Role of Host Cell Factor-1 in MYC-driven gene expression and tumorigenesis

By

Tessa Mary Popay

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Cell and Developmental Biology

January 31, 2021

Nashville, Tennessee

Approved:

Andrea Page-McCaw, Ph.D.

Jason A. MacGurn, Ph.D.

Jin Chen, M.D./Ph.D.

P. Anthony Weil, Ph.D.

William P. Tansey, Ph.D.

Acknowledgements

Dr. Bill Tansey, my mentor for the past five and a half years, has persevered to make me a better scientist, even considering my slow start. He has continued to guide me towards the best experiments, whilst simultaneously challenging and pushing my ideas. I believe his patience and expertise has enabled me an exciting future, and for that I will be eternally grateful. Secondly, I particularly want to thank Dr. April Weissmiller, who has essentially functioned as a second mentor to me during my time in the Tansey laboratory. She has been a guiding light during this process, providing scientific direction, and regularly expressing her belief in me. I additionally would like to acknowledge the other members of the Tansey laboratory, both past and present, for the continued support, collaboration, and friendship, as well as collaborators, who have made important contributions to my dissertation. Finally, mentorship from my committee members, Dr. Andrea Page-McCaw, Dr. Jason MacGurn, Dr. Jin Chen, and Dr. Tony Weil, has encouraged me to improve my scientific approach and involvement.

My time at Vanderbilt has been made substantially easier by friends. Those from New Zealand have kept me grounded, and made my time away from home infinitely easier. The friends I have made here at Vanderbilt have become my second family, and, in doing so, provided me a home away from home. They have supplied many a welcome escape from research.

Lastly, my family are the primary reason I am here. My parents, Dr. Alison Popay and Dr. Ian Popay, and my siblings, Kate Skipper and (almost Dr.) Simon Popay, have inspired me to reach new heights. If there ever was a reason for my ambition, my competitiveness, my intellect, it is them. The things you have all achieved, what you have seen of the world, how much you have done for others, I cannot help but feel my accomplishments dwarf in comparison.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Basic helix-loop-helix — bHLH

Burkitt lymphoma — BL

Chromatin immunoprecipitation — ChIP

Cyclin-dependent kinase — CDK

Deoxyribonucleic acid — DNA

DNA binding domain — DBD

Enhancer box — E-box

Genomic DNA — gDNA

Green fluorescent protein — GFP

HCF-1 binding motif — HBM

Herpes simplex virus — HSV

Histone acetyltransferase — HAT

Histone deacetylase — HDAC

Histone methyltransferase — HMT

Host cell factor-1 — HCF-1

Immunoglobulin heavy — IgH

Immunoprecipitation — IP

Leucine zipper — LZ

Messenger RNA — mRNA

MYC box — Mb

Nuclear localization signal — NLS

Nuclear respiratory factor — NRF1

O-GlcNAc transferase — OGT

Reactive oxygen species — ROS

Ribonucleic acid — RNA

Ribosomal DNA — rDNA

Ribosomal RNA — rRNA

RNA polymerase — RNAP

Transactivation domain — TAD

Transcription start site — TSS

Transfer RNA — tRNA

Tricarboxylic acid cycle — TCA cycle

Ubiquitin-proteasome system — UPS

VP16 induced complex — VIC

WDR5 binding motif — WBM

Wild-type — WT

Chapter I

Introduction

MYC

MYC family of proteins

c-MYC (hereafter 'MYC'), L-MYC, and N-MYC form a family of evolutionarily conserved transcription factors that are commonly overexpressed in cancer [1]. First identified through homology to a gene, v-*myc*, that was co-opted by the avian myelocytomatosis virus and drives tumor development in chickens [2], these genes have since been extensively studied for their role in oncogenic cell function. The different family members were likely derived from two gene duplication events, the first occurring shortly after the appearance of vertebrates hundreds of millions of years ago [3]. Thus, a single copy of *myc* is present in a number of invertebrates, including *Drosophila melanogaster* (*dmyc* or *diminutive, dm*) [4], and can also be found in the closest living relatives of animals, choanoflagellates [5].

In mice, the loss of either *Myc* or *N-Myc* causes embryonic lethality, with both showing extensive proliferation and developmental defects [6, 7], while mice lacking *L-Myc* do not show an overt phenotype [8]. Despite this, *Myc* and *N-Myc* can substitute for each other in development, in that, if *N-Myc* is expressed from the *Myc* locus, embryonic lethality is largely rescued [9]. This suggests that it is the regulation of these genes, resulting in distinct spatio-temporal expression patterns, that causes differential function. Indeed, *Myc* is expressed in nearly all proliferating cells [10], and expression levels vary during development to alter cell growth and proliferation, such as during forebrain development [11]. Both *N-* and *L-Myc* show more restricted temporal and spatial expression patterns [12].

The overall sequence homology of proteins in the MYC family is poor, both within and between species (Figure 1-1A). In vertebrates, the primary exceptions to this include the carboxy-terminal, basic helix-

loop-helix (bHLH) and leucine zipper (LZ) DNA binding domain (DBD), and regions of stronger conservation called MYC boxes (Mb) that are distributed through the amino-terminal transactivation domain and the central portion [13]. Primary sequence conservation of the DBD and MYC boxes likely reflects the critical role of these regions in MYC function, and this has been confirmed repeatedly through deletion experiments [14, 15]. As of now, six MYC boxes have been identified, and these are numbered from amino- to carboxy-termini as Mb0, MbI, MbII, MbIIIa, MbIIIb, and MbIV (Figure 1-1B). Compared to the other family members, L-MYC is relatively short (364 amino acids in humans compared to 439 for MYC and 464 for N-MYC), and is devoid of MbIIIa and a nuclear localization signal (NLS; Figure 1-1B) [1]. Despite this, L-MYC is predominantly localized to the nucleus [16]. The function of these MYC boxes, which is largely distinct, will be addressed below.

MYC proteins are intrinsically disordered and alone are unable to bind to DNA, either *in vitro* or *in vivo*, and instead must first heterodimerize with their obligate binding partner MAX, an interaction that occurs through the leucine zipper region of the MYC DBD [17]. This provides structure to the MYC DBD and grants MYC:MAX dimers an innate capacity to bind to enhancer boxes (E-boxes) in DNA, which have the canonical sequence CACGTG [17, 18]. E-boxes occur at a frequency far greater than reflected by MYC binding to chromatin, indicating that additional factors, including chromatin structure [19], RNA polymerase (RNAP) occupancy [20], and binding partners [21], contribute to specificity in this process.

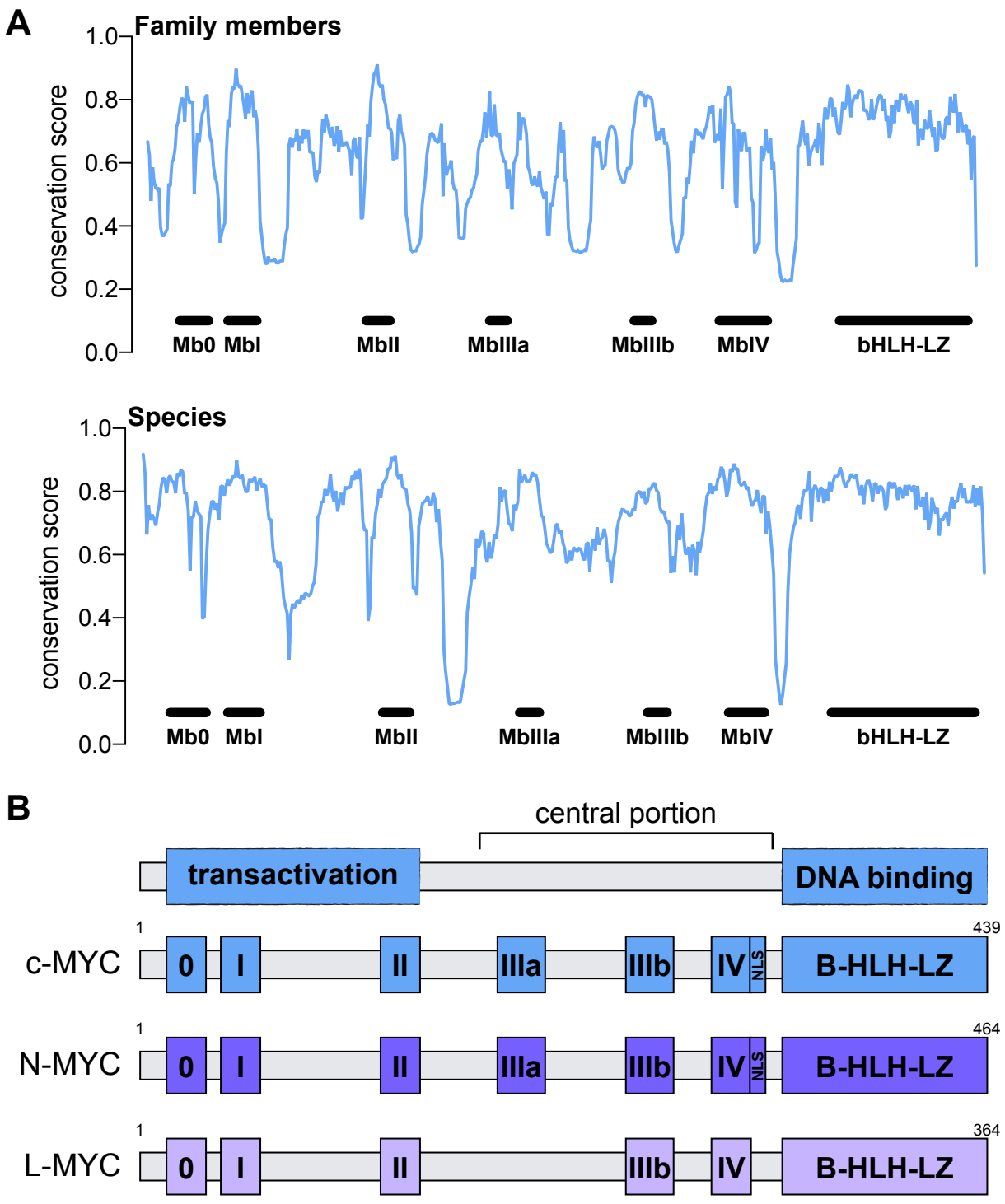


Figure 1-1: MYC family of proteins
(A) Conservation score of primary sequence for the MYC family of proteins in humans (c-MYC, N-MYC, and L-MYC; top) and for c-MYC across ten vertebrate species (bottom). Multiple sequence alignment was conducted using MUSCLE [22] and conservation scores by Jenson-Shannon divergence [23]. The ten vertebrate species aligned were human, mouse, sheep, chicken, goldfish, pig, groundhog, cat, zebrafish, and Rhesus macaque. **(B)** MYC family members and their distribution of MYC boxes in relation to the transactivation domain, central portion, and DNA binding domain.

Overexpression of MYC in cancer

MYC proteins are overexpressed in the majority of cancers, including both as initial and subsequent drivers of tumorigenesis [1, 24, 25]. Although this overexpression can occur through a number of different mechanisms, the proteins themselves are mutated only in a subset of cancers, and only in conjunction with genomic rearrangement [1]. Overexpression typically occurs at the transcriptional level, including through translocation, amplification, and epigenetic regulation [25]. While MYC family members are implicated in a wide variety of cancer types, there is some specificity shown, particularly for N-MYC and L-MYC. Indeed, N-MYC is commonly associated with cancers that have a neuroectodermal origin [26], such as neuroblastoma, and L-MYC can be amplified in small-cell lung cancers [27]. On the other hand, MYC is overexpressed in diverse cancer types, including both blood and solid cancers [27]. Although L-MYC has sequence homology with MYC and N-MYC, the absence of MbIIIa has been proposed as a possible cause of its reduced oncogenicity compared to the other family members [28].

Burkitt lymphoma (BL) is an example of a classic, MYC-driven cancer in which reciprocal translocation of *MYC* to the immunoglobulin heavy (IgH), kappa (Igκ), or lambda (Igλ) locus is the initiating mutation that results in overexpression of *MYC* [29]. Of these, the IgH *t(8;14)* translocation is the most common, accounting for ~80% of cases [30]. By placing *MYC* within the immunoglobulin locus, the allele additionally becomes subject to somatic hypermutation, a process that normally contributes to immunoglobulin diversity [31]. Indeed, a high proportion of BL cases have mutations within the translocated *MYC* locus [29], with up to 70% resulting in an amino acid substitution [32]. These mutations center on the transactivation domain, particularly around Threonine-58 [29], and have been found to enhance the oncogenic potential of MYC *in vivo* [33]. Mutation of Threonine-58 (e.g. T58A) can increase the half-life of MYC protein, and MYC proteins are commonly stabilized in BL [34].

Regulation of MYC expression and stability

The expression of MYC proteins is tightly controlled at all levels. At the transcriptional stage, *MYC* is downstream of various pathways, including WNT and NOTCH signaling [25], can be influenced by local and global chromatin structures [35], and is controlled by a host of transcription factors, such as the E2F family and BRD4 [36, 37]. *MYC* mRNA has a half-life of only ten minutes [38], and is subject to multiple levels of control prior to translation. These include targeting by a number of microRNAs [25], mitogen-regulated export from the nucleus [39], and translational repression [40]. MYC proteins are rapidly turned-over by an intricate mechanism of hierarchical phosphorylation, ubiquitylation, and degradation [41-43]. A number of regions within MYC appear to contribute to its turnover, but the primary degron is within the MYC transactivation domain [44]. Although these extensive systems enable tight control over expression of MYC, they also create multiple steps at which control can be lost.

One very well-understood mode of stability regulation through the MYC transactivation domain involves sequential phosphorylation events in MbI, specifically at Serine-62 by ERKs and Threonine-58 by glycogen synthase kinase 3β (GSK3β), enabling binding by the ubiquitin ligase complex SCF^{FBW7}, and subsequent poly-ubiquitylation and proteasomal degradation of MYC [45]. Serine-62 phosphorylation stabilizes MYC [46], but also primes MYC for phosphorylation at Threonine-58, possibly through regulating GSK3β activity [47]. Serine-62 phosphorylation also increases the interaction between MbI and the catalytic site of the prolyl isomerase PIN1, leading to a *cis* to *trans* proline isomerization of MYC that is necessary for dephosphorylation of Serine-62 [48]. This interaction also regulates MYC binding to, and transactivation from, a subset of MYC target genes [49]. Dephosphorylation of Serine-62 is catalyzed by protein phosphatase 2A (PP2A) [50], with MYC–PP2A interaction regulated by MYC phosphorylation; interaction is increased with the S62D phospho-mimetic and the T58A phospho-null mutations [51]. The ubiquitin ligase SCF complex, which includes SKP1, Cullin, RBX1, and an F-box protein, has multiple points of contact with MYC. The F-box protein present in the complex interacts with a target protein to determine specificity. SCF^{FBW7}, which includes the F-box protein FBXW7, is able

to ubiquitylate MYC and promote its degradation [43, 52]. A fragment that encompasses the majority of MbI is sufficient for interaction with FBXW7, but only if it is phosphorylated at Threonine-58 and Serine-62 [52]. Turnover of MYC protein is amongst the most detailed, mechanistic understandings in the MYC field. This is in part due to the potential of exploiting this process when targeting MYC-driven cancers. Indeed, activation of PP2A or inhibition of PIN1 are both considered good candidates for increasing turnover and reducing the functionality of MYC, respectively [53]. Various additional MYC interaction partners, both within and beyond MbI (Figure 1-2), have been found to influence this process at different stages.

More recently, the bromodomain-containing protein BRD4 has been reported to phosphorylate Threonine-58 and reduce MYC stability, independent of Serine-62 phosphorylation [54]. Like an increasing number of MYC interactions, this regulation is long-range, with BRD4 appearing to directly bind MYC just outside of MbII [54]. This relationship is particularly curious due to the capacity of BRD4 to also promote transcription of *MYC* [37]. While possible BRD4-driven *MYC* transcription and MYC turnover occur under different circumstances or as part of a feedback loop, it has previously been demonstrated that MYC turnover stimulates its transcriptional activity [55]. Whether this applies to BRD4 and MYC, specifically at endogenous MYC target genes, remains to be investigated. The phosphorylation of Threonine-58 by either GSK3β or BRD4 can be competed with by O-GlcNAcylation by O-GlcNAc transferase (OGT) at this residue, thereby stabilizing MYC protein [56-58]. The activity of OGT is subject to upstream regulation based on the metabolic state of the cell [59], and levels of this protein are reportedly increased in a number of cancers [60, 61].

Aurora kinases (AURK) have an integral role in cell division, but have also been implicated in post-translational modification of MYC. An association between N-MYC and AURKA was first identified based on the contribution of AURKA to growth of N-MYC-driven neuroblastoma cell lines [62, 63], and was later mapped as direct with regions in Mb0 and the carboxy-terminus of MbI [64]. This interaction stabilizes N-MYC [62], and this stabilization is independent of the catalytic activity of AURKA [63, 65].

Instead, AURKA has been proposed to compete with for SCF^{FBW7} binding to unphosphorylated N-MYC, thereby impairing efficient ubiquitylation and degradation of N-MYC [64]. Although much of the work on this interaction has focussed on N-MYC, the interaction is conserved in MYC and similarly influences its stability [66]. Furthermore, a region that encompasses MbIIIb, called the Proline (P), Glutamic acid (E), Serine (S), and Threonine (T) (PEST) domain, has recently been found to interact with AURKB [67]. The authors propose that this interaction causes MYC phosphorylation at Serine-67, which impairs interaction with GSK3β and reduces phosphorylation at Threonine-58 [67]. Because Threonine-58 phosphorylation enables interaction with the ubiquitin ligase complex SCF^{FBW7}, the MYC–AURKB interaction then stabilizes MYC [67].

MYC ubiquitylation and destabilization can also occur through interaction with SKP2, as part of the SCF^{SKP2} complex [55]. Both MbII and a region within the MYC DBD are able to interact with SKP2, and the SKP2 association with, and ubiquitylation of, MbII increases transcriptional activation driven by this region [55]. This interaction and subsequent ubiquitylation is independent of the hierarchical phosphorylation in MbI [55]. Finally, deleting a twenty-residue region that includes MbIIIa (“D element”) does not impact MYC ubiquitylation, but does increase protein stability, possibly by contributing to interaction with the proteasome [68].

Transcriptional and non-transcriptional functions of MYC

While the majority of MYC function in the cell is derived from its role as a transcription factor [69], a number of non-transcriptional obligations have been reported, although these have, in general, been poorly studied. As a transcription factor, MYC can function as both an activator and a repressor to regulate a wide variety of genes, many of which underpin its potency as a oncoprotein [69]. Full-length MYC only weakly activates transcription [70, 71]. This is despite having regions within the transactivation domain that can function as powerful transactivators when fused to the Gal4 DBD in Chinese hamster ovary cells [72], indicating that the structure of MYC auto-regulates its transactivation potency. What MYC lacks in strength, it makes up for with a broad set of target genes.

There are two overarching models of how MYC controls transcription; gene-specific and global amplification. The former of these dictates that MYC has distinct target genes, shows specificity in binding, and causes activation or repression of an explicit, well-defined subset of genes [73, 74]. Altered expression of these genes by MYC can then indirectly lead to global changes in mRNA levels [75]. In contrast, in the amplifier model, MYC is able to directly bind to, and promote transcription of all actively transcribed genes without showing any individual gene specificity [76, 77]. With this model, what we observe as specific, MYC-driven transcriptional programs is actually a consequence of additional, indirect changes in transcription that occur secondary to global amplification by MYC.

It has been proposed that the types of genes regulated by MYC expands with its level of expression, as determined by its increasing capacity to bind low-affinity sites [74]. Although MYC favors binding to canonical E-boxes [18, 78], it does have the ability to bind to low-affinity, non-canonical E-boxes [18]. This typically occurs when MYC is present at high levels, is commonly referred to as promoter and enhancer invasion, and influences the type of genes that are responsive to MYC [74]. Genes containing non-canonical E-boxes are commonly associated with promoting metabolism and tumor microenvironment, and are normally bound by other bHLH-LZ transcription factors, including hypoxia inducible factor (HIF)-1α, nuclear respiratory factor (NRF) 1, CLOCK, and sterol regulatory element-binding protein (SREBP) [25, 79]. Whether MYC competes with, or complements, the work of these transcription factors remains unclear [80, 81]. MYC also demonstrates preference for a subset of canonical E-box-containing genes, indicating that an E-box is not sufficient for MYC binding to chromatin, and additional cofactors can define this [74]. The promoter and enhancer invasion that occurs when MYC is overexpressed involves both canonical and non-canonical E-boxes [82]. Thus, MYC likely has a core set of target genes, but can amplify expression from additional genes when it is overexpressed [74, 82]. The level of amplification is, in turn, determined by the strength with which MYC is able to bind. Invasion has been proposed as a possible way in which MYC causes global amplification of transcription [76], but it has also been reported that this amplification occurs even with

physiological levels of MYC [77], and that invasion is separable from amplification [83]. Indeed, Lorenzin *et al.* [74] propose that it is the differential affinity of MYC to chromatin that accounts for gene-specific programs observed with MYC overexpression. In this sense, as MYC levels cross certain thresholds, the genes and resulting processes it regulates expands until a point of saturation, whereby the totality of MYC function, both direct and indirect, can be observed. This totality includes, for the most part, the epitome of processes crucial for the uncontrollable cell growth that defines cancer, including protein synthesis, metabolism, proliferation, genome stability, angiogenesis, and apoptosis [1].

Perhaps the most clearly defined, definitively direct, and baseline function of MYC is its role in ribosome biogenesis [74, 84]. MYC regulates this process through controlling and coordinating the transcription of ribosomal DNA (rDNA), ribosomal protein genes, and components in the processing and assembly of ribosomes [85]. To achieve this, MYC promotes transcription catalyzed by all three RNAPs [25]. Nucleoli are dense, protein-rich compartments within the nucleus that form around a fraction of the rDNA copies in the genome for RNAPI-driven transcription [86]. MYC is able to localize to nucleoli, and increase levels of precursor and mature ribosomal RNA (rRNA) [87]. A single 47S precursor rRNA (pre-rRNA) is transcribed from rDNA and is progressively processed into 18S, 5.8S, and 28S rRNA in the nucleus [85]. An additional rRNA fragment, 5S, is synthesized outside of the nucleolus through the activity of RNAPIII, which is also involved in the transcription of transfer RNA (tRNA), and is similarly activated by MYC [88]. Processing of pre-rRNA and ribosome assembly requires both accessory proteins and structural components (ribosomal protein genes), all of which are transcribed by RNAPII [85] and extensively regulated by MYC [21, 89]. An additional layer of MYC-driven regulation of protein synthesis occurs through the transcription of translation initiation factors, including EIF4E and EIF2α [90]. The combined influence that MYC has on virtually all stages of protein synthesis is consistent with the finding that overexpression of MYC can double the rate of protein synthesis in B-lymphocytes, leading to an increase in biomass and cell size [91], and emphasizing how high levels of MYC can contribute to driving cell growth in cancer. Indeed, the stages of ribosome biogenesis that MYC

promotes are also dependent on its expression level, with ribosomal protein genes considered to be the most high affinity, immediate-early of targets [74].

To support the energy-consuming processes it drives, MYC proteins are additionally involved in regulating cell metabolism and mitochondrial function [92]. As mentioned above, these processes are often considered to be gained targets of overexpressed MYC through invasion to non-canonical E-boxes. As ribosome biogenesis is the most energy consuming process in the cell [93], it seems unsurprising that MYC-driven acceleration of it would be accompanied by mechanisms to match the increased energy requirement. Indeed, MYC can promote expression of nuclear-encoded genes for metabolic pathways and mitochondrial biogenesis, thereby contributing to ATP production by the tricarboxylic acid (TCA) cycle and oxidative phosphorylation [92]. Perhaps one of the more curious roles of MYC in the control of metabolism is its connection to the Warburg effect, a cancer cell phenotype in which glycolysis becomes the favored source of ATP in a cell (anaerobic glycolysis) [94]. Compared to the ~36 molecules of ATP generated through the TCA cycle and oxidative phosphorylation, glycolysis alone results in a net gain of only two ATP molecules [95]. This in itself seems counterproductive to the growth-inducing goals of MYC, but glycolysis also produces building blocks that can be used for cell growth and proliferation, such as carbon and NADPH for fatty acid and nucleotide synthesis [96, 97]. While cancer cells typically increase both glycolytic- and TCA cycle/ oxidative phosphorylation-dependent energy production, the Warburg effect is the altered balance of these to increase both precursor and energy availability [97]. MYC contributes to both aspects of energy production by stimulating expression of enzymes in the glycolytic pathway [98] and inducing mitochondrial biogenesis [99]. Cancer cells also commonly demonstrate a dependency on the non-essential amino acid glutamine [100]. In addition to being necessary for translation, glutamine can be used as an energy source alongside glucose, and is a precursor for the biosynthesis of nucleotides and other non-essential amino acids, including glutamic acid, proline, and arginine [101]. MYC is, at least in part, responsible for this dependency through its ability to promote protein synthesis. However, MYC also enables glutamine addiction through regulating the expression of genes involved in glutamine

uptake and catabolism [101]. Thus, while MYC overexpression can create deficits in the metabolic state of the cell, it also has the means to meet these increased demands. It is this dual functionality that makes MYC such a powerful agent of oncogenesis.

The regulation of mitochondrial function by MYC has been proposed to generate excess reactive oxygen species (ROS) in cells, and this can lead to both nuclear and mitochondrial DNA damage [102-104]. Overexpression of MYC can additionally cause DNA damage independent of ROS production [105]. One mechanism of this is the transcription-independent role of MYC in the inappropriate recruitment of licensing factors, including the MCM complex, for the initiation of DNA replication [106]. In addition to this, MYC can cause replicative stress by increasing expression of replication machinery [107] and impairing cell cycle checkpoints [108], both of which contribute to MYC-driven genomic instability. MYC as a DNA-damaging agent is compounded by its ability to disrupt the DNA damage repair process [109, 110]. Genomic instability is one of the hallmarks of cancer [111], and is the primary cause of genetic heterogeneity within tumors [112]. This process can be important for tumor initiation, and can enable tumor maintenance under a diverse and changing microenvironment or with outside therapeutic pressure [112].

Beyond this point, it becomes difficult to disentangle MYC-driven biomass accumulation from the direct role of MYC in the regulation of additional processes. The loss of MYC causes impaired cell proliferation, both *in vitro* and *in vivo* [113, 114]. Consistent with this, stimulating cells with growth factors results in rapid induction of MYC expression [115]. MYC can accelerate cell cycle progression, specifically through the G_1 to S [116, 117] and S to G_2/M [118] transitions. While MYC's role in regulating cell growth undoubtedly influences the cell cycle, it has also been implicated in regulating expression of cyclins and cyclin-dependent kinases (CDKs) [114], including cyclins D1 [119] and D2 [120], and CDK4 [121]. The complex of CDK4 and cyclin D1/D2 contributes to the G_1 to S transition [114]. Under certain conditions, MYC is also able to promote re-entry of quiescent cells into the cell cycle [122, 123]. Finally, MYC was initially identified as one of the four 'Yamanaka factors', alongside

Oct3/4, Sox2, and Klf4, that are sufficient to revert somatic cells to their pluripotent state [124]. However, in this context, overexpression of MYC also results in a high rate of tumor formation [125], and has since been reported to be dispensable for the induction of pluripotency [126]. Despite the identification of target genes involved in the cell cycle, I doubt MYC's regulation of these would be sufficient to alone promote cell cycle progression. Indeed, this process would be unachievable were it not for the ribosomal and mitochondrial biogenesis that precedes it.

In the case of solid tumors, MYC can also function to adapt the microenvironment, including promoting angiogenesis and recruitment of accessory cells [127, 128]. The vasculature that develops in tumors is distinct from that formed with normal development and repair, particularly in that it is leaky, has irregular branching and wall thickness, is disorganized, and includes enlarged vessels [129]. The 'angiogenic switch' that enables this uncontrolled, abnormal, and inappropriate vascularization is controlled by a set of pro- and anti-angiogenic factors, with components of both able to be influenced by MYC [1]. Vascular endothelial growth factor (VEGF) is a prototypical pro-angiogenic factor whose levels are induced by MYC [130], including through interleukin 1β (IL-1β)-induced release of sequestered VEGF [127] or stimulation of VEGF translation initiation [131]. MYC can also indirectly decrease levels of the anti-angiogenic factor thrombospondin (TSP)-1 [132], by increasing expression of the microRNA miR-17-92 [133]. The amplification of glycolysis and mitochondria makes increased nutrient accessibility in MYC-driven cancers a necessity that can be met through angiogenesis. Akin to MYC enhancing mitochondrial biogenesis to account for increased ribosomal biogenesis, promotion of angiogenesis is a tertiary level of self-satiation by MYC, in that it is its own enabler. Under normoxic conditions, the miR-17-92 cluster also targets HIF-1α [134], which can functionally compete with MYC under certain circumstances [135]. Indeed, under conditions of hypoxia, which typically occurs in the absence of nearby vasculature, HIF-1α is stabilized, leading to a reduction in MYC levels and and impairment of MYC-driven cell growth and proliferation [135]. Another microRNA activated by MYC is miR-9, which in turn can regulate levels of E-cadherin to promote the epithelial-mesenchymal transition for local and systemic invasion of cancer cells [136, 137]. Combined with its regulation of various genes that

facilitate cell motility and adhesion [138, 139], MYC promotes the metastatic phenotype of cancer cells. While MYC can be overexpressed in solid tumors, it also plays an important role in blood-based tumors, whereby these processes are not applicable.

Another set of genes activated by MYC are those involved in apoptosis [140], meaning that overexpression of MYC can promote programmed cell death [141-143]. This is perhaps counterintuitive to the general function of MYC in boosting growth and proliferation, but it does allow for an important failsafe when MYC is overexpressed. MYC-driven apoptosis can be mediated through a number of different pathways [140]. MYC can induce expression of *ARF* and *p53*, leading to apoptosis [144]. In this pathway, p53 is continuously turned over through its interaction with MDM2, but this process is prevented in the presence of ARF, leading to stabilization of p53 [145-147]. Transcriptional activation of pro-apoptotic target genes by p53 then leads to apoptosis [146, 148]. Similarly, MYC itself can alter the balance of pro- and anti-apoptotic genes, including repression of the anti-apoptotic BCL-XL and BCL-2 [149], and induction of the pro-apoptotic BIM [150], Additionally, MYC promotes activity of the pro-apoptotic Bax, independent of transcriptional alterations [151]. Based on these mechanisms, the pro-apoptotic function of MYC can be overcome through additional genetic changes, either to the MYC protein or apoptotic pathways in the cell [140]. In an elegant experiment that I think is amongst the clearest in demonstrating the overall contribution of MYC to a process, Eischen *et al.* [152] used the development of B-cell lymphoma in a transgenic mouse model for overexpression of *Myc* in B-cells, whereby the disease appears in 6-8 months, to determine mechanisms by which MYC-driven cancers can evade MYC-driven apoptosis. All tumors that developed in this model were found to contain undetectable, low, or mutant (inactive) p53, sufficient to impair the MYC-driven apoptosis [152]. Where p53 was not mutated, deletion of $p19^{ARF}$ or overexpression of *Mdm2* were largely responsible for the reduction in p53 levels [152]. In this context, MYC again could be considered to have dual functionality in that it causes the issue of increased apoptosis, but is also likely a means to an end through genomic instability. Curiously, mutations within the MYC transactivation domain, including the T58A mutation mentioned above, have been reported to overcome the requirement for disruption of the ARF-MDM2-

p53 pathway in protection from MYC-driven apoptosis [33]. Instead, these mutations limit the ability of MYC to up-regulate the expression of *Bim*, leading to inhibition of Bcl2. [33].

While specific MYC-regulated genes have been identified for their ability to convey many of these functions, the act of singling out target genes may indeed be a consequence of convenience, rather than of totality. Furthermore, the identification of target genes is commonly based on their altered expression in response to overexpression of MYC, and may reflect both direct or indirect consequences of MYC function. Indeed, there are many complexities of MYC that an absence of appropriate technologies has meant we misunderstand or misattribute, but the availability of tools such as next-generation sequencing and genome editing creates an opportunity to redefine MYC function in both physiological and oncogenic contexts. There is no doubt that MYC can promote these processes I have described, but when and how it achieves this, including in regards to specific genes, cofactors, and overarching models, is a question that is only just beginning to be answered.

MYC function is dependent on cofactors

While MAX is the most well-defined MYC interaction partner, it is only one of many that are important to MYC function. For the most part, these interactions have been mapped to the conserved regions of MYC, including the DBD and MYC boxes [13]. The extent to which the function of the MYC boxes is understood varies, with those in the transactivation domain more extensively studied than those the central portion [1], and each appears to make multiple direct and indirect interactions (Figure 1-2) [15]. The majority of the MYC boxes, but particularly MbI, have also been implicated in regulating MYC stability. For the most part, the interaction partners responsible for this, which are described above, do not contribute to MYC function beyond controlling its turnover. While almost all of our understanding of MYC function relates to its role as a transcriptional activator, there are pockets of evidence indicating that it can also contribute to transcriptional repression. With the intrinsically disordered structure of the MYC protein itself, understanding each of the MYC boxes and their respective cofactors is considered

our best opportunity to impair MYC function in cancer. In an ideal world, this would mean inhibiting only the functions gained by MYC in its overexpression, but whether this is achievable remains to be seen.

MYC-driven transcriptional activation is conveyed through two primary means, the first of which is the modification of histone proteins. MbII has repeatedly been shown to be critical to MYC-driven tumor growth [15]. Specifically, its interaction with TRRAP contributes to MYC-driven transformation [153] by bridging interaction with the HATs GCN5 or TIP60/KAT5 [154, 155]. Curiously, both MbI and MbII appear to impart this interaction, individually demonstrating necessity and sufficiency [153, 156]. Consistent with this, Kalkat *et al.* [15] found that deletion of MbI impacted co-immunoprecipitation of the STAGA (GCN5) complex, while deletion of MbII affected co-immunoprecipitation of both the STAGA and TIP60 complexes. MYC is able to recruit TRRAP and its associated HATs to chromatin, leading to acetylation of histones H3 and/or H4 [155, 157-159]. Effects of histone acetylation are two-fold. First, acetyl groups counteract the positive charge of histone tails, thereby altering the relationship between histones and DNA, resulting in chromatin de-compaction and increasing the accessibility of chromatin for transcriptional machinery [160]. Secondly, acetyl groups can be recognized and bound by bromodomain-containing proteins and complexes, which can further contribute to the opening of the chromatin structure [160]. The stage of the transcriptional program at which MYC recruits TRRAP/HATs is context dependent, with various reports suggesting this occurs prior to the recruitment of RNAPI [161] and RNAPIII [162], and after the recruitment of RNAPII [155]. Furthermore, *RPL11* is a MYC target gene and the encoded protein is a component of the 60S ribosome subunit [163]. When overexpressed, free RPL11 is able to compete with TRRAP for binding to MbII at chromatin [163]. This reduces histone acetylation and leads to decreased RNAPI and II-driven transcription [163], suggesting RPL11 negatively impacts transactivation by MYC to maintain balanced ribosome biogenesis. Thus, while the MYC–TRRAP interaction is likely the primary role of MbII in terms of contributing to overall MYC function and tumorigenic potential, the association with RPL11 is context-dependent and could only influence transcriptional output where MYC is able to interact with TRRAP. Indeed, one would expect the MYC–RPL11 interaction to be particularly relevant during MYC overexpression, when a

balanced, but accelerated, rate of ribosome biogenesis must be achieved. Imbalances in ribosomal proteins or ribosomal RNA can lead to activation of p53 and subsequent apoptosis [164].

The second primary mechanism by which transcriptional activation by MYC is conveyed is through the release of paused polymerase and transcriptional elongation. Following initiation of transcription, RNAPII can become stabilized on chromatin by the DSIF and negative elongation factor (NELF) complexes, causing a transient 'pausing' of RNAP and transcriptional elongation [165]. Release of paused polymerase occurs through phosphorylation of NELF, DSIF, and the carboxy-terminal domain (CTD) of RNAPII by positive transcription elongation factor b (P-TEFb), a complex of CDK9 and cyclin T1/T2 [166]. While phosphorylated NELF is released from the transcription machinery, phosphorylated DSIF becomes a positive elongation factor by stabilizing RNAPII [166]. MYC has been reported to contribute to pause release through multiple mechanisms. The first of these is a direct interaction between the MYC transactivation domain (MbI and MbII) and cyclin T1, leading to recruitment of P-TEFb and phosphorylation of the RNAP CTD to promote elongation [167-170]. Degradation of MYC also promotes transcriptional elongation, with inhibition of its turnover limiting the recruitment of P-TEFb [171]. This is proposed to occur due to an interaction between MYC and the transcriptional elongation complex PAF1, thereby preventing interaction of PAF1 with RNAPII and inhibiting elongation until MYC is turned-over to release PAF1 [171]. A component of the PAF1 complex, LEO1, has also been implicated in the recruitment of MYC to chromatin, at least in *D. melanogaster* [172]. Similarly, MYC binds to and recruits SPT5, a component of DSIF, and transfers it to RNAPII in a CDK7-dependent manner [173]. Because modified DSIF promotes elongation, this function of MYC enables RNAPII processivity [173]. MYC is also reported to increase levels of CDK7 and CDK9 by induction of mRNA cap methylation promoting EIF4F-dependent recruitment of the mRNA to the ribosome, polysome formation, and translation rate [174, 175]. Finally, the most amino-terminal MYC box, Mb0, is also proposed to promote transcriptional elongation [15]. Kalkat *et al.* [15] found that deletion of Mb0 from MYC causes a decrease in cell proliferation, *in vitro* transformation proficiency, and tumorigenesis. This region was implicated in transactivation of upwards of 100 MYC target genes, and found to enable

association of MYC with various factors involved in transcriptional elongation, including a direct interaction with the general transcription factor TFIIF [15]. This transcription factor was initially linked to formation of the RNAPII pre-initiation complex, but there are increasing reports of a function in transcriptional elongation [176, 177]. While the evidence entangling MYC with transcriptional elongation by P-TEFb and SPT5 is definitive, whether this is also the relevant function of the MYC–TFIIF interaction remains to be investigated. Because the two general mechanisms of MYC-driven transcriptional activation occur following recruitment of RNAPII, it would be interesting to know if both are active at all MYC target genes, and, if so, whether they have a defined temporal relationship in regulating transcription. Regardless, both of these are viable means by which MYC could cause global amplification of transcription, particularly where MYC is responsible for the recruitment of relevant cofactors.

These functions of the MYC transactivation domain, and those involved in control over its stability, have been repeatedly investigated, interrogated, and are well-characterized. In contrast, functions relating to the central portion of MYC have only recently come to light. MbIIIa, MbIIIb, and MbIV are the most conserved regions of the central portion. MbIIIa appears to contribute to both activation and repression [14], only the latter of which has been investigated and will be addressed below. The most well-defined central portion interaction is that between MbIIIb and the scaffolding protein WDR5 [178, 179]. WDR5 is a member of various histone modifying complexes, including the SET/MLL family of histone methyltransferases (HMT) and the non-specific lethal (NSL) HAT complex [180]. More recently, WDR5 has been found to recruit MYC to chromatin [178], specifically to a subset of ribosomal protein genes [21]. This highlights that, although MYC is competent for DNA binding, this is not necessarily sufficient for chromatin binding in the complexity of the cell. Facilitated recruitment of MYC by WDR5 contributes to transcriptional activation at these target genes, and disrupting the interaction impairs MYC-driven tumor growth and maintenance [21]. These findings are particularly exciting because of the idea that the ribosomal protein genes constitute core target genes of MYC. The recruitment of MYC by WDR5 to these genes is pivotal for their MYC-dependent transcription, suggesting that, at least at these sites and

with the assistance of WDR5, gene-specific transcriptional regulation by MYC is an accurate model. While many of these binding sites contain canonical E-boxes, a surprisingly high proportion of them do not [21]. Furthermore, perturbation of the MYC–WDR5 interaction eliminates most MYC from these genes, but leaves some residual binding [21]. For binding sites that do not contain a canonical E-box, facilitated recruitment by WDR5 is a logical way in which MYC binding could otherwise be increased at a low-affinity site. However, for sites that do contain a canonical E-box and therefore are high-affinity for MYC, it is difficult to reconcile how genome-wide promoter and enhancer invasion by MYC could occur in the absence of additional cofactors for facilitated recruitment.

Despite a number of attempts, very little progress has been made towards understanding the role of MbIV, with this remaining the only MYC box without an assigned function. An interaction partner of MbIV, host cell factor (HCF)-1, has only recently been identified [179]. Previous work has attempted to interrogate the function of MbIV in the context of N-MYC using deletion of this region [181]. In this context, MbIV was found to contribute to DNA and chromatin binding, apoptosis following serum deprivation, the G_2 to M cell cycle phase transition, and both transcriptional activation and repression [181]. While interaction with upwards of 100 proteins may be influenced by this equivalent deletion in MYC [15], HCF-1 is the only validated MbIV interaction partner [179]. This dissertation and associated manuscript [182] aim to dispense this missing characterization of MbIV.

MYC is predominantly involved in transcriptional activation, but can contribute to repression, in part through its interaction with Miz-1 [183, 184]. The mechanism of this has been attributed to recruitment of, and competition between, MYC and the histone acetyltransferase (HAT) p300 for interaction with Miz-1, leading to altered capacity for transactivation by Miz-1. [184]. This enables MYC to impair Miz-1-driven activation of CDK inhibitors, including $p21^{Cip1}$, $p15^{INK4b}$, and $p57^{KIP2}$ [184-186]. The interaction between MYC and Miz-1 occurs through the MYC DBD, and whether this competes for binding to MAX is disputed [187, 188]. Also contributing to MYC-driven repression of transcription, specifically at $p21^{Cip1}$, is an interaction between MbI/II and DNA methyltransferase 3A (DNMT3A), creating a

DNMT3A–MYC–Miz-1 ternary complex that is responsible for DNA methylation at the $p21^{Cip1}$ promoter [189]. The tumor suppressor SNF5 also interacts with the MYC DBD [190], limiting MYC binding to chromatin and MYC-driven transcriptional activation [191]. Unlike other MYC interaction partners, association with SNF5 appears to suppress global MYC activity. Furthermore, while the function of MbIIIa is not particularly well understood, it is reported to be involved in transcriptional repression by MYC [14]. Specifically, this region interacts with histone deacetylase 3 (HDAC3), leading to a reduction in histone acetylation and variable effects on RNAPII recruitment [192]. The role of MbIIIa in MYC-driven transformation *in vitro* is context-dependent [14, 15]. However, Herbst *et al.* [14] reported that MbIIIa deletion promotes apoptosis and impairs *in vitro* transformation by MYC, with the latter rescued by overexpression by the anti-apoptotic BCL-2. These consequences are consistent with the model whereby reduced oncogenicity of L-MYC is caused by the absence of MbIIIa [28], as this would predict L-MYC overexpression is more likely to cause apoptosis upon overexpression. The concept of MYC acting as a transcriptional repressor does not immediately agree with the amplifier model. Indeed, it has been proposed that the different strengths with which MYC binds different promoters and enhancers, leading to different levels of transcriptional activation, means inappropriate normalization of data can make weak activation appear as repression [79]. On the other hand, MYC may also be able to more delicately fine-tune transcriptional output from specific target genes.

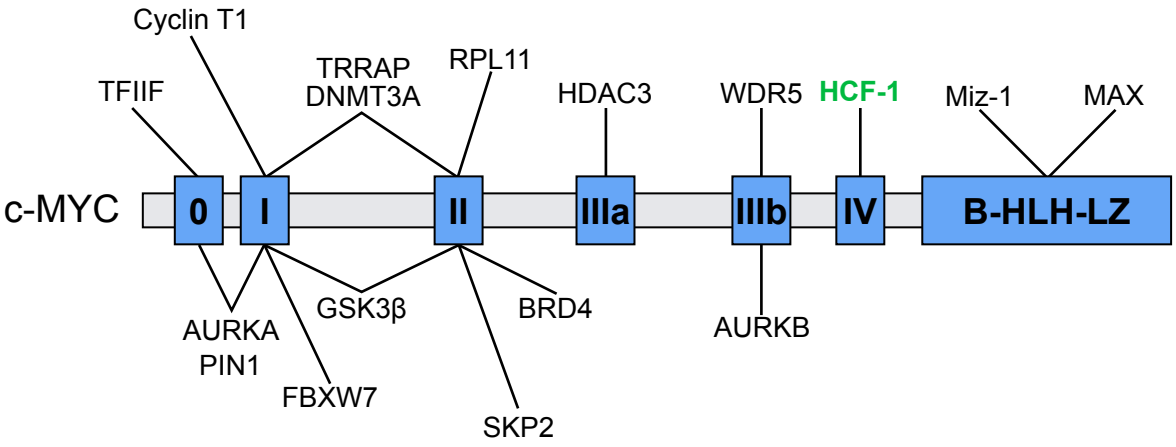


Figure 1-2: MYC interaction partners and the region through which binding is mediated
Above MYC are cofactors that contribute to MYC-dependent transcriptional activation or repression. Below MYC are proteins that associate with MYC to regulate its stability.

Targeting MYC-driven cancers

The overexpression of MYC proteins in the majority of cancers would, in theory, make it an ideal target for the development of anti-cancer therapies. However, as an intrinsically disordered protein, MYC itself is largely considered undruggable [193]. Instead, there are two broad approaches that are proposed; decreasing MYC levels or decreasing MYC function. Ideally, these approaches would differentiate between the normal, physiological function of MYC and the gained function of overexpressed, oncogenic MYC. This could mean reducing MYC levels to baseline, or impairing only a subset of MYC function. Achieving the latter of these is crucially dependent on understanding the contribution of MYC cofactors to its function and tumorigenic potential. And, as it stands, we know the function of only a handful of MYC cofactors, despite hundreds of proteins appearing to be associated with MYC [15].

Controlling MYC levels could be achieved by regulating its transcription, translation, or stability. The first of these might be possible through stabilization of G-quadruplex structures in the *MYC* promoter, which are reported to cause repression of *MYC* transcription [194, 195], or through inhibiting pathways that regulate MYC expression. Bromodomain and extra-terminal motif (BET) protein inhibitors, such as JQ1, were inadvertently found to suppress *MYC* transcription [37], and have demonstrated efficacy in various cancer types [196]. Development of proteolysis-targeting chimeras (PROTACs), which create a molecular link to a protein of interest and the UPS for targeted degradation, based on BET inhibitors, have also proved fruitful as anti-cancer agents in xenograft mouse models [197, 198]. The recent finding that BRD4 also destabilizes MYC [54] is an important aspect to consider when considering BET inhibitors and PROTACs. Indeed, Devaiah *et al.* [54] noted that JQ1 decreases *MYC* transcription, but does not affect MYC stability. On the other hand, MZ1 is a JQ1-based PROTAC that decreases MYC stability, but increases *MYC* transcription [54]. The more important question, though, would be which approach is more adept at reducing MYC function? Similar to the use of BET inhibition, inhibitors of CDKs have been used to reduce MYC transcription [199, 200]. Specifically targeting *MYC* mRNA has been interrogated [201, 202], as has limiting MYC translation through inhibition of translation factors

[203]. The well-understood mechanism of MYC protein turnover is also a strong potential target for reducing MYC levels, including through indirect activation of PP2A [204] or inhibition of AURKA [63].

Our ability to block the function of MYC is dependent on understanding how its oncogenic function is achieved. However, only a few MYC cofactors have been characterized sufficiently for this approach to be exploited. One encouraging example of this is WDR5, which is responsible for the recruitment of MYC to a subset of ribosomal protein genes [21]. Mutation of the MYC interaction surface or inhibition of WDR5 causes a reduction of MYC levels at chromatin and a subsequent impairment in the transcription of these genes [21, 205]. Because mutating the MYC interaction surface also results in diminished tumor growth and tumor regression, it is anticipated that inhibiting WDR5 would have similar consequences on MYC-driven tumors. However, this also risks blocking the baseline function of MYC in normal cells. Perhaps the most well-described and promising inhibitors of MYC function is the dominant-negative, “mini-protein” OmoMYC, which is a mutant form of the MYC DBD that is able to homodimerize and compete with MYC:MAX dimers for binding to chromatin [82, 206]. One of the more interesting aspects of OmoMYC is the finding that it preferentially blocks MYC binding to low-affinity sites, leaving the high-affinity sites intact [82]. In this sense, OmoMYC has the potential to leave the normal functions of MYC largely unscathed, and instead impair expression of the gained targets of oncogenic MYC. OmoMYC has been shown to cause regression of tumors in mouse models of lung cancer [207, 208], pancreatic islet cancer [209], and glioma [210], with minimal and reversible side effects. More recently, OmoMYC has been found to penetrate cells via endocytosis and cause lung adenocarcinoma regression in mice following intranasal administration [211]. Most strikingly though, the tumors that OmoMYC has been found to regress are not MYC-driven, highlighting that MYC is an important contributor to the oncogenic phenotype regardless of the driver and the potential of using OmoMYC in broad tumor types.

For many of these potential candidates, neither the gene encoding MYC nor the protein itself are their only targets. Thus, the development of small-molecule inhibitors to reduce the levels or function of MYC

are faced with not only targeting the gained function of MYC, but also limiting the effect on non-MYC targets.

Host Cell Factor-1

Discovery, family, and proteolysis

Host cell factor (HCF)-1 was first identified for its role in the transcription of viral genes. The herpes simplex virus (HSV) tegument transcription factor VP16 is capable of potent transcriptional activation [212], but shows only weak DNA binding ability alone [213, 214]. To overcome this deficit, VP16 complexes with host cell proteins, including the POU-domain-containing transcription factor Oct-1 (POU2F1) [215]. While Oct-1 is able to bind to DNA, it is not sufficient to complex with VP16, and instead requires an additional factor present in cell lysates [215]. This factor, named host cell factor (HCF) as a consequence, interacts with VP16, priming it for interaction with Oct-1 [216]. The resulting VP16-induced complex (VIC) binds to a TAATGARAT DNA motif (R = purine) present in viral immediate-early genes [217]. While the TAAT region of this motif resembles part of the Octamer motif to which Oct-1 binds [215], VIC formation depends on an intact GARAT motif [218] and is proposed to stabilize disordered regions of VP16 to promote DNA binding [219]. The absence of a strong DNA binding component of VP16, and the requirement for formation of this complex, is thought to enable control of the switch between latent and lytic modes of infection for HSV-1 [217, 220]. Furthermore, the capacity of VP16 to bind HCF-1 and Oct-1 in this manner is likely a consequence of viral mimicry, and suggests that HCF-1 may normally function to enable formation of trimeric complexes. In addition to its role in viral infection, HCF-1 is also essential in normal cells, particularly during development. The complete loss of *HCFC1*, the gene encoding HCF-1, is embryonic lethal [221] and leads to a failure of endoderm migration and primitive streak formation [222]. *HCFC1* is present on the X-chromosome, and knockout of one *HCFC1* allele leads to skewed X-inactivation during embryonic development of females [221].

An additional HCF family member, HCF-2, was identified based on homology to the amino-terminus of HCF-1 [223]. These proteins are vastly different in size, with HCF-1 being 2,035 amino acids in length, and HCF-2 only 792 amino acids. HCF-2 can substitute for HCF-1 in VIC formation, but is much less efficient in this role, both *in vitro* and *in vivo* [223]. The amino-terminal VIC domain of HCF-1, which is both necessary and sufficient for VIC formation [224], is largely conserved in HCF-2 (Figure 1-3A), with 69% identity and 83% similarity [223]. Substitution of the more divergent carboxy-terminus of the VIC domain in HCF-2, with that from HCF-1, is sufficient to rescue this deficit in VIC formation [223]. This may suggest a generic capacity for HCF-2 to substitute for HCF-1, where the defined role of HCF-2 in this context is to be less-productive than HCF-1. Overlapping with the VIC domain is a single fibronectin type-3 (Fn) repeat that is present in both HCF-1 and HCF-2 [225]. Beyond the four hundred amino-terminal residues, HCF-1 and HCF-2 have only patches of conservation, including an extended region at the carboxy-terminus of each protein, corresponding to two additional Fn3 repeats [225]. HCF-1, but not HCF-2, is cleaved at one of six proteolytic (PRO) repeats into its amino- and carboxy-termini [226]. The mechanism of this cleavage will be addressed below. At least in the case of HCF-1, self-association sequences (SAS) present in the Fn3 repeats interdigitate to enable the termini to remain non-covalently associated following cleavage [225, 227]. Although HCF-2 is not cleaved [223], the carboxy-terminal Fn3 repeats instead contribute to a nucleolar localization of the protein [228]. In addition to the basic and acidic regions, HCF-2 also lacks the NLS that is present at the very carboxy-terminus of HCF-1 [229]. For HCF-1, this NLS is necessary for nuclear localization, exclusion from the nucleolus, and VIC formation with full-length VP16 [228, 229].

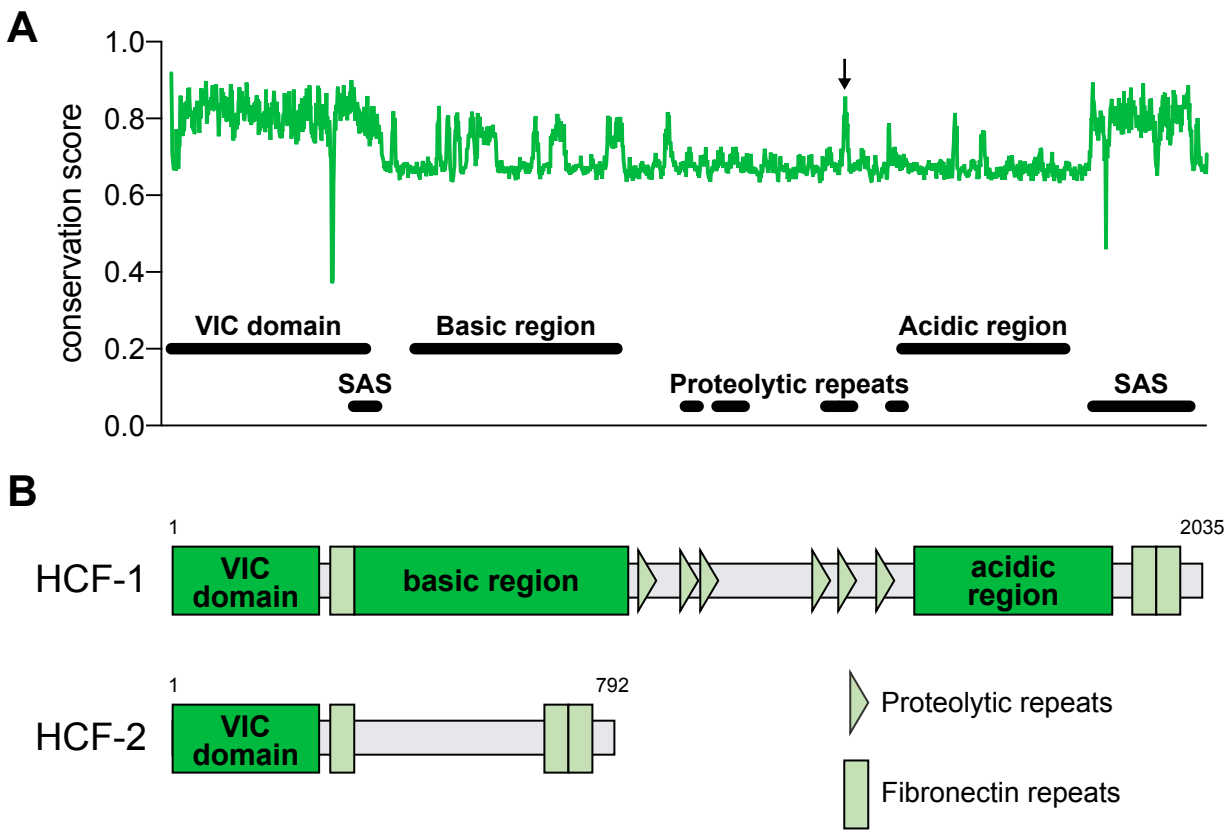


Figure 1-3: Primary sequence conservation of HCF-1 and HCF-2
(A) Conservation score of primary sequence for HCF-1 and HCF-2. Multiple sequence alignment was conducted using MUSCLE [22] and conservation scores by Jenson-Shannon divergence [23]. Arrow highlights a short region of conservation at the center of one proteolytic repeat. VIC = VP16 induced complex; SAS = self-association sequence, part of the fibronectin repeats. **(B)** Comparison of the domains present in, and length of, HCF-1 and HCF-2.

The appearance of multiple HCF family members corresponds with vertebrate development, with only a single member present in invertebrates, including *D. melanogaster* (*Hcf*) and *Caenorhabditis elegans* (*hcf-1*) [228]. For the most part, HCF-2 in vertebrates is very similar to that in humans [228]. One prominent exception to this is *Callorhinchus milii* (ghostshark or *makorepe*) and related species, in which Hcf-1 and Hcf-2 both contain VIC domains, Fn3 repeats, basic and acidic regions, an NLS, and two proteolytic repeats [228]. Both of these proteolytic repeats are degenerate in Hcf-2, while only one is in Hcf-1, and the acidic and basic regions of Hcf-2 are weakened compared to those in Hcf-1 [228]. The HCF homologs in *C. elegans* and *D. melanogaster* both contain the VIC domain, Fn3 repeats, and an NLS, but fruit flies also contain weakened acidic and basic regions, and a Taspase-1 cleavage site [228]. Based largely on the nature of the Hcf paralogs in *C. milii*, Gudkova *et al*. [228] have proposed the ancestral vertebrate HCF was more similar to HCF-1, and that, following gene duplication, HCF-2 progressively lost the acidic and basic regions, and the proteolytic repeat precursors. HCF-1 is ubiquitous, and HeLa cells have been reported to contain up to a hundred thousand molecules of HCF-1 [230]. In mice and humans, levels are highest in proliferating cells, and are expressed throughout the body during development, postnatally, and in adult tissues, including in skeletal muscle, kidney, and lungs [221, 231]. HCF-2 is similarly widely expressed, but is particularly high in the testis [223].

