Supertertiary structure analysis and comparison of BET proteins upon bivalent binding



Ailefeila Aihemaiti

Jesus College

University of Oxford

November 2020

Abstract

BET proteins (BRD2/3/4/T) contain two bromodomains (BD1 and BD2) and an extra terminal (ET) domain. They act as nuclear scaffolds by recognising acetylated histones and recruiting transcriptional complexes. Although all BETs share high levels of sequence conservation, each has different biological functions. Available BET inhibitors do not distinguish between BET members, leading to their toxicity effects. Here, my preliminary data suggest that the domain topology of BETs affects how they engage small molecule inhibitors and histone peptides. Secondary structure prediction analysis reveals that BET conserved domains are linked by flexible regions of variable length. Simultaneous engagement of both BDs by bivalent inhibitors differentially distorts each protein. Chimeric linker mutants show changes in binding behaviour upon bivalent inhibitor engagement. While the BDs of the BRDT simultaneously engage a histone H4 peptide bearing seven acetylated lysine residues, replacement of the inter-domain linker causes loss of binding by the BD2. These observations led to the designing of a set of tandem peptidic inhibitors containing histone and viral peptides that can bind to both BD and ET domains to differentiate each BET. Biophysical evidence showed that one of the tandem peptidic inhibitors binds to BRD4 with nanomolar affinity in *cis* binding mode while showing micromolar affinity towards BRD2. These observations offer a new molecular perspective on the importance of BET multidomain engagement and give insight into designing BET selective bivalent inhibitors.

Contents

Abstract **2**

Statement of contributions **7**

Acknowledgements **8**

Abbreviations **11**

Chapter1. Introduction **16**

1.1. Histone acetylation in transcription 16

1.1.1. Epigenetics and histone acetylation 16

1.1.2. Bromodomains - readers of histone acetyl marks 19

1.2. Bromo and extra terminal (BET) family 23

..... 24

1.2.1. BET bromodomains 24

1.2.2. BET extra terminal (ET) domain 27

1.2.3. BET C-terminal motif (CTM) 29

1.2.4. BET inter-domain linker region 30

1.3. Biological relevance of BET proteins 31

1.3.1. BETs in transcription 31

1.3.2. BETs in cancer 33

1.3.3. BETs in inflammation 35

1.3.4. BETs in viral infection 36

1.4. Discovery of BET bromodomain inhibitors 37

1.4.1. BET bromodomain monovalent inhibitors 37

1.4.2. BET bromodomain bivalent inhibitors 40

1.4.3. BET proteins and PROTACs 42

1.5. Clinical efficacy of BET bromodomain inhibitors 43

1.5.1. Toxicity effect of BET bromodomain monovalent inhibitors 46

1.5.2. Resistance of BET bromodomain monovalent inhibitors 47

1.6. Summary and project aim 48

Chapter 2. BET topology and conformational change upon small molecule binding **50**

2.1. Introduction 50

2.2. BET proteins contain multiple domains linked by a variable flexible linker 55

2.3. BET tandem BDs adopt an extended conformation in solution 56
2.4. BET tandem BDs are differentially compacted when engaged by bivalent inhibitors. 60
2.4.1. Bivalent inhibitors simultaneously bind two BDs 60
2.4.2. Bivalent inhibitor binding mode on BET tandem BDs 62
2.4.3. BvX-n inhibitor led to BET tandem BD differential compaction 66
2.5. Bivalent inhibitor PEG linker length and warheads are both important determinants that changed BRD2(1:2) binding mode 70
2.5.1. Compound BvA-7, but not BvA-3, leading to BRD2(2) oligomerization is the main cause of BRD2(1:2) oligomerization with BvA-7 but not with BvA-3 70
2.5.2. Selectivity of BvB-1 one warhead towards BRD2(2) comparing to the other warhead might be the main reason for why BvB-1 can compact the BRD2(1:2) to the greatest extend 75
2.6. Switching the BRD2(1:2) linker region changes the compound binding behaviour 79
2.7. Discussion 80

Chapter 3. Binding mode of BET tandem bromodomains on human histone H4 tail 84

3.1. Introduction 84
3.2. BRD2(1:2) employs both BDs binding to Kac-XXXXXX-Kac histone H4 peptide ... 86
3.2.1. BRD2(1:2) has improved binding to distant diacetylated H4 peptide compared with that by either BRD2(1:2)$^{WT/NF}$ or BRD2(1:2)$^{NF/WT}$ 86
3.2.2. BRD2(1:2) binds to distant diacetylated H4 peptide in a 1:2 molar ratio 87
3.3. Each BET tandem BDs engages with polyacetylated histone H4 peptide in different modes 90
3.3.1. BRDT(1:2) and BRD4(1:2) bind to polyacetylated H4 peptide bivalently while BRD2(1:2) binds in 1:2 protein:ligand binding mode 90
3.3.2. The BD linker region contributes to bivalent binding in polyacetylated H4 engagement 95
3.3.3. BRDT(1) and the linker region play the dominant role in the bivalent binding of BRDT(1:2) to polyacetylated H4 peptide 97
3.4. Discussion 100

Chapter 4. Characterisation of selective BET-BD2ET bivalent peptidic inhibitors 103

4.1. Introduction 103
4.2. Discovery of bivalent peptidic inhibitor that simultaneously tethers BD2 and ET 104

4.2.1. Bivalent peptidic inhibitor designing and production....104
4.2.2. Bivalent peptidic inhibitor with a 9 or 12 amino acid linker length leads to BRD4-BD2ET compaction....108
4.2.3. Bivalent peptidic inhibitor $H4_{10}$Vn series which only contains K5, K8 cannot compact BRD4-BD2ET....110
4.3. Bivalent peptidic inhibitor $H4_{20}$V4_ac is selective towards BRD4-BD2ET relative to BRD2-BD2ET....113
4.3.1. $H4_{20}$V4_ac compacts BRD4-BD2ET but did not compact BRD2-BD2ET....113
4.3.2. $H4_{20}$V4_ac shows 100-fold more selectivity to BRD4-BD2ET over BRD2-BD2ET....116
4.4. Discussion....122
Chapter 5. Conclusion and future work....127
5.1. Conclusion....127
5.2. Future work....130
Chapter 6. Material and Methods....134
6.1. Cloning....134
6.1.1. Cloning of tandem BET BD and BET-BD2ET constructs....134
6.1.2. Cloning of tandem histone H4 - viral peptides....134
6.2. Protein expression and purification....135
6.2.1. Protein expression and purification of tandem BET BDs....135
6.2.2. Protein expression and purification of BET-BD2ET constructs....139
6.2.3. Protein expression, purification and acetylation of histone H4 peptide and tandem histone H4 - viral peptides....143
6.3. Isothermal titration calorimetry....153
6.3.1. ITC experiments on BET tandem BDs with bivalent compounds (Chapter 2)....153
6.3.2. ITC experiments on BET tandem BDs with acetylated peptides (Chapter3)....154
6.3.3. ITC experiments on BET tandem BDs with tandem histone H4 – viral peptides (Chapter4)....155
6.4. Analytical ultracentrifugation....155
6.4.1. AUC experiments on BET tandem BDs with bivalent compounds (Chapter 2)....155
6.4.2. AUC experiments on BET tandem BDs with acetylated peptides (Chapter3)....156

6.4.3. AUC experiments on BET tandem BDs with tandem histone H4–viral peptides (Chapter4) 157
6.5. Small-angle X-ray scattering (SAXS) 157
6.6. SPOT Array 158
6.7. Biolayer interferometry 161
6.8. Custom Peptide Synthesis 162
Reference. 163
Appendix 174

Statement of contributions

I, Ailefeila Aihemaiti, am the sole author for this thesis that was written under the supervision of my primary supervisor Panagis Filippakopoulos and was not written for publication.

All the BET constructs used in the experiments were designed and cloned by Ailefeila using SGC entry clone full-length BET protein constructs as templates. The tandem peptide constructs were also designed and cloned by Ailefeila using SGC entry clone full-length histone H4 constructs and MLV constructs. The purification of the proteins and tandem peptides used in the experiments was performed by Ailefeila at the SGC university of Oxford. All the cloning and purification processes were conducted under the supervision of Dr. Pavel Stavisky and Dr. Sarah Picaud.

Experiment setup, data collection, and analysis of all the AUC, ITC, BLI, and SPOT array assays were conducted by Ailefeila under the supervision of Dr. Panagis Filippakopoulos. Data collection and data analysis of the SAXS assay were conducted by Ailefeila under the supervision of Dr. Joseph Newman.

The BET bivalent inhibitors (BvA-n and BvB-n) of variable PEG linker length used in the experiments were designed and synthesised by Dr. Minoru Tanaka at the Jun Qi lab, Harvard University. Alpha Screen assay of the bivalent inhibitors toward BRDT(1) was designed and performed by Dr. Justin M Roberts at the Jun Qi lab, Harvard University.

Acknowledgements

I am truly grateful for the scholarship offered by NDM studentship and Chinese scholarship council for giving me the opportunity to study abroad. Thank you to my primary supervisor, Panagis Filippakopoulos, for accepting me to work on this exciting project and providing valuable insight throughout my DPhil. I am grateful to my second supervisor, Opher Gileadi, who encouraged me through my hard times and reminded me to see the bigger picture. Thank you to Sarah Picaud for showing me how to be a dedicated, meticulous, and hardworking scientist. I have great respect for her work. I am also grateful to my other two supervisors, Martin Walsh and Patra Lukacik at diamond, who took care of me when I was partially working at diamond. I especially want to thank our head of graduate study Robert Gilbert for being supportive and carefully listening to what students' needs.

I have been privileged to work in SGC at Oxford. SGC is like a big family, and everyone is so supportive of each other. Projects are discussed through internal seminars, and a variety of helpful advice is proposed. Ideas are shaped through interactions. People are really nice.

I also deeply appreciate the people who train me for experimental instruments. Tomas Christott was incredibly helpful in teaching me how to run ITC experiments and BLI, MSD experiments. Joe Newman was really helpful in showing me how to process SAXS data. My primary supervisor helped me a lot in running experiments as well. I particularly appreciate the patience of Panagis, teaching me how to run AUC experiment and process the data, even helping me with coding R scripts.

The input from all of the members of our group has been amazing over the past few years. Pavel Savitsky taught me how to clone, Sarah taught me how to purify the proteins. Beth and Chen yi are always important emotional support. I couldn't have done this work without the help of my group members. I am also grateful to have Angeline and Lia in the GI group as my

friends. They are both strong, independent, and kind women that I look up to. They helped me go through my weakest moments.

Oxford University sports centre added an incredible dimension to my time in Oxford. My frustrations from work or life difficulties always get soothed by going swimming. I learnt to swim at Oxford, which was a great achievement for me. I thank my friends Zhuoyao, Zongfu, and Laiyin for accompanying me almost every swimming time. I also enjoyed university vibe cycling classes, which were an effective way of relieving stresses and the benefits from college life, which gave me the opportunities to get know different students studying different subjects.

As the main source of my inner strength, I especially want to thank my parents. They created a good environment for my pursuit of academic study. I gained my interest in science from my father, who was a Physics professor at Xinjiang University back home. He always told me to be curious, be strong but don't forget to enjoy life. After he suddenly passed away during the second year of my DPhil, I felt like I had lost my entire world. I saw how devastated my mom was. But she was strong. She comforted me to be strong and finish my studies when I at first decided to drop off and go back to be with her. I was afraid she might feel lonely. It was hard, but we have other family members to support my mom, which I am greatly thankful for. I thank my older brother, my four aunts, and all the other family members for being supportive and loving during our hard times. I am also grateful to my then-boyfriend, now-husband, for being so supportive and loving throughout the year, even though we were separated by a long distance. I am so lucky to have these wonderful people in my life, and they taught me the importance of family and love.

I learnt a lot during these four years, not only from the aspects of work but also from social life. I am so grateful to experience the life of the UK and meet different people all over the world. These experiences gave me valuable insights into human rights. In some places in the world, basic human rights are sabotaged, and the sad part is the people living there have to put up with

it until the end of their lives, not knowing what basic human rights are. I hope one day, I can be capable enough to contribute to human society not only from the aspect of natural sciences but also from the aspect of humanity.

Abbreviations

AML	Acute myeloid leukaemia
AP-MS	Affinity purification mass spectrometry
ASCL1	Achaete-Scute family BHLH transcription factor 1
ATAD2	ATPase family, AAA domain containing protein 2
AUC	Analytical Ultracentrifugation
BAZ2A/B	Bromodomain adjacent to zinc finger domain protein 2A/B
BCL2	B-cell lymphoma 2
BD	Bromodomain
BET	Bromodomain and extra terminal domain
BLI	Bio-layer interferometry
B-ALL	B-cell acute lymphoblastic leukaemia
BRD2/3/4/7/9	Bromodomain-containing protein 2/3/4/9
BRD2(1)	BRD2 N-terminal first bromodomain
BRD2(2)	BRD2 N-terminal second bromodomain
BRD2(1:2)	BRD2 tandem bromodomain
$BRD2(1:2)^{WT/NF}$	BRD2 tandem bromodomain with second bromodomain mutation
$BRD2(1:2)^{NF/WT}$	BRD2 tandem bromodomain with first bromodomain mutation
BRET	Bioluminescence resonance energy transfer
BRPF1	Bromodomain and PHD Finger-Containing protein 1
BRDT	Bromodomain testis-specific protein
BRDT(1)	BRDT N-terminal first bromodomain
BRDT(2)	BRD2 N-terminal second bromodomain
BRDT(1:2)	BRDT tandem bromodomain

BRDT(1:2)$^{WT/NF}$	BRDT tandem bromodomain with second bromotomain mutation
BRDT(1:2)$^{NF/WT}$	BRDT tandem bromodomain with first bromotomain mutation
BRWD3	Bromodomain and WD repeat domain containing 3
BSA	Bovine serum albumin
CCAT1	Colon cancer-associated transcript-1
CDK9	Cyclin-dependent kinase 9
CECR2	Cat eye syndrome chromosome region, candidate 2
CHD4	Chromodomain helicase DNA binding protein 4
ChIP-seq	Chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing
CREBBP	cAMP-response element binding protein (CREB)-binding protein
CRL2/4	Cullin-ring E3 ubiquitin ligase 2/4
CTM	C-terminal domain
DAMMIN	Dummy atom model minimisation
DMSO	Dimethyl sulfoxide
EBV	Epstein-Barr virus
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
ES	Electrospray
ET	Extra terminal
FD	Fabry disease
FALZ	Alternative name BPTF, bromodomain PHD finger transcription factor
FOXG1	Forkhead box G1
Gb3	Globotriaosylceramide
GCN5L	General control of amino-acid synthesis 5-like

GLTSCR1	Glioma tumour suppressor candidate region gene 1protein
GSK	GlaxoSmithKline
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
HER3	Human epidermal growth factor receptor 3
HIV	Human immunodeficiency virus
HPLC	High-performance liquid chromatography
HPV	Human papillomavirus
IL-6	Interleukin-6
IC_{50}	Half maximal inhibitory concentration
IPTG	Isopropyl-β-d-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
JAK/STAT	Janus kinase/signal transducer and activator of transcription
JMJD6	Jumonji domain containing 6 protein
Kac	Acetylated lysine residue
K_D	Dissociation constant
KIAA1240	Alternative name ATAD2B, AAA domain containing 2B
KSHV	Kaposi's sarcoma-associated herpesvirus
LANA	Latency-associated nuclear antigen
LB	Lysogeny broth
LIC	Ligation independent cloning
MBD	Methyl-CpG-binding domain
MLL	Mixed lineage leukaemia
MLV	Murine leukaemia virus

Mw	Molecular weight
MWCO	Molecular weight cut-off
MS	Mass spectrometry
MST	Microscale thermophoresis
Mw	Molecular weight
NF-κB	Nuclear Factor-κ-light-chain-enhancer of activated B cells
NIPBL	Nipped-B-like protein
NMC	NUT midline carcinoma
NMR	Nuclear magnetic resonance
NSD3	Nuclear receptor binding SET domain protein 3
NUT	Nuclear protein in testis
OD	Optical density
P-TEFb	Positive transcription elongation factor
PB1	Polybromo-1
PCAF	P300-CREBBP associated factor
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with tween
PDB	Protein data bank
PEG	Polyethylene glycol
PHD	Plant homeodomain
PHIP	Pleckstrin homology domain interacting protein
PP2A	Protein phosphatase 2
PI3K	Phosphoinositide 3-kinases
Pol II	RNA polymerase II

PP2A	Protein phosphatase 2
PROTACs	Proteolysis targeting chimaeras
PTM	Post-translational modification
RUNX3	Runt-related transcription factor 3
SAXS	Small angle X-ray scattering
SEC	Size exclusion chromatography
seRNA	Super-enhancer RNA
SGC	Structural Genomic Consortium
SMARCA2	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 2
SMARCA4	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 4
SOX9	SRY-Box transcription factor 9
SPR	Surface plasmon resonance
SWI/SNF	Switch/sucrose non-fermentable
TAF1	TATA-box binding protein associated factor 1
TGF-β	Transforming growth factor beta
Tip60	Tat-interactive protein 60
TRIM28	Tripartite motif containing 28
TRIM66	Tripartite motif containing 66
TIF1α	Transcription intermediary factor α
WDR9	WD repeat domain 9
WHSC1	Wolf-Hirschhorn syndrome candidate 1
ZMYND11	Zinc finger MYND domain-containing protein 11

Chapter1. Introduction

1.1. Histone acetylation in transcription

1.1.1. Epigenetics and histone acetylation

Epigenetics is the study of changes in heritable gene expression that does not involve alteration in the DNA sequence[1]. Remodelling of chromatin is currently recognised as the main way of regulating gene expression without a change in the DNA code, and hence is often mentioned in epigenetics. Chromatin is a complex of histone proteins and DNA. The positively charged nature of histone proteins facilitates their interaction with negatively charged DNA. Changes in this interaction state, usually caused by the remodelling process of chromatin, will result in a change of accessibility to DNA by transcription factors, thereby causing a change in gene expression. There are several ways to achieve chromatin remodelling. For example, post-transcriptional modifications (PTMs) of histones, ATP-dependent chromatin remodelling and DNA methylation. PTMs are chemical modifications of a protein made by enzymes and this usually refers to the addition of small functional groups to the protein via acetylation, methylation, phosphorylation and biotinylation[2]. PTMs of histones usually occur on the N-terminal histone tails that protrude from the packed core of the nucleosome[3]. Addition of these functional groups to the histone tail may change the overall shape of the nucleosome, resulting in the change in the accessibility of DNA. It is possible that during mitosis, modified histones may act as template and influence the newly formed histones to retain their altered form, and by doing so, ensure that the divided new cells keep the maintained transcription program. However, the exact mechanism of how the newly divided cells inherit the histone-modified state is poorly understood[4]. ATP-dependent chromatin remodelling refers to translocation of nucleosomes by ATP-dependent chromatin-remodelling complexes. A common feature of these protein complexes is that they all contain a subunit capable of hydrolysing ATP. Energy from the hydrolysis of ATP allows these remodelling complexes to reposition nucleosomes

along the DNA and facilitate exchange of histone variants by ejecting or assembling histones on the DNA, resulting in the change in accessibility of DNA. The examples of known chromatin remodelling complexes are the switch/sucrose non-fermentable (SWI/SNF), ISWI family proteins, and Mi-2 protein families. DNA methylation is the addition of methyl groups to the DNA, converting cytosine to 5-methylcytosine. This usually happens at CpG sites, and, when promoter regions of genes are methylated, it causes the repression of gene expression by blocking the binding of transcription factors to these sites[5,6]. In addition, methylated promoter regions can tether the proteins containing methyl-CpG-binding domain (MBD) to indirectly interact with histone deacetylase complexes, leading to further gene expression silencing (**Figure 1.1**).

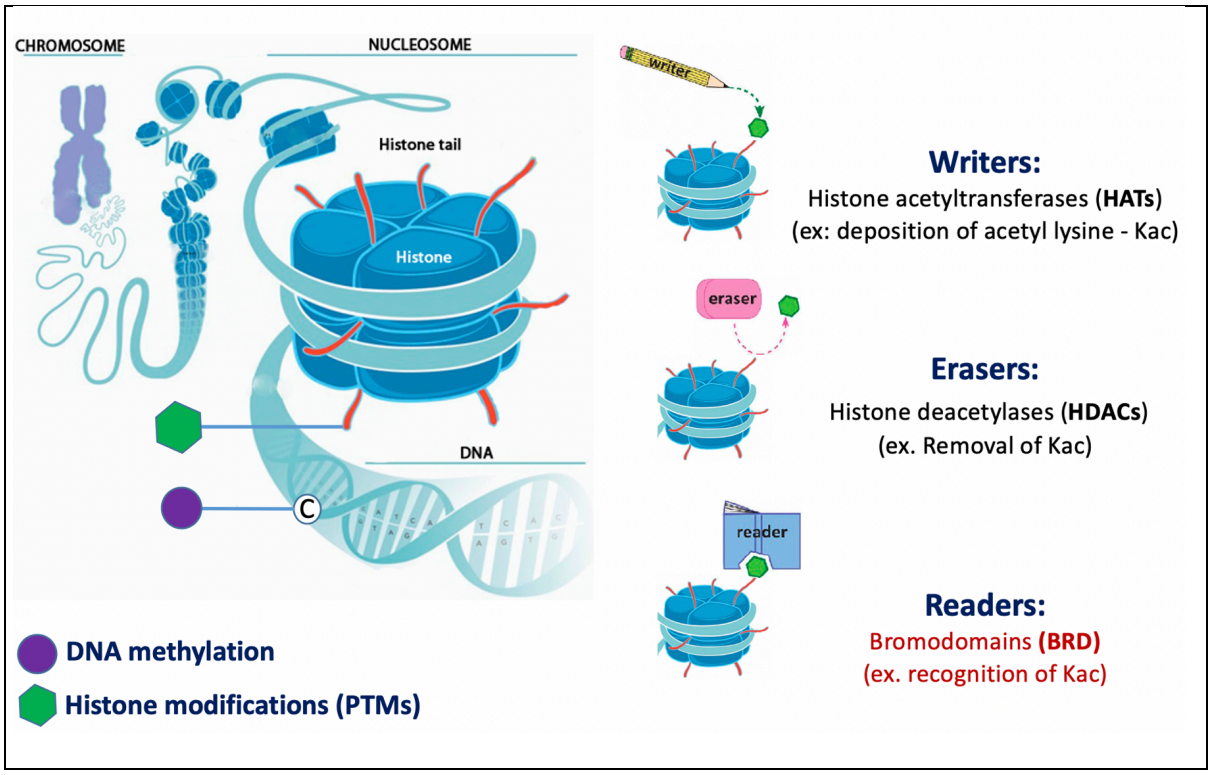


Figure 1.1. Chromosome structure and dynamic regulation of histone modifications. (Adapted from Richard E. Ballermann, http://www.mrballermann.com, all rights reserved). The protruded N terminal tails of histone proteins (red lines) are subjected to post-translational modifications by 'Writers', these modifications can be recognised by "Readers", and can be removed by "Erasers".

Chromatin remodelling by PTMs of histones by addition of phosphate, acetyl or methyl groups on histone tails is the most well understood system. The PTMs of histone are dynamically regulated by two enzymes referred to as 'writers' and 'erasers'. 'Writers' are enzymes that catalyse the addition of acetyl, methyl or other chemical groups to histone tails, for example SUV39h or NSD1, which have histone methyltransferase (HMT) activity. 'Eraser' are enzymes that are involved in the removal of modifications from histone tails, for example LSD1 or Jumonji domain containing 3 protein (JMJD3), which have histone demethylase (HDMs) activity. 'Readers' are proteins, that are able to recognise and bind to specific modifications, for example, proteins that contain a plant homeodomain (PHD), such as TCF19P or HF1. Each modification may have several 'writers', 'readers' and 'erasers' (**Table 1.1**). Emerging evidence has demonstrated that PTMs of histones play a vital role in gene expression both in normal and pathological conditions, and different inhibitors have been designed to block the 'writer' or 'eraser' activities. Among them, the histone deacetylase (HDAC) HMT and DMT inhibitors are the best studied. 'Readers' have also since emerged as a therapeutic target.

Chromatin modification	Residues modified	Writer	Reader	eraser
Acetylation	H3 (K9, K14, K18, K23, K56) H4 (K5, K8, K12, K16)	HATs (e.g. HAT1, CBP/p300, TIP60, PCAF)	Bromodomains	HDAC (e.g. HDAC1, HDAC2, SIRT1)
Methylation	H3 (K4, K9, K27, K36, K79, R2, R8, R17, R26) H4 (K20, R3)	HMTs (e.g. SUV39H, SET1, NSD1)	PHD domain, Chromodomain Tudor domain, MBT domain	DMT (e.g. LSD1, JHMD1, JMJD3)
Phosphorylation	H3 (T3, S28) H4 (S1)	Kinase (e.g. Haspin, MSK1, MSK2)	14-3-3	Phosphatase

Table 1.1. Different classes of modifications identified on histones.

Among the PTMs, the addition of the acetyl group to lysine residues is one of the most commonly occurring PTMs[7] in proteins. A research article published in 1976 was the first study to reveal that acetylation of lysine residues (Kac) on histone tails is associated with an open chromatin architecture[8]. Kac neutralizes the positively charged histone proteins leading to loss of the tight interaction with DNA, and therefore remodels chromatin to an open architecture[9]. When this happens, transcription factors can directly bind to the DNA and initiate the transcription. In this mode, acetylation of lysine residues on histones has a direct effect on DNA accessibility. In the other mode, acetylation of lysine residues on histones could have an indirect effect where modified histones create a binding site for transcription machinery to activate the gene expression or a binding site for histone deacetylase complexes to repress the gene expression. In later studies, not surprisingly, acetylation of lysine residues on histones has been reported to be involved in protein-protein interaction[2], DNA repair[10] and compaction of chromatin[11]. Deregulation of the histone acetylation pattern has been linked to diverse diseases and driving of the aberrant expression of specific genes resulting in tumorigenesis. Both 'writers' and 'readers' of histone acetyl marks appeared as attractive cancer targets in earlier studies, especially 'erasers' (HDAC) inhibitors, which were successfully developed as drugs that are currently in clinical trials or approved for treatment of cancer and other conditions[12-16]. However, a problem arose regarding the specificity of approved 'eraser' (HDAC) inhibitors as they are largely non-selective. Since then, additional studies have emerged that target the 'readers' as a new therapeutic target, and among these, bromodomains have been successfully targeted by small molecule inhibitors[17-21].

1.1.2. Bromodomains - readers of histone acetyl marks

The bromodomain (BD) was named after the discovery of the Drosophila gene Brahma in 1992[22]. As one of the 'readers', a BD initiates the recognition of Kac and is the only 'reader'

module that interacts with Kac. All BD modules are highly conserved with a structural four-helix motif that is composed of 120 residues that vary in sequence. The BD family in the human proteome consists of eight subfamilies of 46 diverse proteins that contain a total of 61 BDs (**Figure 1.2**). BD proteins have been reported to be involved in the regulation of transcription programmes, and aberrant expression of these proteins could lead to the development of several aggressive types of cancer.

In the structure-based classification of the BD family, subfamily 1 contains chromatin remodelling factor (CECR2)[23], acetyltransferase containing proteins (PCAF)[24], transcription factor (FALZ)[25] and the general control of amino-acid synthesis protein 5-like 2 (GCN5L)[26]. CECR2 has been shown to interact with histone H3 peptides (K9ac and K14ac)[27]; this interaction can be attenuated by phosphorylation at S10, a PTM for mitotic progression. PCAF has also been shown to interact with histone H3 (K9ac, K14ac, and K36ac)[27,28] and histone H4 (K8ac, K16ac, and K20ac)[27,28]. In addition, PCAF is known to be involved in stimulation of the integrated human immunodeficiency virus (HIV)-1 genome transcription process by interacting with HIV-1 trans-activator of transcription (Tat)[28]. Subfamily 2 contains the bromo and extra terminal (BET) proteins that are the focus of our project and will be discussed in detail in the next section. Subfamily 3 contains the HAT enzyme CREB binding protein (CREBBP), EP300 protein[29], transcription regulator (BRD8B)[30], pleckstrin homology domain interacting protein (PHIP) (domain 2) protein, which is related to insulin signalling[31], and BD and WD repeat domain containing 3 (BRWD3) protein, which is related to the Janus kinase/signal transducer and activator of transcription (JAK/STAT) pathway[32]. CREBBP has been shown to interact with histone H2B[27], histone H3 (K9ac, K14ac, K36ac and K56ac)[33] and histone H4 (K12ac, K20ac and K44ac)[27,28]. CREBBP has also been shown to exhibit affinity towards acetylated p53 (K382ac)[28] and to be involved in DNA damage response by inducing p53-dependent transcriptional activation of the cyclin-dependent kinase inhibitor p21[34].

Subfamily 4 contains BD-containing protein 1 (BRD1)[35], transcription regulators (BRD7)[36], BD and PHD finger-containing protein (BRPF1)[37], KIAA1240 protein, ATPase family, AAA domain containing protein 2 (ATAD2) protein[38], and BD-containing 9 (BRD9) protein. Proteins of this family have been described to be involved in disease, but their interactions with acetylated histones are poorly studied. The proteins in subfamily 5 all contain tandem a PHD–BD that is reported to interact with peptides to cooperatively help increase binding affinity[39]. The members of this family include the transcriptional repressor tripartite motif containing 66 (TRIM66)[40], BD adjacent to zinc finger domain protein 2A/B (BAZ2A)[41], BAZ2B[42], the transcriptional regulator transcription intermediary factor α (TIF1α)[39], SP100[43] and nuclear body protein SP140[44]. Subfamily 6 contains histone methyltransferase MLL[45] and the transcriptional coregulator TRIM28[46]. However, no acetylated lysine dependent interactions were reported for this family. Subfamily 7 contains transcriptional repressor zinc finger MYND domain-containing protein 11 (ZMYND11)[47], transcription initiators RNA polymerase II TATA box binding protein associated factor (TAF1)[48], TAF1L[49], insulin signalling related protein PHIP domain 1[31] and chromatin remodelling factors WD repeat domain 9 (WDR9) domain 1[50]. TAF1 is the first reported protein to have the tandem BD architecture and can interact with multiple acetylated lysine marks on histone H4 (K8ac/K16ac, K5ac/K12ac and K5ac/K8ac/K12acK16ac)[51] cooperatively. TAF1 has also been shown to be involved in p53 dependent transcription activation of p21 by interacting with p53 diacetyl mark (K372ac/K382ac)[52]. Subfamily 8 contains the methyltransferase ash1 like protein (ASH1L)[53], the chromatin remodelling factor SWI/SWF-related proteins SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 2 (SMARCA2)[54], SMARCA4[55] and polybromo1 protein PB1[56].

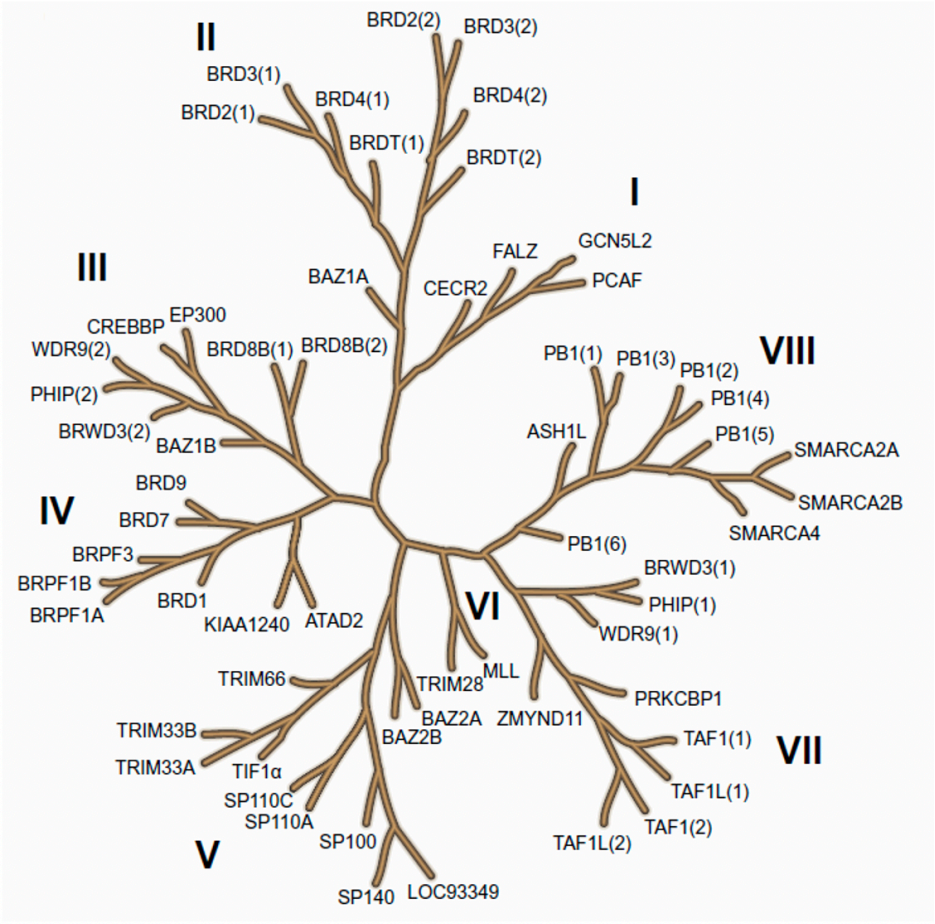


Figure 1.2. Phylogenetic tree of the human bromodomain family constructed on a structure-based alignment. The different families are named by Roman numbers[57]. Where more than one domain exists in a single protein, the bromodomain are numbered according to their position from N-terminus of the protein. (Reprinted with permission form John Wiley and Sons. This figure was published as figure 1c from FEBS Letters 586 (2012) pp: 2692-2704).

In the study of substrate specificity and binding mechanics of BDs, methods such as SPOT arrays, pull-downs, isothermal titration calorimetry (ITC), surface plasmon resonance (SPR), nuclear magnetic resonance (NMR) and co-crystallization have been used[57]. Studies have suggested that several single BDs not only recognise one acetylated mark, but also bind to closely spaced multiple acetylated marks with increased affinity, revealing a new cooperative substrate recognition mechanism, first shown with BRDT binding to histone H4 K5ac/K8ac crystal structure[58]. This cooperative binding mechanics differs from the tandem BD recognition on two separated acetyl marks reported in the TAF1 study discussed earlier. Both ways of binding mechanics add diversity to the overall binding mode of BD family proteins on multiple

acetyl marks, considering that these proteins contain multiple domains and usually contain more than one BD. Simultaneous binding of tandem domains to a long peptide with multiple acetyl marks could increase the affinity for this specific substrate and present a favourable avidity effect[57]. In this case, the linker between the two domains and the two (or multiple) acetyl marks plays an essential role in substrate specificity for effectively exploiting the cooperative binding mechanism. This observation could be further studied for the development of specific multivalent inhibitors for the multidomain reader proteins.

1.2. Bromo and extra terminal (BET) family

The BET family proteins belong to the BD subfamily 2 and consist of four proteins, including BRD2, BRD3, BRD4 and BRDT. BET proteins are multidomain proteins comprised of two conserved N-terminal BDs, an extra terminal protein recruiting domain (ET), and in the case of BRD4 and BRDT, an additional C-terminal motif (CTM), a docking site for other interactions (**Figure 1.3**). The modular architecture of BET creates a platform for assembling transcriptional machinery and recruiting these to the acetylated histone marks. Since molecular architecture and modular arrangement of a multidomain protein are critical factors for determining its function, we will be first looking at each of the BET domains and the interactions they have initiated.

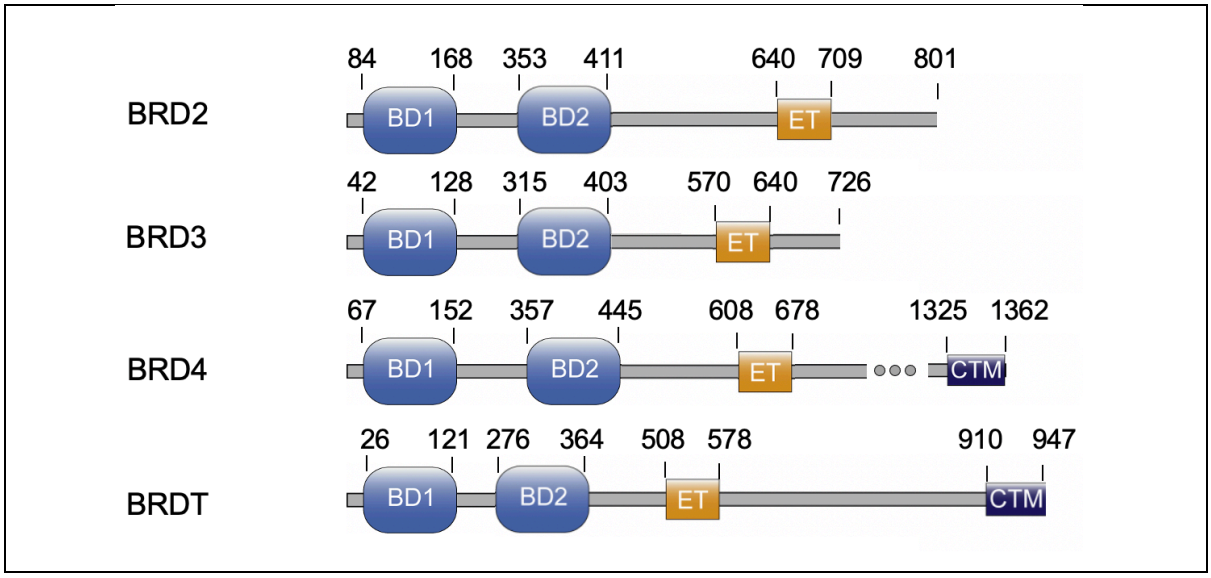


Figure 1.3. Domain organisation of bromo and extra-terminal (BET) proteins. BET proteins contain two N-terminal tandem BDs (shown in blue) responsible for acetyl-lysine recognition and an extra-terminal protein-protein interaction domain ET domain (shown in yellow). BRDT and BRD4 contain a C-terminal motif CTM (shown in dark blue) responsible for other interactions. Domain boundaries are annotated in amino acids.

1.2.1. BET bromodomains

All BDs share a conserved left-handed bundle of four alpha helix motifs (αZ, αA, αB and αC), linked by diverse loop regions (ZA and BC loops) (**Figure 1.4a**) composed of a variable sequence of 120 residues[24]. The Kac binds to a hydrophobic pocket at one end of the helical bundle. Although there is a considerable sequence variation between BD modules, the residues involved in Kac recognition are the most highly conserved residues in all BDs. The BD interacts with the Kac residue via direct hydrogen bonding between the acetyl carbonyl oxygen atom and the NH_2 group of the asparagine amide found at the beginning of the BC loop. The high-resolution crystal structure of the first BD of BRD2 bound to histone H4 peptide at the K12ac mark showed that the acetylated lysine inserted into the hydrophobic pocket and interacted with the N156 conserved asparagine (**Figure 1.4a**). The surface electrostatic property of the hydrophobic pocket surroundings is the main driver for the protein-peptide interaction and thus has a role in binding specificity (**Figure 1.4b**)[59]. Unlike the sequence variation observed in the other BDs, those in the BET family share around 80% sequence

conservation between the first and second BDs in these four proteins. Sequence differences within the loop (ZA and BC loops) region between the BET BDs and other BDs have become a vital factor for developing a BET BD specific inhibitor. In addition to binding to a single acetylated lysine mark, BET BDs have been reported to bind to closely spaced double Kac marks and were found to be conserved between the eight BET BDs. In this binding mode, the N-terminal Kac mark is responsible for the interaction with the conserved asparagine, and the second Kac is involved in a water-mediated interaction that further stabilises the binding[58].

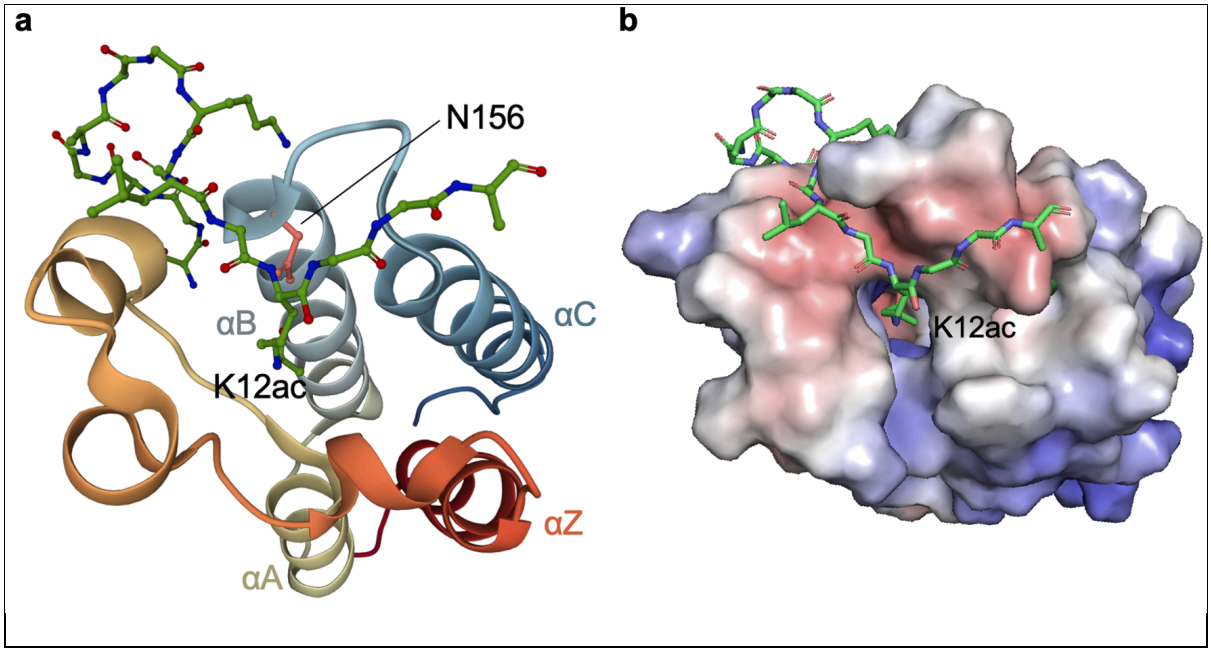


Figure 1.4. Bromodomain overall fold and substrate binding mode. (**a**) Structure of the BRD2(1) (PDB ID: 2DVQ[59]), the interaction with histone H4 peptide with K12ac mark is also shown. (**b**) Surface of the BRD2(1) illustrating the electrostatic potentials around the Kac binding cavity.

An increased binding sensitivity of BET BDs towards the peptides bearing multiple Kac sites compared with that of the peptides with a single Kac was first fully described in a binding pattern study of BET BDs on histone H3 and H4 N-terminal tails using SPOT array[27]. The BRD4 second BD did not interact with an H3 peptide with a single Kac mark at K4, but it did interact strongly with an H3 peptide with diacetyl marks at K4 and K9. An ITC experiment also confirmed the 10-fold increased affinity of the first BD of BRD4 to the H4 peptide with

K5ac and K8ac marks (K_D of 6.8 μM) compared with that of an H4 peptide with a single K8ac mark (K_D of 84.7 μM). Further study suggested that a two-residue spacer residue (Kac-XX-Kac) was required to obtain the highest increased binding affinity of BET for diacetylated H4 peptide. Moreover, the first BD of BRD2 and BRD4 interacts with H4 peptide that has four Kac marks at the K5, K8, K12 and K16 sites with single-digit micromolar K_D values (3.7 and 2.8 μM for BRD2 and BRD4, respectively), which is at least a 20-fold increase compared with that of a single Kac mark interaction. When comparing binding of the first and second BET BDs to an H4 peptide with the K5ac, K8ac, K12ac and K16ac marks, the first BD showed a 10-fold higher affinity, whereas the second BET BD showed higher affinity towards H3 peptide with acetyl marks, suggesting the first BET BD is the main initiator of the interaction with the polyacetylated H4 tail, and the second BD may bind to the H3 in the same nucleosome or to adjacent nucleosomes.

Apart from histones, other proteins with Kac marks have also been reported as important BET BD substrates. These proteins include RelA subunit of nuclear factor-*κ*-light-chain-enhancer of activated B cells (NF-*κ*B), transcription factor GATA1, epithelial–mesenchymal transition (EMT) transcription factor TWIST and cyclin-T1. Acetylation of the RelA subunit of NF-*κ*B, at lysine K310 by histone acetyl transferase CBP/p300 can be recognised by the BDs of BRD4. This allows BRD4 to recruit the NF-*κ*B complex to the promoter region of NF-*κ*B target genes such as proinflammatory cytokine tumour necrosis factor alpha (TNF-α), stimulating their transcription[60]. This study demonstrated that deletion of BD1, BD2 or both of these significantly compromised the BRD4 interaction with RelA, indicating that both BDs are required for this, as evidenced by pull-down assays and in vivo immunoprecipitation assays. Structural studies also confirmed the interaction via BD1 binding to K310ac peptide through the N140 conserved asparagine and BD2 binding to K310ac peptide through the N433 conserved asparagine[61]. The BD1 of BRD3 has been reported to interact with diacetylated

transcription factor GATA1 that is acetylated at the C-terminal tail lysine K312 and K315 sites. This interaction results in the recruitment of the BRD3/GATA1 complex to GATA1 responsive genes, thereby regulating their expression[62]. The structural analysis of this interaction revealed that BD1 of BRD3 bound to the diacetylated GATA1 in a similar way as shown for the interactions between BET BDs with diacetylated histone H4 peptides[63]. However, BD2 of BRD3 is not involved in the interaction, despite the existence of four acetylated lysine sites in GATA1. In a similar fashion, BD2 of BRD4 has been shown to interact with diacetylated transcription factor TWIST that is acetylated at lysine K73 and K76 by Tat-interactive protein 60 (TIP60). This interaction allows recruiting of the TWIST/BRD4/positive transcription elongation factor (P-TEFb) complex to the chromatin, stimulating the expression of TWIST target genes[64]. For example, *WNT5A* is one of the target genes regulated by BRD4/TWIST and the expressed protein is a critical ligand involved in the Wnt pathway, which is related to tumorigenicity of basal-like breast cancer cells, cancer stem cell-like properties and invasion by cancer cells. This study also demonstrated that the two BDs of BRD4 have distinct functions where in this case, BD1 of BRD4 was responsible for binding to histone H4 while BD2 of BRD4 was responsible for binding to TWIST, suggesting the role of tandem BDs as cooperative partners. BD2 of BRD4 has also been shown to interact with acetylated cyclin-T1 that is a component of P-TEFb, resulting in the transcription of P-TEFb-targeted genes[65].

1.2.2. BET extra terminal (ET) domain

While the BET tandem BDs have been shown to tether BETs and other acetylated proteins to chromatin, the ET domain is a vital recruitment point for a variety of cellular proteins such as the transcriptional regulator histone lysine N-methyltransferase NSD3 (NSD3), ATPase family AAA domain-containing protein 5 (ATAD5), glioma tumour suppressor candidate region gene 1 protein (GLTSCR1), JMJD6, chromodomain helicase DNA-binding protein 4 (CHD4)[66,67]

and cohesion-loading factor nipped-B-like protein (NIPBL)[68], as well as viral proteins such as Kaposi's sarcoma-associated herpesvirus (KSHV) latent nuclear antigen (LANA) [69] and murine leukaemia virus (MLV) integrase[70,71]. Moreover, recent studies using AP-MS with a recombinant BRD4-ET domain identified 151 high-confidence interactors[72]. The interaction between BRD4 and NSD3 has been linked to an aggressive midline carcinoma caused by BRD4 fusion with nuclear protein in testes[73]. BRD4 has been shown to regulate transcriptional activation in a p-TEFb-independent manner through interacting with the above transcription regulators[66].

The ET domain consists of a conserved three-helix bundle fold. Binding occurs via an amphipathic motif through a two-strand antiparallel ß sheet that forms in the loop region between the second helix and third helix (**Figure 1.5a**). An NMR study of the BRD4 ET domain with LANA of KAHV highlighted the importance of the electrostatic nature of the ET domain binding to its effector protein[74] (**Figure 1.5b**). The ET domain-binding motif was further profiled using a peptide SPOT array, and a putative motif of Φ[+]Φ[+] (where Φ is one of M, L, V, I, F and [+] is K or R) for ET domain recognition was established. For example, the ET domain binding site in LANA is QSSIVKFKKPLP, and in the MLV integrase is TWRVQRSQNPLKIRLTR. The two [+] residues in the binding motif can form an electrostatic zip with the negatively charged E651, E653 and D655 of the ET domain (**Figure 1.5b**). After the successful mapping of the ET domain binding motif, the idea of a small peptide inhibitor was developed[75]. A small peptide containing 17 residues surrounding the ET domain binding motif from MLV integrase bound to the ET domain with an affinity of 160 nM (established via ITC), and was shown to compete with NSD3 for the BRD4 ET domain[75]. These studies highlighted a new way of disrupting the interactions between BET proteins and their effector partners in transcriptional activation, cancer and viral integration.

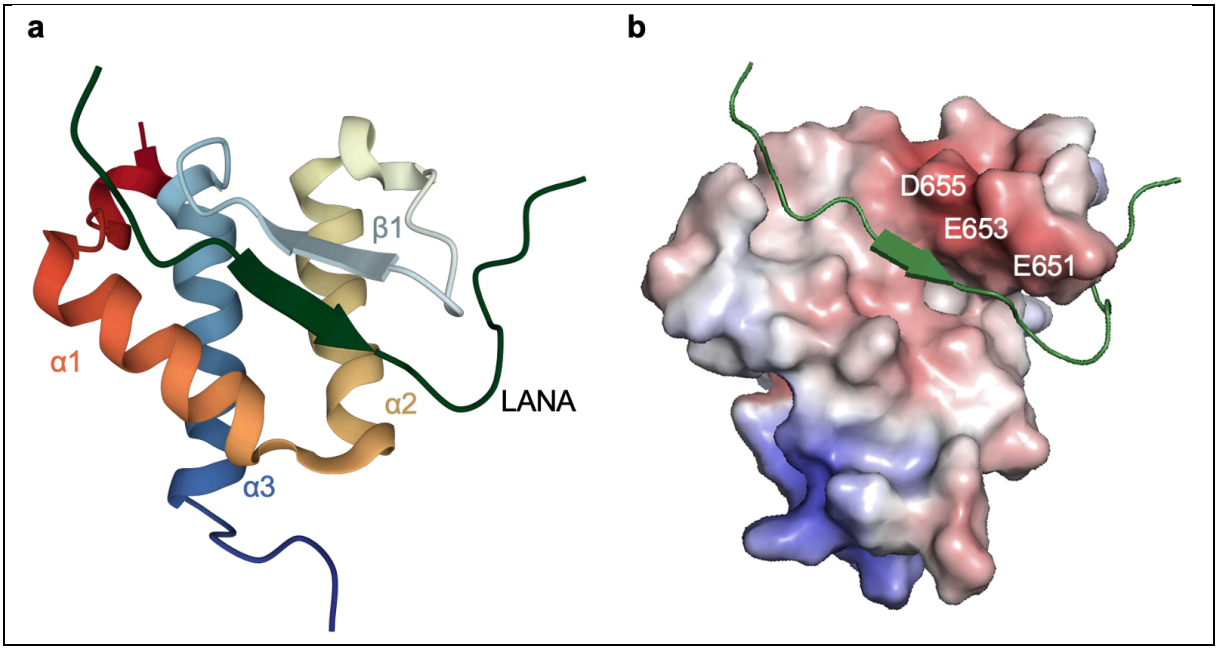


Figure 1.5. BRD4 ET domain recognition of LANA. (**a**) Ribbon representation of average minimised NMR structure of the BRD4 ET domain and LANA peptide complex (PDB ID: 2ND0[74]). (**b**) Electrostatic potential surface representation of the BRD4 ET bound to LANA peptide in cartoon depiction.

1.2.3. BET C-terminal motif (CTM)

Only BRDT and the long isoform of BRD4 contain the C-terminal motif (CTM). The CTM of BRD4 interacts with positive transcription elongation factor b (P-TEFb), which is a heterodimer composed of cyclin T1, T2 or K, and cyclin-dependent kinase 9 (CDK9). The S175 serine in CDK9 was reported to be vital for binding of BRD4 to P-TEFb. This interaction leads to the release of the inhibitory double-stranded RNA binding protein HEXIM, enabling the active form of P-TEFb and transcriptional activity[65]. The CTM of BRD4 also interacts with the human papillomavirus (HPV) E2 protein, leading to the recruitment of viral genomes to host chromosomes supporting replication of the viral genome[76,77] and transcriptional repression of the viral E6 and E7 oncoproteins, which then suppress p53 target gene transcription and related tumour-suppressor activity[78].

1.2.4. BET inter-domain linker region

The domains of the BET proteins are connected by long linker regions. These linker regions are most likely disordered regions that lack a stable structure, thus making BET proteins intrinsically disordered proteins. Intrinsic disorder is reported to be more abundant in eukaryotic proteins than in prokaryotic proteins, indicating the possible involvement in advanced biological functions. The average eukaryotic protein is 450 residues long and contains 145 disordered residues compared with an average prokaryotic protein that is 300 residues long and contains 32–37 disordered residues[79]. The disordered regions are generally composed of a small number of hydrophobic residues and a large number of charged residues[80]. The intrinsically disordered regions of proteins are reported to be involved in protein-protein interactions, including one-to-many protein domain binding, many-to-one protein domain binding and mutual folding where two intrinsically disordered regions interact with each other to yield an intertwined structured domain[81,82]. Therefore, the linker region of BET proteins needs to be carefully investigated. There are several techniques that can be used to characterise intrinsically disordered proteins, for example, computational methods for sequence-based prediction of disorder. Numerous computational algorithms have been developed to predict if a protein is likely to be ordered or disordered under physiological conditions based on the protein sequence. GlobPlot is one of these. GlobPlot defines a propensity for each of the 20 common amino acid residues to be in an ordered or disordered sequence region[83]. A more advanced class of algorithms uses machine learning to distinguish ordered from disordered sequence[84,85]. Moving beyond the general prediction of order or disorder, there are several experimental methods for the structural characterisation of intrinsically disordered regions. The combination of small angle X-ray scattering (SAXS) and NMR is most commonly used. The size distribution of proteins in solution can be obtained from SAXS, providing a measure of the compactness of the solution structure. However, in the study of BET proteins, there have

not been many investigations performed on the linker region. Therefore, the possible roles of the BET linker region need to be carefully evaluated.

As evidenced from all of the above sections, the modular architecture of BET has created a platform for assembling different binding partners, tethered by the ET domain and CTM, as well as binding of BD to chromatin, that tightly participate in controlling transcriptional programmes. It is not hard to imagine that any deficiency in the BET proteins would lead to deregulation of entire related transcription programmes and, therefore, a variety of diseases[86]. Next, we will talk about the roles of BET proteins in different types of diseases.

1.3. Biological relevance of BET proteins

1.3.1. BETs in transcription

BET proteins play an important role in transcription by tethering various proteins to acetylated histones. The roles that BETs have participated in include spermatogenesis, neuronal differentiation, alternative splicing, cell cycle progression, and erythropoiesis.

BRD2 plays a role in neuronal development through an E2F1 transcription factor, which directly binds BRD2 and controls gene expression for cell cycle progression and exit during mouse CNS development[87]. BRD2 has also been reported to affect alternative splicing of multiple genes, including some involved in the response to growth factor stimuli in HeLa cells[88].

BRD3 has been shown to play a role in erythropoiesis by interaction with diacetylated transcription factor GATA1, which has been acetylated at C-terminal tail lysine K312 and K315 sites[63]. This interaction results in the recruitment of the BRD3/GATA1 complex to erythroid-specific genes regulating their expression[62].

BRD4 plays a critical role in transcription through recruiting the P-TEFb complex to chromatin. The CTM of BRD4 interacts with P-TEFb, which is a heterodimer composed of cyclin T1, T2

or K, and cyclin-dependent kinase 9 (CDK9)[89,90]. The S175 serine in CDK9 was reported to be vital for binding of BRD4 to P-TEFb[91]. This interaction between BRD4 and P-TEFb leads to the release of the inhibitory double-stranded RNA binding protein HEXIM1, enabling the active form of P-TEFb, which in turn phosphorylates and activates RNA Pol II[65] (**Figure 1.6**). BRD4 has also been reported to participate in the re-expression of genes during MEF reprogramming to induced pluripotent stem cells[92]. This process needs the binding of BRD4 to the super enhancers of the pluripotential gene that are likely to mediate chromatin remodelling[93]. The shorter variant of BRD4 lacking the CTM has been reported to play a role in DNA damage-response as an endogenous inhibitor and depletion of this variant showed rapid cell-cycle checkpoint recovery[94].

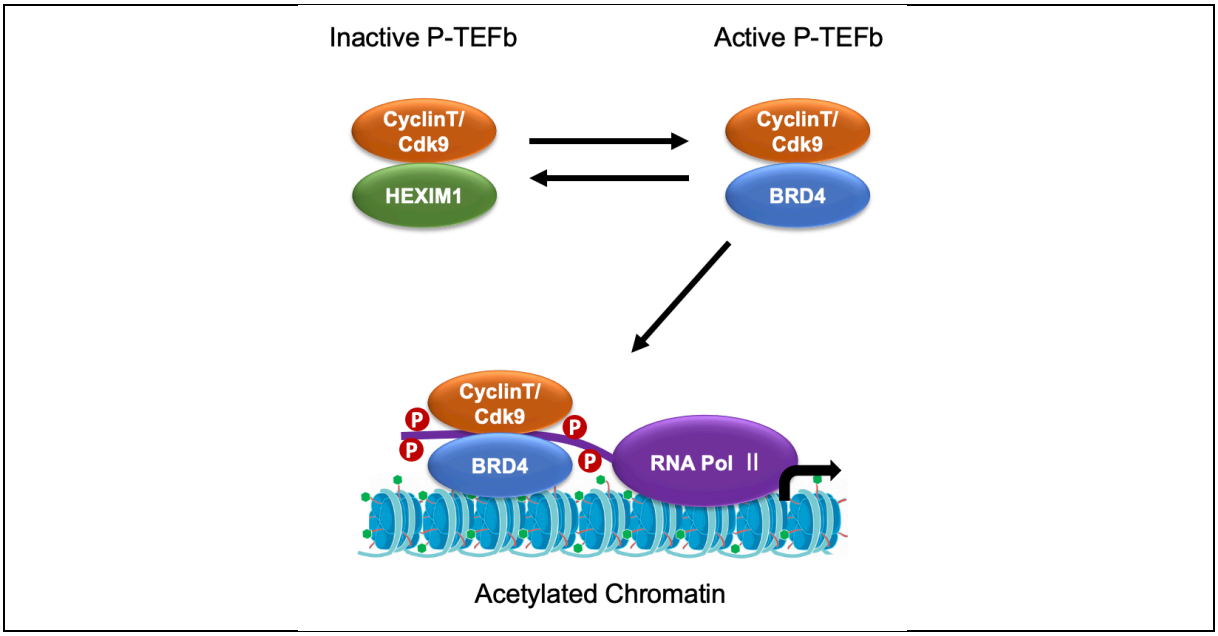


Figure 1.6. A model for BRD4 and P-TEFb interaction. By binding to acetylated chromatin, BRD4 facilitates the recruitment of the active form of P-TEFb to promoters and enhances their transcription by activating RNA Pol II.

