**Overview:**

The cellular machinery responsible for the acetylation of lysines is critically implicated in cancer biology. In particular the only protein domain that recognizes acetylated lysines, the bromodomain, is present in multiple proteins implicated in chromatin biology and regulation of gene expression, and many of these proteins are deregulated in human tumors. Structurally (see Appendix 1), the recognition of acetylated peptides by the bromodomain makes it a druggable target. Through the work of one of the co-applicants (PF), the first anti-cancer bromodomain inhibitor (JQ1; specific for the BET family of bromodomains) was developed and shows promise against a number of tumor types. However, this work also highlighted the need for more information regarding the cellular context in which bromodomain-containing proteins and acetylation enzymes act, and what are the cellular targets for bromodomain recognition of acetylated peptides. Here, we propose to systematically define the specificity and function of human bromodomains through identification of their binding partners and acetylated targets. This project builds on the strengths and the synergy of the research team, and will accelerate the selection and development of new chemical probes to target this family.

**Introduction**

*The acetylation machinery*. Lysine acetylation is the process by which an acetyl group is transferred (from Acetyl Coenzyme A) to the epsilon amine of a lysine residue, a modification that is catalyzed by a family of lysine acetyltransferases (KATs; 18 genes in human). Similar to protein phosphorylation in which kinases and phosphatases oppose each other for the addition (“writing”) and removal (“erasing”) of the phosphate group, acetyl groups (KAc) added by KATs can be removed by lysine deacetylases (KDACs, histone-specific KDACs are known as HDACs, histone deacetylases)[1](#_ENREF_1), [2](#_ENREF_2). KDACs are represented by two main classes which differ in their cofactor dependence, the KDACs proper, 11 genes, and the sirtuins, 7 genes. KAc marks are recognized by a single dedicated protein domain, the bromodomain, which acts as a “reader” of the acetylation mark and provides context-specific recognition (Fig 1). As detailed below, bromodomains are encoded in 44 genes (some proteins contain more than one bromodomain). Most of the bromodomain-containing proteins are comprised of several modular domains[3](#_ENREF_3), suggesting roles as molecular scaffolds (Fig 1c). Though not all bromodomain-containing proteins have been characterized in detail, the majority of those studied have roles in chromatin biology and regulation of gene expression (Table 1; please see[4](#_ENREF_4) for a recent review).

Lysine acetylation first came to the forefront as a histone modification[5](#_ENREF_5). Histone tails are modified by a number of post-translational modifications (PTMs) that present different cues to the cellular machinery (the so-called histone code[6](#_ENREF_6). In general, acetylation of histone tails acts to open the chromatin structure, activating transcription, although certain acetylation marks are associated instead with chromatin compaction and with other processes as well, such as metabolism and DNA repair[7](#_ENREF_7). Regulation of gene expression by histone acetylation status has been the driving force behind the development of KDAC inhibitors[8](#_ENREF_8), as detailed below. However, while first studied intensively in the context of the histone code, lysine acetylation is now recognized as a widespread PTM occurring on a large proportion of the proteome[8](#_ENREF_8), [9](#_ENREF_9). So far, 18330 sites on 6870 proteins have been reported in total in human cells in a PTM repository (http://www.phosphosite.org), which is likely still an underestimation (as detailed below, we consistently detect new sites by mass spectrometry). The prevalence of acetyl lysine on thousands of cellular proteins underlies the vital importance of this modification and also highlights the lack of information regarding which KATs and KDACs are responsible for these acetylation events, and, just as importantly, which sites on which proteins are “read” by the different bromodomains.

*The acetylome in disease*. Tight control of the acetylation system (here referred to as the “acetylome”) is clearly critical to cellular homeostasis. This is evidenced by the fact that many of the components of the acetylome have been associated with disease, particularly cancer (Table 1). Acetylome alterations in disease include expression changes, but also a number of genetic alterations, most notably deletions, point mutations and translocations. For example, the bromodomain-containing acetyltransferase KAT3A (also known as CREBBP, CREB binding protein) is a transcriptional co-activator mutated in Rubinstein-Taybi syndrome, a multiple congenital anomaly syndrome characterized by mental retardation, dysmorphic facial features, postnatal growth deficiency and predisposition to cancer. KAT3A is also translocated, to a variety of fusion partners (including other components of the acetylation machinery such as MLL, KAT6A, KAT6B; Fig 2), in acute monocytic leukemia (AML)[10-12](#_ENREF_10). Mutations in the bromodomain-containing protein BRWD3 are associated with X-linked mental retardation. BRWD3 is also disrupted by translocation in B-cell chronic lymphocytic leukemia[13](#_ENREF_13). The overexpression of multiple KDACs in cancers prompted the testing of KDAC inhibitors as anticancer therapeutics. To date, two KDAC (HDAC) inhibitors have been approved for clinical use. Vorinostat, a broad specificity small molecule inhibitor targeting KDACs 1-11, was approved for the treatment of persistent or recurrent cutaneous T-cell lymphoma (CTCL)[14](#_ENREF_14). Recently, the cyclic peptide Romidepsin, targeting KDAC1 and KDAC2, was approved for treatment of CTCL[15](#_ENREF_15). Though these inhibitors show clinical promise, their cellular mode of action is still not clear. These compounds are also bound to be somewhat non-specific: besides the fact that most of the currently used inhibitors can target more than one KDAC, the number of direct acetylation targets for each of the KDACs (let alone their identity) is still largely unknown. Furthermore, most KDAC inhibitors have been shown in clinical trials to be ineffective as monotherapies[8](#_ENREF_8), reinforcing the need for improved therapeutic agents targeting the acetylome.

*BRD3, BRD4 and NUT midline carcinomas*. The bromodomain-containing protein BRD4 (and more rarely BRD3) is fused to the NUT (Nuclear protein in testis) protein (gene name C15ORF55) in a rare and aggressive form of squamous cell carcinoma called NUT midline carcinoma (NMC)[16](#_ENREF_16). So far, less than 100 NMC cases have been reported[17](#_ENREF_17). However, this may be due in part to misdiagnosis associated – until recently - with a lack of diagnostic tools and registry, according to our collaborator and leading NMC expert, Dr. C French (letter attached). The cell of origin is unknown for NMCs, and NMCs occur in a variety of sites, most commonly in the upper digestive track and the mediastinum. NMCs have a devastating clinical course, and though initial response to chemotherapy is observed, recurrence is rapid and insensitivity to subsequent treatment results in an average survival time of less than 10 months[18](#_ENREF_18). By contrast to most carcinomas, which have very complex karyotypes, NMCs usually harbor a single reciprocal [t(15;19)(q14;p13.1)] chromosomal translocation which results in fusing the promoter and most of the BRD4 coding sequence (including all modular domains) to the nearly complete coding sequence of NUT (BRD4-NUT; Fig 2)[17](#_ENREF_17). NUT expression is normally restricted to testis: consistent with this, the reciprocal translocation product (NUT-BRD4) has not been detected in NMC. The role of the NUT protein in testis is not completely clear, though it was reported to possess an acidic domain enabling interaction with KAT3B (EP300)[19](#_ENREF_19). On the basis of this information, it was suggested that in spermatids, NUT binds and activates KAT3B, leading to global increases in histone acetylation.

The biological function of BRD4 (and of the rarer NUT fusion partner BRD3) is better understood than that of NUT. These proteins belong to a subgroup of bromodomain-containing proteins known as the BET family (for bromodomain and extraterminal) that share the same structure: two bromodomains, followed by the extraterminal (ET) domain which may be involved in protein/DNA recognition. BET proteins regulate transcriptional elongation by providing a docking site for the P-TEFb complex. The kinase component of P-TEFb, CDK9, phosphorylates the C-terminal tail of RNA polymerase II, resulting in transcriptional upregulation[20](#_ENREF_20). BRD4 is implicated in post-mitotic gene transcription, and recruits P-TEFb to mitotic chromosomes resulting in the expression of growth-promoting genes; it is expected to contribute to epigenetic memory[21](#_ENREF_21). Two major BRD4 transcript variants are annotated that encode “long” and “short” isoforms differing in their C-termini. Both of these variants, as well as the BRD4 fragment fused to NUT, contain the tandem bromodomains and the ET domain; the longest form appears to have an extended interactome, which include proteins from viral pathogens such as papillomaviruses[22](#_ENREF_22). The link to viral oncogenic proteins, as well as the control of Myc expression[23](#_ENREF_23), [24](#_ENREF_24), implicates BRD4 in other cancer-related events outside of the NMCs, and BRD4 was recently found in an RNAi screen to be a therapeutic target in AML[25](#_ENREF_25).