The association of HCF-1 with VP16 is dependent on the amino-terminal VIC domain [224]. This region is also often referred to as the Kelch domain, due to the presence of a repeating motif that was first observed in the *D. melanogaster* Kelch protein [224]. HCF-1 contains six Kelch repeats, suggesting that this region of HCF-1 forms a six-bladed β-propeller structure [224]. A temperature-sensitive mutation in this region, P134S, was identified in a BHK21 hamster cell line tsBN67, and found to cause a G_1 cell-cycle arrest after 1-2 days when incubated at the non-permissive temperature, leading to the conclusion that HCF-1 is essential for cell cycle progression [232]. A subset of these cells were also multi-nucleated, suggesting a defect in cytokinesis also occurred [233]. Furthermore, this mutation impairs the VP16–HCF-1 interaction, VIC formation, and transcriptional activation by VP16 [224, 232].

While these VP16-associated phenotypes can be rescued by expression of only the HCF-1$_{VIC}$ domain, this region alone is not sufficient to rescue the growth defects observed in the tsBN67 cells [224]. Instead, the entire amino-terminus of HCF-1, from residues one to 1011, is necessary to rescue cell cycle arrest caused by the P134S mutation [224]. Neither cleavage of HCF-1 nor interaction with the carboxy-terminus of HCF-1 is required for this rescue, and the amino-terminus cannot be expressed as two independent fragments, highlighting a scaffolding function and cooperation between the VIC domain and basic region of HCF-1 [224]. HCF-1 itself cannot bind directly to DNA, but its interaction partners enable it to be chromatin associated [230, 234]. Although it is the VIC domain that is both necessary and sufficient for chromatin association, the interaction between amino- and carboxy- termini of HCF-1 causes both subunits to precipitate with chromatin [230, 235]. The growth defects, but not HCF-1 chromatin binding, of tsBN67 cells can be rescued by expression of the SV40 large T antigen, which inactivates members of the retinoblastoma protein family [236].

Interestingly, forced overexpression of HCF-2 causes its redistribution from the nucleolus and slowed cell growth [228]. This may give credence to the idea that HCF-2 can compete with HCF-1 in the formation of complexes around its VIC domain. With the knowledge that HCF-1 is necessary for cell growth, if HCF-2 is available to bind in place of HCF-1, even if this association is relatively inefficient and perturbs only a fraction of HCF-1 function, this may be sufficient to cause the growth defect observed by Gudkova *et al.* [228]. This may also explain the vastly different sizes of the HCF family members; HCF-2 is localized to the nucleolus by its Fn3 repeats, the VIC domain is the only other region that would be required to compete with most of HCF-1 function, and the entire amino-terminus of HCF-1 is necessary for rescue of growth defects caused by *HCFC1* knockdown. Indeed, HCF-2 may be sequestered to the nucleolus to prevent it interfering with HCF-1 function, but under certain circumstances, such as cell stress or a stage of the cell cycle, levels of HCF-2 may increase to a threshold that instead causes its exclusion from the nucleolus and perturbation to HCF-1 complexes.

HCF-1 is synthesized as a 2,035 amino acid precursor, but undergoes post-translational, proteolytic cleavage [237]. Between the basic and acidic regions of HCF-1 are six conserved 26 amino acid repeats [226] that can be independently subject to cleavage [238, 239], and two additional divergent repeats that are resistant to cleavage (Figure 1-3A). This cleavage was initially mapped to a core CE/TH sequence within each repeat [238, 239], with the glutamic acid critical to cleavage and later found to be eliminated during the process [240]. Although HCF-2 is not proteolytically processed, in humans it does contain a series of conserved residues within the fifth proteolytic repeat of HCF-1 (Figure 1-3A; arrow), including the ETH cleavage sequence. While difficult to make a definitive conclusion on this, it is possible that this is a remnant of a proteolytic cleavage site. The resulting amino- and carboxy-termini of HCF-1 vary in size, ranging from 1,018 to 1,422 amino acids for the amino-terminus and 612 to 1,016 amino acids for the carboxy-terminus, and have been suggested to have different functions within the cell [241]. These proteolytic repeats are unique to HCF-1 [226], and this is perhaps unsurprising considering the mechanism of cleavage is also unique to HCF-1.

Proteolytic cleavage of HCF-1 is catalyzed by OGT [242], an enzyme that is normally responsible for O-linked glycosylation (O-GlcNAcylation) of proteins, using uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) as a donor substrate [243]. Small molecule inhibition of OGT [242, 244] or mutation of its active site, specifically of a residue that binds to and activates UDP-GlcNAc [240], is sufficient to impair cleavage of HCF-1. Indeed, both OGT and UDP-GlcNAc must be present for cleavage to occur, and the structure of the proteolytic repeats within the OGT active site points to an interaction between the proteins that is akin to that which occurs during O-GlcNAcylation [242]. In this sense, OGT is thought to be performing its normal function of O-GlcNAcylation, but the nature of the proteolytic repeats in HCF-1 leads to their cleavage. Additionally, HCF-1 is heavily O-GlcNAcylated by OGT within its amino-terminus, and the interaction between OGT and the PRO repeats is necessary for this to happen [242]. Even after glycosylation, OGT interacts with the basic region [245], and it has been reported that about half of nuclear OGT is stably associated with HCF-1 [246]. Similar to HCF-1 in

humans, Hcf in *D. melanogaster* (dHcf) is proteolytically cleaved into its amino- and carboxy-termini, and the resulting termini remain associated [247]. However, dHcf lacks the proteolytic repeats present in HCF-1, and is instead cleaved through the activity of taspase-1 [248]. The dependency of HCF-1 cleavage by OGT on UDP-GlcNAc is striking, particularly as the levels of this metabolite can be dictated by the nutrient status of the cell [243]. While the extent of O-GlcNAcylation of HCF-1_N is determined by glucose levels [249], there is not yet evidence that cleavage of HCF-1 is affected in a similar manner. Indeed, it may be that availability of UDP-GlcNAc is greater, or actively prioritized, where HCF-1 cleavage occurs. Thus, creating a buffer that enables continuous processing of HCF-1. While it is reported that cleavage and O-GlcNAcylation can occur essentially simultaneously [250], the latter could be reversed following HCF-1 cleavage through the activity of O-GlcNAcase (OGA), allowing flexibility in the O-GlcNAcylation of HCF-1.

Despite the post-cleavage interaction between the amino- and carboxy-termini of HCF-1, each terminus is reported to have a different contribution to the function of HCF-1 in cells [241]. Similar to tsBN67 cells, knockdown of HCF-1 using RNA interference in HeLa cells causes a G_1 cell cycle arrest and multi-nucleation [241]. Only the former of these phenotypes can be completely rescued by overexpression of HCF-1_N, with the latter rescued only with HCF-1_C [241]. Although the rescue of the multi-nucleation by HCF-1_C appears to be inconsistent with the presence of this phenotype in tsBN67 cells, in which endogenous HCF-1_C is still present, the P134S mutation in these cells causes cytosolic distribution of both HCF-1_N and HCF-1_C [230]. Curiously, expression of cleavage-resistant HCF-1 is unable to rescue the multi-nucleation phenotype [241], suggesting cleavage is necessary for independent functions of the termini of HCF-1. Consistent with this, cleavage has been reported to alter interactions made by HCF-1 [251]. While the understanding of the HCF-1 cleavage mechanism is both intriguing and extensive, specifically what the effect on HCF-1 function is and how this effect is achieved, why it is catalyzed by OGT, why there are six different cleavage sites, and what the benefit is of the termini remaining associated, is definitively unclear. Indeed, what we do understand of HCF-1 cleavage appears to be the lowest hanging fruit, and has, perhaps, created more questions than it has

given answers. One of the first steps towards answering at least a subset of these questions would be to interrogate the independent functions of the amino- and carboxy-termini. While this has been done to an extent using knockdown of *HCFC1* mRNA and add-back of either the amino-terminus or full-length protein, these types of experiments have failed to address the molecular function of the termini.

HCF-1 as a scaffolding protein

HCF-1 interacts with a number of different proteins, the majority of which bind through the VIC domain and are involved in transcriptional regulation (Figure 1-4B). Proteins that interact with the VIC domain, including VP16, do so through an HCF-1 binding motif (HBM), which has the canonical sequence $^{D}/_{E}$-H-x-Y [252]. This motif has been identified and validated in many proteins, including transcription factors and histone modifying enzymes (Figure 1-5B). As a scaffolding protein, the primary responsibility of HCF-1 appears to be tethering transcription factors to enzymes or enzyme complexes, leading to modulation of transcriptional output. A number of transcription factors have been implicated in the recruitment of HCF-1 to specific genes, with the idea that HCF-1, either subsequently or simultaneously, recruits additional complexes. While this is one aspect of HCF-1 function, its relationship with VP16 would also suggest that it can force proteins to make additional interactions. Whether this can occur with endogenous interaction partners in a similar manner, remains to be seen. Regardless of the mechanism, HCF-1 has primarily been implicated in regulating the expression of genes involved in the cell cycle, mitochondrial biogenesis, and metabolism.

The mechanism of recruitment of HCF-1 to E2F-responsive cell cycle genes has been attributed to interaction with both THAP domain-containing transcriptional regulators [234] and E2F1 [253]. HCF-1 interacts with multiple members of the E2F family of transcription factors, including the activating E2F1 and the repressive E2F4, and this is believed to underpin much of the influence of HCF-1 over the cell cycle [254]. These interactions are cell cycle phase-dependent, with the E2F1 interaction occurring primarily during late G_1 and S phases, and the E2F4 interaction during early G_1 and S phases [254]. HCF-1 is predominantly enriched at E2F-responsive genes during early G_1, G_1 to S, and S phases, and co-binds with E2F family members at phases relevant to their interaction (i.e. with E2F1 during G_1 to S

and S, and with E2F4 during early G_1 and S) [254]. Knockdown of HCF-1 results in attenuated expression from these E2F target genes [253, 254], and impairing this interaction reduces E2F1-dependent DNA damage and apoptosis [253]. Tyagi and Herr [253] credit interaction with E2F1 for the recruitment of HCF-1, with overexpression of an E2F1 HBM mutant causing a reduction of HCF-1 at chromatin compared to overexpression of wild-type E2F1. In contrast, Parker *et al.* [234] found that knockdown of E2F1 does not influence HCF-1 recruitment to chromatin. Instead, they identified a strong correlation of chromatin binding levels for HCF-1, THAP11, and ZNF143, and discovered interdependent recruitment of these proteins to E2F-responsive genes [234]. HCF-1 interacts with various THAP family members in an HBM-dependent manner [255, 256], and a number of these interactions, including with THAP1, THAP3, and THAP11, have been attributed to the recruitment of HCF-1 to chromatin and transcriptional activation [234, 256, 257]. Tyagi and Herr [253] and Parker *et al.* [234] use two different approaches to reach two different conclusions about the role of E2F1 in the recruitment of HCF-1 to chromatin. One possible cause of this distinction is that E2F1 can recruit HCF-1, but not when it is expressed at physiological levels. How ever HCF-1 is localized to E2F-responsive genes, it is clear that its association with both E2F family members and THAP11 contributes to the transcriptional output from these genes, consistent with the effect that loss of HCF-1 has on cell cycle progression.

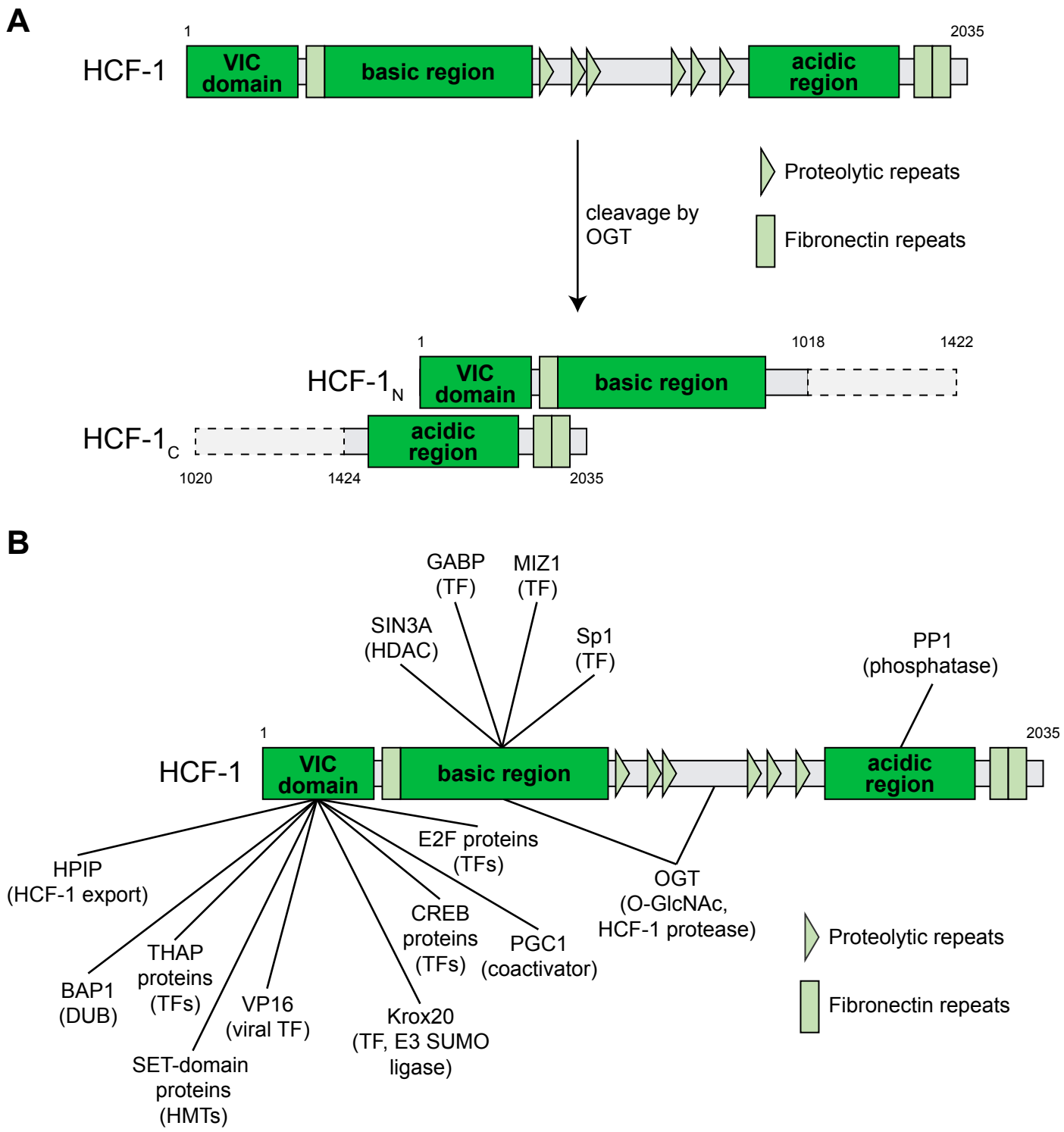


Figure 1-4: Proteolytic cleavage and interaction partners of host cell factor-1
(A) HCF-1 contains six conserved proteolytic repeats that are recognized by O-GlcNAc transferase (OGT), leading to cleavage into amino- and carboxy-termini. These termini remain associated through self-association sequences in the fibronectin repeats. **(B)** HCF-1 interacts with a diverse set of proteins, including transcription factors (TF), histone methyltransferases (HMT), deubiquitylating enzymes (DUB), and histone deacetylases (HDAC).

Little is known about how HCF-1 turnover is controlled, with the deubiquitylating enzyme BRCA associated protein 1 (BAP1) the only interaction partner linked to the UPS. HCF-1 is polyubiquitylated in both its amino- and carboxy-termini, including lysine-48 and lysine-63-linked chains [258, 259]. BAP1 interacts through the HCF-1_{VIC} domain [258], and appears to prioritize deubiquitylation of lysine-48-linked polyubiquitylation [259]. This type of polyubiquitylation is typically considered a marker for degradation, but how knockdown of BAP1 affects HCF-1 stability is disputed. Relatively short-term knockdown experiments have suggested that BAP1 either does not influence steady-state HCF-1 [260, 261] or destabilizes it [259]. However, conditional knockout of BAP1 has been reported to cause a near-complete loss of HCF-1 protein [262]. One would expect that a deubiquitylating enzyme, such as BAP1, would lead to the stabilization of a target protein, making the finding that BAP1 knockdown reduces HCF-1 protein levels more expected. However, it is also possible that HCF-1 is not a target of BAP1 enzymatic activity, and instead their association has a different function. Any previously observed changes to HCF-1 levels in response to BAP1 loss may instead be an indirect effect. BAP1 forms the catalytic core of the polycomb repressive-deubiquitinase (PR-DUB) complex, which additionally includes ASXL regulatory subunits and is responsible for removal of ubiquitin from lysine-119 of histone-H2A (H2AK119Ub) [263]. While the exact mechanism has yet to be elucidated, the interaction between HCF-1 and BAP1 has also been shown to contribute to BAP1 function, including its regulation of cell growth [258]. Thus, it remains possible that HCF-1 scaffolds BAP1 and PR-DUB to target genes through its additional interactions. It has also been reported that interaction between HCF-1_{VIC} and heat-shock protein 90 (HSP90) impairs autophagy-dependent turnover of HCF-1, and enables transcriptional regulation of HCF-1 target genes [264].

In addition to BAP1, HCF-1 has been found to interact with both repressive and activating histone modifying complexes. Amongst these is an interaction between HCF-1_{VIC} and HMT SET-domain containing proteins, including SETD1A and MLL1 [245]. Complexes containing these proteins are typically associated with transcriptional activation through methylation of lysine-4 of histone-H3 (H3K4me) [265], and the HCF-1 amino-terminus co-immunoprecipitates this form of methyltransferase

activity [245]. Indeed, HCF-1 recruits SETD1A- and MLL1-containing complexes to a subset of target genes, contributing to their transcriptional activation [254]. HCF-1 is able to simultaneously interact with SETD1A and the HDAC SIN3A [266]. With SIN3A a repressive regulator, SETD1A and SIN3A have opposite effects on transcriptional output, highlighting the differential influence that HCF-1 can have on gene expression and suggesting that it is able to fine-tune transcription. The ability of HCF-1 to interact with SETD1A and SIN3A is influenced by additional interactions made by HCF-1. VP16 and E2F1 show a preference for interacting with HCF-1 when it is in complex with SETD1A, whereas E2F4 favors HCF-1 complexed with SIN3A, reflecting the repressive function of E2F4 [254, 266]. This may be a second important reason for the interaction between VP16 and HCF-1; knockdown of SETD1A reduces expression of HSV-1 immediate early genes [267], suggesting that HCF-1 may also contribute to expression of these genes through recruitment of HMT activity. During cell cycle phases whereby E2F1 and HCF-1 interact, there is enrichment of SET complex members at E2F-responsive genes [254]. HCF-1 has also been identified in association with the HAT complexes non-specific lethal (NSL), STAGA, and ATAC [268, 269], although if and how it contributes to the function of these is unclear. However, HCF-1 may similarly scaffold these complexes to transcription factors and their target genes.

A number of HCF-1 interaction partners are involved in mitochondrial function and metabolism, and for the majority of these, HCF-1 serves as a mediator for association with OGT. This is particularly striking as OGT is considered to be a metabolic integrator, with its function depending on the levels of UDP-GlcNAc, the levels of which are in-turn dictated by the nutrient-sensing hexosamine biosynthetic pathway [270]. Through its basic region, HCF-1 interacts with GA-binding protein (GABP) [271]. GABP a multi-subunit (GABPα and β) transcription factor that regulates expression of genes involved in mitochondrial biogenesis and energy production in response to mitogen and stress signaling [272, 273], with this interaction occurring through a region of GABPβ that is necessary for its transactivation [271]. PGC-1α (*PPARGC1A*), PGC-1β (*PPARGC1B*), and PGC-1 related cofactor (PRC) are critical regulators of mitochondrial biogenesis and oxidative phosphorylation [274], and can form an HBM- and HCF-1-dependent complex with GABPβ [275, 276]. HCF-1 is also able to form a bridge between PGC-1α and

OGT, with subsequent O-GlcNAcylation of PGC-1α promoting interaction with BAP1 leading to its deubiquitylation and stabilization [277]. Consistent with this, HCF-1 and OGT promote PGC-1α-dependent expression of gluconeogenesis genes [277]. There are a number of reports identifying interaction of HCF-1 with nuclear factor erythroid 2-like factor 1 (NFE2L1), with one suggesting this occurs through a non-canonical HBM in NFE2L1 [278], and the other finding that HCF-1_C is responsible for their association [279]. Regardless of the interaction region, this is proposed to tether OGT to NFE2L1, enabling its O-GlcNAcylation and stabilization by competing for ubiquitylation-promoting phosphorylation events [278, 279]. OGT-induced stabilization of NFE2L1 promotes the expression of antioxidant responsive [278] and proteasome subunit [279] genes. More recently, Lane *et al.* [249] performed an extensive interrogation of the interaction between HCF-1 and ChREBP, a transcription factor involved in *de novo* lipogenesis. In this work, they found O-GlcNAcylation of HCF-1 by OGT under high-glucose primes it for interaction with ChREBP, and additionally leads to O-GlcNAcylation of ChREBP [249]. Glucose stimulation results in ChREBP-dependent recruitment of HCF-1, and subsequent HCF-1-mediated deposition of H3K4me3 marks [249]. This sequence of events occurs at lipogenic genes and promotes their expression under conditions conducive to lipogenesis. Through these interactions, HCF-1 appears to contribute to the expression of a vast array of genes involved in various metabolic and mitochondrial processes. These interactions and corresponding target genes may be the primary reason that half of nuclear OGT remains associated with HCF-1 following its cleavage, as it would enable the cellular nutrient state to influence their transcription. For PGC-1α, NFE2L1, and ChREBP, association with HCF-1 and, as a result, OGT, has a positive effect on their function, consistent with good nutrient levels promoting OGT function to enable expression of metabolic genes.

While the cofactors mentioned up until this point have undergone reasonable interrogation, HCF-1 also has ample interaction partners that are substantially less understood, or not understood at all. Amongst these are HBM-dependent interactions with the cyclic AMP response element binding (CREB) protein family, including CREB3 and CREBZF [252, 280, 281]. The interaction of HCF-1 with CREBZF is

thought to limit transcriptional activation by CREB3 [282] and of HSV immediate-early genes by VP16 [281], possibly by mutually exclusive interaction with the VIC domain. Also through the VIC domain, HCF-1 interacts with, and promotes transcriptional activation by, Krox20 [283], and can be shuttled between the cytoplasm and nucleus by HCF-1 regulator 1 (HCFC1R1) [284]. While the latter of these associations is intriguing, and suggests that HCF-1 activity might be regulated by its sub-cellular distribution, the context in which this occurs is unknown. HCF-1 contributes to the nuclear localization of VP16, and Mahajan *et al.* [284] proposed that nuclear export of HCF-1 by HCFC1R1 would create a pool of cytoplasmic HCF-1 for this function. As this is also likely a consequence of mimicry by VP16, this instead suggests that a cytoplasmic pool of HCF-1 is important for nuclear localization of endogenous proteins. Finally, an interaction between HCF-1 and Miz-1 impairs Miz-1-dependent transcriptional activation [285]. One of the two regions in Miz-1 that associates with HCF-1 is also responsible for interaction with MYC, and their mechanism of opposing Miz-1 function (p300 recruitment) appears to be analogous [285]. Yeast two-hybrid assays have also identified interaction between the HCF-1 basic region and Sp1 [286], although the function of this interaction in normal cells has not been extensively interrogated. Finally, HCF-1_C directly binds with protein phosphatase 1 (PP1) catalytic subunit (PPP1CC) and this complex co-immunoprecipitates phosphatase activity [287], but its function is similarly unknown.

Here I have summarized the majority of what is known of HCF-1, both of its interaction partners and the genes it targets. The extent to which these are understood varies greatly, but the general theme appears to be that HCF-1 tethers transcription factors to either histone modifying complexes or to OGT. The former has mostly been linked to cell cycle genes, and the latter to genes for metabolism and mitochondrial biogenesis. I also expect that by understanding additional HCF-1 cofactors, we would gain a better grasp of its target genes, which currently appears to be an underestimation.

Role of HCF-1 in disease

While the HCF-1 protein was discovered for its role in viral infection, the gene encoding HCF-1, *HCFC1* (or *MRX3*), was identified for being a driver of X-linked intellectual disability [288]. Indeed, a mutation within the 5' untranslated region (UTR) of *HCFC1* affects binding of the transcription factor Yin-Yang-1 (YY1), leading to overexpression of *HCFC1 [289]*. This can impair the expression of genes involved in mitochondrial function, including those involved in mitochondrial ribosome biogenesis, and is thought to increase production of astrocytes and decrease neurite growth [289]. Neurological defects are also associated with an X-linked cobalamin disorder caused by mutations within the HCF-1_{VIC} domain [290]. These mutations are proposed to alter the structure of HCF-1_{VIC}, leading to impaired transcriptional activation of *MMACHC*, which encodes an enzyme involved in cobalamin metabolism [290].

Multiple reports have also implicated HCF-1 in non-alcoholic steatohepatitis (NASH) [249, 291], a severe form of non-alcoholic fatty liver disease (NAFLD) that can precede cirrhosis and hepatocellular carcinoma (HCC). However, the role of HCF-1 in NASH is somewhat conflicting, with Lane *et al.* [249] finding that HCF-1 levels are increased in NASH biopsies, while Minocha *et al.* [291] demonstrated that conditional knockout of HCF-1 from hepatocytes causes development of NASH in mice. Lane *et al.* [249] focussed on the contribution of overexpressed ChREBP to NAFLD progression, and how its interaction with HCF-1 promotes expression of lipogenic genes. Minocha *et al.* [291] also identified metabolic genes regulated by HCF-1, but instead emphasized its role in stabilizing PGC-1α to promote expression of genes involved in mitochondrial function, with the loss of HCF-1 and subsequently PGC-1α leading to malformed mitochondria and NAFLD. While HCF-1 may contribute to the development of cancer through NASH, it has additionally been found to be overexpressed in a variety of different cancer types [292] and this may be correlated with a poor prognosis [293]. Indeed, HCF-1 is associated with a number of well-defined proto-oncogenes, including MYC and MLL1, and tumor suppressors, such as BAP1. Furthermore, its known roles in cell cycle progression and promoting the expression of genes for mitochondrial biogenesis would make it an excellent mediator in the oncogenic

phenotype. Despite these, the contribution of HCF-1 to cancer, particularly on a molecular level, has not yet been investigated.

Interaction between MYC and Host Cell Factor-1

The Tansey laboratory first detected an association between MYC and HCF-1 through a proteomics screen using the MYC central portion [179]. Deletion mapping determined the carboxy-terminus of the central portion, which includes MbIV, was responsible for this interaction [179]. This region contains a strongly conserved series of amino acids, Q-H-N-Y-A-A (Figure 1-5A), that resembles a non-canonical HCF-1 binding motif (HBM) [179]. Although the canonical HBM has been defined as $^{D}/_{E}$-H-x-Y [252], asparagine is the first residue of the BAP1 HBM (Figure 1-5B) [294], and flexibility of the first residue has not been thoroughly interrogated. Thus, we proposed that the Q-H-N-Y sequence in MbIV is a putative HBM. Validation of the MYC HBM would both warrant and enable investigation into the nature of canonical versus non-canonical HBMs, a concept that has evaded appropriate interrogation. Specifically, this could be exploited to determine if the HBM definition should be expanded, or if non-canonical HBMs cause a weakened interaction with HCF-1.

We validated the MYC–HCF-1 interaction by viral transduction and co-immunoprecipitation of MYC containing either a deleted (Δ4) or alanine substituted (4A) putative HBM, finding that both of these were sufficient to disrupt interaction with HCF-1 (Figure 1-5C) [179]. While this provided initial validation of the MYC HBM, we had yet to provide evidence that this association occurred through the HCF-1_{VIC} domain, which is the interaction surface utilized by HBM-containing proteins. Co-immunoprecipitation of endogenous MYC additionally revealed this interaction occurs in the absence of forced over-expression (Figure 1-5D) [179]. Furthermore, the interaction between Myc and Hcf is conserved in *D. melanogaster*, but occurs independently of a canonical HBM present in Myc, and is proposed to

contribute to transcriptional regulation by Myc through the recruitment of histone modifying complexes [295].

While we prepared this first manuscript for publication [179], Dingar *et al.* [296] also highlighted HCF-1 as a possible MYC interaction partner based on BioID and overlap of ENCODE datasets. This was later validated by proximity-ligation assay (PLA), with the MYC–HCF-1 interaction found to be largely nuclear [297]. Kalkat *et al.* [15] also performed BioID, comparing the impact of MYC box deletion on interaction partners, and found the MbIV deletion reduced association with HCF-1 and various HCF-1 cofactors, including NFE2L1, GABP, E2F3, and PPP1CC. I mentioned above that, as a scaffolding protein, HCF-1 is commonly found to tether transcription factors to enzymes, such as the HMT SETD1A, the HDAC SIN3A, or OGT. Although both SIN3A and OGT co-immunoprecipitated with MYC, neither were affected by deletion of MbIV [15]. While it is difficult to draw specific conclusions from this in the absence of validation, it may suggest that the function of HCF-1 in the context of its interaction with MYC is distinct from its previously identified roles.

We assessed the contribution of the MYC–HCF-1 interaction to tumor growth by overexpressing MYC, WT or mutant, by viral transduction of mouse 3T3 fibroblasts, and injecting the cells into the flanks of nude mice [179]. Two mutations in MYC were utilized in these experiments; the 4A mutation described above and a WBM mutation that is known to perturb the association of MYC and WDR5, without affecting the MYC–HCF-1 interaction [179]. In this context, we found that the MYC 4A-expressing cells grew slower (Figure 1-5E) and resulted in a smaller tumor mass (Figure 1-5F). While this work was the first to determine an interaction partner for MbIV and verify the importance of this region to MYC-driven tumorigenesis, the functional significance of the MYC–HCF-1 connection remains unexplored. Even with the accumulating evidence surrounding the MYC–HCF-1 interaction, there is a distinct absence of an interrogation into the elements defining their association, and no molecular, transcriptional, or cellular functions have been determined for their relationship. By assigning a function to the MYC–HCF-1 interaction, I hope to gather the when, where, and why this interaction would occur, and gauge

how it might contribute to both normal cell growth and MYC-driven tumorigenesis. Small-molecule inhibitors against cofactors of MYC are likely our best opportunity to disable MYC function itself. HCF-1 is, in theory, a possible target in this respect, but understanding its function with MYC must first be achieved.

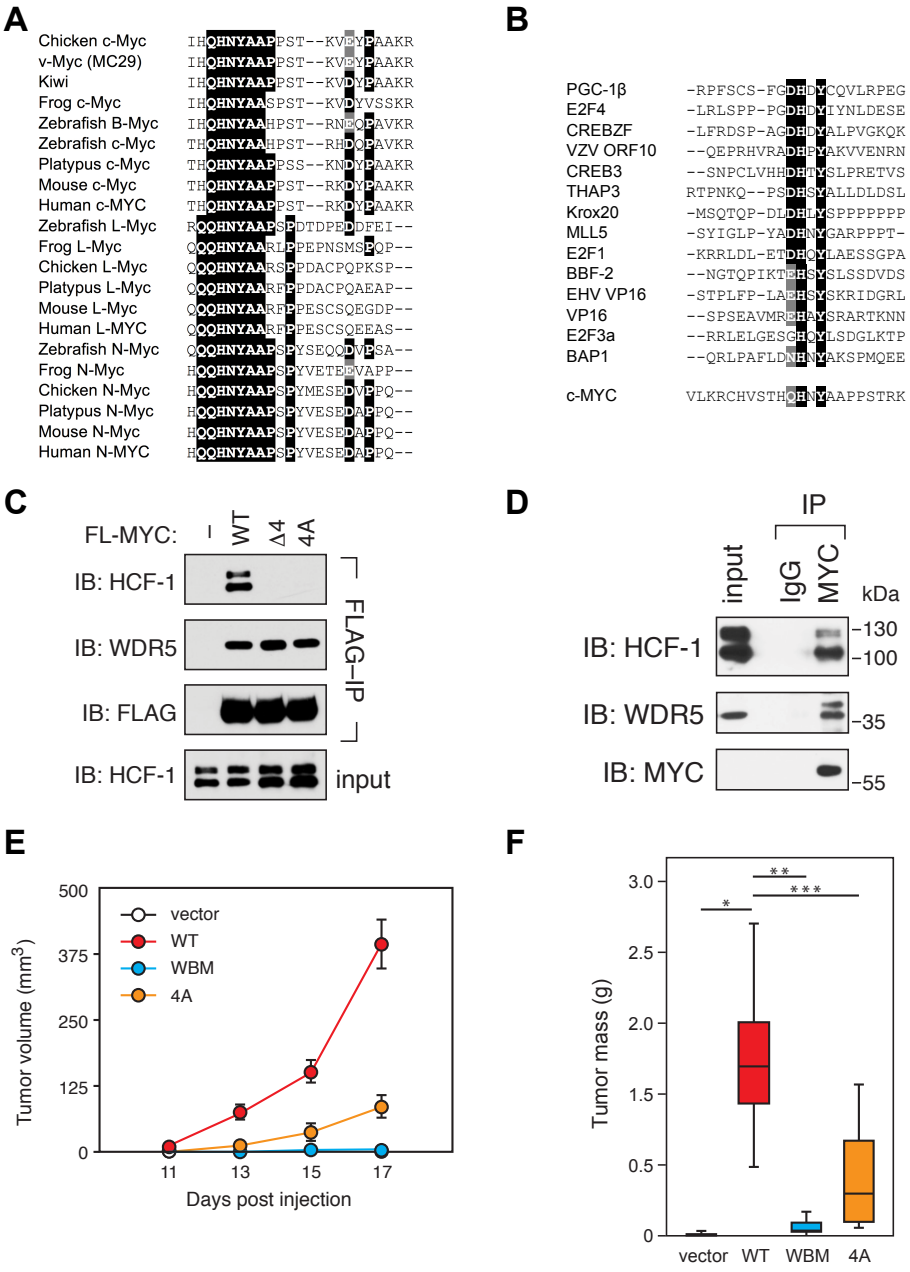


Figure 1-5: HCF-1 is a MYC interaction partner
Figures are from Thomas *et al.* [179]. **(A)** Alignment of MbIV across MYC family members in different species. **(B)** Examples of proteins containing validated HCF-1 binding motifs (HBMs), and a comparison to the putative HBM present in MbIV. The canonical HBM sequence is $^{D}/_{E}$-H-x-Y. **(C)** FLAG-tagged full-length MYC, wild-type (WT) or with deletion (Δ4) or alanine substitution (4A) of the HBM, was stably transduced into HEK293 cells, and M2 agarose were used to recover MYC and associated proteins. Both the deletion and alanine substitution mutations reduced interaction with HCF-1 without affecting interaction with WDR5. **(D)** Endogenous MYC was immunoprecipitated from MG132-treated HEK293 cells, and the resulting eluate probed for HCF-1_C and WDR5 by western blotting. **(E)** and **(F)** NIH3T3 fibroblasts stably transduced with MYC, WT or containing the WDR5 binding motif (WBM) or 4A mutations, were injected into the flanks of athymic nude mice. Shown is the tumor volume (E) and the tumor mass (F).

Chapter II

Materials and methods

The following materials and methods are also described in my manuscript [182].

Primers and cloning

See Table 2-1 for primer sequences. PCRs were performed using either Q5 DNA Polymerase (NEB, Ipswich, Massachusetts, M0491) or OneTaq DNA Polymerase (NEB M0480). Gibson assemblies were performed using Gibson Assembly Cloning Kit (NEB E5510). More specific details about cloning steps can be found in the relevant sections.

Table 2-1: Primer sequences
Oligonucleotides used for cloning, Southern blot probe generation, and EMSA.

Primer name	Sequence	Primer name	Sequence
pCRIS-HCFC1N_F	TCTAGAATGCTGATGGGC	H307A_F	atcaggcCAACTACGCAGCGCCTCC
pCRIS_R	tactcggtcatggatccgggTATGTAACGCGGAACTCC	H307A_R	gagtactGACGTGGCACCTCTTGAG
Puro-1	cccggatccATGACCGAGTACAAGCCC	Y309A_F	TCCCTCCACTaGGAAGGACTAT
Puro-FLAG_R	cttatcgtcgtcatccttgtaatcAGGTCCAGGGTTCTCCTC	Y309A_R	GGCGCTGCGgcGTTGTGCTGATGTGTGGAG
FLAG-FKBP_F	gattacaaggatgacgacgataagGGAAGCGGAGGAGTGCAGG	V264G_F	CccCATCGATTTCTTCCTCATCTTCTTG
FKBP_R	tagcccatcagcattctagaAGATCCGCCGCCACCCGA	V264G_R	TcTCTGTGGAAAAGAGGCAGG
pCRIS-HCFC1N_R	actgtcgagcgcctaggctcaagaaGGATCCGGGTATGTAACG	MYC-151_F	TCCTACCAGGCTGCGCGC
HCFC1N_F	TTCTTGAGCCTAGGCGCTCG	MYC-151_R	GGATCCGCCCGGGCTCTTATC
HCFC1N-5'Hom_R	tgttatcctcctcgcccttgctcacGTTGGCGGGCGACACGGC	MYC-delta331_F	TAAGAATTCGATATCAAGCTTATC
mCherry_F	GTGAGCAAGGGCGAGGAG	MYC-delta331_R	GACACTGTCCAACTTGAC
mCherry_R	CTTGTACAGCTCGTCCAT	VP16 HBM:H307G_F	gcgtatGCAGCGCCTCCCTCCACT
HCFC1-mCherry-FKBP_F	cggcggcatggacgagctgtacaagGGAAGCGGAGCTACTAACTTCAGCC	VP16 HBM:H307G_R	gccctcATGTGTGGAGACGTGGCACC
HCFC1-mCherry-FKBP_R	attagctggactgacagcgctggcAGATCCGCCGCCACCCGA	5'_F	CGATTTCGATTCCTCTGC
HCFC1N-3'Hom_F	gccagcgctgtcagtccagctaatTTGCCAGCGGTGCTTCTG	5'_R	CCAGATCTGCTATCTCTCC
HCFC1N-3'Hom_R	tttgctagcccatcagcattctagaTCGGCGGATGTTAAACTCTG	GFP_F	gggCCATGGTGAGCAAGGG
HCFC1N-sgRNA_F	gctggcaagtGTTTTAGAGCTAGAAATAGCAAG	GFP_R	gggagatctGTACAGCTCGTCCATGC
HCFC1N-sgRNA_R	ggtgcttctgGGTGTTTCGTCCTTTCCAC	T7-HCF1_F	tggccaacagatgggtCGCCATATGGAGCTCCTC
4A_F	ggccgcATGTGTGGAGACGTGGCACCTC	T7-HCF1_R	cctgtcatcgatgccatCCCTTCGATATGGTGATGG
4A_R	gccgcgGCAGCGCCTCCCTCCACTCG	E-box_F-biotin	GCTCAGGGACCACGTGGTCGGGGATC-biotin
VP16 HBM_F	gcgtatGCAGCGCCTCCCTCCACT	E-box_F	GCTCAGGGACCACGTGGTCGGGGATC
VP16 HBM_R	gtgctcATGTGTGGAGACGTGGCACC	E-box_R	GATCCCCGACCACGTGGTCCCTGAGC
GUIDE MYC-1A	cttgtgcgtaGTTTTAGAGCTAGAAATAGCAAG	Mutant E-box_F	GCTCAGGGACCAGCTGGTCGGGGATC
GUIDE MYC-1B	agttccgtagCGGTGTTTCGTCCTTTC	Mutant E-box_R	GATCCCCGACCAGCTGGTCCCTGAGC
H307G_F	gACACATCAGggCAACTACGCAGC		
H307G_R	GACACGTGGCACCTCTTG		

Cell culture

293T cells were maintained in DMEM with 4.5 g/l glucose, L-Glutamine, and sodium pyruvate (Corning, Corning, New York, 10-013-CV), and supplemented with 10% fetal bovine serum (FBS, Denville Scientific, Metuchen, New Jersey, FB5001-H), and 1% penicillin/streptomycin (P/S, Gibco, Waltham, Massachusetts, 15140122). Ramos cells were obtained from the ATCC (Manassas, Virginia, CRL-1596) and maintained in RPMI 1640 with L-glutamine (Corning 10-040-CV), and supplemented with 10% FBS (Denville Scientific FB5001-H), and 1% P/S (Gibco 15140122).

Antibodies

Rabbit anti-$HCF1_C$ polyclonal (Bethyl Laboratories, Montgomery, Texas, Cat# A301-399A); rabbit anti-$HCF1_N$ polyclonal [258]; rabbit anti-c-MYC monoclonal (Abcam, Cambridge, United Kingdom, Cat# ab32072); rabbit anti-HA (C29F4) monoclonal (Cell Signaling Technology, Danvers, Massachusetts, Cat# 3724); rabbit anti-WDR5 (D9E1I) monoclonal (Cell Signaling Technology Cat# 13105); rabbit anti-IgG polyclonal (Cell Signaling Technology Cat# 2729); mouse anti-FLAG monoclonal, HRP-conjugated (Sigma-Aldrich, St. Louis, Missouri, Cat# A8592); mouse anti-T7 monoclonal, HRP conjugated (Millipore, Burlington, Massachusetts, Cat# 69048); mouse anti-GAPDH monoclonal, HRP conjugated (Thermo Fisher Scientific, Waltham, Massachusetts, Cat# MA5-15738-HRP); rabbit anti-GAPDH (D16H11) monoclonal, HRP conjugated (Cell Signaling Technology Cat# 8884); goat anti-rabbit IgG monoclonal, HRP conjugated (Thermo Fisher Scientific Cat# 31463); mouse anti-rabbit IgG monoclonal, light chain specific, HRP conjugated (Jackson ImmunoResearch Laboratories, West Grove, Pennsylvania, Cat# 211-032-171); rabbit anti-histone H3 (D1H2) XP monoclonal, HRP conjugated (Cell Signaling Technology Cat# 12648); rabbit anti-α-Tubulin (11H10) monoclonal, HRP conjugated (Cell Signaling Technology Cat# 9099); goat anti-rabbit IgG polyclonal, Alexa Fluor 594 conjugated (Thermo Fisher Scientific A11012).

Transient transfection, western blotting and immunoprecipitation

pCGT-$HCF\text{-}1_{VIC}$ and pFLAG-$VP16_N$ constructs are described in Thomas *et al.* [179]. Q5 site-directed mutagenesis of pFLAG-MYC WT [178] was used to generate the following plasmids: pFLAG-MYC 4A

(4A_F and 4A_R), pFLAG-MYC H307G (H307G_F and H307G_R), pFLAG-MYC H307A (H307A_F and H307A_R), pFLAG-MYC Y309A (Y309A_F and Y309A_R), pFLAG-MYC VP16 HBM (VP16 HBM_F and VP16 HBM_R), and pFLAG-MYC VP16 HBM:H307G (VP16 HBM:H307G_F and VP16 HBM:H307G_R). pFLAG-MYC WBM is from Thomas *et al.* [178]. Q5 site-directed mutagenesis of FGH-MYC CP WT [179] was used to generate the following plasmids: FGH-MYC CP 4A (4A_F and 4A_R), FGH-MYC CP H307G (H307G_F and H307G_R), FGH-MYC CP H307A (H307A_F and H307A_R), FGH-MYC CP Y309A (Y309A_F and Y309A_R), FGH-MYC CP V264G (V264G_F and V264G_R), FGH-MYC CP VP16 HBM (VP16 HBM_F and VP16 HBM_R), and FGH-MYC CP VP16 HBM:H307G (VP16 HBM:H307G_F and VP16 HBM:H307G_R). The transactivation (MYC-151_F and MYC-151_R) or DNA binding (MYC-delta331_F and MYC-delta331_R) domains were deleted from pFLAG-MYC WT, 4A, and VP16 HBM plasmids by Q5 site-directed mutagenesis.

For transient transfection, plasmid (19 µg) was prepared with 0.25 M $CaCl_2$, incubated for 10 minutes with 1X HBS (2X HBS: 140 mM NaCl, 1.5 mM Na_2HPO_4, 50 mM HEPES, pH to 7.05), then applied drop-wise to 293T cells. Cells were grown for two days before harvesting (see below).

Cell lysates for western blotting or immunoprecipitation (IP) were prepared by rinsing cells twice in ice-cold 1X PBS, and harvesting in Kischkel Buffer (50 mM Tris pH 8.0, 150 mM NaCl, 5 mM EDTA, 1% Triton X-100) + protease inhibitor cocktail (PIC, Roche, Basel, Switzerland, 05056489001). Cells were sonicated at 25% power for 10 s (Cole-Parmer, Vernon Hills, Illinois, GE 130PB-1), debris removed by centrifugation, and protein concentration determined using Protein Assay Dye (BioRad 500-0006) against a bovine serum albumin (BSA) standard. For western blotting, lysate was diluted in 5X Laemmli Buffer (375 mM Tris pH 6.8, 40% glycerol, 10% SDS, bromophenol blue, 2-Mercaptoethanol). For IP, the concentration of samples was balanced using Kischkel Buffer. Antibody or anti-FLAG M2 Affinity Gel (Millipore A2220) was added, and samples rotated overnight at 4°C. For unconjugated antibodies, 20 µl bed volume of Protein A Agarose (Roche 11134515001) was added the following day to each

sample, and rotated for 2-4 hours at 4°C. Samples were then washed 4X with Kischkel buffer (2X 4°C and 2X at room-temperature), and incubated in Laemmli buffer for 5 minutes at 95°C.

Protein from lysates and IPs were separated out by SDS-polyacrylamide gel electrophoresis (PAGE) in running buffer (25 mM Tris, 192 mM glycine, 0.2% SDS). Wet transfer to PVDF (PerkinElmer, Waltham, Massachusetts, NEF1002) was carried out in Towbin Transfer Buffer (25 mM Tris, 192 mM glycine, and 10% methanol). Membrane was blocked in 5% milk in TBS-T (20 mM Tris pH 7.6, 140 mM NaCl, 0.1% Tween-20), hybridized overnight in primary antibody (or 1 hour for HRP-conjugated), and for 1 hour in HRP-conjugated secondary antibody (if required). ECL substrates, SuperSignal West Pico (Pierce, Waltham, Massachusetts, 34080), Pico+ (Pierce 34580), and Femto (Pierce 34095), were used in various combinations for detection of bands by exposure to film.

***In vitro* binding assays**

pSUMO-MYC WT-FLAG containing 6XHis- and FLAG-tagged full-length MYC [178] was used for Q5 site-directed mutagenesis to substitute in the 4A (4A_F and 4A_R) and VP16 HBM (VP16 HBM_F and VP16 HBM_R) mutations. Rosetta cells (Millipore 70954) were transformed with these plasmids, grown overnight, and induced the following day for three hours with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) at 30°C. Resulting cell pellets were resuspended in Buffer A (100 mM NaH_2PO_4, 10 mM Tris, 6 M GuHCl, 10 mM imidazole) + PIC (Roche 05056489001). Cells were sonicated 3X at 25% power for 10s (Cole-Parmer GE 130PB-1), and debris removed by centrifugation. 150 µl bed volume Ni-NTA agarose (QIAGEN, Venlo, Netherlands, 30210) was added to the samples and rotated for two hours at 4°C. Beads were sequentially washed 1X with Buffer A, 1X with Buffer A/TI (1:3 Buffer A:Buffer TI), 1X with Buffer TI (25 mM Tris-HCl pH 6.8, 20 mM imidazole), and 1X with SUMO wash buffer (3 mM imidazole, 10% glycerol, 1X PBS, 2 mM DTT) + PIC (Roche 05056489001). Recombinant MYC was eluted from the beads using SUMO elution buffer (250 mM imidazole, 10% glycerol, 1X PBS, 2 mM DTT) + PIC (Roche 05056489001).

HCF-1_{VIC} (residues 1-380) from FGH-$HCF1_{VIC}$ [179] was cloned into pT7-IRES His-N (Takara 3290) using BamHI-HF (NEB R3136) and SalI-HF (NEB R3138), and Q5 site-directed mutagenesis was used to add an N-terminal T7 tag using the primers T7-HCF1_F and T7-HCF1_R. pT7-IRES His-T7-HCF-1_{VIC} was *in vitro* transcribed/translated using the TnT Quick Coupled Transcription/Translation System (Promega, Madison, Wisconsin, L1171). Two milligrams recombinant MYC and 12 µl T7-HCF-1_{VIC} were rotated overnight at 4°C in Kischkel buffer + PIC (Roche 05056489001). Anti-FLAG M2 Affinity Gel (Sigma-Aldrich), blocked with 10 µg BSA, was added to each sample and rotated for 2 hours at 4°C. Beads were washed 4X in Kischkel buffer + 2 µg/ml Aprotinin (VWR, Radnor, Pennsylvania, 97062-752) + 1 µg/ml Pepstatin (VWR 97063-246) + 1 µg/ml Leupeptin (VWR 89146-578), and incubated in 1X Laemmli buffer for 5 minutes at 95°C.

Generation of switchable *MYC* allele Ramos cell lines

Q5 site-directed mutagenesis of MYC-WT targeting vector from Thomas *et al.* [21] was used to create MYC-4A (4A_F and 4A_R) and MYC-VP16 HBM (VP16 HBM_F and VP16 HBM_R). The pGuide plasmid described by Thomas *et al.* [21] was used as a backbone to introduce the sgRNA sequence GCTACGGAACTCTTGTGCGTA (pGuide-MYC1) by Q5 site directed mutagenesis with the primers GUIDE MYC-1A and GUIDE MYC-1B.

For the generation of switchable cells, 10 million Ramos cells stably expressing CRE-ER^{T2} [21] were electroporated (BioRad Gene Pulser II, 220 V and 950 µF) with 10 µg of relevant targeting vector (MYC-4A or MYC-VP16 HBM), 15 µg pGuide-MYC1, and 15µ g pX330-U6-Chimeric_BB-CBh-hSpCas9 (gift from Feng Zhang, AddGene plasmid #42230) [298]. WT and ∆264 cell lines were the same as those used in Thomas *et al.* [21]. Cells were treated with 150 ng/ml puromycin (Sigma-Aldrich P7255) and 100 µg/ml hygromycin (Corning 30240CR), selecting for the switchable *MYC* cassette and CRE-ER^{T2} recombinase, respectively. Following selection, single cells, stained using propidium iodide (PI, Sigma-Aldrich P4864) for viability, were sorted by the Vanderbilt Flow Cytometry Shared Resource using a BD FACSAria III flow cytometer into a 96-well plate to generate clonal cell lines under

puromycin and hygromycin selection. Individual clones were expanded, and initially validated by switching for 24 hours using 20 nM (Z)-4-Hydroxytamoxifen (4-OHT, Tocris, Minneapolis, Minnesota, 3412), and flow cytometry (see below) for GFP expression. Further validation was performed by western blotting after 24 hours 20 nM ±4-OHT (see below), and by Southern blotting (see below). The 4A cell line (4A-1) used for the majority of experiments is haploinsufficient for part of chromosome 11, approximately between co-ordinates 118,685,194 and 134,982,408. This region of the genome was excluded from genomic analyses. For all experiments, switching was performed by treatment with 20 nM 4-OHT for 2 or 24 hours (see relevant method or figure legend).

Generation of dTAG Ramos cell lines

To create pCRIS-mCherry-FLAG-dTAG-HCFC1, pCRIS-PITChv2-Puro-dTAG (BRD4) (gift from James Bradner, AddGene plasmid #91793) [299] was first modified to remove the BRD4 homology arms and replace the 2XHA tags with a FLAG tag. This was done by Gibson assembly of the vector (Q5 amplification using pCRIS-HCFC1N_F and pCRIS_R), puromycin cassette (Q5 amplification using Puro-1 and Puro-FLAG_R), and $FKBP12^{FV}$ (Q5 amplification using FLAG-FKBP_F and pCRIS-FKBP_R). The resulting vector was again modified using Gibson assembly by combining the vector (Q5 amplification using pCRIS-HCFC1N_F and pCRIS-HCFC1N_R), an upstream 271bp HCFC1 5' homology arm (hg19 chrX:153236265-153236535, OneTaq amplification using HCFC1N_F and HCFC1N-5'Hom_R), mCherry (Q5 amplification using mCherry_F and mCherry_R), $FKBP12^{FV}$ (Q5 amplification using HCFC1-mCherry-FKBP_F and HCFC1-mCherry-FKBP_R), and a downstream 800bp HCFC1 3' homology arm (hg19 chrX:153235465-153236264, OneTaq amplification using HCFC1N-3'Hom_F and HCFC1N-3'Hom_R). The pGuide plasmid described by Thomas *et al.* [300] was used as a backbone to introduce the guide RNA sequence CAGAAGCACCGCTGGCAAGT (pGuide-HCFC1-N) by Q5 site-directed mutagenesis with the primers HCFC1N-sgRNA_F and HCFC1N-sgRNA_R.

Fifteen micrograms pGuide-HCFC1-N and 10 µg pCRIS-mCherry-FLAG-dTAG-HCFC1 were electroporated (BioRad, Hercules, California, Gene Pulser II, 220 V and 950 µF) into Ramos cells,

alongside 15 µg pX330-U6-Chimeric_BB-CBh-hSpCas9 (gift from Feng Zhang, AddGene plasmid #42230) [298] into 10 million Ramos cells. Because HCFC1 is on the X chromosome and Ramos cells are XY [301], only a single copy of HCFC1 is present for targeting using CRISPR/Cas9. Following electroporation, cells were expanded and a population of mCherry-positive cells, stained using Zombie NIR viability dye (BioLegend, San Diego, California, 423105), was sorted by the Vanderbilt Flow Cytometry Shared Resource using a FACSAria III flow cytometer (Becton Dickinson (BD), Franklin Lakes, New Jersey). This population of cells was expanded further before validation by western blotting. All experiments were conducted using this population, and were treated with either DMSO (Sigma-Aldrich D2650) or 500 nM dTAG-47, which was synthesized by the Vanderbilt Institute of Chemical Biology Synthesis Core.

Flow cytometry and cell cycle analysis

Cells were filtered into 35 µm nylon mesh Falcon round bottom test tubes for flow cytometry, which was performed in the Vanderbilt Flow Cytometry Shared Resource. Single cells were gated based on side and forward scatter using the stated instruments.

To determine the proportion of switching of the switchable *MYC* allele Ramos cell lines, cells were fixed in 1% formaldehyde (FA, Thermo Fisher 28908) in PBS for 10 minutes at room temperature. The number of GFP-positive cells was determined using a BD LSR II flow cytometer.

For cell cycle analysis of the switchable *MYC* allele Ramos cell lines, cells were treated with 4-OHT for 2 hours, at which point the media was replaced. Cells were maintained for 7 days, at which point $1x10^6$ cells were collected and fixed in 1% FA (Thermo Fisher 28908) in PBS for 10 minutes at room temperature, then washed 2X with PBS. Permeabilization and staining was done using cell cycle staining buffer (PBS, 10 µg/ml PI (Sigma-Aldrich P4864), 100 µg/ml RNAse A, 2 mM $MgCl_2$, 0.1% Triton X-100) for 25 minutes at room temperature, then stored overnight at 4°C. PI staining of at least 10,000 single cells for each the GFP-negative (GFP-) and GFP-positive (GFP+) populations was measured using a BD LSR Fortessa, and cell cycle distribution determined using BD FACSDIVA software

For cell cycle analysis of the $FKBP^{FV}$-HCF-1_N Ramos cells, cells were treated with DMSO or 500 nM dTAG for 24 hours. One million cells were collected and fixed in ethanol overnight. After fixation, cells were stained overnight at 4°C in cell cycle staining buffer (PBS, 10 µg/ml PI (Sigma-Aldrich P4864), 100 µg/ml RNAse A, 2 mM $MgCl_2$). PI staining of at least 10,000 single cells was measured using a BD LSR Fortessa, and cell cycle distribution determined using BD FACSDIVA software.

See section '*In vivo* studies: Tumor formation and maintenance assays' for details regarding flow cytometry experiments conducted on cells extracted from mice tumors.

Southern blotting

Genomic DNA (gDNA) was prepared from parental and unswitched switchable *MYC* cells (WT, 4A, and VP16 HBM) [302]. Briefly, cells were rinsed in ice-cold 1X PBS, and resuspended in DNA extraction buffer (10 mM Tris pH 8.1, 400 mM NaCl, 10 mM EDTA, 1% SDS, 50 µg/ml proteinase K (PK, Macherey-Nagel, Düren, Germany, 740506)). Lysis was performed overnight in a rotisserie at 56°C, before gDNA was extracted using ethanol precipitation. Southern blot was performed similarly to that described by Southern *et al.* [303]. gDNA (10 µg) was digested using XbaI (NEB R0145) and run out on a 1% agarose gel. DNA was transferred overnight to Hybond-N+ nylon membrane (GE Healthcare, Chicago, Illinois, RPN303B) by capillary action in transfer buffer (0.5 M NaOH, 0.6 M NaCl). The following day the membrane was immersed in neutralization buffer (1 M NaCl, 0.5 M Tris pH 7.4), UV cross-linked, and pre-hybridized overnight at 42°C in pre-hybridization buffer (50% formamide, 5X SSCPE (20X SSCPE: 2.4 M NaCl, 0.3 M Na citrate, 0.2 M KH_2PO_4, 0.02 M EDTA), 5X Denhardt's solution (Invitrogen, Waltham, Massachusetts, 750018), 0.5 mg/ml salmon sperm DNA (Agilent, Santa Clara, California, 201190), 1% SDS). Templates for probe generation were prepared by Q5 amplification from MYC-WT targeting vector using primers GFP_F and GFP_R (GFP template) and from parental Ramos cell gDNA using primers 5'_F and 5'_R, followed by gel purification. Probes were prepared by random priming of corresponding PCR products (5' and GFP templates) in the presence of [αP32]CTP (PerkinElmer BLU513H100UC). Unincorporated nucleotides were removed using a

Sephadex G-50 column (GE Healthcare 28-9034-08). The membrane was incubated overnight at 42°C with probe in hybridization buffer (50% formamide, 5X SSCPE, 5X Denhardt's solution (Invitrogen 750018), 0.1 mg/ml salmon sperm DNA (Agilent 201190), 1% SDS, 10% Dextran solution). Membrane was washed three times in 2X SSC/0.1% SDS (20X SSC: 3 M NaCl, 0.3 M Na Citrate, pH 7.0) and twice in 0.2X SSC/0.1% SDS, then exposed to a phosphor screen and developed using a phosphorimager (GE Healthcare Typhoon).