BRDT plays a role in chromatin remodelling during spermatogenesis and has been suggested as a target of male contraception[95]. BRDT has been reported to be a central regulator of both meiotic divisions and post-meiotic genome packaging through driving and determining the developmental timing of a testis-specific gene expression program[96]. The shorter variant of

BRDT lacking the CTM has been shown to stimulate chromatin compaction, suggesting a possible role of the longer BRDT variant in chromatin decompaction[97].

1.3.2. BETs in cancer

BET proteins play a critical role in transcription by recruiting transcription machinery to chromatin. However, the deregulation of this process could lead to various diseases. BET proteins have been reported to be involved in both solid and liquid tumours. A key example of this is their expression as oncogenic fusions together with nuclear protein in testis (NUT), resulting in very aggressive carcinomas that originate mainly from the NUT midline carcinoma (NMC)[98]. This disease is extremely aggressive, with a large sample study[99] reporting a median overall survival of 6.7 months. In NMCs, the N-terminal region of BRD4 (including tandem BDs and ET but not CTM) fuses to the complete NUT protein. The fusion protein blocks differentiation and retains tumour growth through a mechanism that is still being investigated[98]. These studies suggest that the BET-NUT fusion protein causes hyperacetylation of transcriptionally inactive sites by recruitment of the acetyltransferase p300 protein. p300 has an essential role in regulating p53-dependent cellular response by acetylating p53. Therefore, the existence of BET-NUT fusion protein leads to inadequate acetylation of p53, resulting in the inactivation of p53-dependent antiproliferation cellular response. Other studies have suggested that BET-NUT fusion protein promotes the expression of oncogene-like MYC, which has been reported as a downstream target of BET-NUT[100]. Chromatin immunoprecipitation with massively parallel DNA sequencing (ChIP-seq) data showed that BRD4 is sitting on large numbers of enhancer regions that are responsible for the promoted expression of oncogenes such as MYC[101]. Although there are different opinions on the exact mechanism of how fusion proteins block differentiation and retain tumour growth, inhibition of the BET BD interaction in NMC cell lines was reported to restore the cellular function in all

cases[102]. Significantly, in a report studying the BET inhibition in four NMC patients, tumour regression and symptomatic relief were observed in two patients. The overall survival was also extended to 19 months in these two patients[103].

Promotion of expression of oncogene MYC by BET proteins has been reported in other cancers as well, such as acute myeloid leukaemia (AML)[104-107], mixed lineage leukaemia (MLL)-fusion leukaemia[18], multiple myeloma, Burkitt's lymphoma[108], B-cell acute lymphoblastic (BLL) leukaemia[109], medulloblastoma[110-112], glioblastoma[113], prostate cancer[114-119], lung cancer[120,121] and non-small cell lung cancer[122], ovarian cancer[123,124], breast cancer[125], and skin cancers[126,127]. In glioblastoma, for example, aberrant epidermal growth factor receptor (EGFR) signalling induces expression of transcription factors SRY-Box transcription factor 9 (SOX9) and forkhead box G1 (FOXG1), which activate expression of BRD4, subsequently promoting expression of MYC expression to promote tumour growth. Depletion of BRD4 expression causes reduced glioblastoma cell cycle progression[128]. Similarly, inhibition of BET proteins led to cell arrest in neuroblastoma[129].

Apart from increased oncogene expression caused by BET proteins in cancers, other irregulated pathways have been reported relating to BET proteins as well. For example, in lung adenocarcinoma, a complex formed with BRD2 via the first BD and acetylated runt-related transcription factor 3 (RUNX3) can induce the expression and function of the cell cycle inhibitor p21 in K-Ras-dependent tumours. Abrogation of the interaction between BRD2 and RUNX3 abolishes the p21 expression, resulting in deregulation of the cell cycle[130]. Similarly, there are other examples of BET protein roles in cancers that affect cell cycle regulation. In Hela cells, BRD2 was reported to affect alternative splicing of several genes, including genes that are involved in MAPK signaling[88]. BRD2 also can bind to cyclin A promoter in B cells and regulate the expression of cyclin A, controlling the S-phase progression[131]. BRD4 has been reported to bind to M/G1 genes that are expressed at the end of mitosis[132]. BRD4 also has a

role in cell cycle continuation in P19 embryonic carcinoma cells. Nevertheless, BET proteins have been reported to be involved in other diseases as well.

1.3.3. BETs in inflammation

BET proteins have been reported to function in inflammatory diseases that include a vast array of disorders that feature inflammation response, such as allergy, asthma, autoimmune diseases, inflammatory bowel disease and transplant rejection. BET proteins were shown to interact with NF-κB, a vital regulator of the cellular immune response, facilitating the transcription of NF-κB target genes[133], including those encoding cytokines and chemokines. The NF-κB family is composed of five structurally related proteins, p50 (also named NF-κB1), p52 (also named NF-κB2), RelA, RelB and c-Rel[134]. Members of this family of proteins dimerize to form, in principle, 15 unique homo-and heterodimers, with differing affinities for dimerization. Among NF-κB homo-and heterodimers, RelA and p50 are the most commonly found heterodimer complex and are the functional components participating in the activation of NF-κB. BET proteins have been shown to bind with RelA acetylated at lysine K310 via its BDs. Blocking the interaction using BET inhibitors resulted in protection from the bacteria-induced sepsis and LPS-induced endotoxic shock in mice[20]. Furthermore, BRD4 inhibition using BET inhibitors causes ubiquitylation and degradation of the active form of RelA[61]. ChIP-seq data analysis revealed that NF-κB formed super enhancers to promote proinflammatory gene expression in a BRD4-dependent manner, and that modulating BRD4 levels at enhancers may influence the pathogenesis of inflammatory diseases[135]. BRD4 was also shown to involve in super-enhancer RNA (seRNA) transcription in the macrophage model system, and seRNAs were induced at most of the inflammation driver genes[136]. Apart from the role of BRD4 in inflammation, BRD2 was also reported to be involved in inflammation response. In a metabolism study, BRD2-mutated mice showed severe obesity and reduced inflammation response[137], suggesting a

positive correlation between BRD2 and cytokine production. BRD2 has also been involved in the regulation of inflammatory gene expression in the development of rheumatoid arthritis[138].

1.3.4. BETs in viral infection

BET proteins have also been reported to function in viral infection. Earlier studies have established the role of BETs in viral gene integration to chromosome. These viruses are murine leukaemia virus (MLV), human papillomavirus (HPV), Epstein-Barr virus (EBV), and Kaposi's sarcoma-associated herpesvirus (KSHV)[69,139-142]. The Chip-Seq study regarding the impact of the interaction between BET proteins with MLV revealed that MLV integration site selection was mostly at enriched acetylated histone regions, suggesting BET protein guidance through their BDs, and the C terminus of MLV integrase was responsible for specific binding to the ET domain[71,141]. Furthermore, BRD4 has been shown to interact with HPV viral E1 and E2 proteins, facilitating the tethering of viral DNA to the host genome[140] and has been shown to be required for HPV transcription and replication to maintain latent virus. In maintaining virus protein stability in host cells, BET interaction with HPV viral proteins has been shown to be critical for preventing viral protein ubiquitylation and subsequent degradation[143]. ChIP-seq data analysis showed the interaction and co-occupancy of EBV nuclear antigens and BRD4 on a super-enhancer site with a H3K27ac mark. These sites are critical regulators of B-cell growth and survival genes such as *BCL2* and *MYC*[144]. In HIV, BRD2 and BRD4 have been shown to enforce virus latency by suppressing HIV transcription via different mechanisms. BRD2 recruits the chromatin-remodelling factors, HDACs to the HIV promoter to repress transcription[145], whereas BRD4 competes with P-TEFb and with HIV trans-activator of transcription (Tat) protein at the HIV promoter to repress transcription[146]. In this case, it has been suggested that using BET inhibitors to release BET proteins from the HIV promoter may

help reactivation of virus latency. The combination of antiretroviral treatments with latency reactivation leads to depletion of latent reservoirs and suppression of the diseases.

Therefore, there are pressing needs for therapy that may involve BET proteins involved in cancer, inflammation and viral infection. BET proteins have become a compelling and ideal target for the development of qualified probes.

1.4. Discovery of BET bromodomain inhibitors

1.4.1. BET bromodomain monovalent inhibitors

The first potent BET BD inhibitors identified were triazolobenzodiazepines (based on patent literature from 2009) and were followed by the discovery in 2010 of the broadly known and fully characterised compound (+)-JQ1 that contains a triazolothienodiazepine core scaffold[17]. Currently, a large number of BET BD inhibitors have now been published. All these compounds fall into two main categories: the acetylated lysine mimetic class of small molecules that form the hydrogen bonds with the conserved asparagine residue similar to the binding mode of acetylated lysine, and the non-acetylated lysine mimetic class of small molecules that bind to the BD acetylated lysine binding pocket without forming the hydrogen bond with the conserved asparagine. The non-acetylated lysine mimetic molecules are usually weak inhibitors. One of the examples in this class is BIDI that has a K_D value of 28 μM as measured by SPR. The crystal structures of BIDI binding to the BRD2 first BD demonstrated that this engagement is established mainly through hydrophobic interaction instead of forming a hydrogen bond with the conserved asparagine[147].

(+)-JQ1 belongs to the first class that mimics the binding of acetyl-lysine to BET BDs by forming hydrogen bonds with the conserved asparagine residue. (+)-JQ1 is highly selective towards BET BDs with an affinity for each BET BD that is below 190 nM. The crystal structure of (+)-JQ1 binding to BET BDs showed that the triazolo ring of the inhibitor inserts deep into

the hydrophobic pocket and initiates a hydrogen bond with the conserved asparagine[17] (**Figure 1.7a**). Similar acetyl-lysine mimic compounds include the thienodiazepine core compounds MS417[148] and JQ1, benzodiazepine core compounds I-BET762 and GSK525768A[149], benzotriazepine core compounds BzT-7[150], isoxazole core compound I-BET151[151], dihydroquinazoline-one PFI-1[152] and quinazolone core compound RVX-208[153] (**Figure 1.7**).

All BET inhibitors have been reported to be able to dissociate BETs from chromatin and thereby affect the respective transcription programs involved. The discovery of these compounds vastly contributed to the exploration and validation of BET function in various diseases. (+)-JQ1 inhibited BRD4-NUT fusion protein and showed a strong efficiency in reducing the tumour size in murine models of NMC[17]. In leukaemia, (+)-JQ1 also exerted significant antiproliferative effects. (+)-JQ1 treatment repressed the expression of oncogene MYC. This was first described in multiple myeloma[154], followed by demonstration of a tumour growth suppressive effect in several mouse models such as acute myeloid leukaemia (AML) and MLL[18,104] and acute lymphoblastic leukaemia[109]. MS417 has been reported to suppress TNF-induced transcriptional activation of NF-κB and demonstrated a potential anti-inflammatory effect[155]. I-BET151 was reported to block the clonogenic growth of MLL-fusion-driven leukaemia cells by regulating the cell-cycle progression transcriptional programme[18]. PFI-1 also showed antiproliferative effects in clonogenic growth assays on several leukaemia cell lines. PFI-1 sensitive cell lines showed a G1 cell cycle arrest and induction of apoptosis[156]. Similarly, other BET BD inhibitors developed after (+)-JQ1 also exhibited significant antitumour activity in murine models of NMC, multiple myeloma[157], MLL and ALL[158,159], acute lymphoblastic leukaemia[160,161], glioblastoma[162], melanoma[126] and prostate cancer[163]. In viral infection, these inhibitors also showed promising efficacy in inhibiting viral DNA integration to the host genome[141,142]. In inflammation, these inhibitors reduced the expression of proinflammatory genes and protected mice from bacteria induced sepsis[104].

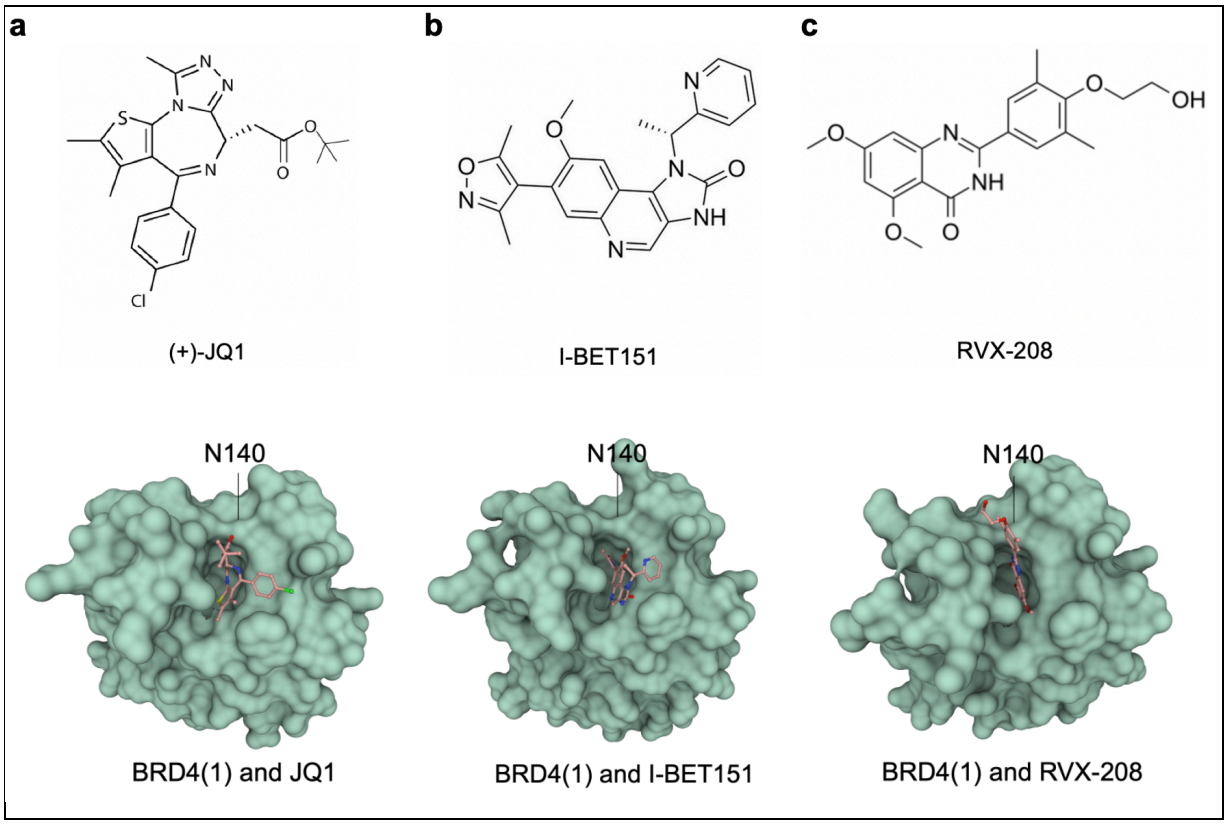


Figure 1.7. Structures of key BET bromodomain inhibitors and binding orientations of inhibitors to bromodomain modules. (**a**) Compound (+)-JQ1 and JQ1 bound to BRD4(1) (PDB ID: 3MXF). (**b**) Compound I-BET151 and I-BET151 bound to BRD4(1) (PDB ID: 3ZYU). (**c**) Compound RVX-208 and RVX-208 bound to BRD4(1) (PDB ID: 4MR4).

RVX-208 was the first BET inhibitor to be reported to preferentially bind to the second BD (BD2) of BETs with up to 20-fold selectivity over the first BD (BD1) [164] and was the first BET BD inhibitor to enter clinical trials. In a recent study, a novel inhibitor ABBV-744 was described that binds 290-fold more potently to the second BD than to the first BD[119] with K_D values of 3300 and 2.1 nM for the BRD4 BD1 and BD2 domains, respectively. RNA and ChIP-seq data analysis following the second BD inhibition suggested that ABBV-744 has a lower impact on global transcription while showing significant inhibition towards AR-dependent transcription in comparison with that of a pan-BET BD inhibitor in LNCaP cells. In a mouse xenograft model, ABBV-744 caused a delay in tumour growth, and also showed improved outcomes for toxicity compared with those of other BET inhibitors. However, there has been minimal study on the diverse function of two BDs in one BET protein so far. A recent paper

explored the distinct functions of the two individual BDs in the BET family proteins[165], suggesting that selective targeting of one of the BDs may result in a different efficacy and tolerability profile compared with that of existing BET inhibitors. In this study, two inhibitors, iBET-BD1 and iBET-BD2, were used to selectively inhibit BET BD1 and BD2, respectively. In affinity measurements via SPR using BD1 and BD2 of BRD4, iBET-BD1 showed a 130-fold selectivity towards BD1 and iBET-BD2 showed a 300-fold selectivity towards BD2. The transcriptome study using these two inhibitors showed that iBET-BD2 treatment did not induce marked transcriptional changes, whereas iBET-BD1 treatment induced transcriptomic features that resembled those induced by I-BET151 in a cancer cell line. These results corresponded with the finding that iBET-BD1 can displace BET proteins from chromatin as well as I-BET151 can, whereas iBET-BD2 cannot displace BET proteins from chromatin. However, in the inflammatory stimuli response study, iBET-BD2 showed effective inhibition of stimuli induced expression of corresponding genes, suggesting the important role of BD2 in the recruitment of the BET proteins for the induction of gene expression. This effect of iBET-BD2 inhibitor makes it a novel and useful therapeutic probe to overcome various immunoinflammatory diseases. Even though this was considerable progress, BET monovalent inhibitors only block a single BET BD, leaving the ET protein recruiting domain open and with no selectivity towards any BET members. This has urged scientists to develop new probes that can differentiate between BET members and take into account the effect of the ET domain.

1.4.2. BET bromodomain bivalent inhibitors

In the drug discovery history of small-molecule inhibition, one small molecule targeting per module binding site is preferred. This type of small molecule is relatively easy to understand, design, and handle. However, if we look at the natural way of developing binders, we find that multivalence is frequently used by exploiting the sum of many weak interactions leading to a

strong overall binding. This is the 'avidity effect'. Giving the existence of two BDs in one BET protein, a bivalent approach to an inhibitor of enhanced avidity has become a favourable idea. The first bivalent inhibitors, which can simultaneously tether the first and second BDs of the BETs, were developed in 2016[166] (**Figure 1.8**).

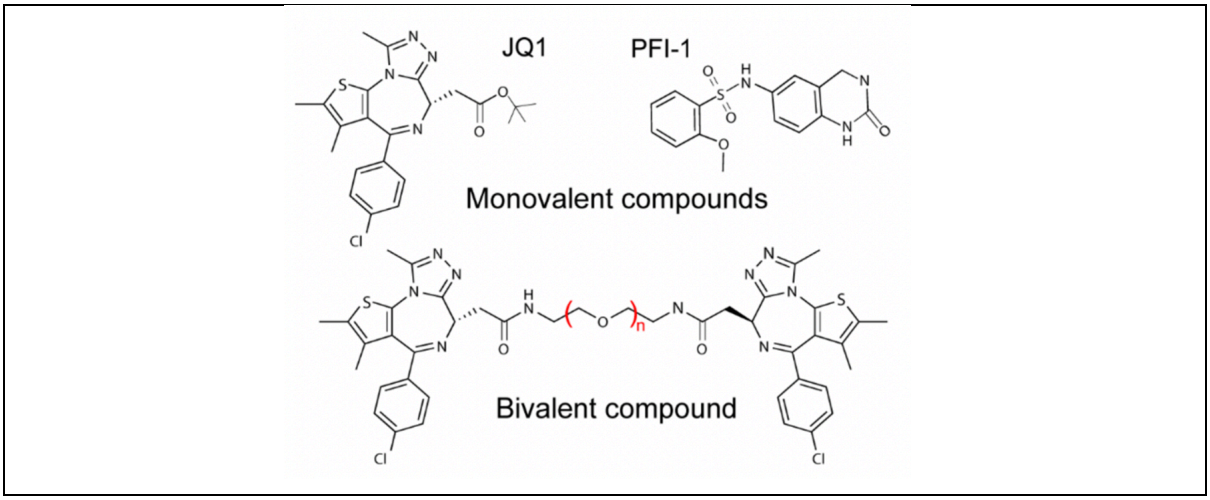


Figure 1.8. Representative monovalent and bivalent BET inhibitors.

In these studies, the bivalent compound was rationally designed, where the monovalent BET inhibitors (+)-JQ1 and I-BET151 were conjugated in all possible combinations and with different lengths of flexible polyethylene glycol (PEG) linker regions. In the cellular compound activity (pIC_{50}) assay, all these compounds showed greater potency in comparison with that of JQ1. Compound MT1 was selected from these compounds for further experiments and had two JQ1 warheads conjugated with a PEG7 linker length. MT1 was proven to bind to BET in a bivalent mode, simultaneously engaging two BDs, with a picomolar affinity towards BRD4, which is 100-fold more potent compared with that of the monovalent inhibitors. In a leukaemia xenograft model, the bivalent new probe significantly delayed leukaemia progression in comparison with the vehicle and JQ1. At the same time, this new bivalent inhibitor also only showed increased potencies over the on-target proteins of the monovalent scaffold; in contrast, the inhibition for off-targets did not increase. Another similar article was also reported in the same journal regarding the discovery of BET bivalent inhibitor compound 6 while developing

AR down regulators[167]. Compound 6 also demonstrated binding to BRD4 tandem BDs in bivalent mode as measured by analytical ultracentrifugation (AUC), SAXS, and bioluminescence resonance energy transfer (BRET) assays. Moreover, treatment with compound 6 in cells showed an enhanced potency in cancer-cell-killing activity. Although this bivalent engagement approach did not answer the question of finding selectivity between BRD2, BRD3, BRD4 and BRDT, this did enhance our understanding of multivalent engagement and provided new ideas for developing a series of probes with distinct advantages over previously described monovalent compounds.

1.4.3. BET proteins and PROTACs

PROteolysis Targeting Chimaeras (PROTACs) are bifunctional molecules that bring together a target protein and an E3 ubiquitin ligase to trigger degradation of the target protein[168]. A typical PROTAC consists of a linker region and two warheads, one for binding of the target protein and the other for recruiting an E3 ligase. An E3 ligase is a protein that tethers an E2 ubiquitin conjugating enzyme and assists the transferring of ubiquitin from E2 to the substrate. Using PROTACs brings target proteins to the E3 ligase complex for ubiquitination and subsequent degradation by the proteasome. There have been several different E3 ligases employed for the designing of PROTACs. CRL2 and CRL4 from cullin-ring E3 ubiquitin ligase (CRL) multi-protein complexes have frequently been employed in the design of PROTACs. In addition to the critical warheads, the linker length can control the proximity to the target protein and influence the ubiquitination efficiency. The development of PROTACs was considered necessary, especially in multidomain proteins, as they can not only bind to the target proteins but also degrade them, leaving no open domains for further interactions.

The first BET PROTACs, dBET1, was reported in 2015, which conjugates (+)-JQ1 to CRL4[169], and has been shown to effectively degrade BRD4 in the human AML cell line. Similarly, other

newly developed BET degraders including dBET6, BETd-246/260[170] and ARV-825/ARV771[171,172] have been reported to successfully degrade BET proteins and suppress the expression of oncogenes regulated by BET proteins much more effectively than the corresponding BET monovalent inhibitors. For example, in glioblastoma cells, dBET6 treatment showed depletion of BET proteins occupancy on chromatin and impaired the E2F1 regulated transcription program[173]. In AML cells, BETd-246 treatment resulted in transcriptional downregulation of MYC[170,174,175]. In osteosarcoma cell lines, BETd-260 treatment potently suppressed the cell viability and showed 1000-fold increased activity compared with that of (+)-JQ1[176]. Moreover, the employment of these PROTACs has led to significant inhibition of cell growth and induction of apoptosis in mice models of solid tumours and hematologic malignancies with no overt toxicity[169,171]. These results were expected since the BET inhibitor cannot influence BD independent functions of BETs by leaving the ET dependent interactions open, whereas PROTAC-based molecules can degrade the whole protein. However, in the evaluation of their effects on global gene expression, the toxicity of BET degraders should be carefully assessed since degraders still cannot distinguish between BET members. Therefore, it would be ideal if a BET member selective inhibitor was designed and used as one of the warheads of PROTAC.

1.5. Clinical efficacy of BET bromodomain inhibitors

Various BET inhibitors have been studied with *in vitro* and *in vivo* testing as well as mouse models of different diseases. In tamoxifen-resistant breast cancer, treatment with BET inhibitor in combination with fulvestrant, showed potent inhibition of tumour growth in tamoxifen-resistant xenograft mouse[177]. BET proteins have been demonstrated to contribute to *ERα* gene inappropriate expression by recruiting Wolf–Hirschhorn syndrome candidate 1 (WHSC1), a methyltransferase, to the promoter region of the *ERα* gene, and BET inhibition suppresses the

aberrant expression of ERα in tamoxifen-resistant cells[177]. In the lapatinib-resistant breast cancer, resistance was shown to result from overexpression of human epidermal growth factor receptor 3 (HER3), and treatment with BET inhibitors suppressed the genes that were overexpressed by lapatinib resistance, thereby leading to growth inhibition of the lapatinib-resistant cells[178]. In small cell lung cancer, treatment with BET inhibitors blocked BETs binding to the enhancer region of Achaete-Scute family BHLH transcription factor 1 (*ASCL1*), a gene whose expression is upregulated in small cell lung cancer; BET inhibitors also decreased the ASCL1 expression level in lung tumour cells[179]. In colon cancer, treatment with BET inhibitors repressed the expression of colon cancer-associated transcript-1 (CCAT1), a long non-coding RNA that promotes tumour formation, with inhibition of tumour growth in xenograft mouse[180]. In glioblastoma, treatment with BET inhibitors decreased the proliferation of glioblastoma cancer cells, promoted apoptosis and blocked glioblastoma progression in an orthotopic tumour model[181]. In SMO-inhibitor resistant medulloblastoma, treatment with BET inhibitors blocked cell viability in medulloblastoma cell lines and tumour growth in xenograft mouse models[182,183]. In pancreatic cancer, treatment with BET inhibitors inhibited tumour growth in patient-derived xenograft models[184]. Targeting BRD4 has been demonstrated to decrease interleukin (IL)-6 production and STAT signalling[185], which are pathways found to regulate pancreatic ductal adenocarcinoma progression in mouse models[186]. In castration-resistant prostate cancer, treatment with BET inhibitors induced apoptosis, suppressed expression of antiapoptotic factors in prostate cell lines and reduced the transcription of AR target genes, suggesting the involvement of BET protein in an AR-mediated transcriptional program[114]. In mouse models, BET inhibition decreased the growth of castration-resistant prostate cancer tumours while not affecting normal prostate growth[114].

Overall, this evidence suggests that BET inhibitors may be of distinct therapeutic benefit to different diseases and that treatment using them is a good rationale for translational clinical

research. The great success of BET inhibition in NMC animal models initiated two Phase I trials targeting NMC. One was for exploring the clinical efficacy of I-BET762 and the other that of TEN-010. The Phase I clinical trial of I-BET762 has recently been completed, and results showed that among the 19 patients with NMC that participated in the study, four were tumour progression-free for more than 6 months while progression was stabilised in a further eight patients[187]. Moreover, in the clinical trial of OTX015 on BET inhibition in four NMC patients, tumour regression and symptomatic relief were observed in two patients. The overall survival was also extended to 19 months for these two patients, which is longer than the median overall survival of 6.7 months reported for advanced NMC patients[103]. Clinical trials of RVX-208 in Fabry disease (FD) were initiated based on its beneficial effects on several pathways downstream of globotriaosylceramide (Gb3) accumulation. FD is a disorder caused by a deficiency in enzyme alpha-galactosidase resulted in progressive accumulation of its substrate Gb3, leading to severe complications in cardiovascular-related issues, stroke and peripheral neuropathy. However, the exact mechanism by which BET proteins are involved in the Gb3 downstream pathways is unclear, and further mechanistic studies are required to clarify the role of BET proteins in FD, although the trials showed promising positive effects on FD patients. The registered active clinical trials that are investigating BET inhibition are listed below (**Table 1.2**; and https://clinicaltrials.gov/ for further details on BET inhibitor trials).

Compound	Company	Phase	Indications in development	ClinicalTrials.gov identifier (status)
RVX-208	Resverlogix	III	Diabetes Mellitus, Type 2; Cardiovascular diseases	NCT02586155 (active)
		II	Atherosclerosis	NCT01058018 (completed)
		II	Fabry disease	NCT03228940 (active)

		II	Kidney failure, Chronic	NCT03160430 (active)
I-BET762	GlaxoSmithKline	I	NUT midline carcinoma (NMC); Haematological malignancies	NCT01587703 (completed)
		II	Neoplasms	NCT02964507 (active)
		I	Solid tumours	NCT03150056 (active)
OTX015	OncoEthix	I	Diffuse large B-cell lymphoma; Acute lymphoblastic leukemia; Multiple myeloma	NCT01713582 (completed)
		I	Acute myeloid leukaemia	NCT02698189 (active)
CPI-0610	Constellation Pharmaceuticals	II	Myelofibrosis; Myelodysplastic syndrome (MDS)	NCT02158858 (active)
		I	Progressive lymphoma	NCT01949883 (completed)
TEN-010	Tensha Therapeutics	I	NUT midline carcinoma (NMC)	NCT01987362 (completed)
ZEN-3694	Zenith Epigenetics	II	Metastatic prostate adenocarcinoma; Metastatic prostate small cell carcinoma	NCT04471974 (active)
BMS-986158	Stand up to cancer	I	Solid tumor; Lymphoma; Brain tumour	NCT03936465 (active)
	Bristol-Myers squibb	II	Advanced tumors	NCT02419417 (active)

Table 1.2. Summary of BET BD inhibitors currently undergoing clinical evaluation.

1.5.1. Toxicity effect of BET bromodomain monovalent inhibitors

Even though pan-BET inhibitors achieved great successes in preliminary studies and initial clinical trials, the fact that different BET proteins are responsible for different functions should not be overlooked. BRD2 is involved in the closure of the embryo neural tubes and promotion of cell cycles[188-190]. BRD3 was shown to play a role in the maturation regulation of erythroid in haematopoietic cells[62,63]. BRD4 has been shown to play a role in transcriptional

activation[91,191], and BRDT was shown to play a role in transcription of gene regulation during spermatogenesis[192].

However, there is no currently available small molecule that is selective for any individual BET protein. The consequences of inhibition of individual BET proteins in differentiated human cells are unknown, and the potential side effect profile of using pan BET inhibitors in a complex system is frightening to imagine. In several BET inhibitor treatment studies, a range of on-target toxicity effects unrelated to cancer has been observed. BRD4 transgenic knockdown mice showed a higher probability of developing radiation damage to their intestines with delay in recovery of normal function[193]. BRD4 knockdown also resulted in the reduction of the number of functional stem cells in the intestine. Similarly, a BET inhibitor suppressed intestinal stem cell differentiation[194]. Apart from the non-specificity of the BET compounds towards individual BET members, recent interest has focused more on the transcription co-activator activity of BET proteins, particularly for cell cycle and proliferation genes in cancer. However, it has long been clear that BET proteins also function as co-repressors of transcription, although this aspect has not been appreciated in drug development. Given the broad involvement of BET proteins in transcriptional regulation, there is an urgent need to develop new compounds with improved selectivity towards individual BET proteins.

1.5.2. Resistance of BET bromodomain monovalent inhibitors

Apart from the toxicity of the inhibitors, several studies have reported resistance emerging after (+)-JQ1 treatment. For example, the Wnt/β-catenin signalling pathway was shown to be increased in human leukaemia BET inhibitor resistance model[195]. Whole-exome sequencing analysis showed no mutations in BETs and or copy number aberrations in these resistant cell lines, but a decrease in chromatin-bound BET was observed. However, critical BRD4 target genes such as *MYC* were equally expressed in resistant cells despite the loss of BRD4 from the

MYC enhancer region. Global transcriptome analyses showed several transcriptional changes in resistant cells. Gene set enrichment analyses revealed that the NF-κB pathway was significantly downregulated, whereas both the transforming growth factor beta (TGF-β) and Wnt/β-catenin pathways were significantly upregulated in the resistant cells, thus facilitating MYC re-expression[195,196]. Blocking Wnt targets sensitised leukaemia cells to (+)-JQ1 and led to re-suppression of c-MYC. Overall, these results demonstrated that resistant cells continued to proliferate via c-MYC expression, but not via BRD4-mediated MYC expression and instead used Wnt/β-catenin pathway signalling to maintain the MYC level.

Similar studies have been done in other types of resistant cancer cell lines[197] where, in resistant triple-negative breast cancer cells, protein phosphatase 2 (PP2A) activity was shown to decrease and lead to hyperphosphorylation of BRD4, which then binds more strongly to MED1, facilitating a BD-independent chromatin recruitment mechanism. In Luminal/HER2+ cells, mutation of *PIK3CA* also caused resistance to BET inhibitors. Overexpression of PI3KCA switched sensitive breast cancer cells to become resistant cells. Blocking the PI3K signalling pathway sensitised resistant cells to JQ1 treatment[198]. Together, these results demonstrated that treatment with BET inhibitor could lead to resistance in different disease models and targeting resistance pathways could help to further increase the clinical efficacy of BET inhibitors.

1.6. Summary and project aim

Many proteins possess an arrangement of ordered functional domains joined by extended disordered regions. Individual weak interactions can be strongly enhanced through multivalent interactions, a phenomenon known as the avidity effect[58,199,200]. The development of bivalent or multivalent compounds for those multi-domain proteins can also act as a method of using the avidity effect to increase potency. Past attempts to produce bivalent binders for BDs and other domains showed the feasibility of this approach[167]. However, these multivalent

compounds do not address longstanding challenges in the BET field: lack of selectivity between BRD2, BRD3, BRD4 and BRDT. But why? From my perspective, the one critical factor that has not been taken into account when comparing all four BETs is the extended disordered linker region. Do we really understand how much a flexible linker region can contribute to the molecular function of the multi-domain proteins? What do we know about the linker region?

The objective of my study is to uncover if the differences in the inter-domain linker region of BET proteins cause their differences in supertertiary structure[201] upon bivalent binding and if it is possible to define a BET member selective multivalent inhibitor by studying the differences in supertertiary structure of BET proteins upon bivalent binding.

Chapter 2. BET topology and conformational change upon small molecule binding

2.1. Introduction

In the biological world, the structure of a protein plays a key role in determining its function. To explore the longstanding question of the basis for functional differences in BET members, we need to first examine their differences at the supertertiary structure levels. BET is a multidomain protein with two conserved N-terminal BDs and an extra terminal protein recruiting domain (ET). Previous studies mostly focused on the role of the single domain and failed to identify the difference in BETs. In this chapter, BETs were explored as whole multidomain protein where the linker regions between domains were compared and their conformation with and without binding of bivalent inhibitors were studied to answer if there is any supertertiary structure difference present in BETs and if it is possible to elicit the conformational difference in BETs with bivalent compounds. The set of bivalent inhibitors introduced in the study contained different warheads and different PEG linker lengths to answer if there is an optimum PEG linker length for bivalent engagement of multidomain proteins. For further exploration, the affinity of certain bivalent inhibitors with the special binding mode on BETs that caused the conformational change of the protein was then examined by ITC. To determine if the linker region in the protein plays any role, linker chimaeras were cloned, and the biophysical properties of these chimaera proteins upon binding of bivalent inhibitors were studied in a similar manner to how we studied the wild-type BETs.

Conformation of the proteins or the protein-compound complexes in this chapter was explored by AUC and SAXS. These are powerful biophysical methods to study the overall shape and size of biological macromolecules in solution. In an AUC experiment, the sedimentation profile of a profile is monitored by detecting and recording the movement of the protein boundary during centrifugation in real-time using an ultraviolet light absorption detection system. The radial position of the protein boundary and absorbance data stored by the instrument can then

be used for further mathematical analysis via direct boundary modelling to the Lamm equation[202] for obtaining sedimentation coefficient distributions $c(s)$ and translational frictional ratio (f/f_0) data (**Figure 2.1**). The sedimentation coefficient value (s) describes the moving rate of the species present. A higher sedimentation coefficient value is expected for a protein with a larger mass and spherical shape compared with that of a protein with a smaller mass and elongated shape. The translational frictional ratio (f/f_0) is the parameter used for describing the shape of the species. When two proteins have the same mass, the more spherical protein is expected to have a higher sedimentation coefficient value. A protein reaches its maximum s value (s_{sphere}) and minimum frictional coefficient (f_0) when it is absolutely spherical. The ratio of the maximum s-value to the observed s-value, $s_{sphere}/s_{w,20}$, is equal to the ratio of the experimental frictional coefficient to the minimum frictional coefficient (f/f_0), named the translational frictional ratio, which measures the maximum shape asymmetry from a sphere[203]. The shape of a protein can be inferred by comparing the frictional coefficient to that of a spherical protein with the same mass, where an f/f_0 equals to 1 means a protein is spherical; an f/f_0 value of 1.4 infers that the protein is globular while a value greater than 1.4 means the protein shape is elongated. In order to compare the extend of shape change manifested in each protein following the binding of a bivalent compound, we coined the term 'shape change percentage' or SC%, defined as:

$$SC\% = \frac{\left(s^0_{w,20}\right)_{complex} - \left(s^0_{w,20}\right)_{apo}}{\left(s^0_{w,20}\right)_{apo}} \times 100$$

where $s^0_{w,20}$ represents the sedimentation coefficient value corrected to water at 20 ºC.

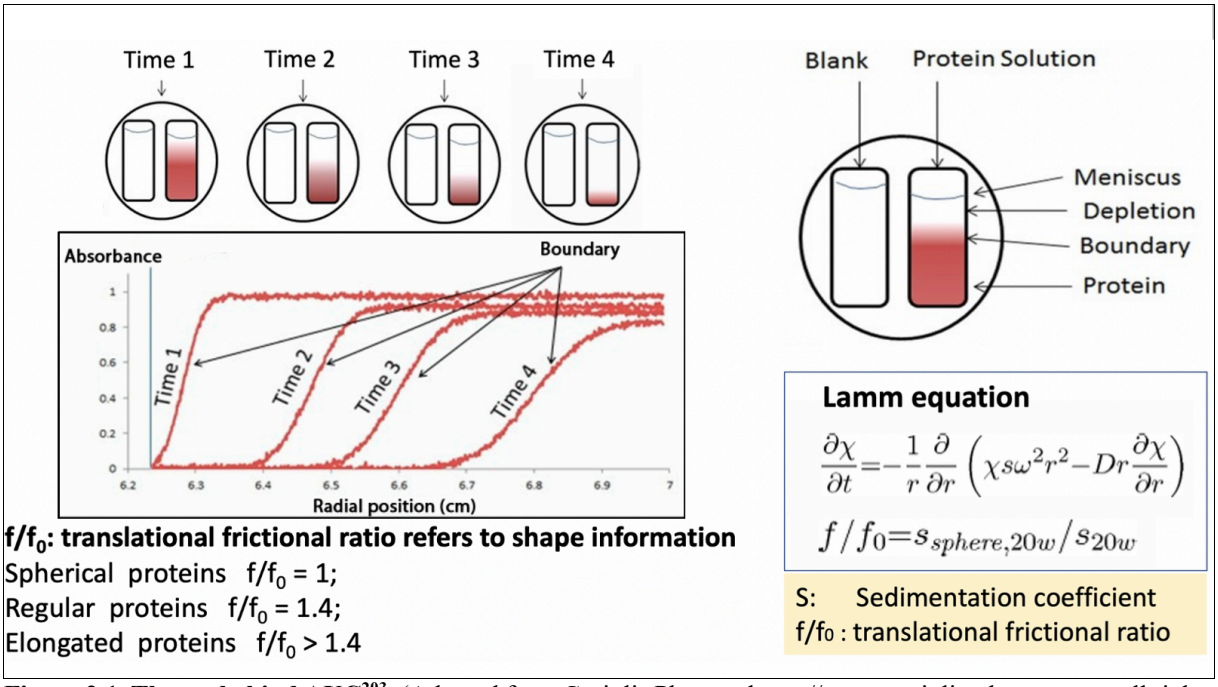
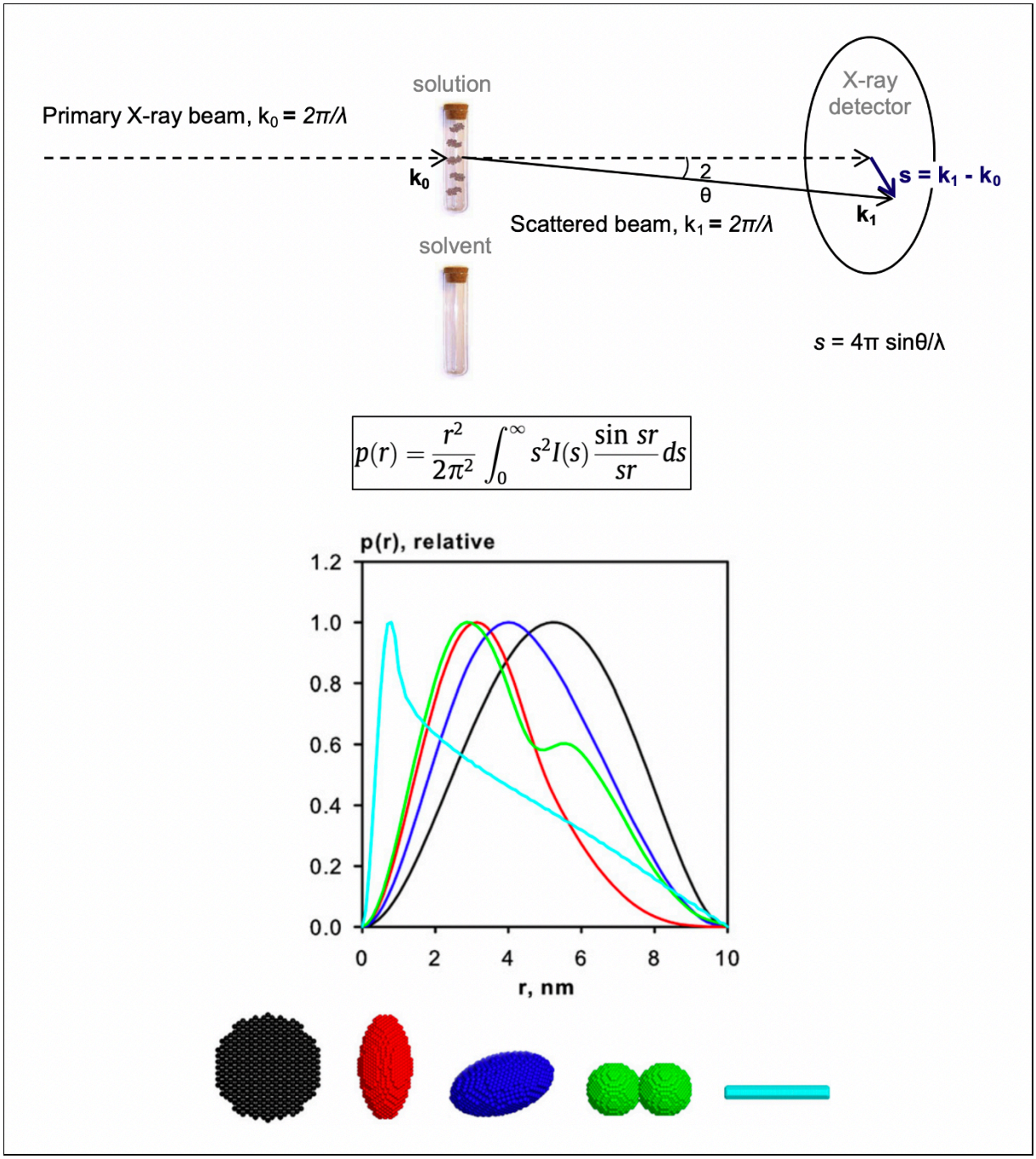


Figure 2.1. Theory behind AUC[203]. (Adapted from Coriolis Pharma, https://www.coriolis-pharma.com, all rights reserved)

In a SAXS experiment, a protein solution placed in a quartz capillary is exposed to a collimated monochromatic X-ray beam, and the scattered intensity is recorded by a detector as a function of the scattering angle[204]. The scattering pattern of the pure solvent is also collected and subtracted from the scattering of the protein solution. The resulting scattering pattern is related to the overall shape and size of the protein under investigation[205,206]. The advanced biological SAXS experiments are usually conducted on synchrotron sources providing high brilliance X-rays; in our case, the experiments were conducted at Diamond Light Source (Didcot, UK). Several characteristic parameters of the investigated protein can be obtained from the scattering pattern. The protein sample can be described by the pair distance distribution function *p(r)*, which is a real-space representation of the scattering data and allows one to graphically display the peculiarities of the particle shape. The *p(r)* function can be obtained from the experimental scattering data using indirect Fourier transformation[207] (**Figure 2.2**). Globular compact proteins have a *p(r)* function with a symmetric bell-shaped single peak, whereas elongated flexible proteins have an extended tail at large r values and can have multiple peaks[208]. Furthermore, an *ab initio* approach can be used for the reconstruction of a low-resolution 3D

model from scattering data[209], which is now a standard procedure for rapid characterisation. The most popular *ab initio* bead-modelling program in current use is DAMMIN (Dummy Atom Model Minimisation)[210]. However, the scattering curves can only fit an *ab initio* approach within a restricted range; therefore, the results should be complemented with other methods. In our experiments, the SAXS characterisation of the protein shape was studied and complemented with AUC experiments.



$$p(r) = \frac{r^2}{2\pi^2}\int_0^\infty s^2 I(s)\frac{\sin sr}{sr}ds$$

Figure 2.2. Schematic representation of a typical SAXS experiment and *p(r)* calculated for typical geometric shape. Solid sphere (black), prolate ellipsoid (red), oblate ellipsoid (blue), two-domain (green) and elongated (cyan)[206,211]. (Reprinted with permission form John Wiley and Sons, and Elsevier. This figure was published as figure 1 from FEBS Letters 589 (2015) pp: 2570-2577 and figure 3 from Journal of Structural Biology 172 (2010) pp: 128-141)

The affinity measurement between the proteins and compounds was studied by ITC. This is a powerful biophysical method to determine the thermodynamic parameters (the binding affinity, K_a; enthalpy changes, ΔH; and binding stoichiometry, n) of the interaction between two or more molecules in solution. ITC, therefore, remains as the 'gold standard' for evaluating small molecule–BD interactions[17,150,156,164,212-214]. In an ITC experiment, the heat evolved or absorbed in sample solution as a result of mixing precise amounts of reactants is directly measured. A spinning syringe is utilised for injecting and mixing of reactants. From these initial measurements, the Gibbs free energy changes ΔG and entropy changes ΔS can be determined using the relationship:

$$\Delta G = -RT \ln Ka = \Delta H - T\Delta S$$

Where R is the gas constant and T is the absolute temperature.

ITC is particularly useful as this gives not only the binding affinity but also the thermodynamic parameters of the binding. This allows us to further study detailed binding processes and the designing of small molecules.

2.2. BET proteins contain multiple domains linked by a variable flexible linker

In order to better understand the functional difference between BETs, the structural organisation was first compared, and the difference in the linker size between the conserved domains was observed (**Figure 2.3**). Disorder prediction analysis using the GlobPlot disorder propensity algorithm[83] showed that linker regions are disordered.

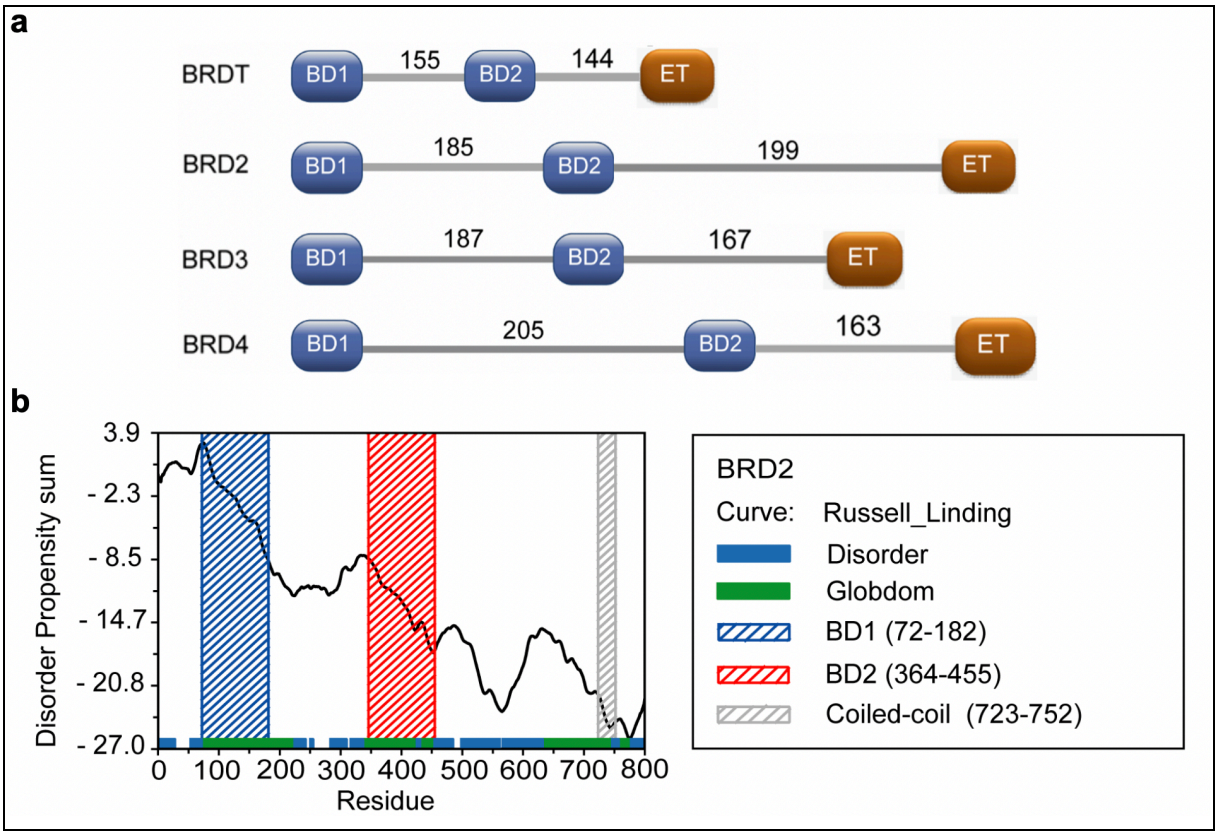


Figure 2.3. Domain organisation of BET proteins. (**a**) BET proteins contain two N-terminal tandem bromodomain (BD1, BD2) responsible for acetyl-lysine recognition and an extra-terminal protein-protein interaction domain (ET). Domain distances between tandem BDs and BD2, ET for all family members are annotated in the figure. (**b**) BRD2 full length disorder prediction using GlobPlot (http://globplot.embl.de/) is shown in the figure. BD and ET domain is predicted to be ordered and globular while the linker regions in between is predicted to be disordered.

This type of domain organisation suggested that BET family proteins are flexible and can adopt different conformations in solution. To test this hypothesis, we next chose to study the shape, size and conformational differences of BET constructs carrying two domains linked by their flexible linker in solution.

2.3. BET tandem BDs adopt an extended conformation in solution

To study the conformation of BET proteins constructs carrying more than one domain, each family member of recombinant BET tandem BDs, which contain both the BDs and linker region in between them, was cloned, expressed and purified in *E. coli* (**Figure 2.4**).

Target	Included domains	Mass (Da)	Protein PI	Construct ID
BRDT (1:2)	BD1 & BD2	42590.2	9.35	BRDTA-c389
BRD2 (1:2)	BD1 & BD2	43431.5	9.85	BRD2A-c019
BRD3 (1:2)	BD1 & BD2	44491.5	9.55	BRD3A-c923
BRD4 (1:2)	BD1 & BD2	47243.6	9.45	BRD4A-c323

Table 2.1. BET tandem BD proteins. Properties and construct IDs of BET tandem BDs proteins used in our experiments are shown.

For all the BET tandem BD proteins, a four-step purification was applied. The first step was immobilised metal affinity chromatography (IMAC) using nickel-charged columns as all the proteins contained a His-tag. Although the whole purification process was conducted at 4 °C, the presence of the flexible linker region resulted in some degree of cleavage and degradation of the protein in this region (extra lower bands in the **Figure 2.4a** SDS-PAGE analysis), and we also suspected some of the cleavages happened during the expression process. To prevent extra degradation during the purification process, the whole purification was conducted as quickly as possible. Next, we used size exclusion chromatography (SEC) to remove the smaller sized cleavage proteins (**Figure 2.4**). Although cleaved proteins were successfully separated from the intact proteins, the overall purity of the target protein after SEC was not ideal. We then applied cation-exchange chromatography to further purify the target protein and eventually obtained the relatively high-purity target proteins. Finally, we used mass spectrometry to confirm the identity of the target protein. All the four tandem BD proteins were purified as described above.

After successful producing and purifying the four BET tandem BDs (**Table 2.1**) (see section 6.2.1 for purification processes of all BET tandem BDs), AUC and SAXS were used to interrogate the shape of these recombinant BET tandem BDs in solution (**Figure 2.5**).

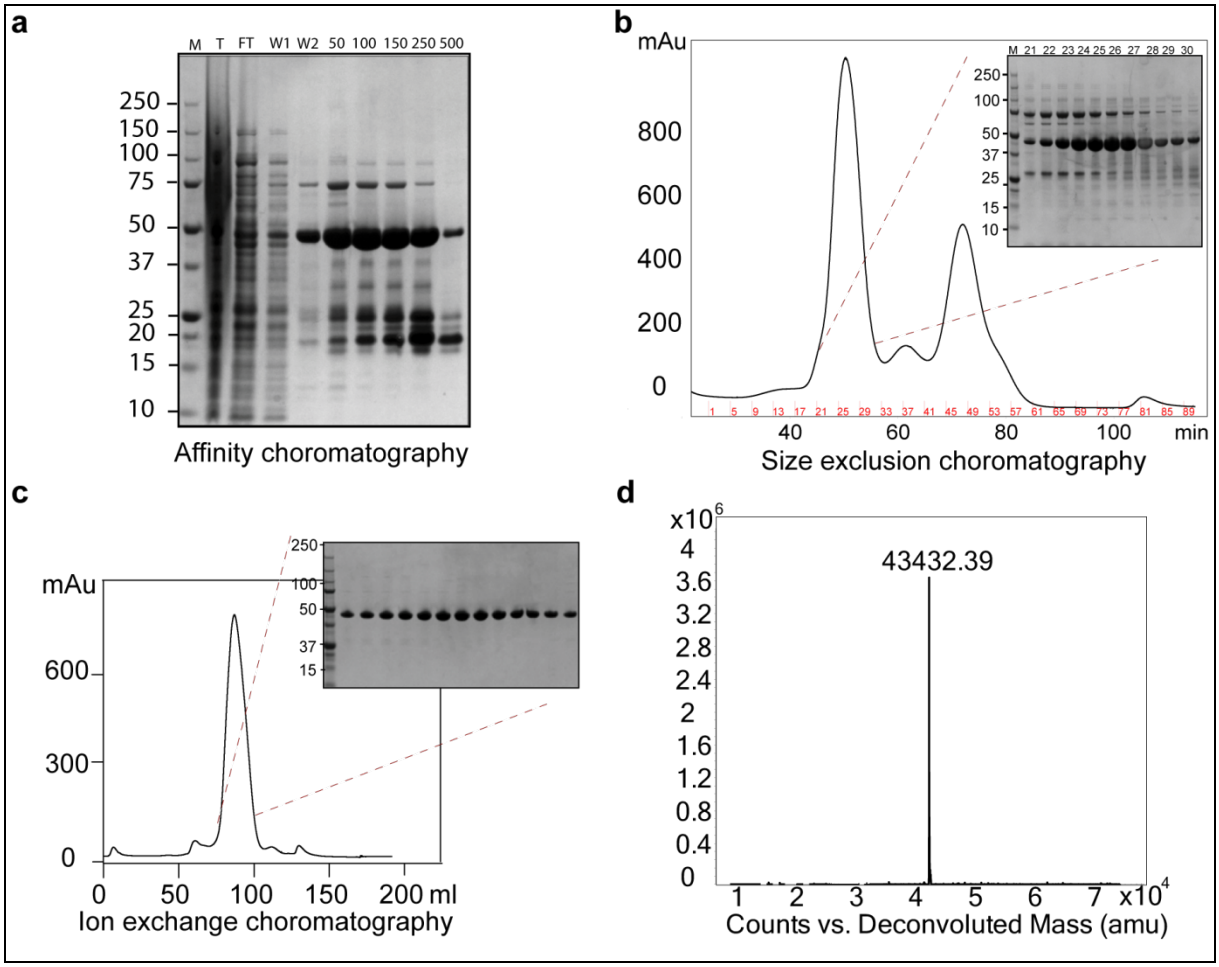


Figure 2.4. Typical protein purification process (here shown for BRD2(1:2)). (a) SDS-PAGE analysis of BRD2(1:2) after affinity chromatography purification. (NuPAGE™ NovexTM 4-12% Bis-Tris protein gel) (**b**) Size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus system in buffer (10 mM HEPES, 500 mM NaCl and 5% glycerol) and SDS-PAGE analysis of the collected fractions (**c**) Cation-exchange chromatography on HiTrap™ SP-HP column on an ÄktaPure system eluted with 1M NaCl in HEPES pH 7.5 buffer and SDS-PAGE analysis of the collected fractions. (**d**) Mass spectrometry confirmation of purified BRD2(1:2) protein in Quad Time-Of-Flight mass spectrometry systems.

Analysis of AUC data followed boundary fitting employing the LAMM equation in SEDFIT, which allowed for the estimation of translational frictional ratio (f/f_0) of each construct tested. This analysis yielded differences in f/f_0 values for the four BET tandem BDs (**Figure 2.5c**), consistent with dispersed monomers of highly extended conformation as well as with the linker size of each protein where BRDT(1:2), with the shortest linker region among four BET tandem BDs, had the smallest f/f_0 value, and BRD4(1:2), with the longest linker region, had the largest f/f_0 value.

As an orthogonal method, we conducted *ab initio* bead modelling approach of SAXS data to construct a low-resolution 3D model of the BET tandem BDs. The results showed the inter-domain distances of 103 Å and 116 Å for BRDT(1:2) and BRD2(1:2) constructs, respectively (**Figure 2.5d**). According to the obtained distances, several systematic studies[215-217] based on Flory's theory of polymer physics[218] were used as references to determine if the inter-domain linker region behaves more as a disordered flexible region or a compact region[219]. From Flory's theory, the expected radius of gyration of a protein containing N residues follows a power law, $R_g=R_0N^{\nu}$, where R_0 and ν depend on the structural behaviour of the protein in the solvent. R_0 and ν have been estimated in different studies depending on whether the studied protein is in globular or disordered conformation (**Table 2.2**). Thus, by using these values and Flory's theory, we can calculate the theoretical linker length of a protein that contains the same number of residues as BRDT(1:2) or BRD2(1:2) and compare the calculated length to the experimentally obtained length to determine if the linker region is more flexible or compact. The calculated results showed that a folded globular protein with 155 residues would have a length of 31.6 Å and with 185 residues would have a length of 33.8 Å. A disordered protein with 155 residues would have a length of 70.6 Å, and with 185 residues would have a length of 77.6 Å (**Table 2.2**). The experimental inter-domain distances of 103 Å and 116 Å (**Figure 2.5d**) were closer to the values predicted for a disordered protein than they are to a globular protein, suggesting the inter-domain linker regions of both BRDT(1:2) and BRD2(1:2) behave more as a disordered flexible region.