The driving role of the BRD4-NUT or BRD3-NUT fusions in NMCs is well established: in cell lines derived from patients with NMCs, knockdown of BRD4-NUT or BRD3-NUT induces rapid terminal squamous differentiation and arrests proliferation[26](#_ENREF_26). This also strongly indicates that a block in differentiation drives NMCs. How exactly the fusion proteins drive NMC is not completely clear, but different lines of evidence have suggested that they may sequester KAT3B from differentiation specific genes[19](#_ENREF_19).

*Recognition of acetylated lysines by bromodomains*. Despite variation in primary amino acid sequence, all bromodomains (a domain of ~120 aa) share a common fold, and bind to acetylated lysines in a similar manner[27](#_ENREF_27). The bromodomain consists of a left-handed bundle of four alpha helices that are linked by more variable loop insert regions[28](#_ENREF_28). The acetylated lysine is recognized in a central deep and largely hydrophobic cavity, as determined by co-crystal structures of bromodomains complexed with acetylated lysine-containing peptides[29](#_ENREF_29). Anchoring of the acetylated lysine (or lysines, see below) is provided by hydrogen bonding to a conserved asparagine residue (Fig 3; other hydrogen bond donors such as the side chains of threonine or tyrosine can replace the asparagine). How bromodomains specifically recognize distinct acetylated peptides is beginning to be understood, in large part due to our group’s efforts to systematically determine the structures of multiple bromodomains (Appendix 1; Filippakopoulos et al., Cell*, in press*). For example, while the structural fold of all bromodomains is similar, the electrostatic potential of the surfaces surrounding the KAc binding site is diverse (Appendix 1), suggesting that they recognize different sequences (e.g. positively charged histones may not be favored by bromodomains that have a positive charge, such as that of the first bromodomain of PBRM1). Note that differences in surface charge have been observed even within the same protein, including the first and second domains of BRD4 (Fig 3), suggesting that different bromodomains within a protein may target different sites, an hypothesis that we addressed directly for BRD4, as described below. Additional selectivity elements are provided by the diverse loop regions that distinguish the different bromodomains.

*Targeting bromodomains for therapies*. Disrupting the interaction between a bromodomain and its substrates could in theory provide a much more specific control over the disease-associated acetylome than is currently afforded by KDAC inhibitors. Structurally, the deep and relatively hydrophobic cavity for binding acetylated lysine residues offers a good target for the development of small molecule inhibitors of the bromodomain-KAc interaction[4](#_ENREF_4). Given that the surface and loops of the bromodomains are diverse, it is expected that highly specific inhibitors can be designed (Table 2). We (PF) developed with our collaborators the first anti-cancer bromodomain inhibitor, (+)-JQ1, based on a triazolodiazepine scaffold, which binds to and inhibits all BET-family bromodomains (Fig 4)[30](#_ENREF_30). (+)-JQ1, but not its stereoisomer, binds in the KAc pocket of the bromodomain. The fact that BRD4-NUT and BRD3-NUT fusions are driving oncogenes in NMCs provided an excellent test case for the development of bromodomain inhibitors in an attempt to treat these rare and aggressive tumors. In cell culture, the cell permeable (+)-JQ1 inhibitor displaces BRD4-NUT from chromatin, induces squamous differentiation, and reduces proliferation[30](#_ENREF_30). Importantly, in xenograft models of NMCs, the well-tolerated (+)-JQ1 inhibitor led to tumor regression and improved mouse survival[30](#_ENREF_30). This demonstrates not only the feasibility of developing specific inhibitors for bromodomain-KAc binding, but also that these inhibitors are of potential clinical value. While initially tested on NMCs, which offered a clean genetic model for BRD4-driven cancers, bromodomain-specific inhibitors have therapeutic value in a wider range of cancers. For example, after identifying BRD4 as a potential therapeutic target in AML by RNA interference (RNAi), (+)-JQ1 was shown to exhibit robust anti-leukemic effects both *in vitro* and *in vivo*, and these were accompanied with terminal differentiation and elimination of the leukemic stem cells[25](#_ENREF_25). This therapeutic effect is also evident in multiple myelomas, and occurs through regulation of the Myc oncogenic driver in these tumors[23](#_ENREF_23). In addition, bromodomain inhibitors also have potential as anti-inflammatories, e.g. the BRD4 specific I-BET inhibitor (structurally related to JQ1) reduces expression of pro-inflammatory genes in activated macrophages[31](#_ENREF_31). Inhibitors to bromodomains outside of the BET family are also being developed, for example against KAT3B (CREBBP[32](#_ENREF_32), [33](#_ENREF_33)) and PCAF[34](#_ENREF_34), and this will remain a very active area of research, given the involvement of the acetylome in disease, and the striking results obtained with the first generation bromodomain inhibitors.

*How is the acetylome organized?* There is a pressing need to characterize systematically and in an unbiased manner the function of each of the components of the acetylation machinery and the relationships between the writer-reader-eraser modules. As mentioned above, structurally, bromodomains offer attractive targets for therapies and are the topic of an intense drug discovery program, yet very little is known regarding the cellular context in which many of the bromodomain-containing proteins reside or their specificity for their targets *in vivo*. As shown in Appendix 1, we have used synthetic libraries of histone-derived peptides to provide information regarding the specificity of isolated bromodomains *in vitro*. This expanded the number of bromodomain substrates, but also led to the realization that flanking PTMs (especially phosphorylation and acetylation) have a strong influence on the recognition of marks, indicating that bromodomains often recognize combinations of marks rather than isolated acetylated lysines. This observation was previously made for BRDT which requires the presence of several acetylation sites for high-affinity binding to histone tails[35](#_ENREF_35). Structural determination of BRD4 with different diacetylated peptides derived from histone H4 showed that both KAc groups bind within the same pocket, and that the different sequences exhibit the same mode of binding (Appendix 1). However, our studies have also outlined the need to use alternative approaches to the synthetic libraries to identify specific sequences to which each of the bromodomains are associated in a cellular context. This is important especially in light of the therapeutic design of inhibitors: the inhibitors should be able to displace relevant (and presumably higher affinity) KAc peptides from the bromodomain. As described in this proposal, we have assembled a team of co-applicants and collaborators to provide *in vivo* context for the acetylome and accelerate development and testing of bromodomain-specific inhibitors.

**Preliminary data:**

**Towards mapping the interactome for the acetylation machinery**

Our research team has a long-standing interest in understanding how PTMs affect protein-protein interactions[36-46](#_ENREF_36), and we (TP, ACG) have developed robust and sensitive experimental approaches to systematically identify interactions using affinity purification coupled to mass spectrometry[47-57](#_ENREF_47) (e.g. Appendix 2). An important factor for success of the AP-MS approach is the expression of the epitope-tagged protein at near endogenous levels. While several approaches can be employed to ensure appropriate expression levels, a system that we have found particularly useful is the Flp-In T-REx system (Invitrogen) in which isogenic stable integrants are generated in a HEK293 derivative, and recombinant protein expression is under the control of a tetracycline inducible promoter[58](#_ENREF_58), [59](#_ENREF_59). We have demonstrated that this system is ideal for the expression of allelic series (e.g. wild type proteins and proteins in which the bromodomain has been mutated; e.g. Fig 17). This expression system was selected for the generation of the interactome for the acetylation system (Fig 5). We have also elected to use proteins tagged at their N-termini with a 3xFLAG tag and a magnetic bead purification strategy, based on our previous work using different tagging approaches.

