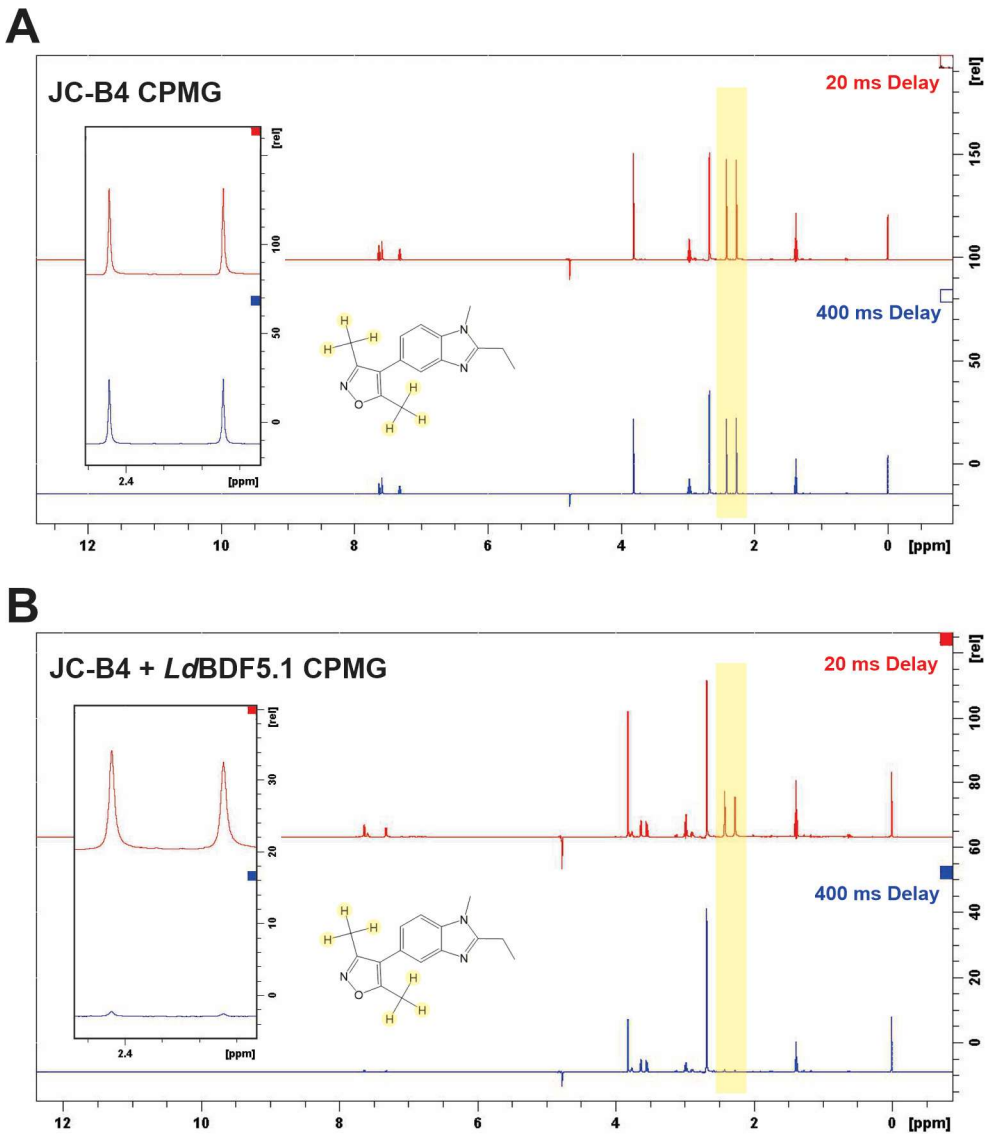


Figure 6.5. CPMG NMR spectra showing binding of JC-B4 to *Ld*BDF5.1. Spectra for **(A)** JC-B4 alone and **(B)** JC-B4 with *Ld*BDF5.1, comparing long 400 ms (blue) and short 20 ms (red) relaxation delays. Representative ligand peaks are highlighted and shown as an inset, which correspond to the methyl groups highlighted in the JC-B4 structure.

6.2.3. Investigating JC-B4 analogues

Following the identification of JC-B4 as a hit compound binding to *Ld*BDF5.1, analogues of this compound were synthesised by Jennifer Carter at the University of Oxford. The JC-B4 compound constitutes the core group of the human bromodomain inhibitor SGC-CBP30, also shown to bind *Ld*BDF5.1 (Chapter 5). In light of this, JC-B4 analogues were designed with different groups substituted at the R^1 and R^2 positions (Figure 6.6).

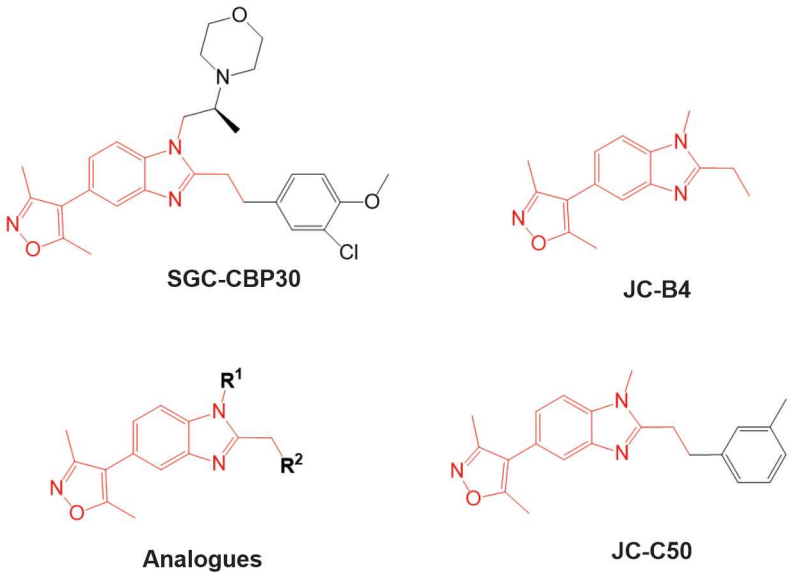


Figure 6.6. JC-B4 hit compound and analogues sharing the SGC-CBP30 core. Structure of the bromodomain inhibitor SGC-CBP30 with core group highlighted corresponding to the compound JC-B4 and was further developed into analogues with different groups at the R^1 and R^2 positions including JC-C50.

The first analogue studied was JC-C50 (Figure 6.6) which had shown improved affinity over JC-B4 in assays at the University of Oxford. This compound was designed with a pendant 3-methylbenzene substitution at the R^2 position to increase interactions of the compound with the acetyllysine binding pocket of *Ld*BDF5.1. The optimised FP assay was used to screen both JC-C50 and JC-B4 alongside SGC-CBP30 and an unmodified negative control peptide (H2B-15mer-N) against *Ld*BDF5.1, with displacement of the H2B-7mer-Ac-F probe measured to calculate IC_{50} values.

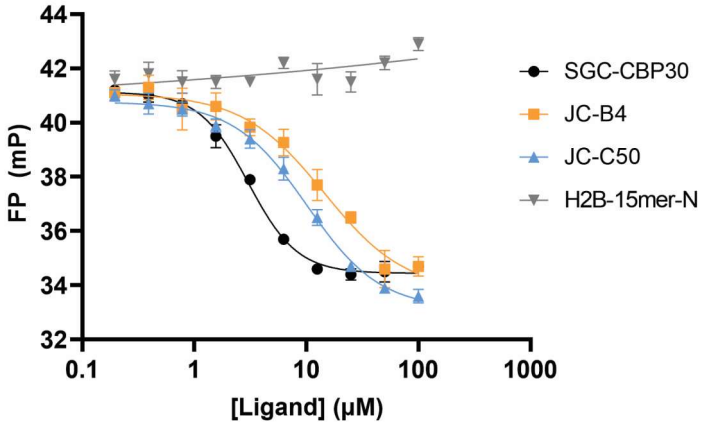


Figure 6.7. FP displacement curves for competition of JC-B4 and JC-C50 with H2B-7mer-Ac-F probe for *Ld*BDF5.1 binding. FP response for the binding of SGC-CBP30, JC-B4 and JC-C50 to *Ld*BDF5.1, compared with negative control competitor peptide H2B-15mer-N. Average, blank-corrected FP values plotted against ligand concentration, and data fitted with [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; error bars representing SD (n = 3).

FP confirmed binding of both JC-B4 and JC-C50 to *Ld*BDF5.1 (Figure 6.7) and all three compounds gave micromolar IC_{50} values (Table 6.1). SGC-CBP gave the lowest IC_{50} value which correlated well with previous FP data (Table 5.1). The two new acetyllysine compounds both showed similar potencies with the JC-C50 analogue exhibiting a slight improvement over the original JC-B4 hit.

Table 6.1. IC_{50} values for the binding of SGC-CBP30, JC-B4 and JC-C50 to *Ld*BDF5.1 as determined by FP. Values calculated by plotting average blank-corrected mP values against [inhibitor] and fitting [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; reported as IC_{50} ± standard error (n = 3) with R^2 for model fitting.

Compound	IC_{50} (µM)	R^2
SGC-CBP30	3.02 ± 0.19	0.984
JC-B4	14.4 ± 3.0	0.953
JC-C50	10.1 ± 0.83	0.988

Ligand-observed NMR experiments were also performed in an attempt to validate the binding of JC-C50 to *Ld*BDF5.1 and investigate the binding epitope, however despite trialling several protein and compound concentrations, these experiments were unsuccessful. This may indicate that the JC-C50 binding affinity is too strong to be detected by waterLOGSY, STD or CPMG which rely on fast protein-ligand exchange and are known to show false negatives for high affinity binders.

A further eight SAR-optimised JC-B4 analogues (Table G.3) which superseded JC-C50 were subsequently acquired and screened in the FP assay with *Ld*BDF5.1 alongside JC-C50. These compounds all showed clear displacement curves confirming binding to *Ld*BDF5.1 (Figure 6.8). The new analogues also showed improved potency over JC-C50 (IC_{50} = 16.6 ± 3.5), with IC_{50} values ranging from 3.61 to 7.00 µM (Table 6.2).

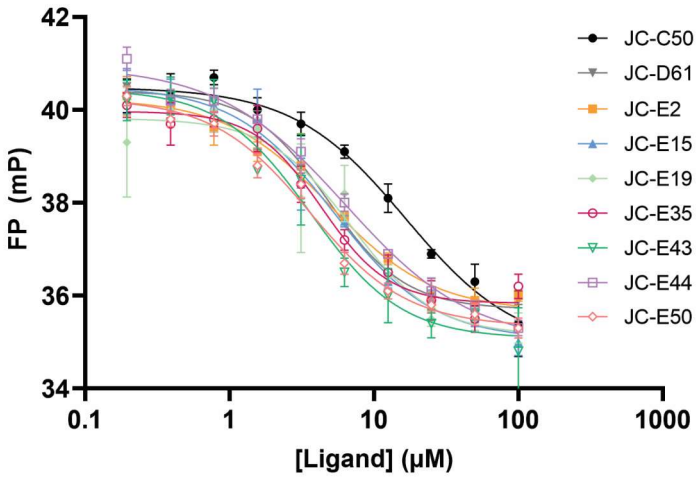


Figure 6.8. FP competition curves showing displacement of H2B-7mer-Ac-F probe by JC-B4 analogues for *Ld*BDF5.1 binding. FP response for the binding of nine JC-B4 analogues to *Ld*BDF5.1. Average, blank-corrected FP values plotted against ligand concentration, and data fitted with [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; error bars representing SD (n = 3).

Table 6.2. IC_{50} values for the binding of JC-B4 analogues to *Ld*BDF5.1 as determined by FP. Values calculated by plotting average blank-corrected mP values against inhibitor concentration and fitting [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; reported as IC_{50} ± standard error (n = 3) with R^2 for model fitting. Heat map display shows 'hot' colours corresponding to lower IC_{50} values.

Compound	IC_{50} (µM)	Compound	IC_{50} (µM)
JC-C50	16.6 ± 3.5 (R^2 = 0.967)	JC-E35	4.32 ± 0.50 (R^2 = 0.957)
JC-D61	4.79 ± 0.37 (R^2 = 0.983)	JC-E43	3.67 ± 0.65 (R^2 = 0.939)
JC-E2	4.92 ± 0.69 (R^2 = 0.968)	JC-E44	6.34 ± 0.77 (R^2 = 0.985)
JC-E15	5.30 ± 0.64 (R^2 = 0.974)	JC-E50	3.61 ± 0.51 (R^2 = 0.967)
JC-E19	7.00 ± 1.5 (R^2 = 0.902)		

However, at this point a limitation of the FP assay began to emerge. Whilst the assay was able to confirm binding interactions, it became apparent that the relatively low affinity of the probe (K_D = 77.5 ± 0.95 µM) limited the sensitivity of the assay and that these high affinity SAR-optimised compounds may have true IC_{50} values below the detection limit of the assay. Affinities calculated for same compounds in ITC and MST screens at the University of Oxford suggested they bind with higher and more varied affinities than those inferred from the FP IC_{50} assays; all showed improved potency compared with SGC-CBP30. Therefore, for the screening of high affinity compounds, the FP assay may prove more useful as a qualitative indicator of binding and IC_{50} values should be interpreted with this in mind. Future work may look to develop a new, higher binding affinity probe based

on a small molecule hit compound for further optimisation of the FP assay to characterise *Ld*BDF ligands.

6.2.4. Top four acetyllysine mimics hits binding to *Ld*BDF5 tandem bromodomains

Based on both the FP results and binding affinities determined in simultaneous ITC and MST screens conducted at the University of Oxford, four compounds were selected as the 'top hits' from the library of JC-B4 analogues (Figure 6.9). Three of these compounds have a methyl group at the R^1 position, and one, JC-E15, has an *N*-ethylmorpholine group at this position. All compounds contained a benzene ring at the R2 position with various different substitutions. Notably, the benzene rings in both JC-E35 and JC-E43 contain nitro group substituents.

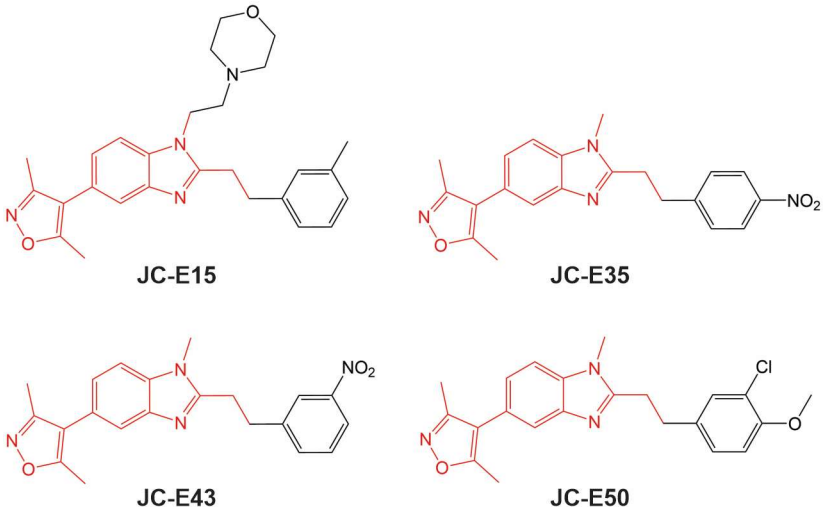


Figure 6.9. Structures of the top four hit acetyllysine mimics. JC-B4 analogues that showed greatest binding to *Ld*BDF5.1 in biophysical assays selected for further characterisation and inclusion in phenotypic assays.

Whilst FP and NMR experiments had established binding of the acetyllysine mimics to the recombinant *Ld*BDF5.1 bromodomain protein, it was also important to consider whether they would also bind the native full-length protein. TSA was employed to begin this assessment by screening the top hits against the recombinant tandem bromodomain protein, *Ld*BDF5T alongside *Ld*BDF5.1. The four top hits alongside the JC-C50 precursor compound, the original JC-B4 lead fragment and SGC-CBP30 were screened at 50 and 250 µM against *Ld*BDF5T and *Ld*BDF5.1 (3 µM) with 3x sypro orange dye. As anticipated, at the higher

concentration of 250 µM, compounds showed large negative thermal shifts. However, at 50 µM positive shifts were seen for all compounds with both proteins (Figure 6.10). In correlation with the FP assay, the top four analogues produced larger thermal shifts for *Ld*BDF5.1 than both JC-B4 and JC-C50 (Figure 6.10A). The compounds JC-E43 and JC-E50 also exhibited higher thermal shifts than SGC-CBP30, with the highest value for JC-E50 (ΔT_m = 3.74 ± 0.22 °C). Whilst *Ld*BDF5T gave smaller and a narrower range of thermal shifts, it was pleasing to see that the compounds bound to the tandem bromodomain protein, not just *Ld*BDF5.1 (Figure 6.10B). As with *Ld*BDF5.1, the top two thermal shifts were obtained with JC-E43 (ΔT_m = 1.58 ± 0.29 °C) and JC-E50 (ΔT_m = 1.39 ± 0.10 °C). This correlates with the FP data, in which JC-E43 and JC-E50 yielded the lowest IC_{50} values.

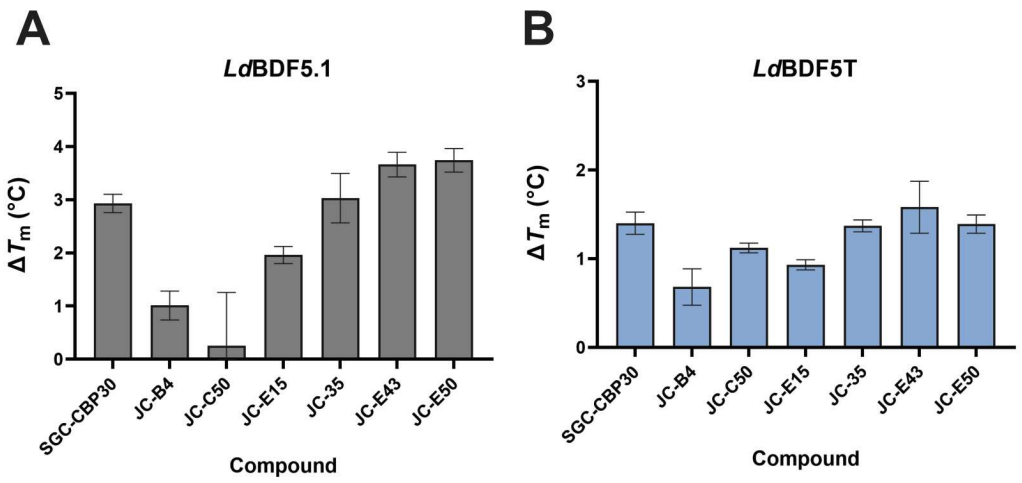


Figure 6.10. Thermal shifts of *Ld*BDF5.1 and 5T incubated with the top four hit acetyllysine mimic fragments at 50 µM. ΔT_m of **(A)** *Ld*BDF5.1 and **(B)** *Ld*BDF5T calculated as the difference between the average melting temperature compared with deuterated DMSO reference samples, with positive shifts in blue. Error bars represent standard deviation (SD) of the replicate samples (n = 6)

6.2.5. Co-crystallisation screens for acetyllysine mimics with *Ld*BDF5.1

Attempts were made to co-crystallise *Ld*BDF5.1 with several acetyllysine mimic lead compounds generated throughout the pipeline. These were JC-B4, JC-C50, JC-E43 and JC-E65 (Table G.3), a new SAR-optimised compound most recently obtained which has exceeded prior iterations as the highest affinity compound in assays performed by Jennifer Carter at the University of Oxford. This compound shares the same structure as JC-E43 with bioisosteric replacement of the NO_2 with a CN. Crystallisation conditions varied in the screens included *Ld*BDF5.1 concentration (18 – 24 mg/ml), compound concentration (1.5 – 6.0 mM) and drop ratio, within different commercial screening plates (full screening conditions are listed in Appendix F).