Protein conformation	R_0 (Å)	ν	N=155		N=185	
			R_g	Length	R_g	Length
Globular[215]	2.32	0.38	15.8 Å	31.6 Å	16.9 Å	33.8 Å
Denatured[216]	1.93	0.598	39.4 Å	78.8 Å	43.8 Å	87.6 Å
Disordered[217]	2.54	0.522	35.3 Å	70.6 Å	38.8 Å	77.6 Å

Table 2.2. Predicted length of proteins in different conformation. Lengths of proteins with the same residue numbers as linker region of BRDT(1:2) and BRD2(1:2) (N=155, N=185) were predicted. The R_0 and ν values were taken from the cited corresponding studies. R_g was calculated based on power law $R_g=R_0N^{\nu}$. The length was estimated to be two times the R_g value.

Furthermore, the pair distance distribution function *p(r)* of BRDT(1:2) and BRD2(1:2) exhibited an extended tail at large *r*, which is the characteristic shape of an elongated species with a long flexible linker (**Supplementary figure 1, Supplementary table 1**). The *ab initio* shapes of all four BET tandem BD constructs were also published in a recent paper[72], demonstrating their extended shape. These SAXS analyses further supported the hypothesis that BET tandem BDs adopt extended conformations in solution. Overall, it is safe to state that these BET tandem BDs are in an extended conformation.

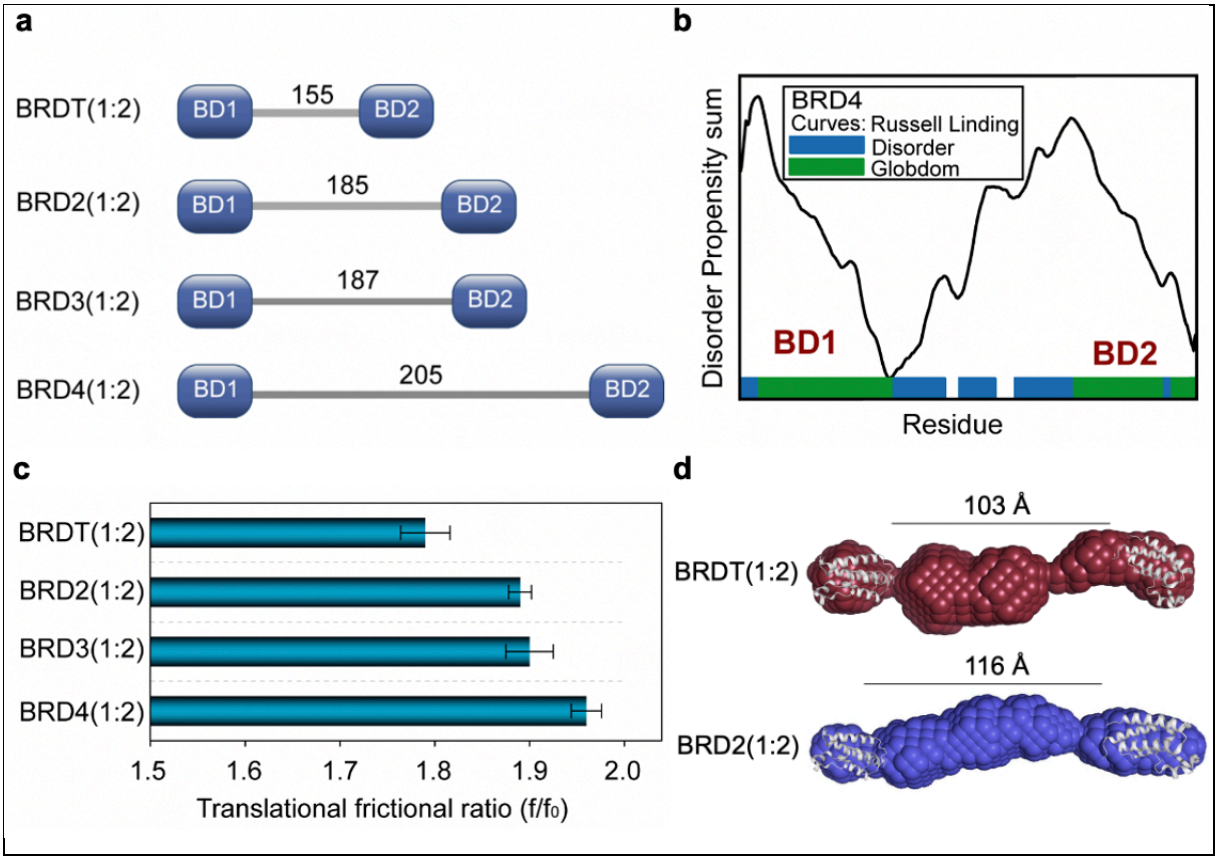


Figure 2.5. Conformational study of BET tandem BDs. (**a**) Domain distances between tandem BDs for all four family members are annotated in the figure. (**b**) BRD4 tandem BDs linker disorder propensity prediction is shown in the figure, the linker region between the BDs is predicted to be disordered. (**c**) Translational frictional ratio (f/f_0) values for all four tandem BDs obtained from the in-solution characterization by sedimentation velocity (SV) AUC are shown, suggesting the elongated shape of each protein and in agreement with the short to long order of the tandem BD linker length shown in (a) (Experiment repeated 15 times for each apo BET tandem BDs, error bars represent mean ± standard deviation). Experiments were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 ºC. (**d**) *Ab initio* shapes of the BRDT(1:2) and BRD2(1:2) constructs restored from the SASX data; mean distances are shown next to the models, in agreement with hydrodynamic analysed elongated shape in (c).

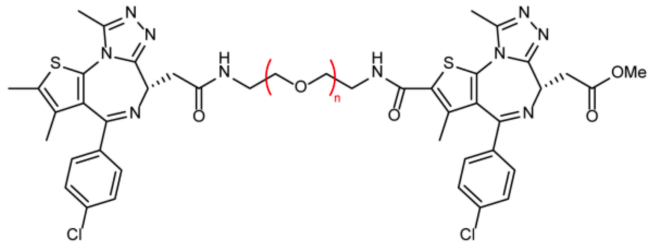
Taken together, the AUC and SAXS data suggest that BET tandem BD constructs adopt extended conformations in solution, as shown by their calculated f/f_0 value (> 1.7). Intriguingly, the translational frictional ratios (f/f_0) follow the same order dictated by the linker length connecting the two folded BD domains within its protein (BRDT < BRD2 ~ BRD3 < BRD4). These observations were orthogonally confirmed for the smallest construct BRDT(1:2) as well as for BRD2(1:2), which were found to have an estimated inter-domain distance of 103 and 116 Å, respectively, following *ab initio* reconstruction of the volume occupied by these proteins in SAXS experiments. It is tempting to speculate that these supertertiary structure differences may form a basis to harness selectivity towards BET proteins by designing inhibitors that would exploit these subtle differences. To this end, two series of bivalent compounds that theoretically have the capacity to interact with both BDs at the same time were used in the following experiments.

2.4. BET tandem BDs are differentially compacted when engaged by bivalent inhibitors

2.4.1. Bivalent inhibitors simultaneously bind two BDs

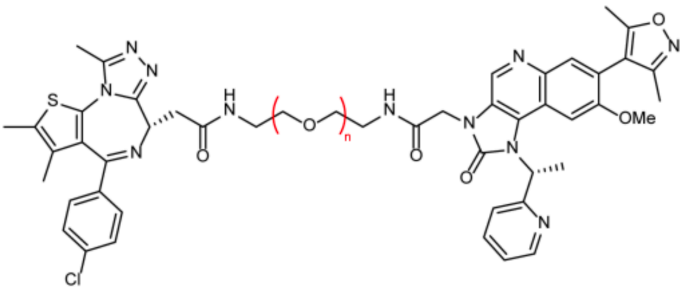
Two series of bivalent inhibitors (BvA-n and BvB-n **Table 2.3 and 2.4**) carrying a flexible PEG linker of variable length were synthesised (in collaboration with Jun Qi's lab, DFCI/Harvard University) to explore the simultaneous engagement of both BDs within the BET tandem BD constructs. All these bivalent inhibitors were designed and synthesized by Dr. Minoru Tanaka, and the Alpha Screen assay of the bivalent inhibitors toward BRDT(1) was designed and performed by Dr. Justin M Roberts at the Jun Qi lab, Harvard University. BvA-n (**Table 2.3**) represents a series of synthetic compounds linking two identical warheads based on the thienodiazepine scaffold of JQ1[17] tethered via a variable PEG linker from the C3 position of the first to the C6 position of the second. BvB-n (**Table 2.4**) represents a series of synthetic compounds linking two distinct warheads (a thienodiazepine one based on the JQ1

chemotype[17] and an isoxazole one based on the IBET-151 chemotype[18]) via a variable PEG linker. Both series yield asymmetric species, one with the same warhead and the other with distinct warheads. Given parameters in the table for these two series of compounds are PEG linker number, PEG linker length, molecular weight (MW), total polar surface area (tPSA), and alpha screen potencies towards BRDT(1). The alpha screen is a bead based amplified luminescent proximity homogenous assay. When the donor and acceptor beads are brought into proximity by biological interaction, singlet oxygen evolved from donor beads travels to the spatially close acceptor beads after laser illumination, resulting in chemiluminescence. This assay can be used for the detection of protein-protein, protein-peptide interaction. The alpha screen competition binding assay can also be used to determine small molecule inhibition efficacy, which is represented by the alpha screen potencies in our tables below. Alpha screen potency here describes how many of these bivalent inhibitors are needed to inhibit BRDT(1) by 50%. The total polar surface area (tPSA) is the surface sum over all polar atoms in a molecule and is a metric for the optimization of a drug's ability to permeate cells.



ID	n	Linker length (Å)	MW	tPSA	Alpha: IC_{50} (nM) BRDT(1)
BvA - 0	0	0	870	165	2.36
BvA - 1	1	2.8	914	174	0.574
BvA - 2	2	5.3	958	184	1.84
BvA - 3	3	7.6	1002	193	0.878
BvA - 4	4	10.2	1046	202	1.09
BvA - 7	7	17.6	1178	230	1.59

Table 2.3. BvA-n bivalent inhibitors. Chemical representation and compound properties with different PEG size, including molecular weight (Mw), the total polar surface area (tPSA) and alpha screen potencies towards BRDT(1) are shown.



ID	n	Linker length (Å)	MW	tPSA	Alpha: IC_{50} (nM) BRDT(1)
BvB - 0	0	0	898	178	148
BvB - 1	1	2.8	942	187	12.1
BvB - 2	2	5.3	987	196	10
BvB - 3	3	7.6	1031	205	5.87
BvB - 4	4	10.2	1075	215	6.2
BvB - 7	7	17.6	1207	242	6.1

Table 2.4. BvB-n bivalent inhibitors. Chemical representation and compound properties with different PEG size, including molecular weight (Mw), the total polar surface area (tPSA) and alpha screen potencies towards BRDT(1) are shown.

The designing process, the biochemical and biophysical studies of these bivalent inhibitors towards the single domains BRDT(1) and BRD4(1) were described in detail by Jun Qi's paper[166]. In their affinity study, ITC experiments showed that JQ1 and BvA-1 bound with a similar range of K_D (40 and 17 nM, respectively) to BRD4(1), but the binding stoichiometry was different. JQ1 bound to BRD4(1) in a 1:1 ratio, whereas BvA-1 bound to BRD4(1) in a 1:2 ratio, which demonstrated the capacity of these bivalent inhibitors to bind two BDs simultaneously. In their crystallography study, the cocrystal structure of one of the bivalent inhibitors bound to two BRD4(2) (Protein databank [PDB]: 5JWM) further supported the 'one inhibitor simultaneously bound to two BDs' statement.

2.4.2. Bivalent inhibitor binding mode on BET tandem BDs

While the binding stoichiometry of bivalent inhibitors against a BET single BD can easily be visualized by two BDs independently engaging each warhead of the ligand, binding to BET tandem BDs may be more complicated.

Given the presence of two BDs in BET tandem BD constructs, there are several binding modes for such bivalent inhibitors. For example, one ligand may engage both domains, resulting in a 1:1 *in cis* binding complex (**Figure 2.6a**) or engage just one domain, yielding the same stoichiometry in solution (**Figure 2.6b**). Additional modes may include two ligands engaging the tandem BD constructs from independent (**Figure 2.6c**) or the same (**Figure 2.6d**) warheads, leading to a 2:1 binding complex, one ligand engaging two tandem BD constructs, leading to a 1:2 binding complex (**Figure 2.6e**) or two ligands engaging two tandem BD constructs in a 2:2 (or 1:1) binding complex (**Figure 2.6f**). In addition, higher-order multimeric species may form. The relative population of each is likely to vary with both protein and compound concentration. In particular, the 2:1 binding complex (**Figure 2.6c&d**) is likely to dominate at a very high concentration of the compound as the protein becomes saturated with the ligand.

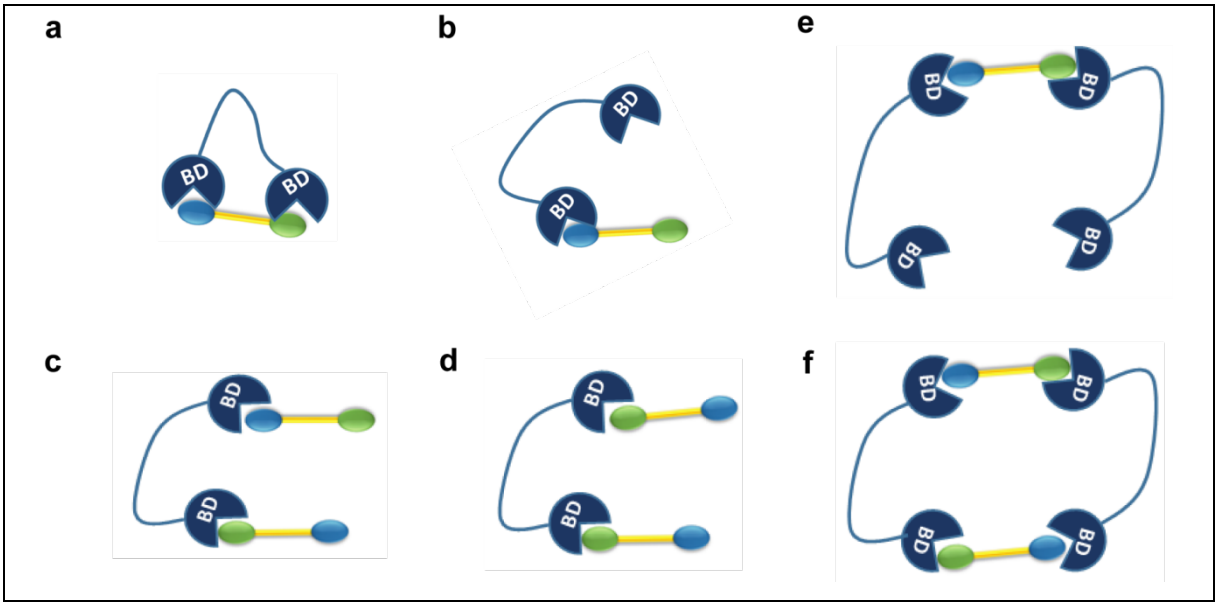


Figure 2.6. Possible binding mode for bivalent inhibitors. (**a**) 1:1 *in cis* bivalent binding. (**b**) 1:1 binding to one domain. (**c**)(**d**) 2:1 binding where two compounds binding to one tandem BD with different or same warhead binding to each BD. (**e**) 1:2 *in trans* binding where one compound binding to two tandem BD proteins. (**f**) 1:1 *in trans* binding.

To address different binding modes of BET tandem BDs to BvX-n inhibitors, we examined the topology of BET tandem BD constructs in the presence of bivalent inhibitors in solution. Here,

we described the supertertiary structure analysis of BRDT(1:2) in the presence of BvX-3 as an example.

Hydrodynamic analysis of the tandem BRDT(1:2) in the presence of BvX-3 at 1:1 protein:ligand ratio suggested that compound engagement results in faster sedimentation (larger sedimentation coefficient value s) (red curve in **Figure 2.7a**), when compared with that of the apo protein BRDT(1:2) (dashed curve in **Figure 2.7a**), suggesting that the 1:1 *in cis* binding complex may be present (see **Figure 2.6a**), which would justify a much more compacted ensemble. Interestingly, at higher ligand to protein ratios (e.g. at 1:3 protein: ligand ratio), isolation of a single peak in the *c*(*s*) analysis (green curve in **Figure 2.7a**), manifesting between the apo-BRDT(1:2) and the 1:1 *in cis* binding complex, indicates the potential saturation of the protein binding sites, resulting in a 2:1 binding complex as illustrated in **Figure 2.6c, d**. Intrigued by this shift in complex consistency, we next performed the same experiment employing sub-stoichiometric amounts of the inhibitor, using a 1:0.5 protein to ligand ratio. This led to a mixture of two states (blue curve in **Figure 2.7a**), suggesting the coexistence of both apo-BRDT(1:2) and the 1:1 *in cis* binding complex where the first peak (of the blue curve) resulted from the slower sedimentation of apo-BRDT(1:2) and the second peak resulted from the faster sedimentation of 1:1 *in cis* binding complex.

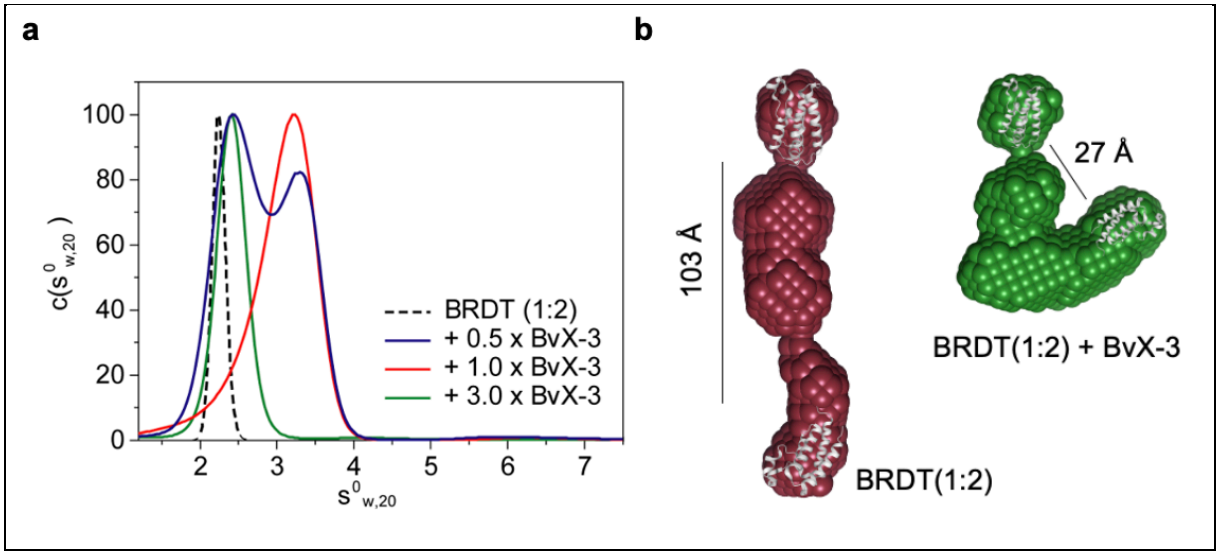


Figure 2.7. Biophysical characterisation of bivalent inhibitors binding mode. (**a**) Sedimentation coefficient distribution corrected to water at 20 °C for BRDT(1:2) the presence of bivalent inhibitor BvX-3 at 1:0.5 (blue line), 1:1 (red line) and 1:3 (green line) protein: ligand ratio. Obtained from AUC sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (**b**) *Ab initio* shapes of the BRDT (1:2) constructs restored from the SAXS data for apo-BRDT and 1:1 complex with bivalent inhibitor BvX-3; mean distances are shown next to the models.

To orthogonally confirm these observations, we next studied the behaviour of BRDT(1:2) in solution employing SAXS in the presence of the same bivalent inhibitor BvX-3. SAXS experiments performed with a 1:1 protein to ligand ratio resulted in a largely monomeric species at a low concentration of 1 mg/mL, whereas at higher concentrations (5 and 10 mg/mL) the protein tended to aggregate. Intriguingly, the protein became progressively and noticeably more compact compared with apo-BRDT(1:2) as exemplified by the *ab initio* reconstruction of the calculated volume (**Figure 2.7b**), further demonstrating the existence of a 1:1 *in cis* binding complex. Indeed, simultaneous engagement of both domains would yield a more compact species given the short length between the two warheads (7.6 Å), consistent with the hydrodynamic results. However, the minimum inter-domain distance measured on the SAXS bead model was around 27 Å, which is much larger than the length (7.6 Å) between the two warheads of the inhibitor. This could be explained by the fact that the bead model generated by the SAXS experiment represents an average conformation. In our experiment, most likely, there is a population of fully compact species and a population of less compact species with

only one warhead engaged, resulting in a larger average inter-domain distance than it is supposed to be.

Having established the possibility of obtaining a compact species with a 1:1 protein to ligand ratio, we decided to use this stoichiometry and explore the effect of the PEG linker length in BET tandem BD binding. We hypothesised that different series of compounds that exploit the distance between the two warheads would yield differential binding, allowing a harnessing of selectivity between BET proteins.

2.4.3. BvX-n inhibitor led to BET tandem BD differential compaction

We established in 2.4.2 that bivalent inhibitors engage BET tandem BDs in solution from both BDs, resulting in distortion of the extended arrangement of the two domains and leading to a more compacted structure. Several possibilities arise from these observations: first, it is possible that the length between the two warheads can be attenuated to differentially affect the compaction of the tandem BET BDs; second, the type of warheads used could be engineered to promote binding towards the four tandem BET-BDs when combined with the appropriate PEG linker.

In the following experiments, we added the indicated bivalent inhibitors to the proteins at 1:1 protein:ligand ratio to the BET tandem BDs, and extracted the translational frictional ratio (f/f_0), corrected sedimentation coefficient ($s^0_{w,20}$) and calculated molecular weight (Mw) to see the effect of all of the bivalent inhibitors to the BET tandem BDs (**Table 2.5 & Table 2.6**). In most of the cases, compound engagement resulted in smaller f/f_0 values and larger sedimentation coefficient values, suggesting the successful formation of a compacted complex. However, compound BvA-7 induced two peaks or two different sedimentation coefficient values (**Table 2.5** in red) in the case of BRD2(1:2) and BRD4(1:2) constructs where the second sedimentation coefficient value was not only high but also had a very high f/f_0 (~3.2) value, indicating the

presence of chain like dimer state of the protein (illustrated in **Figure 2.8**). Interestingly, this behaviour was restricted to the BvA-n series and specifically to the compound carrying a PEG linker size of n=7, but not when engaging the BRDT(1:2) construct that has the shorter length linker among all of the BET tandem BDs.

Protein	Compound	AUC				
	BvA - n	f/f_0	s	s^0_{20w}	**MW(Da)**	**rmsd(OD)**
BRDT(1:2)	-	1.80	1.42 ± 0.08	2.37	42682	0.0070
	BvA - 0	1.69	1.61 ± 0.07	2.70	43621	0.0104
	BvA - 1	1.59	1.69 ± 0.13	2.82	44856	0.0101
	BvA - 2	1.69	1.61 ± 0.12	2.70	45231	0.0105
	BvA - 3	1.66	1.63 ± 0.14	2.74	44879	0.0087
	BvA - 4	1.71	1.61 ± 0.10	2.69	45807	0.0090
	BvA - 7	1.73	1.59 ± 0.13	2.66	47389	0.0089
BRD2(1:2)	-	1.89	1.31 ± 0.07	2.19	43540	0.0076
	BvA - 0	1.81	1.53 ± 0.10	2.56	45807	0.0076
	BvA - 1	1.79	1.55 ± 0.10	2.59	45730	0.0075
	BvA - 2	1.82	1.53 ± 0.15	2.57	48037	0.0078
	BvA - 3	1.69	1.60 ± 0.14	2.67	45931	0.0078
	BvA - 4	1.67	1.61 ± 0.17	2.70	45559	0.0083
	BvA - 7	1.82	1.51 ± 0.10	2.55	48452	0.0079
		3.21	2.03 ± 0.12	3.40	82357	
BRD4(1:2)	-	1.96	1.38 ± 0.07	2.32	46892	0.0067
	BvA - 0	1.78	1.61 ± 0.10	2.68	47477	0.0105
	BvA - 1	1.76	1.62 ± 0.09	2.70	46667	0.0055
	BvA - 2	1.74	1.63 ± 0.12	2.72	48236	0.0052
	BvA - 3	1.69	1.65 ± 0.13	2.76	50933	0.0105
	BvA - 4	1.72	1.64 ± 0.16	2.74	47428	0.0053
	BvA - 7	1.83	1.58 ± 0.12	2.65	51582	0.0104
		3.52	2.24 ± 0.08	3.74	91675	

Table 2.5. In solution evaluation of bivalent inhibitor BvA-n binding to BET tandem BDs by AUC. Hydrodynamic values including translational frictional ratio (f/f_0), apparent sedimentation coefficient (s), corrected sedimentation coefficient ($s_{20,w}$), and system mass (Mw) values of each BET tandem BDs in the presence of compound BvA-n with indicated PEG size at 1:1 molar ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (Shown errors on s values represent fitted s distribution value peak area, rmsd is the root-mean-square deviation of detected protein boundary from the boundary generated by Lamm equation modelling).

Protein	Compound	AUC				
	BvB - n	f/f_0	s	s^0_{20w}	MW(Da)	rmsd(OD)
BRDT(1:2)	-	1.80	1.42 ± 0.08	2.37	42682	0.0070
	BvB - 0	1.64	1.62 ± 0.11	2.72	43336	0.0076
	BvB - 1	1.61	1.64 ± 0.11	2.75	44022	0.0063
	BvB - 2	1.63	1.63 ± 0.11	2.73	43562	0.0060
	BvB - 3	1.70	1.61 ± 0.11	2.70	43731	0.0072
	BvB - 4	1.72	1.61 ± 0.11	2.70	44974	0.0065
	BvB - 7	1.71	1.60 ± 0.09	2.69	43773	0.0074
BRD2(1:2)	-	1.89	1.31 ± 0.07	2.19	43540	0.0076
	BvB - 0	1.73	1.58 ± 0.12	2.65	44689	0.0083
	BvB - 1	1.61	1.70 ± 0.14	2.85	44816	0.0079
	BvB - 2	1.70	1.61 ± 0.13	2.69	45346	0.0080
	BvB - 3	1.75	1.56 ± 0.11	2.61	44102	0.0079
	BvB - 4	1.69	1.60 ± 0.13	2.67	45425	0.0077
	BvB - 7	1.68	1.59 ± 0.13	2.67	44126	0.0080
BRD4(1:2)	-	1.96	1.38 ± 0.07	2.32	46892	0.0067
	BvB - 0	1.67	1.64 ± 0.12	2.75	46199	0.0067
	BvB - 1	1.62	1.68 ± 0.12	2.82	46687	0.0055
	BvB - 2	1.73	1.63 ± 0.12	2.73	47439	0.0052
	BvB - 3	1.79	1.60 ± 0.13	2.67	47284	0.0065
	BvB - 4	1.79	1.62 ± 0.16	2.69	48518	0.0053
	BvB - 7	1.67	1.65 ± 0.13	2.77	45952	0.0065

Table 2.6. In solution evaluation of bivalent inhibitor BvB-n binding to BET tandem BDs by AUC. Hydrodynamic values including translational frictional ratio (f/f_0), apparent sedimentation coefficient (s), corrected sedimentation coefficient ($s_{20,w}$), and system mass (Mw) values of each BET tandem BDs in the presence of compound BvB-n with indicated PEG size at 1:1 molar ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (Shown errors on s values represent fitted s distribution value peak area, rmsd is the root-mean-square deviation of detected protein boundary from the boundary generated by Lamm equation modelling).

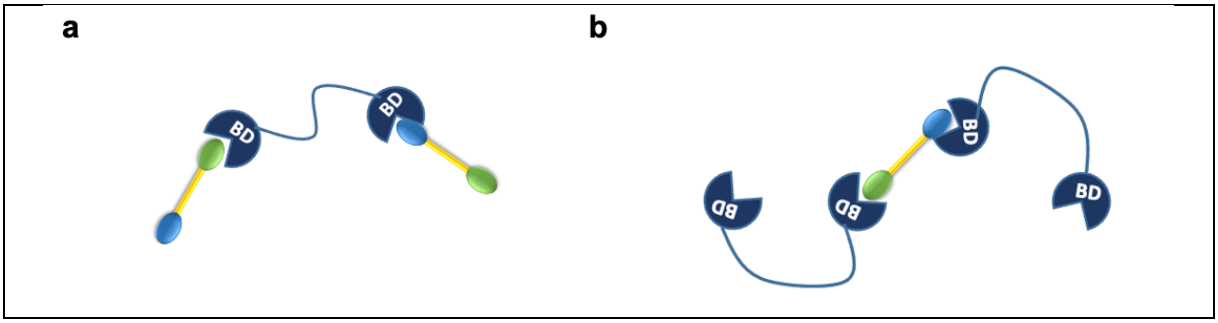


Figure 2.8. Two possible binding modes of BvA-7 engaging BRD4(1:2) at 1:1 protein:ligand molar ratio. Same for BRD2(1:2). (**a**) Species number one presented in the solution with the lower apparent sedimentation coefficient (s) value (**b**) Species number two presented in the solution with the higher apparent sedimentation coefficient (s) value and high translational frictional ratio (f/f_0) value.

To further compare and evaluate the extent of shape change manifested in each protein following engagement of a bivalent inhibitor, we calculated the 'shape change percentage' value of each protein upon indicated compound engagement (**Figure 2.9**). Inhibitors from both series yielded shape change (%) values > 12% with all BET tandem BDs, suggesting that a minimal distortion takes place. Except from oligomer formation with BvA-7 compound on BRD2(1:2) and BRD4(1:2) discussed above, there was no clear trend present between the different sized PEG compounds in either the BvA-n series and BvB-n series. However, compound BvB-1 showed relatively high compaction towards BRD2(1:2) construct by 30% shape change in comparison with that of other compounds.

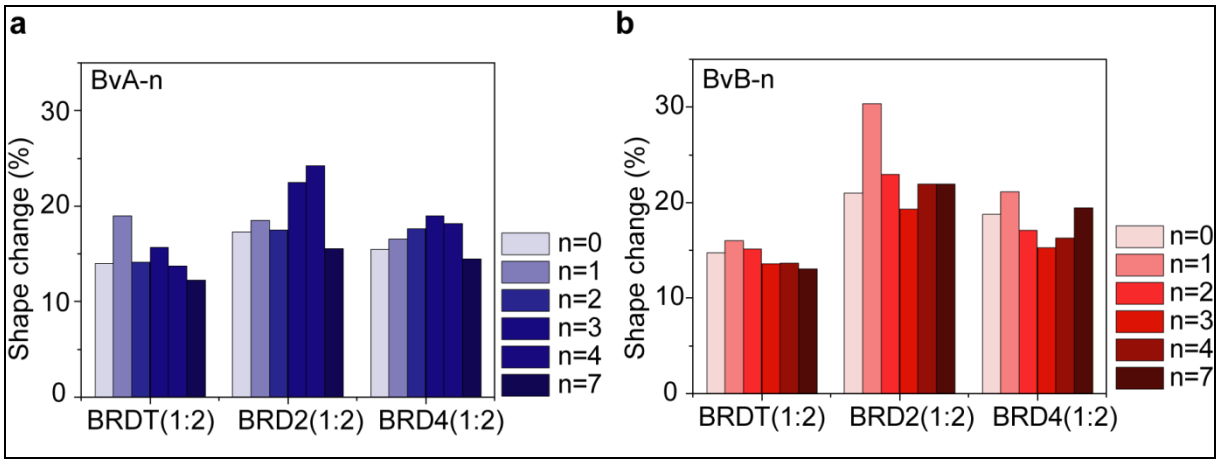


Figure 2.9. Compacting efficacy of bivalent inhibitors on tandem BET BD proteins (**a**) Shape change percentage of BET tandem BDs in the presence of compound BvA-n at 1:1 protein:ligand ratio (**b**) Shape change percentage of BET tandem BDs in the presence of compound BvB-n at 1:1 protein:ligand ratio. Shape change percentage is calculated from $s_{20,w}$ values of each complex obtained from above tables, equals to $\frac{(s^0_{w,20})_{complex}-(s^0_{w,20})_{apo}}{(s^0_{w,20})_{apo}} \times 100\%$ (This comparison experiment was only performed once).

To further study the underlying mechanics of why BvA-7 compound caused oligomerisation of BRD2(1:2) construct while other BvA-n compounds did not and to answer why BvB-1 led to relatively high compaction of BRD2(1:2) construct but not to others, ITC experiments were combined with AUC experiments and analysed in detail.

2.5. Bivalent inhibitor PEG linker length and warheads are both important determinants that changed BRD2(1:2) binding mode

2.5.1. Compound BvA-7, but not BvA-3, leading to BRD2(2) oligomerization is the main cause of BRD2(1:2) oligomerization with BvA-7 but not with BvA-3

To investigate the causes leading to different binding behaviour of compounds within the same series with different PEG sizes, the interaction with BRD2(1:2) as well as with BRD2(1) and BRD2(2) were further studied in solution with AUC and ITC. For the BvA-n series, BvA-3, which did not cause oligomerisation of BRD2(1:2), and BvA-7, which did cause oligomerisation, were compared and characterised first.

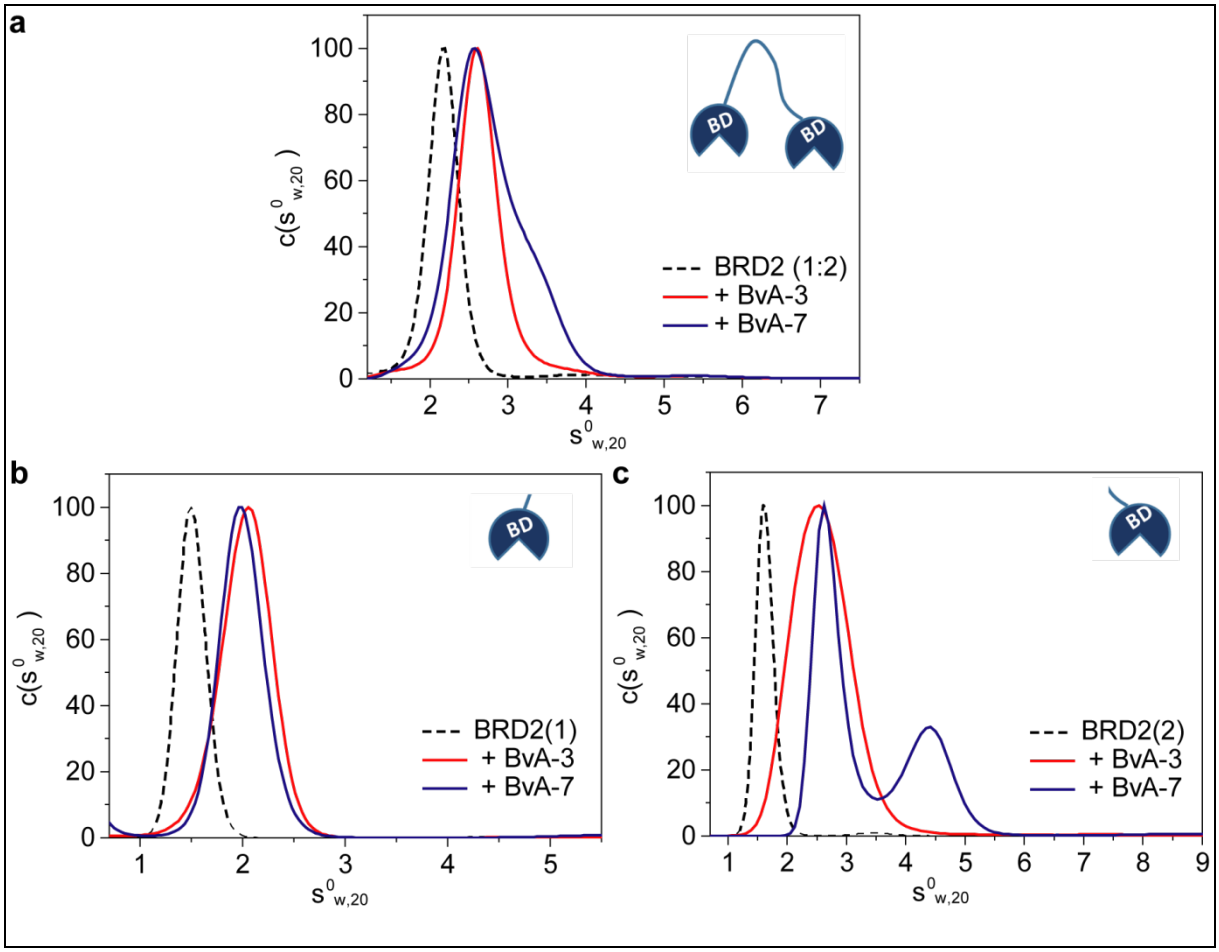


Figure 2.10. In solution evaluation of bivalent inhibitor BvA-3 and BvA-7 binding to BRD2 BDs by AUC. (**a**) Sedimentation coefficient distribution corrected to water at 20 °C for BRD2(1:2) in the presence of compounds BvA-3 (red line) or BvA-7 (blue line) at 1:1 protein:ligand ratio. (**b**)(**c**) Sedimentation coefficient distribution corrected to water at 20 °C for BRD2(1) or BRD2(2) in the presence of compounds BvA-3 or BvA-7 at 1:1

protein:ligand ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C.

From the sedimentation distribution plots in **Figure 2.10** (see **Table 2.7** for extracted values), a comparison of BvA-3 and BvA-7 addition to BRD2(1:2) clearly demonstrates the existence of multiple sedimenting species in the case of BvA-7, as exemplified by the formation of a shoulder in the sedimentation distribution plot (**Figure 2.10a**), which suggests the existence of an additional faster sedimenting species (albeit in small amount) and indicates the formation of protein oligomer. There were no differences observed between the addition of BvA-3 and BvA-7 to BRD2 first bromodomain BRD2(1), and both of them formed BRD2(1) dimers, as expected (**Figure 2.10b**). However, in the case of the BRD2 second bromodomain BRD2(2), the addition of BvA-7 formed two separate peaks in the sedimentation distribution trace (**Figure 2.10c**). The second peak corresponds to a much faster sedimenting species with a calculated molecular weight of 57,155 (**Table 2.7**), suggesting the formation of a tetramer or higher oligomer of BRD2(2), whereas the addition of BvA-3 to BRD2(2) only formed one peak indicative of the forming of a BRD2(2) dimer, as expected. The formation of a tetramer of BRD2(2) with the addition of BvA-7 was not expected since BvA-n compounds only have two warheads that can capture only two BDs.

Protein	Compound	AUC				
	BvX - n	f/f_0	s	s^0_{20w}	**MW(Da)**	**rmsd(OD)**
BRD2(1:2)	-	1.89	1.31 ± 0.07	2.19	43540	0.0076
	BvA - 3	1.69	1.60 ± 0.14	2.67	45931	0.0078
	BvA - 7	1.82	1.51 ± 0.10	2.55	48452	0.0079
		3.21	2.03 ± 0.12	3.40	82357	
BRD2(1)	-	1.42	0.92 ± 0.08	1.54	15667	0.0160
	BvA – 3	1.43	1.25 ± 0.13	2.09	29606	0.0154
	BvA - 7	1.39	1.22 ± 0.12	2.04	28123	0.0145
BRD2(2)	-	1.30	0.99 ± 0.09	1.66	14378	0.0139
	BvA – 3	1.21	1.58 ± 0.27	2.64	27835	0.0136
	BvA - 7	1.34	1.64 ± 0.12	2.74	28535	0.0141

			2.68 ± 0.18	4.48	57155	

Table 2.7. in solution evaluation of bivalent inhibitor BvA-3 and BvA-7 binding to BRD2 BDs by AUC. Hydrodynamic values including translational frictional ratio (f/f_0), apparent sedimentation coefficient (s), corrected sedimentation coefficient ($s_{20,w}$), and system mass (Mw) values of BRD2 BDs in the presence of compound BvA-3/7 at 1:1 molar ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (Shown errors on s values represent fitted *s* distribution value peak area, rmsd is the root-mean-square deviation of detected protein boundary from the boundary generated by Lamm equation modelling).

To shed more light into this behaviour, we next performed in solution thermodynamic analysis of the BvA-3 and BvA-7 interactions with BRD2 constructs, seeking to understand the underpinning mechanics of protein-compound engagement. ITC experiments were performed as previously described (section 2.1.2), by injecting solutions of the protein into the compound (reverse titration) to avoid signal saturation due to DMSO into water dilution.

Titration of BRD2(1:2) into BvA-3 resulted in 1:1 molar ratio binding. BvA-7 binding to BRD2(1:2) was roughly in a 1:1 molar ratio but was in a non-regular sigmoidal curve (**Figure 2.11**). When comparing the thermodynamic parameters between the BvA-3 and BvA-7 binding to BRD2(1:2), positive $-T\Delta S$ value was observed for the BvA-7 binding while $-T\Delta S$ value was negative for BvA-3 binding (**Table 2.8**), suggesting the interaction between BRD2(1:2) and BvA-7 but not BvA-3 caused a loss of conformational freedom. BvA-7 bound to BRD2(1) in 1:2 molar ratio, while the interaction between BvA-3 and BRD2(1) showed a biphasic isotherm with different thermodynamic parameters (**Table 2.8**). The first binding site of BvA-3 on BRD2(1) had an *N* value of 1.14 and the second had a value of 0.63. These values were not expected, considering BRD2(1) was titrated into BvA-3 in our experiment. Assuming protein concentration and compound concentrations were measured as close to the real concentration as possible, the exact reasons for the unexpected *N* values need further experimental clarification. In terms of other parameters, the first binding event has both favourable enthalpy and entropy contributions, while the second binding event was driven mainly by much larger favourable enthalpy and cost a large penalty on conformational freedom lost. These differences in the enthalpy and entropy between the first binding event and second binding event suggested

the binding mechanism difference of BRD2(1) to the two sites of BvA-3, where during the second binding there were more hydrogen or other non-covalent bonds formed and there was a loss of conformational freedom compared to the first binding event. However, these differences were not shown in the interaction between BRD2(1) and BvA-7, suggesting the certain role of the inhibitor's PEG linker size.

Interestingly, BRD2(2) binding to BvA-3 and BvA-7 yielded a non-sigmoidal titration curve (**Figure 2.11e, f**). The interaction between BRD2(2) and BvA-3 (**Figure 2.11e**) showed a biphasic isotherm where the initial integrated heats of injection showed a trend towards positive, and later data trended negatively until saturation reached. They have two inflexion points, indicating the existence of at least two different binding events. The interaction between BRD2(2) and BvA-7 (**Figure 2.11f**) was found to be triphasic, with the first two phase integrated heats of injection trending positively but with different inflexion points, and the third phase integrated heats of injection trending negatively, indicating the existence of at least three different binding events. This observation correlates well with the previous hydrodynamic observations in the AUC assay (**Figure 2.10**), where the interaction between BvA-7 and BRD2(2) induced BRD2(2) tetramerization, suggesting the existence of more than two binding events. We speculated that in the interaction between BRD2(2) and BvA-7, two copies of the BRD2(2) dimers (**Figure 2.12a**) associated to form a tetramer (**Figure 2.12b**). The precise spatial relationship of the two BRD2(2) within the dimer (**Figure 2.12a**) might just happen to favour the association of the two dimers to form a tetramer. This might be the reason why BvA-7, not BvA-3 or other BvA-n compounds, caused the formation of the BRD2(2) tetramer. However, the exact mechanism of how the dimers were associated to form the tetramer is not clear. **Figure 2.12** is just a speculation that needs further experimental validation. Altogether, these results demonstrated that, in this particular compound series BvA-n, PEG linker size in

the compound has a substantial effect on protein binding behaviour. If the PEG linker is too long, the risk of forming an oligomer is higher when binding to a certain protein.

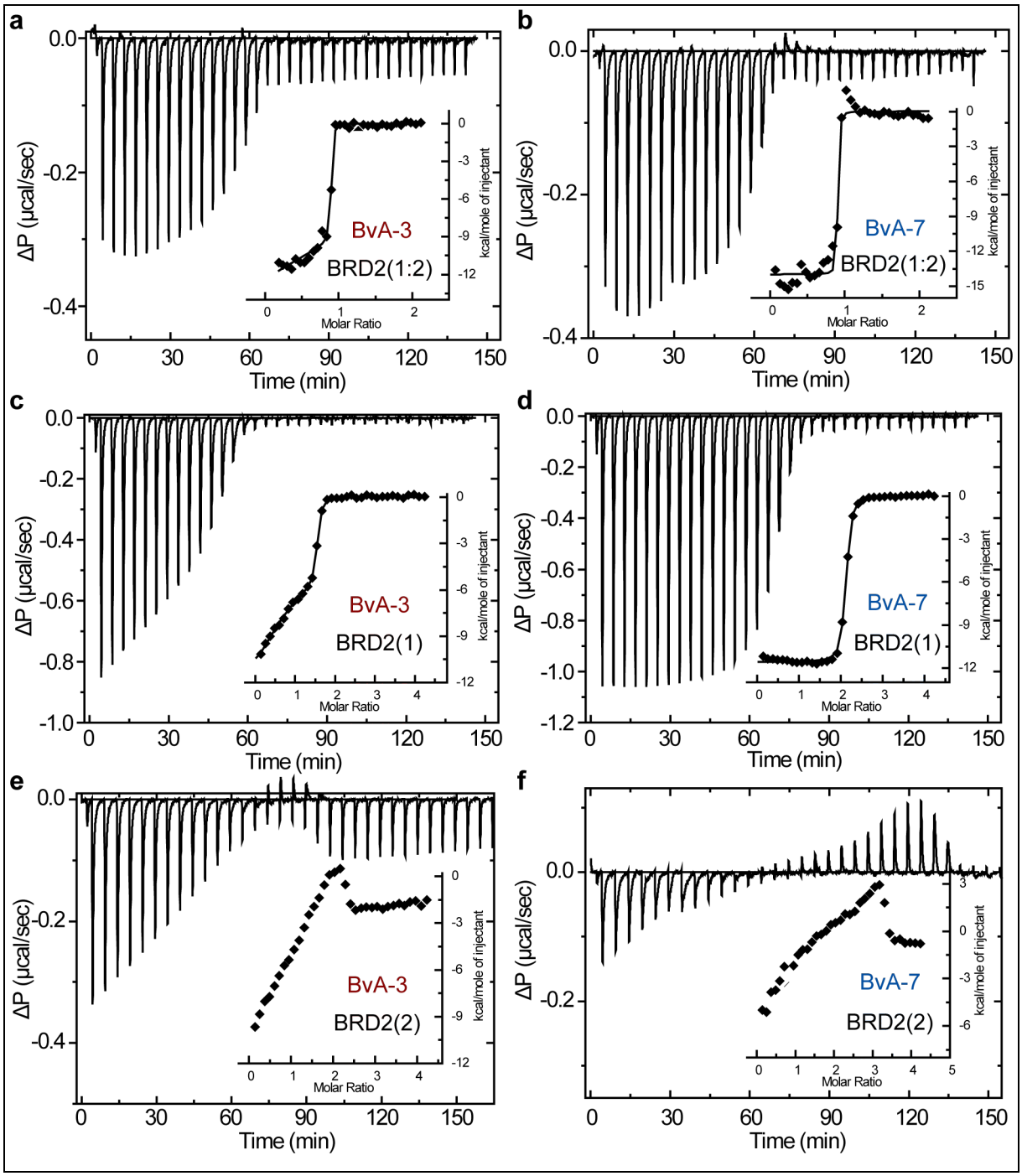


Figure 2.11. In solution evaluation of bivalent inhibitor BvA-3 and BvA-7 binding to BRD2 BDs by ITC. Raw injection heats for titration of indicated proteins into a solution of indicated compounds and the normalized binding enthalpies corrected for the heat of protein dilution are shown. Solid lines represent a nonlinear least square fit using a multiple site binding model. Titrations were performed in 50 mM HEPES pH 7.5, 100 mM NaCl, and 15 °C while stirring at 295 rpm. (**a**) ITC evaluation of BRD2(1:2) binding to BvA-3, and (**b**) BvA-7. (**c**) ITC

evaluation of BRD2(1) binding to BvA-3, and (**d**) BvA-7. (**e**) ITC evaluation of BRD2(2) binding to BvA-3, and (**f**) BvA-7

Protein	BvX-n	ITC				
		K_D (nM)	*N*	Δ*H* (kcal/mol)	-*T*Δ*S* (kcal/mol)	Δ*G* (kcal/mol)
BRD2(1:2)	BvA-3	3.4 ± 0.4	0.92 ± 0.14	-10.66 ± 0.95	-1.32 ± 0.62	-11.98 ± 1.17
	BvA-7	2.9 ± 0.9	0.93 ± 0.09	-14.51 ± 1.11	3.06 ± 0.91	-11.45 ± 1.02
BRD2(1)	BvA-3	36.2 ± 12.4	1.14 ± 0.13	-3.40 ± 0.74	-7.49 ± 1.22	-10.89 ± 1.26
		18.5 ± 9.5	0.63 ± 0.21	-19.72 ± 1.91	9.26 ± 1.74	-10.46 ± 1.15
	BvA-7	28.5 ± 10.3	2.11 ± 0.05	-13.08 ± 2.41	2.16 ± 0.70	-10.92 ± 1.11
BRD2(2)	BvA-3	-	-	-	-	-
	BvA-7	-	-	-	-	-

Table 2.8. In solution evaluation of bivalent inhibitor BvA-3 and BvA-7 binding to BRD2 BDs by ITC. Thermodynamic values including dissociation constant (K_D), binding molar ratio (*N*), Enthalpy change (Δ*H*), and Entropy change (Δ*S*), obtained from Figure 2.11. ITC fitting. Experiments were performed in duplicate. (Shown errors correspond to the mean value ± the range.)

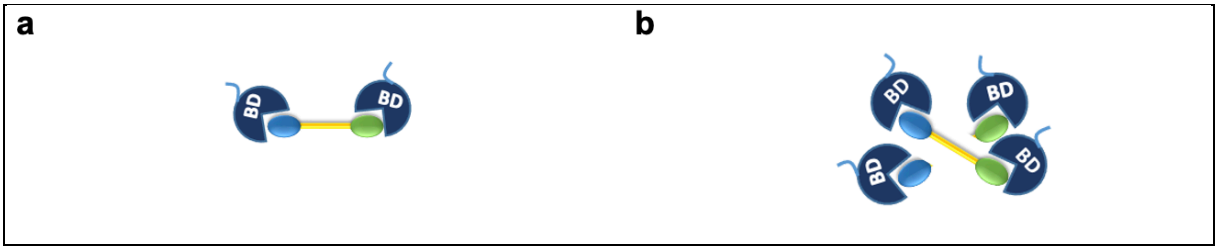


Figure 2.12. Possible binding mode of BvA-7 engaging BRD2(2). (**a**) Species number one (dimer) presented in the solution with the lower apparent sedimentation coefficient (*s*) value (**b**) Species number two (tetramer) presented in the solution with the higher apparent sedimentation coefficient (*s*) value.

2.5.2. Selectivity of BvB-1 one warhead towards BRD2(2) comparing to the other warhead might be the main reason for why BvB-1 can compact the BRD2(1:2) to the greatest extend

To investigate the causes leading to the high degree of compaction of BRD2(1:2) when engaging compound BvB-1 compared with that found with other compounds, the BvB-1 interactions with BRD2(1:2) as well as BRD2(1) and BRD2(2) were further studied in solution with AUC and ITC. All the AUC and ITC experiments were conducted as previously described.

From the sedimentation distribution plots in **Figure 2.13** (see **Table 2.9** for extracted values), the formation of a single peak with a higher sedimentation coefficient value was observed with BRD2(1:2) upon addition of BvB-1 compound suggesting compaction of the protein. The BRD2 first bromodomain, BRD2(1), formed a dimer, as expected. However, in the case of BRD2(2), the addition of BvB-1 formed two separate peaks in the sedimentation distribution trace (**Figure 2.13e**). Interestingly, this time, the second peak corresponds to the dimer form of the BRD2(2) while the first peak corresponds to the BRD2(2) monomer (**Table 2.9**), suggesting insufficient binding of BRD2(2) to one (or both) BvB-1 warheads. This observation was further explained by ITC experiments where titration of BRD2(2) into compound BvB-1 resulted in two binding events with largely different affinities: one with nanomolar affinity, the other with micromolar affinity (**Table 2.10**). The difference in the affinity of two warheads towards BRD2(2) resulted in the insufficient formation of BRD2(2). In contrast, the interaction of BvB-1 with BRD2(1) had a nanomolar affinity yielding a 1:2 molar ratio binding, corresponding with the dimer formation in the AUC experiment. Interestingly, the interaction of BvB-1 with BRD2(1:2) resulted in higher affinity binding (**Table 2.10**) compared with the interactions between BRD2(1) and BRD2(2), suggesting the successful exploitation of the avidity effect.

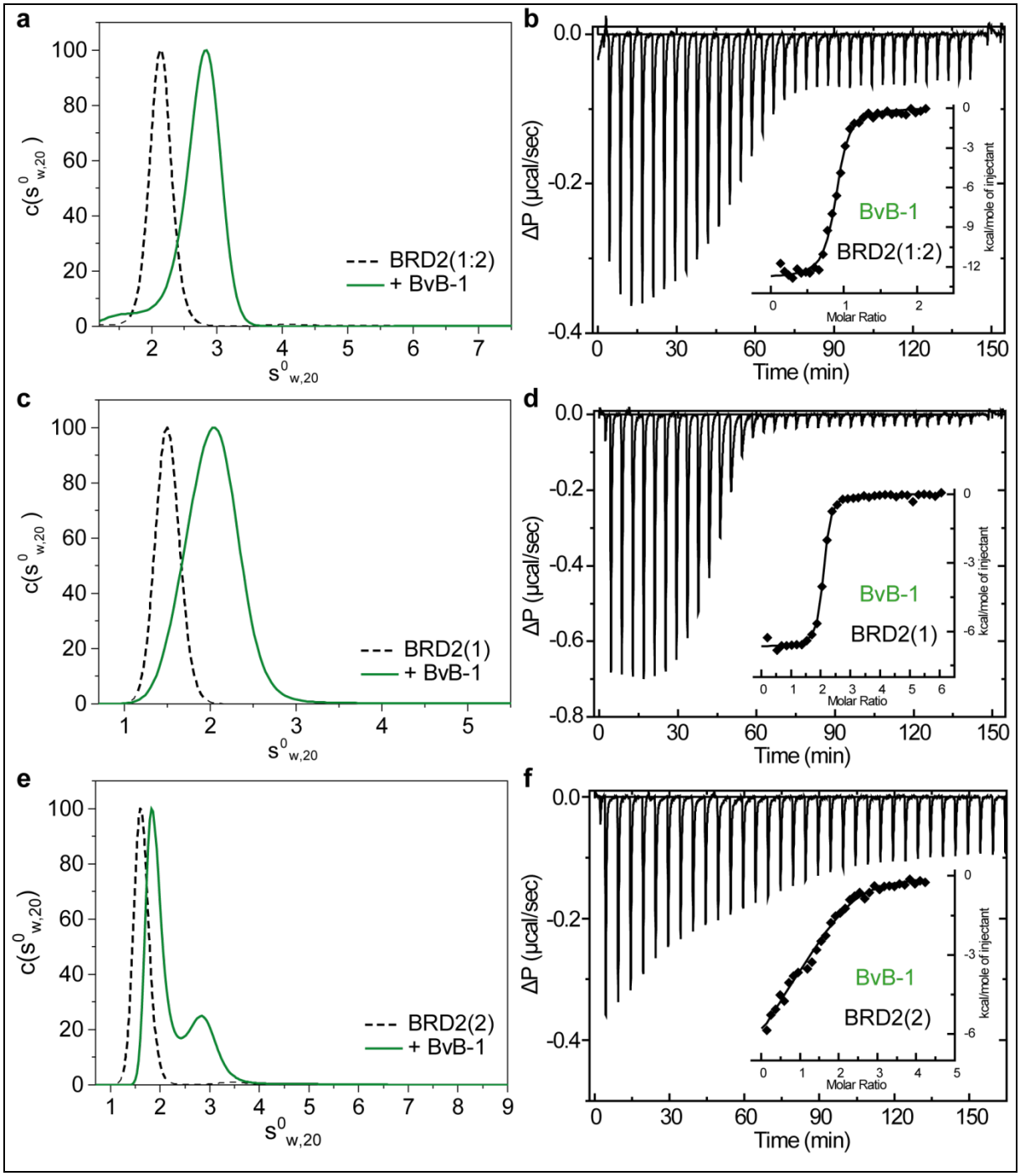


Figure 2.13. In solution evaluation of bivalent inhibitor BvA-3 and BvA-7 binding to BRD2 BDs by AUC and ITC. (**a**)(**c**)(**e**) Sedimentation coefficient distribution corrected to water at 20 °C for BRD2(1:2) or BRD2(1) or BRD2(2) in the presence of compounds BvB-1 at 1:1 protein:ligand ratio. (**b**)(**d**)(**f**) ITC evaluation of indicated protein biding to compound BvB-1.

Protein	Compound	AUC				
	BvX - n	f/f_0	s	s^0_{20w}	MW(Da)	rmsd(OD)
BRD2(1:2)	-	1.89	1.31 ± 0.07	2.19	43540	0.0076
	BvB - 1	1.61	1.70 ± 0.14	2.85	44816	0.0079
BRD2(1)	-	1.42	0.92 ± 0.08	1.54	15667	0.0160
	BvB - 1	1.47	1.24 ± 0.16	2.07	28850	0.0157
BRD2(2)	-	1.30	0.99 ± 0.09	1.66	14378	0.0139
	BvB - 1	1.42	1.14 ± 0.07	1.81	18463	0.0149
			1.70 ± 0.13	2.85	32948	

Table 2.9. In solution evaluation of bivalent inhibitor BvB-1 binding to BRD2 BDs by AUC. Hydrodynamic values including translational frictional ratio (f/f_0), apparent sedimentation coefficient (s), corrected sedimentation coefficient ($s_{20,w}$), and system mass (Mw) values of BRD2 BDs in the presence of compound BvB-1 at 1:1 molar ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (Shown errors on s values represent fitted s distribution value peak area, rmsd is the root-mean-square deviation of detected protein boundary from the boundary generated by Lamm equation modelling).