The standard affinity purification protocol developed in our labs for different projects was developed for soluble proteins; however, multiple components of the acetylation machinery are known to be associated with chromatin. In addition, key substrates, including histones, are essentially insoluble under standard protocols, and normally not recovered in AP-MS approaches[60](#_ENREF_60). To circumvent this problem, postdoctoral fellow Jean-Philippe Lambert developed during his PhD an approach for AP-MS of chromatin associated proteins in yeast, which he termed mChIP (modified chromatin immunoprecipitation[61](#_ENREF_61), [62](#_ENREF_62)), which he has now optimized for mammalian AP-MS (Fig 6; reviewed in Appendix 3). He demonstrated that this approach enabled recovery of multiple additional interactors for chromatin-associated proteins. In some cases, this is linked to a better solubilization of the bait protein (e.g. BAZ2A), while in other cases (e.g. BRD2), this is most likely due to better solubilization of the interactors (Fig 7). He also demonstrated that for this type of data, the sensitivity of the mass spectrometer is important: using new generation instruments (here an Orbitrap Velos) results in the identification of significantly more interactions than was previously possible.

The mChIP approach works because it solubilizes chromatin-associating proteins by shearing and cleaving the DNA using a combination of sonication and nuclease treatment. At the same time, this solubilization – especially in the context of single step purification and very sensitive mass spectrometers – creates a significant level of background (proteins which either interact with the affinity matrix or with the epitope tag independently of the bait). As in our previous studies, however, we found that the background was quantitatively reproducible and that the quantitative information embedded in each mass spectrometry experiment could be used to provide a confidence value to each potential protein-protein interaction detected here. To provide meaningful scoring of the interactions, we developed a series of software tools that we are distributing freely to the scientific community. ProHits tracks all mass spectrometry data, and functions as an important tool to annotate, visualize and interpret mass spectrometry data[63](#_ENREF_63). Recently, we fully integrated within the ProHits system the SAINT statistical tool. SAINT (Significance Analysis of INTeractome) uses quantitative information (e.g. spectral counts, which is how often peptides for a given protein have been detected in the mass spectrometer) for the recovery of each potential interactor across a series of negative control purifications as compared to its recovery in association with the different baits in the dataset[48](#_ENREF_48), [58](#_ENREF_58), [64](#_ENREF_64) (Fig 8; Appendix 4). We have found that a particularly successful procedure, ensuring the identification of true interactors, was to perform each experiment in at least two biological replicates, to score each replicate independently with SAINT (SAINT scores are between 0 and 1) and then average the independent scores (AvgP). By selecting AvgP ≥ 0.8, all interactions included are high-confidence, and have been detected in each of the biological replicates (this is the cutoff used for the data presented here).

Using the approaches described above, the components of the acetylation system (wt proteins) are each being expressed stably in the Flp-In T-REx cell lines, and AP-MS is being performed in biological duplicates. Together, the team (ACG, TP, PF) has already assembled a collection containing all lysine acetyltransferases and deacetylases, and 41 of the bromodomain-containing proteins. We are well on our way to transferring all of these cDNAs to the expression vector, generating stable cell lines, and performing AP-MS (the status of this project is represented on Fig 9). So far, 57 bait proteins have been analyzed by mass spectrometry in at least biological duplicates. This has led to the recovery of ~23000 pairs of unfiltered interactions, and ~1500 interactions amongst ~450 proteins at a SAINT AvgP ≥0.8 (Fig 10). As expected, we recover many published interactions, though the vast majority of the interactions detected here are novel. Which interactions are mediated by the bromodomain-acetyl-lysine interaction will be addressed in the proposal using a collection of point mutants of the critical asparagine residue (Fig 3) in the background of the full-length proteins (mutants already available are depicted in Fig 9). While we are still in the data acquisition phase of the project, we have begun analyzing in more depth some of the therapeutically-relevant interactions, such as those established by the BET family of bromodomain-containing proteins (Fig 11) and KAT3A/KAT3B (CREBBP; Fig 12). We have also detected a previously unknown interaction between the poorly characterized HDAC11 and the DICER1 protein (implicated in miRNA processing[65](#_ENREF_65)), which we validated (Fig 13).

**Determining specificity of bromodomains for acetyl-lysine residues**

Our interactome map is critical to understand the processes in which bromodomain-containing proteins and other acetylome components are acting; however, how the recognition of the acetylated lysine by the bromodomain is related to the interactome is unclear at this point. To begin to characterize interactome specificity, we are proposing a multi-pronged approach to identify acetylation-dependent interactions, both in the context of recombinantly-expressed isolated bromodomains (tested both on synthetic peptides and on cell extracts), but also within the context of the full length proteins expressed in human cells. A key element of our strategy is to be able to identify acetylation sites directly from immunoprecipitates of bromodomain-containing proteins. To start, we have modified our mChIP strategy to add, after the tryptic digest, an additional affinity step consisting of an anti-acetyl lysine pull-down at the peptide level and analysis on a high mass resolution mass spectrometer (Fig 14; we validated the specificity of the anti-KAc antibody on acetylated BSA). We first tested this approach on the bromodomain-containing protein BRPF3, a component of the MOZ/MORF histone acetyl-transferase complex[66](#_ENREF_66), [67](#_ENREF_67). In addition to the known BRPF3 interaction partners, our interaction proteomics approach (see above) identified new interactors for this protein. After the acetyl-lysine enrichment approach, we also found acetylation sites on BRPF3 itself, and on many of its interactors (Fig 15A, B). While several of these sites were present in a public repository (PhosphoSitePlus.org), 15 are new (and have been manually validated), highlighting the sensitivity of our approach. We also optimized the mass spectrometric identification of acetylated sites by analyzing fractions of the same purification on different mass spectrometers (and in the case of the Orbitrap Velos by employing different modes of fragmentation). Importantly, we show that – in comparison to a proteome-wide study of acetylation[68](#_ENREF_68) – our targeted study of the acetylome facilitates the identification of polyacetylated peptides (Fig 15C, D). This is an important result, especially in light of our recent study where we demonstrated that at least some of the bromodomains exhibit a preference for polyacetylated sites, both *in vitro* and *in vivo*, and that this preference may contribute to the recognition of histone code marks (Fig 16). It is also noteworthy that many of the lysine acetyltransferases (especially those of the MYST family) require auto-acetylation for activity. As such, being able to identify and study these critical residues is important. As detailed in the proposal, and highlighted in Appendix 1 and Fig 16, we have also developed approaches to determine the specificity of recombinant isolated bromodomains against synthetic peptides and histone preparations. The list of the recombinant bromodomains currently available for our studies is in Fig 9.

**Development of new methods for accurate mass spectrometry-based quantification**

An important aspect of the proposal is our ability to accurately quantify interactions regulated by the recognition of the acetylation site (e.g. from a parallel purification of a wild-type bromodomain-containing protein and a mutant in which the acetyl-lysine recognition is prevented). Furthermore, we wish to monitor the enrichment of acetylated peptides and quantify their relative abundance. While we and others have used several different quantitative proteomics approaches, including isotopic labeling approaches such as SILAC[69](#_ENREF_69), and label-free approaches such as spectral counting[48](#_ENREF_48), [70](#_ENREF_70), [71](#_ENREF_71) and intensity measurement of the precursor ion in the mass spectrometer (so-called MS1 quantitation[72](#_ENREF_72)), here we propose to use quantification methods based on the intensity of the product ions in the MS/MS spectrum. We have previously used single reaction monitoring (SRM) to monitor changes in interactomes accurately and with high sensitivity[52](#_ENREF_52) (Appendix 5). Since SRM method building (i.e. the selection of the best precursor ions and daughter ions to monitor) is time-consuming, we have also developed (in collaboration industrial partner AB SCIEX, letter appended) a modified approach called MRMHR, which accelerates the development of quantification assays while providing the same accuracy on quantification (Zheng et al., submitted). Lastly, we have implemented a data-independent quantification approach called SWATH[73](#_ENREF_73) which enables us to determine – again with high accuracy – the relative abundance of every identified species in a given dataset (Fig 17). Since the SWATH approach is very recent[73](#_ENREF_73), we benchmarked it here to determine its efficiency in measuring interactome differences on cancer-associated mutations[74](#_ENREF_74); this type of sample is very similar to what we are proposing to monitor in this proposal. As shown in Fig. 17, AP-MS with SWATH quantification reveals differences between the interactions established by wild type CDK4 (cyclin dependent kinase 4) and two point mutants identified in melanoma that act in a dominant manner[74](#_ENREF_74), [75](#_ENREF_75). As expected, we could readily detect the disruption of the interaction between the mutants and the cyclin dependent inhibitors of the INK family. Interestingly, however, we could also quantify increased interactions between the mutants and components of the HSP90 chaperone complex (suggesting folding issues with the mutants), and with another class of CDK inhibitor. These novel changes were validated by immunoblotting, confirming the usefulness of the approach. Importantly, quantification is accurate at the peptide level, which will enable us to monitor changes in acetylation. Further, the instrument on which we are performing SWATH measurements is the same as that used for acetylation site mapping, the AB-SCIEX TripleTOF 5600, ensuring the compatibility of our platform.

