STRUCTURAL CHARACTERIZATION AND CHEMICAL INHIBITION OF THE
ARNT/TACC3 COMPLEX

APPROVED BY SUPERVISORY COMMITTEE

NOTE: The top line is for the
Supervising Professor's name. There
should be as many lines as there are
members of the committee. All
signatures must be original and in ink.
Adjust "Approved by Supervisory
Committee" line upward if the
committee list is very large.

_____
Dr. Kevin H. Gardner

_____
Dr. Daniel M. Rosenbaum

_____
Dr. Michael G. Roth

_____
Dr. Luke M. Rice

DEDICATION

I greatly thank my mentor, Dr. Kevin H. Gardner, for his considerable support and guidance over the past five years. His passion in science, creativity and constructive criticism was the critical driving force for this project, and will always influence me and guide me to be a better scientist.

I thank my thesis committee, Drs. Daniel Rosenbaum, Michael Roth and Luke Rice, who were more than generous with their time, expertise and warm encouragement throughout the entire process.

Thank you to people who assisted with this project, especially Dr. Carrie Partch who guided me through the early stage of my graduate career and generously passed all her knowledge on this project to me; Drs. Thomas Scheuermann and Fernando Correa, who kindly helped me on developing my technical skills on ITC and NMR spectroscopy; Drs. Arati Ramesh, Dominika Borek, Jason Key and Diana Tomchick, who took off countless hours to help me with X-ray crystallography, and Dr. Shuguang Wei and Melissa Bennion for screening thousands of compounds. Also, I thank all members of the Gardner lab for always being there, giving me company and advice. You guys made it so much fun to work

STRUCTURAL CHARACTERIZATION AND CHEMICAL INHIBITION OF THE
ARNT/TACC3 COMPLEX

by

YIRUI GUO

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

December, 2014

STRUCTURAL CHARACTERIZATION AND CHEMICAL INHIBITION OF THE
ARNT/TACC3 COMPLEX

Publication No. _____

Yirui Guo (Ph. D.)

The University of Texas Southwestern Medical Center at Dallas, 2014

Supervising Professor: Kevin H. Gardner

My research project focuses on mechanistic studies of a new group of transcriptional coactivators (coiled-coil coactivators: TACC3, TRIP230, CoCoA), involved in cancer development and progression. Normally, these coactivators play an essential role in the HIF hypoxia response, directly interacting with the ARNT subunit of HIF in a novel and promoter-specific way. However, misregulation by overexpression or activating fusions (for example, FGFR-TACC3) is sufficient for transformation and associated with the development of glioblastoma, renal cell carcinoma and other cancers. In light of this

connection between coiled-coil coactivators (CCCs) and HIF signaling, tools that inhibit HIF/CCC complex formation might present opportunities to interrogate the linkage between different CCC-containing pathways and may offer a novel route to blocking cancer formation and progression.

As a member of this new group of transcription coactivators, knowing how TACC3 interacts with ARNT is critical in understanding the general role of CCCs in HIF-dependent transcription machinery. In the first half of my study, I characterized the ARNT/TACC3 complex with various biophysical and biochemical methods including solution NMR, X-ray crystallography, circular dichroism, luminescence proximity and numerous cell-based assays. A 3.15 Å ARNT/TACC3 crystal structure was solved by molecular replacement, revealing details of this protein complex and providing a structural function for coactivator recruitment in HIF signaling pathway.

The second half of this study focuses on the search for ARNT/TACC3 inhibitors with *in vitro* screens to regulate ARNT/CCCs activity in a rapid and flexible way. From a fragment-based NMR screen, I identified small molecules that specifically bound within the second PER-ARNT-SIM (PAS) domain of ARNT and perturb its interaction with TACC3. However, these small molecules have drawbacks, such as low potency or unclear modes of action. To identify higher potency small molecules targeting ARNT/TACC3 complexes, I developed an AlphaScreen-based high throughput screen. Hopefully the discovery of artificial ligands with known mode-of-action that inhibit this typically "undruggable" protein complex will provide new perspectives in small molecule inhibitor development, and also serve as a very useful tool in cell biology for studying pathways utilizing ARNT/TACC3.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Dor Salomon, Yirui Guo, Lisa N. Kinch, Nick V. Grishin, Kevin H. Gardner and Kim Orth (2013) Effectors of animal and plant pathogens use a common domain to bind host phosphoinositides. *Nature Communications* **4**, 2473

Yirui Guo, Carrie L. Partch, Jason Key, Paul B. Card, Victor Pashkov, Anjana Patel, Richard K. Bruik, Heiko Wurdak and Kevin Gardner (2013) Regulating the ARNT/TACC3 axis: Multiple approaches to manipulating protein/protein interactions with small molecules. *ACS Chemical Biology* **8**, 626-635

LIST OF FIGURES

CHAPTER 4

CHAPTER 5

LIST OF TABLES

LIST OF APPENDICES

CHAPTER 1
INTRODUCTION

MAMMALIAN bHLH-PAS TRANSCRIPTION FACTOR CONTROLLED GENE EXPRESSION

Transcription is the initial step of gene expression and is crucial for many cellular processes. Well-controlled transcription is important for normal functions of the cell while mis-regulated transcription is constantly associated with many diseases. Transcription can be triggered by many signals, such as environmental stimuli, and regulation of this triggering process is often controlled by transcription factors. My work centers on one class of these proteins, the basic helix-loop-helix PAS (bHLH-PAS) proteins, where in most cases the C-terminal PAS domain is responsible for sensing signal molecules and the N-terminal bHLH motif directly interacts with the promoter DNA (1). Studying the function and regulation of the bHLH-PAS transcription factors will not only reveal fundamental aspects of gene transcription, but could also provide promising therapeutic routes for anti-cancer treatment given the roles that certain bHLH-PAS proteins play in tumor formation and development (HIF, AhR) (2).

i. bHLH-PAS family proteins are powerful regulators in gene transcription

The basic helix-loop-helix PAS (bHLH-PAS) proteins are critical transcriptional factors involved in the transcription of many cellular processes, often launching robust

responses to several environmental stimuli such as aromatic hydrocarbons and oxygen tensions (3). These responses allow cells to adapt to environmental stress and metabolic state of the whole organism. bHLH-PAS transcription factors can be controlled at many levels, including post-translational modifications and recruitment of other cofactors (1,4,5). Misregulation of the bHLH-PAS protein function usually leads to a wide variety of diseases (2).

a. General domain arrangement of bHLH-PAS proteins

bHLH-PAS proteins belong to the superfamily of bHLH transcription factors, which are characterized by a bHLH motif followed by two tandem PAS domains (PAS-A and PAS-B) (1,2).

bHLH motif - The bHLH motif is commonly found in many eukaryotes from yeast to human. This motif has two distinct, well-conserved domains. The N-terminal region is the basic domain, which binds to a consensus hexa-nucleotide DNA sequence (CANNTG) called the E-box with high specificity (6). Each E-box sequence can be recognized by a corresponding bHLH protein complex. The C-terminal HLH domain can facilitate homo- or heterodimerization of bHLH motif-containing proteins. Different combinations of bHLH proteins and E-boxes control diverse cellular functions by transcription regulation (7).

Figure 1-1 shows examples of typical bHLH structures (8,9). MyoD is a characteristic member of bHLH proteins that controls myogenic transcription. MyoD forms

a homodimer when associated with the consensus DNA sequence CAGCTG, chiefly homodimerizing via the helix-loop-helix domain. This arrangement leaves the basic region of the MyoD dimer free to directly interact with DNA in the major groove (Fig. 1-1a) (8). Complementing the MyoD homodimer, bHLH proteins can also form heterodimers, as exemplified by MYC/MAX (Fig.1-1b) (9). MYC plays an important role in transcription activation, cellular transformation and apoptosis. De-regulation of MYC is associated with the development of several human cancers such as lymphoma, neuroblastomas and small cell lung cancers. Similar to MyoD, the MYC/MAX complex forms major contacts with the DNA major groove using the basic region and dimerizes through the C-terminal HLH domain. In addition to the bHLH domain, the MYC/MAX complex also has an extra leucine zipper domain at the C-terminus (9). This leucine zipper is not only an extension of the helical domain from bHLH that facilitates protein dimerization, but also plays an important role in the selection of dimerization partner. For instance, electrostatic interactions between MYC amino acid residues R423 and R424 and MAX residues Q91 and Q92 provide a tighter complex than the MAX/MAX homodimer. A similar strategy is used in the MAD/MAX heterodimer, where MAD E125 forms a hydrogen bond with MAX N92 (9).

Taken together, the bHLH proteins constitute a large superfamily of more than 13,800 members with diverse functions in cells such as neurogenesis, myogenesis, heart development and hematopoiesis (10,11). Despite the high sequence conservation at the dimerization interface (12,13), the bHLH proteins never-the-less demonstrate high selectivity in E-box binding and distinct transcription functions, which makes fusing a bHLH motif with other protein domains a good strategy to expand the utility of this motif.

3
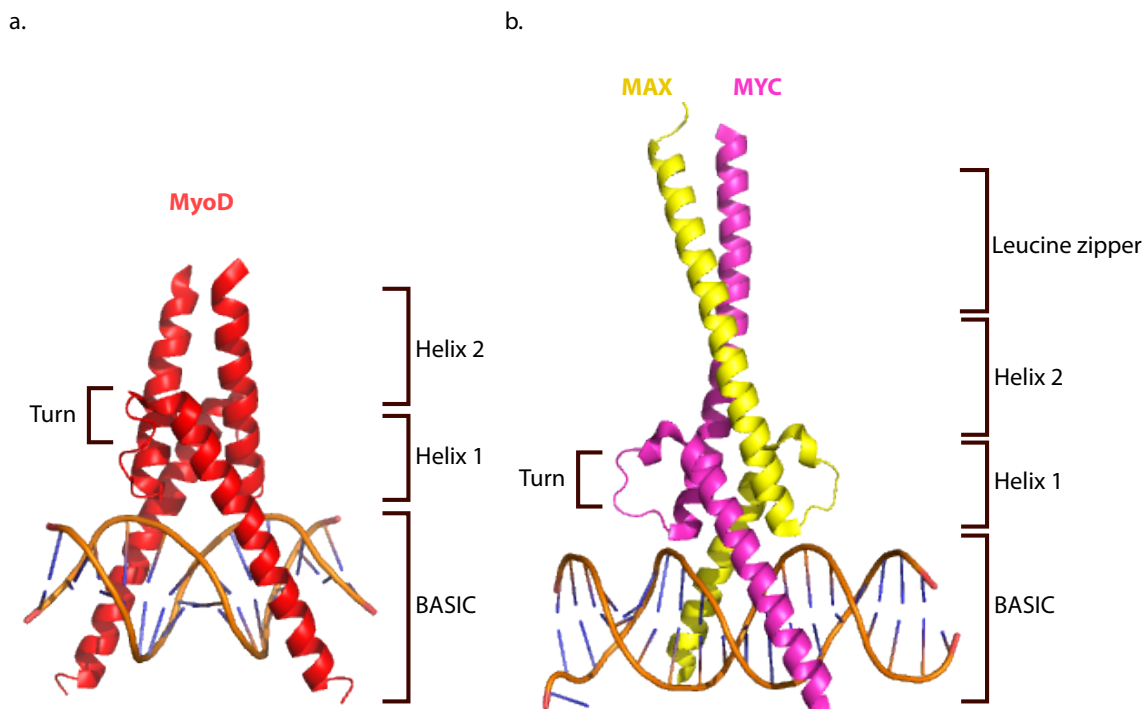
Figure 1-1. Structures of bHLH/Leucine zipper protein dimer bound to DNA. a) MyoD homodimer bound to E-box (PDB: 1MDY) (8). b) MAX/MYC heterodimer bound to E-box (PDB: 1NKP) (9).

PAS domain - PAS is short for PER-ARNT-SIM, which is named after three proteins in which it was first identified: the *Drosophila* period clock protein (PER) (14-16), the vertebrate aryl hydrocarbon receptor nuclear translocator (ARNT) (17) and the *Drosophila* single-minded protein (SIM) (18). Although it was initially discovered in animals, PAS domain mediated signaling pathways are widely spread among all three kingdoms of life (*Bacteria, Archaea and Eucarya*) (1).

PAS domains are found in many organisms and share low sequence conservation (~20-30%) (19). However, the tertiary structures of different PAS domains are extremely similar. A typical PAS domain contains ~100 residues and adopts a mixed α/β fold as

4

described in the first PAS structure, the photoactive yellow protein (PYP), which featured 6 anti-parallel β strands flanked by several α helices (Fig. 1-2a) (20). With the availability of additional PAS domain structures, the canonical PAS fold was found in general to consist of a central β-sheet with 5 anti-parallel β strands (Aβ, Bβ, Gβ, Hβ and Iβ) in B-A-I-H-G topological order surrounded on side by several α helices (Cα, Dα, Eα and Fα) (Fig. 1-2b) (3,21). The region from Aβ to Iβ is referred to the "PAS core", and may be accompanied by N- or C-terminal extensions (3). The central β-sheet is the most conserved region in different PAS domains with an average backbone root-mean-square deviation = 1.9 ± 0.6 Å (3). Tandem PAS domains from one protein are usually named alphabetically from N- to C-terminus such as PAS-A and PAS-B (3). Notably, PAS domains share remarkable similarity with GAF domains both in terms of tertiary structure and types of effector domains that linked to them, suggesting a common evolutionary origin between the two domains (22,23).

PAS domains are versatile folds that are involved in many cellular functions depending on their abilities to bind small molecule ligands such as heme and flavin (24). The major sensory function of PAS domains could be classified mainly as redox state/oxygen sensing and light sensing.
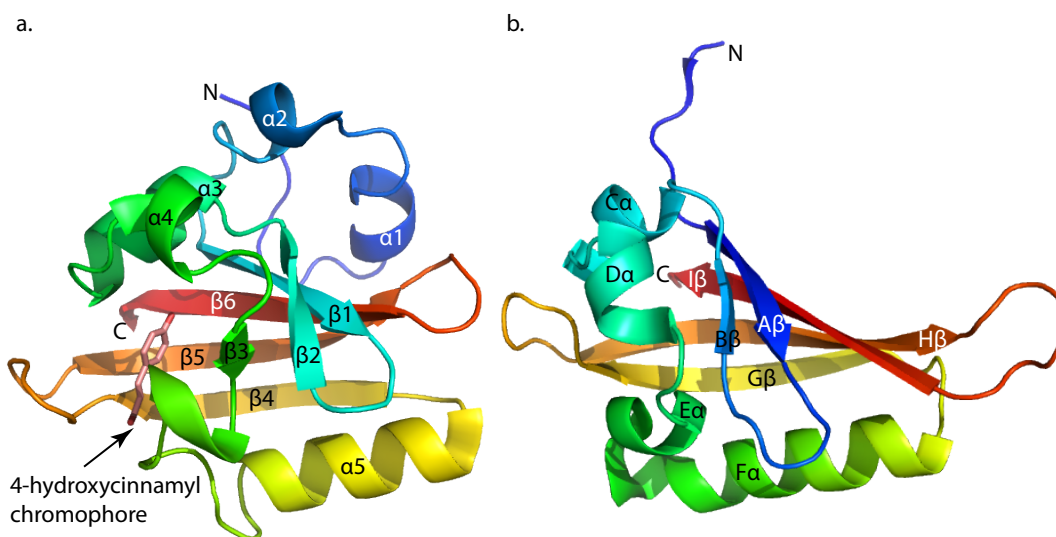
Figure 1-2. General PAS fold. a) Crystal structure of photoactive yellow protein (PYP) (PDB ID: 2PHY). Secondary elements and N-, C-terminus are marked in the structure. The bound chromophore is displayed in sticks and highlighted by arrow. b) Crystal structure of ARNT PAS-B (PDB ID: 4EQ1).

A good example of a redox state/oxygen sensing PAS domain is from the FixL sensor histidine kinase (HK), part of the FixL/FixJ two-component system that controls gene expression under micro-aerobic or anaerobic conditions in some bacteria (1). FixL directly senses oxygen via a heme b cofactor bound inside its PAS domain (25-27) (Fig. 1-3a). In the crystal structure of this domain, the heme iron was coordinated with a histidine residue on Fα. Heme propionate side chains were stabilized by contacting positively charged side chains of arginine and histidine in on the Fα and FG-loop, respectively. O_2 binding makes the heme b cofactor more planar, thus inducing conformational changes that lead to the inhibition of autophosphorylation and phoshoryl transfer between histidine kinase module and the cognate response regulator FixJ (27). A related heme b factor binding PAS domain is the PAS1 domain of the *Escherichia coli* direct oxygen sensor DosP (28) where the heme iron was

again liganded to the conserved Fα histidine (29), suggesting a common evolutionary origin between FixL and DosP. However, DosP PAS1 forms more extensive contacts with the heme propionate side chains through several backbone amide interactions, providing evidence for evolutionary divergence in heme binding PAS domains (30). Besides heme, PAS domains also utilize FAD cofactors to sense redox state of the cell by monitoring electron transport (31).

PAS domains are also visible light sensors when bound to cofactors such as flavins (Fig. 1-3b). One group of such photosensory PAS domains is specifically referred as LOV (light, oxygen, voltage) domains (24). When exposed to blue light, a new covalent bond is formed between a conserved cysteine residue on the Eα helix and the flavin C4a position (24,32,33). This light-dependent crosslinking induces conformational changes that perturb the interaction between the LOV core and flanking regions such as the Jα helix in AsLOV2 and an HTH domain in EL222 (34-36), resulting in the rearrangement of domains in full length proteins and releasing the previously locked effector domain for function. Similar mechanisms are also seen in other LOV containing proteins (33). This light-driven conformational change is the key step in the LOV domain signaling mechanism, which has been utilized in engineered photoswitches. For example, multiple light inducible promoter systems have been engineered by fusing activation domains or DNA binding motifs with LOV domains to control gene expression *in vivo* (37,38).
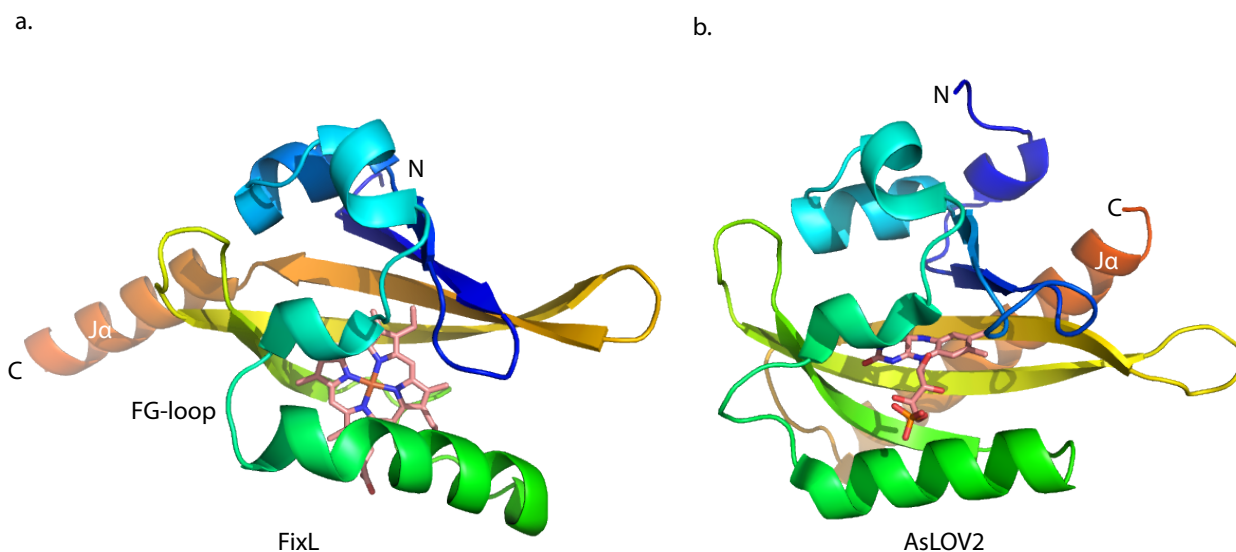
Figure 1-3. PAS domain binds to different cofactors. a) Structure of heme b bound FixL (PDB ID: 1DRM). b) Structure of FMN bound AsLOV2 (PDB ID: 2V1B)

Besides oxygen and light sensing, PAS sensor kinases can regulate di- and tricarboxylate metabolism through direct interactions with C4 and C6 carboxylate-containing substrates, leading to the enhancement of sensor kinase activity and regulation of genes required for C4 and C6 transport (39-41). PAS sensor domain in the PhoQ histidine kinase also utilizes negatively charged surface residues on the β-sheet to sense divalent cations such as Ca^{2+}, Ni^{2+} and Mg^{2+} in order to suppress PhoQ kinase activity (42-44).

PAS domains are important sensors of environmental changes that are critical for the survival of an organism. To achieve many diverse functions with a limited number of structural scaffolds, nature has utilized the strategy of combining chemically diverse cofactors and a wide variety of effector domains with the PAS sensory domain. Although it is unclear if all PAS domains have a common ancestor and essentially the same signaling mechanism, examples have shown that in the context of a chimeric protein, an effector

domain can still function after being fused with an input PAS domain that senses different cofactors (3,45). This finding makes PAS domains extremely useful in synthetic biology, potentially allowing the construction of genetically encoded protein switches responding to multiple signals (3,33).

bHLH-PAS protein - Similar to bacteria, eukaryotes also use PAS domains as important environmental sensors. Although PAS domains serve as input domain of some eukaryotic C-terminal sensor kinases (46,47), PAS domains in eukaryotes are often associated with bHLH DNA binding motifs and serve as transcription factors regulating gene expression in cells or whole organisms in response to environmental changes. This particular arrangement of PAS domain and bHLH motifs belongs to the bHLH superfamily and is referred as the bHLH-PAS transcription factor family (7). As mentioned previously, the typical bHLH-PAS transcription factor contains an N-terminal bHLH motif that is responsible for E-box binding and two C-terminal PAS repeats, namely PAS-A and PAS-B. bHLH-PAS transcription factors can be divided in two major classes based on their dimerization partners: class I bHLH-PAS proteins do not homodimerise or heterodimerize with other class I factors, whereas class II bHLH-PAS transcription factors normally require dimerization with a class I bHLH-PAS protein for function, often using the two tandem PAS domains (Fig. 1-4) (2,7). It has been observed that in dioxin signaling, deleting the PAS-A domains in Aryl hydrocarbon receptor (AhR) severely reduced its interaction with ARNT, reducing gene transcription initiated by AhR/ARNT (19). Similar cases were found in the hypoxia response pathway involving HIFα and ARNT (48,49) and in circadian clock

signaling between NPAS2 and BMAL1 (50,51). Moreover, in addition to mediating protein/protein interactions, PAS-B domains of the bHLH-PAS transcription factor often bind small molecule ligands. In xenobiotic signaling, AhR is stabilized in the cytoplasm by chaperones such as Hsp90, p23 and Ara9 through direct interactions with the PAS-B domain. Ligand binding to AhR PAS-B domain activates the AhR signaling pathway by disrupting the AhR/chaperone interaction, allowing AhR to translocate into the nucleus and partner with ARNT (52-54). Cavities are also observed in the PAS-B domains of other bHLH-PAS proteins such as HIF1α, HIF2α and ARNT (21,55,56). Although natural ligands for those PAS-B domains have not been identified yet, several studies have developed artificial small molecules to bind inside the cavity and disrupt PAS-B mediated protein/protein interaction (57,58). In addition to the bHLH and PAS domains, most class I bHLH-PAS transcription factors also contain a few transcriptional activation or repression domains (TAD or TRD) following the two PAS domains that are responsible for the recruitment of transcription coactivators or corepressors (7,59-61). Therefore, we could conclude that a functioning bHLH-PAS transcription factor complex requires three major contacts: the bHLH motif/DNA, PAS-AB dimerization and C-terminal TAD/transcription coactivators.
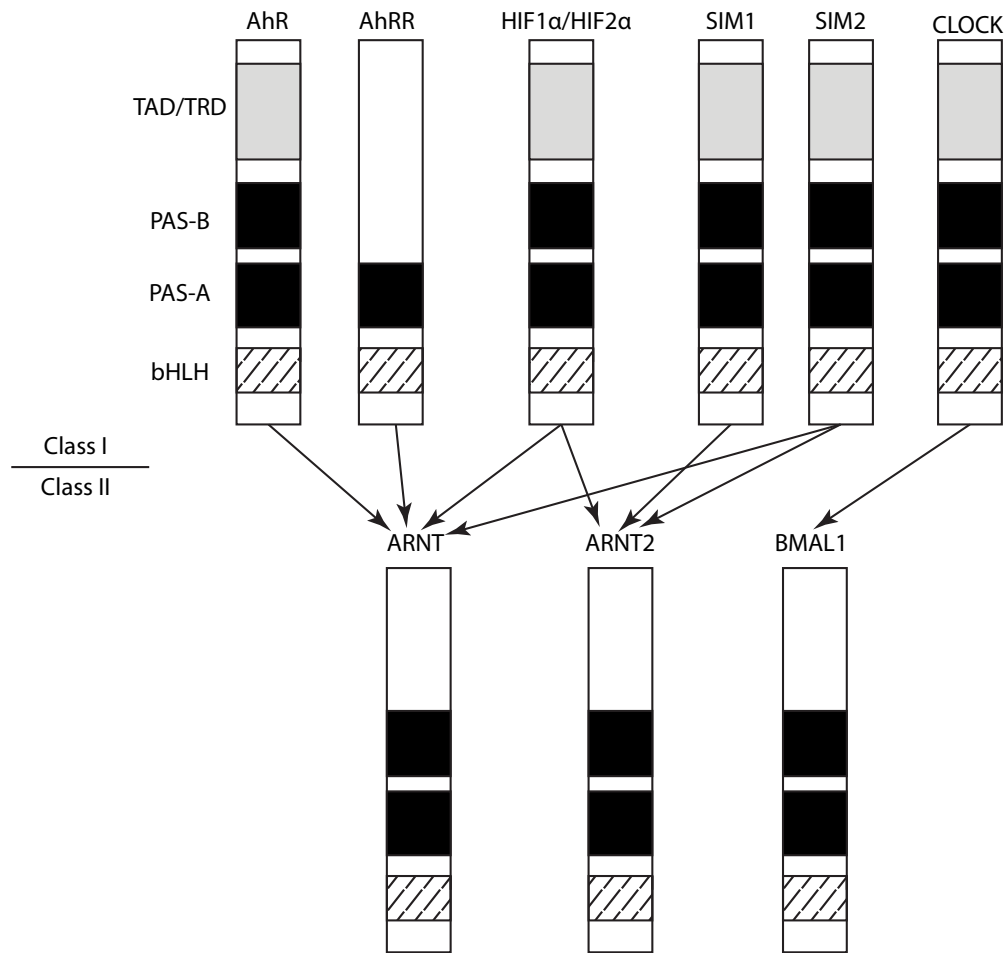
Figure 1-4. bHLH-PAS factor classification. (Adapted from Bersten D, *et al. Nature Reviews.* 2013). bHLH-PAS proteins are classified into two classes: class I (upper panel) and class II (lower panel). Proteins from class I do not dimerize with other factors from class I but only class II proteins for function. Class II proteins can either heterodimerize with class I protein or homodimerize with itself. Most times class II factors are common partners for multiple class I proteins. Heterodimers between bHLH-PAS partners are matched with arrows.

b. bHLH-PAS protein signaling

bHLH-PAS proteins are potent transcription factors that control gene expression in response to a wide variety of physiological signals. Misregulation of bHLH-PAS factors is associated with several forms of cancer (2). As such, studies of bHLH-PAS pathways are important to both understand the mechanism of bHLH-PAS controlled gene transcription and provide possible therapeutic routes for cancer treatment. Next, I will introduce three well-studied bHLH-PAS protein pathways associated with cell sensing and adaption to different stimuli (aromatic hydrocarbons, oxygen level and circadian rhythms) as examples to further understand the mechanisms of bHLH-PAS signaling pathways.

The aryl hydrocarbon receptor (AhR) - The AhR bHLH-PAS transcription factor controls responses to xenobiotic compounds. As one of the earliest discovered bHLH-PAS proteins, the signaling mechanism of AhR has been characterized by many approaches (Fig. 1-5). In its latent state, ligand-free AhR is stabilized in the cytoplasm by several chaperone proteins Hsp90, p23 and hepatitis B virus X-associated protein (XAP2/AIP/Ara9) (7). When exposed to either natural or synthetic xenobiotics, such as the combustion-derived polycyclic aromatic hydrocarbon benzo [a] pyrene (B[a]P) and 2,3,7,8-tetrachlorodibenzodioxin (TCDD) (62), the AhR/chaperone interaction is weakened by ligand-induced conformational changes of the PAS-B domain. Dissociated AhR protein then translocates to the nucleus and partners with a class II bHLH-PAS transcription factor ARNT. This AhR/ARNT heterodimer binds to the XRE (xenobiotic responsive element) site in the promoters of AhR-regulated

12

genes, initiating a classic detoxification pathway by enhancing the expression of multiple enzymes involved in xenobiotic metabolism such as cytochrome P4501A1, P4501A2, P4501B1 (for phase I xenobiotic metabolism), along with conjugating enzyme glutathione-*S*-transferase and glucuronyl transferase (for phase II xenobiotic metabolism) (2).
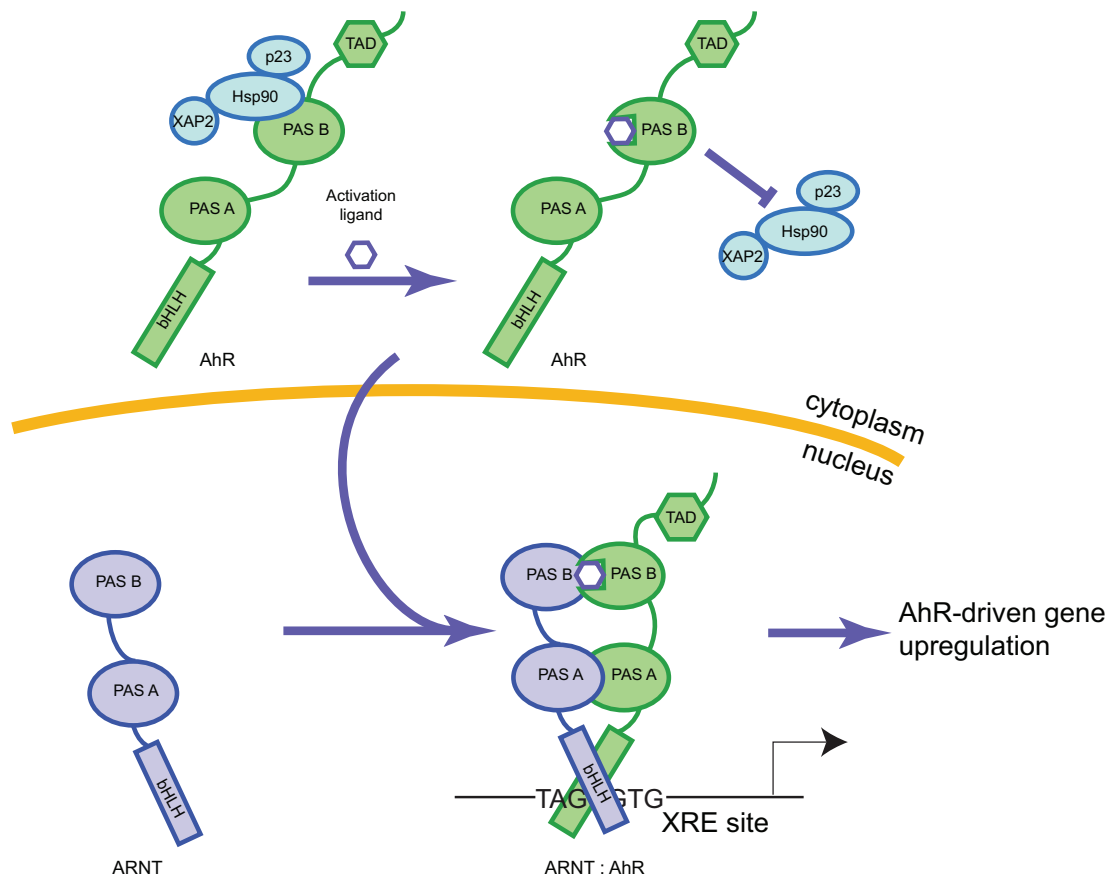
Figure 1-5. AhR signaling pathway. In its latent state, ligand-free AhR is stabilized in cytoplasm by several chaperone proteins. When exposed to either natural or synthetic xenobiotics, the AhR/chaperone interaction is weakened by ligand-induced conformational changes of the PAS-B domain. Dissociated AhR protein then translocates to the nucleus and partners with a class II bHLH-PAS transcription factor ARNT. This AhR/ARNT heterodimer binds to the XRE (xenobiotic responsive element) site in the promoters of AhR-regulate genes.

This signaling pathway is negatively controlled by three avenues: 1) excretion of activating ligands; 2) expression of AhRR, an AhR repressor that lacks essential transcription activation domains and can compete with AhR for ARNT dimerization (63); 3) proteasome-dependent degradation of AhR (2).

In addition to its role in xenobiotic responses, AhR is critical in many physiological processes. These include fetal liver angiogenesis, female reproductive capacity, haematopoietic stem cell maintenance and immunomodulation via B/T cell development (64). However, studies searching for the full complement of AhR target genes revealed little consensus throughout different cell lines and tissues, suggesting additional complexity in AhR control of multiple physiological roles (2). Turning to the role of AhR in cancer, this system is linked to tumorigenesis via two different mechanisms. The first one is associated with the carcinogenic side products produced by downstream enzymes whose expression levels are controlled by xenobiotic response. For example, it has been observed that cytochrome P4501A1, which is upregulated in B[a]P signaling, can generate B[a]P diol epoxides (65). These molecules are strong DNA-damaging reagents that could be one of the major causes of high mutation rates in lung tumor cells (66,67). The second way of AhR-stimulated tumorigenesis is by stimulating cell proliferation (2). Unlike the first method, the exact mechanism of this second tumorigenesis pathway is not fully defined. However, multiple lines of evidence were found to support this hypothesis. It has been reported that in gliomas the highly activated tryptophan 2,3-dioxygenase (TDO) can increase the metabolic rate of turning tryptophan to the AhR agonist kynurenine, leading to autocrine and paracrine mechanisms in tumor formation (68). Analysis of gene expression across 947 human cancer

14

cell lines also demonstrated that high expression levels of AhR are correlated with proliferation in many NRAS-mutant cell lines (2,69). Interestingly, besides its role in tumorigenesis, AhR is also believed to have anti-cancer functions. Several mouse models have shown that *Ahr$^{-/-}$* mice develop β-catenin-mediated colon tumors due to the lack of AhR's inherent ubiquitin ligase activity and its role to promote anti-inflammation (70). As such, AhR activation is like a double-edged sword that can both promote or suppress tumorigenesis under different circumstances. Accordingly, the development of highly selective AhR modulators has become a promising route for anti-cancer treatment.

The hypoxia inducible factor (HIF) - HIF is a heterodimeric bHLH-PAS transcription factor complex that controls the genomic response to hypoxia. This complex comprises an oxygen regulated α subunit composed of one of three orthologs (HIF1α, 2α and 3α) and a common β subunit ARNT, which is also a binding partner for AhR (71,72). The α subunit is sensitive to oxygen through an oxygen degradation domain (ODD), located at the C-terminus of the two PAS domains. Under normoxia (~20% O_2), two conserved proline residues in the cytoplasmic HIFα ODD domain (P402, P564 in HIF1α) are selectively hydroxylated by three O_2-dependent prolyl hydroxylase enzymes (PHD1-3), generating a binding site for the Von Hippel-Lindau tumor repressor (VHL) and associated an E3 ubiquitin-protein ligase complex that targets cytoplasmic HIFα for proteasome-dependent degradation (73-76). In parallel, HIFα that is already localized in the nucleus is hydroxylated at a C-terminal transactivation domain (C-TAD) asparagine site (N803 in HIF1α) by a specific asparaginyl hydroxylase (FIH-1), preventing the HIFα TAD from interacting with transcriptional coactivators (77,78)

(Fig. 1-6). During hypoxia (less than 5% O_2), these hydroxylases are deactivated, allowing the HIFα subunit to be stabilized and translocated to the nucleus. HIFα then heterodimerizes with ARNT via bHLH and PAS domain interactions and binds to the hypoxia response element (HRE) site at the promoter region (7). Recruitment of transcription coactivators such as CBP/p300 to the HIF complex enhances transcription of hundreds of genes that allow cells to adapt to low oxygen levels (7) (Fig. 1-6). Hypoxia mimics such as Co^{2+}, Ni^{2+} and iron chelators can also induce similar response by interfering with hydroxylase function (79,80).

Besides O_2 levels, variation in metabolite production can also alter HIF activity. For example, an increase in succinate levels due to mutations in succinate dehydrogenase can inhibit PHDs, thus increasing HIF levels and leading to the development of specific types of cancer (81-83). Conversely, mutations in isocitrate dehydrogenase 1 (IDH1) and IDH2 can increase (R)-2-hydroxyglutarate levels, which activates PHDs and thus reduces HIF level (84-87).

Regardless of mechanism, deregulation of HIF assists in tumor formation and progression at the transcription level. Many HIF-driven genes are associated with cell proliferation, glycolytic metabolic response, angiogenesis and metastasis (88). High levels of HIF are often found in solid tumors and can assist tumor progression (89). As expected from this correlation, *Hif1α*$^{-/-}$ and *Hif2α*$^{-/-}$ knockouts showed decreased tumorigenesis in mammary or liver/colon models, respectively (90,91). However, HIF may also play anti-tumorigenic roles under certain conditions. For example, researchers have found that *Hif2α*$^{-/-}$ in a KRAS-initiated lung cancer model actually promotes tumor formation (92). Those observations

suggest that HIF function is highly dependent on the context of cellular condition and the HIF-cancer correlation is very complicated.
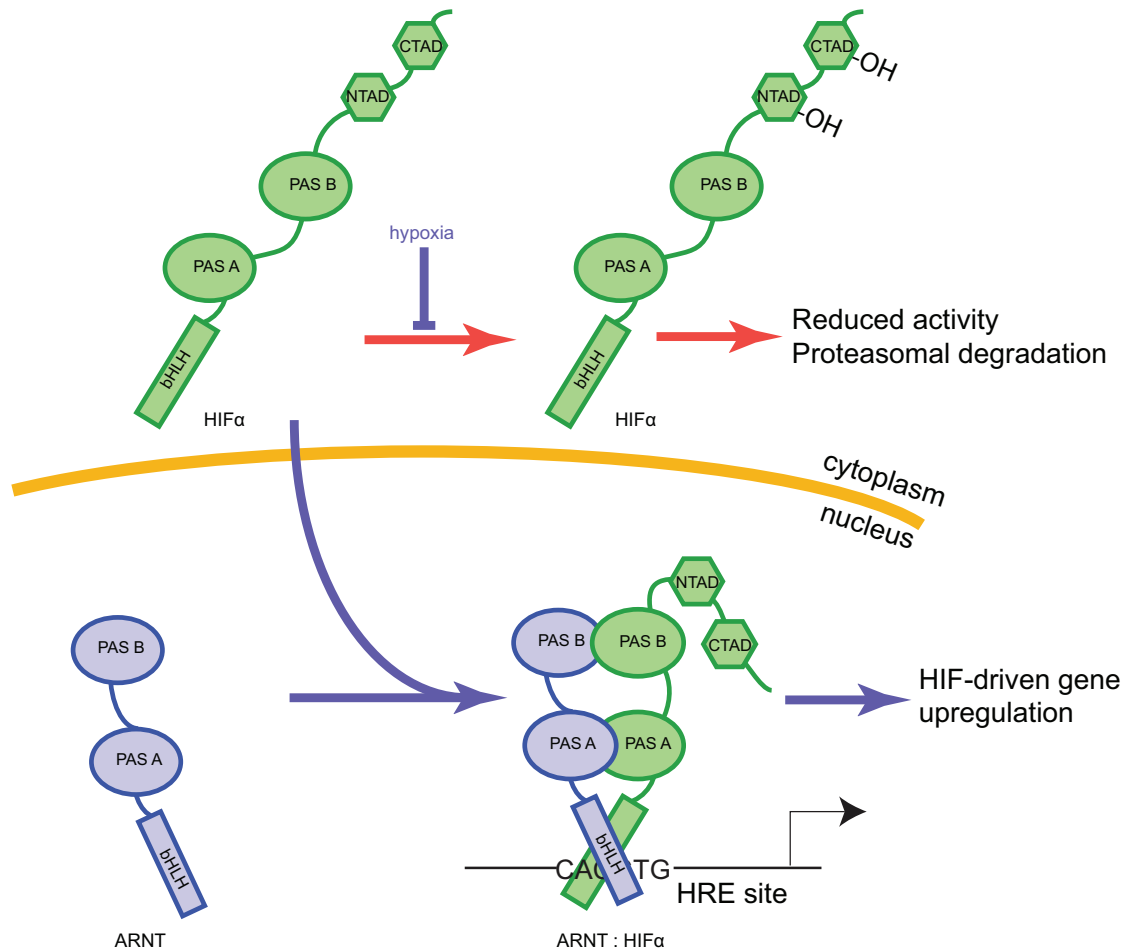
Figure 1-6. HIF signaling pathway. Under normoxia, O_2-dependent hydroxylation of HIFα decreases its abundance and activity. Hypoxia stops these modifications, allowing HIFα to accumulate in the nucleus and dimerize with ARNT. This heterodimer binds to hypoxia responsive enhancer (HRE) sites, controlling target gene transcription.

Despite the apparently contradictory roles of HIF in tumorigenesis, the tight association with cancer still makes it an important therapeutic target. Several small molecule ligand screens have been launched to look for HIF inhibitors (93,94). Although the HIF full length structure is unavailable, the structures of isolated domains still provide valuable

guidance in ligand design and mechanistic studies (56,95). Because HIF is a non-enzymatic protein that lacks a canonical "active site", researchers developed strategies to target protein/protein interaction between HIF and its partner proteins. The first promising application of this approach was during the development of HIF/p300 inhibitor. After the structure of a minimal interaction HIF1α/p300 complex (HIF1α C-TAD and p300 CH1 domain) was solved by NMR (95), Kung *et al* utilized a fluorescent-based high throughput assay to identify compounds that disrupted this minimal complex *in vitro* and discovered a small molecule, chetomin, that bound to the CH1 domain and disrupted HIF1α and HIF2α interactions with p300 (95,96). Another successful application of this strategy is in HIF2α inhibitor development. A potent artificial ligand with known mode-of-action was identified by a combination of *in vitro* protein/protein screen and chemical modification (57,58). The small molecule discovered in this study specifically interacted with HIF2α PAS-B and was able to disrupt full length HIF2α/ARNT interaction *in vitro* and downregulate HIF2α downstream genes in cell based assays. The high-resolution crystal structure of HIF2α PAS-B revealed a 290 Å^3 water-binding cavity inside the core (56). Co-crystallization of HIF2α PAS-B and this compound further confirmed the cavity as the compound-binding site, suggesting potential dynamics of the PAS-B domain consistent with NMR data but which could not be observed in the crystal structure (58). This HIF2α PAS-B structure exhibited another example of the ligand-controlled PAS domain, and provided the structural basis for HIF inhibitor design. A similar PAS-B cavity was also observed in ARNT PAS-B domain but with a much smaller size (~100 Å^3), providing a complementary route to target all three

18

HIF orthologs and ARNT-binding coactivators. This ARNT PAS-B cavity will be discussed in detail in Chapter 3 (21).

The circadian rhythm regulators - Unlike the AhR and HIF pathways, circadian oscillation is mediated by a negative feedback loop that involves several circadian clock modulators including the PAS-containing period circardian protein (PER); the bHLH-PAS proteins circadian locomotor output cycles kaput (CLOCK), brain muscle ARNT-like 1 (BMAL1), and the blue light sensing protein cryptochrome (CRY). During the daytime, the CLOCK/BMAL1 bHLH-PAS heterodimer upregulates several circadian clock proteins, resulting a cytoplasmic accumulation of the PER/CRY (97). At night, PER/CRY complexes translocate to the nucleus, binding to CLOCK/BMAL1 and suppressing their own transcription (98). The PER/CRY complex is then phosphorylated by casein kinase 1ε (CK1ε) or CK1δ and removed by F-box binding protein Fbxl3 assisted proteosomal-dependent degradation to restore the cycle (Fig. 1-7) (99). One remarkable feature about the clock is its highly specific response to light but not other environmental perturbations, which is achieved by a powerful genetic and intercellular regulatory network (100,101). For example, the transcription of BMAL1 at ROR-binding elements (ROREs) can both be suppressed or enhanced by nuclear hormone receptor superfamily proteins serving as repressor (Nr1d1 (nuclear receptor subfamily 1, group D, member 1)) (102) or activator (Nr1f1 (nuclear receptor subfamily 1, group F, member 1)) (103,104). Other metabolites and proteins are also involved in this regulatory network to help keeping the accurate response of clock by directly interacting with the CLOCK/BMAL1 complex (19). NAD$^+$ is one key

19

metabolite that could feedback to the circadian pathway by either directly interacting with the CLOCK/BMAL1 complex to attenuate its interaction with DNA, or indirectly through NAD^+-dependent enzymes such as NAD^+-dependent deacetylase sirtuin 1 (SIRT1) (105). SIRT1 stimulates clock function in two parallel mechanisms: 1) activating BMAL1 gene expression via an indirect pathway involving the retinoic acid-related orphan receptors (RORs); 2) deacetylating BMAL1 and PER2, which weakens the CLOCK/BMAL1 interaction with the repressor PER/CRY and leads to the degradation of PER2 (106).
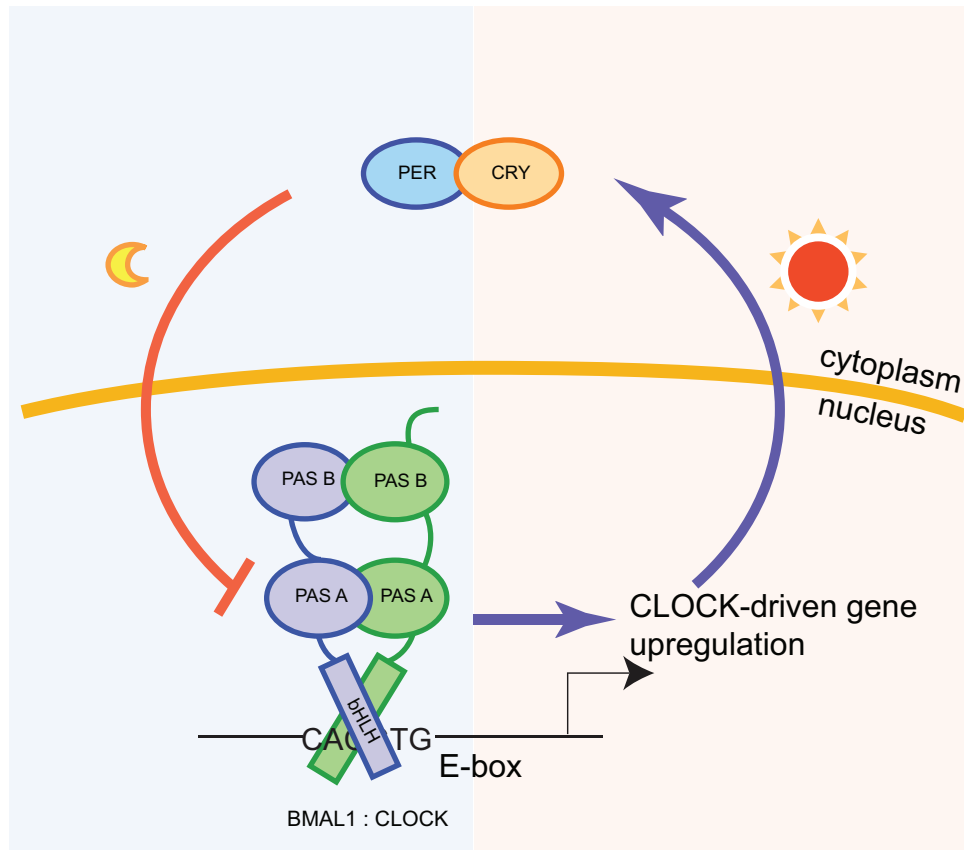
Figure 1-7. Circadian clock signaling pathway. In daytime, the CLOCK/BMAL1 bHLH-PAS heterodimer upregulates several circadian clock proteins, resulting in cytoplasmic accumulation of the PER/CRY. At night, PER/CRY complexes translocate to the nucleus, binding to CLOCK/BMAL1 and suppressing their own transcription.