Chromatin fractionation

Chromatin fractionation was performed, with slight modification, as described by Mendez and Stillman [304]. Switchable *MYC* Ramos cells that had been treated for 24 hours with 4-OHT were washed in ice-cold PBS, resuspended in 200 µl Buffer A (10mM HEPES pH 7.9, 10mM KCl, 1.5mM $MgCl_2$, 0.34 M sucrose, 10% glycerol, 1mM DTT) + PIC + PMSF + 0.1% Triton X-100, and incubated on ice for 10 minutes. The resulting lysate was centrifuged 1,300 x g for 5 minutes at 4°C, with the pellet (P1) containing the nuclei. The supernatant (S1) was centrifuged at 20,000 x g for 10 minutes at 4°C, giving pellet P2 and supernatant S2. P2 was discarded and S2, corresponding to the soluble portion of the total cell extract, was diluted out in Laemmli buffer. P1 was gently washed in Buffer A, resuspended by pipetting up and down in Buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT) + PIC + PMSF, and incubated on ice for 30 minutes. The lysed nuclei were centrifuged at 1,700 x g for 5 minutes at 4°C to give soluble nuclear proteins (S3) and chromatin-bound proteins (P3). S3 was diluted out in Laemmli buffer. P3 was gently washed in Buffer B, resuspended in Laemmli buffer, and sonicated for 15 seconds at 25% power. All samples (S2, S3, P3) were incubated at 95°C for 3 minutes. Proteins were separated out by SDS-PAGE and transferred to PVDF, as described above, and probed for HCF-1_C, HA, tubulin, and H3.

Immunofluorescence

This protocol was performed by G. Caleb Howard of Cell and Developmental Biology at Vanderbilt University. Following a 24 hour treatment with 4-OHT, 10^5 switchable *MYC* Ramos cells were attached to slides by CytoSpin (800 RPM, 3 minutes), fixed for 10 minutes with 3% methanol-free FA (Thermo

Fisher Scientific 28908) diluted in PBS, and washed three times with PBS. Cells were permeabilized for 10 minutes with Permeabilization Solution (0.1% Triton X-100 in PBS), blocked for 1 hour with Blocking Solution (2.5% BSA in Permeabilization Solution), and incubated with Blocking Solution containing anti-HA antibody (1:500, Cell Signaling 3724) for 1 hour. Cells were washed three times with PBS and incubated with Blocking Solution containing Goat anti-Rabbit IgG Alexa Fluor 594 antibody (1:350, Thermo Fisher Scientific A11012) for 1 hour. Cells were washed three times with PBS, and coverslips mounted with ProLong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific P36941). Slides were imaged by wide-field fluorescent microscopy on a Nikon Eclipse Ti equipped with a Nikon Plan Apo λ 100x/1.45 Oil objective, Nikon DS-Qi2 camera, and Excelitas X-Cite 120LED illuminator using identical settings for each sample and representative images shown.

Chromatin immunoprecipitation and library preparation

Chromatin immunoprecipitation (ChIP) was performed, with slight modification, as described by Thomas *et al.* [178]**.** Cells were first treated with 20nM 4-OHT for 24 hours, then crosslinked in 1% methanol-free FA (Thermo Fisher 28908) for 10 minutes and quenched using 0.125 mM glycine. The cells were then rinsed twice in ice-cold 1X PBS, and lysed in formaldehyde lysis buffer (FALB: 50 mM HEPES pH 7.5, 140 mM NaCl, 1mM EDTA, and 1% Triton X-100) + 1% SDS + PIC (Roche 05056489001). Sonication was performed in a BioRuptor (Diagenode, Denville, New Jersey) for 25 minutes, 30s on/30s off, and debris removed by centrifugation. To enable ChIP efficiency to be determined by qPCR, a 1:50 (2%) sample of chromatin was removed (input) prior to antibody addition.

For anti-HA, anti-rabbit IgG or anti-$HCF1_N$ ChIP, antibody was added to chromatin prepared from $12x10^6$ Ramos cells, and samples rotated overnight at 4°C. Protein A Agarose (Roche 11134515001), blocked with 10µg BSA, was added to each sample and rotated at 4°C for 2-4 hours. Washes were performed with Low Salt Wash Buffer (20 mM Tris pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100), High Salt Wash Buffer (20 mM Tris pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100), Lithium Chloride Wash Buffer (10 mM Tris pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% Triton X-100), and

twice with TE (10 mM Tris pH 8.0, 1 mM EDTA). Input and ChIP sample crosslinking were reversed at 65°C overnight in 50 µl TE + 200 mM NaCl + 0.1% SDS + 20-40 µg PK (Macherey-Nagel 740506).

For sequencing, three independent ChIPs (anti-IgG from CRE-ER^{T2} parental cells, anti-HCF1$_N$ from MYC-WT, 4A, or VP16 HBM cells, or anti-HA from CRE-ER^{T2} parental cells, MYC-WT, 4A, or VP16 HBM cells) performed using the same antibody with the same chromatin were combined and DNA was extracted with phenol:chloroform:isoamyl alcohol, and ethanol precipitated with glycogen (Roche 10901393001) or GlycoBlue (Invitrogen AM9516). Libraries were prepared using NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB E7645S) and NEBNext Multiplex Oligos for Illumina (NEB Set 1 E7335, Set 2 E7500S, NEB Unique Dual Index E6440S). Additional AMPure clean-ups at the start and the end of library preparation were included. Sequencing was carried out by Vanderbilt Technologies for Advanced Genomics using 75 bp paired-end sequencing on Illumina NextSeq 500 for anti-IgG and anti-HCF1$_N$ ChIPs or 150 bp paired-end sequencing on Illumina NovaSeq for anti-HA. The total number of sequencing reads for each replicate is shown in Table 2-2.

For qPCR, samples (either input or ChIP) were brought up to a final volume of 200 µl using TE. Each reaction was performed in a final volume of 15 µl, containing 2X SYBR FAST qPCR Master Mix (Kapa, Wilmington, Massachusetts, KK4602), 300 nM of each primer, and 2 µl of diluted sample. Three technical replicates were performed for each sample, and the mean Ct of these was used for calculating percent input. The mean Ct value for input was adjusted using the equation Ct(input)-$\log_2(50)$. Percent input was then calculated using the equation 100*2^(adjustedCt-Ct(ChIP)). Three biological replicates of ChIP-qPCR were performed using primers to amplify across the genes EXOSC5, UTP20, POLR1A, EIF2S1, EIF4G3, and HBB (β-Globin).

Table 2-2: Next-generation sequencing read counts

Experiment	Sample	Total Reads	Experiment	Sample	Total Reads
Parental dTAG RNA-Seq	DMSO rep1	45286657	Tumor RNA-Seq	4A2 rep3	52286744
Parental dTAG RNA-Seq	DMSO rep2	42180049	Tumor RNA-Seq	4A2 rep4	52314725
Parental dTAG RNA-Seq	DMSO rep3	46846258	Tumor RNA-Seq	delta264 rep1	52248487
Parental dTAG RNA-Seq	dTAG rep1	51532685	Tumor RNA-Seq	delta264 rep2	52798759
Parental dTAG RNA-Seq	dTAG rep2	51751441	Tumor RNA-Seq	delta264 rep3	52770341
Parental dTAG RNA-Seq	dTAG rep3	48532954	Tumor RNA-Seq	delta264 rep4	52681554
FKBPFV-HCF-1N RNA-Seq	DMSO rep1	137074323	Switchable MYC cells HCF-1N ChIP-Seq	Parental IgG rep1	61857969
FKBPFV-HCF-1N RNA-Seq	DMSO rep2	97856532	Switchable MYC cells HCF-1N ChIP-Seq	Parental IgG rep2	62287619
FKBPFV-HCF-1N RNA-Seq	DMSO rep3	98517439	Switchable MYC cells HCF-1N ChIP-Seq	Parental IgG rep3	71279380
FKBPFV-HCF-1N RNA-Seq	dTAG rep1	97407294	Switchable MYC cells HCF-1N ChIP-Seq	WT rep1	69241747
FKBPFV-HCF-1N RNA-Seq	dTAG rep2	66587719	Switchable MYC cells HCF-1N ChIP-Seq	WT rep2	65745408
FKBPFV-HCF-1N RNA-Seq	dTAG rep3	83248853	Switchable MYC cells HCF-1N ChIP-Seq	WT rep3	69388788
Switchable MYC cells RNA-Seq	WT rep1	45540235	Switchable MYC cells HCF-1N ChIP-Seq	4A rep1	69785267
Switchable MYC cells RNA-Seq	WT rep2	43916750	Switchable MYC cells HCF-1N ChIP-Seq	4A rep2	64115684
Switchable MYC cells RNA-Seq	WT rep3	45294510	Switchable MYC cells HCF-1N ChIP-Seq	4A rep3	74811512
Switchable MYC cells RNA-Seq	4A rep1	51877019	Switchable MYC cells HCF-1N ChIP-Seq	VP16 HBM rep1	62199784
Switchable MYC cells RNA-Seq	4A rep2	47122065	Switchable MYC cells HCF-1N ChIP-Seq	VP16 HBM rep2	63942557
Switchable MYC cells RNA-Seq	4A rep3	41380579	Switchable MYC cells HCF-1N ChIP-Seq	VP16 HBM rep3	64738443
Switchable MYC cells RNA-Seq	VP16 HBM rep1	40905512	Switchable MYC cells HA ChIP-Seq	Parental rep1	65083068
Switchable MYC cells RNA-Seq	VP16 HBM rep2	44677502	Switchable MYC cells HA ChIP-Seq	Parental rep2	65894596
Switchable MYC cells RNA-Seq	VP16 HBM rep3	40651373	Switchable MYC cells HA ChIP-Seq	Parental rep3	67616570
Tumor RNA-Seq	WT rep1	52435549	Switchable MYC cells HA ChIP-Seq	WT rep1	71303212
Tumor RNA-Seq	WT rep2	52282103	Switchable MYC cells HA ChIP-Seq	WT rep2	74829473
Tumor RNA-Seq	WT rep3	52600096	Switchable MYC cells HA ChIP-Seq	WT rep3	62868116
Tumor RNA-Seq	WT rep4	52173879	Switchable MYC cells HA ChIP-Seq	4A rep1	64784838
Tumor RNA-Seq	4A1 rep1	52647667	Switchable MYC cells HA ChIP-Seq	4A rep2	71217193
Tumor RNA-Seq	4A1 rep2	52816902	Switchable MYC cells HA ChIP-Seq	4A rep3	71209691
Tumor RNA-Seq	4A1 rep3	52295999	Switchable MYC cells HA ChIP-Seq	VP16 HBM rep1	75586322
Tumor RNA-Seq	4A1 rep4	52908454	Switchable MYC cells HA ChIP-Seq	VP16 HBM rep2	82572915
Tumor RNA-Seq	4A2 rep1	52150641	Switchable MYC cells HA ChIP-Seq	VP16 HBM rep3	55442880
Tumor RNA-Seq	4A2 rep2	52678027			

Electrophoretic mobility shift assays

MYC:MAX dimers were purified and prepared as described by Farina *et al.* [305]. pRSET-6XHis-MYC WT, a gift from Ernest Martinez, was used as a template for Q5 site directed mutagenesis to substitute in the 4A (4A_F and 4A_R) and VP16 HBM (VP16 HBM_F and VP16 HBM_R) mutations. The resulting plasmids, or pET-His-MAX, also from Ernest Martinez, were transformed into Rosetta cells (Millipore 70954), grown overnight, and induced the following day for three hours with 1 mM IPTG at 30°C. Resulting bacterial cell pellets were washed 1X with ice-cold wash buffer (10 mM Tris pH 7.9, 100 mM NaCl, 1 mM EDTA), then resuspended in lysis buffer (20 mM HEPES pH 7.9, 500 mM NaCl, 10% glycerol, 0.1% NP-40, 10 mM BME, 1 mM PMSF) and sonicated. Centrifugation was used to separate out the insoluble (pellet) and soluble (supernatant) fractions. For MAX, the recombinant protein is present in the supernatant, and was then purified using Ni-NTA agarose (see below). For MYC samples, the supernatant was discarded, and the insoluble pellet was resuspended in E-buffer (50 mM HEPES pH 7.9, 5% glycerol, 1% NP-40, 10% Na-DOC, 0.5 mM BME), lysed using a Dounce homogenizer and B-pestle, and centrifuged. The pelleted inclusion bodies were lysed overnight in S-buffer (10 mM HEPES pH 7.9, 6 M GuHCl, 5 mM BME) by shaking at 25°C, and debris spun out by centrifugation.

Both the MAX supernatant and MYC supernatant from lysed inclusion bodies were adjusted to 5 mM imidazole. MYC and MAX were then bound to 75 µl bed volume Ni-NTA agarose (QIAGEN 30210). Successive washes were performed for 5 minutes each at 4°C: 3X with S-buffer + 5 mM imidazole, 3X with BC500 (20 mM Tris pH 7.9, 20% glycerol, 500 mM KCl, 0.05% NP-40, 10 mM BME, 0.2 mM PMSF) + 7 M urea + 5 mM imidazole, 1X with BC100 (20 mM Tris pH 7.9, 20% glycerol, 100 mM KCl, 0.05% NP-40, 10 mM BME, 0.2 mM PMSF) + 7 M urea + 15 mM imidazole, 1X with BC100 + 7 M urea + 30 mM imidazole. Elution was performed using BC100 + 7 M urea + 300 mM imidazole. Concentration of recombinant MYC and MAX was determined by running samples out on a 12% acrylamide gel alongside a BSA standard, and staining using Coomassie stain (50% methanol, 10% acetic acid, 0.1% w/v Coomassie Brilliant Blue).

For renaturation, 1.5 µg recombinant MAX was combined with 15 µg recombinant MYC (1:3 molar ratio), and brought up to a final volume of 150 µl using BC100 + 7 M urea. Dialysis was performed using a “Slide-A-Lyzer” (Thermo Fisher 66383), with each dialysis step done for 2 hours stirring in the following solutions: BC500 + 0.1% NP-40 + 4 M urea at room temperature, BC500 + 0.1% NP-40 + 2 M urea at room temperature, BC500 + 0.1% NP-40 + 1 M urea at room temperature, BC500 + 0.1% NP-40 + 0.5 M urea at room temperature, BC500 at 4°C, and twice BC100 at 4°C. The product was centrifuged to remove debris, and BSA was added to a final concentration of 500 ng/µl.

Double-stranded labeled E-box probe (biotin group at the 3’ end) and unlabeled competitors were prepared with dsDNA buffer (30 mM Tris pH 7.9, 200 mM KCl), and incubated at 95°C for 5 minutes. The E-box sequence used was 5′-GCTCAGGGACCACGTGGTCGGGGATC-3′ and the mutant E-box sequence used was 5′-GCTCAGGGACCAGCTGGTCGGGGATC-3′ (IDT). Double-stranded probe, and the specific and non-specific competitors were prepared by combining 25 µM of each strand in dsDNA buffer (30 mM Tris pH 7.9, 200 mM KCl). For the probe, the forward strand carried a 3’ biotin group. 20 fmol of labeled probe was bound to 0.55 pmol MYC:MAX or 0.06 pmol MAX:MAX dimers in the presence of 20 ng poly(dI-dC) (Thermo Fisher 20148E) in binding reaction buffer (15 mM Tris pH 7.9, 15% glycerol, 100 mM KCl, 0.15 mM EDTA, 0.075% NP-40, 7.5 mM BME, 375 ng/µl BSA) for 30 minutes at room temperature. For reactions involving unlabeled specific or non-specific competitor, these were included in the binding reaction at a 100-fold excess over the biotinylated probe. EMSA gel loading solution (Thermo Fisher 20148K) was added to each sample and these were loaded onto a pre-run 6% polyacrylamide gel in 0.5X TBE (45 mM Tris, 45 mM boric acid, 1 mM EDTA). The gel was transferred to Hybond-N+ nylon membrane (GE Healthcare RPN303B) in 0.5X TBE for 30 minutes at 100 V. The remainder of the protocol was performed using LightShift Chemiluminescent EMSA Kit (Thermo Fisher 20148) according to manufacturer’s instructions.

RNA preparation, RT-qPCR, and RNA-Seq

Cell pellets were resuspended in 1 ml TRIzol (Invitrogen 15596026), and RNA was prepared according to the manufacturer's instructions. For switchable *MYC* allele Ramos cells, cells were treated with 20 nM 4-OHT for 24 hours, and harvested. Prepared RNA was submitted to GENEWIZ (South Plainfield, NJ) for DNAse treatment, rRNA depletion, library preparation, and 150 bp paired-end sequencing on Illumina HiSeq. For untagged or $FKBP^{FV}$-$HCF\text{-}1_N$ Ramos cells, cells were treated with DMSO or 500 nM dTAG-47 for 3 hours, and prepared RNA was DNAse treated prior to submission to Vanderbilt Technologies for Advanced Genomics for rRNA depletion, library preparation, and 150 bp paired-end sequencing on Illumina NovaSeq 6000. The total number of sequencing reads for each replicate is shown in Table 2-2.

For validation of RNA-Seq by reverse transcriptase qPCR (RT-qPCR), RNA was prepared as above and 1 µg was converted to cDNA using M-MLV reverse transcriptase (Promega M1701) in the presence of random hexamers (Invitrogen N8080127), RNase inhibitor (Thermo Fisher Scientific N8080119), and dNTPs (NEB N0446S). The resulting cDNA was brought up to a final volume of 160 µl using water. qPCR was performed in a final volume of 15 µl, containing 2X SYBR FAST qPCR Master Mix (Kapa), 300 nM of each primer, and 2 µl of diluted sample. Three technical replicates were performed for each sample, and the mean Ct of these was used for calculating fold-change. The mean Ct value for the gene of interest (GOI) was normalized to GAPDH (ΔCt) using the equation Ct(GOI)-Ct(GAPDH). For switchable *MYC* allele Ramos cells, ΔΔCt was calculated between treated (+4-OHT) and untreated (-4-OHT) cells. For $FKBP^{FV}$-$HCF\text{-}1_N$ Ramos cells, ΔΔCt was calculated between dTAG-47-treated and DMSO-treated cells. Fold-change was then calculated using the equation 2^(-ΔΔCt). Three biological replicates of RT-qPCR were performed.

Table 2-3: Quantitative PCR primer sequences

Primer name	Sequence	Primer name	Sequence
MYCP-4	AaGAaAAcGTgAAacGca	EXOSC5-2	AGGATCACTTCGAGTGTGGC
MYCP-5	gAGaGCgAAgAAGgatCtc	UTP20-1	TCCGTGGATACCAGGTCCAT
SNHG15_F	CGCCACTGAACCCAATCC	UTP20-2	TCCGACCTGCAGCTTATTGG
SNHG15_R	TCTAGTCATCCACCGCCATC	MARS2-1	TCTCTCCTGGAGGACTTCGG
HBB-1	GGCTGTCATCACTTAGACCTC	MARS2-2	CATCACAAGCATCATCGCCG
HBB-2	GGTTGCTAGTGAACACAGTTG	EIF2S1-1	AAGCATGCAGTCTCAGACCC
EIF2S1_F	TTCGTCCTTGTTCTCGGAGG	EIF2S1-2	GTGGGGTCAAGCGCCTATTA
EIF2S1_R	CGCAAGCGCTGAACGAATAG	POLRMT-1	GCTACAGGAAGGGCCTGAC
EXOSC5_F	TTTGGCGTCAGTATGCGTCT	POLRMT-2	CACACCTGGTTCATGACGGA
EXOSC5_R	CTGGTTACGCAGCCTGTTTG	SUCLA2-1	TGGTACGGTTACAAGGTACACG
UTP20_F	ACACCTCACCCAAACGCC	SUCLA2-2	TGCTTGCTTCGCTAAGGTCA
UTP20_R	GGGTCACGTGGACGGAAAA	WDR3-1	TATGATCAGCTAGGAGGCAG
POLR1A_F	GTTCCACTCACCACCTGACT	WDR3-2	CTTTCCCACTAGTAACTAGC
POLR1A_R	AAATGTCGTTCCTGCGACCG	RRP9-1	GACATTCGCGTTTTACGGGG
EIF4G3-1	CCTTTCACGGCAATATCCTC	RRP9-2	CACTCTCCACGCTCCACTTA
EIF4G3-2	TGAGTGAAGAAAATCCACCG	GAPDH-1	AAGGTGAAGGTCGGAGTCAAC
EXOSC5-1	CTTCCTTCCTGCAAGGTGACA	GAPDH-2	GTTGAGGTCAATGAAGGGGTC

Next generation sequencing analyses

The following analysis was conducted by Jing Wang and Qi Liu at the Department of Biostatistics and Center for Quantitative Sciences, Vanderbilt University Medical Center. After adapter trimming by Cutadapt [306], RNA-Seq reads were aligned to the hg19 genome using STAR [307] and quantified by featureCounts [308]. Differential analysis were performed by DESeq2 [309], which estimated the log2 fold changes, Wald test P-values, and adjusted P-value (false discovery rate, FDR) by the Benjamini-Hochberg procedure. The significantly changed genes were chosen with the criteria FDR < 0.05. ChIP-Seq reads were aligned to the hg19 genome using Bowtie2 [310] after adapter trimming. Peaks were called by MACS2 [311] with a q-value of 0.01. Calculation of ChIP read counts and identification of differential binding peaks were done using DiffBind [312]. Peaks were annotated using Homer command annotatePeaks, and enriched motifs were identified by Homer command findMotifsGenome (http://homer.ucsd.edu/homer/). All genomics data were deposited at GEO with the accession number GSE152385.

Cell growth and glutamine deprivation assays

To measure the impact of HCF-1_N degradation on cell growth, $FKBP^{FV}$-HCF-1_N Ramos cells were plated at a density of 20,000 cells/ml with either DMSO or 500 nM dTAG-47. Cells were then counted every 24 hours for the following four days, without replacement of the compound or changing of the media.

To measure the impact of altering the MYC−HCF-1 interaction on cell growth, switchable *MYC* allele Ramos cell lines were treated with 20 nM 4-OHT for 2 hours to create a 50/50 mix of GFP-negative (WT) to GFP-positive (mutant) cells. Cells were sampled 24 hours later, and every three days following. Sampled cells were fixed in 1% FA in PBS for 10 minutes at room temperature, and the proportion of GFP-positive cells was determined using a BD LSR II flow cytometer at the Vanderbilt Flow Cytometry Shared Resource. To account for variation in the proportion of GFP-positive cells between replicates, each replicate was normalized to the proportion of GFP-positive cells at 24 hours post-treatment with 4-OHT.

To measure the impact of altering the MYC−HCF-1 interaction on glutamine dependence, switchable *MYC* allele Ramos cell lines were treated with 20 nM 4-OHT for 2 hours, and allowed to recover for three days. Cells were then split into RPMI 1640 without L-Glutamine (Corning 15-040-CV), supplemented with 10% dialyzed FBS (Gemini Bio, West Sacramento, California, 100-108), and 1% P/S (Gibco 15140122), and grown for 16 hours with or without supplemental 2mM glutamine (Gibco 25030081). Glutamine was added back to the cells that were deprived, grown for three days, and fixed in 1% FA in PBS for 10 minutes at room temperature. The proportion of GFP-positive cells was determined using a BD LSR II flow cytometer, and normalized to the proportion of GFP-positive cells prior to being grown with or without supplemental glutamine. For the 4A and VP16 HBM cells, the proportion of GFP-positive cells was normalized to that in WT for glutamine supplementation (Gln+) or deprivation (Gln-).

Metabolomics

Sample preparation and analysis for metabolomics was performed by the Vanderbilt Center for Innovative Technology.

Sample Preparation

Global, untargeted metabolomics was performed on switchable *MYC* allele Ramos cell lines treated with 20 nM 4-OHT for 24 hours. Individual cell pellet samples were lysed using 200 µl ice cold lysis buffer (1:1:2, Acetonitrile : Methanol : Ammonium Bicarbonate 0.1 M, pH 8.0, LC-MS grade) and sonicated using a probe tip sonicator, 10 pulses at 30% power, cooling down on ice between samples. A bicinchoninic acid protein assay was used to determine the protein concentration for individual samples, and adjusted to 200 µg total protein in 200 µl of lysis buffer. Isotopically labeled standard molecules, Phenylalanine-D8 and Biotin-D2, were added to each sample to assess sample extraction quality. Samples were subjected to protein precipitation by addition of 800 µl of ice cold methanol (4X by volume), and incubated at -80°C overnight. Samples were centrifuged at 10,000 rpm for 10 minutes to eliminate precipitated proteins and supernatant(s) were transferred to a clean microcentrifuge tube and dried down *in vacuo*. Samples were stored at -80°C prior to LC-MS analysis.

Global untargeted LC-MS/MS analysis

For mass spectrometry analysis, individual samples were reconstituted in 50 µl of appropriate reconstitution buffer (HILIC: acetonitrile/ H_2O, 90:10, v/v, RPLC: acetonitrile/H_2O with 0.1% formic acid, 3:97, v/v). Samples were vortexed well to solubilize the metabolites and cleared by centrifugation using a benchtop mini centrifuge to remove insoluble material. Quality control samples were prepared by pooling equal volumes from each sample. During final reconstitution, isotopically labeled standard molecules, Tryptophan-D3, Carnitine-D9, Valine-D8, and Inosine-4N15, were spiked into each sample to assess LC-MS instrument performance and ionization efficiency.

High resolution (HR) MS and data-dependent acquisition (MS/MS) analyses were performed on a high resolution Q-Exactive HF hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a Vanquish UHPLC binary system and autosampler (Thermo Fisher Scientific, Germany) at the Vanderbilt Center for Innovative Technology. For HILIC analysis metabolite extracts (10 µl injection volume) were separated on a SeQuant ZIC-HILIC 3.5-µm, 2.1 mm × 100 mm column (Millipore Corporation, Darmstadt, Germany) held at 40°C. Liquid chromatography was performed at a 200 µl min^{-1} using solvent A (5 mM Ammonium formate in 90% H_2O, 10% acetonitrile) and solvent B (5 mM Ammonium formate in 90% acetonitrile, 10% H_2O) with the following gradient: 95% B for 2 min, 95-40% B over 16 min, 40% B held 2 min, and 40-95% B over 15 min, 95% B held 10 min (gradient length 45 minutes). For the RPLC analysis metabolite extracts (10 µl injection volume) were separated on a Hypersil Gold, 1.9 µm, 2.1 mm x 100 mm column (Thermo Fisher) held at 40°C. Liquid chromatography was performed at a 250 µl/min using solvent A (0.1% formic acid in H_2O) and solvent B (0.1% formic acid in acetonitrile) with the following gradient: 5% B for 1 minute, 5-50% B over 9 minutes, 50-70% B over 5 minutes, 70-95% B over 5 minutes, 95% B held 2 minutes, and 95-5% B over 3 minutes, 5% B held 5 minutes (gradient length: 30 minutes).

Full MS analyses were acquired over a mass range of m/z 70-1050 using electrospray ionization positive mode. Full mass scan was used at a resolution of 120,000 with a scan rate of 3.5 Hz. The automatic gain control (AGC) target was set at 1×10^6 ions, and maximum ion injection time was at 100 ms. Source ionization parameters were optimized with the spray voltage at 3.0 kV, and other parameters were as follows: transfer temperature at 280 °C; S-Lens level at 40; heater temperature at 325°C; Sheath gas at 40, Aux gas at 10, and sweep gas flow at 1.

Tandem mass spectra were acquired using a data dependent scanning mode in which one full MS scan (m/z 70-1050) was followed by 2, 4 or 6 MS/MS scans. MS/MS scans were acquired in profile mode using an isolation width of 1.3 m/z, stepped collision energy (NCE 20, 40), and a dynamic exclusion of 6 s. MS/MS spectra were collected at a resolution of 15000, with an AGC target set at 2×10^5 ions, and maximum ion injection time of 100 ms. The retention times and peak areas of the isotopically labeled standards were used to assess data quality.

Metabolite data processing and analysis

LC-HR MS/MS raw data were imported, processed, normalized and reviewed using Progenesis QI v.2.1 (Non-linear Dynamics, Newcastle, UK). All MS and MS/MS sample runs were aligned against a quality control (pooled) reference run, and peak picking was performed on individual aligned runs to create an aggregate data set. Unique ions (retention time and m/z pairs) were grouped (a sum of the abundance of unique ions) using both adduct and isotope deconvolutions to generate unique “features” (retention time and m/z pairs) representative of unannotated metabolites. Data were normalized to all features using Progenesis QI. Compounds with <25% coefficient of variance (%CV) were retained for further analysis. Variance stabilized measurements achieved through log normalization were used with Progenesis QI to calculate P-values by one-way analysis of variance (ANOVA) test and adjusted P-values (false discovery rate, FDR). Significantly changed metabolites were chosen with the criteria FDR < 0.05 & |FC| > 1.5.

Tentative and putative identifications were determined within Progenesis QI using accurate mass measurements (<5 ppm error), isotope distribution similarity, and fragmentation spectrum matching based on database searches against Human Metabolome Database (HMDB) [313], METLIN [314], the National Institute of Standards and Technology (NIST) database [315], and an in-house library. In these experiments, the level system for metabolite identification confidence was utilized [316]. Briefly, many annotations were considered to be tentative (level 3, L3) and/or putative (level 2, L2); in numerous circumstances a top candidate cannot be prioritized, thus annotations may represent families of molecules that cannot be distinguished.

In vivo studies

Tumor formation and maintenance assays

These assays were performed by Clare Adams in the laboratory of Christine Eischen at Thomas Jefferson University. Six-week-old athymic nude mice (female *Foxn1*$^{nu/nu}$; Envigo, Indianapolis, IN) were injected subcutaneously into one flank with 10^7 switched or unswitched WT, ∆264, 4A-1 or 4A-2 cells at the Thomas Jefferson Research Animals Shared Resource Core. 4A-1 and 4A-2 carry the same 4A mutation, but were independent clones obtained from the same population. To facilitate lymphoma cell seeding, mice received whole-body irradiation (6 Gy) 24 hours prior to cell injection. For tumor formation studies, lymphoma cells were treated *in vitro* with 4-OHT for 24 hours to induce the switchable *MYC* cassette and expanded for two days prior to being injected into mice. For tumor maintenance studies, mice were injected with unswitched cells and allowed to form palpable tumors. Once tumors reached approximately 200 mm^3, mice received intraperitoneal injections of tamoxifen (2 mg in corn oil, Sigma-Aldrich T5648) once daily for 3 consecutive days to induce the switchable *MYC* cassette *in vivo*. Digital calipers were used to measure tumors and volumes calculated using the ellipsoid formula. Mice were sacrificed at humane endpoints based on tumor volume. Kaplan-Meier survival analyses were compared by log-rank tests to determine statistical significance. For lymphoma cell apoptosis evaluation, a cohort of mice with size-matched tumors prior to tamoxifen injection were sacrificed 48 and 96 hours following the first administration of tamoxifen. Flow cytometry (BD LSRII) at Thomas Jefferson University Flow Cytometry Shared Resource was used to measure fragmented (sub-

G_1) apoptotic DNA with propidium iodide (Sigma-Aldrich P4170), Annexin V/7AAD (BD Pharmingen 559763), and Caspase 3 activity (BD Pharmingen) in the lymphoma cells isolated from the tumors, as we previously reported [317] and as per manufacturer's protocols. Two-tailed t-tests were used to determine significance when comparing two groups. All mouse experiments were approved by the Institutional Animal Care and Use Committee at Thomas Jefferson University and complied with state and federal guidelines.

Tumor gDNA Analysis

To determine the proportion of cells that remain switched in the resulting tumors, gDNA was prepared from tumors using the PureLink Genomic DNA Mini Kit (Invitrogen K182002) according to manufacturer's instructions. As described in Thomas *et al.* [21], gDNA from 0% switched (switchable *MYC* WT cells grown in puromycin) and 100% switched (permanently switched clonal cells) was also prepared in this manner for normalization. The resulting gDNA was diluted down to 50 ng/µl, so that 2 µl (100 ng) was loaded per well for qPCR. qPCR was performed using 2X SYBR FAST qPCR Master Mix (Kapa, Wilmington, Massachusetts, KK4602) in a Bio-Rad CFX96 Real-Time System. For each primer set (MYCP-4 and MYCP-5; SNHG15_F and SNHG15_R) and each independent tumor replicate, the means of three technical replicates was used. MYCP-4 and MYCP-5 primer set only amplifies gDNA from unswitched cells, whereas SNHG15_F and SNHG15_R amplifies gDNA from both unswitched and switched cells. For each gDNA sample, ΔCt was calculated as the difference between MYCP and SNHG15. gDNA from the 0% and 100% switched cells was used to calculate $\Delta\Delta$Ct, which was then used to normalize the ΔCt for the tumor gDNA to estimate the proportion of switched cells.

Tumor RNA-Seq

Tumors from mice sacrificed 48 hours after the first tamoxifen administration were submitted to GENEWIZ for RNA extraction, DNAse treatment, rRNA depletion, library preparation, and 150 bp paired-end sequencing on Illumina HiSeq. Four tumors for each WT, 4A-1, 4A-2, and Δ264 were submitted. The total number of sequencing reads for each replicate is shown in Table 2-2.

Pathway and gene ontology analysis, and figure generation

Classification of annotated metabolites was extracted from HMDB [318] and LIPID MAPS [319]. Metabolite pathway analyses were performed using MetaboAnalyst 4.0 [320], and gene ontology (GO) analyses using Metascape [321] or DAVID [322, 323]. Pathways were created using Cytoscape [324], and bubble plots using ggplot2 [325].

Statistical analysis and replicates for non-high throughput data

Unless otherwise stated, all experiments were conducted with at least three independent, biological replicates, and statistical tests were carried out using PRISM 8 (GraphPad).

Chapter III

Interrogation of the MYC HCF-1 binding motif

Introduction

The HCF-1_{VIC} domain is proposed to form a six-bladed β-propeller structure based on the presence of six Kelch repeats in this region [326]. This domain is responsible for the majority of interactions made by HCF-1, and is critical for most of the known function of HCF-1. Indeed, a point mutation, P134S, within this region was discovered in the tsBN67 cell line derived from a mutagenesis screen, and found to cause temperature-sensitive cell cycle arrest, loss of protein-protein interactions, and dissociation of HCF-1 from chromatin [230, 232]. Based on the primary sequence of the HCF-1_{VIC} domain, this mutation was predicted to be at the base of a loop connecting each blade of the proposed β-propeller structure, leading to identification of equivalent proline residues at each repeat [327]. Substitution of these for serine residues has varying effects on interaction with VP16, with half disrupting this association [327]. Despite this, all of these point mutations caused impaired transcriptional activation by HCF-1 in a reporter assay and failed to rescue the cell growth defect in tsBN67 cells [327].

Proteins that interact with the HCF-1_{VIC} domain contain an HCF-1 binding motif (HBM), a tetrapeptide motif with the canonical sequence $^{D}/_{E}$-H-x-Y [252]. Most validated HBMs adhere to this definition, with only one transcription factor, E2F3, containing a non-canonical HBM [254]. There has been limited exploration into the implications of flexibility in the HBM sequence outside of the invariant second and fourth residues of the motif. Indeed, substitution of the canonical HBM in E2F1 (D-H-Q-Y) for the canonical HBM from VP16 (E-H-A-Y) affects E2F1 function without causing overt changes in its interaction with HCF-1 [253, 254]. The Tansey laboratory previously identified an interaction between MYC and HCF-1, mapping it to a putative HBM within MbIV that is conserved in MYC family members and vertebrates [179]. However, we had yet to determine whether this interaction occurs through the

HCF-1_{VIC} domain, which the presence of an HBM predicts, or if this interaction is sensitive to mutations in this region. Understanding the mode of association between MYC and HCF-1 can provide context for this interaction, such as whether a MYC–HCF-1 could form part of a complex, and may help us exploit and target it for the development of small-molecule inhibitors. Furthermore, the Q-H-N-Y sequence in MYC appears to be non-canonical, posing the question of whether this is influential over MYC function or if we should expand the definition of a canonical HBM to include that found in MYC. While on one hand this clarification could be pivotal in predicting HCF-1-associated proteins, it would also provide insight into the functional consequences and evolutionary drivers of non-canonical binding motifs.

Results

HCF-1 mutations

Based on the presence of a putative HBM, I expected MYC to interact with the HCF-1_{VIC} domain. Mutations in HCF-1_{VIC}, including the P134S mutation from tsBN67 cells, have previously been used to challenge HBM-dependent interactions, including the prototypical HCF-1 interaction partner VP16 [258, 327]. I first confirmed how four mutations, P134S, P197S, P252S, and P319S, affected interaction with VP16 by performing transient co-transfection of T7-HCF-1_{VIC} and FLAG-$VP16_N$ into HEK393T cells. With FLAG immunoprecipitation (IP) of VP16, I found that three out of the four mutations impaired co-IP of T7-HCF-1_{VIC} (Figure 3-1A). The P252S mutation had minimal effect on the HCF-1–VP16 interaction, consistent with what has previously been observed [327]. I then tested these mutations in the context of the MYC–HCF-1 interaction, again by transient co-transfection of T7-HCF-1_{VIC} and FLAG-Gal4-HA-MYC central portion (CP), and FLAG IP. The HCF-1_{VIC} domain was sufficient for interaction with MYC CP, and inclusion of the MYC CP 4A mutation disrupted co-IP of HCF-1_{VIC} (Figure 3-1B). Similar to VP16, only the P252S mutation was unable to affect interaction with MYC CP, with the other mutations causing extensive disruption to their binding (Figure 3-1B). Thus, the MYC–HCF-1 interaction is HBM-dependent and occurs through the HCF-1_{VIC} domain.

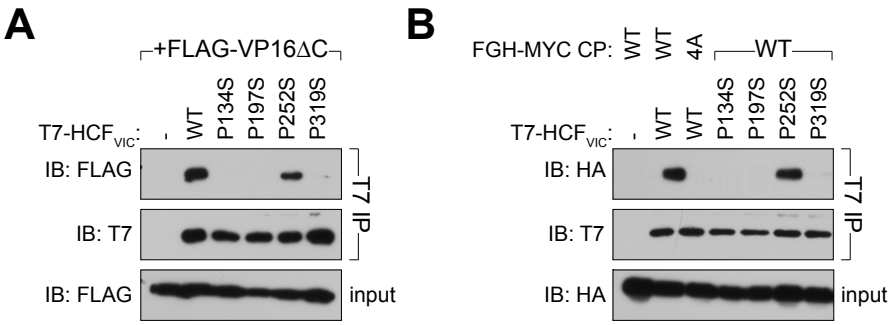


Figure 3-1: MYC interacts with the VIC domain of HCF-1
(**A**) and (**B**) T7-tagged HCF-1$_{VIC}$, either WT or mutant, was transiently co-transfected with FLAG-tagged VP16ΔC (A) or HA- and FLAG-tagged MYC central portion (CP; B) into HEK293T cells. T7 immunoprecipitation (IP) was performed and the resulting eluates probed for pull-down of VP16ΔC or MYC CP.

Further validation of the MYC HBM

According to primary sequence conservation of validated HBM-containing proteins (Figure 1-5B) and the proposed canonical HBM ($^{D}/_{E}$-H-x-Y), the second and fourth residues of the HBM are invariant and likely necessary for interaction with HCF-1. To provide further validation that the Q-H-N-Y sequence in MbIV is an HBM, I introduced glycine and/or alanine substitution mutations at these residues in the context of both CP and full-length MYC, and compared these to WT, and WDR5 binding motif mutations (V264G and WBM; Figure 3-2A). I performed transient transfection and co-IP of these constructs from HEK293T cells, and found that mutation of either the invariant second (H307G, H307A) or fourth (Y309A) residues of the HBM robustly perturbed interaction with endogenous HCF-1 (Figure 3-2B and 3-2C). Cleavage of HCF-1 by OGT results in multiple fragments of varying size [328], with the expression levels, and resulting co-IP, of these approximately equal in HEK293T cells (Figure 3-2B and 3-2C). In full-length MYC, HBM mutations also appeared to have a slight effect on WDR5 binding, but this was relatively minor compared to the impact of the the WBM mutation, which also had little-to-no effect on HCF-1 binding (Figure 3-2C).

The MYC transactivation domain is subject to O-GlcNAcylation by OGT [57], and about half of nuclear OGT is reported to interact with HCF-1 [246]. Based on this, I proposed that OGT would be associated with at least two regions of MYC, including through the transactivation domain and via HCF-1 at MbIV. By probing the IP of MYC CP, I found that OGT is associated with the CP and this is impacted by HBM mutations to a similar degree as HCF-1 (Figure 3-2B). In contrast, the effect of these mutations in full-length MYC was only modest (Figure 3-2C), indicating the OGT interacts with MYC both within and outside of the MYC CP.

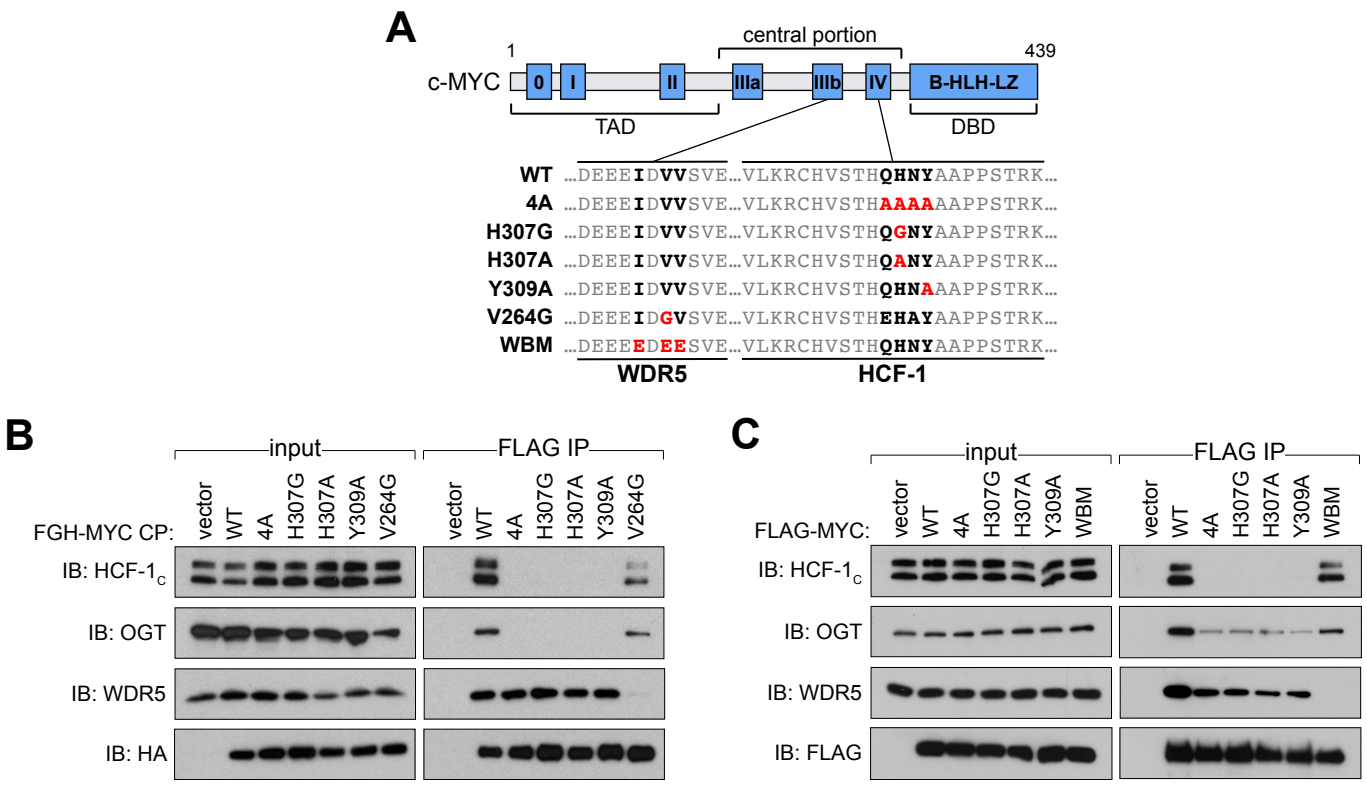


Figure 3-2: MbIV contains a validated HBM
(**A**) Schematic of MYC, depicting the location of the six MYC boxes (Mb0–MbIV). MbIIIb carries a WDR5-binding motif (WBM). MbIV contains an HCF-1 binding motif (HBM). Residues relevant to the WBM or HBM are in bold, and residues mutated in these experiments are in red. (**B**) and (**C**) HA and FLAG-tagged central portion (B) or FLAG-tagged full-length (C) MYC carrying the mutations described in (A) were transiently expressed in HEK293T cells, recovered by anti-FLAG IP. The input and IP eluates were probed for the presence of HCF-1_C, OGT, WDR5, and HA or FLAG-tagged proteins by western blotting.

Exploring the "non-canonical" MYC HBM

Non-canonical HBMs have been validated in a number of proteins, including BAP1 and E2F3, but if and how these influence their interaction with HCF-1 has not been investigated. I addressed this in the context of the MYC–HCF-1 interaction by substituting the MYC HBM for the prototypical VP16 HBM (Q-H-N-Y to E-H-A-Y, Figure 3-3A). As a control for any effect that this substitution had, I additionally included the H307G point mutation in this context (VP16 HBM: H307G). CP or full-length MYC containing the mutations described in Figure 3-3A were transiently transfected into HEK293T cells, and FLAG IP was performed. With MYC CP, I found that the VP16 HBM substitution had no overt effect on the interaction between MYC and HCF-1 when compared to WT (Figure 3-3B). The 4A, H307G, and VP16 HBM:H307G mutations all similarly disrupted the MYC–HCF-1 interaction (Figure 3-3B). Thus, in this context, the VP16 HBM can appropriately substitute for the MYC HBM. However, IP of full-length MYC with the VP16 HBM resulted in a substantial increase in the amount of HCF-1 and OGT pulled-down (Figure 3-3C). The specificity of this increased binding is reflected in unperturbed co-IP of WDR5, and no detectable interaction of HCF-1 with the VP16 HBM:H307G mutation (Figure 3-3C). The contrast in findings between CP and full-length MYC suggests that regions outside of the CP may influence its association with HCF-1, and inclusion of the VP16 HBM is sufficient to overcome this.

To determine which regions present in full-length MYC affect the interaction with HCF-1, I generated MYC constructs whereby either the transactivation domain or DBD was deleted, and included the 4A or VP16 HBM mutations in these contexts. Transient transfection and FLAG IP found that, with deletion of the MYC DBD, the VP16 HBM mutation still caused increased binding to HCF-1 compared to WT (Figure 3-3D). In contrast, the VP16 HBM mutation was equivalent to WT when the transactivation domain was deleted (Figure 3-3D). Finally, the amount of HCF-1 co-immunoprecipitated by ΔTAD/WT HBM appears equivalent to that with ΔDBD/VP16 HBM (Figure 3-3D), suggesting that the transactivation domain causes decreased HCF-1 binding to MbIV, with this suppressed by the VP16 HBM mutation, rather that the VP16 HBM causing increased binding.

The interaction between MYC and HCF-1 is likely direct, based on my characterization of the HBM, but I confirmed this using an *in vitro* binding assay. I performed FLAG IP of recombinant full-length FLAG-MYC, WT, 4A, or VP16 HBM, and *in vitro* transcribed/translated T7-HCF-1_{VIC}. With this system, I validated a direct interaction between MYC and HCF-1, and found that, akin to what I observed in cells, the 4A mutation disrupts and the VP16 HBM mutation increases this association (Figure 3-3E). These mutations can thereby function as loss- and gain-of-function tools in studying the direct interaction between MYC and HCF-1.

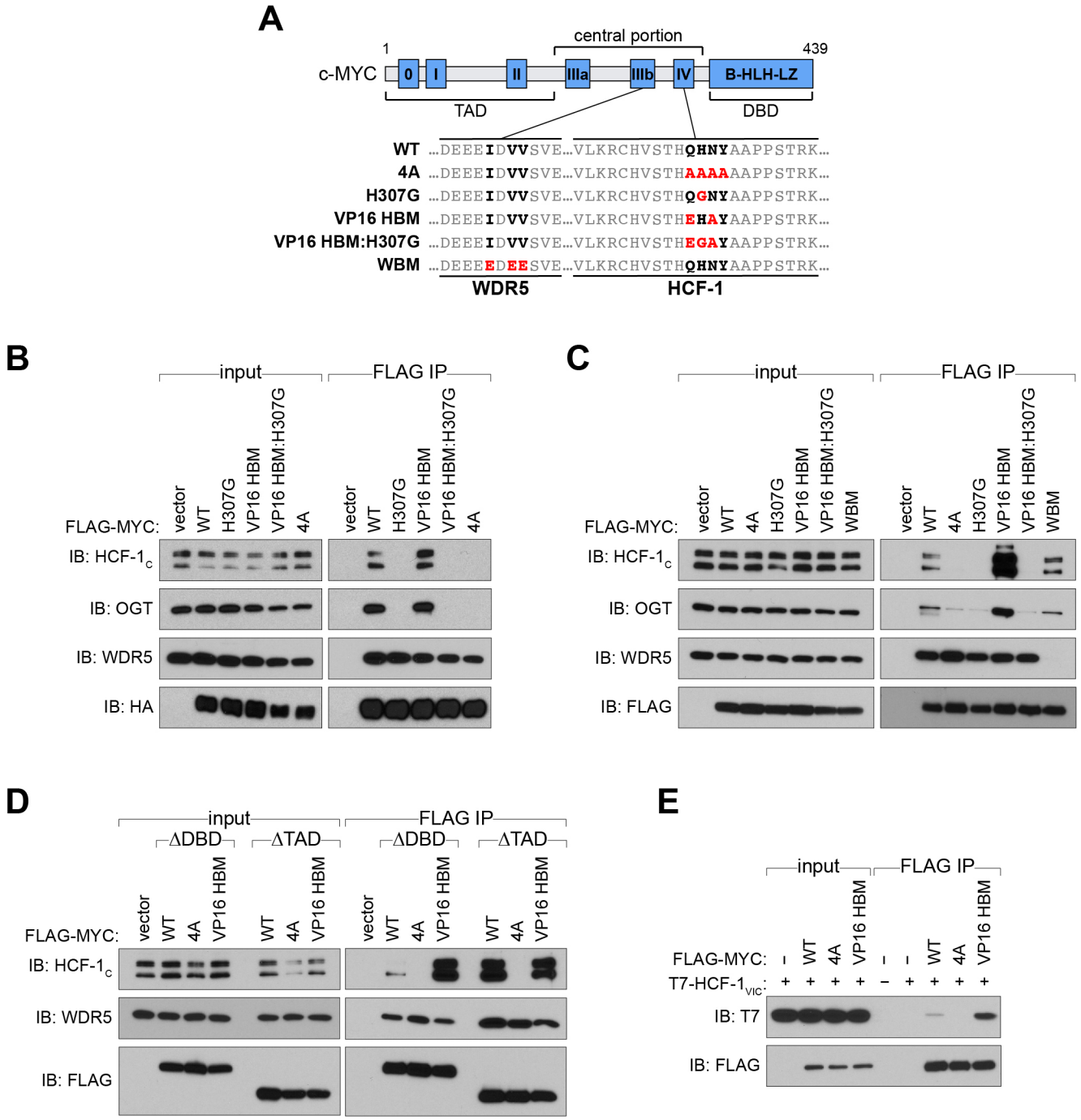


Figure 3-3: Modulating the MYC–HCF-1 interaction
(**A**) Schematic of MYC, depicting the location of the six MYC boxes (Mb0–MbIV). MbIIIb carries a WDR5-binding motif (WBM). MbIV contains an HCF-1 binding motif (HBM). Residues relevant to the WBM or HBM are in bold, and residues mutated in these experiments are in red. (**B**) and (**C**) HA and FLAG-tagged central portion (B) or FLAG-tagged full-length (C) MYC carrying the mutations described in (A) were transiently expressed in 293T cells, recovered by anti-FLAG IP. Input and IP eluates were probed for the presence of HCF-1_C, OGT, WDR5, and HA or FLAG-tagged proteins by western blotting. (**D**) The transactivation (TAD) or DNA binding (DBD) domains were deleted from the WT, 4A, and VP16 HBM constructs used in (C), and used for transient transfection of 293T cells. IP was performed using anti-FLAG M2 agarose, and western blots of the input lysate and the IP eluate were

probed using antibodies against HCF-1_C, WDR5, and FLAG. (**E**) *In vitro* transcribed/translated T7-tagged HCF-1_{VIC} was incubated with recombinant full-length, FLAG-tagged MYC, either WT or mutant (4A or VP16 HBM), and IP performed using anti-FLAG M2 agarose. Western blot of the input lysate and the IP eluate was probed using antibodies against the T7 and FLAG tags.

Discussion

Numerous groups have reported on the interaction between MYC and HCF-1, finding they overlap on chromatin by ENCODE data [296], are co-localized in the nucleus by proximity ligation assay [329], and depend on OGT for association [330]. However, none of these works have sought to uncover the molecular elements defining the MYC–HCF-1 interaction. Such an understanding, which I have conveyed here both from the perspective of MYC and HCF-1, is necessary for ascertaining the context in which this interaction functions, and is an essential accomplishment when considering the potential of targeting the MYC–HCF-1 interaction with small-molecule inhibitors.

The specific proline residues within the HCF-1_{VIC} domain that I targeted for mutation are understood to be localized to loops that connect the Kelch repeats, and these mutations have been proposed to cause disruption to the β-propeller structure [327]. This is plausible considering mutation of at least three of these loops is sufficient to impair interaction with HBM-containing proteins, meaning our ability to clearly define interaction surfaces is obstructed. It has been reported that multiple HBM-containing proteins can interact with the HCF-1_{VIC} domain simultaneously [266], pointing to the possibility of multiple similar interaction surfaces or to a mechanism by which these interactions are not mutually exclusive. Although I did not investigate whether additional HBM-containing proteins are associated with MYC by way of HCF-1, I did find that HCF-1 tethers OGT to MYC. While a similar tethering by HCF-1 has been observed previously [249], the functional relevance of the MYC–HCF-1–OGT complex may be clouded by the additional interaction of OGT with the MYC transactivation domain. Based on my interrogation of these interactions, I believe the binding of the MYC transactivation domain to OGT occurs independently of HCF-1, and the OGT involved in this interaction is unable to simultaneously associate with HCF-1. This may stem from differential localization, with the MYC–HCF-1 complex predominantly nuclear [297] and O-GlcNAcylation of MYC proposed to promote its nuclear localization [331].

I found that substitution of the MYC HBM for that from VP16 can increase co-immunoprecipitation of HCF-1, suggesting that the wild-type MYC HBM is relatively weak in its ability to interact with HCF-1. This is context dependent, with the gain-of-function nature of this mutation lost with deletion of the transactivation domain of MYC, but not the DBD. Indeed, in the context of the ΔTAD mutation, the interaction of WT MYC with HCF-1 was equivalent to that of the VP16 HBM mutation with ΔDBD. These data indicate that the MYC transactivation domain may influence the interaction between MYC and HCF-1, but this influence is eliminated with the VP16 HBM mutation, suggesting the MYC HBM has evolved specifically to enable governance by the MYC transactivation domain. Curiously, the finding that substitution of the wild-type HBM for the VP16 HBM, or deletion of the transactivation domain increases interaction with HCF-1 is not unique. In a yeast two-hybrid assay, Tyagi *et al.* [254] found that E2F1 is unable to interact with HCF-1 unless its transactivation domain is deleted or its canonical HBM (D-H-Q-Y) is replaced with that from VP16. In contrast, E2F1 with the VP16 HBM mutation does not cause an overt increase in binding to HCF-1 in transient transfection of U2OS cells and co-IP, but does increase E2F1-driven DNA damage and apoptosis [253]

Based on my observations and the precedent set by E2F1, I hypothesize that the MYC transactivation domain is auto-inhibitory of its interaction with HCF-1, and this can be overcome by either deleting the transactivation domain or by substituting in the VP16 HBM. These findings also suggest that the non-canonical HBM in MYC is not innately more efficient in binding to HCF-1, and rather the influence of the MYC transactivation domain makes it less efficient. This idea is further emphasized by evidence that deletion of the VP16 transactivation domain also increases VP16 induced complex formation [332], evoking a model in which the HBM sequence present is specifically adept at enabling transactivation domain-modulated auto-inhibition of interaction with HCF-1. Beyond this work, long-range intra-molecular interactions and influences have been increasingly identified with MYC, such as the phosphorylation of MbI by AURKB, which interacts at MbIIIb [67]. From a different perspective, this may additionally indicate that MbIV is auto-inhibitory of the MYC transactivation domain. The transactivation potential of a GAL4/MYC fusion is amplified with the exclusion of exon 3 [72], suggesting that this

region, which includes MbIV, influences the activity of the MYC transactivation domain. Thus, it could be hypothesized that HCF-1 helps fine-tune MYC-driven transcription by regulating these long-range interactions.

The innate capacity of the “non-canonical” MYC HBM to bind with a similar strength to HCF-1 as a canonical HBM, as evidenced by the equivalent binding of the MYC central portion to HCF-1 regardless of the HBM sequence, draws three final conclusions. The first is that the canonical HBM should be re-defined as $^{B}/_{Z}$-H-x-Y, where B is aspartic acid or asparagine, and Z is glutamic acid or glutamine. While this does not encompass all validated HBMs, with E2F3 remaining an exception (Figure 1-5B), it does incorporate the MYC and BAP1 HBMs as being canonical, and opens the door for the prediction and identification of additional HCF-1 cofactors and HBM-containing proteins. The second conclusion is that HCF-1-associated transcription factors likely specifically evolve a relationship between the HBM sequence and the transactivation domain. MYC is the third such transcription factor in which deletion of the transactivation domain has been found to influence association with HCF-1, and the second whereby replacement of the wild-type HBM for the prototypical VP16 HBM has a similar outcome. Thus, although we should expand the definition of the canonical HBM, we should also appreciate that canonical HBMs are not immediately interchangeable, with substitution of one canonical HBM for another likely to have functional consequences. Finally, if the default interaction of MYC and HCF-1 is weak, then it could be hypothesized that the amount of complexed MYC–HCF-1 only becomes *functionally* relevant when MYC is overexpressed, or when more MYC is competent for interaction with HCF-1. Perhaps the biggest consequence of this hypothesis is that it suggests there is a therapeutic window for targeting MYC-driven cancers through HCF-1; specifically inhibiting their association might only have ramifications when MYC levels are high, meaning that cells in which MYC overexpression is not present, would face minimal, or at least fewer, consequences. This may also suggest that interaction with MYC may be overtly sensitive to inhibition of the HCF-1_{VIC} domain.