Protein	ITC				
	K_D (nM)	N	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
BRD2(1:2)	26.4 ± 12.5	1.00 ± 0.07	-12.04 ± 1.11	2.61 ± 0.98	-9.43 ± 1.07
BRD2(1)	109.8 ± 28.4	2.07 ± 0.11	-7.12 ± 0.89	-3.09 ± 1.02	-10.21 ± 1.11
BRD2(2)	1533.3 ± 651.1	1.77 ± 0.08	-1.64 ± 0.21	-7.53 ± 2.51	-9.17 ± 1.21
	129.8 ± 44.2	0.57 ± 0.18	-6.64 ± 0.46	-3.21 ± 0.87	-9.85 ± 1.33

Table 2.10. In solution evaluation of bivalent inhibitor BvB-1 binding to BRD2 BDs by ITC. Thermodynamic values including dissociation constant (K_D), binding molar ratio (N), enthalpy change (ΔH), and entropy change (ΔS), obtained from Figure 2.13. ITC fitting. Experiments were performed in duplicate. (Shown errors correspond to the mean value ± the range).

From the results, we speculated that the spacing and relative orientation between the two warheads of BvB-1 enable the two BDs within the BRD2(1:2) protein to engage in favourable inhibitor and BD interaction. This effect is specific to BvB-1 but not with other BvB-n compounds, and the preferential binding orientation towards BD2 might contribute to this effect. However, this speculation needs further experimental validation.

2.6. Switching the BRD2(1:2) linker region changes the compound binding behaviour

The previous section established that bivalent compounds that can engage both BDs within BET tandem BD constructs can bind to all BETs, which results in distortion of an extended and elongated apo protein and leads to more compacted species. Importantly, our data suggested that bivalent engagement leads to differential compaction of the BET tandem BD constructs. It is, therefore, tempting to speculate that the protein linker may also play a role in the BET tandem BD constructs engage bivalent inhibitors. To investigate this hypothesis, BRD2 tandem BD chimaeras were designed so that they either carried the shorter linker from BRDT named as BRD2(1:2)^{LT} or the longer linker from BRD4 named as BRD2(1:2)^{L4} (**Figure 2.14a**). These constructs were cloned, expressed, and purified as described in the Material and Methods section.

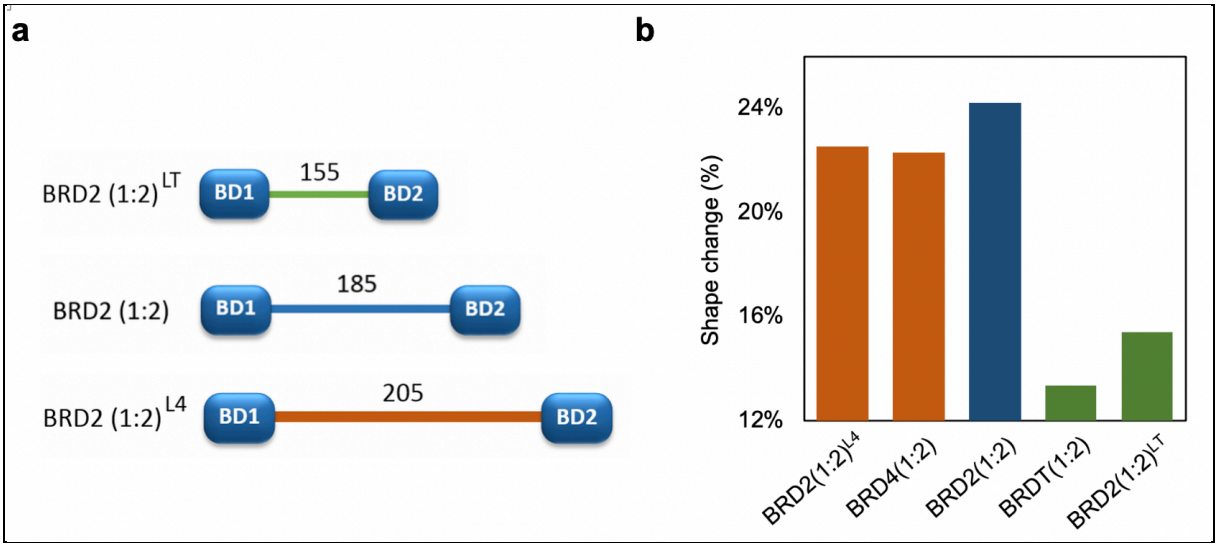


Figure 2.14. Chimaera tandem BDs binding behaviour to bivalent compounds (**a**) Chimaera tandem BDs architecture. BRD2(1:2)^{LT} comprised of BDs from BRD2(1:2) and linker region from BRDT(1:2); BRD2(1:2)^{L4} comprise of BDs from BRD2(1:2) and linker region from BRD4(1:2). (**b**) Shape change percentage of chimaera BET tandem BDs in the presence of BvA-3 bivalent compounds at 1:1 protein:ligand ratio. (This comparison experiment was only performed once).

Similar experiments were performed with wild-type BET tandem BDs, and the shape change percentage was calculated upon adding the bivalent compounds and compared with wild-type BET tandem BD constructs. Adding of compound BvA-3 caused 25% compaction of

BRD2(1:2), 22% compaction of BRD4(1:2), and 13% compaction of BRDT(1:2) (**Figure 2.14b**). BvA3 caused approximately 21% compaction of BRD2(1:2)^{L4}, which is similar shape change as BRD4(1:2), and caused around 15% compaction of BRD2(1:2)^{LT}, which is similar shape change as BRDT(1:2). This result showed that changing of linker region induces different binding behaviour, suggesting this linker region between the two domains may play an important role. However, the underlying mechanics for this observation remain unclear, which raises a number of questions: Do domains in tandem protein interact with adjacent linker region upon binding to small molecules inhibitors? Is the small difference in the linker sequence rigidity playing a role? Were the small molecules simply interacting with the protein linker region? For further exploration, cloning and testing a BET tandem BD protein with a random sequenced linker region with different length and different level of rigidity would be ideal. If the new proteins show similar binding behaviour to BET tandem BD constructs with corresponding linker length, we can assume linker length is the core determinant for the overall conformational change of each BET. If not, the sequence would be the only other factor.

2.7. Discussion

Despite their high degree of sequence similarity as well as sharing the same domain architecture, BET proteins adopt specific cellular functions, contributing to complicated transcriptional programmes. However, we do not understand what the specific structural determinant underpins their functional role(s). A common pattern in the literature is the *in vitro* study of interactions by the use of single domains. These studies have uncovered important information, identifying, for example, interactions of BET BDs with two adjacent Kac sites on histones[27,58] or with KacY motifs[72], and thereby laying down the groundwork for the development of potent and specific inhibitors that target the Kac site. In the context of longer protein sequences, however, individual domains appear to have different behaviours in vitro

and in cells[220]. To better understand domain specific contributions, one could study multiple domains in recombinant systems, seeking to identify specific patterns of behaviour that would discriminate between these highly conserved proteins.

In this chapter, we examined the relative topology of the two N-terminal bromodomain (BD) modules of the four BET proteins in solution, in the absence or presence of bivalent inhibitors that carried variable linker lengths. Secondary structure prediction analysis revealed that the BET conserved domains are linked by flexible region of variable length following the general trend BRDT < BRD2 ~ BRD3 < BRD4. Hydrodynamic analysis by AUC sedimentation velocity AUC established that tandem constructs bearing the two N-terminal BDs adopted extended and elongated conformations, and this was orthogonally confirmed by SAXS.

Protein	AUC	ITC	SAXS	Comment
BRDT(1:2)	f/f_0=1.80 s =1.42	-	Elongated shape R_g = 48 Å D_{max}=183 Å	Extended shape is consistent with BDs flexibly linked by a disordered linker
BRDT(1:2) + BvA-n	f/f_0 decreased to 1.59-1.73 s increased to 1.59-1.69 12-18% compacting efficiency	Not tested	Not tested	BvA-n compounds induce compaction, consistent with bivalent binding.
BRDT(1:2) + BvB-n	f/f_0 decreased to 1.61-1.72 s increased to 1.61-1.64 12-16% compacting efficiency	Not tested	V-shaped model with BvB-3 R_g = 44 Å D_{max}=135 Å	BvB-n compounds induce compaction, consistent with bivalent binding.
BRD2(1:2)	f/f_0=1.89 s =1.31	-	Elongated shape R_g = 55 Å D_{max}=190 Å	Extended shape is consistent with BDs flexibly linked by a disordered linker
BRD2(1:2) + BvA-n	f/f_0 decreased to 1.67-1.81 s increased to1.53-1.61 Dimer with BvA-7 s =3.21 15-24% compacting efficiency	With BvA-3 K_D =3.4 nM $-T\Delta S$ = -1.32 With BvA-7 K_D =2.9 nM $-T\Delta S$ = 3.06	Not tested	BvA-n compounds induce compaction, consistent with bivalent binding. Dimer observed with BvA-7
BRD2(1:2) + BvB-n	f/f_0 decreased to 1.61-1.73 s increased to1.56-1.70	With BvB-1 K_D =26.4 nM	Not tested	BvB-n compounds induce compaction, consistent

	18-30% compacting efficiency			with bivalent binding. Best compaction with BvB-1 with 30% compacting efficiency
BRD4(1:2)	f/f_0 =1.96 s =1.38	-	Not tested	Extended shape is consistent with BDs flexibly linked by a disordered linker
BRD4(1:2) + BvA-n	f/f_0 decreased to 1.69-1.78 s increased to 1.61-1.64 Dimer with BvA-7 s =3.52 13-16% compacting efficiency	Not tested	Not tested	BvA-n compounds induce compaction, consistent with bivalent binding. Dimer observed with BvA-7
BRD4(1:2) + BvB-n	f/f_0 decreased to 1.62-1.79 s increased to 1.60-1.68 15-20% compacting efficiency	Not tested	Not tested	BvB-n compounds induce compaction, consistent with bivalent binding

Table 2.11. Overview of the hydrodynamic and thermodynamic results done on BET tandem BDs with the BvX-n compounds.

To examine the properties of the linker bridging the two N-terminal BDs within each BET, we employed two series of bivalent BET inhibitors carrying two high affinity warheads linked by flexible PEG linkers of variable length. The hydrodynamic analysis demonstrated that engagement of these compounds results in differential compaction of BET tandem BDs, albeit with small differences (**Table 2.11**). The effect of PEG linker length was then studied, with different binding behaviour being evident between BvA-3 and BvA-7 on BRD2(1:2). BvA-7 induced dimerization on BRD2(1:2) constructs with a very high f/f_0 (~3.2) value, indicating the presence of a chain like dimer state of the protein. However, this behaviour was restricted to the BvA-n series and specifically to the compound carrying a PEG linker size of n=7, suggesting the importance of the PEG linker size selection. In the case of BvB-1, selectively binding of the one warhead towards BRD2(2) suggested to be the main reason for the formation of the most compacted BRD2(1:2) when engaging BvB-1 with 30% compacting efficiency, suggesting the importance of rational designing of the warheads. Finally, chimaera tandem BD constructs were studied and showed that alteration of protein linker region changes the bivalent

compound binding behaviour, demonstrating that protein linker regions are the vital determinant for compound binding behaviour of each BET tandem BD construct.

The study in this chapter led to several new questions concerning the binding behaviour of each BET protein in nature: How do they engage with the nucleosome via their tandem BDs? Is there also a conformation difference between BET members upon histone peptide engagement? These questions will be answered in the next chapter.

Chapter 3. Binding mode of BET tandem bromodomains on human histone H4 tail

3.1. Introduction

The binding of histone acetylated lysine residues by BDs is a crucial step in epigenetic regulation, which involves transcriptional activation/deactivation processes. Although histone acetylation sites bound by a single BD have been largely explored, the mechanisms of how tandem BDs engage on chromatin are unknown. In the last chapter's study, we observed different binding behaviour of BET tandem BDs upon bivalent compound engagement. In this chapter, in order to investigate how BET tandem BDs interact with acetylated histone peptides, SPOT array and bio-layer interferometry (BLI) were performed on BRD2(1:2), and its mutants in which either BD1 or BD2 was inactivated to compare their binding patterns with multiple acetylated histone peptides.

A SPOT array is a form of array where synthesised peptides, 10 to 20 amino acids in length, are covalently linked to a membrane. The studied protein is then screened against the array. If an interaction exists, the protein stays on the membrane bound to the peptide and can be detected by an antibody specific to the bound protein. In our case, acetylated H3 and H4 peptides of 20 amino acids in length were linked to the membrane. His-tagged BRD2 tandem BD proteins were then screened against the array and detected through HPR-conjugated anti-His antibody. This assay offers a rapid technology for screening protein peptide interactions and is usually used to scan protein sequences to determine the binding motif in studying the interacting partners[221]. The early applications for peptide arrays were in (1) identifying and characterising binding interactions between protein and peptides; (2) screening for the active substrate of enzymes; and (3) epitope mapping for antibody binding. The technology was developed in 2001[222] and has recently been applied to study human BD family interactions with histone tails[27] as well as methyl-lysine reader domain interactions with histone tails[223,224]. In the study of BD with histone tails, BDs were not only screened against single acetylated

peptides that cover all possible acetylation sites in histones but also against multiple acetylated peptides with other PTM marks to explore the influence of neighbouring PTMs on BD interactions with acetylated histone peptides. Recent progress in array technology allows peptide densities of 1,000,000 spots per square cm of solid support[225]. The advantage of using this technique is the ability for high throughput screening and the large number of variable sequence peptides that can be scanned in one membrane. The disadvantage is that even though the binding intensity can be roughly compared, the affinity cannot be determined. Therefore, this assay should only be used for the initial identification of binding partners.

BLI was used as an orthogonal validation method to the SPOT array in this chapter. BLI is an optical analytical technique that compares the interference pattern of light reflected from two surfaces. One layer is composed of immobilised protein, which in our case was biotinylated histone H4 peptide with acetyl marks on the biosensor tip, and the other layer is an internal reference. Changes in the number of molecules bound to the sensor cause shifts in the interference pattern, and therefore detects the binding intensity of the target of interest. The advantage of this method is label-free detection; however, in our experiment, only the binding signal intensity was compared, and the affinity was not measured.

However, both of these assays cannot determine the binding mode of BET tandem BDs on the acetylated peptide. For the characterisation of this, AUC was used as described in the previous chapter. Double acetylated peptides were synthesised, and polyacetylated peptides were produced in house. The conformations of BET tandem BDs were assessed and compared upon the addition of acetylated peptides by AUC. An ITC experiment was then performed on the tandem BDs with polyacetylated peptides to validate if there is any potential bivalent binding present. Overall, in this chapter, the binding of BET tandem BDs on acetylated histone peptide was explored and compared between BET members by biophysical methods.

3.2. BRD2(1:2) employs both BDs binding to Kac-XXXXXX-Kac histone H4 peptide

3.2.1. BRD2(1:2) has improved binding to distant diacetylated H4 peptide compared with that by either BRD2(1:2)$^{WT/NF}$ or BRD2(1:2)$^{NF/WT}$

To assess whether two BDs in one tandem BD protein could engage independent target sites, BRD2(1:2) mutants, BRD2(1:2)$^{NF/WT}$ and BRD2(1:2)$^{WT/NF}$, which lack the binding capacity of the first or second BD, respectively, because the conserved asparagine (N) has been mutated to phenylalanine (F) in BRD2(1:2), were produced and compared with wild-type BRD2(1:2) via SPOT array. BET BDs have been shown to preferentially bind to histone H4 compared with binding to other histones[27,58,59,226]. Accordingly, we focused on H4 interactions in the experiment, and our SPOT array covered all the possible combinations of multiple acetylated lysine sites of the human histone H4 over a 20 amino acid length. (The original SPOT array membranes are shown in **Figure 6.16 and Table 6.3**). As an orthogonal method, the BLI experiment was performed with the same multiple acetylated lysine pattern (**Figure 3.1**). (The original BLI data are shown in **Figure 6.17**).

As expected, both the wild-type and mutant BRD2(1:2) gave strong dark spots and high BLI signal binding to polyacetylated lysine peptides, corresponding with previous findings that showed greater affinity of single BD to closely located diacetylated lysine sites[27]. The study suggested that a spacer of two residues between the two acetylated lysine sites was optimal for the increased affinity of a BD to diacetylated lysine sites compared with that of a single acetylated lysine site or diacetylated lysine sites with a longer spacer. However, dark spots and high BLI signal were observed with BRD2(1:2) on distantly located diacetylated lysine sites: H4K5acK12ac and H4K5acK16ac, whereas BRD2(1:2)$^{NF/WT}$ and BRD2(1:2)$^{WT/NF}$ showed weak binding signals on these peptides (**Figure 3.1c, d**), suggesting that both wild-type BDs in BRD2(1:2) interacted with distant diacetylated peptides and therefore achieved better binding stability compared with that of the mutants.

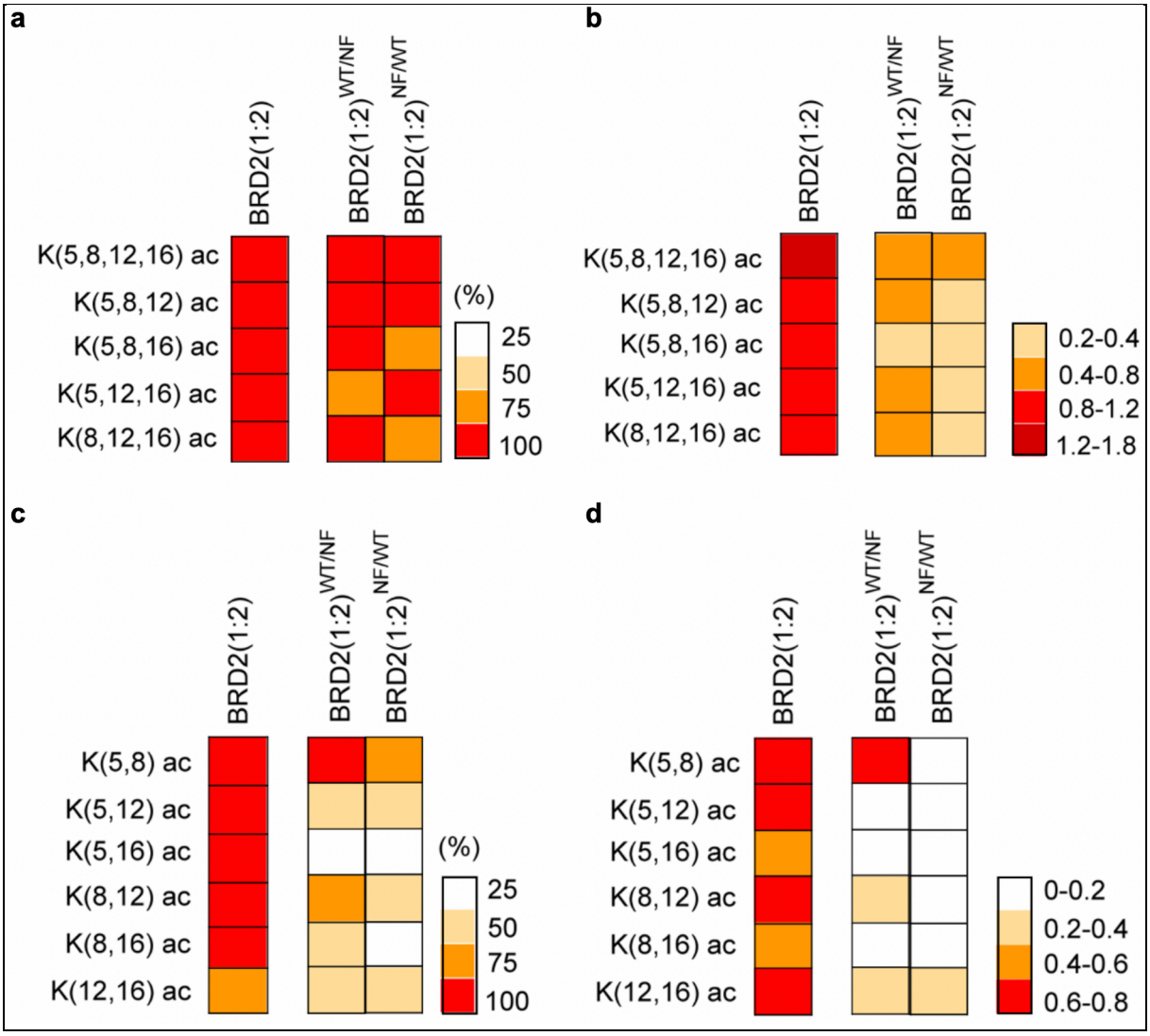


Figure 3.1. BRD2(1:2) has improved binding to distant multiple acetylated H4 peptide in comparison with that of mutated forms. (a) Multiple Kac marks bound by tandem BDs mutants in SPOT array. (b) Multiple Kac marks bound by tandem BDs mutants in BLI. (c) Double Kac marks bound by tandem BDs mutants in SPOT array. (d) Double Kac mark bound by tandem BDs mutants in BLI. Experiments were performed only once.

3.2.2. BRD2(1:2) binds to distant diacetylated H4 peptide in a 1:2 molar ratio

To address the different binding modes of BRD2(1:2) on distant diacetylated H4 peptide, we examined the conformation of BRD2(1:2) in solution in the presence of these peptides using AUC. Peptides H4K5acK12ac or H4K5acK16ac (20 amino acids in length) were custom synthesised (**Table 3.1**). Peptides were then added to the protein at a 1:1 molar ratio in the AUC experiment. If BRD2(1:2) bound to H4K5acK12ac or H4K5acK16ac in bivalent mode,

a compacted conformation of the protein would be expected. If this was not a bivalent interaction, an extended conformation of the protein would be observed (**Figure 3.2a** [A][B]).

	Sequence	Mw
H4-K5acK12ac	SGRGKacGGKGLGKacGGAKRHRY	2111.3
H4-K5acK16ac	SGRGKacGGKGLGKGGAKacRHRY	2111.3

Table 3.1. Synthesised acetylated peptide sequence and molecular weight.

Hydrodynamic analysis showed that there was no major increase in the *s* value observed with BRD2(1:2) in the presence of diacetylated peptides (**Figure 3.2b**), suggesting that no conformational change had occurred. For further validation, a $c(s, f/f_0)$ model was applied to the binding of the K(5,12)ac peptide, and the contour plot showed that the species in the solution had an f/f_0 value of around 2 (**Figure 3.2c**), which was larger than that of the apo BRD2(1:2) (f/f_0=1.89) (**Figure 2.5c**), demonstrating that an extended conformation resulted from non-bivalent binding, where two peptides bound to one BRD2(1:2) protein. To further validate the binding mode and binding stoichiometry, an ITC experiment was performed (**Figure 3.2d**) (**Table 3.2**). As expected, the ITC results showed a 1:2 molar ratio binding of BRD2(1:2) to the peptide, suggesting that the bivalent binding and the subsequent conformational change to a compacted form was not favoured by BRD2(1:2) when engaging distant diacetylated peptides.

We then hypothesised that the bivalent binding of BET tandem BDs to H4 peptide could be observed when the peptide was polyacetylated and 50 amino acids long since both of the BDs of BRD2(1:2) were employed in binding to distant diacetylated H4 peptides. Polyacetylation would increase the affinity of each BD to the peptide and with a longer peptide length allowing for greater spacing, which together may have a favourable effect on fighting the entropy penalty caused by the conformational change from the bivalent binding. However, one thing that needs to be noted is that binding of a BET protein to such a long peptide might not be physiologically

relevant if the binding specifically happens on the last two lysine residues (K31 and K44) as these two residues are partly buried in the nucleosome.

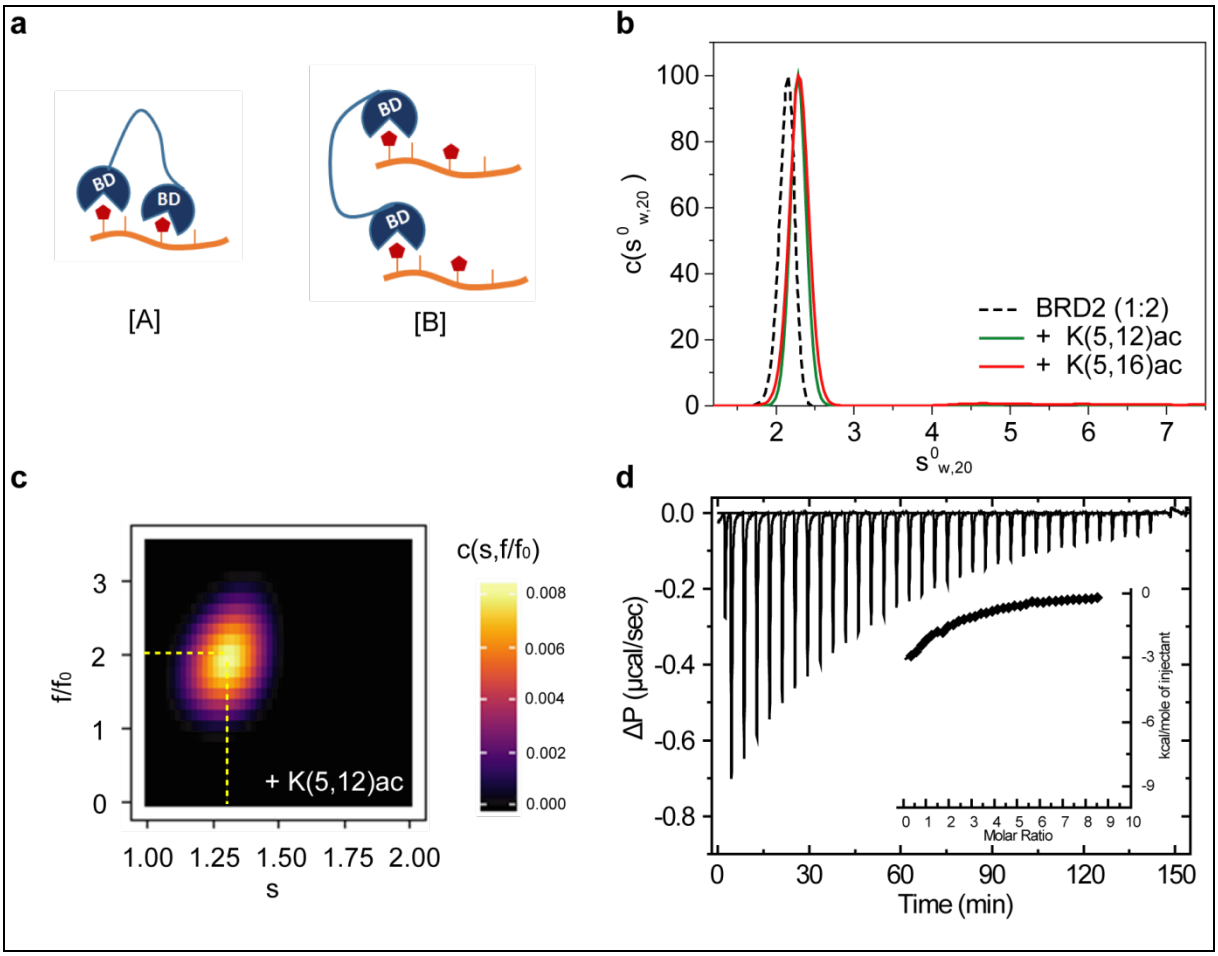


Figure 3.2. BRD2(1:2) binds to distant diacetylated H4 peptide in 1:2 protein:ligand binding mode. (**a**) Representative figure of possible engagement mode of BRD2(1:2) with double acetylated H4 peptide. (**b**) Sedimentation coefficient distribution corrected to water at 20 °C for BRD2(1:2) in the presence of diacetylated peptide K(5,12)ac and K(5,16)ac, respectively, at 1:1 protein:peptide ratio. (**c**) Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of BRD2 (1:2) in the presence of diacetylated peptide K(5,12)ac at 1:1 protein:peptide ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (**d**) Representative ITC trace of BRD2(1:2) binding to diacetylated peptide K(5,12)ac peptide. Raw injection heats for titration of peptide into a solution of BRD2(1:2) and the normalised binding enthalpies corrected for the heat of protein dilution are shown. Solid lines represent a nonlinear least square fit using a single site binding model. Titrations were performed in 50 mM HEPES pH 7.5, 100 mM NaCl, and 15 °C while stirring at 295 rpm.

Protein	*N*	K_D (μM)	Δ*H* (kcal/mol)	-*T*Δ*S* (kcal/mol)	Δ*G* (kcal/mol)
BRD2(1:2)	1.80 ± 0.16	62.7 ± 18.9	-8.11 ± 1.23	2.24 ± 0.77	- 5.88 ± 1.02

Table 3.2. **In solution evaluation of K(5,12)ac peptide binding to BRD2(1:2) by ITC.** Thermodynamic values including dissociation constant (K_D), binding molar ratio (*N*), enthalpy change (Δ*H*), and entropy change (Δ*S*), obtained from Figure 3.2 d ITC fitting. The experiment was performed in duplicate. (Shown errors correspond to the mean value ± the range).

3.3. Each BET tandem BDs engages with polyacetylated histone H4 peptide in different modes

3.3.1. BRDT(1:2) and BRD4(1:2) bind to polyacetylated H4 peptide bivalently while BRD2(1:2) binds in 1:2 protein:ligand binding mode

To study the binding mode of BET tandem BDs on polyacetylated H4 peptide, polyacetylated H4 peptide in 50 amino acids long was produced in our lab as described in section 6.2.3 (**Figure 6.6**), and the conformations of BET tandem BDs were examined in the presence of peptide at a 1:1 and 1:10 molar ratio. The polyacetylated histone H4 sequence purified in our lab was: SMSGRGKGGKGLGKGGAKRHRKVLRDNIQGITKPAIRRLARRGGVKRISY (**Table 3.3**).

	Acetylation state	Mw
H4ac	Polyacetylated	5400.4
H4	Non-acetylated	5736.4

Table 3.3. Purified histone H4 peptide molecular weight with or without acetylation.

The AUC results for BRDT(1:2) suggested that in the presence of acetylated peptide in a 1:1 ratio, peptide engagement resulted in faster sedimentation (larger sedimentation coefficient value s) when compared with that of the apo BRDT(1:2) as shown in **Figure 3.3a**, indicating the presence of compacted species. To further validate the conformational changes, the data was analysed using a $c(s, f/f_0)$ model. The contour plot showed that the species in the solution

had an f/f_0 value of 1.5, confirming the compaction of the protein resulted from bivalent binding (**Figure 3.3c**). In contrast, at a high peptide to protein ratio (at 1:10 protein:peptide ratio), the presence of a single peak in the $c(s)$ analysis, manifesting between the apo BRDT(1:2) and the 1:1 complex species, indicated the potential saturation of the protein binding sites; this was assumed to be a 1:2 protein:peptide complex, which should be in extended shape and sediments slower than the compacted protein (**Figure 3.3a**). To further validate this assumption, the $c(s, f/f_0)$ model was used for analysis. The contour plot showed that the species in the solution had an f/f_0 value over 2, demonstrating an extended conformation of the complex (**Figure 3.3d**). As a control experiment, the sedimentation coefficient value s of BRDT(1:2) in the presence of non-acylated peptide was also studied. Theoretically, tandem BDs should not bind to non-acetylated H4 peptides, and therefore there should be no change in s value or conformational change. As expected, the presence of non-acetylated peptide did not shift the peak in the sedimentation coefficient distribution plot compared with that of the apo BRDT(1:2) (**Figure 3.3b**).

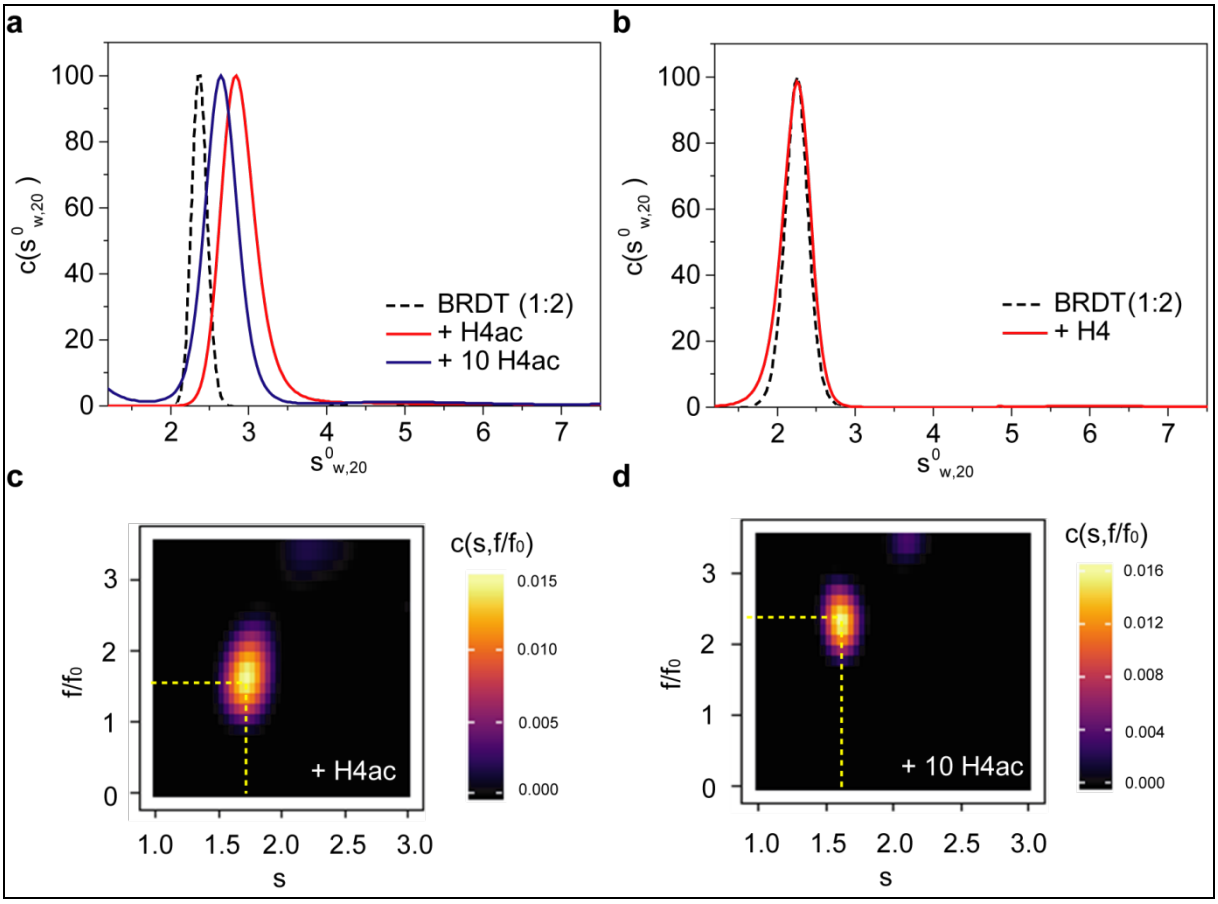


Figure 3.3. BRDT(1:2) binds to polyacetylated H4 peptide bivalently. (**a**) Sedimentation coefficient distribution corrected to water at 20 °C for BRDT (1:2) in the presence of polyacetylated peptide at 1:1 and 1:10 protein:peptide ratio. (**b**) Sedimentation coefficient distribution corrected to water at 20 °C for BRDT(1:2) in the presence of non-acetylated peptide at 1:1 protein:peptide ratio. (**c**) Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of BRDT(1:2) in the presence of polyacetylated peptide at 1:1 protein: peptide ratio (**d**) and at 1:10 molar ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C.

The same experiment and analysis were performed on BRD2(1:2) and BRD4(1:2). Interestingly, AUC analysis on BRD2(1:2) suggested that in comparison with the presence of acetylated peptide at the 1:10 ratio, the presence of acetylated peptide at a 1:1 ratio peptide engagement resulted in slower sedimentation, indicating no conformational change occurred at this ratio (**Figure 3.4a**). For further validation, the $c(s, f/f_0)$ model was applied, and the contour plot showed that the species in the solution have an f/f_0 value of around 2 for both 1:1 and 1:10 molar ratio experiments (**Figure 3.4c, e**), demonstrating an extended conformation resulted from non-bivalent binding. However, BRD4(1:2), exhibited similar results to those of BRDT(1:2) where the presence of acetylated peptide at a 1:1 ratio resulted in faster

sedimentation of BRD4(1:2) compared to that at a 1:10 ratio (**Figure 3.4b**). These results, together with the f/f_0 value of 1.5 from the contour plot in the presence of acetylated peptide in 1:1 ratio (**Figure 3.4d**) and the f/f_0 value of around 2 from the contour plot in the presence of acetylated peptide in 1:10 ratio (**Figure 3.4f**), suggest bivalent binding to the acetylated peptide. The result for BRD4(1:2) was further confirmed by a recent paper that demonstrated the bivalent binding of BRD4(1:2) to a polyacetylated peptide by the use of NanoBRET assay with HaloTag-BRD4(1:2)-NanoLuc protein[227]. In this assay, a HaloTag was added to the N-terminal side of the first BD, and a NanoLuc tag was added to the C-terminal side of the second BD in BRD4(1:2) construct (**Figure 3.5a**). Bivalent engagement of HaloTag-BRD4(1:2)-NanoLuc construct by a single ligand would bring the NanoLuc tag and fluorophore-labelled HaloTag in physical proximity, resulting in increased BRET signal. Here, the addition of histone H4-diacetylated K5ac/K8ac peptide or K5ac/K8ac/K12ac peptide did not show increased NanoBRET signal, whereas the addition of K5ac/K8ac/K20ac peptide did show increased NanoBRET signal (**Figure 3.5b**), suggesting that greater spacing and multiple acetylations in a histone H4 peptide allows for the bivalent binding of BRD4(1:2) protein.

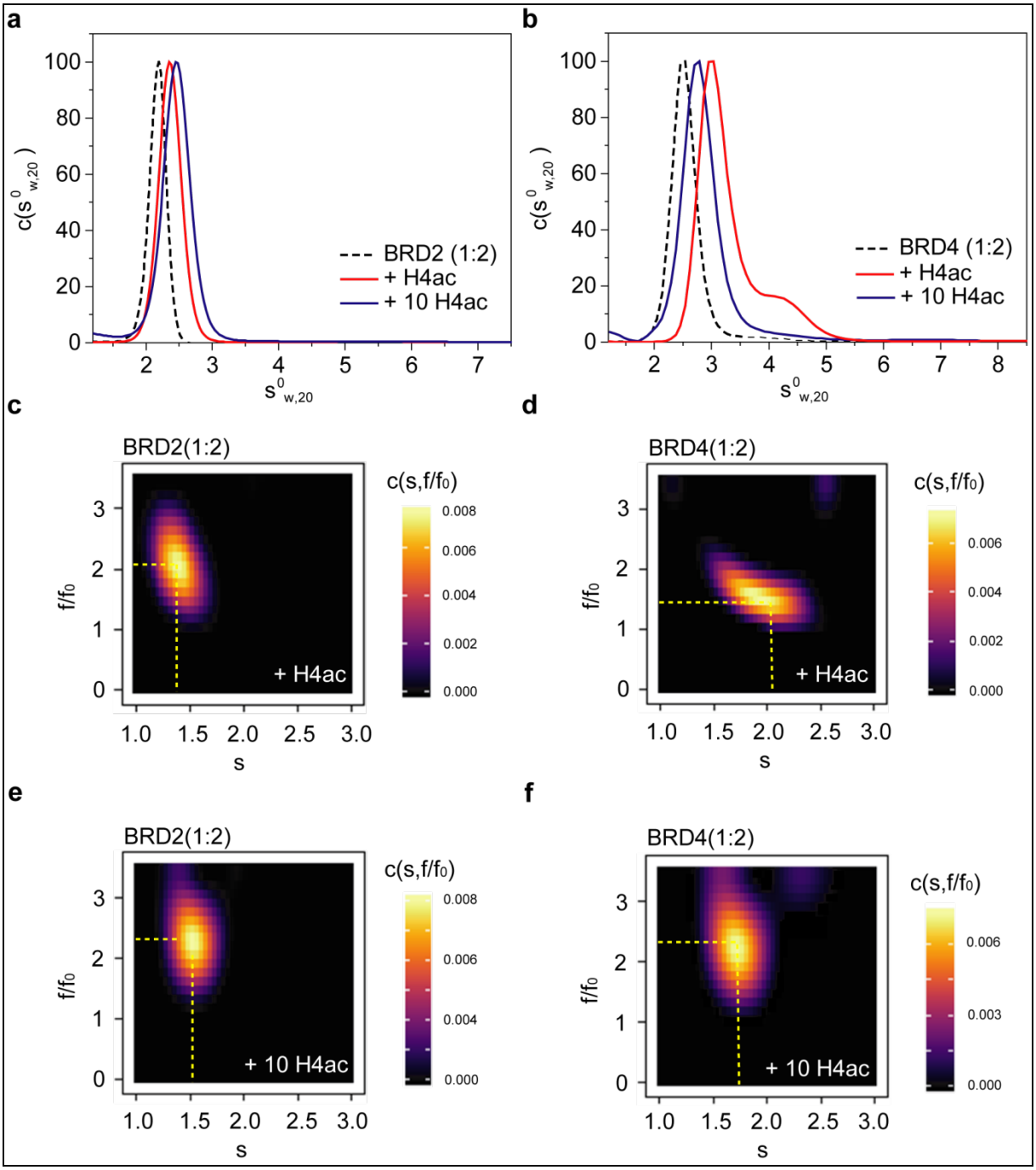


Figure 3.4. BRD4(1:2) binds to polyacetylated H4 peptide bivalently while BRD2(1:2) binds non-bivalently. Sedimentation coefficient distribution corrected to water at 20 °C for (**a**) BRD2(1:2) and (**b**) BRD4(1:2) in the presence of polyacetylated peptide at 1:1 and 1:10 protein:peptide ratio. (**c**) Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of BRD2(1:2) in the presence of polyacetylated peptide at 1:1 and (**e**) 1:10 protein:peptide ratio (**d**) Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of BRD4(1:2) in the presence of polyacetylated peptide at 1:1 and (**f**) 1:10 protein:peptide ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C.

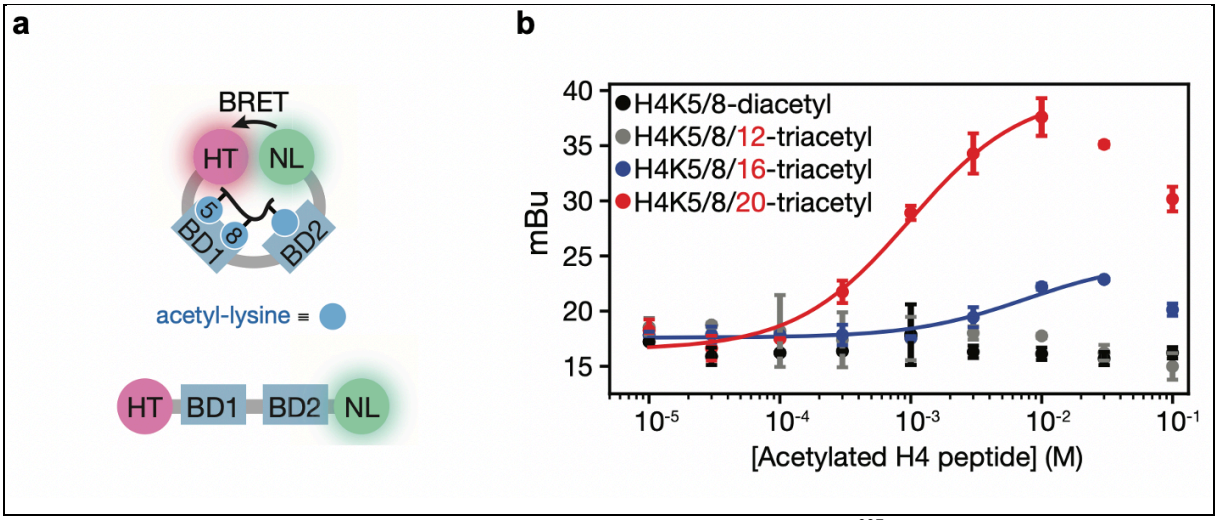


Figure 3.5. BRD4(1:2) binds to H4-K5ac/K8ac/K20ac peptide bivalently[227]. (**a**) Illustration of HaloTag-BRD4(1:2)-NanoLuc construct that bivalently bound to triacetylated peptide. (**b**) NanoBRET signal arising from HaloTag-BRD4(1:2)-NanoLuc titration with histone H4 K5ac/K8ac (black), K5ac/K8ac/K12ac (grey), K5ac/K8ac/K16ac (blue), and K5ac/K8ac/K20ac (red) peptides. Error bars represent standard deviation, n=3.

3.3.2. The BD linker region contributes to bivalent binding in polyacetylated H4 engagement

To assess the role of the inter-domain linker region of tandem BDs in peptide binding, a BRDT(1:2)^{L2} chimaera mutant with BDs from BRDT and the linker region from BRD2 were studied with the same assay. In the previous chapter, we demonstrated that the change of the linker region changes the binding behaviour of the tandem BDs to bivalent compounds. Here, we also hypothesised that switching the linker region would result in a change in the peptide binding mode. As expected, unlike BRDT(1:2) that bound to acetylated peptide bivalently, changing the linker region in BRDT(1:2)^{L2} produced similar non-bivalent binding results as those for BRD2(1:2), where the species in the solution have f/f_0 values of around 2 for both 1:1 and 1:10 molar ratios (**Figure 3.6a, c**), demonstrating an extended conformation resulted from non-bivalent binding; this suggested the change of the linker region of BRDT(1:2) to BRDT(1:2)^{L2} led to alteration of the binding behaviour from BRDT(1:2) to that found with BRD2(1:2). Another chimaera mutant, BRD2(1:2)^{LT}, however, had similar initial non-bivalent binding results as those of BRD2(1:2), where the species in the solution have the f/f_0 value of around 2 for both 1:1 and 1:10 molar ratio experiments (**Figure 3.6b, d**), and did not change

into the bivalent binding mode even with the change in the linker region, suggesting that the combined existence of tandem BDs and a linker region between BDs make each of the BETs unique. One thing that needs to be noted is that this chimaera mutant BRD2(1:2)^{LT}, showed similar binding behaviour as BRDT(1:2) when tested with the bivalent inhibitor BvA-3 (**Figure 2.14b**), suggesting the importance of the linker region in binding behaviour. However, in this case, when tested with the polyacetylated peptide, it did not alter to the similar binding behaviour as BRDT(1:2). It might be caused by the differences in affinities of BET BDs to the peptide and BvA-3 inhibitor. Bivalent inhibitor BvA-3 binds to both BRD2 BDs and BRDT BDs in nanomolar affinity but to polyacetylated peptide in micromolar affinity. If the interaction between the binding pocket and the ligand is relatively weak, the small changes in the peripheral residues or the linker region may completely change the binding behaviour. Therefore, BET BD's ability of how strongly it can bind to the ligand seems also important for the bivalent binding behaviour, suggesting the combined existence of tandem BDs and linker region between BDs make each of the BETs unique.

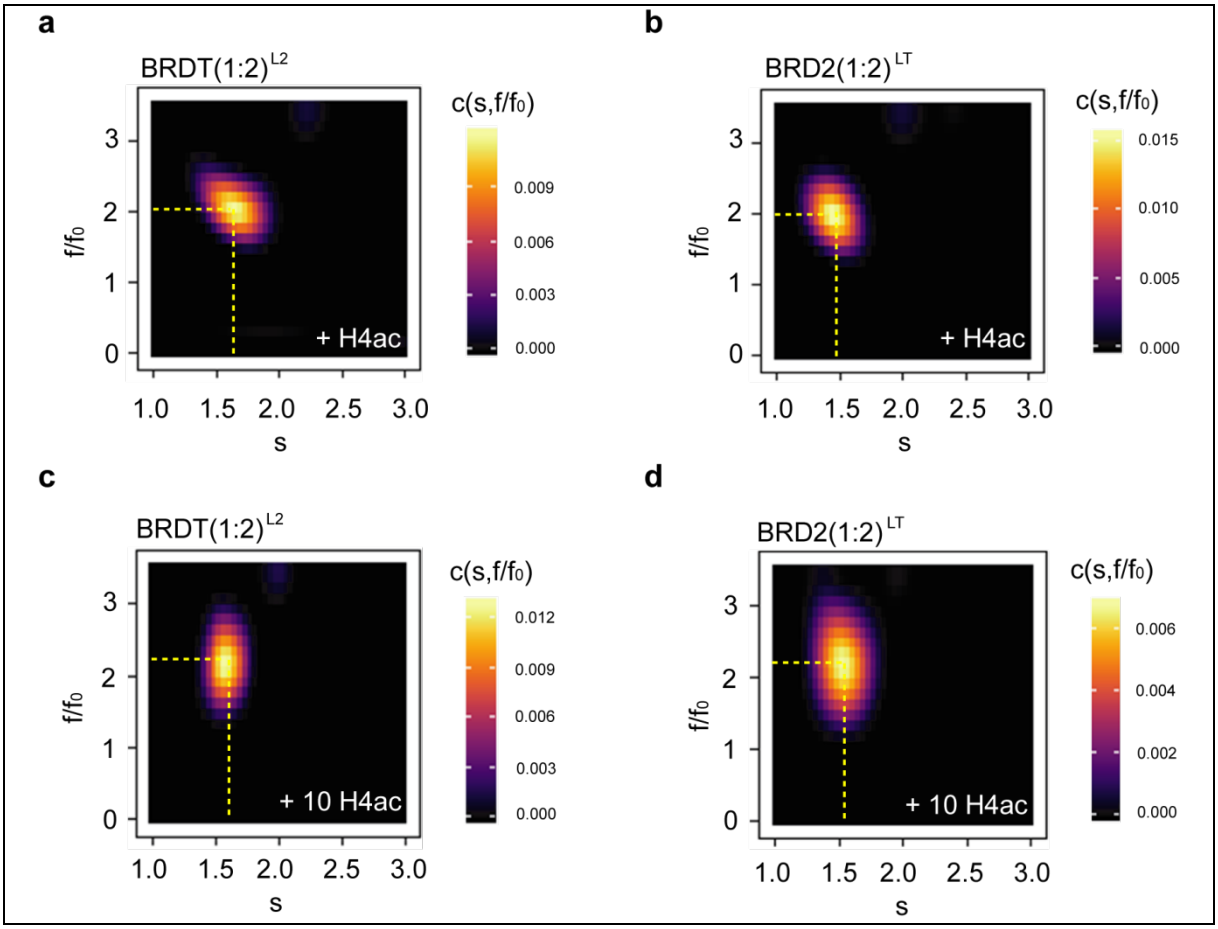


Figure 3.6. Switching the BRDT(1:2) linker region to BRD2(1:2) causes a change from bivalent to non-bivalent binding in polyacetylated H4 engagement. (**a**) Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of BRDT(1:2)^{L2} in the presence of polyacetylated peptide at 1:1 and (**c**) 1:10 protein:peptide ratio. (**b**) Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of BRD2(1:2)LT in the presence of polyacetylated peptide at 1:1 and (**d**) 1:10 protein:peptide ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C.

3.3.3. BRDT(1) and the linker region play the dominant role in the bivalent binding of BRDT(1:2) to polyacetylated H4 peptide

To shed more light into the bivalent and non-bivalent behaviour in polyacetylated H4 engagement, in solution thermodynamic analysis of BRDT(1:2), BRDT(1:2)$^{WT/NF}$, BRDT(1:2)$^{NF/WT}$, and BRDT(1:2)^{L2} interactions with polyacetylated H4 peptide were performed and compared (**Figure 3.7**). ITC experiments were performed as previously described by injecting solutions of the peptide into the protein. These results show that BRDT(1:2) bound to the acetylated peptide with a 5.4 μM affinity at a molar ratio of 1, which

was consistent with the bivalent binding AUC results of BRDT(1:2). The two mutants BRDT(1:2)$^{WT/NF}$ and BRDT(1:2)$^{NF/WT}$ bound to the peptide with affinities of 6.9 μM and 25.8 μM, respectively, at a molar ratio of 1 (**Table 3.4**). Even though BRDT(1:2) showed bivalent binding to the peptide from AUC analysis, there was no apparent increase in the affinity in the ITC experiment compared with that of the BRDT(1:2)$^{WT/NF}$ binding. However, the thermodynamic parameters suggested different binding mechanics between BRDT(1:2) and BRDT(1:2)$^{WT/NF}$. There was an increase in the entropy for the binding process of the two mutants, whereas there was a decrease in entropy for BRDT(1:2), suggesting an unfavourable conformational change to a more ordered state. This means between BRDT(1:2) and BRDT(1:2)$^{WT/NF}$ engaging to histone H4 peptide, even though there was no change in binding, the presence of second active BD in the tandem BD constructs definitely influences the binding mechanics. However, we don't know if the entropic penalty of the interaction between the BRDT(1:2) and the H4ac peptide is caused by the bivalent binding mode of the BRDT(1:2) on the peptide. To answer this, we proceed to compare the results with chimaera mutant BRDT(1:2)^{L2}. Since, unlike BRDT(1:2), BRDT(1:2)^{L2} was the non-bivalent binder to the acetylated peptide, we expected that the interaction between BRDT(1:2)^{L2} with the H4ac peptide would show an increase in the entropy. However, the results showed an even larger entropy decrease than BRDT(1:2). This discrepancy was perplexing, and we could not determine the reasons for the entropic penalty of the two interactions. Comparing all four interactions, we speculated that the presence of two active BDs in a tandem BD construct, rather than one active BD, might be the reason for the entropy penalty and that, in this case, the linker regions of the constructs with the two active BDs also participated in the interactions by forming secondary structures. The hydrogen bond formed during secondary structure formation could be the reason for the large decrease in enthalpy, and the change from a

disordered conformation to a secondary structure could account for the large decrease in enthalpy.

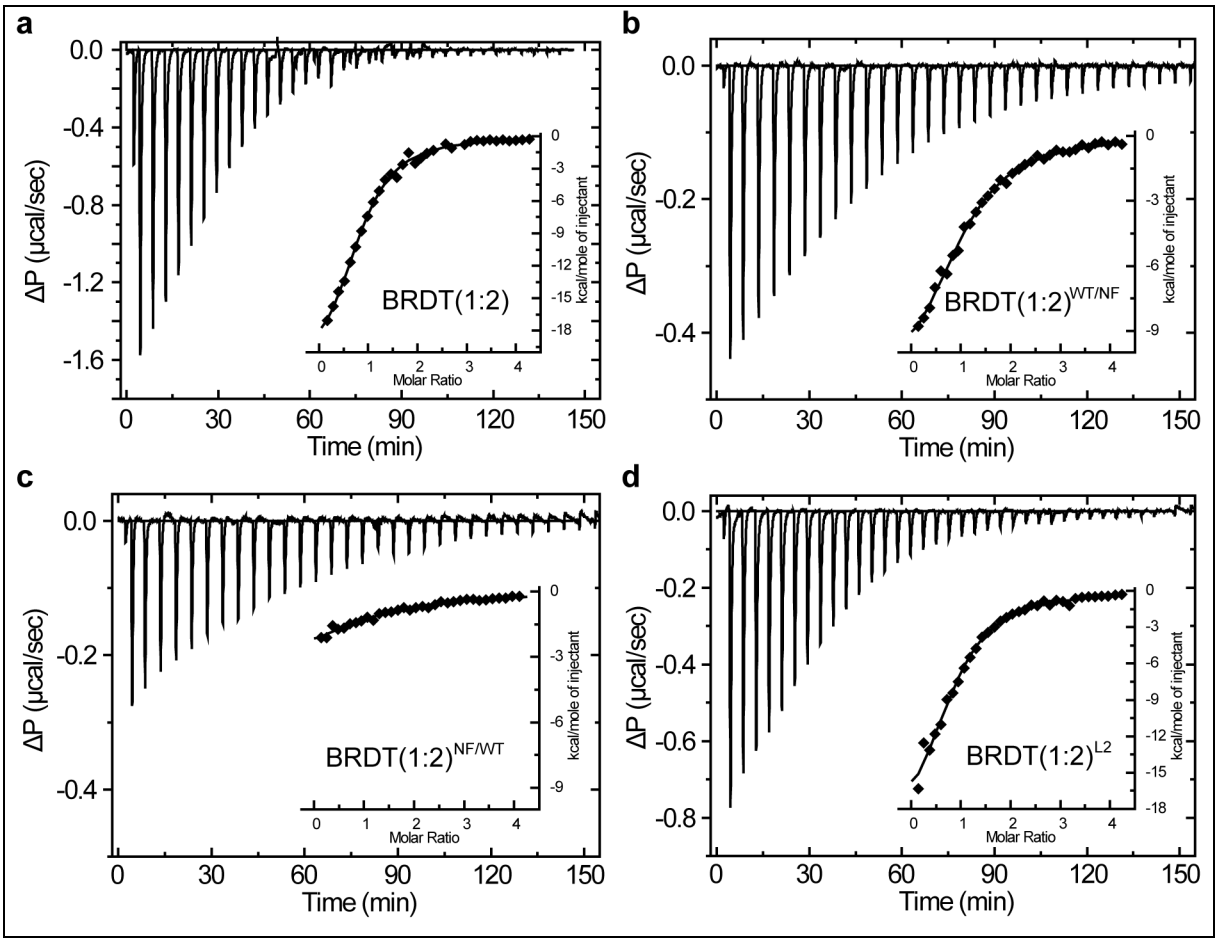


Figure 3.7. BRDT(1) and the linker region play the dominant role in the bivalent binding of BRDT(1:2) to polyacetylated H4 peptide. Representative ITC trace of polyacetylated H4 binding to (**a**) BRDT(1:2), (**b**) BRDT(1:2)$^{WT/NF}$, (**c**) BRDT(1:2)$^{NF/WT}$ and (**d**) BRDT(1:2)^{L2}. Raw injection heats for titration of peptide into a solution of BRDT(1:2) and the normalised binding enthalpies corrected for the heat of protein dilution are shown. Solid lines represent a nonlinear least square fit using a single site binding model. Titrations were performed in 50 mM HEPES pH 7.5, 100 mM NaCl, and 15 °C while stirring at 295 rpm

Protein	*N*	K_D (µM)	Δ*H* (kcal/mol)	-*T*Δ*S* (kcal/mol)	Δ*G* (kcal/mol)
BRDT(1:2)	1.07 ± 0.04	5.4 ± 0.4	- 19.1 ± 1.3	13.7 ± 4.3	- 5.4 ± 1.1
BRDT(1:2)$^{WT/NF}$	1.16 ± 0.08	6.9 ± 0.9	-6.1 ± 0.6	- 1.2 ± 0.2	- 7.3 ± 0.9
BRDT(1:2)$^{NF/WT}$	1.28 ± 0.11	25.8 ± 5.2	-3.5 ± 0.4	- 3.4 ± 0.9	- 6.9 ± 1.3
BRDT(1:2)^{L2}	0.79 ± 0.13	4.8 ± 0.3	- 27.3 ± 3.7	21.2 ± 5.7	- 6.1 ± 1.4

Table 3.4. In solution evaluation of polyacetylated H4 peptide binding to BRDT(1:2) by ITC. Thermodynamic values including dissociation constant (K_D), binding molar ratio (*N*), enthalpy change (Δ*H*), and entropy change (Δ*S*), obtained from Figure 3.7 ITC fitting. Experiments were performed in duplicate. (Shown errors correspond to the mean value ± the range).