Like other bHLH-PAS transcription factors, the misregulation of circadian rhythm PAS factors is directly associated with cancer. Researchers have found that PER1 and PER2 can serve as tumor repressors in mice (107-109). Mice with PER2 deficiency suffer deregulation of several cell cycle regulators and tumor suppressors (e.g. cyclin D1, cyclin A, MDM2, GADD45α) and upregulation of oncogenes (MYC), making those mice more prone to several forms of tumors (108). Mice with defects/alterations in other circadian rhythm genes (*bmal1, cry1, cry2*) are also more susceptible to spontaneous or radiation-induced tumorigenesis (110). Due to the importance of the circadian rhythm pathway, several small molecules that can alter circadian function have been identified via high throughput screening using circadian rhythm sensitive reporters (111-115). Undoubtedly, more advanced drugs will be designed with the aid of structural characterization of the CLOCK/BMAL1 heterodimer and other circadian rhythm proteins (116).

ii. The unique transcription coactivators associated with bHLH-PAS signaling - the coiled-coil coactivator (CCC) family

Transcription coactivators are transcription factor associated proteins that recruit transcription machineries to boost gene expression. In bHLH-PAS signaling, each bHLH-PAS transcription factor complex can utilize a combination of multiple coactivators to tune the downstream genes' expression levels. The most commonly involved transcription coactivators for bHLH-PAS transcription factors include, but are not limited to CBP/p300

and members of the nuclear receptor coactivator family such as SRC1 or NCoA, which can lead to histone acetylation and Pol II recruitment (5,117,118). These transcription coactivators usually directly bind to bHLH-PAS transcription factors at the C-terminal transactivation domain (TAD). However, a unique family of transcription coactivator called the coiled-coil coactivators (CCCs) started to attract researchers attention in the past decade due to their distinctive way of interacting with the bHLH-PAS transcription factor ARNT. There are currently three proteins in the CCC family: transforming acidic coiled-coil 3 (TACC3) (119), coiled-coil coactivator (CoCoA) (120) and thyroid hormone receptor interacting protein 230 (TRIP230) (121). This classification is not established on overall sequence similarity, but is based on the ARNT-binding coiled-coil domain found in all three proteins and their ability to affect ARNT-involved gene expression (4,122).

TACC3 - TACC3 belongs to the TACC family that contains a highly conserved predicted coiled-coil region at the C-terminus, namely the TACC domain. TACC homologues have been found in a few species: *S. pombe* (Alp7), *Caenorhabditis elegans* (TAC-1), *Drosophila* (D-TACC), *Xenopus laevis* (maskin) and mammals (TACC1, TACC2 and TACC3) (119,123-129) (Fig. 1-8). The three mammalian orthologs of TACC proteins could suggest functional redundancy in higher organisms. Interestingly, each human *TACC* gene sits close to a fibroblast growth factor receptor (FGFR) gene inside the genome, suggesting that those TACC orthologs share a common ancestor (129,130). Other splicing variants have also been isolated for mouse TACC3 (mERIC-1, mTACC3b) (126,129) and human TACC2 (Azu1, ECTACC) (123,131).

22

The first TACC protein (TACC1) was discovered in a genomic search looking for regions that are amplified in breast cancer, and was named based on its acidic nature and capability to stimulate cellular transformation (128). The other two human TACC proteins, TACC2 and TACC3, were mapped to chromosomes 10q26 and 4p16, respectively, implicated with breast cancer and multiple myeloma (130).
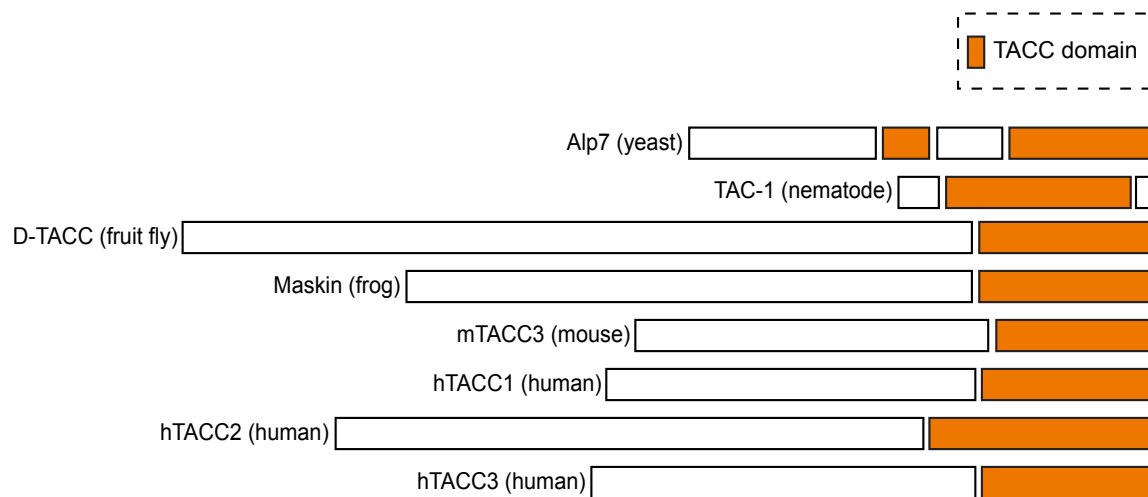
Figure 1-8. The TACC family. TACC proteins are found in a few eukaryotes. The key feature of TACC family proteins is the conserved TACC domain at the C-terminal (orange), despite their diverse N-terminal regions. TACC domains contain roughly 200 amino acids. Sequence analysis suggests that TACC domains have high contents of coiled-coil secondary structure.

This close connection between mis-regulated TACC proteins and cancer is largely due to the diverse functions observed in TACC family members. One key function of TACC proteins is their ability to regulate microtubule assembly by centrosome and mitotic spindle interactions during mitosis (132). This function was first discovered in the characterization of the *Drosophila* D-TACC protein. It has been shown that the centrosomal localized D-TACC protein can interact with microtubules and is essential for mitosis in *Drosophila* embryo

(124). Similar functions are also observed in other TACC family members including TACC3, demonstrating microtubule binding as a common function for TACC domain throughout the species (133).

Besides this characteristic microtubule binding function for all TACC domain-containing proteins, TACC3 also has a unique function in gene regulation via direct interaction with transcription factors (119). TACC3 was first discovered as an ARNT binding partner by yeast-two-hybrid assay, providing another example of a bHLH-PAS protein interacting with non-PAS members. Because ARNT is a common partner for many class I bHLH-PAS transcription factors such as AhR and HIF, interference with ARNT can affect multiple bHLH-PAS signaling pathways (xenobiotic and hypoxic signaling pathways, respectively) (2). Therefore, as an ARNT-binding protein, TACC3 has become a potential transcription cofactor for ARNT-involved gene expression. This cofactor role was quickly confirmed by monitoring HIF target gene expression under TACC3 protein overexpression or depletion conditions (4,119). Interestingly, overexpression of TACC3 gave positive feedback on HRE-regulated transcripts while TACC3 knockdowns with siRNA always resulted reduced HIF downstream gene transcription. The high correlation between target gene expression levels and TACC3 protein levels suggested that TACC3 acts as a transcription enhancer rather than suppressor (134). In combination with the coiled-coil nature of the TACC domain, TACC3 is classified as a characteristic representative of the coiled-coil coactivator (CCC) family.

Further studies have been conducted to reveal the ARNT/TACC3 interaction at atomic level. With the help of previously identified minimum binding domains (ARNT PAS-

B domain and C-terminus of TACC domain) (134), we have solved the crystal structure of the ARNT/TACC3 complex in their minimum interacting domains, demonstrating that the TACC3 coiled-coil domain forms direct contacts with the ARNT PAS-B β-sheet (4). The crystal structure will be discussed in more detail in Chapter 2. Given the importance of the cavity in ARNT PAS-B as a potential drug-binding site (21), searching for ARNT PAS-B ligands targeting ARNT/TACC3 will be as equally important as finding inhibitors for the ARNT/HIF interaction, since both routes provide complementary ways to target HIF orthologs. Ligand searches and characterization will be further discussed in Chapters 3 and 4.
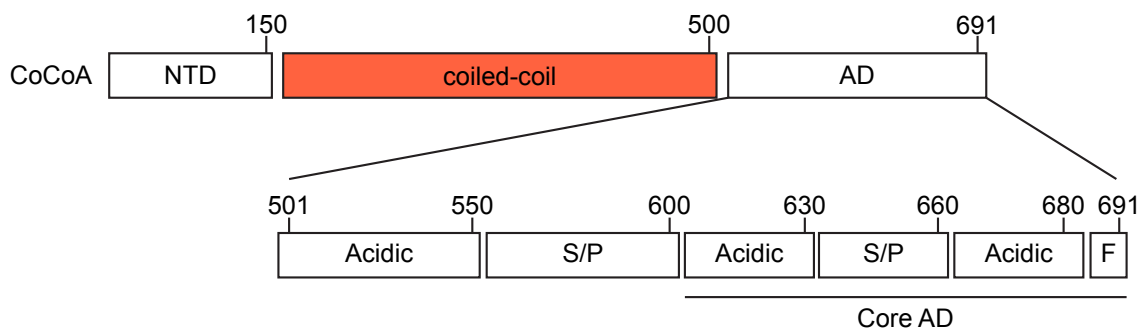
Figure 1-9. Schematic of CoCoA domain arrangements. (Adapted from Kim J, *et al. Nucleic Acids Research*. 2006). CoCoA contains three major domains: a N-terminal domain (NTD), a coiled-coil domain in middle of the protein and a transcriptional activation domain at the C-terminus (AD). A fine mapping of the AD domain exhibits three acidic subdomains (Acidic), two serine/proline-rich subdomains (S/P) and a phenylalanine-rich domain (F) at the C-term.

CoCoA - After TACC3 was discovered as the first member of the CCC family, more proteins behaving in a similar manner have been identified and isolated. The CoCoA (coiled-coil coactivator) is the second member of the CCC family. Similar to TACC3, CoCoA can be

dissected into three major domains and the center portion of the protein has high coiled-coil content based on secondary structure prediction (Fig. 1-9) (135).

CoCoA was first isolated by using the GRIP1 (glutamate receptor interacting protein 1) bHLH-PAS domain as a bait in a yeast two-hybrid screen and was later identified as a novel nuclear receptor coactivator that functions via direct interaction with the N-terminal bHLH-PAS domain of p160, GRIP1, CARM1 (coactivator-associated arginine methyltransferase 1) and p300 with its coiled-coil domain, serving as a secondary coactivator to enhance the transcription by estrogen receptor and other nuclear receptors (NR) mediated transcription (135-137).

The discovery of CoCoA as a secondary transcription coactivator and its ability to interact with the bHLH-PAS domain of several NR coactivators raises an interesting question: based on the high structural similarity between the bHLH-PAS proteins, is CoCoA also capable of interacting directly with other bHLH-PAS transcription factors as a transcription coactivator? To address this question, Kim and Stallcup pursued biochemical characterizations of CoCoA interaction with AhR and ARNT in AhR-mediated signaling (120). Pulldown assays on full length protein and truncated domains proved that CoCoA is physically associated with the activation domain of AhR and the bHLH-PAS region of ARNT. Using XRE-luciferase assays to monitor AhR/ARNT-driven transcript levels revealed a strongly positive correlation with CoCoA protein expression level, suggesting that CoCoA is working as a transcription coactivator for the AhR/ARNT signaling pathway (120). This discovery drastically broadens the potential function of CoCoA: besides being the secondary coactivator for NR-mediated transcription, CoCoA can

also be a primary coactivator for AhR/ARNT, or theoretically all other bHLH-PAS transcription factors that dimerize with ARNT (122). Further biochemical studies demonstrated that CoCoA interacts with ARNT at the PAS-B domain, similarly to the ARNT/TACC3 interaction, adding another general feature of the CCC protein family (122).

TRIP230 - TRIP230 as a transcription coactivator was first reported by yeast two-hybrid assay using the N-terminal truncated form of the retinoblastoma protein (Rb) as a probe and was later characterized as a thyroid hormone receptor (TR) coactivator based on the linkage of a genetic alteration of the TRIP230 gene loci (14q31) with abnormal thyroid hormone responses (138). Therefore, TRIP230 was named after its function in TR-mediated transcription and its 230 kDa molecular weight. TRIP230 can also be referred as GMAP-210 (Golgi-Microtubule Associated Protein 210 kDa), a earlier discovered *cis*-Golgi network-associated protein (identified by expression library screen with a human auto-antiserum), and TRIP11, a protein that binds to the ligand-binding domain of TR in a T3-dependent way (identified by yeast two-hybrid assay using TR as a bait) (139,140). TRIP230 was later identified as an ARNT coactivator in a yeast two-hybrid screen, making it the third member of the CCC family (121).

The extended coiled-coil domain predicted at the TRIP230 N-terminus is the signature feature of CCC-mediated transactivation (Fig. 1-10). The GRAB domain, GA1 and proline-rich regions at the C-terminus can also assist with protein/protein or protein/membrane interactions under different circumstances (Fig. 1-10).
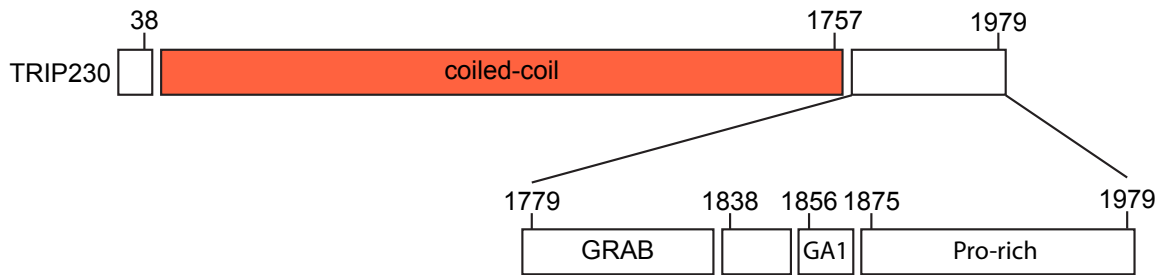
Figure 1-10. Schematic show of TRIP230 domain arrangement. (Adapted from *Barr FA & Egerer J. J. Cell Biol.* 2005 and *Gillingham AK, et al. JCB.* 2004) (141,142). TRIP230 is predicted to contain a major coiled-coil domain, followed by a GRAB (GRIP-related Arf-binding) domain, a GRAB-associated region (GR1) and a proline-rich hydrophobic region (Pro-rich) at the very C-terminus. The GRAB domain and GA1 region are required for Golgi interaction under assistance with other factors.

TRIP230 is a protein with diverse functions including Golgi-association and gene transcription. The Golgi association of TRIP230 was initially found in EM studies using antisera staining on samples from Sjoegren's disease patients, where TRIP230 is localized peripherally to the *cis*-Golgi (140). Additional evidence further supported TRIP230 being a potential Golgi matrix protein (141,142). The TRIP230/Golgi interaction involves multiple levels of contacts. The major Golgi targeting motif was found at the N-terminal region of TRIP230 (residue 1-375) (143). This was later refined to residues 1-38 as it was predicted as an ALPS (ArfGAP1 lipid packing sensor)-like motif that senses membrane curvatures in a bioinformatics search. TRIP230 can also associate with the Golgi indirectly via interaction with other Golgi-bound proteins (143,144). It was observed that the yeast homolog of TRIP230, Rud3p, could be recruited to the *cis*-Golgi by binding to the Golgi-bound GTPase Arf1p (142). However, this mechanism is under debate because the Rud3p/Arf1p interaction is not observed in human homologs, plus the GRAB domain is also characterized as microtubule interacting domain that could potentially help nucleate of non-centrosomal

28

microtubules from the Golgi (145,146). As a protein that is closely associated with the Golgi apparatus, perturbation of TRIP230 expression levels changes many Golgi-involved cellular functions. Overexpression or depletion of TRIP230 can generate perturbed Golgi morphology, disturbed microtubule network and obstructed intracellular protein transport (140,147).

Besides binding to Golgi apparatus, TRIP230 also acts as a transcription coactivator in controlling gene expression. Because TRIP230 is mainly localized on Golgi complexes, it is puzzling to tie a connection between TRIP230 as a Golgi-associated protein and transcription coactivator. Luckily, this gap is nicely filled by the observation of subtle but consistent transport of TRIP230 from Golgi to the nucleus during S phase (148). T3-dependent phosphorylation on TRIP230 was later determined as the nuclear import signal for this process. The TRIP230 co-localization with TR also suggests a potential transcription function of TRIP230 (148). This hypothesis was later supported by several cell-based reporter assays and biochemical characterization, identifying the GRAB domain and GA1 region in TRIP230 as the direct TR-binding domains (138). TRIP230 can also utilize other domains to recruit transcription factors. In XRE and HRE signaling pathways, TRIP230 interacts with ARNT PAS-B domains via its coiled-coil domain rather than the GRAB domain (121). Competition on ARNT PAS-B binding was observed among TRIP230, TACC3 and CoCoA, suggesting a very similar mode-of-action within the CCC family members when interacting with ARNT bHLH-PAS transcription factor (4,122).

Although the three CCC family members have low sequence similarity and diverse cellular functions, it is stunning that they still share a conserved role as a transcription coactivator for ARNT function. Adjusting CCC interaction with ARNT could be a potentially efficient way to tune gene expression in many ARNT-containing signaling pathways. However, details about ARNT/CCC interaction and functions are still largely unknown. The crystal structure of ARNT/TACC3 complex described in Chapter 2 will partly address this problem, giving a partial answer to the ARNT/CCC interaction at atomic level. However, structural characterization and functional studies of the other two coiled-coil proteins are still required to deconvolute each CCC member's distinct role in ARNT downstream gene activation in the future.

CHAPTER 2

STRUCTURAL INSIGHTS OF COILED-COIL COACTIVATOR RECRUITMENT IN HIF SIGNALING PATHWAY

The hypoxia inducible factor complex (HIFα/ARNT) requires the association of several transcriptional coactivators to response to hypoxic stress. In addition to the conventional global transcription coactivator CBP/p300 that binds to HIFα transactivation domain (TAD) (95), a new group of transcriptional coactivators called the coiled-coil coactivators (CCCs) interact directly with the second PER-ARNT-SIM (PAS) domain of ARNT (ARNT PAS-B). These less studied transcription coactivators play essential roles in HIF-dependent hypoxia responses and CCCs misregulation is associated with several forms of cancer. To better understand CCC protein recruitment by the HIF complex, we used X-ray crystallography, NMR spectroscopy and biochemical methods to investigate the structure of the ARNT PAS-B domain in complex with the C-terminal fragment of a coiled-coil coactivator protein: transforming acidic coiled-coil coactivator 3 (TACC3). We found that the HIF2α PAS-B domain also directly interacts with TACC3, motivating an NMR data-derived model suggesting a mean by which TACC3 could form a ternary complex with HIF2α PAS-B and ARNT PAS-B via the β-sheet/coiled-coil interactions. These findings suggest that TACC3 could be recruited as a mediator protein that bridges between the HIF2α PAS-B/ARNT PAS-B complex, therefore participating more directly in HIF-dependent gene transcription than previously anticipated.

i. Introduction

Hypoxia inducible factor (HIF) proteins are the central regulators of the mammalian hypoxia response (149), consisting of an O_2-regulated α subunit (HIF1α, 2α and 3α) and the stably expressed β subunit (ARNT or HIFβ) (150). Under hypoxia, stabilized HIFα and ARNT subunits dimerize through the N-terminal bHLH (basic Helix Loop Helix) and two PAS (Per-ARNT-Sim) domains. This heterodimer binds to the HRE (hypoxia response element) promoter with its N-terminal bHLH domains, and controls the transcription of hundreds of target genes such as pro-angiogenic factors and metabolic enzymes (150,151).

HIF target gene regulation depends on the participation of several transcription cofactors. The intrinsically unfolded C-terminal transactivation domain (TAD) of HIFα subunit plays a major role in this process by directly interacting with the global transcription coactivator CBP/p300 (95). Interestingly, recent studies have shown that a newly described group of transcription coactivators involved in cancer development and progression, namely the coiled-coil coactivators (CCCs), could be recruited in a TAD-independent manner (4,122,134). Three CCC family members have been described to date: coiled-coil coactivator (CoCoA) (120), thyroid hormone receptor interacting protein 230 (TRIP230) (121) and transforming acidic coiled-coil 3 (TACC3) (119). Under normal situations, these coactivators play an essential role in the hypoxia response by directly interacting with the ARNT subunit in a promoter-specific way. However, misregulation by overexpression or activating fusions caused by chromosomal translocations (eg. FGFR-TACC3) is sufficient for transformation

and is associated with the development of glioblastoma, renal cell carcinoma and other cancers (152).

To investigate the molecular details of the HIF/CCC complex, we first characterized the structure of a minimal interacting fragments of ARNT/TACC3 using X-ray crystallography and solution NMR spectroscopy. We previously reported that ARNT PAS-B utilizes its Fα-helix as the main interface to directly bind to the C-terminus of TACC3 *in vitro* (134). However, crystal structures solved in this study directly demonstrate that the β-sheet of ARNT PAS-B serves as the TACC3 binding site. Subsequent solution NMR measurements reveal that the HIF2α PAS-B β-sheet also binds to the TACC3 C-terminus. Based on these observations, we generated a ternary complex model of ARNT PAS-B/TACC3/HIF2α PAS-B that is supported by several lines of experimental data. The ternary complex formation among TACC3, ARNT PAS-B and HIF2α PAS-B described here could shed light on the general mechanism of the CCC protein recruitment in HIF signaling and provide a structural framework to inform future anti-cancer therapies.

ii. Results

Overall structure of ARNT/TACC3 - The ARNT PAS-B domain and the last 20 amino acids of the TACC3 parallel dimer (residues 610-631) of the TACC domain have been identified as the minimum binding fragments that are necessary and sufficient to form the ARNT/TACC3 complex (Fig. 2-1a) (119,122,134). However, structural studies of the wild-
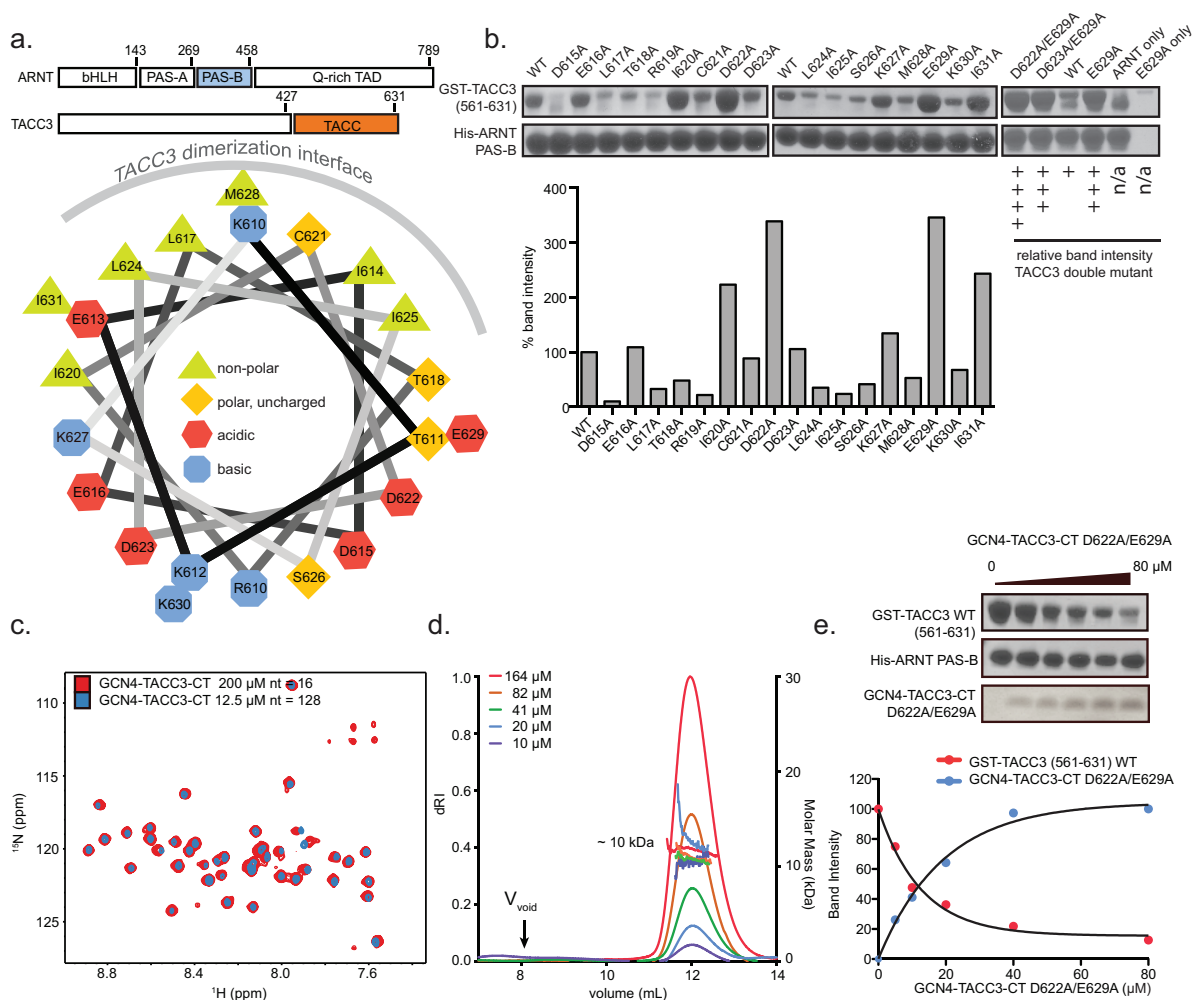
Figure 2-1. GCN4-TACC3-CT D622A/E629A is a dimeric TACC3 mutant that enhances ARNT PAS-B binding. a) Upper panel: Schematic of full length ARNT and TACC3. Direct interacting domains of the two proteins are highlighted in blue (ARNT PAS-B) and orange (TACC domain). Lower panel: Helical wheel of TACC3 (610-631). b) Alanine scan on GST-TACC3 (561-631) by pulldown assay. TACC3 D622A and E629A show dramatically enhanced affinity with ARNT PAS-B. TACC3 D622A/E629A double mutant enhanced the ARNT PAS-B binding further. c) ^{1}H-^{15}N HSQC spectra of 200 μM (red) and 12.5 μM GCN4-TACC3-CT D622A/E629A (blue). Peaks under both conditions are well dispersed and superimposed well, indicating that this TACC3 construct is well-folded and homogeneous under high and low concentration. d) Superdex 75 size exclusion chromatography with inline multi-angle laser light scattering (SEC-MALS) shows that this TACC3 construct forms a constitutive dimer (theoretical molar mass = 9.7 kDa for dimer) at concentrations between 10 - 164 μM. e) Ni pulldown assay shows that GCN4-TACC3-CT D622A/E629A is able to compete with GST-TACC3 (561-631) WT for ARNT PAS-B binding, indicating that GCN4-TACC3-CT D622A/E629A shares the same ARNT PAS-B binding interface with GST-TACC3 (561-631) WT.

type ARNT PAS-B/TACC3 complex are complicated by the moderate affinity (low micromolar) observed for this protein/protein interaction.

To optimize the ARNT/TACC3 complex for crystallography, an alanine scan of the TACC3 C-terminus identified multiple mutations appearing to provide tighter complexes with ARNT PAS-B, as accessed by co-precipitation (Fig. 2-1ab). Among all the mutants, mutations to two surface exposed charged residues (D622A and E629A) exhibited the best effect, while a TACC3 D622A/E629A double mutant provided a still stronger result (Fig. 2-1b). The N-terminal GCN4 fusion with TACC3-CT (610-631) D622A/E629A locks the TACC3 double mutant as a homogeneous dimer (153,154) (Fig. 2-1cd). Well-dispersed peaks representing 43 out of 44 residues in the ^{1}H-^{15}N TROSY spectra demonstrated that this fusion protein is well-folded in solution (Fig. 2-1c). Nicely superimposed peaks in the ^{1}H-^{15}N TROSY spectrum and the constant molar mass determined by size-exclusion multi-angle laser light scattering (SEC-MALS) confirm the GCN4-fusion protein homogeneity in solution at various concentrations (Fig. 2-1cd). The GCN4-TACC3-CT D622A/E629A protein can also compete away the WT GST-TACC3 (561-631) from a pre-formed complex with ARNT PAS-B in a pulldown assay, demonstrating that this stabilized dimeric high affinity TACC3 construct shares the same binding interface on ARNT PAS-B as WT TACC3 (Fig. 2-1e).

The high affinity ARNT PAS-B/GCN4 -TACC3-CT D622A/E629A complex readily crystallized, ultimately providing a 3.15 Å dataset that was phased by molecular replacement using a crystal structure of the isolated ARNT PAS-B domain (PDB ID: 4EQ1) as a search model (Table 2-1) (Fig. 2-2a, PDB ID: 4LPZ) (21). GCN4-TACC3-CT D622A/E629A

Table 2-1. X-ray Crystallography Data Processing and Refinement Statistics.

Data collection		
Crystal	ARNT PAS-B/ GCN4-TACC3-CT D622A/E629A	ARNT PAS-B E362R/ TACC3•HIF2a PAS-B R247E
Space group	C2	P4$_1$2$_1$2
Wavelength (Å)	0.97951	0.97951
Resolution range (Å)	44.3-3.15 (3.20-3.15)	48.07-3.20 (3.26-3.20)
Unique reflections	8,354 (278)	14,853 (705)
Multiplicity	3.5 (2.3)	7.2 (4.8)
Data completeness (%)	94.9 (65.1)	99.9 (98.7)
R_{merge} (%)[a]	10.5 (46.2)	14.8 (100.0)
R_{pim} (%)[b]	—[d]	6.8 (62.9)
I/σ(I)	14.8 (1.55)	11.6 (1.3)
Wilson B-value (Å^2)	73.6	48.3
Refinement statistics		
Resolution range (Å)	44.3-3.15 (3.35-3.15)	48.07-3.20 (3.26-3.20)
No. of reflections R_{work}/R$_{free}$	8,332/832 (954/106)	12,842/626 (1,782/93)
Data completeness (%)	94.2 (73.0)	86.8 (89.0)
Atoms (non-H protein/ions)	2,108/NA	4,082/10
R_{work} (%)	24.7 (27.9)	23.1 (32.3)
R_{free} (%)	27.2 (31.2)	27.0 (34.0)
R.m.s.d. bond length (Å)	0.004	0.008
R.m.s.d. bond angle (°)	0.759	0.93
Mean B-value (Å^2) (protein/ions)	65.1/NA	56.1/52.7
Ramachandran plot (%) (favored/additional/disallowed)[c]	95.9/4.1/0.0	97.5/2.3/0.2
Maximum likelihood coordinate error	0.31	0.47
Missing residues	A: 356-360, 447-454, 465-470. B: 356-360, 447-453, 465-470. C: 626-631.	A: 356-360, 449-451, 465-470. C: 630-631. D: 356-360, 450-452, 465-470. G: 235-238, 345-350.

Data for the outermost shell are given in parentheses.

[a]R_{merge} = 100 $\Sigma_h\Sigma_i|I_{h,i}— \langle I_h\rangle|/\Sigma_h\Sigma_i \langle I_{h,i}\rangle$, where the outer sum (h) is over the unique reflections and the inner sum (i) is over the set of independent observations of each unique reflection.

[b] R_{pim} = 100 $\Sigma_h\Sigma_i [1/(n_h - 1)]^{1/2}|I_{h,i}— \langle I_h\rangle|/\Sigma_h\Sigma_i \langle I_{h,i}\rangle$, where n_h is the number of observations of reflections **h**.
[c]As defined by the validation suite *MolProbity (155)*
[d] Not reported in this version of HKL3000 when processing the data
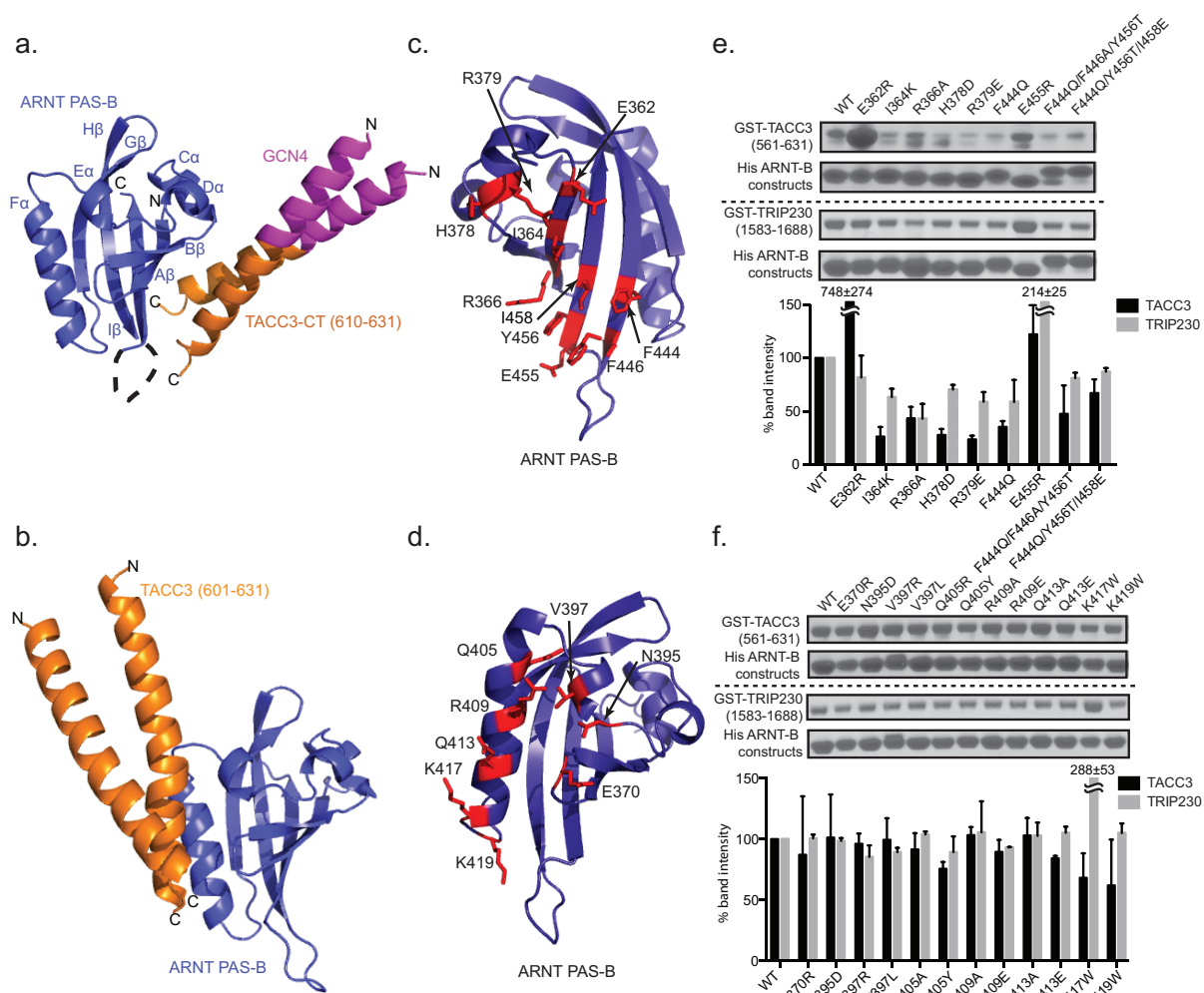
Figure 2-2. Crystal structure and validation of ARNT PAS-B/GCN4-TACC3-CT D622A/E629A complex. a) Crystal structure of ARNT and TACC3 with minimum interacting fragments (PDB ID: 4LPZ): ARNT PAS-B (blue)/GCN4-TACC3-CT D622A/E629A (orange). Dimerization tag GCN4 is shown in magenta. TACC3 is binding to the β-sheet side of the PAS domain in this structure. b) HADDOCK model of ARNT PAS-B/TACC3 complex generated in previous study showing TACC3 binding to ARNT PAS-B on the Fα-helix (134). c) Mutations on ARNT PAS-B β-sheet. d) Mutations on ARNT PAS-B Fα-helix. e) Ni pulldown assay and band quantification (mean ± standard deviation) shows that His-ARNT PAS-B β-sheet mutation perturbs its binding affinity with WT GST-TACC3 (561-631) by either enhancing or reducing the effect, indicating that the β-sheet of ARNT PAS-B is the direct TACC3 binding interface. Each data point is measured in duplicates. Similar to TACC3 interaction, Ni pulldown assay and band quantification (mean ± standard deviation) shows His-ARNT PAS-B β-sheet mutation changes its binding affinity with WT GST-TRIP230 (1583-1688), indicating that the β-sheet of ARNT PAS-B is the major interface for TRIP230 binding. f) Ni pulldown assay and band quantification shows that His-ARNT PAS-B Fα-helix mutation is unable to alter its affinity with WT GST-TACC3 (581-631) and WT GST-TRIP230 (1583-1688), indicating that the Fα-helix is not the major CCC binding interface.

crystalized as a parallel coiled-coil dimer. The TACC3 portion of the fusion protein (orange, residue 610-631) formed major contacts with the β-sheet of ARNT PAS-B with a buried surface area of 2350 Å^2. The GCN4-tag (magenta) does not appear to directly contact with the ARNT PAS-B domain. The extended HI loop of the ARNT PAS-B domain and the last five residues of TACC3 are omitted from this model on account of poor and incomplete electron density (Fig. 2-2a). Interestingly, the ARNT/ TACC3 binding interface shown in this structure is drastically different from a previous NMR-guided computational model where TACC3 makes primary contacts with ARNT PAS-B on the Fα-helix (Fig. 2-2b). This observation makes it critical to further validate the current crystal structure.

Validating the binding interface in ARNT PAS-B/GCN4-TACC3-CT D622A/E629A crystal structure by mutagenesis - To confirm the ARNT/TACC3 binding interface revealed in the ARNT PAS-B/GCN4-TACC3-CT D622A/E629A crystal structure, we extensively mutated the ARNT/TACC3 binding interface on the ARNT PAS-B side. As expected, single and triple mutants on the β-sheet of the PAS domain attenuated TACC3 binding (Fig. 2-2ce), while mutations on the Fα-helix, a surface that is not involved in the crystallography-defined binding interface, didn't show comparable TACC3 binding affinity reduction (Fig. 2-2df), confirming the ARNT PAS-B β-sheet as the major TACC3 binding interface.

We additionally tested the ARNT PAS-B complex with another CCC protein, TRIP230, inquiring if ARNT PAS-B can recruit a different CCC protein in a similar manner as observed before by competition assay (134). As was the case for TACC3, most β-sheet mutants weakened the ARNT/TRIP230 interaction while the Fα-helix mutants showed

minimum effects, demonstrating that the ARNT PAS-B β-sheet is essential for CCC protein recruitments (Fig. 2-2ef).

Electrostatic interaction is critical for ARNT/TACC3 interaction - A close look at the binding interface of ARNT PAS-B and GCN4-TACC3-CT D622A/E629A revealed many interprotein contacts involving charged residues. ARNT PAS-B mutants H378D and R379E markedly reduced complexation with TACC3 in pulldown assays (Fig. 2-2ce), likely by removing an interface-spanning salt bridge by exchanging a +/- charged residue pair for a -/- charged pair (Fig. 2-3a). However, ARNT PAS-B E362R and E455R showed enhanced complex formation with TACC3 (Fig. 2-2ce), likely due to the formation of a new salt bridge with TACC3 D622 and D623 respectively by flipping a -/- charged residue pair to a +/- charged pair (Fig. 2-3a). ^{1}H-^{15}N HSQC spectra of all four mutants overlapped nicely with WT ARNT PAS-B spectrum, demonstrating that the binding perturbations observed in pulldown assays were not caused by protein unfolding or other artifacts (Fig. 2-3b).

Likewise, the high affinity mutants used in crystallography appear to work by alleviating electrostatic repulsion present at the ARNT/TACC3 interface by removing one negatively charged residue in this -/- charged residue pair consisting of TACC3 D622 and ARNT PAS-B E362. To test this hypothesis, a pulldown assay with different combinations of charged residues at the ARNT E362/TACC3 D622 positions was conducted to examine the binding affinity alterations by those mutants. As expected, the highest affinity enhancement was observed in the ARNT E362R/TACC3 WT, ARNT WT/TACC3 D622R and ARNT
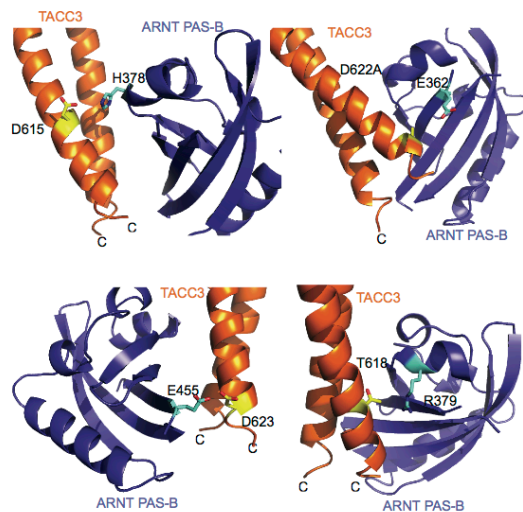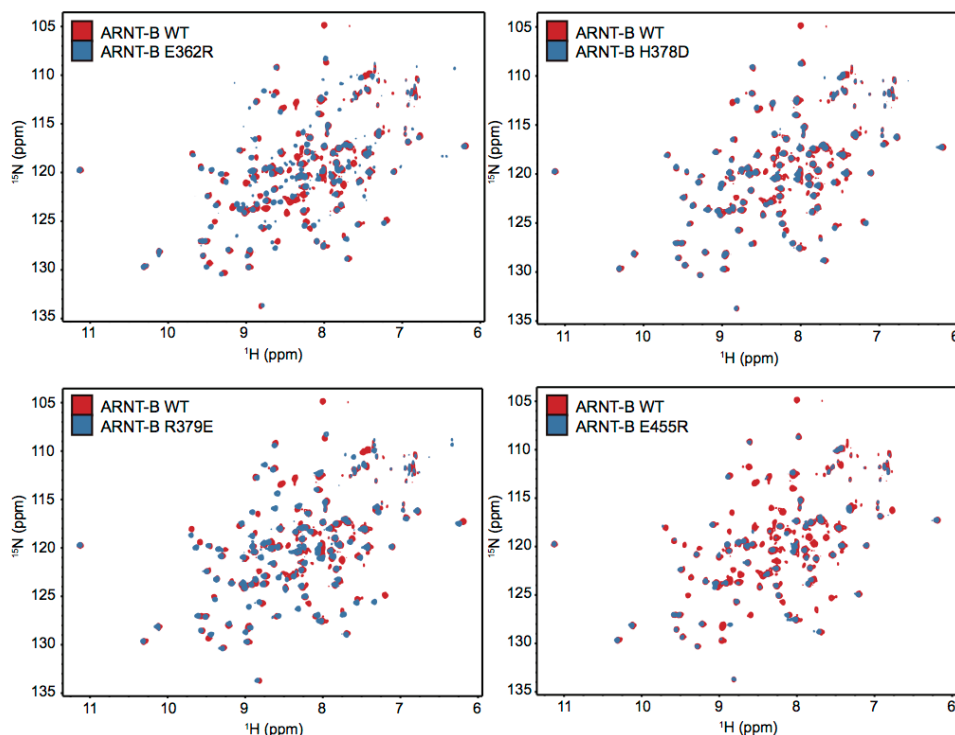
Figure 2-3. Illustration of four key ARNT PAS-B/TACC3 interface residue pairs. a) Most of those pairs consist of charged residues. Changing electrostatic status of those residues could severely affect the ARNT PAS-B/TACC3 interaction (Fig. 2-2ce). b) ^{1}H-^{15}N HSQC spectra of four key ARNT PAS-B mutants shown in panel a. Well-dispersed peaks and good overlap with WT ARNT PAS-B spectrum indicates that those mutants are well-folded and adopt similar structure with WT ARNT PAS-B, further confirming that the effects observed in Fig. 2-2ce are not due to protein misfolding.

WT/D622K complex, where a pair of +/- residues were facing each other at the binding interface. Moderate affinity enhancement was shown in the ARNT WT/TACC3 D622A complex, where one of the negatively charged residues was replaced with a neutral residue. Minimal effects were observed in both the -/- pair ARNT WT/TACC3 WT complex, and the +/+ pair complexes such as ARNT E362R/TACC3 D622R and ARNT E362R/TACC3 D622K (Fig. 2-4ac). Salt bridge formation was further observed in a 3.5 Å crystal structure of ARNT PAS-B E362R/TACC3 (585-631) WT (Fig. 2-4b, Fig. 2-5 PDB ID: 4PKY, Table 2-1). Superimposition of this structure with the ARNT PAS-B/GCN4-TACC3-CT D622A/E629A crystal structure resulted backbone RMSD = 0.9 Å over 153 aligned Cα (Fig. 2-5b). This result demonstrated that ARNT E362 and TACC3 D622 are critical contact spots at the binding interface. Changing the electrostatics of these positions could severely affect the binding. It also confirmed that the β-sheet of ARNT PAS-B, where residue 362 is located, is directly associated in TACC3 complexation. The affinity enhancement by TACC3 E629A is via a similar charge-charge interaction mechanism (data not shown).