Chapter IV

Function of the MYC–HCF-1 interaction

Introduction

Having defined the molecular elements underlying the MYC–HCF-1 interaction, I next sought to exploit these to determine the function of their association. The region in MYC through which HCF-1 interacts, MbIV, is the last remaining MYC box with an unknown function. Based on the transcriptional roles of MYC and HCF-1 independently, and their co-localization in the nucleus [297] and on chromatin [329], I anticipate that the functional contribution of their interaction is in the control of transcription. Because of this, the ramifications of perturbing their association, with either the loss- or gain-of-function mutations described in the previous chapter, would be two-fold; gene expression changes would be expected to manifest as downstream phenotypic consequences. To understand the totality of the function of the MYC–HCF-1 interaction, I interrogated both of these aspects using a switchable *MYC* exon system in the context of a MYC-driven cancer cell line.

Our understanding of MYC, including its contribution to transcription and the processes it drives, has largely been derived from overexpression systems. I also noted in the previous chapter that, due to the tempered nature of their association, the MYC–HCF-1 interaction may only achieve functional relevance when MYC levels are high. Thus, to enable me to make comparisons to what is already known in the MYC field, and to ensure the MYC–HCF-1 interaction is applicable, I chose to focus on a cell line in which MYC is overexpressed. Burkitt lymphoma (BL) is a classic MYC-driven cancer that primarily results from translocation of *MYC* to the IgH chain locus, resulting in increased transcription of *MYC* [29]. Specifically, I employed Ramos cells, a BL line that contains this aforementioned translocation [333], due to the absence of additional, population-scale mutations in the translocated *MYC* allele [31], and the Epstein-Barr virus-negative status of these cells [334]. BL also bears many of

the characteristics typical of a MYC-driven cancer, including reprogrammed metabolism [335, 336], altered cell cycling [337], partial resistance to apoptosis [338, 339], and a capacity to form solid tumors in mice [340].

Traditionally, introduction of exogenous genes or knockdown have been applied to probe the roles of MYC and HCF-1 in cell function, largely because they are encoded by essential genes. To overcome many of the limitations associated with these approaches, I instead applied a switchable exon system to *MYC* to enable rapid substitution of a wild-type exon 3 for one that contains either the loss- or gain-of-function mutations shown in Figure 3-3A. This system, akin to that described in Thomas *et al.* [21], is an efficient means to assess how specifically altering the MYC–HCF-1 interaction affects cells over a relatively short timescale, without introducing exogenous MYC or relying on long-term consequences. Additionally, in Ramos cells, only the translocated *MYC* allele is expressed, while the other allele is inactivated, possibly through heterochromatinization [31]. Thus, by replacing the endogenous exon 3 in these cells with the switchable allele, I can create a cell line in which the only form of MYC being expressed carries the mutations of interest.

MYC and HCF-1 have broad functions in cells, with both typically linked to transcriptional regulation, but their dysfunction can lead to widespread changes, including in growth, cell cycle, and metabolism. For each of these proteins, the mechanisms by which they promote transcriptional activation has been characterized at only a fraction of their targets, and I expect my interrogation of the MYC–HCF-1 interaction to expand our knowledge of the genes they modulate, and may unveil novel roles for MYC and HCF-1 in transcriptional regulation. While I anticipate the loss- and gain-of-function mutations in MYC to affect only a subset of MYC target genes, this can help guide our understanding of the specific contribution of HCF-1 to MYC function on a cellular level and to MYC-driven tumorigenesis, and can provide a proof-of-concept for how inhibiting this interaction may facilitate the treatment of cancer.

Results

Inducible exon switch of MYC

To understand the function of the MYC–HCF-1 interaction, I used CRISPR/Cas9-directed homologous recombination to replace the endogenous *MYC* exon 3 with a switchable allele in Ramos cells stably expressing CRE-ER^{T2} (Figure 4-1A). This was done using co-transfection of a targeting vector containing the switchable allele flanked by homology arms up- and downstream of *MYC* exon 3, a plasmid encoding a guide RNA against the end of the *MYC* open reading frame, and a Cas9-expressing plasmid. The switchable allele contains a wild-type (WT) exon 3 and a P2A-linked puromycin selectable marker, both flanked by loxP sites (Figure 4-1A). Activation of CRE-ER^{T2} by addition of 4-hydroxytamoxifen (4-OHT) causes excision of the WT exon 3, and this is replaced by a downstream mutant exon 3, with WT, 4A, or VP16 HBMs (Figure 4-1A). Following puromycin selection and cell sorting, individual clones were validated by Southern blot (Figure 4-1B and 4-1C), flow cytometry (Figure 4-1D), western blot (Figure 4-1E), and co-IP (Figure 4-1F and 4-1G). A green fluorescent protein (GFP) marker is linked to the mutant exon 3, and by this measure, I estimate at least 85% of cells are switched to the mutant exon within 24 hours (Figure 4-1D). Also corresponding to a switch to the mutant exon 3 is appearance of an HA-tagged species and a shift in the molecular weight of MYC (Figure 4-1E). Importantly, this switching caused no overt change in MYC or HCF-1 levels (Figure 4-1E), and the loss- and gain-of-function effects of the 4A (Figure 4-1F) and VP16 HBM (Figure 4-1G) were maintained in this context.

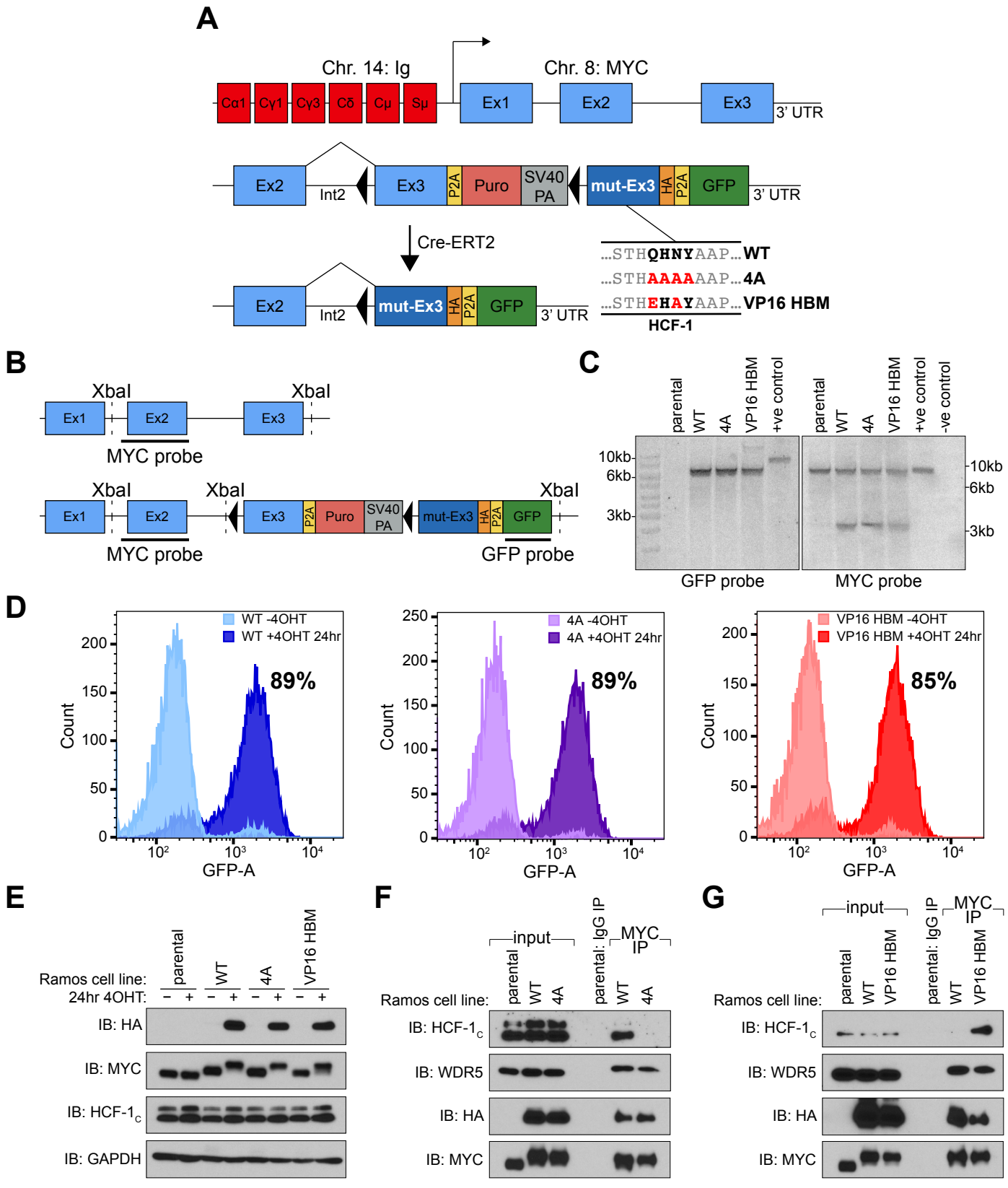

A
Chr. 14: Ig
Chr. 8: MYC
Cα1 Cγ1 Cγ3 Cδ Cμ Sμ
Ex1 Ex2 Ex3 3' UTR
Ex2 Int2 Ex3 P2A Puro SV40 PA mut-Ex3 HA P2A GFP 3' UTR
Cre-ERT2
Ex2 Int2 mut-Ex3 HA P2A GFP 3' UTR
...STHQHNYAAP... WT
...STHAAAAAAP... 4A
...STHEHAYAAP... VP16 HBM
HCF-1
B
XbaI XbaI
Ex1 Ex2 Ex3
MYC probe
XbaI XbaI XbaI
Ex1 Ex2 Ex3 P2A Puro SV40 PA mut-Ex3 HA P2A GFP
MYC probe
GFP probe
C
parental WT 4A VP16 HBM +ve control -ve control
10kb 6kb 3kb
GFP probe
MYC probe
D
WT -4OHT
WT +4OHT 24hr
89%
4A -4OHT
4A +4OHT 24hr
89%
VP16 HBM -4OHT
VP16 HBM +4OHT 24hr
85%
Count
GFP-A
E
Ramos cell line:
24hr 4OHT:
parental WT 4A VP16 HBM
IB: HA
IB: MYC
IB: HCF-1C
IB: GAPDH
F
input
MYC IP
Ramos cell line:
parental WT 4A parental: IgG IP WT 4A
IB: HCF-1C
IB: WDR5
IB: HA
IB: MYC
G
input
MYC IP
Ramos cell line:
parental WT VP16 HBM parental: IgG IP WT VP16 HBM
IB: HCF-1C
IB: WDR5
IB: HA
IB: MYC

Figure 4-1: Switchable MYC allele in Ramos cells

(Previous page)

(**A**) The translocated *MYC* locus from Ramos cells is depicted at top, with chromosome 14 (red) and 8 (blue) elements indicated. Beneath is a representation of the locus modification, in either the unswitched (middle) or switched (bottom) states. This switchable allele contains a wild-type (WT) exon 3, a P2A-linked puromycin cassette, and a SV40 polyadenylation (SV40 PA) signal, all of which are flanked by LoxP sites (black triangles). Downstream of the LoxP-flanked region is an HA-tagged mutant exon 3 (mut-Ex3) and a P2A-linked GFP cassette, with the endogenous 3' untranslated region (UTR) intact. Activation of CRE-ER^{T2} results in excision of WT exon 3 and its replacement with mutant exon 3 which carries sequences encoding either WT or mutant (4A or VP16 HBM) MYC protein. (**B**) Comparison of the structure of the parental (non-modified) *MYC* allele (top) compared to the switchable *MYC* allele (bottom). XbaI sites used for digestion of genomic DNA in Southern blot are highlighted, as are the complementary sites for the MYC and GFP probes. The expected products of XbaI digestion for the parental line are 6,784 bp for the MYC probe (which detects both the translocated and non-translocated alleles) and nothing for the GFP probe; for correctly-engineered lines the expected sizes are 6,784 bp and 2,942 bp for the MYC probe, and 6,624 bp for the GFP probe. (**C**) Southern blot using GFP and MYC probes on XbaI-digested gDNA from unswitched parental or switchable cells (WT, 4A, and VP16 HBM), with digested positive and negative control plasmids. (**D**) Switchable cells were treated with or without 20 nM 4-OHT (24 hours), fixed using 1% formaldehyde, and subject to flow cytometry. The GFP profiles of the -4-OHT and +4-OHT cells are shown overlaid onto the same axes, with the approximate percentage of GFP-positive cells for 24 hours +4-OHT shown. (**E**) Western blot of lysates from parental (CRE-ER^{T2}) or switchable Ramos cells (WT, 4A, or VP16 HBM) ± 20 nM 4-OHT for 24 hours. Blots were probed with antibodies against the HA tag, c-MYC, HCF-1$_C$, and GAPDH. (**F**) and (**G**) Parental (CRE-ER^{T2}) or switchable Ramos cells (WT, 4A, or VP16 HBM) were treated with 20 nM 4-OHT for 24 hours, lysates prepared, and IP performed using anti-IgG or anti-MYC antibodies. Input lysates and IP eluates were probed using antibodies against HCF-1$_C$, WDR5, HA tag, and MYC by western blotting.

Growth and glutamine defects of MYC mutants

Both MYC and HCF-1 have well-documented contributions to cell growth and proliferation, leading me to ask whether their interaction is also important to these processes. The switchable system I have used here enabled me to look at the effect these mutants have when they are expressed as the only form of MYC in a cell, while the GFP marker is used to track which cells were switched and express the mutants. First, I confirmed that the distribution of MYC and HCF-1 was unaffected by these mutants using chromatin fractionation (Figure 4-2A) and immunofluorescence (Figure 4-2B). I then exploited the GFP marker to interrogate how altering the MYC–HCF-1 interaction affects cell growth, specifically in competition between unswitched, WT cells and their switched, mutant counterparts. I did this by switching approximately 50% of cells using a 2 hour treatment with 4-OHT, and tracking the proportion of GFP-positive cells over time. Consistent with the effect the 4A and VP16 HBM mutants have on interaction with HCF-1, I found that these mutants also had opposite effects on long-term Ramos cell growth, with the 4A causing a selective disadvantage and the VP16 HBM a selective advantage compared to their WT counterparts (Figure 4-2C).

In cancer, including BL, MYC is known to contribute to a glutamine addiction phenotype [101, 336] and to cell cycle progression [337]. The selective disadvantage of the 4A and selective advantage of the VP16 HBM was reversed if cells were also deprived of glutamine overnight, which instead caused a selective advantage for the 4A cells and a selective disadvantage for the VP16 HBM cells (Figure 4-2D). On day seven of an equivalent experiment to that in Figure 4-2C, I evaluated the cell cycle distribution of GFP-negative and GFP-positive cells using propidium iodide incorporation. With this I observed a small, but statistically significant decrease of 4A and an increase of VP16 HBM cells in the G_2/M phase, with the VP16 HBM cells also showing a reduction in the proportion of cells in the G_1 phase (Figure 4-2E). The proportion of sub-G_1 cells was unaffected in either of these mutants (Figure 4-2E), suggesting that the effects observed in Figure 4-2C were due to an altered growth rate of Ramos cells, rather than cell death.

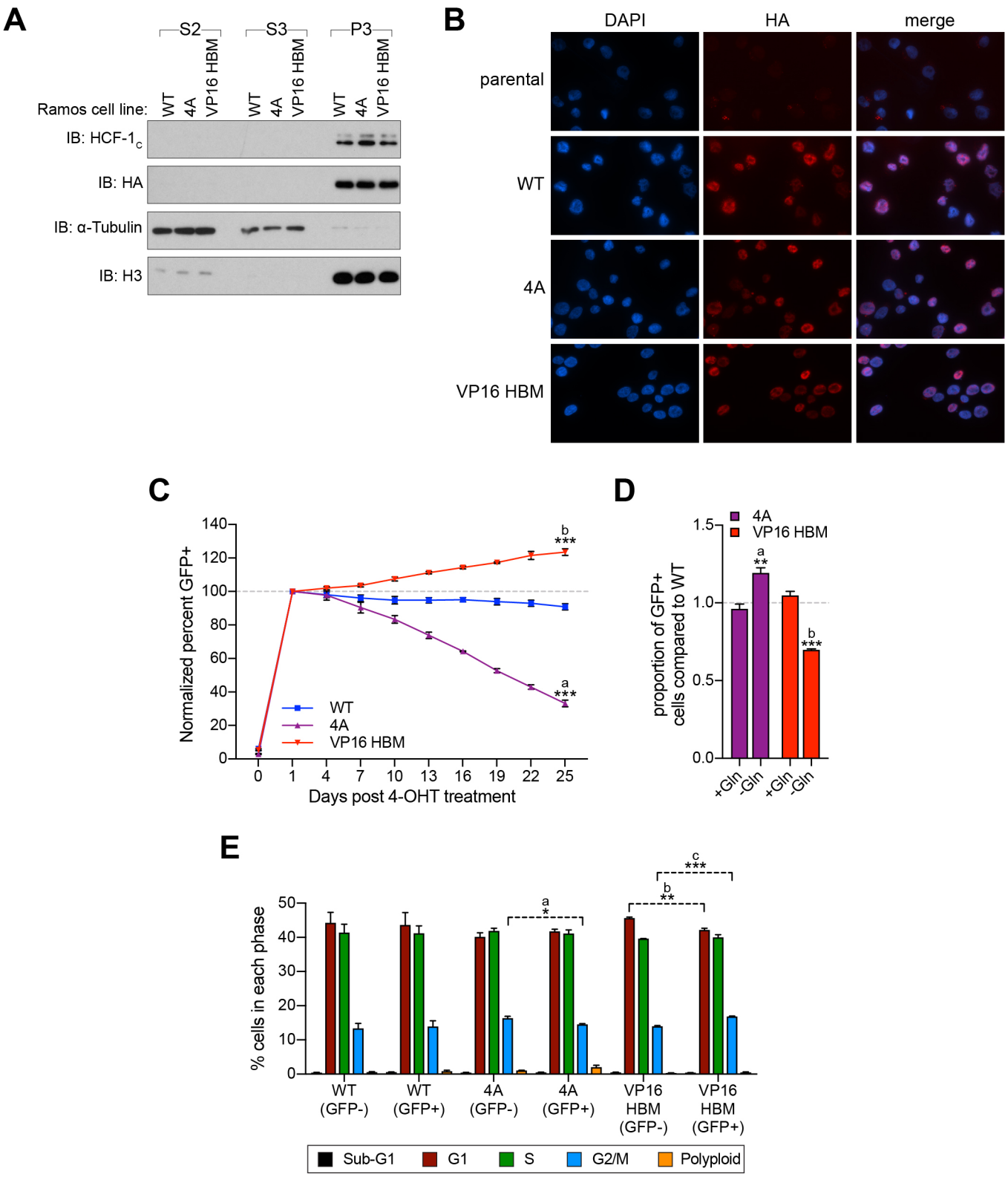


Figure 4-2: Perturbation of glutamine-dependent and independent growth
(A) Chromatin fractionation of switchable *MYC* Ramos cells was performed after a 24 hour treatment with 4-OHT. S2 reflects the cytosolic, S3 the non-chromatin nuclear, and P3 the chromatin-bound fraction. Each fraction was

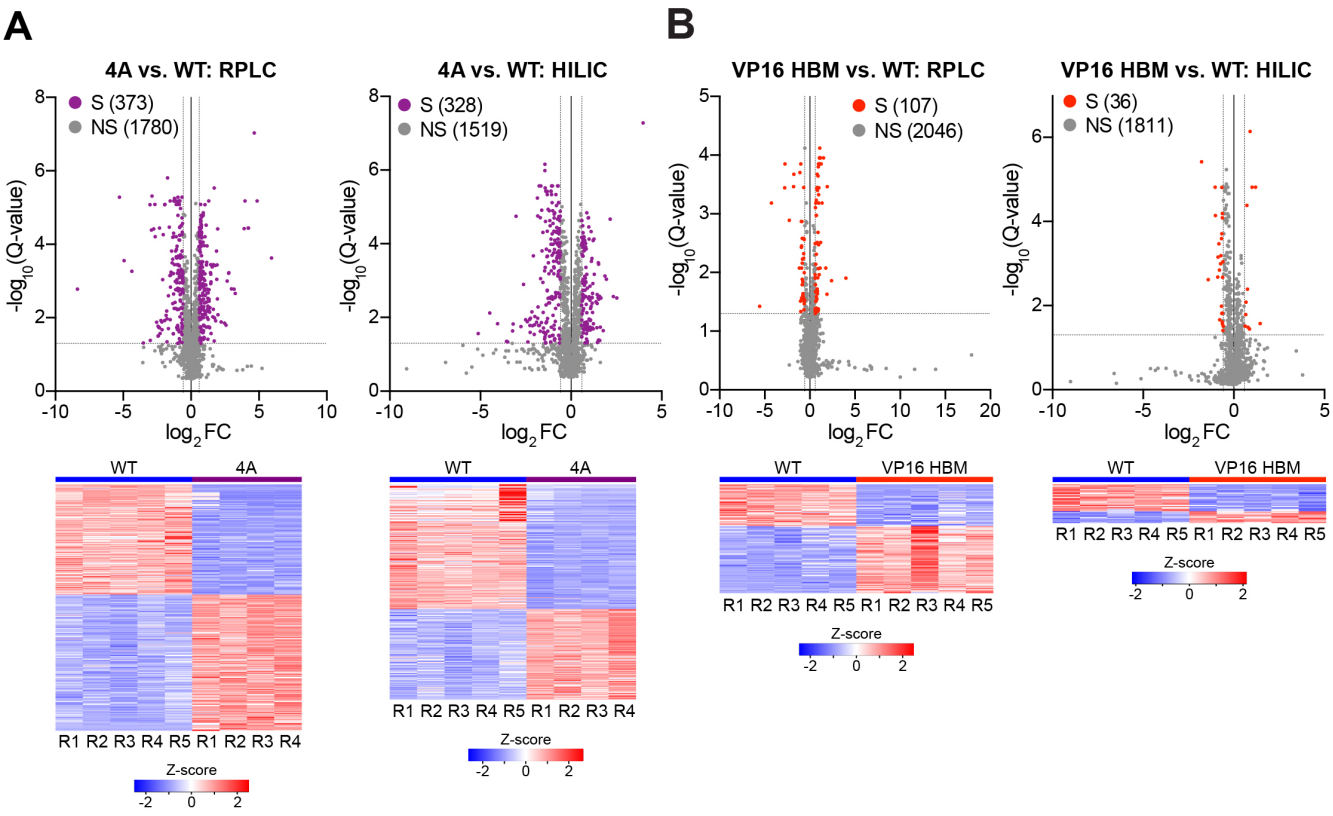
probed using antibodies against HCF-1_C, HA, α-Tubulin, and histone-H3. **(B)** Immunofluorescence of cells that had been switched for 24 hours, using a DAPI stain and antibody against the HA tag on MYC. (**C**) Switchable Ramos cells were pulsed with 20 nM 4-OHT for two hours to switch approximately 50% of cells, and the proportion of GFP-positive cells measured by flow cytometry 24 hours after treatment and every three days following. For each of the replicates, the proportion of GFP-positive cells is normalized to that on day one. Shown are the mean and standard error for three biological replicates. Student's t-test between WT and each of the mutants at day 25 was used to calculate P-values; a=0.000028, b=0.00026. (**D**) Switchable Ramos cell lines were pulsed with 4-OHT as in (A), propagated for three days, and grown for 16 hours in media with or without glutamine. The impact of glutamine deprivation was measured by flow cytometry to determine the proportion of GFP-positive (switched) cells. For each of the mutants, the proportion of GFP-positive cells was normalized to that for WT cells. Shown are the mean and standard error for three biological replicates. Student's t-test between +Gln and -Gln was used to calculate P-values; a = 0.0066, b = 0.0002. (**E**) Switchable Ramos cells were pulsed with 4-OHT as in (A), grown for seven days, and cell cycle distribution determined by propidium iodide (PI) staining and flow cytometry, binning cells according to whether they expressed GFP (GFP+, switched) or not (GFP–, unswitched). Shown are the mean and standard error for three biological replicates. Student's t-test between GFP- and GFP+ cells was used to calculate P-values; a = 0.033, b = 0.0041, c = 0.0006.

Widespread changes in metabolites

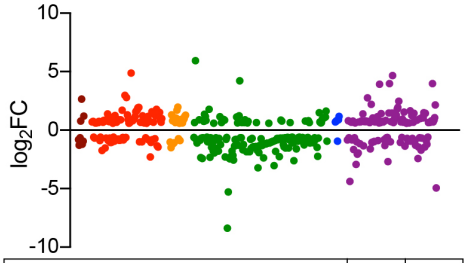
The known function of both MYC and HCF-1 in metabolism and mitochondrial function, combined with my finding that altering the MYC–HCF-1 interaction affects glutamine-dependency of Ramos cells, suggested that this interaction may contribute to metabolic processes in these cells. To address this, I formed a collaboration with the Vanderbilt Center for Innovative Technology, and investigated whether the 4A and VP16 HBM MYC mutants caused global effects on metabolite levels. We performed untargeted metabolomics using two liquid chromatography methods, reverse phase liquid chromatography (RPLC) and hydrophilic interaction liquid chromatography (HILIC), and tandem mass spectrometry on switchable *MYC* cells that had been treated with 4-OHT for 24 hours.

Both RPLC and HILIC approaches enabled detection of ~2,000 metabolites (Figures 4-3A to 4-3B), although only a little over half of these could be matched with a level of confidence to a known metabolite. Regardless of the approach used, more metabolites were significantly changed in the 4A mutant compared to the VP16 HBM mutant, and the magnitude of these changes was typically larger for the 4A mutant (Figures 4-3A to 4-3D). For those metabolites that were significantly affected, the replicates showed strong consistency in levels (Figures 4-3A to 4-3B, bottom row).

The type of metabolites impacted by altering the MYC–HCF-1 interaction was assessed by broad classification and pathway analyses. For both 4A and VP16 HBM mutants, the primary classes of metabolites affected were linked to amino acids and lipids (Figures 4-3C and 4-3D). This was further emphasized by pathway analysis, finding that the most significant and impacted pathways were those related to amino acid metabolism, aminoacyl-tRNA biosynthesis, and glycerophospholipid metabolism (Figures 4-3E and 4-3F). In general, enrichment of pathways was stronger with the 4A mutant than the VP16 HBM mutant, likely due to the disproportionate number of significant metabolite changes. The similarities in the classes of metabolites affected raises the unsurprising possibility that these mutants influence the same pathways in cells.

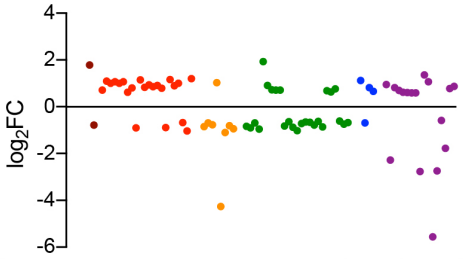


C



4A vs. WT	Down	Up
Amines	5	3
Amino acids, peptides, and analogs	31	61
Carbohydrates and conjugates	6	15
Lipids	132	43
Nucleosides, nucleotides, and analogs	1	5
Other	42	73

D



VP16 HBM vs. WT	Down	Up
Amines	1	1
Amino acids, peptides, and analogs	4	18
Carbohydrates and conjugates	7	1
Lipids	17	8
Nucleosides, nucleotides, and analogs	1	3
Other	6	11

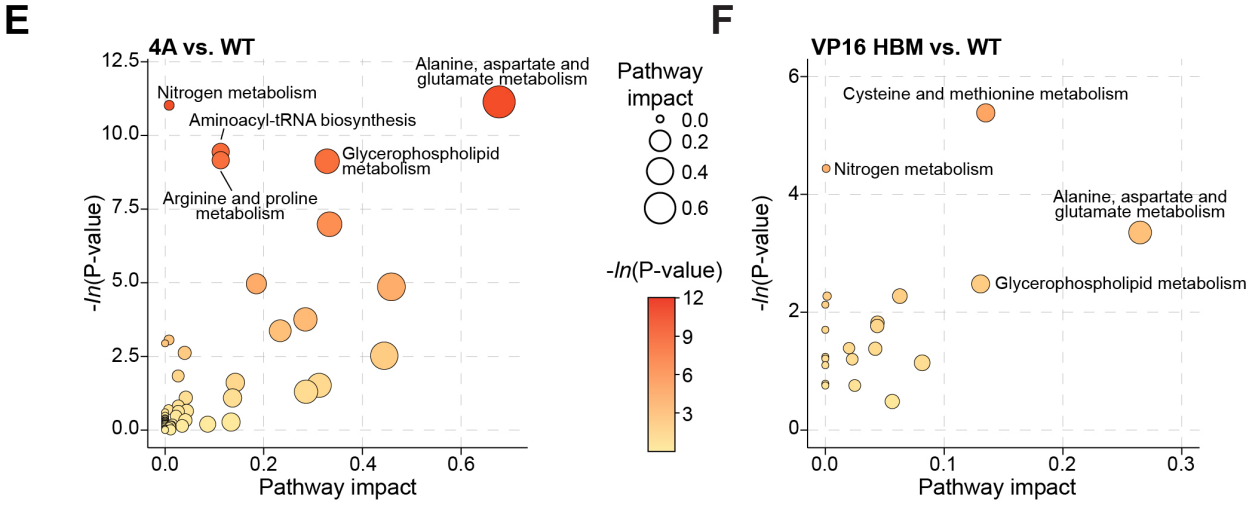


Figure 4-3: Shift in the metabolic landscape caused by MYC HBM mutants
(Previous page)
(**A**) and (**B**) Volcano plots of metabolites detected by RPLC (left) or HILIC (right) for 4A vs. WT (A) and VP16 HBM vs. WT (B) switchable Ramos cells treated for 24 hours with 20 nM 4-OHT. Metabolites that were significantly (S) changed (false discovery rate, FDR < 0.05 & |FC| > 1.5) with the 4A MYC or VP16 HBM MYC mutant compared to WT are colored. Non-significant (NS) changes are in grey. Five biological replicates for WT and VP16 HBM, and four biological replicates for 4A were used to calculate FDR and fold-changes (FC). (**C**) and (**D**) Classification of metabolites that were significantly changed (FDR < 0.05 & |FC| > 1.5) with the 4A (C) or VP16 HBM (D) mutants compared to WT cells. (**E**) and (**F**) Enrichment analysis of KEGG pathways using a combined list of significantly changed metabolites (FDR < 0.05 & |FC| > 1.5) with 4A (E) or VP16 HBM (F) from HILIC and RPLC. Pathway impact reflects centrality and the number of matched metabolites.

Metabolomics: Relationship between 4A and VP16 HBM mutants

The opposite effects that the 4A and VP16 HBM mutations have on the MYC–HCF-1 interaction and on the growth of Ramos cells suggests that additional resulting phenotypes in these cells would also be inversely related. To determine if this was the case for the changes I observed in metabolite levels, I compared metabolites that were significantly changed under both conditions. Curiously, this relationship differed substantially between HILIC and RPLC (Figure 4-4A), and may be a consequence of these separation approaches being more apt at the detection of certain molecules [341]. For RPLC, the majority of shared metabolites changed in opposite directions, whereas the majority for HILIC changed in the same direction (Figure 4-4A).

Pathway analysis of correlated and anti-correlated metabolites identified more significant pathways with anti-correlated than correlated metabolites for both techniques (Figures 4-4B and 4-4C). Correlated metabolites were enriched for glycerophospholipid metabolism (Figure 4-4B), consistent with the high numbers of lipids that were significantly changed in both mutants (Figure 4-3C and 4-3D). Conversely, anti-correlated metabolites were particularly enriched for pathways related to amino acids (Figure 4-4C), including alanine, aspartate and glutamate metabolism. The amino acids and various intermediates affected in this pathway emphasize the opposite effect that the 4A and VP16 HBM MYC mutants had on metabolism (Figure 4-4D). Amongst these was an increase in L-Glutamine for 4A and a decrease for VP16 HBM (Figure 4-4D), pointing to a possible mechanism by which glutamine dependency of Ramos cells was affected by these mutants. All amino acids apart from L-Cysteine were detected, and all detected amino acids were significantly changed with the 4A mutant (Table 4-1). Aside from L-Aspartic Acid, amino acids were increased with the 4A mutant compared to WT (Table 4-1). Amino acids that were also significantly changed with the VP16 HBM mutant were altered in an opposite direction to what I observed with the 4A mutant, including L-Aspartic Acid (Table 4-1). Thus, I have identified amino acids to be the most consistent and predominant metabolite subclass affected by modulating the MYC–HCF-1 interaction.

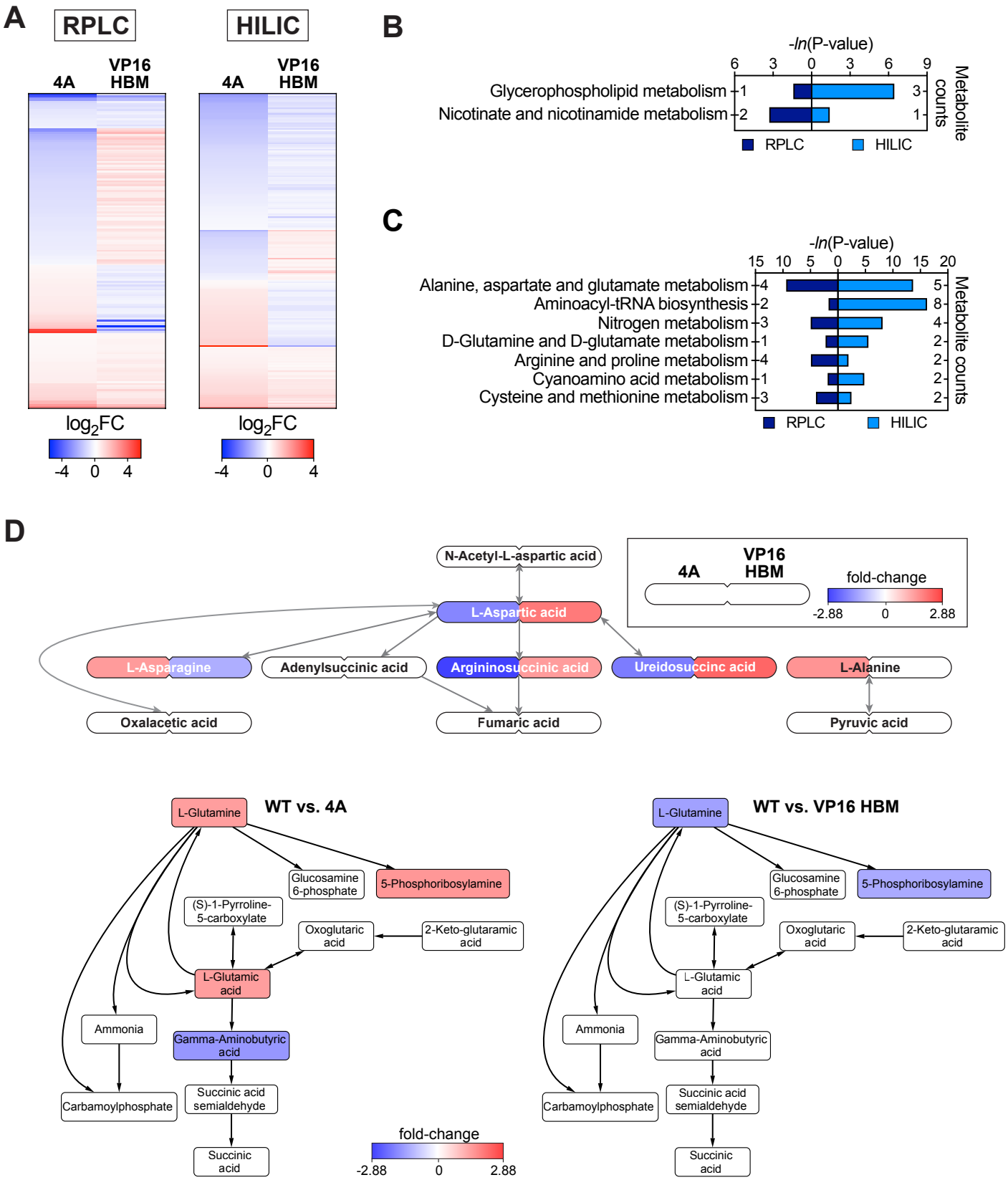

A
RPLC
HILIC
4A
VP16 HBM
log2FC
-4 0 4
B
-ln(P-value)
6 3 0 3 6 9
Glycerophospholipid metabolism
Nicotinate and nicotinamide metabolism
Metabolite counts
RPLC
HILIC
C
-ln(P-value)
15 10 5 0 5 10 15 20
Alanine, aspartate and glutamate metabolism
Aminoacyl-tRNA biosynthesis
Nitrogen metabolism
D-Glutamine and D-glutamate metabolism
Arginine and proline metabolism
Cyanoamino acid metabolism
Cysteine and methionine metabolism
Metabolite counts
RPLC
HILIC
D
N-Acetyl-L-aspartic acid
L-Aspartic acid
L-Asparagine
Adenylsuccinic acid
Argininosuccinic acid
Ureidosuccinc acid
L-Alanine
Oxalacetic acid
Fumaric acid
Pyruvic acid
4A
VP16 HBM
fold-change
-2.88 0 2.88
WT vs. 4A
L-Glutamine
Glucosamine 6-phosphate
5-Phosphoribosylamine
(S)-1-Pyrroline-5-carboxylate
Oxoglutaric acid
2-Keto-glutaramic acid
L-Glutamic acid
Ammonia
Gamma-Aminobutyric acid
Carbamoylphosphate
Succinic acid semialdehyde
Succinic acid
fold-change
-2.88 0 2.88
WT vs. VP16 HBM
L-Glutamine
Glucosamine 6-phosphate
5-Phosphoribosylamine
(S)-1-Pyrroline-5-carboxylate
Oxoglutaric acid
2-Keto-glutaramic acid
L-Glutamic acid
Ammonia
Gamma-Aminobutyric acid
Carbamoylphosphate
Succinic acid semialdehyde
Succinic acid

Figure 4-4: MYC–HCF-1 interaction influences amino acid levels
(Previous page)
(**A**) Heatmap showing metabolites detected with RPLC (left) or HILIC (right) that are significantly (FDR < 0.05) changed for both the 4A and VP16 HBM mutants, compared to WT MYC. Metabolites are clustered according to the relationship between the two mutants, and ranked by the $\log_2$FC of 4A versus WT. (**B**) Clusters of annotated metabolites from (A) that were changed in the same direction for the 4A and VP16 HBM mutants were independently subjected to pathway enrichment analysis. Pathways with P-value < 0.05 for either RPLC and HILIC are shown. (**C**) As in (B), except for of annotated metabolites from (A) that were changed in opposite directions in the 4A and VP16 HBM mutants. (**D**) Metabolites (FDR < 0.05) in the "alanine, aspartate, and glutamate metabolism" pathway that were impacted by the 4A (left) and VP16 HBM (right) MYC mutants. Node color represents the fold-change over WT.

Table 4-1: Modulating the MYC–HCF-1 interaction affects intracellular amino acid levels
All data are derived from switchable Ramos cells treated with 20 nM 4-OHT for 24 hours. Amino acid levels were measured following separations by HILIC. Q-value and fold-changes (FC) were calculated between the mutants and WT. Five biological replicates for WT and VP16 HBM and four biological replicates for 4A were analyzed. Q-value < 0.05 are highlighted in green, FC > 0 in red, and FC < 0 in blue. Confidence levels reflect the confidence in metabolite identification; L1 is validated, L2 is putative, and L3 is tentative. ND=not detected; NS=not significant.

	4A vs. WT		VP16 HBM vs. WT		Confidence level
	Q-value	FC	Q-value	FC	
Glycine	2.52E-04	1.31	3.11E-05	-1.48	L3
L-Alanine	9.97E-05	1.62	NS		L2
L-Arginine	2.31E-02	1.18	NS		L1
L-Asparagine	6.67E-05	1.50	2.02E-03	-1.23	L2
L-Aspartic acid	1.24E-03	-1.37	1.54E-05	2.01	L2
L-Cysteine	ND		ND		
L-Glutamic acid	2.53E-03	1.46	NS		L2
L-Glutamine	2.33E-05	1.46	9.38E-05	-1.41	L2
L-Histidine	8.70E-05	1.64	NS		L2
L-Isoleucine	1.10E-03	1.39	2.36E-02	-1.21	L2
L-Leucine	8.63E-04	1.33	2.55E-02	-1.18	L2
L-Lysine	1.66E-02	1.19	NS		L2
L-Methionine	6.51E-05	1.61	4.39E-03	-1.27	L2
L-Phenylalanine	4.71E-04	1.48	3.46E-02	-1.22	L2
L-Proline	1.01E-02	1.17	NS		L2
L-Serine	7.48E-05	1.59	NS		L1
L-Threonine	9.63E-05	1.54	1.25E-02	-1.17	L2
L-Tryptophan	1.56E-04	1.64	NS		L1
L-Tyrosine	1.22E-04	1.59	1.9E-02	-1.18	L2
L-Valine	4.41E-02	1.14	NS		L3

Regulation of gene expression

The majority of functions for MYC and HCF-1 are attributed to their ability to influence transcriptional output, leading to the hypothesis that their interaction has a similar contribution and that the effects on growth and metabolism I observed in the switchable *MYC* Ramos cells are consequences of altered gene expression. To understand specifically which genes were involved, I performed RNA-Seq following a 24-hour treatment with 4-OHT. For both the 4A and VP16 HBM mutants, I observed extensive alterations in transcript levels, with both cell lines showing ~4,000 significantly changed transcripts over wild-type (Figure 4-5A) that were strongly conserved across replicates (Figure 4-5B). The magnitude of these changes varied dramatically, but the vast majority were small in size for both mutants (Figure 4-5C), consistent with what is commonly found with the contribution of MYC to transcriptional regulation [77].

Gene ontology analysis of significantly changed transcripts revealed a striking relationship between the 4A and VP16 HBM mutants (Figures 4-5D and 4-5E). While the ontology of transcripts that were increased with the 4A mutant bore minimal resemblance to those that were decreased with the VP16 HBM mutant (Figures 4-5D, right and 4-5E, left), that of transcripts decreased with 4A or increased with the VP16 HBM mutant were near-identical (Figures 4-5D, left and 4-5E, right). The latter of these gene ontology analyses were enriched primarily for protein synthesis and mitochondrial function categories, including ribosome biogenesis, translation, tRNA metabolic process, and mitochondrial matrix (Figures 4-5D, left and 4-5E, right). The strength and consistency of these categories suggests that HCF-1 likely promotes transcriptional activation by MYC at a process-specific set of target genes, with this function being impaired in the 4A mutant and amplified in the VP16 HBM mutant.

I then confirmed that the gene expression changes identified by comparing the WT to mutant cell lines in RNA-Seq recapitulated when comparing unswitched cells to switched cells by reverse transcriptase quantitative PCR (RT-qPCR). By focusing on genes that fell into the most consistent gene ontology categories, I validated a number of transcripts. For all tested the 4A and VP16 HBM mutants

respectively caused decreased and increased transcript levels over their wild-type counterparts (Figure 4-5F).

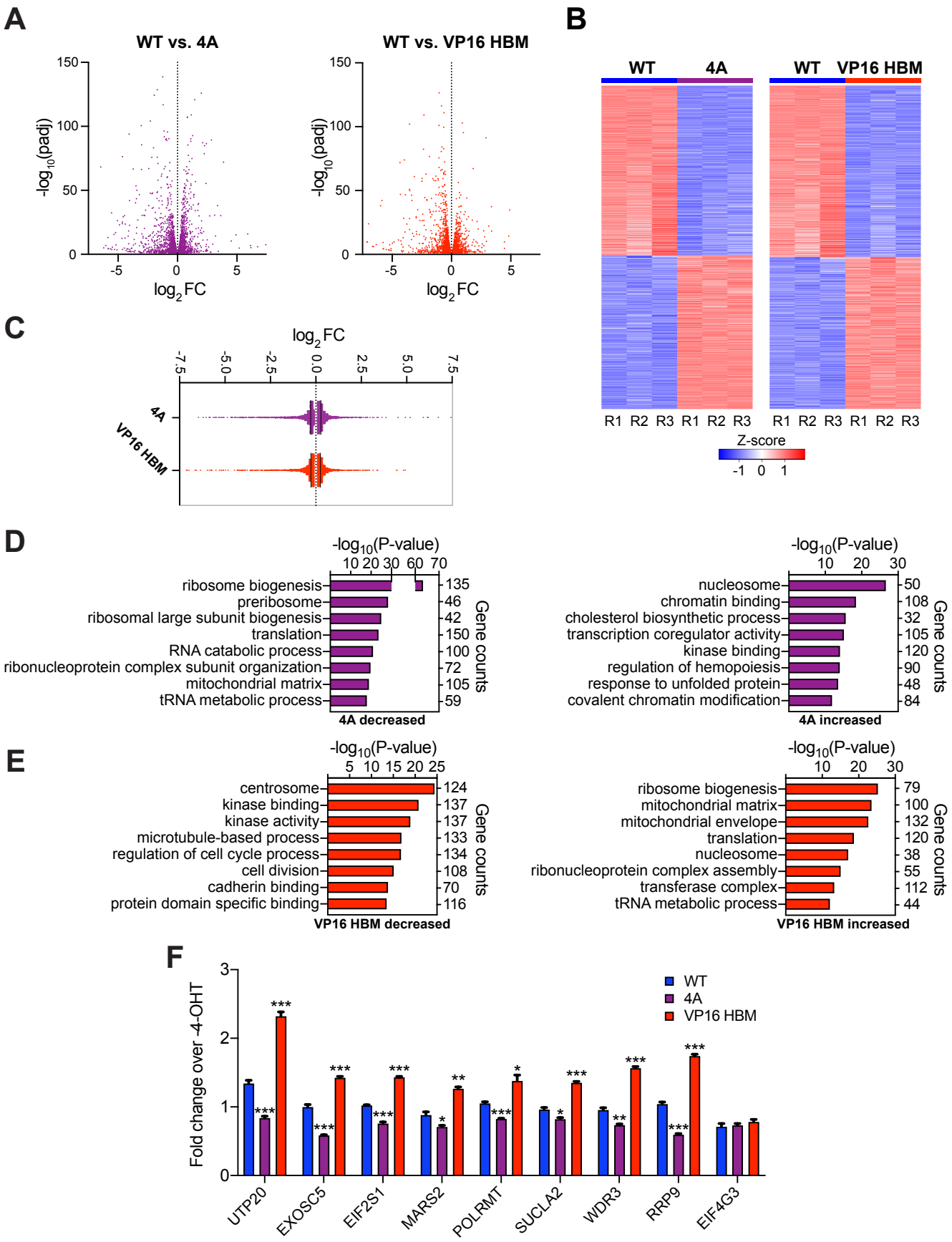

A
WT vs. 4A
WT vs. VP16 HBM
-log10(padj)
log2FC
B
WT
4A
WT
VP16 HBM
R1 R2 R3 R1 R2 R3 R1 R2 R3 R1 R2 R3
Z-score
C
log2FC
4A
VP16 HBM
D
-log10(P-value)
ribosome biogenesis
preribosome
ribosomal large subunit biogenesis
translation
RNA catabolic process
ribonucleoprotein complex subunit organization
mitochondrial matrix
tRNA metabolic process
Gene counts
4A decreased
nucleosome
chromatin binding
cholesterol biosynthetic process
transcription coregulator activity
kinase binding
regulation of hemopoiesis
response to unfolded protein
covalent chromatin modification
4A increased
E
centrosome
kinase binding
kinase activity
microtubule-based process
regulation of cell cycle process
cell division
cadherin binding
protein domain specific binding
VP16 HBM decreased
ribosome biogenesis
mitochondrial matrix
mitochondrial envelope
translation
nucleosome
ribonucleoprotein complex assembly
transferase complex
tRNA metabolic process
VP16 HBM increased
F
Fold change over -4-OHT
WT
4A
VP16 HBM
UTP20
EXOSC5
EIF2S1
MARS2
POLRMT
SUCLA2
WDR3
RRP9
EIF4G3

Figure 4-5: Gene expression changes induced by the 4A and VP16 HBM mutants
(Previous page)
(**A**) Volcano plots showing the significant (FDR < 0.05) gene expression changes for the 4A (left) and VP16 HBM (right) mutants. For clarity, some data points were excluded. (**B**) Heatmap visualizing consistency amongst replicates of RNA-Seq for switchable *MYC* allele cell lines at 24 hours, by Z-transformation and ranking by FC. Genes that were significantly (FDR < 0.05) impacted compared to WT are shown. Three biological replicates for each WT, 4A, and VP16 HBM were used to calculate FDR and fold-changes. (**C**) Scatter plot showing the distribution of log_2FC of significant (FDR < 0.05) RNA-Seq changes with the 4A and VP16 HBM MYC mutants, compared to the WT switch. Solid lines represents the median log_2FC for decreased (4A: -0.2858; VP: -0.2747) and increased (4A: 0.281; VP: 0.2558) genes compared to WT. For clarity, some data points were excluded. (**D**) and (**E**) Categories from the top eight families in GO enrichment analysis of significant (FDR < 0.05) gene expression changes under each condition (D: 4A; E: VP16 HBM). (**F**) RNA-Seq changes in the switchable cells were validated using reverse transcriptase quantitative PCR (RT-qPCR). Cells were grown with or without 4-OHT for 24 hours to confirm changes in RNA levels detected by RNA-seq, with Ct values normalized to those for GAPDH and fold-changes calculated over -4-OHT samples. Shown are the mean and standard error for three biological replicates. Student's t-test between WT and mutant cells was used to calculate P-values; *P<0.05, **P<0.01, ***P<0.001.

RNA-Seq: Relationship between 4A and VP16 HBM mutants

Although I anticipate the 4A and VP16 HBM mutants affect expression of a similar set of genes, the congruent gene ontology analyses of the mutants may also be a consequence of altered expression of different gene sets that bear the same function. By comparing transcripts whose expression was changed under both conditions, I found that almost half of the significantly changed transcripts were reciprocated between the two datasets, and of these, a slight majority had an inverse fold-change (Figure 4-6A and 4-6B). Excluding the direction, the magnitude of changes were similar between the two mutants, but the bulk of shared transcripts had only small alterations in transcript levels compared to wild-type (Figure 4-6B). Thus, both metabolites and transcripts demonstrated a mix of correlated and anti-correlated changes with the 4A and VP16 HBM mutants. Based on the loss- and gain-of-function nature of these mutants in regards to their association with HCF-1, I propose that primarily the anti-correlated changes are a consequence of specifically modulating this interaction, with the correlated changes more likely to occur due to both mutants disrupting interaction with a different MbIV cofactor.

I performed gene ontology analysis of the specific clusters shown in Figure 4-6B. As expected according to the categories identified in Figures 4-5D and 4-5E, transcripts that were decreased with the 4A mutant and increased with the VP16 HBM mutant were enriched in functions related to protein synthesis and mitochondria (Figure 4-6C, left). The remaining clusters, although showing some enrichment for previously identified categories, including kinase binding (Figure 4-6D, right), generally had substantially weaker enrichment than the ‘4A down/VP16 HBM up’ cluster. The primary exception to this was a strong over-representation of the nucleosome category, which consists of 17 histone-encoding transcripts that were increased with both the 4A and VP16 HBM mutants (Figure 4-6D, right). These findings provide further clarification that HCF-1 activates MYC-driven gene expression, and appears to specifically contribute to the expression of genes involved in ribosome biogenesis, translation, and mitochondrial function.

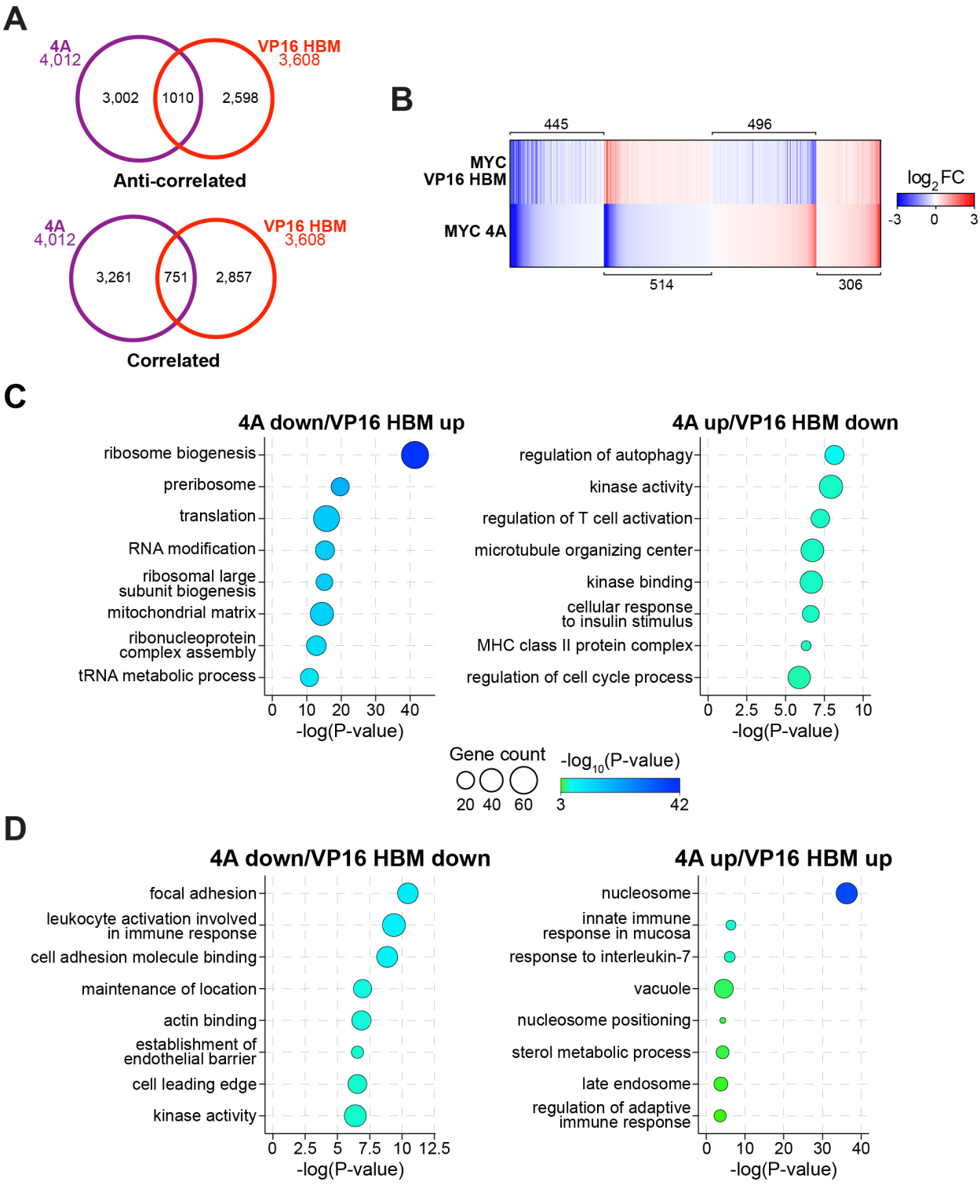


Figure 4-6: Correlative and anti-correlative gene expression changes
(**A**) Significant (FDR < 0.05) gene expression changes that are anti-correlative (left) and correlative (right) in direction between the 4A and VP16 HBM mutants. (**B**) Heatmap showing the log_2FC of significantly (FDR < 0.05) changed genes that are shared between the 4A and VP16 HBM mutants. Genes are clustered according to the relationship in expression changes between the 4A and VP16 HBM mutants, and ranked by the log_2FC for the 4A

mutant. Scale of heatmap is limited to [-3,3]. (**C**) and (**D**) Categories from the top eight families in GO enrichment analysis of the anti-correlated (C) and correlated (D) gene clusters shown in (B), for genes that were decreased (left) or increased (right) with the 4A mutant. The P-value of categories is represented by the bubble color, which is scaled across these figures, and the number of genes present in a category is represented by the bubble size.

Relationship between gene expression and metabolic changes

There are multiple mechanisms possibly responsible for the alterations in amino acid levels I observed in switchable *MYC* cell lines. By performing a combined analysis of metabolites and transcripts that were significantly changed in their respective assays (Figures 4-7A and 4-7B), I took a closer look at which gene expression changes may be responsible for the observed alterations in metabolites. As expected based on the metabolomics analyses alone, and the strong influence of the 4A and VP16 HBM mutants on amino acids, the predominant KEGG pathways that emerged involve amino acids, particularly for the 4A mutant (Figure 4-7A). Pathways enriched for the VP16 HBM mutant were more diverse, but generally much weaker, and included pyrimidine metabolism (Figure 4-7B).

Curiously, the preeminent pathway to emerge from combined analyses was aminoacyl-tRNA biosynthesis, which consists of all of the amino acids and the enzymes, commonly referred to as tRNA ligases, synthetases, or transferases, that catalyze the transfer of amino acids to uncharged tRNAs. The majority of these ligases have multiple paralogs, and these usually reflect cellular distribution (e.g. MARS1 is cytoplasmic and MARS2 is mitochondrial). Looking specifically at this pathway, there is a distinct anti-correlation between the effects of these mutants on amino acid levels and on the transcripts encoding the tRNA ligases (Figures 4-8A and 4-8B). With the 4A mutant, where the levels of almost all amino acids were increased, the tRNA ligase transcripts are largely decreased (Figure 4-8A). The opposite was true for the VP16 HBM mutant (Figure 4-8B). The relationship between amino acids and tRNA ligases was not necessarily consistent or opposite, with some amino acids changing without a detectable alteration in tRNA ligase (e.g. L-Serine for 4A), or vice versa (e.g. CARS1/2 for 4A).

While this relationship is intriguing, changes in the levels of tRNA ligases are unlikely to directly lead to alterations in amino acid levels. However, one would anticipate that impacting the levels of charged tRNAs would reduce the incorporation of amino acids into proteins, as would the observed changes in the expression of ribosome biogenesis components. In addition to these, I also noted enrichment of amino acid transporters in the RNA-Seq data, with these predominantly increased in the 4A mutant and

decreased in the VP16 HBM mutant (Figure 4-8C). Thus, the mechanism of amino acid variations when perturbing the MYC–HCF-1 interaction is likely multi-faceted, including both increased uptake and decreased utilization.