***Here, we propose to systematically define the specificity and function of human bromodomains through identification of their binding partners and acetylated targets. This knowledge will assist in the development of new inhibitors of bromodomains, in order to accelerate the development of new therapeutics****.*

**We propose the following specific aims:**

1) Identify the interactome for all components of the human acetylation machinery, including translocation products and mutated proteins

2) Define the specificity for each bromodomain-containing protein towards acetylated lysines

3) Reconstitute the acetylome specificity map

4) Test new bromodomain inhibitors

**Aim 1: Identify the interactome for all components of the human acetylation machinery, including translocation products and mutated proteins**

*As described in the preliminary data (and figures 9-13) we are well on our way to creating a map of the steady-state interactome for all full-length components of the acetylation machinery. Here, we will complete and validate this map, and expand it to the study of fusion proteins and point mutants involving the components of the acetylation machinery. Note that, while we strive to provide a global map of acetylome specificity, our team and our collaborators have a particular interest in the BET family of bromodomains, and we will therefore start all aims by studying this important group of proteins.*

* 1. Complete the interactome map for wt proteins

We will continue generating stable cell lines, and performing two biological replicate AP-MS analyses for each of the baits detailed in Figure 9; note that, based on the comparison of the instrument types displayed in Fig. 7, we will generate a complete dataset using more sensitive mass spectrometers (Orbitrap-Velos or 5600 TripleTOF). We are making a deliberate decision not to artificially increase the acetylation level in these cells via the addition of KDAC inhibitors; see Aim 2.1 for experiments in which these inhibitors will be used. When the entire dataset is generated, the data will be analyzed using our bioinformatics pipeline (with ProHits and SAINT) as detailed in the preliminary data section. While some of the cDNAs have been difficult to obtain due to their length, we (PF, collaboration with Stefan Knapp) have initiated a gene synthesis program to complete the dataset, and we therefore do not expect any major difficulty with this sub-aim. In parallel to the generation of the interaction data, we will be determining the subcellular localization in our stable cells for each of the proteins, as we are doing for our other projects; this is important, as many of the acetylome components have been completely uncharacterized so far. All data will be further analyzed for functional enrichment[76-78](#_ENREF_76) (with Dr. Gary Bader, letter attached) to provide additional information regarding the biological roles of the acetylation system (see Appendix 2 for an example of the type of analysis we are proposing to do). From the preliminary data that we have obtained thus far, we have detected interesting and previously unsuspected links between acetylome components and other cellular proteins. For example, HDAC11 interacts specifically with DICER1, raising the interesting hypothesis that acetylation may be implicated in the regulation of miRNA biology; as detailed in Aim 2.2, experiments are in progress to map potential acetylation sites on DICER1 and components of the miRNA pathway. It is intriguing that DICER1 has previously been linked to acetylation regulation but that the mechanisms are unknown[79](#_ENREF_79). In addition to generating a large interactome map in exponentially growing cells, we will, when appropriate, investigate condition-specific interactions. In the case of the BET family of bromodomain-containing proteins that are recruited to chromatin during mitosis[21](#_ENREF_21), [80](#_ENREF_80), we will also perform comparative analysis of the interactions in asynchronous cells, and in cells arrested with paclitaxel (Taxol; 5ug/ml for 16 hours) or nocodazole (100ng/ml for 16 hours); this may assist in determining regulated interactions and provide a more dynamic view of the protein-protein interactions mediated by acetylation changes (these experiments will be performed using quantitative proteomics, as described in Fig 17).

* 1. Validate interactions and define proteins complexes interacting with acetylome components

While we have now demonstrated that our experimental and bioinformatics pipeline enables us to accurately identify bona fide interaction partners, as opposed to contaminants, in AP-MS data (we are initially using especially stringent filtering criteria for the acetylome project presented here to ensure that follow-up experiments are performed on valid interactions), we are interested in defining how many independent protein complexes can associate with the interactome. We have previously determined[49](#_ENREF_49), [51](#_ENREF_51), [70](#_ENREF_70) that iterative AP-MS (in which a prey identified in a round of AP-MS is in turn cloned and used as a bait in the next round of AP-MS) is an efficient method to define the modularity of the interactome, and identify mutually exclusive and cooperative interactions[81](#_ENREF_81). Here, up to 50 hits from the first round of AP-MS (selected based on the biological functions and the complexity of the interactomes) will be used for this purpose, to further define the assembly of the acetylation system. Additional interactions of biological interest (see, e.g. Fig. 12) will be further validated by IP-western, and co-localization in human cells, as we have done extensively in the past. Biological follow-up and structure/function analysis of new protein complexes will be performed as we recently did for a kinase/phosphatase complex linked to Cerebral Cavernous Formations, STRIPAK[70](#_ENREF_70), [72](#_ENREF_72), [82](#_ENREF_82).

1.3 Determine the consequences of genetic aberration of the acetylome on the recruitment of interaction partners.

We have established a quantitative proteomics approach (which we refer to as AP-SWATH) that is particularly suited for unbiased interactome analysis of wt and mutant proteins. AP-SWATH was developed to speed-up the study of the functional consequences of a mutation on the interactome (Lambert et al., *in prep*; Fig 17). Our central hypothesis in developing this project is that the mutant or fusion protein may either acquire or lose interactions which were detected with the wild type protein, and that these interactions will help to understand the function of the protein in healthy and diseased cells (such mutations have been referred to as “edgetic” mutations[74](#_ENREF_74)). Here, AP-SWATH will be applied to compare the interactome of the acetylome components, of their point mutants, of fusion products involving the acetylome components, and of their translocation partners. We have already begun generating stable lines of BRD4, the BRD4-NUT fusion protein, and its fusion partner NUT, a poorly characterized protein (implicated in aggressive carcinoma; collaboration with C.A. French). This will be expanded to other known translocation products within the acetylome (Fig 2; Table 1), such as the KAT6B-KAT3A fusion[4](#_ENREF_4) which our team is committed to clone, as we have done with the full length, wild-type, proteins. A caveat to the use of the fusion proteins is that their expression levels may differ from that of the wild-type protein; in fact, we have already observed this phenomenon with the BRD4-NUT fusions (not shown). To prevent inducing or losing interactions based on expression alone, we will perform parallel AP-SWATH experiments in which the induction of protein expression with tetracycline will be performed using different concentrations of tetracycline or different expression times[59](#_ENREF_59). Robustness of the interactome identified after driving different expression levels will be monitored for those fusion proteins (and their wt counterparts) for which expression level differences are more dramatic. We will continue collaborating with Alexey Nesvizhskii (letter attached) on scoring these potentially more difficult interactions. The biological consequences of interaction with translocation products or mutant proteins will be investigated in appropriate systems. For example, for fusion proteins such as BRD4-NUT which is essential for transformation in midline carcinoma, we will test whether the interactions are essential for the transformation phenotype by performing transient or stable knockdowns of the interaction partners in NMC cells (we have access through CA French to several patient-derived lines: 8645[83](#_ENREF_83), 11060[30](#_ENREF_30), 10326 (a BRD3-NUT line[84](#_ENREF_84)), 14169, and 879).

1.4 Map the interactome for wild-type and cancer-associated mutant bromodomain-containing proteins expressed at endogenous levels.