ARNT/TACC3 interaction in solution is similar to the crystal structure - To characterize the ARNT PAS-B/TACC3 complex in solution, TACC3 (585-631) was titrated into ^{15}N-labeled ARNT PAS-B, with complex formation monitored by ^{1}H-^{15}N HSQC spectra. Intermediate exchange kinetics were observed in the course of TACC3 titration, suggesting a low to mid-micromolar range affinity. The most severely broadened residues were localized to the β-sheet of ARNT PAS-B (Fig. 2-6ab) (21,156), consistent with the TACC3 binding surface. Analyzing these data with the protein rigid body docking program
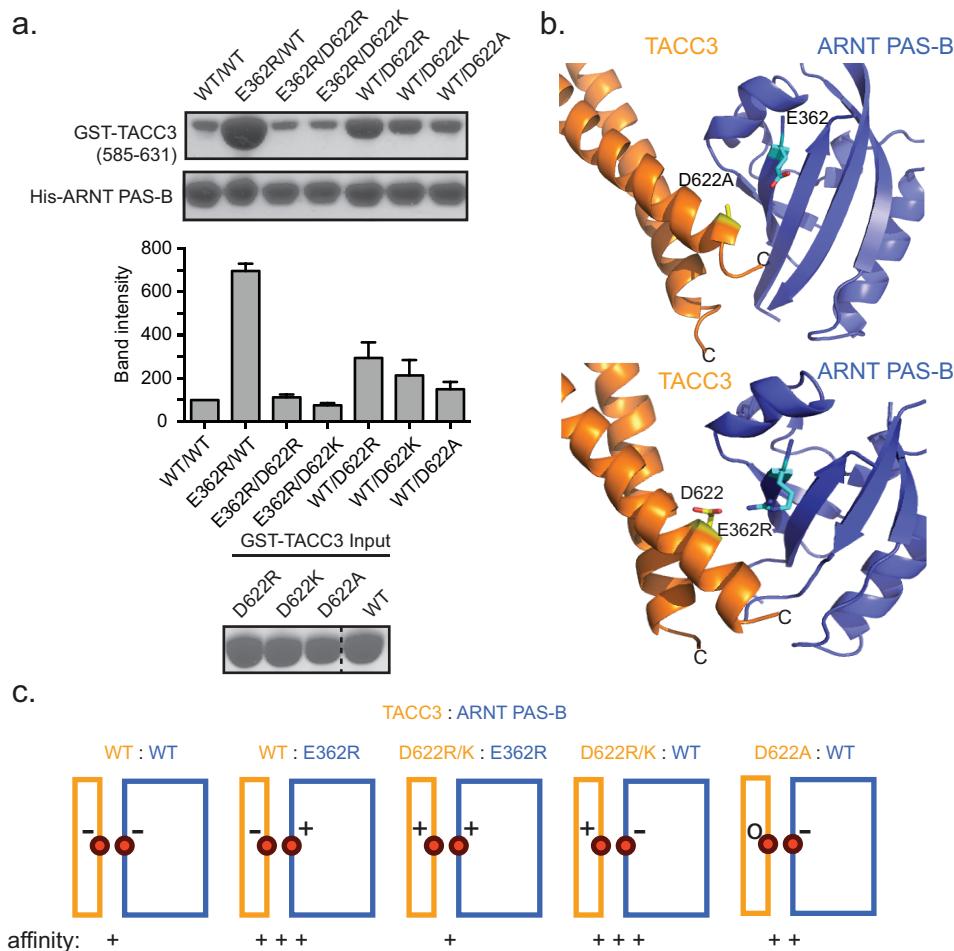
Figure 2-4. Charge swap between residue pairs on ARNT/TACC3 further confirms the interacting interface. a) Ni pulldown assay to quantify binding affinity between residue pair His-ARNT PAS-B E362 and GST-TACC3 D622 mutants. The +/- or -/o charged residue pair (E362R/WT, WT/D622R, WT/D622A) increases the binding affinity, while the -/- or +/+ charged residue pair (E362R/D622R, E362R/D622K) reduces the binding affinity. GST-TACC3 input gel shows equal loading amount. b) Crystal structures of ARNT PAS-B/GCN4-TACC3-CT D622A/E629A (upper panel) and ARNT PAS-B E362R/TACC3 (585-631) WT (lower panel). ARNT PAS-B E362 is at close proximity with TACC3 D622 in both structures. The ARNT/TACC3 affinity improvement in both mutant complexes is possibly caused by the alleviation of -/- electrostatic repulsion in WT proteins. c) Cartoon illustration of charge-charge interaction - affinity relationship at the ARNT/TACC3 binding interface (TACC3: orange; ARNT PAS-B: blue)
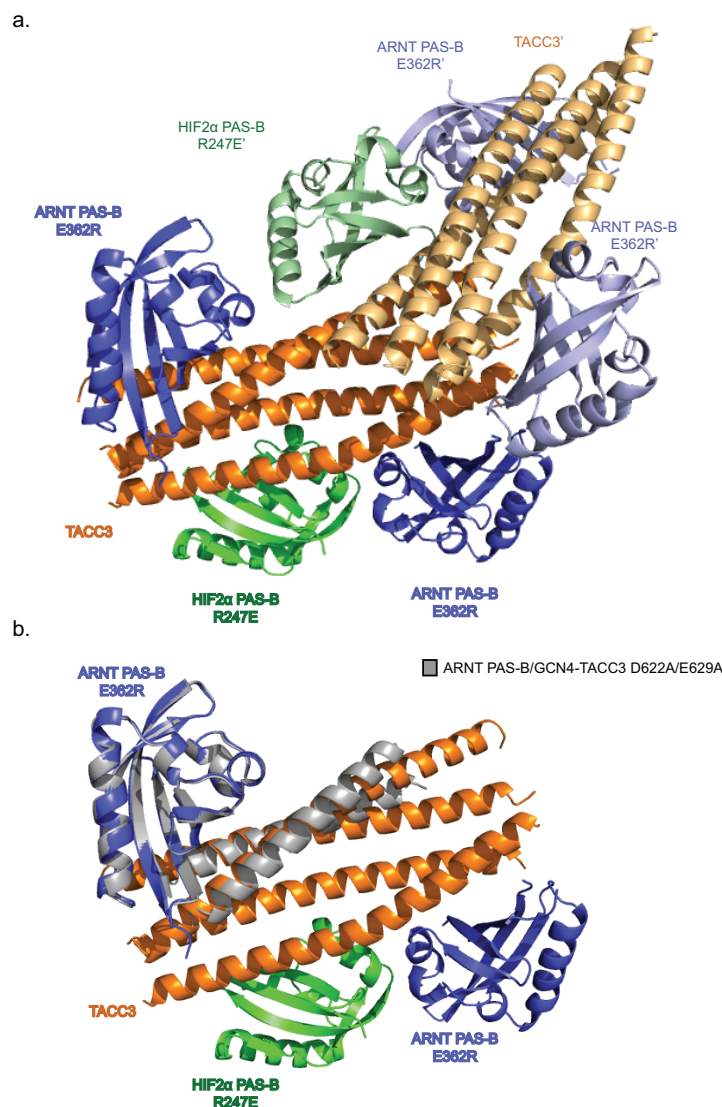
Figure 2-5. ARNT PAS-B E362R/TACC3•HIF2α PAS-B R247E structure adopts similar ARNT/TACC3 interface to the ARNT PAS-B/GCN4-TACC3-CT D622A/E629A structure. a) ARNT PAS-B E362R/ TACC3•HIF2α PAS-B R247E structure with symmetry mates (PDB ID: 4PKY). ARNT PAS-B E362R, TACC3 and HIF2α PAS-B R247E are highlighted in blue, orange and green, respectively. Molecules from single NCS are shown in darker color and bold text. Symmetry related molecules are shown in lighter color. In this structure, ARNT PAS-B E362R forms major contacts with TACC3 (585-631) C-terminus via β-sheet, which is consistent with the ARNT PAS-B/GCN4-TACC3-CT D622A/E629A crystal structure (Fig. 2-2a). HIF2α PAS-B R247E is showing minor contact with the N-terminus part of TACC3 (585-631) with limited interface area (< 500 Å^2), suggesting that is contact is more likely to be created by crystal packing rather than protein/protein interaction. b) Superimposition of ARNT/TACC3 portion of the ARNT PAS-B E362R/ TACC3•HIF2α PAS-B R247E structure with ARNT PAS-B/ GCN4-TACC3-CT D622A/E629A structure (grey). Two structures overlaid very well with a backbone RMSD = 0.9 Å over 153 aligned Cα.
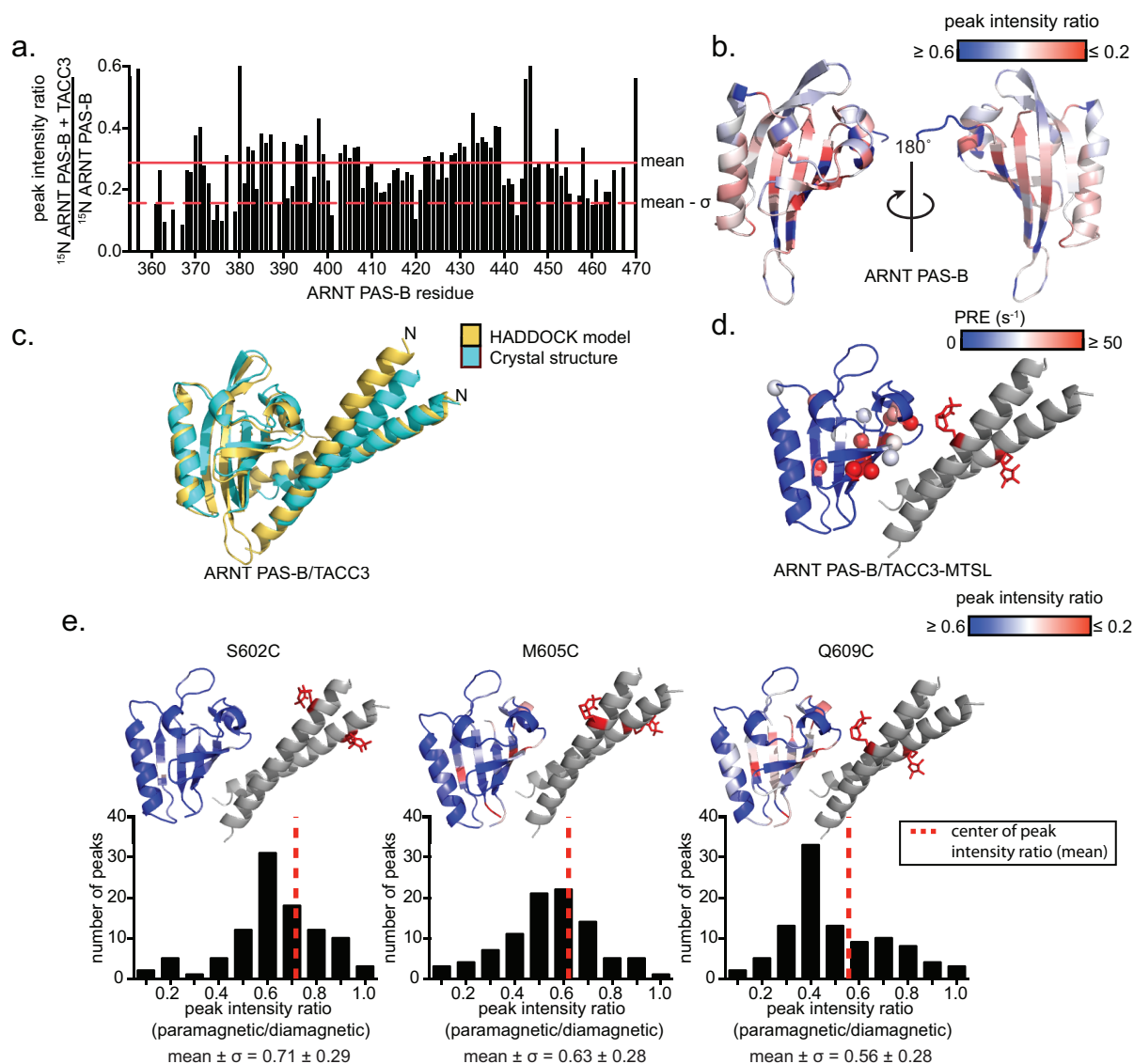
Figure 2-6. ARNT/TACC3 binding in solution is consistent with the crystal structure. a) ARNT PAS-B/TACC3 interaction observed by solution NMR. 225 μM TACC3 (585-631) WT monomer was titrated into 100 μM ^{15}N ARNT PAS-B. Peak broadening ratio of ARNT PAS-B residues caused by TACC3 titration is plotted in bar graph (Mean intensity ratio: red line; mean intensity ratio - σ: red dashed line). b) ARNT PAS-B peak intensity ratio mapped on ARNT PAS-B crystal structure (peak broadening colored from less (blue) to more (red) with the largest broadening effect shown on the β-sheet). c) HADDOCK modeling using the most severely broadened residues (peak intensity ratio ≤ mean - σ) as active residues. The top scored HADDOCK model (lowest energy) is shown in yellow. The superimposition of the HADDOCK model and the ARNT PAS-B E362R/TACC3 (585-631) crystal structure (cyan) resulted a backbone RMSD = 2.1 Å. d) PRE effect of TACC3 (585-631) Q609C-MTSL on ^{15}N ARNT PAS-B is mapped on ARNT PAS-B/GCN4-TACC3-CT D622A/E629A crystal structure. MTSL is shown in red sticks. ARNT PAS-B residues with high PRE effect are highlighted in spheres and the overall PRE effect is colored from less (blue) to more (red). Residues with the large PRE effects (red spheres) are mostly localized near

the spin label.　　e) PRE experiments with multiple TACC3 spin label sites demonstrate consistent relative ARNT/TACC3 orientation in solution and solid state. Upper panel: individual spin label positions on TACC3 (S602, M605, Q609, shown as red sticks) and their corresponding peak broadening effect mapped on ^{15}N ARNT PAS-B (blue: less broadening; red: more broadening). TACC3 Q609C-MTSL induces the most peak broadening effect, whereas TACC3 S602C-MTSL induces the least peak broadening effect. Lower panel: peak broadening ratio distribution induced by the corresponding spin labels. Quantitative analysis shows that the center of mean peak intensity ratio histogram shift to a lower number (to the left, more broadening) when moving the spin label towards the C-terminus of TACC3, demonstrating the spin label to ARNT PAS-B distance (r) follows $r_{S602C} > r_{M605C} > r_{Q609C}$.

HADDOCK (Fig. 2-6c, Table. 2-2a) (157) found TACC3 to interact with the ARNT PAS-B β-sheet in a fashion similar to that observed in the crystal structure. Superposition of the solution NMR-guided model with the ARNT PAS-B E362R/TACC3 (585-631) crystal structure revealed good agreement between solid state and solution measurements (backbone RMSD = 2.1 Å over 163 aligned Cα). Notably, this model is also different from the previous NMR-guided model generated using a similar method that identified the Fα as the principal TACC3 binding site (Fig. 2-2b, Table 2-2c) (134). Nevertheless, based on the fact that Fα-helix mutants fail to attenuate TACC3 binding, as well as the NMR broadening effects mostly clustered on the β-sheet but not the Fα-helix, we have a high degree of confidence that the current structure represents the actual ARNT PAS-B/TACC3 complex in solution.

In further support of the present HADDOCK-calculated model, we examined the relative orientation of the TACC3 dimer in complex with ARNT PAS-B by paramagnetic relaxation enhancement (PRE) (158). In this experiment, the TACC3 (585-631) Q609C position was crosslinked with the spin label (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL) by forming disulfide bond between the spin label and the TACC3 cysteine (Fig. 2-6d). T_2 relaxation rates were measured from ^{15}N ARNT PAS-B residues by TROSY-HSQC in the presence of TACC3 (585-631) Q609C-MTSL under both oxidative condition (paramagnetic status) and reducing condition (diamagnetic status, negative control). Net relaxation rates changed between the oxidative and reducing conditions were converted to PRE values (s^{-1}). Residues in close proximity to the MTSL unpaired electron (approx. < 25 Å) (159) should have faster relaxation rates (enhanced peak broadening) and a corresponding larger PRE value (highlighted in sphere), whereas residues

that are far away from the spin label should exhibit slower relaxation rates (less peak broadening) and a smaller PRE values. For TACC3 Q609C-MTSL, ^{15}N ARNT PAS-B amide groups demonstrating the largest PRE values map to ARNT PAS-B sites proximal to TACC3 Q609 in the crystal structure (Fig 2-6d). MTSL labeling at other TACC3 sites (S602C, M605C) more distal to ARNT PAS-B in the crystal structure attenuated the observed PRE effects (Fig. 2-6e). These solution NMR experiments collected from wild-type proteins provide independent confirmation of complex arrangement observed in the ARNT PAS-B/ TACC3 crystal.

HIF2α PAS-B directly interacts with TACC3 C-terminus - TACC3-dependent transactivation of HIF genes requires the ternary complex formation among HIF2α, ARNT and TACC3 (134). This ternary complex was observed by microscale thermophoresis (MST) assay where HIF2α PAS-B binds tighter to ARNT PAS-B/TACC3 complex than ARNT PAS-B alone (Fig. 2-7ab). This interaction can also be disrupted by a ligand (compound 2) that targeting HIF2α PAS-B cavity, demonstrating that the ternary complex formed among HIF2α PAS-B, ARNT PAS-B and TACC3 is not due to non-specific aggregation (Fig. 2-7c) (58). While the existence of this three-way complex was verified by MST assay, the structural arrangement of this three-protein complex still remains unclear. Previous studies show that HIF2α PAS-B and ARNT PAS-B dimerize through the β-sheet (56,160) (Fig. 2-7d, left panel). Disruption of this PAS-B/PAS-B interaction by mutagenesis or truncation seriously reduces HIF transcription factor activity (49). Unexpectedly, our present data
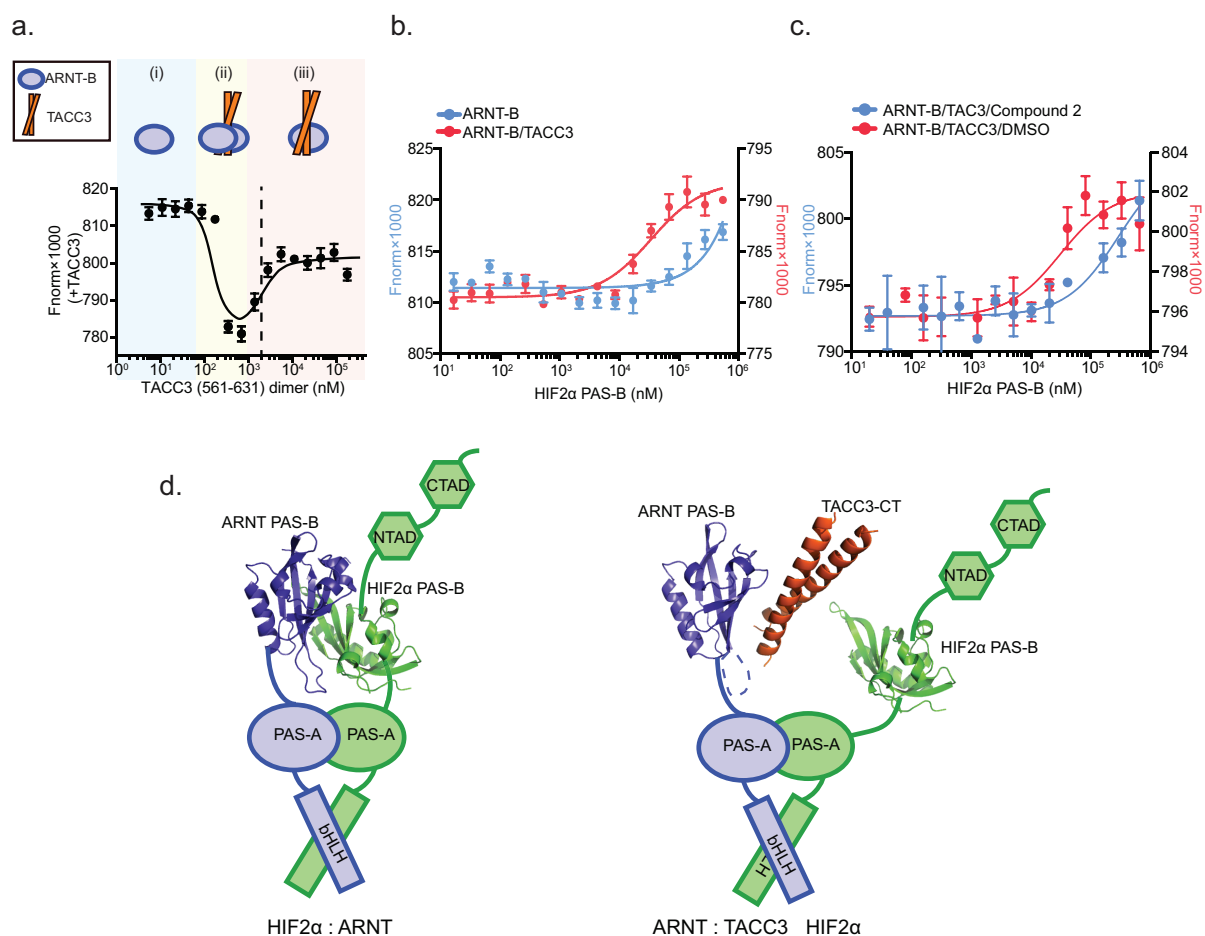
Figure 2-7. HIF2α PAS-B forms ternary complex with ARNT PAS-B/TACC3. a) TACC3 forms a complex with ARNT PAS-B E362R. TACC3 (561-631) dimer is titrated into ARNT PAS-B E362R-fluor. Binding events were characterized by change in thermophoresis. With the increase of TACC3 concentration, three binding stages are observed: (i) ARNT PAS-B E362R alone (blue); (ii) two ARNT PAS-B E362R bind to one TACC3 dimer (yellow); (iii) one ARNT PAS-B E362R binds to one TACC3 dimer (red). The following experiments were conducted at 2 μM TACC3 dimer concentration where the complex is formed at 1:1 ratio (dashed line). b) HIF2α PAS-B forms ternary complex with ARNT PAS-B E362R/TACC3 complex. Titrating HIF2α PAS-B into ARNT PAS-B E362R-fluor/TACC3 complex (red) showed curve saturation at ~100 μM HIF2α PAS-B concentration, indicating that HIF2α PAS-B is able to bind to ARNT PAS-B E362R-fluor/TACC3 and forms a ternary complex with an apparent K_d at 30 μM. Titrating HIF2α PAS-B into ARNT PAS-B E362R-fluor alone (blue) does not show saturation, which is consistent with the low affinity between HIF2α PAS-B and ARNT PAS-B E362R. c) Small molecule targeting HIF2α PAS-B β-sheet (compound 2) reduces HIF2α PAS-B ability to form ternary complex with ARNT/TACC3. Titrating equimolar mixture of HIF2α PAS-B/compound 2 into ARNT PAS-B E362R-fluor/TACC3 showed at least 10-fold larger apparent K_d compared to the DMSO control (apparent $K_d \approx 30$ μM), indicating reduced affinity between HIF and ARNT/TACC3 in the presence of compound 2. This result also demonstrated that HIF2α PAS-B β-sheet is critical for this ternary complex formation. d) Left panel: HIF2α and ARNT dimerize via three major contacts: bHLH, PAS-A and PAS-B. The

48

PAS-B domain contacts are formed mainly through the β-sheet interaction according to a crystal structure solved in previous study (PDB ID: 3F1O).   Right panel: TACC3 occupies the β-sheet of ARNT PAS-B, which requires HIF2α PAS-B to find a new binding interface other than the ARNT PAS-B β-sheet in order to form the ternary structure and initiate transcription.
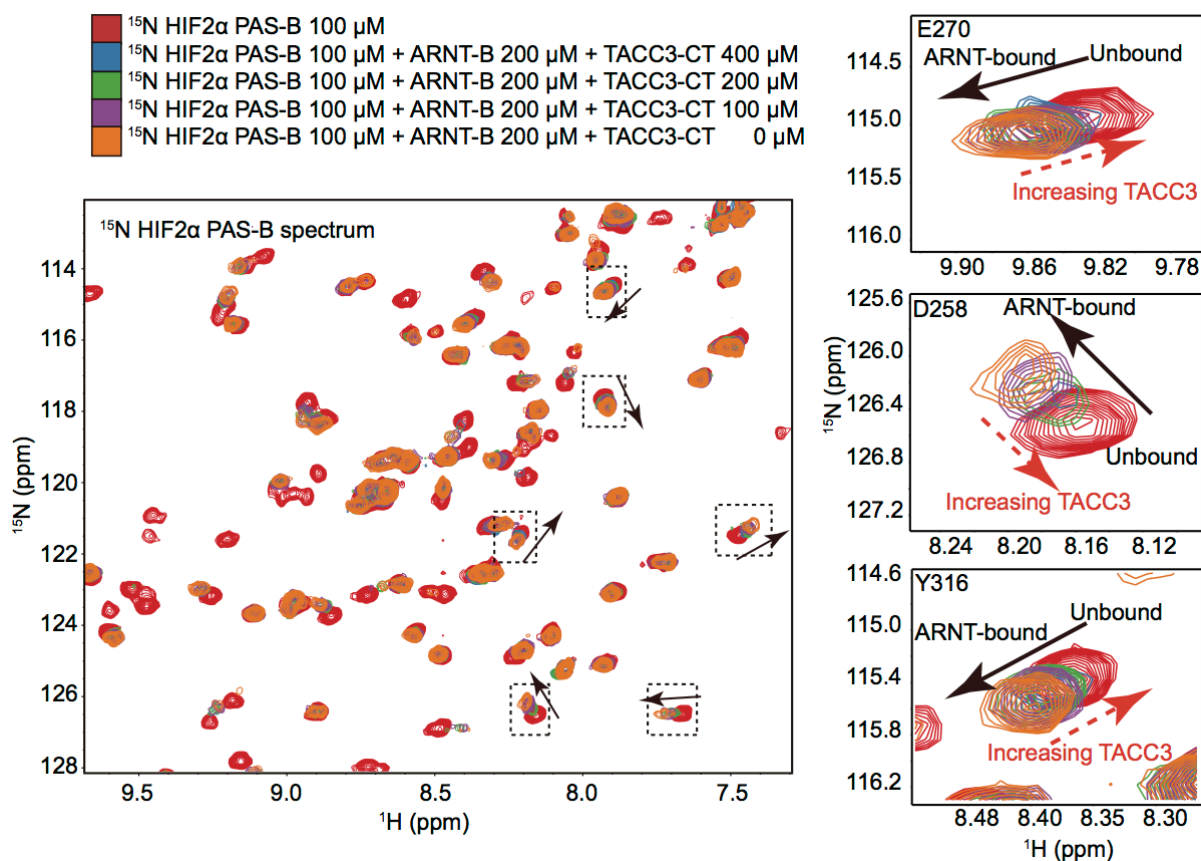
Figure 2-8. TACC3 substitutes HIF2α PAS-B from a pre-formed [15]**N HIF2α PAS-B/ARNT PAS-B complex.** 100 µM ^{15}N HIF2α PAS-B was incubated with 200 µM ARNT PAS-B to form heterodimer under equilibrium condition. Direct interaction with ARNT PAS-B caused both peak broadening and peak shifting. To simplify the analysis, only shifted peaks are studied (direction of peak shift is highlighted with black arrow). Titrating TACC3 (585-631) to this heterodimer makes the ^{15}N HIF2α PAS holo peaks progressively return to their apo state in a more broadened manner (direction of peak return is highlighted in red dashed arrow), indicating that TACC3 interacts and replaces HIF2α PAS-B from direct ARNT PAS-B binding.
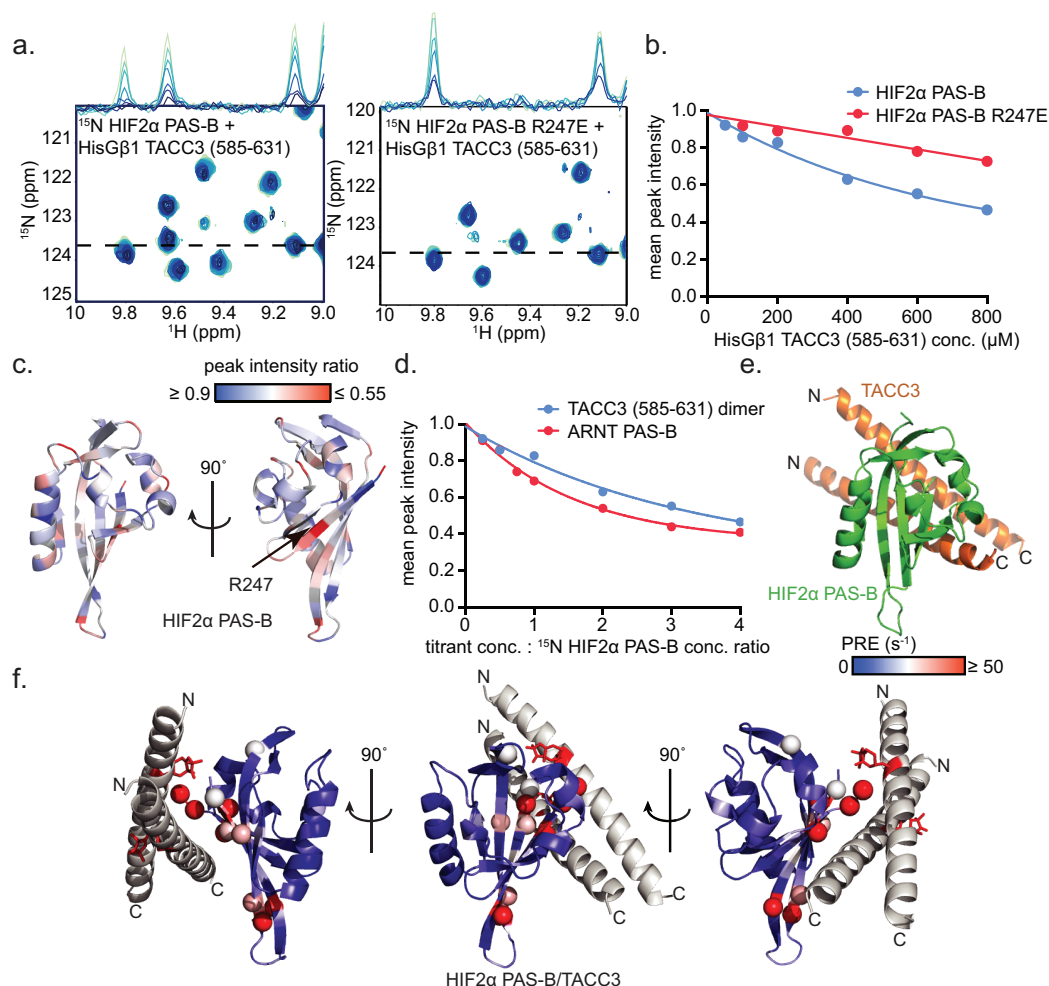
Figure 2-9. HIF2α PAS-B directly interacts with TACC3. a) HisGβ1 TACC3 (585-631) titration induces peak broadening effect on ^{15}N HIF2α PAS-B WT and ^{15}N HIF2α PAS-B R247E, indicating direction interaction between HIF2α PAS-B and TACC3. TACC3 concentration-dependent peak broadening is illustrated as 1D traces at the ^{1}H dimension. b) Quantification of panel a shows weaker peak broadening effect on HIF2α PAS-B R247E compared with HIF2α PAS-B WT, suggesting that residue R247 is critical for TACC3 binding. c) ^{15}N HIF2α PAS-B peak broadening ratio is mapped on HIF2α PAS-B crystal structure (PDB ID: 3F1O, chain A) (peak broadening colored from less (blue) to more (red) with the largest broadening shown on the β-sheet). Residue R247 was severely affected by the TACC3 binding (black arrow). d) ^{15}N HIF2α PAS-B/HisGβ1 TACC3 (585-631) interaction shows comparable peak broadening effect with ^{15}N HIF2α PAS-B/ARNT PAS-B interaction, demonstrating that the TACC3-induced peak broadening is not due to non-specific aggregation. e) HIF2α PAS-B/TACC3 complex HADDOCK model (most populated) was generated utilizing the top 20 broadened residues in ^{15}N HIF2α PAS-B/TACC3 titration experiment as active residues (see Table 2-2b). The most abundant cluster from the prediction is shown here, where TACC3 is binding to HIF2α PAS-B at the β-sheet. f) TACC3-Q609C-MTSL induces significant PRE effect at positions that are adjacent to the spin label as predicted in the HIF2α PAS-B/TACC3 HADDOCK model, demonstrating that the model is truly representing the complex formation in solution.

51

shows that TACC3 binds to ARNT PAS-B on the same β-sheet that is used by HIF2α PAS-B. This competitive binding mechanism raises the prospect of TACC3 disrupting the PAS-B/PAS-B contacts within the heterodimeric HIF transcription factor and displacing HIF2α PAS-B from ARNT PAS-B binding (Fig. 2-7d, right panel). This replacement was observed by solution NMR using ^{15}N HIF2α PAS-B as a probe (Fig. 2-8). The progressive, TACC3-dependent retention of ^{15}N HIF2α PAS-B peaks from ARNT-bound to apo position suggests that TACC3 has substituted HIF2α PAS-B from the initial HIF2α PAS-B/ARNT PAS-B complex.

Interestingly, TACC3 not only displaces HIF2α from ARNT binding, but also causes peak broadening effects on HIF2α PAS-B residues, suggesting direct interactions between TACC3 and HIF2α PAS-B. Titrating HisGβ1-TACC3 (585-631) into ^{15}N HIF2α PAS-B revealed TACC3-dependent peak broadening (Fig. 2-9a, left panel, Fig. 2-9b, blue curve) mapping predominantly to HIF2α PAS-B β-sheet (Fig. 2-9c). HIF2α PAS-B residue R247 was particularly broadened (Fig. 2-9c, right panel). Mutating HIF2α R247 into glutamic acid attenuated the overall peak broadening effect (Fig. 2-9a, right panel, Fig. 2-9b, red curve), further confirming HIF2α PAS-B β-sheet as the TACC3 binding site. Moreover, levels of ^{15}N HIF2α PAS-B WT peak intensity changes induced by TACC3 was comparable to the peak broadening induced by ARNT PAS-B in HIF2α PAS-B/ARNT PAS-B interaction (160), indicating that the HIF2α PAS-B/TACC3 interaction observed in this study is not due to non-specific aggregation (Fig. 2-9d). Based on NMR mapping data, we generated a

Table 2-2. Active residues in HADDOCK modeling.

a. Active residues in ARNT PAS-B/TACC3 (585-631) dimer docking.

ARNT PAS-B	TACC3 dimer
374, 420, 443, 401, 379, 365, 460, 375, 361, 389, 396, 462	610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 1610, 1611, 1612, 1613, 1614, 1615, 1616, 1617, 1618, 1619, 1620, 1621, 1622, 1623, 1624, 1625, 1626, 1627, 1628, 1629, 1630

b. Active residues in HIF2α PAS-B/TACC3 (585-631) docking dimer docking

HIF2α PAS-B	TACC3 dimer
245, 247, 248, 250, 255, 268, 274, 282, 284, 310, 312, 313, 321, 323, 335, 337, 338, 342, 343, 344	602, 603, 605, 606, 609, 610, 612, 613, 616, 617, 619, 620, 623, 624, 626, 627, 630, 1604, 1605, 1607, 1608, 1611, 1612, 1614, 1615, 1618, 1619, 1621, 1622, 1625, 1626, 1628, 1629

c. Active residues used in previous reported ARNT PAS-B/TACC3 (585-631) monomer docking (Partch & Gardner, *PNAS.* 2011) (134)

ARNT PAS-B	TACC3 monomer
409, 370, 369, 417, 397, 398, 430, 395	595, 596, 597, 598 ,599, 600, 601,602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630

HADDOCK model of TACC3 C-terminus dimer binds to HIF2α PAS-B on the β-sheet (Fig. 2-9e, Table 2-2b).

HIF2α PAS-B/TACC3 HADDOCK model was further validated by PRE experiments where TACC3 (585-631) Q609C-MTSL induced PRE effects on several ^{15}N HIF2α PAS-B β-sheet residues that are adjacent to the spin label (red stick) (Fig. 2-9f).

Model of ARNT/TACC3/HIF ternary complex - The biophysical characterization and computational model of HIF2α PAS-B/TACC3 complex described above provides the structural basis for ARNT/TACC3/HIF ternary complex assembly. By comparing the TACC3 binding mode with both HIF2α and ARNT PAS-B, it is highly possible that TACC3 could use both sides of the coiled-coil dimer to recruit HIF2α and ARNT at the same time via their PAS-B β-sheets. A starting model for ARNT/TACC3/HIF ternary complex was generated by superimposing the TACC3 dimer from the ARNT PAS-B/TACC3 crystal structure and HIF2α PAS-B/TACC3 HADDOCK model (Fig. 2-10a). In this model, ARNT PAS-B and HIF2α PAS-B β-sheets are making contacts with opposite surfaces of the TACC3 dimer. To validate this model, distance restraints between ARNT PAS-B S451C-CMTSL and ^{15}N HIF2α PAS-B residues were measured by PRE under conditions with or without TACC3 (Fig. 2-10b). In the absence of TACC3, only a small portion of the HIF2α PAS-B residues located at the beginning of Aβ is affected by the spin label (Fig. 2-10c). This observation is in agreement with the anti-parallel HIF2α PAS-B/ARNT PAS-B assembly in a previous determined crystal structure (Fig. 2-10d) (56). However, the PRE effect on ^{15}N HIF2α PAS-B was dramatically enhanced in the presence of TACC3, especially on the Hβ and Iβ,
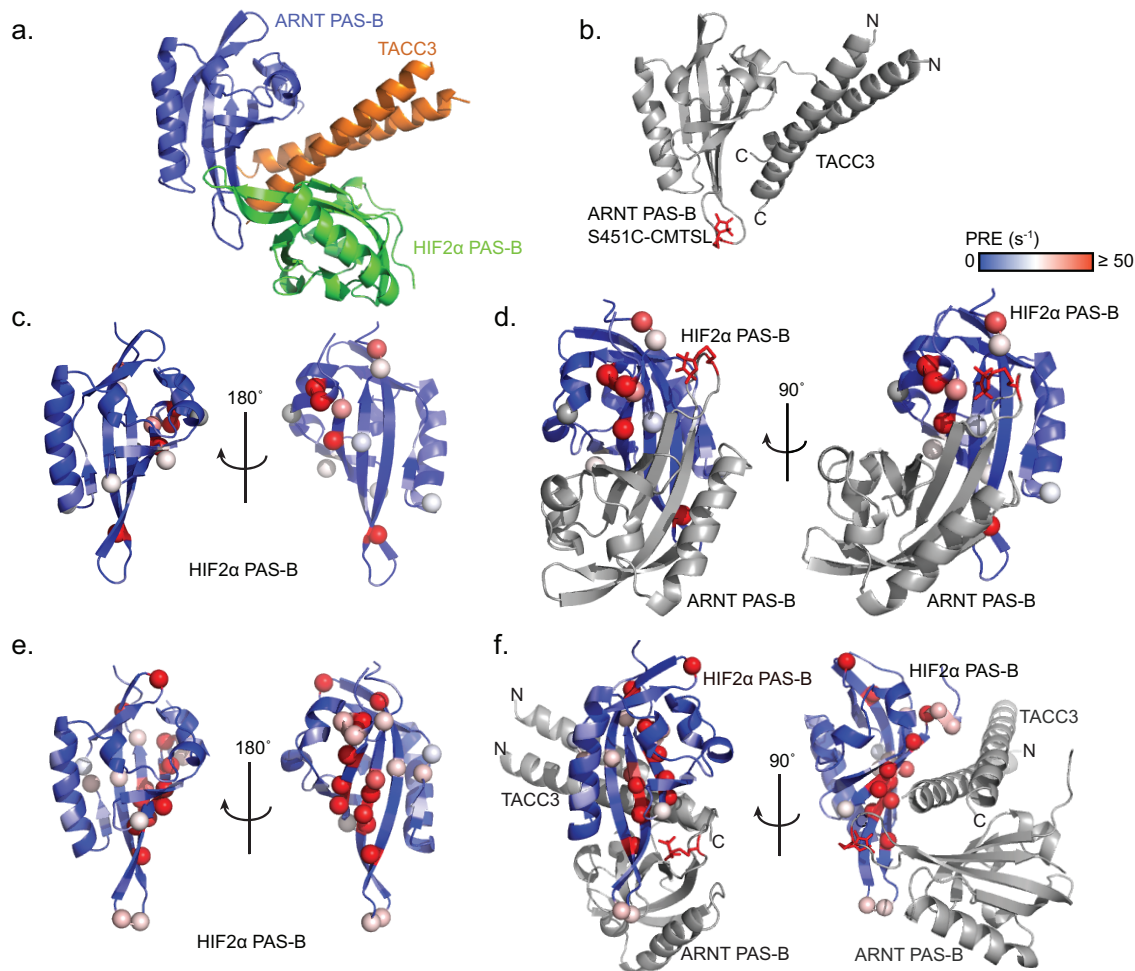
Figure 2-10. Characterization of the ARNT PAS-B/TACC3/HIF2α PAS-B ternary complex. a) HADDOCK model-based ARNT PAS-B/TACC3/HIF2α PAS-B ternary complex. b) ARNT PAS-B/TACC3 complex labeled with CMTSL at ARNT residue 451 (red sticks). c) Titrating ARNT PAS-B S451C-CMTSL into ^{15}N HIF2α PAS-B induces several PRE effect on the β-sheet of HIF2α PAS-B. Highly affected HIF2α PAS-B residues are highlighted in spheres and overall PRE effect is colored from less (blue) to more (red). d) Fitting the PRE effect in Fig. 2-10c on HIF2α PAS-B/ARNT PAS-B crystal structure (PDB ID: 3F1O). CMTSL is shown in red sticks. The highly affected residues are localized close to the CMTSL spin label. e) Adding TACC3 into ^{15}N HIF2α PAS-B/ARNT PAS-B S451C-CMTSL heterodimer induces stronger PRE effect on ^{15}N HIF2α PAS-B β-sheet, indicating a shorter distance between CMTSL label and HIF2α PAS-B residues and a TACC3-dependent complex rearrangement. f) TACC3-dependent PRE effects in Fig. 2-10e are mapped on the ternary structure model in Fig. 2-10a (blue - red: low to high PRE effect), highlighting the most affected residues in red spheres. Most spheres are surrounded to the S451C-CMTSL spin label (red sticks), suggesting a good agreement between the experimental data and the ternary structure model.

indicating a closer proximity with the ARNT PAS-B S451C-CMTSL spin label and also a different protein complex arrangement (Fig. 2-10e). This result could fit very well in the ternary structure model where most of the spheres are clustered in the vicinity of ARNT PAS-B S451C-CMTSL spin label (red sticks) (Fig. 2-10f). The PRE effect observed on the Aβ and the loop between Gβ and Hβ could be possibly from the HIF2α PAS-B/ARNT PAS-B fraction of the mixture. The importance of HIF2α PAS-B β-sheet in ternary complex formation is further confirmed by MST experiment using a small molecule (compound 2) that induces conformational change on HIF2α PAS-B β-sheet (Fig. 2-7c) (58). Taken together, the ternary complex model is validated via different experimental methods and could accurately represent the in-solution association among these three proteins.

iii. Discussion

Our ARNT PAS-B/TACC3 structural study reveals a unique approach to recruit coiled-coil transcription coactivators in HIF signaling pathway comparing to the conventional TAD-dependent mechanism. In this study, we solved two crystal structures of ARNT PAS-B/TACC3 complexes and characterized the wild-type ARNT/TACC3 interaction in solution. Those results consistently show that the β-sheet of ARNT PAS-B is the direct interface for TACC3 binding. Interestingly, this β-sheet has also been found critical for many other intermolecular interactions, such as with HIFα PAS-B (56), raising the need to include TACC3 or other CCC protein recruitment with the existing HIFα/ARNT signaling model where HIFα and ARNT dimerize through the same PAS-B β-sheet (49). Disrupting
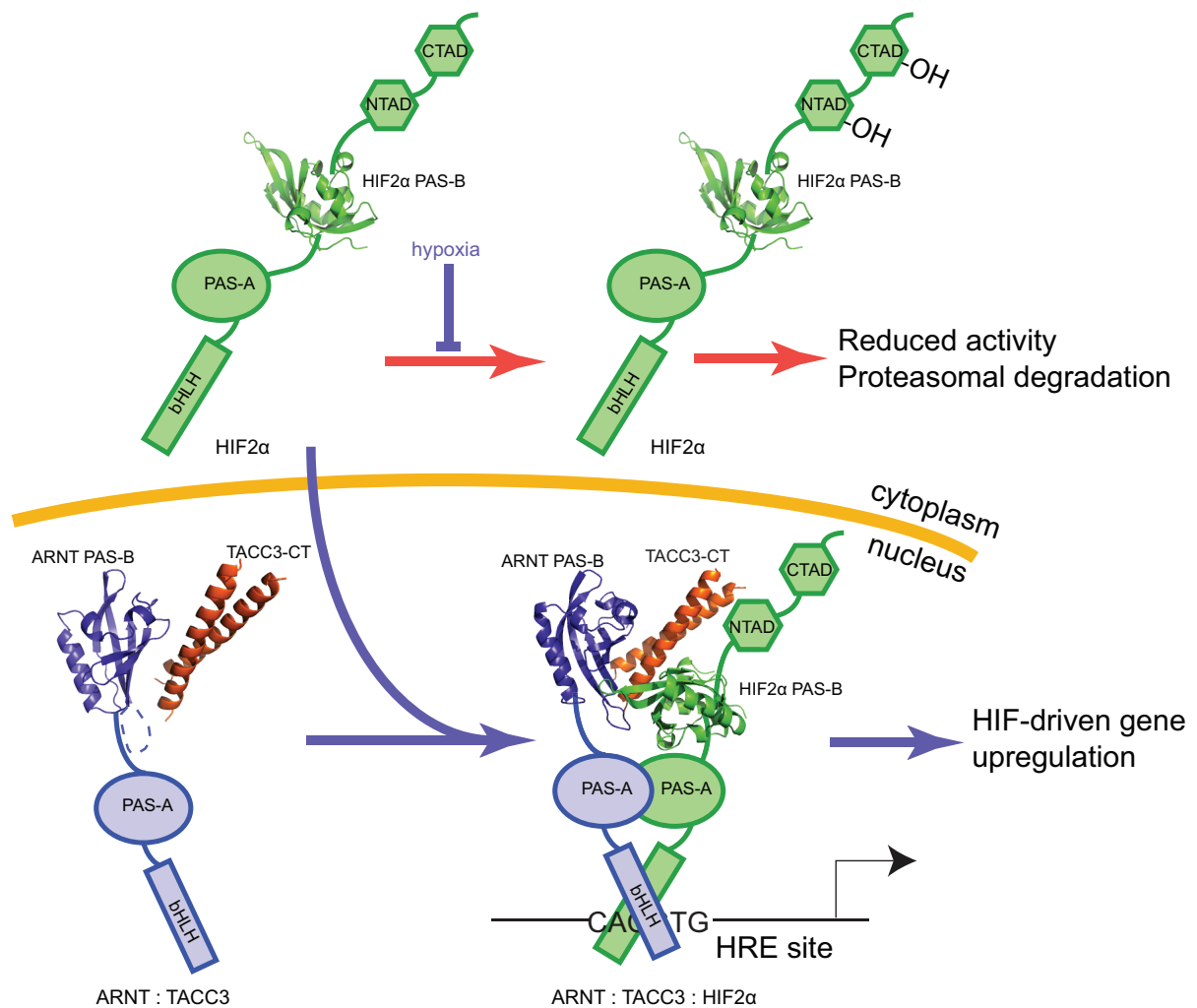
Figure 2-11. TACC3-participated HIF2α transcription activation model. HIF complexes are bHLH-PAS heterodimers that include an O_2-sensitive HIFα subunit and a constitutive ARNT subunit. In addition to binding HIF2α, ARNT PAS-B directly recruits CCC proteins such as TACC3 (122,134). Under normoxia, O_2-dependent hydroxylation of HIF2α decreases its abundance and activity (red arrows) (150). Such modifications do not happen on ARNT and TACC3 so the ARNT/TACC3 complex is readily assembled in the nucleus. Hypoxia stops the hydroxylation, allowing HIF2α to accumulate in the nucleus and complex with ARNT/TACC3. This heterotrimer binds to hypoxia responsive enhancer (HRE) sites, controlling target gene transcription (blue arrows).

this PAS-B/PAS-B interface would result in the dissociation of the complex and reduced HIF activity. To clarify this phenomenon, we discovered a linkage between the ARNT/TACC3 crystal structure and the HIFα/ARNT signaling mechanism, which could be represented in an ARNT/TACC3/HIF ternary complex model. Our model suggest that although TACC3 occupies the usual HIF2α PAS-B binding site on ARNT PAS- B, a new binding interface for HIF2α PAS-B will be created on the opposite side of the TACC3 dimer. TACC3 serves as a bridge to bring the two PAS-B domains together to form the ternary complex. As a result, recruiting TACC3 to the β-sheet of ARNT PAS-B will not impair the integrity of the HIFα/ARNT complex and their transcription activity (Fig. 2-11).

Notably, the TACC3 D622N mutation has been discovered as a somatic mutation associated with cancer (161). We postulated that this mutation could function in a similar manner as TACC3 D622A in our crystallography study, which increases the affinity with ARNT PAS-B and therefore enhances the HIF-dependent gene upregulation.

To summarize, ARNT/CCC interaction plays an important role in mammalian hypoxia response. Since ARNT is the common partner for all three HIF paralogs, blocking ARNT/CCC could affect all HIF containing transcriptions. This makes targeting ARNT/CCC complex a more practical way to regulate HIF activity in cancer therapy compared to the previously described HIFα-specific methods (57,58). *In vitro* small molecule inhibitors have already been designed to target the water-binding cavity inside ARNT PAS-B in order to induce vast conformational changes and disrupt the full length ARNT/TACC3 interaction (21). Continuing development of inhibitors that have higher potency *in vivo* could be a

promising therapeutic route for HIF-dependent cancer treatment and powerful tool for studying CCC function in hypoxia signaling pathways.

iv. Acknowledgements

CHAPTER 3

REGULATING THE ARNT/TACC3 INTERACTION BY SMALL MOLECULES

For several well-documented reasons, it has been challenging to develop artificial small molecule inhibitors of protein/protein complexes. Such reagents are of particular interest for transcription factor complexes given links between their mis-regulation and disease. Here we report parallel approaches to identify regulators of a hypoxia signaling transcription factor complex, involving the ARNT subunit of the HIF (Hypoxia Inducible Factor) activator and the TACC3 (Transforming Acidic Coiled-coil Containing Protein 3) coactivator. In one route, we used *in vitro* NMR and biochemical screening to identify small molecules that selectively bind within the ARNT PAS (Per-ARNT-Sim) domain that recruits TACC3, identifying KG-548 as an ARNT/TACC3 disruptor. A parallel, cell-based screening approach previously implicated the small molecule KHS101 as an inhibitor of TACC3 signaling. Here, we show that KHS101 works indirectly on HIF complex formation by destabilizing both TACC3 and the HIF component HIF1α. Overall, our data identify small molecule regulators for this important complex and highlight the utility of pursuing parallel strategies to develop protein/protein inhibitors.

i. Introduction

While many small molecule enzyme inhibitors and receptor ligands that modulate cellular signaling pathways have been discovered for research and therapeutic use, comparable reagents that target non-enzymatic protein/protein interactions are relatively rare. While such compounds are available for several systems, technical issues - from the suitability of compounds in screening libraries to the difficulty of predicting "druggable" sites (162,163) - complicate the development of specific inhibitors of targeted protein/protein interactions.

Such inhibitors have been particularly sought for transcription factors and their associated regulatory proteins (162,164), given well-validated links between misregulation of these proteins and disease. Here we focus on one such complex as a model: hypoxia inducible factor (HIF), the central regulator of the mammalian hypoxia response (149). HIF is a heterodimer of two bHLH-PAS (basic Helix Loop Helix - Per-ARNT-Sim) subunits, including a HIFα paralog (HIF1α, 2α, 3α) and aryl hydrocarbon receptor nuclear translocator (ARNT, also known as HIFβ) (Fig. 3-1a). While O_2-dependent post-translational hydroxylation normally lowers both of HIFα abundance and activity, these modifications are reduced under hypoxia and allow HIFα to accumulate in the nucleus (150).