Crystal growth was observed in the JC-B4 co-crystal screens using a JCSG+ screen with 19 mg/ml *Ld*BDF5.1 and 2.5 mM JC-B4 using a 150:150 nl drop ratio. The most promising crystals were seen in three wells, all of which contained zinc acetate dihydrate (0.2 M) salt. Though small, these crystals were fished and tested by X-ray diffraction. Only one crystal produced diffraction, and data were collected, however data analysis revealed the cell dimensions were too small to be the bromodomain protein and was more likely the JC-B4 compound alone. Optimisation of the screening conditions was unable to generate larger crystals.

Co-crystallisation attempts with JC-C50 yielded crystals in a JCSG+ screen using 150:150 nl drops with 20 mg/ml *Ld*BDF5.1 and JC-C50 at an excess of 4.5x. Crystals that grew in wells D6 (0.2 M magnesium chloride hexahydrate, 0.1 M Tris, 20% w/v PEG 8000, pH 8.5) and E8 (1 M ammonium phosphate dibasic, 0.1 M sodium acetate, pH 4.5) showed promising levels of polarisation under a polarised light filter (Figure 6.11). These crystals were fished, but unfortunately did not produce X-ray diffraction. Optimisation using varied JC-C50 concentration and drop ratios did not yield any promising protein crystals.

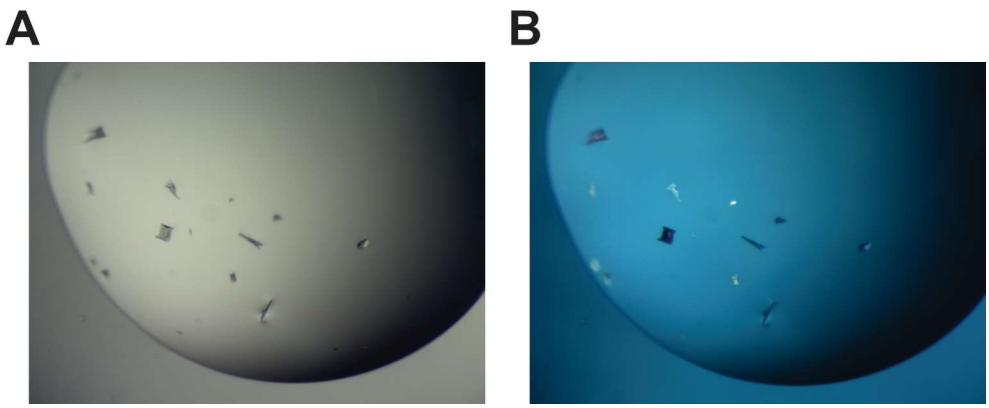


Figure 6.11. Co-crystals of *Ld*BDF5.1 with JC-C50. JCSG+ screen hit well D6 images under **(A)** white light and **(B)** using a polarised light filter, several of which were fished for X-ray diffraction along with crystals from well E8.

In the first round of JC-E43 co-crystallisation screens, small crystals formed in lots of wells of JCSG+ and Morpheus screens which used 2-methyl-2,4-pentanediol (MPD) as the precipitant. This guided a second round of screening using the commercial MPD screen, alongside a second Morpheus screen with different drop ratios and larger drop volumes up to 1 µl, with *Ld*BDF5.1 at 20 – 24 mg/ml and JC-E43 at an excess of 1.5x. The new screens also yielded crystals in numerous wells, but these did not show improved size over the initial screens. Nevertheless, a total of 11 small crystals were fished from the Morpheus and MPD screens but unfortunately, none of these crystals diffracted. No crystals grew in Morpheus or MPD co-crystallisation screens for *Ld*BDF5.1 with JC-E65.

6.3. Ligands of *Ld*BDF5.2 identified in high-throughput screens

Using *Ld*BDF2, 5.1 and 5.2 recombinant proteins produced by Rebecca Preece in the University of York Technology Facility, high-throughput screens (ASMS, sulfonyl fluoride screen and SO_2F displacement screens, described below) were performed and follow-up analysis of hits conducted by Jacob Bush, Arron Aatkar and Emma Grant at GSK Stevenage, UK; and Raquel Gabarro-Carrion and Félix Calderón at GSK Tres Cantos, Spain.

6.3.1.1. Affinity selection mass spectrometry screen

A library of around 175,000 compounds was screened against recombinant *Ld*BDF2, 5.1 and 5.2 bromodomain proteins using the ASMS method (Section 6.1). Proteins were incubated with compounds and complexes separated from the mixture by size exclusion chromatography or spin column whereby unbound ligands are retained on the column whilst complexes flow through. Proteins and bound ligands were then separated by denaturation and liquid chromatography mass spectrometry (LC-MS) applied to discern the identity of the bound compounds. Both the screen and subsequent analysis was performed by the aforementioned collaborators at GSK. 181 hits were identified and further refined to six compounds provided for validation and characterisation. The six hits included one *Ld*BDF2 hit (GSK2A), four *Ld*BDF5.1 hits (GSK5.1A, B, C, D) and one *Ld*BDF5.2 hit (GSK5.2A). Validation of the *Ld*BDF5 hits was performed using TSA.

All six compounds were screened against *Ld*BDF5.2 at 100 µM using optimised protein and dye concentrations (Appendix Table E.1). GSK5.2A which was identified as a hit for *Ld*BDF5.2 produced a thermal shift of 1.30 ± 0.20 °C, whilst all *Ld*BDF5.1 hits gave negative shifts (Figure 6.12A & C). Interestingly, GSK2A also showed a small positive shift of 0.49 ± 0.21 °C. The five *Ld*BDF5 ASMS hits were then screened against *Ld*BDF5T at the same concentration using the optimised protein and dye concentrations. GSK5.2A showed a similar shift of 1.66 ± 0.05 °C, and all four *Ld*BDF5.1 hits also gave positive shifts of 0.52 – 0.76 °C (Figure 6.12B, & D). These screens validated the ASMS hits, confirmed selectivity and also showed that the *Ld*BDF5.1 and 5.2 hit compounds were able to bind both the single and tandem bromodomain recombinant proteins.

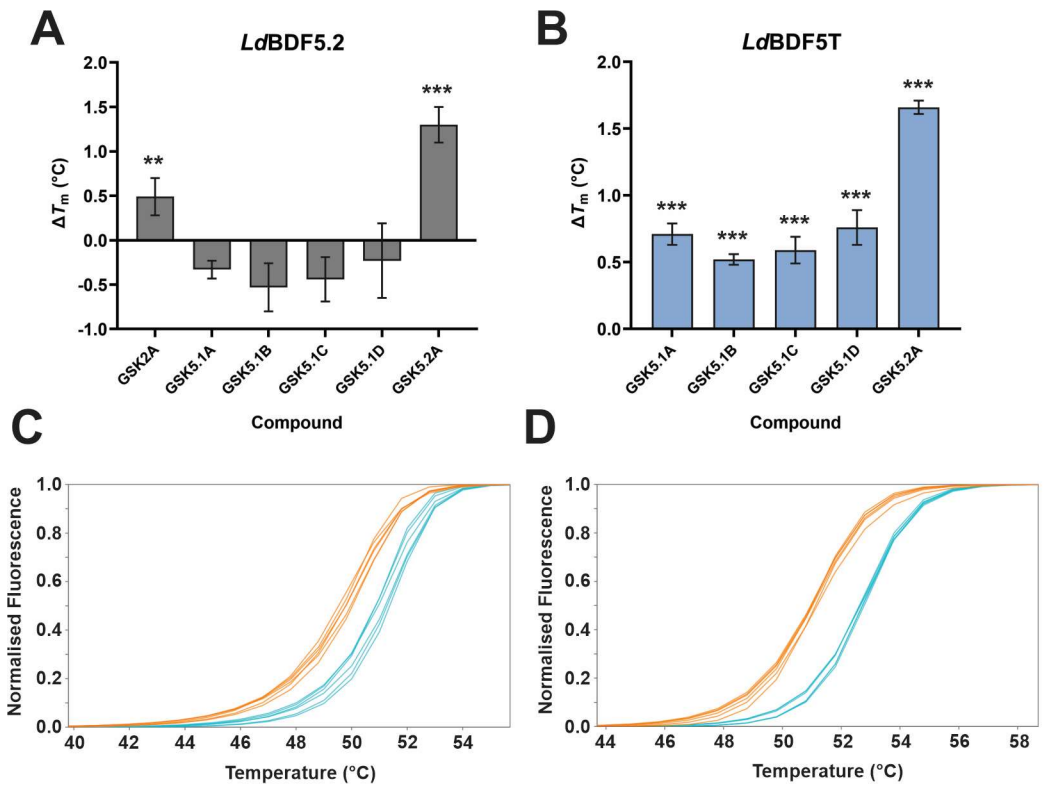


Figure 6.12. TSA screens for ASMS hits with *Ld*BDF5.2 and *Ld*BDF5T. ΔT_m values calculated as the difference between the average melting temperature of **(A)** *Ld*BDF5.2 or **(B)** *Ld*BDF5T with different ASMS hit compounds (100 µM) compared with DMSO reference samples. Error bars represent standard deviation (SD) of the replicate samples (n ≥ 5). P values were calculated for positive shifts using an unpaired t-test, * P < 0.05, ** P < 0.01, *** P < 0.001 vs respective protein + DMSO sample, as indicated. Melting curves for GSK5.2A shown for **(C)** *Ld*BDF5.1 and **(D)** *Ld*BDF5T, showing normalised curves for five parameter sigmoid equation model fitting for each of the six replicates of each sample of protein with GSK5.2A (blue) alongside DMSO control samples (orange); produced using the online JTSA tool.

The ability to selectively target one bromodomain in a tandem bromodomain protein has been highly sought after in bromodomain inhibitor research, particularly in the field of human BET bromodomains (Faivre *et al.*, 2020; Gilan *et al.*, 2020). Here, *Ld*BDF5 compounds GSK5.1A, B, C and D have been identified which show selectively for the first bromodomain of *Ld*BDF5 over the second, as the compounds elicited thermal shifts for the tandem *Ld*BDF5T bromodomain protein but not the *Ld*BDF5.1 bromodomain. Binding of GSK5.2A to *Ld*BDF5.2 was also confirmed using the FP assay, where the compound caused displacement of the H2B-7mer-Ac-F probe and IC_{50} was calculated as 8.72 ± 0.87 µM (R^2 = 0.980) (Figure 6.13).

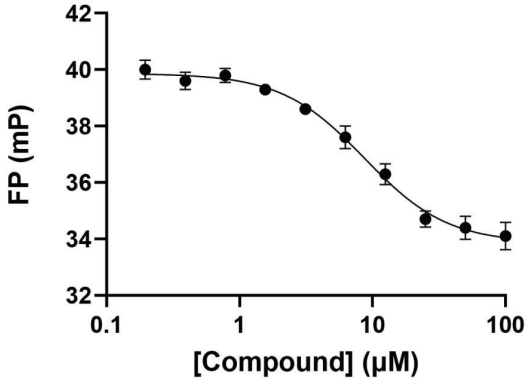


Figure 6.13. FP displacement curve for competition of GSK5.2A with H2B-7mer-Ac-F probe for *Ld*BDF5.2 binding. Binding of GSK5.2A to *Ld*BDF5.2, with blank-corrected FP values plotted against ligand concentration, and data fitted with [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; error bars representing SD (n = 3).

6.3.2. Sulfonyl fluoride reactive fragment screen

As discussed above, covalent compounds hold potential both as tool compounds and within drug discovery. Sulfonyl fluoride groups (Figure 6.14A) can be used as warheads which react with multiple nucleophilic amino acid residues (Gilbert *et al.*, 2022). This reaction consists of an initial reversible association between the protein and the reactive fragment, followed by covalent bond formation which involves cleavage of the S-F bond and release of HF (Figure 6.14B).

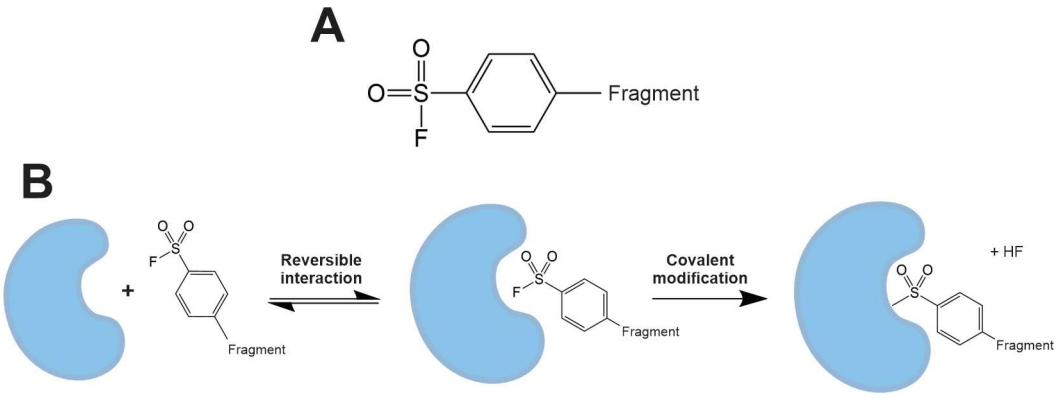


Figure 6.14. Sulfonyl fluoride reactive fragments and their covalent protein binding. (A) Structure of the sulfonyl fluoride fragments screened against *Ld*BDF bromodomains. **(B)** Covalent modification of proteins by SO_2F reactive fragments involving a reversible association followed by covalent binding and release of HF.

A library of 175 commercially-available reactive fragments (Enamine) containing the SO_2F warhead were screened against *Ld*BDF2, 5.1 and 5.2 recombinant proteins by Arron Aatkar and Jacob Bush at GSK Stevenage, UK. Fragments were incubated with the three recombinant *Ld*BDF proteins and the human BRD4 BD1 protein as a control for selectivity. Following an incubation period of 24 hours at 4°C, intact protein LC-MS was used to identify proteins which had undergone covalent modification, exhibiting a change in molecular weight corresponding with the addition of the fragment minus the loss of HF. *Ld*BDF5.2 yielded the highest number of hits, followed by *Ld*BDF5.1, with very few hits identified for *Ld*BDF2. Hits were refined to 15 compounds of interest, including two compounds, named SO2FA and SO2FB, which showed high levels of crosslinking to both *Ld*BDF5 bromodomains. Follow-up analysis by Arron Aatkar at GSK established that the compounds were reacting with tyrosine residues Y37 and Y96 proximal to the binding pocket in *Ld*BDF5.1.

These two hit compounds were purchased alongside a negative control sulfonyl fluoride fragment, SO2FNeg1, which showed no binding to *Ld*BDF5.1 or 5.2 in the screen. To validate the binding of SO2FA and SO2FB, the three compounds were screened at 100 µM against *Ld*BDF5.2 and 5T in the TSA assay using optimised protein and dye concentrations (Table E.1).

TSA data confirmed that SO2FA and SO2FB bound to both *Ld*BDF5.2 and *Ld*BDF5T (ΔT_m values 1.31 – 2.96 °C), whereas the negative control compound did not bind either protein (Figure 6.15A). Both hit compounds gave similar thermal shifts when screened against each protein, where the largest thermal shift was elicited by SO2FB with *Ld*BDF5T (Figure 6.15B).

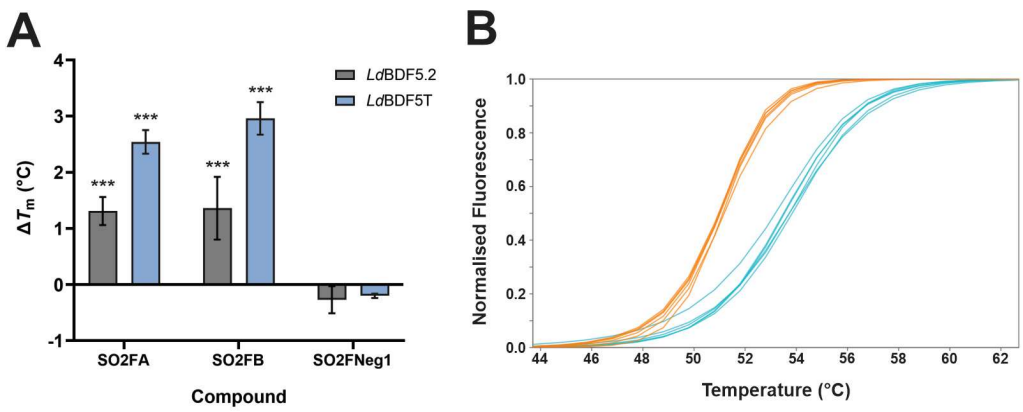


Figure 6.15. TSA screens for SO_2F reactive fragment hits with *Ld*BDF5.2 and *Ld*BDF5T. **(A)** ΔT_m values calculated as the difference between the average melting temperature of *Ld*BDF5.2 and *Ld*BDF5T with the three SO_2F compounds (100 µM) compared with DMSO reference samples. Error bars represent standard deviation (SD) of the replicate samples ($n \geq 4$). P values were calculated using an unpaired t-test, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ vs respective protein + DMSO sample as indicated. **(B)** Representative melting curves for SO2FB with *Ld*BDF5T showing normalised curves for five parameter sigmoid equation model fitting for the six replicates of *Ld*BDF5T with SO2FB (blue) alongside DMSO control samples (orange); produced using the online JTSA tool.

6.3.3. Displacement screen using a sulfonyl fluoride tool compound

A new sulfonyl fluoride compound was subsequently identified which superseded SO2FA and SO2FB. The new reactive fragment, SO2FC, was utilised as a tool compound within a new high-throughput displacement screen to identify higher affinity reversible ligands of *Ld*BDF5.2. The screen, performed by GSK, investigated a subset of compounds from different bromodomain ASMS screens. Compounds within this library were incubated with the recombinant *Ld*BDF5.2 before addition of the SO2FC tool compound which competed for binding and produced covalent modification of *Ld*BDF5.2 proportional to the affinity of the competing ASMS compound. As before, LC-MS analysis was performed and competition assessed by measuring reduction in SO2FC binding in comparison with a control sample with no ASMS compound. Thirteen compounds were identified with potentially improved affinity for *Ld*BDF5.2 compared with the original ASMS screen hits. These 13 compounds (named GSK5.2B – GSK5.2N) were provided for further

characterisation along with SO2FC, three negative control reversible compounds (GSK5.2Neg1, 2, 3) and a second negative control sulfonyl fluoride (SO2FNeg2).

The FP assay was used to validate and further characterise binding of the new ASMS hits to *Ld*BDF5.2. All 13 compounds were screened against *Ld*BDF5.2, alongside negative controls and the original ASMS hit (GSK5.2A), with displacement of the H2B-7mer-Ac-F probe measured to quantify binding affinity.