3.4. Discussion

Despite their high degree of sequence similarity as well as sharing the same domain architecture, BET proteins adopt specific cellular functions, contributing to complicated transcriptional programmes. BET proteins are known to interact with acetylated histone tails via their BDs, and previous studies have hypothesised several possible models for how BET BDs interact with acetylated peptides[226] based on the affinity of each single BD towards acetylated histone tails. However, we do not yet understand or have experimentally proven how each BET tandem BD engages with chromatin, nor do we know what specific structural determinants underpin their functional role(s).

In this chapter, we studied the interaction of the BET tandem BDs with acetylated peptides and the different supertertiary structure changes of each BET tandem BD in the presence of acetylated histone peptide. SPOT array and BLI experiments on BRD2(1:2), BRD2(1:2)$^{WT/NF}$ and BRD2(1:2)$^{NF/WT}$ showed that BRD2(1:2) employs both of the BDs binding to H4K5acK12ac and H4K5acK16ac. The hydrodynamic and thermodynamic analysis further revealed that binding of BRD2(1:2) to these distant diacylated peptides was in 1:2 protein:ligand binding mode (**Figure 3.8**).

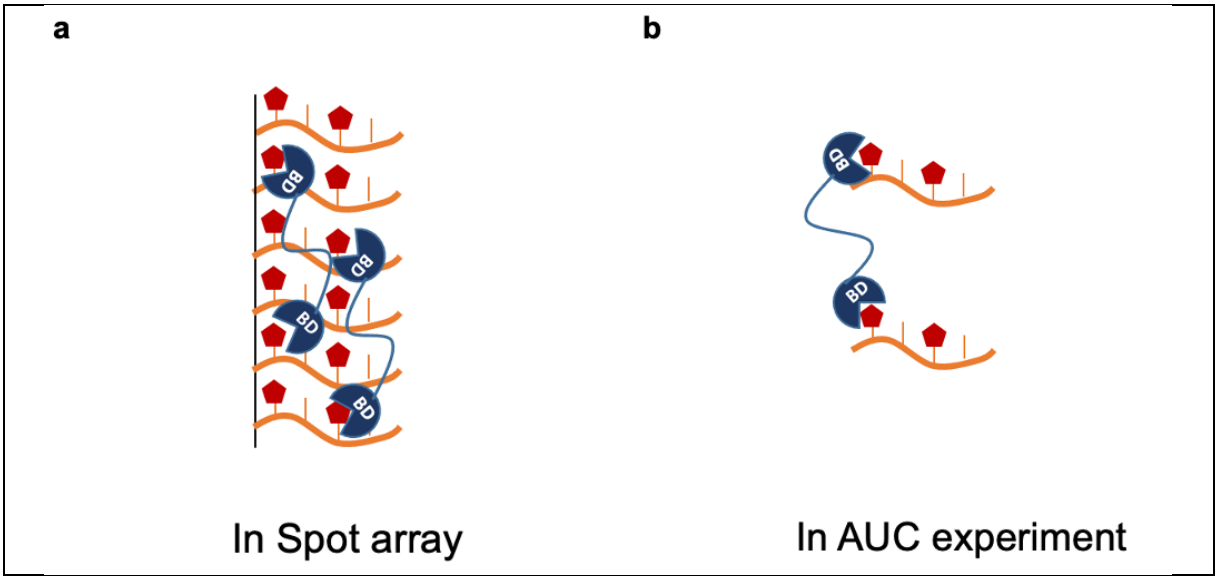


Figure 3.8. Representative figure of engagement mode of BRD2(1:2) with distant diacetylated histone H4 peptide in SPOT array and AUC experiment. **(a)**1:2 protein:ligand binding mode in SPOT array. **(b)**1:2 protein:ligand binding mode in AUC experiments.

The binding modes of BET tandem BDs on polyacetylated peptides of 50 amino acids in length were then studied and compared (**Table 3.5**) (**Supplementary figure 2**). The Hydrodynamic analysis demonstrated that BRDT(1:2) and BRD4(1:2) bound to polyacetylated H4 peptide bivalently while BRD2(1:2) bound non-bivalently. Similar to work in the previous chapter, linker chimaera mutants were also analysed and showed that switching the BRDT(1:2) linker region to that of BRD2(1:2) causes a change from bivalent to non-bivalent binding in polyacetylated H4 engagement. To further explore the basis of the observed bivalent and non-bivalent behaviour, thermodynamic analysis was performed on BRDT(1:2) and derived mutants. These results revealed that bivalent binding of BRDT(1:2) to polyacetylated H4 did not result in tighter affinity compared with that of BRDT(1:2)$^{WT/NF}$ but did show altered thermodynamic parameters resulting from different binding mechanics, suggesting the dominant role of BRDT(1) in the binding specificity and the importance of the BRDT(1:2) linker region in the bivalent binding of BRDT(1:2) to polyacetylated H4 peptide.

Protein	Peptides	AUC				
		f/f_0	s	s^0_{20w}	MW(Da)	rmsd(OD)
BRDT(1:2)	-	1.80	1.42 ± 0.08	2.37	42682	0.0070
	+ H4ac	1.50	1.77 ± 0.12	2.95	45241	0.0092
	+ 10H4ac	2.25	1.62 ± 0.13	2.71	51138	0.0088
BRD2(1:2)	-	1.89	1.31 ± 0.07	2.19	43540	0.0076
	+ H4 K(5,12)ac	1.93	1.40 ± 0.06	2.34	46301	0.0102
	+ H4ac	2.10	1.44 ± 0.10	2.42	47106	0.0109
	+ 10H4ac	2.34	1.51 ± 0.13	2.52	55776	0.0093
BRD4(1:2)	-	1.96	1.38 ± 0.07	2.32	46892	0.0067
	+ H4ac	1.45	1.86 ± 0.11	3.11	50510	0.0104
	+ 10H4ac	2.31	1.72 ± 0.13	2.83	57965	0.0068
BRDT(1:2)^{L2}	-	1.92	1.38 ± 0.09	2.30	45722	0.0098
	+ H4ac	2.04	1.58 ± 0.13	2.64	52163	0.0100
	+ 10H4ac	2.21	1.62 ± 0.12	2.75	58011	0.0082
BRD2(1:2)^{LT}	-	1.86	1.41 ± 0.07	2.36	42065	0.0086
	+ H4ac	1.98	1.49 ± 0.10	2.40	47022	0.0077
	+ 10H4ac	2.18	1.52 ± 0.11	2.53	50012	0.0101

Table 3.5. In solution evaluation of acetylated peptides binding to BET tandem BDs by AUC. Hydrodynamic values including translational frictional ratio (f/f_0), apparent sedimentation coefficient (s), corrected sedimentation coefficient ($s_{20,w}$), and system mass (Mw) values of each BET tandem BDs in the presence of peptides. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 100 mM NaCl at 4 °C. (Shown errors on s values represent fitted s distribution value peak area, rmsd is the root-mean-square deviation of detected protein boundary from the boundary generated by Lamm equation modelling)

The results described in this chapter encouraged us to consider that the different binding behaviour of BET tandem BDs on chromatin may be the natural way of achieving specificity and differences in functional roles. These results also inspired us to develop a novel multivalent peptidic inhibitor that could bivalently or multivalently capture one of the BETs, using both the BD region and ET domain, and therefore achieve specificity for BET binding. Is it possible to make this type of peptidic inhibitor? How long should the linker region be? What is the evaluation process? These questions will be answered in the next chapter.

Chapter 4. Characterisation of selective BET-BD2ET bivalent peptidic inhibitors

4.1. Introduction

In nature, multivalent recognition between multidomain proteins is commonly exploited, where many weak interactions between each domain lead to exceptionally potent apparent binding of two proteins. This 'avidity' effect is greater than the sum of the affinities contributed from each single interaction in isolation and therefore achieves higher specificity towards that particular protein. In the last chapter's study, we observed different binding behaviours of BET tandem BDs upon polyacetylated histone H4 peptide engagement where BRDT(1:2) showed bivalent binding to the acetylated peptide while BRD2(1:2) did not. This result inspired us to develop a novel multivalent peptidic inhibitor that can capture BETs both from the BD region and ET domain region to achieve selectivity towards one of the four BET proteins. In the rational designing process of this type of multivalent peptide, we chose acetylated histone H4 peptide that can bind to BDs and a potent viral peptide sequence that has been previously reported[74,75] to bind to ET domain with nanomolar affinity as two of the essential binding components (**Table 4.1**). As for the peptidic inhibitor linker region, we used a flexible peptide sequence of different lengths based on the experiences of BvX-n bivalent inhibitors that we studied in Chapter 2. The histone H4 and the viral peptides were then fused with different flexible linker lengths, cloned into constructs that contained His- and Venus-tags, and then expressed in bacteria and purified. Venus-tag (a variant of yellow fluorescent protein) was used to improve the stability of the peptides during bacterial growth and for the convenience of the purification process. Without the Venus tag, the expressed peptides tend to disappear during the first step of purification IMAC.

	Sequence
Histone H4	SMSGRGKGGKGLGKGGAKRHRKVL
Viral	TWRVQRSQNPLKIRLTR
Aim	SMSGRGKGGKGLGKGGAKRHRKVL + (linker) + TWRVQRSQNPLRIRLTR

Table 4.1. Sequence components of proposed bivalent peptide.

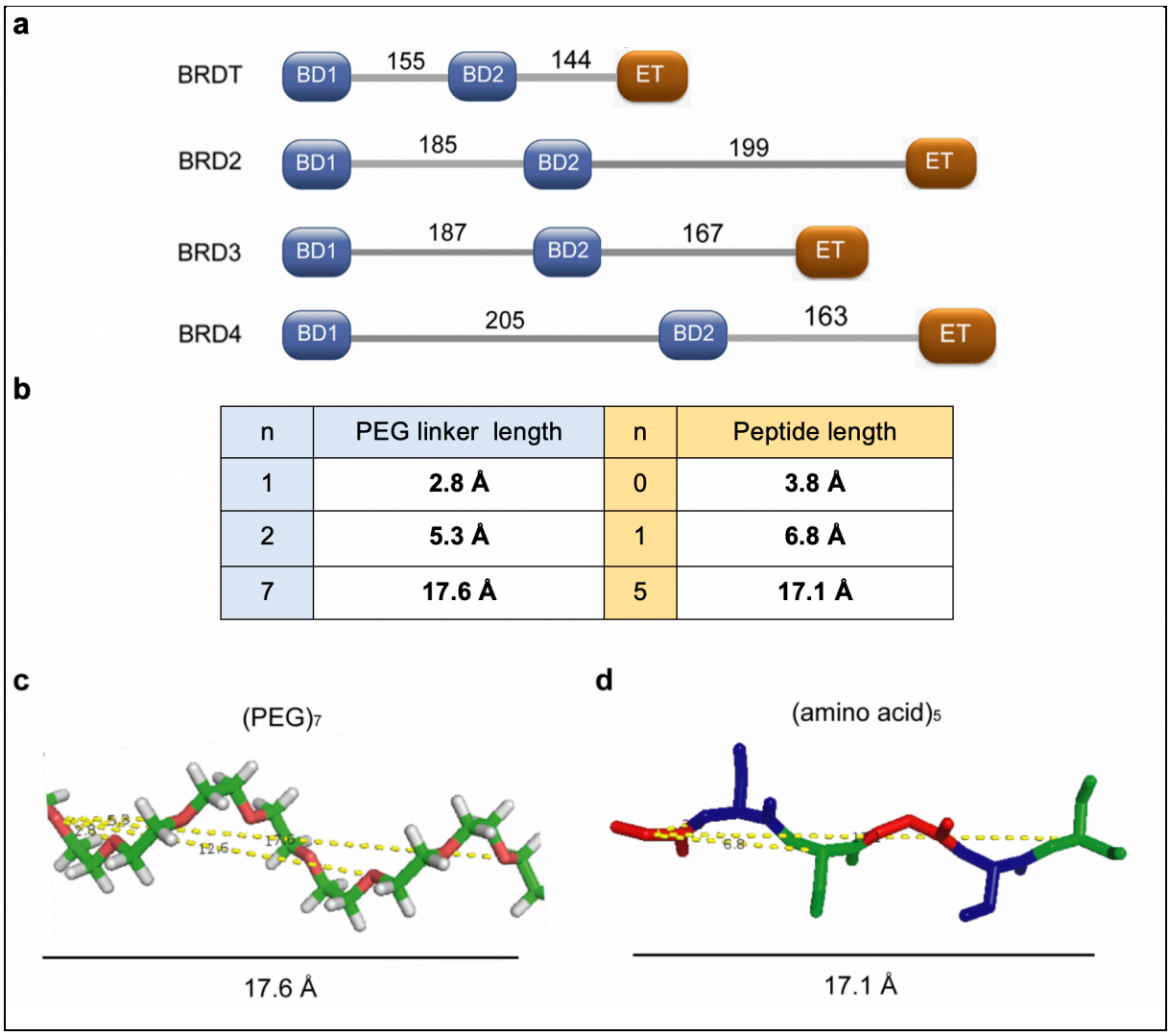
BET protein truncations that span the second bromodomain BD2 and ET domains, BET-BD2ETs, were then cloned and purified for characterisation of binding mode engaging with the various bivalent peptides that we had designed. Similar to the previous chapters, all the conformational changes caused by the different binding modes were assessed with AUC. An ITC experiment was then performed to confirm or validate the avidity effect of multivalent binding. Overall, in this chapter, various multivalent peptidic inhibitors were produced and biophysically characterised upon binding to BET-BD2ETs, in order to achieve selectivity towards one of the BET proteins.

4.2. Discovery of bivalent peptidic inhibitor that simultaneously tethers BD2 and ET

4.2.1. Bivalent peptidic inhibitor designing and production

The three basic components of bivalent peptidic inhibitors were an acetylated histone H4 peptide part for binding to the BD of BETs, a viral peptide part for binding to ET of BETs, and the linker region in between. But how long should the linker regions be? To answer this question, we looked back to the tandem BDs and the PEG size of BvX-n compounds that successfully compacted them. In Chapter 2, we demonstrated that BvB-n compounds from n equal to 1 to 7 could all tether both BDs simultaneously and cause compaction of the tandem BD proteins. This means that a bivalent probe linker length of 2.8 to 17.6 Å is sufficient for tethering the two domains of lengths of 150 to 200 amino acids simultaneously. The length between BD2 and ET domains of BETs varied from 144 to 199 amino acids, which is in a

similar range as the length between tandem BDs of BETs. By comparison, we found it reasonable to use a bivalent peptidic inhibitor linker length from around 2 to 30 Å, which is equal to a peptide of up to 12 amino acids in length (**Figure 4.1**).



n	PEG linker length	n	Peptide length
1	**2.8 Å**	0	**3.8 Å**
2	**5.3 Å**	1	**6.8 Å**
7	**17.6 Å**	5	**17.1 Å**

Figure 4.1. Rational designing of bivalent peptidic inhibitor linker length. (**a**) Domain organisation of BET proteins, domain distances between tandem BDs and BD2, ET for all family members are annotated in the figure. (**b**) Comparison between the PEG linker length and peptide length. (**c**) Length of a PEG7 linker and length of a short peptide linker with five amino acids are illustrated.

Accordingly, five different lengths (0 to 4) of the linker region were designed in three amino acid length increments. We then annotated our bivalent peptidic inhibitors as $H4_{20}Vn$, where $H4_{20}$ is the human histone H4 peptide with 20 residues, including the K5, K8, K12, K16 and K20 residues that were necessary for BD bindings; V is the viral peptide for the binding of ET,

and n (n equals 0 to 4) is the length of linker region that varied from 0 to 12 amino acids in length in 3 amino acid increments. To avoid affecting the binding capacity of the final product of the viral peptide region to ET domain after acetylation, the lysine (K) residue of the viral peptide was mutated to arginine (R) in the first step via PCR. For the constitution of the flexible linker region, glycine (G) and serine (S) residues were used, as these are most frequently used according to published research[228]. To add the diversity for the linker region, one or two flexible amino acids, asparagine (N) and aspartic acid (D), were also used. The designed linker regions were DGS, GSSDGS, NSSGSSDGS, and NSSNSSGSSDGS (**Table 4.2**).

Peptide	sequence	linker length	MW (Da)	ac-MW (Da)
$H4_{20}V0$	SMSGRGKGGKGLGKGGAKRHRKVL + TWRVQRSQNPLRIRLTR	0 aa	4748	5000
$H4_{20}V1$	SMSGRGKGGKGLGKGGAKRHRKVL + (DGS) + TWRVQRSQNPLRIRLTR	3 aa	5007	5259
$H4_{20}V2$	SMSGRGKGGKGLGKGGAKRHRKVL + (GSS)(DGS) + TWRVQRSQNPLRIRLTR	6 aa	5239	5491
$H4_{20}V3$	SMSGRGKGGKGLGKGGAKRHRKVL + (NSS)(GSS)(DGS) + TWRVQRSQNPLRIRLTR	9 aa	5364	5616
$H4_{20}V4$	SMSGRGKGGKGLGKGGAKRHRKVL + (NSS)(NSS)(GSS)(DGS) + TWRVQRSQNPLRIRLTR	12 aa	5652	5904

Table 4.2. Sequence of the bivalent peptidic inhibitor series $H4_{20}Vn$. Bivalent peptidic inhibitor comprised of histone H4 sequence region including K5, K8, K12, K16, and K20 which when acetylated can be recognised by BD, linked by GSS linker region with different lengths of up to 12 amino acids (aa), followed by the viral sequence which can be recognised by ET.

After successful cloning of designed bivalent peptidic inhibitors, expression in bacteria cells, purification and polyacetylation was performed (**Figure 4.2**). The process of purification and acetylation was quite straightforward, as described in section 6.2.3. Notably, mass spectrometry confirmation prior to TEV cleavage indicated a mass of 34,358.22 Da, which was

around 18 Da lower than the calculated expected mass of 34,376.6 Da (**Figure 4.2b**). This was expected as in the maturation of the yellow fluorescent protein, one water molecule needs to be lost to form the mature chromophore. Other than this, no precipitation was observed after cleavage of either His- or Venus-tags, which was confirmed by mass spectrometry (**Figure 4.2c**), suggesting there was no degradation as well. Each acetyl group adds a mass of 42 Da to the total peptide mass, resulting in an increase of 42 × 6 = 252 Da to the actual peptide mass after acetylation (**Figure 4.2d**). The other five peptides were all successfully produced and acetylated, as confirmed by mass spectrometry (**Table 4.2**). All the purification processes for these peptides and their mass spectrometry confirmation are listed in Chapter 6: Materials and Methods (**Figure 6.7-6.15**).

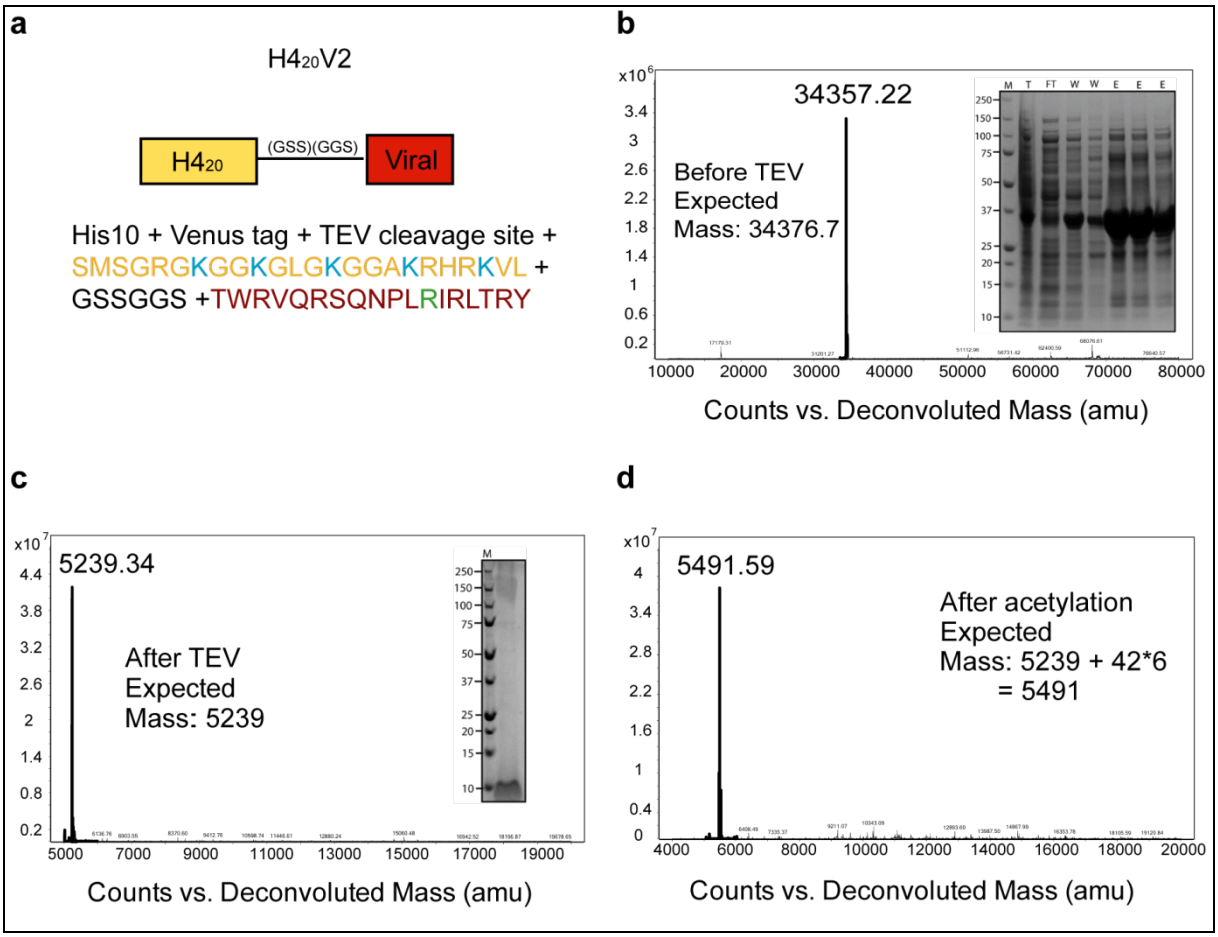


Figure 4.2. Typical bivalent peptidic inhibitor purification and acetylation process (here shown for H4$_{20}$V2). (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{20}$V2 comprised of histone H4 sequence region recognised by BD, linked by GSSGGS linker region, followed by the viral sequence recognised by ET. **(b)** SDS-PAGE analysis of H4$_{20}$V2 after affinity chromatography purification. (NuPAGE^{TM} NovexTM 4-12% Bis-Tris

protein gel) and mass spectrometry confirmation. (**c**) SDS-PAGE analysis of $H4_{20}V2$ after TEV cleavage and mass spectrometry confirmation. (**d**) Mass spectrometry confirmation of acetylated peptide.

4.2.2. Bivalent peptidic inhibitor with a 9 or 12 amino acid linker length leads to BRD4-BD2ET compaction

To study the binding mode of BRD4-BD2ET with bivalent peptides, sedimentation velocity experiments were conducted with AUC. The bivalent peptidic inhibitors were individually added to the protein solution at a 1:1 molar ratio. Hydrodynamic analysis of $c(s, f/f_0)$ showed that apo BRD4-BD2ET has an apparent sedimentation coefficient value s of 1.5 and a translational frictional ratio value f/f_0 of 2.1, suggesting that this was a monomeric elongated species (**Figure 4.3a**). The addition of $H4_{20}V0$_ac peptide at a 1:1 ratio did not significantly affect the sedimentation coefficient value s nor translational frictional ratio value f/f_0, indicating that this was a monomeric and elongated complex and suggesting the failure of simultaneous tethering of the BD2ET two domains (**Figure 4.3b**). However, with the addition of $H4_{20}V1$_ac peptide at a 1:1 ratio, there was a gradual change in both of the s and f/f_0 values of the complex as the bivalent peptide linker length increased. (**Figure 4.3c and d**). The sedimentation coefficient value s was shifted to the right, and the f/f_0 value shifted to below 2, suggesting the formation of faster moving and compacted species. Following the addition of $H4_{20}V3$_ac peptide, the changes in s and f/f_0 value became much more apparent, and the sedimentation coefficient value s shifted from 1.5 to 1.9 while the translational frictional ratio value f/f_0 lowered from 2.1 to 1.6, indicating the formation of a more compacted protein-peptide complex (**Figure 4.3e**). Similarly, a more defined and compacted species also appeared in the presence of $H4_{20}V4$_ac peptide at a 1:1 molar ratio, with the sedimentation coefficient value s of 1.9 and translational frictional ratio value f/f_0 of 1.5 (**Figure 4.3f**).

Overall, from the hydrodynamic study, only the addition of peptide $H4_{20}V3$_ac and $H4_{20}V4$_ac, the peptides with linker lengths of 9 and 12 amino acids, respectively, led to compaction of

BRD4-BD2ET protein that resulted from successful simultaneous tethering of BD2ET two domains bivalently, whereas all the other peptides with the same binding motifs but shorter linker lengths failed to do so.

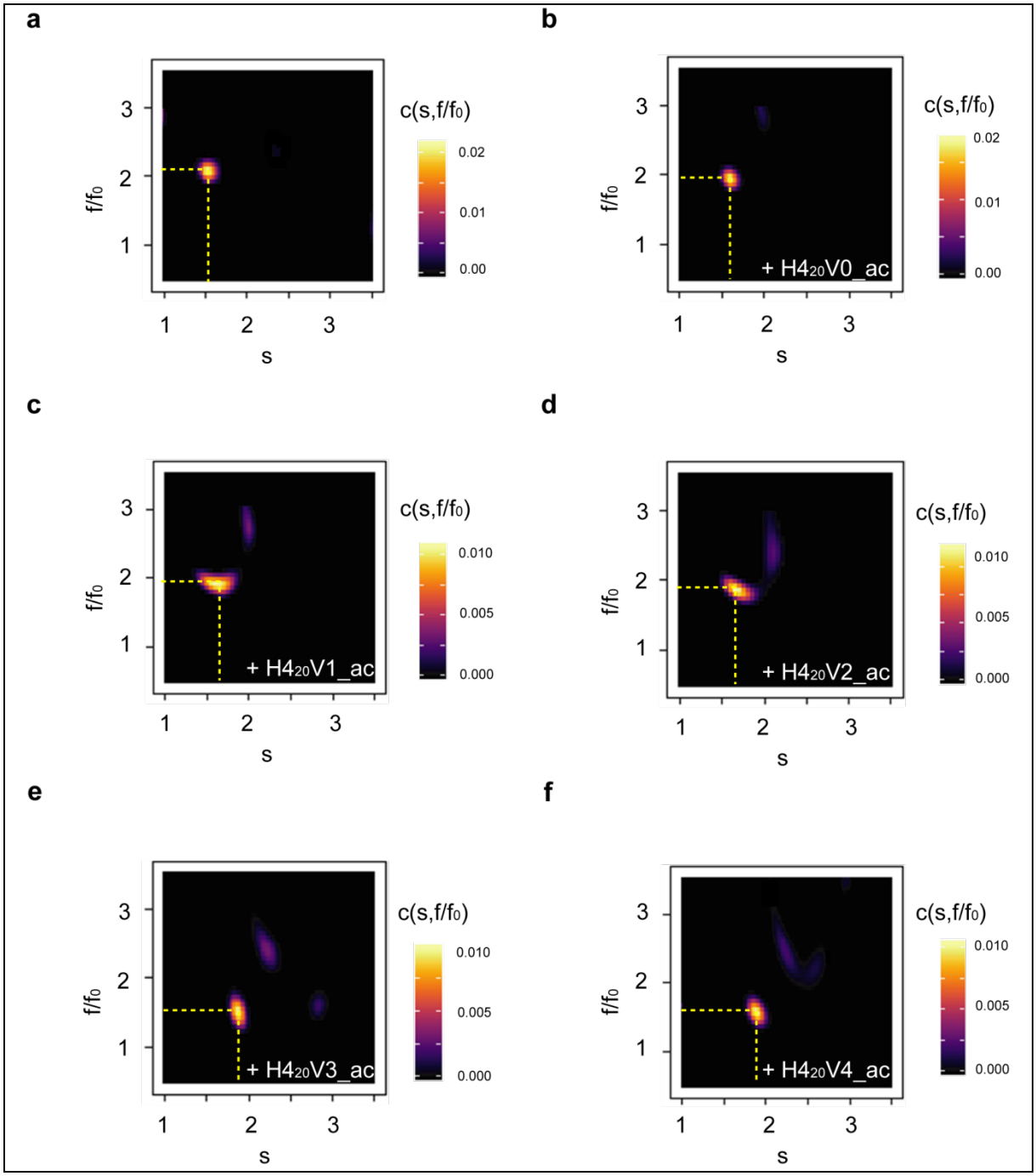


Figure 4.3. Bivalent peptidic inhibitors H4$_{20}$V3_ac and H4$_{20}$V4_ac lead to BRD4-BD2ET compaction. Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of (**a**) apo-BRD4-BD2ET and (**b**)(**c**)(**d**)(**e**)(**f**) in the presence of indicated polyacetylated peptide series H4$_{20}$Vn at 1:1 protein:peptide ratio. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 300 mM NaCl at 4 °C.

4.2.3. Bivalent peptidic inhibitor $H4_{10}Vn$ series which only contains K5, K8 cannot compact BRD4-BD2ET

The bivalent peptidic inhibitor $H4_{20}V4_ac$ that contains K5ac, K8ac, K12ac, K16ac and K20ac, showed simultaneous tethering of BRD4-BD2ET two domains; therefore, we asked which part of the histone H4 sequence was critical in the BD2 binding of tandem BD2ET protein to the bivalent peptide $H4_{20}V4_ac$. To explore this question, a shorter version of the bivalent peptide was designed and produced similarly to the production of $H4_{20}Vn_ac$ peptides. This set of new shorter $H4_{10}Vn$ peptides contained the same set of linker regions and the viral region as $H4_{20}Vn$ peptides but was fused with a shorter histone H4 sequence that only included K5ac and K8ac (**Table 4.3**).

Peptide	sequence	linker length	MW (Da)	ac-MW (Da)
$H4_{10}V0$	SMSGRGKGGKGLG + TWRVQRSQNPLRIRLTR	0 aa	3517	3643
$H4_{10}V1$	SMSGRGKGGKGLG + (DGS) + TWRVQRSQNPLRIRLTR	3 aa	3776	3902
$H4_{10}V2$	SMSGRGKGGKGLG + (GSS)(DGS) + TWRVQRSQNPLRIRLTR	6 aa	3844	3970
$H4_{10}V3$	SMSGRGKGGKGLG + (NSS)(GSS)(DGS) + TWRVQRSQNPLRIRLTR	9 aa	4132	4259
$H4_{10}V4$	SMSGRGKGGKGLG + (NSS)(NSS)(GSS)(DGS) + TWRVQRSQNPLRIRLTR	12 aa	4420	4546

Table 4.3. Sequence of the bivalent peptidic inhibitor series $H4_{10}Vn$. Bivalent peptidic inhibitor comprises of histone H4 sequence region including K5 and K8, which when acetylated can be recognised by BD, linked by GSS linker region with different length up to 12 amino acids, followed by the viral sequence recognised by ET.

Similar to the $H4_{20}Vn$ peptides binding mode study, AUC sedimentation velocity was conducted, and all the bivalent peptidic inhibitors were individually added to the protein solution at a 1:1 molar ratio. However, this did not affect the sedimentation coefficient value *s* nor translational frictional ratio value f/f_0, indicating a monomeric and elongated complex was still formed and suggesting the failure in the simultaneous tethering of BD2ET two domains (**Figure 4.4**). The gradual change in both of the *s* and f/f_0 values along with the bivalent peptide linker length change that happened with $H4_{20}Vn_ac$ was also not observed with the set of $H4_{10}Vn_ac$ peptide binding studies.

This finding suggested that either BD2 of BRD4-BD2ET tandem protein either preferentially bound to K12ac, K16ac and K20 ac, rather than binding to K5ac and K8ac or that the $H4_{10}Vn_ac$ peptides, even with $H4_{10}V4_ac$ peptide (**Figure 4.4f**), which has the longest linker region, was too short to tether the BD2ET two domains of BRD4 simultaneously and compact the protein. In either case, we learned that $H4_{20}V4_ac$ was the most optimal bivalent peptidic inhibitor with the most optimal linker length and histone H4 sequence to achieve BRD4-BD2ET protein compaction or bivalent engagement.

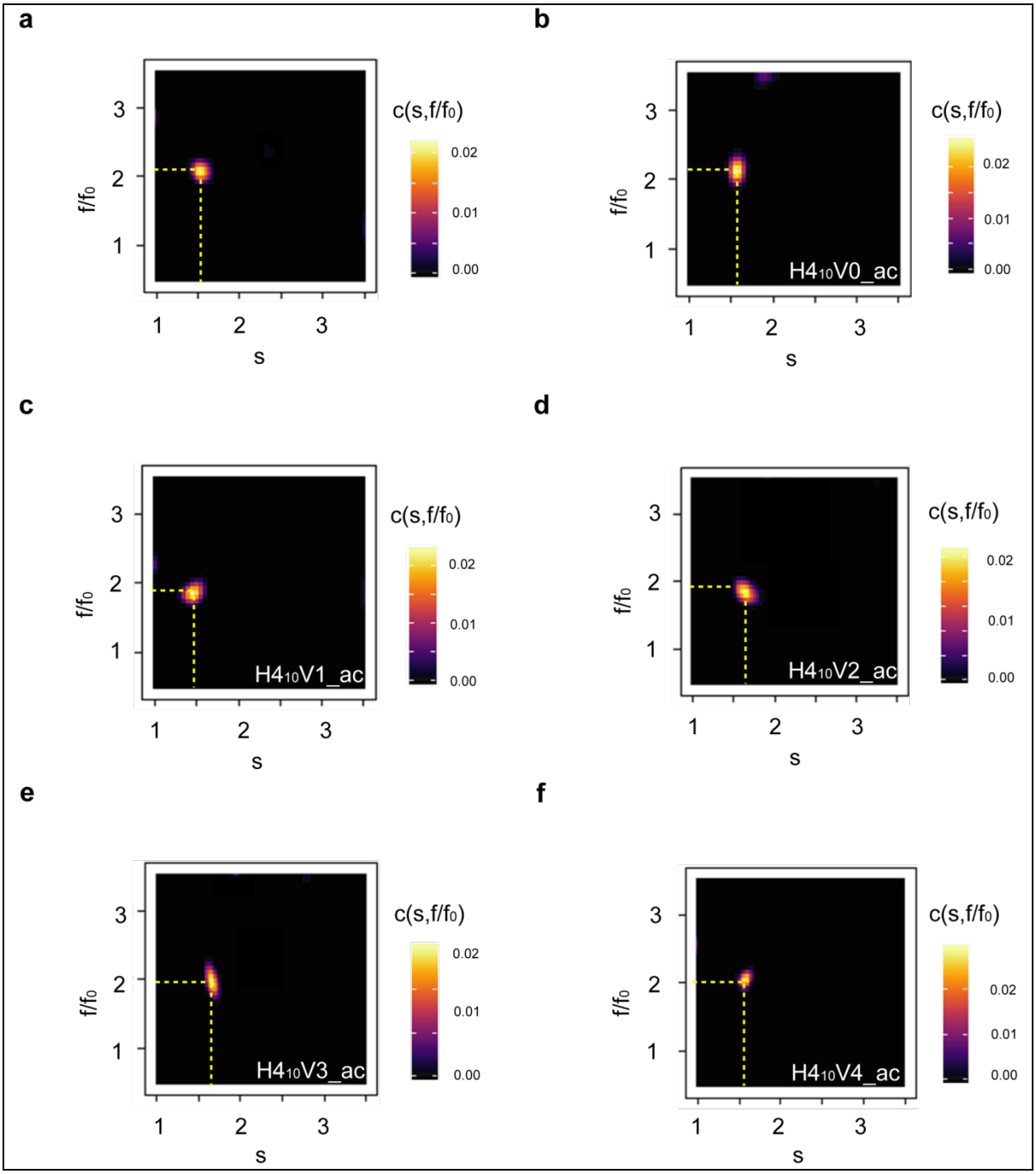


Figure 4.4. Bivalent peptidic inhibitor H4$_{10}$Vn_ac series cannot compact BRD4-BD2ET. Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of (**a**) apo BRD4-BD2ET and (**b**)(**c**)(**d**)(**e**)(**f**) in the presence of indicated polyacetylated peptide series H4$_{10}$Vn_ac at 1:1 protein:peptide ratio. Obtained from 280 nm absorbance sedimentation velocity experiments conducted in 50 mM HEPES, pH 7.5, 300 mM NaCl at 4 °C.

4.3. Bivalent peptidic inhibitor $H4_{20}V4_ac$ is selective towards BRD4-BD2ET relative to BRD2-BD2ET

4.3.1. $H4_{20}V4_ac$ compacts BRD4-BD2ET but did not compact BRD2-BD2ET

To further validate the bivalent binding mode of $H4_{20}V4_ac$ peptide to BRD4BD2ET, the mutated viral sequence regions of the $H4_{20}V4$ peptide were cloned and produced in a similar way as the other bivalent peptides. In this mutant peptide, named $H4_{20}V4LIL$, the crucial three residues leucine (L), isoleucine (I), and leucine (L), in recognition of ET domain binding motif were all mutated to alanine (A) to abolish the ET domain binding ability (**Table 4.4**).

Peptide	sequence	linker length	MW (Da)	ac-MW (Da)
$H4_{20}V4$	SMSGRGKGGKGLGKGGAKRHRKVL + (NSS)(NSS)(GSS)(DGS) + TWRVQRSQNPLRIRLTR	12 aa	5652	5904
$H4_{20}V4$ LIL	SMSGRGKGGKGLGKGGAKRHRKVL + (NSS)(NSS)(GSS)(DGS) + TWRVQRSQNPARARATR	12 aa	5526	5778

Table 4.4. Sequence of the bivalent peptidic inhibitor $H4_{20}V4$ and its mutant $H4_{20}V4_LIL$. Bivalent peptidic inhibitor comprises of histone H4 sequence region including K5, K8, K12, K16 and K20, which when acetylated can be recognised by BD, linked by 12 amino acid (aa) long linker region, followed by the viral sequence recognised by ET ($H4_{20}V4$), or the viral sequence region mutant, which abolished the ET domain binding ability.

Non-acetylated $H4_{20}V4$ peptide and acetylated $H4_{20}V4LIL_ac$ peptide were then added to BRD4-BD2ET protein at a 1:1 molar ratio as controls for the $H4_{20}V4_ac$ peptide addition to BRD4-BD2ET. Theoretically, non-acetylated $H4_{20}V4$ peptide cannot be recognised by the BDs of BET but can still bind to the ET domain, whereas $H4_{20}V4LIL_ac$ peptide can still be recognised by BDs of BET but cannot bind to the ET domain. Without the ability to tether the BD2ET two domains simultaneously, the addition of these two peptides to BRD4-BD2ET should result in a complex with an extended conformation, while $H4_{20}V4_ac$ peptide should compact the protein properly.

As expected, the addition of these two control peptides to the protein at a 1:1 ratio had minimal effect on the sedimentation coefficient value *s* and on translational frictional ratio value f/f_0, suggesting they failed to simultaneous tether both domains in BD2ET (**Figure 4.5c, d**). However, the addition of $H4_{20}V4_ac$ peptide shifted the apparent sedimentation coefficient value *s* from 1.5 to 1.9 and translational frictional ratio value f/f_0 from 2.1 to 1.5. These results further confirmed the formation of a compacted complex with $H4_{20}V4_ac$ and BRD4-BD2ET was due to the direct interaction between the peptide and the two domains.

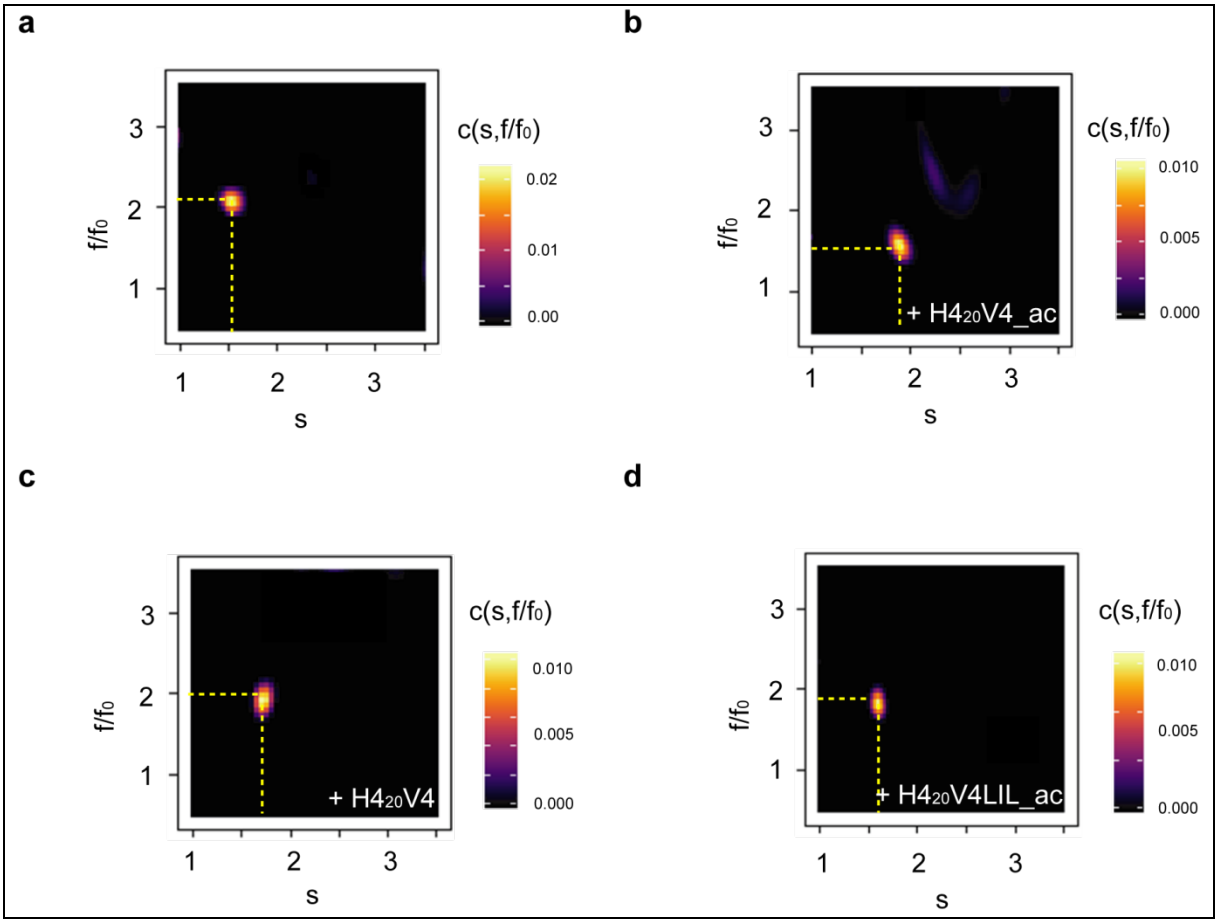


Figure 4.5. In solution evaluation of bivalent peptidic inhibitor $H4_{20}V4$ binding to BRD4-BD2ET by AUC. Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of (**a**) apo-BRD4-BD2ET. (**b**) in the presence of polyacetylated peptide $H4_{20}V4_ac$ at 1:1 protein:peptide ratio, (**c**) in the presence of non-acetylated peptide $H4_{20}V4$, and (**d**) in the presence of polyacetylated viral sequence mutated peptide $H4_{20}V4LIL_ac$. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 300 mM NaCl at 4 °C.

Next, we started to study the binding mode of BRD2-BD2ET engaging bivalent peptides. Similarly, sedimentation velocity experiments were conducted with AUC. All the bivalent peptidic inhibitors were individually added to the protein solution at a 1:1 molar ratio. Hydrodynamic analysis of the $c(s, f/f_0)$ model showed that apo BRD2-BD2ET had an apparent sedimentation coefficient value s of 2.0 and a translational frictional ratio value f/f_0 of 1.9, suggesting an elongated species (**Figure 4.6a**). However, the addition of the set of $H4_{20}$Vn_ac peptides at a 1:1 ratio did not affect the sedimentation coefficient value s nor the translational frictional ratio value f/f_0, suggesting the failure in simultaneous tethering of the two domains in BRD2-BD2ET (**Figure 4.6**). The gradual change in both of the s and f/f_0 values along that occurred with the increase in the bivalent peptide linker length in $H4_{20}$Vn_ac with BRD4-BD2ET was also not observed in the $H4_{20}$Vn_ac peptide to BRD2-BD2ET binding complex. This finding suggested that $H4_{20}$V4_ac was the most optimal bivalent peptidic inhibitor to achieve BRD4-BD2ET protein compaction and did not affect the conformation of BRD2-BD2ET.

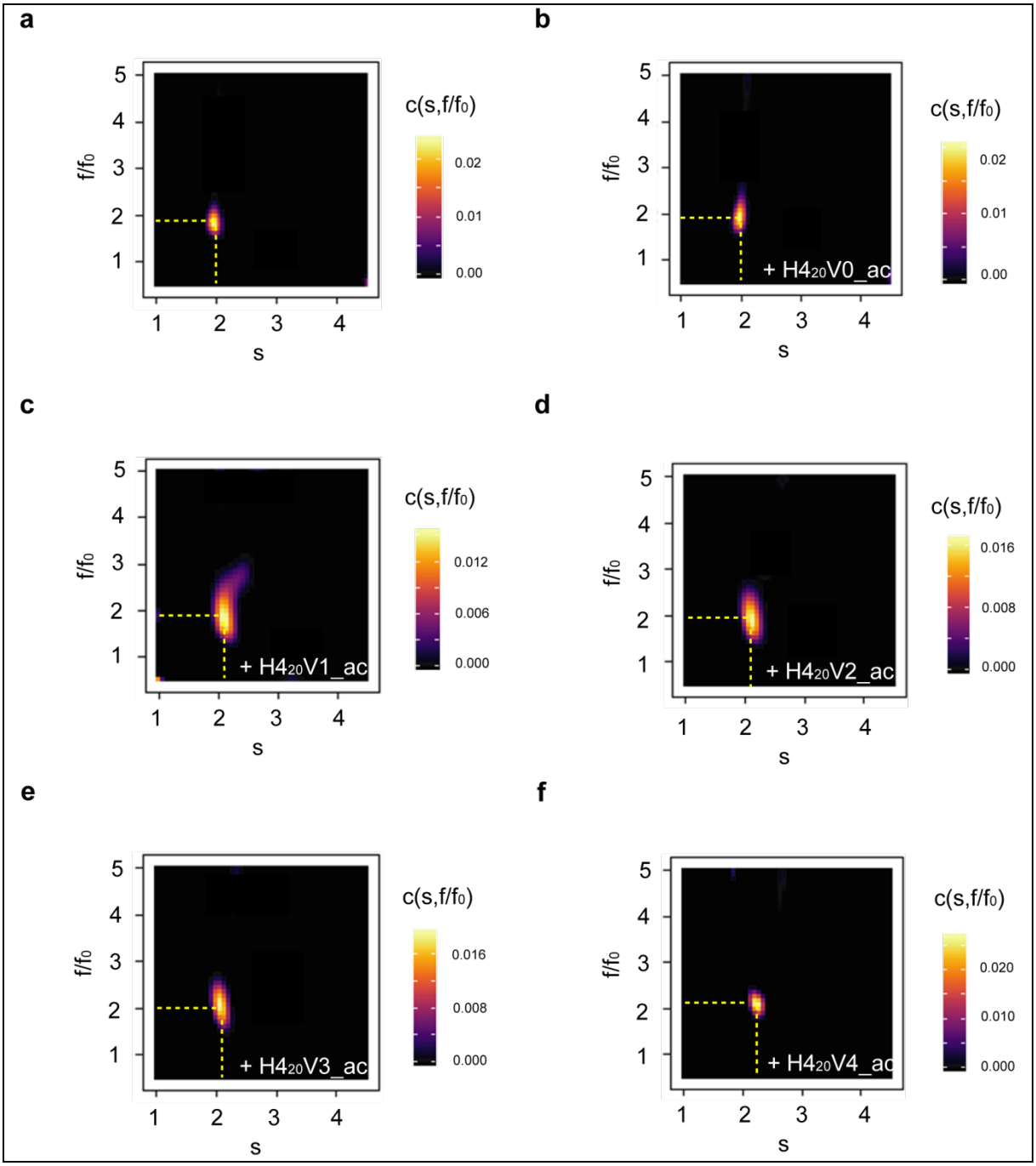


Figure 4.6. Bivalent peptide inhibitor $H4_{20}Vn$ series cannot compact BRD2BD2ET. Sedimentation coefficient and translational frictional ratio distributions overlaid contour heat map of (**a**) apo-BRD2BD2ET and (**b**)(**c**)(**d**)(**e**)(**f**) in the presence of indicated polyacetylated peptide series $H4_{20}Vn$ at 1:1 protein:peptide ratio. Obtained from sedimentation 280 nm absorbance velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 300 mM NaCl at 4 °C.

4.3.2. $H4_{20}V4_ac$ shows 100-fold more selectivity to BRD4-BD2ET over BRD2-BD2ET

To further understand the binding mode specificity of $H4_{20}V4_ac$ for BRD4-BD2ET in comparing with that for BRD2-BD2ET, we performed in solution thermodynamic analysis of

BRD4-BD2ET and BRD2-BD2ET interactions with $H4_{20}V4$_ac peptide and its mutants were performed (**Figure 4.7**). ITC experiments performed as described in the method section by injecting solutions of peptides into the protein solution. The thermodynamic parameters for the interactions obtained from ITC are listed in **Table 4.5**. The interaction data of BRD4-BD2ET with $H4_{20}V4$_ac peptide did not appear to accurately fit to the single binding site model but showed clear two binding events enthalpy trace that fitted well to the more complex two binding sites model (**Figure 4.7a**). The solid smooth line represents the best fit of the experimental data using the two binding sites model. For data analysed using the two binding sites model, the higher-affinity interaction event had a dissociation constant (K_D) of 28.6 nM, whereas the lower-affinity interaction event had a K_D of 327.6 nM. Comparatively, the interaction data of BRD2-BD2ET with $H4_{20}V4$_ac peptide appeared to validly fit to the single binding site model (**Figure 4.7b**). The data analysed using this model showed a binding molar ratio of approximately 1 and a K_D of 2895 nM, which is a 100-fold weaker affinity in comparison with that of BRD4-BD2ET with $H4_{20}V4$_ac peptide interaction (**Table 4.5**). This result supported the binding mode study that showed that $H4_{20}V4$_ac peptide binds to BRD4-BD2ET bivalently while binding to BRD2-BD2ET non-bivalently.

To shed more light into the mechanics of the binding events, the same thermodynamic analysis was conducted on the non-acetylated $H4_{20}V4$ peptide and the ET domain binding motif mutant $H4_{20}V4LIL$_ac peptide. Theoretically, the $H4_{20}V4$ peptide should bind to both BRD4-BD2ET and BRD2-BD2ET in the double-digit nanomolar range since the only possible existing binding available is the interaction between the ET domain and the viral sequence of the peptide. $H4_{20}V4LIL$_ac peptide should bind to both BRD4-BD2ET and BRD2-BD2ET in a similar single-digit micromolar range since the only possible existing binding event in these peptides is the interaction between BD and H4 acetylated lysine sequence. As expected, the interaction data of BRD4-BD2ET with $H4_{20}V4$ peptide exhibited a typical sigmoid curve shaped isotherm

with an inflexion point at a molar ratio of approximately 1:1 and a K_D of 74.9 nM. However, the interaction data of BRD2-BD2ET with $H4_{20}V4$ peptide did not properly fit, indicating a far weaker association, which was unexpected. In the case of $H4_{20}V4LIL_ac$ peptide interactions with BRD4-BD2ET and BRD2-BD2ET, the K_D of each were both in the single-digit micromolar range, as expected.

From the thermodynamic parameters of the interaction between BRD4-BD2ET with $H4_{20}V4_ac$, $H4_{20}V4$ and $H4_{20}V4LIL_ac$ peptides (**Table 4.5**), we observed that the entropy change of the first binding event in the interaction of BRD4-BD2ET with $H4_{20}V4_ac$ peptide had a similar entropy change to that of the BRD4-BD2ET and $H4_{20}V4LIL_ac$ peptide interaction, which had a positive $-T\Delta S$ value, whereas the second binding event had a similar entropy change trend to that of the BRD4-BD2ET and $H4_{20}V4$ peptide interaction, which had a negative $-T\Delta S$ value, suggesting that the first binding event might correspond to the binding of BD to acetylated H4 sequence and the second binding event might correspond to the binding of ET to the viral sequence. However, the K_D and ΔH values are largely different between the first binding event (K_D = 28.6 nM, ΔH= -17.54 kcal/mol) and the interaction of BRD4-BD2ET with $H4_{20}V4LIL_ac$ peptide (K_D = 8689 nM, ΔH= -12.38 kcal/mol). From the affinity and enthalpy values, we speculated that the first binding event might be the result of the bivalent binding of $H4_{20}V4_ac$ peptide to BRD4-BD2ET that includes not only the binding of BD to acetylated H4 sequence but also includes the binding of ET to the viral sequence. This bivalent binding might be the reason for the significant increase in affinity of the first binding event. During this bivalent binding, some enthalpically favourable process might also have occurred and contributed to the drop in enthalpy and the increase in affinity. The second binding event might only include the binding of ET to the viral sequence, and the slight decrease in the binding affinity might be caused by the competition between the excessive free peptide and the peptides already bivalently bound to BRD4-BD2ET.

From the thermodynamic parameters of the interaction between BRD2-BD2ET with $H4_{20}V4$_ac, $H4_{20}V4$ and $H4_{20}V4LIL$_ac peptides (**Table 4.5**), we observed that the entropy change of the BRD2-BD2ET and $H4_{20}V4$_ac peptide interaction had a similar entropy change trend with the BRD2-BD2ET and $H4_{20}V4LIL$_ac peptide interaction, and no second binding events were detected. This observation suggested that there was only one binding event which was the binding of BD to the acetylated H4 sequence, and no binding of ET to the viral sequence as occurred in the interaction of BRD2-BD2ET with the $H4_{20}V4$_ac peptide.

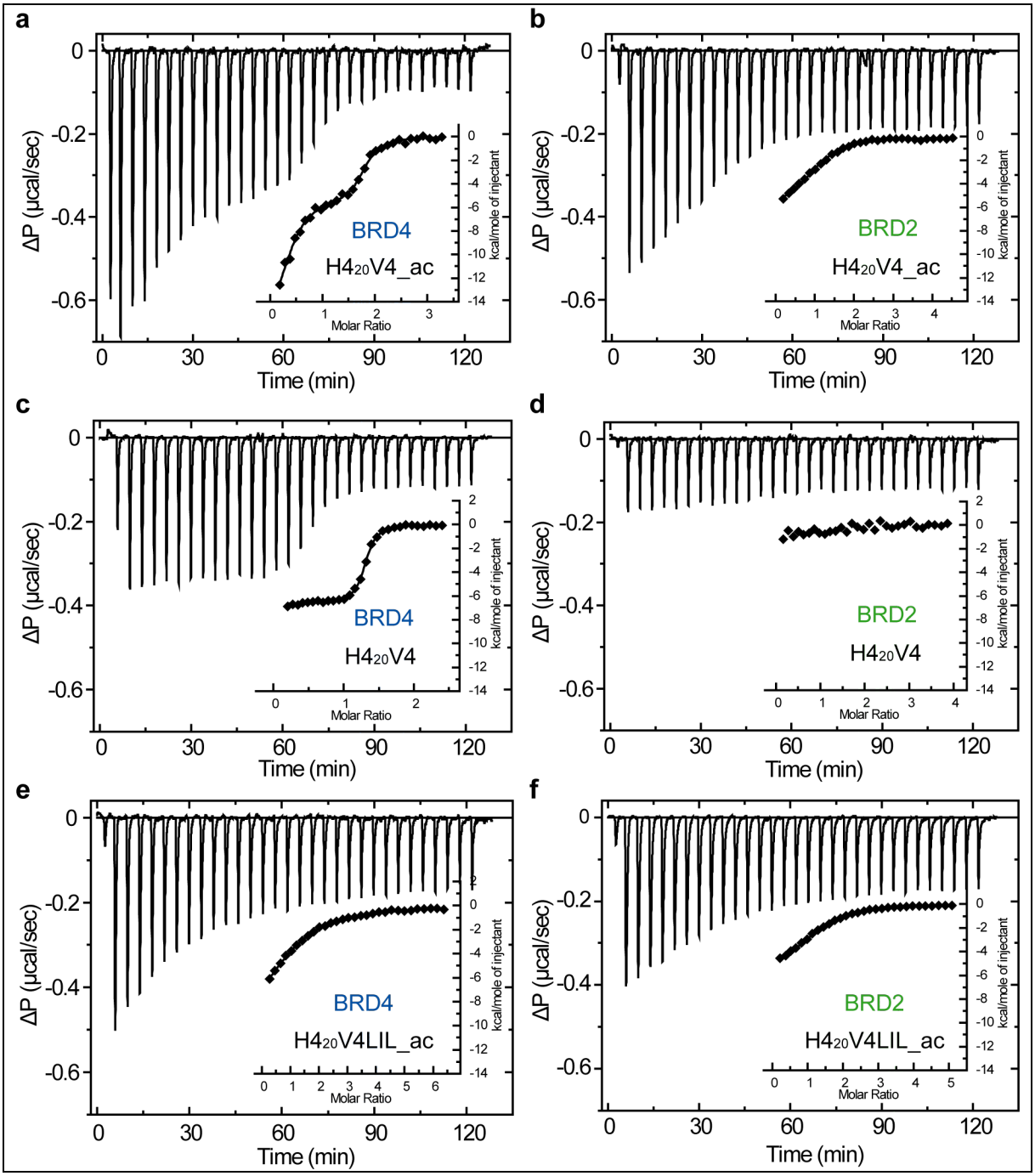


Figure 4.7. In solution evaluation of bivalent peptidic inhibitor H4$_{20}$V4_ac binding to BRD4-BD2ET and BRD2-BD2ET by ITC. Raw injection heats for titration of indicated peptides into a solution of indicated protein and the normalised binding enthalpies corrected for the heat of protein dilution are shown. Solid lines represent a nonlinear least square fit using a multiple site binding model. Titrations were performed in 50 mM HEPES pH 7.5, 300 mM NaCl, and 15 °C while stirring at 295 rpm. (**a**) ITC evaluation of H4$_{20}$V4_ac binding to BRD4-BD2ET, and (**b**) BRD2-BD2ET. (**c**) ITC evaluation of H4$_{20}$V4 binding to BRD4-BD2ET, and (**d**) BRD2-BD2ET. (**e**) ITC evaluation of H4$_{20}$V4LIL_ac binding to BRD4-BD2ET, and (**f**) BRD2-BD2ET.