Subsequently, HIF complexes form and control the expression of several hundred genes, including potent angiogenic and growth factors (151). As such, abnormally high levels of HIF correlate with several forms of cancer, suggesting that HIF inhibitors could potentially block tumor formation and progression (150). Such inhibition might be achieved
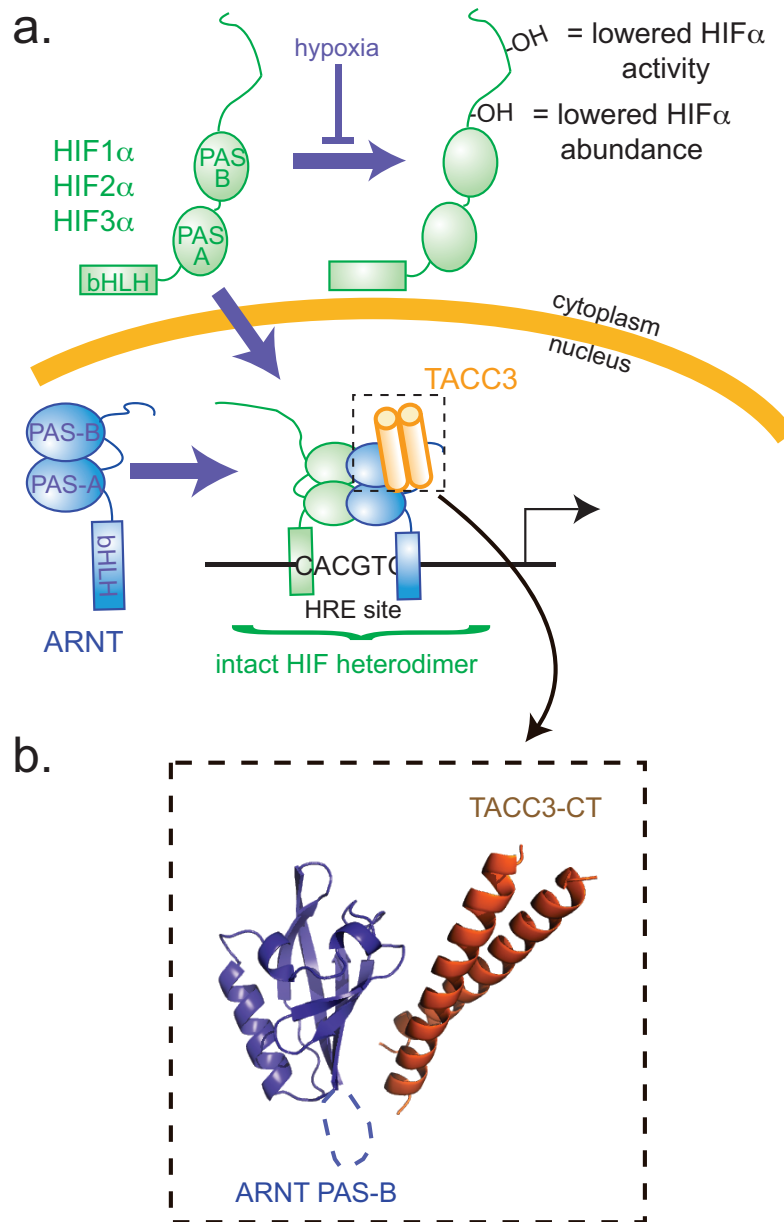
Figure 3-1. Overview of the ARNT/TACC3 complex. a) Schematic of HIF complexes, which are bHLH-PAS heterodimers that include an O_2-sensitive HIFα subunit and a constitutive ARNT subunit. Under normoxia, O_2-dependent hydroxylation of HIFα decreases its abundance and activity (150). Hypoxia stops these modifications, allowing HIFα to accumulate in the nucleus and dimerize with ARNT. This heterodimer binds to hypoxia responsive enhancer (HRE) sites, controlling target gene transcription. In addition to binding HIFα, ARNT PAS-B directly recruits CCC proteins (122,134). b) Structural model of an ARNT/CCC complex showing how the TACC3 coiled-coil interacts with the β-sheet of ARNT PAS-B, similar to where HIFα PAS-B binds.

by blocking the HIFα and ARNT interaction, which uses interchain contacts between bHLH and PAS (Per-ARNT-Sim) domains (56,160,165-167). While we have successfully found inhibitors that use this approach by exploiting a ligand-binding cavity within one of the HIF2 PAS domains (56,166), differences among HIFα sequences suggest that this route is paralog-specific.

To simultaneously inhibit all HIF complexes, we considered targeting interactions between the ARNT subunit, shared among these complexes, with transcriptional coactivators. This strategy is predicated on the ARNT PAS-B domain (Fig. 3-1) directly recruiting coiled-coil coactivators (CCCs) to HIF for proper transcriptional regulation (4,122,134). By depleting endogenous proteins or overexpressing mutants, we found that HIF complexes differentially utilize several CCC proteins including thyroid hormone receptor interacting protein 230 (TRIP230 (121)), Coiled-Coil Coactivator (CoCoA (120)) and transforming acidic coiled-coil 3 (TACC3 (119)) at different promoters (134). Combining biophysical and mutagenesis data, we generated a structural model of the ARNT/TACC3 complex, showing that CCC proteins use part of the coiled-coil domain to bind to ARNT PAS-B on the β-sheet (Fig. 3-1b) (122,134). Notably, the CCC-binding surface on ARNT is near where other PAS domains bind cofactors that modulate their protein/protein interactions, leading us to hypothesize that artificial ARNT-binding compounds might similarly control ARNT PAS-B/CCC interactions to regulate HIF activity.

Here we characterize the mechanisms of action of two small molecule inhibitors of ARNT/TACC3 signaling, identified from independent *in vitro* target-based and cell-based phenotypic screens. The first approach took advantage of our NMR studies of ARNT PAS-B

(168), letting us use this method to screen over 760 compounds (169,170) for protein binding. One ARNT-binding compound, KG-548, binds in a cavity adjacent to the TACC3 binding site and displaces CCCs from ARNT *in vitro*. In parallel, we investigated KHS101, a thiazole derivative originally identified in a cell-based phenotypic screen (171) and that specifically induces neuronal differentiation (172). Based on KHS101 affecting aspects of the neuronally-expressed ARNT homolog, ARNT2, in neural progenitor cells (NPCs) and crosslinking to TACC3 in NPC lysates, the question arises as to whether KHS101 directly disrupts the ARNT2/TACC3 interaction. We show that KHS101 does not directly bind to the minimal ARNT or TACC3 interacting domains, but instead promotes TACC3 protein turnover within cells, indicating an indirect mode of function. Notably, KHS101 also promotes the turnover of the HIF1α subunit itself and interferes with HIF-driven transcription in living cells under hypoxia, as anticipated by these effects on the levels of both activator (HIF1α) and coactivator (TACC3) components. Taken together, KG-548 and KHS101 provide useful molecules for studies of HIF signaling, and more broadly, valuable examples of different ways to control protein complex formation with small molecules.

ii. Results

Identifying direct inhibitors of ARNT/TACC3 interactions by NMR screening - We have solved a new 1.6 Å resolution X-ray diffraction structure of ARNT PAS-B (Table 3-1; Fig. 3-2a), revealing two cavities near the regulatory cofactor binding sites in other PAS domains (*e.g.* flavins in photosensors; heme in oxygen sensors (173,174)). The larger 65 Å^3

Table 3-1. X-ray Crystallography Data Processing and Refinement Statistics.

Data collection	
Space group	C2
Cell dimensions	
a, b, c (Å)	93.3, 61.7, 55.5
β (°)	124.6
Resolution (Å)	20.7 to 1.6 (1.63 to 1.60)
R_{sym} or R_{merge}	4.7 (43.9)
I / σ_I	38.8 (2.2)
Observed reflections	155,415
Unique reflections	34402
Completeness % (highest shell)	99.9 (100)
Refinement	
Resolution (Å)	22.5 to 1.6 (1.63 to 1.60)
R_{work} / R_{free}	20.15 / 22.65
No. atoms	
Protein	1832
Water	206
Avg. B-factor	
Protein	20.3
Water	35.4
R.M.S. deviations	
Bond lengths (Å)	0.014
Bond angles (°)	1.45
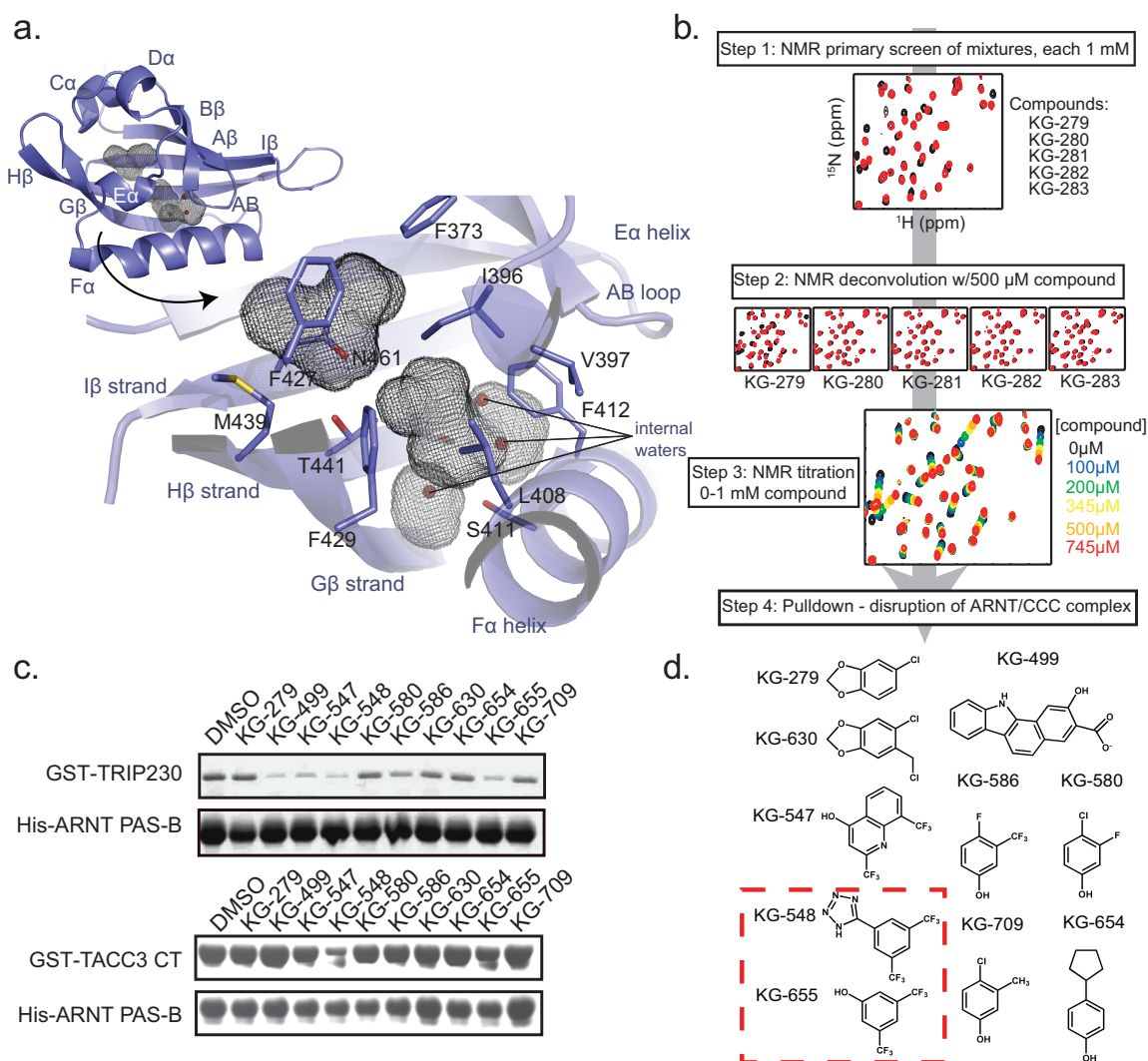
Figure 3-2. **Screening ARNT PAS-B for small molecule effectors of the ARNT/TACC3 interaction.** a) Diagram of the ARNT PAS-B crystal structure, shown inset with secondary structure designations and internal cavities (grey mesh). The larger cavity contains three waters; the smaller cavity is primarily hydrophobic. b) Schematic of the NMR-based screen for compounds that bind ARNT PAS-B. Initial ^{15}N/^{1}H HSQC spectra were acquired with 250 μM ^{15}N-labeled ARNT PAS-B with a mixture of five compounds (1 mM each); mixtures producing large chemical shift perturbations (compared to DMSO) were deconvoluted as shown. c) Lead compounds (500 μM each) from NMR-based screen were tested for their ability to disrupt complexes of ARNT PAS-B with CCC fragments of TRIP230 (1583-1716) and TACC3 (561-631 = TACC3-CT). d) Summary of ARNT PAS-B binding and ARNT/CCC disruption of tested compounds. All ten compounds generated chemical shift perturbations when titrated into ARNT PAS-B; KG-548 and KG-655 (red box) also disrupted TRIP230 and TACC3 binding to ARNT PAS-B *in vitro*.

cavity is flanked by the Eα / Fα helices, Gβ / Hβ / Iβ strands and the AB loop, and includes several polar residues (*e.g.* S411, T441 and S443) that facilitate the binding of three waters at typical cofactor sites. The second cavity is slightly smaller (40 A^3) and comparable to the chromophore binding site in photoactive yellow protein (175). While these cavities are relatively small, their locations (potential to merge into a single larger cavity) raised the possibility for them to bind artificial compounds and modulate CCC binding, similarly to other PAS domains.

To identify such compounds, we used solution NMR to search over 760 small molecules (56,169,170) for ARNT PAS-B binding. This library consists of low molecular weight fragments (average MW: 203 ± 73 Da, Appendix table 1-1) containing "privileged" moieties enriched in protein-binding compounds (176). Further, this collection has previously provided us several compounds (or analogs) which bind to other PAS domains (56,166,169) and disruptors of protein/protein interactions in other transcriptional regulators (56,170). We quickly evaluated these compounds for ARNT PAS-B binding using protein-detected NMR (Fig. 3-2b). In the primary screen, ^{15}N/^{1}H HSQC spectra were acquired on ^{15}N-labeled ARNT PAS-B samples mixed with five candidate compounds (250 μM protein, 1 mM each compound) or DMSO. Compound mixtures that produced changes in peak locations or intensities indicated that one or more compounds bound the protein target; sixteen such mixtures were subsequently deconvoluted by acquiring spectra on ARNT PAS-B with individual compounds. Eighteen hits from these steps were tested in titrations at concentrations up to 1 mM to establish binding potency and location; ten of these (Fig. 3-2c) were soluble throughout this concentration range and further studied.
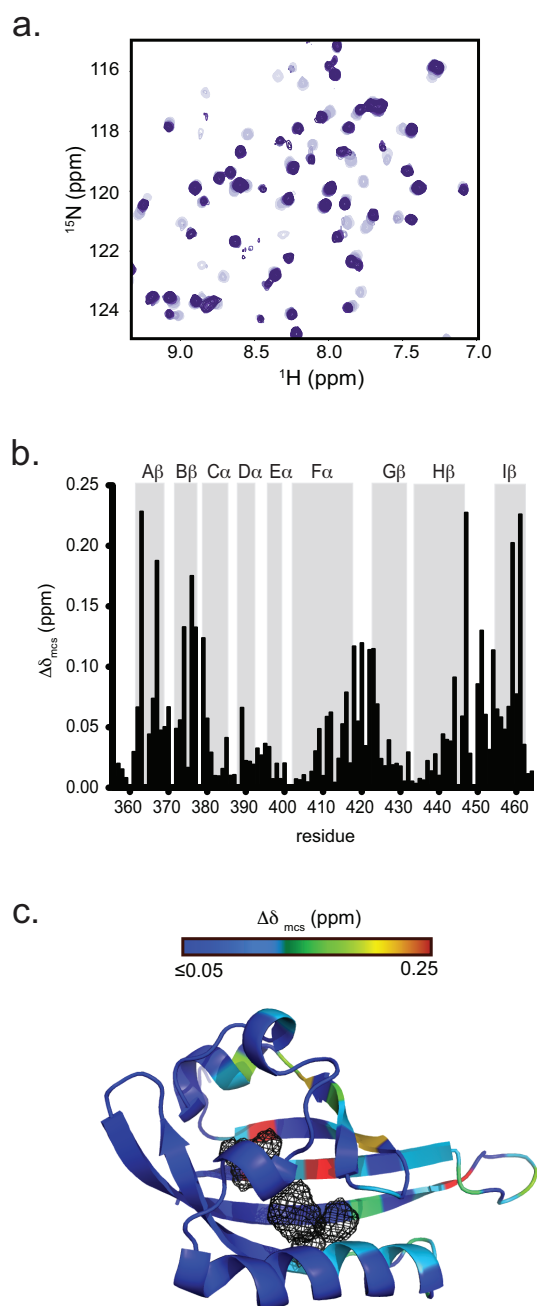
Figure 3-3. KG-548 appears to bind within the ARNT PAS-B cavities. a) ^{15}N/^{1}H HSQC spectra of a KG-548 titration (0-1 mM from light to dark crosspeaks) into 320 μM ^{15}N ARNT PAS-B. Slow exchange behavior was observed, indicated by the disappearance of apo- crosspeaks and concomitant appearance of new peaks. b) Minimum chemical shift analysis(177) of KG-548 titration into ARNT PAS-B, mapped onto the sequence and secondary structure. c) Chemical shift mapping suggests that KG-548 binds the ARNT PAS-B cavities, as shown by a heat map of KG-548-induced chemical shift changes on the ARNT PAS-B crystal structure ($\Delta\delta_{mcs}$ colored from low (blue) to high (red)) with the largest changes near the internal cavities (mesh).

To examine whether these ARNT-binding compounds affected the stability of ARNT PAS-B/CCC complexes, we initially checked their disruption of ARNT-mediated pulldowns of fragments of the TRIP230 and TACC3 coactivators (Fig. 3-2c). The fragments of both CCC proteins contained the coiled-coil that binds ARNT PAS-B (122,134), facilitating robust interaction without compounds present. However, several chemicals markedly interfered with ARNT binding to one or both coactivators. Among these, KG-548 exhibited the greatest reduction of ARNT/CCC complex formation for both coactivators, with a smaller effect seen for the structurally related fragment KG-655 (Fig. 3-2d).

KG-548 binds to the cavity of ARNT PAS-B - Next, we further characterized compound-mediated ARNT/CCC disruption with solution NMR to identify ligand-binding sites within ARNT PAS-B. Using KG-548 as the most effective disruptor of this complex, we observed slow exchange behavior in compound titrations monitored by ARNT PAS-B $^{15}N/^{1}H$ HSQC spectra (Fig. 3-3a), suggesting micromolar (or tighter) dissociation constants. We used our prior chemical shift assignments (168) to map ligand-induced changes onto the ARNT PAS-B structure using minimum chemical shift difference analyses, assuming correlations between the nearest pairs of peaks in apo- and KG-548 saturated spectra (Fig. 3-3b). Residues that were most strongly affected by KG-548 addition were located close to the internal ARNT PAS-B cavities (Fig. 3-3b,c) and analogous to where other PAS domains often bind ligands (3). Notably, these ligand-mediated structural and functional effects were sensitive to minor changes in compound structure. Two KG-548 variants with small changes - CF_3 to Cl substitutions on the phenyl ring, and addition of methyl or ethyl moieties on the
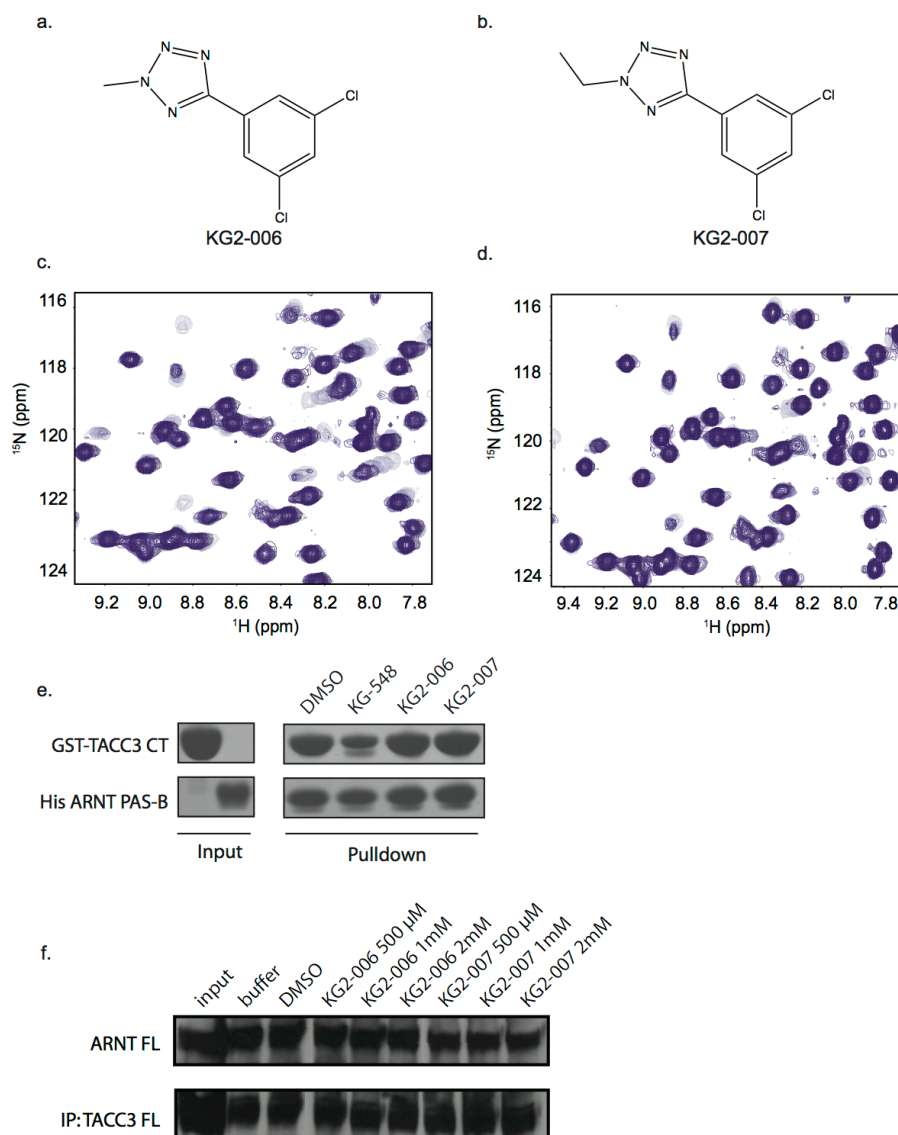
Figure 3-4. KG-548 analogs do not disrupt ARNT PAS-B/TACC3-CT complex. a) Structure of KG2-006 (Maybridge). b) Structure of KG2-007 (Maybridge). c) Titration of KG2-006 into ^{15}N ARNT PAS-B (increasing ligand concentrations from light to dark purple: 0, 62.5, 125, 250, 500, 750, 1000 µM; protein concentration: 200 µM). Fast exchange behavior is observed in some of the peaks, indicating relatively weak binding affinity compared with KG-548. d) Titration of KG2-007 into ^{15}N ARNT PAS-B (same concentrations as panel c). Spectrum was not perturbed by the addition of KG2-007, indicating an even weaker binding to ARNT PAS-B. e) Pulldown experiment with KG-548, KG2-006 and KG2-007 (500 µM of all compounds; 5 µM ARNT, 10 µM TACC3). Neither of the KG2 compounds is able to break up the ARNT PAS-B/TACC3-CT complex. f) Co-immunoprecipitation experiment of KG2-006 and KG2-007 on full length ARNT/TACC3 in whole cell lysate. At concentrations up to 2 mM, neither of the KG2 compounds is able to break up the ARNT/TACC3 complex as KG-548 does, consistent with lower binding affinities of these derivatives.

71

tetrazole-bound ARNT more weakly and were unable to disrupt ARNT/TACC3 complexes (Fig. 3-4a-e). Coupled with the limited number of leads from the library screen, these data demonstrate specificity of the ARNT/ligand interaction.

KG-548 binding shows selectivity among PAS domain target - Having demonstrated specificity from the ligand perspective, we next explored the potential for other bHLH-PAS PAS-B domains to bind KG-548. We started by examining ARNT2, a closely-related ARNT homolog that is chiefly expressed in neuronal and kidney tissue (178,179) (Fig. 3-5). Using solution NMR methods, we verified that the ARNT and ARNT2 PAS-B domains adopt comparable structures, as expected from the 80% sequence identity between them. Similarities in chemical shifts and TALOS+-derived secondary structures (180) (Fig. 3-6a,b) strongly indicate ARNT2 adopts the same overall fold as seen in our crystal ((Fig. 2 and ref. 56)) and solution (168) structures of ARNT PAS-B, giving us confidence in a ARNT2 PAS-B homology model to suggest the placement of residues involved in ARNT/CCC interactions (134) (Fig. 3-6c).

To functionally test these structural similarities, we examined the ability of ARNT2 PAS-B to bind coactivator fragments and KG-548. Pulldown experiments demonstrated that ARNT2 PAS-B directly interacts with GST-tagged TACC3-CT (=C-terminal 70 aa of TACC3, including residues 561-631) *in vitro* (Fig. 3-6d), as seen for ARNT (Fig. 3-2c). Mutations to ARNT2 PAS-B residues E372 and K391 weakened this interaction (Fig. 3-6d), chosen for their similarity to ARNT E398 and K417 (Fig. 3-6c) (134). ^{15}N/^{1}H HSQC spectra of both ARNT2 E372A and K391A PAS-B mutants were similar to wildtype protein (Fig. 3-
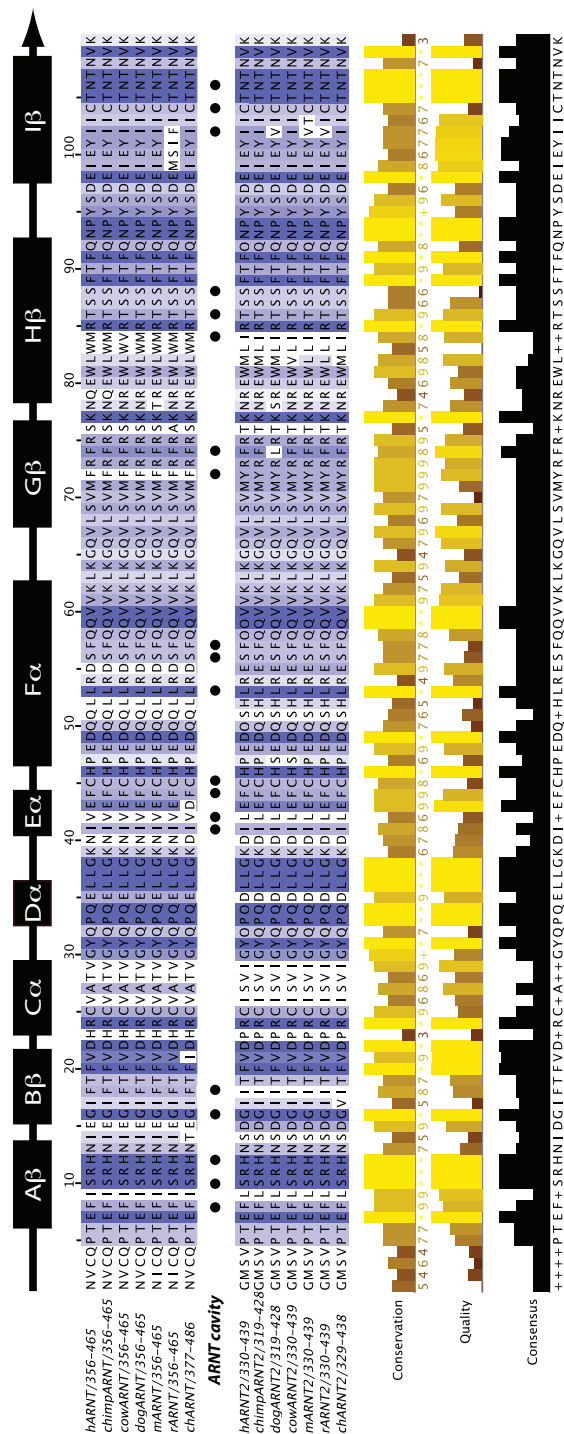
Figure 3-5. Sequence alignment of ARNT homologs on the PAS-B domain. The PAS-B domain of ARNT shows a high degree of sequence similarity to the corresponding ARNT2 PAS-B domain. Notably, residues adjacent to the small cavities inside of ARNT PAS-B (black dots) are also chiefly conserved in ARNT2 PAS-B.

73
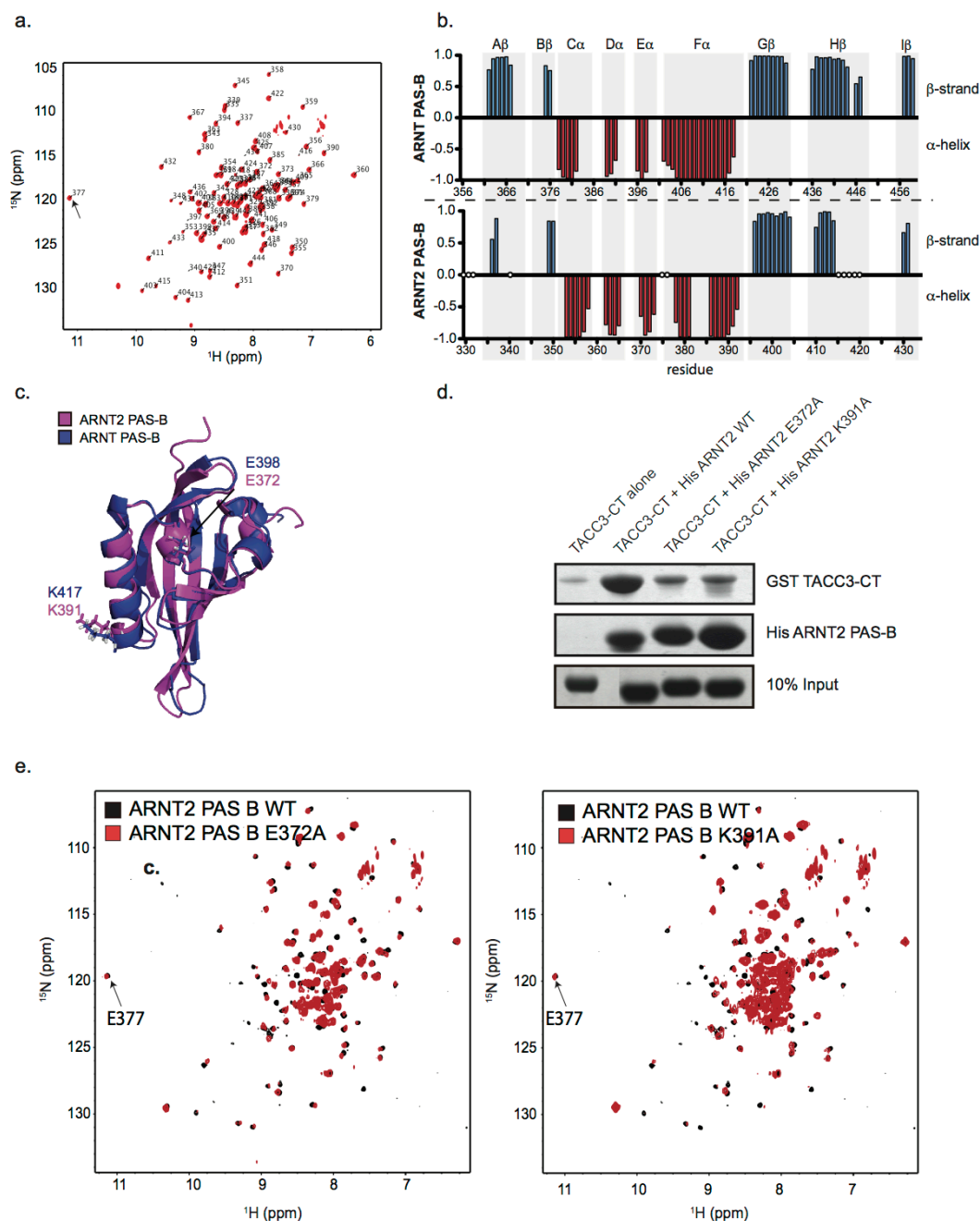
Figure 3-6. The PAS-B domains from ARNT and ARNT2 have similar structures and TACC3 binding modes. a) ^{15}N/^{1}H HSQC spectrum of ARNT2 PAS-B, with peaks corresponding to backbone amides labeled with the residue numbers identified in the backbone chemical shift assignment (35°C, pH 7.3). Peaks comparable to the far downfield peak (ca. 11.2 ppm ^{1}H, 119.5 ppm ^{15}N; assigned to residue E377) have been observed for other PAS domains; the distinctive shifts of this peak arise from the tertiary structure near the N-terminus of the Fα helix and provide a qualitative verification of a PAS domain tertiary structure. b) TALOS+ secondary structure analysis based on backbone chemical shifts (180). Residues assigned to α-helix (red, negative) and β-sheet (blue,

positive) are mapped along the amino acid sequence of ARNT PAS-B (upper) and ARNT2 PAS-B (lower), showing very similar secondary structures in both proteins. Residues in ARNT2 PAS-B indicated with white circles are unassigned, which we attribute to problems with line broadening caused by transient dimerization, as we have observed to varying degrees with several PAS domains (168,181).   c) A SWISS-MODEL (182) homology model of ARNT2 PAS-B (purple) aligns well with ARNT PAS-B crystal structure (blue).   Residues demonstrated to be important for TACC3 binding are also conserved in the tertiary structure (E398, K417 for ARNT PAS-B (134); E372, K391 for ARNT2 PAS-B).   d) Ni-NTA pulldown experiments demonstrate that ARNT2 PAS-B interacts with TACC3-CT, and that the affinity for this interaction is lessened by point mutations to residues analogous to those involved in ARNT/TACC3 binding (Fig. 3-6c).   e) ^{15}N/^{1}H HSQC spectra of ARNT2 PAS-B E372A and K391A (red), each overlaid with spectra of wildtype ARNT2 PAS-B (black). The minimal effects of the E372A and K391A mutations on the patterns of peaks in the ^{15}N/^{1}H HSQC spectra, including the location of the E377 peak reflecting PAS domain tertiary structure (Fig. 3-6a), strongly suggest that these changes have limited effects on the structure of this domain.
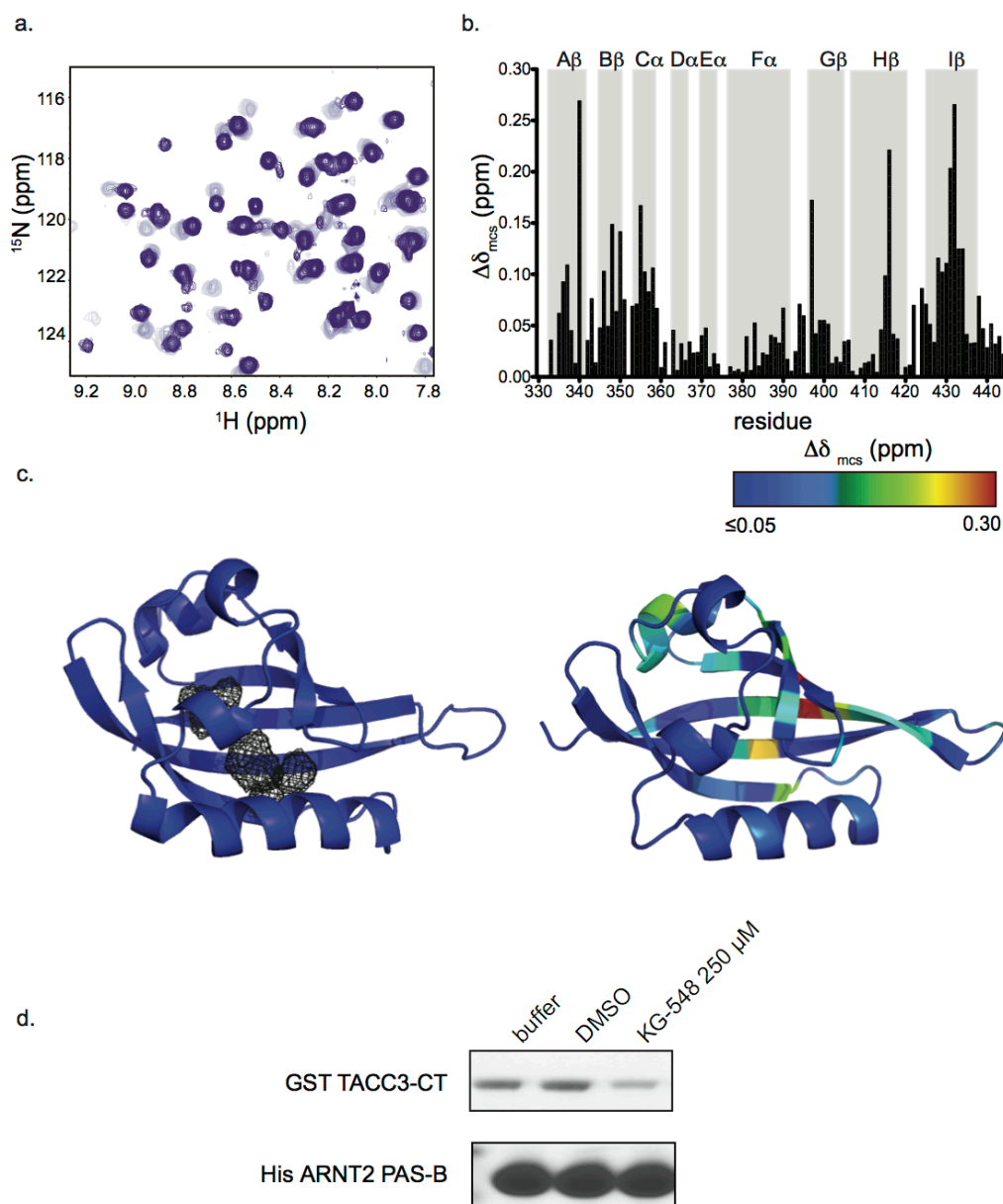
Figure 3-7. KG-548 binds similarly to ARNT2 PAS-B and ARNT PAS-B, and disrupts ARNT2/TACC3 interactions. a) ^{15}N/^{1}H HSQC spectra of 100 µM ^{15}N-labeled ARNT2 PAS-B titrated with increasing concentrations of KG-548 (from none (light purple) to 100 µM (dark purple)). We observed slow exchange behavior, analogous to the ARNT/KG-548 titration (Fig. 3-3a). b) Minimum chemical shift perturbation analysis (177) of KG-548 effects on ARNT2 PAS-B, showing similar patterns of shifts as observed with ARNT PAS-B (Fig. 3-3b). c) Side-by-side comparison of ARNT PAS-B structure (including internal cavities) and chemical shift perturbation data mapped onto a homology model of ARNT2 PAS-B (scale as indicated). Most of the affected residues are close to the location of the cavity inside of ARNT PAS-B. d) Binding of KG-548 can disrupt ARNT2 PAS-B/TACC3-CT interaction by pulldown assay.
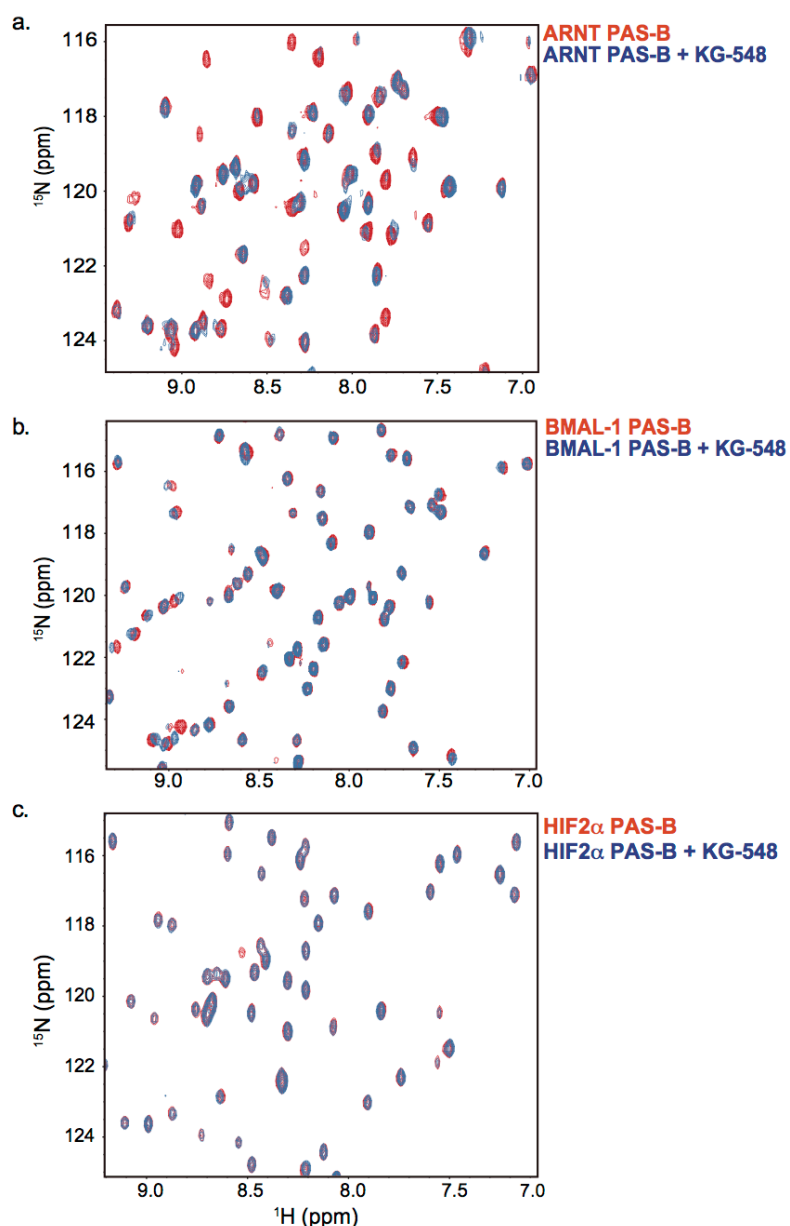
Figure 3-8. KG-548 binds to different bHLH/PAS PAS-B domains with differential affinities.
a) ^{15}N/^{1}H HSQC spectra of 100 μM ^{15}N-labeled ARNT PAS-B in the absence and presence of 100 μM KG-548. A large number of residues are affected by compound addition, with correspondingly large chemical shift changes, consistent with KG-548 binding to ARNT.   b) Superposition of ^{15}N/^{1}H HSQC spectra of 100 μM ^{15}N-labeled BMAL1 PAS domain titrated with 100 μM KG-548. Very limited chemical shift changes were observed, suggesting that there is less interaction between protein and ligand at these concentrations.   c) Superposition of ^{15}N/^{1}H HSQC spectra of 100 μM ^{15}N-labeled HIF2α PAS-B titrated with 100 μM KG-548. No chemical shift changes were observed, suggesting that there is effectively no interaction between protein and ligand at these concentrations.

6e), suggesting that both point mutations had minimal structural effects. Next, we asked if KG-548 could also bind to ARNT2 PAS-B and compete away TACC3. Ligand titrations monitored with $^{15}N/^{1}H$ HSQC spectra showed the same slow exchange (Fig. 3-7a) observed with ARNT PAS-B, with the largest effects (Fig. 3-7b,c) clustered in the same internal sites (Fig. 3-5). Finally, as expected from the similar CCC and ligand binding modes, KG-548 also disrupts ARNT2 PAS-B/TACC3 interactions in pulldown assays (Fig. 3-7d), underscoring the high degree of similarity between ARNT and ARNT2.

In contrast, similar titrations of KG-548 into two more widely diverged PAS-B domains from the BMAL-1 and HIF2α bHLH-PAS proteins (38% and 34% identity to ARNT PAS-B) showed specificity in the protein/ligand interaction. We observed very few chemical shift changes caused by KG-548 addition to BMAL-1, and virtually none with HIF2α (Fig. 3-8) despite the large cavities within both PAS domains (56,116,166). We interpret these data to indicate that these two PAS-B domains have much lower affinities for KG-548 than ARNT PAS-B, showing little to no interaction at the tested concentrations and thus establishing that observed ligand-binding specificity is not solely based on simple accessibility. More broadly, we have not observed KG-548 binding to other PAS and non- PAS targets (169,170), consistent with the specificity that can be observed in small fragments from other libraries (183).

KG-548 breaks up the ARNT/TACC3 complex in vitro and in cell lysate - To further characterize KG-548 induced disruption of ARNT/TACC3, we examined the dose dependence of this effect in two ways. Pulldown assays using His-ARNT PAS-B and GST-
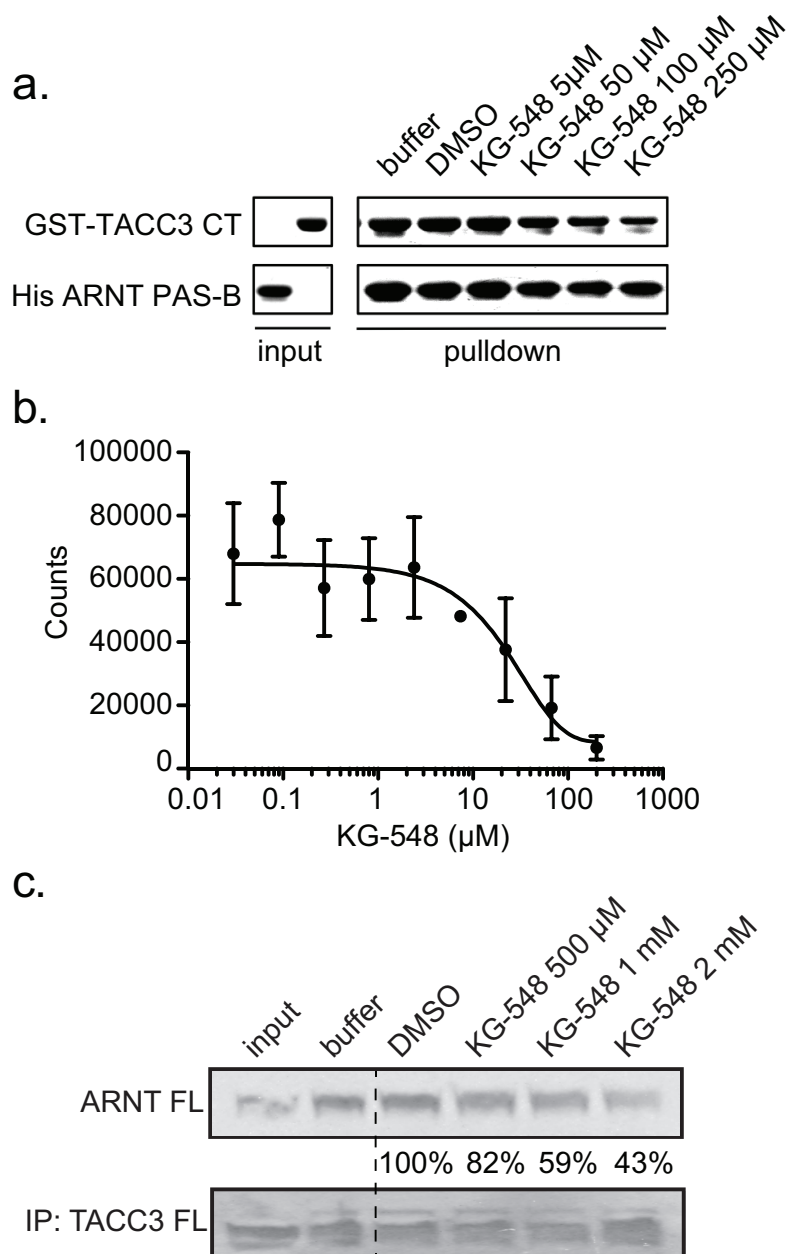
Figure 3-9. KG-548 disrupts *in vitro* ARNT/TACC3 interactions. a) Titration of KG-548 into an *in vitro* pulldown assay of minimal ARNT PAS-B and TACC3-CT interacting fragments shows a dose-dependent reduction in ARNT/TACC3 complex formation. b) Quantification of KG-548 potency for disrupting the ARNT PAS-B/TACC3-CT interaction as provided by AlphaScreen, showing an apparent IC_{50} of 25 µM. c) Co-immunoprecipitation assays of full length ARNT and TACC3 proteins in HEK293T cell lysates show that KG-548 weakens the ARNT/TACC3 interaction as demonstrated by the dose-dependent decrease in the intensity of the ARNT protein band (quantitated with % remaining compared to DMSO control) associated with immunoprecipitated TACC3 protein.