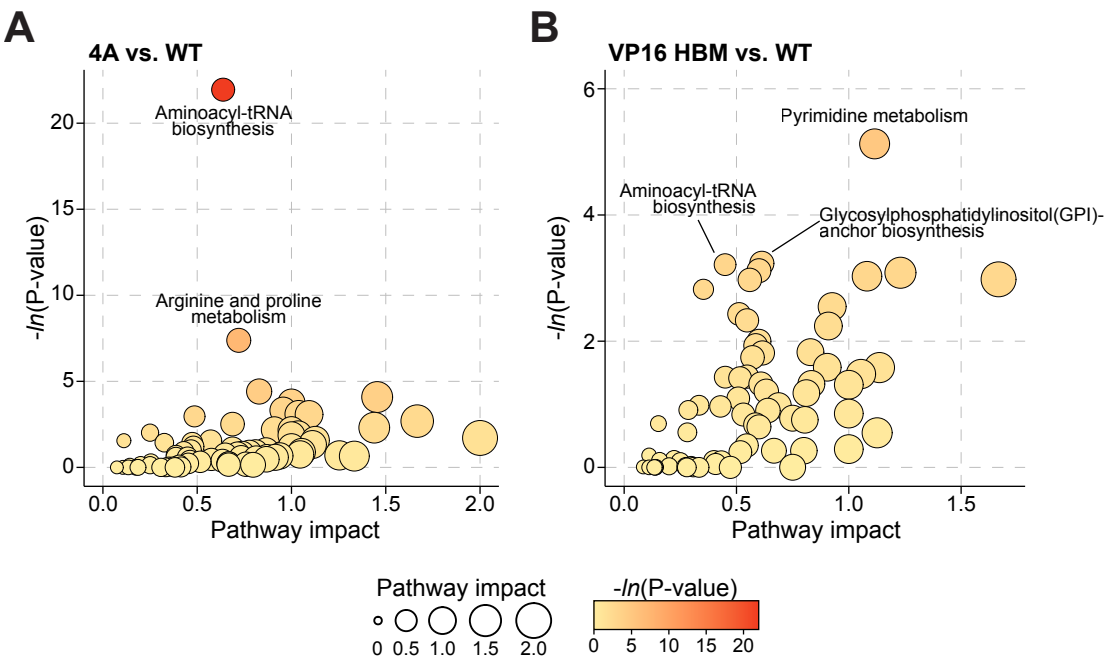


Figure 4-7: Pathways driving amino acid accumulation
(**A**) and (**B**) Enrichment analysis of KEGG pathways using a combined list of significantly changed metabolites (FDR < 0.05 & |FC| > 1.5) and transcripts (FDR < 0.05) with the 4A (A) and VP16 HBM (B) mutants. Pathway impact reflects centrality and the number of matched metabolites.

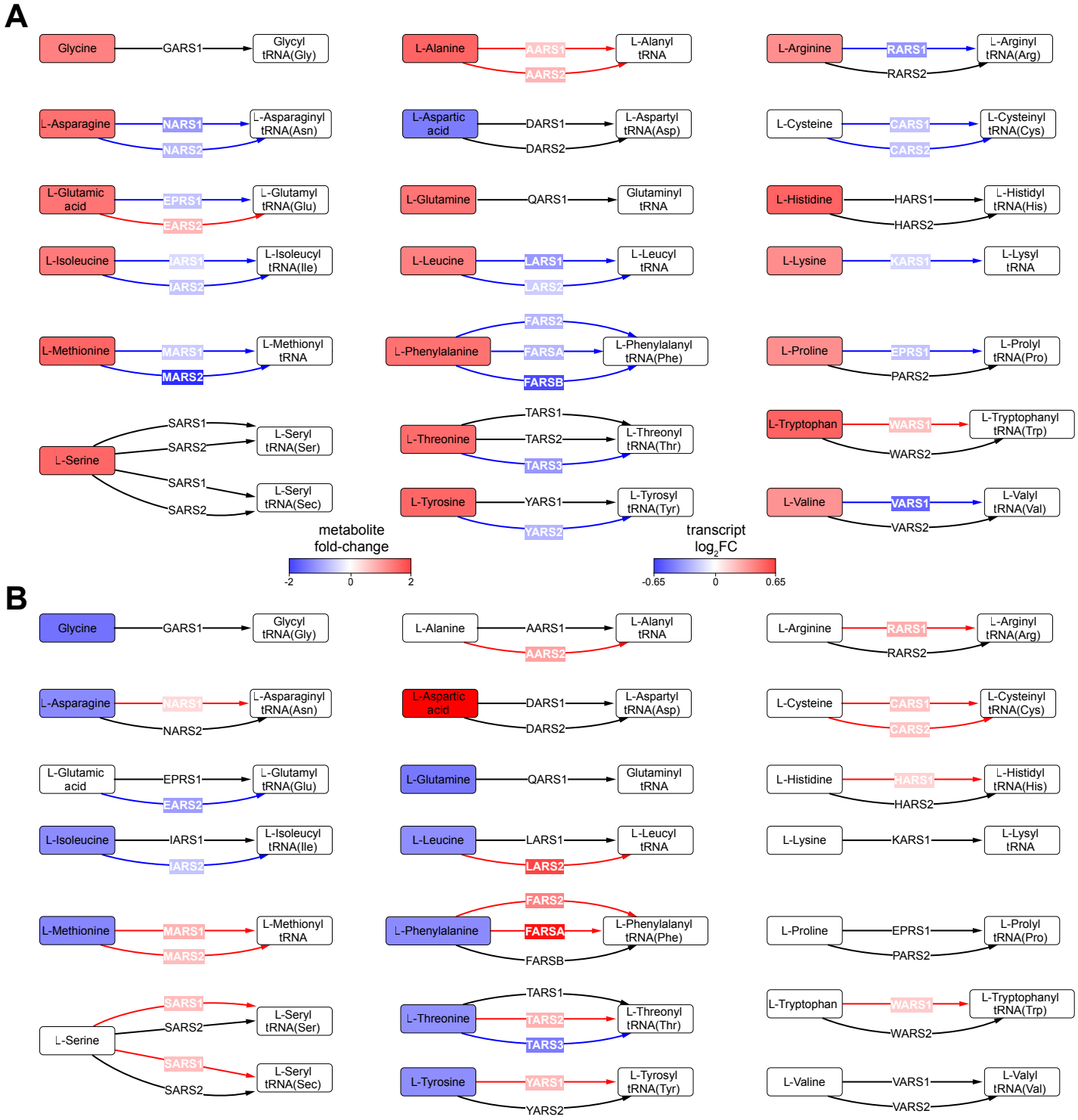

A
Glycine
GARS1
Glycyl tRNA(Gly)
L-Alanine
AARS1
AARS2
L-Alanyl tRNA
L-Arginine
RARS1
RARS2
L-Arginyl tRNA(Arg)
L-Asparagine
NARS1
NARS2
L-Asparaginyl tRNA(Asn)
L-Aspartic acid
DARS1
DARS2
L-Aspartyl tRNA(Asp)
L-Cysteine
CARS1
CARS2
L-Cysteinyl tRNA(Cys)
L-Glutamic acid
EPRS1
EARS2
L-Glutamyl tRNA(Glu)
L-Glutamine
QARS1
Glutaminyl tRNA
L-Histidine
HARS1
HARS2
L-Histidyl tRNA(His)
L-Isoleucine
IARS1
IARS2
L-Isoleucyl tRNA(Ile)
L-Leucine
LARS1
LARS2
L-Leucyl tRNA
L-Lysine
KARS1
L-Lysyl tRNA
L-Methionine
MARS1
MARS2
L-Methionyl tRNA
L-Phenylalanine
FARS2
FARSA
FARSB
L-Phenylalanyl tRNA(Phe)
L-Proline
EPRS1
PARS2
L-Prolyl tRNA(Pro)
L-Serine
SARS1
SARS2
L-Seryl tRNA(Ser)
L-Seryl tRNA(Sec)
L-Threonine
TARS1
TARS2
TARS3
L-Threonyl tRNA(Thr)
L-Tryptophan
WARS1
WARS2
L-Tryptophanyl tRNA(Trp)
L-Tyrosine
YARS1
YARS2
L-Tyrosyl tRNA(Tyr)
L-Valine
VARS1
VARS2
L-Valyl tRNA(Val)
metabolite fold-change
-2
0
2
transcript log2FC
-0.65
0
0.65
B
C
MYC 4A
MYC VP16 HBM
log2FC
-1
0
1
SLC7A11
SLC38A5
SLC7A3
SLC7A5
SLC1A4
SLC7A1
SLC3A2
SLC1A5
SLC38A1
SLC7A6
SLC43A2
SLC25A29
SLC16A10
SLC38A2
SLC43A1
SLC6A6
SLC7A8

Figure 4-8: Amino acids and their cognate tRNA-ligases and transporters
(Previous page)
Impact of the 4A **(A)** and VP16 HBM **(B)** mutants on amino acid levels (HILIC) and on the expression of aminoacyl tRNA ligases (RNA-Seq). (**C**) Transcripts encoding amino acid transporters that were significantly changed in either the 4A or VP16 HBM mutants.

Discussion

By implementing and validating a system for perturbing the MYC–HCF-1 interaction in the BL Ramos cell line, I have determined the functional contribution of MbIV, both in terms of the genes it regulates and the effected phenotypes. The most prominent, recurring gene expression changes resulting from mutation to this region are linked to ribosome biogenesis, translation, and mitochondrial function, leading to consequences in cellular metabolic profile, glutamine-dependency, and growth. Thus, I have found the MYC–HCF-1 interaction to modulate process-specific genes that are recurring targets of MYC and are thought to be important contributors to its oncogenic potency.

The 4A and VP16 HBM mutants that I know manifest as loss- and gain-of-function effects, respectively, on the association between MYC and HCF-1, also caused both correlative and anti-correlative effects on gene expression and metabolism. Regardless of these, the dominant phenotype that prevailed in cell growth was similarly a loss and a gain, suggesting that the most important and defining changes were anti-correlative. The influence of MbIV deletion on the MYC interactome has previously been interrogated [15], with this affecting the association of MYC with upwards of 100 proteins, many of which are transcriptional regulators. The most strongly conserved residues within MbIV in MYC family members correspond to the HBM (Figure 1-5A), and one would anticipate the mutants utilized in my experiments would also affect, in addition to HCF-1, at least a handful of these interactions. I hypothesize that perturbation of the MYC–HCF-1 interaction is responsible for only a subset of the metabolite and gene expression changes I observed, specifically the subset that is anti-correlated between 4A and VP16 HBM mutants. Because of the unique nature of the gain-of-function VP16 HBM, in that it is specific to HCF-1, I additionally propose that the metabolite and gene expression changes correlated between these mutants were instead the consequence of both mutants disrupting the association of MYC with additional cofactors. Because the end-goal of this work would be the development of small-molecule inhibitors against HCF-1 to prevent its association with MYC, these “off-target” effects would likely not be relevant in this context. However, concluding this with any conviction

would also require assessment of the functional contribution of the MYC–HCF-1 interaction from the perspective of HCF-1.

I have shown here that disrupting the MYC–HCF-1 interaction slows Ramos cell growth, whereas increasing the interaction appears to accelerate it. The rate of change for these mutants is different, with the growth change for the 4A mutant occurring more quickly than that for the VP16 HBM mutant. Similarly, both approaches for untargeted metabolomics found fewer metabolites were significantly changed with the VP16 HBM mutation, and those that were significantly changed, typically did so with a smaller magnitude. For both growth and metabolism, the difference between the mutants may reflect a "ceiling" reached in the growth of an already-aggressive cancer cell with the VP16 HBM mutant, driven by enzymatic or physical constraints of metabolism in Ramos cells [342]. These findings imply that, in their "wild-type" form, Ramos cells are not maximizing their growth or metabolic potential, but instead retain room to expand these processes. This further supports the idea that the HBM present in MYC is weakened, but triggers the question of why. One possible explanation is that acceleration of metabolism and growth are important only in certain contexts, or perhaps when the nutrient status of the cell allows. The converse of this is that the VP16 HBM is disadvantageous under certain conditions, or that the MYC HBM is necessary for the totality of normal MYC function. Indeed, the latter of these suggestions is supported by the above-mentioned idea that the VP16 HBM disrupts association with additional MbIV cofactors.

Of the genes that did demonstrate an inverse relationship between the 4A and VP16 HBM mutants, the tRNA ligases were particularly striking, not only because of the corresponding effects on amino acid levels, but because they are increasingly considered to be fundamental contributors to MYC-driven cell growth and tumorigenesis [343-345]. The genes encoding tRNA ligases are reported to be direct targets of MYC [344, 345], and inhibiting them, only in combination with MYC overexpression, is sufficient to cause programmed cell death [345]. While I did not observe apoptosis with decreased expression of tRNA ligases in the 4A mutant, there is likely a threshold that I did not cross with the

modest transcript changes I observed. One of the modes through which MYC is proposed to control expression of tRNA ligases is activation of the integrated stress response [344]. The acceleration of protein synthesis by MYC, compounded by MYC-driven RNAPIII synthesis of tRNAs, has been shown to cause an accumulation of uncharged tRNA [344], leading to activation of the protein kinase GCN2 and subsequent induction of the integrated stress response [346, 347]. This pathway results in reduced translation initiation of most mRNAs, and translational up-regulation of the activating transcription factor 4 (ATF4), which promotes the expression of genes encoding amino acid transporters, tRNA ligases, and metabolic enzymes [347]. However, somewhat circuitously, MYC co-binds with ATF4 at these genes to control their output [344]. If genes encoding tRNA ligases are direct targets of the MYC–HCF-1 interaction, it is possible that I have interfered with the compensatory mechanisms through which MYC normally enables balanced, accelerated cell growth. Furthermore, I would anticipate that reducing the expression of tRNA ligases causes accumulation of uncharged tRNA, thereby leading to an increased, but disabled, compensatory response that may account for the widespread gene expression changes I observed. This, compounded by reduced expression of ribosome biogenesis and translation components, may be responsible for the accumulation of amino acids in the 4A cell line.

Overexpression of MYC can increase the rate of protein synthesis two-fold [91]. Its ability to achieve this in a balanced, productive manner likely depends on the genome-wide function of MYC. The most energy-intensive process in cells [93], ribosome biogenesis is a crucial step in controlling protein synthesis, and MYC plays a critical role in regulating each phase of it [85]. The MbIIIb cofactor WDR5 specifically contributes to MYC-driven regulation of the structural components of ribosomes, whereas here I have found the MYC–HCF-1 interaction to be primarily important to the expression of genes for rDNA transcription, rRNA processing, and ribosome assembly. Indeed, this points to MYC making specific interactions to drive the expression of subsets of genes that contribute to the same processes. Why? Ribosomal protein genes regulated by MYC–WDR5 are primary, baseline targets for MYC [74, 85]. While MYC has previously been found to control the ribosome biogenesis process, this is observed with MYC overexpression [89]. Thus, rising MYC levels could enlist HCF-1 to additionally contribute to

ribosome biogenesis, meaning that their interaction is most relevant only under certain contexts. This suggests that MYC is responsible for orchestrating all stages of ribosome biogenesis when it forces a cell to surpass its physiological protein synthesis and growth rates.

That MYC is both a cause and a cure, such as that occurring in the integrated stress response, is not a novel idea. Increasing MYC levels not only extends its control into distinct ribosome biogenesis genes [85], its regulation of metabolism and mitochondrial biogenesis is also generally considered to be a gained function of overexpressed MYC [25, 74]. Mitochondria and ribosome biogenesis correspond to the biggest sources and sinks of energy in the cell, respectively [93], suggesting increasing MYC levels enables for a coordinated increase in both processes and an intrinsic compensatory response. My works defines the MYC–HCF-1 interaction as a central contributor to this coordination. Thus, it is possible that, when MYC is expressed at a certain level, its association with HCF-1 is responsible for controlling ribosome biogenesis genes, but, upon increased overexpression, their target genes are expanded to include genes for mitochondrial function. In this sense, while overexpression of MYC causes amplification of ribosome biogenesis, it is simultaneously exploiting HCF-1 to increase the expression of genes for mitochondrial biogenesis to keep up with the corresponding energy demand. This would suggest that the MYC–HCF-1 interaction is relevant and pivotal primarily when MYC levels in the cell are high, which occurs both during development and cancer.

Chapter V

Inducible degradation of HCF-1

Introduction

First identified for its role in viral infection [348], HCF-1 has since been found to function as a cofactor for various chromatin-bound proteins, and contributes to the regulation of genes involved in cell cycling and metabolism [291, 349]. Through the activity of OGT [328], HCF-1 is cleaved into its amino- and carboxy-termini, with the termini remaining tightly associated following cleavage [225]. HCF-1 is commonly referred to as a scaffolding protein due to its ability to to interact with a variety of different proteins [217]. The majority of these interactions occur through the amino-terminus, particularly with the highly conserved VIC domain. Much of the function of HCF-1 has been defined based on its interaction partners or through long-term knockdown/knockout experiments. These approaches have found HCF-1_N contributes to the transition from the G_1 to S phase [350], stabilizes PGC-1α to regulate metabolic and mitochondrial genes [275], undergoes co-dependent recruitment with ZNF143 and THAP11 to modulate E2F-responsive genes [349], enables chromatin-association of insulin receptor for transcription of insulin-responsive genes [351], and various other functions. However, in the absence of a comprehensive assessment of HCF-1 target genes, its function on a shorter timescale and genome-wide level, beyond its previously interrogated interaction partners, is unknown.

Here, I have proposed a role for HCF-1 in contributing to the activation of protein synthesis and mitochondrial genes through its interaction with MYC. Indeed, neither of these would constitute entirely novel functions for HCF-1, with conditional knockout of HCF-1 found to affect transcripts of genes involved in mitochondrial ribosome biogenesis and aminoacyl tRNA transferases [291]. The authors of this work implied that stabilization of PGC-1α by HCF-1 is, at least in part, responsible for these expression changes [291], consistent with what has previously been demonstrated with knockdown of

PGC-1-related coactivator [352]. However, a role for HCF-1 in the regulation of ribosome biogenesis outside of the mitochondria, which I suggest based on the loss- and gain-of-function mutants in the MYC HBM, would be a novel finding.

As mentioned in the previous chapter, I believe the gene expression changes found by perturbing the MYC HBM may be a consequence of affecting MYC's interaction with both HCF-1 and additional, unconfirmed cofactors. The 4A and VP16 HBM mutations within the MYC HBM constitute a unique set of genetic tools that, I believe, are unlikely to also be loss- and gain-of function mutations for interaction with non-HCF-1, MbIV binders. Based on this, I concluded that the transcripts that were anti-correlated between the 4A and VP16 HBM mutants are due to specifically perturbing the MYC–HCF-1 interaction. To confirm that HCF-1 does regulate these genes, I applied inducible degradation of HCF-1_N using the dTAG system. This system involves fusing a mutant form of the FKBP12 protein (FKBP12^{F36V}) to a protein of interest, in this case the amino-terminus of HCF-1. A PROTAC molecule, dTAG, can then be added to cells and one end of this specifically binds to the mutant form of FKBP12 [299]. The other end of dTAG molecule binds to the E3 ligase cereblon, leading to polyubiquitylation and degradation of the protein of interest. Thus, having the ability to directly target a specific protein, the dTAG system would allow for rapid and highly specific loss of HCF-1_N. This approach is particularly important not only because it enables us to focus on the function of the amino-terminus alone, but it avoids the complications of long-term knockdown, particularly secondary transcriptional changes in response to the loss of an essential protein.

Ultimately, the undruggable, intrinsically disordered nature of MYC means that our best path forward for perturbing the MYC–HCF-1 interaction is through HCF-1 itself. While removing the entirety of HCF-1_N from a cell does not immediately replicate what would occur in the presence of a small-molecule inhibitor, by combining this information with that obtained through mutation of the MYC HBM, I hope to gain a representation of the consequences of impairing a fraction of HCF-1 function.

Results

Inducible degradation of HCF-1_N

To interrogate the role of HCF-1 in phenotypes observed when I alter the MYC HBM in Ramos cells, I also generated a cell line for inducible degradation of HCF-1_N. I used CRISPR/Cas9-directed homologous recombination to fuse $FKBP12^{F36V}$ to the amino-terminus of HCF-1 in Ramos cells (Figure 5-1A), which contain only a single allele of the X-linked *HCFC1* gene. I additionally included a P2A-linked mCherry marker, and used this for fluorescence-activated cell sorting to obtain a population of mCherry-positive cells. Compared to the parental Ramos cells, this population had a shifted molecular weight of HCF-1_N, and a FLAG-tagged species, both of which were undetectable by western blot after a 24 hour treatment with dTAG-47 (Figure 5-1B). Degradation of HCF-1_N using this system occurred rapidly, with about half of the protein gone within one hour and little-to-none detectable within eight hours (Figure 5-1C). This approach also enabled specific targeting of HCF-1_N, with only a minimal effect on the levels of HCF-1_C (Figure 5-1C). Longer treatments with dTAG-47 also caused a reduction in MYC protein levels (Figure 5-1C). Combined with the switchable *MYC* allele, inducible degradation of HCF-1_N serves as a strong approach for interrogating the function of the MYC–HCF-1 interaction from the perspective of both proteins.

HCF-1 is an essential protein, with multiple reports demonstrating its knockdown or knockout leads to slowed growth or cell death [221, 350]. I recapitulated and extended these findings here, demonstrating that the specific loss of HCF-1_N caused a rapid reduction in proliferative capacity and subsequent cell death (Figure 5-1D). Although cell numbers appeared to be mostly unaffected at 24 hours (Figure 5-1D), cell cycle analysis at this timepoint revealed a modest accumulation of sub-G_1 cells, and a reduction of cells in the S-phase (Figure 5-1E). Thus, this system enables rapid degradation of HCF-1_N and can be used to assess early phenotypes associated with its loss.

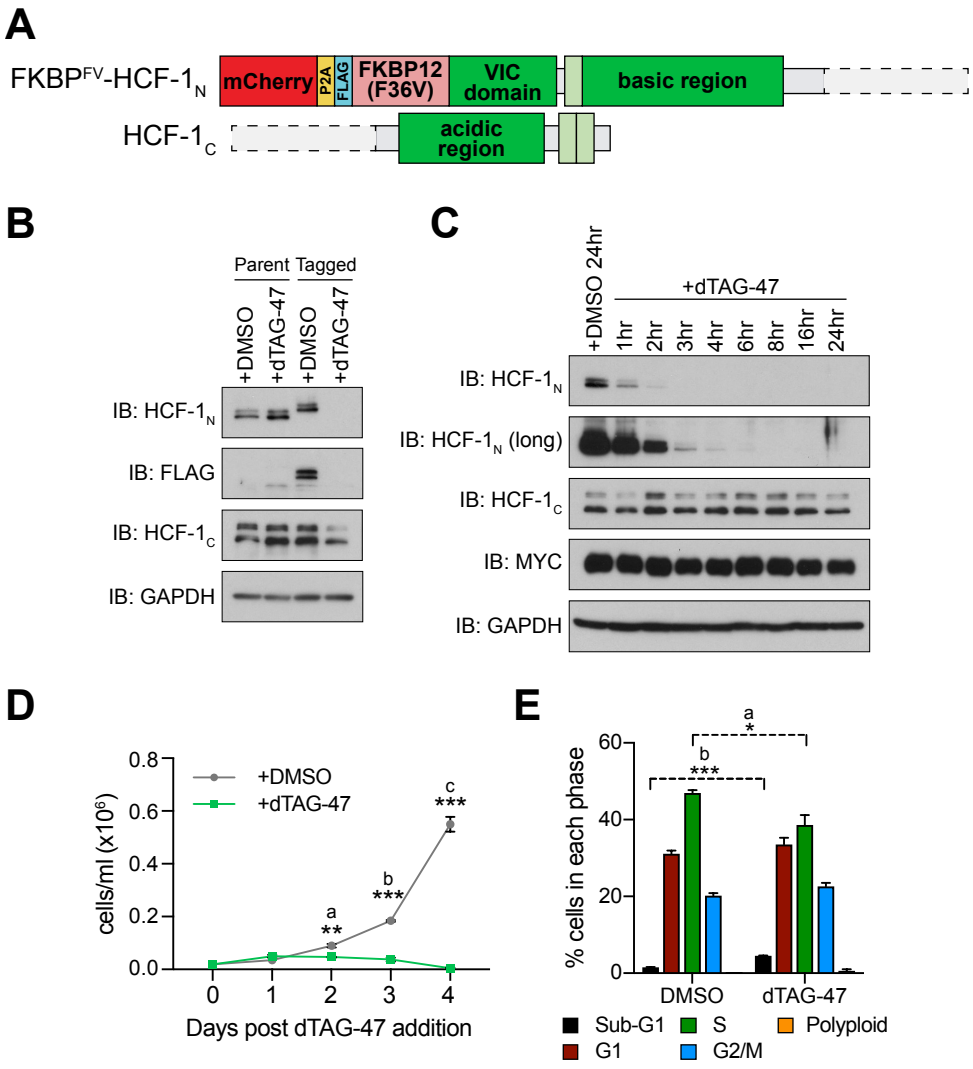


Figure 5-1: Applying the dTAG system to HCF-1$_N$
(**A**) Schematic of the HCF-1 fusion protein in Ramos FKBP^{FV}-HCF-1$_N$ Ramos cells, which were generated using CRISPR/Cas9 homology-directed repair. Fused to the N-terminus of the VP16 induced complex (VIC) domain is mCherry linked by P2A to FLAG-tagged FKBP12^{FV}. (**B**) Western blot, comparing effect of treating untagged, parental cells or FKBP^{FV}-HCF-1$_N$ Ramos cells with DMSO or 500 nM dTAG-47 for 24 hours. Blots for HCF-1$_N$, FLAG tag, HCF-1$_C$, and GAPDH are shown. (**C**) Western blot of lysates from FKBP^{FV}-HCF-1$_N$ Ramos cells treated with 500 nM dTAG-47 for varying times, compared to cells treated with DMSO for 24 hours. Shown are short and long exposures of HCF-1$_N$, and HCF-1$_C$, with a GAPDH loading control. (**D**) Growth curve of FKBP^{FV}-HCF-1$_N$ Ramos cells treated with DMSO or 500 nM dTAG-47. Cells were counted every 24 hours for four days after plating. Shown are the mean and standard error for three biological replicates. Student's t-test was used to calculate P-values; a = 0.0029, b = 0.000051, c = 0.000040. (**E**) Cell cycle distribution was determined using propidium iodide (PI) staining of FKBP^{FV}-HCF-1$_N$ Ramos cells treated with DMSO or 500 nM dTAG-47 for 24 hours. Shown are the mean and standard error for three biological replicates. Student's t-test was used to calculate P-values; a = 0.037, b < 0.0001.

Impact of HCF-1 degradation on gene expression

Application of the dTAG system for inducible degradation of HCF-1_N provided an opportunity for a temporal understanding of transcriptional regulation by HCF-1. To exploit the rapid degradation that occurs and minimize secondary consequences of HCF-1_N loss, I performed RNA-Seq on cells that had been treated with dTAG-47 for 3 hours. Even at this early timepoint I observed extensive changes in gene expression, with ~2,100 and ~2,400 transcripts significantly decreased and increased, respectively (Figure 5-2A). Similar to the switchable *MYC* cells, the magnitude of these changes varied, but the vast majority were small (Figure 5-2B). Although more transcripts were increased than decreased, the median fold-change was greater for transcripts that were decreased (Figure 5-2B). While it has been reported that dTAG molecule does not cause off-target effects [299], I treated the parental Ramos cells with dTAG molecule for 3 hours, finding that ~100 transcripts were significantly affected (Figure 5-2C). These transcripts were predominantly increased (Figure 5-2C), and a surprisingly high proportion involved in cholesterol biosynthesis. The bulk of these transcripts were also changed when FKBP-HCF-1_N cells were treated with dTAG molecule, both in magnitude and direction (Figure 5-2D).

The transcripts that were significantly decreased by the loss of HCF-1_N were enriched for functions including ribosome biogenesis, transferase complex, and tRNA metabolic process, while those that were increased are involved in cell cycle, RNA catabolic process, and translation (Figure 5-2E). Thus, categories of both decreased and increased transcripts resemble those that were repeatedly represented in the switchable *MYC* cells. Targeting genes in these categories, I additionally validated a selection of changes by RT-qPCR (Figure 5-2F). Finally, gene set enrichment analysis (GSEA) re-emphasized the role of HCF-1 in the cell cycle, with the hallmark gene sets G_2/M checkpoint and E2F targets both significantly enriched (Figure 5-2G).

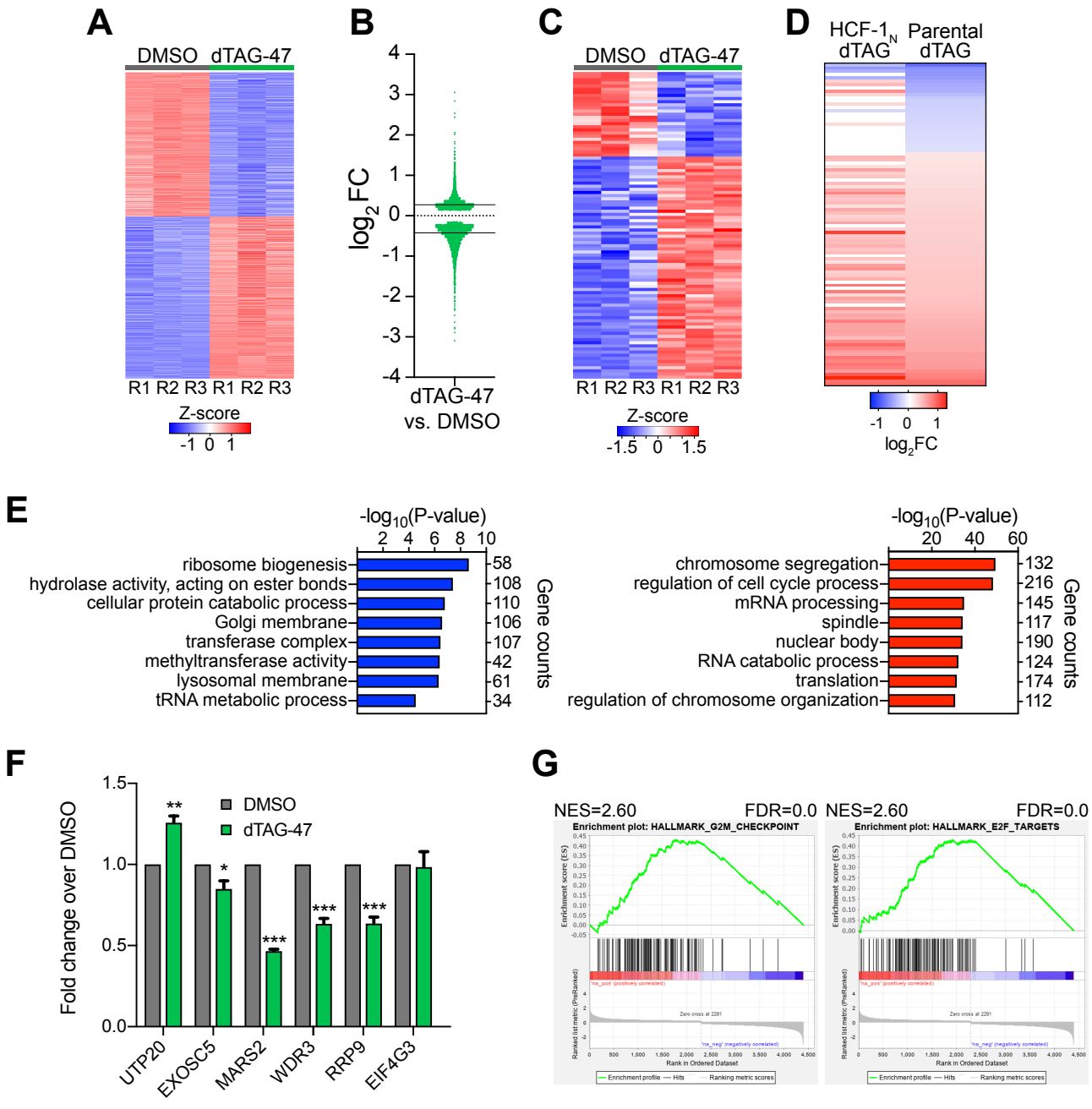


Figure 5-2: Gene expression changes in response to HCF-1$_N$ degradation
(**A**) Heatmap of significantly (FDR < 0.05) changed genes in FKBP^{FV}-HCF-1$_N$ Ramos cells treated with DMSO or dTAG-47 for three hours. Shown are the z-transformed expression data for three biological replicates of RNA-Seq, ranked by FC. (**B**) Scatter plot showing the distribution of log$_2$FC in RNA-Seq comparing DMSO to three hours of 500 nM dTAG-47 treatment. Solid lines represent the median log$_2$FC for decreased (-0.425655) and increased (0.270428) genes. For clarity, one data point was excluded. (**C**) Heatmap of significantly (FDR < 0.05) changed transcripts in parental Ramos cells treated with DMSO or 500 nM dTAG-47 for three hours. Shown are the z-transformed expression data for three biological replicates of RNA-Seq, ranked by FC. (**D**) Heatmap showing the log$_2$FC of all transcripts that were significantly (FDR < 0.05) changed when untagged Ramos cells were treated with DMSO or dTAG-47 for three hours, and the corresponding log$_2$FC data in FKBP^{FV}-HCF-1$_N$ Ramos cells. (**E**) GO enrichment analysis of genes significantly (FDR < 0.05) decreased (left) and increased

(right) in expression following treatment of $FKBP^{FV}$-$HCF\text{-}1_N$ Ramos cells with dTAG-47 for three hours. Excluded from this analysis are genes that were significantly changed when parental Ramos cells were treated with dTAG-47 for three hours. (**F**) RNA-Seq changes in the $FKBP^{FV}$-$HCF\text{-}1_N$ Ramos cells were validated using RT-qPCR. Cells were grown with DMSO or dTAG-47 for three hours, with Ct values normalized to those for GAPDH and fold-changes calculated over DMSO-treated samples. Shown are the mean and standard error for three biological replicates. Student's t-test between DMSO and dTAG-47-treated cells was used to calculate P-values; *P<0.05, **P<0.01, ***P<0.001. (**G**) GSEA of significantly (FDR < 0.05) changed transcripts in $FKBP^{FV}$-$HCF\text{-}1_N$ Ramos cells. Shown are the NES, and FDR for each gene set.

Gene expression changes shared between with MYC HBM mutants and HCF-1 degradation

The goal of inducible degradation of HCF-1_N was to gain a better understanding of how HCF-1 influences the expression of MYC target genes, and to identify which transcripts affected by both the MYC 4A and VP16 HBM mutants were the result of specifically perturbing the MYC–HCF-1 interaction. I first compared the transcripts that were significantly changed with the HCF-1_N degradation, and with the 4A and VP16 HBM mutants (Figure 5-3A). More than half of the transcripts altered with HCF-1_N degradation were also impacted by either the 4A or VP16 HBM mutants, with ~700 transcripts affected under all three conditions (Figure 5-3A). While all of these overlapping transcripts, and more, may be regulated by the MYC–HCF-1 interaction, I was particularly interested in those that were anti-correlated between 4A and VP16 HBM, and also changed with HCF-1_N degradation. This strict criteria amounts to more than 450 transcripts, and for the majority of these, the direction of change is conserved between the 4A mutant and HCF-1_N degradation (Figure 5-3B). The relationship between the 4A mutant and HCF-1_N degradation is somewhat stronger for transcripts that were increased under both conditions, but the median $\log_2$FC hovers at a very similar value for transcripts that were both decreased (median: -0.2276) and increased (median: 0.2349) with HCF-1_N degradation (Figure 5-3C).

Finally, ontology analysis of the gene expression changes in the 4A and VP16 HBM mutants had highlighted a role for this interaction in specifically regulating genes involved in protein synthesis and mitochondrial function, including the categories “ribosome biogenesis” and “mitochondrial matrix”. Focusing on transcripts that were decreased with the 4A mutant and increased with the VP16 HBM mutant, I determined that many of these were also affected by HCF-1_N degradation (Figure 5-3D and Figure 5-3E). Amongst those affected in “mitochondrial matrix” are a number additionally involved in mitochondrial ribosome biogenesis and protein synthesis, including structural components of mitochondrial ribosomes (Figure 5-3E).

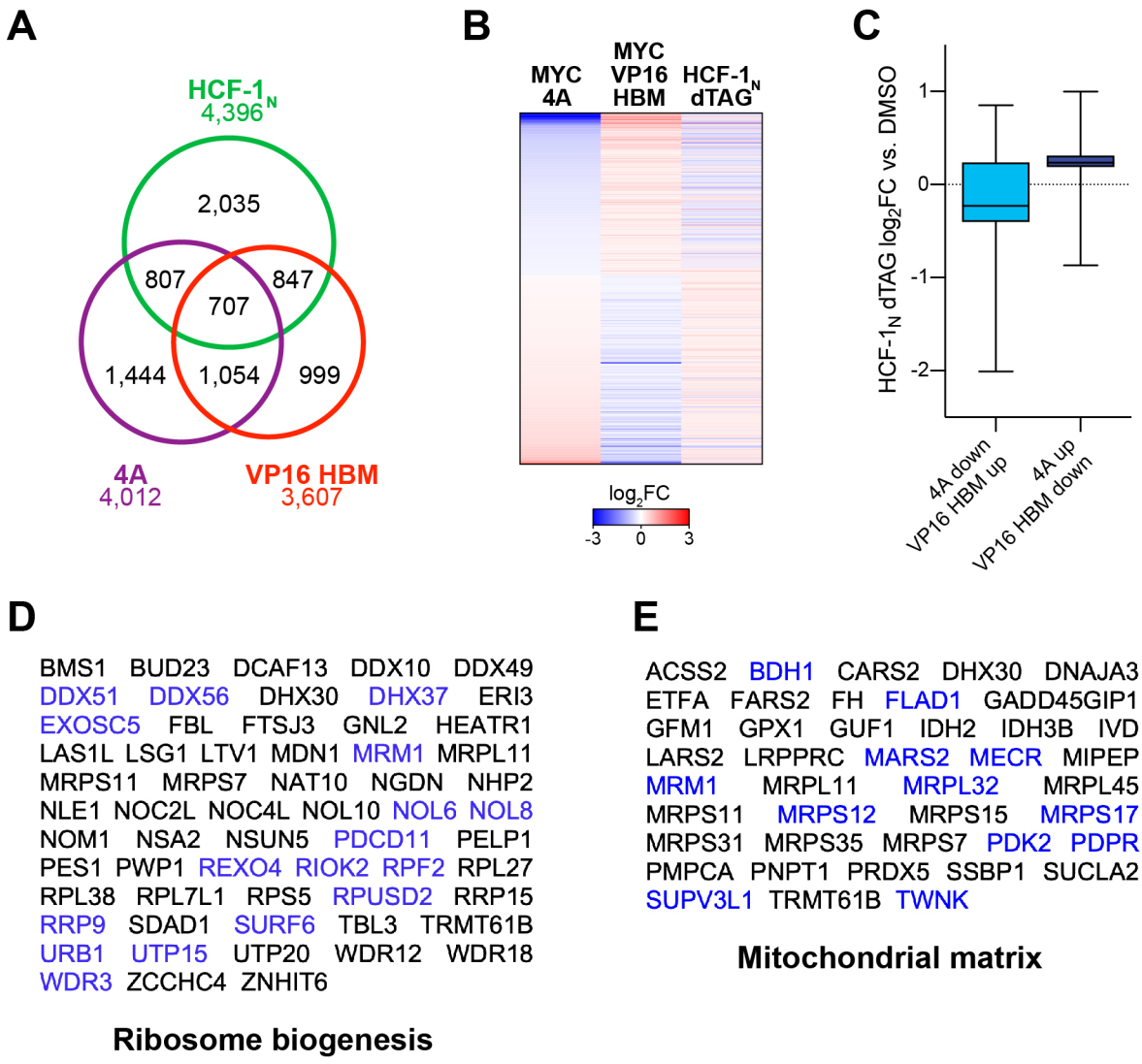


Figure 5-3: Comparison of RNA-Seq between switchable *MYC* allele and HCF-1$_N$ degradation
(**A**) Venn diagram showing overlap of gene transcripts that were significantly (FDR < 0.05) changed in RNA-Seq with the 4A and VP16 HBM mutation, and with degradation of HCF-1$_N$ in Ramos cells. (**B**) Heatmap showing the log$_2$FC of genes with significantly (FDR < 0.05) changed expression, as measured by RNA-Seq, under all conditions (switchable *MYC* alleles, 4A and VP16 HBM, and FKBP^{FV}-HCF-1$_N$ Ramos cells). Genes are clustered according to the relationship in expression changes between the 4A and VP16 HBM mutants, and ranked by the log$_2$FC for the 4A mutant. Scale of heatmap is limited to [-3,3]. (**C**) Box-and-whisker plot showing the relationship between genes that are anti-correlated between the 4A and VP16 HBM MYC mutants, and significantly changed with the degradation of HCF-1$_N$. Box denotes the 25th to 75th percentile, whiskers extend from minimum to maximum point, and middle line marks the median (4A down/VP16 HBM up: -0.2276; 4A up/VP16 HBM down: 0.2349). (**D**) and (**E**) Transcripts that were decreased with the 4A MYC mutant and increased with the VP16 HBM MYC mutant and fall into the “Ribosome biogenesis” (D) or “Mitochondrial matrix” (E) gene ontology categories. Highlighted in blue are the genes that are also significantly decreased with degradation of HCF-1$_N$.

Discussion

In the absence of a definitive set of known target genes and functions of HCF-1 in a cell, we cannot distinguish the specific purpose of each of its interactions. Interrogating this aspect of HCF-1 is a necessary step towards understanding how and why HCF-1 is so important to cell function, and what the outcome would be if only a subset of its interactions were perturbed. In the context of my work, I am particularly interested in the contribution of HCF-1 to the function of MYC, which I propose to be the transcription of genes involved in ribosome biogenesis, translation, and mitochondrial function.

The termini of HCF-1 are thought to have distinct contributions to cell cycle progression, specifically that the amino-terminus is necessary for the G_1 to S phase transition and the carboxy-terminus for cytokinesis [350]. Inducible degradation of HCF-1_N recapitulated this reported phenotype, with the proportion of cells in the S phase significantly decreased. While G_0/G_1 arrest in the absence of HCF-1 can be rescued by add-back of HCF-1_N [350], I did not explore this phenotype in the context of Ramos cells due to the accompanying accumulation of cells in sub-G_1, suggesting irreversible cell death in response to the loss of HCF-1_N occurred in this context. One possible reason for this difference is that MYC protein levels are impacted by HCF-1 loss within 16 hours, and this likely has severe downstream consequences on cell growth and survival. Based on my finding that MYC and HCF-1 levels are largely unaffected by mutation of the MYC HBM (Figure 4-1E), the effect of HCF-1 degradation on MYC protein levels is probably not a consequence of the MYC–HCF-1 interaction regulating MYC stability. Instead, HCF-1 may contribute to the transcription of the *MYC* allele, but whether this occurs specifically at the translocated allele, or also at non-translocated *MYC* in other contexts, would need to be investigated. Cell cycle phenotypes observed with knockdown of MYC are cell type-dependent [353], meaning that HCF-1-induced loss of MYC may also occur in other contexts with a different outcome. Regardless, if blocking HCF-1 reduces both *MYC* expression and MYC function in multiple contexts (i.e. with and without translocation), the benefit of having a small-molecule inhibitor against it is essentially doubled.

The distinct nature of the switchable *MYC* allele and the dTAG systems makes it near-impossible to identify a timepoint where any effects on transcription or downstream cellular phenotypes are equivalent. For both systems, the timepoint used was intended to maximize switching/degradation and minimize non-transcriptional phenotypes, but likely meant a subset of target genes were missed. Despite this, hundreds of genes were affected by both the loss- and gain-of-function MYC mutants, and with HCF-1_N degradation, suggesting that these, or at least a fraction of these, are directly regulated by the interaction of MYC and HCF-1. The relationship between these changes varied, but there was an overt correlation in the direction of changes that occurred with the 4A mutant and HCF-1_N degradation. Previously I suggested that, based on the gene ontology analysis of transcripts that were decreased with 4A and increased with VP16 HBM, HCF-1 likely activates MYC-driven transcription. However, their combined function does not necessarily appear to be transcriptional activation, with more than half of the transcripts shown in Figure 5-3B being increased with both the 4A mutant and HCF-1_N degradation. HCF-1 is found in complexes involved in transcriptional repression and activation [245], and I hypothesize that either of these complex classes may be recruited to MYC target genes through association with HCF-1.

In addition to recapitulating known HCF-1 transcriptional targets using the dTAG system, I have defined ribosome biogenesis, distinct from that occurring in the mitochondria, as a novel gene set targeted by HCF-1_N. In addition to ribosome biogenesis, previously identified mitochondrial function target genes also link MYC and HCF-1, suggesting multiple factors are responsible for conveying HCF-1 regulation of these genes. However, MYC is predominantly recognized for its baseline function in ribosome biogenesis [85], while regulation of mitochondrial biogenesis and metabolism are functions gained with overexpression [25, 354]. Thus, it is possible that MYC–HCF-1 modulation of these latter functions is primarily a gained consequence of MYC overexpression, with the former possibly occurring regardless of MYC levels in the cell.

My work emphasizes the absolute importance of HCF-1, not just to cell growth, but also to cell viability. This is in contrast to the MYC–HCF-1 interaction itself, which appears to be relevant only to rapid proliferation of Ramos cells, suggesting that it is the other interactions made by HCF-1_N, or the synergistic necessity of these, that make HCF-1_N essential. Based on this, I can generally conclude that, with the goal of inhibiting MYC function through HCF-1, a way to specifically target this interaction would be necessary to maintain the majority of HCF-1 function. In the absence of this specificity, we may lose any therapeutic opportunity that the MYC–HCF-1 interaction would provide and cause unwanted, off-target cell death. As it stands, we do not currently know how the binding strength of MYC compares to different HCF-1_{VIC} interaction partners. However, it stands to reason that, if overexpression of MYC is necessary for functional relevance of this interaction and for its complete breadth of target genes (i.e. ribosome biogenesis and mitochondrial function), blocking just a fraction of MYC–HCF-1 interactions using a relatively weak inhibitor may be sufficient to return the HCF-1-dependent function of MYC to baseline levels without imparting deficits on off-target interactions and cells.

Chapter VI

Distribution of MYC and HCF-1 on chromatin

Introduction

Up to this point, I have identified a set of genes regulated by the MYC–HCF-1 interaction (Figure 5-3B), and while we know the MYC–HCF-1 interaction occurs in the nucleus [297], I have yet to assess the relationship of MYC and HCF-1 on chromatin. The vast majority of MYC function is linked to its role as a transcription factor, and I anticipate that the mechanism of HCF-1's contribution to MYC is similarly embedded at chromatin. Furthermore, the co-regulation by MYC and HCF-1 of genes involved in ribosome biogenesis, translation, and mitochondrial function may reflect both direct targets, and indirect, secondary transcriptional consequences. Determining the co-binding of MYC and HCF-1 on chromatin will enable me to differentiate between these direct and indirect targets, which has two layers of significance. First, I might better predict the direct consequences of perturbing the association of MYC and HCF-1 in a cell-type independent manner. Second, I would have a subset of genes whereby I am highly confident MYC and HCF-1 contribute to transcriptional output, and at which I could use to determine a specific mechanism of regulation by this interaction.

By comparing ENCODE data, Dingar *et al.* [296] have previously suggested that ~80% of MYC binding sites in K562 cells are co-bound HCF-1_C. However, the termini of HCF-1 are thought to have somewhat distinct functions within cells [350], and Michaud *et al.* [235] previously demonstrated that fewer than 60% of HCF-1_C peaks overlap with HCF-1_N. The chromatin association of HCF-1 is driven by its amino-terminus, yet much of research into its distribution on chromatin has employed antibodies against the carboxy-terminus [254, 349]. Regardless, a number of factors that interact with HCF-1_N have been shown to contribute to the recruitment of HCF-1_C to chromatin, including E2F1 [254] and THAP11/ ZNF143 [349], albeit not on a genome-wide level. Unlike HCF-1, MYC is competent for DNA binding in

the presence of its obligate binding partner MAX. While a number of cofactors, including WDR5 [21], BPTF [355], and Leo1 [172], are involved in facilitated recruitment of MYC to chromatin, this may not be a necessary step at all target genes. The genetic tools I have created to interrogate the MYC–HCF-1 interaction, including placing mutant MYC as the only MYC in the cell, and with both loss- and gain-of-function mutants, enable me to determine the specific contribution of their association to the genome-wide recruitment of either MYC or HCF-1 to chromatin. Finally, this understanding may provide an initial insight into the mode by which the MYC–HCF-1 interaction has such widespread and disruptive effects on Ramos cells.

Results

HCF-1 in Ramos cells

To determine the distribution of the amino-terminus of HCF-1 on chromatin, I performed ChIP-Seq using an polyclonal HCF-1_N-specific antibody in Ramos cells. This antibody was a gift from Yuichi Machida at Mayo Clinic, and I had validated it for HCF-1_N specificity using the dTAG system (Figure 5-1B). In these cells, I identified ~2,000 peaks, and the distribution of these was strongly promoter-proximal (Figure 6-1A). More than three-quarters of these peaks were between -1000bp and +500bp, with the majority upstream of the TSS (Figure 6-1A). Gene ontology analysis of all peaks emphasized previously known functions of HCF-1, such as mitochondrial function and cell cycle, and less-well established binding sites, including transferase complex and RNA splicing (Figure 6-1B). I also looked at known and *de novo* motif analyses of these peaks to investigate whether this was a determinant for HCF-1_N binding. Both approaches showed a strong over-representation of the NRF1 motif (Figure 6-1C). This corresponds to a sequence associated with nuclear respiratory factor (NRF) 1, but it is also considered a non-canonical E-box and can be bound by MYC [18, 356, 357].

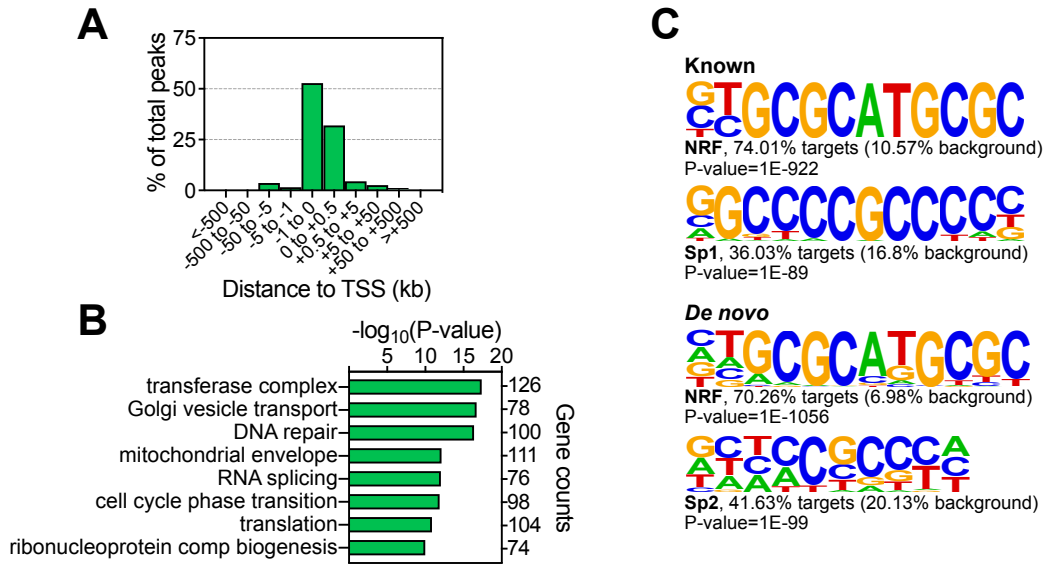


Figure 6-1: ChIP-Seq of HCF-1$_N$ in Ramos cells
(**A**) Distribution of HCF-1$_N$ peaks in Ramos cells in relation to the nearest TSS, as determined by ChIP-Seq. (**B**) GO categories strongly represented in genes nearest to HCF-1$_N$ peaks in Ramos cells. (**C**) Known and *de novo* motif analysis of HCF-1$_N$ peaks in Ramos cells. Two of the most highly enriched motifs are shown, as well the percentage of target and background sequences with the motif, and the P-value.

Co-binding of MYC and HCF-1

While enrichment of the NRF1 motif suggests that MYC in Ramos cells may co-bind with HCF-1_N, I analyzed this further by looking at the intersection of my HCF-1_N ChIP-Seq data with a previously published MYC ChIP-Seq dataset from this same cell line [21]. In Ramos cells, there are significantly more binding sites for MYC than there were for HCF-1_N, with almost all HCF-1_N sites showing overlap with MYC (Figures 6-2A and 6-2B). Furthermore, MYC peaks are generally much wider than those for HCF-1_N, and may have a slightly stronger signal intensity at co-bound regions compared to non-co-bound (Figures 6-2A). Looking specifically at the ~1,600 overlapping peaks shown in Figure 6-2B, these strongly resembled the HCF-1_N data alone, including promoter-proximal distribution (Figure 6-2C), enrichment at genes involved in translation, mitochondrial function, and ribonucleoprotein complex biogenesis (Figure 6-2D), and prevalence of the NRF1 motif in both known and *de novo* analyses (Figure 6-2E).

The apparent higher ChIP-Seq signal of MYC at co-bound sites observed in Figure 6-2A is reflected in a greater fragment depth for MYC where it overlaps with HCF-1_N compared to where there is no overlap (Figure 6-2F). Although more modest, HCF-1_N also appears to have slightly stronger signal when co-binding with MYC is found (Figure 6-2G), but this may be skewed by a low number of non-co-bound regions. Focusing specifically on genes whose roles are in protein synthesis and mitochondrial function, and were affected at the RNA level by the 4A and VP16 HBM mutants, I found a tight overlap of MYC and HCF-1_N on chromatin (Figure 6-2H). For most of the genes shown, both the MYC and HCF-1 peaks tended towards being upstream of the TSS, with POLR1A the exception to this (Figure 6-2H). Furthermore, many of these peaks show a strong alignment with the CATGCG motif element (Figure 6-2H), corresponding to that bound by NRF1.

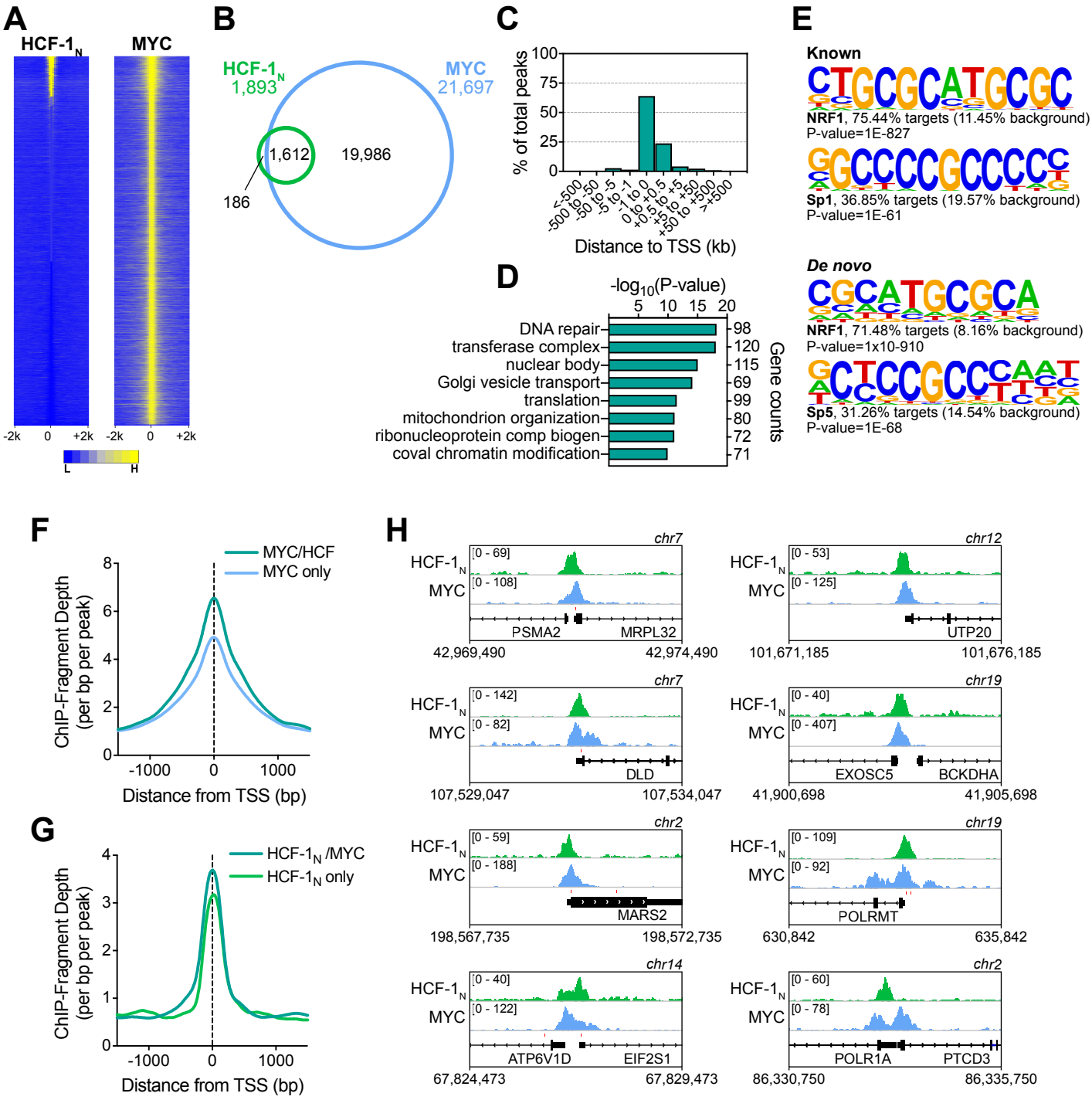


Figure 6-2: Co-binding of MYC and HCF-1$_N$

(**A**) Heatmap of all MYC peaks in Ramos cells (from GSE126207) and the corresponding region in Ramos HCF-1$_N$ ChIP-Seq, representing the combined average of normalized peak intensity in ± 2 kb regions surrounding the peak centers with 100 bp bin sizes. Ranking is by peak intensity in HCF-1$_N$. (**B**) Venn diagram showing HCF-1$_N$ and MYC peaks in Ramos cells, and the number of regions that overlap between the datasets. ChIP-Seq data for MYC are from GSE126207. (**C**) Distribution of regions that overlap between MYC and HCF-1$_N$ peaks in Ramos cells in relation to the nearest TSS. (**D**) GO enrichment analysis of genes nearest to regions of overlap between MYC and HCF-1$_N$ peaks. (**E**) Known and *de novo* motif analysis of regions that overlap between MYC

and HCF-1_N peaks in Ramos cells. For each motif represented, percentage of target and background sequences with the motif, and P-value are shown. (**F**) Normalized MYC ChIP-Seq fragment counts where peaks overlap with HCF-1_N (MYC/HCF), compared to where they do not overlap (MYC) in Ramos cells. Data are smoothed with a cubic spline transformation. (**G**) Normalized HCF-1_N ChIP-Seq fragment counts where peaks overlap with MYC (HCF-1_N/MYC) compared to where they do not overlap (HCF-1_N only). Data are smoothed with a cubic spline transformation. (**H**) Example Integrative Genomics Viewer (IGV) screenshots of regions that have overlapping peaks for MYC and HCF-1_N in Ramos cells. The red line below the peaks represents the NRF1 CATGCG motif.