Protein	Peptide	ITC				
		K_D (nM)	N	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
BRD4-BD2ET	$H4_{20}V4_ac$	28.6 ± 6.2	0.41 ± 0.12	-17.54 ± 1.12	7.33 ± 0.71	-10.21 ± 2.12
		327.6 ± 43.2	1.47 ± 0.09	-5.54 ± 0.42	-3.77 ± 1.1	-9.31 ± 1.23
	$H4_{20}V4$	74.9 ± 12.2	1.31 ± 0.07	-6.93 ± 0.52	-2.89 ± 0.4	-9.82 ± 1.44
	$H4_{20}V4LIL_ac$	8689 ± 621	1.13 ± 0.10	-12.38 ± 1.13	5.58 ± 0.9	-6.80 ± 2.82
BRD2-BD2ET	$H4_{20}V4_ac$	2895 ± 132	1.05 ± 0.05	-6.46 ± 0.62	- 1.41 ± 0.51	-7.87 ± 2.02
	$H4_{20}V4$	-	-	-	-	
	$H4_{20}V4LIL_ac$	4741 ± 233	1.27 ± 0.08	-6.18 ± 0.66	-1.53 ± 0.72	-7.71 ± 1.98

Table 4.5. In solution evaluation of bivalent peptidic inhibitor $H4_{20}V4$ binding to BRD4-BD2ET and BRD2-BD2ET by ITC. Thermodynamic values including dissociation constant (K_D), binding molar ratio (N), enthalpy change (ΔH), and entropy change (ΔS), obtained from Figure 4.7 ITC fitting. Experiments were performed in duplicate. (Shown errors correspond to the mean value ± the range.)

Next, to investigate the reasons for why there was no binding of ET to the viral sequence was detected in the interaction of BRD2-BD2ET with $H4_{20}V4_ac$ peptide, in solution thermodynamic analysis of BRD4-BD2ET (as control) and BRD2-BD2ET interaction with the viral peptide was performed (**Figure 4.8**). ITC experiments were performed as previously described. The thermodynamic parameters for the interactions obtained from ITC are listed in **Table 4.6**. The interaction data of BRD4-BD2ET with the viral peptide exhibited a typical sigmoid curve shaped isotherm with an inflexion point at a molar ratio of approximately 1:1 and a K_D of 83.1 nM. However, the interaction data of BRD2-BD2ET with the viral peptide showed a far weaker association with a K_D of 1809 nM. This result suggested that binding of the viral peptide to the BRD2 ET domain in the context of BRD2-BD2ET tandem protein was impaired with weak micromolar interaction. The reduction of the affinity of the viral sequence to the BRD2 ET domain was further impaired when the viral peptide sequence was present in the $H4_{20}V4_ac$ peptide. This may be the main reason why $H4_{20}V4_ac$ peptide can compact

BRD4-BD2ET by simultaneous tethering of BD and ET domain, whereas no conformational change was observed on BRD2-BD2ET in the presence of $H4_{20}V4_ac$ peptide.

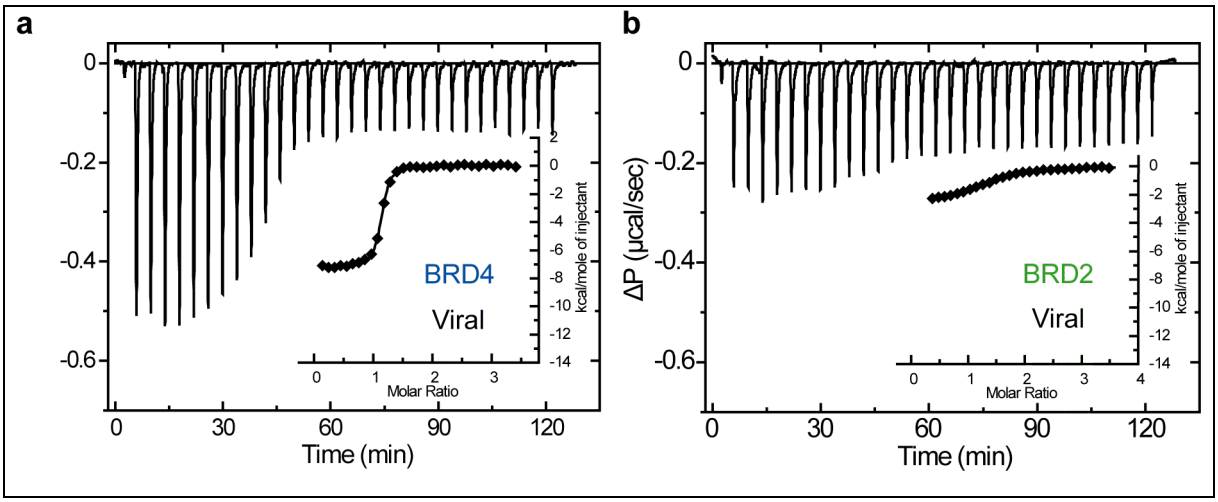


Figure 4.8. In solution evaluation of viral peptide binding to BRD4-BD2ET and BRD2-BD2ET by ITC. Raw injection heats for titration of viral peptide into a solution of indicated proteins and the normalised binding enthalpies corrected for the heat of protein dilution are shown. Solid lines represent a nonlinear least square fit using a multiple site binding model. Titrations were performed in 50 mM HEPES pH 7.5, 300 mM NaCl, and 15 °C while stirring at 295 rpm. (**a**) ITC evaluation of viral peptide binding to BRD4-BD2ET and (**b**) BRD2-BD2ET.

Protein	Peptide	ITC				
		K_D (nM)	*N*	Δ*H* (kcal/mol)	-*T*Δ*S* (kcal/mol)	Δ*G* (kcal/mol)
BRD4-BD2ET	viral	83.1 ± 20.6	1.14 ± 0.09	-7.80 ± 1.23	-2.54 ± 0.62	-10.34 ± 1.22
BRD2-BD2ET	viral	1809 ± 98.2	1.35 ± 0.06	-2.76 ± 0.33	-6.37 ± 2.76	-9.13 ± 1.05

Table 4.6. In solution evaluation of viral peptide binding to BRD4-BD2ET and BRD2-BD2ET by ITC. Thermodynamic values including dissociation constant (K_D), binding molar ratio (*N*), enthalpy change (Δ*H*) and entropy change (Δ*S*), obtained from Figure 4.8 ITC fitting. Experiments were performed in duplicate. (Shown errors correspond to the mean value ± the range.)

4.4. Discussion

In the past few years, intensive studies have focused on targeting BET BDs for drug development. Selective and potent small-molecule BET BD inhibitors have been successfully generated, and their usage in scientific exploration has revealed distinct mechanistic insights

into how BET BD inhibition impacts transcription programs. Recently, two BD inhibitors, iBET-BD1 and iBET-BD2, that can selectively inhibit BD1 or BD2 have been developed[119,165]. This further enriched the BET probes toolbox for understanding the distinct function of BET proteins in various biology. However, the longstanding challenge of selectively targeting each BET member has not been addressed so far.

In this chapter, we tried to use the avidity effect by simultaneously tethering of BD and ET domains of BET proteins to achieve the selectivity between the BET members. Unlike bivalent compounds, we designed peptidic inhibitors that are composed of three essential parts: an acetylated histone H4 sequence that can bind to BDs; a viral sequence that has been reported to bind to ET at nanomolar affinity; and a different length of flexible linker region in between these domains. After successful production of the peptidic inhibitors, we used hydrodynamic analysis to show that among these inhibitors, the $H4_{20}V3_ac$ and $H4_{20}V4_ac$ peptides, with linker lengths of 9 and 12 amino acids, respectively, led to BRD4-BD2ET compaction while all the other peptides with linkers of different lengths failed to do so (**Table 4.7**) (**Supplementary figure 3**). Then we produced shorter peptidic inhibitors $H4_{10}Vn_ac$ that contained only the K5ac and K8ac of histone H4 to investigate what part of the histone H4 sequence played a critical role in binding the BD2 domain in the tandem BD2ET protein with the $H4_{20}V4_ac$ peptide. This demonstrated that $H4_{20}V4_ac$ was the most optimal bivalent peptidic inhibitor with the most optimal linker length and histone H4 sequence to achieve BRD4-BD2ET bivalent engagement (**Table 4.8**).

Protein	Peptides	AUC				
		f/f_0	s	s^0_{20w}	MW(Da)	rmsd(OD)
BRD4-BD2ET	-	2.04	1.52 ± 0.092	2.59	40964	0.0030
	+ $H4_{20}V0_ac$	2.00	1.58 ± 0.088	2.62	42070	0.0029
	+ $H4_{20}V1_ac$	1.98	1.66 ± 0.125	2.85	47887	0.0033
	+ $H4_{20}V2_ac$	1.91	1.65 ± 0.101	2.82	45045	0.0032
	+ $H4_{20}V3_ac$	1.50	1.92 ± 0.099	3.12	45888	0.0034

	+ $H4_{20}V4_ac$	1.48	1.98 ± 0.095	3.18	47796	0.0032
	+ $H4_{20}V4$	2.01	1.72 ± 0.099	2.81	45297	0.0026
	+ $H4_{20}V4LIL_ac$	1.98	1.61 ± 0.083	2.70	44387	0.0028
	+ $H4_{10}V0_ac$	2.09	1.59 ± 0.077	2.66	43776	0.0032
	+ $H4_{10}V1_ac$	1.98	1.49 ± 0.089	2.56	45386	0.0029
	+ $H4_{10}V2_ac$	1.98	1.71 ± 0.100	2.83	44778	0.0027
	+ $H4_{10}V3_ac$	1.99	1.70 ± 0.099	2.82	45895	0.0030
	+ $H4_{10}V4_ac$	2.00	1.53 ± 0.075	2.62	43996	0.0029
BRD2-BD2ET	-	1.99	2.03 ± 0.063	3.12	50804	0.0032
	+ $H4_{20}V0_ac$	1.98	2.00 ± 0.089	3.09	53599	0.0034
	+ $H4_{20}V1_ac$	1.98	2.08 ± 0.115	3.16	54556	0.0028
	+ $H4_{20}V2_ac$	2.01	2.14 ± 0.127	3.21	56878	0.0029
	+ $H4_{20}V3_ac$	2.03	2.12 ± 0.103	3.19	57209	0.0022
	+ $H4_{20}V4_ac$	2.11	2.24 ± 0.098	3.31	58002	0.0038

Table 4.7. In solution evaluation of peptidic inhibitors binding to BRD4-BD2ET and BRD2-BD2ET by AUC. Hydrodynamic values including translational frictional ratio (f/f_0), apparent sedimentation coefficient (s), corrected sedimentation coefficient ($s_{20,w}$), and system mass (MW) values of each BET tandem BDs in the presence of peptides. Obtained from 280 nm absorbance sedimentation velocity experiments which were conducted in 50 mM HEPES, pH 7.5, 300 mM NaCl at 4 °C. (Shown errors on s values represent fitted s distribution value peak area, rmsd is the root-mean-square deviation of detected protein boundary from the boundary generated by Lamm equation modelling)

Protein	AUC	ITC	Comment
BRD4-BD2ET	f/f_0 =2.04 s =1.52	-	Extended shape is consistent with BDET domains flexibly linked by a disordered linker
BRD4-BD2ET + $H4_{20}Vn_ac$	f/f_0 decreased to ~1.50 with $H4_{20}V3_ac$ and $H4_{20}V4_ac$ peptides s increased to 1.98	Tested with $H4_{20}V4_ac$, K_D= 28 nM	Only the $H4_{20}V3_ac$ and $H4_{20}V4_ac$ peptides induced compaction, consistent with bivalent binding
BRD4-BD2ET + $H4_{10}Vn_ac$	No obvious change in f/f_0 and s values	Not tested	$H4_{10}Vn_ac$ peptides cannot induce compaction of BRD4-BD2ET
BRD2-BD2ET + $H4_{20}Vn_ac$	No obvious change in f/f_0 and s values	Tested with $H4_{20}V4_ac$, K_D= 2895 nM	$H4_{20}Vn_ac$ peptides cannot induce compaction of BRD2-BD2ET

Table 4.8. Overview of the hydrodynamic and thermodynamic results done on BRD4-BD2ET and BRD2-BD2ET with the bivalent peptidic inhibitors.

The same set of $H4_{20}Vn_ac$ peptides was then tested with BRD2-BD2ET. Hydrodynamic and thermodynamic analysis revealed that $H4_{20}V4_ac$ peptide only compacted BRD4-BD2ET,

whereas none of all the $H4_{20}Vn_ac$ peptide set compacted BRD2-BD2ET. The $H4_{20}V4_ac$ peptide showed a 100-fold selectivity for BRD4-BD2ET over BRD2-BD2ET with a K_D, obtained from ITC, of 28 nM compared with that of 2895 nM for BRD2-BD2ET (**Table 4.8**). The thermodynamic parameters were further analysed, and an impaired binding ability of the viral peptide sequence to the ET domain in the context of BRD2-BD2ET tandem protein was discovered. This discovery was deemed to be the main reason why the $H4_{20}V4_ac$ peptide could compact BRD4-BD2ET by simultaneous tethering of the BD and ET domains but failed to produce any conformational change with BRD2-BD2ET. For further investigation, the interaction of the viral peptide with BRD2-BD2ET was assessed by ITC and the reduced binding affinity was detected. However, the affinity of viral peptide to BRD4-BD2ET remained in the nanomolar range. But why? The ET domain is highly conserved between BET proteins. Where is the difference coming from? There are two possible explanations, either the linker region between the BD and ET domain of BRD2 was disrupting the interaction or the fourth helix present in the BRD2 ET domain but not in the BRD4 ET domain was disrupting the interaction. In a recent paper on comparative structure-function analysis of BET proteins[220], the fourth helix in the BRD2 ET domain was suggested to be the critical element for the functional difference in specific gene expression between the BRD2 and BRD4. To confirm this possible explanation, BRD2-BD2ET protein either without the fourth helix or with linker region of BRD4 could be produced and hydrodynamic/thermodynamic analysis performed with the $H4_{20}V4_ac$ peptide to see if there is any change in comparison with that of the BRD2-BD2ET protein.

Overall, in this chapter, the selective bivalent peptidic inhibitor $H4_{20}V4_ac$ towards BRD4 over BRD2 was designed, produced, and biophysically validated. This information provided a novel perspective for developing BET inhibitors or even specific inhibitors for BET members.

However, a great amount of work remains to be done at the cellular level to assess the value of this type of bivalent peptidic inhibitors.

Chapter 5. Conclusion and future work

5.1. Conclusion

The work in Chapter 2 described the relative topology of four BET tandem BDs in solution with or without small molecule bivalent inhibitors in order to identify specific patterns of binding behaviour that would differentiate between the highly conserved BET family members. The results not only unravelled the subtle supertertiary structure differences between the four BET tandem BDs but also demonstrated the feasibility of using biophysical methods such as AUC, ITC and SAXS to investigate conformational change differences between the tandem BDs upon inhibitor engagement. Based on the different binding behaviours between BvA-3, BvA-7 and BvB-1 with BRD2(1:2), the optimal linker length and warhead affinity towards each BD for such bivalent inhibitors for simultaneous tethering of both BDs were discussed. Selective binding of one warhead towards BRD2(2) over the lower affinity of the other warheads towards BRD2(2) proved to be the main reason for the formation of the most compact BRD2(1:2) after engaging bivalent inhibitor BvB-1. It would be preferable to use a bivalent inhibitor with iBET-BD1 and iBET-BD2 as two warheads that are selective towards BD1 and BD2, respectively. In this case, we would have an improved starting point for studying the effect of PEG linker lengths on tandem BD bivalent engagement. Hopefully, following the publication of iBET-BD1 and iBET-BD2 compounds earlier this year[165], we believe this form of bivalent inhibitor will be described soon. Although we did not find the exact correlation between the PEG linker length and tandem BD protein linker length in this study, we did discover the inappropriate PEG linker length in bivalent inhibitors that could cause oligomerisation of tandem BDs when engaged (**Figure 5.1**). Our study of chimaera tandem BD constructs revealed that the protein linker regions are the vital determinant for the compound binding behaviour of each BET tandem BD construct. All the studies done in this chapter

indicate that the binding behaviour difference between the four BETs is the main reason for their functional difference in nature.

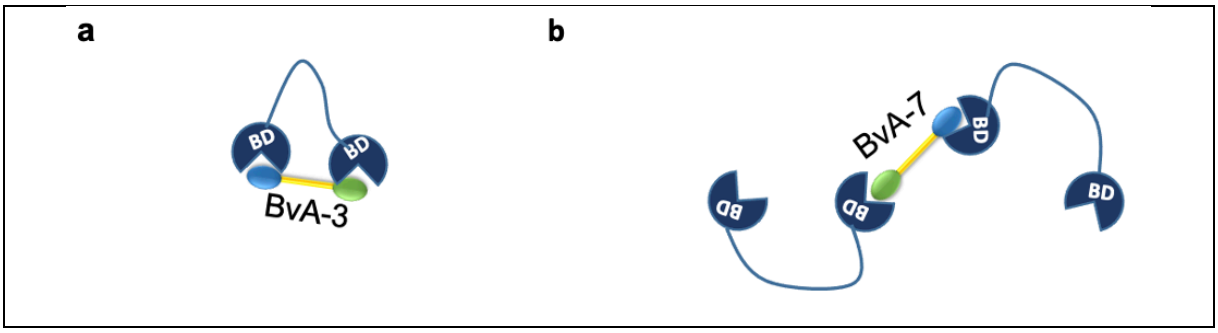


Figure 5.1. Representative figure of engagement mode of BRD2(1:2) with inhibitor BvA-3 and BvA-7 in AUC experiment. (**a**) BRD2(1:2) binds to the shorter inhibitor BvA-3 in bivalent mode (**b**) While dimerizes with the longer inhibitor BvA-7.

In Chapter 3, we used the biophysical tools described in Chapter 2 to explore the binding behaviour of each BET tandem BDs on acetylated histone peptides in order to identify the basis for the functional differences that exist between BETs despite their high degree of sequence similarity. Initially, we used SPOT array and BLI experiments on BRD2(1:2), BRD2(1:2)$^{WT/NF}$, and BRD2(1:2)$^{NF/WT}$ to demonstrate that BRD2(1:2) employs both BDs to bind to distant diacetylated histone H4 peptides. Then we compared the binding mode of BET tandem BDs on polyacetylated histone H4 peptides. The hydrodynamic analysis demonstrated that BRDT(1:2) and BRD4(1:2) bound to polyacetylated histone H4 peptide bivalently with compaction while BRD2(1:2) bound non-bivalently (**Figure 5.2**). This was followed by the thermodynamic analysis of BRDT(1:2) and a series of mutants that confirmed the bivalent binding mechanics of BRDT(1:2) on polyacetylated histone H4 peptide. All the binding behaviour studies of BET tandem BDs conducted in this chapter using the hydrodynamic and thermodynamic methods inspired us to develop the novel bivalent peptidic inhibitor described in chapter 4.

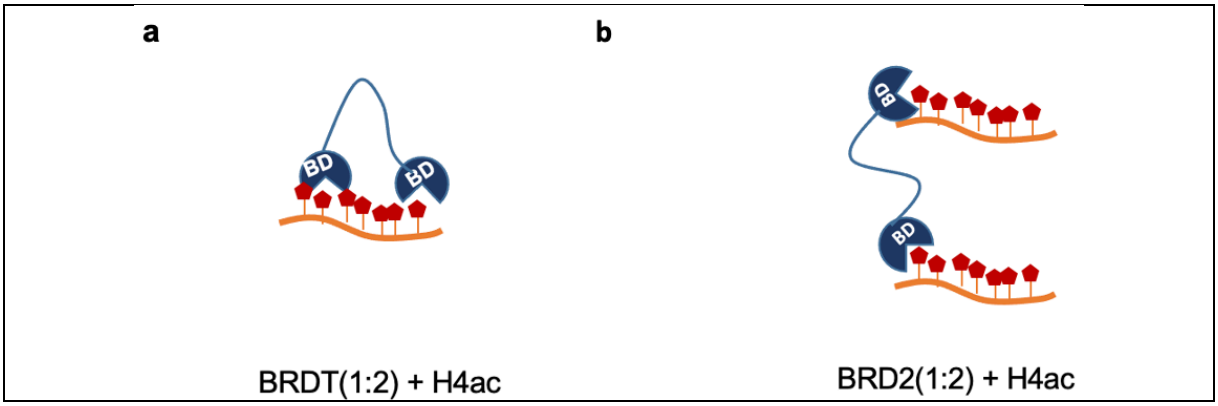


Figure 5.2. Representative figure of engagement mode of BRDT(1:2) and BRD2(1:2) with polyacetylated histone H4 peptide in AUC experiment. (**a**) BRDT(1:2) binds to the polyacetylated histone H4 peptide in bivalent compaction mode (**b**) BRD2(1:2) binds to the polyacetylated histone H4 peptide in 1:2 protein:ligand binding mode.

In Chapter 4, a peptidic inhibitor that could bivalently capture one of the BET proteins via both the BD region and the ET domain was designed. This series of inhibitors was composed of three essential parts: acetylated histone H4 sequence that can bind to BDs, a viral sequence that has been reported to bind to ET domain in nanomolar affinity, and a different length flexible linker region in between. Out of the 12 different bivalent peptides produced, only $H4_{20}$V3_ac and $H4_{20}$V4_ac peptides, which have linker lengths of 9 and 12 amino acids, respectively, led to compaction of BRD4-BD2ET but not that of BRD2-BD2ET. This demonstrated that $H4_{20}$V4_ac was the bivalent peptidic inhibitor with the most optimal linker length and histone H4 sequence that achieved BRD4-BD2ET bivalent engagement (**Figure 5.3**). The further thermodynamic study revealed that $H4_{20}$V4_ac peptide was 100-fold more selective for BRD4-BD2ET over BRD2-BD2ET with a K_D of 28 nM to BRD4-BD2ET compared with that of 2895 nM to BRD2-BD2ET. The low affinity of $H4_{20}$V4_ac peptide to BRD2-BD2ET was later discovered as the cause of impaired affinity of the viral peptide for BRD2-BD2ET. Comparatively, the affinity of the viral peptide for BRD4-BD2ET remained in the nanomolar range as expected.

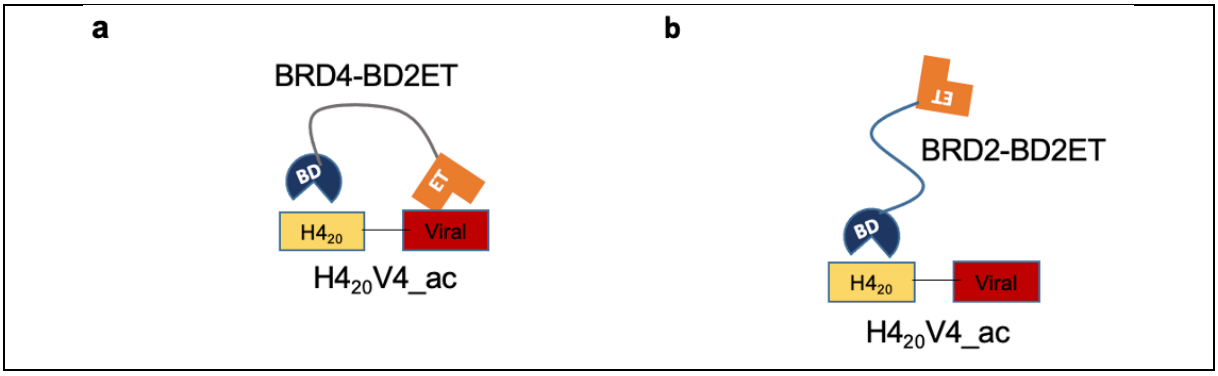


Figure 5.3. Representative figure of engagement mode of BRD4-BD2ET and BRD2-BD2ET with $H4_{20}V4$ in AUC experiment. **(a)** BRD4-BD2ET binds to the $H4_{20}V4$ peptide in bivalent compaction mode **(b)** BRD2-BD2ET binds to the $H4_{20}V4$ peptide non bivalently with a 100-fold weaker affinity.

Producing and biophysical assessment of BRDT-BD2ET and BRD3-BD2ET would be ideal in order to further characterise the $H4_{20}Vn_ac$ peptide binding with the remaining two BETs. This was planned to finish as part of my thesis before Covid-19 began. If the $H4_{20}V4_ac$ peptide can only compact and bind to BRD4-BD2ET at nanomolar affinity while interacting with other BETs (BRD2-BD2ET, BRD3-BD2ET, and BRDT-BD2ET) with low affinity, this would suggest we have successfully developed a BRD4-selective bivalent peptidic inhibitor *in vitro*. Currently, however, it is safe to assert that we have developed a bivalent peptidic inhibitor that is selective for BRD4 over BRD2.

5.2. Future work

Our study of chimaera tandem BD constructs revealed that the protein linker regions are the vital determinant for the compound binding behaviour of each BET tandem BD construct. However, to understand the underlying mechanics of this observation, producing a BET tandem BD construct with a random sequenced linker region of different lengths would be ideal. If the new constructs follow the binding behaviour of wild-type BET tandem BD

constructs with corresponding linker length, we can assume that linker length is the core determinant for the overall conformational change of each BET.

The study on the binding mode of BET tandem BDs on polyacetylated histone H4 peptide revealed that BRD2(1:2), BRD4(1:2), and BRDT(1:2) each have different binding behaviour. To study if the different binding mode on acetylated histone H4 peptide determines the physiological functional differences between each BET tandem BDs, several cellular level experiments are required; these would express linker chimaera BET mutants in the wild-type BET-depleted cell lines to detect any transcriptional changes. For example, using stable BRD2 knockout cells that express the $BRD2^{L4}$ construct, which contains the BD from BRD2 and linker region from BRD4, at a similar level with wild-type BRD2 to compare the expression of *BRD2*-based transgenes[220] and control genes (*ACTB* and *GAPDH*). If the expression level of the BRD2 target genes changes, this result would support the idea of the different binding modes determining functional differences.

The study on the binding of $H4_{20}$V4_ac peptide to BRD2-BD2ET and BRD4-BD2ET revealed that $H4_{20}$V4_ac peptide was 100-fold more selective for BRD4-BD2ET over BRD2-BD2ET with a K_D of 28 nM to BRD4-BD2ET compared with that of 2895 nM to BRD2-BD2ET. To further explore the mechanics behind the reduced affinity of viral peptide to BRD2-BD2ET, more experiments are required to confirm how the linker region of BRD2-BD2ET disrupts the interaction between the viral peptide and ET domain. For example, a chimaera BRD2-BD2ET construct with the linker region of BRD4 or a deletion mutant BRD2-BD2ET construct without the fourth helix in the ET domain could be produced and used to characterise the interaction of the new construct with the $H4_{20}$V4_ac peptide. If the new chimaera BRD2-BD2ET constructs restore the nanomolar affinity with the viral peptide, this would support the idea of the linker region between BD2 and ET of BRD2 disrupting the interaction between the ET domain and viral peptide and resulting in lower affinity.

Discovering a peptidic inhibitor that selectively inhibits BRD4-BD2ET over BRD2-BD2ET *in vitro* is definitely great news. However, a great amount of work remains to be done *in vivo* to evaluate the effects of this type of bivalent peptidic inhibitor, and this may be problematic for cellular evaluation as peptidic inhibitors usually present poorer pharmacokinetic characteristics and are generally considered to be poor drug molecules [229]. Peptides tend to be rapidly metabolized and usually do not cross the cell membrane as some small molecules do. Regarding this potential downside of our peptidic inhibitor, the finding of iBET-BD1 and iBET-BD2 selective compounds are definitely great news for us. iBET-BD2 could replace the histone H4 sequence, a cyclic peptide could replace the viral sequence in our bivalent peptidic inhibitor $H4_{20}V4$. It has been reported that cyclic peptides show better biological activity compared to linear peptides due to conformational rigidity. They are resistant to hydrolysis and can cross the cell membrane[229]. Conjugating the two elements (IBET-BD inhibitors and cyclic peptide) by equivalent PEG linker region to construct a new probe would significantly improve the pharmacokinetic features of our bivalent peptidic inhibitor. The study data on iBET-BD1 and iBET-BD2 selective inhibitors suggested that inhibiting BD2 while not affecting BD1 in BET proteins could be a useful therapeutic strategy to counter the immunoinflammatory damage. Our bivalent inhibitor, which can bind BD2 and ET domain as well as selective towards BRD4 over BRD2, would definitely be a novel and useful therapeutic strategy if successfully developed.

There are three different designing strategies for it. The probe can only contain iBET-BD2 and a cyclic viral peptide region linked by a PEG linker (**Figure 5.4a**). The advantage of this design is that it won't affect the binding of BD1, as BD1 and BD2 have their own functions, leaving BD1 not affected can be beneficial in immunoinflammatory damage. Not only that, but this probe is also a BRD4 selective inhibitor over BRD2 (of course, more validations are needed here), selectivity would undoubtedly add more benefits. Another way of designing the probe is

to contain iBET-BD1 and cyclic viral peptide, leaving BD2 open (**Figure 5.4b**). Similar to the above, the advantage of this design is that it won't affect the binding of BD2 and could be a BRD4 selective inhibitor. However, because of the long distance between the BD1 and ET domain, also the free BD2 domain, the probe might not work or bind to the protein as we would expect. The third way is to contain all three elements for binding all three domains (**Figure 5.4c**). This probe would have iBET-BD1, iBET-BD2, and the cyclic viral peptide. This strategy would be great if only it binds to the BETs in multivalent mode and selective towards only one of them. The possible disadvantages of this type of probe are that the probe might not bind to the protein as we would expect and randomly binds to the other BETs in a 1:3 protein:ligand ratio. However, all the drugs or inhibitors need careful evaluation. The above discussions on designing multivalent probes are just speculation and experimental work needed for the accurate assessment.

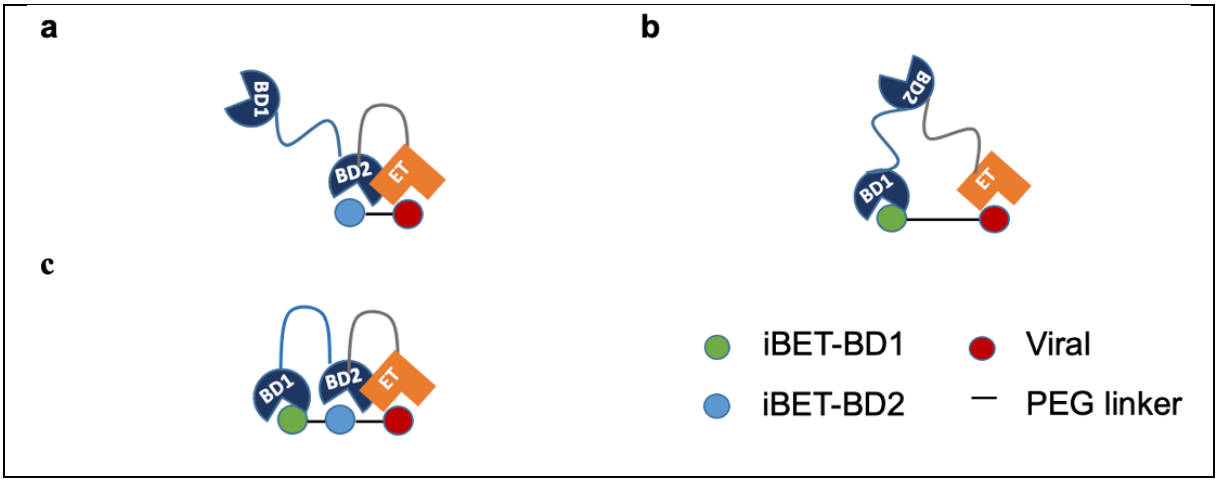


Figure 5.4. Representative figure of engagement mode of BET proteins with novel multivalent peptidic inhibitors. (a) Inhibitor contains iBET-BD2 for capturing BD2 and viral region for capturing the ET domain. (**b**) Inhibitor contains iBET-BD1 for capturing BD1 and viral region for capturing the ET domain. (**c**) Inhibitor contains both iBET-BD1 and iBET-BD2 for capturing BD1, BD2 and viral region for capturing the ET domain.

Chapter 6. Material and Methods

6.1. Cloning

6.1.1. Cloning of tandem BET BD and BET-BD2ET constructs

SGC entry clone constructs were used as templates to amplify BD regions by PCR. PCR products were purified and subcloned into pNIC28-Bsa4 or pNIC-CTHF, by ligation independent cloning[230] (LIC). For cloning of chimeric protein, Gibson assembly master mix kit (NEB Inc) was used. Constructs were transformed into competent MACH1 cells to obtain the final plasmid DNA and were verified by sequencing.

6.1.2. Cloning of tandem histone H4 - viral peptides

SGC constructs were used as templates to amplify the histone H4 region and viral regions by PCR. Viral peptide contains one lysine residue; in order to not affect the binding capacity of final product viral peptide region to ET domain after acetylation, this lysine (K) residue of the viral peptide was mutated to arginine (R) in the first step PCR. Products were then purified and subjected to next step PCR for adding the linker region. The final PCR products subcloned into pNIC-CTHF vector by LIC. Constructs were transformed into competent MACH1 cells to obtain the final plasmid DNA and were verified by sequencing (**Table 6.1**).

	Sequence
Histone H4	SMSGRGKGGKGLGKGGAKRHRKVL
Viral	TWRVQRSQNPLKIRLTR
Aim	SMSGRGKGGKGLGKGGAKRHRKVL + **(linker)** + TWRVQRSQNPLRIRLTR

Table 6.1. Sequence components of proposed bivalent peptide.

For constitution of the flexible linker region, glycine (G) and serine (S) residues are used to the most according published research[228]: A conformational flexibility scale for amino acids in peptides. To add the diversity for the linker region, one or two top scaled flexible amino acids

asparagine (N) and aspartic acid (D) residues were also used. The designed linker regions were DGS, GSSDGS, NSSGSSDGS, and NSSNSSGSSDGS.

6.2. Protein expression and purification

Protein and peptides	Figure
BRDT(1:2)	Figure 6.1
BRD2(1:2)	Figure 2.4
BRD4(1:2)	Figure 6.2
BRD2(1:2)^{LT}	Figure 6.3
BRD4-BD2ET	Figure 6.4
BRD2-BD2ET	Figure 6.5
$H4_{50}$ peptide	Figure 6.6
$H4_{20}V0$	Figure 6.7
$H4_{20}V1$	Figure 6.8
$H4_{20}V2$	Figure 4.2
$H4_{20}V3$	Figure 6.9
$H4_{20}V4$	Figure 6.10
$H4_{10}V0$	Figure 6.11
$H4_{10}V1$	Figure 6.12
$H4_{10}V2$	Figure 6.13
$H4_{10}V3$	Figure 6.14
$H4_{10}V4$	Figure 6.15

Table 6.2. Summary of the protein and peptides purified with corresponding figures.

6.2.1. Protein expression and purification of tandem BET BDs.

Bacterial expression**:** Constructs were transformed into competent BL21 (DE3)-R3-pRARE2 cells. Cells were grown at 37 °C with agitation speed of 180 rpm in lysogeny broth (LB) supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol from overnight cultures. Protein expression was induced overnight with 0.1 mM isopropyl-b-D-thiogalactopyranoside (IPTG) when the optical density OD_{600} nm when reached 0.8 and the culture temperature was lowered to 18 °C. Bacterial cultures were pelleted by centrifugation (5000 rpm for 15 min) using a JLA8.1 rotor Beckman centrifuge and kept at -80 °C until protein purification.

Protein purification: The cell pellets were resuspended in lysis buffer (50 mM HEPES, pH 7.5, 1.5M NaCl, 5% glycerol in the presence of 1:1000 (v/v) Protease Inhibitor Cocktail III and 1 mM TCEP) at 20 °C. Cells were lysed at 4 °C using a Constant Systems cell disruptor at 25 k.s.i and the DNA was removed by precipitation on ice for 30 min with 0.15% (v/v) of PEI (polyethyleneimine). Lysates were centrifuged (21,000 rpm for 45 min at 4 °C, JA 25.50 rotor) and the supernatant was applied to a gravity-flow Ni-NTA resin column. Before application of the supernatant, the Ni-NTA resin column was first rinsed with 10 bed volumes of lysis buffer. After the column was subsequently loaded with the supernatant, columns were washed once with 20 mL of lysis buffer to remove non-specifically bound material. Proteins were then eluted using a step elution of imidazole in lysis buffer (50, 100, 150, 250 mM Imidazole) in 5mL volume. All fractions were collected and checked by SDS-PAGE. The eluted proteins were treated overnight at 4 °C with TEV protease to remove the His-tag and were further purified by size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on ÄktaPrime plus systems in Buffer 10 mM HEPES, 500 mM NaCl and 5% glycerol at a constant flow rate (1 mL/min) and filtrates were collected in 1 mL fractions. The quality of SEC-purified protein sample was collected and checked by SDS-PAGE. The purified samples were further purified by cation-exchange chromatography on HiTrap™ SP-HP column on ÄktaPure system eluted with 1M NaCl, HEPES pH 7.5 buffer. Fractions were checked by SDS-PAGE to confirm their quality. Proteins in 50 mM HEPES pH 7.5, 300 mM NaCl and 5% glycerol were concentrated to 10 mg/mL with concentrators with 30 kDa molecular weight cut-offs (30 MWCO), frozen in liquid nitrogen and stored at -80 °C. Protein concentrations were determined by a NanoDrop 1000.

All of the buffers used in the purification procedures were pre-filtered through 0.2 μm pore size filter papers under vacuum to remove precipitates. Buffer used in ion-exchange were

cation exchange buffer A: 50 mM HEPES, pH 7.5**;** cation exchange buffer B: 50 mM HEPES, pH 7.5, 1M NaCl.

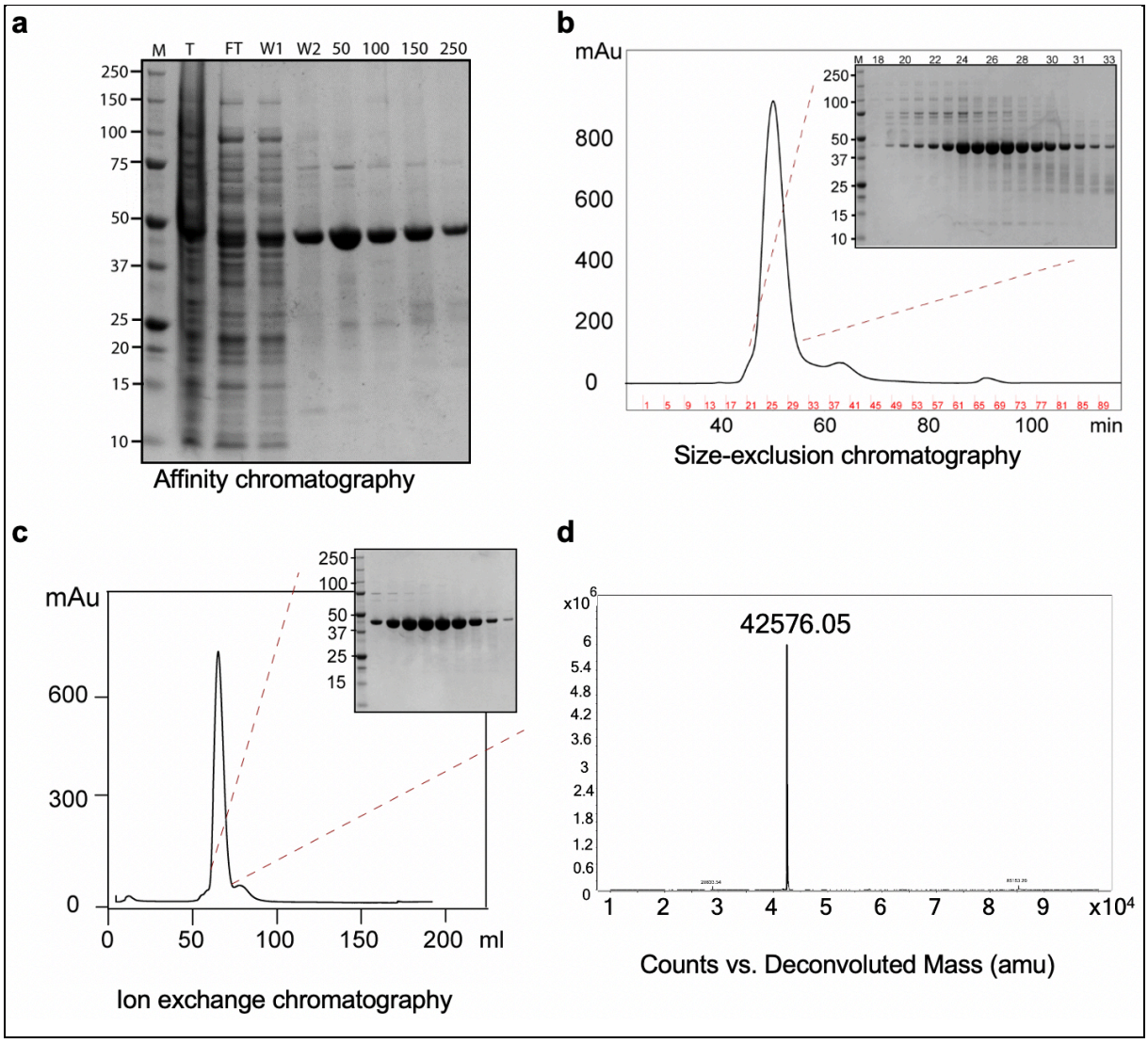


Figure 6.1. Protein purification process of BRDT(1:2). (a) SDS-PAGE analysis of BRDT(1:2) after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel) (**b**) Size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus system in buffer (10 mM HEPES, 500 mM NaCl and 5% glycerol) and SDS-PAGE analysis of the collected fractions (**c**) Cation-exchange chromatography on HiTrap™ SP-HP column on an ÄktaPure system eluted with 1M NaCl in HEPES pH 7.5 buffer and SDS-PAGE analysis of the collected fractions. (**d**) Mass spectrometry confirmation of purified BRDT(1:2) protein in Quad Time-Of-Flight mass spectrometry systems.

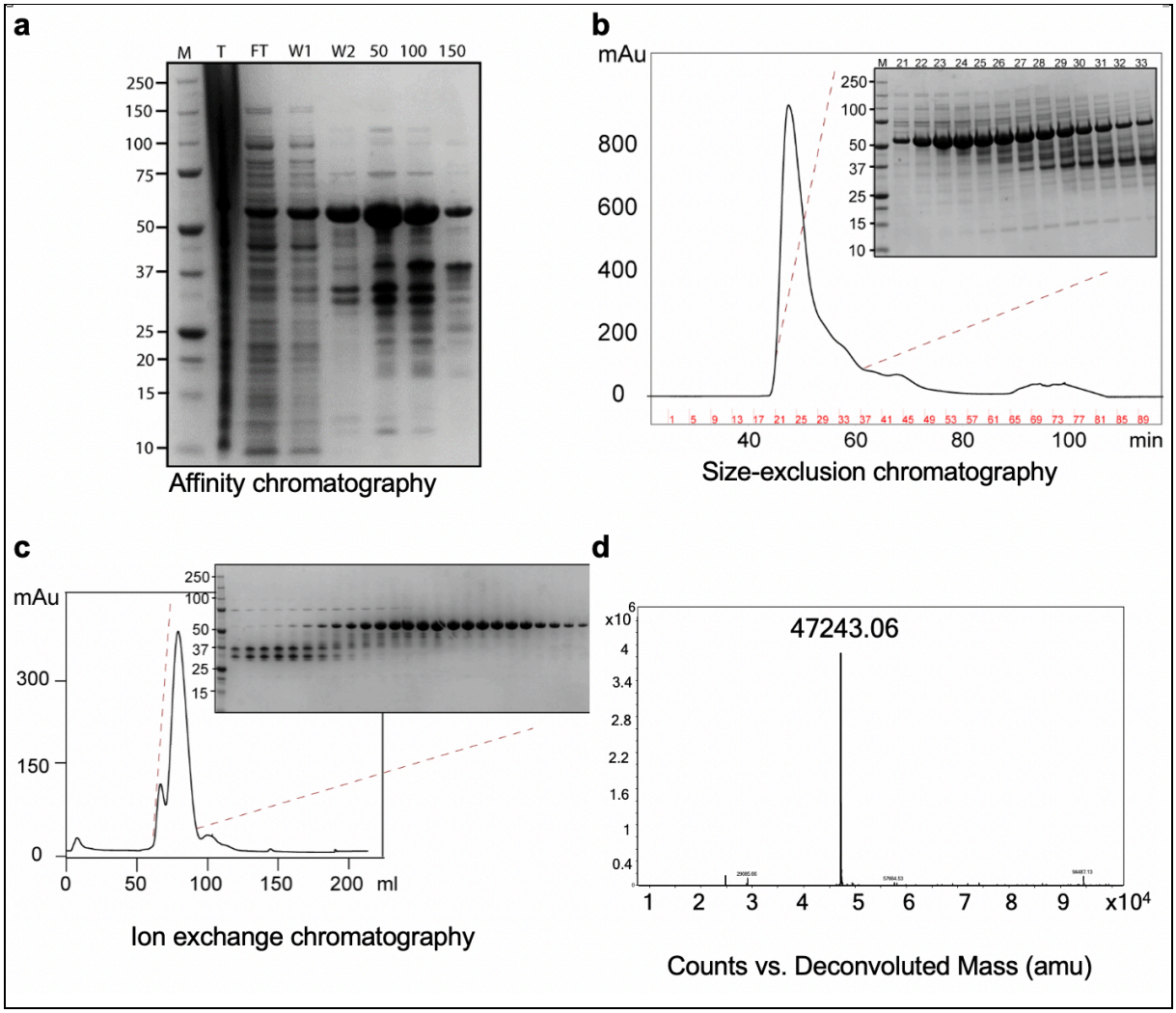


Figure 6.2. Protein purification process of BRD4(1:2). (a) SDS-PAGE analysis of BRD4(1:2) after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel) (**b**) Size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus system in buffer (10 mM HEPES, 500 mM NaCl and 5% glycerol) and SDS-PAGE analysis of the collected fractions (**c**) Cation-exchange chromatography on HiTrap™ SP-HP column on an ÄktaPure system eluted with 1M NaCl in HEPES pH 7.5 buffer and SDS-PAGE analysis of the collected fractions. (**d**) Mass spectrometry confirmation of purified BRD4(1:2) protein in Quad Time-Of-Flight mass spectrometry systems.

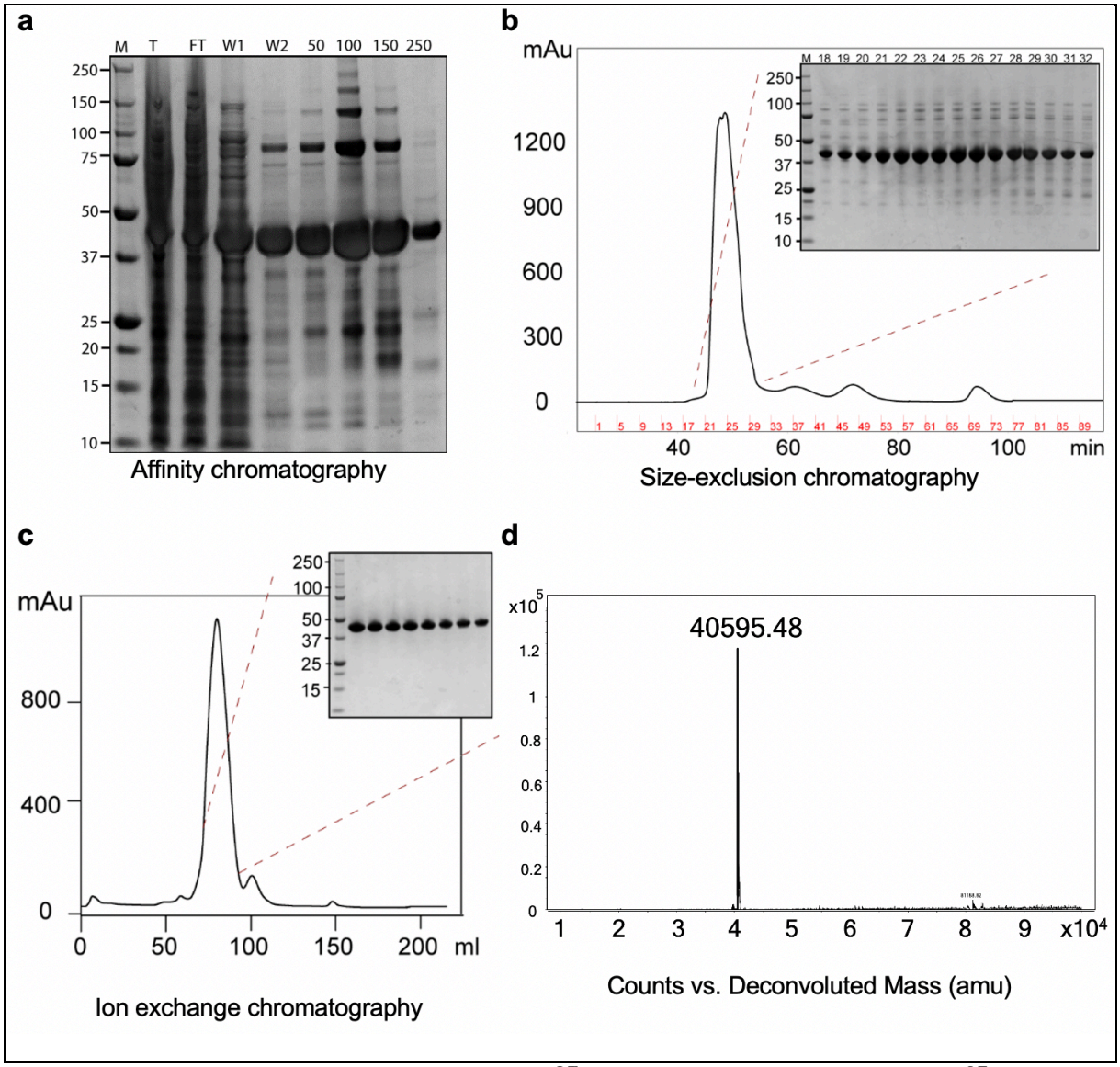


Figure 6.3. Protein purification process of BRD2(1:2)^{LT}. (a) SDS-PAGE analysis of BRD2(1:2)^{LT} after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) (**b**) Size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus system in buffer (10 mM HEPES, 500 mM NaCl and 5% glycerol) and SDS-PAGE analysis of the collected fractions (**c**) Cation-exchange chromatography on HiTrap™ SP-HP column on an ÄktaPure system eluted with 1M NaCl in HEPES pH 7.5 buffer and SDS-PAGE analysis of the collected fractions. (**d**) Mass spectrometry confirmation of purified BRD2(1:2)^{LT} protein in Quad Time-Of-Flight mass spectrometry systems.

6.2.2. Protein expression and purification of BET-BD2ET constructs

The expression and purification process of BRD2-BD2ET and BRD4-BD2ET constructs were similar to the expression and purification process of tandem BET BDs, except from some of the buffer compositions. Constructs were transformed into competent BL21 (DE3)-R3-

pRARE2 cells. Cells were grown at 37 °C with agitation speed of 180 rpm in LB supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol from overnight cultures. Protein expression was induced overnight with 0.1 mM IPTG when the optical density OD_{600} nm when reached 0.8, and the culture temperature was lowered to 18 °C. Bacterial cultures were pelleted by centrifugation (5000 rpm for 15 min) using a JLA8.1 rotor Beckman centrifuge and kept at -80 °C until protein purification.

Protein purification: The cell pellets were resuspended in lysis buffer (50 mM Tris, pH 8 at 20 °C, 1 M NaCl, 5% glycerol) in the presence of 1:1000 (v/v) Protease Inhibitor Cocktail III and 1 mM TCEP. Cells were lysed at 4 °C using a Constant Systems cell disruptor at 25 k.s.i and the DNA was removed by precipitation on ice for 30 min with 0.15% (v/v) of PEI. Lysates were centrifuged (21,000 rpm for 45 min at 4 °C, JA 25.50 rotor), and the supernatants were applied to gravity-flow Ni-NTA resin column. Before application of the supernatant, the Ni-NTA resin column was first rinsed with 10 bed volumes of lysis buffer. After the column was subsequently loaded with the supernatant, columns were washed once with 20 mL of lysis buffer to remove non-specifically bound material. Proteins were then eluted using a step elution of imidazole in lysis buffer (50, 100, 150, 250 mM Imidazole) in 5 mL volume. All fractions were collected and checked by SDS-PAGE. The eluted proteins were treated overnight at 4 °C with TEV protease to remove the His-tag and were further purified by size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus systems in Buffer 10 mM Tris, 500 mM NaCl and 5% glycerol at a constant flow rate (1 mL/min) and filtrates were collected in 1 mL fractions. The quality of SEC-purified protein sample was collected and checked by SDS-PAGE. The purified samples were further purified by anion-exchange chromatography on HiTrap™ Q-HP column on an ÄktaPure system and eluted with 1M NaCl, Tris pH 8 buffer. Fractions then checked by SDS-PAGE for the quality confirm. Proteins in 50 mM Tris pH8, 300 mM NaCl and 5% glycerol were concentrated to 10

mg/mL with concentrators with 30 kDa cut-offs (30 MWCO), frozen in liquid nitrogen and stored at - 80 °C. Protein concentrations were determined by a NanoDrop 1000.

All of the buffers used in the purification procedures were pre-filtered through 0.2 µm pore size filter papers under vacuum to remove precipitates. Buffer used in ion-exchange were cation exchange buffer A: 50 mM Tris, pH 8**;** Cation exchange buffer B: 50 mM Tris, pH 8, 1M NaCl.

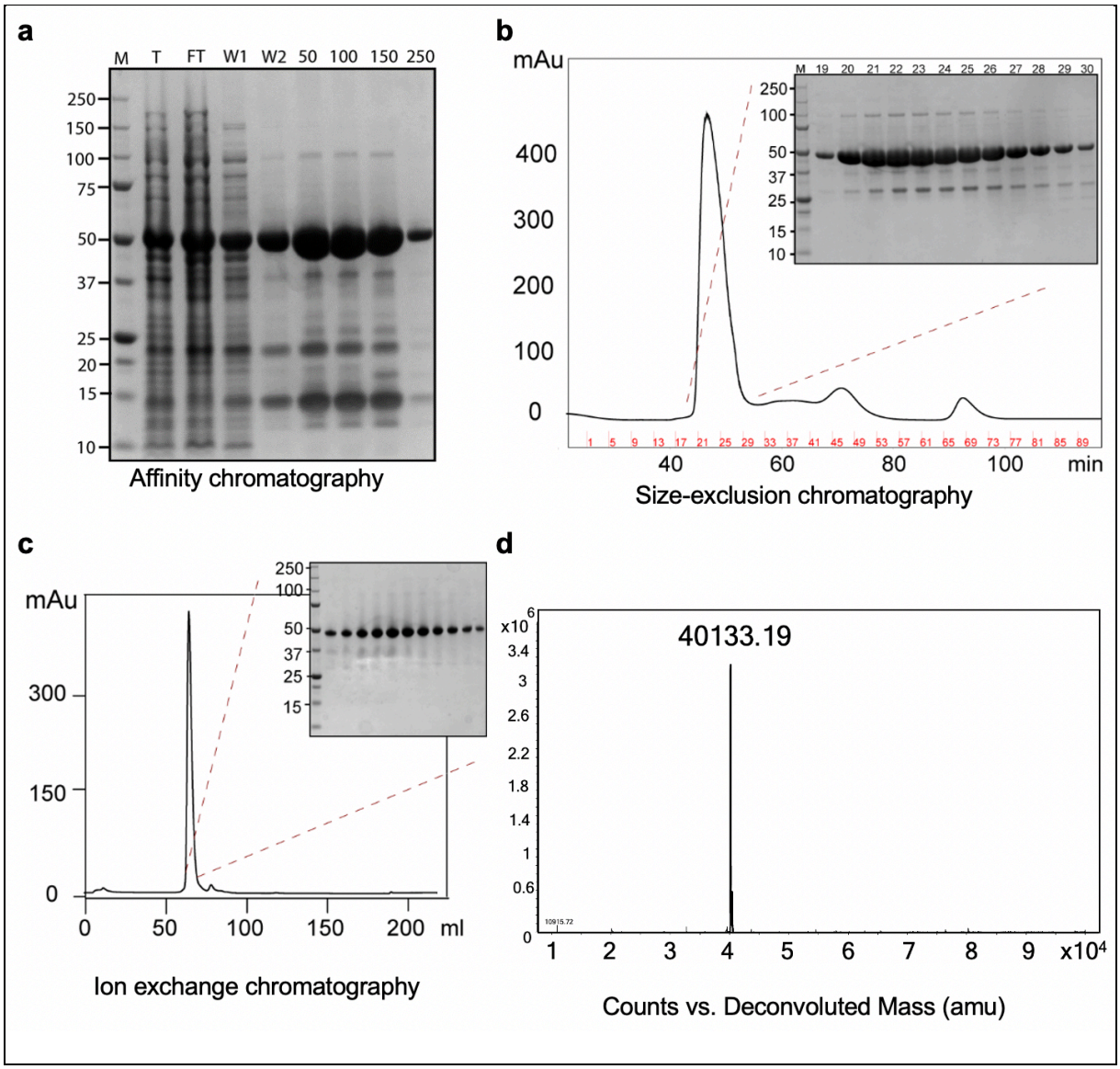


Figure 6.4. Protein purification process of BRD4-BD2ET. (a) SDS-PAGE analysis of BRD4-BD2ET after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel) (**b**) Size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus system in buffer (10 mM Tris, 500 mM NaCl and 5% glycerol) and SDS-PAGE analysis of the collected fractions. (**c**) Anion-exchange chromatography on HiTrap™ Q-HP column on an ÄktaPure system eluted with 1M NaCl in Tris pH 8 buffer and

SDS-PAGE analysis of the collected fractions. (**d**) Mass spectrometry confirmation of purified BRD4-BD2ET protein in Quad Time-Of-Flight mass spectrometry systems.