79
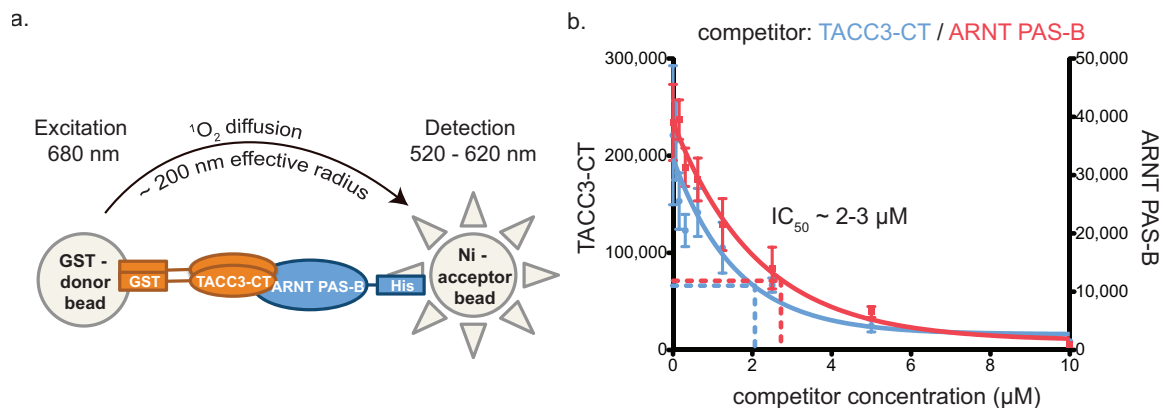
Figure 3-10. Validation of ARNT PAS-B/TACC3-CT AlphaScreen protein/protein interaction assay. a) Schematic of the AlphaScreen protein/protein interaction assay used to estimate the affinity of the ARNT/TACC3-CT complex. GST and His$_6$-tagged proteins will recruit donor and acceptor AlphaScreen beads into close enough proximity that a luminescence signal will be detected. b) Titration of either untagged ARNT PAS-B (red, right y-axis) or TACC3 (blue, left y-axis) into a preformed ARNT/TACC3-CT complex competes with the corresponding proteins in complex, decreasing the AlphaScreen luminescence signal. Dose-dependence curves generated from these titrations suggest that each protein has a 2-3 µM IC$_{50}$ for complex disruption.

TACC3-CT, tagged versions of the minimal interacting components of both proteins (134), showed that levels of GST-TACC3-CT pulled down by His-ARNT PAS-B decreased in a KG-548 dose-dependent manner (Fig. 3-9a) at concentrations that perturbed ARNT PAS-B NMR spectra. We quantified the potency of KG-548 using AlphaScreen, a luminescence proximity assay of complex formation between the two tagged fragments (Fig. 3-10a). The assay was validated by showing that untagged ARNT PAS-B or TACC3 competed against complex formation between tagged proteins with IC_{50}~2-3 µM (Fig. 3-10b). Similar AlphaScreen assays with KG-548 showed a 25 µM IC_{50} inhibition (Fig. 3-9b), consistent with the apparent ARNT PAS-B/KG-548 affinity seen in NMR-based titrations.

To examine how well these results from isolated domains translate to full length proteins, we surveyed the ability of KG-548 to disrupt the complex between full length human ARNT and mouse TACC3 proteins. While cell-based qPCR and HRE reporter assays of this effect were hampered by issues with cell toxicity at mid-micromolar KG-548 concentrations, an alternative was provided by immunoprecipitation experiments in lysates of HEK 293T cells transfected with expression vectors for both proteins. Within this system, the ARNT/TACC3 interaction can clearly be observed by co-IP without ligand present; addition of increasing concentrations of KG-548 leads to a progressive decrease in the amount of TACC3-associated ARNT protein (Fig. 3-9c). While the apparent potency of KG-548 is lower in this more complex setting, our data clearly demonstrate that this ARNT-binding compound can inhibit ARNT/TACC3 complex formation in truncated or full length proteins. Notably, two negative control compounds KG2-006 and KG2-007 did not show such
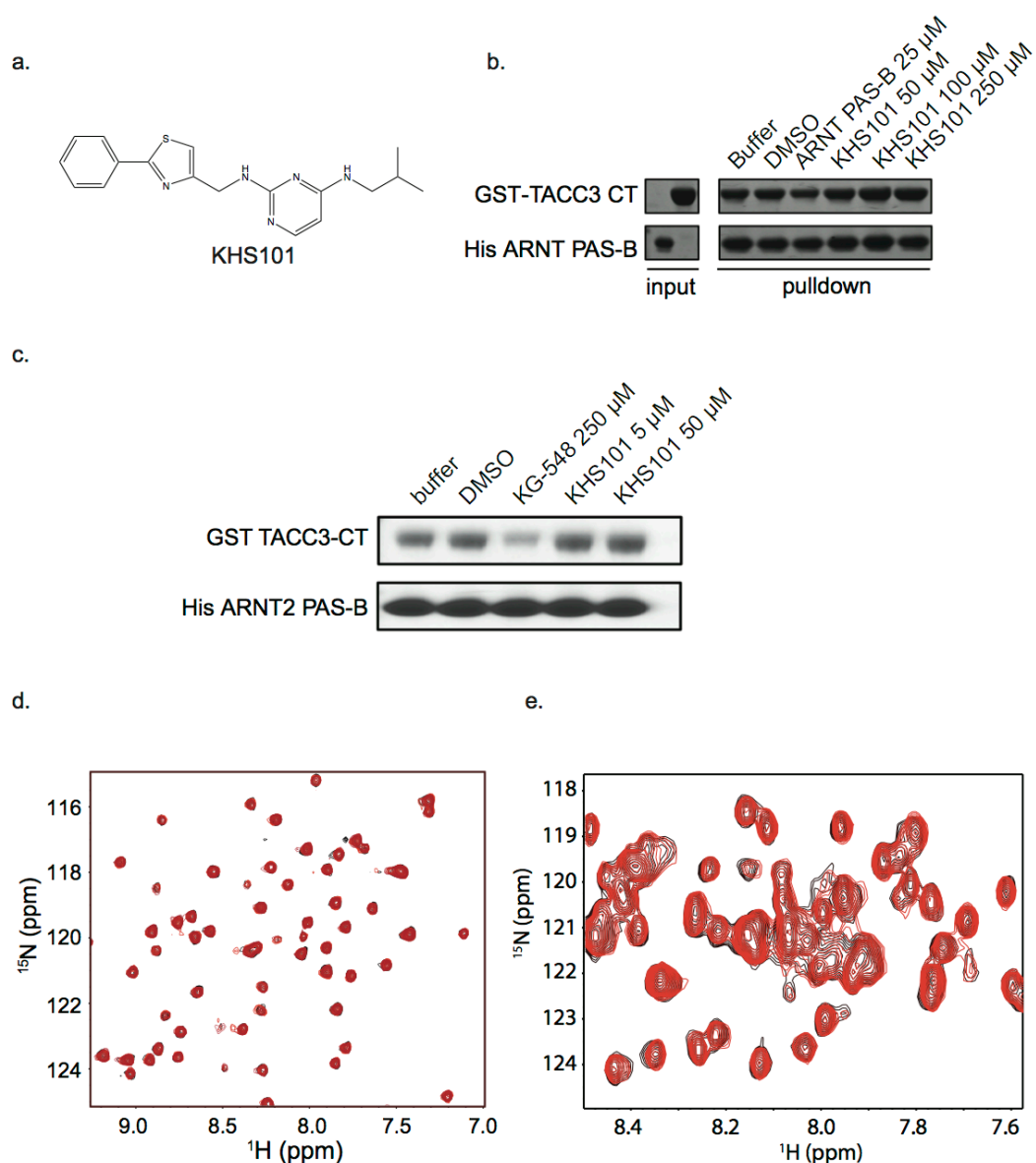
Figure 3-11. KHS101 does not directly regulate the formation of a minimal ARNT/TACC3 complex. a) Structure of KHS101 (172). b) At concentrations up to 250 µM, KHS101 fails to disrupt the complex of minimal His-ARNT PAS-B and GST-TACC3-CT fragments by *in vitro* pulldown assays. c) KHS101 is unable to interfere with *in vitro* interactions between ARNT2 PAS-B/TACC3-CT. In contrast to KG-548, addition of KHS101 (up to concentrations of 50 µM) into an *in vitro* Ni-NTA pulldown assay shows no significant effect on the stability of a complex between ARNT2 PAS-B and TACC3-CT. d) ^{15}N/^{1}H HSQC spectra of 100 µM ^{15}N ARNT PAS-B in the absence (black) or presence (red) of 100 µM KHS101 show no ligand-induced effects, indicating no direct binding of KHS101 to ARNT PAS-B. e) ^{15}N/^{1}H HSQC spectra of 100 µM GCN4-TACC3-CT (610-631) dimer titrated with 100 µM KHS101 (apo: black, holo: red)). No significant peak perturbation is observed, indicating no direct binding of KHS101 to the TACC3-CT.

inhibition at comparable concentrations, consistent with the isolated domain results (Fig. 3-4f).

Discovery of KHS101 as a TACC3 inhibitor - Complementing our targeted *in vitro* screening, we examined a presumed disruptor of ARNT/CCC interactions provided by the HTS-derived compound KHS101 (Fig. 3-11a) that accelerates NPC differentiation in the adult rat (172). Crosslinking data initially associated TACC3 with KHS101 by showing that a benzophenone derivative bound TACC3 at an uncharacterized location. This linkage between KHS101 and TACC3 was underscored by the observation of similar cellular effects with either KHS101 treatment or anti-TACC3 siRNAs in several settings (172,184), suggesting that this compound is a general TACC3 inhibitor. However, this study did not characterize the mechanism of KHS101 action on TACC3, leading us to ask if KHS101: 1) directly interferes with TACC3 binding to its partners ARNT or ARNT2 PAS-B *in vitro*; 2) alters TACC3 protein levels in cells; or 3) affects transcription from ARNT/CCC-reliant promoters, such as HIF-driven genes (134).

KHS101 is not a direct regulator of the ARNT/TACC3 complex - To address how KHS101 regulates ARNT/TACC3, we performed *in vitro* pulldown experiments in the presence of KHS101 using the minimal constructs of both proteins. We did not observe any substantial KHS101-dependent effect on TACC3 pulldowns with either ARNT or ARNT2 (Fig. 3-11b,c), suggesting that KHS101, in contrast to KG-548, works indirectly to disrupt ARNT/TACC3 function. To test this hypothesis, we used solution NMR spectroscopy to

examine KHS101 binding to ARNT PAS-B and TACC3. ^{15}N/^{1}H HSQC spectra of ARNT PAS-B titrated with KHS101 showed only minor changes (Fig. 3-11d; compare to Fig. 3-3a for KG-548), indicating no binding. Analogous spectra using a ^{15}N-labeled minimal version of TACC3 with both the ARNT-interacting C-terminal 21 residues of TACC3 and a stabilizing GCN4 coiled-coil (GCN4-TACC3-CT) also showed no differences between DMSO and KHS101-treated samples (Fig. 3-11e) at concentrations that trigger biological responses ((vide infra and ref. 172)). These negative data strongly suggested that KHS101 perturbs the ARNT/TACC3 complex in a different, likely indirect, mechanism than KG-548.

KHS101 reduces intracellular TACC3 stability - Coupling prior data implicating KHS101 interference with ARNT/TACC3 (172) with our observation that KHS101 does not directly block the minimal ARNT PAS-B/TACC3-CT interaction, we hypothesized that KHS101 might function indirectly by modulating TACC3 stability. To test this, we examined TACC3 protein stability in HEK293T cells using pulse chase experiments conducted with the translation inhibitor cycloheximide (CHX) and either 5 μM KHS101 or DMSO control. Cells were harvested at different times post-CHX treatment and TACC3 protein levels were monitored by immunoblot (Fig. 3-12a). KHS101 treatment decreased the stability of TACC3 compared to DMSO controls (Fig. 3-12a,b), with substantial differences after 6 hr of treatment. We observe a statistically-significant difference at 6 hr and a net three-fold drop in TACC3 protein levels at 8 hr (relative levels of 9.3 [KHS101] and 29.9 [DMSO], Fig. 3-12b), whereas ARNT levels were barely affected (Fig. 3-13a,b). In contrast, cells treated with KG-548 or an inactive KHS101 analog, KHS91 (172), showed no change in TACC3 level
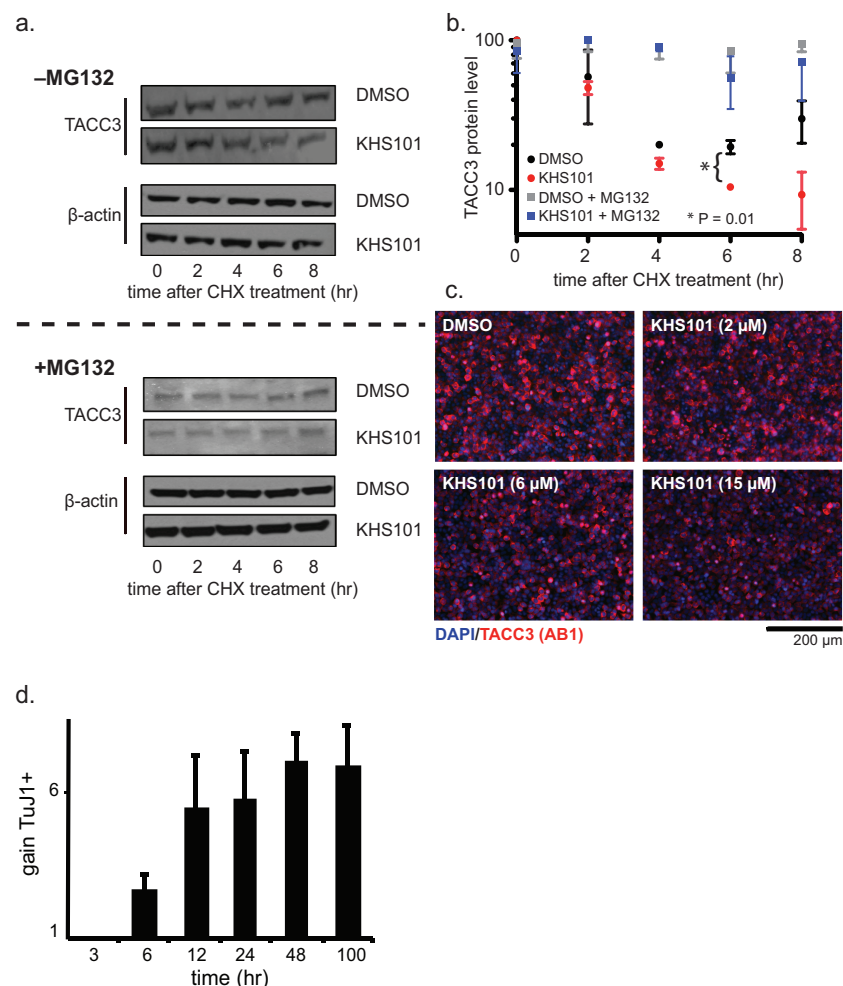
Figure 3-12. KHS101 decreases TACC3 levels in cells and regulates HIF gene expression. a) KHS101 facilitates TACC3 degradation in a proteasome-dependent manner. HEK293T cells were treated with 5 μM KHS101, 100 μg/ml cycloheximide (CHX) (upper panel); additional cells were similarly treated with KHS101 and CHX plus 20 μM MG132 (lower panel). Cells were harvested 0-8 hr post-treatment and prepared for TACC3 immunoblot analyses. b) Quantification of TACC3 protein levels from data shown in panel a. Without MG132, TACC3 protein levels decreased, with greater drops observed in KHS101-treated cells (compared to DMSO) after 6 hr incubation. A statistically-significant difference was observed 6 hr post-treatment ($p < 0.01$ by Student's t-test); the 8 hr timepoint also shows a substantial decrease in TACC3 levels, but this is not statistically significant due to large variations in data values. In the presence of MG132, we observed little decrease in TACC3 levels with no KHS101-dependent effects, implicating a proteasomal-dependent degradation pathway. c) Steady state treatment with KHS101 reduces TACC3 levels. HEK293T cells were treated with KHS101 (0-15 μM) and immunostained with TACC3 AB1 after 14 hr. TACC3 intensity was negatively affected by KHS101. d) KHS101 induces cell differentiation with a minimum of 6 hr exposure. Adult rat NPCs were exposed to 5 μM KHS101 (or DMSO) for the indicated times, after which media were replaced with compound-free versions and incubated for a total of 100 hr. Neuronal differentiation was assessed by expression of the TuJ1 marker, showing upregulation after times consistent with TACC3 levels falling in CHX-treated cells (panels a, b).
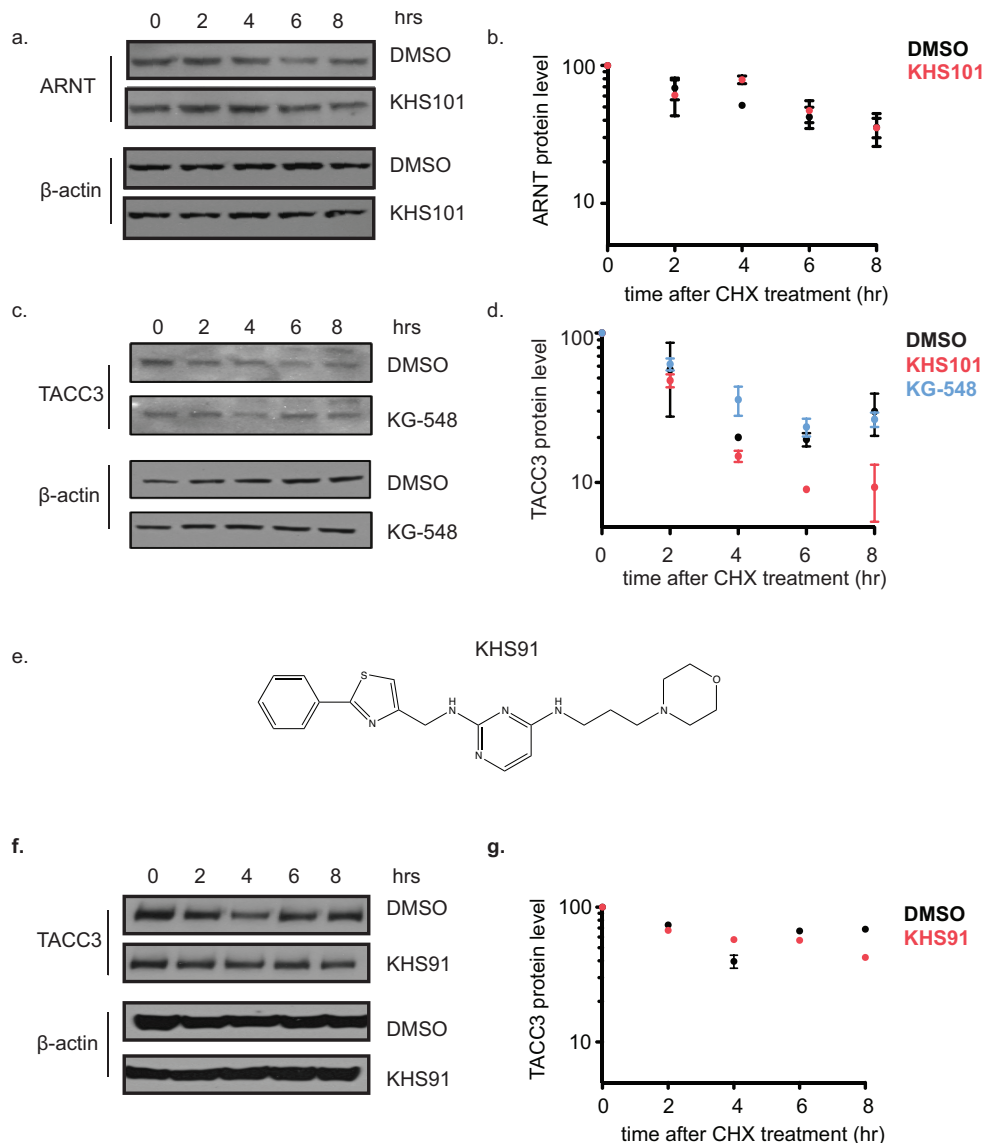
Figure 3-13. Specificity of KHS101 effects on TACC3 levels. a) Blot of ARNT and β-actin protein levels in HEK293T cells after 5 μM KHS101 treatment as described before. ARNT level was not changed in the presence of KHS101, indicating its effect in protein turnover is specific to TACC3. b) Quantification of data presented in panel a, showing no effect of KHS101 on the ARNT turnover rate. c) Blot of TACC3 and β-actin protein levels in HEK293T cells after treatment with either DMSO or 100 μM KG-548 in the presence of 100 μg/ml cycloheximide. d) Quantification of data presented in panel c (DMSO, KG-548 treatments) or Fig. 3-12a (KHS101 treatment) demonstrates that turnover effects are specific to KHS101. e) Structure of KHS91, an inactive analog of KHS101 (172). f) Blot of TACC3 and β-actin protein levels in HEK293T cells after treatment with either DMSO or 5 μM KHS91 in the presence of 100 μg/ml cycloheximide. g) Quantification of data presented in panel e, showing no effect of KHS91 on the TACC3 turnover rate.
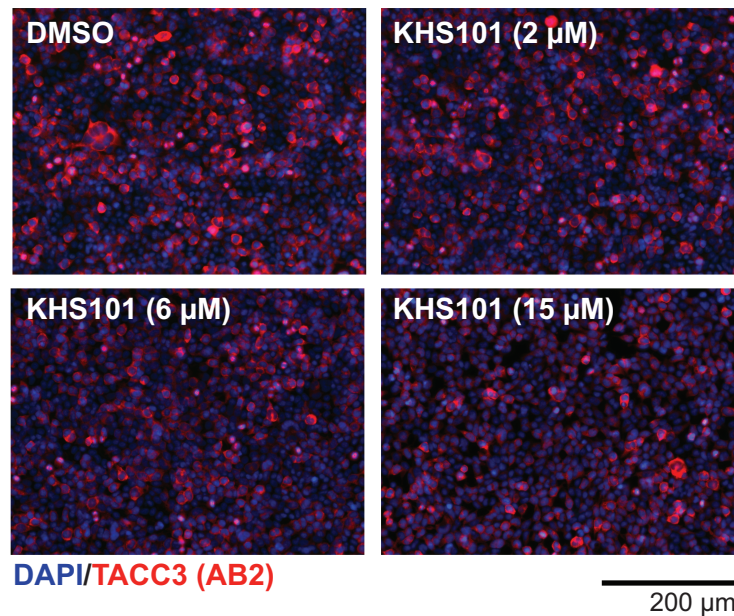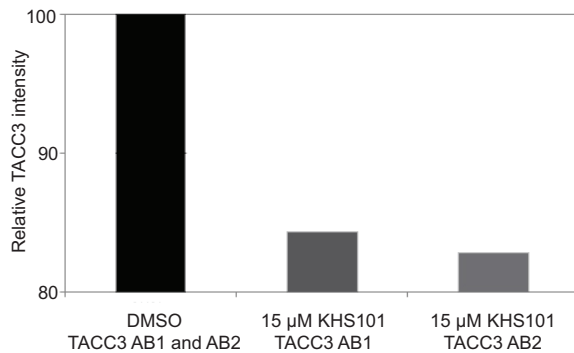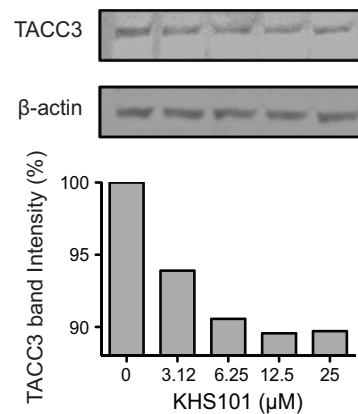
Figure 3-14. TACC3 level is reduced by KHS101 in a dose-dependent manner. a) HEK293T cells were treated with KHS101 (0, 2, 6 and15 μM) and immunostained with TACC3 AB2 after 14 hr. TACC3 intensity was negatively affected by KHS101, consistent with results shown in Fig. 3-12c using the independent TACC3 AB1. b) Quantification of TACC3 intensity as determined by the anti-TACC3 immunostaining shown in Fig. 3-12c (AB1) and Fig. 3-14a (AB2). Overall TACC3 staining intensity levels were reduced to approximately 90% after KHS101 treatment measured by two different TACC3 antibodies. c) HEK293T cells were treated with KHS101 (0-25 μM) and harvested after 16 hr. Quantification of TACC3 levels by immunoblots in the lysates of these cells show a dose-dependent reduction in TACC3 levels with an approximate IC_{50} of 3 μM and final reduction to approximately the same 90% level as seen in cells (Fig. 3-14b).

compared with DMSO, indicating specificity in the ligand-induced degradation (Fig. 3-13c-g). Parallel experiments including the proteasome inhibitor MG132 muted this loss of TACC3 protein, without any marked difference between KHS101 and control groups (Fig. 3-12a,b). Finally, we verified that KHS101 affects steady state TACC3 levels in the absence of CHX, showing a dose-dependent and saturable drop in TACC3 protein levels with increasing concentrations of KHS101 (Fig. 3-12c, 3-14a-c). While this drop was smaller than we observed with CHX treatment, the consistent observation of a 10% drop in TACC3 levels in independent experiments with different sample types (using either intact cells or cell lysates) and different primary antibodies (Fig. 3-14) gives us confidence in this trend. Taken together, our results suggest that KHS101 treatment leads to increased proteasome-mediated degradation of TACC3 protein in cells, conceivably mediated by interactions outside of the ARNT-binding motif.

KHS101 treatment affects two independent TACC3-containing pathways in cells - To evaluate the functional implications of the KHS101-triggered drop in TACC3 levels, we looked for correlations between KHS101 effects on TACC3 levels and the activities of two TACC3-dependent pathways. We initially examined the kinetics of compound-induced NPC differentiation, which involves both TACC3 and ARNT2 (172). Here we used KHS101 washout experiments, exposing rat NPCs to KHS101 for various times before switching to compound-free media for the remainder of the 100 hr incubation. Afterwards, differentiation was assessed using immunofluorescence-based detection of the pan-neuronal TuJ1 marker. We saw a KHS101-dependent increase in TuJ1 expression only after 6 hr of treatment,
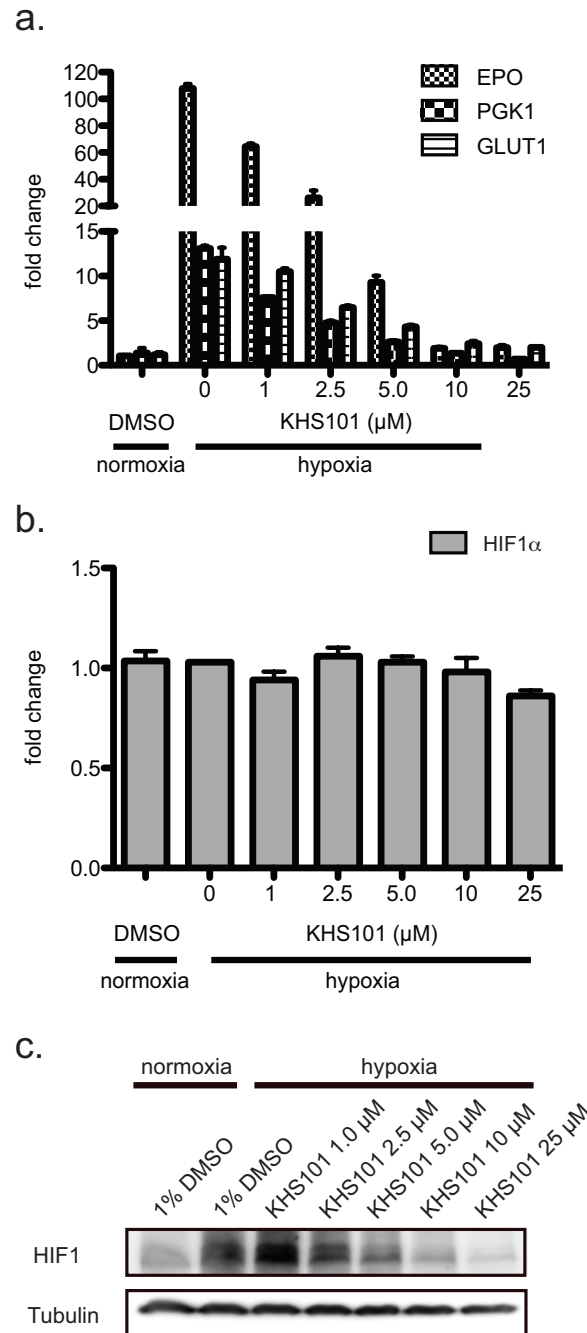
Figure 3-15. KHS101 inhibits HIF target gene expression and decreases HIF1α protein levels.
a) KHS101 treatment potently reduces HIF target gene expression. Hep3B cells were treated with KHS101 (0-25 μM) and incubated under hypoxia (1% O_2) for 16 hr. The expression of three HIF-driven genes (EPO, PGK1, GLUT1) were assessed by qPCR, demonstrating KHS101 inhibiting transcription with $IC_{50} < 5$ μM.   b) HIF1α mRNA level was measured by qPCR, showing no significant change with increasing KHS101 concentration.   c) HIF1α protein levels are reduced in a KHS101 dose-dependent manner, using anti HIF1α Western blot.

maximizing after 12-24 hr of exposure (Fig. 3-12d). A reasonable hypothesis for this delay in KHS101 efficacy is that its proneurogenic effects are directly related to the time required to alter TACC3 protein levels (4-8 hr by pulse chase; 16 hr under steady state conditions).

To independently assess KHS101 effects in another TACC3-dependent activity, we examined the dose dependence of KHS101 on the transcription of endogenous HIF target genes using qPCR (134). Since KHS101 decreases steady state TACC3 levels, we suspected that compound treatment would lower TACC3 participation in hypoxia-induced HIF complexes, analogously to TACC3 knockdown and ARNT point mutations which weaken TACC3 binding (134). To test this possibility, we measured mRNA levels for three HIF-responsive genes in Hep3B cells, which utilize ARNT in both HIF1 and HIF2 signaling. Though some genes are regulated by both HIF paralogs, others are controlled only by either HIF1 (*e.g.* PGK-1) or HIF2 (*e.g.* Epo) (185,186). As expected, levels of these HIF-regulated transcripts increase (from 5- to 130-fold) upon exposure to hypoxia (Fig. 3-15a). Concomitant treatment with KHS101 and hypoxia led to dose-dependent reductions in the levels of all three transcripts, with apparent IC_{50} values below 5 µM (Fig. 3-15a), correlating the dose dependencies of KHS101 on TACC3 levels (Fig. 3-12c). To further evaluate the mechanism of this effect, we quantitated HIF1α mRNA and protein levels by qPCR and Western blot. While HIF1α mRNA levels were unaffected by KHS101 treatment (Fig. 3-15b), HIF1α protein levels were drastically decreased in a KHS101 dose-dependent manner (Fig. 3-15c). We suspect HIF2α is similarly destabilized in KHS101-treated hypoxic cells, but quantification of this effect is complicated by technical issues with available anti-HIF2α antibodies; compound treatment has no effect on HIF2α mRNA levels (data not shown).

While further studies are needed to fully characterize the breadth of KHS101-induced protein destabilization with respect to several parameters (protein target, cell type, growth conditions), our functional data provide a mode of action for KHS101 and demonstrate its efficacy in two distinct pathways depending on cellular context.

iii. Discussion

Protein/protein interactions are often difficult to modulate with chemical reagents. Here we describe two compounds that affect the ARNT/TACC3 complex, an important component of HIF. These chemicals work by two different mechanisms: KG-548 directly interferes with ARNT/TACC3 complex formation by competing with TACC3 for binding to the ARNT PAS-B domain while KHS101 modulates the abundance of both TACC3 and the HIF component, HIF1α. The different origins of these compounds underscores the merits of parallel *in vitro* target-based and cell-based phenotypic screens, each having strengths and weaknesses in drug discovery. *In vitro* target-based methods are appealing in their use of mechanism-driven hypotheses and focused searches for inhibitors of disease-associated targets, such as HIF or TACC3 (187,188). However, this route often cannot address issues of potency, specificity and metabolism that are essential for cellular applications, and which can be problematic to subsequently incorporate into lead compounds. In contrast, cell-based phenotypic methods have tremendous power to identify new inhibitors of biological activities, but require downstream mechanistic studies to clarify modes of action. In our case, the linkage between ARNT2/TACC3 interactions and neuronal differentiation might have

remained cryptic without the phenotypic screen and initial characterization of KHS101 (171,172). While KHS101 demonstrates the potential of TACC3 destabilization to alter HIF transcriptional activation, we suggest that a target-based approach embodied by KG-548 can provide compounds that work more specifically than by simply destabilizing TACC3 and HIF1α, which may have potential secondary effects. Retrospective analyses of drug discovery successes underscore the utility of combining phenotypic and targeted approaches: the former still generate the majority of first-in-class new molecular entities but are often followed by target-based screens which identify many more candidates using the foundation established by the phenotypic efforts (189).

From the standpoint of small molecule HIF inhibitors, blocking ARNT/TACC3 interactions may hold several advantages over previously described strategies (56,96,190,191). Most notably, this strategy targets a mechanistically-defined interaction (122,134) common among all three HIF paralogs, allowing a single compound to simultaneously block transcription from multiple HIFs given their shared use of ARNT. While our approach shares the general concept of disrupting activator/coactivator interactions with the HIF inhibitor chetomin (96), it is worth emphasizing that chetomin targets a completely different complex (HIFα C-terminal transcriptional activation domain with the CH1 domain of the p300/CBP coactivator) that lacks the small-molecule binding pockets that confer specificity to ARNT PAS-B targeting compounds.

To close, integrating two screening strategies has provided us with small molecules that are useful tools for continuing studies of the critical HIF signaling pathway. This should also refine our understanding of the roles of CCC proteins in HIF-driven gene activation and

could potentially lead to development of new therapeutic routes for HIF-dependent cancers. Finally, we hope that this parallel direct/indirect inhibitor approach provides another example in the relatively limited number of small molecule inhibitors of protein/protein interactions.

CHAPTER 4

SEARCHING FOR HIGH POTENCY SMALL MOLECULE REGULATORS FOR ARNT/CCC

ARNT participates in several key transcriptional factor complexes, where it modulates cellular responses to hypoxia, xenobiotic exposure and other environmental perturbations. As described in Chapter 3, searching for higher potency small molecules to regulate protein/protein interactions between ARNT and CCCs in living cells is of immediate importance.

To obtain this goal, we needed to 1) screen a larger number of compounds with more complexity and diversity than the initial NMR-based screen; 2) switch to a different assay format that allows us to detect a functional readout of disrupting protein/protein interaction instead of simply protein/ligand binding; 3) perform the assay in high throughput format to ensure the screening of large compound libraries can be completed in a competitive time scale.

i. Assay setup and validation

In order to fulfill the robust and interaction-specific criteria for this high throughput screening, I developed a primary screen assay based on an *in vitro* luminescence proximity assay (AlphaScreen, Perkin-Elmer). This assay relies on the singlet O_2-mediated

luminescence signal generated between the AlphaScreen donor and acceptor beads that are brought together within 200 nm by protein complex formation. Compound-mediated signal decreases will suggest candidate disruptor.

    a. AlphaScreen assay setup

Protein reagents GST-TACC3 (561-631) E629A and His-ARNT PAS-B were acquired in sufficient quantities from *E.coli* BL21(DE3) expression and standard column chromatography purification. In 384-well plate format, both proteins were mixed at 1:1 ratio of monomers at a final concentration of 400 nM in 10 μl AlphaScreen buffer (50 mM Tris, pH 7.5; 100 mM NaCl; 1 mM DTT) with 0.2 μl compounds or DMSO controls (DMSO 1.3% v/v). After a 1 hr incubation at 4 °C, AlphaScreen beads (GST AlphaScreen donor beads, Ni AlphaLISA acceptor beads) were added to the protein mixture at 10 μg/ml final concentration (in a total volume 25 μl) under dim green light followed by a 3.5 hr room temperature incubation in a dark, humidified chamber. Protein/protein interactions between His-ARNT PAS-B (blue) and GST-TACC3 (561-631) E629A (orange) bring donor beads and acceptor beads into close proximity and generates detectable chemiluminescence (Fig. 4-1a). Titration of untagged TACC3 or the positive control compound KG-548 perturbed this interaction and resulted in a dose-dependent chemiluminescence signal decrease with IC_{50} values comparable to other biochemical and biophysical experiments (Fig. 4-1bc) (21), suggesting that this assay is interaction-specific and is suitable for small molecule screening.

b. Assay-to-screen transition

The AlphaScreen assay described above was extensively optimized to achieve a successful assay-to-screen transition. The original AlphaScreen buffer was supplement with additives (10% glycerol; 0.02 % Tween 20; 0.1 % BSA) to promote protein stability when using automatic liquid handling. We also lowered the concentration of AlphaScreen beads from 10 μg/ml to 5 μg/ml final concentration, reducing the expenses without impairing the outstanding dynamic range of the assay (~45 fold difference from the ARNT PAS-B-only negative control) (Fig. 4-2a). The optimized condition resulted a robust primary screen assay with great plate-to-plate and day-to-day reproducibility (Z' values ranging from 0.6-0.7; ~50 fold dynamic range) in a 384-well plate format that was amenable to large scale compound screening (Fig. 4-2b).

ii. In-house test screen

A test screen in collaboration with HTS core facility at UT Southwestern Medical Center was conducted to check the compatibility of the assay with the microliter plate format and automated liquid handling. In this screen, we surveyed a library of 11,072 isolated compounds at 5 μM and 2,688 mixtures of natural products obtained from fractionated marine bacterial extracts at 1 mg/ml. A total of 317 hits were identified from the primary screen (2.3% of all samples = 2.1% of isolated compounds + 3.0% natural products fractions)

97
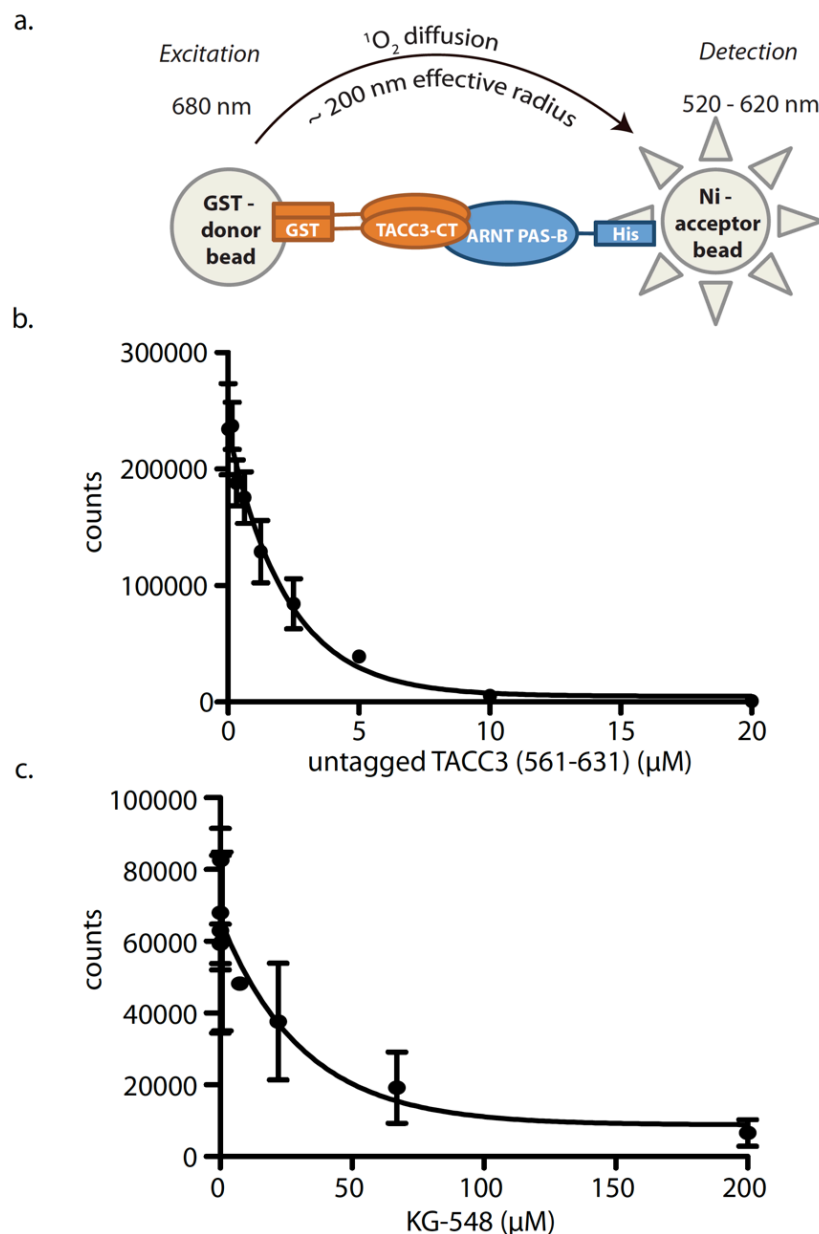
Figure 4-1. AlphaScreen assay setup. a) Schematic of the AlphaScreen assay mechanism. This *in vitro* assay relies on luminescence proximity between "donor" and "acceptor" beads. When two beads are localized within 200 nm by a protein/protein interaction, the singlet oxygen will diffuse from donor beads to acceptor beads, generating detectable chemiluminescence signal. This signal could be disrupted by both competition of untagged protein b) and small molecule inhibitor targeting the protein/protein interaction c) in a dose-dependent manner. The sensitivity of this assay to small molecule inhibitor demonstrates that the signal is not from non-specific interaction between beads and also allows this assay to be used in compound screening.
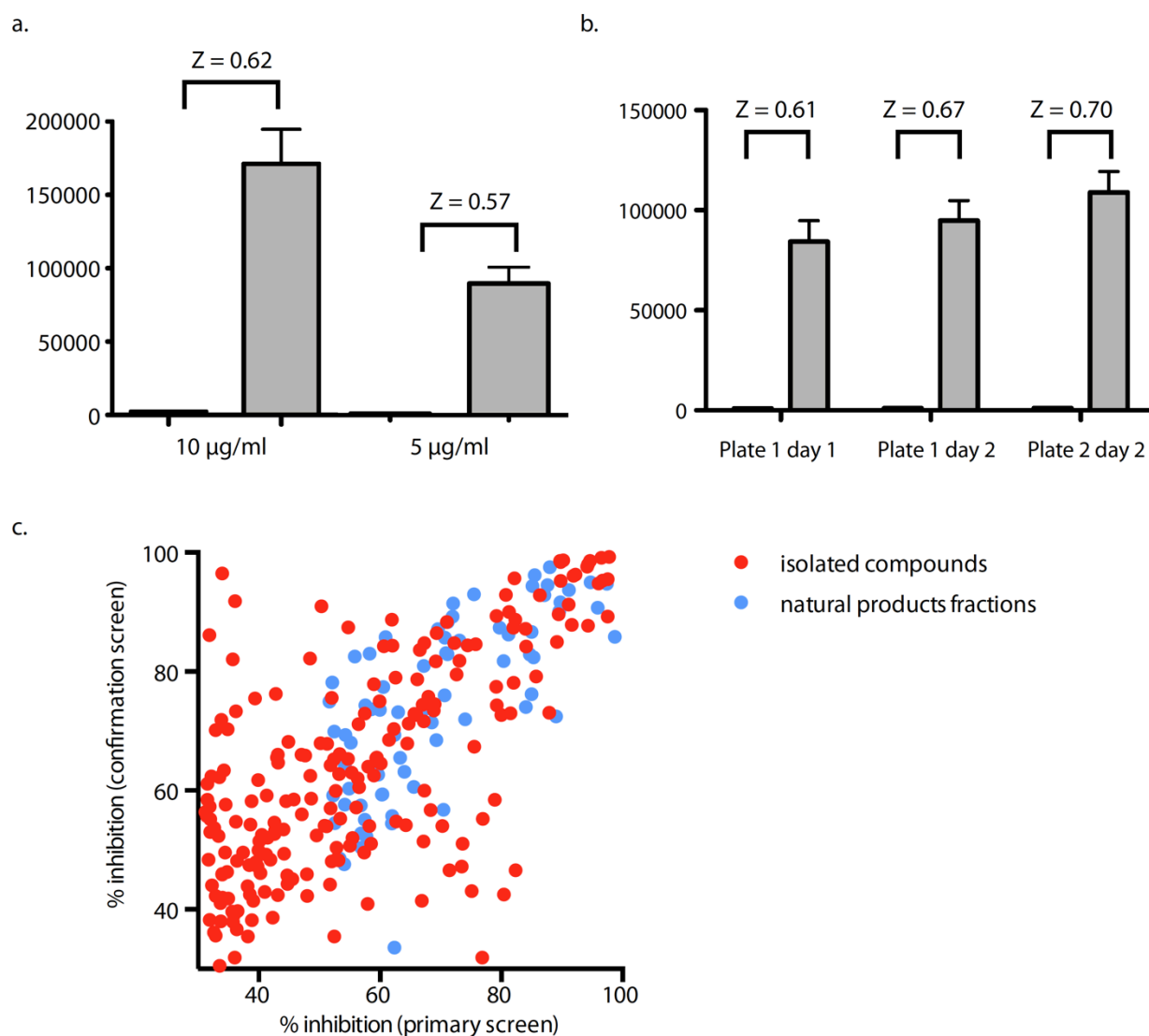
Figure 4-2. Assay automation validation. a) Signal to noise ratio (S/N) and Z score measured using different final concentrations of AlphaScreen beads. There is no significant difference of S/N and Z score by using 10 μg/ml or 5 μg/ml beads (S/N ~ 100). To make the screening more economical, 5 μg/ml final beads concentration was selected for future experiments. b) Day-to-day and plate-to-plate variance of the assay. AlphaScreen assay was conducted on two consecutive days and two independent 384-well plates, showing no significant variance on both S/N and Z score. These data suggest that this assay is robust and suitable for high throughput screening. c) Reproducibility of the AlphaScreen assay. A test screening of over 13,000 single compounds and natural product mixtures identified several inhibitors for ARNT/TACC3 interaction. Retest of over 300 initial hits show a high degree of correlation between replicates.

based on a 3σ cut-off rule (raw signal ≤ plate mean signal - 3σ). Cherry-picked candidates were re-tested in triplicate at the above concentration in a confirmation screen. A high degree of correlation was observed between the primary screen and confirmation screen with 305 compounds out of the 317 hits showing > 30% inhibition in both screens, demonstrating the reproducibility and reliability of this assay (Fig. 4-2c).

iii. Chemical probe development plan - Broad Institute

With the assay well validated, we decided to move on to a larger library to identify high potency small molecule inhibitors/chemical probes for ARNT/TACC3 complex. This was conducted with the collaboration with the Broad Institute. The potentially higher potency, higher affinity probes in this small molecule screen will be useful in cell-based investigation of ARNT/CCCs-dependent transcription in responses to hypoxia, xenobiotic and other environmental perturbations, and in complementary biophysical studies of mechanism of disrupting protein/protein interaction.

a. Compound identification critical path flow

A critical path flow chart was generated as a general guideline for the entire probe development plan, which defines specific experiment in every step of the screen and the cut-off range to identify potential candidate for each assay (Fig. 4-3). Three stages of assays are included in this flow: 1) Primary assay: including a primary AlphaScreen-based high
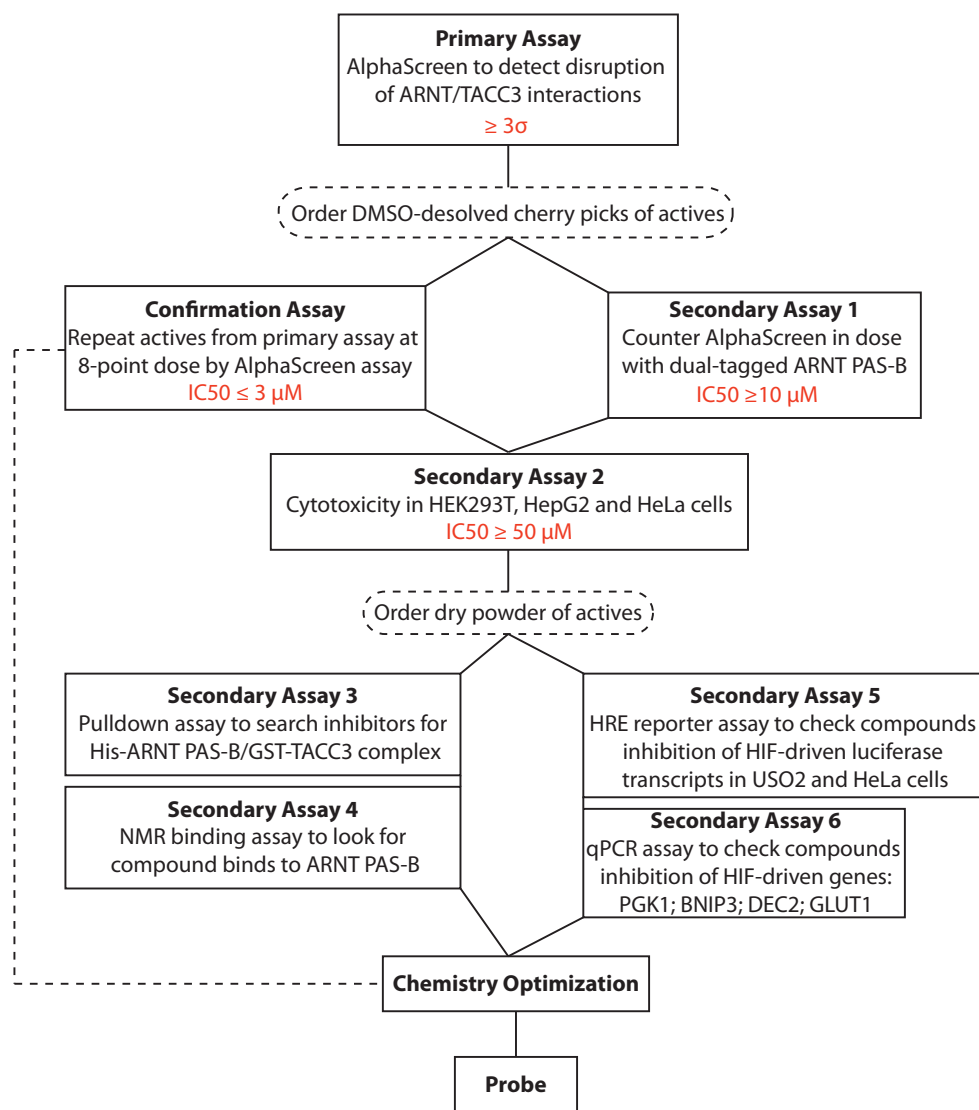
Figure 4-3. Critical path flow of chemical screening and followup assays.

throughput screening over the entire library to detect candidate compounds that can disrupt ARNT/TACC3 interaction at a single dose, and a confirmation AlphaScreen assay to re-test the primary candidates in dose-dependent manner; 2) Secondary assays: including a counter screen assay using GST-ARNT PAS-B-His to rule out false positives that interfere with the AlphaScreen beads, cell-based assays testing compound cytotoxicity and HIF-driven gene inhibition, and *in vitro* pulldown or NMR binding assay to reveal the mode-of-action of the hits. Candidates pass all the steps will be followed up by chemistry optimization that leads to the final probe.

b. Chemical probe development in progress

1) Initial screens searching for small molecule candidates

A primary screen over 392,872 compounds from MLPCN (Molecular Libraries Probe Production Centers Network) and DOS (diversity-oriented synthesis) libraries identified 1338 compounds with > 60% inhibition (0.34% hit rate), including 259 out of 79,632 compounds from DOS library (0.32% hit rate) and 1079/313,180 compounds from MLPCN library (0.34% hit rate). 1247 out of 1338 hits (988 in MLPCN and 259 in DOS) were available for cherry pick and confirmation screens in 8-point doses, from 0-80 μM, generating 153 hits which had IC$_{50}$ < 3 μM (96 in MLPCN and 57 in DOS). 25 out of 153 hits (25 in MLPCN and 0 in DOS) passed the counter screen demonstrated IC$_{50}$ > 10 μM for the single protein, dual tagged GST-ARNT PAS-B-His counter AlphaScreen (Fig. 4-4,4-5).
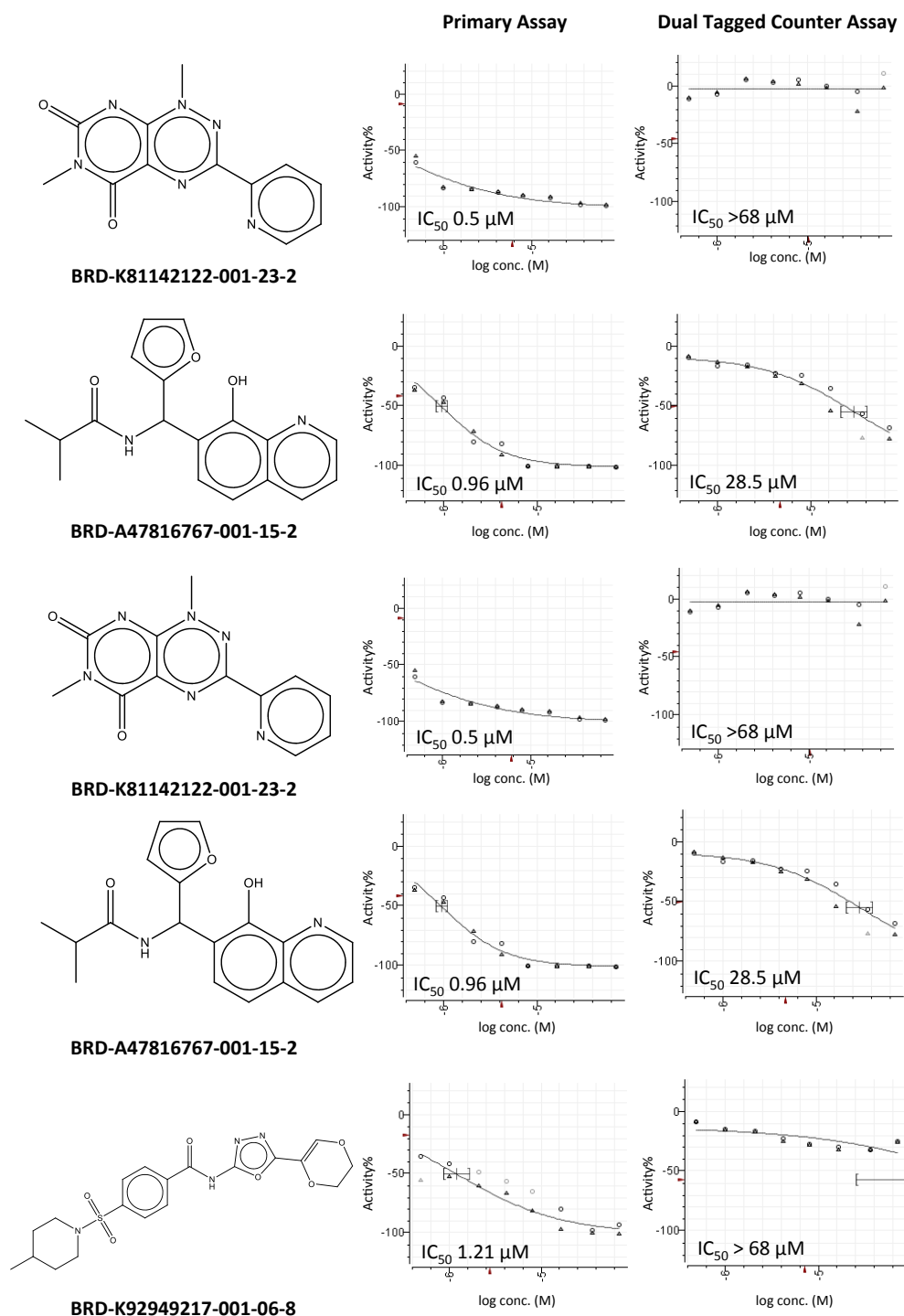
Figure 4-4. Primary assay and dual tagged counter assay results for the top five candidates with the lowest primary assay IC$_{50}$.