While our growing experience indicates that HEK293s constitute an excellent system for interaction mapping, they may not express all proteins expressed in different types of cancers. Also, while the interactions detected by AP-MS in this cell model are usually robust, it is important to be able to validate that the interactions do occur amongst endogenous proteins. We have therefore established a collaboration with Sachdev Sidhu (letter attached) to generate phage affinity reagents to different biotinylated and soluble bromodomains (we have successfully employed such reagents for AP-MS). Currently, 59 high-affinity Fabs have been generated against 14 bromodomains (Fig 9), and this collection is growing. After determining which of these Fabs are inhibitory to the bromodomain-KAc interaction and which are not, non-inhibitory sequences will be transferred into a full length IgG scaffold to use as affinity reagents. Ideally, we would want to test in parallel several affinity reagents per target, first on the stable cell lines, each expressing a full length tagged bromodomain-containing protein (obtaining multiple reagents is important for discerning the background from the true interaction partners in AP-MS experiments; inhibitory Fabs will be used in Aim 4). Once the affinity reagents are validated, they will be tested for AP-MS on endogenous protein interactions, first in an HEK293 cell lysate, next on other cellular sources, including cancer cell lines (including the NMC cells described above) which our groups (TP, ACG) use routinely, as appropriate. Other antibodies to endogenous proteins are available, including to the NUT fusion partner (different antibodies are available, both commercially, and through CA French). As we have done extensively in the past, we will also systematically validate new and interesting interactions by performing immunoprecipitation followed by immunoblotting with endogenous proteins, using either the phage affinity reagents described here, or standard antibodies.

**Aim 2: Define the specificity for each bromodomain-containing protein toward acetylated lysines**

*Bromodomains are the only known domain that recognize the KAc group, and are expected to add specificity to the acetylation system, yet, little is known regarding their individual specificity for acetylated peptides. As detailed in Appendix 1, we have used pull-downs with biotinylated recombinant bromodomains to define the specificity for acetylated peptides on spot arrays, in ITC measurements, and after enrichment from histone preps (in combination with mass spectrometric analysis). We have also optimized an approach to combine affinity purification with acetyl-lysine enrichment and mass spectrometry (AP-KAc-MS; Fig 14). Here, we will use a multi-pronged approach to define the specificity of each bromodomain-containing protein for specific acetylated lysine peptides.*

2.1 Define which interactions are mediated by the bromodomains

While Aim 1 will provide a global map of the interactome for bromodomain-containing proteins, some or many of the interactions we are detecting may not be mediated by a bromodomain-KAc interaction. Indeed, as shown in Fig 2, many of the bromodomain-containing proteins contain several additional protein interaction domains; furthermore, some of them contain more than a single bromodomain. As shown on Fig 9, we have generated stable cell lines for the expression of point mutants within bromodomains which disrupt the interaction with the acetylated target proteins. Guided by the results of Aim 1 (using the bromodomain-containing proteins that recover rich interactomes in our cell system), we will generate additional mutants and cells lines. As detailed in Aim 1.3, we will perform AP-SWATH on pairs of wt and mutant bromodomains (all within the context of the full-length proteins). In order to potentially increase the number of KAc-mediated interactions, the steady-state level of KAc in the cells will be increased in some of the experiments by the addition of Trichostatin A (TSA; 200nM for 6 hours); parallel purifications of the treated and untreated cells both in the context of the wt and the bromodomain mutant protein will be performed. This will enable the definition of the bromodomain-associated interactome, i.e. those interactions that are mediated by the direct interaction of a bromodomain with an acetyl-lysine substrate.

2.2 Identify acetylation sites on acetylome-associated proteins.

As shown in Fig 15 (with BRPF3), we have been successful at developing an approach to systematically map acetylated sites on binding partners of KATs, KDACs and bromodomain-containing proteins. We are now expanding this analysis to each member of the BET family (BRD2, BRD3, BRD4, BRDT; each of these proteins contains two bromodomains). Initially, the samples will be lysed in the presence of sodium butyrate, a generic deacetylase inhibitor to stabilize KAc sites prior to their purification (again, in some experiments, acetylation will be enhanced by treatment of the cells with KDAC inhibitors). Since the BET-bromodomain specific inhibitor JQ1 efficiently displaces those interactions that depend on the KAc-bromodomain interaction, parallel experiments will be performed on cells treated with the (+)-JQ1 and the inactive (-)-JQ1 compounds. Proteins detected in the absence of JQ1, but not in its presence will be deemed likely bromodomain-KAc-mediated interactors; their KAc sites will be inspected closely, and consensus sites will be determined (see 2.3). In parallel to the purification of the BET-associated KAc peptides, we will determine the acetylation sites in purifications of other acetylome components. As detailed above, we are particularly interested in determining whether DICER1 itself is acetylated (based on the recovery of DICER1 in the purifications of HDAC11); this will initially be monitored by immunoblotting of affinity-purified DICER1 using anti-acetyl-lysine antibodies. If we indeed do detect reactivity with the KAc antibody, the role of HDAC11 in mediating deacetylation will be tested by silencing HDAC11 (and other HDACs as controls) and monitoring DICER acetylation level. Sites will be mapped using the method defined above.

2.3 Define the specificity of bromodomains in isolation for peptides and proteins.

As shown in Fig 15 and Appendix 1, we have used biotinylated bromodomains to define KAc specificity preference for BRD4 bromodomain 1 and BRD4 bromodomain 2. This approach will be repeated, starting from purified histone preparations, first on each of the bromodomains from the BET family. In addition to using purified histones as starting material, we will employ tryptic peptides generated from a mixture of cell lines (e.g. K562, HeLa, U2OS and MCF-7), to insure a diverse source of starting material. We expect that if we perform the pull-down at the peptide level (rather than at the protein level like we have done so far) we may not need to couple our approach with anti-KAc enrichment; if this is not the case, we will perform a dual purification). To help identify relevant targets, we will incorporate a SILAC isotopic-labeling approach to generate pools of heavy and light lysates which will be used for these experiments. As controls, we will use validated inhibitors when available (e.g. the JQ1 inhibitor which we have used for BRD4; also see Aim 3), bromodomain inhibitory Fabs (if available) (Aim 4) and mutants in the bromodomains themselves. This is important, as peptide interactions can likely occur with surfaces outside of the KAc-binding pocket. Resulting enriched peptide sequences for each of the isolated bromodomains will be analyzed for the presence of consensus motifs, in collaboration with Gary Bader[85](#_ENREF_85). We are initially focusing on the BET family of bromodomains since these are the targets of our intensive structural analysis and drug development (PF, collaboration S Knapp). Note that our mass spectrometric experiments enable us to look at combinatorial modifications on the KAc-peptides, providing that they are in close proximity; all the mass spectrometry data will be searched for potential multi-modifications (KAc, Lysine methylation, phosphoS/T/Y, etc.) and the motif enrichment analysis will also take combinatorial modifications into account. This will likely be important for the other bromodomains that exhibit preference for peptides that are multiply modified (Appendix 1). If we detect a bias for peptides harboring a different modification (e.g. phosphorylation), we could combine the approach with another affinity step for the particular modification (e.g. IMAC enrichment for phosphorylation[86](#_ENREF_86)). Once the data analysis on the BET family is completed, we will systematically apply the approach to all other bromodomains, again, prioritizing the analysis to targets of biological interest in cancer (clear candidates are the KAT3A/KAT3B bromodomains). Importantly, this work will be done in the context of a parallel effort (PF) to characterize specificity based on *in vitro* spot assays and isothermal calorimetry measurements (see Appendix 1). Advantages of the pull-down from lysates include the possibility of detection of combinatorially modified peptides, and detection of peptides from real proteins, accelerating the study of the functional roles of the bromodomains. Peptides corresponding to the new sequences determined by mass spectrometry will be synthesized (PF) for SPOT assays followed by biophysical measurements and structural analysis, as appropriate. For bromodomains for which the substrate is completely unknown, this approach will enable the team (PF) to set up primary screens for small molecule binding (by ALPHA screen or surface plasmon resonance) utilizing the peptides identified and displacing them using small molecule scaffolds[30](#_ENREF_30).