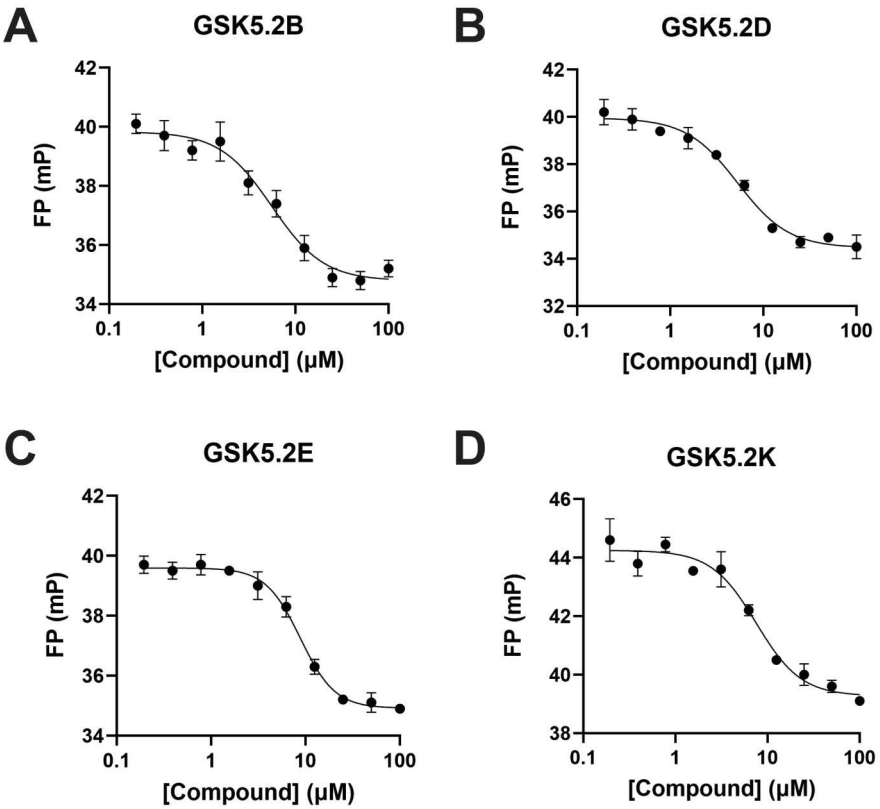


Figure 6.16. FP competition curves for binding of new ASMS hits to *Ld*BDF5.2. FP response for the binding of **(A)** GSK5.2B, **(B)** GSK5.2D, **(C)** GSK5.2E and **(D)** GSK5.2K to *Ld*BDF5.2 with average, blank-corrected FP values plotted against ligand concentration, and data fitted with [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; error bars representing SD (n = 3).

Table 6.3. IC_{50} values for binding of new ASMS compounds to *Ld*BDF5.2 as determined by FP. Values calculated by plotting average blank-corrected mP values against inhibitor concentration and fitting [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression, reporting IC_{50} ± standard error, with R^2 for model fitting. Heat map display shows 'hot' colours corresponding to lower IC_{50} values.

Compound	IC_{50} (µM)	Compound	IC_{50} (µM)
GSK5.2A	9.44 ± 1.4 (R^2 = 0.959)	GSK5.2F	10.1 ± 1.0 (R^2 = 0.979)
GSK5.2B	5.58 ± 1.07 (R^2 = 0.911)	GSK5.2G	16.5 ± 1.5 (R^2 = 0.974)
GSK5.2C	10.4 ± 0.70 (R^2 = 0.976)	GSK5.2H	14.9 ± 6.6 (R^2 = 0.957)
GSK5.2D	5.30 ± 0.49 (R^2 = 0.977)	GSK5.2J	11.9 ± 0.90 (R^2 = 0.969)
GSK5.2E	8.82 ± 0.60 (R^2 = 0.977)	GSK5.2K	7.55 ± 0.85 (R^2 = 0.960)
		GSK5.2M	11.1 ± 1.7 (R^2 = 0.944)

The three negative controls showed no clear binding to *Ld*BDF5.2. All hit compounds caused displacement of the probe and yielded binding curves, with the exception of GSKI GSKL and GSKN. These three compounds showed unexpected results potentially arising from auto-fluorescence or interaction with the probe. For the 11 remaining compounds, IC_{50} values were calculated for the interaction with *Ld*BDF5.2 (Table 6.3), highlighting four which showed improved affinity over the original GSK hit; GSK5.2A, were GSK5.2B, D, E, and K (Figure 6.16). Additionally, all eleven compounds showed improved affinity over the two human bromodomain inhibitors previously tested in the FP assay; bromosporine (22.1 ± 6.5 µM) and

I-BRD9 (24.7 ± 6.2 µM). As noted for the acetyllysine compounds, it is possible that the FP assay sensitivity is limited which may cause the IC_{50} values for these high affinity compounds to underestimate their true potency.

The FP displacement assay was also performed with the new SO2FC reactive fragment. In this experiment, FP readings were taken every 15 minutes for 2 hours and at 20 hours due to the time-dependency of the covalent bond formation. SO2FNeg2 was included as a negative control and did not show binding with *Ld*BDF5.2. Over the initial two hours, SO2FC showed displacement of the probe but the competition curves did not plateau and associated IC_{50} values fell within a wide range (89.0 – 986 µM). In comparison, the binding curve at 20 hours showed completeness and model fitting yielded an IC_{50} value of 15.6 ± 1.8 µM (R^2 = 0.967) (Figure 6.17).

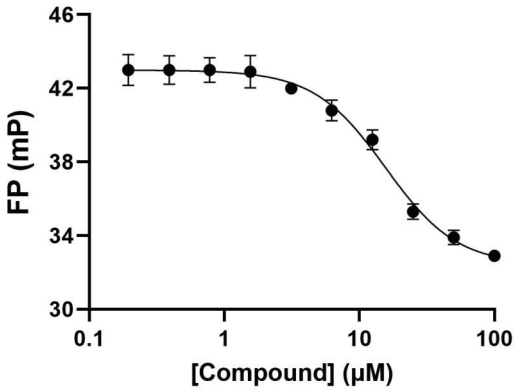


Figure 6.17. FP competition curve for binding of SO2FC to *Ld*BDF5.2. FP response for the binding of the sulfonyl fluoride compound SO2FC to *Ld*BDF5.2 after 20 hours incubation, with average, blank-corrected FP values plotted against ligand concentration, and data fitted with [Inhibitor] vs. response -- Variable slope (four parameters) non-linear regression; error bars representing SD (n = 3).

6.3.4. Co-crystallisation screens with sulfonyl fluoride reactive fragments

To visualise the binding interaction of hit sulfonyl fluoride fragments to *Ld*BDF5.1 and 5.2, attempts were made to co-crystallise these proteins with the compounds. Sitting drop co-crystallisation screens were first set up for *Ld*BDF5.1 with SO2FA,

using five commercial screens with 20 mg/ml protein and 5 mM SO2FA, and 150:150 or 200:100 nl drops. Hits were seen in several conditions of the Hampton HT and JCSG+ plates, with the largest crystal grown in 0.17 M ammonium sulfate, 25% w/v PEG 4000 (Figure 6.18A). This crystal showed minor polarisation under polarised light (Figure 6.18B) and was therefore fished but did not show any diffraction.

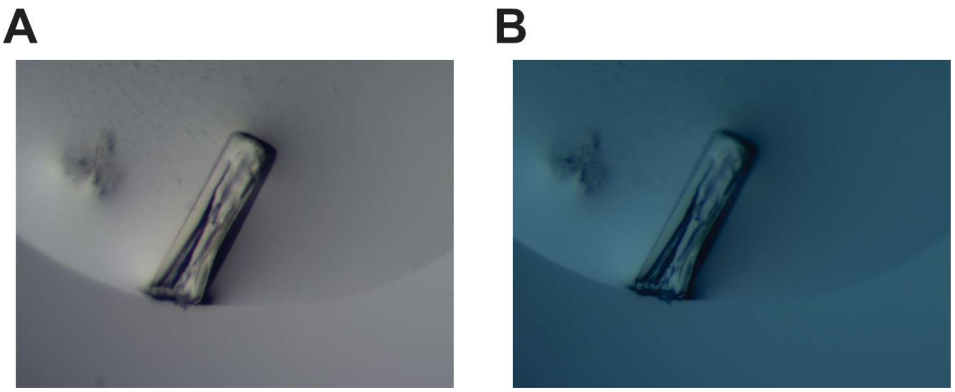


Figure 6.18. Co-crystals of *Ld*BDF5.1 with SO2FA. JCSG+ screen hit well D9 images under **(A)** white light and **(B)** polarised light, showing minor polarisation.

After emergence of SO2FC as the lead covalent ligand of *Ld*BDF5.2, screens were also set up for this compound with *Ld*BDF5.2. For this compound, a different approach was used whereby the compound was incubated with SO2FC at low concentrations (1.78 mg/ml *Ld*BDF5.2, 0.2 mM SO2FC) for 24 hours at 4°C. The protein-ligand complex was then concentrated to a higher concentration (20 or 22 mg/ml) and used within co-crystallisation screens. Four different commercial screens were used, with 150:150 and 200:100 nl drops, however these conditions did not yield protein crystals suitable for fishing and diffraction.

6.3.5. *Leishmania* promastigote single point-screen

As an additional screening step to refine the new ASMS hits into a smaller subset of hits for follow-up characterisation and development, a point-screen cell viability assay was performed using *Leishmania* promastigotes. Compounds which induced a significant effect on parasite growth at 10 µM could then be further investigated in a dose-response assay over a range of concentrations. The reversible compounds GSK5.2A – GSK5.2N and the three sulfonyl fluoride fragments, SO2FA, B, C were screened alongside negative control compounds, SGC-CBP30 and miltefosine.

Compounds were screened against *L. donovani* and *L. mexicana* promastigotes in two separate biological replicate experiments, with 5-day exposure and cell viability quantified by measuring fluorescence associated conversion of resazurin to resorufin. Average, blank-corrected fluorescence measurements were normalised to give values as % growth of DMSO control. A heat map of the full set of results which indicated good correlation between the replicates (Figure 6.19) is shown alongside results from the first biological replicate (Figure 6.20).

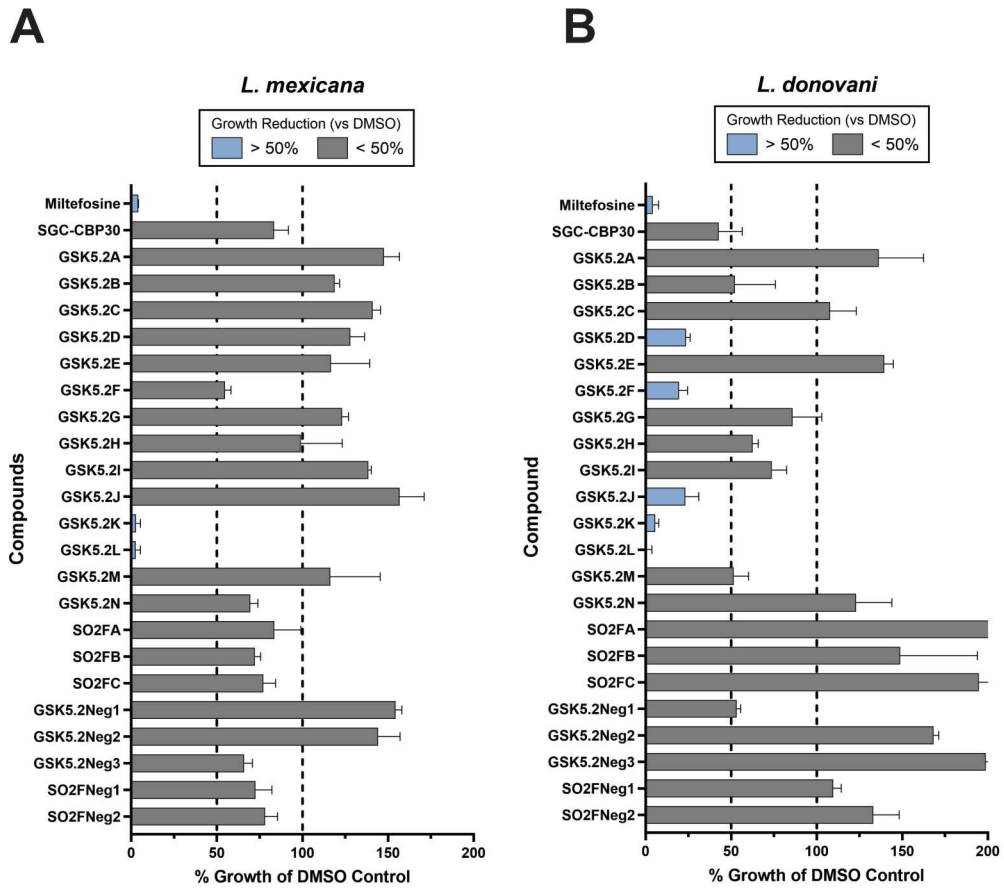


Figure 6.19. Point-screen of high-throughput screen compounds against *Leishmania* promastigotes. Data from the first biological replicate of compounds (10 µM) screened against **(A)** *L. mexicana* and **(B)** *L. donovani* promastigotes, with average, blank-corrected fluorescence readings normalised to give values as % growth of DMSO control plotted with SEM shown as error bars (n = 3).

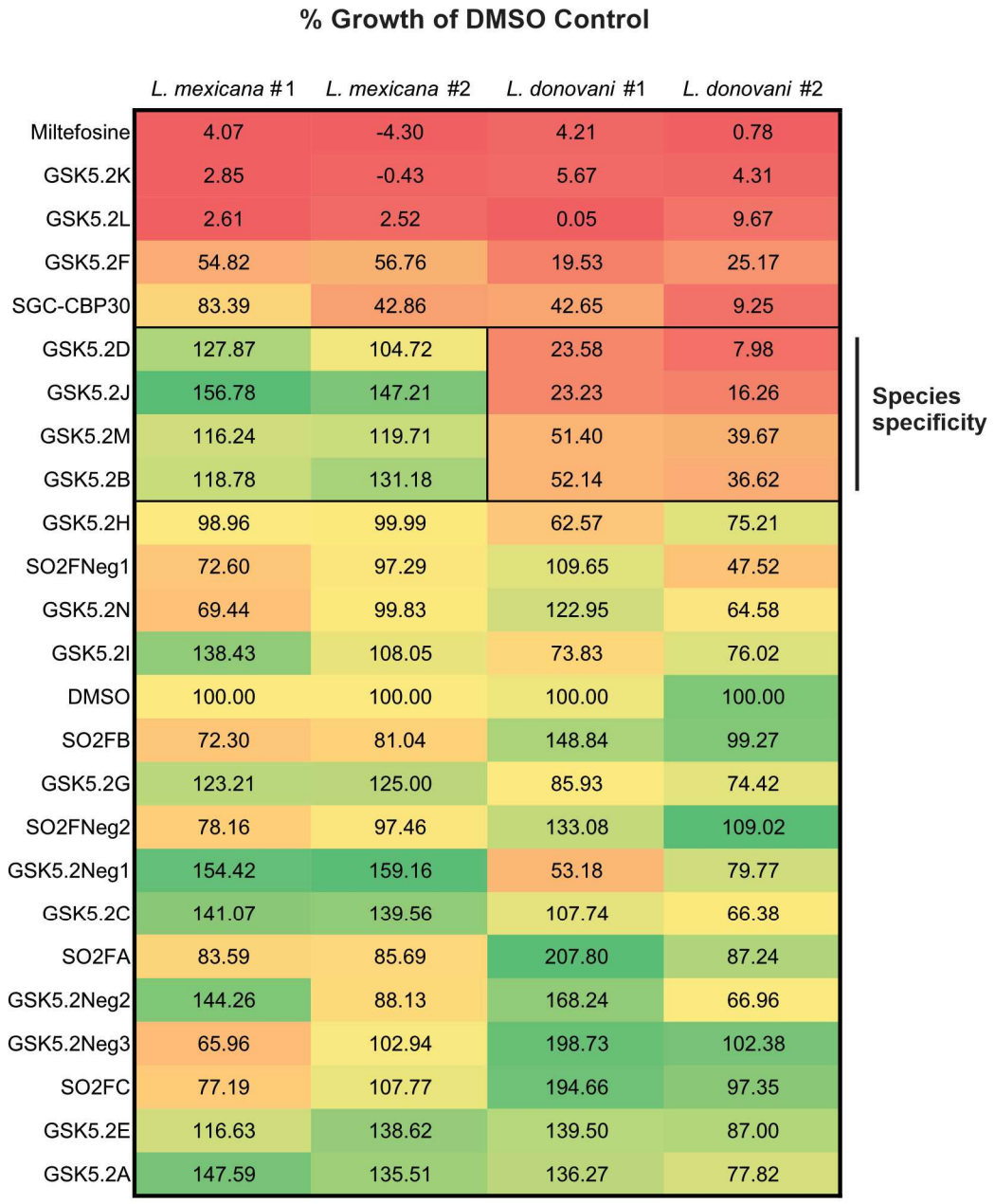


Figure 6.20. Activity of high-throughput screen hit compounds against *Leishmania* promastigotes. Results from the two replicate point-screens of *L. mexicana* and *L. donovani* promastigotes exposed to compounds at 10 µM for 5 days. Average, blank-corrected fluorescence readings normalised to give values as % growth of DMSO with heat map display shows 'hot' colours corresponding to compounds with greater effect on parasite growth. Compounds which exhibited clear species-specificity are indicated.

Seven compounds caused a greater than 50% reduction in parasite growth in *L. donovani* and two of these, GSK5.2K and GSK5.2L, also caused a greater than 50% reduction in *L. mexicana*. These two compounds exhibited particularly high potency, comparable to miltefosine. There is some degree of correlation in binding affinity and effect on the parasite, as GSK5.2B, D and K were top hits in both cell viability and FP assays. This was a pleasing result as it indicated such compounds were not only performing well in target-based assays but were also able to reach and bind these targets in the parasite and elicit an effect on cell growth. Interestingly, a species-specificity was observed for GSK5.2B, D, M and J, with markedly higher activity against *L. donovani* over *L. mexicana* (Figure 6.20). As these compounds were identified in a screen using recombinant *L. donovani* BDF5, it is conceivable that they would perform better against this species. Based on these data, five 'top hits' were selected for further analysis in a dose-response assay; GSK5.2K, GSK5.2L, GSK5.2F, GSK5.2D and GSK5.2J, where the latter two showed species-specificity.

6.4. Dose-response activity of hit compounds against *Leishmania* promastigotes

Using the biophysical assays described above, a subset of the most promising inhibitors of *Ld*BDF5 was defined from the two separate veins of ligand identification; SAR-optimised acetyllysine mimic compounds targeting *Ld*BDF5.1, and hit compounds from high-throughput screens targeting *Ld*BDF5.2. These included four *Ld*BDF5.1 hits; JC-E15, JC-E35, JC-E43, JC-E50 and five *Ld*BDF5.2 hits; GSK5.2D, GSK5.2F, GSK5.2J, GSK5.2K, GSK5.2L. To characterise and quantify the *in vitro* activity of these compounds against *Leishmania* promastigotes, a dose-response assay was performed by exposing the parasites to each compound (0 – 30 µM) for five days and measuring fluorescence emitted by resorufin to determine cell viability as before.

All compounds showed higher activity against *L. donovani* than *L. mexicana* (Table 6.4; Figure 6.21). As seen previously, compounds GSK5.2D and GSK5.2J showed

strong species-specificity, as did all acetyllysine mimic compounds with the exception of JC-E15 which showed low cytotoxicity to both species. As the acetyllysine mimic compounds were SAR-optimised based on biophysical assays using the *L. donovani* BDF5.1 protein, this may again explain this observation.