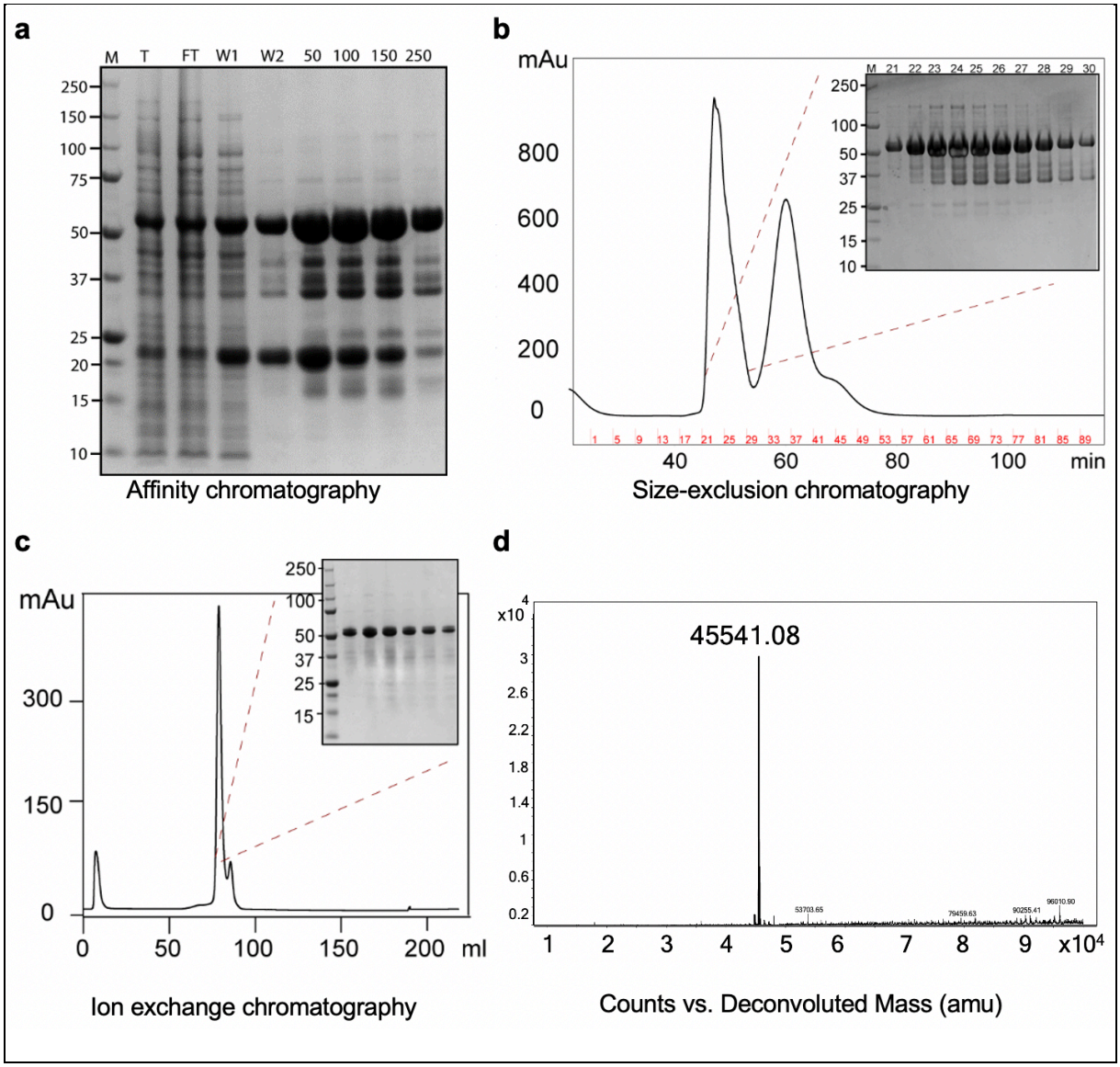


Figure 6.5. Protein purification process of BRD2-BD2ET. (a) SDS-PAGE analysis of BRD2-BD2ET after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel). (**b**) Size exclusion chromatography on Superdex 75 16/60 HiLoad gel filtration columns on an ÄktaPrime plus system in buffer (10 mM Tris, 500 mM NaCl and 5% glycerol) and SDS-PAGE analysis of the collected fractions (**c**) Anion-exchange chromatography on HiTrap™ Q-HP column on an ÄktaPure system eluted with 1M NaCl in Tris pH 8 buffer and SDS-PAGE analysis of the collected fractions. (**d**) Mass spectrometry confirmation of purified BRD2-BD2ET protein in Quad Time-Of-Flight mass spectrometry systems.

6.2.3. Protein expression, purification and acetylation of histone H4 peptide and tandem histone H4 - viral peptides

Expression: Peptide DNA were designed as fused with yellow fluorescent protein (Venus-tag) DNA in expression constructs in order to increase the stability of the 50-mer peptides during expression. Constructs were transformed into competent BL21 (DE3)-R3-pRARE2 cells and grown at 37 °C in terrific broth medium supplemented with 50 µg/mL kanamycin and 34 µg/mL chloramphenicol from overnight cultures. Protein expression was induced overnight with 0.5 mM IPTG at 18 °C at an OD_{600} nm of 2.5, and the culture temperature was lowered to 25 °C. Cultures were pelleted by centrifugation (5000 rpm for 15 min) using a JLA8.1 rotor Beckman centrifuge Beckman centrifuge and kept at -80 °C until protein purification.

Purification: The cell pellets were resuspended in lysis buffer (50 mM HEPES, pH 7.5, 1.5 M NaCl in the presence of 1:1000 (v/v) Protease Inhibitor Cocktail III and 1 mM TCEP) at 20 °C. Cells were lysed, disrupted, centrifugated, and passed through a Ni-NTA column as previously described. Proteins were eluted in one step in 50 mM HEPES pH 7.5, 1500 mM NaCl and 500 mM imidazole. Fractions were checked by SDS-PAGE, and the cleanest fraction pooled and were treated overnight at 37 °C with TEV protease to remove the Venus-His-tag. Next day, the peptide solution was centrifuged for 10 min at 21,000 rpm to remove precipitated Venus-tag; the retentate was transferred to an Amicon concentrator (MWCO 50 kDa) and spin at 4000 rpm until the majority of the sample volume had been filtered and the sample concentrated in the retentate. This step was repeated with MWCO 30 kDa and MWCO 10 kDa concentrator, until MWCO 3kDa peptide were then collected and concentrated to ~ 2 mg/mL. Peptides were then placed in a beaker.

Acetylation: Acetic anhydride was added to the peptide solution drop by drop while monitoring the solution pH and stabilised the pH to 7.0 by addition of NaOH. After checking if all the lysine were successfully acetylated by mass spectrometry (in addition to the lysine present in

the peptide sequence and the N-terminal residue, acetic anhydride can also acetylate amino acids containing hydroxyl groups, they are tyrosine, serine and threonine), the pH was adjusted to pH11 with NaOH to hydrolyse the esterified residues for 2 hours. The final product was confirmed by mass spectrometry, then buffer exchanged using a PD Minitrap G10 column.

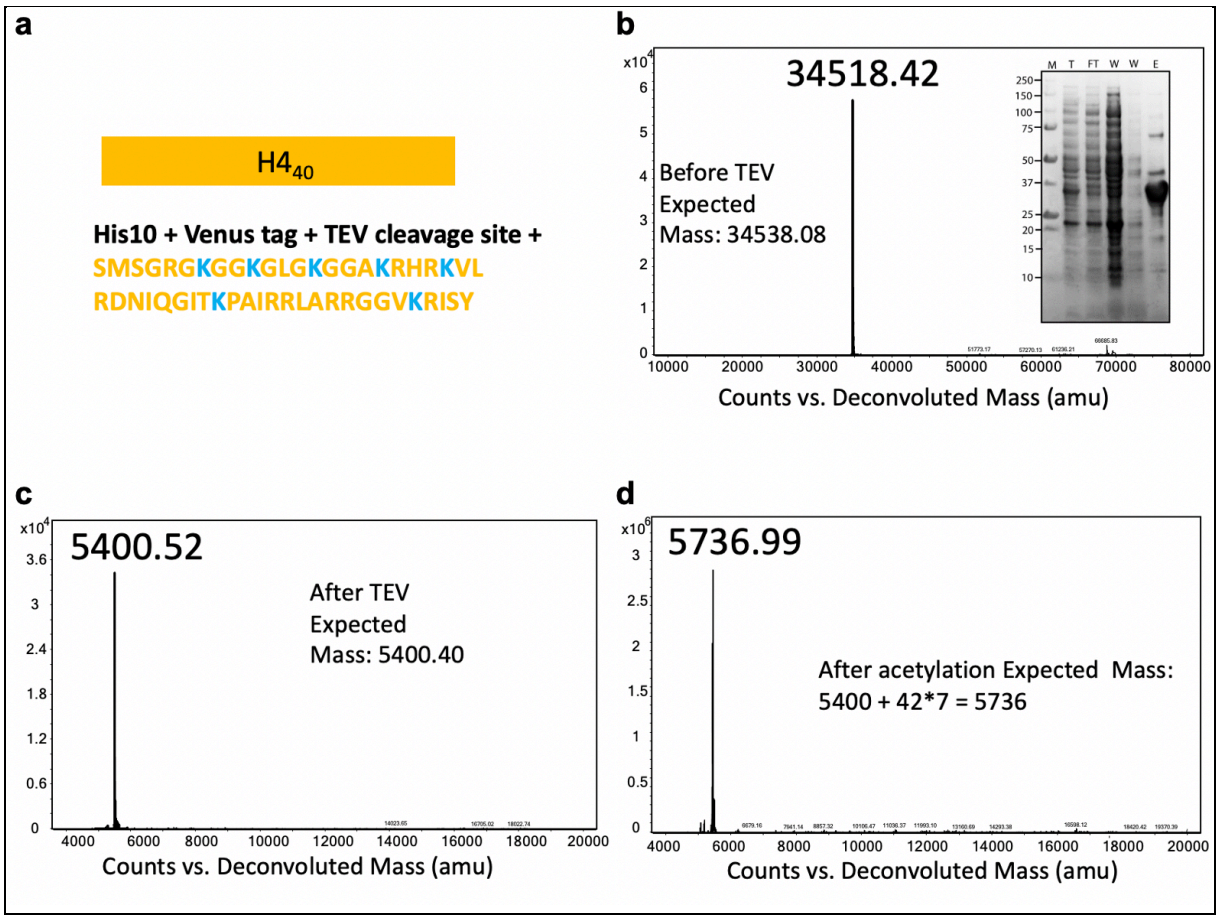


Figure 6.6. Purification and acetylation process of H4$_{50}$ (**a**) Sequence organisation of the peptide. H4$_{50}$ comprise of histone H4 sequence region which can be recognised by tandem BDs. **(b)** SDS-PAGE analysis of H4$_{50}$ after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{40}$ after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

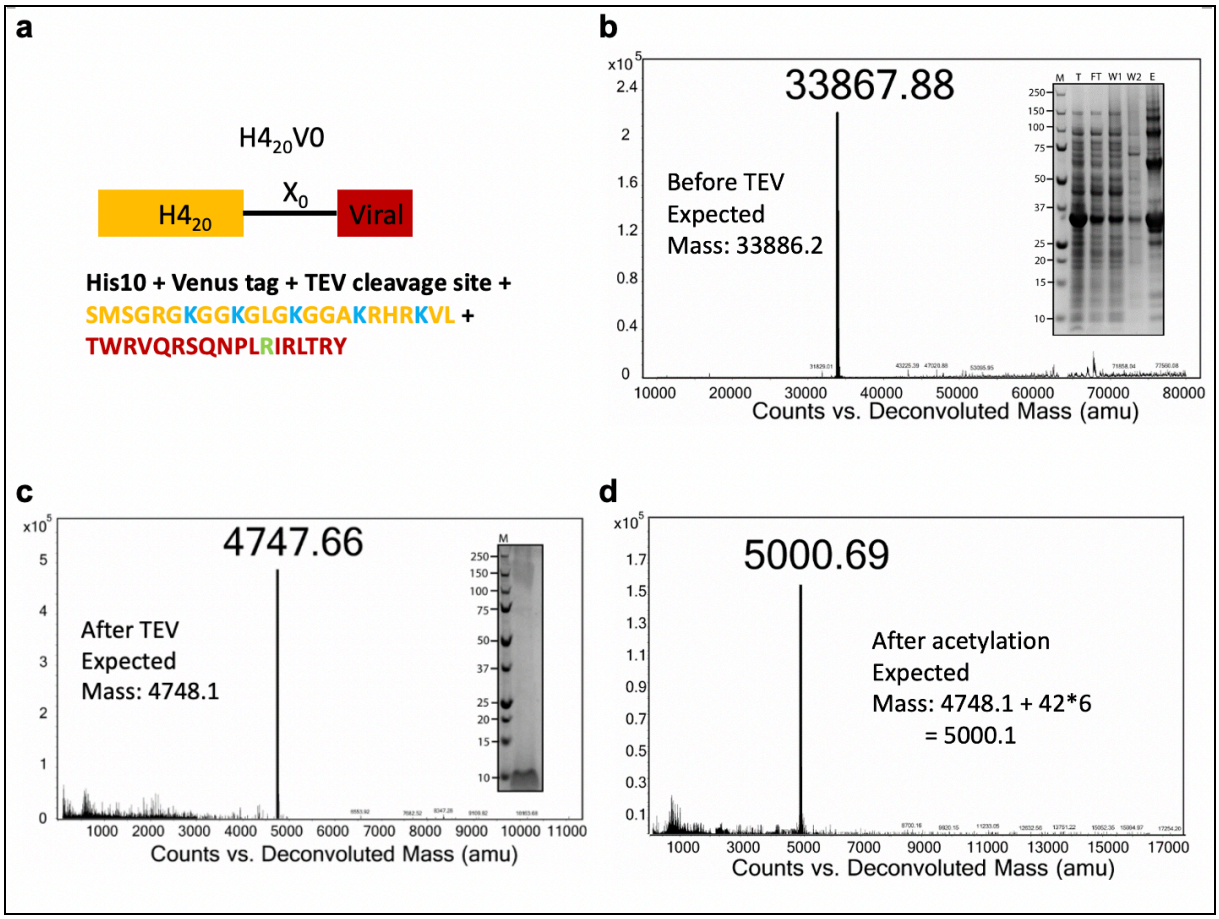


Figure 6.7. Purification and acetylation process of H4$_{20}$V0 (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{20}$V0 comprise of histone H4 sequence region which can be recognised by BD directly linked with the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{20}$V0 after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) SDS-PAGE analysis of H4$_{20}$V0 after TEV cleavage and mass spectrometry confirmation. (**d**) Mass spectrometry confirmation of acetylated peptide.

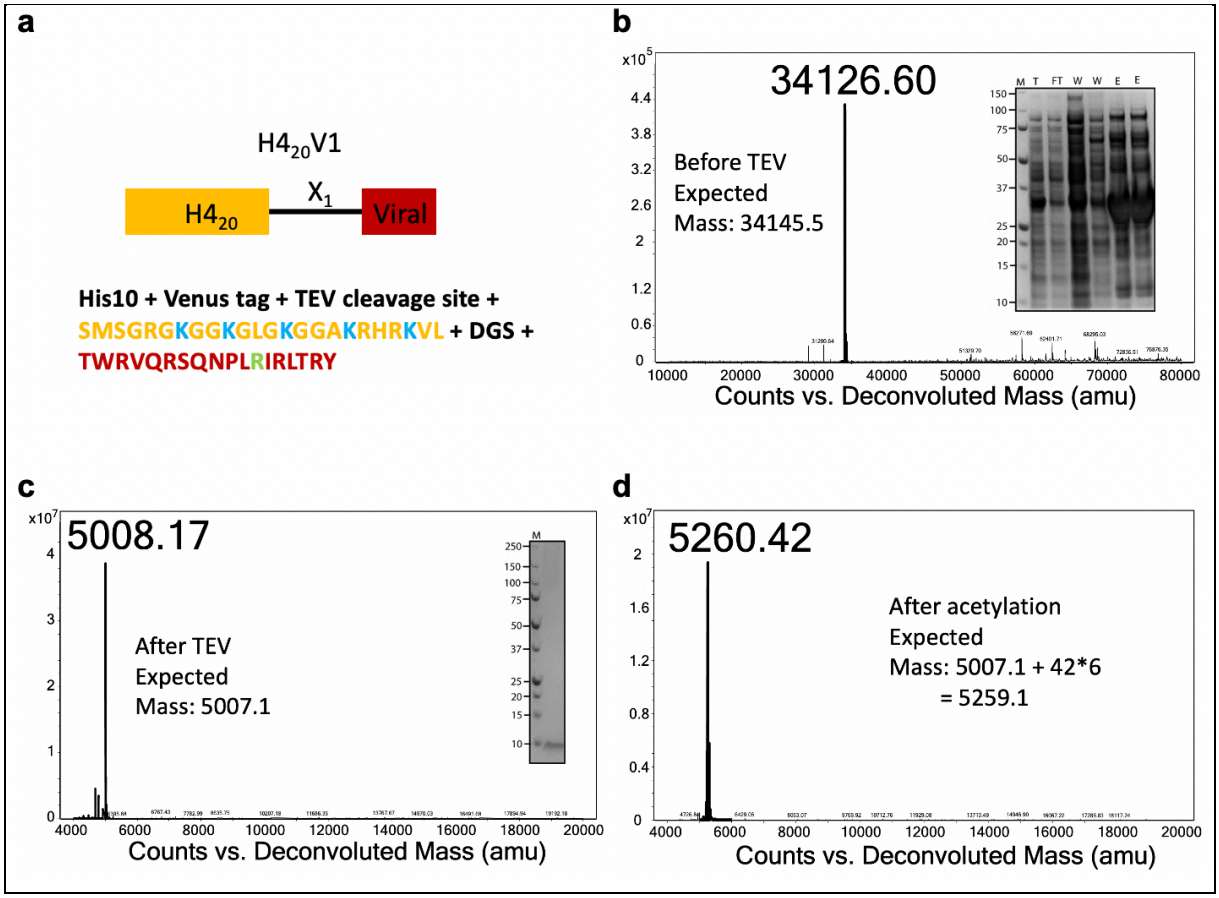


Figure 6.8. Purification and acetylation process of $H4_{20}V1$. (**a**) Sequence organisation of the bivalent peptidic inhibitor. $H4_{20}V1$ comprised of histone H4 sequence region which can be recognised by BD, linked by DGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of $H4_{20}V1$ after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) SDS-PAGE analysis of $H4_{20}V2$ after TEV cleavage and mass spectrometry confirmation. (**d**) Mass spectrometry confirmation of acetylated peptide.

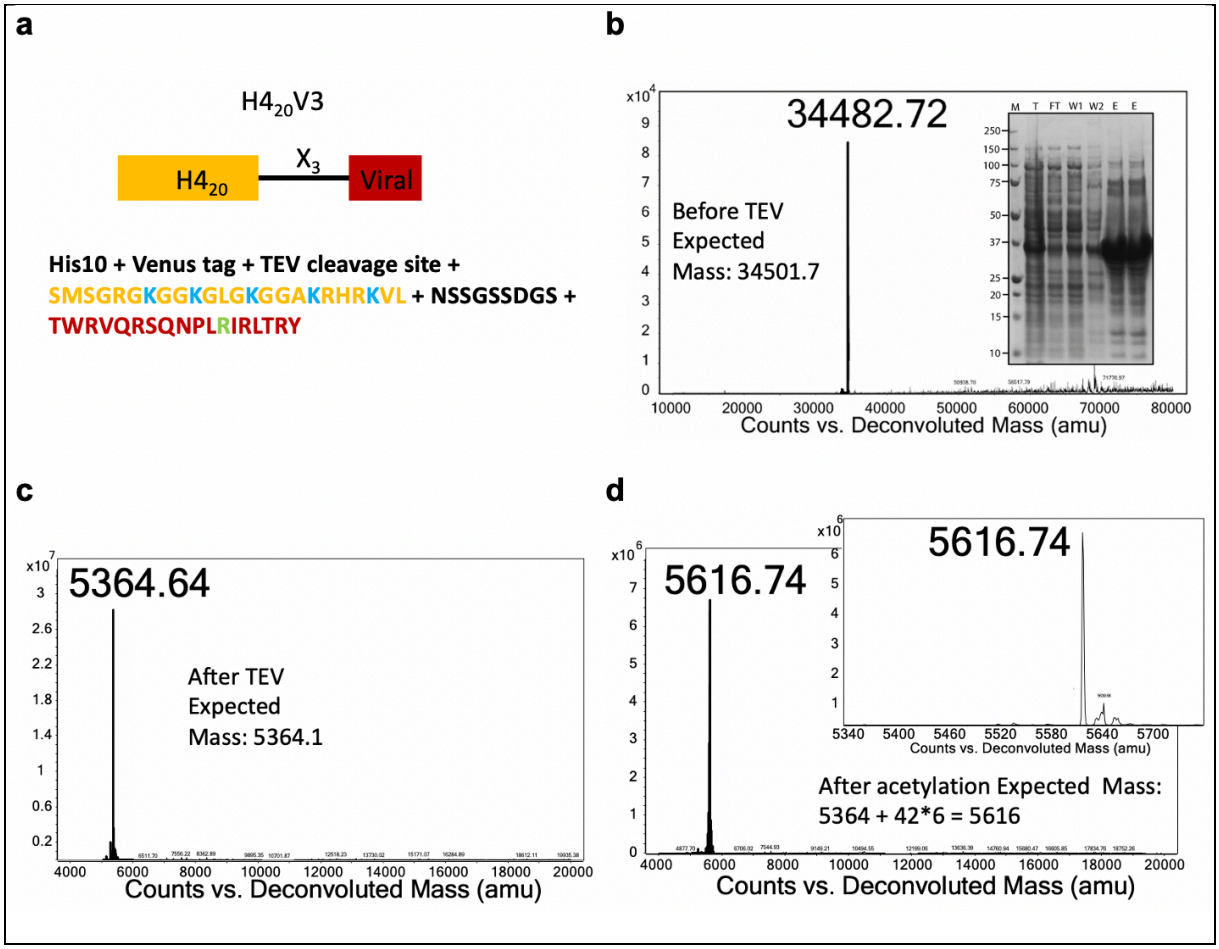


Figure 6.9. Purification and acetylation process of H4$_{20}$V3. (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{20}$V3 comprised of histone H4 sequence region which can be recognised by BD, linked by NSSGSSDGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{20}$V3 after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{20}$V3 after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

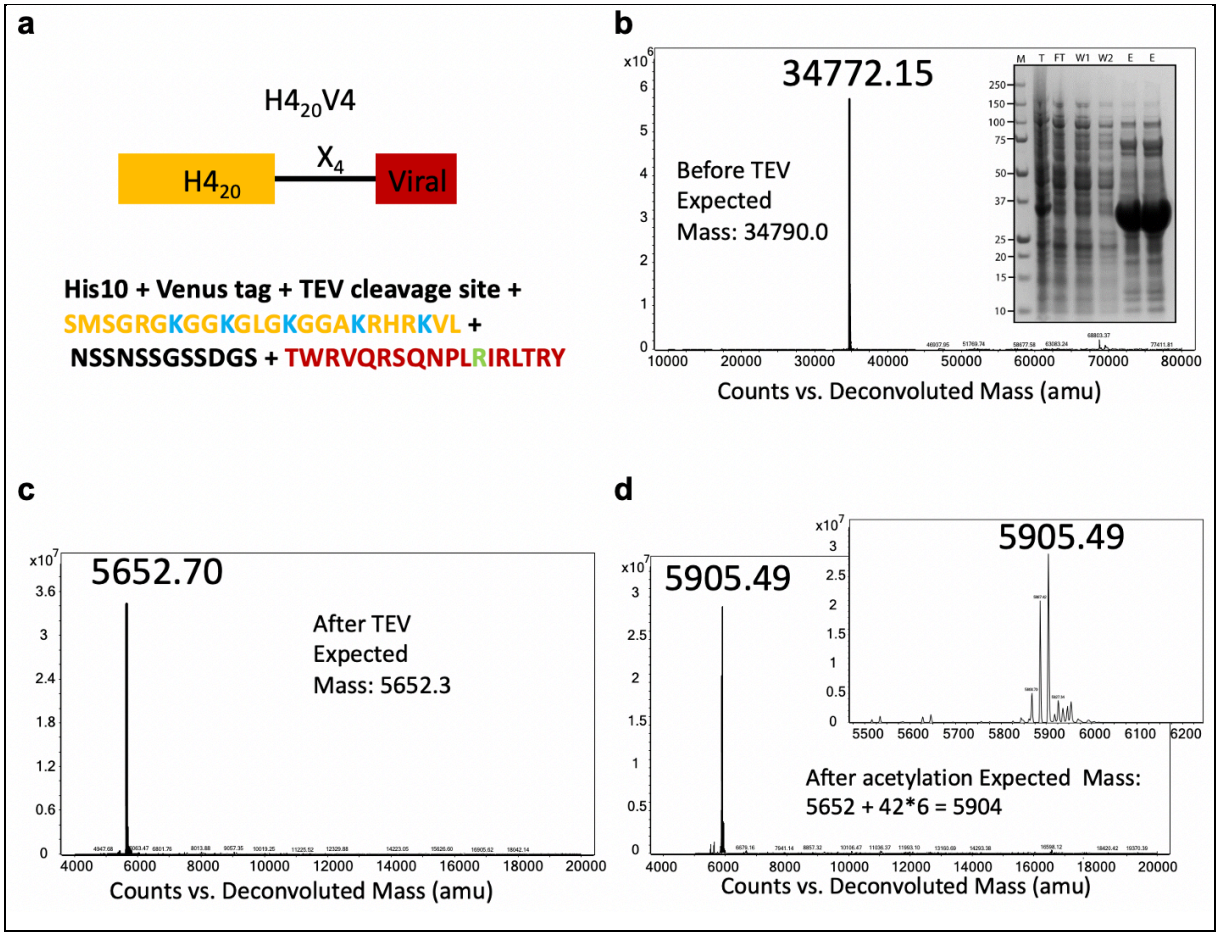


Figure 6.10. Purification and acetylation process of H4$_{20}$V4. (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{20}$V4 comprised of histone H4 sequence region which can be recognised by BD, linked by NSSNSSGSSDGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{20}$V4 after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{20}$V4 after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

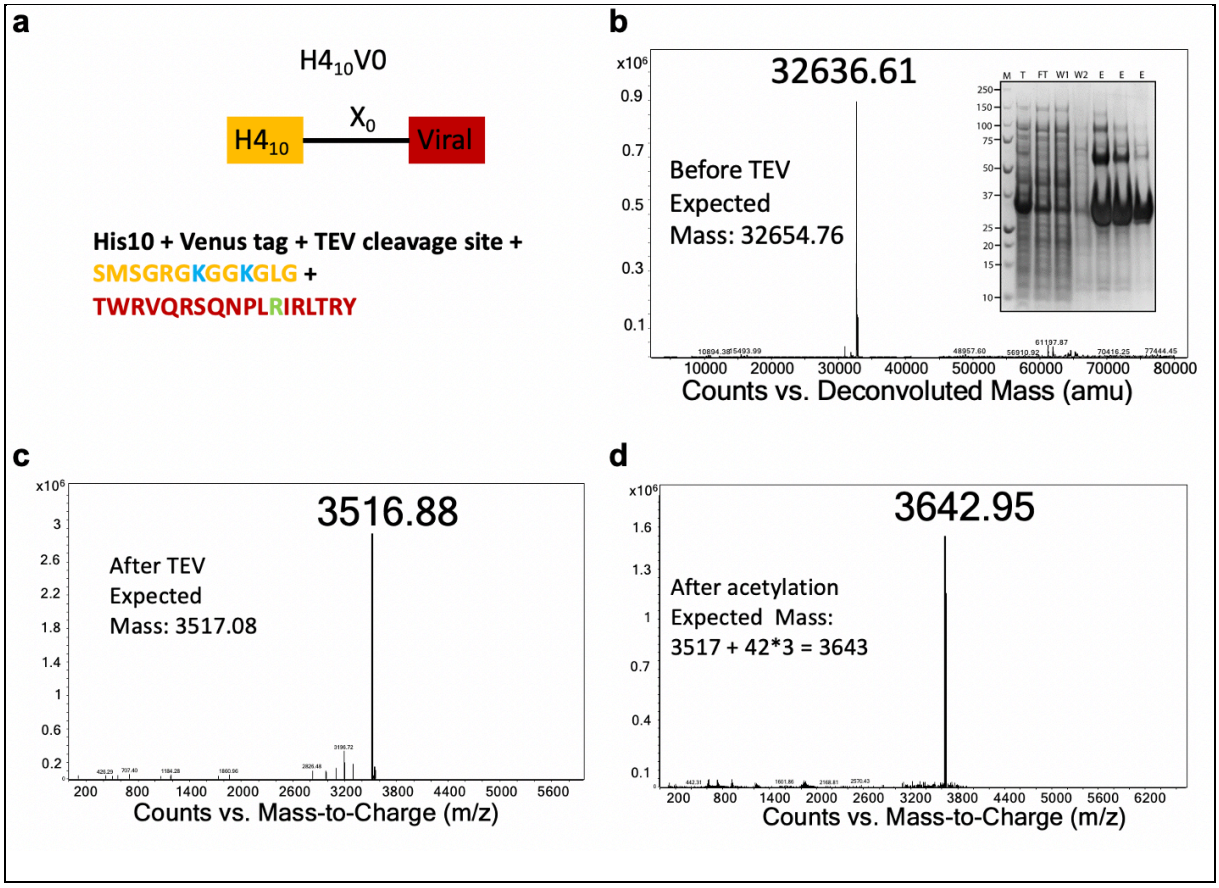


Figure 6.11. Purification and acetylation process of H4$_{10}$V0. (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{10}$V0 comprised of histone H4 sequence region which can be recognised by BD directly followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{10}$V2 after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{10}$V2 after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

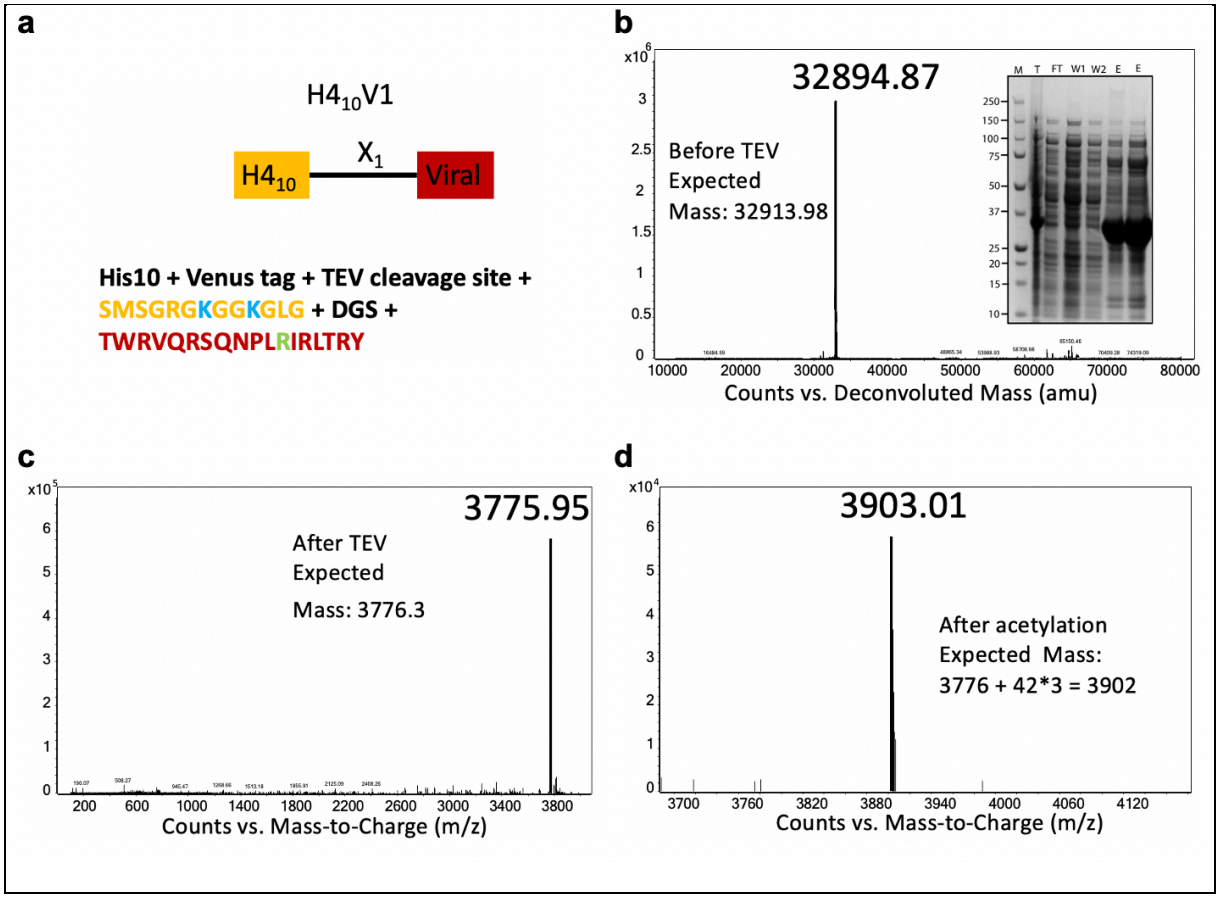


Figure 6.12. Purification and acetylation process for H4$_{10}$V1. (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{10}$V1 comprised of histone H4 sequence region which can be recognised by BD, linked by DGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{10}$V1 after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{20}$V2 after TEV cleavage. (**d**) Mass s spectrometry confirmation of acetylated peptide.

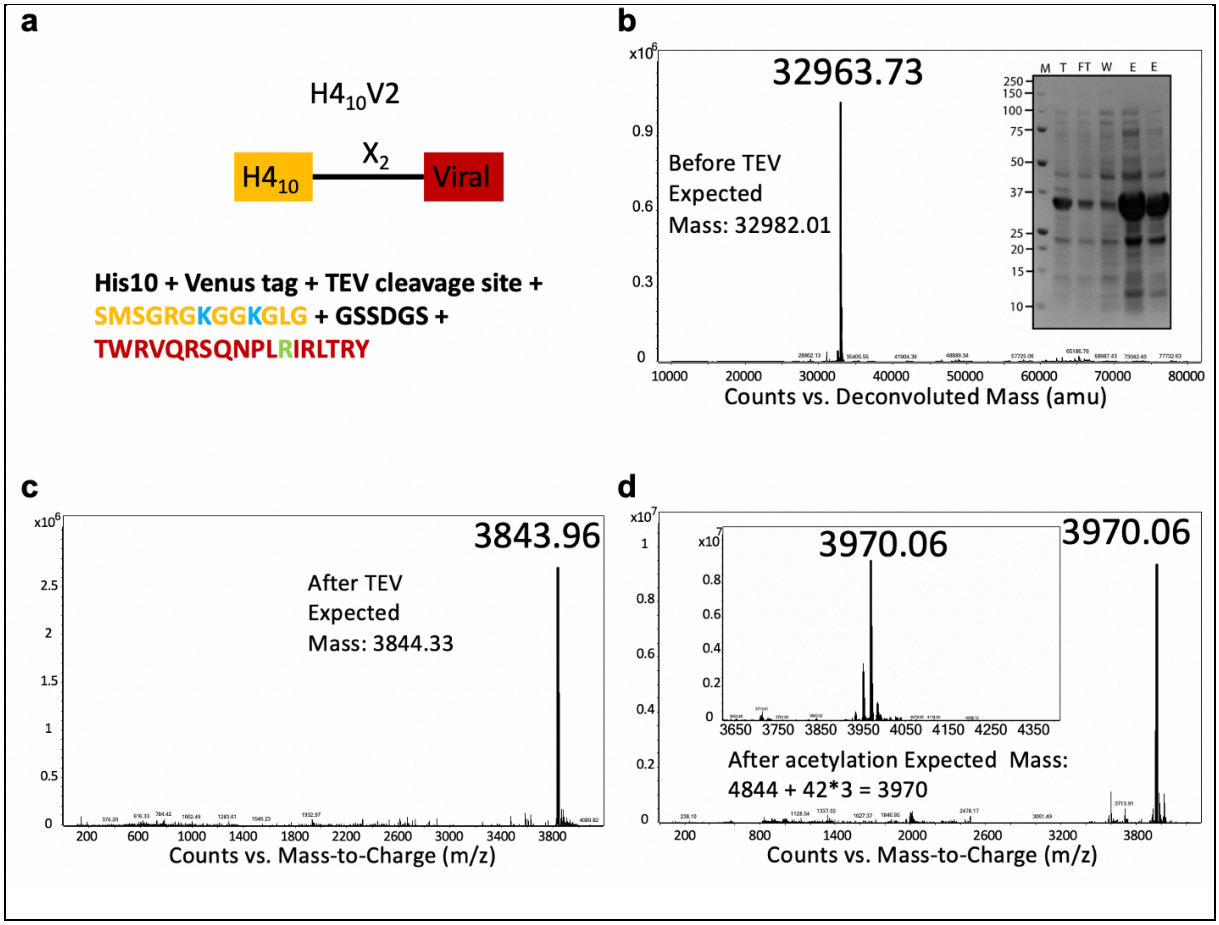


Figure 6.13. Purification and acetylation process for H4$_{10}$V2 (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{10}$V2 comprised of histone H4 sequence region which can be recognised by BD, linked by GSSDGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{10}$V2 after affinity chromatography purification. (NuPAGE™ Novex™ 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{10}$V2 after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

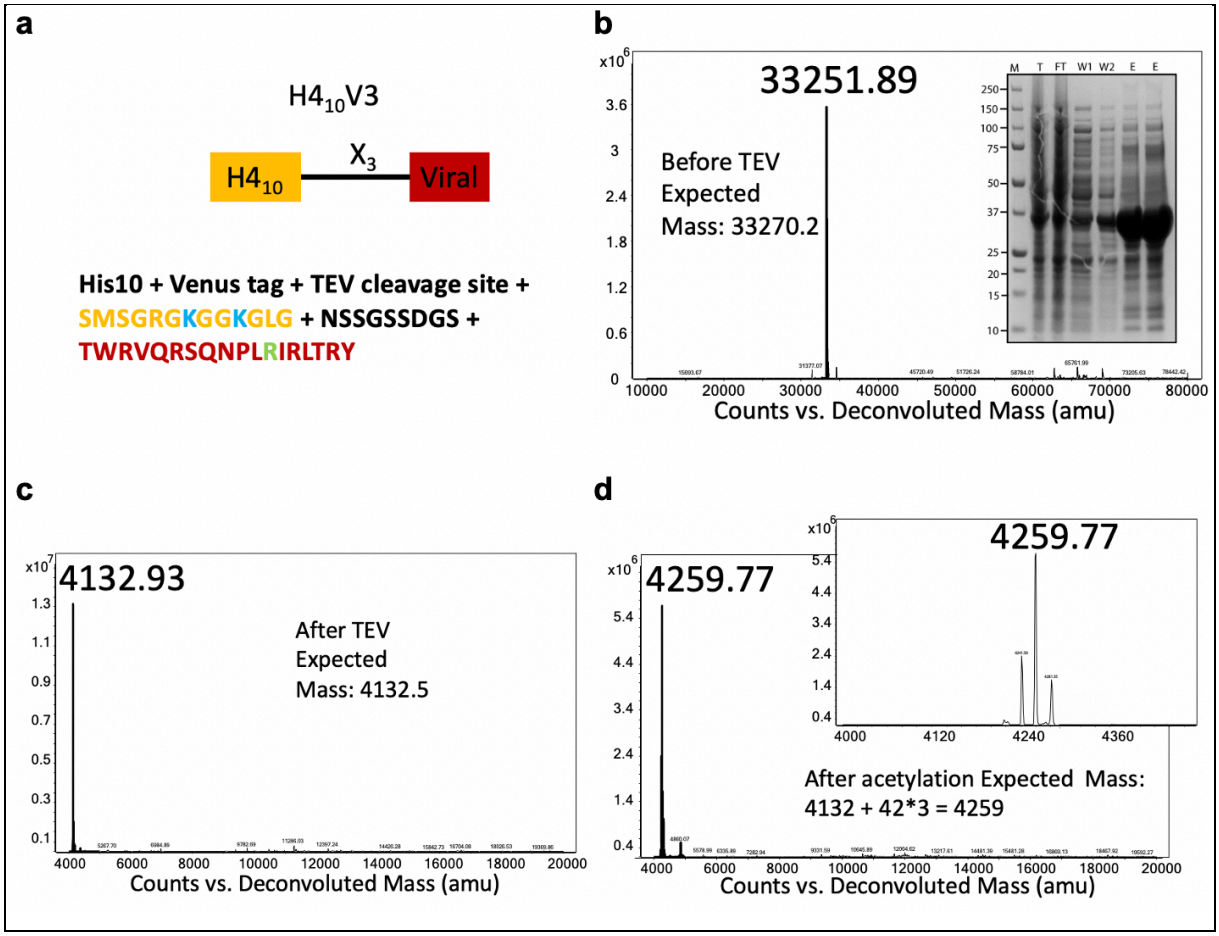


Figure 6.14. Purification and acetylation process for H4$_{20}$V3. (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{10}$V3 comprise of histone H4 sequence region which can be recognised by BD, linked by NSSGSSDGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{10}$V3 after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation H4$_{10}$V3 after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

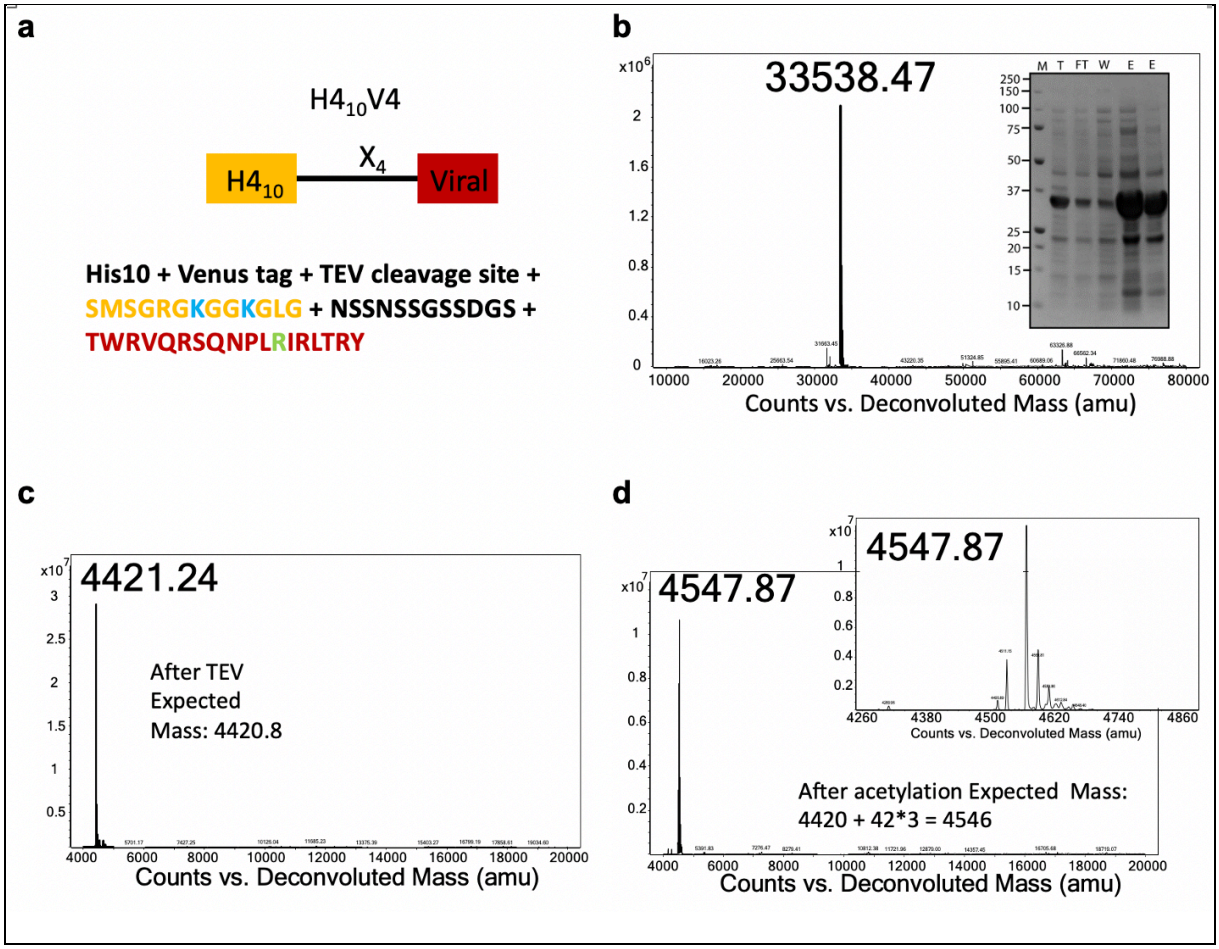


Figure 6.15. Purification and acetylation process for H4$_{10}$V4 (**a**) Sequence organisation of the bivalent peptidic inhibitor. H4$_{10}$V4 comprised of histone H4 sequence region which can be recognised by BD, linked by NSSNSSGSSDGS linker region, followed by the viral sequence which can be recognised by ET. **(b)** SDS-PAGE analysis of H4$_{10}$V4 after affinity chromatography purification. (NuPAGE^{TM} Novex^{TM} 4-12% Bis-Tris protein gel) and mass spectrometry confirmation. (**c**) Mass spectrometry confirmation of H4$_{10}$V4 after TEV cleavage. (**d**) Mass spectrometry confirmation of acetylated peptide.

6.3. Isothermal titration calorimetry

6.3.1. ITC experiments on BET tandem BDs with bivalent compounds (Chapter 2)

Experiments were performed on a VP-ITC titration microcalorimeter, with a cell volume of 1.4 mL and a 250 µL syringe. All experiments were conducted at 15 ºC with stirring rate of 295 rpm, in 50 mM HEPES, pH 7.5 and 100 mM NaCl buffer. The syringe was loaded with of protein sample (100 µM for tandem BET BDs, 200 µM for single BET BD) and was carefully inserted into the cell which was filled with 10 µM of bivalent compounds. A total of 35

titrations were set using an initial injection of 2 μL followed by 34 injections of 8 μL with a duration of 16 sec and spacing of 250 sec between injections. The heat of dilution for the proteins were corresponded to the heat observed from last injection, following saturation of ligand binding. The collected data were deconvoluted using the MicroCal Origin software to yield enthalpies of binding (H) and dissociation constants (K_D). Thermodynamic parameters were calculated using the basic equation of thermodynamics ($\Delta G = \Delta H - T\Delta S$, where ΔG, ΔH and ΔS are the changes in the free energy, enthalpy, and entropy of binding, respectively). In all cases a single binding site model, supplied with the MicroCal Origin software package was employed.

6.3.2. ITC experiments on BET tandem BDs with acetylated peptides (Chapter3)

Experiments were performed on a VP-ITC titration microcalorimeter, with a cell volume of 1.4 mL and a 250 μL syringe. All experiments were conducted at 15 °C with stirring rate of 295 rpm, in 50 mM HEPES, pH 7.5 and 100 mM NaCl buffer. The syringe was loaded with of peptide sample (500 - 1000 μM in ITC buffer) and was carefully inserted into the cell which was filled with 20 μM of proteins. A total of 35 titrations were set using an initial injection of 2 μL followed by 34 injections of 8 μL with duration of 16 sec and spacing of 250 sec between injections. The heat of dilution for the proteins were corresponded to the heat observed from last injection, following saturation of ligand binding. The collected data were deconvoluted using the MicroCal Origin software to yield enthalpies of binding (H) and dissociation constants (K_D). Thermodynamic parameters were calculated using the basic equation of thermodynamics ($\Delta G = \Delta H - T\Delta S$, where ΔG, ΔH and ΔS are the changes in the free energy, enthalpy, and entropy of binding, respectively). In all cases a single binding site model, supplied with the MicroCal Origin software package was employed.

6.3.3. ITC experiments on BET tandem BDs with tandem histone H4 – viral peptides (Chapter4)

Experiments were performed on a Nano-ITC titration microcalorimeter, with a cell volume of 1 mL and a 250 μL syringe. All experiments were conducted at 15 °C with stirring rate of 286 rpm, in 50 mM HEPES, pH 7.5 and 300 mM NaCl buffer. The syringe was loaded with of peptide sample (300 – 500 μM in ITC buffer) and was carefully inserted into the cell which was filled with 25 μM of proteins. A total of 31 titrations were set using an initial injection of 4 μL followed by 30 injections of 8 ul with duration of 16 sec and spacing of 240 sec between injections. The heat of dilution for the proteins were corresponded to the heat observed from last injection, following saturation of ligand binding. The collected data were deconvoluted using the Nano Analyze™ software to yield enthalpies of binding (H) and dissociation constants (K_D). Thermodynamic parameters were calculated using the basic equation of thermodynamics ($\Delta G = \Delta H - T\Delta S$, where ΔG, ΔH and ΔS are the changes in the free energy, enthalpy, and entropy of binding, respectively). In most of the cases a single binding site model, supplied with the Nano Analyze™ software package was employed.

6.4. Analytical ultracentrifugation

6.4.1. AUC experiments on BET tandem BDs with bivalent compounds (Chapter 2)

Conformation of the BET proteins with or without compounds were characterised by sedimentation velocity AUC. Samples of protein were initially analysed and buffer-exchanged on a Superdex 200 size exclusion chromatography (SEC) column which was equilibrated in 50 mM HEPES, pH 7.5 and 100 mM NaCl. This performed buffer exchange and allowed analysis of sample heterogeneity. All compounds were separately dissolved to a concentration of 1 mM in 100% DMSO. Protein samples were diluted in 50 mM HEPES, pH 7.5 and 100 mM NaCl to final concentrations of 8 – 13 μM. The final concentration of compounds was adjusted to compound: protein molar ratios of 2:1, 1:1, 0.5:1 and 0:1 by addition to the protein samples. A

400µL of each of the above-mentioned samples was loaded into two-sector peon centrepieces with sapphire glass and centrifuged at 40,000 rpm in a Beckman XL-I ultracentrifuge at 4 °C, 300 scans (280 nm) were collected until sedimentation of the protein was reached. The software package SEDNTERP was used to calculate the solvent density, viscosity, and the partial specific volume of the protein sample.

Data were analysed using the sedimentation distribution modelling software SEDFIT[202] whereby sedimentation coefficient distributions $c(s)$ were obtained by direct boundary modelling to Lamm Equation. In order to count as good fitting, for a signal of 1.0 OD, the generated rmsd should be below 0.01 OD and as low as possible. In our dataset, experiments done with BET tandem BDs all generated rmsd values lower than 0.01 OD. However, experiments done with BRD2(1) and BRD2(2) generated rmsd values higher than 0.01 (highest 0.016). It is mainly because these single domain experiments were added at a later time, and our AUC machine (Beckman XLI) at that time had some lamp problems producing noisier data. These single domain experiments should have been repeated on a machine with a better lamp quality.

6.4.2. AUC experiments on BET tandem BDs with acetylated peptides (Chapter3)

Conformation of the BET proteins with or without the peptides were characterised by sedimentation velocity AUC. All the experiments were conducted in the buffer condition of 50 mM HEPES, pH 7.5 and 100 mM NaCl. The final protein final concentration was at 10 - 15 µM and the peptides were added in 1:1 and 1:10 molar ratio. 400µL of each of the above-mentioned samples was loaded into two-sector peon centrepieces with sapphire glass and centrifuged at 40,000 rpm in a Beckman XL-I ultracentrifuge at 4 °C, and 300 scans (280 nm) were collected until sedimentation of the protein was reached. Data were analysed same as 6.4.1.

6.4.3. AUC experiments on BET tandem BDs with tandem histone H4–viral peptides (Chapter4)

Conformation of the BET proteins with or without the peptides were characterised by sedimentation velocity AUC. All the experiments were conducted in the buffer condition of 50 mM HEPES, pH 7.5 and 300 mM NaCl. The final protein final concentration was at 25 μM and the peptides were added in 1:1 molar ratio. 400μL of each of the above-mentioned samples was loaded into two-sector peon centrepieces with sapphire glass and centrifuged at 40,000 rpm in a Beckman Optima ultracentrifuge at 4 °C, and 300 scans (280 nm) were collected until sedimentation of the protein was reached. Data were analysed same as 6.4.1.

6.5. Small-angle X-ray scattering (SAXS)

For the orthogonal way of validating the conformation of BET proteins, the synchrotron radiation X-ray scattering data from solutions of BRD2 and BRDT construct containing two BDs were collected at beamline B21 at Diamond Light Source, Didcot, UK. The apo protein solutions and those containing test compounds were measured at concentrations of 1, 2, 5 and 10 mg/mL in buffer 20 mM HEPES pH 7.5 and 150 mM NaCl using a continuous flow cell capillary. The scattering of the buffer was subtracted, and the difference curves were scaled for protein concentration. The forward scattering I(0), the radius of gyration Rg, the distance distribution function of the particle *p(r)* and the maximum dimension Dmax were derived using the SAXS data analysis pipeline[231]. The *p(r)* functions were further recomputed manually with the program GNOM[232]. Ab initio shape models of the tandem constructs were generated using the bead modelling program DAMMIN[210]. Multiple runs were performed to generate 15–20 models. All samples were measured at concentrations of 1, 2, 5 and 10 mg/mL.

6.6. SPOT Array

Membrane peptides for the arrays were synthesised on cellulose membranes using a MultiPep SPOT peptide arrayer (INTAVIS) according to the method provided by the manufacturer. The array contains histone H4 covering all possible lysine acetylated sites and different combination, with a peptide size of 20 amino acid. Before the assay began, membranes were washed three times with 5% PBST for 5 minutes and blocked with 5% BSA for 8 hours in order to minimise nonspecific binding of the proteins to the membrane. After two washes with PBST followed by a wash with PBS, His-tagged proteins were added to final concentration of 1 μM and the membrane were incubated overnight at 4 °C in PBS. Next day, each membrane was washed three times in PBST, block with 5% BSA for 1 hour, and washed again three times with PBST. Secondary antibodies were added in 5% to PBST solution at a dilution of 1:2000, after one-hour incubation, membrane was washed three times 20 minutes in PBST. The assay was performed using hexa-His-tagged BRD2 recombinant tandem domains BRD2(1:2), BRD2(1:2)$^{WT/NF}$, BRD2(1:2)$^{NF/WT}$, BRD2(1:2)^{LT}, and BRD2(1:2)^{L4} in five identical membranes. These proteins were purified as the same way as previously described in Chapter 2 without the cleavage of His-tag. Proteins bound to peptides were detected using HPR-conjugated anti-His antibody (Novagene) and the Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific). Chemiluminescence was detected with an image reader using an exposure time of 5 min to 20 min. The intensity of each spot was detected and divided into four intensity degrees (25%, 50%, 75%, and 100%)

A1	PolyHis	D6	H3(13-32)-K23ac/K27ac
A2	PolyHis	D7	H4(1-20)-nonacetylated
A3	Blank	D8	H4(1-20)-K5ac/K8ac/K12ac/K16ac
A4	H3(1-20)-nonacetylated	D9	H4(1-20)-K5ac/K8ac/K12ac
A5	H3(1-20)-K4ac/K9ac/K14ac/K18ac	D10	H4(1-20)-K5ac/K8ac/K16ac
A6	H3(1-20)-K4ac/K9ac/K14ac	D11	H4(1-20)-K5ac/K12ac/K16ac

A7	H3(1-20)-K4ac/K9ac/K18ac	E1	H4(1-20)-K8ac/K12ac/K16ac
A8	H3(1-20)-K4ac/K14ac/K18ac	E2	H4(1-20)-K5ac/K8ac
A9	H3(1-20)-K9ac/K14ac/K18ac	E3	H4(1-20)-K5ac/K12ac
A10	H3(1-20)-K4ac/K9ac	E4	H4(1-20)-K5ac/K16ac
A11	H3(1-20)-K4ac/K14ac	E5	H4(1-20)-K8ac/K12ac
B1	Blank	E6	H4(1-20)-K8ac/K16ac
B2	Blank	E7	H4(1-20)-K12ac/K16ac
B3	H3(1-20)-K4ac /K18ac	E8	H4(5-24)-nonacetylated
B4	H3(1-20)-K9ac/K14ac	E9	H4(5-24)-K8ac/K12ac/K16ac/K20ac
B5	H3(1-20)-K9ac/K18ac	E10	H4(5-24)-K8ac/K16ac/K20ac
B6	H3(1-20)-K14ac/K18ac	E11	H4(5-24)-K8ac/K12ac/K20ac
B7	H3(4-23)-nonacetylated	F1	H4(5-24)-K8ac/K16ac/K20ac
B8	H3(4-23)-K9ac/K14ac/K18ac	F2	H4(5-24)-K12ac/K16ac/K20ac
B9	H3(4-23)-K14ac/K18ac	F3	H4(5-24)-K8ac/K12ac
B10	H3(4-23)-K9ac/K18ac	F4	H4(5-24)-K8ac/K16ac
B11	H3(4-23)-K14ac/K18ac	F5	H4(5-24)-K8ac/K20ac
C1	H3(11-30)-nonacetylated	F6	H4(5-24)-K12ac/K16ac
C2	H3(11-30)-K14ac/K18ac/K23ac/K27ac	F7	H4(5-24)-K12ac/K20ac
C3	H3(11-30)-K14ac/K18ac/K23ac	F8	H4(5-24)-K16ac/K20ac
C4	H3(11-30)-K14ac/K18ac/K27ac	F9	H4(7-26)-nonacetylated
C5	H3(11-30)-K14ac/ K23ac/K27ac	F10	H4(7-26)-K12ac/K16ac/K20ac
C6	H3(11-30)-K18ac/K23ac/K27ac	F11	H4(7-26)-K12ac/K16ac
C7	H3(11-30)-K14ac/K18ac	G1	H4(7-26)-K12ac/K20ac
C8	H3(11-30)-K14ac/K23ac	G2	H4(7-26)-K16ac/K20ac
C9	H3(11-30)-K14ac/K27ac	G3	H4(9-28)-nonacetylated
C10	H3(11-30)-K18ac/K23ac	G4	H4(9-28)-K16ac/K20ac
C11	H3(11-30)-K18ac/K23ac/K27ac	G5	H4(16-35)-nonacetylated
D1	H3(11-30)-K23ac/K27ac	G6	H4(16-35)-K20ac/K31ac
D2	H3(13-32)-nonacetylated	G7	Blank
D3	H3(13-32)-K18ac/K23ac/K27ac	G8	Blank
D4	H3(13-32)-K18ac/K23ac	G9	Blank
D5	H3(13-32)-K18ac/K27ac	G10	Blank

Table 6.3. SPOT membrane layout, positions and corresponding peptides.