Direct targets for regulation by MYC and HCF-1

Using my extensive interrogation of the contribution of the MYC–HCF-1 interaction to regulating gene expression, combined with identification of their co-bound genes, I determined which genes are directly regulated by this interaction in Ramos cells. Of the genes co-bound by MYC and HCF-1, about one-third encoded transcripts that were changed with the 4A or VP16 HBM mutants (Figure 6-3A). While similar proportions of these transcripts were increased and decreased, slightly more were decreased for the 4A mutant and slightly more were increased with the VP16 HBM mutant (Figure 6-3A). If I then focus only on transcripts that were anti-correlated between the 4A and VP16 HBM mutants, 112 of these were bound by both MYC and HCF-1, with ~60% decreased with the 4A mutant (Figure 6-3A). Although these anti-correlated sets contained relatively few genes, ontology analysis of those from "4A down/VP16 HBM up" showed a strong enrichment of categories that have repeatedly been emphasized in my research, including ribosome biogenesis and mitochondrial matrix (Figure 6-3B).

Finally, I added the criteria of also being affected by HCF-1_N degradation, leading to a list of almost 60 genes that I classify as high-confidence, direct targets of the MYC–HCF-1 interaction in Ramos cells (Figure 6-3C). The majority of these were decreased with the 4A mutant and increased with the VP16 HBM mutant, with a definitive correlation between the 4A mutant and HCF-1_N degradation. Based on the repeated appearance of genes relating to mitochondrial function and protein synthesis, I summarized the contribution of each of the direct targets to these processes. There was a definitive enrichment for these functions, predominantly for transcripts that were down with 4A and up with VP16 HBM (Figure 6-3C). A number of genes bridge both functions, including those involved in mitochondrial ribosome biogenesis (MRPS12 and MRPL32) and aminoacyl-tRNA biosynthesis (MARS2). Furthermore, a number of the transcripts that were increased with the 4A, such as MBD2 and CBX5 (heterochromatin protein 1, HP1), encode proteins that typically repress transcription of rDNA [358, 359], and are therefore consistent in the overall effect on these processes. Finally, all of these direct target genes experienced only relatively small shifts in expression in response to either altering the MYC–HCF-1 interaction or degradation of HCF-1_N.

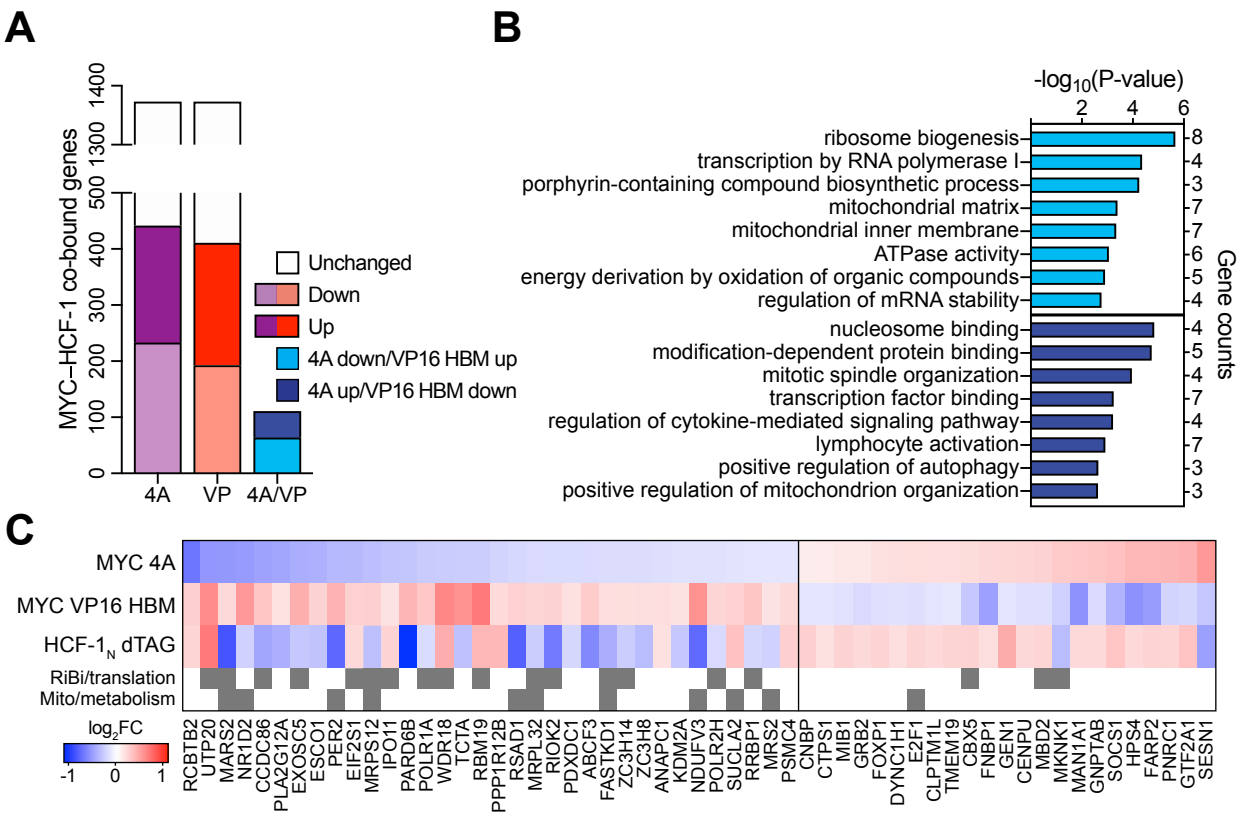


Figure 6-3: Identification of direct MYC–HCF-1 target genes
(**A**) Relationship between protein-coding genes that are co-bound by promoter-proximal MYC and HCF-1$_N$ by ChIP-Seq, and are significantly (FDR < 0.05) decreased or increased in response to the 4A or VP16 HBM mutations. Also shown are genes where the expression changes were anti-correlated between the 4A and VP16 HBM mutants. (**B**) Ontology analysis of genes that are co-bound by MYC and HCF-1 and whose transcripts were anti-correlated between the 4A and VP16 HBM mutants. Analysis of transcripts that were decreased with 4A and increased with VP16 HBM is shown at the top, and analysis of transcripts that were increased with 4A and decreased with VP16 HBM is shown at the bottom. (**C**) Heatmap showing genes that are co-bound by promoter proximal MYC and HCF-1$_N$ in Ramos cells, have anti-correlative gene expression changes between the 4A and VP16 MYC mutants, and have significant gene expression changes with HCF-1$_N$ degradation. Genes that fall into GO categories relating to ribosome biogenesis or translation (RiBi/translation), and mitochondrial function or metabolism (Mito/metabolism) are highlighted.

Independent function of HCF-1 and WDR5 with MYC

MYC is recruited to a subset of ribosomal protein genes through its MbIIIb-dependent interaction with the scaffolding protein WDR5 [21]. HCF-1 and WDR5 are found in a number of the same complexes, but their interaction with MYC is independent of each other, as demonstrated by me (Figures 3-2, 3-3, and 4-1) and Thomas *et al.* [21, 179]. To confirm their functional independence in regards to MYC, I compared their relationship on chromatin, finding only 88 regions of overlap between MYC, HCF-1, and WDR5 (Figure 6-4, left). Only three of these correspond to sites at which MYC undergoes WDR5-dependent recruitment (Figure 6-4, bottom-right). These data further emphasize the independent functions of WDR5 and HCF-1 in regards to their relationship with MYC.

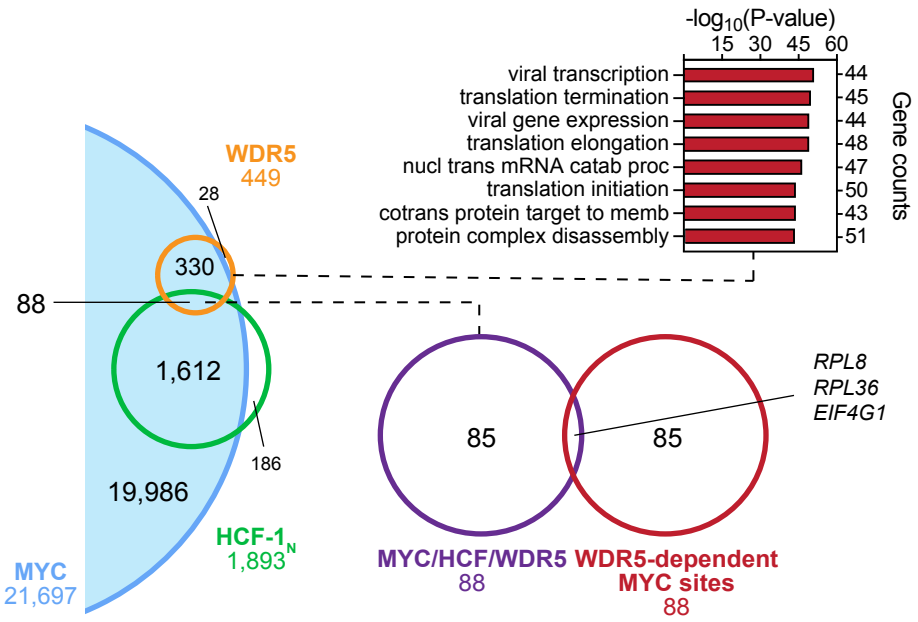


Figure 6-4: Independent function of HCF-1 and WDR5 with MYC
Venn diagram showing relationship between HCF-1_N, MYC, and WDR5 peaks in Ramos cells, and the overlap between co-bound genes and genes for which WDR5 is responsible for MYC recruitment. GO enrichment analysis of genes co-bound by MYC and WDR5—taken from Thomas *et al.* [21]—is also shown.

Impact of MYC–HCF-1 interaction on recruitment to chromatin

Numerous factors have been identified for both MYC and HCF-1 recruitment to chromatin, and this process is widely accepted to be the primary mechanism by which protein-protein interactions regulate transcription [360]. Cowling *et al.* [181] previously suggested that deletion of MbIV from N-MYC impairs its ability to bind to DNA, as measured by electrophoretic mobility shift assay (EMSA). However, the Tansey laboratory has also found that MYC:MAX dimers carrying the 4A mutation and derived from cell lysates are equally as efficient at binding to DNA as wild-type MYC [179]. To confirm this and to explore whether the VP16 HBM mutant is similarly efficient, I purified recombinant MYC (WT, 4A, or VP16 HBM) and MAX independently from *Escherichia coli*, and combined these to generate MYC:MAX dimers (Figure 6-5A). Using EMSAs, I found that the 4A and VP16 HBM mutations caused no overt disruption to DNA-protein complex formation, and these are effectively disentangled in the presence of a specific competitor (Figure 6-5B).

I also determined whether these MYC HBM mutants affected binding of both MYC and HCF-1_N to chromatin through ChIP-seq of switchable *MYC* cells that had been treated with 4-OHT for 24 hours. For HCF-1_N, the mutants had very similar ChIP-seq signal intensity (Figure 6-5C), with fewer than ten significant changes for either of the mutants, none of which were shared. In contrast, MYC-HA ChIP-seq had extensive alterations in signal with the 4A and VP16 HBM mutants, but almost all of these peaks were increased compared to wild-type (Figure 6-5D). Indeed, none of the significant changes that were present in both mutants were anti-correlative, suggesting that HCF-1 is not responsible for the variations observed. This is reflected in Figure 6-5E, which highlights the significantly changed peaks that exceed an absolute fold-change of 1.62 for both mutants; the vast majority of these were increased for both mutants compared to wild-type. These ramifications for both MYC and HCF-1, or lack thereof, can additionally be seen by looking at screenshots of individual peaks (Figure 6-5F).

Specifically focusing on genes that I determined to be direct targets of the MYC–HCF-1 interaction, I confirmed that neither the 4A or VP16 HBM mutants caused a significant change in the detection of

MYC or HCF-1_N binding to chromatin (Figures 6-4G and 6-4H). There was, however, a modest increase in MYC binding with the VP16 HBM, even at genes that are not direct targets (Figures 6-4G). Thus, the mechanism by which the MYC–HCF-1 interaction affects gene expression is likely co-recruitment-independent.

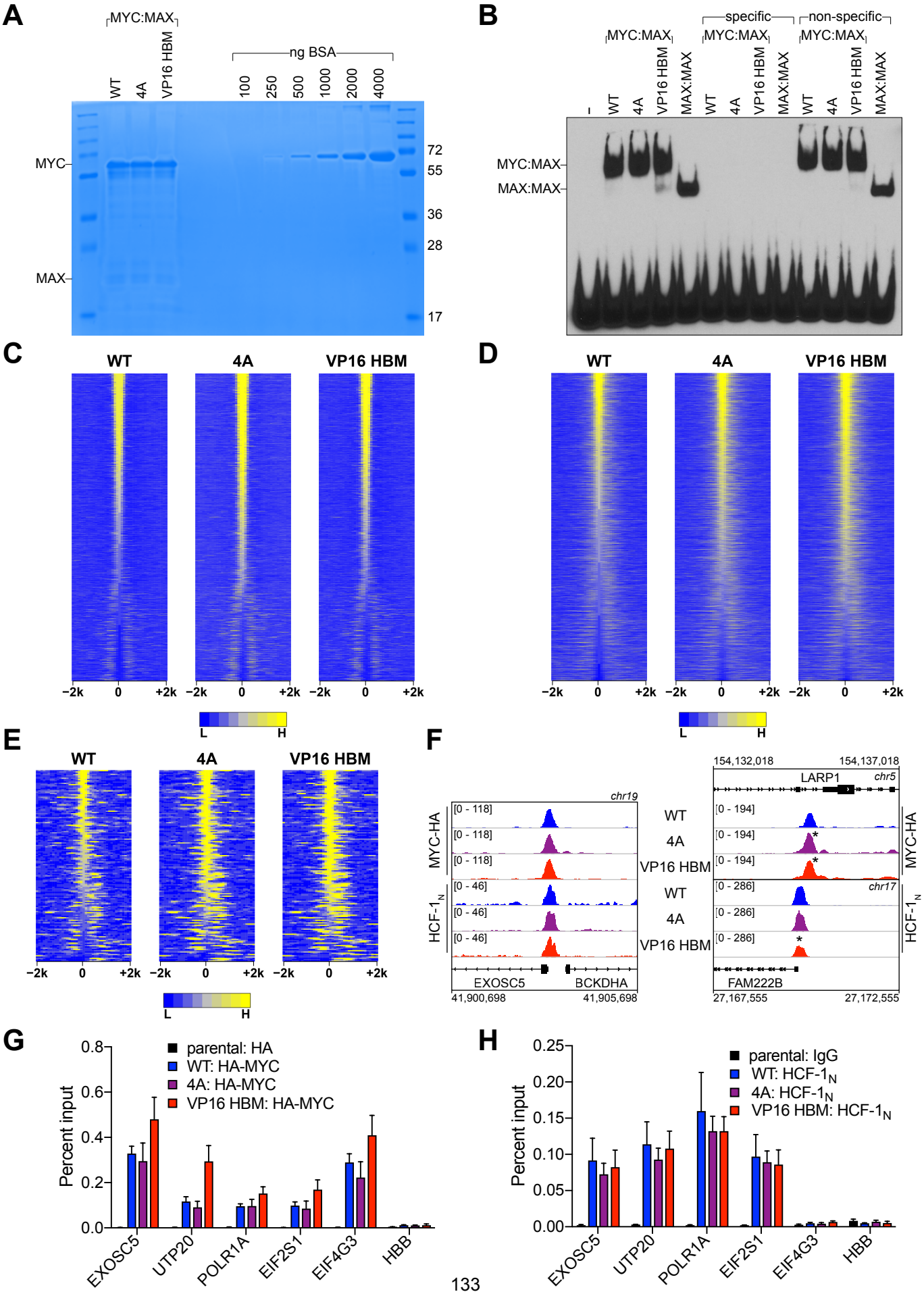

A
MYC:MAX
WT
4A
VP16 HBM
ng BSA
100
250
500
1000
2000
4000
MYC
MAX
72
55
36
28
17
B
MYC:MAX
specific
non-specific
MYC:MAX
MYC:MAX
-
WT
4A
VP16 HBM
MAX:MAX
MYC:MAX
MAX:MAX
C
WT
4A
VP16 HBM
−2k
0
+2k
L
H
D
WT
4A
VP16 HBM
E
WT
4A
VP16 HBM
F
chr19
MYC-HA
HCF-1N
[0 - 118]
[0 - 46]
WT
4A
VP16 HBM
EXOSC5
BCKDHA
41,900,698
41,905,698
154,132,018
154,137,018
LARP1
chr5
[0 - 194]
chr17
[0 - 286]
FAM222B
27,167,555
27,172,555
G
Percent input
parental: HA
WT: HA-MYC
4A: HA-MYC
VP16 HBM: HA-MYC
EXOSC5
UTP20
POLR1A
EIF2S1
EIF4G3
HBB
H
parental: IgG
WT: HCF-1N
4A: HCF-1N
VP16 HBM: HCF-1N

Figure 6-5: Impact of HBM mutations on DNA and chromatin binding
(Previous page)
(**A**) Recombinant MYC (WT, 4A, or VP16 HBM):MAX dimers used in these assays. Dimers were separated by SDS-PAGE alongside a BSA standard, and stained using Coomassie Brilliant Blue. (**B**) Electrophoretic mobility shift assay (EMSA) using recombinant MYC:MAX and MAX:MAX dimers incubated with a biotinylated E-box probe (5′-GCTCAGGGACCACGTGGTCGGGGATC-3′). Binding reactions were conducted with or without 100-fold excess of unlabeled specific (as above) or non-specific (5′-GCTCAGGGACCAGCTGGTCGGGGATC-3′) competitors. (**C**) and (**D**) Heatmap of the combined average of normalized peak intensity in 100 bp bins for all HCF-1_N (C) or MYC-HA (D) ChIP-seq peaks that were within ±2 kb of a TSS, and scaled across each of the cell lines. (**E**) Heatmap of the combined average of normalized peak intensity in 100 bp bins for MYC-HA peaks that were significantly changed (FDR < 0.05 & $|\log_2 FC| > 0.7$) for both the 4A and VP16 HBM mutants, and were within ±2 kb of a TSS. (**F**) Example IGV screenshots of regions that had non-significant (left) or significant (right) changes for MYC-HA or HCF-1_N by ChIP-seq. Asterisks mark the peaks that were significantly changed compared to WT. (**G**) and (**H**) ChIP, using anti-HA (G) or anti-HCF-1_N (H) antibody, was performed on parental or switchable Ramos cells treated for 24 hour with 20 nM 4-OHT. Enrichment of genomic DNA was monitored by qPCR using primers that amplify across peaks. *HBB* is a negative locus for HA-MYC and HCF-1_N, *EIF4G3* is a negative locus for HCF-1_N. ChIP efficiency was measured based on the percent recovery from input DNA. Shown are the mean and standard error for three biological replicates.

Discussion

In this chapter, I have looked at the relationship between MYC and HCF-1_N on chromatin in Ramos cells, and whether their interaction influences their levels at chromatin. Understanding these aspects of the MYC–HCF-1 interaction are important for determining their direct target genes and identifying the mechanism by which they facilitate transcription, respectively. Compared to what has previously been found with this HCF-1_N antibody in HeLa cells, I found substantially fewer binding sites in Ramos cells and a higher proportion of promoter-proximal sites. The cause of these differences may be largely technical, or simply may reflect cell type differences. Despite this, I conclude that almost all of the HCF-1_N binding sites overlapped with MYC binding sites in Ramos cells.

The NRF1 motif that is over-represented with both HCF-1_N ChIP-Seq and at sites of MYC–HCF-1_N co-binding is found in a number of genes involved in mitochondrial function, with the NRF1 transcription factor proposed to coordinate gene expression across the mitochondrial and nuclear genomes [361]. NRF1 achieves this in concert with the PGC family [274], and while an interaction between NRF1 and HCF-1 has not been identified, all PGC family members can interact with the HCF-1_{VIC} domain [275, 276]. As a non-canonical E-box, MYC is also capable of binding to the NRF1 motif *in vitro* [18] and *in vivo* [356], and can regulate transcription from genes containing this motif [357]. Indeed, one of the proposed mechanisms through which MYC affects the expression of such a wide array of genes when it is overexpressed, is through invasion to low-affinity binding sites and non-canonical E-boxes, including the NRF1 motif [25, 74].

While expression of only a subset of MYC–HCF-1 co-bound genes was affected by the HBM mutants and HCF-1_N degradation, there are various limitations that may have affected this outcome. Perhaps the most influential of these is the strict criteria used to identify target genes of this interaction, including anti-correlative changes between the 4A and VP16 HBM mutations, and alterations at an early timepoint in response to HCF-1_N degradation. The rigidity of such criteria is compounded by the low

number of HCF-1_N peaks detected by ChIP-Seq, which may be a consequence of only detecting stable association of HCF-1_N with chromatin. Indeed, this may be particularly relevant to the histone modifying role of HCF-1, with recent work demonstrating that histone modifications may occur with only transient ("hit and run") binding of the modifying complex [362]. The high-confidence, direct MYC–HCF-1 targets that I did identify show a definitive functional relationship, with the majority being involved in controlling protein synthesis and mitochondrial function. While neither of these functions are unique to MYC, the presence of both is intriguing, and the specific role of HCF-1 in contributing to the MYC-driven transcriptional regulation of these genes is a novel finding. Among the direct targets of the MYC−HCF-1 interaction are genes that catalyze rate-limiting steps in both ribosome biogenesis (rDNA transcription by POLR1A) and translation (initiator tRNA binding to start codon by EIF2S1 [EIF-2α]) [363, 364]. MYC regulates all stages of ribosome biogenesis, including transcription of rDNA, rRNA processing proteins, and ribosomal protein genes [85]. Interestingly, even with the direct MYC–HCF-1 targets classified as mitochondrially-connected, we see links to protein synthesis—MARS2, MRPS12, and MRPL32, for example, are specifically involved in the synthesis of mitochondrial proteins, including those required for oxidative phosphorylation. Furthermore, these target genes are independent of those regulated by MYC–WDR5, with this interaction showing specificity for ribosomal protein genes [21], rather than the various ribosome biogenesis factors that I identified in this work.

I also observed no consistent influence of the MYC HBM mutants on binding of either MYC or HCF-1 to chromatin. This points to a co-recruitment-independent mechanism for regulation of the identified target genes. While alternative mechanisms will be explored in the final chapter, this finding has previously been observed for the Myc–HCF interaction in *D. melanogaster*, despite this conserved interaction occurring through an alternative mode (i.e. HBM-independent) [295]. In this context, the authors did however suggest that multiple contact points on Myc may be responsible for association with HCF, and further hypothesized that the mechanism of Myc–HCF transcriptional regulation involved the formation of a trimeric complex [295]. Thus, not only is the ability HCF-1 to contribute to MYC-driven growth and transcriptional activation conserved, including of ribosome biogenesis and rRNA processing genes

[295], the mechanism by which it achieves this may also be conserved. Additionally, knockdown of HCF-1 does not affect the recruitment of E2F1, and vice versa [349]. These emphasize that, for some HCF-1 interaction partners, the capacity of the interaction to regulate transcriptional output is a two-step system, with the arrival at chromatin being the first step, which occurs independent of the interaction, and the modification to the transcription process being the second, which occurs dependent on the interaction.

The combination of direct target genes involving mitochondrial biogenesis and identification of a non-canonical motif in MYC–HCF-1 co-bound sites further suggests that this interaction may be most relevant for transcriptional regulation when MYC is overexpressed. While MYC may demonstrate only minimal binding to the NRF1 motif when at physiological levels, its ability to activate transcription from the associated genes might only occur when its levels are high and when there is association with HCF-1. Furthermore, genes involved in mitochondrial function and metabolism are commonly considered to be gained MYC targets through its binding to low-affinity sites when overexpressed [354], but it is also possible that the auxiliary ribosome biogenesis factors I have identified are gained targets. Indeed, this seems like a strong possibility considering that a number of these targets are rate limiting for their respective processes, meaning that MYC overexpression orchestrates accelerated ribosome biogenesis, translation, and mitochondrial biogenesis above baseline to enable rapid cell growth. To address this, one could create a *leaky* inducible system for overexpression of MYC, WT or mutants, so that genes that are bound by MYC and HCF-1, and regulated by their interaction, could be determined in the context of low and high levels of MYC protein.

Chapter VII

MYC−HCF-1 interaction in MYC-driven tumor engraftment and maintenance

Introduction

The overexpression of MYC in the majority of cancers makes it a promising target for the treatment of cancer. However, there is an absence of feasible MYC inhibitors, largely due to the intrinsically disordered nature of the protein. Instead, targeting MYC interaction partners, such as HCF-1, is perhaps the best opportunity available. The Tansey laboratory has previously demonstrated that cells overexpressing the MYC 4A mutant have a reduced capacity for tumor formation in a mouse xenograft assay compared to wild-type MYC [179]. While this approach was a promising first step to understanding the MYC–HCF-1 contribution to cancer, it is a not necessarily a good marker for how effective inhibiting this interaction will be for anti-cancer therapies. To make conclusions about this, we must determine if perturbation of the MYC–HCF-1 interaction, following tumor formation, results in regression.

Inactivation of MYC, even for a short period of time, is often sufficient for tumor regression [365], even in cancers not driven by MYC [207]. A number of specific gene expression changes in response to the loss of MYC have been identified as possible mechanisms of regression, including the induction of senescence by re-expression of $p15^{INK4b}$ and $p21^{CIP}$ [366], reduced expression of ribosome biogenesis genes [367], and impairment of metabolism [368]. My interrogation of the MYC–HCF-1 interaction has revealed that this interaction directly regulates the expression of genes involved in ribosome biogenesis, translation, and mitochondrial function, suggesting that their association is likely essential for tumor growth and maintenance *in vivo*. However, this may be somewhat confounded by reports suggesting that intact, functional p53 is necessary for tumor regression following MYC inactivation [366]. While the mutant p53 status of Ramos cells may affect the capacity of the 4A mutant to cause

tumor regression in this context, the Tansey laboratory has previously demonstrated that inactivation of MYC in Ramos cells is sufficient to cause rapid and sustained tumor regression *in vivo* [21].

In culture, disruption of the MYC–HCF-1 interaction caused a slowing of cell growth, but there was no observable increase in apoptotic cells. This is a commonly observed phenotype, even for inhibition of MYC, and will still result in apoptosis and tumor regression *in vivo* [369]. The pressures faced by tumors growing *in vivo* is greater than cells confront in culture, and totality of MYC function is more likely to be wholly applicable. Interrogating the MYC–HCF-1 interaction in this context will reveal not only their contribution to MYC-driven tumor growth and maintenance, but also enable identification of their specific target genes that remain essential *in vivo*.

Results

MYC–HCF-1 interaction in tumor engraftment and maintenance

In a collaboration with Clare Adams and Christine Eischen at Thomas Jefferson University, we first looked at how tumor engraftment of Ramos cells is impacted by the 4A mutant compared to wild-type. For this series of experiments, I employed an additional clone carrying the switchable *MYC* 4A allele (4A-2) and a switchable *MYC* Δ264 cell line, in which the excised exon 3 is not replaced upon switching, creating a truncated MYC protein [21]. In this additional cell line, MYC loses its ability to both interact with HCF-1 and to associate with chromatin. With this approach, cells were switched in culture for 24 hours using 4-OHT and injected into the flanks of nude mice (Figure 7-1A). We found that, while WT cells rapidly formed tumors, this was delayed significantly with each of the mutant cell lines (Figure 7-1B and 7-1C). Indeed, while the WT cell-injected mice all had to be sacrificed within 21 days, this time was increased by 50% with the mutants (Figure 7-1D). I purified genomic DNA from tumors extracted from mice to determine the proportion of switched cells present, finding that for each of the mutant tumors, the majority of cells were unswitched/WT (Figure 7-1E). This suggests that the majority

of the tumor outgrowth observed for the mutant cell lines was driven by the presence of remaining unswitched cells.

Next, we exploited the inducible nature of the switchable *MYC* allele system to address how disrupting the MYC–HCF-1 interaction affects tumor maintenance. In this assay, unswitched cells were injected into the flanks of nude mice, allowed to form tumors, and then treated with tamoxifen to induce switching following engraftment (Figure 7-1F). Prior to *in vivo* switching, all of the cell lines rapidly formed tumors at similar rates (Figures 7-1G and 7-1H). However, following treatment with tamoxifen, only WT tumors continued to grow, with the mutant tumors demonstrating a rapid and prolonged regression (Figures 7-1G and 7-1H). Thus, the WT cells required all mice to be sacrificed within 29 days, with the mice in which cells were switched to mutant MYC surviving until completion of the experiment (Figure 7-1I). We also looked at the emergence of cell death in the days following injection with tamoxifen, finding an accumulation of annexin V-positive cells (Figure 7-1J), caspase activity (Figure 7-1K), and sub-G_1 cells (Figure 7-1L) in the mutants compared to WT cells within 48 hours. These data indicate cell death was the primary driving force behind tumor regression, with the specific mechanism of this currently unclear.

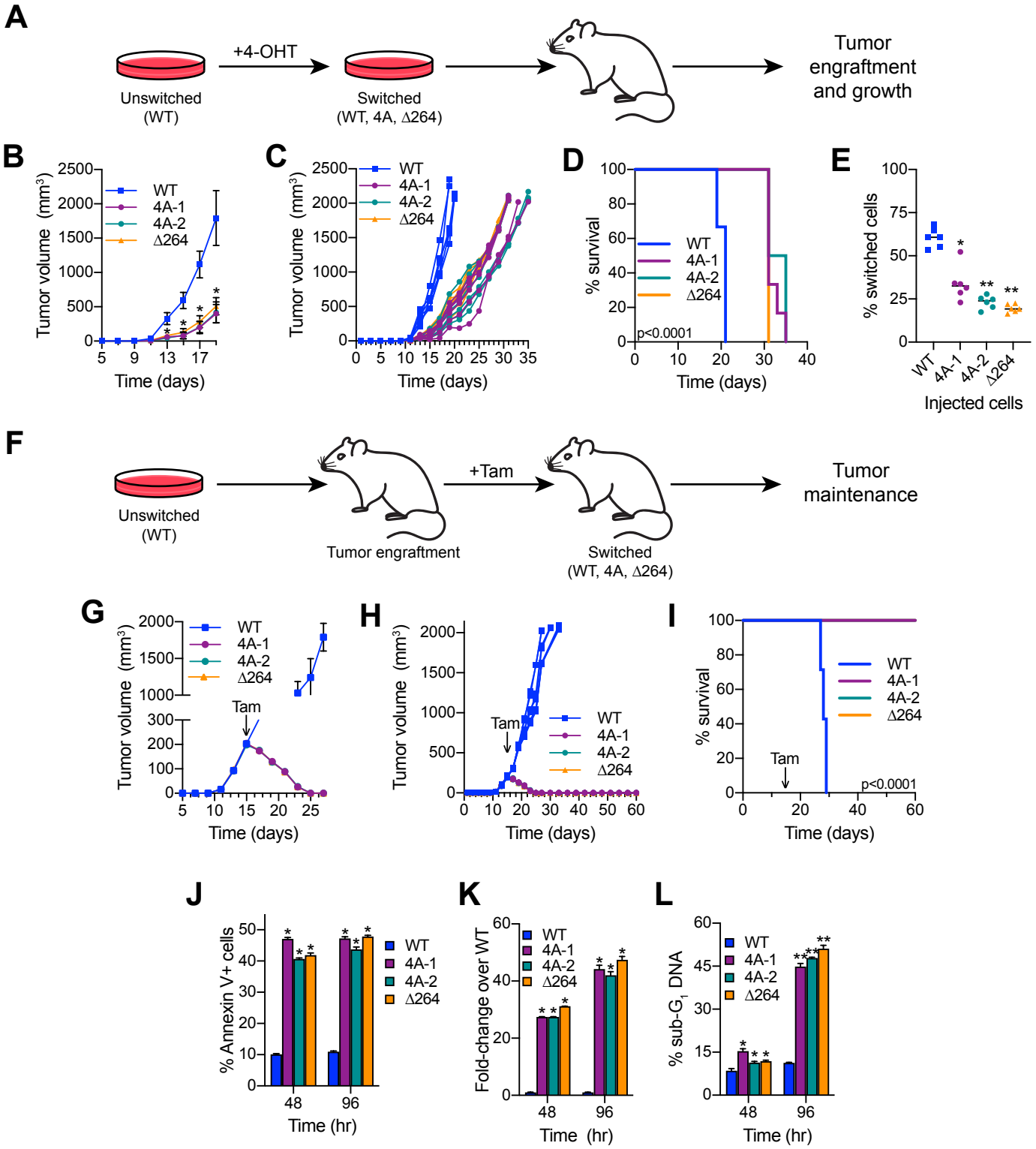


Figure 7-1: The MYC–HCF-1 interaction is required for tumor engraftment and maintenance
(**A**) Tumor engraftment schema: WT, 4A-1, 4A-2, and Δ264 cells are switched in culture prior to injection into the flanks of nude mice to test the impact of the mutations on tumor engraftment and growth. (**B**) Average tumor volume over time following injection of switched cells. Shown are the mean and standard error for six mice each. Only days five to 19 are shown here; the full course of the experiment is depicted in (C). Student's t-test between

WT and each of the mutants was used to calculate P-value; *P < 0.000043. (**C**) Tumor volumes for individual mice in the tumor engraftment assay. (**D**) Kaplan-Meier survival curves of mice (n=6 of each) injected with switched cells. Log-rank test was used to calculate P-value (< 0.0001) from six biological replicates. (**E**) PCR assays of genomic DNA was used to determine the proportion of switched cells present in each tumor after sacrifice. Each dot represents an individual tumor, and the line indicates the mean for each group. Student's t-test between WT and each of the mutants was used to calculate P-values; a = 0.0002, b < 0.0001. (**F**) Tumor maintenance schema: Unswitched WT, 4A-1, 4A-2, and ∆264 cells were injected into the flanks of nude mice. Tumors were grown until day 15, at which point mice received tamoxifen injections (one/day for three days) to induce switching of the cells. (**G**) Average tumor volume before and after cells were switched. The day at which tamoxifen (Tam) administration was initiated is indicated with an arrow. Shown are the mean and standard error for seven mice for WT and six mice for 4A-1, 4A-2, and ∆264 cells. Plots for 4A-1, 4A-2, and ∆264 are stacked on top of each other. (**H**) Tumor volumes for individual mice in the tumor maintenance assay. The day at which tamoxifen (Tam) administration was initiated is indicated with an arrow. Plots for 4A-1, 4A-2, and ∆264 are stacked on top of each other. (**I**) Kaplan-Meier survival curves of mice in the tumor maintenance assay (n=7 for WT, and n=6 for 4A-1, 4A-2, and ∆264). The day at which tamoxifen (Tam) administration was initiated is indicated with an arrow. Log-rank test was used to calculate P-value (< 0.0001). Curves for 4A-1, 4A-2, and ∆264 are stacked on top of each other. (**J**) Annexin V staining and flow cytometry were performed on cells isolated from tumors at 48 and 96 hours following the first tamoxifen administration to determine the extent of apoptosis. Shown are the mean and standard error for four mice each. Student's t-test between WT and each of the mutants was used to calculate P-value; *P < 0.000001. (**K**) Caspase activity in cells isolated from tumors (n=4 each) at 48 and 96 hours following the first tamoxifen administration was measured. Shown are the mean and standard error for the fold-change (FC) over WT for each. Student's t-test between WT and each of the mutants was used to calculate P-value; *P < 0.000001. (**L**) Propidium staining and flow cytometry were performed on cells isolated from four tumors each at 48 and 96 hours following the first tamoxifen administration to determine the extent of cell death. Shown are the mean and standard error. Student's t-test between WT and each of the mutants was used to calculate P-value; *P < 0.000628, **P < 0.000001.

Gene expression changes with tumor regression

To investigate the underlying transcriptional changes associated with the tumor regression, I performed RNA-Seq on tumors that were excised 48 hours following treatment with tamoxifen. As to be expected, the number of significantly changed transcripts was extensive with each of the mutants, including a similar number of transcripts decreased and increased (Figure 7-2A). Comparing the transcripts that were decreased (Figure 7-2B) and increased (Figure 7-2C) in the context of all three mutants identified a list of ~3,000 that were shared. The gene ontology of those that were decreased included ribosome biogenesis, translation, mitochondrial envelope, and tRNA metabolic process (Figure 7-2D). On the other hand, transcripts that were increased showed enrichment for functions such as transcription coregulator activity, kinase binding, and vacuole (Figure 7-2E).

The presence of cell death *in vivo* likely compounded the extent of gene expression effects observed with the mutants. To narrow these changes down to those caused specifically by altering the MYC–HCF-1 interaction, I related the *in vivo* transcripts to those identified *in vitro*. More than 70% of the transcripts that were decreased with the 4A mutant *in vitro* were also decreased with either or both of the 4A clones *in vivo* (Figure 7-2F). These ~1,000 shared transcripts are involved in ribosome biogenesis, translation, nucleolar part, and mitochondrial gene expression (Figure 7-2G). Although similar numbers of transcripts were increased as decreased with both the 4A-1 and 4A-2 clones, fewer than 600 of increased transcripts were also increased *in vitro* (Figure 7-2H). These transcripts showed much less enrichment by gene ontology, but included the categories chromatin binding, response to topologically incorrect protein, and kinase binding (Figure 7-2I). These findings again emphasize the dominant effect of disrupting the MYC–HCF-1 interaction is down-regulation of gene expression, and the apparent role of this interaction in controlling genes involved protein synthesis and mitochondrial function.

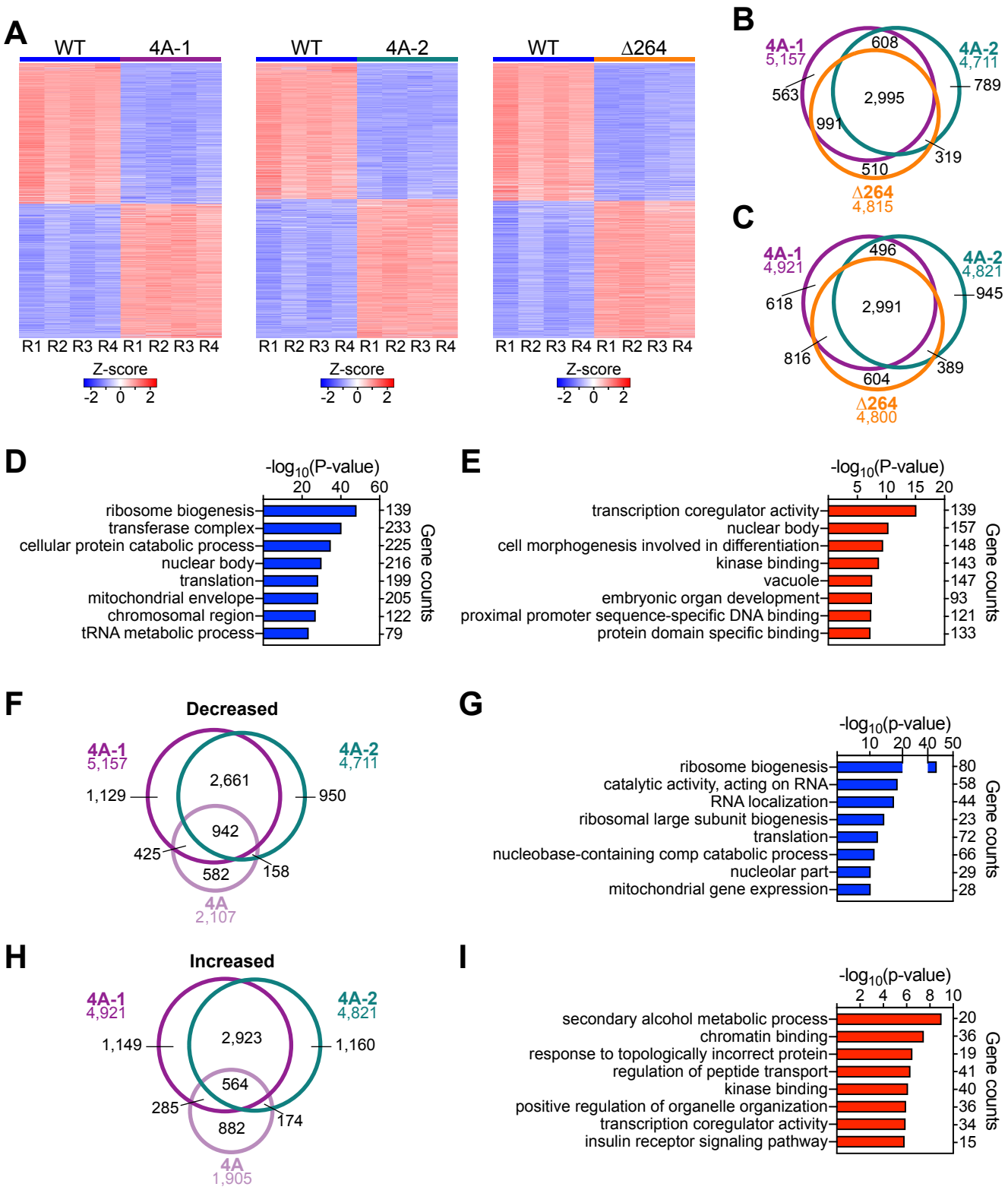


Figure 7-2: Gene expression changes during tumor regression

(**A**) Z-transformed RNA-Seq data from tumors (n=4) extracted after switching, ranked by FC. Genes that were significantly (FDR < 0.05) impacted compared to WT are shown. (**B**) Venn diagram showing the overlap of genes significantly (FDR < 0.05) decreased in the 4A-1, 4A-2, and Δ264 tumors. (**C**) Venn diagram showing the relationship of genes significantly (FDR < 0.05) increased in the 4A-1, 4A-2, and Δ264 tumors. (**D**) GO enrichment

analysis of genes with significantly (FDR < 0.05) decreased expression in the 4A-1, 4A-2, and Δ264 tumors. (**E**) GO enrichment analysis of genes significantly (FDR < 0.05) increased in the 4A-1, 4A-2 , and Δ264 tumors. (**F**) Venn diagram showing the relationship between genes significantly (FDR < 0.05) decreased in the 4A cell line and the 4A-1 and 4A-2 tumors. (**G**) Gene ontology enrichment analysis of genes significantly (FDR < 0.05) decreased in the 4A cell line, and the 4A-1 and 4A-2 tumors. (**H**) Venn diagram showing the overlap of genes significantly (FDR < 0.05) increased in the 4A cell line, and the 4A-1 and 4A-2 tumors. (**I**) Gene ontology enrichment analysis of genes significantly (FDR < 0.05) increased in the 4A cell line, and the 4A-1 and 4A-2 tumors.

MYC–HCF-1 target genes are disrupted in tumor regression

Finally, almost all of the genes that I identified as direct target genes of MYC–HCF-1 interaction that were decreased with the 4A mutant *in vitro*, were also impacted in the 4A-1 and 4A-2 tumors during regression (Figure 7-3). For the most part, these changes occurred in the same direction both *in vitro* and *in vivo*, but this relationship appears to be more consistent for genes involved in ribosome biogenesis and mitochondrial function. Indeed, this may emphasize that these are the more important targets of the MYC–HCF-1 interaction. On the other hand, fewer of these direct target genes that were increased with the 4A mutant *in vitro*, were also affected *in vivo* (Figure 7-3). For all of these genes, it is possible that, following perturbation of the MYC–HCF-1 interaction, they faced immediate disruption to their expression that was consistent with what I observed *in vitro*, but subsequent tumor regression affected what I ultimately detected by RNA-seq.

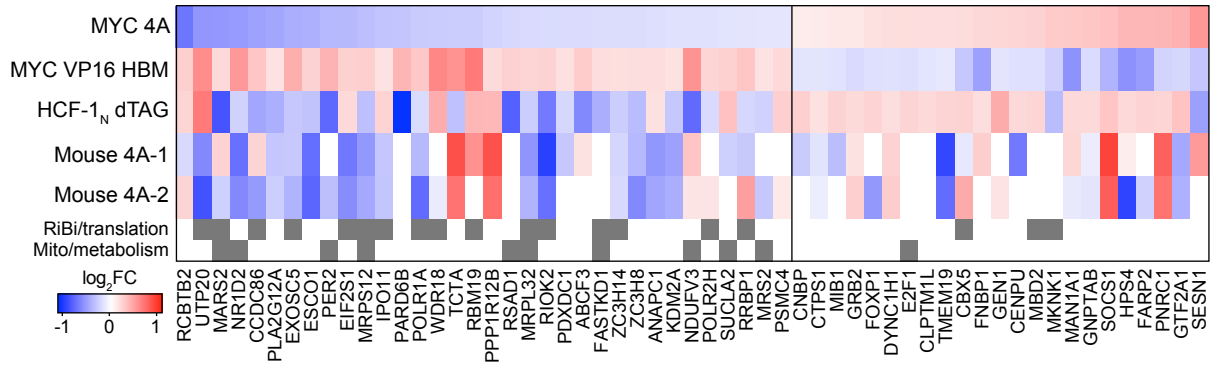


Figure 7-3: High-confidence MYC–HCF-1 target genes in regressing tumors
Heatmap showing genes that are co-bound by promoter proximal MYC and HCF-1_N in Ramos cells, have anti-correlative gene expression changes between for the 4A and VP16 MYC mutants, and have significant gene expression changes with HCF-1_N degradation. The effect on encoded transcripts 48 hours following switching *in vivo* is also shown. Genes that fall into GO categories relating to ribosome biogenesis or translation (RiBi/translation), and mitochondrial function or metabolism (Mito/metabolism) are highlighted.

Discussion

Using the switchable *MYC* cell lines, I was able to confirm the role of the MYC-HCF-1 interaction in tumor engraftment, and interrogate the importance of this interaction for tumor maintenance. Understanding how this interaction contributes to these processes is a necessary step in pursuing it for anti-cancer therapy. Furthermore, I have identified genes regulated by the MYC–HCF-1 interaction that are impacted during tumor regression. While preliminary, this work is a promising investigation into how development of MYC–HCF-1 inhibitors may progress, and highlights the potential in targeting MYC through its cofactors.

Disrupting the MYC–HCF-1 interaction was as effective at perturbing MYC-driven tumor engraftment and maintenance as deletion of the carboxy-terminus of MYC. While the MYC Δ264 mutation impairs interaction with HCF-1, it also prevents MYC binding to DNA and likely affects a more diverse gene set than disrupting only the MYC HBM. This equivalency may reflect similar mechanisms, or a saturation of cell stress or death pathways between these mutants. The cell death observed *in vivo* for the 4A mutant contrasted with what occurred *in vitro*, whereby cell growth appeared to only slow. This may be a consequence of different selective pressures in these contexts, with the changes that occurred in the 4A mutant creating a vulnerability that is most relevant, and with the biggest implications, in solid tumors *in vivo*. Cells that were switched *in vitro* to measure tumor engraftment had incomplete switching, as revealed by the outgrowth of unswitched cells during tumor development. Conversely, when cells were switched *in vivo*, tumor regression was complete and long-term. For both the engraftment and maintenance assays, including the outgrowth of wild-type cells in the engraftment assay, similar phenotypes have previously been observed when perturbing the MYC–WDR5 interaction using this same switchable system [21]. While possible, albeit unlikely, that switching was more efficient *in vivo*, the mechanism is difficult to attribute specifically, but cell switching may create an environment that is untenable for survival of the remaining unswitched cells. While apoptosis might first be triggered in the switched cells, neighboring unswitched cells may then undergo apoptosis-induced apoptosis

[370]. Such bystander apoptosis has previously been observed in treating mixed populations containing cancer cells both with (target) and without (bystander) an antigen [371]. In this context, both the target and bystander cells, the latter of which included Ramos cells, would undergo apoptosis in response to treatment [371], suggesting that this may be a possible mechanism by which unswitched Ramos cells undergo tumor regression alongside their switched, mutant counterparts. While this is an interesting conceptually, it might also suggest that neighboring, non-cancerous mouse cells would also be vulnerable. However, this issue is particularly relevant in the context of a mixed population of switchable cells, and likely would not be applicable with the end-goal of using small-molecular inhibitors to block the MYC–HCF-1 interaction in *all* MYC overexpressing cells.

The specific consequences of MYC inactivation are cell-type dependent, but, in general, inactivation leads to rapid and prolonged tumor regression [365], akin to what I have observed here. These consequences can include differentiation [372], senescence [366], apoptosis, or a combination of these [373]. In the case of lymphoma cells, cell-autonomous restoration of cell senescence pathways and subsequent apoptosis with MYC inactivation have been reported, but this is also dependent on functional p53 [366], at least in part due to its induction of anti-angiogenic factors [374]. This suggests that, either the mutant p53 present in Ramos cells is sufficient for the tumor regression I observed, or a different pathway is responsible in this context. The latter of these may be a consequence of disabling MYC, rather than reducing its levels, and directly affecting the expression of only a subset of its target genes.

How might regression and cell death, then, be specifically initiated with the 4A mutant? I propose that disruption to both mitochondria and protein synthesis is a one-two hit for cancer cells. The uncontrollable growth that is characteristic of cancer cells is dependent on sufficient protein production and metabolic output, and both of these are compromised by disrupting the MYC–HCF-1 interaction. Coordination of the various stages of ribosome biogenesis is necessary for normal and productive protein synthesis, with perturbation of this process causative of proteotoxic stress and activation of

stress response pathways [375]. Disrupting various stages of ribosome biogenesis, including rDNA transcription, rRNA processing, and nuclear import/export pathways, can lead to activation of p53 and initiation of apoptosis [376]. MYC itself has been implicated in this process [377]. While there are mixed reports of the p53 status of Ramos cells [378, 379], it has been suggested that p53-independent mechanisms may also be involved in the ribosomal stress response [375]. Mitochondria are also susceptible to proteotoxic stress [380], and this is particularly relevant here based on the disruption of genes involved in mitochondrial protein synthesis. Amongst the pathways activated with mitochondrial dysfunction is the integrated stress response, which can dictate the delicate balance of cell survival and death [381]. I mentioned this pathway previously as a possible driver of amino acid accumulation in the 4A mutant. While the primary goal of the integrated stress response is to enable cell survival in deleterious conditions, more intense or longer disturbances can limit a cell's capacity to achieve homeostasis, leading to cell death by apoptosis, necrosis, or autophagy [382]. It seems possible then, that *in vitro* the imbalance caused by the 4A mutant can be compensated for by the integrated stress response, but *in vivo*, Ramos cells are unable to achieve this sufficiently.

The process-specific genes that I have identified as being directly regulated by the MYC–HCF-1 interaction have recurred *in vivo*. The extent of gene expression changes observed in regressing tumors suggests a catastrophic collapse occurring in both the 4A and Δ264 cell lines. For 4A, I expect a subset of these were immediately impacted by disrupting the MYC–HCF-1 interaction, with the remainder affected by the consequent initiation of cell death pathways. To take into account the confounding stress/death gene expression changes, I specifically focussed on the transcripts that were also affected *in vitro*. This emphasized many of the recurring gene ontology categories that I have observed in this work, including those related to protein synthesis and mitochondrial function. My findings suggest that even the modest transcriptional consequences of perturbing the MYC–HCF-1 interaction is sufficient to reverse the cancer phenotype.

The development of HCF-1 inhibitors that limit its association with MYC would be the next step for this work. One aspect of this would be to specifically target cells in which MYC is overexpressed, and return its function to baseline. By using the switchable allele system, I have demonstrated the consequences of eliminating the totality of MYC–HCF-1 interaction. The processes that I have characterized as targets are often considered to be gained [74], and this is supported by my hypothesis that MYC overexpression is required for functional relevance of its interaction with HCF-1. If this is the case, the switchable allele is an accurate predictor of the consequences of small molecular inhibitors targeting MYC–HCF-1. Furthermore, the multitude of regulatory capabilities I have attributed to the MYC–HCF-1 interaction and overexpression of MYC are defining characteristics of most tumor types and necessary for rampant cell growth [93, 111]. While MYC is not necessarily the only driver of these processes, it is certainly a recurring, aggressive, and well-defined one. I anticipate then, that the MYC–HCF-1 interaction broadly contributes to the growth of all cancers in which MYC is overexpressed.

Chapter VIII

General Discussion

Here I have performed an extensive characterization of the interaction between the oncoprotein transcription factor MYC and the scaffolding protein HCF-1, including defining the necessary molecular elements, identifying a set of target genes, and finding a contribution to both tumor growth and maintenance. Each of these aspects of understanding I have gained embed the MYC–HCF-1 interaction as a promising target for the development of anti-cancer agents. At the same time, I have also made a number of interesting observations in my experiments that I think warrant further discussion and, ultimately, further investigation. While not necessary to targeting this interaction with small-molecule inhibitors, I believe additional insight into these observations will create more modes through which we can specifically target MYC-driven tumorigenesis through its interaction with HCF-1.

MYC transactivation domain is auto-inhibitory of MbIV–HCF-1 interaction

Based on my finding that the MYC–HCF-1 interaction is influenced by the MYC transactivation domain, I propose that this region of MYC folds back onto the central portion, resulting in an auto-inhibitory conformation that dampens the ability of MbIV to interact with HCF-1. Combined with the knowledge that MYC is O-GlcNAcylated in its transactivation domain [56, 57] and a recent finding that inhibition or knockdown of OGT impairs association of MYC and HCF-1 [330], I additionally suggest that O-GlcNAcylation of the MYC transactivation domain relieves this auto-inhibition to make MbIV more accessible and competent for HCF-1 binding (Figure 8-1). Furthermore, deletion of the MYC transactivation domain or substitution of the MYC HBM with the VP16 HBM is also sufficient to relieve this auto-inhibition, and both of these promote interaction with HCF-1. While O-GlcNAcylation of the MYC transactivation domain enables initial accessibility to MbIV, the interaction of HCF-1 may further stabilize an open conformation of MYC. While intrinsically disordered proteins such as MYC do not adhere to a single conformation, a subset of structures may be preferentially stabilized by protein-

protein interactions [383, 384], favoring additional interactions. Small-angle scattering (SAS) has been used previously to investigate how protein-protein interactions influence structures of intrinsically disordered proteins [385], and this could be applied to the MYC–HCF-1 interaction.

This hypothesis leads directly into another interesting idea: that the efficiency of the MYC–HCF-1 interaction is attuned to the cell's metabolic state. O-GlcNAcylation by OGT requires the substrate UDP-GlcNAc, which is generated through the hexosamine biosynthetic pathway, and is dependent on the levels of glucose and glutamine in the cell (Figure 8-1; [243]). Although we have not tested the MYC–HCF-1 interaction in other species, conservation of the MYC transactivation domain and HBM in vertebrates (Figure 1-5A) would suggest that this nutrient-dependent regulation could also be conserved. The activity of OGT is reported to stabilize MYC [386], likely through competition with the hierarchical phosphorylation that contributes to its turnover, and MYC levels are also reduced during glucose deprivation [387]. I anticipate that OGT-dependent stabilization of MYC is independent of the MYC–HCF-1 interaction, based on my findings that the MYC 4A and VP16 HBM mutants have no overt effect on MYC levels, but that changing availability of glucose to cells would affect both MYC stability and its association with HCF-1. Similarly, the interaction between HCF-1 and ChREBP has previously been shown to be sensitive to glucose levels and OGT activity [249]. However, this has been attributed to O-GlcNAcylation of HCF-1 itself [249].

Finally, the gene set that I identified as direct targets for regulation by the MYC–HCF-1 interaction are primarily involved in ribosome biogenesis and mitochondrial function, the former of which is the most energy intensive process in a cell [93]. Thus, under conditions of nutrient deprivation, the resulting impairment of this interaction would lead to a coordinated reduction in the expression of genes responsible for the processes that utilize [93] and generate the most energy in a cell (Figure 8-1). The mammalian target of rapamycin complex (mTORC) and integrated stress response pathways have both been recognized for their role in inhibiting mRNA translation under such conditions [388], while mTORC is also reported to regulate rDNA transcription as the rate-limiting step in ribosome biogenesis [389].

The gene set regulated by the MYC–HCF-1 is distinct, and would allow for a rapid transcriptional response for longer-term compensation to a changing environment. This model provides relevant detail as to why, at least in this context, the MYC HBM appears to be sub-optimal. The gain-of-function VP16 HBM mutant would likely be devoid of the upstream regulation that I propose modulates the MYC–HCF-1 interaction and its downstream gene targets, thereby preventing a favorable transcriptional program to compensate for nutrient deprivation. Thus, while we did not challenge the tumor growth and maintenance properties of the VP16 HBM mutant, the advantageous behavior of this mutant may not necessarily translate *in vivo*. I additionally propose that, given the appropriate conditions, the enhanced cell growth I demonstrated with the VP16 HBM mutant could also be reproduced with wild-type MYC.