It is also interesting to note that these hits were developed as inhibitors of either the first or second bromodomain of *Ld*BDF5 and have exhibited selectivity in assays performed by collaborators, therefore these results indicate that selective inhibition of either BDF5.1 or BDF5.2 may be sufficient to inhibit parasite growth. This observation contrasts with previous findings that mutation of the single bromodomains can be tolerated by the parasite, but mutation of both bromodomains caused growth defects (Jones *et al.*, 2022). However, in this study the conserved asparagine residues were mutated to phenylalanine in order to occlude binding in the hydrophobic pocket, and as seen in Chapter 5, *Ld*BDF5.1N90F carrying this same mutation was still able to bind to the human bromodomain inhibitor SGC-CBP30 in TSA. Although, it should also be noted that potential off-target effects of these compounds with other BDFs, or indeed other non-bromodomain proteins in the parasite, cannot be ruled out.

Whilst all compounds were phenotypically active to some degree, the greatest reductions in parasite viability were obtained with JC-E50 (highest EC_{50} = 1.52 ± 0.35 µM), GSK5.2D (highest EC_{50} = 1.64 ± 0.095 µM) which were hit compounds for *Ld*BDF5.1 and 5.2, respectively. All compounds showed improved activity over the human bromodomain inhibitors bromosporine and I-BRD9, and many also showed higher potency than SGC-CBP30 which gave EC_{50} values of 4.43 – 8.37 µM (Table 5.3). This indicates that the compounds identified in these drug discovery efforts not only have activity against the parasite but show comparable and often improved cytotoxicity compared with an optimised human bromodomain inhibitor.

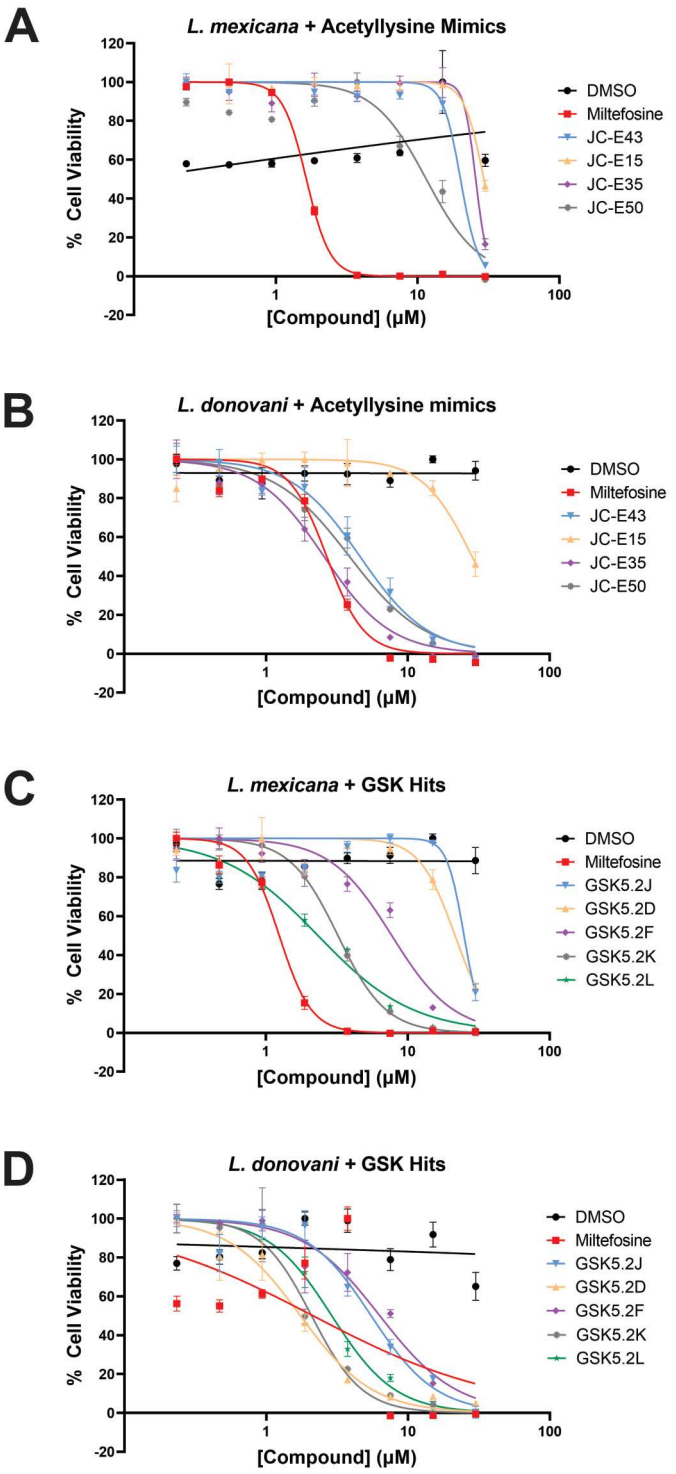


Figure 6.21. Dose-response curves for *Leishmania* promastigotes with hit compounds. Graphs from the first cell viability assay replicate experiment, with 5-day incubation of acetyllysine mimic hit compounds with **(A)** *L. mexicana* and **(B)** *L. donovani* promastigotes, and GSK hit compounds with **(C)** *L. mexicana* and **(D)** *L. donovani* promastigotes. Average, blank-corrected fluorescence intensity measurements were normalised and fitted to an [Inhibitor] vs. normalized response -- variable slope dose-response model. Error bars represent SEM (n = 3).

Table 6.4. EC_{50} values for hit compounds against *Leishmania* promastigotes. EC_{50} values calculated from dose-response activity of hit compounds and miltefosine (taken from acetyllysine experiments) against *L. mexicana* and *L. donovani* promastigotes, reported as EC_{50} ± standard error (n = 3). Heat map display shows 'hot' colours corresponding to lower EC_{50} values; #1 and #2 denote the two biological replicates; dashes are EC_{50} that could not be reliably determined.

Compound	*L. mexicana* #1 EC_{50} (µM)	*L. mexicana* #2 EC_{50} (µM)	*L. donovani* #1 EC_{50} (µM)	*L. donovani* #2 EC_{50} (µM)
Miltefosine	1.64 ± 0.019 (R^2 = 1.00)	0.959 ± 0.50 (R^2 = 0.605)	2.68 ± 0.43 (R^2 = 0.976)	1.67 ± 0.099 (R^2 = 0.993)
JC-E15	29.3 ± 0.49 (R^2 = 986)	-	28.4 ± 2.8 (R^2 = 0.880)	-
JC-E35	25.7 ± 9.0 (R^2 =0..965)	-	2.57 ± 0.17 (R^2 = 0.991)	2.32 ± 0.31 (R^2 = 0.966)
JC-E43	20.1 ± 1.6 (R^2 = 0.973)	-	4.61 ± 0.41 (R^2 = 0.983)	6.39 ± 1.36 (R^2 = 0.902)
JC-E50	11.5 ± 2.0 (R^2 = 0.865)	-	3.88 ± 0.39 (R^2 = 0.978)	1.52 ± 0.35 (R^2 = 0.912)
GSK5.2D	21.5 ± 2.1 (R^2 = 0.899)	-	1.74 ± 0.17 (R^2 = 0.979)	1.64 ± 0.095 (R^2 = 0.991)
GSK5.2F	7.72 ± 0.91 (R^2 = 0.960)	11.9 ± 0.52 (R^2 = 0.993)	6.33 ± 0.72 (R^2 = 0.971)	4.32 ± 0.54 (R^2 = 0.966)
GSK5.2J	24.9 ± 6.1 (R^2 = 0.710)	-	5.50 ± 0.66 (R^2 = 0.964)	4.14 ± 0.27 (R^2 = 0.991)
GSK5.2K	3.22 ± 0.036 (R^2 = 1.00)	3.89 ± 0.085 (R^2 = 0.999)	2.09 ± 0.15 (R^2 = 0.987)	2.37 ± 0.17 (R^2 = 0.990)
GSK5.2L	2.33 ± 0.26 (R^2 = 0.977)	4.27 ± 0.47 (R^2 = 0.973)	2.90 ± 0.33 (R^2 = 0.970)	3.80 ± 0.27 (R^2 = 0.989)

6.4.1. Species-specific cytotoxicity of hit compounds

To further probe the species-specificity observed for several compounds within the cell viability assays, the homology between BDF5 in *L. mexicana* and *L. donovani* was investigated. Sequence alignment identified differences in two residues in the first bromodomain and seven in the second bromodomain. However, mapping these onto the co-crystal structures of the *L. donovani* bromodomain proteins (PDB codes 6BYA and 5TCK) showed none of these residues were within the acetyllysine binding pockets and were not in close proximity to the bound inhibitors (SGC-CBP30 and bromosporine) in these structures (Figure 6.22).

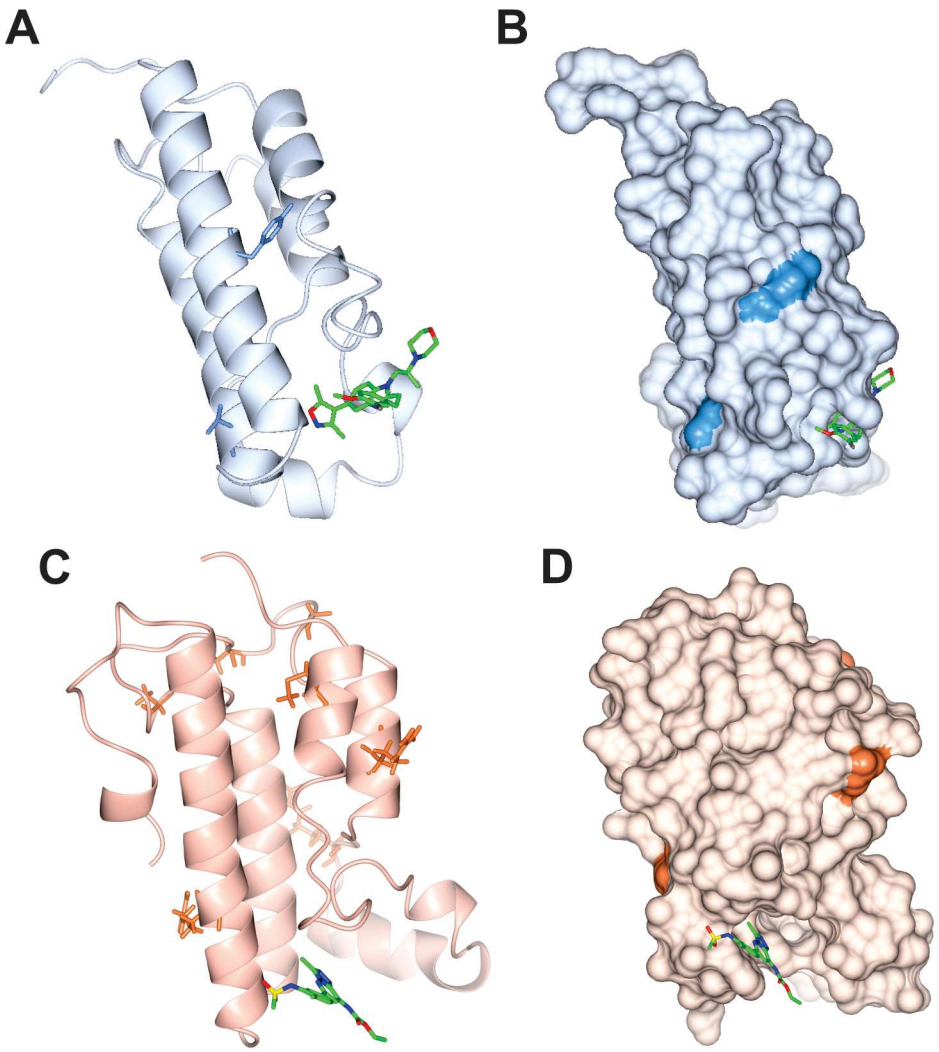


Figure 6.22. Amino acid residue differences between *L. donovani* and *L. mexicana* BDF5 bromodomains. Ribbon and surface representations of **(A) & (B)** *Ld*BDF5.1 in complex with SGC-CBP30 (PDB code 6BYA), and **(C) & (D)** *Ld*BDF5.2 in complex with bromosporine (PDB code 5TCK). Amino acid residues in the *L. donovani* structures which are not conserved in *L. mexicana* are highlighted. Images created using CCP4mg software.

This analysis would suggest that the selectivity these compounds have for *L. donovani* is not due to sequence differences between BDF5 in the two species, unless the residues, located outside the binding pocket, somehow convey a difference in the structure of the binding cavity. Also intriguing is the observation that SGC-CBP30, which shares the same core as the acetyllysine mimic compounds, did not exhibit any species-specificity in the cell viability assay (Table 5.3). If the species specificity is not due to differences in the BDF5 proteins in the two species, it may be related to other interactions between the compounds and the parasites, such as cell permeability. Further work could examine this possibility and also compare the activity in other species such as *L. major*.

6.5. Discussion

The collaborative drug discovery efforts outlined in this chapter led to the identification of several *Ld*BDF5 inhibitors with low micromolar affinity for the first and second bromodomains of the protein. These include compounds with selectivity for one of the bromodomains (such as GSK5.1A) as well as compounds with affinity for both the bromodomains (such as SO2FA). In addition to the reversible compounds identified in high-throughput screens and through SAR-optimisation, covalent compounds have also been identified which react with *Ld*BDF5, one of which has already been employed as a tool compound in drug-discovery screens. An array of biophysical assays was employed to validate and characterise the binding of these compounds to *Ld*BDF5 which fed back into the pipeline for further optimisation of the inhibitors. Additionally, cell viability assays were able to confirm that these compounds caused a phenotypic effect on the parasite *in vitro*.

There is clearly scope for the further development of these lead compounds, and various avenues may be explored in search of high affinity inhibitors, such as combining *Ld*BDF5.1 and 5.2 hits to generate compounds with dual bromodomain-binding ability. Additionally, whilst the reactive sulfonyl fluoride fragments have already shown applicability as tool compounds, they may also be explored for their therapeutic potential. Historically there have been concerns around the use of

covalent inhibitors in medical applications due to off-target toxicity, however recent studies have begun to evaluate their safety and highlight the untapped potential these high potency compounds possess (Sutanto, Konstantinidou and Dömling, 2020).

The work described in this chapter provides confirmation that targeted inhibition of bromodomains in *Leishmania* is a viable drug-discovery route for the identification of new antileishmanials, and provides a strong foundation for the future optimisation of potent inhibitors.

7. Chapter 7: Discussion and Future Work

The search for new treatments for leishmaniasis is of pressing international importance and epigenetic modules such as bromodomains present attractive targets for the development of new antileishmanials, with genetic target validation already beginning to unveil the potential of *Leishmania* BDF proteins (Jones *et al.*, 2022). The aim of this work was to investigate bromodomain proteins in *Leishmania*, both to better understand their roles within the parasite, and to validate these proteins as chemically tractable drug targets for the treatment of leishmaniasis.

Using DNA manipulation and recombinant protein production methods, soluble *L. donovani* BDF bromodomain proteins were produced (Chapter 3). These were then applied in biophysical assays to investigate histone binding interactions using specifically designed peptides based on the *L. donovani* histone sequences. Using thermal shift assay (TSA) and fluorescence polarisation (FP) approaches, two histone binding sites of *Ld*BDF5 were identified in $H2B_{17-23}$ acetylated at K19 and K21, and $H4_{1-15}$ acetylated at K2, K4, K10 and K14 (Chapter 4). This latter discovery was particularly gratifying, as it is consistent with the previous work suggesting these lysine residues are acetylated by HAT proteins in the parasite (Kumar *et al.*, 2012; Maity and Saha, 2012; Kumar and Saha, 2015; Chandra *et al.*, 2017; Jha *et al.*, 2017). These findings provide insight into the histone binding sites and associated biological function of BDF5 in *Leishmania*, with further implications in the wider field of parasite epigenetics.

The highest affinity peptide was then utilised as a probe for subsequent FP competition assays to investigate the binding of inhibitor compounds to the recombinant *Ld*BDF5 bromodomain proteins. This assay was used alongside TSA, and ligand-observed NMR experiments to first evaluate binding of human bromodomain inhibitors to the *Ld*BDF bromodomains. Bromosporine, SGC-CBP30 and I-BRD9 were initially identified as ligands of *Ld*BDF2, 5.1 and 5.2, and subsequent analysis further characterised the interactions of SGC-CBP30 with *Ld*BDF5.1, and bromosporine and I-BRD9 with *Ld*BDF5.2. This characterisation included calculation of IC_{50} values for the binding interactions by FP, whilst NMR

experiments also provided insight into the binding epitope of SGC-CBP30. Furthermore, promastigote cell viability assays showed that SGC-CBP30 produced a significant cytotoxic effect on the parasite (Chapter 5). This provided evidence that *Leishmania* BDF5 may be a chemically tractable drug target, with an inhibitor compound both able to bind to the protein, and to exert detrimental effects on the parasite.

Building upon these observations, subsequent work sought to identify novel compounds specifically targeting the *Leishmania* BDF proteins. In a collaborative effort with the University of Oxford and GSK, a series of hit compounds were identified for validation and further characterisation. These included both SAR-optimised acetyllysine mimic fragments, hits from high-throughput screens, and covalent sulfonyl fluoride fragments. Employing the orthogonal biophysical assays, a pipeline was established for the screening and subsequent refinement of these compound libraries to define smaller cohorts of higher affinity *Ld*BDF5.1 and *Ld*BDF5.2 ligands which showed low micromolar affinity in the FP assay. Nine of the most promising compounds were screened against *Leishmania* promastigotes which pleasingly showed that not only were the majority of these compounds phenotypically active, but they displayed potency (up to EC_{50} = 1.52 ± 0.35 µM) which was comparable with the established antileishmanial, miltefosine (Chapter 6). These findings have established a strong foundation for the future generation of optimised compounds which target *Leishmania* BDF5.

One limitation of this work was the inability of the FP assay to confidently determine binding affinities for high potency compounds, due to the weak binding of the peptide probe to the protein. Ideally, a higher affinity compound would be fluorescently labelled and used as the probe in this assay. Indeed, previous attempts were made by collaborators to develop a probe based on the human bromodomain inhibitor BI 2536 which co-crystallised with *Ld*BDF5.2, however this proved unsuccessful in FP assays with *Leishmania* bromodomains (Laurin, 2019). Another limitation of this work was the absence of crystal structures for the *Ld*BDF bromodomains in complex with peptides or inhibitors. Future work should aim to co-crystallise *Ld*BDF5 with optimised inhibitors, and with the acetylated H2B and H4 histone peptides identified here. Visualising the latter interactions would help

dissect the specific acetyllysine residues involved, and could provide further insight into the tandem bromodomain peptide recognition mechanism. Whilst binding properties and inhibition of tandem bromodomains are poorly understood, the work described here has uncovered details surrounding the possible functions and distinct binding interactions of the two bromodomains in *Ld*BDF5. The tandem bromodomain binding mechanism could be further investigated in regard to the observation that the first bromodomain of *Ld*BDF5 bound to H2B and H4 peptide substrates with a higher affinity than the second. A potential auto-regulation mechanism of BDF5 could be explored, specifically in relation to PTMs on the protein itself such as acetylation of the five lysine residues proximal to the tandem bromodomains (Section 4.6.1), and phosphorylation of the linker between the two bromodomains (Jones *et al*., 2022). In addition to this, solving a crystal structure of the tandem *Ld*BDF5T protein would be an extremely valuable addition to our wider understanding of tandem bromodomain architecture, considering the notable shortage of published tandem bromodomain structures.