BRD-K81142122-001-23-2
IC$_{50}$ = 0.5 μM

BRD-A47816767-001-15-2
IC$_{50}$ = 0.96 μM

BRD-K92949217-001-06-8
IC$_{50}$ = 1.21 μM

BRD-K85597988-004-09-9
IC$_{50}$ = 1.47 μM

BRD-A97888922-001-06-1
IC$_{50}$ = 1.47 μM

BRD-K95473893-001-07-5
IC$_{50}$ = 2.07 μM

BRD-K00097598-001-05-8
IC$_{50}$ = 2.08 μM

BRD-K46071822-001-06-5
IC$_{50}$ = 2.17 μM

BRD-K36074284-001-06-2
IC$_{50}$ = 2.37 μM

BRD-K41646294-001-05-0
IC$_{50}$ = 2.51 μM

BRD-K32342297-001-08-8
IC$_{50}$ = 2.56 μM

BRD-K17909254-001-05-0
IC$_{50}$ = 2.57 μM

BRD-K43909206-001-07-3
IC$_{50}$ = 2.57 μM

BRD-K33857288-001-10-1
IC$_{50}$ = 2.60 μM

BRD-K30213699-001-05-3
IC$_{50}$ = 2.62 μM

BRD-K59544019-001-06-8
IC$_{50}$ = 2.68 μM

BRD-K56711065-001-05-1
IC$_{50}$ = 2.78 μM

BRD-K11140968-001-05-5
IC$_{50}$ = 2.79 μM

BRD-K00617228-001-05-1
IC$_{50}$ = 2.86 μM

BRD-K69085837-001-09-4
IC$_{50}$ = 2.86 μM

BRD-K65837708-001-06-4
IC$_{50}$ = 2.93 μM

BRD-A16145642-001-06-3
IC$_{50}$ = 3.0 μM

IBRD-K54140179-001-05-8
IC$_{50}$ = 3.0 μM

BRD-K46204722-001-05-5
IC$_{50}$ = 3.05 μM

BRD-K01716187-001-05-9
IC$_{50}$ = 3.05 μM

Figure 4-5. Small molecule candidates from MLPCN library that passed primary screen and confirmation assay. Compounds are listed in ascending order of IC$_{50}$ values. Candidates used in the next round of followup assays are highlighted in yellow.
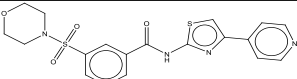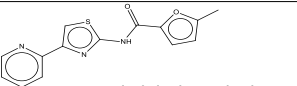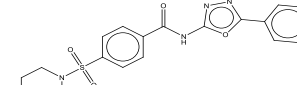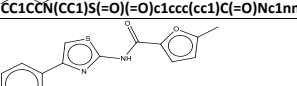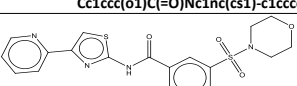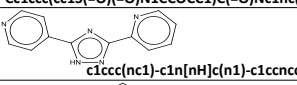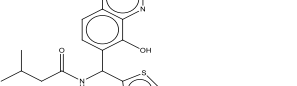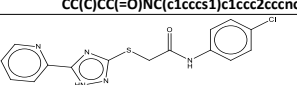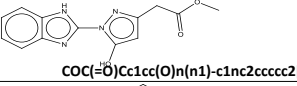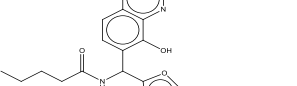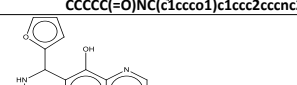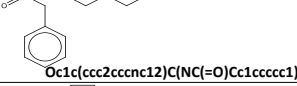
2) *In vitro* pulldown and NMR binding assay confirming mode-of-action

Dry powders of 40 compounds including 16 BRD hits and 24 analogues (9 BRD compounds out of the total 25 hits contained unstable functional groups and were removed from followup assays, Table 4-1, compounds not highlighted in yellow) were ordered for several *in vitro* assays to further confirm the ability of those potential candidates in disrupting the ARNT/TACC3 interaction and determine direct protein binding. Compounds BRD ID was renamed to "YG" ID for simplification (Table 4-1, original hits are highlighted in yellow).

Pulldown assays were conducted by incubating 500 μM final concentration of compounds into pre-formed His-ARNT PAS-B/GST-TACC3 (561-631) E629A complex (5 μM His-ARNT PAS-B + 10 μM GST-TACC3 (561-631) E629A) overnight at 4 °C. ARNT/TACC3 complex was precipitated with 15 μl Ni-NTA beads and eluted protein was resolved in SDS-PAGE and stained with Coomassie blue stain. Compounds capable of disrupting the ARNT/TACC3 interaction will result a weakened band for GST-TACC3 (561-631) E629A compared to DMSO control. Overall, 11 out of 40 compounds showed ≥ 50% inhibition (weakened TACC3 bands) and were marked as "pulldown effective ligands" (Fig. 4-5, Table 4-1).

To test the binding mode between inhibitor candidates and the ARNT/TACC3 complex, ^{1}H-^{15}N HSQC spectra were was used to monitor perturbations of U-^{15}N ARNT PAS-B spectra upon compound addition to reveal direct compound/ARNT PAS-B interaction (87 μM ^{15}N ARNT PAS-B ± 500 μM compound). Unlike KG-548, most BRD hits exhibited only slight peak broadening behavior (Fig. 4-7a). Mapping the top 20 affected residues on the
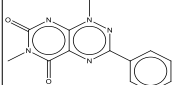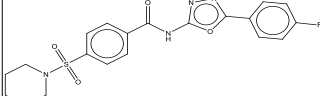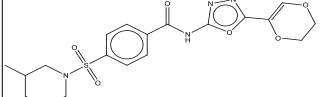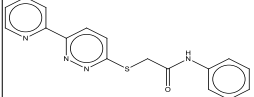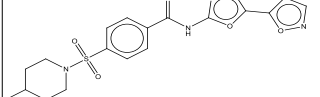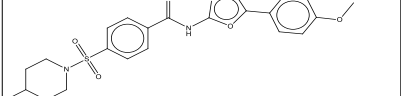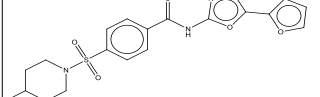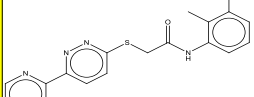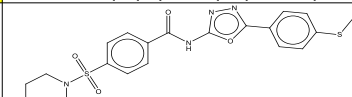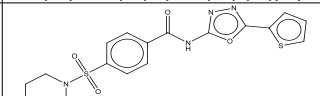
Table 4-1. 40 candidates from primary AlphaScreen hits and analogues.

Rename	BRD ID	SMILES	MW	IC50 primary (µM)	IC50 counter screen (µM)	Pulldown effective	NMR effective
YG-1	BRD-K00073103-001-01-6	O=C(Nc1nc(cs1)-c1ccncc1)c1cccc(c1)S(=O)(=O)N1CCOCC1	430.501				
YG-2	BRD-K95473893-001-09-1	Cc1ccc(o1)C(=O)Nc1nc(cs1)-c1ccccn1	285.321	2.07	22.8		
YG-3	BRD-K55728210-001-01-9	CC1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)-c1ccccc1	426.489				
YG-4	BRD-K18737048-001-01-4	Cc1ccc(o1)C(=O)Nc1nc(cs1)-c1ccccc1	284.333				
YG-5	BRD-K36074284-001-08-8	Cc1ccc(cc1S(=O)(=O)N1CCOCC1)C(=O)Nc1nc(cs1)-c1ccccn1	444.527	2.37	23.5	✔	✔
YG-6	BRD-K11140968-001-07-1	c1ccc(nc1)-c1n[nH]c(n1)-c1ccncc1	223.233				
YG-7	BRD-A58369652-001-12-3	CC(C)CC(=O)NC(c1cccs1)c1ccc2cccnc2c1O	340.439				
YG-8	BRD-K00097598-001-07-4	Clc1ccc(NC(=O)CSc2n[nH]c(n2)-c2ccccn2)cc1	345.807	2.08	>68	✔	
YG-9	BRD-K59544019-001-08-4	COC(=O)Cc1cc(O)n(n1)-c1nc2ccccc2[nH]1	272.259	2.68	60.2	✔	
YG-10	BRD-A07520004-001-01-1	CCCCC(=O)NC(c1ccco1)c1ccc2cccnc2c1O	324.374				
YG-11	BRD-A09774522-001-01-5	Oc1c(ccc2cccnc12)C(NC(=O)Cc1ccccc1)c1ccco1	358.39				
YG-12	BRD-A47816767-001-17-8	CC(C)C(=O)NC(c1ccco1)c1ccc2cccnc2c1O	310.347	0.96	28.5		✔

* Compounds that belong to the original 25 hits (Fig. 4-5) are highlighted in yellow.
* Compounds that are both pulldown effective and NMR effective are highlighted in green at far right column.
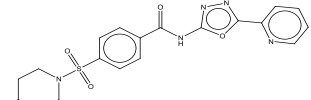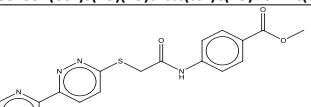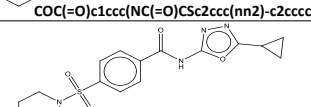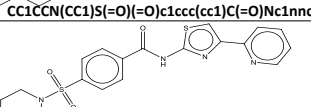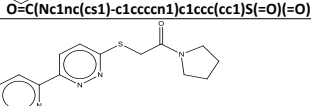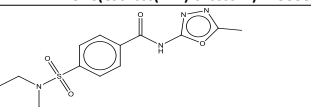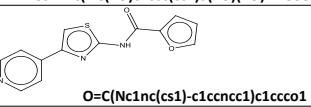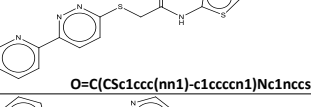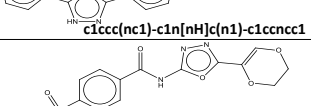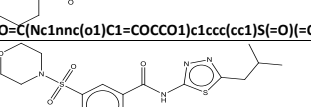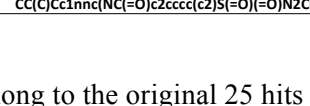
Table 4-1. 40 candidates from primary AlphaScreen hits and analogues (continued).

Rename	BRD ID	SMILES	MW	IC50 primary (μM)	IC50 counter screen (μM)	Pulldown effective	NMR effective
YG-13	BRD-K42767609-001-01-2	CCn1nc(nc2c1nc(=O)n(C)c2=O)-c1cccnc1	284.273				✓
YG-14	BRD-K68444589-001-01-1	Fc1ccc(cc1)-c1nnc(NC(=O)c2ccc(cc2)S(=O)(=O)N2CCCCC2)o1	430.453				
YG-15	BRD-A68913459-001-08-9	C1CCCN(C1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)C1=COCCO	434.466			✓	
YG-16	BRD-K39348181-001-08-7	O=C(CSc1ccc(nn1)-c1ccccn1)Nc1ccccc1	322.384			✓	✓
YG-17	BRD-K91226758-001-01-8	CC1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)-c1ccno1	417.439			✓	✓
YG-18	BRD-K94732626-001-01-2	Oc1ccc(cc1)-c1nnc(NC(=O)c2ccc(cc2)S(=O)(=O)N2CCC(C)CC2)o	456.515				
YG-19	BRD-K34246556-001-06-9	CC1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)-c1ccco1	416.451				
YG-20	BRD-K00617228-001-07-7	Cc1cccc(NC(=O)CSc2ccc(nn2)-c2ccccn2)c1C	350.437	2.86	30.4		
YG-21	BRD-K89567764-001-01-1	Sc1ccc(cc1)-c1nnc(NC(=O)c2ccc(cc2)S(=O)(=O)N2CCC(C)CC2)o	472.58				
YG-22	BRD-K19271421-001-01-0	CC1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)-c1cccs1	432.517				

* Compounds that belong to the original 25 hits (Fig. 4-5) are highlighted in yellow.
* Compounds that are both pulldown effective and NMR effective are highlighted in green at far right column.
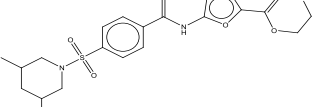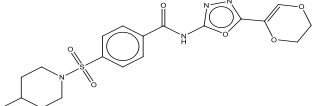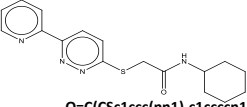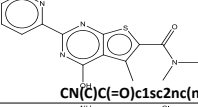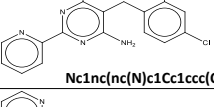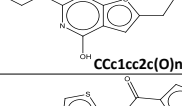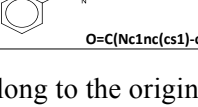
Table 4-1. 40 candidates from primary AlphaScreen hits and analogues (continued).

Rename	BRD ID	SMILES	MW	IC50 primary (μM)	IC50 counter screen (μM)	Pulldown effective	NMR effective
YG-23	BRD-K14600582-001-07-7	CC1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)-c1ccccn1	427.477			✔	✔
YG-24	BRD-K17909254-001-07-6	COC(=O)c1ccc(NC(=O)CSc2ccc(nn2)-c2ccccn2)cc1	380.42	2.57	>68		
YG-25	BRD-K87583716-001-07-6	CC1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)C1CC1	390.457				✔
YG-26	BRD-K47123877-001-01-8	O=C(Nc1nc(cs1)-c1ccccn1)c1ccc(cc1)S(=O)(=O)N1CCOCC1	430.501				✔
YG-27	BRD-K46204722-001-07-1	O=C(CSc1ccc(nn1)-c1ccccn1)N1CCCC1	300.379	3.5	48.1	✔	✔
YG-28	BRD-K64432103-001-01-7	Cc1nnc(NC(=O)c2ccc(cc2)S(=O)(=O)N2CCCCC2)o1	350.393			✔	✔
YG-29	BRD-K42922901-001-08-3	O=C(Nc1nc(cs1)-c1ccncc1)c1ccco1	271.294				✔
YG-30	BRD-K65837708-001-08-0	O=C(CSc1ccc(nn1)-c1ccccn1)Nc1nccs1	329.4	2.93	60.3		✔
YG-31	BRD-K11140968-001-08-9	c1ccc(nc1)-c1n[nH]c(n1)-c1ccncc1	223.233	2.79	13.9		✔
YG-32	BRD-K11283112-001-06-6	O=C(Nc1nnc(o1)C1=COCCO1)c1ccc(cc1)S(=O)(=O)N1CCCCC1	420.44			✔	✔
YG-33	BRD-K19112325-001-05-4	CC(C)Cc1nnc(NC(=O)c2cccc(c2)S(=O)(=O)N2CCOCC2)s1	410.511				✔

* Compounds that belong to the original 25 hits (Fig. 4-5) are highlighted in yellow.
* Compounds that are both pulldown effective and NMR effective are highlighted in green at far right column.

Table 4-1. 40 candidates from primary AlphaScreen hits and analogues (continued).

Rename	BRD ID	SMILES	MW	IC50 primary (μM)	IC50 counter screen (μM)	Pulldown effective	NMR effective
YG-34	BRD-A35765283-001-01-0	1CC(C)CN(C1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)C1=COCCO	448.493				
YG-35	BRD-K92949217-001-08-4	C1CCN(CC1)S(=O)(=O)c1ccc(cc1)C(=O)Nc1nnc(o1)C1=COCCO	434.466	1.21	>68		✔
YG-36	BRD-K32342297-001-10-4	O=C(CSc1ccc(nn1)-c1ccccn1)NC1CCCCC1	328.432	2.56	25.7	✔	✔
YG-37	BRD-K56711065-001-07-7	CN(C)C(=O)c1sc2nc(nc(O)c2c1C)-c1ccccn1	314.362	2.78	23	✔	✔
YG-38	BRD-K33857288-001-12-7	Nc1nc(nc(N)c1Cc1ccc(Cl)cc1Cl)-c1ccccn1	346.214	2.6	25.7		
YG-39	BRD-K69085837-001-11-0	CCc1cc2c(O)nc(nc2s1)-c1ccccn1	257.311	2.86	50.8		
YG-40	BRD-K46071822-001-08-1	O=C(Nc1nc(cs1)-c1ccccn1)c1ccco1	271.294	2.17	26.1		✔

* Compounds that belong to the original 25 hits (Fig. 4-5) are highlighted in yellow.
* Compounds that are both pulldown effective and NMR effective are highlighted in green at far right column.
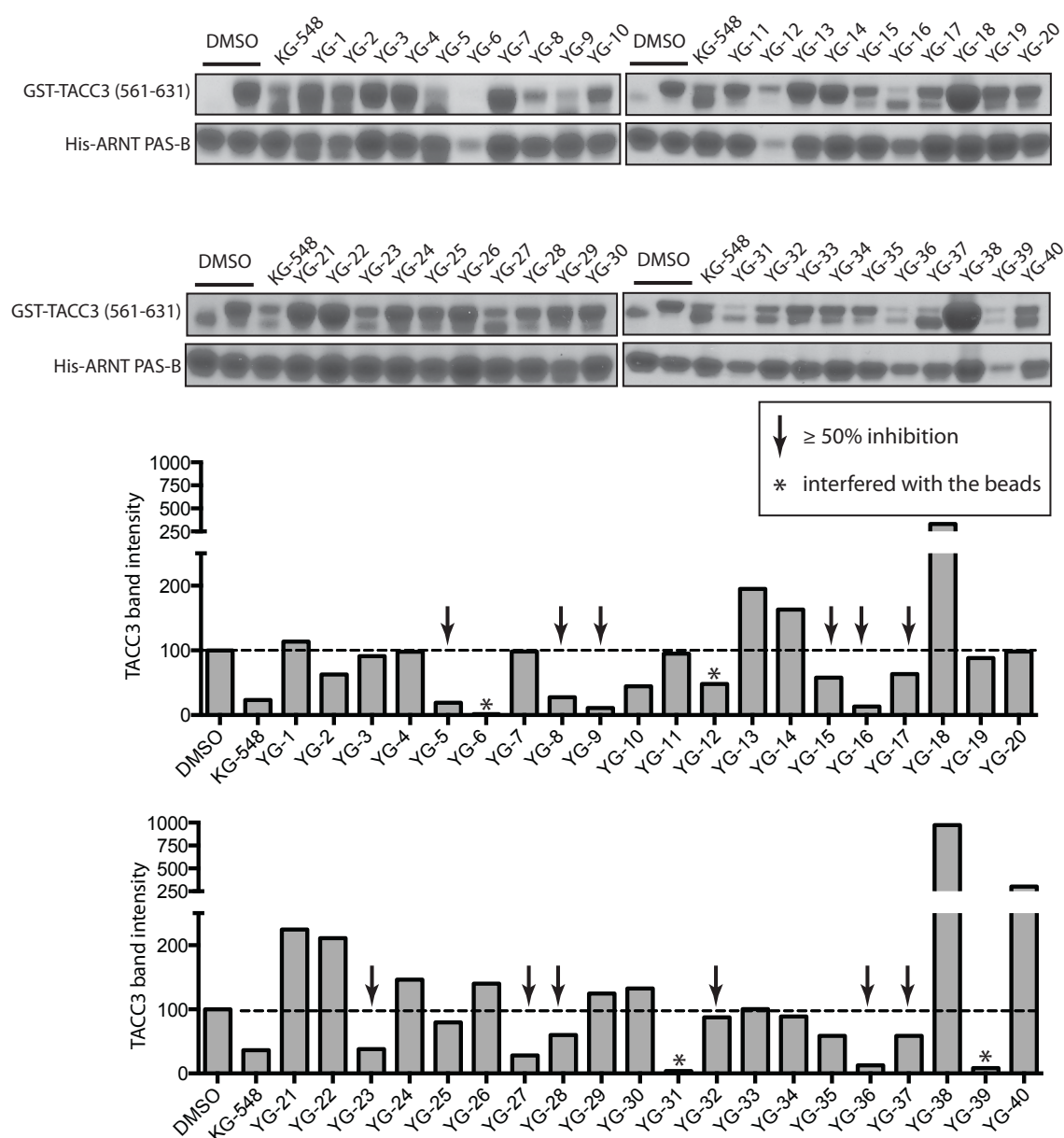
Figure 4-6. Ni pulldown assay testing 40 compounds' ability in disrupting ARNT/TACC3 complex *in vitro*. Upper panel: Ni pulldown assay testing 40 compounds against His-ARNT PAS-B/GST-TACC3 complex. Lower panel: quantification of Ni pulldown assay. Black arrow indicates compounds with ≥ 50% inhibition. Compounds interfering with Ni beads are marked with *.

110

a.

-YG-5
+YG-5

b. top 20 broadened residues

shared by both groups of ligands
pulldown effective ligands only

pulldown effective
pulldown non-effective

c.

d.

peak intensity ratio

≥ 1.0   ≤ 0.35

pulldown effective   pulldown non-effective

Figure 4-7. Small molecule candidates show different ARNT PAS-B binding ability. a) Overlay of ^{1}H-^{15}N HSQC spectra of 87 μM ^{15}N ARNT PAS-B +/- 500 μM YG-5. b) Top 20 residues broadened from compound binding. c) Histogram of average broadening effect on ^{15}N ARNT PAS-B by different compound. The 0.7 cut-off is shown as dashed line and "pulldown effective ligands" are colored red. d) "Pulldown effective ligands" showed stronger average peak broadening effect than "pulldown non-effective ligands".

111

ARNT PAS-B structure (Fig. 4-7b), with residues broadened by both "pulldown effective ligands" and "pulldown non-effective ligands" in red and residues broadened by "pulldown effective ligands" alone in cyan, we observed several trends. Although the "pulldown effective ligands" caused a wider range of peak broadening compared to the "pulldown non-effective ligands", the high degree of overlap between the affected residues (red) suggests a similar binding mechanism with ARNT PAS-B (Fig. 4-7b). Based on this observation, we used an average peak intensity ratio to establish whether a compound is a hit (peak intensity ratio < 0.7 is considered "NMR effective"). Among the 40 BRD hits plus analogs tested by NMR, 19 compounds exhibited strong peak broadening effects and 9 of these "NMR effective ligands" also belong to the "pulldown effective ligands" group, suggesting direct ARNT PAS-B binding is correlated with disrupting protein/protein interactions involving ARNT PAS-B (Fig. 4-7c, Table 4-1).

Worth noticing, both "pulldown non-effective ligands" and "pulldown effective ligands" broadened peaks from residues surrounding the two ARNT PAS-B water-binding cavities. However, the extent of peak broadening caused by these two groups of ligands exhibited some differences: "pulldown effective ligands" induced much larger peak broadening on ^{15}N ARNT PAS-B than "pulldown non-effective ligands" (Fig. 4-7d). One possible explanation behind this phenomenon is that in order to perturb protein/protein interaction by small molecules, a conformational change (here observed as perturbations to the NMR spectra) that is large enough to directly interfere with the binding interface must be introduced.

To further examine the ligand interaction with ARNT PAS-B domain, isothermal calorimetry (ITC) experiments were conducted to establish the binding affinity of potential hits to ARNT PAS-B. To quickly identify compounds amenable to ITC characterization, 9 YG compounds (diluted from 10 mM DMSO stock) (Table 4-1, highlighted in green) were tested by a single, 280 µl injection of 1 mM ARNT PAS-B into 1.4 ml of a 100 µM compound solution. Protein injection into DMSO and KG-548 solutions served as negative and positive controls, respectively. Heats generated in this single injection ITC experiment are listed in Table 4-2, where % difference is defined as the difference in heats evolved in compound titrations divided by heats in the DMSO control. Since all compounds demonstrated very small heat differences in comparison with the DMSO control, we selected the three compounds which exhibited ~5% heat changes compared to DMSO for full titration: YG-17, YG-27 and YG-32.

Table 4-2. Heat and % heat change generated by single injecting of ARNT PAS-B into compounds.

Compound name	Heat generated (kcal/mol)	% Difference (with DMSO)
DMSO	-305.5	-
KG-548	-400.7	31.2
YG-05	-303.1	-0.8
YG-16	-319.2	4.5
YG-17	-299.8	-1.9
YG-23	-302.9	-0.9
YG-27	-279.2	-8.6
YG-28	-310.9	1.8
YG-32	-289.9	-5.1
YG-36	-293.6	-3.9
YG-37	-299.5	-2.0
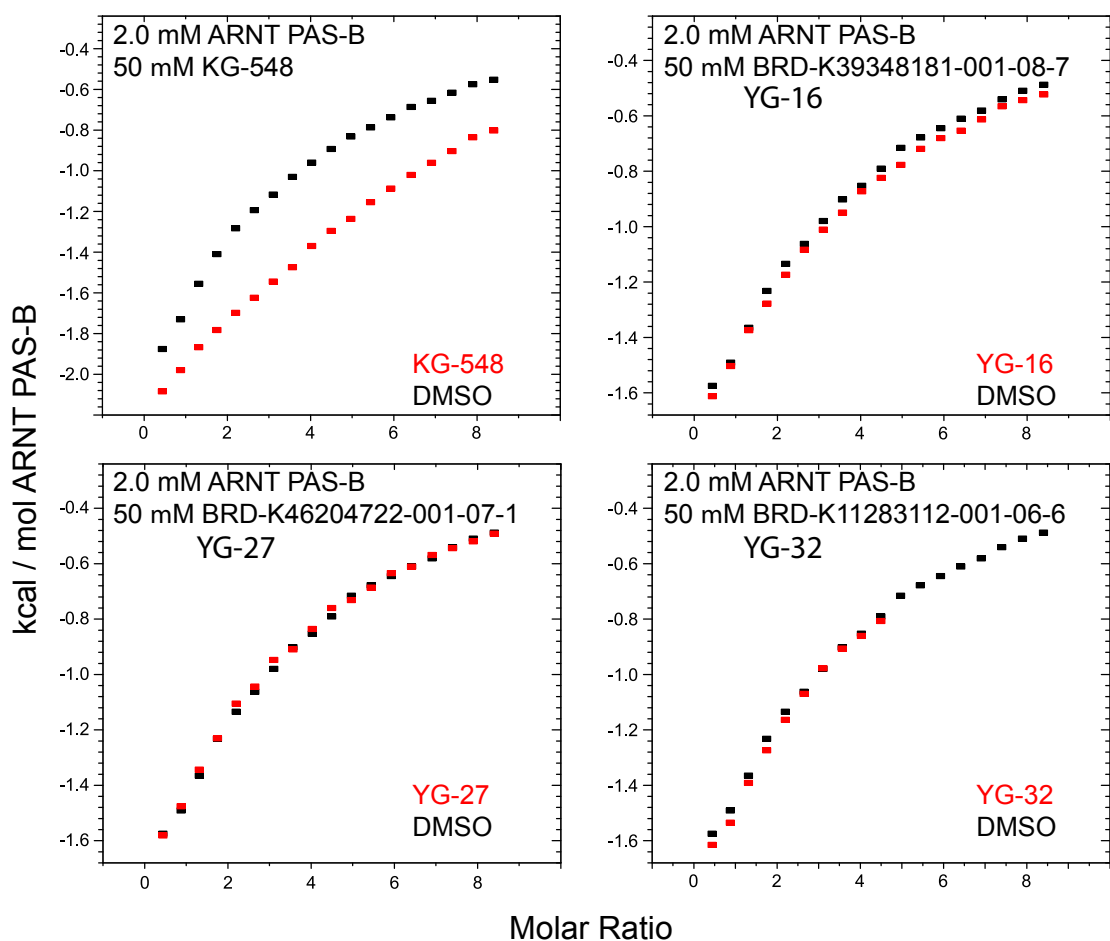
Figure 4-8. ITC titrations of potential ARNT PAS-B binding compounds. Heat evolved from ARNT PAS-B full titrations (15 μl/injection, 19 injections) into 1.4 ml, 2 mM KG-548 (positive control) or YG compounds (red) and DMSO control (black). While titration of ARNT PAS-B into KG-548 provided small magnitude heats of binding distinct from titration in to a no compound solution, YG compound complexes provided heats indistinguishable from the control experiment.

Conventional ITC titrations were conducted as 19, 15 μl injections of 2 mM ARNT PAS-B into 50 μM YG-compound solutions, in a buffer supplemented with 5% DMSO to promote free compound solubility. Unfortunately, the heats evolved from titrating ARNT PAS-B into YG compounds were not substantially different from the heat evolved in the DMSO control experiment (Fig. 4-8). This result suggested that either these complexes bear weak affinity relative to the concentration of material in the calorimeter cell ($K_d > 50$ μM) or that these complexes possess very small heats of binding under the current experimental conditions. If the heat capacity of these complexes is stable over a broad temperature range and the true K_d of these complexes are less than 50 μM, changing assay temperature could provide measurable heats of association.

While the YG compounds resisted quantitative analysis, to better understand the ligand binding role of ARNT PAS-B, we turned attention to previously identified drug fragment hits (KG-548 and KG-655). Due to the poor solubility of compounds, experiments were conducted by titrating a total of 285 μl, 1 mM ARNT PAS-B into 1.4 ml, 100 μM solution (Fig. 4-9). This ITC titration provided an unusual isotherm with the largest magnitude heats evolving at the end of the titration. To better understand this phenomenon, we repeated the titration with higher protein concentration (2 mM ARNT PAS-B) and lower compound concentration (50 μM compound solution) to see if we could saturate this effect (Fig. 4-10).

Interestingly, the net heat measured with high ARNT PAS-B concentration titration exhibited a bell shaped curve instead of the usual hyperbolic or sigmoidal curve observed in successful ITC experiments (Fig. 4-9, 4-10). We hypothesized that this phenomena was
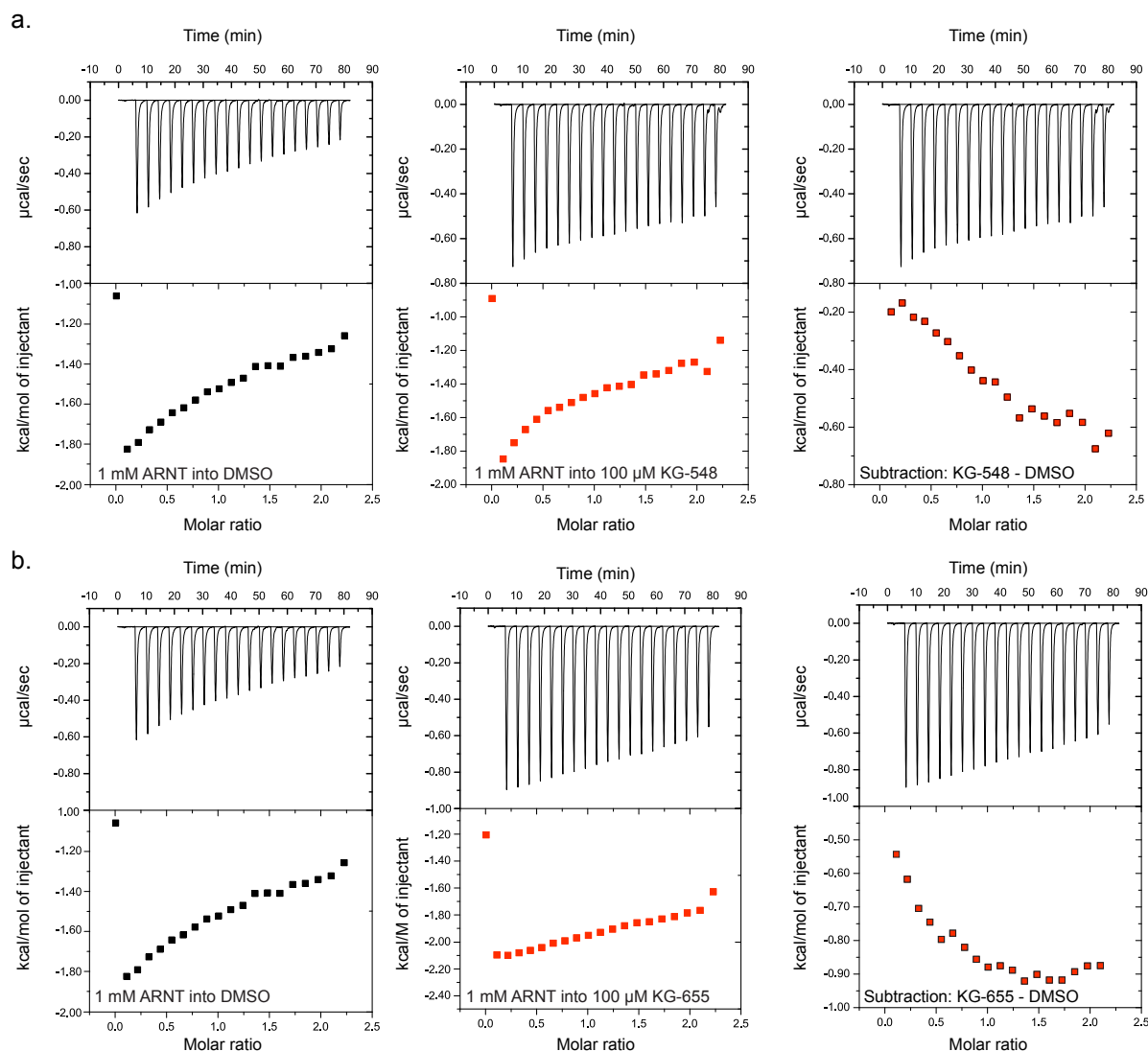
Figure 4-9. ITC experiments with KG-548 and KG-655 resulted unusual isotherm. These complexes were tested by 19, 15 μl injections of 1 mM ARNT PAS-B into 1.4 ml, 100 μM a) KG-548 and b) KG-655. Left panel: thermogram of titrations into DMSO. Middle panel: heats evolved from titration of protein into compound. Right panel: net heats of ARNT PAS-B/compound titration (compound - DMSO). Unlike isotherm, heats of compound binding were greatest at the end of titration.
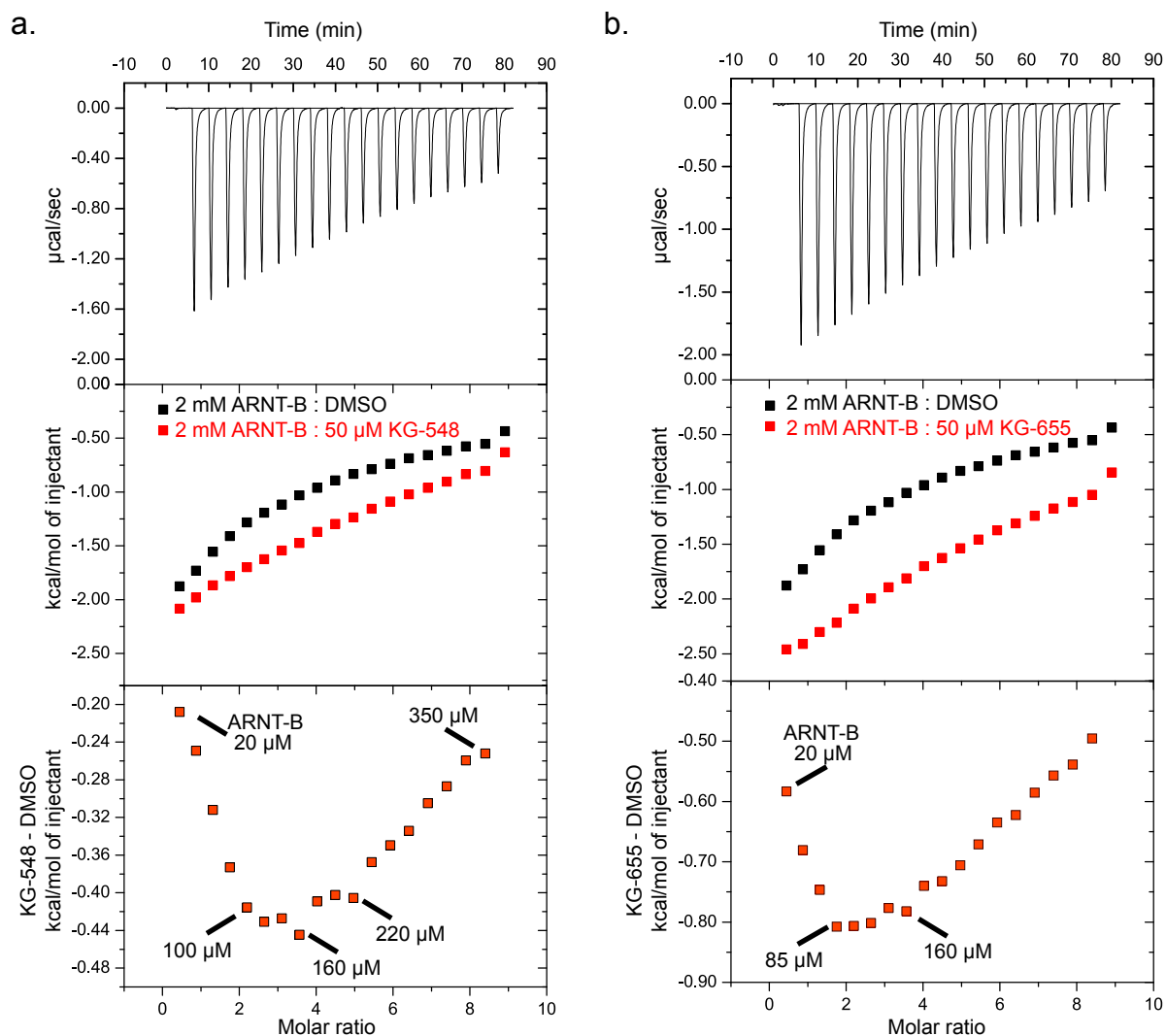
Figure 4-10. KG-548 and KG-655 may preferentially bind to ARNT homodimer. a) ARNT PAS-B/KG-548 complex. Upper panel: thermogram of 19, 15 µl injections of 2 mM ARNT PAS-B titrated into 1.4 ml, 50 µM KG-548 or DMSO-only control. Middle panel: comparison of heats evolved from titration into KG-548 (red) and DMSO control solutions (black). Bottom panel: net heats of ARNT PAS-B/KG-548 (KG548 - DMSO) titration. b) ARNT PAS-B/KG-655 complex. Upper panel: thermogram of 19, 15 µl injections of 2 mM ARNT PAS-B titrated into 1.4 ml, 50 µM KG-655 or DMSO-only control. Middle panel: comparison of heats evolved from titration into KG-655 (red) and DMSO control solutions (black). Bottom panel: net heats of ARNT PAS-B/KG-655 titration (KG-655 - DMSO).
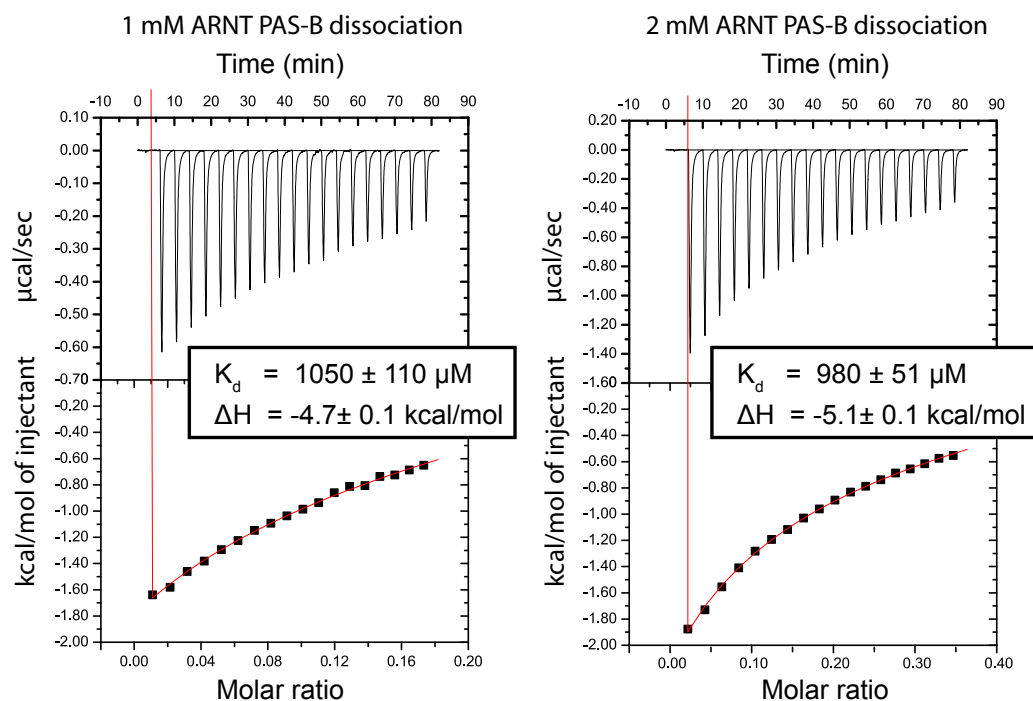
Figure 4-11. ARNT PAS-B homodimer affinity measured by ITC experiments. The affinity of the ARNT PAS-B homodimer is estimated by dilution of either 1 mM (left panel) or 2 mM (right panel) ARNT PAS-B into buffer. Both experiments were conducted as 19 injections of 15 μl protein and were analyzed with the "dissociation" model implemented in the vendor-supplied version of Origen.

associated with ARNT PAS-B homodimerization. The ARNT PAS-B homodimer was previously detected in solution by chemical shift difference observed in serially-diluted ARNT PAS-B NMR samples with Kd ~500 μM (156). This dissociation constant was measured again by ITC in this study, giving Kd ~1 mM (Fig. 4-11). Taken together with the unconventional isotherms collected in the presence of compound, these data raise the possibility that the compound may preferentially bind to or even stabilize the ARNT PAS-B homodimer. In such a binding mechanism, the initial injections of ARNT PAS-B into the calorimeter cell will provide low ARNT PAS-B concentrations, limited ARNT homodimer accumulation in the cell, and very small heats associated with compound binding. As the concentration of ARNT PAS-B in the calorimeter cell approaches the ARNT PAS-B homodimer K_d, ARNT homodimers will accumulate, which will be correlated with increasing heats of compound binding (Fig. 4-10, bottom panel) (molar ratio = 2). With additional ARNT PAS-B titration, binding heat evolved from compound/ARNT PAS-B dimer interaction was then attenuated when un-bound compound concentration dropped to a certain limit (Fig. 4-9, 4-10). Conversely, the same compound can attenuate ARNT PAS-B heterodimerization with HIF2α PAS-B. Therefore, due to the uncertain biological significance of ARNT PAS-B homodimer and the opposite effect that the drug fragments have on homo- and heterodimers, this hypothesis needs further modeling to generate the thermodynamic cycle for this process as well as biophysical characterization to reveal the structural details of how compounds distinguish different ARNT PAS-B oligomers.