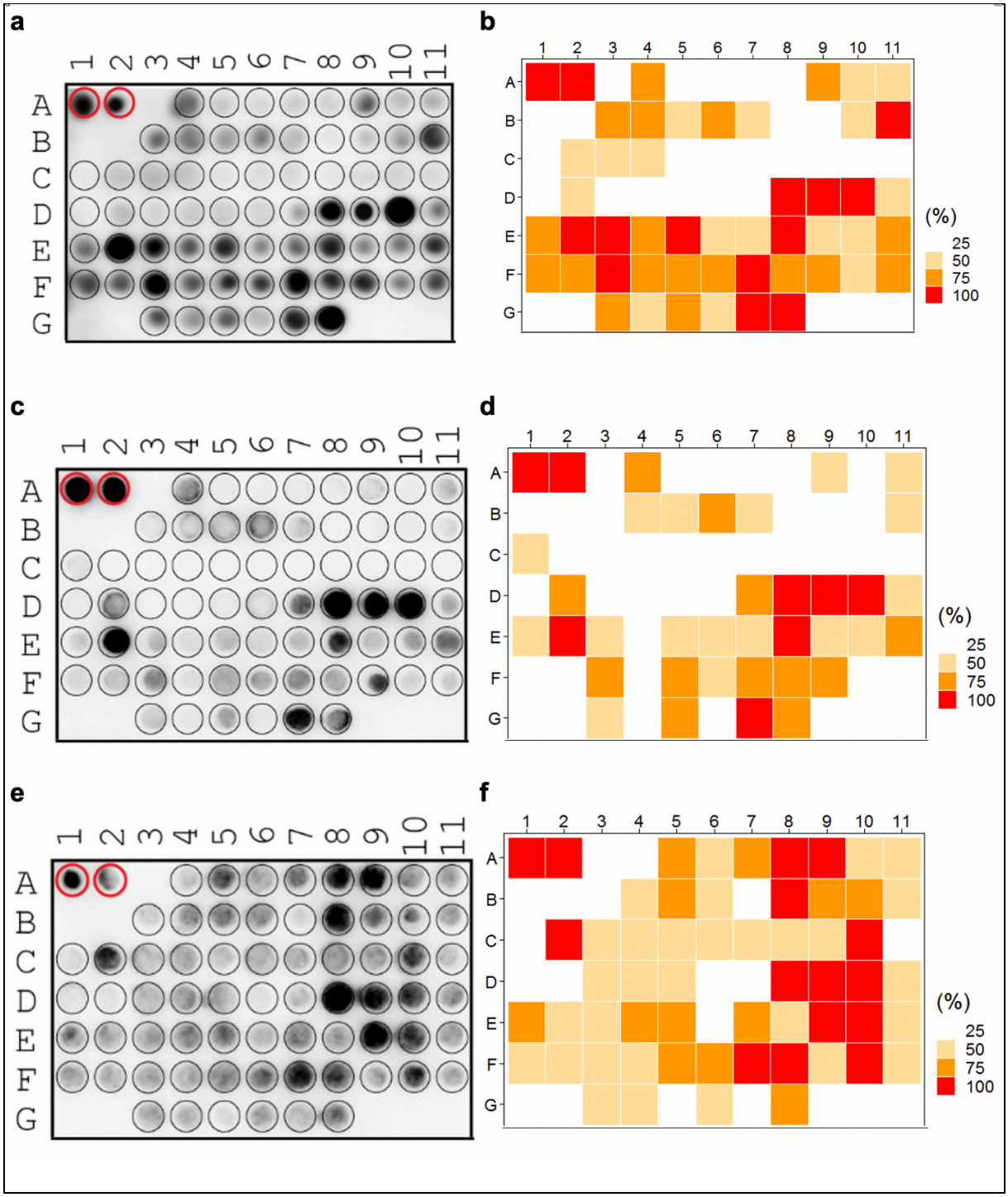


Figure 6.16. Spot array of multiple Kac peptide with tandem bromodomains related to Figure 3.1. Membranes used for probing the interaction between BRD2 tandem BD and multiple acetylated histone peptides (**a**) BRD2(1:2) probed membrane. (**b**) BRD2(1:2) corresponding spot intensity heatmap. (**c**) BRD2(1:2)$^{WT/NF}$ probed membrane. (**d**) BRD2(1:2)$^{WT/NF}$ corresponding spot intensity heatmap. (**e**) BRD2(1:2)$^{NF/WT}$ probed membrane. (**f**) BRD2(1:2) $^{NF/WT}$ corresponding spot intensity heatmap

6.7. Biolayer interferometry

Experiments were performed on an Octet RED384 System (FortéBio) using commercial libraries (AltaBioSciences) at 25 °C in 50 mM HEPES (pH 7.5), 100 mM NaCl, and 0.5 mM tris using the FortéBio data acquisition software V.7.1.0.100. Biotinylated histone H4 peptides were immobilised onto Super Streptavidin biosensors, preequilibrated in the BLI buffer, and quenched in a solution of 5 mM biotin (baseline equilibration for 60 s, peptide loading for 240 s, and quenching for 60 s; shake speed of 1000 rpm at 25 °C). The immobilised peptides were subsequently used in association and dissociation measurements performed within a time window of 600 s (baseline equilibration for 60 s, association for 600 s, and dissociation for 600 s; shake speed of 1000 rpm at 25 °C) with 5 μM concentration of BRD2(1:2), BRD2(1:2)$^{WT/NF}$, BRD2(1:2)$^{NF/WT}$, BRD2(1:2)^{LT}, and BRD2(1:2)^{L4}. Interference patterns from peptide-coated biosensors without peptide were used as controls. After referencing corrections, the subtracted binding interference data were analysed using the FortéBio data analysis software provided with the instrument following the manufacturer's protocols. Response of each interaction were divided into four degrees of intensity. (0.2-0.4, 0.4-0.8, 0.8-1.2, and 1.2-1.8).

The histone H4 sequence used in this assay with different combination of lysine acetylation: SGRGKGGKGLGKGGAKRHRKV

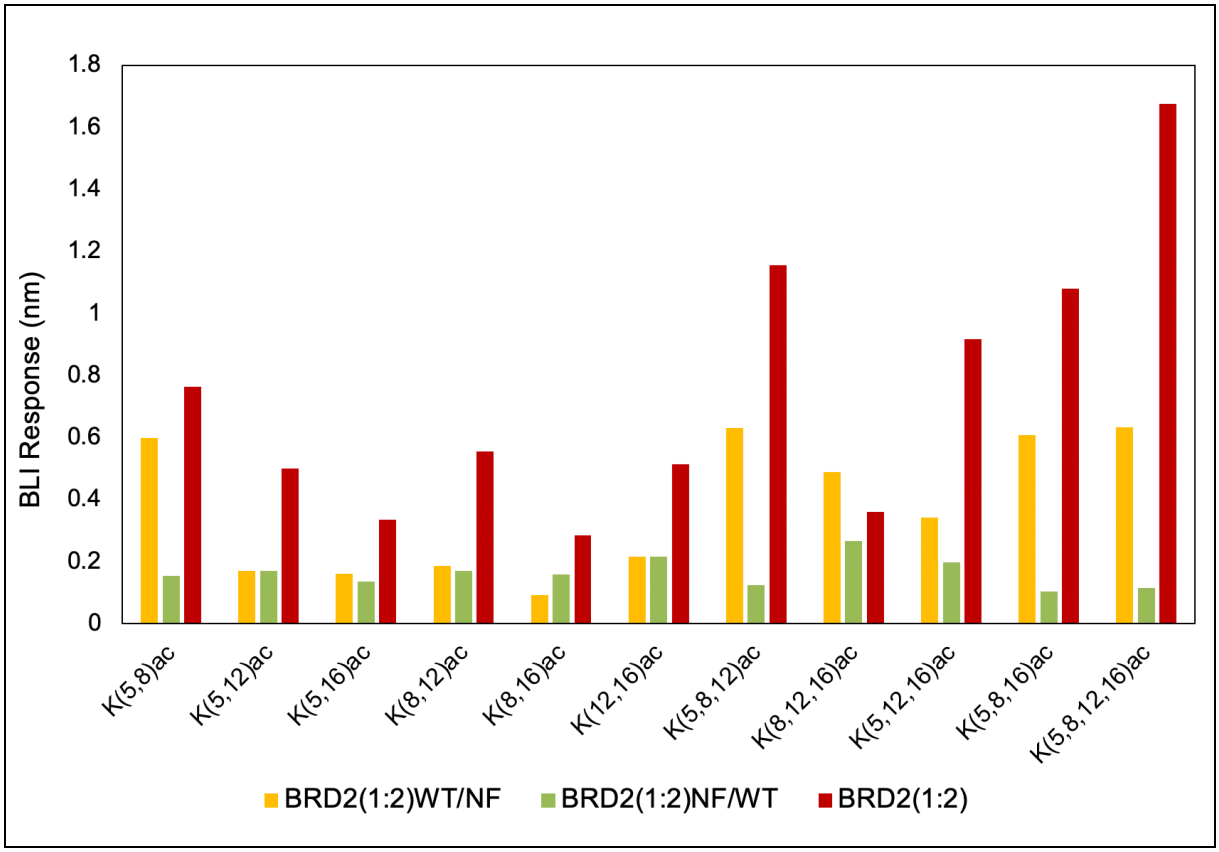


Figure 6.17. BLI response of interaction between BRD2 tandem BDs and histone H4 peptide.

6.8. Custom Peptide Synthesis

Double acetylated histone H4 peptides K5acK12ac and K5acK16ac used in AUC and ITC assays were synthesised by the TUFTS Core facility, on a 0.1 mmol scale with one round of HPLC purification. Before use, peptides were re-suspended into buffer 50 mM HEPES pH 7.5 and 100 mM NaCl based on buffer consistency with the experiments. Peptides then further purified using PD MiniTrap G-10 columns (GE Healthcare Life sciences), to remove remaining chemical residuals.

Reference.

1. Holliday, R. The inheritance of epigenetic defects. *Science* **238**, 163 (1987).
2. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).
3. Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-60 (1997).
4. Jenuwein, T. & Allis, C.D. Translating the histone code. *Science* **293**, 1074-80 (2001).
5. Kumar, S., Chinnusamy, V. & Mohapatra, T. Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. *Front Genet* **9**, 640 (2018).
6. Greenberg, M.V.C. & Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* **20**, 590-607 (2019).
7. Choudhary, C. et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-40 (2009).
8. Marushige, K. Activation of chromatin by acetylation of histone side chains. *Proc Natl Acad Sci U S A* **73**, 3937-41 (1976).
9. Kouzarides, T. Acetylation: a regulatory modification to rival phosphorylation? *Embo j* **19**, 1176-9 (2000).
10. Celic, I. et al. The sirtuins hst3 and Hst4p preserve genome integrity by controlling histone h3 lysine 56 deacetylation. *Curr Biol* **16**, 1280-9 (2006).
11. Shogren-Knaak, M. et al. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844-7 (2006).
12. Bolden, J.E., Peart, M.J. & Johnstone, R.W. Anticancer activities of histone deacetylase inhibitors. *Nat Rev Drug Discov* **5**, 769-84 (2006).
13. Sanaei, M. & Kavoosi, F. Histone Deacetylases and Histone Deacetylase Inhibitors: Molecular Mechanisms of Action in Various Cancers. *Adv Biomed Res* **8**, 63 (2019).
14. Garber, K. HDAC inhibitors overcome first hurdle. *Nat Biotechnol* **25**, 17-9 (2007).
15. Kazantsev, A.G. & Thompson, L.M. Therapeutic application of histone deacetylase inhibitors for central nervous system disorders. *Nat Rev Drug Discov* **7**, 854-68 (2008).
16. Thurn, K.T., Thomas, S., Moore, A. & Munster, P.N. Rational therapeutic combinations with histone deacetylase inhibitors for the treatment of cancer. *Future Oncol* **7**, 263-83 (2011).
17. Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067-73 (2010).
18. Dawson, M.A. et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature* **478**, 529-33 (2011).
19. Gilan, O. et al. Selective targeting of BD1 and BD2 of the BET proteins in cancer and immunoinflammation. *Science* (2020).
20. Nicodeme, E. et al. Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119-23 (2010).
21. Wu, Q. et al. A chemical toolbox for the study of bromodomains and epigenetic signaling. *Nat Commun* **10**, 1915 (2019).
22. Tamkun, J.W. et al. brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68**, 561-72 (1992).
23. Fairbridge, N.A. et al. Cecr2 mutations causing exencephaly trigger misregulation of mesenchymal/ectodermal transcription factors. *Birth Defects Res A Clin Mol Teratol* **88**, 619-25 (2010).

24. Dhalluin, C. et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-6 (1999).
25. Cavellán, E., Asp, P., Percipalle, P. & Farrants, A.K. The WSTF-SNF2h chromatin remodeling complex interacts with several nuclear proteins in transcription. *J Biol Chem* **281**, 16264-71 (2006).
26. Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H. & Nakatani, Y. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**, 319-24 (1996).
27. Filippakopoulos, P. et al. Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214-31 (2012).
28. Zeng, L., Zhang, Q., Gerona-Navarro, G., Moshkina, N. & Zhou, M.M. Structural basis of site-specific histone recognition by the bromodomains of human coactivators PCAF and CBP/p300. *Structure* **16**, 643-52 (2008).
29. Kalkhoven, E. CBP and p300: HATs for different occasions. *Biochem Pharmacol* **68**, 1145-55 (2004).
30. Cai, Y. et al. Identification of new subunits of the multiprotein mammalian TRRAP/TIP60-containing histone acetyltransferase complex. *J Biol Chem* **278**, 42733-6 (2003).
31. Podcheko, A. et al. Identification of a WD40 repeat-containing isoform of PHIP as a novel regulator of beta-cell growth and survival. *Mol Cell Biol* **27**, 6484-96 (2007).
32. Müller, P., Kuttenkeuler, D., Gesellchen, V., Zeidler, M.P. & Boutros, M. Identification of JAK/STAT signalling components by genome-wide RNA interference. *Nature* **436**, 871-5 (2005).
33. Kouskouti, A. & Talianidis, I. Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *Embo j* **24**, 347-57 (2005).
34. Mujtaba, S. et al. Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. *Mol Cell* **13**, 251-63 (2004).
35. Ullah, M. et al. Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes. *Mol Cell Biol* **28**, 6828-43 (2008).
36. Kaeser, M.D., Aslanian, A., Dong, M.Q., Yates, J.R., 3rd & Emerson, B.M. BRD7, a novel PBAF-specific SWI/SNF subunit, is required for target gene activation and repression in embryonic stem cells. *J Biol Chem* **283**, 32254-63 (2008).
37. Laue, K. et al. The multidomain protein Brpf1 binds histones and is required for Hox gene expression and segmental identity. *Development* **135**, 1935-46 (2008).
38. Ciró, M. et al. ATAD2 is a novel cofactor for MYC, overexpressed and amplified in aggressive tumors. *Cancer Res* **69**, 8491-8 (2009).
39. Tsai, W.W. et al. TRIM24 links a non-canonical histone signature to breast cancer. *Nature* **468**, 927-32 (2010).
40. Khetchoumian, K. et al. TIF1delta, a novel HP1-interacting member of the transcriptional intermediary factor 1 (TIF1) family expressed by elongating spermatids. *J Biol Chem* **279**, 48329-41 (2004).
41. Zhou, Y. et al. Reversible acetylation of the chromatin remodelling complex NoRC is required for non-coding RNA-dependent silencing. *Nat Cell Biol* **11**, 1010-6 (2009).
42. Jones, M.H., Hamana, N., Nezu, J. & Shimane, M. A novel family of bromodomain genes. *Genomics* **63**, 40-5 (2000).
43. Yordy, J.S. et al. SP100 inhibits ETS1 activity in primary endothelial cells. *Oncogene* **24**, 916-31 (2005).
44. Zong, R.T., Das, C. & Tucker, P.W. Regulation of matrix attachment region-dependent, lymphocyte-restricted transcription through differential localization within promyelocytic leukemia nuclear bodies. *Embo j* **19**, 4123-33 (2000).
45. Dou, Y. et al. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. *Cell* **121**, 873-85 (2005).

46. Rowe, H.M. et al. KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* **463**, 237-40 (2010).
47. Masselink, H. & Bernards, R. The adenovirus E1A binding protein BS69 is a corepressor of transcription through recruitment of N-CoR. *Oncogene* **19**, 1538-46 (2000).
48. Wassarman, D.A. & Sauer, F. TAF(II)250: a transcription toolbox. *J Cell Sci* **114**, 2895-902 (2001).
49. Wang, P.J. & Page, D.C. Functional substitution for TAF(II)250 by a retroposed homolog that is expressed in human spermatogenesis. *Hum Mol Genet* **11**, 2341-6 (2002).
50. Huang, H., Rambaldi, I., Daniels, E. & Featherstone, M. Expression of the Wdr9 gene and protein products during mouse development. *Dev Dyn* **227**, 608-14 (2003).
51. Jacobson, R.H., Ladurner, A.G., King, D.S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**, 1422-5 (2000).
52. Li, A.G. et al. An acetylation switch in p53 mediates holo-TFIID recruitment. *Mol Cell* **28**, 408-21 (2007).
53. Gregory, G.D. et al. Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. *Mol Cell Biol* **27**, 8466-79 (2007).
54. Harikrishnan, K.N. et al. Brahma links the SWI/SNF chromatin-remodeling complex with MeCP2-dependent transcriptional silencing. *Nat Genet* **37**, 254-64 (2005).
55. Rada-Iglesias, A. et al. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279-83 (2011).
56. Xue, Y. et al. The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. *Proc Natl Acad Sci U S A* **97**, 13015-20 (2000).
57. Filippakopoulos, P. & Knapp, S. The bromodomain interaction module. *FEBS Lett* **586**, 2692-704 (2012).
58. Moriniere, J. et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**, 664-8 (2009).
59. Umehara, T. et al. Structural basis for acetylated histone H4 recognition by the human BRD2 bromodomain. *J Biol Chem* **285**, 7610-8 (2010).
60. Huang, B., Yang, X.D., Zhou, M.M., Ozato, K. & Chen, L.F. Brd4 coactivates transcriptional activation of NF-kappaB via specific binding to acetylated RelA. *Mol Cell Biol* **29**, 1375-87 (2009).
61. Zou, Z. et al. Brd4 maintains constitutively active NF-κB in cancer cells by binding to acetylated RelA. *Oncogene* **33**, 2395-404 (2014).
62. Lamonica, J.M. et al. Bromodomain protein Brd3 associates with acetylated GATA1 to promote its chromatin occupancy at erythroid target genes. *Proc Natl Acad Sci U S A* **108**, E159-68 (2011).
63. Gamsjaeger, R. et al. Structural basis and specificity of acetylated transcription factor GATA1 recognition by BET family bromodomain protein Brd3. *Mol Cell Biol* **31**, 2632-40 (2011).
64. Shi, J. et al. Disrupting the interaction of BRD4 with diacetylated Twist suppresses tumorigenesis in basal-like breast cancer. *Cancer Cell* **25**, 210-25 (2014).
65. Schröder, S. et al. Two-pronged binding with bromodomain-containing protein 4 liberates positive transcription elongation factor b from inactive ribonucleoprotein complexes. *J Biol Chem* **287**, 1090-9 (2012).
66. Rahman, S. et al. The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol Cell Biol* **31**, 2641-52 (2011).
67. Shen, C. et al. NSD3-Short Is an Adaptor Protein that Couples BRD4 to the CHD8 Chromatin Remodeler. *Mol Cell* **60**, 847-59 (2015).

68. Luna-Peláez, N. et al. The Cornelia de Lange Syndrome-associated factor NIPBL interacts with BRD4 ET domain for transcription control of a common set of genes. *Cell Death & Disease* **10**, 548 (2019).
69. You, J. et al. Kaposi's sarcoma-associated herpesvirus latency-associated nuclear antigen interacts with bromodomain protein Brd4 on host mitotic chromosomes. *J Virol* **80**, 8909-19 (2006).
70. Aiyer, S. et al. Altering murine leukemia virus integration through disruption of the integrase and BET protein family interaction. *Nucleic Acids Res* **42**, 5917-28 (2014).
71. Larue, R.C. et al. Bimodal high-affinity association of Brd4 with murine leukemia virus integrase and mononucleosomes. *Nucleic Acids Res* **42**, 4868-81 (2014).
72. Lambert, J.P. et al. Interactome Rewiring Following Pharmacological Targeting of BET Bromodomains. *Mol Cell* **73**, 621-638.e17 (2019).
73. French, C.A. et al. NSD3-NUT fusion oncoprotein in NUT midline carcinoma: implications for a novel oncogenic mechanism. *Cancer Discov* **4**, 928-41 (2014).
74. Zhang, Q. et al. Structural Mechanism of Transcriptional Regulator NSD3 Recognition by the ET Domain of BRD4. *Structure* **24**, 1201-8 (2016).
75. Crowe, B.L. et al. Structure of the Brd4 ET domain bound to a C-terminal motif from gamma-retroviral integrases reveals a conserved mechanism of interaction. *Proc Natl Acad Sci U S A* **113**, 2086-91 (2016).
76. Abbate, E.A., Voitenleitner, C. & Botchan, M.R. Structure of the papillomavirus DNA-tethering complex E2:Brd4 and a peptide that ablates HPV chromosomal association. *Mol Cell* **24**, 877-89 (2006).
77. You, J., Croyle, J.L., Nishimura, A., Ozato, K. & Howley, P.M. Interaction of the bovine papillomavirus E2 protein with Brd4 tethers the viral DNA to host mitotic chromosomes. *Cell* **117**, 349-60 (2004).
78. Wu, S.Y. et al. Brd4 links chromatin targeting to HPV transcriptional silencing. *Genes Dev* **20**, 2383-96 (2006).
79. Basile, W., Salvatore, M., Bassot, C. & Elofsson, A. Why do eukaryotic proteins contain more intrinsically disordered regions? *PLoS Comput Biol* **15**, e1007186 (2019).
80. Mao, A.H., Crick, S.L., Vitalis, A., Chicoine, C.L. & Pappu, R.V. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci U S A* **107**, 8183-8 (2010).
81. Ebert, M.O., Bae, S.H., Dyson, H.J. & Wright, P.E. NMR relaxation study of the complex formed between CBP and the activation domain of the nuclear hormone receptor coactivator ACTR. *Biochemistry* **47**, 1299-308 (2008).
82. Demarest, S.J. et al. Mutual synergistic folding in recruitment of CBP/p300 by p160 nuclear receptor coactivators. *Nature* **415**, 549-53 (2002).
83. Linding, R., Russell, R.B., Neduva, V. & Gibson, T.J. GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res* **31**, 3701-8 (2003).
84. Linding, R. et al. Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453-9 (2003).
85. Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F. & Jones, D.T. The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138-9 (2004).
86. Belkina, A.C. & Denis, G.V. BET domain co-regulators in obesity, inflammation and cancer. *Nat Rev Cancer* **12**, 465-77 (2012).
87. Tsume, M. et al. Brd2 is required for cell cycle exit and neuronal differentiation through the E2F1 pathway in mouse neuroepithelial cells. *Biochem Biophys Res Commun* **425**, 762-8 (2012).
88. Hnilicová, J. et al. The C-terminal domain of Brd2 is important for chromatin interaction and regulation of transcription and alternative splicing. *Mol Biol Cell* **24**, 3557-68 (2013).

89. Peng, J., Zhu, Y., Milton, J.T. & Price, D.H. Identification of multiple cyclin subunits of human P-TEFb. *Genes Dev* **12**, 755-62 (1998).
90. Fu, T.J., Peng, J., Lee, G., Price, D.H. & Flores, O. Cyclin K functions as a CDK9 regulatory subunit and participates in RNA polymerase II transcription. *J Biol Chem* **274**, 34527-30 (1999).
91. Yang, Z. et al. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* **19**, 535-45 (2005).
92. Liu, L. et al. Transcriptional pause release is a rate-limiting step for somatic cell reprogramming. *Cell Stem Cell* **15**, 574-88 (2014).
93. Di Stefano, B. et al. C/EBPα creates elite cells for iPSC reprogramming by upregulating Klf4 and increasing the levels of Lsd1 and Brd4. *Nat Cell Biol* **18**, 371-81 (2016).
94. Floyd, S.R. et al. The bromodomain protein Brd4 insulates chromatin from DNA damage signalling. *Nature* **498**, 246-50 (2013).
95. Matzuk, M.M. et al. Small-molecule inhibition of BRDT for male contraception. *Cell* **150**, 673-84 (2012).
96. Gaucher, J. et al. Bromodomain-dependent stage-specific male genome programming by Brdt. *Embo j* **31**, 3809-20 (2012).
97. Pivot-Pajot, C. et al. Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* **23**, 5354-65 (2003).
98. French, C. NUT midline carcinoma. *Nat Rev Cancer* **14**, 149-50 (2014).
99. Bauer, D.E. et al. Clinicopathologic features and long-term outcomes of NUT midline carcinoma. *Clin Cancer Res* **18**, 5773-9 (2012).
100. Grayson, A.R. et al. MYC, a downstream target of BRD-NUT, is necessary and sufficient for the blockade of differentiation in NUT midline carcinoma. *Oncogene* **33**, 1736-1742 (2014).
101. Hnisz, D. et al. Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934-47 (2013).
102. Reynoird, N. et al. Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. *Embo j* **29**, 2943-52 (2010).
103. Stathis, A. et al. Clinical Response of Carcinomas Harboring the BRD4-NUT Oncoprotein to the Targeted Bromodomain Inhibitor OTX015/MK-8628. *Cancer Discov* **6**, 492-500 (2016).
104. Zuber, J. et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524-8 (2011).
105. Call, S.G. et al. Targeting Oncogenic Super Enhancers in MYC-Dependent AML Using a Small Molecule Activator of NR4A Nuclear Receptors. *Sci Rep* **10**, 2851 (2020).
106. Hultmark, S. et al. Combinatorial molecule screening identifies a novel diterpene and the BET inhibitor CPI-203 as differentiation inducers of primary acute myeloid leukemia cells. *Haematologica* (2020).
107. Shi, Y. et al. Low-dose triptolide enhances antitumor effect of JQ1 on acute myeloid leukemia through inhibiting RNA polymerase II in vitro and in vivo. *Mol Carcinog* **59**, 1076-1087 (2020).
108. Mertz, J.A. et al. Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc Natl Acad Sci U S A* **108**, 16669-74 (2011).
109. Ott, C.J. et al. BET bromodomain inhibition targets both c-Myc and IL7R in high-risk acute lymphoblastic leukemia. *Blood* **120**, 2843-52 (2012).
110. Bandopadhayay, P. et al. BET bromodomain inhibition of MYC-amplified medulloblastoma. *Clin Cancer Res* **20**, 912-25 (2014).
111. Henssen, A. et al. BET bromodomain protein inhibition is a therapeutic option for medulloblastoma. *Oncotarget* **4**, 2080-95 (2013).
112. Venkataraman, S. et al. Inhibition of BRD4 attenuates tumor cell self-renewal and suppresses stem cell signaling in MYC driven medulloblastoma. *Oncotarget* **5**, 2355-71 (2014).

113. Rajagopalan, V., Vaidyanathan, M., Janardhanam, V.A. & Bradner, J.E. Pre-clinical analysis of changes in intra-cellular biochemistry of glioblastoma multiforme (GBM) cells due to c-Myc silencing. *Cell Mol Neurobiol* **34**, 1059-69 (2014).
114. Asangani, I.A. et al. Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature* **510**, 278-82 (2014).
115. Cho, H. et al. RapidCaP, a novel GEM model for metastatic prostate cancer analysis and therapy, reveals myc as a driver of Pten-mutant metastasis. *Cancer Discov* **4**, 318-33 (2014).
116. Gao, L. et al. Androgen receptor promotes ligand-independent prostate cancer progression through c-Myc upregulation. *PLoS One* **8**, e63563 (2013).
117. Welti, J. et al. Targeting Bromodomain and Extra-Terminal (BET) Family Proteins in Castration-Resistant Prostate Cancer (CRPC). *Clin Cancer Res* **24**, 3149-3162 (2018).
118. Zhang, P. et al. Intrinsic BET inhibitor resistance in SPOP-mutated prostate cancer is mediated by BET protein stabilization and AKT-mTORC1 activation. *Nat Med* **23**, 1055-1062 (2017).
119. Faivre, E.J. et al. Selective inhibition of the BD2 bromodomain of BET proteins in prostate cancer. *Nature* **578**, 306-310 (2020).
120. Tsai, L.H. et al. The MZF1/c-MYC axis mediates lung adenocarcinoma progression caused by wild-type lkb1 loss. *Oncogene* **34**, 1641-9 (2015).
121. Gobbi, G. et al. The Hippo pathway modulates resistance to BET proteins inhibitors in lung cancer cells. *Oncogene* **38**, 6801-6817 (2019).
122. Shimamura, T. et al. Efficacy of BET bromodomain inhibition in Kras-mutant non-small cell lung cancer. *Clin Cancer Res* **19**, 6183-92 (2013).
123. Baratta, M.G. et al. An in-tumor genetic screen reveals that the BET bromodomain protein, BRD4, is a potential therapeutic target in ovarian carcinoma. *Proc Natl Acad Sci U S A* **112**, 232-7 (2015).
124. Qiu, H. et al. JQ1 suppresses tumor growth through downregulating LDHA in ovarian cancer. *Oncotarget* **6**, 6915-30 (2015).
125. Bihani, T. et al. Resistance to everolimus driven by epigenetic regulation of MYC in ER+ breast cancers. *Oncotarget* **6**, 2407-20 (2015).
126. Segura, M.F. et al. BRD4 sustains melanoma proliferation and represents a new target for epigenetic therapy. *Cancer Res* **73**, 6264-76 (2013).
127. Shao, Q. et al. BET protein inhibitor JQ1 attenuates Myc-amplified MCC tumor growth in vivo. *Cancer Res* **74**, 7090-102 (2014).
128. Pastori, C. et al. BET bromodomain proteins are required for glioblastoma cell proliferation. *Epigenetics* **9**, 611-20 (2014).
129. Puissant, A. et al. Targeting MYCN in neuroblastoma by BET bromodomain inhibition. *Cancer Discov* **3**, 308-23 (2013).
130. Lee, Y.S. et al. Runx3 inactivation is a crucial early event in the development of lung adenocarcinoma. *Cancer Cell* **24**, 603-16 (2013).
131. Belkina, A.C., Blanton, W.P., Nikolajczyk, B.S. & Denis, G.V. The double bromodomain protein Brd2 promotes B cell expansion and mitogenesis. *J Leukoc Biol* **95**, 451-60 (2014).
132. Dey, A., Nishiyama, A., Karpova, T., McNally, J. & Ozato, K. Brd4 marks select genes on mitotic chromatin and directs postmitotic transcription. *Mol Biol Cell* **20**, 4899-909 (2009).
133. Pasparakis, M. Regulation of tissue homeostasis by NF-kappaB signalling: implications for inflammatory diseases. *Nat Rev Immunol* **9**, 778-88 (2009).
134. Oeckinghaus, A. & Ghosh, S. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol* **1**, a000034 (2009).
135. Brown, J.D. et al. NF-κB directs dynamic super enhancer formation in inflammation and atherogenesis. *Mol Cell* **56**, 219-231 (2014).
136. Hah, N. et al. Inflammation-sensitive super enhancers form domains of coordinately regulated enhancer RNAs. *Proc Natl Acad Sci U S A* **112**, E297-302 (2015).

137. Wang, F. et al. Brd2 disruption in mice causes severe obesity without Type 2 diabetes. *Biochem J* **425**, 71-83 (2009).
138. Mahdi, H. et al. Specific interaction between genotype, smoking and autoimmunity to citrullinated alpha-enolase in the etiology of rheumatoid arthritis. *Nat Genet* **41**, 1319-24 (2009).
139. Lin, A., Wang, S., Nguyen, T., Shire, K. & Frappier, L. The EBNA1 protein of Epstein-Barr virus functionally interacts with Brd4. *J Virol* **82**, 12009-19 (2008).
140. Sakakibara, N. et al. Brd4 is displaced from HPV replication factories as they expand and amplify viral DNA. *PLoS Pathog* **9**, e1003777 (2013).
141. Sharma, A. et al. BET proteins promote efficient murine leukemia virus integration at transcription start sites. *Proc Natl Acad Sci U S A* **110**, 12036-41 (2013).
142. Gupta, S.S. et al. Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. *J Virol* **87**, 12721-36 (2013).
143. Li, J., Li, Q., Diaz, J. & You, J. Brd4-mediated nuclear retention of the papillomavirus E2 protein contributes to its stabilization in host cells. *Viruses* **6**, 319-35 (2014).
144. Zhou, H. et al. Epstein-Barr virus oncoprotein super-enhancers control B cell growth. *Cell Host Microbe* **17**, 205-16 (2015).
145. Boehm, D. et al. BET bromodomain-targeting compounds reactivate HIV from latency via a Tat-independent mechanism. *Cell Cycle* **12**, 452-62 (2013).
146. Li, Z., Guo, J., Wu, Y. & Zhou, Q. The BET bromodomain inhibitor JQ1 activates HIV latency through antagonizing Brd4 inhibition of Tat-transactivation. *Nucleic Acids Res* **41**, 277-87 (2013).
147. Ito, T. et al. Real-time imaging of histone H4K12-specific acetylation determines the modes of action of histone deacetylase and bromodomain inhibitors. *Chem Biol* **18**, 495-507 (2011).
148. Zhang, G. et al. Down-regulation of NF-κB transcriptional activity in HIV-associated kidney disease by BRD4 inhibition. *J Biol Chem* **287**, 28840-51 (2012).
149. Chung, C.W. et al. Discovery and characterization of small molecule inhibitors of the BET family bromodomains. *J Med Chem* **54**, 3827-38 (2011).
150. Filippakopoulos, P. et al. Benzodiazepines and benzotriazepines as protein interaction inhibitors targeting bromodomains of the BET family. *Bioorg Med Chem* **20**, 1878-86 (2012).
151. Bamborough, P. et al. Fragment-based discovery of bromodomain inhibitors part 2: optimization of phenylisoxazole sulfonamides. *J Med Chem* **55**, 587-96 (2012).
152. Fish, P.V. et al. Identification of a chemical probe for bromo and extra C-terminal bromodomain inhibition through optimization of a fragment-derived hit. *J Med Chem* **55**, 9831-7 (2012).
153. Khmelnitsky, Y.L. et al. In vitro biosynthesis, isolation, and identification of predominant metabolites of 2-(4-(2-hydroxyethoxy)-3,5-dimethylphenyl)-5,7-dimethoxyquinazolin-4(3H)-one (RVX-208). *Eur J Med Chem* **64**, 121-8 (2013).
154. Delmore, J.E. et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904-17 (2011).
155. Devaiah, B.N. et al. BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain. *Proc Natl Acad Sci U S A* **109**, 6927-32 (2012).
156. Picaud, S. et al. PFI-1, a highly selective protein interaction inhibitor, targeting BET Bromodomains. *Cancer Res* **73**, 3336-46 (2013).
157. Chaidos, A. et al. Potent antimyeloma activity of the novel bromodomain inhibitors I-BET151 and I-BET762. *Blood* **123**, 697-705 (2014).
158. Fiskus, W. et al. Highly active combination of BRD4 antagonist and histone deacetylase inhibitor against human acute myelogenous leukemia cells. *Mol Cancer Ther* **13**, 1142-54 (2014).

159. Herrmann, H. et al. Small-molecule inhibition of BRD4 as a new potent approach to eliminate leukemic stem- and progenitor cells in acute myeloid leukemia AML. *Oncotarget* **3**, 1588-99 (2012).
160. Da Costa, D. et al. BET inhibition as a single or combined therapeutic approach in primary paediatric B-precursor acute lymphoblastic leukaemia. *Blood Cancer J* **3**, e126 (2013).
161. Roderick, J.E. et al. c-Myc inhibition prevents leukemia initiation in mice and impairs the growth of relapsed and induction failure pediatric T-ALL cells. *Blood* **123**, 1040-50 (2014).
162. Cheng, Z. et al. Inhibition of BET bromodomain targets genetically diverse glioblastoma. *Clin Cancer Res* **19**, 1748-59 (2013).
163. Wyce, A. et al. Inhibition of BET bromodomain proteins as a therapeutic approach in prostate cancer. *Oncotarget* **4**, 2419-29 (2013).
164. Picaud, S. et al. RVX-208, an inhibitor of BET transcriptional regulators with selectivity for the second bromodomain. *Proc Natl Acad Sci U S A* **110**, 19754-9 (2013).
165. Gilan, O. et al. Selective targeting of BD1 and BD2 of the BET proteins in cancer and immunoinflammation. *Science* **368**, 387-394 (2020).
166. Tanaka, M. et al. Design and characterization of bivalent BET inhibitors. *Nat Chem Biol* **12**, 1089-1096 (2016).
167. Waring, M.J. et al. Potent and selective bivalent inhibitors of BET bromodomains. *Nat Chem Biol* **12**, 1097-1104 (2016).
168. Sakamoto, K.M. et al. Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation. *Proc Natl Acad Sci U S A* **98**, 8554-9 (2001).
169. Winter, G.E. et al. DRUG DEVELOPMENT. Phthalimide conjugation as a strategy for in vivo target protein degradation. *Science* **348**, 1376-81 (2015).
170. Bai, L. et al. Targeted Degradation of BET Proteins in Triple-Negative Breast Cancer. *Cancer Res* **77**, 2476-2487 (2017).
171. Raina, K. et al. PROTAC-induced BET protein degradation as a therapy for castration-resistant prostate cancer. *Proc Natl Acad Sci U S A* **113**, 7124-9 (2016).
172. Lu, J. et al. Hijacking the E3 Ubiquitin Ligase Cereblon to Efficiently Target BRD4. *Chem Biol* **22**, 755-63 (2015).
173. Xu, L. et al. Targetable BET proteins- and E2F1-dependent transcriptional program maintains the malignancy of glioblastoma. *Proc Natl Acad Sci U S A* **115**, E5086-e5095 (2018).
174. Qin, C. et al. Discovery of QCA570 as an Exceptionally Potent and Efficacious Proteolysis Targeting Chimera (PROTAC) Degrader of the Bromodomain and Extra-Terminal (BET) Proteins Capable of Inducing Complete and Durable Tumor Regression. *J Med Chem* **61**, 6685-6704 (2018).
175. Zhou, B. et al. Discovery of a Small-Molecule Degrader of Bromodomain and Extra-Terminal (BET) Proteins with Picomolar Cellular Potencies and Capable of Achieving Tumor Regression. *J Med Chem* **61**, 462-481 (2018).
176. Shi, C. et al. PROTAC induced-BET protein degradation exhibits potent anti-osteosarcoma activity by triggering apoptosis. *Cell Death Dis* **10**, 815 (2019).
177. Feng, Q. et al. An epigenomic approach to therapy for tamoxifen-resistant breast cancer. *Cell Res* **24**, 809-19 (2014).
178. Stuhlmiller, T.J. et al. Inhibition of Lapatinib-Induced Kinome Reprogramming in ERBB2-Positive Breast Cancer by Targeting BET Family Bromodomains. *Cell Rep* **11**, 390-404 (2015).
179. Lenhart, R. et al. Sensitivity of Small Cell Lung Cancer to BET Inhibition Is Mediated by Regulation of ASCL1 Gene Expression. *Mol Cancer Ther* **14**, 2167-74 (2015).
180. McCleland, M.L. et al. CCAT1 is an enhancer-templated RNA that predicts BET sensitivity in colorectal cancer. *J Clin Invest* **126**, 639-52 (2016).
181. Liu, F. et al. EGFR Mutation Promotes Glioblastoma through Epigenome and Transcription Factor Network Remodeling. *Mol Cell* **60**, 307-18 (2015).

182. Long, J. et al. The BET bromodomain inhibitor I-BET151 acts downstream of smoothened protein to abrogate the growth of hedgehog protein-driven cancers. *J Biol Chem* **289**, 35494-502 (2014).
183. Tang, Y. et al. Epigenetic targeting of Hedgehog pathway transcriptional output through BET bromodomain inhibition. *Nat Med* **20**, 732-40 (2014).
184. Garcia, P.L. et al. The BET bromodomain inhibitor JQ1 suppresses growth of pancreatic ductal adenocarcinoma in patient-derived xenograft models. *Oncogene* **35**, 833-45 (2016).
185. Mazur, P.K. et al. Combined inhibition of BET family proteins and histone deacetylases as a potential epigenetics-based therapy for pancreatic ductal adenocarcinoma. *Nat Med* **21**, 1163-71 (2015).
186. Fukuda, A. et al. Stat3 and MMP7 contribute to pancreatic ductal adenocarcinoma initiation and progression. *Cancer Cell* **19**, 441-55 (2011).
187. Piha-Paul, S.A. et al. Phase 1 Study of Molibresib (GSK525762), a Bromodomain and Extra-Terminal Domain Protein Inhibitor, in NUT Carcinoma and Other Solid Tumors. *JNCI Cancer Spectr* **4**, pkz093 (2020).
188. Denis, G.V. et al. Identification of transcription complexes that contain the double bromodomain protein Brd2 and chromatin remodeling machines. *J Proteome Res* **5**, 502-11 (2006).
189. Gyuris, A. et al. The chromatin-targeting protein Brd2 is required for neural tube closure and embryogenesis. *Biochim Biophys Acta* **1789**, 413-21 (2009).
190. Shang, E., Wang, X., Wen, D., Greenberg, D.A. & Wolgemuth, D.J. Double bromodomain-containing gene Brd2 is essential for embryonic development in mouse. *Dev Dyn* **238**, 908-17 (2009).
191. Jang, M.K. et al. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol Cell* **19**, 523-34 (2005).
192. Berkovits, B.D. & Wolgemuth, D.J. The first bromodomain of the testis-specific double bromodomain protein Brdt is required for chromocenter organization that is modulated by genetic background. *Dev Biol* **360**, 358-68 (2011).
193. Bolden, J.E. et al. Inducible in vivo silencing of Brd4 identifies potential toxicities of sustained BET protein inhibition. *Cell Rep* **8**, 1919-1929 (2014).
194. Nakagawa, A. et al. Selective and reversible suppression of intestinal stem cell differentiation by pharmacological inhibition of BET bromodomains. *Sci Rep* **6**, 20390 (2016).
195. Fong, C.Y. et al. BET inhibitor resistance emerges from leukaemia stem cells. *Nature* **525**, 538-42 (2015).
196. Rathert, P. et al. Transcriptional plasticity promotes primary and acquired resistance to BET inhibition. *Nature* **525**, 543-547 (2015).
197. Shu, S. et al. Response and resistance to BET bromodomain inhibitors in triple-negative breast cancer. *Nature* **529**, 413-417 (2016).
198. Marcotte, R. et al. Functional Genomic Landscape of Human Breast Cancer Drivers, Vulnerabilities, and Resistance. *Cell* **164**, 293-309 (2016).
199. Ruthenburg, A.J. et al. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**, 692-706 (2011).
200. Ruthenburg, A.J., Li, H., Patel, D.J. & Allis, C.D. Multivalent engagement of chromatin modifications by linked binding modules. *Nat Rev Mol Cell Biol* **8**, 983-94 (2007).
201. Tompa, P. On the supertertiary structure of proteins. *Nat Chem Biol* **8**, 597-600 (2012).
202. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys J* **78**, 1606-19 (2000).
203. Lebowitz, J., Lewis, M.S. & Schuck, P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci* **11**, 2067-79 (2002).

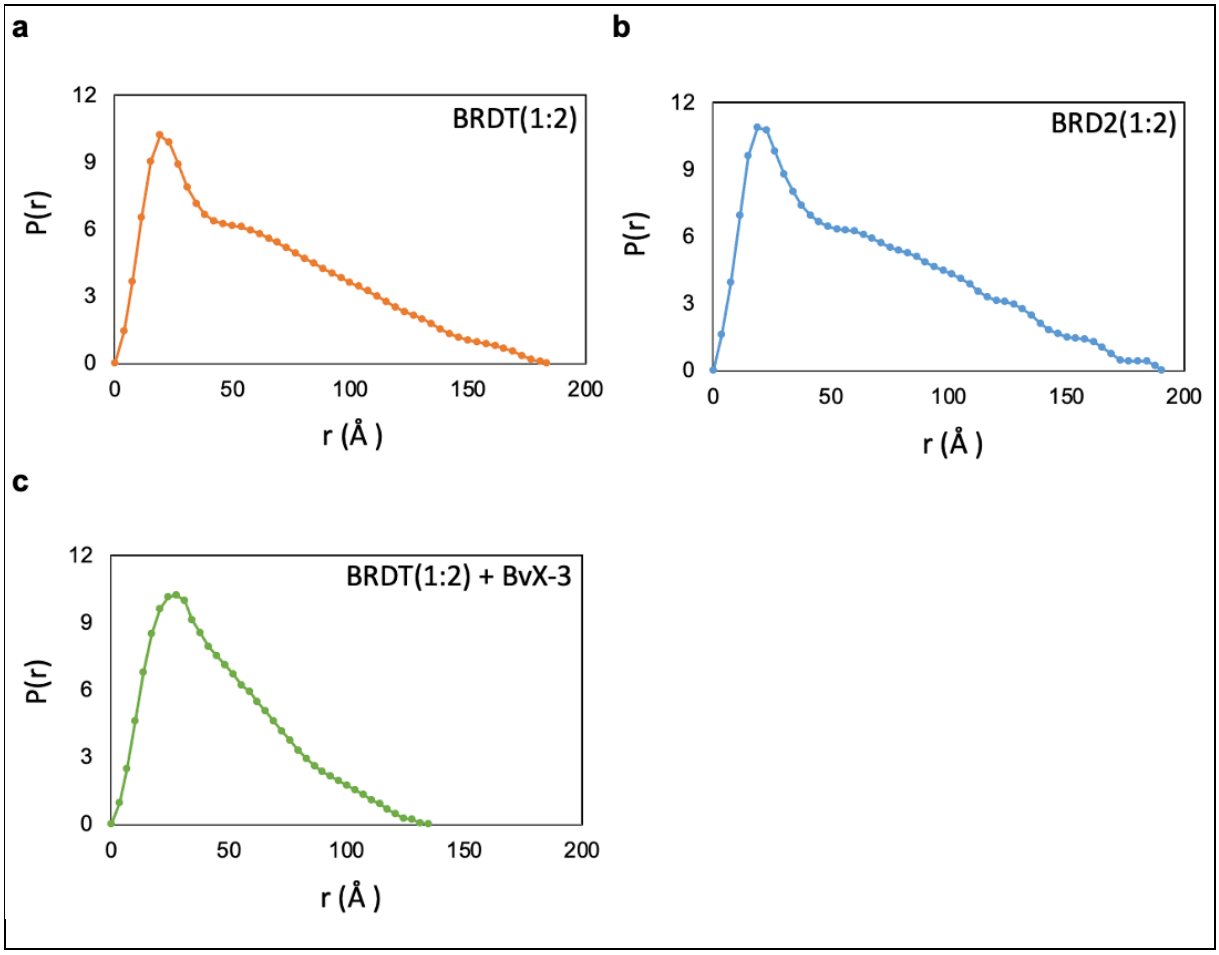
204. Yang, C., van der Woerd, M.J., Muthurajan, U.M., Hansen, J.C. & Luger, K. Biophysical analysis and small-angle X-ray scattering-derived structures of MeCP2-nucleosome complexes. *Nucleic Acids Res* **39**, 4122-35 (2011).
205. Jacques, D.A. & Trewhella, J. Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls. *Protein Sci* **19**, 642-57 (2010).
206. Mertens, H.D. & Svergun, D.I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J Struct Biol* **172**, 128-41 (2010).
207. Glatter, O. Data evaluation in small angle scattering: calculation of the radial electron density distribution by means of indirect Fourier transformation. *Acta Physica Austriaca* **47**, 83-102 (1977).
208. Putnam, C.D., Hammel, M., Hura, G.L. & Tainer, J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q Rev Biophys* **40**, 191-285 (2007).
209. Bada, M., Walther, D., Arcangioli, B., Doniach, S. & Delarue, M. Solution structural studies and low-resolution model of the Schizosaccharomyces pombe sap1 protein. *J Mol Biol* **300**, 563-74 (2000).
210. Svergun, D.I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys J* **76**, 2879-86 (1999).
211. Kikhney, A.G. & Svergun, D.I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett* **589**, 2570-7 (2015).
212. Chung, C.-w. et al. Discovery and Characterization of Small Molecule Inhibitors of the BET Family Bromodomains. *Journal of Medicinal Chemistry* **54**, 3827-3838 (2011).
213. McLure, K.G. et al. RVX-208, an inducer of ApoA-I in humans, is a BET bromodomain antagonist. *PLoS One* **8**, e83190 (2013).
214. Flynn, E.M. et al. A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. *Structure* **23**, 1801-1814 (2015).
215. Millett, I.S., Doniach, S. & Plaxco, K.W. Toward a taxonomy of the denatured state: small angle scattering studies of unfolded proteins. *Adv Protein Chem* **62**, 241-62 (2002).
216. Kohn, J.E. et al. Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc Natl Acad Sci U S A* **101**, 12491-6 (2004).
217. Bernadó, P. & Blackledge, M. A self-consistent description of the conformational behavior of chemically denatured proteins from NMR and small angle scattering. *Biophys J* **97**, 2839-45 (2009).
218. Flory, P.J. *Principles of polymer chemistry*, (1953).
219. Receveur-Brechot, V. & Durand, D. How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr Protein Pept Sci* **13**, 55-75 (2012).
220. Werner, M.T. et al. Comparative structure-function analysis of bromodomain and extraterminal motif (BET) proteins in a gene-complementation system. *J Biol Chem* **295**, 1898-1914 (2020).
221. Cushman, I. Utilizing peptide SPOT arrays to identify protein interactions. *Curr Protoc Protein Sci* **Chapter 18**, Unit 18.10 (2008).
222. Reineke, U., Volkmer-Engert, R. & Schneider-Mergener, J. Applications of peptide arrays prepared by the SPOT-technology. *Curr Opin Biotechnol* **12**, 59-64 (2001).
223. Nady, N. et al. Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *J Biol Chem* **286**, 24300-11 (2011).
224. Nady, N., Min, J., Kareta, M.S., Chédin, F. & Arrowsmith, C.H. A SPOT on the chromatin landscape? Histone peptide arrays as a tool for epigenetic research. *Trends Biochem Sci* **33**, 305-13 (2008).
225. Szymczak, L.C., Kuo, H.Y. & Mrksich, M. Peptide Arrays: Development and Application. *Anal Chem* **90**, 266-282 (2018).

226. Miller, T.C. et al. A bromodomain-DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT. *Nat Commun* **7**, 13855 (2016).
227. Olp, M.D., Jackson, V. & Smith, B.C. Nucleosome scaffolding by Brd4 tandem bromodomains in acetylation-dependent chromatin compartmentalization. *bioRxiv*, 699967 (2019).
228. Huang, F. & Nau, W.M. A conformational flexibility scale for amino acids in peptides. *Angew Chem Int Ed Engl* **42**, 2269-72 (2003).
229. Joo, S.H. Cyclic peptides as therapeutic agents and biochemical tools. *Biomol Ther (Seoul)* **20**, 19-26 (2012).
230. Stols, L. et al. A new vector for high-throughput, ligation-independent cloning encoding a tobacco etch virus protease cleavage site. *Protein Expr Purif* **25**, 8-15 (2002).
231. Franke, D., Kikhney, A.G. & Svergun, D.I. Automated acquisition and analysis of small angle X-ray scattering data. *Nuclear instruments & methods in physics research. Section A, Accelerators, spectrometers, detectors and associated equipment* **689**, 52-59 (2012).
232. Svergun, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *Journal of Applied Crystallography $V 25* (1992).

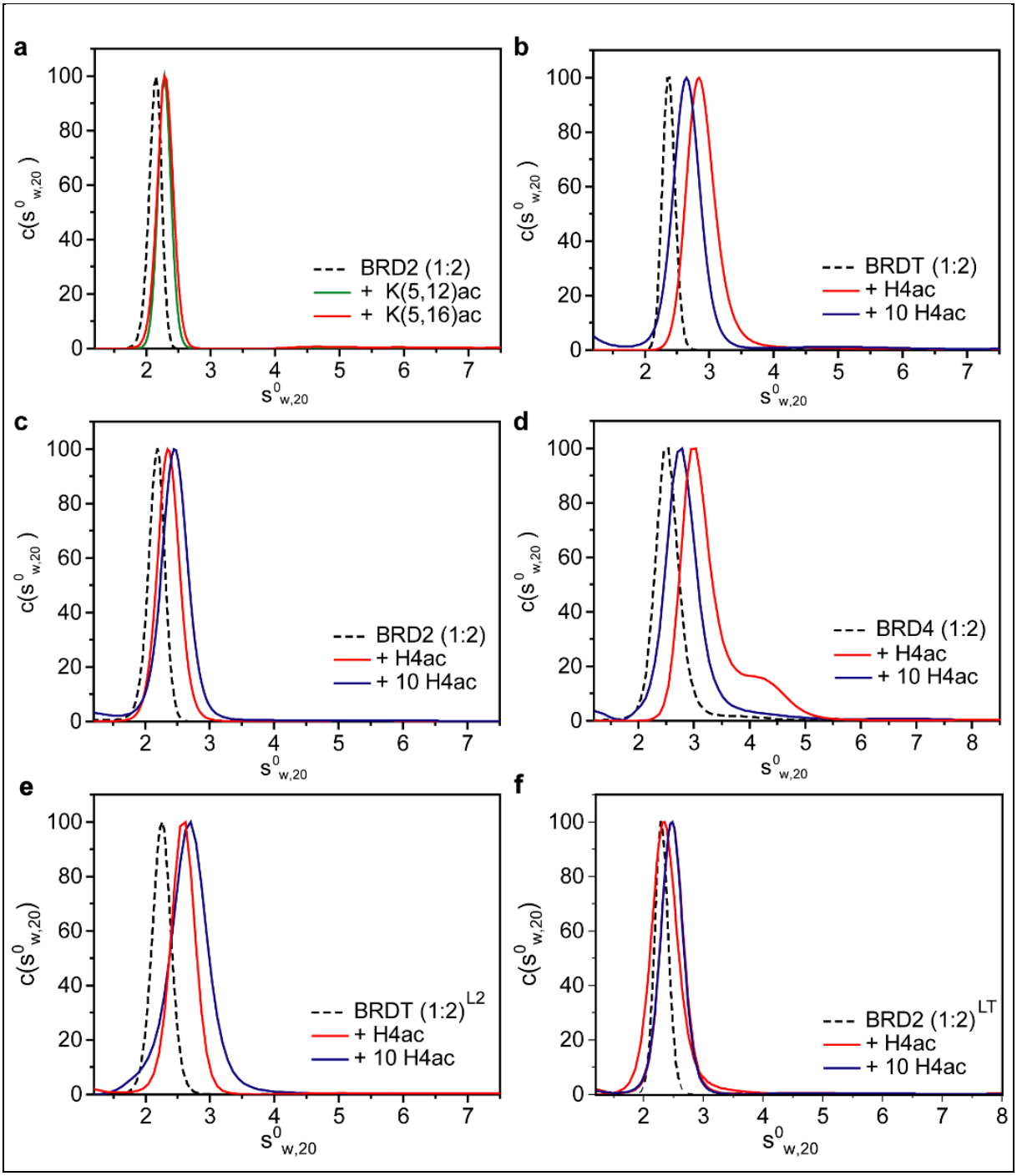
Appendix

Sample	R_g (Å)	D_{max} (Å)
BRDT(1:2)	48 ± 1.3	183 ± 20
BRD2(1:2)	55 ± 2.6	190 ± 20
BRDT(1:2) + BvX-3	44 ± 1.8	135 ± 20

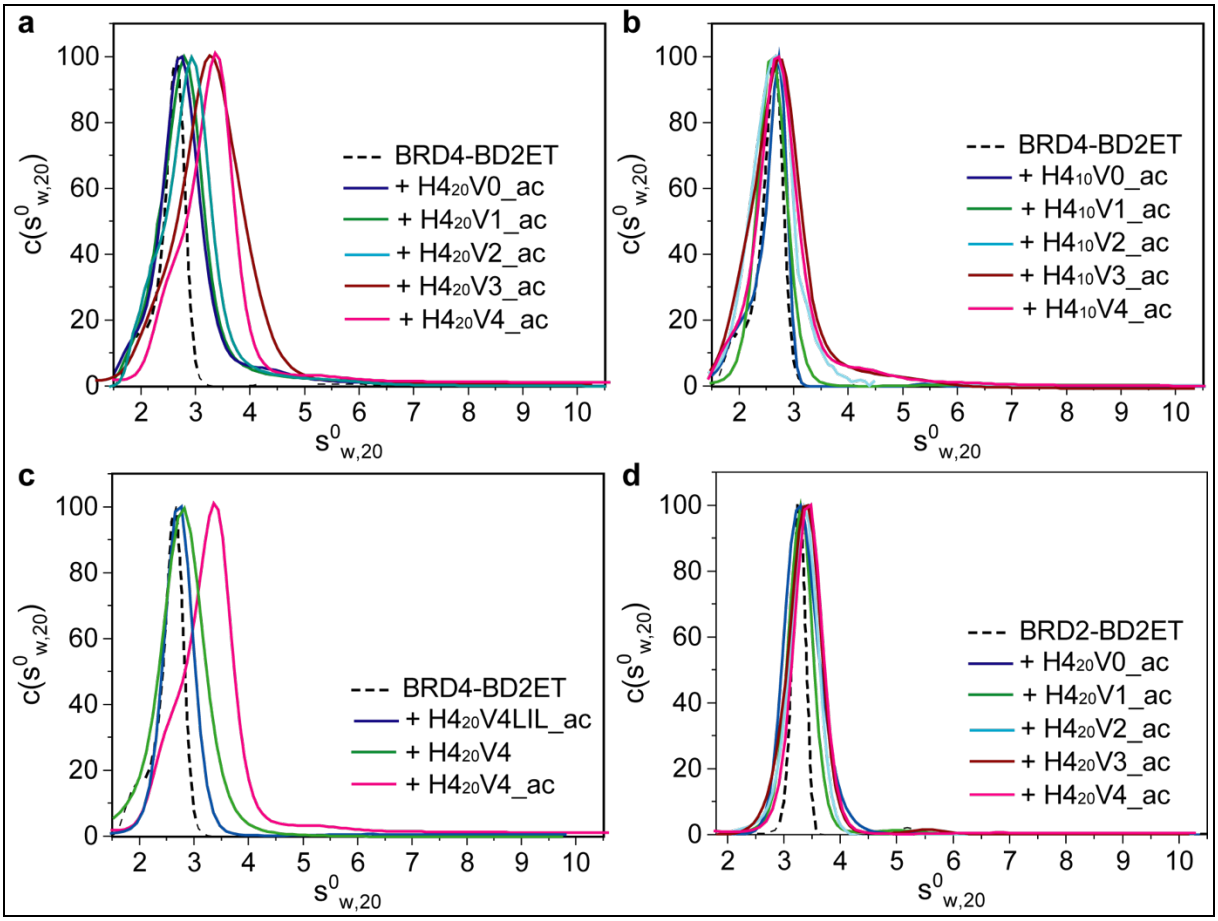
Supplementary table 1. Overall parameters of the BRDT(1:2) and BRD4(1:2) with or without the bivalent compound derived from SAXS experiments.



Supplementary figure 1. (**a**) The pair distance distribution function *p(r)* of BRDT(1:2). **(b)** The pair distance distribution function *p(r)* of BRD2(1:2). (**c**) The pair distance distribution function *p(r)* of BRDT(1:2) with the compound BX-3 at 1:1 molar ratio.



Supplementary figure 2. Sedimentation coefficient distribution corrected to water at 20 °C for BET tandem BDs in the presence of polyacetylated peptide at 1:1 and 1:10 protein:peptide ratio. (**a**) BRD2(1:2) in the presence of distant diacetylated peptide. (**b**) BRDT(1:2) in the presence of polyacetylated peptide. (**c**) BRD2(1:2) in the presence of polyacetylated peptide (**d**) BRD4(1:2) in the presence of polyacetylated peptide. (**e**) BRDT(1:2)^{L2} in the presence of polyacetylated peptide. (**f**) BRD2(1:2)^{LT} in the presence of polyacetylated peptide



Supplementary figure 3. Sedimentation coefficient distribution corrected to water at 20 °C for BET-BD2ETs in the presence of peptidic inhibitors at 1:1 protein: peptide ratio. (**a**) BRD4-BD2ET in the presence of H_{20}Vn_ac peptides (**b**) BRD4-BD2ET in the presence of H_{10}Vn_ac peptides. (**c**) BRD4-BD2ET in the presence of H_{20}V4_ac wild-type and mutant peptides (**d**) BRD2-BD2ET in the presence of H_{20}Vn_ac peptides