The findings presented here form part of a larger overall process of validating BDF5 as a viable drug target in *Leishmania* and assessing the tractability of BDF5 inhibitors. Whilst this work focused on the chemical validation of BDF5, it is also important to consider the other elements within the overall pharmacological target validation process. As described in other work, there are several important criteria to consider when pursuing a drug target such as BDF5 (Forte *et al*., 2021).

Genetic validation is another important early step when evaluating a potential drug target. One key requirement is that the protein must be essential for parasite survival and transmission. For BDF5, previous work established that the protein is essential for both promastigote and amastigote stages of *Leishmania*, and for murine infection (Jones *et al*., 2022). Understanding the biological function of the target is also important in order to understand the mode of action of inhibitors and the downstream effects resulting from inhibition. Whilst the role of BDF5 in *Leishmania* has not been fully characterised, it is known to be essential for RNA polymerase II-mediated transcription and has essential involvement in regulation of cell growth (Jones *et al*., 2022). Furthermore, as reported here (Chapter 4), hypothetical interactions of BDF5 with acetylated histones H2B and H4 have also

been uncovered. Therefore, it may be theorised that inhibition of BDF5 disrupts interactions with acetylated H2B or H4 histones, causing dysregulation of RNA polymerase II-mediated transcription with downstream effects on cell growth resulting in parasite death. Future work should aim to further elucidate the interaction of BDF5 with H2B and H4, and how this relates to the biological function of BDF5.

The work described here addresses elements of the chemical validation of BDF5, bridging the gap between genetic validation and further downstream pharmacological validation steps. It was established that soluble protein can be expressed for the *Ld*BDF5 bromodomains, and that such protein can be applied within biophysical assays. Furthermore, druggability of BDF5 was addressed which led to the identification of small molecule inhibitors which were able to bind to the protein, and disrupt binding to acetylated histone peptides as demonstrated in the FP assay. Furthermore, novel compounds were identified which displayed low micromolar affinity binding to the BDF5 bromodomains. Future optimisation of these compounds will aim to generate higher affinity inhibitors, which should be further evaluated within the context of BDF5 target validation. This will include gathering structural information pertaining to the protein and its inhibition. As discussed (Section 1.5.2), structures have previously been solved for the bromodomains of *Ld*BDF5 in complex with human bromodomain inhibitors, granting insight into the possible binding mechanisms and inhibition strategies. Whilst attempts to co-crystallise BDF5 bromodomains with novel compounds in the present work were ultimately unsuccessful, future co-crystallisation efforts with higher affinity ligands may prove more successful and in turn aid in future SAR-optimisation efforts. In this work, cell viability assays were applied to demonstrate that novel BDF5 inhibitor compounds were active against promastigote stage *Leishmania*. Future work must also confirm that compounds are active against the clinically relevant, intracellular amastigote stage of the parasite. It should also be established that the compounds are acting on-target within the parasites. Techniques including use of an overexpression library, or thermal proteome profiling may be applied to this end, and have been used successfully to confirm other antileishmanial compounds are acting on-target (Corpas-Lopez *et al.*, 2019).

Another element within the future target validation of BDF5 will be predicting resistance potential. *In vitro* resistance selection experiments may be used to generate parasites with mutations conveying resistance to BDF5 inhibitors, followed by whole-genome sequencing to identify the mutations responsible (Leprohon *et al*., 2015). This will enable resistance mechanisms to be predicted and assessment of how manageable the resistance risk is in the context of pursuing BDF5 inhibitors. The selectivity of future lead compounds will also need to be assessed to establish whether they bind selectively to *Leishmania* BDF5 and not to human bromodomain proteins. BLAST and DALI homology searches (Section 1.5.3) did not identify any human bromodomains with high sequence or structure homology to *Ld*BDF5 bromodomains, and human bromodomains (namely BRD4) were included at various points in initial compound screens to allow exclusion of promiscuous binders. Nevertheless, confirming selectivity of lead compounds for *Leishmania* BDF5 over proteins in the human host will be a crucial checkpoint on the trajectory of these compounds.

Together, these pieces of information will culminate to provide a thorough evaluation of BDF5 as a tractable drug target in *Leishmania*, with the long-term goal of establishing BDF5 inhibitors as candidates for pre-clinical development and clinical trials. Though there is still much to be learnt about the *Leishmania* bromodomains, the advancements described here provide crucial information for the validation of these proteins as tractable drug targets in ongoing efforts to identify new antileishmanial treatments which are so urgently needed to relieve the burden of this devastating neglected tropical disease.

Appendices

A. Appendix A: Native and recombinant *Ld*BDF proteins

Table A.1. Details of the native full-length *Ld*BDF proteins. Molecular weight was calculated using the Protparam online tool within the Expasy Swiss bioinformatics resource portal. PDB codes are given for bromodomains for which structures have been deposited for the proteins in complex with inhibitors.

Protein	Gene Name	UniProt ID	Amino Acid Residues	Molecular Weight (Da)	PDB Code(s)
*Ld*BDF1	LdBPK_367210	E9BV77	286	33163	-
*Ld*BDF2	LdBPK_363130	E9BU22	236	26770	5C4Q
*Ld*BDF3	LdBPK_363520	E9BU61	223	25906	5FEA
*Ld*BDF4	LdBPK_140360	E9BBS1	1041	110186	-
*Ld*BDF5	LdBPK_091320	E9BA17	652	69833	6BYA, 5TCM, 5TCK

Sequences of recombinant *Ld*BDF bromodomain proteins are listed below, followed by length and molecular weight information (Table A.2). Sequences include His-tags underlined, His-tag cleavage site in bold and mutations in grey. Amino acid residue positions corresponding to the full-length *Ld*BDF proteins (Table A.1) are given in brackets. *Ld*BDF2, 3, 5.1, 5.2 and 5T vectors were derived from a pET-15-MHL backbone plasmid, produced by Dr Cheryl Arrowsmith at the Structural Genomics Consortium (Addgene plasmid #26092); plasmids gifted by Dr Raymond Hui at the Structural Genomics Consortium, Toronto. *Ld*BDF1, 2, 3 and 4 were cloned into pET-YSBLLIC3C (Chapter 3) using the pET-15 *Ld*BDF2 and 3 plasmids and pEX-A128 and pUC-SP backbone containing *Ld*BDF1 and 4 sequences, respectively, designed by Dr Nathaniel Jones and synthesised externally (Eurofins & DC Biosciences). *Ld*BDF2 (N87F), 5.1(N90F) and 5.2(N257F) were produced using site-directed mutagenesis by Dr Nathaniel Jones at the University of York using the pET-15 template plasmids.

*Ld*BDF1a (pET-YSBLLIC3C) (8-131):

MGSSHHHHHHSSG**LEVLFQGP**ALVDELRSRLKRRRADIVRYLDGLFYDECRRIIECVSELDRD
NLFLRNPSSLPDYAQHVTRPMYWELIQRKLQRYEYRTAADFMADMRAVVNNCYLYNGIQA
PASKLARTMEVLMEDRFVTELRA

*Ld*BDF1b (pET-YSBLLIC3C) (18-131):

MGSSHHHHHHSSG**LEVLFQGP**ARRRADIVRYLDGLFYDECRRIIECVSELDRDNLFLRNPSSL
PDYAQHVTRPMYWELIQRKLQRYEYRTAADFMADMRAVVNNCYLYNGIQAPASKLARTM
EVLMEDRFVTELRA

*Ld*BDF1c (pET-YSBLLIC3C) (28-131):

MGSSHHHHHHSSG**LEVLFQGP**ADGLFYDECRRIIECVSELDRDNLFLRNPSSLPDYAQHVTR
PMYWELIQRKLQRYEYRTAADFMADMRAVVNNCYLYNGIQAPASKLARTMEVLMEDRFV
TELRA

*Ld*BDF2 (pET-YSBLLIC3C) (1-125):

MGSSHHHHHHSSG**LEVLFQGP**AMDVSKRPREEFHKEQCLSFVKKLWAADTLAMFHYPVS
ATEVPGYYDVVDTPMDLSTIRKNIEQGKYRTDTEVENDVVLMLSNALDFNEKGSQWHDLAK
QLKKRYLTLAQESGLSFDADQAFIPTK

*Ld*BDF3 (pET-YSBLLIC3C) (25-160):

MGSSHHHHHHSSG**LEVLFQGP**AHPRPLPAGKHAHRQSLETIPEVAELYHCIYKLYNEEESSV
WFREPVNALAQEIFTYYDVVKSPMSLRHILDNIVKGDTYSTALQVMEDVELIWKNCITFNGA
NSLLATEAGKCRSALDRIRRAYQDDQRITVEEAE

*Ld*BDF4 (pET-YSBLLIC3C) (473-629):

MGSSHHHHHHSSG**LEVLFQGP**ASIRVPTATRMLSLLHVLDVLWNTLPTSIPFRMPVSEMEA
PGYYRSTRDPVSLCQLYEEAFCGMTAPTHRACVRQRALITRRLLEVQAYPPSRIGSGNAVGEE
VQSCFFSPTHLRERLTTLKANCVDYNGGGSELSRQASALLSAGVKALRDSQAEED

*Ld*BDF2 (pGL2772 pET-15) (1-125):

MHHHHHHSSGR**ENLYFQG**MDVSKRPREEFHKEQCLSFVKKLWAADTLAMFHYPVSATEV
PGYYDVVDTPMDLSTIRKNIEQGKYRTDTEVENDVVLMLSNALDFNEKGSQWHDLAKQLKK
RYLTLAQESGLSFDADQAFIPTK

*Ld*BDF5.1 (pGL2776 pET-15) (5-122):

MHHHHHHSSGR**ENLYFQG**PKLYNEADVAALVRSLDRAEDHHIFAVDVLETYPYLAESYTKV
CPRRCDLATAAQKALEGAYSYDLRLEGLKADIALMASNCVAYNGPTSAYAETAAKFERYALE
QIDAFVLEHNGGC

*Ld*BDF5.2 (pGL2775 pET-15) (162-300):

MHHHHHHSSGR**ENLYFQG**AAPPSTREMVQLVDSLNRREDGGAFSVDVAEAYPDLRDSYR
KICPRPMNLILMRQRAKEGYYTSGSATVYGDTVAASLTRLREDIELLVRNCITFNVKVESWVTL
ARSFQAFAHRRVDDFVLRHAAFLRGTTMGAEVYE

*Ld*BDF5T (pGL2774 pET-15) (1-300):

MHHHHHHSSGR**ENLYFQG**MSATPKLYNEADVAALVRSLDRAEDHHIFAVDVLETYPYLAES
YTKVCPRRCDLATAAQKALEGAYSYDLRLEGLKADIALMASNCVAYNGPTSAYAETAAKFERY
ALEQIDAFVLEHNGGCRVSRLRLPRASASQEHASADGTAPKKGSAGTSAAHKTAAAAPPSTR
EMVQLVDSLNRREDGGAFSVDVAEAYPDLRDSYRKICPRPMNLILMRQRAKEGYYTSGSAT
VYGDTVAASLTRLREDIELLVRNCITFNVKVESWVTLARSFQAFAHRRVDDFVLRHAAFLRGT
TMGAEVYE

*Ld*BDF2(N87F) (pET-15) (1-125):

MHHHHHHSSGR**ENLYFQG**MDVSKRPREEFHKEQCLSFVKKLWAADTLAMFHYPVSATEV
PGYYDVVDTPMDLSTIRKNIEQGKYRTDTEVENDVVLMLSNALDFFEKGSQWHDLAKQLKK
RYLTLAQESGLSFDADQAFIPTK

*Ld*BDF5.1(N90F) (pET-15) (5-122):

MHHHHHHSSGR**ENLYFQG**PKLYNEADVAALVRSLDRAEDHHIFAVDVLETYPYLAESYTKV
CPRRCDLATAAQKALEGAYSYDLRLEGLKADIALMASNCVAYFGPTSAYAETAAKFERYALEQ
IDAFVLEHNGGC

*Ld*BDF5.2(N257F) (pET-15) (162-300):

MHHHHHHSSGR**ENLYFQG**AAPPSTREMVQLVDSLNRREDGGAFSVDVAEAYPDLRDSYR
KICPRPMNLILMRQRAKEGYYTSGSATVYGDTVAASLTRLREDIELLVRNCITFFVKVESWVTL
ARSFQAFAHRRVDDFVLRHAAFLRGTTMGAEVYE

Table A.2. Details of recombinant *Ld*BDF proteins. MW and ε_{280} determined using the Protparam tool in the Expasy Swiss bioinformatics resource portal. Values for His-tag cleaved proteins are also listed where His-tags were removed by TEV protease.

Protein	Amino Acid Residues	MW (Da)	ε_{280} ($M^{-1}cm^{-1}$)
His_6-*Ld*BDF1a (pET-YSBLLIC3C)	146	17274	17545
His_6-*Ld*BDF1b (pET-YSBLLIC3C)	136	16063	17545
His_6-*Ld*BDF1c (pET-YSBLLIC3C)	126	14764	16055
His_6-*Ld*BDF2 (pET-YSBLLIC3C)	147	16793	18450
His_6-*Ld*BDF3 (pET-YSBLLIC3C)	158	18053	21555
His_6-*Ld*BDF4 (pET-YSBLLIC3C)	179	19730	13200
His_6-*Ld*BDF2 (pET-15)	143	16615	19940
*Ld*BDF2 (pET-15)	126	14478	18450
His_6-*Ld*BDF5.1 (pET-15)	136	15191	15150
His_6-*Ld*BDF5.2 (pET-15)	157	17931	16055
His_6-*Ld*BDF5T (pET-15)	318	35185	29715
*Ld*BDF5T (pET-15)	301	33048	28225
His_6-*Ld*BDF2(N87F) (pET-15)	143	16648	19940
*Ld*BDF2(N87F) (pET-15)	126	14511	18450
His_6-*Ld*BDF5.1(N90F) (pET-15)	136	15224	15150
His_6-*Ld*BDF5.2(N257F) (pET-15)	157	17964	16055

B. Appendix B: Primers and plasmids used for cloning *Ld*BDF1, 2, 3 and 4

Table B.1. Sequences of primers used for the amplification of DNA by PCR in cloning. In the primers used for DNA insert preparation lower case letters indicate the overhangs and upper case letters indicate regions homologous to the *Ld*BDF template DNA.

Primer	Sequence
pETYSBLLIC3C F	5'-AGGCGCGCCTTCTCCTCACATATG-3'
pET-YSBLLIC3C R	5'-TGCTGGTCCCTGGAACAGAACTT-3'
BD1a F	5'-ttctgttccagggaccagcaCTCGTTGACGAACTGCGGAGTC-3'
BD1b F	5'-ttctgttccagggaccagcaCGCCGTCGTGCGGACATTGT-3'
BD1c F	5'-ttctgttccagggaccagcaGATGGCTTGTTCTACGACGAAT-3'
BD1 R	5'-tgtgaggagaaggcgcgcctctaAGCACGCAATTCGGTCACAAAG-3'
BD2 F	5'-ttctgttccagggaccagcaATGGACGTCAGCAAGCGGCCA-3'
BD2 R	5'-tgtgaggagaaggcgcgcctctaTTTCGTAGGGATGAACGCCTG-3'
BD3 F	5'-ttctgttccagggaccagcaCACCCACGGCCGCTGCCA-3'
BD3 R	5'-tgtgaggagaaggcgcgcctctaCTCAGCCTCCTCGACGGTGAT-3'
BD4 F	5'-ttctgttccagggaccagcaTCGATCCGCGTGCCTACCG-3'
BD4 R	5'-tgtgaggagaaggcgcgcctctaGTCCTCTTCTGCCTGGCTAT-3'
T7 F	5' TAATACGACTCACTATAGGG 3'
T7 R	5' TAGTTATTGCTCAGCGGTGG 3'
T7 R (Eurofins)	5' CTAGTTATTGCTCAGCGGT 3'

Table B.2. Plasmids and bromodomain regions used for cloning bromodomains of *Ld*BDF1, 2, 3 and 4. Names and sources of the plasmids including bromodomain residue positions, and residues of the cloned regions, with full cloned sequences listed Appendix A. Primers refer to sequences in Table B.1.

***Ld*BDF1a, 1b, 1c**	
Template Plasmid	pEX-A128 backbone
Template Source	Synthesised externally (Eurofins)
Template *Ld*BDF Nucleic Acid Sequence	ATGAGCTCACGTGATCAGTTTCTCGTTGACGAACTGC GGAGTCGCCTGAAACGCCGTCGTGCGGACATTGTGC GCTATCTGGATGGCTTGTTCTACGACGAATGTCGTC GCATCATCGAGTGCGTTTCGGAGTTAGATCGCGACA ACCTGTTTCTTCGGAATCCGTCCAGCTTACCCGATTA TGCGCAGCATGTAACCCGTCCGATGTATTGGGAACT GATTCAGCGCAAACTGCAACGCTATGAGTATCGTAC AGCAGCCGATTTCATGGCCGATATGCGCGCAGTCGT GAACAACTGCTACCTGTACAATGGGATTCAAGCCCCT GCGTCGAAACTCGCCCGCACTATGGAAGTCCTGATG GAAGATCGCTTTGTGACCGAATTGCGTGCTGCTCCG GTACCGCCAGCAGAAGTGAAGAAAGCGTGTACGGG TATGAGCTCTGCGGAT
Template *Ld*BDF Residues	LdBPK_367210 1-150 (codon optimised for E. coli)
Cloned Residues	8-131, 18-131, 28-131
Cloning Primers	BD1a F & BD1 R, BD1b F & BD1 R, BD1c F & BD1 R

*Ld*BDF2	
Template Plasmid	pGL2772 pET-15
Template Source	Gifted by Dr Raymond Hui - Structural Genomics Consortium, Toronto
Template *Ld*BDF Nucleic Acid Sequence	ATGGACGTCAGCAAGCGGCCACGCGAGGAATTCCAC AAGGAGCAGTGTCTCTCCTTCGTGAAGAAGCTTTGG GCGGCCGACACGCTCGCCATGTTTCACTATCCGGTG AGCGCCACCGAGGTGCCCGGCTACTACGACGTCGTA GATACACCTATGGACCTGTCCACGATTCGGAAGAAC ATCGAGCAGGGCAAGTACAGAACGGACACTGAGGT CGAGAATGACGTGGTGCTCATGCTGTCGAACGCTCT GGACTTTAACGAGAAAGGCTCGCAATGGCATGATTT GGCGAAGCAGCTCAAGAAGCGGTACCTGACCCTTGC GCAGGAGTCGGGGCTGTCCTTCGACGCAGACCAGG CGTTCATCCCTACGAAA
Template *Ld*BDF Residues	LdBPK_363130 1-125
Cloned Residues	1-125
Cloning Primers	BD2 F & BD2 R

*Ld*BDF3	
Template Plasmid	pGL2773 pET-15
Template Source	Gifted by Dr Raymond Hui - Structural Genomics Consortium, Toronto
Template *Ld*BDF Nucleic Acid Sequence	CACCCACGGCCGCTGCCAGCCGGCAAGCACGCGCAT CGACAGTCCCTAGAGACCATCCCCGAGGTGGCTGAG CTGTATCACTGTATCTACAAGCTCTACAATGAGGAAG AGAGCTCTGTGTGGTTTCGCGAGCCCGTCAACGCGC TGGCGCAGGAAATTTTCACGTACTACGACGTGGTCA AGTCACCCATGAGCCTGCGACACATTCTCGACAACAT TGTAAAGGGCGACACCTATTCCACCGCCCTACAGGT GATGGAAGACGTGGAGCTCATCTGGAAAAACTGCAT TACCTTCAACGGTGCCAACTCGCTCTTGGCCACCGAG

	GCGGGCAAGTGCCGATCCGCGCTAGACCGAATCCGT CGCGCCTACCAGGACGATCAGCGCATCACCGTCGAG GAGGCTGAG
Template *Ld*BDF Residues	LdBPK_363520 25-160
Cloned Residues	25-160
Cloning Primers	BD3 F & BD3 R
***Ld*BDF4**	
Template Plasmid	pUC-SP backbone
Template Source	Synthesised externally (DC Bioscience)
Template *Ld*BDF Nucleic Acid Sequence	AAAGACGCTGCAGCAGGACCGTTACCGACAGATGG CCTGTCGATCCGCGTGCCTACCGCAACTCGGATGCT GAGTTTGCTGCACGTGCTGGATGTGCTGTGGAACAC ATTGCCGACCTCTATCCCGTTTCGCATGCCGGTTTCG GAGATGGAAGCGCCAGGCTATTACCGGTCCACCCGT GATCCCGTCAGCCTGTGTCAGCTCTACGAGGAAGCG TTTTGCGGGATGACCGCGCCAACCCATCGTGCCTGT GTACGCCAACGTGCCCTGATTACGCGTCGTTTGCTG GAAGTCCAGGCTTATCCGCCGAGTCGCATTGGGTCG GGTAACGCTGTTGGCGAGGAAGTGCAGAGCTGCTTC TTCTCTCCCACCCACTTACGCGAACGCCTGACTACGC TCAAAGCCAATTGCGTGGACTACAATGGCGGTGGTT CCGAACTGAGCCGTCAAGCATCAGCCCTTCTTAGCG CGGGTGTCAAGGCGTTACGCGATAGCCAGGCAGAA GAGGACTCACTGACGGCGTATCAGGTTCGCGAAGAA GCCCATGTACGCGAAGCCCCTCCAGCGGCG
Template *Ld*BDF Residues	LdBPK_140360 460-629 (codon optimised for *E. coli*)
Cloned Residues	473-629
Cloning Primers	BD4 F & BD4 R

C. Appendix C: *L. donovani* histone sequences

Table C.1. Details of the native *L. donovani* histone proteins. Information listed for histones H2A.Z, H2B and H4 for which peptides were designed based on their sequences.