Taken together, by combining the hits from pulldown assay and NMR binding assay, 9 hits that can disrupt ARNT/TACC3 interaction via direct ARNT PAS-B binding were selected for further evaluation.

3) Cytotoxicity and gene expression inhibition assay

All 9 candidates were tested at the Broad Institute using their cytotoxicity panel on three cell lines: HEK293T, HepG2 and HeLa. No severe cytotoxicity was observed in all three cell-lines, making them all suitable for cell-based assays to test whether they are able to inhibit HIF-driven gene expression under hypoxia (Table 4-3). Two parallel types of transcriptional assays were used: HRE-luciferase reporter assay and qPCR assay. KHS101 was used as a positive control.

Table 4-3. Cell-based data summary for AlphaScreen candidates.

compound ID	Cytotoxicity IC$_{50}$ (µM)			HRE-luc reporter AC$_{50}$ (µM)		qPCR AC$_{50}$ (µM)			
	HEK293T	HepG2	HeLa	U2OS	HeLa	PGK1	BNIP3	DEC2	GLUT1
YG-5	> 50	> 50	> 50	3.6	3.5	62	> 100	25	76
YG-16	25.8	> 50	> 50	12	30	> 100	31	75	53
YG-17	> 50	> 50	> 50	> 60	> 60	> 100	> 100	50	0.4
YG-23	> 50	> 50	> 50	> 60	49	> 100	-	> 100	43
YG-27	> 50	> 50	> 50	26	30	88	> 100	> 100	> 100
YG-28	> 50	> 50	> 50	> 60	> 60	7.8	> 100	88	51
YG-32	> 50	> 50	> 50	> 60	55	> 100	> 100	> 100	> 100
YG-37	> 50	> 50	> 50	30	> 60	> 100	> 100	> 100	88
YG-36	> 50	> 50	> 50	12	46	> 100	-	50	83

HRE-luciferase assay were done in two stable cell-lines with HRE-driven luciferase reporters incorporated in the genome (U2OS and HeLa). In this arrangement, the luciferase signal can only be induced by HIFα subunit binding to the HRE site under hypoxia or

AlphaScreen Assay

Primary Assay

Dual Tagged Counter Assay

IC_{50} = 1.1 µM

IC_{50} = 10.0 µM

YG-5

Cytotoxicity Assay

Hek293T Cytotox

HepG2 Cytotox

HeLa Cytotox

IC_{50} > 52 µM

IC_{50} > 52 µM

IC_{50} > 52 µM

HRE-Luc Reporter Assay

U2OS HRE-Luc

HeLa HRE-Luc

AC_{50} = 3.6 µM

AC_{50} = 3.5 µM

qPCR Assay

PGK1

BNIP3

DEC2

GLUT1

AC_{50} = 62 µM

AC_{50} > 100 µM

AC_{50} = 25 µM

AC_{50} = 76 µM
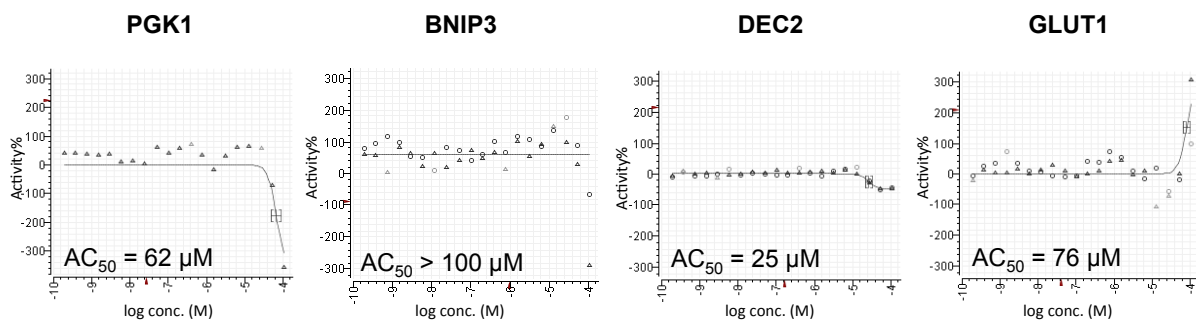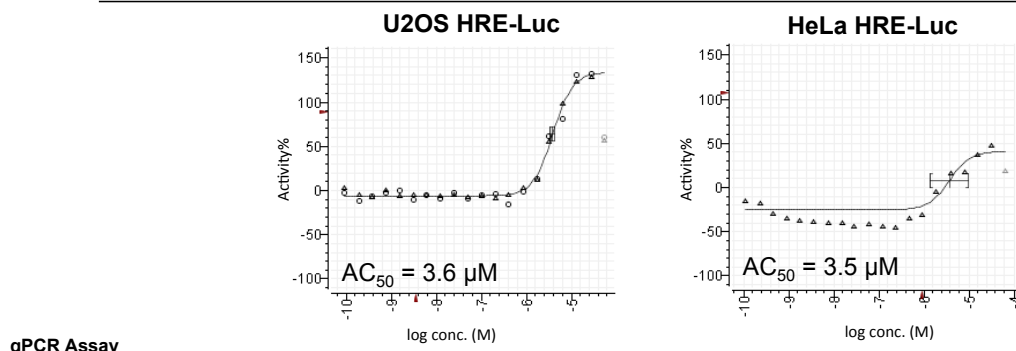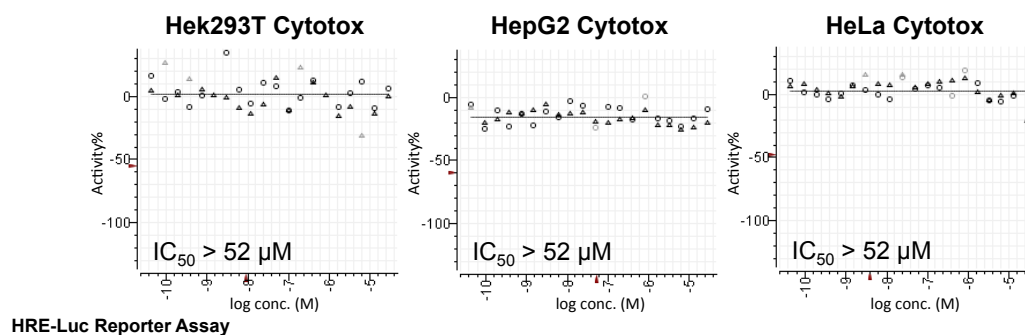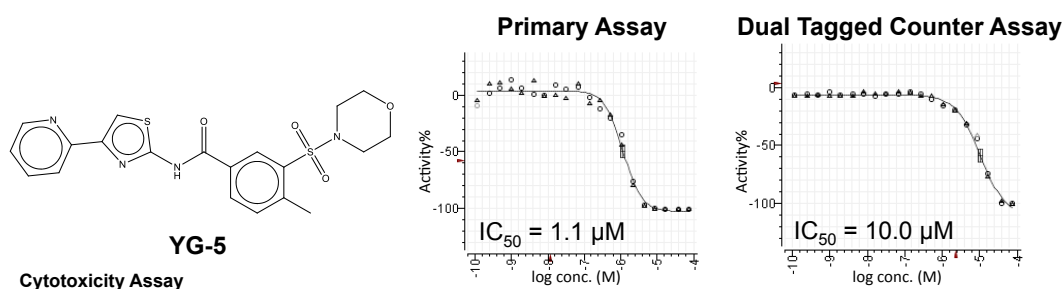
Figure 4-12. YG-5 primary and followup assay results.

121

chemical mimics and is a good read-out for HIF-driven gene upregulation. Compounds capable of blocking the luciferase signal upregulation will be considered as good candidates for follow-up optimization. Results of compound YG-5 were shown as examples in Figure 4-12. Data for other compounds were summarized in Table 4-3. Unfortunately, only one out of 9 compounds met the pre-established cut-off value from the critical path with AC50 < 10 μM (Table 4-3). Since HRE-luciferase assay represents expression from a single promoter that has been artificially constructed, qPCR experiments in Hep3B cells could provide more information on compounds' effect on native gene expression. Four TACC3-dependent HIF-driven genes were examined here (PGK1, BNIP3, DEC2 and GLUT1) at dose-dependent manner. Unfortunately, except KHS101, none of the 9 candidates were able to show potent AC_{50} to all four genes (Table 4-3). Due to this reason, the chemical probe development plan with Broad Institute paused here. All results generated from this screen were updated to PubChem (BioAssay ID: AID 623900) and can be found at:

http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=623900.

iv. Summary and future directions

In this chapter, I described the generation of an assay suitable for high throughput screening for small molecule inhibitors of ARNT/TACC3 transcription activator/coactivator interactions. Several validation assays and test screen demonstrated that this assay is reproducible and robust for screening for large libraries.

The high throughput screening in collaboration with Broad Institute failed to generate useful high potency small molecule candidate for chemical optimization and lead to early termination of the project. I suspect that one major reason for this result was that the size of the compounds used in the screen exceeds the binding capacity of ARNT PAS-B cavity. Compared to the previous NMR fragment-based screen where the average molar mass of the compound library is ~203±70 Da, compounds used in this screen is much bulkier with multiple heterocyclic rings (final 40 BRD hits have an average molar mass ~ 360±70 Da). Considering the cavity size of ARNT PAS-B domain measure from the crystal structure is only ~ 100 $Å^3$, it is very hard to imagine such bulky ligands being able to fit into this small cavity and induce conformational changes. Additionally, many of the MLPCN compounds exhibited very low solubility in aqueous solutions. The solubility issue makes it difficult to reach high enough concentration of the ligands in cell culture media and harder for cell to take them in, possibly explaining why high AC_{50} values observed in Luc-reporter assay and qPCR.

Although this initial chemical probe development plan didn't result high potency inhibitors for ARNT/TACC3 complex that could be used in cell-based studies, it still provided several meaningful aspects. First of all, the AlphaScreen assay developed in this study was proven to be robust and reliable. The successful transfer from in-house HTS core facility to the HTS platform at the Broad Institute, achieved without making major changes to the protocol, demonstrated that this assay is reproducible and could be repeated on screening platforms. It is suitable to switch to other libraries in the future to look for inhibitors. Secondly, the 9 compounds identified in this screen that could disrupt

ARNT/TACC3 interaction *in vitro* may provide structural information for ARNT PAS-B inhibitor development if it is possible to acquire the ARNT PAS-B/ligand complex structures by soaking compounds with pre-formed ARNT PAS-B crystal.

Taken together, finding small molecule inhibitors by this *in vitro* target-based method is still an exciting idea due to its mechanism-driven hypothesis. The successful application of this method will be another important proof of concept for anti-cancer potentials in ligands that inhibit transcription factor/coactivator interaction in HIF complex. With a well-validated, robust screening method in hand, I believe that potent inhibitors targeting ARNT/TACC3 with known mode-of-function will eventually be discovered.

v. Acknowledgements

CHAPTER 5

CONCLUSION AND FUTURE DIRECTIONS

i. Conclusions

Coiled-coil coactivators are important transcription coactivators in ARNT-involved signaling pathways. Misregulation of CCCs can perturb ARNT downstream gene expression and is associated with several types of cancer. Due to this significant biomedical relevance, it is critical to understand the recruiting mechanism and the transactivation function of CCCs in ARNT-controlled transcription. Here we use TACC3 as an example, hoping to reveal the linkage between TACC3 recruitment and HIF function.

To address the problem of CCC recruitment, we first characterized the ARNT/TACC3 complex in multiple biochemical and biophysical methods, confirming the minimal interaction domains of the two proteins. Later on, mutagenesis was conducted to generate mutations that are capable to stabilize the ARNT/TACC3 interaction to assist structural studies. Two well-superimposed crystal structures of ARNT PAS-B/TACC3-CT (the minimum interaction domains) were solved at 3.15 Å and 3.5 Å resolution, utilizing different set of mutants and crystallization conditions, confirming that TACC3 is recruited to ARNT PAS-B through a β-sheet interaction. Detailed analysis on two structures in combination with mutagenesis and NMR spectroscopy further demonstrated that the crystal

structures holds true in solution. We also observed direct interactions between TACC3 and HIF2α PAS-B for the first time, letting us generate NMR models of the HIF2α PAS-B/TACC3-CT heterodimer and the HIF2α PAS-B/TACC3-CT/ARNT PAS-B ternary complex. Those models were both confirmed by PRE studies in solution. This discovery provides the structural basis for TACC3 recruitment in HIF signaling pathway and a good reference to understand the recruitment of other CCCs.

As mentioned before, misregulation of CCC proteins shows a strong linkage with cell transformation and the development of glioblastoma, renal cell carcinoma and other cancers. Inhibiting ARNT/CCC interaction with small molecule inhibitors could possibly become a novel and rapid route to blocking cancer formation and progression. Thus, in the second half of the study, we developed two *in vitro* compound screens (fragment-based NMR screen, AlphaScreen-based high throughput screen) to identify small molecules that specifically bind within the cavity of ARNT PAS-B to perturb TACC3 interactions. Compound KG-548 was discovered in the fragment-based NMR screen, which exhibits specific ARNT PAS-B binding and is capable of breaking ARNT/TACC3 complex both in minimum interaction domain and full length protein from cell lysate. Although KG-548 is unable to regulate gene transcription in cells, it still serves as a good example of ARNT PAS-B inhibitor with known mode-of-action, which can be useful in future inhibitor design. To pursue inhibitors with higher potency, we developed a different screen protocol, the AlphaScreen-based high throughput screen, to screen larger libraries in a more efficient manner. While the collaboration to use this screen with the Broad Institute didn't result high potency inhibitors so far, this well-validated screen method could be easily translated to other screening

platforms, making it possible to continue the search. The potent ARNT/CCC inhibitors discovered in this method can serve as good cell-biology tools and precursors for potential anti-cancer drugs.

ii. Future directions

In this study, we made several progresses in understanding the ARNT/TACC3 recruitment and developing artificial inhibitors for this complex. However, several questions regarding ARNT/CCC interactions in physiological conditions and the role of CCC proteins in HIF signaling still remain unknown. In this regards, future experiments should be conducted to address these problems in the following aspects.

a. Biochemical and biophysical studies on longer protein constructs

Structures and models described in Chapter 2 are based on minimum interaction domains of the three proteins. Since ARNT and HIF2α are both multi-domain transcription factors, it is very likely that other protein/protein contacts between TACC3 and the HIF complex will exist. Preliminary NMR titration data have already shown some limited direct interaction between CCC and the PAS-A domain of ARNT (Fig. 5-1, Carrie Partch unpublished data). Those contacts are probably weak so they could not be detected using truncated domains by pulldown or other biochemical methods. However, those interactions could still play critical roles to stabilize the ternary complex formation in a cooperative way.

Therefore, structural studies with longer constructs or even full length proteins would help us to look at CCC recruitment with the HIF complex with a more physiological relevant view and to discover more drug target sites.
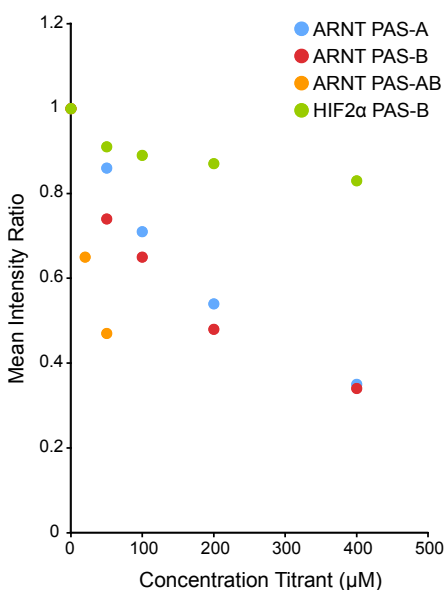
Figure 5-1. ^{15}N HSQC titration tests protein binding with ^{15}N/^{13}C TRIP230 (1583-1619). (Adapted from Carrie Partch unpublished data). Up to 400 μM of different combinations of PAS domain from ARNT and HIF2α were titrated into 100 μM ^{15}N/^{13}C TRIP230. Binding was measured by dose-dependent peak broadening upon titration. ARNT PAS-A titration (blue) demonstrated similar broadening effect with ARNT PAS-B (red), suggesting direct interaction between TRIP230 and ARNT PAS-A.

b. Transactivation functions of the CCC protein in HIF signaling

In vitro studies have shown a competitive relationship among all CCC members in association with ARNT. Disrupting each ARNT/CCC complex also exhibits distinctive effects on the expression levels of different HIF-driven genes (134). These observations could suggest a strategy that cells use for rapid adaptation to environmental changes by

choosing different CCCs. Hence, it will be of great interest to look at individual transactivation functions for each CCC member in HIF signaling. A critical step to address this question is to be able to selectively block individual ARNT/CCC interactions or to block all three ARNT/CCC interactions at the same time, so that we could use RNAseq to compare the gene expression profile under hypoxia or normoxia conditions to find out distinct role for each CCC. Selectively blocking ARNT/CCC interactions can be easily done by siRNA knockdown; however, it is difficult to block all three ARNT/CCC interactions at once. To achieve this goal, several attempts were tried to develop a TACC3-CT based dominant negative peptide (DNP) that has a higher affinity with ARNT PAS-B than WT CCCs but doesn't have the transactivation domains at the N-terminus. Theoretically, this DNP would compete away all endogenous CCCs but not have the transactivation activity. Unfortunately, while several high affinity DNPs were successfully made, they all failed to downregulate HIF target genes under hypoxia. This could possibly be explained that the DNP also tightens HIFα/ARNT interactions so that the DNP effect on CCC transactivation is cancelled out by the enhanced transcription caused by this tighter HIF complex. Another approach to accomplish this task is to develop ARNT/CCC inhibitors that selectively block the ARNT/CCC interaction, but does not affect HIFα/ARNT. This could be done by continue the high throughput screens described in Chapter 4. To summarize, this initial attempt to distinguish each CCC's function in HIF signaling will help us to find out how cells balance between different coactivator recruitments under different circumstances. In combination with structural studies on HIF complexes with CoCoA and TRIP230, we may also find out the mechanism for HIF selectivity over CCC proteins.

c. Translate the AlphaScreen-based high throughput screen to other screening platforms and continue search for ARNT/CCC inhibitors

Finding high potent inhibitors for ARNT/TACC3 is important from many perspectives, including cell biology tool development and drug design. The AlphaScreen-based high throughput screening protocol described in Chapter 4 is well-validated by several assays and is easily translated from the in-house HTS core to the Broad Institute screening platform. I suspect that the BRD library screen didn't produce useful hits due to the incompatible compound size and ARNT PAS-B cavity. For future applications, we should carefully choose compound libraries and continue searching for inhibitors with higher potency.

CHAPTER 6

MATERIALS AND METHODS

i. Common methods for all chapters

Plasmids - For bacterial expression, human ARNT PAS-B domain (356-470), human ARNT2 PAS-B domain (330-444) and mutants were cloned into pHis-parallel and pHisGβ1-parallel vector (192). Human HIF2α PAS-B domain (240-350) and mutants were cloned into pHis-parallel vector and pGST-parallel vector. Mouse TACC3 (585-631) was cloned into pGST-parallel vector. GCN4 -TACC3-CT was made by fusing *S. cerevisiae* GCN4 (264-281) (154) to the N-terminus of mouse TACC3 (610-631). GCN4-TACC3-CT was cloned into pHisGβ1- parallel vector.

Protein purification - Proteins were expressed in BL21(DE3) cells (New England Biolabs), grown in LB or M9 minimal media (with 1 g/L $^{15}NH_4Cl$ for U-^{15}N labeling; with 3 g/L U-^{13}C glucose U-^{13}C labeling) at 37°C and induced with 500 μM IPTG at an OD_{600} of 0.70. After overnight growth at 20°C, cells were harvested by centrifuging at 4°C, 4600 g for 40 min. Pellets were resuspended in appropriate buffers for affinity chromatography (Ni^{2+} affinity: 50 mM Tris pH 7.5, 150 mM NaCl, 20 mM imidazole, 5 mM β-mercaptoethanol; glutathione affinity: 1× phosphate buffered saline) and stored at -80°C. Purification was conducted using affinity chromatography (Ni Sepharose High

Performance or Glutathione Sepharose 4B, GE Healthcare). Affinity tags (GST, His or HisGβ1) were cleaved by His6-TEV protease (193) overnight at 4°C, followed by size-exclusion chromatography (Superdex 75, GE Healthcare). Samples of ARNT2 PAS-B used for triple resonance NMR experiments for backbone chemical shift assignment were kept in a buffer of 25 mM Tris (pH 7.3), 17 mM NaCl, 2.5 mM TCEP and 10% (v/v) D2O. Other NMR experiments were done in 25 mM Tris (pH 7.5), 17 mM NaCl, 5 mM β-mercaptoethanol and 10% (v/v) D2O.

.

ii. Methods for Chapter 2

ARNT PAS-B/GCN4-TACC3-CT D622A/E629A X-ray crystallography - Single crystals of ARNT PAS-B/GCN4-TACC3-CT D622A/E629A were grown by hanging drop vapor diffusion against 1.0 M succinic acid (pH 6.5) in 25 mM Tris pH 7.5, 17 mM NaCl, 5 mM β-mercaptoethanol buffer. Drops contained 2 μl of 2.5 mg/ml ARNT PAS-B/GCN4-TACC3-CT D622A/E629A (molar ratio = 1:2) were mixed with 0.7 μl 1.0 M succinic acid (pH 6.5) (reservoir solution) and 0.3 μl 30% sorbitol (additive). Crystals were observed within 2 days and reached maximum size within 1 week at 20 °C. The crystals exhibited C2 space group symmetry with cell dimensions of a = 116.85 Å, b = 59.82 Å, c = 73.51 Å, b = 97.85° and contained two molecules of ARNT PAS-B, one molecule of GCN4-TACC3-CT D622A/E629A per asymmetric unit, and diffracted to a minimum Bragg spacing (d_{min}) of 3.15 Å when exposed to synchrotron radiation.

The ARNT PAS-B E362R/TACC3 heterodimer protein complex was formed by mixing equal volumes of 300 μM ARNT PAS-B E362R with 300 μM of the TACC3 S-S linked dimer at C621 position. The ARNT PAS-B E362R/TACC3 heterodimer and HIF2α PAS-B R247E protein were further purified separately by size exclusion chromatography on a Superdex 75 column (GE Healthcare) equilibrated with 50 mM Tris (pH 7.5), 17 mM NaCl. The final ARNT PAS-B E362R/TACC3•HIF2α PAS-B R247E sample was formed by mixing equimolar amounts of all proteins and concentrated using an Amicon ultrafiltration device (Millipore) to 8 mg/ml for crystallization. Crystals of ARNT PAS-B E362R /TACC3•HIF2α PAS-B R247E were grown using the sitting-drop vapor-diffusion method from drops containing 0.2 μl protein (8 mg/ml) and 0.2 μl of reservoir solution (1.6 M MgSO$_4$, 0.1 M MES (pH 6.5)) and equilibrated over reservoir solution at 20˚C. Cryoprotection was performed by transferring the crystals to a final solution of 1.6 M MgSO$_4$, 0.1 M MES (pH 6.5), 20% ethylene glycol, and were flash-cooled in liquid nitrogen. The crystals exhibited P4$_1$2$_1$2 space group symmetry with cell dimensions of a = 96.14 Å, c = 182.84 Å, contained two molecules of ARNT PAS-B E362R, two molecules of TACC3 S-S linked dimer, and one molecule of HIF2α PAS-B R247E per asymmetric unit, and diffracted to a minimum Bragg spacing (d$_{min}$) of 3.20 Å when exposed to synchrotron radiation.

Diffraction data for both complexes were collected at the Advanced Photon Source beamline 19-ID, and were indexed, integrated and scaled using the *HKL-3000 (194)* program package. Data collection statistics are provided in Table 2-1. Including data for the ARNT PAS-B/GCN4-TACC3-CT D622A/E629A crystals to a higher

resolution limit did not substantially improve calculated electron density maps and provided poorer refinement results, likely due to anisotropy that in turn limited data completeness in the highest resolution shells.

Phases for ARNT PAS-B/GCN4-TACC3-CT D622A/E629A were obtained via molecular replacement in the program *Phaser (195)* using a search model derived from the previously determined ARNT PAS-B domain crystal structure (PDB ID: 4EQ1). Two copies of ARNT PAS-B were located in the cell and inspection of the electron density maps revealed density corresponding to the coiled-coil domain, which was initially modeled as a pair of idealized poly-alanine helices. Identification of the inter-helical disulfide bond allowed accurate assignment of GCN4-TACC3 density for residues 595 through 626.

Refinement was performed to a resolution of 3.15 Å with non-crystallographic symmetry restraints using the program *Phenix (196)* with a random 10% of all data set aside for an R_{free} calculation. The current model contains two ARNT PAS-B monomers, included are residues 361-446 and 455-464 of monomers A and B; and one TACC3 dimer, included are residues 595-625 of monomer C, and residues 596-626 of monomer D. The R_{work} is 0.247, and the R_{free} is 0.272. A Ramachandran plot generated with *MolProbity (155)* indicated that 95.9% of all protein residues are in the most favored regions, and none are in the disallowed regions. Phases for ARNT PAS-B E362R/TACC3•HIF2α PAS-B R247E crystals were obtained via molecular replacement in the program *Phaser*. Search models for the PAS domains were constructed from the previously determined ARNT PAS-B E362R/HIF2α PAS-B R247E heterodimer

coordinates (PDB ID: 3F1O) by removal of residues 356-361, 448-453 from the ARNT PAS-B E362R, and residues 236-242, 345-349 from the HIF2α PAS-B R247E. The search model for the TACC3 domain was constructed from the coordinates for the GCN4-TACC3-CT D622A/E629A domain by conversion of residues 595-609 to polyalanine. Two copies of the ARNT PAS-B E362R and two copies of the TACC3 domain were located in the cell and inspection of the electron density maps revealed density corresponding to additional helical density at both N- and C-termini of the TACC3. Phases were further improved by density modification with two-fold non-crystallographic symmetry averaging in the program *Parrot (197)* resulting in a figure-of-merit of 0.745. A model containing 92.7% of all residues in the two TACC3 domains was automatically generated in the program *Buccaneer (198)*, although the initial sequence assignment was incorrect. The inter-helical disulfide bond used to confirm the register of the helices was intact in monomers E and F, but was reduced in monomers B and C. Inspection of electron density maps after rigid body refinement of the ARNT PAS-B E362R/TACC3 heterodimer using the program *Phenix* revealed density for an additional PAS domain in the lattice. Identification of this as the HIF2α PAS-B R247E domain was confirmed via the statistics from molecular replacement in *Phaser* and inspection of kicked omit maps following refinement in *Phenix*.

Additional residues for ARNT PAS-B E362R/TACC3•HIF2α PAS-B R247E were manually modeled in the programs *O (199)* and *Coot (200)*. Refinement was performed to a resolution of 3.20 Å using the program *Phenix* with a random 10% of all data set aside for an R_{free} calculation. Non-crystallographic symmetry averaging and

reference model restraints were used in initial rounds of refinement but were removed once the R_{free} dropped below 0.30 and the geometry stabilized. Due to the low resolution of the data, grouped isotropic as well as TLS atomic displacement parameters were refined. Mean atomic displacement parameters were 43.1 - 56.4 Å^2 for the domains of the ARNT PAS-B E362R/TACC3 heterodimer and 83.3 Å^2 for the HIF2α PAS-B R247E domain. The current model contains two ARNT PAS-B E362R monomers, included are residues 361-448 and 452-464 of monomer A, and residues 361-449, 451-464 of monomer D; two TACC3 dimers, included are residues 582-631 of monomers B, E and F, and residues 582-629 of monomer C; one monomer of HIF2α PAS-B R247E, included are residues 239-344; and two sulfate anions. The R_{work} is 0.231, and the R_{free} is 0.270. A Ramachandran plot generated with *Molprobity* indicated that 97.5% of all protein residues are in the most favored regions, and 0.2% (one residue) in the disallowed regions.

Microscale Thermophoresis (MST) - MST experiments were conducted on NanoTemper Microscale Thermophoresis instrument (Monolith NT.115, Nanotemper). ARNT PAS-B E362R was labeled with blue fluorescent dye NT-495-NHS following the Monolith NT Protein Labeling Kit BLUE-NHS protocol (NanoTemper) and buffer exchanged to working buffer 25 mM Tris (pH 7.5), 17 mM NaCl, 0.05% Tween 20 and 5 mM β-mercaptoethanol. Titrations were conducted with 500 nM fluorophore labeled ARNT PAS-B (ARNT PAS-B E362R-fluor) in standard coated capillaries.

Paramagnetic relaxation enhancement (PRE) - Paramagnetic relaxation enhancement was measured by comparing the transverse (T_2) relaxation rates of amide proton in paramagnetic state and diamagnetic state. Paramagnetic labeling with (1-oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) methanethiosulfonate (MTSL) or (1-oxyl-2,2,5,5- tetramethylpyrroline-3-yl) carbamidoethyl methanethiosulfonate (CMTSL) (Toronto Research Chemicals) at specific cysteine site of target protein was conducted by incubating protein and spin label with molar ratio 1:2 in non-reducing condition and rotating overnight at room temperature. Labeled protein was purified by gel filtration through a Superdex 75 column. Diamagnetic controls were generated by adding 10-fold sodium dithionite and purging N_2 for 30 min to quench the spin label. ^{15}N T_2 relaxation rates were determined using a modified ^{15}N/^{1}H TROSY-HSQC experiment with delay times of 0.0065, 0.008, 0.01, 0.013, 0.016, 0.026 seconds. Rates were calculated using NMRViewJ (Version 9.0.0, One Moon Scientific).

HADDOCK modeling - A TACC3 C-terminal peptide (residues 585–631) was modeled as a parallel coiled-coil dimer based on the GCN4-TACC3-CT crystal structure using Rosetta (157). Complexes of ARNT PAS-B/TACC3 (585-631) and HIF2α PAS-B/ TACC3 (585-631) were generated by rigid body docking (HADDOCK 2.0 Web server) using TACC3 (585-631) dimer model, ARNT PAS-B crystal structure (PDB ID: 4EQ1) and HIF2α PAS-B crystal structure (PDB ID: 3F1O, chain A) (56). Active residues used for docking are listed in Table 2-2. Passive residues were automatically defined around the active residues. Docking and refining were done following the default parameters,

including semiflexible simulated annealing of all proteins in the 200 lowest-intermolecular energy solutions and refinement in explicit water. C2 symmetry restraints were turned on to lock TACC3 as a parallel dimer in the docking process.

iii. Methods for Chapter 3

Cell culture and transfection - HEK293T and HeLa cells were maintained in Dulbecco's modified Eagle's medium (Thermo Scientific) supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 U/ml penicillin and 100 μg/ml streptomycin. Human hepatocellular carcinoma Hep3B cells were grown in DMEM/high glucose media (HyClone, # SH30022.01) supplemented with 10% of fetal bovine serum (Atlanta Biologicals), 20 mM HEPES buffer (pH 7.4), 1 mM sodium pyruvate, 100 U/ml penicillin and 100 μg/ml streptomycin (Invitrogen). Hep3B cells at ~80% confluence were incubated with KHS101 for 2 hr prior to incubation under normoxic (~20% O_2) or hypoxic (1% O_2) conditions using a hypoxic incubator (COY Laboratory Products Inc.). Plasmids were transfected using Lipofectamine 2000 (Invitrogen). Cells were harvested in IP lysis buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.5% (v/v) NP-40 and freshly added EDTA-free protease inhibitors (Roche)). Cell lysates were clarified by centrifugation for 10 min at 21,000 g at 4°C and used for immunoprecipitation or Western blot.

Immunoblot, pulldown and immunofluorescence assays - For immunoblots, proteins resolved from SDS-PAGE gels were transferred to PVDF membranes (GE Healthcare) and immunoblotted with these antibodies: anti-ARNT (A-3), anti-TACC3 (T-17), anti-goat IgG-HRP (Santa Cruz Biotechnology); anti-TACC3 (ab56595) (abcam); anti-Myc (9E10), anti-HIF1α (BD Transduction Laboratories); anti-mouse IgG-HRP antibodies (Sigma). For pulldown experiments, 5 μM purified His-ARNT/ARNT2 PAS-B and 8-15 μM TACC3-CT were incubated with 15 μl Ni-NTA beads overnight at 4°C. Eluted protein was resolved in SDS-PAGE and stained with Coomassie Blue stain. Immunofluorescence studies of steady state TACC3 levels used HEK293 cells stained with an anti-TACC3 rabbit polyclonal antibody (Sigma-Aldrich HPA005781; here called TACC3 AB1).

Immunofluorescence - HEK293 cells were grown to 70% confluency and treated with KHS101 compound at concentrations indicated in the figures. Incubation was for 14 hr in a humidified incubator at 37 °C and 5% CO_2. Next, cells were fixed in 4% (v/v) paraformaldehyde/PBS for 15 minutes at room temperature. Cells were washed in PBS and permeabilized with a PBS-based buffer containing 0.3% Triton X-100 and 5% FBS. Incubation was for one hour at room temperature to prevent nonspecific binding of antibodies. Cells were incubated at 4°C overnight with the following primary antibodies: TACC3 AB1 (Rabbit polyclonal HPA005781, Sigma-Aldrich), TACC3 AB2 (Rabbit polyclonal 5640-1, Epitomics). After washing with PBS, cells were incubated for 40 min at room temperature with a Cy3–conjugated rabbit secondary antibody (Jackson

ImmunoResearch Laboratories) and washed with PBS in the presence of DAPI (0.1 ng/ml). Fluorescent cells were visualized using a digital inverted microscope and similar exposure time (EVOS, Advanced Microscopy Group (AMG)). ImageJ quantification of TACC3 staining intensity was normalized to DAPI signal with DMSO.

ARNT PAS-B X-ray crystallography - Single crystals of ARNT PAS-B were grown by hanging drop vapor diffusion against 2 - 2.2 M ammonium sulfate, 3% (wt/vol) PEG 400, 2% (wt/vol) polyethyleneimine, 100 mM Tris (pH 8.5) solutions. Drops contained 5 µl of 500 µM ARNT PAS-B mixed with 5 µl of precipitant solution. Crystals were observed within minutes and reached their maximum sizes within 1 day.

Diffraction data were obtained at the Structural Biology Center at the Advanced Photon Source (Argonne National Laboratory, Argonne, IL). All data were indexed and scaled with HKL2000 (201). Molecular replacement was performed using phaser (202) using the ARNT subunit of the HIF2α/ARNT PAS-B heterodimer as a search model (chain B of PDB 3F1P) (56).Refinement was completed using PHENIX and COOT (203,204). Coordinates have been deposited with the RCSB database and assigned accession number 4EQ1. Cavities were identified using the program VOIDOO using a 1.2 Å radius probe (205).

NMR spectroscopy - NMR experiments for backbone and sidechain chemical shift assignments of ARNT2 PAS-B were carried out at 35°C on Varian Inova 600 and 800 MHz spectrometers using a 400 µM sample of uniformly ^{15}N/^{13}C labeled ARNT2

PAS-B. Backbone assignments were based on data from standard 3D triple resonance experiments: HNCACB, CBCA(CO)NH, HNCO, H(CCO)HN-TOCSY and (H)C(CO)HN-TOCSY (206). Sidechain chemical shift assignments were based on HCCH-TOCSY and simultaneous $^{15}N,^{13}C$-edited NOESY data. Chemical shift assignments of ARNT PAS-B were used as previously established (168). All NMR data were processed with NMRpipe/NMRDraw (207) and analyzed with NMRviewJ (Version 8.2.29, One Moon Scientific). Backbone assignment results were uploaded to TALOS+ web server (208) for chemical-shift based secondary structure analysis. Heat maps of minimum chemical shift perturbations and peak broadening onto the ARNT PAS-B crystal structure and an ARNT2 PAS-B homology model (generated by SWISS-MODEL (182) from the ARNT PAS-B crystal structure) were generated by PyMOL (Version 1.5.0.4, Schrödinger, LLC).

Compound sources - KHS101 was kindly provided by Dr. Peter Schultz (The Scripps Research Institute), synthesized as described (172). KG-548 was purchased from Fluorochem and resupplied from Matrix Scientific; the composition of the entire fragment-based library is described in Appendix table 1-1.

AlphaScreen assay - 400 nM His-ARNT PAS-B and 400 nM GST-TACC3-CT E629A were mixed in a 384-well plate with a total volume of 15 µl and incubated at 4°C for 1 hr. AlphaScreen glutathione donor beads and AlphaLISA Ni chelate acceptor beads (5 µg/ml each; Perkin Elmer) were added under dim green light in 50 mM Tris (pH 7.5),

100 mM NaCl, 0.02% Tween 20 and 1 mM DTT. After incubating in a dark and humidified chamber for 3.5 hr, the plate was read using an Envision plate reader (Perkin Elmer).

KHS101 washout assay - Adult rat hippocampal NPCs were differentiated upon treatment with 5 µM KHS101 or a DMSO negative control as previously described (172). After incubation for certain time, the culture medium was replaced with fresh (compound-free) media. After a total 100 hr incubation period, cells were fixed and stained for the neuronal marker TuJ1 (172). Neuronal differentiation of KHS101-treated rat hippocampal NPCs was analyzed by microscopy, the percentage of TuJ1-positive cells was determined, and data were normalized to the DMSO control values.

TACC3 turnover assay - HEK293T cells were treated with 5 µM KHS101 and 100 µg/ml cycloheximide (CHX) with or without 20 µM MG132. DMSO treatment served as a negative control. Cells were harvested at several time points from 0 to 8 hr post-treatment and prepared for TACC3, ARNT and β-actin immunoblot analyses.

qPCR - Cells were collected with Trizol (Invitrogen) and total RNA was extracted using the RNeasy Mini Kit (Qiagen). Following DNase treatment, cDNA was synthesized from 1 µg of RNA using SuperScript II Reverse Transcriptase (Invitrogen) in final volume of 50 µl. Real-time PCR was performed (with target gene primers listed in Supporting Information) on 1.25 µl of cDNA in triplicate using Power SYBR Green PCR

Master Mix and the 7900HT Fast Real-Time PCR System (Applied Biosystems). Target gene expression levels were normalized to the expression level of cyclophilin B in the same sample. Data were analyzed using the comparative C_T method ($2^{-\Delta\Delta C_T}$) (209) and normalized to normoxia/1% DMSO conditions.

qPCR primers - The following primer sets were used to amplify target genes: Cyclophilin B, forward 5'- tgccatcgccaaggagtag - 3' and reverse 5'-tgcacagacggtcactcaaa - 3'; HIF2α, forward 5'- gcgacaatgacagctgacaa -3' and reverse 5'- cagcatcccgggacttct -3'; EPO, forward 5'- gaggccgagaatatcacgacggg -3' and reverse 5'- tgcccgacctccatcctcttccag - 3'; GLUT1, forward 5'- cttttctgttggggggcatgat -3' and reverse 5'- ccgcagtacacaccgatgat - 3'; HIF1α, forward 5'- tgccacatcatcaccatatagaga -3' and reverse 5'- tccttttcctgctctgtttgg - 3'; PGK1, forward 5'- ttaaagggaagcgggtcgtta -3' and reverse 5'- tccattgtccaagcagaatttga - 3'.

iv. Methods for Chapter 4

Isothermal calorimetry - Thermodynamic parameters of small-molecule binding were determined using a MicroCal VP-ITC calorimeter. Protein solutions were extensively dialyzed against buffer (50 mM Tris (pH 7.5), 20 mM NaCl and 5 mM β-mercaptoethanol), which was subsequently used to prepare a matched compound solution by dilution from a 10 mM compound stock in 100% DMSO. ITC data were collected in 5.0% DMSO to improve compound solubility. Each isotherm was recorded by injecting

ARNT PAS-B 1-2 mM protein (syringe) into 50-100 μM compound solutions (cell), accounting for dilution heats by subtracting data from a control titration of ARNT PAS-B into a matched DMSO solution.

Mammalian Cytotoxicity Assay - HEK293T, HepG2 or HeLa cells were grown in Triple flasks (NUNC) to ~95% confluence (TrypLE Phenol Red free) and resuspended for dispensing at 125,000 cells/ml of DMEM, 10% FBS/Pen/Strep/L-Glutamine (using the TAP Compact SelecT automated cell culture system).

Day 1: Plate cells at 5000 per well in 40 μl media (DMEM/10% FBS/Pen/Strep/ L-Glutamine) using Corning 8867BC 384 well plates; incubate in standard TC conditions (5% CO_2; 95% humidity, 37 °C) for 24 hours (Compact SelecT).

Day 2: Add 100 nl compound per well at dose into 40 ul assay volume using a pin tool (CyBi Well). Pin 100 nl cytotoxic compounds, mitoxantrone to positive control wells to a final concentration of 10 μM (100 nl 6.4 mM DMSO stock). Incubate for 72 hours at 37 °C in Liconic incubator, 95% humidity 5% CO_2.

Day 4: Remove plate from incubator, cool for 15 minutes to room temperature; add 20 μl 50% Promega CellTiter-Glo (diluted 1:1 with PBS, pH 7.4) with Thermo Combi. Incubate at RT for 5 minutes. Read plates on a Perkin-Elmer EnVision plate reader with standard luminescence settings for 0.1 sec per well.

U2OS and HeLa HRE-luciferase assay - The 3×HRE-Luc cell line was propagated in DMEM media supplemented with 10% fetal bovine serum, 1% penicillin/

145

streptomycin/glutamine and 200 μg/ml G418 at 37 °C in CO_2 incubators with 5% CO_2, 21% O_2, and 95% humidity.

Day 1 (Cell plating): 3×HRE-Luc cells were harvested and resuspended in DMEM without phenol red with 10% fetal bovine serum (Invitrogen, 16140089), 1% penicillin/ streptomycin/glutamine. 3×HRE-Luc cells (from an initial cell suspension of 120,000 cells/ml) were dispensed using a MultiDrop Combi/Standard tube dispensing cassette (Thermo Scientific) in white bottom 384 well assay plates (Corning, Cat.No. 8867BC) at a final density of 6,000 cells per well in final volume of 50 μl. The assay plates were incubated for 24 hours at 37 °C in the Liconic CO_2 incubator (Liconic Instruments) calibrated at 5% CO_2, 21% O_2, and 95% humidity.

Day 2 (Compound pinning into assay plate): The dose response compound plate with the positive control were pinned using 384 well pin tool (250 nl) on pin table (Walkup Cybi Well) and transferred to assay plate. Pins were washed with methanol and DMSO between each pinning. The assay plates treated with compounds were moved to Hypoxia incubator (1% O_2) to be incubated for an additional 24 hours.

Day 3 (Reading luminescence from assay plates with Envision): Each assay plate was pulled out of the incubator and cooled down at room temperature for 10 minutes. 20 μl/well (384 well) of Steady-Glo luciferase reagent (diluted in H2O) (Promega, E2550) was dispensed using the MultiDrop Combi dispensing cassette from Thermo Scientific. The assay plate incubated at room temperature for 10 minutes to allow a complete cellular lysis. Luminescence was measured (0.2 second/well) using the ultra sensitive luminescence detector (384-well aperture, 0.5 mm height) in each well using the Envision

146

plate reader (Perkin Elmer)(Corning plate setting).

$qPCR$ - Hep3B cells were cultured in EMEM complete media (EMEM: cat. no. ATCC 30-2003, 10% fetal bovine serum, 1% penicillin/streptomycin/glutamine) at 37 °C in CO_2 incubators with 5% CO_2, 21% O_2, and 95% humidity.

Day 1 (Cell Plating): 30 µl Hep3B cells were added to each well at 4,000/well in a Corning White, 384-well, tissue culture plate using a standard Combi liquid dispenser (Thermo). The plated cells were then incubated overnight in a tissue culture incubator (Thermo) at at 37 °C in CO_2 incubators with 5% CO_2, 21% O_2, and 95% humidity.

Day 2 (Compound Pinning): The following day, 100 nl of compound was added to each well using the Cy-Bi Well pin transfer array. After addition of compound, the plates were moved to Hypoxia incubator (1% O_2) to be incubated overnight.

Day 3 (cDNA Synthesis): Cells were harvested following the Cells-to-CT kits (Ambion cat.no. 4391851C) protocol. Reverse transcription (RT) was done following the Cells-to-CT RT kit (Ambion cat. no. #4402958) instructions with 10 µl total reaction volume. cDNA is stored at -80°C until ready for qPCR analysis.

Day 4 (qPCR analysis): Preparing qPCR master mix following the Light Cycler Probes master (Roche cat. no.4887301001) instructions (Table 6-1).

Table 6-1. qPCR master mix setup.

Component	Amount per reaction
2× Roche Master Mix (Probes Master)	2.5 µl
*20× FAM Taqman probe/primers set	0.125 µl
*20× VIC Taqman probe/primers set	0.125 µl
PCR H_2O	1.25 µl

4 μl/well of PCR master mix and 1 μl/well of RT DNA was dispensed in PCR plate (Roche Light Cycler 480 MultiWell Plate 384, Cat# 04 729 749 001) using the Multidrop Combi-nl (Thermo Scientific). PCR is performed using ThermoCycler (Roche Light Cycler 480 II) with Macro Probe Protocol with setup in Table 6-2:

Table 6-2. qPCR amplification setup.

Step	Temperature	Time	# Cycles
1	95°C	10 minutes	1
2	95°C	10 seconds	55
3	60°C	30 seconds	
4	40°C	30 seconds	1

Target gene expression levels were normalized to the expression level of GADPH in the same sample. Data were analyzed using the comparative C_T method ($2^{-\Delta\Delta C_T}$) (209) and normalized to normoxia/1% DMSO conditions.

AlphaScreen assay for high throughput screening - written in a protocol format:

Materials:

1. Ni^{2+} acceptor beads, 5 mg/ml (Perkin-Elmer. Inc., Nickel Chelate AlphaLISA® Acceptor Beads 5 mg, Product number: AL108M)

2. GST donor beads, 5 mg/ml (Perkin-Elmer. Inc., AlphaScreen® Glutathione Donor Beads, Product number: 6765301 for 5 mg pack, Product number: 6765302 for 25mg pack)

3. Alpha buffer: 50 mM Tris, pH 7.5; 100 mM NaCl; 10% glycerol; 1 mM DTT (freshly added); 0.02 % Tween 20 (freshly added); 0.1% BSA (freshly added)

4. Plates used: OptiPlate-384, White Opaque 384-well MicroPlate, Product number: 6007290 for 50/box pack, Product number: 6007299 for 200/box pack

5. His-ARNT PAS-B (In Alpha buffer without BSA)

6. GST-TACC3 E629A dimer (In Alpha buffer without BSA)

7. GST-ARNT PASB-His (In Alpha buffer without BSA)

Procedure:

For 384 well format:

1. Mix His-ARNT PAS-B and GST-TACC3 E629A dimer, 15 μl/well with final His-ARNT PAS-B = 668 nM and GST-TACC3 E629A dimer = 334 nM.
   (Note: For positive controls, add KG-548 later, or use 668 nM His-ARNT PAS-B only; for counter screens, use 668 nM GST-ARNT PAS-B-His)

2. Add protein mixture to 384 well plates, 15 μl/well. Spin down gently (~30 ×g in for 1 min.

3. Add DMSO or DMSO dissolved compound to each well at 0.2 μl/well (1.3% of the total volume) and spin again. (Note: For negative controls, add DMSO only)

4. Cover plates with lid or foil and incubate plates on ice or 4 °C for 1 hour.

****

ATTENTION:

The following steps must be done under **DARK** or **DIM GREEN SAFETY LIGHT** (It is extremely important **NOT TO** expose alpha beads under regular light!)

5. Dilute Alpha donor beads and acceptor beads to 12.5 μg/ml.

   (Note: for one 384-well plate, total volume for beads = 10 μl/well × 384 wells = 3840 μl, for more calculations see Table 6-1)

Table 6-3. AlphaScreen beads dilution spreadsheet.