**Aim 3. Reconstitute a specificity map for the acetylome**

*In Aim 2, we focused specifically on the reader component of the acetylation machinery, the bromodomain, as these proteins are expected to be key to the specificity of the acetylome, and are the target of our small molecule generation efforts. However, we recognize that to fully define the acetylome, the specificity of each of the KATs and KDACs for their substrates will need to be better defined; this is particularly relevant as we and others have identified several protein-protein interactions amongst the writer-eraser-reader components of the system (Fig 18), suggesting an interplay between these modules, and/or a coordination of the acetylation system (interestingly, we previously noted such an enrichment of kinase-kinase interactions in a systematic analysis of the phosphorylation system in yeast – Appendix 2). Here we will attempt to further refine the relationships between the acetylome components.*

3.1. Identification of targets of KATs and KDACs. Whenever possible, we will use literature-curated data as they become available to identify the targets of KATs and KDACs (this is an active area of research, and we expect to see several reports on this throughout the duration of the grant). However, it is possible that for one or multiple KATs or KDACs of interest, no data will be available. If this were the case, we would employ the same approach employed recently for finding SIRT1 substrates[87](#_ENREF_87), and identify the target proteins by quantitative mass spectrometry. Briefly, cell extracts will be treated with N‐succinimidyl N‐methylcarbamate to cap free lysines, the sample split into two equal aliquots and recombinant FLAG-tagged KDAC from human cells added to one aliquot. Newly exposed lysines will then be labeled with Sulfo‐NHS‐SS‐Biotin, the peptides purified with streptavidin beads and the precise acetylation sites detected by mass spectrometry. To confirm novel KAT substrates, we will employ *in vitro* KAT assays, in which FLAG-tagged recombinant KAT will be incubated with affinity-purified putative substrates and heavy labeled acetyl-CoA (containing two C13 atoms), as was performed by Dr. Lambert during his PhD (unpublished work). Following mass spectrometry analysis, we will be able to identify both *in vivo* (light KAc) and *in vitro* (heavy KAc) acetylation sites on a given protein (a caveat is that if the sites are fully occupied, we may need to strip them first, e.g. by employing deacetylases). Finally the intersection between the list of KAT / KDAC substrates and bromodomain-specific interactions will be analyzed computationally to provide an overview of the acetylome.

3.2 Mapping of KDAC/bromodomain relationships. HDAC inhibitors are used in the clinic, but there still remains much to learn regarding their specificity and functional consequences on cells. As shown in Fig 18, we have identified several physical interactions between KDACs and bromodomain-containing proteins in our network. Here we propose to further study these relationships by investigating the consequence(s) of depletion of specific KDACs using RNAi, or their inhibition by specific KDAC/HDAC inhibitors, on the interactions established by individual bromodomains. For example, we detected a physical interaction between BRD3 and HDAC2; HDAC2 will be silenced by RNAi, or inhibited by romidepsin (a Class I HDAC inhibitor), and the interactions established by FLAG-BRD3 will be quantified by AP-SWATH. If the physical interaction between HDAC2 and BRD3 is functional, we expect depletion of HDAC2 to lead to an increased association of HDAC2 targets to BRD3. We will integrate the data from Aims 1-3 and the literature (as in[39](#_ENREF_39), [88](#_ENREF_88)), and build functional association maps for the writer-eraser-reader system. This should provide a much improved view of the acetylation system in human cells, better setting the stage for target identification for future drug design.

**Aim 4. Test new small molecule inhibitors of bromodomains**

*As shown in Fig 4, we (PF, collaboration with S Knapp) have already demonstrated that new inhibitors targeted to bromodomains can show selectivity and exquisite efficacy in a mouse tumor model. We have therefore embarked upon a systematic effort to generate new bromodomain inhibitors, guided in this by the crystal structures we have determined (Appendix 1). What these inhibitors have in common is that they target the KAc binding pocket, thereby disrupting interactions with bromodomain substrates. Here, we will capitalize on the team expertise to accelerate the testing of first generation molecules. Inhibitors targeting KAT3A (BDOIA298; 105 nM by SPR), BAZ2B (BDG00021335; 96nM IC50 on an ALPHA assay displacing a H3K14 peptide) and BRD9 (BDOTC000133a; 300nM by SPR) will be available for testing within the next 6 months (1 or 2 distinct scaffolds per protein), and the team aims to generate and test inhibitors for 10 targets during the duration of this grant. These small molecule chemical probes have in vitro affinity in the order of 100 nM and in vivo cell efficacy at 1-10μM and are being developed as part of the SGC’s epigenetics probe program (http://www.thesgc.org/scientists/chemical\_probes). The testing platform developed here will also be employed to monitor inhibitory Fabs as they become available.*

4.1 *In vitro* testing of interactome disruption. As the specificity of each bromodomain for KAc-peptides is determined (Aim 2.3) and validated *in vitro*, synthetic peptides derived from these target sites will first be employed for affinity measurements and initial testing of the different drug-like molecules using our standard platforms (Appendix 1[30](#_ENREF_30), [89](#_ENREF_89)). Specificity will be determined by employing soluble bromodomains from related and distant families. We (PF) have extensive expertise in this area and do not anticipate major difficulties. This initial screening will be followed by confirmatory pull-down experiments in a cell lysate (as in Aim 2.3) that interactions with all KAc-peptides is indeed abrogated (quantification by AP-SWATH will be performed). If the drug works as expected, there should be a marked reduction of KAc-peptide interaction, as we demonstrated for each of the isolated domains from BRD4 after addition of the JQ1 inhibitor (Fig 11, 16; Appendix 1). In addition to the small molecule inhibitors developed by the team, Fabs generated in collaboration with Sachdev Sidhu (letter appended) will be tested in the same system. Inhibitory Fabs will be further tested in Aim 4.3.

4.2 Disruption of the interactome within the context of full-length proteins. The SGC pipeline strives at making potent, selective and cell permeable chemical probes, and the compounds we will be testing here will have already been subjected to extensive testing in Oxford. If the inhibitors are membrane permeable (as is JQ1), they will be used to treat stable cell lines expressing the target bromodomain, and parallel purification and LC-MS/MS analysis of cells treated or not with the inhibitor will be performed (as in Fig 16). This is important to ensure that within the context of a living cell, the inhibitor is sufficient for dissociation. In the event that the first generation inhibitor is not cell permeable, we will test whether it can disrupt complex formation in digitonin-permeabilized cells[90](#_ENREF_90). Alternatively, we will determine whether the inhibitor can disrupt interactions pre-established in a cell lysate. For cell permeable inhibitors, we will perform parallel measurements of proliferation rates with different concentrations of inhibitors and AP-MS or IP/western analysis. Taken together, these results should enable the team to refine the drug design and/or focus on the most promising chemical probes.

4.3 Intracellular scFvs as molecular probes. Fab sequences demonstrated to be potent for the disruption of KAc-bromodomain interactions and proven to be specific for a given bromodomain would constitute outstanding probes, if delivered within the cell, to accelerate the understanding of the molecular functions of the KAc-bromodomain interaction. Our collaborator Sachdev Sidhu will transfer the sequences we validate in Aim 4.1 into an expression system suitable for the expression of single-chain Fv molecules in mammalian cells; these ScFvs will be affixed to an epitope tag (we will use a dual epitope comprised of HA and StreptagII; SH tag), and transfected into our stable cell lines expressing the cognate FLAG-tagged bromodomain-containing protein (initially, we will perform transient transfection). We will be monitoring first that the SH-tagged scFv can precipitate the FLAG-tagged protein by IP/western. If this is the case, we will assess whether it exhibits specificity by comparing the interactions established by the SH-scFv to those of the FLAG-bromodomain-containing protein (they should be the same). After this validation, parallel comparisons of the interactomes of the FLAG-tagged protein in the presence or absence of the SH-scFv protein will reveal which portion of the interactome is displaced by the scFv. The biological role of the scFv will then be directly tested by analyzing the consequences of its overexpression using appropriate readouts for the given bromodomain-KAc interactions being disrupted (e.g. changes in proliferation, etc.).