*Ld*H2A.Z	
Gene Name	LdBPK_170330
UniProt ID	E9BD54
Amino Acid Sequence	MSYTGEESTGMPPQPPMMGPGSATADQTSIVSGGKLGGK GKGKGKGKGKRGGKTGGKAGRRDRMSRAARAELNFPVGRI HSRLKDGLYRKQRCGASAAIYCAALLEYLTTEVIELSGAAAKA QKTERIKPRHLLLAIRGDEELNQVVKATISRGGVVPNVHKAL EKKSKKKSAKRAA
*Ld*H2B	
Gene Name	LdBPK_171320
UniProt ID	E9BDF4
Amino Acid Sequence	MATRRRPSCKPSNSHKSHRKPKRSWNVYVCRSLKAINAQM SMSHRTMKIVNSYVSDVMERICTEAAVLVRTNKKRTLGARE VQTAVRVVLPAELAKHAMAEGTKAVSKMCH
*Ld*H4	
Gene Name	LdBPK_150010
UniProt ID	E9BC42
Amino Acid Sequence	MAKGKRSADAKSSQKRQKKVLRDNIRGITRGCVRRMARRG GVKRISSEVYEEVRRVLKAYVEDIVRCSTAYTEYARKKTVTAC DVVNALRKQGHILYGYA

D. Appendix D: Buffer and SDS-PAGE gel compositions

Table D.1. Buffers used for cloning, protein production and biophysical assays. DTT was added to protein purification buffers following column equilibration.

Buffer	Recipe
50x TAE buffer	242 g Tris-HCl pH 8, 18.6 g EDTA, 57.1 ml acetic acid made up to 1 L with distilled water
SDS-PAGE resolving gel buffer	1.5 M Tris base, 0.4% SDS, pH 8.8
SDS-PAGE stacking gel buffer	0.5 M Tris base, 0.4% SDS, pH 6.8
SDS-PAGE running buffer	14.4 g glycine (electrophoresis grade), 3 g Tris base, 10 ml 10% SDS made up to 1 L with distilled water
4x SDS-PAGE sample buffer	120 mM Tris HCl pH 6.8, 20% glycerol, 4% SDS, 0.09% bromophenol blue, 10% β-mercaptoethanol
Protein purification buffer A	20 mM HEPES pH 7.5, 500 mM NaCl, 30 mM Imidazole, (1mM DTT), pH 7.5
Protein purification buffer B	20 mM HEPES pH 7.5, 500 mM NaCl, 300 mM Imidazole, (1 mM DTT), pH 7.5
Protein purification buffer C	20 mM HEPES pH 7.5, 500 mM NaCl, (1 mM DTT), pH 7.5
HEPES buffer	20 mM HEPES pH 7.5, 500 mM NaCl, pH 7.5
FP buffer	20 mM HEPES, 500 mM NaCl, 1 mg/ml BSA, pH 7.5
NMR buffer	20 mM NaPi, 100 mM NaCl, pH 7.5

Table D.2. SDS-PAGE gel compositions. Volumes are given for the preparation of two gels. Resolving and stacking gel buffer recipes are given in Table D.1. Ammonium perfulfate (APS) and tetramethylethylenediamine (TEMED) were added last before pouring gels.

Component	Resolving Gel 12%	Resolving Gel 20%	Stacking Gel 4%
Resolving gel buffer	2.5 ml	2.5 ml	-
Stacking gel buffer	-	-	1.3 ml
30% (w/v) Acrylamide	4.2 ml	6.7 ml	0.5 ml
10% APS	50 µl	50 µl	25 µl
TEMED	8 µl	8 µl	8 µl
0.1% Bromophenol blue	-	-	10 µl
Milli-Q water	3.2 ml	0.7 ml	3.2 ml

E. Appendix E: Ligand-binding assay conditions and optimisation

Table E.1. Optimised TSA concentrations. Protein and dye concentrations used in TSA experiments, established by optimisation experiments. For the mutant proteins, new optimisation experiments were not performed and instead the same concentrations as the native versions were used to allow for comparison.

Protein	Protein Concentration	SYPRO Orange Dye Concentration
*Ld*BDF2	2.0 µM	3x
*Ld*BDF2(N87F)	2.0 µM	3x
*Ld*BDF5.1	3.0 µM	3x
*Ld*BDF5.1(N90F)	3.0 µM	3x
*Ld*BDF5.2	3.0 µM	2x
*Ld*BDF5T	2.1 µM	2 x

NMR acquisition parameters:

1D proton experiments used a spectral width of 16 ppm and acquisition time of 1.5 seconds and recycle delay of 1 second.

For STD experiments, the stddiffesgp pulse program was used (Mayer and Meyer, 1999) with the protein was selectively saturated using a pulse centred at either 8.5 or 9 ppm and a Gaussian pulse shape used.

WaterLOGSY experiments used the ephogsygpno pulse program (Dalvit *et al.*, 2001) and a mixing time of 1.5 seconds.

For CPMG experiments, the standard Bruker cpmg_esgp2d pulse program was used with a short relaxation delay of 20 ms, and a long relaxation delay of 400 ms.

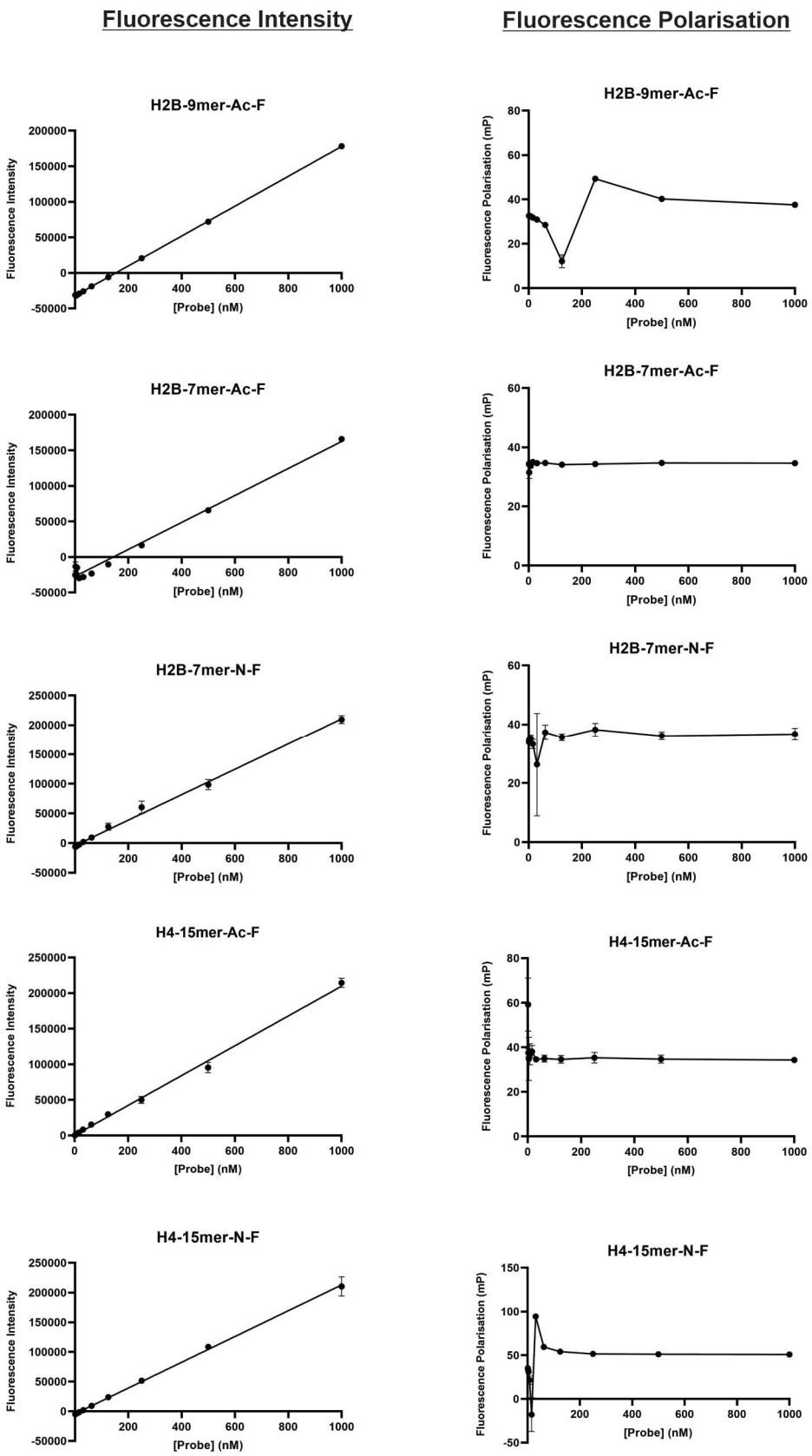


Figure E.1. FP probe optimisation. Fluorescence intensity and polarisation readings for the 5 histone peptide probes named. Average values are plotted ± SD and fluorescence intensity data fitted with linear regression.

F. Appendix F: Protein crystallisation conditions

Table F.1. Co-crystallisation screens trialled for *Ld*BDF proteins. Apo and co-crystal screens in 96-well plates using commercial screening solutions; JCSG+, PACT, Morpheus, PDB (Molecular Dimensions); $AmSO_4$ & MPD (Qiagen); Hampton HT (Hampton Research). Protein and ligand concentrations are given alongside % DMSO for the pre-mixed solutions used to setup sitting-drops.

*Ld*BDF Protein	Ligand	Screens	[Protein] (mg/ml)	[Ligand] (mM)	% DMSO	Drops (nl)
5.1	SO2FA	Hampton HT PACT $AmSO_4$	20	5	5	150:150
5.1	SO2FA	JCSG+ PDB	20	5	5	150:150 200:100
5.1	JC-B4	JCSG+ Morpheus	19	2.52	2.52	150:150 200:100
5.1	JC-B4	Morpheus	21	2.52	2.52	150:150
5.1	JC-C50	JCSG+	20	6	6	150:150 200:100
5.1	H2B-7mer-Ac	JCSG+ PACT	19	1.89 1.51	3.78 3.02	150:150
5.1	JC-E43	Morpheus $AmSO_4$ JCSG+	18 22	2.92 2.40	2.92 2.40	150:150
5.1	JC-E43	MPD	22	2.19	2.19	150:150 200:100
5.1	JC-E43	Morpheus MPD	22	2.19	2.19	250:250 300:200 500:500 600:300

5.1	JC-E43	Morpheus MPD	20 24	1.58 1.87	1.58 1.87	300:200
5.1	JC-E65	Morpheus MPD	20 24	1.58 1.87	1.58 1.87	300:200
5.2	I-BRD9	Hampton HT PACT $AmSO_4$	21	20	4	150:150
5.2	I-BRD9	Hampton HT	10	15	3	150:150
5T	I-BRD9	$AmSO_4$	4.68	0.75	5	150:150
5T	I-BRD9	$AmSO_4$ Hampton HT PACT JCSG+	8.50	11	6.5	150:150

Table F.2. Co-crystallisation screens trialled for *Ld*BDF5.2 with the covalent inhibitor SO2FC. Sitting-drop 96-well plate commercial screens PACT, Morpheus (Molecular Dimensions); $AmSO_4$ (Qiagen); Hampton HT (Hampton Research). Protein concentrations are given for the both the initial sample and final concentrated sample.

*Ld*BDF Protein	Ligand	Screens	Starting/ Final [Protein] (mg/ml)	Starting [Ligand] (mM)	Starting % DMSO	Drops (nl)
5.2	SO2FC	Morpheus Hampton HT PACT $AmSO_4$	1.78/ 19.70	0.2	2	150:150 200:100
5.2	SO2FC	Morpheus Hampton HT	1.78/ 22.0	0.2	2	150:150 200:100

Table F.3. Conditions for the optimisation of *Ld*BDF co-crystals. Optimisation by sitting-drop vapour diffusion in 48-well plates. Protein and ligand concentrations are given alongside % DMSO for the pre-mixed solutions used to setup sitting-drops.

	*Ld*BDF5.1 + JC-B4 Optimisation
[Protein] (mg/ml)	19, 21
[Ligand] (mM)	2.52
% DMSO	2,52
Drops (nl)	200:400, 300:300, 400:200
Screen Solutions	0.2 M zinc acetate dihydrate, 0.1 M sodium acetate pH 4.5, 10% w/v PEG 3000 0.2 M zinc acetate dihydrate, 0.1 M imidazole pH 8.0, 20% w/v PEG 3000 0.2 M zinc acetate dihydrate, 0.1 M sodium cacodylate pH 6.5, 10% v/v 2-propanol
	*Ld*BDF5.1 + JC-C50 Optimisation
[Protein] (mg/ml)	20
[Ligand] (mM)	5, 6
% DMSO	5, 6
Drops (nl)	500:500, 600:400, 670:330
Screen Solutions	0.2 M $MgCl_2$ hexahydrate, 0.1 M Tris pH 8.5, 10% w/v PEG 800

G. Appendix G: Acetyllysine mimic compound library

Table G.1. Structures of the first 11 compounds within the acetyllysine mimic library. Compounds screened by TSA for binding to *Ld*BDF bromodomain proteins.

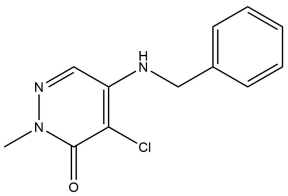
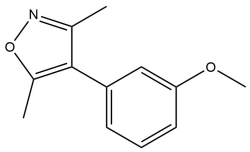
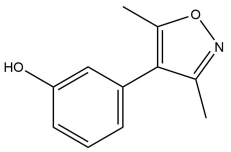
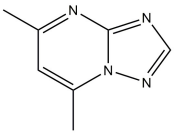
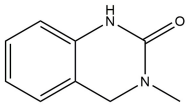
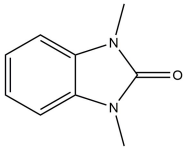
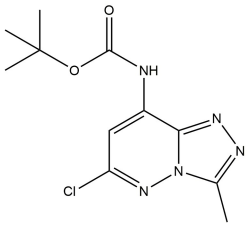
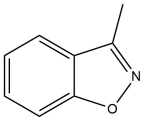
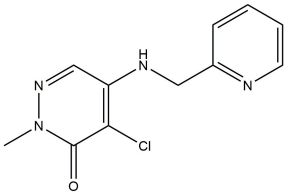
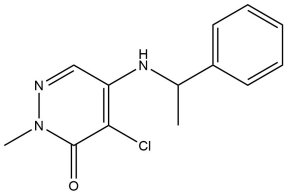
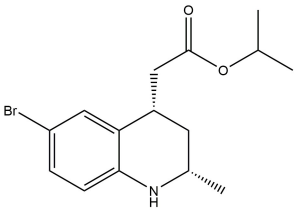
JC-A1 	JC-A2 	JC-A3 
JC-A4 	JC-A12 	JC-A15 
JC-A22 	JC-A29 	JC-A37 
JC-A38 	LLS14 	

Table G.2. Structures of the second 20 compounds within the acetyllysine mimic library.

Compounds screened by TSA and characterised in subsequent biophysical assays.