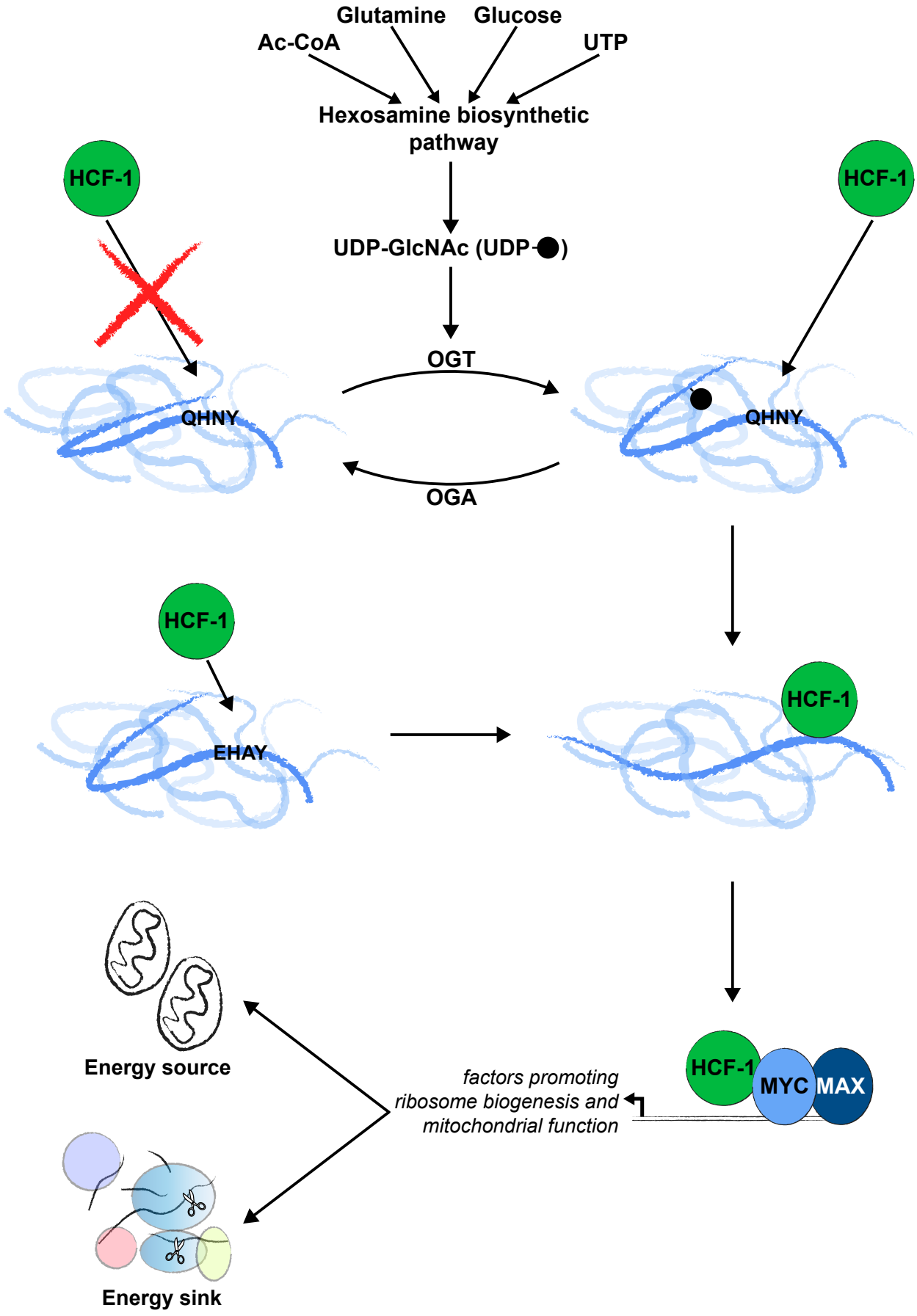

Glutamine
Glucose
Ac-CoA
UTP
Hexosamine biosynthetic pathway
HCF-1
UDP-GlcNAc (UDP-●)
OGT
QHNY
OGA
HCF-1
QHNY
HCF-1
EHAY
HCF-1
Energy source
factors promoting ribosome biogenesis and mitochondrial function
HCF-1
MYC
MAX
Energy sink

Figure 8-1: Modulation of the MYC–HCF-1 interaction by the cellular nutrient state
(Previous page)
Model in which the MYC–HCF-1 interaction is influenced by O-GlcNAcylation of the MYC transactivation domain, which is in-turn dependent on glucose and glutamine levels, and activity of the hexosamine biosynthetic pathway. Deletion of the MYC transactivation domain or substitution of the MYC HBM (QHNY) for the VP16 HBM (EHAY) overcomes the requirement for modulation. By tuning this interaction to the cell's accessibility to nutrients, there can be a coordinated gene expression pattern that forces cell growth to match that supported by its environment.

Co-recruitment-independent transcriptional regulation

My finding that the MYC–HCF-1 interaction regulates the expression of a defined set of target genes, but does not play a role in the recruitment of either protein to chromatin, is somewhat surprising. Recruitment is widely considered the primary mechanism by which protein-protein interactions modulate transcriptional output [360]. This co-recruitment-independent transcriptional regulation has been reported previously, including for HCF-1 [295, 349], but the underlying mechanism is rarely interrogated.

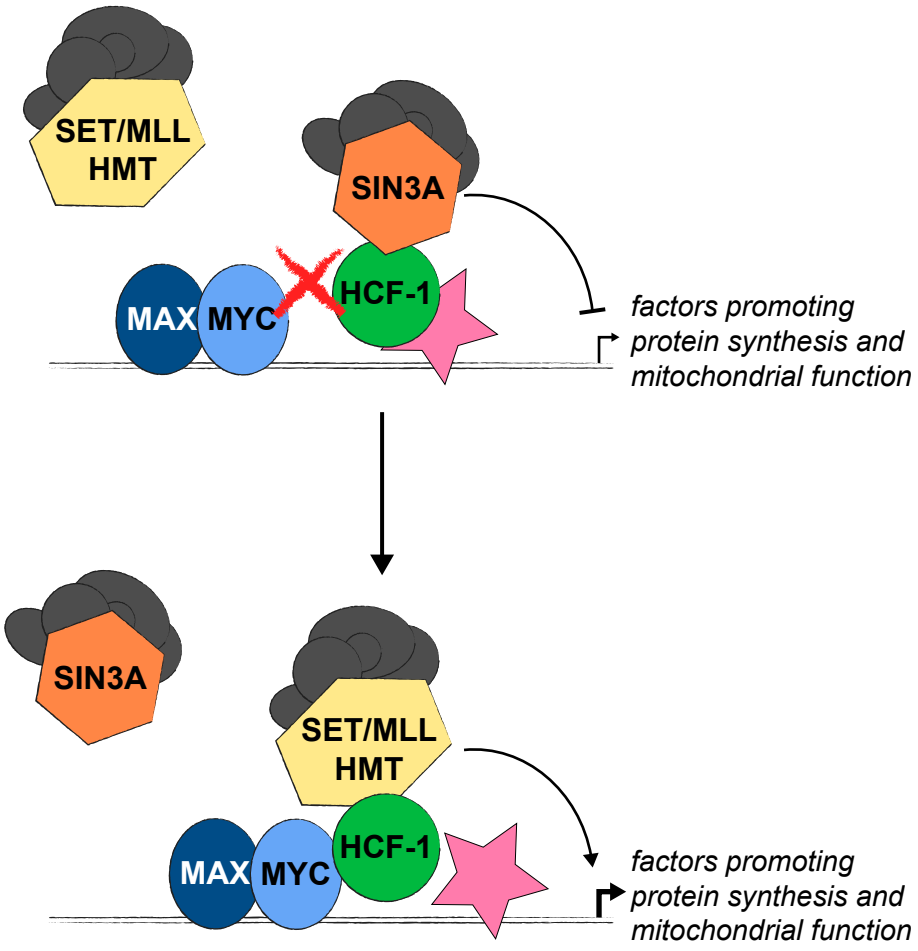
Both MYC and HCF-1 can undergo facilitated recruitment through interaction with other cofactors, including with WDR5 [21, 178] and THAP11 [349], respectively. WDR5-dependent recruitment of MYC to chromatin occurs only at a subset of MYC target genes [21], and only a few of these also correspond to MYC–HCF-1 co-bound sites (Figure 6-4). This suggests that a different MYC interaction partner is involved in its stabilization at HCF-1 binding sites, or that MYC:MAX dimers alone are sufficient for co-localization. On the other hand, recruitment of HCF-1 by THAP family members has not been assessed globally, and this interaction may remain relevant at sites of MYC and HCF-1 co-binding. At MYC–HCF-1 target genes, I hypothesize that for transcriptional activation specifically driven by this interaction there are two necessary steps, neither of which is sufficient on its own. The first step is independent arrival of MYC and HCF-1 to chromatin, as catalyzed by other cofactors. While facilitated recruitment itself may contribute to transcription, a second step consisting of interaction between MYC and HCF-1 amplifies this process post-recruitment. Here, I provide some possible mechanisms by which this second step influences transcriptional output from MYC target genes.

An alternative mode of regulation by the MYC–HCF-1 interaction is stabilization or destabilization of a transcriptional activator or repressor, respectively. As a scaffolding protein, HCF-1 has a multitude of interaction partners, including both activating and repressive transcription factors and histone modifiers. Previous evidence has suggested that the type of complex that forms around HCF-1 is determined by the associated transcription factor. For example, an E2F1–HCF-1 interaction favors binding to the

SETD1A HMT complex, whereas an E2F4–HCF-1 interaction favors binding to the SIN3A HDAC complex [254]. In the context of MYC–HCF-1, their post-recruitment interaction may cause the ejection of SIN3A from, and/or the recruitment of SETD1A to, chromatin (Figure 8-2A). HCF-1 is also associated with deubiquitylating (DUb) enzyme BAP1, which is a member of the polycomb repressive DUb (PR-DUb) complex and contributes to transcriptional activation [390]. Thus, the levels of a number of proteins at chromatin could be affected by perturbing the MYC–HCF-1 interaction, leading to the transcriptional changes I observed.

Similarly, interaction with HCF-1 may affect MYC interaction partners. The flip-side of my proposed auto-inhibitory modulation of the MYC–HCF-1 interaction is that the MYC transactivation domain would be less accessible to the cofactors it utilizes to stimulate transcription. Consistent with this, Kato *et al.* [72] have demonstrated that deletion of exon 3 from *MYC* causes an almost nine-fold increase in MYC-driven transcriptional activation, suggesting that this region, which includes MbIIIb and MbIV, can influence activity of the MYC transactivation domain. I hypothesize that, akin to deletion, association of MbIV with HCF-1 impairs auto-inhibition to promote interaction of MYC with transcriptional activators. In contrast to the MYC central portion, the MYC transactivation domain has been subject to extensive interrogation and found to contribute to activation of transcription through a number of interactions, including with TRRAP for histone acetylation [154], TFIIF for transcriptional initiation/elongation [15], and P-TEFb for RNAP CTD phosphorylation and transcriptional elongation [167, 168]. This then poses the question, how does MYC regulate transcription at sites devoid of HCF-1? The gene expression changes I observed with the MYC 4A and VP16 HBM mutants were modest, at best, and auto-inhibition between MbIV and the MYC transactivation domain is unlikely to be all-or-nothing. This is emphasized by MYC being an intrinsically disordered protein, with certain pro-transcription conformations possibly stabilized by interaction with HCF-1. That said, additional MYC cofactors may associate with its central portion to relieve auto-inhibition. Finally, understanding how MYC influences HCF-1 interaction partners, or vice versa, could be determined using proteomics screens with the loss- and gain-of-function mutants developed in this work.

A: HCF-1-centric



B: MYC-centric

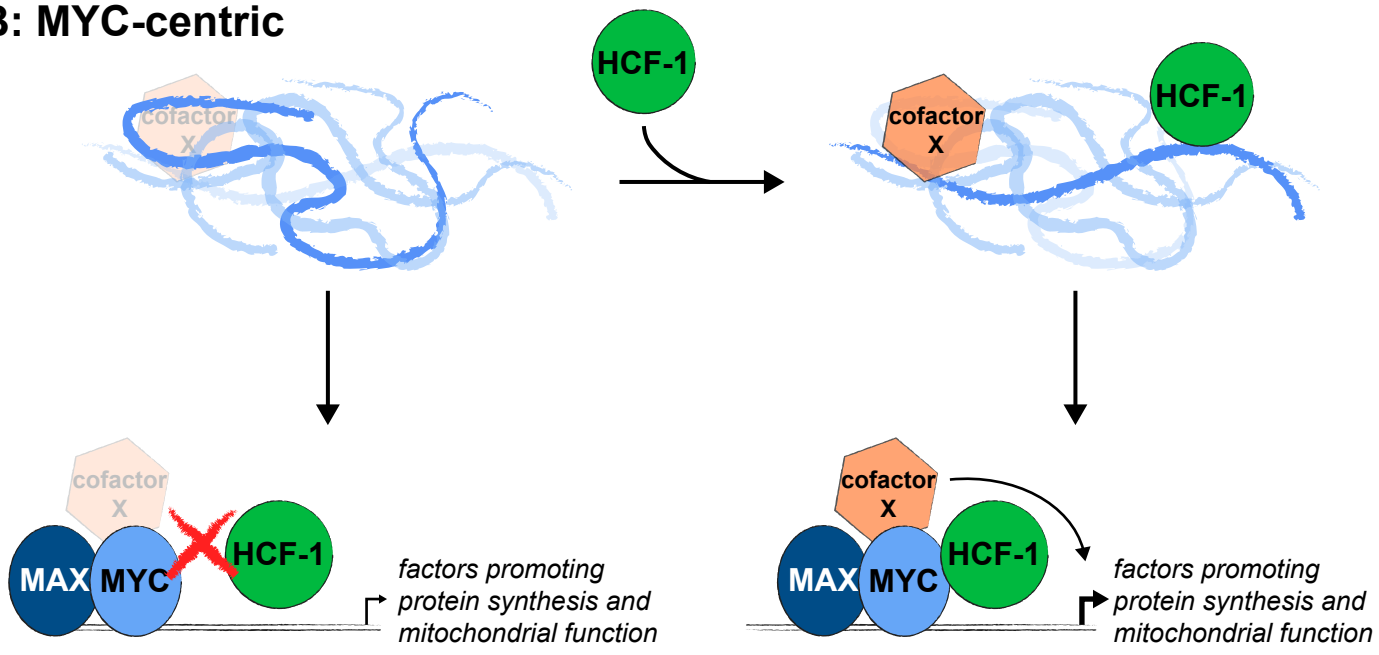


Figure 8-2: Mechanism of co-recruitment-independent transcriptional regulation
(Previous page)
(**A**) HCF-1-centric model whereby, in the absence of an interaction with MYC, HCF-1 contributes to transcriptional repression, possibly through interaction with the histone deacetylase SIN3A. When the MYC–HCF-1 interaction does occur, SIN3A is excluded from chromatin, and replaced by interaction with a transcriptional activator, such as one of the SET/MLL complexes. (**B**) MYC-centric model in which interaction with HCF-1 stabilizes a certain conformation of the intrinsically disordered MYC, thereby favoring interaction with additional MYC cofactors. These cofactors promote transcriptional activation at MYC–HCF-1 target genes.

HCF-1 as a process-specific MYC cofactor

The genes affected when the MYC–HCF-1 interaction is perturbed show a particular enrichment for functions relating to ribosome biogenesis and mitochondrial function, suggesting that HCF-1-dependent transcriptional activation of MYC target genes is primarily confined to certain processes. Indeed, deletion of MbIV has previously demonstrated repression of genes involved in these functions [15], indicating that this interaction regulates the same processes independent of cell type. WDR5 and Miz-1 both also show specificity in the types of genes they regulate in conjunction with MYC, controlling ribosomal protein genes [21] and genes for neuronal projection/differentiation/development [391], respectively.

Despite similarity in the processes modulated by the MYC–WDR5 and MYC–HCF-1 interactions, coexistence of WDR5 and HCF-1 in many histone modifying complexes, and contribution of these interactions to tumor growth and maintenance, the associations of WDR5 and HCF-1 with MYC are independent of each other, and the specific genes they target are distinct. In theory, this separation of function enables separable regulation of ribosomal protein genes by MYC–WDR5 and ribosome biogenesis cofactors by MYC–HCF-1 (Figure 8-3). The latter of these interactions also contributes to expression of the rate-limiting genes in both ribosome biogenesis (POLR1A) and translation initiation (EIF2S1) [363, 364]. Thus, under conditions whereby the MYC–HCF-1 interaction is perturbed by reduced OGT activity, both ribosome biogenesis and translation initiation would be reduced (Figure 8-3). However, I anticipate that the MYC–WDR5 interaction and regulation of ribosomal protein genes would remain intact (Figure 8-3), leaving levels of these proteins unaffected and poised for a rapid recovery when the nutritional state allows.

As MYC levels increase in a cell, the breadth of its target genes is increased through binding to its low-affinity sites on chromatin [74]. High-affinity MYC binding sites typically correspond to a subset of canonical E-box-containing promoters, whereas low-affinity sites include non-canonical E-boxes [74], such as the NRF1 motif strongly over-represented in MYC–HCF-1 co-binding. Many of these low-

affinity sites are linked to metabolism and mitochondrial function [25], meaning that the processes controlled by MYC expands with overexpression and enable cells to keep up with demand [74]. While we have identified target genes of the MYC–HCF-1 interaction in the context of MYC overexpression, whether the associated processes remain relevant when MYC levels are lower is unclear. This is further emphasized by the apparently tempered nature of the default MYC–HCF-1 interaction, which may require high MYC levels for their association to occur in any productive manner. Indeed, this interaction may be most important when MYC is overexpressed (Figure 8-4A), such as during development or cancer.

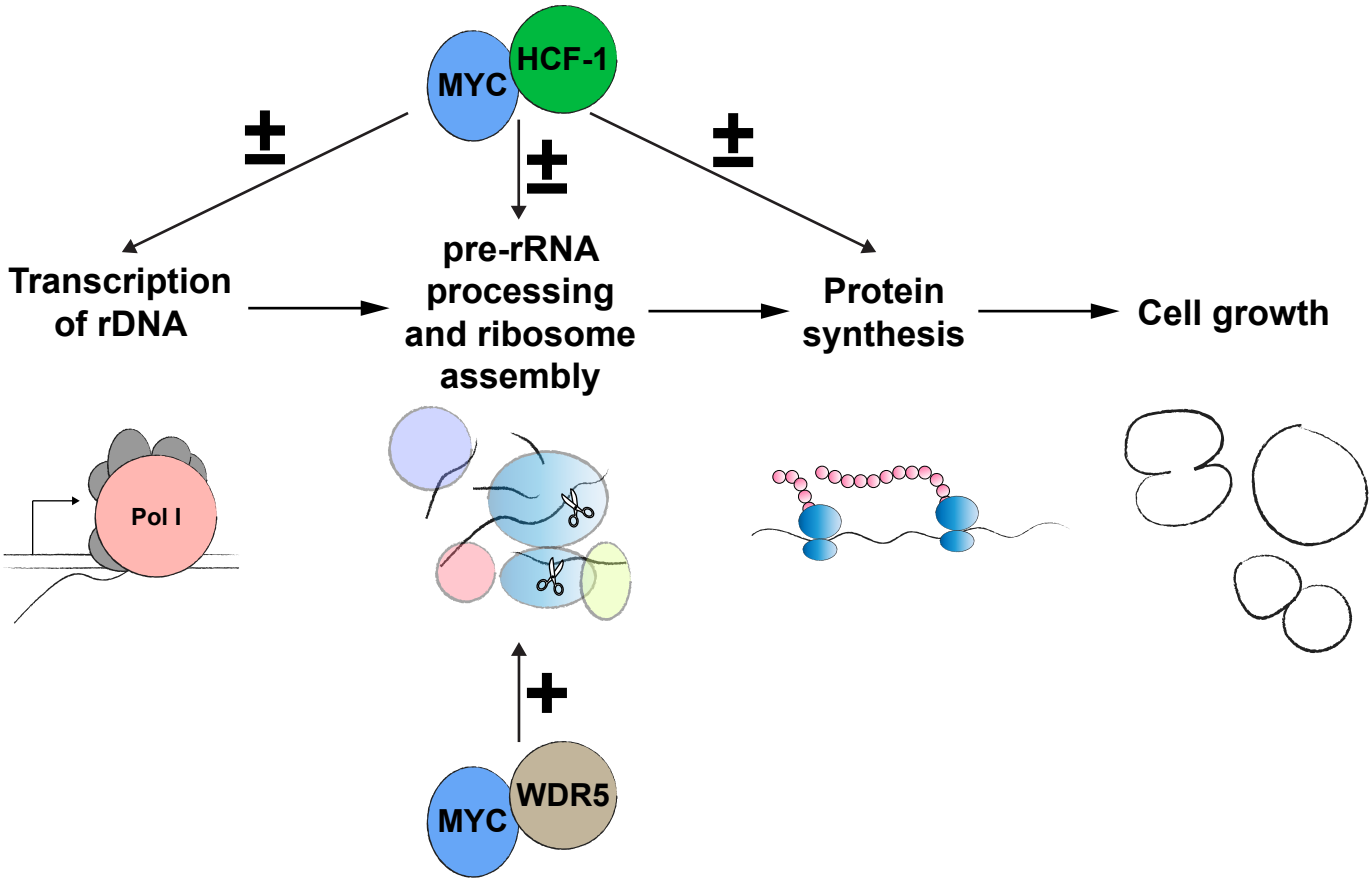


Figure 8-3: Coordinated regulation of ribosome biogenesis
The MYC–HCF-1 interaction controls the expression of genes involved in all stages of ribosome biogenesis and protein synthesis, including the rate-limiting transcription of ribosomal DNA (rDNA) and initiation of translation at the ribosome. Adjustment of this interaction in response to the cellular metabolic state means that these processes are altered accordingly. However, the independence of the MYC–WDR5 interaction means that production of ribosomal proteins is unaffected, leaving levels of these necessary components high and available for rapid induction of ribosome biogenesis when conditions allow.

The MYC–HCF-1 interaction as an anti-cancer target

The potential of targeting MYC directly is limited by its intrinsically disordered structure. By instead interrogating MYC cofactors, we can create opportunities to indirectly disrupt MYC function. Here, I have pursued HCF-1 for this purpose, presenting definitive evidence for the necessity of the MYC–HCF-1 interaction in tumor growth and maintenance. Whilst this in itself is an exciting step towards the development of small-molecule inhibitors against MYC function, my work has additionally revealed a number of insights into the MYC–HCF-1 interaction that elevates it as a possible candidate.

The MYC HBM is responsible for a relatively weak, but possibly tunable, interaction with HCF-1 that is perhaps more likely to have functional relevance when MYC is overexpressed, both due to the nature of their interaction and their target genes. If we assume that baseline function of HCF-1 is largely independent of MYC and instead relies on its other interactions, I propose that the contribution of MYC in the context of this interaction could be disabled using a small-molecule inhibitor with a relatively low binding strength, whilst leaving the majority of other HCF-1 cofactors and function intact (Figure 8-4B). Additionally, as mentioned above, the gene set I have established to be modulated by this interaction, as well as the interaction itself, may be a consequence of high levels of MYC (Figure 8-4A). And, as MYC is overexpressed in most cancers, this interaction may prove to be a therapeutic window for targeting MYC-driven cancers.

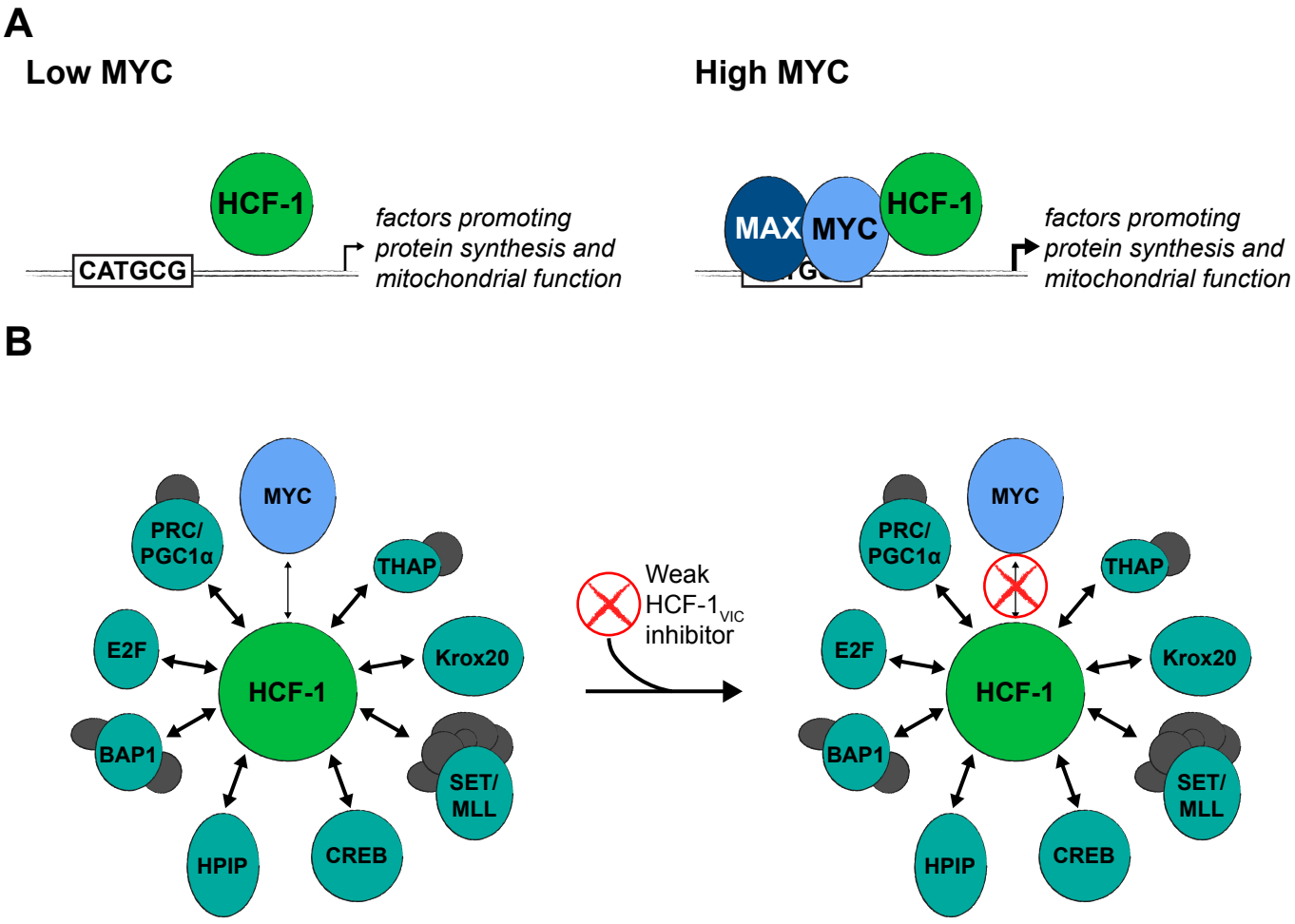


Figure 8-4: MYC–HCF-1 interaction is most relevant when MYC levels are high
(**A**) As MYC levels in a cell increase, its chromatin binding pattern changes as it becomes associated with lower affinity binding sites, such as the NRF1 CATGCG motif. MYC overexpression may influence interaction with HCF-1 at two levels: first, the MYC transactivation domain-weakened interaction may be more likely to force a functionally relevant association with HCF-1 when MYC is overexpressed; second, co-localization of MYC and HCF-1 at non-canonical E-boxes on chromatin is more likely to occur when MYC is overexpressed, meaning that regulation of the MYC–HCF-1 target genes by this interaction primarily occurs when MYC levels are high. (**B**) The above aspects of the MYC–HCF-1 interaction create a therapeutic window in which their association could be impaired specifically when MYC is overexpressed, and use of a weak HCF-1_{VIC} domain inhibitor enables prioritized disruption of the interaction with MYC due to its tempered association. Thus, the majority of HCF-1 function and any baseline function of the MYC–HCF-1 interaction would be retained.

Future directions

Within this chapter, I have mentioned in passing the direction that I would like to see research into the MYC–HCF-1 follow, but I will additionally summarize these here. Firstly, determining whether the MYC–HCF-1 interaction is indeed regulated by nutrient status and OGT can be addressed rather simply in tissue culture by modifying the levels of glucose and/or glutamine in the media, and performing a co-immunoprecipitation experiment. I have also created a Ramos cell line for inducible degradation of OGT to confirm the findings of Itkonen *et al.* [330] in this context. This is particularly interesting because it is an added vulnerability by which the MYC–HCF-1 interaction could be targeted using small-molecule inhibitors, including against OGT itself or further upstream in the synthesis of UDP-GlcNAc.

How its association with HCF-1 influences MYC-driven transcription, on a mechanistic and molecular level, will likely be more difficult to decipher, but a preliminary understanding could be obtained using proteomics of MYC or HCF-1 co-immunoprecipitates in the context of the MYC 4A or VP16 HBM mutants. In this discussion I proposed both HCF-1- and MYC-centric models of how this interaction may favor the formation of specific trimeric complexes, with the third protein interacting either through HCF-1 or through MYC. While many cofactors have been examined for their role in promoting MYC-driven transcriptional activation, there are also a slew of associated proteins that are not understood to any degree [15]. By addressing the formation of a MYC–HCF-1 trimeric complex, we may simultaneously determine the contribution of these cofactors to MYC function and tumorigenesis.

To determine the baseline function, including extent of co-binding on chromatin and direct target genes, of the MYC–HCF-1 interaction in the absence of MYC overexpression, we would need to utilize the switchable *MYC* allele in a different cell line. Indeed, this type of interrogation could be enhanced using an inducible expression system, such as Tet-On, as the consequences of low and high MYC levels could be addressed in the same context, and ideally in the absence of endogenous MYC. The specific goal of this experimental approach would be to determine if the MYC–HCF-1 interaction is still responsible for regulating the expression of genes involved in ribosome biogenesis, translation, and

mitochondrial function, as I have identified here, and to determine if overexpression of MYC would be provide therapeutic specificity for targeting this interaction.

Finally, for the development of small-molecule inhibitors against the HCF-1_{VIC} domain, we must first procure a structure of this region. This depends on expression of folded and functional recombinant HCF-1_{VIC}, which has not been achieved in the thirty years since HCF-1 was first discovered. The recent development of chaperone systems to promote protein folding and solubility in *E. coli*, (for example, [392]) may provide reprieve in this process. However, alternative systems that rely on eukaryotes, such as baculovirus, may instead be a necessary approach, particularly with the post-translational modifications we know afflict HCF-1.

References

1. Tansey, W.P., *Mammalian MYC Proteins and Cancer*. New Journal of Science, 2014. **2014**: p. 1-27.
2. Vennstrom, B., Sheiness, D., Zabielski, J., and Bishop, J.M., *Isolation and characterization of c-myc, a cellular homolog of the oncogene (v-myc) of avian myelocytomatosis virus strain 29*. J Virol, 1982. **42**(3): p. 773-9.
3. Atchley, W.R. and Fitch, W.M., *Myc and Max: molecular evolution of a family of proto-oncogene products and their dimerization partner*. Proc Natl Acad Sci U S A, 1995. **92**(22): p. 10217-21.
4. Gallant, P., Myc/Max/Mad in invertebrates: the evolution of the Max network. Curr Top Microbiol Immunol, 2006. **302**: p. 235-53.
5. Young, S.L., Diolaiti, D., Conacci-Sorrell, M., Ruiz-Trillo, I., Eisenman, R.N., and King, N., *Premetazoan ancestry of the Myc-Max network.* Mol Biol Evol, 2011. **28**(10): p. 2961-71.
6. Charron, J., Malynn, B.A., Fisher, P., Stewart, V., Jeannotte, L., Goff, S.P., Robertson, E.J., and Alt, F.W., *Embryonic lethality in mice homozygous for a targeted disruption of the N-myc gene.* Genes Dev, 1992. **6**(12A): p. 2248-57.
7. Davis, A.C., Wims, M., Spotts, G.D., Hann, S.R., and Bradley, A., *A null c-myc mutation causes lethality before 10.5 days of gestation in homozygotes and reduced fertility in heterozygous female mice*. Genes Dev, 1993. **7**(4): p. 671-82.
8. Hatton, K.S., Mahon, K., Chin, L., Chiu, F.C., Lee, H.W., Peng, D., Morgenbesser, S.D., Horner, J., and DePinho, R.A., *Expression and activity of L-Myc in normal mouse development.* Mol Cell Biol, 1996. **16**(4): p. 1794-804.
9. Malynn, B.A., de Alboran, I.M., O'Hagan, R.C., Bronson, R., Davidson, L., DePinho, R.A., and Alt, F.W., *N-myc can functionally replace c-myc in murine development, cellular growth, and differentiation.* Genes Dev, 2000. **14**(11): p. 1390-9.
10. Marcu, K.B., Bossone, S.A., and Patel, A.J., *myc function and regulation.* Annu Rev Biochem, 1992. **61**: p. 809-60.
11. Chau, K.F., Shannon, M.L., Fame, R.M., Fonseca, E., Mullan, H., Johnson, M.B., Sendamarai, A.K., Springel, M.W., Laurent, B., and Lehtinen, M.K., *Downregulation of ribosome biogenesis during early forebrain development.* Elife, 2018. **7**.
12. DePinho, R.A., Hatton, K.S., Tesfaye, A., Yancopoulos, G.D., and Alt, F.W., The human myc gene family: structure and activity of L-myc and an L-myc pseudogene. Genes Dev, 1987. **1**(10): p. 1311-26.
13. Baluapuri, A., Wolf, E., and Eilers, M., Target gene-independent functions of MYC oncoproteins. Nat Rev Mol Cell Biol, 2020.
14. Herbst, A., Hemann, M.T., Tworkowski, K.A., Salghetti, S.E., Lowe, S.W., and Tansey, W.P., *A conserved element in Myc that negatively regulates its proapoptotic activity.* EMBO Rep, 2005. **6**(2): p. 177-83.
15. Kalkat, M., Resetca, D., Lourenco, C., Chan, P.K., Wei, Y., Shiah, Y.J., Vitkin, N., Tong, Y., Sunnerhagen, M., Done, S.J., Boutros, P.C., Raught, B., and Penn, L.Z., *MYC Protein Interactome Profiling Reveals Functionally Distinct Regions that Cooperate to Drive Tumorigenesis.* Mol Cell, 2018. **72**(5): p. 836-848 e7.
16. De Greve, J., Battey, J., Fedorko, J., Birrer, M., Evan, G., Kaye, F., Sausville, E., and Minna, J., *The human L-myc gene encodes multiple nuclear phosphoproteins from alternatively processed mRNAs.* Mol Cell Biol, 1988. **8**(10): p. 4381-8.
17. Blackwood, E.M. and Eisenman, R.N., *Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc*. Science, 1991. **251**(4998): p. 1211-7.
18. Blackwell, T.K., Huang, J., Ma, A., Kretzner, L., Alt, F.W., Eisenman, R.N., and Weintraub, H., *Binding of myc proteins to canonical and noncanonical DNA sequences.* Mol Cell Biol, 1993. **13**(9): p. 5216-24.

19. Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L., and Amati, B., *Myc-binding-site recognition in the human genome is determined by chromatin context.* Nat Cell Biol, 2006. **8**(7): p. 764-70.
20. Guo, J., Li, T., Schipper, J., Nilson, K.A., Fordjour, F.K., Cooper, J.J., Gordan, R., and Price, D.H., *Sequence specificity incompletely defines the genome-wide occupancy of Myc.* Genome Biol, 2014. **15**(10): p. 482.
21. Thomas, L.R., Adams, C.M., Wang, J., Weissmiller, A.M., Creighton, J., Lorey, S.L., Liu, Q., Fesik, S.W., Eischen, C.M., and Tansey, W.P., *Interaction of the oncoprotein transcription factor MYC with its chromatin cofactor WDR5 is essential for tumor maintenance.* Proc Natl Acad Sci U S A, 2019. **116**(50): p. 25260-25268.
22. Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., and Lopez, R., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic Acids Res, 2019. **47**(W1): p. W636-W641.
23. Capra, J.A. and Singh, M., *Predicting functionally important residues from sequence conservation*. Bioinformatics, 2007. **23**(15): p. 1875-82.
24. Vita, M. and Henriksson, M., *The Myc oncoprotein as a therapeutic target for human cancer*. Seminars in Cancer Biology, 2006. **16**: p. 318-330.
25. Dang, C.V., MYC on the path to cancer. Cell, 2012. **149**(1): p. 22-35.
26. Strieder, V. and Lutz, W., *Regulation of N-myc expression in development and disease*. Cancer Lett, 2002. **180**(2): p. 107-19.
27. Nesbit, C.E., Tersak, J.M., and Prochownik, E.V., *MYC oncogenes and human neoplastic disease.* Oncogene, 1999. **18**(19): p. 3004-16.
28. Barrett, J., Birrer, M.J., Kato, G.J., Dosaka-Akita, H., and Dang, C.V., *Activation domains of L-Myc and c-Myc determine their transforming potencies in rat embryo cells*. Mol Cell Biol, 1992. **12**(7): p. 3130-7.
29. Boxer, L.M. and Dang, C.V., *Translocations involving c-myc and c-myc function*. Oncogene, 2001. **20**(40): p. 5595-610.
30. Janz, S., *Myc translocations in B cell and plasma cell neoplasms*. DNA Repair (Amst), 2006. **5**(9-10): p. 1213-24.
31. Bemark, M. and Neuberger, M.S., *The c-MYC allele that is translocated into the IgH locus undergoes constitutive hypermutation in a Burkitt's lymphoma line*. Oncogene, 2000. **19**(30): p. 3404-10.
32. Bhatia, K., Huppi, K., Spangler, G., Siwarski, D., Iyer, R., and Magrath, I., *Point mutations in the c-Myc transactivation domain are common in Burkitt's lymphoma and mouse plasmacytomas*. Nat Genet, 1993. **5**(1): p. 56-61.
33. Hemann, M.T., Bric, A., Teruya-Feldstein, J., Herbst, A., Nilsson, J.A., Cordon-Cardo, C., Cleveland, J.L., Tansey, W.P., and Lowe, S.W., *Evasion of the p53 tumour surveillance network by tumour-derived MYC mutants.* Nature, 2005. **436**(7052): p. 807-11.
34. Gregory, M.A. and Hann, S.R., *c-Myc proteolysis by the ubiquitin-proteasome pathway: stabilization of c-Myc in Burkitt's lymphoma cells*. Mol Cell Biol, 2000. **20**(7): p. 2423-35.
35. Levens, D., You Don't Muck with MYC. Genes Cancer, 2010. **1**(6): p. 547-554.
36. Oswald, F., Lovec, H., Moroy, T., and Lipp, M., *E2F-dependent regulation of human MYC: trans-activation by cyclins D1 and A overrides tumour suppressor protein functions*. Oncogene, 1994. **9**(7): p. 2029-36.
37. Delmore, J.E., Issa, G.C., Lemieux, M.E., Rahl, P.B., Shi, J., Jacobs, H.M., Kastritis, E., Gilpatrick, T., Paranal, R.M., Qi, J., Chesi, M., Schinzel, A.C., McKeown, M.R., Heffernan, T.P., Vakoc, C.R., Bergsagel, P.L., Ghobrial, I.M., Richardson, P.G., Young, R.A., Hahn, W.C., Anderson, K.C., Kung, A.L., Bradner, J.E., and Mitsiades, C.S., *BET bromodomain inhibition as a therapeutic strategy to target c-Myc.* Cell, 2011. **146**(6): p. 904-17.
38. Dani, C., Blanchard, J.M., Piechaczyk, M., El Sabouty, S., Marty, L., and Jeanteur, P., *Extreme instability of myc mRNA in normal and transformed human cells.* Proc Natl Acad Sci U S A, 1984. **81**(22): p. 7046-50.

39. Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L., *eIF4E is a central node of an RNA regulon that governs cellular proliferation.* J Cell Biol, 2006. **175**(3): p. 415-26.
40. Mazan-Mamczarz, K., Lal, A., Martindale, J.L., Kawai, T., and Gorospe, M., *Translational repression by RNA-binding protein TIAR.* Mol Cell Biol, 2006. **26**(7): p. 2716-27.
41. Rabbitts, P.H., Watson, J.V., Lamond, A., Forster, A., Stinson, M.A., Evan, G., Fischer, W., Atherton, E., Sheppard, R., and Rabbitts, T.H., *Metabolism of c-myc gene products: c-myc mRNA and protein expression in the cell cycle.* EMBO J, 1985. **4**(8): p. 2009-15.
42. Lutterbach, B. and Hann, S.R., *Hierarchical phosphorylation at N-terminal transformation-sensitive sites in c-Myc protein is regulated by mitogens and in mitosis*. Mol Cell Biol, 1994. **14**(8): p. 5510-22.
43. Welcker, M., Orian, A., Jin, J., Grim, J.E., Harper, J.W., Eisenman, R.N., and Clurman, B.E., *The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation*. Proc Natl Acad Sci U S A, 2004. **101**(24): p. 9085-90.
44. Salghetti, S.E., Kim, S.Y., and Tansey, W.P., *Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc*. EMBO J, 1999. **18**(3): p. 717-26.
45. Thomas, L.R. and Tansey, W.P., *Proteolytic control of the oncoprotein transcription factor Myc*. Adv Cancer Res, 2011. **110**: p. 77-106.
46. Sears, R., Nuckolls, F., Haura, E., Taya, Y., Tamai, K., and Nevins, J.R., *Multiple Ras-dependent phosphorylation pathways regulate Myc protein stability.* Genes Dev, 2000. **14**(19): p. 2501-14.
47. Gregory, M.A., Qi, Y., and Hann, S.R., *Phosphorylation by glycogen synthase kinase-3 controls c-myc proteolysis and subnuclear localization*. J Biol Chem, 2003. **278**(51): p. 51606-12.
48. Helander, S., Montecchio, M., Pilstal, R., Su, Y., Kuruvilla, J., Elven, M., Ziauddin, J.M.E., Anandapadamanaban, M., Cristobal, S., Lundstrom, P., Sears, R.C., Wallner, B., and Sunnerhagen, M., *Pre-Anchoring of Pin1 to Unphosphorylated c-Myc in a Fuzzy Complex Regulates c-Myc Activity.* Structure, 2015. **23**(12): p. 2267-2279.
49. Farrell, A.S., Pelz, C., Wang, X., Daniel, C.J., Wang, Z., Su, Y., Janghorban, M., Zhang, X., Morgan, C., Impey, S., and Sears, R.C., *Pin1 regulates the dynamics of c-Myc DNA binding to facilitate target gene regulation and oncogenesis.* Mol Cell Biol, 2013. **33**(15): p. 2930-49.
50. Yeh, E., Cunningham, M., Arnold, H., Chasse, D., Monteith, T., Ivaldi, G., Hahn, W.C., Stukenberg, P.T., Shenolikar, S., Uchida, T., Counter, C.M., Nevins, J.R., Means, A.R., and Sears, R., *A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells.* Nat Cell Biol, 2004. **6**(4): p. 308-18.
51. Arnold, H.K. and Sears, R.C., *Protein phosphatase 2A regulatory subunit B56alpha associates with c-myc and negatively regulates c-myc accumulation*. Mol Cell Biol, 2006. **26**(7): p. 2832-44.
52. Yada, M., Hatakeyama, S., Kamura, T., Nishiyama, M., Tsunematsu, R., Imaki, H., Ishida, N., Okumura, F., Nakayama, K., and Nakayama, K.I., *Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7*. EMBO J, 2004. **23**(10): p. 2116-25.
53. Allen-Petersen, B.L. and Sears, R.C., *Mission Possible: Advances in MYC Therapeutic Targeting in Cancer*. BioDrugs, 2019. **33**(5): p. 539-553.
54. Devaiah, B.N., Mu, J., Akman, B., Uppal, S., Weissman, J.D., Cheng, D., Baranello, L., Nie, Z., Levens, D., and Singer, D.S., *MYC protein stability is negatively regulated by BRD4.* Proc Natl Acad Sci U S A, 2020. **117**(24): p. 13457-13467.
55. Kim, S.Y., Herbst, A., Tworkowski, K.A., Salghetti, S.E., and Tansey, W.P., *Skp2 regulates Myc protein stability and activity.* Mol Cell, 2003. **11**(5): p. 1177-88.
56. Chou, T.Y., Dang, C.V., and Hart, G.W., *Glycosylation of the c-Myc transactivation domain.* Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**: p. 4417-4421.
57. Chou, T.-Y., Hart, G.W., and Dang, C.V., *c-Myc Is Glycosylated at Threonine 58, a Known Phosphorylation Site and a Mutational Hot Spot in Lymphomas*. Journal of Biological Chemistry, 1995. **270**: p. 18961-18965.

58. Itkonen, H.M., Minner, S., Guldvik, I.J., Sandmann, M.J., Tsourlakis, M.C., Berge, V., Svindland, A., Schlomm, T., and Mills, I.G., *O-GlcNAc Transferase Integrates Metabolic Pathways to Regulate the Stability of c-MYC in Human Prostate Cancer Cells.* Cancer Research, 2013. **73**: p. 5277-5287.
59. Gabriel, S.S. and Kallies, A., *Glucose- and glutamine-fueled stabilization of C-Myc is required for T-cell proliferation and malignant transformation*. Cell Death Discov, 2016. **2**: p. 16047.
60. Xu, D., Wang, W., Bian, T., Yang, W., Shao, M., and Yang, H., *Increased expression of O-GlcNAc transferase (OGT) is a biomarker for poor prognosis and allows tumorigenesis and invasion in colon cancer*. Int J Clin Exp Pathol, 2019. **12**(4): p. 1305-1314.
61. Akella, N.M., Le Minh, G., Ciraku, L., Mukherjee, A., Bacigalupa, Z.A., Mukhopadhyay, D., Sodi, V.L., and Reginato, M.J., *O-GlcNAc Transferase Regulates Cancer Stem-like Potential of Breast Cancer Cells.* Mol Cancer Res, 2020. **18**(4): p. 585-598.
62. Otto, T., Horn, S., Brockmann, M., Eilers, U., Schuttrumpf, L., Popov, N., Kenney, A.M., Schulte, J.H., Beijersbergen, R., Christiansen, H., Berwanger, B., and Eilers, M., *Stabilization of N-Myc is a critical function of Aurora A in human neuroblastoma.* Cancer Cell, 2009. **15**(1): p. 67-78.
63. Brockmann, M., Poon, E., Berry, T., Carstensen, A., Deubzer, H.E., Rycak, L., Jamin, Y., Thway, K., Robinson, S.P., Roels, F., Witt, O., Fischer, M., Chesler, L., and Eilers, M., *Small Molecule Inhibitors of Aurora-A Induce Proteasomal Degradation of N-Myc in Childhood Neuroblastoma.* Cancer Cell, 2013. **24**(1): p. 75-89.
64. Richards, M.W., Burgess, S.G., Poon, E., Carstensen, A., Eilers, M., Chesler, L., and Bayliss, R., *Structural basis of N-Myc binding by Aurora-A and its destabilization by kinase inhibitors.* Proc Natl Acad Sci U S A, 2016. **113**(48): p. 13726-13731.
65. Gustafson, W.C., Meyerowitz, J.G., Nekritz, E.A., Chen, J., Benes, C., Charron, E., Simonds, E.F., Seeger, R., Matthay, K.K., Hertz, N.T., Eilers, M., Shokat, K.M., and Weiss, W.A., *Drugging MYCN through an allosteric transition in Aurora kinase A.* Cancer Cell, 2014. **26**(3): p. 414-427.
66. Dauch, D., Rudalska, R., Cossa, G., Nault, J.C., Kang, T.W., Wuestefeld, T., Hohmeyer, A., Imbeaud, S., Yevsa, T., Hoenicke, L., Pantsar, T., Bozko, P., Malek, N.P., Longerich, T., Laufer, S., Poso, A., Zucman-Rossi, J., Eilers, M., and Zender, L., *A MYC-aurora kinase A protein complex represents an actionable drug target in p53-altered liver cancer.* Nat Med, 2016. **22**(7): p. 744-53.
67. Jiang, J., Wang, J., Yue, M., Cai, X., Wang, T., Wu, C., Su, H., Wang, Y., Han, M., Zhang, Y., Zhu, X., Jiang, P., Li, P., Sun, Y., Xiao, W., Feng, H., Qing, G., and Liu, H., *Direct Phosphorylation and Stabilization of MYC by Aurora B Kinase Promote T-cell Leukemogenesis.* Cancer Cell, 2020. **37**(2): p. 200-215 e5.
68. Herbst, A., Salghetti, S.E., Kim, S.Y., and Tansey, W.P., *Multiple cell-type-specific elements regulate Myc protein stability*. Oncogene, 2004. **23**(21): p. 3863-71.
69. Dang, C.V., O'Donnell, K.A., Zeller, K.I., Nguyen, T., Osthus, R.C., and Li, F., *The c-Myc target gene network.* Semin Cancer Biol, 2006. **16**(4): p. 253-64.
70. Kretzner, L., Blackwood, E.M., and Eisenman, R.N., *Myc and Max proteins possess distinct transcriptional activities*. Nature, 1992. **359**(6394): p. 426-9.
71. Eisenman, R.N., *Deconstructing myc.* Genes Dev, 2001. **15**(16): p. 2023-30.
72. Kato, G.J., Barrett, J., Villa-Garcia, M., and Dang, C.V., *An amino-terminal c-myc domain required for neoplastic transformation activates transcription*. Mol Cell Biol, 1990. **10**(11): p. 5914-20.
73. Levens, D., Cellular MYCro economics: *Balancing MYC function with MYC expression*. Cold Spring Harb Perspect Med, 2013. **3**(11).
74. Lorenzin, F., Benary, U., Baluapuri, A., Walz, S., Jung, L.A., von Eyss, B., Kisker, C., Wolf, J., Eilers, M., and Wolf, E., *Different promoter affinities account for specificity in MYC-dependent gene regulation.* eLife, 2016. **5**: p. 1009-1019.
75. Sabo, A. and Amati, B., *Genome recognition by MYC.* Cold Spring Harb Perspect Med, 2014. **4**(2).

76. Lin, C.Y., Loven, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A., *Transcriptional amplification in tumor cells with elevated c-Myc.* Cell, 2012. **151**(1): p. 56-67.
77. Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., Zhao, K., and Levens, D., *c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells.* Cell, 2012. **151**(1): p. 68-79.
78. Blackwell, T.K., Kretzner, L., Blackwood, E.M., Eisenman, R.N., and Weintraub, H., *Sequence-specific DNA binding by the c-Myc protein.* Science, 1990. **250**(4984): p. 1149-51.
79. Wolf, E., Lin, C.Y., Eilers, M., and Levens, D.L., *Taming of the beast: shaping Myc-dependent amplification*. Trends Cell Biol, 2015. **25**(4): p. 241-8.
80. Luscher, B. and Larsson, L.G., *The basic region/helix-loop-helix/leucine zipper domain of Myc proto-oncoproteins: function and regulation*. Oncogene, 1999. **18**(19): p. 2955-66.
81. Bhawe, K. and Roy, D., *Interplay between NRF1, E2F4 and MYC transcription factors regulating common target genes contributes to cancer development and progression*. Cell Oncol (Dordr), 2018. **41**(5): p. 465-484.
82. Jung, L.A., Gebhardt, A., Koelmel, W., Ade, C.P., Walz, S., Kuper, J., von Eyss, B., Letschert, S., Redel, C., d'Artista, L., Biankin, A., Zender, L., Sauer, M., Wolf, E., Evan, G., Kisker, C., and Eilers, M., *OmoMYC blunts promoter invasion by oncogenic MYC to inhibit gene expression characteristic of MYC-dependent tumors.* Oncogene, 2016.
83. Sabo, A., Kress, T.R., Pelizzola, M., de Pretis, S., Gorski, M.M., Tesi, A., Morelli, M.J., Bora, P., Doni, M., Verrecchia, A., Tonelli, C., Faga, G., Bianchi, V., Ronchi, A., Low, D., Muller, H., Guccione, E., Campaner, S., and Amati, B., *Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis.* Nature, 2014. **511**(7510): p. 488-492.
84. Ji, H., Wu, G., Zhan, X., Nolan, A., Koh, C., De Marzo, A., Doan, H.M., Fan, J., Cheadle, C., Fallahi, M., Cleveland, J.L., Dang, C.V., and Zeller, K.I., *Cell-type independent MYC target genes reveal a primordial signature involved in biomass accumulation.* PLoS One, 2011. **6**(10): p. e26057.
85. van Riggelen, J., Yetil, A., and Felsher, D.W., *MYC as a regulator of ribosome biogenesis and protein synthesis*. Nat Rev Cancer, 2010. **10**(4): p. 301-9.
86. Lindstrom, M.S., Jurada, D., Bursac, S., Orsolic, I., Bartek, J., and Volarevic, S., *Nucleolus as an emerging hub in maintenance of genome stability and cancer pathogenesis.* Oncogene, 2018. **37**(18): p. 2351-2366.
87. Grandori, C., Gomez-Roman, N., Felton-Edkins, Z.A., Ngouenet, C., Galloway, D.A., Eisenman, R.N., and White, R.J., *c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I.* Nat Cell Biol, 2005. **7**(3): p. 311-8.
88. Gomez-Roman, N., Grandori, C., Eisenman, R.N., and White, R.J., *Direct activation of RNA polymerase III transcription by c-Myc.* Nature, 2003. **421**(6920): p. 290-4.
89. Schlosser, I., Holzel, M., Murnseer, M., Burtscher, H., Weidle, U.H., and Eick, D., *A role for c-Myc in the regulation of ribosomal RNA processing.* Nucleic Acids Res, 2003. **31**(21): p. 6148-56.
90. Rosenwald, I.B., Rhoads, D.B., Callanan, L.D., Isselbacher, K.J., and Schmidt, E.V., *Increased expression of eukaryotic translation initiation factors eIF-4E and eIF-2 alpha in response to growth induction by c-myc*. Proc Natl Acad Sci U S A, 1993. **90**(13): p. 6175-8.
91. Iritani, B.M. and Eisenman, R.N., *c-Myc enhances protein synthesis and cell size during B lymphocyte development*. Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13180-5.
92. Stine, Z.E., Walton, Z.E., Altman, B.J., Hsieh, A.L., and Dang, C.V., *MYC, Metabolism, and Cancer.* Cancer Discov, 2015. **5**(10): p. 1024-39.
93. Pelletier, J., Thomas, G., and Volarevic, S., *Ribosome biogenesis in cancer: new players and therapeutic avenues*. Nat Rev Cancer, 2018. **18**(1): p. 51-63.
94. Warburg, O., On respiratory impairment in cancer cells. Science, 1956. **124**(3215): p. 269-70.
95. Kim, J.W. and Dang, C.V., *Cancer's molecular sweet tooth and the Warburg effect*. Cancer Res, 2006. **66**(18): p. 8927-30.
96. Vander Heiden, M.G., Cantley, L.C., and Thompson, C.B., *Understanding the Warburg effect: the metabolic requirements of cell proliferation*. Science, 2009. **324**(5930): p. 1029-33.

97. Goetzman, E.S. and Prochownik, E.V., *The Role for Myc in Coordinating Glycolysis, Oxidative Phosphorylation, Glutaminolysis, and Fatty Acid Metabolism in Normal and Neoplastic Tissues*. Front Endocrinol (Lausanne), 2018. **9**: p. 129.
98. Osthus, R.C., Shim, H., Kim, S., Li, Q., Reddy, R., Mukherjee, M., Xu, Y., Wonsey, D., Lee, L.A., and Dang, C.V., *Deregulation of glucose transporter 1 and glycolytic gene expression by c-Myc.* J Biol Chem, 2000. **275**(29): p. 21797-800.
99. Li, F., Wang, Y., Zeller, K.I., Potter, J.J., Wonsey, D.R., O'Donnell, K.A., Kim, J.W., Yustein, J.T., Lee, L.A., and Dang, C.V., *Myc stimulates nuclearly encoded mitochondrial genes and mitochondrial biogenesis.* Mol Cell Biol, 2005. **25**(14): p. 6225-34.
100. Eagle, H., *Nutrition needs of mammalian cells in tissue culture*. Science, 1955. **122**(3168): p. 501-14.
101. Wise, D.R., DeBerardinis, R.J., Mancuso, A., Sayed, N., Zhang, X.Y., Pfeiffer, H.K., Nissim, I., Daikhin, E., Yudkoff, M., McMahon, S.B., and Thompson, C.B., *Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction.* Proc Natl Acad Sci U S A, 2008. **105**(48): p. 18782-7.
102. Factor, V.M., Kiss, A., Woitach, J.T., Wirth, P.J., and Thorgeirsson, S.S., *Disruption of redox homeostasis in the transforming growth factor-alpha/c-myc transgenic mouse model of accelerated hepatocarcinogenesis*. J Biol Chem, 1998. **273**(25): p. 15846-53.
103. Vafa, O., Wade, M., Kern, S., Beeche, M., Pandita, T.K., Hampton, G.M., and Wahl, G.M., *c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability*. Mol Cell, 2002. **9**(5): p. 1031-44.
104. Dang, C.V., Li, F., and Lee, L.A., *Could MYC induction of mitochondrial biogenesis be linked to ROS production and genomic instability?* Cell Cycle, 2005. **4**(11): p. 1465-6.
105. Ray, S., Atkuri, K.R., Deb-Basu, D., Adler, A.S., Chang, H.Y., Herzenberg, L.A., and Felsher, D.W., *MYC can induce DNA breaks in vivo and in vitro independent of reactive oxygen species.* Cancer Res, 2006. **66**(13): p. 6598-605.
106. Dominguez-Sola, D., Ying, C.Y., Grandori, C., Ruggiero, L., Chen, B., Li, M., Galloway, D.A., Gu, W., Gautier, J., and Dalla-Favera, R., *Non-transcriptional control of DNA replication by c-Myc.* Nature, 2007. **448**(7152): p. 445-51.
107. Valovka, T., Schonfeld, M., Raffeiner, P., Breuker, K., Dunzendorfer-Matt, T., Hartl, M., and Bister, K., *Transcriptional control of DNA replication licensing by Myc.* Sci Rep, 2013. **3**: p. 3444.
108. Felsher, D.W. and Bishop, J.M., *Transient excess of MYC activity can elicit genomic instability and tumorigenesis*. Proc Natl Acad Sci U S A, 1999. **96**(7): p. 3940-4.
109. Karlsson, A., Deb-Basu, D., Cherry, A., Turner, S., Ford, J., and Felsher, D.W., *Defective double-strand DNA break repair and chromosomal translocations by MYC overexpression*. Proc Natl Acad Sci U S A, 2003. **100**(17): p. 9974-9.
110. Li, Z., Owonikoko, T.K., Sun, S.Y., Ramalingam, S.S., Doetsch, P.W., Xiao, Z.Q., Khuri, F.R., Curran, W.J., and Deng, X., *c-Myc suppression of DNA double-strand break repair.* Neoplasia, 2012. **14**(12): p. 1190-202.
111. Hanahan, D. and Weinberg, R.A., *Hallmarks of cancer: the next generation*. Cell, 2011. **144**(5): p. 646-74.
112. Burrell, R.A., McGranahan, N., Bartek, J., and Swanton, C., *The causes and consequences of genetic heterogeneity in cancer evolution*. Nature, 2013. **501**(7467): p. 338-45.
113. Mateyak, M.K., Obaya, A.J., Adachi, S., and Sedivy, J.M., *Phenotypes of c-Myc-deficient rat fibroblasts isolated by targeted homologous recombination*. Cell Growth Differ, 1997. **8**(10): p. 1039-48.
114. de Alboran, I.M., O'Hagan, R.C., Gartner, F., Malynn, B., Davidson, L., Rickert, R., Rajewsky, K., DePinho, R.A., and Alt, F.W., *Analysis of C-MYC function in normal cells via conditional gene-targeted mutation.* Immunity, 2001. **14**(1): p. 45-55.
115. Kelly, K., Cochran, B.H., Stiles, C.D., and Leder, P., *Cell-specific regulation of the c-myc gene by lymphocyte mitogens and platelet-derived growth factor*. Cell, 1983. **35**(3 Pt 2): p. 603-10.
116. Heikkila, R., Schwab, G., Wickstrom, E., Loke, S.L., Pluznik, D.H., Watt, R., and Neckers, L.M., *A c-myc antisense oligodeoxynucleotide inhibits entry into S phase but not progress from G0 to G1.* Nature, 1987. **328**(6129): p. 445-9.