**Perspectives**

In summary, our project will provide researchers worldwide with several valuable datasets. First, we will be providing a comprehensive and high quality for all the components of the acetylation machinery. This interaction map will be unique: to the best of our knowledge, all other laboratories performing high-throughput AP-MS are focusing on the soluble fraction, and miss the interactions involving chromatin components, which are essential to understand the function of the acetylation machinery. Importantly, this map will also extend to fusion products mutated in cancers, which could help providing a better understanding of the mechanisms of cellular transformation by these oncogenes. Next, we will provide a map of the specificity for the recognition of acetylated lysine by bromodomains, which will complement the data we published from on peptide array efforts (Appendix 1). This is important, since many of the bromodomains do not currently have any known target. This information should enable us to uncover additional relationships between the components of the acetylation machinery (KAT, KDAC, bromodomain) and crosstalk to cellular signaling. Lastly, we will use our toolset to test new chemical (and Fab) bromodomain inhibitors. The strength of this proposal is that there is a strong synergy between team members and integration between the Toronto (proteomics) and Oxford (structural and chemical biology) groups: Essentially, new interactions or KAc recognition sites discovered here will fuel the determination of additional co-crystal structures, allowing for the development of new small molecule inhibitors. The reagents generated for the structural analysis (e.g. biotinylated recombinant proteins) will in turn be used for additional proteomics analysis here. Importantly, the new small molecule inhibitors of specific bromodomains will be rapidly tested in the cognate cell line / biotinylated protein assay in Toronto, accelerating the development of well-characterized molecular probes that can be used to perturb and better understand the biological systems in which bromodomains participate.

**The team**

The team is led by three dedicated researchers: at the SLRI (Toronto), AC Gingras (start of independent career in 2005) and Distinguished Investigator T Pawson specialize in functional proteomics and their application to the study of interaction specificity and signaling events. P Filippakopoulos (start of independent career in 2011) at the Structural Genomics Consortium in Oxford is an expert in structural biology and rational drug design (see letter of collaboration by the SGC group leader in structural biology, Stefan Knapp). PF has already received funding for his part of the project (structural biology and inhibitor development), and we are therefore only requesting funds for the Toronto component of the research project (proteomics). TP and ACG co-direct the SLRI mass spectrometry facility, and have built through capital equipment grants and an intensive collaboration with the mass spectrometry vendor AB-SCIEX the infrastructure necessary for the completion of this project. All co-applicants (including Dr. Lambert) are co-authors on the appended manuscript in *Cell*, which provides the foundation for much of Aims 2 and 3, and have a very active and open collaboration, which will be maintained via frequent webinars and yearly visits. In Toronto, the work will be carried out by a postdoc co-supervised by TP/ACG, JP Lambert who brings a strong expertise in proteomics of chromatin-associated complexes, by B Badillo (cloning and cell line generation), A Veri (endogenous interactions in cancer cells and pharmacology screens), B Larsen (SWATH mass spectrometry), M Tucholska (mass spectrometric analysis) and J Zhang (data analysis). The team will be rounded up by other SGC-Oxford team members, S Picaud (recombinant protein engineering) and I Felletar (protein production and biophysical assays). Lastly, we have enlisted the help of a number of colleagues to provide additional expertise to enable us to accomplish all the aims in the short time frame (4-years) of funding requested.

**REFERENCES**

1. Kouzarides, T. Acetylation: a regulatory modification to rival phosphorylation? *Embo J* **19**, 1176-1179 (2000).

2. Lim, W.A. & Pawson, T. Phosphotyrosine signaling: evolving a new cellular communication system. *Cell* **142**, 661-667 (2010).

3. Pawson, T. & Nash, P. Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445-452 (2003).

4. Muller, S., Filippakopoulos, P. & Knapp, S. Bromodomains as therapeutic targets. *Expert Rev Mol Med* **13**, e29 (2011).

5. Allfrey, V.G., Faulkner, R. & Mirsky, A.E. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* **51**, 786-794 (1964).

6. Strahl, B.D. & Allis, C.D. The language of covalent histone modifications. *Nature* **403**, 41-45 (2000).

7. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).

8. 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**, 263-283 (2011).

9. Zhao, S. et al. Regulation of cellular metabolism by protein lysine acetylation. *Science* **327**, 1000-1004 (2010).

10. Sobulo, O.M. et al. MLL is fused to CBP, a histone acetyltransferase, in therapy-related acute myeloid leukemia with a t(11;16)(q23;p13.3). *Proc Natl Acad Sci U S A* **94**, 8732-8737 (1997).

11. Troke, P.J., Kindle, K.B., Collins, H.M. & Heery, D.M. MOZ fusion proteins in acute myeloid leukaemia. *Biochem Soc Symp*, 23-39 (2006).

12. Vizmanos, J.L. et al. t(10;16)(q22;p13) and MORF-CREBBP fusion is a recurrent event in acute myeloid leukemia. *Genes Chromosomes Cancer* **36**, 402-405 (2003).

13. Kalla, C. et al. Translocation t(X;11)(q13;q23) in B-cell chronic lymphocytic leukemia disrupts two novel genes. *Genes Chromosomes Cancer* **42**, 128-143 (2005).

14. Mann, B.S., Johnson, J.R., Cohen, M.H., Justice, R. & Pazdur, R. FDA approval summary: vorinostat for treatment of advanced primary cutaneous T-cell lymphoma. *Oncologist* **12**, 1247-1252 (2007).

15. Piekarz, R.L. et al. Phase II multi-institutional trial of the histone deacetylase inhibitor romidepsin as monotherapy for patients with cutaneous T-cell lymphoma. *J Clin Oncol* **27**, 5410-5417 (2009).

16. French, C.A. et al. BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. *Cancer Res* **63**, 304-307 (2003).

17. French, C.A. Pathogenesis of NUT Midline Carcinoma. *Annu Rev Pathol* **7**, 247-265 (2012).

18. French, C.A. Demystified molecular pathology of NUT midline carcinomas. *J Clin Pathol* **63**, 492-496 (2010).

19. Reynoird, N. et al. Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. *Embo J* **29**, 2943-2952 (2010).

20. Wu, S.Y. & Chiang, C.M. The double bromodomain-containing chromatin adaptor Brd4 and transcriptional regulation. *J Biol Chem* **282**, 13141-13145 (2007).

21. Zhao, R., Nakamura, T., Fu, Y., Lazar, Z. & Spector, D.L. Gene bookmarking accelerates the kinetics of post-mitotic transcriptional re-activation. *Nat Cell Biol* **13**, 1295-1304 (2011).

22. Rahman, S. et al. The Brd4 extraterminal domain confers transcription activation independent of pTEFb by recruiting multiple proteins, including NSD3. *Mol Cell Biol* **31**, 2641-2652 (2011).

23. Delmore, J.E. et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell* **146**, 904-917 (2011).

24. Mertz, J.A. et al. Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc Natl Acad Sci U S A* **108**, 16669-16674 (2011).

25. Zuber, J. et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. *Nature* **478**, 524-528 (2011).

26. Yan, J., Diaz, J., Jiao, J., Wang, R. & You, J. Perturbation of BRD4 protein function by BRD4-NUT protein abrogates cellular differentiation in NUT midline carcinoma. *J Biol Chem* **286**, 27663-27675 (2011).

27. Mujtaba, S., Zeng, L. & Zhou, M.M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26**, 5521-5527 (2007).

28. Dhalluin, C. et al. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-496 (1999).

29. Owen, D.J. et al. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *Embo J* **19**, 6141-6149 (2000).

30. Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067-1073 (2010).

31. Nicodeme, E. et al. Suppression of inflammation by a synthetic histone mimic. *Nature* **468**, 1119-1123 (2010).

32. Borah, J.C. et al. A small molecule binding to the coactivator CREB-binding protein blocks apoptosis in cardiomyocytes. *Chem Biol* **18**, 531-541 (2011).

33. Gerona-Navarro, G. et al. Rational design of cyclic peptide modulators of the transcriptional coactivator CBP: a new class of p53 inhibitors. *J Am Chem Soc* **133**, 2040-2043 (2011).

34. Zeng, L. et al. Selective small molecules blocking HIV-1 Tat and coactivator PCAF association. *J Am Chem Soc* **127**, 2376-2377 (2005).

35. Moriniere, J. et al. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**, 664-668 (2009).

36. Chen, C., Nott, T.J., Jin, J. & Pawson, T. Deciphering arginine methylation: Tudor tells the tale. *Nat Rev Mol Cell Biol* **12**, 629-642 (2011).

37. Pawson, T. & Kofler, M. Kinome signaling through regulated protein-protein interactions in normal and cancer cells. *Curr Opin Cell Biol* **21**, 147-153 (2009).