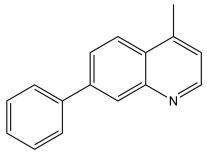
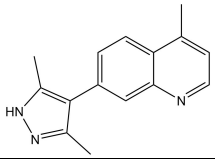
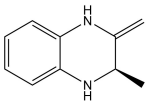
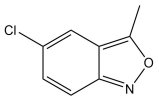
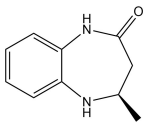
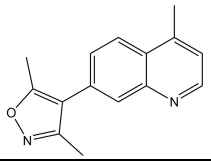
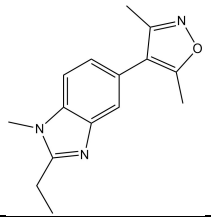
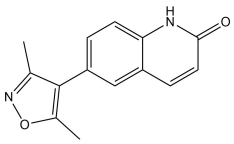
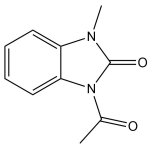
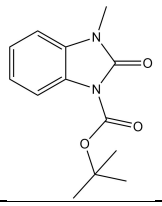
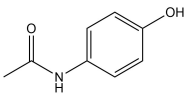
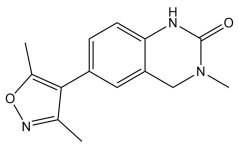
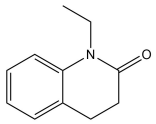
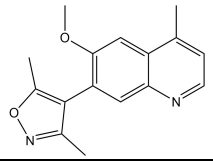
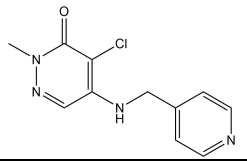
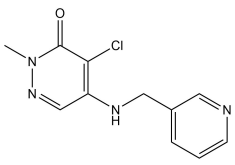
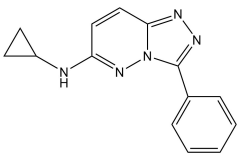
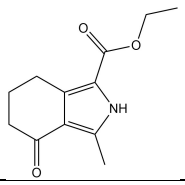
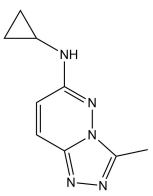
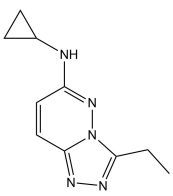
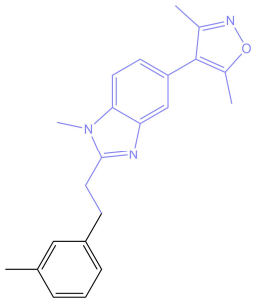
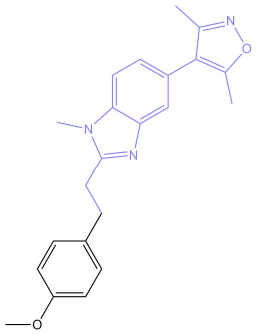
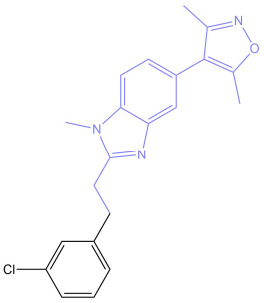
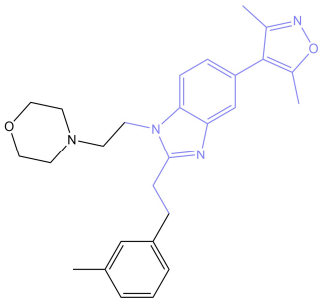
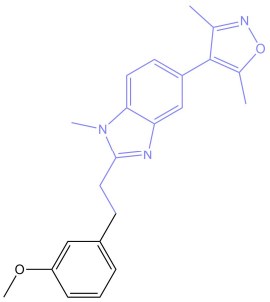
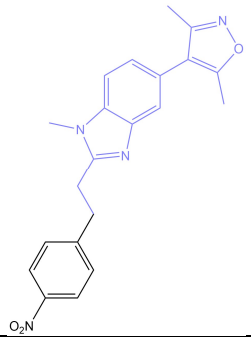
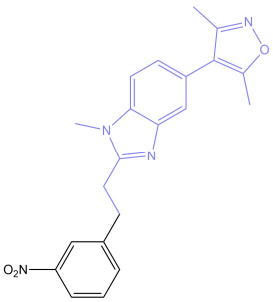
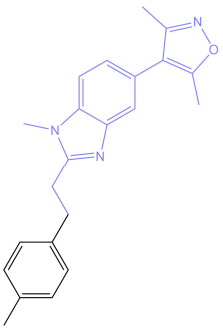
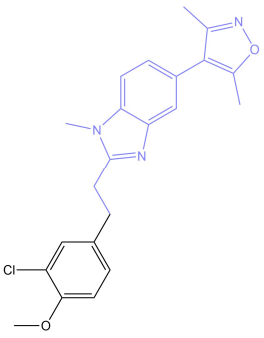
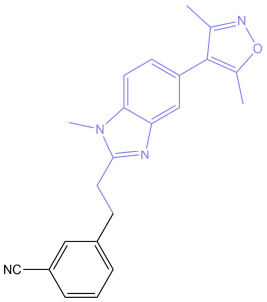
AS-A7 	AS-A8 	JC-A36 
JC-A44 	JC-A51 	JC-A78 
JC-B4 	JC-B28 	JC-B61 
JC-B62 	JC-B65 	JC-B79 
JC-B81 	JC-B84 	JC-B87 
JC-B88 	JC-B92 	JC-B93 
JC-C1 	JC-C2 	

Table G.3. Structures of the JC-B4 acetyllysine mimic analogues. SAR-optimised compounds screened for binding to *Ld*BDF bromodomain proteins in biophysical assays and co-crystallography, with hits assessed for cytotoxicity against *Leishmania* promastigotes in cell viability assays.

JC-C50 	JC-D61 	JC-E2 
JC-E15 	JC-E19 	JC-E35 
JC-E43 	JC-E44 	JC-E50 
JC-E65 		

Abbreviations

5-FAM	5-carboxyfluorescein
ε_{280}	extinction coefficient at 280 nm
Å	Angstrom
A_{280}	absorbance at 280 nm
APS	ammonium persulfate
ASMS	affinity selection mass spectrometry
BET	bromodomain and extra terminal
BLAST	basic local alignment search tool
BDF	bromodomain factor
bp	base pair
BSA	Bovine serum albumin
CPMG	Carr-Purcell-Meiboom-Gill
CV	column volume
Da	Dalton
DMSO	dimethyl sulfoxide
DNDi	Drugs for Neglected Diseases Initiative
DSS	sodium trimethylsilylpropanesulfonate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EC_{50}	half maximal effective concentration

EDTA	ethylenediaminetetraacetic acid
x g	relative centrifugal force
hi-FCS	heat-inactivated fetal calf serum
Fmoc	fluorenylmethyloxycarbonyl
FP	fluorescence polarisation
HAT	histone acetyltransferase
HDAC	histones deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOMEM	hemoflagellate-modified minimum essential medium
HPLC	high-performance liquid chromatography
IC_{50}	half maximal inhibitory concentration
IMAC	immobilised metal affinity chromatography
IPTG	isopropyl β-d-1-thiogalactopyranoside
ITC	isothermal titration calorimetry
kb	kilobase
K_D	dissociation constant
kDa	kilodaltons
LB	Luria-Bertani broth
L. donovani	*Leishmania donovani*
L. mexicana	*Leishmania mexicana*
*Ld*BDF	*Leishmania donovani* bromodomain factor

LCMS	liquid chromatography mass spectrometry
M	molar
mP	millipolarisation
MPD	2-methyl-2,4-pentanediol
NMR	nuclear magnetic resonance
MRG	MORF4-related gene
MS	mass spectrometry
MW	molecular weight
OD_{600}	optical density at 600 nm
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
ppm	parts per million
PTM	post-translational modification
PTU	polycistronic transcription unit
rpm	revolutions per minute
SAR	structure-activity relationship
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC-MALLS	size-exclusion chromatography multi-angle laser light scattering

SEC	size-exclusion chromatography
SEM	standard error of the mean
SD	standard deviation
SOC	super optimal broth with catabolite repression
SPR	surface plasmon resonance
STD	saturation transfer difference
TAE	tris-acetate-EDTA
TEMED	tetramethylethylenediamine
TEV	tobacco etch virus
ΔT_m	thermal shift
T_m	melting temperature
TSA	thermal shift assay
TSR	transcriptional start regions
TTA	transcriptional termination sites
U	enzyme unit
UV	ultraviolet
v/v	volume per unit volume
w/v	weight per unit volume
waterLOGSY	water ligand-observed via gradient spectroscopy

Reference List

Alamer, E. *et al.* (2021) 'Modulation of BRD4 in HIV epigenetic regulation: implications for finding an HIV cure', *Retrovirology*, 18(1), p. 3. doi: 10.1186/S12977-020-00547-9.

Ali, I. *et al.* (2018) 'Lysine Acetylation Goes Global: From Epigenetics to Metabolism and Therapeutics', *Chem Rev*, 118(3), pp. 1216–1252. doi: 10.1021/acs.chemrev.7b00181.

Ali, M. *et al.* (2012) 'Tandem PHD Fingers of MORF/MOZ Acetyltransferases Display Selectivity for Acetylated Histone H3 and Are Required for the Association with Chromatin', *J Mol Biol*, 424(5), pp. 328–338. doi: 10.1016/j.jmb.2012.10.004.

Alonso, V. L. *et al.* (2014) 'Trypanosoma cruzi Bromodomain Factor 3 Binds Acetylated α-Tubulin and Concentrates in the Flagellum during Metacyclogenesis', *Eukaryot Cell*, 13(6), pp. 822–831. doi: 10.1128/EC.00341-13.

Alonso, V. L. *et al.* (2016) 'Overexpression of bromodomain factor 3 in Trypanosoma cruzi (TcBDF3) affects differentiation of the parasite and protects it against bromodomain inhibitors', *FEBS J*, 283(11), pp. 2051–2066. doi: 10.1111/febs.13719.

Altschul, S. F. *et al.* (1990) 'Basic local alignment search tool', *J Mol Biol*, 215(3), pp. 403–410. doi: 10.1016/S0022-2836(05)80360-2.

Alvar, J. *et al.*, (2012) 'Leishmaniasis Worldwide and Global Estimates of Its Incidence', *PLoS ONE*, 7(5), p. e35671. Doi: 10.1371/JOURNAL.PONE.0035671.

Amorim, S. *et al.* (2016) 'Bromodomain inhibitor OTX015 in patients with lymphoma or multiple myeloma: a dose-escalation, open-label, pharmacokinetic, phase 1 study', *Lancet Haematol*, 3(4), pp. e196–e204. doi: 10.1016/S2352-3026(16)00021-1.

Annis, D. A. *et al.* (2007) 'Affinity selection-mass spectrometry screening techniques for small molecule drug discovery', *Curr Opin Chem Biol,* 11(5), pp. 518–526. doi: 10.1016/J.CBPA.2007.07.011.

Baharia, R. K. *et al.* (2014) 'Nucleosomal Histone Proteins of L. donovani: A Combination of Recombinant H2A, H2B, H3 and H4 Proteins Were Highly Immunogenic and Offered Optimum Prophylactic Efficacy against Leishmania Challenge in Hamsters', *PLoS ONE*, 9(6), p. e97911. doi: 10.1371/journal.pone.0097911.

Banfi, D. and Patiny, L. (2008) 'Resurrecting and Processing NMR Spectra On-line', *CHIMIA*, 62(4), pp. 280–281. doi: 10.2533/chimia.2008.280.

Berkovits, B. D. and Wolgemuth, D. J. (2013) 'The role of the double bromodomain-containing BET genes during mammalian spermatogenesis.', *Curr Top Dev Biol*, 102, pp. 293–326. doi: 10.1016/B978-0-12-416024-8.00011-8.

Boehm, D. *et al.* (2013) 'BET bromodomain-targeting compounds reactivate HIV from latency via a Tat-independent mechanism', *Cell Cycle*, 12(3), pp. 452–462. doi: 10.4161/cc.23309.

Bond, P. (2017) *JTSA*. Available at: http://paulsbond.co.uk/jtsa.

Bowman, B. R. *et al.* (2006) 'Multipurpose MRG Domain Involved in Cell Senescence and Proliferation Exhibits Structural Homology to a DNA-Interacting Domain', *Structure*, 14(1), pp. 151–158. doi: 10.1016/J.STR.2005.08.019.

Burza, S., Croft, S. L. and Boelaert, M. (2018) 'Leishmaniasis', *Lancet*, 392(10151), pp. 951–970. doi: 10.1016/S0140-6736(18)31204-2.

Burza, S. *et al*. (2022) 'AmBisome Monotherapy and Combination AmBisome-Miltefosine Therapy for the Treatment of Visceral Leishmaniasis in Patients Coinfected With Human Immunodeficiency Virus in India: A Randomized Open-Label, Parallel-Arm, Phase 3 Trial', *Clin Infect Dis*, 75(8), pp. 1423–1432. doi: 10.1093/CID/CIAC127.

Cammas, F. *et al.* (2002) 'Cell differentiation induces TIF1β association with centromeric heterochromatin via an HP1 interaction Florence', *J Cell Sci*, 115(Pt 17), pp. 3439–3448. doi: 10.1242/jcs.115.17.3439.

Carr, H. Y. and Purcell, E. M. (1954) 'Effects of Diffusion on Free Precession in Nuclear Magnetic Resonance Experiments', *Phys Rev*, 94(3), pp. 630–638. doi: 10.1103/PhysRev.94.630.

Castillo, A. M., Patiny, L. and Wist, J. (2011) 'Fast and accurate algorithm for the simulation of NMR spectra of large spin systems', *J Magn Reson*, 209(2), pp. 123–130. doi: 10.1016/J.JMR.2010.12.008.

Chandra, U. *et al.* (2017) 'Cell cycle stage-specific transcriptional activation of cyclins mediated by HAT2-dependent H4K10 acetylation of promoters in Leishmania donovani', *PLoS Pathog*, 13(9), p. e1006615. doi: 10.1371/JOURNAL.PPAT.1006615.

Chatelain, E. and Ioset, J. R. (2011) 'Drug discovery and development for neglected diseases: the DNDi model', *Drug Des Devel Ther*, 5, pp. 175–181. doi: 10.2147/DDDT.S16381.

Chung, C. *et al.* (2011) 'Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains', *J Med Chem*, 54(11), pp. 3827–3838. doi: 10.1021/jm200108t.

Cimmperman, P. *et al.* (2008) 'A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands', *Biophys J*, 95(7), pp. 3222–3231. doi: 10.1529/BIOPHYSJ.108.134973.

Cochran, A. G., Conery, A. R. and Sims, R. J. (2019) 'Bromodomains: a new target class for drug development', *Nat Rev Drug Discov*, 18(8), pp. 609–628. doi: 10.1038/s41573-019-0030-7.

Conery, A. R. *et al.* (2016) 'Bromodomain inhibition of the transcriptional coactivators CBP/EP300 as a therapeutic strategy to target the IRF4 network in multiple myeloma', *eLife*, 5, p. e10483. doi: 10.7554/eLife.10483.

Conway, S. J. (2012) 'Bromodomains: Are readers right for epigenetic therapy?', *ACS Med Chem Lett*, 3(9), pp. 691–694. doi: 10.1021/ml300221t.

Corpas-Lopez, V. *et al.* (2019) 'Pharmacological Validation of N-Myristoyltransferase as a Drug Target in Leishmania donovani'*, ACS Infect Dis*, 5(1), pp. 111–122. doi: 10.1021/acsinfecdis.8b00226.

Cutter, A. R. and Hayes, J. J. (2015) 'A brief review of nucleosome structure', *FEBS Lett*, 589(20 Pt A), pp. 2914–2922. doi: 10.1016/J.FEBSLET.2015.05.016.

Dalvit, C. *et al.* (2000) 'Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water', *J Biomol NMR*, 18(1), pp. 65–68. doi: 10.1023/a:1008354229396.

Dalvit, C. *et al.* (2001) 'WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability', *J Biomol NMR*, 21(4), pp. 349–359. doi: 10.1023/a:1013302231549.

DNDi (2023) *Visceral leishmaniasis Projects & achievements*. Available at: https://dndi.org/diseases/visceral-leishmaniasis/projects-achievements/ (Accessed: 8 April 2023).

Ember, S. W. J. *et al.* (2014) 'Acetyl-lysine Binding Site of Bromodomain-Containing Protein 4 (BRD4) Interacts with Diverse Kinase Inhibitors', *ACS Chem Biol*, 9(5), pp. 1160–1171. doi: 10.1021/cb500072z.

Faivre, E. J. *et al.* (2020) 'Selective inhibition of the BD2 bromodomain of BET proteins in prostate cancer', *Nature*, 578(7794), pp. 306–310. doi: 10.1038/s41586-020-1930-8.

Filippakopoulos, P. *et al.* (2010) 'Selective inhibition of BET bromodomains.', *Nature*, 468(7327), pp. 1067–73. doi: 10.1038/nature09504.

Filippakopoulos, P. *et al.* (2012) 'Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family', *Cell*, 149(1), pp. 214–231. doi: 10.1016/j.cell.2012.02.013.

Filippakopoulos, P. and Knapp, S. (2012) 'The bromodomain interaction module', *FEBS Lett*, 586(17), pp. 2692–2704. doi: 10.1016/j.febslet.2012.04.045.

Flynn, E. M. *et al.* (2015) 'A subset of human bromodomains recognizes butyryllysine and crotonyllysine histone peptide modifications', *Structure*, 23(10), pp. 1801–1814. doi: 10.1016/j.str.2015.08.004.

Fogg, M. J. and Wilkinson, A. J. (2008) 'Higher-throughput approaches to crystallization and crystal structure determination', *Biochem Soc Trans*, 36(Pt 4), pp. 771–775. doi: 10.1042/BST0360771.

Forte, B. *et al.* (2021) 'Prioritization of Molecular Targets for Antimalarial Drug Discovery', *ACS Infect Dis*, 7(10), pp. 2764–2776. doi: 10.1021/acsinfecdis.1c00322.

Frézard, F. *et al.* (2009) 'Pentavalent Antimonials: New Perspectives for Old Drugs', *Molecules*, 14(7), pp. 2317–2336. doi: 10.3390/molecules14072317.

Fujisawa, T. and Filippakopoulos, P. (2017) 'Functions of bromodomain-containing proteins and their roles in homeostasis and cancer', *Nat Rev Mol Cell Biol*, 18(4), pp. 246–262. doi: 10.1038/nrm.2016.143.

Georgiadou, S. P., Makaritsis, K. P. and Dalekos, G. N. (2015) 'Leishmaniasis revisited: Current aspects on epidemiology, diagnosis and treatment', *J Transl Int Med*, 3(2), pp. 43–50. doi: 10.1515/JTIM-2015-0002.

Giaimo, B. D. *et al.* (2019) 'The histone variant H2A.Z in gene regulation', *Epigenetics Chromatin*, 12(1), p. 37. doi: 10.1186/S13072-019-0274-9.

Gilan, O. *et al.* (2020) 'Selective targeting of BD1 and BD2 of the BET proteins in cancer and immunoinflammation', *Science*, 368(6489), pp. 387–394. doi: 10.1126/science.aaz8455.

Gilbert, K. *et al.* (2022) 'Profiling Sulfur(VI) Fluorides as Reactive Functionalities for Chemical Biology Tools and Expansion of the Ligandable Proteome', *ChemRxiv*. doi: 10.26434/CHEMRXIV-2022-J8B8D-V2.

Gluenz, E., Ginger, M. L. and McKean, P. G. (2010) 'Flagellum assembly and function during the Leishmania life cycle', *Curr Opin Microbiol,* 13(4), pp. 473–479. doi: 10.1016/J.MIB.2010.05.008.

Gossage, S. M., Rogers, M. E. and Bates, P. A. (2003) 'Two separate growth phases during the development of Leishmania in sand flies: implications for understanding the life cycle', *Int J Parasitol*, 33(10), pp. 1027–1034. doi: 10.1016/s0020-7519(03)00142-5.

Hammitzsch, A. *et al.* (2015) 'CBP30, a selective CBP/p300 bromodomain inhibitor, suppresses human Th17 responses', *Proc Natl Acad Sci U S A*, 112(34), pp. 10768–10773. doi: 10.1073/pnas.1501956112.

Harp, J. M. *et al.* (2000) 'Asymmetries in the nucleosome core particle at 2.5 A resolution', *Acta Crystallogr D Biol Crystallogr*, 56(Pt 12), pp. 1513–1534. doi: 10.1107/S0907444900011847.

Hay, D. A. *et al.* (2014) 'Discovery and Optimization of Small-Molecule Ligands for the CBP/p300 Bromodomains', *J Am Chem Soc*, 136(26), pp. 9308–9319. doi: 10.1021/ja412434f.

Hewings, D. S. *et al.* (2011) '3,5-Dimethylisoxazoles Act As Acetyl-lysine-mimetic Bromodomain Ligands', *J Med Chem*, 54(19), pp. 6761–6770. doi: 10.1021/JM200640V.

Hewings, D. S. *et al.* (2013) 'Optimization of 3,5-dimethylisoxazole derivatives as potent bromodomain ligands', *J Med Chem*, 56(8), pp. 3217–3227. doi: 10.1021/jm301588r.