# of plates	Ni acceptor beads (μl)	GST donor beads (μl)	Alpha buffer (μl)	Total Volume (μl)
1 plate	9.6	9.6	3820.8	3840
3 plate	28.8	28.8	11462.4	11520
30 plate	288	288	114624	115200

6. Add diluted beads to each well at 10 μl/well to make total volume = 25 μl and spin again in dark.

   (Note: after adding diluted alpha beads, [His-ARNT PAS-B]$_{final}$ or [GST-ARNT PAS-B-His]$_{final}$ = 400 nM, [GST-TACC3-E629A dimer]$_{final}$ = 200 nM, [Ni^{2+} donor beads]$_{final}$ and [GST acceptor beads]$_{final}$ = 5 μg/ml)

7. Wrap plates with 2 layers of aluminum foil to avoid light.

8. Put plates into a humidified chamber (covered Styrofoam box with wet paper towel at the bottom) and incubate for 3.5 hours.

9. Read plates at an EnVision Multilabel Plate Reader (Perkin-Elmer. Inc) under **dim green light.**

All results generated from this screen were updated to PubChem (BioAssay ID: AID 623900) and can be found at:

http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=623900.

BIBLIOGRAPHY

1.   Taylor, B. L., and Zhulin, I. B. (1999) PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiology and molecular biology reviews : MMBR* **63**, 479-506

2.   Bersten, D. C., Sullivan, A. E., Peet, D. J., and Whitelaw, M. L. (2013) bHLH-PAS proteins in cancer. *Nature reviews. Cancer* **13**, 827-841

3.   Moglich, A., Ayers, R. A., and Moffat, K. (2009) Structure and signaling mechanism of Per-ARNT-Sim domains. *Structure* **17**, 1282-1294

4.   Partch, C. L., and Gardner, K. H. (2010) Coactivator recruitment: a new role for PAS domains in transcriptional regulation by the bHLH-PAS family. *Journal of cellular physiology* **223**, 553-557

5.   Kobayashi, A., Numayama-Tsuruta, K., Sogawa, K., and Fujii-Kuriyama, Y. (1997) CBP/p300 functions as a possible transcriptional coactivator of Ah receptor nuclear translocator (Arnt). *Journal of biochemistry* **122**, 703-710

6.   Voronova, A., and Baltimore, D. (1990) Mutations that disrupt DNA binding and dimer formation in the E47 helix-loop-helix protein map to distinct domains. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 4722-4726

7.   Kewley, R. J., Whitelaw, M. L., and Chapman-Smith, A. (2004) The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *The international journal of biochemistry & cell biology* **36**, 189-204

8.   Ma, P. C., Rould, M. A., Weintraub, H., and Pabo, C. O. (1994) Crystal structure of MyoD bHLH domain-DNA complex: perspectives on DNA recognition and implications for transcriptional activation. *Cell* **77**, 451-459

9.   Nair, S. K., and Burley, S. K. (2003) X-ray structures of Myc-Max and Mad-Max recognizing DNA. Molecular bases of regulation by proto-oncogenic transcription factors. *Cell* **112**, 193-205

10.   Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., and Punta, M. (2014) Pfam: the protein families database. *Nucleic acids research* **42**, D222-230

11.   Sonnhammer, E. L., Eddy, S. R., and Durbin, R. (1997) Pfam: a comprehensive database of protein domain families based on seed alignments. *Proteins* **28**, 405-420

12.   Atchley, W. R., and Fitch, W. M. (1997) A natural classification of the basic helix-loop-helix class of transcription factors. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 5172-5176

13.   Ledent, V., and Vervoort, M. (2001) The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome research* **11**, 754-770

14.   Citri, Y., Colot, H. V., Jacquier, A. C., Yu, Q., Hall, J. C., Baltimore, D., and Rosbash, M. (1987) A family of unusually spliced biologically active transcripts encoded by a Drosophila clock gene. *Nature* **326**, 42-47

152

15. Reddy, P., Jacquier, A. C., Abovich, N., Petersen, G., and Rosbash, M. (1986) The period clock locus of D. melanogaster codes for a proteoglycan. *Cell* **46**, 53-61

16. Jackson, F. R., Bargiello, T. A., Yun, S. H., and Young, M. W. (1986) Product of per locus of Drosophila shares homology with proteoglycans. *Nature* **320**, 185-188

17. Hoffman, E. C., Reyes, H., Chu, F. F., Sander, F., Conley, L. H., Brooks, B. A., and Hankinson, O. (1991) Cloning of a factor required for activity of the Ah (dioxin) receptor. *Science* **252**, 954-958

18. Nambu, J. R., Lewis, J. O., Wharton, K. A., Jr., and Crews, S. T. (1991) The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell* **67**, 1157-1167

19. McIntosh, B. E., Hogenesch, J. B., and Bradfield, C. A. (2010) Mammalian Per-Arnt-Sim proteins in environmental adaptation. *Annual review of physiology* **72**, 625-645

20. Borgstahl, G. E., Williams, D. R., and Getzoff, E. D. (1995) 1.4 A structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore. *Biochemistry* **34**, 6278-6287

21. Guo, Y., Partch, C. L., Key, J., Card, P. B., Pashkov, V., Patel, A., Bruick, R. K., Wurdak, H., and Gardner, K. H. (2013) Regulating the ARNT/TACC3 axis: multiple approaches to manipulating protein/protein interactions with small molecules. *ACS chemical biology* **8**, 626-635

22. Ho, Y. S., Burden, L. M., and Hurley, J. H. (2000) Structure of the GAF domain, a ubiquitous signaling motif and a new class of cyclic GMP receptor. *The EMBO journal* **19**, 5288-5299

23. Podust, L. M., Ioanoviciu, A., and Ortiz de Montellano, P. R. (2008) 2.3 A X-ray structure of the heme-bound GAF domain of sensory histidine kinase DosT of Mycobacterium tuberculosis. *Biochemistry* **47**, 12523-12531

24. Crosson, S., Rajagopal, S., and Moffat, K. (2003) The LOV domain family: photoresponsive signaling modules coupled to diverse output domains. *Biochemistry* **42**, 2-10

25. de Philip, P., Batut, J., and Boistard, P. (1990) Rhizobium meliloti Fix L is an oxygen sensor and regulates R. meliloti nifA and fixK genes differently in Escherichia coli. *Journal of bacteriology* **172**, 4255-4262

26. Gilles-Gonzalez, M. A., Ditta, G. S., and Helinski, D. R. (1991) A haemoprotein with kinase activity encoded by the oxygen sensor of Rhizobium meliloti. *Nature* **350**, 170-172

27. Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., and Chan, M. K. (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 15177-15182

28. Delgado-Nixon, V. M., Gonzalez, G., and Gilles-Gonzalez, M. A. (2000) Dos, a heme-binding PAS protein from Escherichia coli, is a direct oxygen sensor. *Biochemistry* **39**, 2685-2691

29. Park, H., Suquet, C., Satterlee, J. D., and Kang, C. (2004) Insights into signal transduction involving PAS domain oxygen-sensing heme proteins from the X-ray crystal structure of Escherichia coli Dos heme domain (Ec DosH). *Biochemistry* **43**, 2738-2746

30. Henry, J. T., and Crosson, S. (2011) Ligand-binding PAS domains in a genomic, cellular, and structural context. *Annual review of microbiology* **65**, 261-286

31. Taylor, B. L., Zhulin, I. B., and Johnson, M. S. (1999) Aerotaxis and other energy-sensing behavior in bacteria. *Annual review of microbiology* **53**, 103-128

32. Moglich, A., and Moffat, K. (2007) Structural basis for light-dependent signaling in the dimeric LOV domain of the photosensor YtvA. *Journal of molecular biology* **373**, 112-126

33. Herrou, J., and Crosson, S. (2011) Function, structure and mechanism of bacterial photosensory LOV proteins. *Nature reviews. Microbiology* **9**, 713-723

34. Harper, S. M., Neil, L. C., and Gardner, K. H. (2003) Structural basis of a phototropin light switch. *Science* **301**, 1541-1544

35. Nash, A. I., McNulty, R., Shillito, M. E., Swartz, T. E., Bogomolni, R. A., Luecke, H., and Gardner, K. H. (2011) Structural basis of photosensitivity in a bacterial light-oxygen-voltage/helix-turn-helix (LOV-HTH) DNA-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 9449-9454

36. Zoltowski, B. D., Motta-Mena, L. B., and Gardner, K. H. (2013) Blue light-induced dimerization of a bacterial LOV-HTH DNA-binding protein. *Biochemistry* **52**, 6653-6661

37. Wang, X., Chen, X., and Yang, Y. (2012) Spatiotemporal control of gene expression by a light-switchable transgene system. *Nature methods* **9**, 266-269

38. Motta-Mena, L. B., Reade, A., Mallory, M. J., Glantz, S., Weiner, O. D., Lynch, K. W., and Gardner, K. H. (2014) An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature chemical biology* **10**, 196-202

39. Kramer, J., Fischer, J. D., Zientz, E., Vijayan, V., Griesinger, C., Lupas, A., and Unden, G. (2007) Citrate sensing by the C4-dicarboxylate/citrate sensor kinase DcuS of Escherichia coli: binding site and conversion of DcuS to a C4-dicarboxylate- or citrate-specific sensor. *Journal of bacteriology* **189**, 4290-4298

40. Pappalardo, L., Janausch, I. G., Vijayan, V., Zientz, E., Junker, J., Peti, W., Zweckstetter, M., Unden, G., and Griesinger, C. (2003) The NMR structure of the sensory domain of the membranous two-component fumarate sensor (histidine protein kinase) DcuS of Escherichia coli. *The Journal of biological chemistry* **278**, 39185-39188

41. Reinelt, S., Hofmann, E., Gerharz, T., Bott, M., and Madden, D. R. (2003) The structure of the periplasmic ligand-binding domain of the sensor kinase CitA reveals the first extracellular PAS domain. *The Journal of biological chemistry* **278**, 39189-39196

42. Cheung, J., Bingman, C. A., Reyngold, M., Hendrickson, W. A., and Waldburger, C. D. (2008) Crystal structure of a functional dimer of the PhoQ sensor domain. *The Journal of biological chemistry* **283**, 13762-13770

43. Cho, U. S., Bader, M. W., Amaya, M. F., Daley, M. E., Klevit, R. E., Miller, S. I., and Xu, W. (2006) Metal bridges between the PhoQ sensor domain and the membrane regulate transmembrane signaling. *Journal of molecular biology* **356**, 1193-1206

44. Vescovi, E. G., Ayala, Y. M., Di Cera, E., and Groisman, E. A. (1997) Characterization of the bacterial sensor protein PhoQ. Evidence for distinct binding sites for Mg2+ and Ca2+. *The Journal of biological chemistry* **272**, 1440-1443

45. Moglich, A., Ayers, R. A., and Moffat, K. (2009) Design and signaling mechanism of light-regulated histidine kinases. *Journal of molecular biology* **385**, 1433-1444

46. Amezcua, C. A., Harper, S. M., Rutter, J., and Gardner, K. H. (2002) Structure and interactions of PAS kinase N-terminal PAS domain: model for intramolecular kinase regulation. *Structure* **10**, 1349-1361

47. Rutter, J., Michnoff, C. H., Harper, S. M., Gardner, K. H., and McKnight, S. L. (2001) PAS kinase: an evolutionarily conserved PAS domain-regulated serine/threonine kinase. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 8991-8996

48. Chapman-Smith, A., Lutwyche, J. K., and Whitelaw, M. L. (2004) Contribution of the Per/Arnt/Sim (PAS) domains to DNA binding by the basic helix-loop-helix PAS transcriptional regulators. *The Journal of biological chemistry* **279**, 5353-5362

49. Yang, J., Zhang, L., Erbel, P. J., Gardner, K. H., Ding, K., Garcia, J. A., and Bruick, R. K. (2005) Functions of the Per/ARNT/Sim domains of the hypoxia-inducible factor. *The Journal of biological chemistry* **280**, 36047-36054

50. Dioum, E. M., Rutter, J., Tuckerman, J. R., Gonzalez, G., Gilles-Gonzalez, M. A., and McKnight, S. L. (2002) NPAS2: a gas-responsive transcription factor. *Science* **298**, 2385-2387

51. Ishida, M., Ueha, T., and Sagami, I. (2008) Effects of mutations in the heme domain on the transcriptional activity and DNA-binding activity of NPAS2. *Biochemical and biophysical research communications* **368**, 292-297

52. Carver, L. A., LaPres, J. J., Jain, S., Dunham, E. E., and Bradfield, C. A. (1998) Characterization of the Ah receptor-associated protein, ARA9. *The Journal of biological chemistry* **273**, 33580-33587

53. Cox, M. B., and Miller, C. A., 3rd. (2004) Cooperation of heat shock protein 90 and p23 in aryl hydrocarbon receptor signaling. *Cell stress & chaperones* **9**, 4-20

54. Petrulis, J. R., and Perdew, G. H. (2002) The role of chaperone proteins in the aryl hydrocarbon receptor core complex. *Chemico-biological interactions* **141**, 25-40

55. Cardoso, R., Love, R., Nilsson, C. L., Bergqvist, S., Nowlin, D., Yan, J., Liu, K. K., Zhu, J., Chen, P., Deng, Y. L., Dyson, H. J., Greig, M. J., and Brooun, A. (2012) Identification of Cys255 in HIF-1alpha as a novel site for development of covalent inhibitors of HIF-1alpha/ARNT PasB domain protein-protein interaction. *Protein science : a publication of the Protein Society* **21**, 1885-1896

56. Scheuermann, T. H., Tomchick, D. R., Machius, M., Guo, Y., Bruick, R. K., and Gardner, K. H. (2009) Artificial ligand binding within the HIF2alpha PAS-B domain of the HIF2 transcription factor. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 450-455

57. Rogers, J. L., Bayeh, L., Scheuermann, T. H., Longgood, J., Key, J., Naidoo, J., Melito, L., Shokri, C., Frantz, D. E., Bruick, R. K., Gardner, K. H., MacMillan, J. B., and Tambar, U. K. (2013) Development of inhibitors of the PAS-B domain of the HIF-2alpha transcription factor. *Journal of medicinal chemistry* **56**, 1739-1747

58. Scheuermann, T. H., Li, Q., Ma, H. W., Key, J., Zhang, L., Chen, R., Garcia, J. A., Naidoo, J., Longgood, J., Frantz, D. E., Tambar, U. K., Gardner, K. H., and Bruick, R. K. (2013) Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nature chemical biology* **9**, 271-276

59. Bruick, R. K. (2003) Oxygen sensing in the hypoxic response pathway: regulation of the hypoxia-inducible transcription factor. *Genes & development* **17**, 2614-2623

60. Ema, M., Morita, M., Ikawa, S., Tanaka, M., Matsuda, Y., Gotoh, O., Saijoh, Y., Fujii, H., Hamada, H., Kikuchi, Y., and Fujii-Kuriyama, Y. (1996) Two new members of the murine Sim gene family are transcriptional repressors and show different expression patterns during mouse embryogenesis. *Molecular and cellular biology* **16**, 5865-5875

61. Li, B., Qiu, B., Lee, D. S., Walton, Z. E., Ochocki, J. D., Mathew, L. K., Mancuso, A., Gade, T. P., Keith, B., Nissim, I., and Simon, M. C. (2014) Fructose-1,6-bisphosphatase opposes renal carcinoma progression. *Nature*

62. Denison, M. S., Soshilov, A. A., He, G., DeGroot, D. E., and Zhao, B. (2011) Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicological sciences : an official journal of the Society of Toxicology* **124**, 1-22

63. Baba, T., Mimura, J., Gradin, K., Kuroiwa, A., Watanabe, T., Matsuda, Y., Inazawa, J., Sogawa, K., and Fujii-Kuriyama, Y. (2001) Structure and expression of the Ah receptor repressor gene. *The Journal of biological chemistry* **276**, 33101-33110

64. Nguyen, L. P., and Bradfield, C. A. (2008) The search for endogenous activators of the aryl hydrocarbon receptor. *Chemical research in toxicology* **21**, 102-116

65. Denissenko, M. F., Pao, A., Tang, M., and Pfeifer, G. P. (1996) Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* **274**, 430-432

66. Shimizu, Y., Nakatsuru, Y., Ichinose, M., Takahashi, Y., Kume, H., Mimura, J., Fujii-Kuriyama, Y., and Ishikawa, T. (2000) Benzo[a]pyrene carcinogenicity is lost in mice lacking the aryl hydrocarbon receptor. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 779-782

67. Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr., and Kinzler, K. W. (2013) Cancer genome landscapes. *Science* **339**, 1546-1558

68.   Opitz, C. A., Litzenburger, U. M., Sahm, F., Ott, M., Tritschler, I., Trump, S., Schumacher, T., Jestaedt, L., Schrenk, D., Weller, M., Jugold, M., Guillemin, G. J., Miller, C. L., Lutz, C., Radlwimmer, B., Lehmann, I., von Deimling, A., Wick, W., and Platten, M. (2011) An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. *Nature* **478**, 197-203

69.   Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., Wilson, C. J., Lehar, J., Kryukov, G. V., Sonkin, D., Reddy, A., Liu, M., Murray, L., Berger, M. F., Monahan, J. E., Morais, P., Meltzer, J., Korejwa, A., Jane-Valbuena, J., Mapa, F. A., Thibault, J., Bric-Furlong, E., Raman, P., Shipway, A., Engels, I. H., Cheng, J., Yu, G. K., Yu, J., Aspesi, P., Jr., de Silva, M., Jagtap, K., Jones, M. D., Wang, L., Hatton, C., Palescandolo, E., Gupta, S., Mahan, S., Sougnez, C., Onofrio, R. C., Liefeld, T., MacConaill, L., Winckler, W., Reich, M., Li, N., Mesirov, J. P., Gabriel, S. B., Getz, G., Ardlie, K., Chan, V., Myer, V. E., Weber, B. L., Porter, J., Warmuth, M., Finan, P., Harris, J. L., Meyerson, M., Golub, T. R., Morrissey, M. P., Sellers, W. R., Schlegel, R., and Garraway, L. A. (2012) The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607

70.   Kawajiri, K., Kobayashi, Y., Ohtake, F., Ikuta, T., Matsushima, Y., Mimura, J., Pettersson, S., Pollenz, R. S., Sakaki, T., Hirokawa, T., Akiyama, T., Kurosumi, M., Poellinger, L., Kato, S., and Fujii-Kuriyama, Y. (2009) Aryl hydrocarbon receptor suppresses intestinal carcinogenesis in ApcMin/+ mice with natural ligands. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 13481-13486

71.   Wang, G. L., Jiang, B. H., Rue, E. A., and Semenza, G. L. (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 5510-5514

72.   Wang, G. L., and Semenza, G. L. (1995) Purification and characterization of hypoxia-inducible factor 1. *The Journal of biological chemistry* **270**, 1230-1237

73.   Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J. M., Lane, W. S., and Kaelin, W. G., Jr. (2001) HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science* **292**, 464-468

74.   Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., von Kriegsheim, A., Hebestreit, H. F., Mukherji, M., Schofield, C. J., Maxwell, P. H., Pugh, C. W., and Ratcliffe, P. J. (2001) Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science* **292**, 468-472

75.   Lisztwan, J., Imbert, G., Wirbelauer, C., Gstaiger, M., and Krek, W. (1999) The von Hippel-Lindau tumor suppressor protein is a component of an E3 ubiquitin-protein ligase activity. *Genes & development* **13**, 1822-1833

76.   Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., Wykoff, C. C., Pugh, C. W., Maher, E. R., and Ratcliffe, P. J.

(1999) The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature* **399**, 271-275

77.   Lando, D., Peet, D. J., Whelan, D. A., Gorman, J. J., and Whitelaw, M. L. (2002) Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. *Science* **295**, 858-861

78.   Sang, N., Fang, J., Srinivas, V., Leshchinsky, I., and Caro, J. (2002) Carboxyl-terminal transactivation activity of hypoxia-inducible factor 1 alpha is governed by a von Hippel-Lindau protein-independent, hydroxylation-regulated association with p300/CBP. *Molecular and cellular biology* **22**, 2984-2992

79.   Wang, G. L., and Semenza, G. L. (1993) Desferrioxamine induces erythropoietin gene expression and hypoxia-inducible factor 1 DNA-binding activity: implications for models of hypoxia signal transduction. *Blood* **82**, 3610-3615

80.   Yuan, Y., Beitner-Johnson, D., and Millhorn, D. E. (2001) Hypoxia-inducible factor 2alpha binds to cobalt in vitro. *Biochemical and biophysical research communications* **288**, 849-854

81.   Kaelin, W. G., Jr., and McKnight, S. L. (2013) Influence of metabolism on epigenetics and disease. *Cell* **153**, 56-69

82.   King, A., Selak, M. A., and Gottlieb, E. (2006) Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. *Oncogene* **25**, 4675-4682

83.   Pollard, P. J., Wortham, N. C., and Tomlinson, I. P. (2003) The TCA cycle and tumorigenesis: the examples of fumarate hydratase and succinate dehydrogenase. *Annals of medicine* **35**, 632-639

84.   Isaacs, J. S., Jung, Y. J., Mole, D. R., Lee, S., Torres-Cabala, C., Chung, Y. L., Merino, M., Trepel, J., Zbar, B., Toro, J., Ratcliffe, P. J., Linehan, W. M., and Neckers, L. (2005) HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: novel role of fumarate in regulation of HIF stability. *Cancer cell* **8**, 143-153

85.   Koivunen, P., Lee, S., Duncan, C. G., Lopez, G., Lu, G., Ramkissoon, S., Losman, J. A., Joensuu, P., Bergmann, U., Gross, S., Travins, J., Weiss, S., Looper, R., Ligon, K. L., Verhaak, R. G., Yan, H., and Kaelin, W. G., Jr. (2012) Transformation by the (R)-enantiomer of 2-hydroxyglutarate linked to EGLN activation. *Nature* **483**, 484-488

86.   Losman, J. A., and Kaelin, W. G., Jr. (2013) What a difference a hydroxyl makes: mutant IDH, (R)-2-hydroxyglutarate, and cancer. *Genes & development* **27**, 836-852

87.   Selak, M. A., Armour, S. M., MacKenzie, E. D., Boulahbel, H., Watson, D. G., Mansfield, K. D., Pan, Y., Simon, M. C., Thompson, C. B., and Gottlieb, E. (2005) Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-alpha prolyl hydroxylase. *Cancer cell* **7**, 77-85

88.   Semenza, G. L. (2003) Targeting HIF-1 for cancer therapy. *Nature reviews. Cancer* **3**, 721-732

89. Ryan, H. E., Lo, J., and Johnson, R. S. (1998) HIF-1 alpha is required for solid tumor formation and embryonic vascularization. *The EMBO journal* **17**, 3005-3015

90. Doedens, A. L., Stockmann, C., Rubinstein, M. P., Liao, D., Zhang, N., DeNardo, D. G., Coussens, L. M., Karin, M., Goldrath, A. W., and Johnson, R. S. (2010) Macrophage expression of hypoxia-inducible factor-1 alpha suppresses T-cell function and promotes tumor progression. *Cancer research* **70**, 7465-7475

91. Imtiyaz, H. Z., Williams, E. P., Hickey, M. M., Patel, S. A., Durham, A. C., Yuan, L. J., Hammond, R., Gimotty, P. A., Keith, B., and Simon, M. C. (2010) Hypoxia-inducible factor 2alpha regulates macrophage function in mouse models of acute and tumor inflammation. *The Journal of clinical investigation* **120**, 2699-2714

92. Mazumdar, J., Hickey, M. M., Pant, D. K., Durham, A. C., Sweet-Cordero, A., Vachani, A., Jacks, T., Chodosh, L. A., Kissil, J. L., Simon, M. C., and Keith, B. (2010) HIF-2alpha deletion promotes Kras-driven lung tumor development. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 14182-14187

93. Poon, E., Harris, A. L., and Ashcroft, M. (2009) Targeting the hypoxia-inducible factor (HIF) pathway in cancer. *Expert reviews in molecular medicine* **11**, e26

94. Wang, R., Zhou, S., and Li, S. (2011) Cancer therapeutic agents targeting hypoxia-inducible factor-1. *Current medicinal chemistry* **18**, 3168-3189

95. Freedman, S. J., Sun, Z. Y., Poy, F., Kung, A. L., Livingston, D. M., Wagner, G., and Eck, M. J. (2002) Structural basis for recruitment of CBP/p300 by hypoxia-inducible factor-1 alpha. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 5367-5372

96. Kung, A. L., Zabludoff, S. D., France, D. S., Freedman, S. J., Tanner, E. A., Vieira, A., Cornell-Kennon, S., Lee, J., Wang, B., Wang, J., Memmert, K., Naegeli, H. U., Petersen, F., Eck, M. J., Bair, K. W., Wood, A. W., and Livingston, D. M. (2004) Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer cell* **6**, 33-43

97. Lowrey, P. L., and Takahashi, J. S. (2011) Genetics of circadian rhythms in Mammalian model organisms. *Advances in genetics* **74**, 175-230

98. Shearman, L. P., Sriram, S., Weaver, D. R., Maywood, E. S., Chaves, I., Zheng, B., Kume, K., Lee, C. C., van der Horst, G. T., Hastings, M. H., and Reppert, S. M. (2000) Interacting molecular loops in the mammalian circadian clock. *Science* **288**, 1013-1019

99. Shirogane, T., Jin, J., Ang, X. L., and Harper, J. W. (2005) SCFbeta-TRCP controls clock-dependent transcription via casein kinase 1-dependent degradation of the mammalian period-1 (Per1) protein. *The Journal of biological chemistry* **280**, 26863-26872

100. Baggs, J. E., Price, T. S., DiTacchio, L., Panda, S., Fitzgerald, G. A., and Hogenesch, J. B. (2009) Network features of the mammalian circadian clock. *PLoS biology* **7**, e52

101. Ueda, H. R., Hayashi, S., Chen, W., Sano, M., Machida, M., Shigeyoshi, Y., Iino, M., and Hashimoto, S. (2005) System-level identification of transcriptional circuits underlying mammalian circadian clocks. *Nature genetics* **37**, 187-192

102. Preitner, N., Damiola, F., Lopez-Molina, L., Zakany, J., Duboule, D., Albrecht, U., and Schibler, U. (2002) The orphan nuclear receptor REV-ERBalpha controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251-260

103. Akashi, M., and Takumi, T. (2005) The orphan nuclear receptor RORalpha regulates circadian transcription of the mammalian core-clock Bmal1. *Nature structural & molecular biology* **12**, 441-448

104. Sato, T. K., Panda, S., Miraglia, L. J., Reyes, T. M., Rudic, R. D., McNamara, P., Naik, K. A., FitzGerald, G. A., Kay, S. A., and Hogenesch, J. B. (2004) A functional genomics strategy reveals Rora as a component of the mammalian circadian clock. *Neuron* **43**, 527-537

105. Rutter, J., Reick, M., Wu, L. C., and McKnight, S. L. (2001) Regulation of clock and NPAS2 DNA binding by the redox state of NAD cofactors. *Science* **293**, 510-514

106. Asher, G., Gatfield, D., Stratmann, M., Reinke, H., Dibner, C., Kreppel, F., Mostoslavsky, R., Alt, F. W., and Schibler, U. (2008) SIRT1 regulates circadian clock gene expression through PER2 deacetylation. *Cell* **134**, 317-328

107. Chen, S. T., Choo, K. B., Hou, M. F., Yeh, K. T., Kuo, S. J., and Chang, J. G. (2005) Deregulated expression of the PER1, PER2 and PER3 genes in breast cancers. *Carcinogenesis* **26**, 1241-1246

108. Fu, L., Pelicano, H., Liu, J., Huang, P., and Lee, C. (2002) The circadian gene Period2 plays an important role in tumor suppression and DNA damage response in vivo. *Cell* **111**, 41-50

109. Gery, S., Komatsu, N., Baldjyan, L., Yu, A., Koo, D., and Koeffler, H. P. (2006) The circadian gene per1 plays an important role in cell growth and DNA damage control in human cancer cells. *Molecular cell* **22**, 375-382

110. Lee, S., Donehower, L. A., Herron, A. J., Moore, D. D., and Fu, L. (2010) Disrupting circadian homeostasis of sympathetic signaling promotes tumor development in mice. *PloS one* **5**, e10995

111. Chen, Z., Yoo, S. H., Park, Y. S., Kim, K. H., Wei, S., Buhr, E., Ye, Z. Y., Pan, H. L., and Takahashi, J. S. (2012) Identification of diverse modulators of central and peripheral circadian clocks by high-throughput chemical screening. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 101-106

112. Chen, Z., Yoo, S. H., and Takahashi, J. S. (2013) Small molecule modifiers of circadian clocks. *Cellular and molecular life sciences : CMLS* **70**, 2985-2998

113. Hirota, T., Lee, J. W., Lewis, W. G., Zhang, E. E., Breton, G., Liu, X., Garcia, M., Peters, E. C., Etchegaray, J. P., Traver, D., Schultz, P. G., and Kay, S. A. (2010) High-throughput chemical screen identifies a novel potent modulator of cellular circadian rhythms and reveals CKIalpha as a clock regulatory kinase. *PLoS biology* **8**, e1000559

114. Hirota, T., Lee, J. W., St John, P. C., Sawa, M., Iwaisako, K., Noguchi, T., Pongsawakul, P. Y., Sonntag, T., Welsh, D. K., Brenner, D. A., Doyle, F. J., 3rd, Schultz, P. G., and Kay, S. A. (2012) Identification of small molecule activators of cryptochrome. *Science* **337**, 1094-1097

115. Solt, L. A., Wang, Y., Banerjee, S., Hughes, T., Kojetin, D. J., Lundasen, T., Shin, Y., Liu, J., Cameron, M. D., Noel, R., Yoo, S. H., Takahashi, J. S., Butler, A. A., Kamenecka, T. M., and Burris, T. P. (2012) Regulation of circadian behaviour and metabolism by synthetic REV-ERB agonists. *Nature* **485**, 62-68

116. Huang, N., Chelliah, Y., Shan, Y., Taylor, C. A., Yoo, S. H., Partch, C., Green, C. B., Zhang, H., and Takahashi, J. S. (2012) Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. *Science* **337**, 189-194

117. Hosoda, H., Kato, K., Asano, H., Ito, M., Kato, H., Iwamoto, T., Suzuki, A., Masushige, S., and Kida, S. (2009) CBP/p300 is a cell type-specific modulator of CLOCK/BMAL1-mediated transcription. *Molecular brain* **2**, 34

118. Ruas, J. L., Poellinger, L., and Pereira, T. (2005) Role of CBP in regulating HIF-1-mediated activation of transcription. *Journal of cell science* **118**, 301-311

119. Sadek, C. M., Jalaguier, S., Feeney, E. P., Aitola, M., Damdimopoulos, A. E., Pelto-Huikko, M., and Gustafsson, J. A. (2000) Isolation and characterization of AINT: a novel ARNT interacting protein expressed during murine embryonic development. *Mechanisms of development* **97**, 13-26

120. Kim, J. H., and Stallcup, M. R. (2004) Role of the coiled-coil coactivator (CoCoA) in aryl hydrocarbon receptor-mediated transcription. *The Journal of biological chemistry* **279**, 49842-49848

121. Beischlag, T. V., Taylor, R. T., Rose, D. W., Yoon, D., Chen, Y., Lee, W. H., Rosenfeld, M. G., and Hankinson, O. (2004) Recruitment of thyroid hormone receptor/retinoblastoma-interacting protein 230 by the aryl hydrocarbon receptor nuclear translocator is required for the transcriptional response to both dioxin and hypoxia. *The Journal of biological chemistry* **279**, 54620-54628

122. Partch, C. L., Card, P. B., Amezcua, C. A., and Gardner, K. H. (2009) Molecular basis of coiled coil coactivator recruitment by the aryl hydrocarbon receptor nuclear translocator (ARNT). *The Journal of biological chemistry* **284**, 15184-15192

123. Chen, H. M., Schmeichel, K. L., Mian, I. S., Lelievre, S., Petersen, O. W., and Bissell, M. J. (2000) AZU-1: a candidate breast tumor suppressor and biomarker for tumor progression. *Molecular biology of the cell* **11**, 1357-1367

124. Gergely, F., Kidd, D., Jeffers, K., Wakefield, J. G., and Raff, J. W. (2000) D-TACC: a novel centrosomal protein required for normal spindle function in the early Drosophila embryo. *The EMBO journal* **19**, 241-252

125. Le Bot, N., Tsai, M. C., Andrews, R. K., and Ahringer, J. (2003) TAC-1, a regulator of microtubule length in the C. elegans embryo. *Current biology : CB* **13**, 1499-1505

126. McKeveney, P. J., Hodges, V. M., Mullan, R. N., Maxwell, P., Simpson, D., Thompson, A., Winter, P. C., Lappin, T. R., and Maxwell, A. P. (2001) Characterization and localization of expression of an erythropoietin-induced gene,

ERIC-1/TACC3, identified in erythroid precursor cells. *British journal of haematology* **112**, 1016-1024

127. Stebbins-Boaz, B., Cao, Q., de Moor, C. H., Mendez, R., and Richter, J. D. (1999) Maskin is a CPEB-associated factor that transiently interacts with elF-4E. *Molecular cell* **4**, 1017-1027

128. Still, I. H., Hamilton, M., Vince, P., Wolfman, A., and Cowell, J. K. (1999) Cloning of TACC1, an embryonically expressed, potentially transforming coiled coil containing gene, from the 8p11 breast cancer amplicon. *Oncogene* **18**, 4032-4038

129. Still, I. H., Vince, P., and Cowell, J. K. (1999) The third member of the transforming acidic coiled coil-containing gene family, TACC3, maps in 4p16, close to translocation breakpoints in multiple myeloma, and is upregulated in various cancer cell lines. *Genomics* **58**, 165-170

130. Peset, I., and Vernos, I. (2008) The TACC proteins: TACC-ling microtubule dynamics and centrosome function. *Trends in cell biology* **18**, 379-388

131. Pu, J. J., Li, C., Rodriguez, M., and Banerjee, D. (2001) Cloning and structural characterization of ECTACC, a new member of the transforming acidic coiled coil (TACC) gene family: cDNA sequence and expression analysis in human microvascular endothelial cells. *Cytokine* **13**, 129-137

132. Gergely, F. (2002) Centrosomal TACCtics. *BioEssays : news and reviews in molecular, cellular and developmental biology* **24**, 915-925

133. Gergely, F., Karlsson, C., Still, I., Cowell, J., Kilmartin, J., and Raff, J. W. (2000) The TACC domain identifies a family of centrosomal proteins that can interact with microtubules. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 14352-14357

134. Partch, C. L., and Gardner, K. H. (2011) Coactivators necessary for transcriptional output of the hypoxia inducible factor, HIF, are directly recruited by ARNT PAS-B. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7739-7744

135. Kim, J. H., Yang, C. K., and Stallcup, M. R. (2006) Downstream signaling mechanism of the C-terminal activation domain of transcriptional coactivator CoCoA. *Nucleic acids research* **34**, 2736-2750

136. Kim, J. H., Li, H., and Stallcup, M. R. (2003) CoCoA, a nuclear receptor coactivator which acts through an N-terminal activation domain of p160 coactivators. *Molecular cell* **12**, 1537-1549

137. Yang, C. K., Kim, J. H., Li, H., and Stallcup, M. R. (2006) Differential use of functional domains by coiled-coil coactivator in its synergistic coactivator function with beta-catenin or GRIP1. *The Journal of biological chemistry* **281**, 3389-3397

138. Chang, K. H., Chen, Y. M., Chen, T. T., Chou, W. H., Chen, P. L., Ma, Y. Y., YangFeng, T. L., Leng, X. H., Tsai, M. J., OMalley, B. W., and Lee, W. H. (1997) A thyroid hormone receptor coactivator negatively regulated by the retinoblastoma protein. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 9040-9045

139. Lee, J. W., Choi, H. S., Gyuris, J., Brent, R., and Moore, D. D. (1995) Two classes of proteins dependent on either the presence or absence of thyroid hormone for interaction with the thyroid hormone receptor. *Molecular endocrinology* **9**, 243-254

140. Rios, R. M., Tassin, A. M., Celati, C., Antony, C., Boissier, M. C., Homberg, J. C., and Bornens, M. (1994) A peripheral protein associated with the cis-Golgi network redistributes in the intermediate compartment upon brefeldin A treatment. *The Journal of cell biology* **125**, 997-1013

141. Barr, F. A., and Egerer, J. (2005) Golgi positioning: are we looking at the right MAP? *The Journal of cell biology* **168**, 993-998

142. Gillingham, A. K., Tong, A. H., Boone, C., and Munro, S. (2004) The GTPase Arf1p and the ER to Golgi cargo receptor Erv14p cooperate to recruit the golgin Rud3p to the cis-Golgi. *The Journal of cell biology* **167**, 281-292

143. Infante, C., Ramos-Morales, F., Fedriani, C., Bornens, M., and Rios, R. M. (1999) GMAP-210, A cis-Golgi network-associated protein, is a minus end microtubule-binding protein. *The Journal of cell biology* **145**, 83-98

144. Drin, G., Casella, J. F., Gautier, R., Boehmer, T., Schwartz, T. U., and Antonny, B. (2007) A general amphipathic alpha-helical motif for sensing membrane curvature. *Nature structural & molecular biology* **14**, 138-146

145. Efimov, A., Kharitonov, A., Efimova, N., Loncarek, J., Miller, P. M., Andreyeva, N., Gleeson, P., Galjart, N., Maia, A. R., McLeod, I. X., Yates, J. R., 3rd, Maiato, H., Khodjakov, A., Akhmanova, A., and Kaverina, I. (2007) Asymmetric CLASP-dependent nucleation of noncentrosomal microtubules at the trans-Golgi network. *Developmental cell* **12**, 917-930

146. Marsh, B. J., Mastronarde, D. N., Buttle, K. F., Howell, K. E., and McIntosh, J. R. (2001) Organellar relationships in the Golgi region of the pancreatic beta cell line, HIT-T15, visualized by high resolution electron tomography. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2399-2406

147. Pernet-Gallay, K., Antony, C., Johannes, L., Bornens, M., Goud, B., and Rios, R. M. (2002) The overexpression of GMAP-210 blocks anterograde and retrograde transport between the ER and the Golgi apparatus. *Traffic* **3**, 822-832

148. Chen, Y., Chen, P. L., Chen, C. F., Sharp, Z. D., and Lee, W. H. (1999) Thyroid hormone, T3-dependent phosphorylation and translocation of Trip230 from the Golgi complex to the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 4443-4448

149. Majmundar, A. J., Wong, W. J., and Simon, M. C. (2010) Hypoxia-inducible factors and the response to hypoxic stress. *Molecular cell* **40**, 294-309

150. Semenza, G. L. (2012) Hypoxia-inducible factors in physiology and medicine. *Cell* **148**, 399-408

151. Gordan, J. D., and Simon, M. C. (2007) Hypoxia-inducible factors: central regulators of the tumor phenotype. *Curr Opin Genet Dev* **17**, 71-77

152. Singh, D., Chan, J. M., Zoppoli, P., Niola, F., Sullivan, R., Castano, A., Liu, E. M., Reichel, J., Porrati, P., Pellegatta, S., Qiu, K., Gao, Z., Ceccarelli, M., Riccardi, R., Brat, D. J., Guha, A., Aldape, K., Golfinos, J. G., Zagzag, D.,

Mikkelsen, T., Finocchiaro, G., Lasorella, A., Rabadan, R., and Iavarone, A. (2012) Transforming fusions of FGFR and TACC genes in human glioblastoma. *Science* **337**, 1231-1235

153. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **254**, 539-544

154. Greenfield, N. J., Montelione, G. T., Farid, R. S., and Hitchcock-DeGregori, S. E. (1998) The structure of the N-terminus of striated muscle alpha-tropomyosin in a chimeric peptide: nuclear magnetic resonance structure and circular dichroism studies. *Biochemistry* **37**, 7834-7843

155. Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12-21

156. Card, P. B., Erbel, P. J., and Gardner, K. H. (2005) Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization. *Journal of molecular biology* **353**, 664-677

157. de Vries, S. J., van Dijk, M., and Bonvin, A. M. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nature protocols* **5**, 883-897

158. Iwahara, J., and Clore, G. M. (2006) Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* **440**, 1227-1230

159. Battiste, J. L., and Wagner, G. (2000) Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. *Biochemistry* **39**, 5355-5365

160. Erbel, P. J., Card, P. B., Karakuzu, O., Bruick, R. K., and Gardner, K. H. (2003) Structural basis for PAS domain heterodimerization in the basic helix--loop--helix-PAS transcription factor hypoxia-inducible factor. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 15504-15509

161. Forbes, S. A., Tang, G., Bindal, N., Bamford, S., Dawson, E., Cole, C., Kok, C. Y., Jia, M., Ewing, R., Menzies, A., Teague, J. W., Stratton, M. R., and Futreal, P. A. (2010) COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic acids research* **38**, D652-657

162. Koehler, A. N. (2010) A complex task? Direct modulation of transcription factors with small molecules. *Current opinion in chemical biology* **14**, 331-340

163. Wells, J. A., and McClendon, C. L. (2007) Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001-1009

164. Darnell, J. E., Jr. (2002) Transcription factors as targets for cancer therapy. *Nature reviews. Cancer* **2**, 740-749

165. Yang, J., Zhang, L., Erbel, P. J. A., Gardner, K. H., Ding, K. M., Garcia, J. A., and Bruick, R. K. (2005) Functions of the Per/ARNT/Sim (PAS) domains of the hypoxia inducible factor (HIF). *Journal of Biological Chemistry* **280**, 36047-36054

166. Key, J., Scheuermann, T. H., Anderson, P. C., Daggett, V., and Gardner, K. H. (2009) Principles of ligand binding within a completely buried cavity in HIF2alpha PAS-B. *Journal of the American Chemical Society* **131**, 17647-17654

167. Hao, N., Whitelaw, M. L., Shearwin, K. E., Dodd, I. B., and Chapman-Smith, A. (2011) Identification of residues in the N-terminal PAS domains important for dimerization of Arnt and AhR. *Nucleic Acids Research* **39**, 3695-3709

168. Card, P. B., Erbel, P. J. A., and Gardner, K. H. (2005) Structural basis of ARNT PAS-B dimerization: Use of a common beta-sheet interface for hetero- and homodimerization. *Journal of Molecular Biology* **353**, 664-677

169. Amezcua, C. A., Harper, S. M., Rutter, J., and Gardner, K. H. (2002) Structure and interactions of PAS kinase N-terminal PAS domain: model for intramolecular kinase regulation. *Structure* **10**, 1349-1361

170. Best, J. L., Amezcua, C. A., Mayr, B., Flechner, L., Murawsky, C. M., Emerson, B., Zor, T., Gardner, K. H., and Montminy, M. (2004) Identification of small-molecule antagonists that inhibit an activator: coactivator interaction. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 17622-17627

171. Warashina, M., Min, K. H., Kuwabara, T., Huynh, A., Gage, F. H., Schultz, P. G., and Ding, S. (2006) A synthetic small molecule that induces neuronal differentiation of adult hippocampal neural progenitor cells. *Angew Chem Int Ed Engl* **45**, 591-593

172. Wurdak, H., Zhu, S., Min, K. H., Aimone, L., Lairson, L. L., Watson, J., Chopiuk, G., Demas, J., Charette, B., Halder, R., Weerapana, E., Cravatt, B. F., Cline, H. T., Peters, E. C., Zhang, J., Walker, J. R., Wu, C., Chang, J., Tuntland, T., Cho, C. Y., and Schultz, P. G. (2010) A small molecule accelerates neuronal differentiation in the adult rat. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 16542-16547

173. Crosson, S., and Moffat, K. (2001) Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2995-3000

174. Gong, W., Hao, B., Mansy, S. S., Gonzalez, G., Gilles-Gonzalez, M. A., and Chan, M. K. (1998) Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc. Natl. Acad. Sci.* **95**, 15177-15182

175. Genick, U. K., Borgstahl, G. E., Ng, K., Ren, Z., Pradervand, C., Burke, P. M., Srajer, V., Teng, T. Y., Schildkamp, W., McRee, D. E., Moffat, K., and Getzoff, E. D. (1997) Structure of a protein photocycle intermediate by millisecond time-resolved crystallography. *Science* **275**, 1471-1475

176. Hajduk, P. J., Bures, M., Praestgaard, J., and Fesik, S. W. (2000) Priveleged molecules for protein binding identified from NMR-based screening. *Journal of Medicinal Chemistry* **43**, 3443-3447

177. Farmer, B. T., Constantine, K. L., Goldfarb, V., Friedrichs, M. S., Wittekind, M. G., Yanchunas, J. J., Robertson, J. G., and Mueller, L. (1996) Localizing the $NADP^+$ binding site on the MurB enzyme by NMR. *Nat. Struct. Biol.* **3**, 995-997

178. Hirose, K., Morita, M., Ema, M., Mimura, J., Hamada, H., Fujii, H., Saijo, Y., Gotoh, O., Sogawa, K., and Fujii-Kuriyama, Y. (1996) cDNA cloning and tissue-specific expression of a novel basic helix-loop-helix/PAS factor (Arnt2) with close sequence similarity to the aryl hydrocarbon receptor nuclear translocator (Arnt). *Molecular and cellular biology* **16**, 1706-1713

179. Sekine, H., Mimura, J., Yamamoto, M., and Fujii-Kuriyama, Y. (2006) Unique and overlapping transcriptional roles of arylhydrocarbon receptor nuclear translocator (Arnt) and Arnt2 in xenobiotic and hypoxic responses. *The Journal of biological chemistry* **281**, 37507-37516

180. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *Journal of biomolecular NMR* **44**, 213-223

181. Lee, J., Tomchick, D. R., Brautigam, C. A., Machius, M., Kort, R., Hellingwerf, K. J., and Gardner, K. H. (2008) Changes at the KinA PAS-A dimerization interface influence histidine kinase function. *Biochemistry* **47**, 4051-4064

182. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195-201

183. Barelier, S., and Krimm, I. (2011) Ligand specificity, privileged substructures and protein druggability from fragment-based screening. *Curr Opin Chem Biol* **15**, 469-474

184. Cappell, K. M., Sinnott, R., Taus, P., Maxfield, K., Scarbrough, M., and Whitehurst, A. W. (2012) Multiple cancer testis antigens function to support tumor cell mitotic fidelity. *Molecular and cellular biology* **32**, 4131-4140

185. Dioum, E. M., Chen, R., Alexander, M. S., Zhang, Q., Hogg, R. T., Gerard, R. D., and Garcia, J. A. (2009) Regulation of hypoxia-inducible factor 2alpha signaling by the stress-responsive deacetylase sirtuin 1. *Science* **324**, 1289-1293

186. Scortegagna, M., Ding, K., Zhang, Q., Oktay, Y., Bennett, M. J., Bennett, M., Shelton, J. M., Richardson, J. A., Moe, O., and Garcia, J. A. (2005) HIF-2alpha regulates murine hematopoietic development in an erythropoietin-dependent manner. *Blood* **105**, 3133-3140

187. Qing, G., and Simon, M. C. (2009) Hypoxia inducible factor-2alpha: a critical mediator of aggressive tumor phenotypes. *Curr Opin Genet Dev* **19**, 60-66

188. Lauffart, B., Vaughan, M. M., Eddy, R., Chervinsky, D., DiCioccio, R. A., Black, J. D., and Still, I. H. (2005) Aberrations of TACC1 and TACC3 are associated with ovarian cancer. *BMC Womens Health* **5**, 8

189. Swinney, D. C., and Anthony, J. (2011) How were new medicines discovered? *Nature reviews. Drug discovery* **10**, 507-519

190. Park, E. J., Kong, D., Fisher, R., Cardellina, J., Shoemaker, R. H., and Melillo, G. (2006) Targeting the PAS-A domain of HIF-1alpha for development of small molecule inhibitors of HIF-1. *Cell Cycle* **5**, 1847-1853

191. Lee, K., Zhang, H., Qian, D. Z., Rey, S., Liu, J. O., and Semenza, G. L. (2009) Acriflavine inhibits HIF-1 dimerization, tumor growth, and vascularization.

Proceedings of the National Academy of Sciences of the United States of America **106**, 17910-17915

192.   Sheffield, P., Garrard, S., and Derewenda, Z. (1999) Overcoming expression and purification problems of RhoGDI using a family of "parallel" expression vectors. *Protein Expression and Purification* **15**, 34-39

193.   Blommel, P. G., and Fox, B. G. (2007) A combined approach to improving large-scale production of tobacco etch virus protease. *Protein expression and purification* **55**, 53-68

194.   Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 859-866

195.   McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674

196.   Afonine, P. V., Mustyakimov, M., Grosse-Kunstleve, R. W., Moriarty, N. W., Langan, P., and Adams, P. D. (2010) Joint X-ray and neutron refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 1153-1163

197.   Cowtan, K. (2010) Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 470-478

198.   Cowtan, K. (2006) The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1002-1011

199.   Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47 (Pt 2)**, 110-119

200.   Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501

201.   Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enz.* **276**, 307-326

202.   McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674

203.   Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132

204.   Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221

205.   Kleywegt, G. J., and Jones, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr D Biol Crystallogr* **50**, 178-185

206. Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93-158

207. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277-293

208. Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213-223

209. Bookout, A. L., and Mangelsdorf, D. J. (2003) Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nuclear receptor signaling* **1**, e012