38. Filippakopoulos, P. et al. Structural coupling of SH2-kinase domains links Fes and Abl substrate recognition and kinase activation. *Cell* **134**, 793-803 (2008).

39. Linding, R. et al. NetworKIN: a resource for exploring cellular phosphorylation networks. *Nucleic Acids Res* **36**, D695-699 (2008).

40. Hardy, W.R. et al. Combinatorial ShcA docking interactions support diversity in tissue morphogenesis. *Science* **317**, 251-256 (2007).

41. Wells, C.D. et al. A Rich1/Amot complex regulates the Cdc42 GTPase and apical-polarity proteins in epithelial cells. *Cell* **125**, 535-548 (2006).

42. Jones, N. et al. Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. *Nature* **440**, 818-823 (2006).

43. Gingras, A.C. et al. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev* **13**, 1422-1437 (1999).

44. Gingras, A.C., Kennedy, S.G., O'Leary, M.A., Sonenberg, N. & Hay, N. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. *Genes Dev* **12**, 502-513 (1998).

45. Gingras, A.C. et al. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev* **15**, 2852-2864 (2001).

46. Filippakopoulos, P., Muller, S. & Knapp, S. SH2 domains: modulators of nonreceptor tyrosine kinase activity. *Curr Opin Struct Biol* **19**, 643-649 (2009).

47. Dunham, W.H. et al. A cost-benefit analysis of multidimensional fractionation of affinity purification-mass spectrometry samples. *Proteomics* **11**, 2603-2612 (2011).

48. Breitkreutz, A. et al. A global protein kinase and phosphatase interaction network in yeast. *Science* **328**, 1043-1046 (2010).

49. Gingras, A.C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* **8**, 645-654 (2007).

50. Chen, G.I. & Gingras, A.C. Affinity-purification mass spectrometry (AP-MS) of serine/threonine phosphatases. *Methods* **42**, 298-305 (2007).

51. Gingras, A.C. et al. A novel, evolutionarily conserved protein phosphatase complex involved in cisplatin sensitivity. *Mol Cell Proteomics* **4**, 1725-1740 (2005).

52. Bisson, N. et al. Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. *Nat Biotechnol* **29**, 653-658 (2011).

53. Chen, C. et al. Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc Natl Acad Sci U S A* **106**, 20336-20341 (2009).

54. Ingham, R.J. et al. WW domains provide a platform for the assembly of multiprotein networks. *Mol Cell Biol* **25**, 7092-7106 (2005).

55. Smith, M.J., Kulkarni, S. & Pawson, T. FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. *Mol Cell Biol* **24**, 9274-9285 (2004).

56. Jin, J. et al. Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. *Curr Biol* **14**, 1436-1450 (2004).

57. Ho, Y. et al. Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature* **415**, 180-183 (2002).

58. Skarra, D.V. et al. Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5. *Proteomics* **11**, 1508-1516 (2011).

59. Glatter, T., Wepf, A., Aebersold, R. & Gstaiger, M. An integrated workflow for charting the human interaction proteome: insights into the PP2A system. *Mol Syst Biol* **5**, 237 (2009).

60. Lambert, J.P., Pawson, A.J. & Gingras, A.C. Mapping physical interactions within chromatin by proteomic approaches. *Proteomics* **in press** (2012).

61. Lambert, J.P. et al. Defining the budding yeast chromatin-associated interactome. *Mol Syst Biol* **6**, 448 (2010).

62. Lambert, J.P., Mitchell, L., Rudner, A., Baetz, K. & Figeys, D. A novel proteomics approach for the discovery of chromatin-associated protein networks. *Mol Cell Proteomics* **8**, 870-882 (2009).

63. Liu, G. et al. ProHits: integrated software for mass spectrometry-based interaction proteomics. *Nat Biotechnol* **28**, 1015-1017 (2010).

64. Choi, H. et al. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat Methods* **8**, 70-73 (2011).

65. Ji, X. The mechanism of RNase III action: how dicer dices. *Curr Top Microbiol Immunol* **320**, 99-116 (2008).

66. Ullah, M. et al. Molecular architecture of quartet MOZ/MORF histone acetyltransferase complexes. *Mol Cell Biol* **28**, 6828-6843 (2008).

67. Doyon, Y. et al. ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol Cell* **21**, 51-64 (2006).

68. Choudhary, C. et al. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* **325**, 834-840 (2009).

69. Jorgensen, C. et al. Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science* **326**, 1502-1509 (2009).

70. Goudreault, M. et al. A PP2A phosphatase high density interaction network identifies a novel striatin-interacting phosphatase and kinase complex linked to the cerebral cavernous malformation 3 (CCM3) protein. *Mol Cell Proteomics* **8**, 157-171 (2009).

71. Chen, G.I. et al. PP4R4/KIAA1622 forms a novel stable cytosolic complex with phosphoprotein phosphatase 4. *J Biol Chem* **283**, 29273-29284 (2008).

72. Kean, M.J. et al. Structure-function analysis of core STRIPAK Proteins: a signaling complex implicated in Golgi polarization. *J Biol Chem* **286**, 25065-25075 (2011).

73. Gillet, L.C. et al. Targeted data extraction of the MS/MS spectra generated by data independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* (2012).

74. Zhong, Q. et al. Edgetic perturbation models of human inherited disorders. *Mol Syst Biol* **5**, 321 (2009).

75. Zuo, L. et al. Germline mutations in the p16INK4a binding domain of CDK4 in familial melanoma. *Nat Genet* **12**, 97-99 (1996).

76. Merico, D., Isserlin, R. & Bader, G.D. Visualizing gene-set enrichment results using the Cytoscape plug-in enrichment map. *Methods Mol Biol* **781**, 257-277 (2011).

77. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).

78. Warde-Farley, D. et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* **38**, W214-220 (2010).

79. Giles, K.E., Ghirlando, R. & Felsenfeld, G. Maintenance of a constitutive heterochromatin domain in vertebrates by a Dicer-dependent mechanism. *Nat Cell Biol* **12**, 94-99; sup pp 91-96 (2010).

80. Dey, A. et al. A bromodomain protein, MCAP, associates with mitotic chromosomes and affects G(2)-to-M transition. *Mol Cell Biol* **20**, 6537-6549 (2000).

81. Choi, H., Kim, S., Gingras, A.C. & Nesvizhskii, A.I. Analysis of protein complexes through model-based biclustering of label-free quantitative AP-MS data. *Mol Syst Biol* **6**, 385 (2010).

82. Ceccarelli, D.F. et al. CCM3/PDCD10 heterodimerizes with germinal center kinase III (GCKIII) proteins using a mechanism analogous to CCM3 homodimerization. *J Biol Chem* **286**, 25056-25064 (2011).

83. Schwartz, B.E. et al. Differentiation of NUT midline carcinoma by epigenomic reprogramming. *Cancer Res* **71**, 2686-2696 (2011).

84. French, C.A. et al. BRD-NUT oncoproteins: a family of closely related nuclear proteins that block epithelial differentiation and maintain the growth of carcinoma cells. *Oncogene* **27**, 2237-2242 (2008).

85. Tonikian, R. et al. A specificity map for the PDZ domain family. *PLoS Biol* **6**, e239 (2008).

86. Bodenmiller, B. & Aebersold, R. Quantitative analysis of protein phosphorylation on a system-wide scale by mass spectrometry-based proteomics. *Methods Enzymol* **470**, 317-334 (2010).

87. Andersen, J.L. et al. A biotin switch-based proteomics approach identifies 14-3-3zeta as a target of Sirt1 in the metabolic regulation of caspase-2. *Mol Cell* **43**, 834-842 (2011).

88. Linding, R. et al. Systematic discovery of in vivo phosphorylation networks. *Cell* **129**, 1415-1426 (2007).

89. Filippakopoulos, P. et al. Benzodiazepines and benzotriazepines as protein interaction inhibitors targeting bromodomains of the BET family. *Bioorg Med Chem* (2011).

90. Chi, N.C., Adam, E.J., Visser, G.D. & Adam, S.A. RanBP1 stabilizes the interaction of Ran with p97 nuclear protein import. *J Cell Biol* **135**, 559-569 (1996).