Holm, L. and Sander, C. (1993) 'Protein Structure Comparison by Alignment of Distance Matrices', *J Mol Biol*, 233(1), pp. 123–138. doi: 10.1006/jmbi.1993.1489.

Ito, K., Charron, C. E. and Adcock, I. M. (2007) 'Impact of protein acetylation in inflammatory lung diseases', *Pharmacol Ther*, pp. 249–65. doi: 10.1016/j.pharmthera.2007.06.009.

Jacobson, R. H. *et al.* (2000) 'Structure and function of a human TAFII250 double bromodomain module.', *Science*, 288(5470), pp. 1422–5. doi: 10.1126/science.288.5470.1422.

Jain, K. and Jain, N. K. (2013) 'Novel therapeutic strategies for treatment of visceral leishmaniasis', *Drug Discov Today*, 18(23–24), pp. 1272–1281. doi: 10.1016/J.DRUDIS.2013.08.005.

Jha, P. K. *et al.* (2017) 'HAT2 mediates histone H4K4 acetylation and affects micrococcal nuclease sensitivity of chromatin in Leishmania donovani', *PLoS ONE*, 12(5), p. e0177372. doi: 10.1371/JOURNAL.PONE.0177372.

Jones, N. G. *et al.* (2022) 'Bromodomain factor 5 is an essential regulator of transcription in Leishmania', *Nat Commun*, 13(1), p. 4071. doi: 10.1038/s41467-022-31742-1.

Jung, M. *et al.* (2014) 'Affinity map of bromodomain protein 4 (BRD4) interactions with the histone H4 tail and the small molecule inhibitor JQ1.', *J Biol Chem*, 289(13), pp. 9304–9319. doi: 10.1074/jbc.M113.523019.

Kamiński, D. M. (2014) 'Recent progress in the study of the interactions of amphotericin B with cholesterol and ergosterol in lipid environments', *Eur Biophys J*, 43(10–11), pp. 453–467. doi: 10.1007/S00249-014-0983-8.

Karim, R. M. *et al.* (2022) 'Discovery of Dual TAF1-ATR Inhibitors and Ligand-Induced Structural Changes of the TAF1 Tandem Bromodomain', *J Med Chem*, 65(5), pp. 4182–4200. doi: 10.1021/acs.jmedchem.1c01999.

Kougnassoukou Tchara, P. E., Filippakopoulos, P. and Lambert, J. P. (2020) 'Emerging tools to investigate bromodomain functions', *Methods*, 184, pp. 40–52. doi: 10.1016/J.YMETH.2019.11.003.

Kumar, D. *et al.* (2012) 'Histone H4 lysine 14 acetylation in Leishmania donovani is mediated by the MYST-family protein HAT4', *Microbiology*, 158(Pt 2), pp. 328–337. doi: 10.1099/mic.0.050211-0.

Kumar, D. and Saha, S. (2015) 'HAT3-mediated acetylation of PCNA precedes PCNA monoubiquitination following exposure to UV radiation in Leishmania donovani', *Nucleic Acids Res*, 43(11), pp. 5423–5441. doi: 10.1093/NAR/GKV431.

Kutilek, V. D. *et al.* (2016) 'Integration of Affinity Selection–Mass Spectrometry and Functional Cell-Based Assays to Rapidly Triage Druggable Target Space within the NF-κB Pathway', *J Biomol Screen*, 21(6), pp. 608–619. doi: 10.1177/1087057116637353.

Lalli, M. A. *et al.* (2016) 'Haploinsufficiency of BAZ1B contributes to Williams syndrome through transcriptional dysregulation of neurodevelopmental pathways', *Hum Mol Genet*, 25(7), pp. 1294–1306. doi: 10.1093/hmg/ddw010.

Lange, M. *et al.* (2008) 'Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex', *Genes Dev*, 22(17), pp. 2370–2384. doi: 10.1101/gad.471408.

Laurin, C. M. C. (2019) *Chemical Tools To Probe The Role Of Bromodomains In The Parasite Trypanosoma cruzi*. PhD Thesis, University of Oxford.

Laurin, C. M. C. *et al.* (2021) 'Fragment-Based Identification of Ligands for Bromodomain-Containing Factor 3 of Trypanosoma cruzi', *ACS Infect Dis*, 7(8), pp. 2238–2249. doi: 10.1021/acsinfecdis.0c00618.

Leprohon, P. *et al.* (2015) 'Drug resistance analysis by next generation sequencing in Leishmania', *Int J Parasitol Drugs Drug Resist*, 5(1), pp. 26–35. doi: 10.1016/j.ijpddr.2014.09.005.

Li, Y. *et al.* (2014) 'AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation', *Cell*, 159(3), pp. 558–571. doi: 10.1016/J.CELL.2014.09.049.

Lonsdale, R. and Ward, R. A. (2018) 'Structure-based design of targeted covalent inhibitors', *Chem Soc Rev*, pp. 3816–3830. doi: 10.1039/c7cs00220c.

Luque-Ortega, J. R. and Rivas, L. (2007) 'Miltefosine (hexadecylphosphocholine) inhibits cytochrome c oxidase in Leishmania donovani promastigotes', *Antimicrob Agents Chemother*, 51(4), pp. 1327–1332. doi: 10.1128/AAC.01415-06.

Maity, A. K. and Saha, P. (2012) 'The histone acetyl transferase LdHAT1 from Leishmania donovani is regulated by S-phase cell cycle kinase', *FEMS Microbiol Lett*, 336(1), pp. 57–63. doi: 10.1111/J.1574-6968.2012.02656.X.

Malvezzi, F. *et al.* (2021) 'Phosphorylation-dependent BRD4 dimerization and implications for therapeutic inhibition of BET family proteins', *Commun Biol*, 4(1), p. 1273. doi: 10.1038/S42003-021-02750-6.

Mandava, V. *et al.* (2007) 'Histone modifications in Trypanosoma brucei', *Mol Biochem Parasitol*, 156(1), pp. 41–50. doi: 10.1016/J.MOLBIOPARA.2007.07.005.

Maruyama, T. *et al.* (2002) 'A Mammalian bromodomain protein, brd4, interacts with replication factor C and inhibits progression to S phase.', *Mol Cell Biol*, 22(18), pp. 6509–6520. doi: 10.1128/mcb.22.18.6509-6520.2002.

Mayer, M. and Meyer, B. (1999) 'Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy', *Angew Chem Int Ed Engl*, 38(12), pp. 1784–1788. doi: 10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q.

McNicholas, S. *et al.* (2011) 'Presenting your structures: the CCP4mg molecular-graphics software', *Acta Crystallogr D Biol Crystallogr*, 67(Pt 4), pp. 386–394. doi: 10.1107/S0907444911007281.

Meiboom, S. and Gill, D. (1958) 'Modified Spin-Echo Method for Measuring Nuclear Relaxation Times', *Rev Sci Instrum*, 29(8), pp. 688–691. doi: 10.1063/1.1716296.

Milne, T. A. *et al.* (2002) 'MLL targets SET domain methyltransferase activity to Hox gene promoters', *Mol Cell*, 10(5), pp. 1107–1117. doi: 10.1016/s1097-2765(02)00741-4.

Moerke, N. J. (2009) 'Fluorescence Polarization (FP) Assays for Monitoring Peptide-Protein or Nucleic Acid-Protein Binding', *Curr Protoc Chem Biol*, 1(1), pp. 1–15. doi: 10.1002/9780470559277.ch090102.

Mujtaba, S., Zeng, L. and Zhou, M.-M. (2007) 'Structure and acetyl-lysine recognition of the bromodomain', *Oncogene*, 26(37), pp. 5521–5527. doi: 10.1038/sj.onc.1210618.

Muller, S., Filippakopoulos, P. and Knapp, S. (2011) 'Bromodomains as therapeutic targets', *Expert Rev Mol Med*, 13, p. e29. doi: 10.1017/S1462399411001992.

Musa, A. *et al.* (2022) 'LEISH2b - A phase 2b study to assess the safety, efficacy, and immunogenicity of the Leishmania vaccine ChAd63-KH in post-kala azar dermal leishmaniasis [version 1; peer review: 1 approved]', *Wellcome Open Res*, 7, p. 200.

Nicholls, S. J. *et al.* (2011) 'Efficacy and safety of a novel oral inducer of apolipoprotein A-I synthesis in statin-treated patients with stable coronary artery disease: A randomized controlled trial', *J Am Coll Cardiol*, 57(9), pp. 1111–1119. doi: 10.1016/j.jacc.2010.11.015.

Nicodeme, E. *et al.* (2010) 'Suppression of inflammation by a synthetic histone mimic', *Nature*, 468(7327), pp. 1119–1123. doi: 10.1038/nature09589.

Noel, J. K. *et al.* (2013) 'Development of the BET bromodomain inhibitor OTX015', *Mol Cancer Ther*, (12 (11 Suppl)), p. C244. doi: 10.1158/1535-7163.targ-13-c244.

Paulson, C. N. *et al.* (2018) 'Design, Synthesis, and Characterization of a Fluorescence Polarization Pan-BET Bromodomain Probe', *ACS Med Chem Lett*, 9(12), pp. 1223–1229. doi: 10.1021/acsmedchemlett.8b00380.

Philpott, M. *et al.* (2014) 'Assessing cellular efficacy of bromodomain inhibitors using fluorescence recovery after photobleaching', *Epigenetics Chromatin*, 7, p. 14. doi: 10.1186/1756-8935-7-14.

Picaud, S. *et al.* (2013) 'RVX-208, an inhibitor of BET transcriptional regulators with selectivity for the second bromodomain', *Proc Nat Acad Sci U S A*, 110(49), pp. 19754–19759. doi: 10.1073/pnas.1310658110.

Picaud, S. *et al.* (2015) 'Generation of a Selective Small Molecule Inhibitor of the CBP/p300 Bromodomain for Leukemia Therapy', *Cancer Res*, 75(23), pp. 5106–5119. doi: 10.1158/0008-5472.CAN-15-0236.

Picaud, S. *et al.* (2016) 'Promiscuous targeting of bromodomains by bromosporine identifies BET proteins as master regulators of primary transcription response in leukemia', *Sci Adv*, 2(10), pp. 1–16. doi: 10.1126/sciadv.1600760.

Pinto-Martinez, A. K. *et al*. (2018) 'Mechanism of Action of Miltefosine on Leishmania donovani Involves the Impairment of Acidocalcisome Function and the Activation of the Sphingosine-Dependent Plasma Membrane Ca2+ Channel.', *Antimicrob Agents Chemother*, 62(1), pp. e01614-17. doi: 10.1128/AAC.01614-17.

Rakotomanga, M. *et al.* (2007) 'Miltefosine affects lipid metabolism in Leishmania donovani promastigotes', *Antimicrob Agents Chemother*, 51(4), pp. 1425–1430. doi: 10.1128/AAC.01123-06.

Riggs, A. D., Martienssen, R. A. and Russo, V. E. A. (1996) 'Introduction', in *Epigenetic Mechanisms of Gene Regulation*. Cold Spring Harbour, NY : Cold Spring Harbour Laboratory Press, pp. 1–4.

Ringel, A. E., Tucker, S. A. and Haigis, M. C. (2018) 'Chemical and Physiological Features of Mitochondrial Acylation', *Mol Cell*, pp. 610–624. doi: 10.1016/j.molcel.2018.10.023.

Ritagliati, C. *et al.* (2016) 'Glycosomal bromodomain factor 1 from Trypanosoma cruzi enhances trypomastigote cell infection and intracellular amastigote growth.', *Biochem J*, 473(1), pp. 73–85. doi: 10.1042/BJ20150986.

Scarpini, S. *et al.* (2022) 'Visceral Leishmaniasis: Epidemiology, Diagnosis, and Treatment Regimens in Different Geographical Areas with a Focus on Pediatrics', *Microorganisms*, 10(10), p. 1887. doi: 10.3390/MICROORGANISMS10101887.

Schulz, D. *et al.* (2015) 'Bromodomain Proteins Contribute to Maintenance of Bloodstream Form Stage Identity in the African Trypanosome', *PLoS Biol*, 13(12), p. e1002316. doi: 10.1371/journal.pbio.1002316.

Schulz, M. N., Landström, J. and Hubbard, R. E. (2013) 'MTSA - A Matlab program to fit thermal shift data', *Anal Biochem*, 433(1), pp. 43–47. doi: 10.1016/j.ab.2012.10.020.

Shorstova, T., Foulkes, W. D. and Witcher, M. (2021) 'Achieving clinical success with BET inhibitors as anti-cancer agents', *Br J Cancer*, 124(9), pp. 1478–1490. doi: 10.1038/s41416-021-01321-0.

Siegel, T. N. *et al.* (2009) 'Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei', *Genes Dev*, 23(9), pp. 1063–1076. doi: 10.1101/gad.1790409.

Singh, M. *et al.* (2007) 'Structural Ramification for Acetyl-Lysine Recognition by the Bromodomain of Human BRG1 Protein, a Central ATPase of the SWI/SNF Remodeling Complex', *ChemBioChem*, 8(11), pp. 1308–1316. doi: 10.1002/cbic.200600562.

Staneva, D. P. *et al.* (2021) 'A systematic analysis of Trypanosoma brucei chromatin factors identifies novel protein interaction networks associated with sites of transcription initiation and termination', *Genome Res*, 31(11), pp. 2138–2154. doi: 10.1101/gr.275368.121.

Steegmaier, M. *et al.* (2007) 'BI 2536, a Potent and Selective Inhibitor of Polo-like Kinase 1, Inhibits Tumor Growth In Vivo', *Curr Biol*, 17(4), pp. 316–322. doi: 10.1016/j.cub.2006.12.037.

Stonestrom, A. J. *et al.* (2015) 'Functions of BET proteins in erythroid gene expression', *Blood*, 125(18), pp. 2825–2834. doi: 10.1182/blood-2014-10-607309.

Stuart, K. *et al.* (2008) 'Kinetoplastids: related protozoan pathogens, different diseases', *J Clin Invest*, 118(4), pp. 1301–1310. doi: 10.1172/JCI33945.

Su, D. *et al.* (2012) 'Structural basis for recognition of H3K56-acetylated histone H3–H4 by the chaperone Rtt106', *Nature*, 483(7387), pp. 104–107. doi: 10.1038/nature10861.

Sundar, S. and Chakravarty, J. (2015) 'An update on pharmacotherapy for leishmaniasis.', *Expert Opin Pharmacother*, 16(2), pp. 237–52. doi: 10.1517/14656566.2015.973850.

Sunter, J. and Gull, K. (2017) 'Shape, form, function and Leishmania pathogenicity: from textbook descriptions to biological understanding', *Open Biol*, 7(9), p. 170165. doi: 10.1098/RSOB.170165.

Sutanto, F., Konstantinidou, M. and Dömling, A. (2020) 'Covalent inhibitors: a rational approach to drug discovery', *RSC Med Chem*, 11(8), pp. 876–884. doi: 10.1039/D0MD00154F.

Tamkun, J. W. *et al.* (1992) 'brahma: A regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2 SWI2', *Cell*, 68(3), pp. 561–572. doi: 10.1016/0092-8674(92)90191-e.

Tanaka, M. *et al.* (2016) 'Design and characterization of bivalent BET inhibitors', *Nat Chem Biol*, 12(12), pp. 1089–1096. doi: 10.1038/nchembio.2209.

Theodoulou, N. H., Tomkinson, N. C. O., *et al.* (2016) 'Clinical progress and pharmacology of small molecule bromodomain inhibitors', *Curr Opin Chem Biol*, 33, pp. 58–66. doi: 10.1016/J.CBPA.2016.05.028.

Theodoulou, N. H., Bamborough, P., *et al.* (2016) 'Discovery of I-BRD9, a Selective Cell Active Chemical Probe for Bromodomain Containing Protein 9 Inhibition', *J Med Chem*, 59(4), pp. 1425–1439. doi: 10.1021/acs.jmedchem.5b00256.

Thompson, M. (2009) 'Polybromo-1: the chromatin targeting subunit of the PBAF complex', *Biochimie*, 91(3), pp. 309–319. doi: 10.1016/J.BIOCHI.2008.10.019.

Umehara, T. *et al.* (2010) 'Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain', *J Biol Chem*, 285(10), pp. 7610–7618. doi: 10.1074/jbc.M109.062422.

VanDemark, A. P. *et al.* (2007) 'Autoregulation of the Rsc4 Tandem Bromodomain by Gcn5 Acetylation', *Mol Cell*, 27(5), pp. 817–828. doi: 10.1016/J.MOLCEL.2007.08.018.

Villanova, G. V. *et al.* (2009) 'Trypanosoma cruzi bromodomain factor 2 (BDF2) binds to acetylated histones and is accumulated after UV irradiation', *Int J Parasitol*, 39(6), pp. 665–673. doi: 10.1016/J.IJPARA.2008.11.013.

Wadhone, P. *et al.* (2009) 'Miltefosine promotes IFN-gamma-dominated anti-leishmanial immune response', *J Immunol*, 182(11), pp. 7146–7154. doi: 10.4049/JIMMUNOL.0803859.

WHO (2010) 'Control of the leishmaniases: report of a meeting of the WHO Expert Commitee on the Control of Leishmaniases'. Geneva, 22-26 March 2010, pp. 1–186. Available at: https://apps.who.int/iris/bitstream/handle/10665/44412/WHO_TRS_949_eng.pdf?sequence=1&isAllowed=y.

WHO (2022a) *Leishmaniasis Key Facts*, *WHO Fact Sheets*. Available at: https://www.who.int/news-room/fact-sheets/detail/leishmaniasis (Accessed: 14 December 2022).

WHO (2022b) *Neglected Tropical Diseases*, *Neglected Tropical Diseases*. Available at: https://www.who.int/health-topics/neglected-tropical-diseases#tab=tab_1 (Accessed: 14 December 2022).

Wu, Q. *et al.* (2019) 'A chemical toolbox for the study of bromodomains and epigenetic signaling', *Nat Commun*, 10(1), p. 1915. doi: 10.1038/s41467-019-09672-2.

Wyllie, S., Cunningham, M. L. and Fairlamb, A. H. (2004) 'Dual Action of Antimonial Drugs on Thiol Redox Metabolism in the Human Pathogen Leishmania donovani', *J Biol Chem,* 279(38), pp. 39925–39932. doi: 10.1074/JBC.M405635200.

Yadav, Y. *et al.* (2022) 'Uncovering the Domain-Specific Interactome of the TAF1 Tandem Reader Using Site-Specific Azide-Acetyllysine Photochemistry', *Biochemistry*. doi: 10.1021/ACS.BIOCHEM.2C00140.

Yang, X. *et al.* (2017) 'Recognition of hyperacetylated N-terminus of H2AZ by TbBDF2 from Trypanosoma brucei', *Biochem J*, 474(22), pp. 3817–3830. doi: 10.1042/BCJ20170619.

Zaware, N. and Zhou, M.-M. (2019) 'Bromodomain biology and drug discovery', *Nat Struct Mol Biol*, 26(10), pp. 870–879. doi: 10.1038/s41594-019-0309-8.

Zeng, L. *et al.* (2010) 'Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b', *Nature*, 466(7303), pp. 258–262. doi: 10.1038/nature09139.

Zengerle, M., Chan, K. H. and Ciulli, A. (2015) 'Selective Small Molecule Induced Degradation of the BET Bromodomain Protein BRD4', *ACS Chem Biol*, 10(8), pp. 1770–1777. doi: 10.1021/ACSCHEMBIO.5B00216.