117. Eilers, M., Schirm, S., and Bishop, J.M., *The MYC protein activates transcription of the alpha-prothymosin gene*. EMBO J, 1991. **10**(1): p. 133-41.
118. Shibuya, H., Yoneyama, M., Ninomiya-Tsuji, J., Matsumoto, K., and Taniguchi, T., *IL-2 and EGF receptors stimulate the hematopoietic cell cycle via different signaling pathways: demonstration of a novel role for c-myc*. Cell, 1992. **70**(1): p. 57-67.
119. Daksis, J.I., Lu, R.Y., Facchini, L.M., Marhin, W.W., and Penn, L.J., *Myc induces cyclin D1 expression in the absence of de novo protein synthesis and links mitogen-stimulated signal transduction to the cell cycle*. Oncogene, 1994. **9**(12): p. 3635-45.
120. Bouchard, C., Thieke, K., Maier, A., Saffrich, R., Hanley-Hyde, J., Ansorge, W., Reed, S., Sicinski, P., Bartek, J., and Eilers, M., *Direct induction of cyclin D2 by Myc contributes to cell cycle progression and sequestration of p27.* EMBO J, 1999. **18**(19): p. 5321-33.
121. Hermeking, H., Rago, C., Schuhmacher, M., Li, Q., Barrett, J.F., Obaya, A.J., O'Connell, B.C., Mateyak, M.K., Tam, W., Kohlhuber, F., Dang, C.V., Sedivy, J.M., Eick, D., Vogelstein, B., and Kinzler, K.W., *Identification of CDK4 as a target of c-MYC.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2229-34.
122. Kaczmarek, L., Hyland, J.K., Watt, R., Rosenberg, M., and Baserga, R., *Microinjected c-myc as a competence factor.* Science, 1985. **228**(4705): p. 1313-5.
123. Eilers, M., Picard, D., Yamamoto, K.R., and Bishop, J.M., *Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells*. Nature, 1989. **340**(6228): p. 66-8.
124. Takahashi, K. and Yamanaka, S., *Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors*. Cell, 2006. **126**: p. 663-676.
125. Okita, K., Ichisaka, T., and Yamanaka, S., *Generation of germline-competent induced pluripotent stem cells*. Nature, 2007. **448**(7151): p. 313-7.
126. Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R., *c-Myc is dispensable for direct reprogramming of mouse fibroblasts*. Cell Stem Cell, 2008. **2**(1): p. 10-2.
127. Shchors, K., Shchors, E., Rostker, F., Lawlor, E.R., Brown-Swigart, L., and Evan, G.I., *The Myc-dependent angiogenic switch in tumors is mediated by interleukin 1beta.* Genes Dev, 2006. **20**(18): p. 2527-38.
128. Soucek, L., Lawlor, E.R., Soto, D., Shchors, K., Swigart, L.B., and Evan, G.I., *Mast cells are required for angiogenesis and macroscopic expansion of Myc-induced pancreatic islet tumors*. Nat Med, 2007. **13**(10): p. 1211-8.
129. Nagy, J.A., Chang, S.H., Shih, S.C., Dvorak, A.M., and Dvorak, H.F., *Heterogeneity of the tumor vasculature.* Semin Thromb Hemost, 2010. **36**(3): p. 321-31.
130. Baudino, T.A., McKay, C., Pendeville-Samain, H., Nilsson, J.A., Maclean, K.H., White, E.L., Davis, A.C., Ihle, J.N., and Cleveland, J.L., *c-Myc is essential for vasculogenesis and angiogenesis during development and tumor progression.* Genes Dev, 2002. **16**(19): p. 2530-43.
131. Mezquita, P., Parghi, S.S., Brandvold, K.A., and Ruddell, A., Myc regulates VEGF production in B cells by stimulating initiation of VEGF mRNA translation. Oncogene, 2005. **24**(5): p. 889-901.
132. Janz, A., Sevignani, C., Kenyon, K., Ngo, C.V., and Thomas-Tikhonenko, A., *Activation of the myc oncoprotein leads to increased turnover of thrombospondin-1 mRNA*. Nucleic Acids Res, 2000. **28**(11): p. 2268-75.
133. Dews, M., Homayouni, A., Yu, D., Murphy, D., Sevignani, C., Wentzel, E., Furth, E.E., Lee, W.M., Enders, G.H., Mendell, J.T., and Thomas-Tikhonenko, A., *Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster.* Nat Genet, 2006. **38**(9): p. 1060-5.
134. Taguchi, A., Yanagisawa, K., Tanaka, M., Cao, K., Matsuyama, Y., Goto, H., and Takahashi, T., *Identification of hypoxia-inducible factor-1 alpha as a novel target for miR-17-92 microRNA cluster*. Cancer Res, 2008. **68**(14): p. 5540-5.
135. Gordan, J.D., Thompson, C.B., and Simon, M.C., *HIF and c-Myc: sibling rivals for control of cancer cell metabolism and proliferation*. Cancer Cell, 2007. **12**(2): p. 108-13.
136. Onder, T.T., Gupta, P.B., Mani, S.A., Yang, J., Lander, E.S., and Weinberg, R.A., *Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways*. Cancer Res, 2008. **68**(10): p. 3645-54.

137. Ma, L., Young, J., Prabhala, H., Pan, E., Mestdagh, P., Muth, D., Teruya-Feldstein, J., Reinhardt, F., Onder, T.T., Valastyan, S., Westermann, F., Speleman, F., Vandesompele, J., and Weinberg, R.A., *miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis.* Nat Cell Biol, 2010. **12**(3): p. 247-56.
138. Frye, M., Gardner, C., Li, E.R., Arnold, I., and Watt, F.M., *Evidence that Myc activation depletes the epidermal stem cell compartment by modulating adhesive interactions with the local microenvironment*. Development, 2003. **130**(12): p. 2793-808.
139. Yan, S., Zhou, C., Lou, X., Xiao, Z., Zhu, H., Wang, Q., Wang, Y., Lu, N., He, S., Zhan, Q., Liu, S., and Xu, N., *PTTG overexpression promotes lymph node metastasis in human esophageal squamous cell carcinoma.* Cancer Res, 2009. **69**(8): p. 3283-90.
140. McMahon, S.B., *MYC and the control of apoptosis.* Cold Spring Harb Perspect Med, 2014. **4**(7): p. a014407.
141. Wyllie, A.H., Rose, K.A., Morris, R.G., Steel, C.M., Foster, E., and Spandidos, D.A., *Rodent fibroblast tumours expressing human myc and ras genes: growth, metastasis and endogenous oncogene expression*. Br J Cancer, 1987. **56**(3): p. 251-9.
142. Askew, D.S., Ashmun, R.A., Simmons, B.C., and Cleveland, J.L., *Constitutive c-myc expression in an IL-3-dependent myeloid cell line suppresses cell cycle arrest and accelerates apoptosis*. Oncogene, 1991. **6**(10): p. 1915-22.
143. Evan, G.I., Wyllie, A.H., Gilbert, C.S., Littlewood, T.D., Land, H., Brooks, M., Waters, C.M., Penn, L.Z., and Hancock, D.C., *Induction of apoptosis in fibroblasts by c-myc protein.* Cell, 1992. **69**(1): p. 119-28.
144. Zindy, F., Eischen, C.M., Randle, D.H., Kamijo, T., Cleveland, J.L., Sherr, C.J., and Roussel, M.F., *Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization.* Genes Dev, 1998. **12**(15): p. 2424-33.
145. Kamijo, T., Weber, J.D., Zambetti, G., Zindy, F., Roussel, M.F., and Sherr, C.J., *Functional and physical interactions of the ARF tumor suppressor with p53 and Mdm2.* Proc Natl Acad Sci U S A, 1998. **95**(14): p. 8292-7.
146. Sherr, C.J., Tumor surveillance via the ARF-p53 pathway. Genes Dev, 1998. **12**(19): p. 2984-91.
147. Zhang, Y., Xiong, Y., and Yarbrough, W.G., *ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways*. Cell, 1998. **92**(6): p. 725-34.
148. Aubrey, B.J., Kelly, G.L., Janic, A., Herold, M.J., and Strasser, A., *How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression?* Cell Death Differ, 2018. **25**(1): p. 104-113.
149. Eischen, C.M., Woo, D., Roussel, M.F., and Cleveland, J.L., *Apoptosis triggered by Myc-induced suppression of Bcl-X(L) or Bcl-2 is bypassed during lymphomagenesis*. Mol Cell Biol, 2001. **21**(15): p. 5063-70.
150. Egle, A., Harris, A.W., Bouillet, P., and Cory, S., *Bim is a suppressor of Myc-induced mouse B cell leukemia.* Proc Natl Acad Sci U S A, 2004. **101**(16): p. 6164-9.
151. Soucie, E.L., Annis, M.G., Sedivy, J., Filmus, J., Leber, B., Andrews, D.W., and Penn, L.Z., *Myc potentiates apoptosis by stimulating Bax activity at the mitochondria.* Mol Cell Biol, 2001. **21**(14): p. 4725-36.
152. Eischen, C.M., Weber, J.D., Roussel, M.F., Sherr, C.J., and Cleveland, J.L., *Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis*. Genes Dev, 1999. **13**(20): p. 2658-69.
153. McMahon, S.B., Van Buskirk, H.A., Dugan, K.A., Copeland, T.D., and Cole, M.D., *The Novel ATM-Related Protein TRRAP Is an Essential Cofactor for the c-Myc and E2F Oncoproteins*. Cell, 1998. **94**: p. 363-374.
154. McMahon, S.B., Wood, M.A., and Cole, M.D., *The essential cofactor TRRAP recruits the histone acetyltransferase hGCN5 to c-Myc*. Mol Cell Biol, 2000. **20**(2): p. 556-62.
155. Frank, S.R., Parisi, T., Taubert, S., Fernandez, P., Fuchs, M., Chan, H.M., Livingston, D.M., and Amati, B., *MYC recruits the TIP60 histone acetyltransferase complex to chromatin.* EMBO Rep, 2003. **4**(6): p. 575-80.

156. Liu, X., Tesfai, J., Evrard, Y.A., Dent, S.Y., and Martinez, E., *c-Myc transformation domain recruits the human STAGA complex and requires TRRAP and GCN5 acetylase activity for transcription activation*. J Biol Chem, 2003. **278**(22): p. 20405-12.
157. Bouchard, C., Dittrich, O., Kiermaier, A., Dohmann, K., Menkel, A., Eilers, M., and Luscher, B., *Regulation of cyclin D2 gene expression by the Myc/Max/Mad network: Myc-dependent TRRAP recruitment and histone acetylation at the cyclin D2 promoter*. Genes Dev, 2001. **15**(16): p. 2042-7.
158. Frank, S.R., Schroeder, M., Fernandez, P., Taubert, S., and Amati, B., *Binding of c-Myc to chromatin mediates mitogen-induced acetylation of histone H4 and gene activation*. Genes Dev, 2001. **15**(16): p. 2069-82.
159. Nikiforov, M.A., Chandriani, S., Park, J., Kotenko, I., Matheos, D., Johnsson, A., McMahon, S.B., and Cole, M.D., *TRRAP-dependent and TRRAP-independent transcriptional activation by Myc family oncoproteins.* Mol Cell Biol, 2002. **22**(14): p. 5054-63.
160. Bannister, A.J. and Kouzarides, T., *Regulation of chromatin by histone modifications*. Cell Res, 2011. **21**(3): p. 381-95.
161. Arabi, A., Wu, S., Ridderstrale, K., Bierhoff, H., Shiue, C., Fatyol, K., Fahlen, S., Hydbring, P., Soderberg, O., Grummt, I., Larsson, L.G., and Wright, A.P., *c-Myc associates with ribosomal DNA and activates RNA polymerase I transcription.* Nat Cell Biol, 2005. **7**(3): p. 303-10.
162. Kenneth, N.S., Ramsbottom, B.A., Gomez-Roman, N., Marshall, L., Cole, P.A., and White, R.J., *TRRAP and GCN5 are used by c-Myc to activate RNA polymerase III transcription.* Proc Natl Acad Sci U S A, 2007. **104**(38): p. 14917-22.
163. Dai, M.S., Arnold, H., Sun, X.X., Sears, R., and Lu, H., *Inhibition of c-Myc activity by ribosomal protein L11.* EMBO J, 2007. **26**(14): p. 3332-45.
164. Deisenroth, C., Franklin, D.A., and Zhang, Y., *The Evolution of the Ribosomal Protein-MDM2-p53 Pathway*. Cold Spring Harb Perspect Med, 2016. **6**(12).
165. Vos, S.M., Farnung, L., Urlaub, H., and Cramer, P., *Structure of paused transcription complex Pol II-DSIF-NELF*. Nature, 2018. **560**(7720): p. 601-606.
166. Peterlin, B.M. and Price, D.H., *Controlling the elongation phase of transcription with P-TEFb*. Mol Cell, 2006. **23**(3): p. 297-305.
167. Eberhardy, S.R. and Farnham, P.J., *Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter*. J Biol Chem, 2002. **277**(42): p. 40156-62.
168. Kanazawa, S., Soucek, L., Evan, G., Okamoto, T., and Peterlin, B.M., *c-Myc recruits P-TEFb for transcription, cellular proliferation and apoptosis*. Oncogene, 2003. **22**(36): p. 5707-11.
169. Gargano, B., Amente, S., Majello, B., and Lania, L., *P-TEFb is a crucial co-factor for Myc transactivation.* Cell Cycle, 2007. **6**(16): p. 2031-7.
170. Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., Sharp, P.A., and Young, R.A., *c-Myc regulates transcriptional pause release.* Cell, 2010. **141**(3): p. 432-45.
171. Jaenicke, L.A., von Eyss, B., Carstensen, A., Wolf, E., Xu, W., Greifenberg, A.K., Geyer, M., Eilers, M., and Popov, N., *Ubiquitin-Dependent Turnover of MYC Antagonizes MYC/PAF1C Complex Accumulation to Drive Transcriptional Elongation.* Mol Cell, 2016. **61**(1): p. 54-67.
172. Gerlach, J.M., Furrer, M., Gallant, M., Birkel, D., Baluapuri, A., Wolf, E., and Gallant, P., *PAF1 complex component Leo1 helps recruit Drosophila Myc to promoters.* Proc Natl Acad Sci U S A, 2017. **114**(44): p. E9224-E9232.
173. Baluapuri, A., Hofstetter, J., Dudvarski Stankovic, N., Endres, T., Bhandare, P., Vos, S.M., Adhikari, B., Schwarz, J.D., Narain, A., Vogt, M., Wang, S.Y., Duster, R., Jung, L.A., Vanselow, J.T., Wiegering, A., Geyer, M., Maric, H.M., Gallant, P., Walz, S., Schlosser, A., Cramer, P., Eilers, M., and Wolf, E., *MYC Recruits SPT5 to RNA Polymerase II to Promote Processive Transcription Elongation.* Mol Cell, 2019. **74**(4): p. 674-687 e11.
174. Cowling, V.H. and Cole, M.D., *The Myc transactivation domain promotes global phosphorylation of the RNA polymerase II carboxy-terminal domain independently of direct DNA binding*. Mol Cell Biol, 2007. **27**(6): p. 2059-73.
175. Cole, M.D. and Cowling, V.H., *Specific regulation of mRNA cap methylation by the c-Myc and E2F1 transcription factors*. Oncogene, 2009. **28**(9): p. 1169-75.

176. Shilatifard, A., Conaway, R.C., and Conaway, J.W., *The RNA polymerase II elongation complex.* Annu Rev Biochem, 2003. **72**: p. 693-715.
177. Luse, D.S., *Rethinking the role of TFIIF in transcript initiation by RNA polymerase II*. Transcription, 2012. **3**(4): p. 156-9.
178. Thomas, L.R., Wang, Q., Grieb, B.C., Phan, J., Foshage, A.M., Sun, Q., Olejniczak, E.T., Clark, T., Dey, S., Lorey, S., Alicie, B., Howard, G.C., Cawthon, B., Ess, K.C., Eischen, C.M., Zhao, Z., Fesik, S.W., and Tansey, W.P., *Interaction with WDR5 promotes target gene recognition and tumorigenesis by MYC.* Mol Cell, 2015. **58**(3): p. 440-52.
179. Thomas, L.R., Foshage, A.M., Weissmiller, A.M., Popay, T.M., Grieb, B.C., Qualls, S.J., Ng, V., Carboneau, B., Lorey, S., Eischen, C.M., and Tansey, W.P., *Interaction of MYC with host cell factor-1 is mediated by the evolutionarily conserved Myc box IV motif.* Oncogene, 2016. **35**: p. 3613-3618.
180. Guarnaccia, A.D. and Tansey, W.P., *Moonlighting with WDR5: A Cellular Multitasker*. J Clin Med, 2018. **7**(2).
181. Cowling, V.H., Chandriani, S., Whitfield, M.L., and Cole, M.D., *A Conserved Myc Protein Domain, MBIV, Regulates DNA Binding, Apoptosis, Transformation, and G2 Arrest*. Molecular and Cellular Biology, 2006. **26**: p. 4226-4239.
182. Popay, T.M., Wang, J., Adams, C.M., Codreanu, S.G., Sherrod, S.D., McLean, J.A., Thomas, L.R., Lorey, S.L., Machida, Y.J., Weissmiller, A.M., Eischen, C., Liu, Q., and Tansey, W.P., *MYC regulates ribosome biogenesis and mitochondrial gene expression programs through interaction with Host Cell Factor-1.* In submission.
183. Peukert, K., Staller, P., Schneider, A., Carmichael, G., Hanel, F., and Eilers, M., *An alternative pathway for gene regulation by Myc.* EMBO J, 1997. **16**(18): p. 5672-86.
184. Staller, P., Peukert, K., Kiermaier, A., Seoane, J., Lukas, J., Karsunky, H., Moroy, T., Bartek, J., Massague, J., Hanel, F., and Eilers, M., *Repression of p15INK4b expression by Myc through association with Miz-1.* Nat Cell Biol, 2001. **3**(4): p. 392-9.
185. Herold, S., Wanzel, M., Beuger, V., Frohme, C., Beul, D., Hillukkala, T., Syvaoja, J., Saluz, H.P., Haenel, F., and Eilers, M., *Negative regulation of the mammalian UV response by Myc through association with Miz-1.* Mol Cell, 2002. **10**(3): p. 509-21.
186. Adhikary, S., Peukert, K., Karsunky, H., Beuger, V., Lutz, W., Elsasser, H.P., Moroy, T., and Eilers, M., *Miz1 is required for early embryonic development during gastrulation.* Mol Cell Biol, 2003. **23**(21): p. 7648-57.
187. Walz, S., Lorenzin, F., Morton, J., Wiese, K.E., von Eyss, B., Herold, S., Rycak, L., Dumay-Odelot, H., Karim, S., Bartkuhn, M., Roels, F., Wustefeld, T., Fischer, M., Teichmann, M., Zender, L., Wei, C.L., Sansom, O., Wolf, E., and Eilers, M., *Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles.* Nature, 2014. **511**(7510): p. 483-7.
188. Bedard, M., Maltais, L., Montagne, M., and Lavigne, P., *Miz-1 and Max compete to engage c-Myc: implication for the mechanism of inhibition of c-Myc transcriptional activity by Miz-1*. Proteins, 2017. **85**(2): p. 199-206.
189. Brenner, C., Deplus, R., Didelot, C., Loriot, A., Vire, E., De Smet, C., Gutierrez, A., Danovi, D., Bernard, D., Boon, T., Pelicci, P.G., Amati, B., Kouzarides, T., de Launoit, Y., Di Croce, L., and Fuks, F., *Myc represses transcription through recruitment of DNA methyltransferase corepressor.* EMBO J, 2005. **24**(2): p. 336-46.
190. Cheng, S.W., Davies, K.P., Yung, E., Beltran, R.J., Yu, J., and Kalpana, G.V., *c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function*. Nat Genet, 1999. **22**(1): p. 102-5.
191. Weissmiller, A.M., Wang, J., Lorey, S.L., Howard, G.C., Martinez, E., Liu, Q., and Tansey, W.P., *Inhibition of MYC by the SMARCB1 tumor suppressor.* Nat Commun, 2019. **10**(1): p. 2014.
192. Kurland, J.F. and Tansey, W.P., *Myc-mediated transcriptional repression by recruitment of histone deacetylase*. Cancer Res, 2008. **68**(10): p. 3624-9.
193. Dang, C.V., Reddy, E.P., Shokat, K.M., and Soucek, L., *Drugging the 'undruggable' cancer targets.* Nat Rev Cancer, 2017. **17**(8): p. 502-508.

194. Simonsson, T., Pecinka, P., and Kubista, M., *DNA tetraplex formation in the control region of c-myc*. Nucleic Acids Res, 1998. **26**(5): p. 1167-72.
195. Brooks, T.A. and Hurley, L.H., *Targeting MYC Expression through G-Quadruplexes*. Genes Cancer, 2010. **1**(6): p. 641-649.
196. Stathis, A. and Bertoni, F., *BET Proteins as Targets for Anticancer Treatment*. Cancer Discov, 2018. **8**(1): p. 24-36.
197. Shi, C., Zhang, H., Wang, P., Wang, K., Xu, D., Wang, H., Yin, L., Zhang, S., and Zhang, Y., *PROTAC induced-BET protein degradation exhibits potent anti-osteosarcoma activity by triggering apoptosis*. Cell Death Dis, 2019. **10**(11): p. 815.
198. Zhang, H., Li, G., Zhang, Y., Shi, J., Yan, B., Tang, H., Chen, S., Zhang, J., Wen, P., Wang, Z., Pang, C., Li, J., Guo, W., and Zhang, S., *Targeting BET Proteins With a PROTAC Molecule Elicits Potent Anticancer Activity in HCC Cells.* Front Oncol, 2019. **9**: p. 1471.
199. Chipumuro, E., Marco, E., Christensen, C.L., Kwiatkowski, N., Zhang, T., Hatheway, C.M., Abraham, B.J., Sharma, B., Yeung, C., Altabef, A., Perez-Atayde, A., Wong, K.K., Yuan, G.C., Gray, N.S., Young, R.A., and George, R.E., *CDK7 inhibition suppresses super-enhancer-linked oncogenic transcription in MYCN-driven cancer.* Cell, 2014. **159**(5): p. 1126-1139.
200. Christensen, C.L., Kwiatkowski, N., Abraham, B.J., Carretero, J., Al-Shahrour, F., Zhang, T., Chipumuro, E., Herter-Sprie, G.S., Akbay, E.A., Altabef, A., Zhang, J., Shimamura, T., Capelletti, M., Reibel, J.B., Cavanaugh, J.D., Gao, P., Liu, Y., Michaelsen, S.R., Poulsen, H.S., Aref, A.R., Barbie, D.A., Bradner, J.E., George, R.E., Gray, N.S., Young, R.A., and Wong, K.K., *Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor.* Cancer Cell, 2014. **26**(6): p. 909-922.
201. Devi, G.R., Beer, T.M., Corless, C.L., Arora, V., Weller, D.L., and Iversen, P.L., In vivo bioavailability and pharmacokinetics of a c-MYC antisense phosphorodiamidate morpholino oligomer, AVI-4126, in solid tumors. Clin Cancer Res, 2005. **11**(10): p. 3930-8.
202. Li, Y., Zhang, B., Zhang, H., Zhu, X., Feng, D., Zhang, D., Zhuo, B., Li, L., and Zheng, J., *Oncolytic adenovirus armed with shRNA targeting MYCN gene inhibits neuroblastoma cell proliferation and in vivo xenograft tumor growth*. J Cancer Res Clin Oncol, 2013. **139**(6): p. 933-41.
203. Wiegering, A., Uthe, F.W., Jamieson, T., Ruoss, Y., Huttenrauch, M., Kuspert, M., Pfann, C., Nixon, C., Herold, S., Walz, S., Taranets, L., Germer, C.T., Rosenwald, A., Sansom, O.J., and Eilers, M., *Targeting Translation Initiation Bypasses Signaling Crosstalk Mechanisms That Maintain High MYC Levels in Colorectal Cancer.* Cancer Discov, 2015. **5**(7): p. 768-781.
204. Janghorban, M., Farrell, A.S., Allen-Petersen, B.L., Pelz, C., Daniel, C.J., Oddo, J., Langer, E.M., Christensen, D.J., and Sears, R.C., *Targeting c-MYC by antagonizing PP2A inhibitors in breast cancer.* Proc Natl Acad Sci U S A, 2014. **111**(25): p. 9157-62.
205. Aho, E.R., Wang, J., Gogliotti, R.D., Howard, G.C., Phan, J., Acharya, P., Macdonald, J.D., Cheng, K., Lorey, S.L., Lu, B., Wenzel, S., Foshage, A.M., Alvarado, J., Wang, F., Shaw, J.G., Zhao, B., Weissmiller, A.M., Thomas, L.R., Vakoc, C.R., Hall, M.D., Hiebert, S.W., Liu, Q., Stauffer, S.R., Fesik, S.W., and Tansey, W.P., *Displacement of WDR5 from Chromatin by a WIN Site Inhibitor with Picomolar Affinity.* Cell Rep, 2019. **26**(11): p. 2916-2928 e13.
206. Soucek, L., Jucker, R., Panacchia, L., Ricordy, R., Tatò, F., and Nasi, S., *Omomyc, a potential Myc dominant negative, enhances Myc-induced apoptosis.* Cancer research, 2002. **62**: p. 3507-10.
207. Soucek, L., Whitfield, J., Martins, C.P., Finch, A.J., Murphy, D.J., Sodir, N.M., Karnezis, A.N., Swigart, L.B., Nasi, S., and Evan, G.I., *Modelling Myc inhibition as a cancer therapy.* Nature, 2008. **455**(7213): p. 679-83.
208. Soucek, L., Whitfield, J.R., Sodir, N.M., Masso-Valles, D., Serrano, E., Karnezis, A.N., Swigart, L.B., and Evan, G.I., *Inhibition of Myc family proteins eradicates KRas-driven lung cancer in mice.* Genes Dev, 2013. **27**(5): p. 504-13.
209. Sodir, N.M., Swigart, L.B., Karnezis, A.N., Hanahan, D., Evan, G.I., and Soucek, L., *Endogenous Myc maintains the tumor microenvironment.* Genes Dev, 2011. **25**(9): p. 907-16.
210. Annibali, D., Whitfield, J.R., Favuzzi, E., Jauset, T., Serrano, E., Cuartas, I., Redondo-Campos, S., Folch, G., Gonzalez-Junca, A., Sodir, N.M., Masso-Valles, D., Beaulieu, M.E., Swigart, L.B.,

Mc Gee, M.M., Somma, M.P., Nasi, S., Seoane, J., Evan, G.I., and Soucek, L., *Myc inhibition is effective against glioma and reveals a role for Myc in proficient mitosis.* Nat Commun, 2014. **5**: p. 4632.
211. Beaulieu, M.E., Jauset, T., Masso-Valles, D., Martinez-Martin, S., Rahl, P., Maltais, L., Zacarias-Fluck, M.F., Casacuberta-Serra, S., Serrano Del Pozo, E., Fiore, C., Foradada, L., Cano, V.C., Sanchez-Hervas, M., Guenther, M., Romero Sanz, E., Oteo, M., Tremblay, C., Martin, G., Letourneau, D., Montagne, M., Morcillo Alonso, M.A., Whitfield, J.R., Lavigne, P., and Soucek, L., *Intrinsic cell-penetrating activity propels Omomyc from proof of concept to viable anti-MYC therapy.* Sci Transl Med, 2019. **11**(484).
212. Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M., *GAL4-VP16 is an unusually potent transcriptional activator.* Nature, 1988. **335**(6190): p. 563-4.
213. Marsden, H.S., Campbell, M.E., Haarr, L., Frame, M.C., Parris, D.S., Murphy, M., Hope, R.G., Muller, M.T., and Preston, C.M., *The 65,000-Mr DNA-binding and virion trans-inducing proteins of herpes simplex virus type 1.* J Virol, 1987. **61**(8): p. 2428-37.
214. Preston, C.M., Frame, M.C., and Campbell, M.E., *A complex formed between cell components and an HSV structural polypeptide binds to a viral immediate early gene regulatory DNA sequence*. Cell, 1988. **52**(3): p. 425-34.
215. Gerster, T. and Roeder, R.G., *A herpesvirus trans-activating protein interacts with transcription factor OTF-1 and other cellular proteins*. Proc Natl Acad Sci U S A, 1988. **85**(17): p. 6347-51.
216. Kristie, T.M. and Sharp, P.A., *Interactions of the Oct-1 POU subdomains with specific DNA sequences and with the HSV alpha-trans-activator protein*. Genes Dev, 1990. **4**(12B): p. 2383-96.
217. Wysocka, J. and Herr, W., *The herpes simplex virus VP16-induced complex: the makings of a regulatory switch*. Trends Biochem Sci, 2003. **28**(6): p. 294-304.
218. O'Hare, P., Goding, C.R., and Haigh, A., *Direct combinatorial interaction between a herpes simplex virus regulatory protein and a cellular octamer-binding factor mediates specific induction of virus immediate-early gene expression*. EMBO J, 1988. **7**(13): p. 4231-8.
219. Babb, R., Huang, C.C., Aufiero, D.J., and Herr, W., *DNA recognition by the herpes simplex virus transactivator VP16: a novel DNA-binding structure*. Mol Cell Biol, 2001. **21**(14): p. 4700-12.
220. Whitlow, Z. and Kristie, T.M., *Recruitment of the transcriptional coactivator HCF-1 to viral immediate-early promoters during initiation of reactivation from latency of herpes simplex virus type 1*. J Virol, 2009. **83**(18): p. 9591-5.
221. Minocha, S., Sung, T.L., Villeneuve, D., Lammers, F., and Herr, W., *Compensatory embryonic response to allele-specific inactivation of the murine X-linked gene Hcfc1*. Dev Biol, 2016. **412**(1): p. 1-17.
222. Minocha, S., Bessonnard, S., Sung, T.L., Moret, C., Constam, D.B., and Herr, W., *Epiblast-specific loss of HCF-1 leads to failure in anterior-posterior axis specification*. Dev Biol, 2016. **418**(1): p. 75-88.
223. Johnson, K.M., Mahajan, S.S., and Wilson, A.C., *Herpes simplex virus transactivator VP16 discriminates between HCF-1 and a novel family member, HCF-2*. Journal of virology, 1999. **73**: p. 3930-40.
224. Wilson, A.C., Freiman, R.N., Goto, H., Nishimoto, T., and Herr, W., *VP16 targets an amino-terminal domain of HCF involved in cell cycle progression*. Mol Cell Biol, 1997. **17**(10): p. 6139-46.
225. Wilson, A.C., Boutros, M., Johnson, K.M., and Herr, W., *HCF-1 amino- and carboxy-terminal subunit association through two separate sets of interaction modules: involvement of fibronectin type 3 repeats*. Mol Cell Biol, 2000. **20**(18): p. 6721-30.
226. Wilson, A.C., LaMarco, K., Peterson, M.G., and Herr, W., *The VP16 accessory protein HCF is a family of polypeptides processed from a large precursor protein*. Cell, 1993. **74**(1): p. 115-25.
227. Park, J., Lammers, F., Herr, W., and Song, J.J., *HCF-1 self-association via an interdigitated Fn3 structure facilitates transcriptional regulatory complex formation*. Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17430-5.
228. Gudkova, D., Dergai, O., Praz, V., and Herr, W., *HCF-2 inhibits cell proliferation and activates differentiation-gene expression programs*. Nucleic Acids Res, 2019. **47**(11): p. 5792-5808.

229. La Boissière, S., Hughes, T., and O'Hare, P., *HCF-dependent nuclear import of VP16.* EMBO J, 1999. **18**(2): p. 480-9.
230. Wysocka, J., Reilly, P.T., and Herr, W., *Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells*. Mol Cell Biol, 2001. **21**(11): p. 3820-9.
231. Wilson, A.C., Parrish, J.E., Massa, H.F., Nelson, D.L., Trask, B.J., and Herr, W., *The gene encoding the VP16-accessory protein HCF (HCFC1) resides in human Xq28 and is highly expressed in fetal tissues and the adult kidney*. Genomics, 1995. **25**: p. 462-468.
232. Goto, H., Motomura, S., Wilson, A.C., Freiman, R.N., Nakabeppu, Y., Fukushima, K., Fujishima, M., Herr, W., and Nishimoto, T., *A single-point mutation in HCF causes temperature-sensitive cell-cycle arrest and disrupts VP16 function.* Genes Dev, 1997. **11**(6): p. 726-37.
233. Reilly, P.T. and Herr, W., *Spontaneous reversion of tsBN67 cell proliferation and cytokinesis defects in the absence of HCF-1 function*. Exp Cell Res, 2002. **277**(1): p. 119-30.
234. Parker, J.B., Yin, H., Vinckevicius, A., and Chakravarti, D., *Host cell factor-1 recruitment to E2F-bound and cell-cycle-control genes is mediated by THAP11 and ZNF143*. Cell Rep, 2014. **9**(3): p. 967-82.
235. Michaud, J., Praz, V., James Faresse, N., Jnbaptiste, C.K., Tyagi, S., Schutz, F., and Herr, W., *HCFC1 is a common component of active human CpG-island promoters and coincides with ZNF143, THAP11, YY1, and GABP transcription factor occupancy*. Genome Res, 2013. **23**(6): p. 907-16.
236. Reilly, P.T., Wysocka, J., and Herr, W., *Inactivation of the retinoblastoma protein family can bypass the HCF-1 defect in tsBN67 cell proliferation and cytokinesis*. Mol Cell Biol, 2002. **22**(19): p. 6767-78.
237. Wilson, A.C., Peterson, M.G., and Herr, W., *The HCF repeat is an unusual proteolytic cleavage signal. Genes and Development*, 1995. **9**: p. 2445-2458.
238. Kristie, T.M., Pomerantz, J.L., Twomey, T.C., Parent, S.A., and Sharp, P.A., *The Cellular C1 Factor of the Herpes Simplex Virus Enhancer Complex Is a Family of Polypeptides*. Journal of Biological Chemistry, 1995. **270**: p. 4387-4394.
239. Wilson, A.C., Peterson, M.G., and Herr, W., *The HCF repeat is an unusual proteolytic cleavage signal*. Genes Dev, 1995. **9**(20): p. 2445-58.
240. Lazarus, M.B., Jiang, J., Kapuria, V., Bhuiyan, T., Janetzko, J., Zandberg, W.F., Vocadlo, D.J., Herr, W., and Walker, S., *HCF-1 is cleaved in the active site of O-GlcNAc transferase.* Science, 2013. **342**(6163): p. 1235-9.
241. Julien, E. and Herr, W., *Proteolytic processing is necessary to separate and ensure proper cell growth and cytokinesis functions of HCF-1*. EMBO J, 2003. **22**(10): p. 2360-9.
242. Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., Conaway, J.W., Conaway, R.C., and Herr, W., *O-GlcNAc transferase catalyzes site-specific proteolysis of HCF-1.* Cell, 2011. **144**(3): p. 376-88.
243. Ong, Q., Han, W., and Yang, X., *O-GlcNAc as an Integrator of Signaling Pathways*. Front Endocrinol (Lausanne), 2018. **9**: p. 599.
244. Martin, S.E.S., Tan, Z.W., Itkonen, H.M., Duveau, D.Y., Paulo, J.A., Janetzko, J., Boutz, P.L., Tork, L., Moss, F.A., Thomas, C.J., Gygi, S.P., Lazarus, M.B., and Walker, S., *Structure-Based Evolution of Low Nanomolar O-GlcNAc Transferase Inhibitors.* J Am Chem Soc, 2018. **140**(42): p. 13542-13545.
245. Wysocka, J., Myers, M.P., Laherty, C.D., Eisenman, R.N., and Herr, W., *Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1*. Genes & development, 2003. **17**: p. 896-911.
246. Daou, S., Mashtalir, N., Hammond-Martel, I., Pak, H., Yu, H., Sui, G., Vogel, J.L., Kristie, T.M., and Affar, e.B., *Crosstalk between O-GlcNAcylation and proteolytic cleavage regulates the host cell factor-1 maturation pathway.* Proc Natl Acad Sci U S A, 2011. **108**(7): p. 2747-52.
247. Mahajan, S.S., Johnson, K.M., and Wilson, A.C., *Molecular cloning of Drosophila HCF reveals proteolytic processing and self-association of the encoded protein*. J Cell Physiol, 2003. **194**(2): p. 117-26.
248. Capotosti, F., Hsieh, J.J., and Herr, W., *Species selectivity of mixed-lineage leukemia/trithorax and HCF proteolytic maturation pathways*. Mol Cell Biol, 2007. **27**(20): p. 7063-72.

249. Lane, E.A., Choi, D.W., Garcia-Haro, L., Levine, Z.G., Tedoldi, M., Walker, S., and Danial, N.N., *HCF-1 Regulates De Novo Lipogenesis through a Nutrient-Sensitive Complex with ChREBP.* Mol Cell, 2019. **75**(2): p. 357-371 e7.
250. Capotosti, F., Guernier, S., Lammers, F., Waridel, P., Cai, Y., Jin, J., Conaway, J.W., Conaway, R.C., and Herr, W., *O-GlcNAc Transferase Catalyzes Site-Specific Proteolysis of HCF-1.* Cell, 2011. **144**: p. 376-388.
251. Vogel, J.L. and Kristie, T.M., *Site-specific proteolysis of the transcriptional coactivator HCF-1 can regulate its interaction with protein cofactors*. Proc Natl Acad Sci U S A, 2006. **103**(18): p. 6817-22.
252. Freiman, R.N. and Herr, W., *Viral mimicry: Common mode of association with HCF by VP16 and the cellular protein LZIP*. Genes and Development, 1997. **11**: p. 3122-3127.
253. Tyagi, S. and Herr, W., *E2F1 mediates DNA damage and apoptosis through HCF-1 and the MLL family of histone methyltransferases*. EMBO J, 2009. **28**(20): p. 3185-95.
254. Tyagi, S., Chabes, A.L., Wysocka, J., and Herr, W., *E2F activation of S phase promoters via association with HCF-1 and the MLL family of histone H3K4 methyltransferases*. Mol Cell, 2007. **27**(1): p. 107-19.
255. Dejosez, M., Krumenacker, J.S., Zitur, L.J., Passeri, M., Chu, L.F., Songyang, Z., Thomson, J.A., and Zwaka, T.P., *Ronin is essential for embryogenesis and the pluripotency of mouse embryonic stem cells.* Cell, 2008. **133**(7): p. 1162-74.
256. Mazars, R., Gonzalez-de-Peredo, A., Cayrol, C., Lavigne, A.C., Vogel, J.L., Ortega, N., Lacroix, C., Gautier, V., Huet, G., Ray, A., Monsarrat, B., Kristie, T.M., and Girard, J.P., *The THAP-zinc finger protein THAP1 associates with coactivator HCF-1 and O-GlcNAc transferase: a link between DYT6 and DYT3 dystonias.* J Biol Chem, 2010. **285**(18): p. 13364-71.
257. Parker, J.B., Palchaudhuri, S., Yin, H., Wei, J., and Chakravarti, D., *A transcriptional regulatory role of the THAP11-HCF-1 complex in colon cancer cell function*. Mol Cell Biol, 2012. **32**(9): p. 1654-70.
258. Machida, Y.J., Machida, Y., Vashisht, A.A., Wohlschlegel, J.A., and Dutta, A., *The deubiquitinating enzyme BAP1 regulates cell growth via interaction with HCF-1.* J Biol Chem, 2009. **284**(49): p. 34179-88.
259. Misaghi, S., Ottosen, S., Izrael-Tomasevic, A., Arnott, D., Lamkanfi, M., Lee, J., Liu, J., O'Rourke, K., Dixit, V.M., and Wilson, A.C., *Association of C-Terminal Ubiquitin Hydrolase BRCA1-Associated Protein 1 with Cell Cycle Regulator Host Cell Factor 1.* Molecular and Cellular Biology, 2009. **29**: p. 2181-2192.
260. Yu, H., Mashtalir, N., Daou, S., Hammond-Martel, I., Ross, J., Sui, G., Hart, G.W., Rauscher, F.J., Drobetsky, E., Milot, E., Shi, Y., and Affar, e.B., *The ubiquitin carboxyl hydrolase BAP1 forms a ternary complex with YY1 and HCF-1 and is a critical regulator of gene expression.* Mol Cell Biol, 2010. **30**(21): p. 5071-85.
261. Qin, J., Zhou, Z., Chen, W., Wang, C., Zhang, H., Ge, G., Shao, M., You, D., Fan, Z., Xia, H., Liu, R., and Chen, C., *BAP1 promotes breast cancer cell proliferation and metastasis by deubiquitinating KLF5.* Nat Commun, 2015. **6**: p. 8471.
262. Dey, A., Seshasayee, D., Noubade, R., French, D.M., Liu, J., Chaurushiya, M.S., Kirkpatrick, D.S., Pham, V.C., Lill, J.R., Bakalarski, C.E., Wu, J., Phu, L., Katavolos, P., LaFave, L.M., Abdel-Wahab, O., Modrusan, Z., Seshagiri, S., Dong, K., Lin, Z., Balazs, M., Suriben, R., Newton, K., Hymowitz, S., Garcia-Manero, G., Martin, F., Levine, R.L., and Dixit, V.M., *Loss of the tumor suppressor BAP1 causes myeloid transformation.* Science, 2012. **337**(6101): p. 1541-6.
263. Foglizzo, M., Middleton, A.J., Burgess, A.E., Crowther, J.M., Dobson, R.C.J., Murphy, J.M., Day, C.L., and Mace, P.D., *A bidentate Polycomb Repressive-Deubiquitinase complex is required for efficient activity on nucleosomes.* Nat Commun, 2018. **9**(1): p. 3932.
264. Antonova, A., Hummel, B., Khavaran, A., Redhaber, D.M., Aprile-Garcia, F., Rawat, P., Gundel, K., Schneck, M., Hansen, E.C., Mitschke, J., Mittler, G., Miething, C., and Sawarkar, R., *Heat-Shock Protein 90 Controls the Expression of Cell-Cycle Genes by Stabilizing Metazoan-Specific Host-Cell Factor HCFC1.* Cell Rep, 2019. **29**(6): p. 1645-1659 e9.

265. Nagy, P.L., Griesenbeck, J., Kornberg, R.D., and Cleary, M.L., *A trithorax-group complex purified from Saccharomyces cerevisiae is required for methylation of histone H3*. Proc Natl Acad Sci U S A, 2002. **99**(1): p. 90-4.
266. Wysocka, J., Myers, M.P., Laherty, C.D., Eisenman, R.N., and Herr, W., *Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1*. Genes Dev, 2003. **17**(7): p. 896-911.
267. Huang, J., Kent, J.R., Placek, B., Whelan, K.A., Hollow, C.M., Zeng, P.Y., Fraser, N.W., and Berger, S.L., *Trimethylation of histone H3 lysine 4 by Set1 in the lytic infection of human herpes simplex virus 1.* J Virol, 2006. **80**(12): p. 5740-6.
268. Wang, Y.L., Faiola, F., Xu, M., Pan, S., and Martinez, E., *Human ATAC is a GCN5/PCAF-containing acetylase complex with a novel NC2-like histone fold module that interacts with the TATA-binding protein*. Journal of Biological Chemistry, 2008. **283**: p. 33808-33815.
269. Cai, Y., Jin, J., Swanson, S.K., Cole, M.D., Choi, S.H., Florens, L., Washburn, M.P., Conaway, J.W., and Conaway, R.C., *Subunit composition and substrate specificity of a MOF-containing histone acetyltransferase distinct from the male-specific lethal (MSL) complex*. J Biol Chem, 2010. **285**(7): p. 4268-72.
270. Jóźwiak, P., Forma, E., Bryś, M., and Krześlak, A., *O-GlcNAcylation and Metabolic Reprograming in Cancer.* Frontiers in endocrinology, 2014. **5**: p. 145.
271. Vogel, J.L. and Kristie, T.M., *The novel coactivator C1 (HCF) coordinates multiprotein enhancer formation and mediates transcription activation by GABP*. The EMBO journal, 2000. **19**: p. 683-90.
272. Scarpulla, R.C., *Nuclear activators and coactivators in mammalian mitochondrial biogenesis*. Biochim Biophys Acta, 2002. **1576**(1-2): p. 1-14.
273. Yang, Z.F., Drumea, K., Mott, S., Wang, J., and Rosmarin, A.G., *GABP transcription factor (nuclear respiratory factor 2) is required for mitochondrial biogenesis*. Mol Cell Biol, 2014. **34**(17): p. 3194-201.
274. Wu, Z., Puigserver, P., Andersson, U., Zhang, C., Adelmant, G., Mootha, V., Troy, A., Cinti, S., Lowell, B., Scarpulla, R.C., and Spiegelman, B.M., *Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1.* Cell, 1999. **98**(1): p. 115-24.
275. Lin, J., Puigserver, P., Donovan, J., Tarr, P., and Spiegelman, B.M., *Peroxisome proliferator-activated receptor gamma coactivator 1beta (PGC-1beta), a novel PGC-1-related transcription coactivator associated with host cell factor*. J Biol Chem, 2002. **277**(3): p. 1645-8.
276. Vercauteren, K., Gleyzer, N., and Scarpulla, R.C., *PGC-1-related coactivator complexes with HCF-1 and NRF-2beta in mediating NRF-2(GABP)-dependent respiratory gene expression*. J Biol Chem, 2008. **283**(18): p. 12102-11.
277. Ruan, H.B., Han, X., Li, M.D., Singh, J.P., Qian, K., Azarhoush, S., Zhao, L., Bennett, A.M., Samuel, V.T., Wu, J., Yates, J.R., and Yang, X., *O-GlcNAc transferase/host cell factor C1 complex regulates gluconeogenesis by modulating PGC-1α stability.* Cell Metab, 2012. **16**(2): p. 226-37.
278. Han, J.W., Valdez, J.L., Ho, D.V., Lee, C.S., Kim, H.M., Wang, X., Huang, L., and Chan, J.Y., *Nuclear factor-erythroid-2 related transcription factor-1 (Nrf1) is regulated by O-GlcNAc transferase*. Free Radic Biol Med, 2017. **110**: p. 196-205.
279. Sekine, H., Okazaki, K., Kato, K., Alam, M.M., Shima, H., Katsuoka, F., Tsujita, T., Suzuki, N., Kobayashi, A., Igarashi, K., Yamamoto, M., and Motohashi, H., *O-GlcNAcylation Signal Mediates Proteasome Inhibitor Resistance in Cancer Cells by Stabilizing NRF1.* Mol Cell Biol, 2018. **38**(17).
280. Lu, R., Yang, P., O'Hare, P., and Misra, V., *Luman, a new member of the CREB/ATF family, binds to herpes simplex virus VP16-associated host cellular factor*. Mol Cell Biol, 1997. **17**(9): p. 5117-26.
281. Lu, R. and Misra, V., *Zhangfei: a second cellular protein interacts with herpes simplex virus accessory factor HCF in a manner similar to Luman and VP16*. Nucleic Acids Res, 2000. **28**(12): p. 2446-54.

282. Misra, V., Rapin, N., Akhova, O., Bainbridge, M., and Korchinski, P., *Zhangfei is a potent and specific inhibitor of the host cell factor-binding transcription factor Luman*. J Biol Chem, 2005. **280**(15): p. 15257-66.
283. Luciano, R.L. and Wilson, A.C., *HCF-1 functions as a coactivator for the zinc finger protein Krox20*. J Biol Chem, 2003. **278**(51): p. 51116-24.
284. Mahajan, S.S., Little, M.M., Vazquez, R., and Wilson, A.C., *Interaction of HCF-1 with a cellular nuclear export factor.* J Biol Chem, 2002. **277**(46): p. 44292-9.
285. Piluso, D., Bilan, P., and Capone, J.P., *Host cell factor-1 interacts with and antagonizes transactivation by the cell cycle regulatory factor Miz-1*. J Biol Chem, 2002. **277**(48): p. 46799-808.
286. Gunther, M., Laithier, M., and Brison, O., *A set of proteins interacting with transcription factor Sp1 identified in a two-hybrid screening*. Mol Cell Biochem, 2000. **210**(1-2): p. 131-42.
287. Ajuh, P.M., Browne, G.J., Hawkes, N.A., Cohen, P.T., Roberts, S.G., and Lamond, A.I., *Association of a protein phosphatase 1 activity with the human factor C1 (HCF) complex.* Nucleic Acids Res, 2000. **28**(3): p. 678-86.
288. Gedeon, A., Kerr, B., Mulley, J., and Turner, G., *Localisation of the MRX3 gene for non-specific X linked mental retardation*. J Med Genet, 1991. **28**(6): p. 372-7.
289. Huang, L., Jolly, L.A., Willis-Owen, S., Gardner, A., Kumar, R., Douglas, E., Shoubridge, C., Wieczorek, D., Tzschach, A., Cohen, M., Hackett, A., Field, M., Froyen, G., Hu, H., Haas, S.A., Ropers, H.H., Kalscheuer, V.M., Corbett, M.A., and Gecz, J., *A noncoding, regulatory mutation implicates HCFC1 in nonsyndromic intellectual disability.* Am J Hum Genet, 2012. **91**(4): p. 694-702.
290. Yu, H.C., Sloan, J.L., Scharer, G., Brebner, A., Quintana, A.M., Achilly, N.P., Manoli, I., Coughlin, C.R., Geiger, E.A., Schneck, U., Watkins, D., Suormala, T., Van Hove, J.L., Fowler, B., Baumgartner, M.R., Rosenblatt, D.S., Venditti, C.P., and Shaikh, T.H., *An X-linked cobalamin disorder caused by mutations in transcriptional coregulator HCFC1.* Am J Hum Genet, 2013. **93**(3): p. 506-14.
291. Minocha, S., Villeneuve, D., Praz, V., Moret, C., Lopes, M., Pinatel, D., Rib, L., Guex, N., and Herr, W., *Rapid Recapitulation of Nonalcoholic Steatohepatitis upon Loss of Host Cell Factor 1 Function in Mouse Hepatocytes*. Mol Cell Biol, 2019. **39**(5).
292. Abu-Farha, M., Lambert, J.-P., Al-Madhoun, A.S., Elisma, F., Skerjanc, I.S., and Figeys, D., *The tale of two domains: proteomics and genomics analysis of SMYD2, a new histone methyltransferase*. Molecular & cellular proteomics : MCP, 2008. **7**: p. 560-72.
293. Glinsky, G.V., Berezovska, O., and Glinskii, A.B., *Microarray analysis identifies a death-from-cancer signature predicting therapy failure in patients with multiple types of cancer*. J Clin Invest, 2005. **115**(6): p. 1503-21.
294. Misaghi, S., Ottosen, S., Izrael-Tomasevic, A., Arnott, D., Lamkanfi, M., Lee, J., Liu, J., O'Rourke, K., Dixit, V.M., and Wilson, A.C., *Association of C-terminal ubiquitin hydrolase BRCA1-associated protein 1 with cell cycle regulator host cell factor 1.* Mol Cell Biol, 2009. **29**(8): p. 2181-92.
295. Furrer, M., Balbi, M., Albarca-Aguilera, M., Gallant, M., Herr, W., and Gallant, P., *Drosophila Myc interacts with host cell factor (dHCF) to activate transcription and control growth*. Journal of Biological Chemistry, 2010. **285**: p. 39623-39636.
296. Dingar, D., Kalkat, M., Chan, P.-K., Srikumar, T., Bailey, S.D., Tu, W.B., Coyaud, E., Ponzielli, R., Kolyar, M., Jurisica, I., Huang, A., Lupien, M., Penn, L.Z., and Raught, B., *BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors.* Journal of Proteomics, 2015. **118**: p. 95-111.
297. Lundberg, S.M., Tu, W.B., Raught, B., Penn, L.Z., Hoffman, M.M., and Lee, S.-I., *ChromNet: Learning the human chromatin network from all ENCODE ChIP-seq data.* Genome biology, 2016. **17**: p. 82.
298. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F., *Multiplex genome engineering using CRISPR/Cas systems.* Science, 2013. **339**(6121): p. 819-23.

299. Nabet, B., Roberts, J.M., Buckley, D.L., Paulk, J., Dastjerdi, S., Yang, A., Leggett, A.L., Erb, M.A., Lawlor, M.A., Souza, A., Scott, T.G., Vittori, S., Perry, J.A., Qi, J., Winter, G.E., Wong, K.K., Gray, N.S., and Bradner, J.E., *The dTAG system for immediate and target-specific protein degradation.* Nat Chem Biol, 2018. **14**(5): p. 431-441.
300. Venkateswaran, N., Lafita-Navarro, M.C., Hao, Y.H., Kilgore, J.A., Perez-Castro, L., Braverman, J., Borenstein-Auerbach, N., Kim, M., Lesner, N.P., Mishra, P., Brabletz, T., Shay, J.W., DeBerardinis, R.J., Williams, N.S., Yilmaz, O.H., and Conacci-Sorrell, M., *MYC promotes tryptophan uptake and metabolism by the kynurenine pathway in colon cancer.* Genes Dev, 2019. **33**(17-18): p. 1236-1251.
301. Klein, G., Giovanella, B., Westman, A., Stehlin, J.S., and Mumford, D., *An EBV-genome-negative cell line established from an American Burkitt lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by in vitro infection*. Intervirology, 1975. **5**(6): p. 319-34.
302. Miller, S.A., Dykes, D.D., and Polesky, H.F., *A simple salting out procedure for extracting DNA from human nucleated cells*. Nucleic Acids Res, 1988. **16**(3): p. 1215.
303. Southern, E.M., *Detection of specific sequences among DNA fragments separated by gel electrophoresis*. J Mol Biol, 1975. **98**(3): p. 503-17.
304. Mendez, J. and Stillman, B., *Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis*. Mol Cell Biol, 2000. **20**(22): p. 8602-12.
305. Farina, A., Faiola, F., and Martinez, E., *Reconstitution of an E box-binding Myc:Max complex with recombinant full-length proteins expressed in Escherichia coli*. Protein Expr Purif, 2004. **34**(2): p. 215-22.
306. Martin, M., *Cutadapt removes adapter sequences from high-throughput sequencing reads*. EMBnet. journal, 2011. **17**(1): p. 10-12.
307. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
308. Liao, Y., Smyth, G.K., and Shi, W., *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features*. Bioinformatics, 2014. **30**(7): p. 923-30.
309. Love, M.I., Huber, W., and Anders, S., *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.
310. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome*. Genome Biol, 2009. **10**(3): p. R25.
311. Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S., *Identifying ChIP-seq enrichment using MACS.* Nat Protoc, 2012. **7**(9): p. 1728-40.
312. Stark, R. and Brown, G., *DiffBind: differential binding analysis of ChIP-Seq peak data*. Bioconductor, 2011.
313. Wishart, D.S., Jewison, T., Guo, A.C., Wilson, M., Knox, C., Liu, Y., Djoumbou, Y., Mandal, R., Aziat, F., Dong, E., Bouatra, S., Sinelnikov, I., Arndt, D., Xia, J., Liu, P., Yallou, F., Bjorndahl, T., Perez-Pineiro, R., Eisner, R., Allen, F., Neveu, V., Greiner, R., and Scalbert, A., *HMDB 3.0--The Human Metabolome Database in 2013.* Nucleic Acids Res, 2013. **41**(Database issue): p. D801-7.
314. Smith, C.A., O'Maille, G., Want, E.J., Qin, C., Trauger, S.A., Brandon, T.R., Custodio, D.E., Abagyan, R., and Siuzdak, G., *METLIN: a metabolite mass spectral database.* Ther Drug Monit, 2005. **27**(6): p. 747-51.
315. Salvat, F., Jablonski, A., Powell, C.J., and Lee, A.Y., *NIST Electron Elastic-Scattering Cross-Section Database, Version 4.0*. 2016, Gaithersburg, MD: National Institute of Standards and Technology.
316. Schrimpe-Rutledge, A.C., Codreanu, S.G., Sherrod, S.D., and McLean, J.A., *Untargeted Metabolomics Strategies-Challenges and Emerging Directions.* J Am Soc Mass Spectrom, 2016. **27**(12): p. 1897-1905.

317. Adams, C.M., Kim, A.S., Mitra, R., Choi, J.K., Gong, J.Z., and Eischen, C.M., *BCL-W has a fundamental role in B cell survival and lymphomagenesis.* J Clin Invest, 2017. **127**(2): p. 635-650.
318. Wishart, D.S., Feunang, Y.D., Marcu, A., Guo, A.C., Liang, K., Vazquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., and Scalbert, A., *HMDB 4.0: the human metabolome database for 2018.* Nucleic Acids Res, 2018. **46**(D1): p. D608-D617.
319. Fahy, E., Subramaniam, S., Murphy, R.C., Nishijima, M., Raetz, C.R., Shimizu, T., Spener, F., van Meer, G., Wakelam, M.J., and Dennis, E.A., *Update of the LIPID MAPS comprehensive classification system for lipids.* J Lipid Res, 2009. **50 Suppl**: p. S9-14.
320. Chong, J., Wishart, D.S., and Xia, J., Using MetaboAnalyst 4.0 for *Comprehensive and Integrative Metabolomics Data Analysis*. Curr Protoc Bioinformatics, 2019. **68**(1): p. e86.
321. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K., *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.* Nat Commun, 2019. **10**(1): p. 1523.
322. Huang da, W., Sherman, B.T., and Lempicki, R.A., *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.
323. Huang da, W., Sherman, B.T., and Lempicki, R.A., *Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists*. Nucleic Acids Res, 2009. **37**(1): p. 1-13.
324. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.
325. Wickham, H., ggplot2: *Elegant Graphics for Data Analysis*. 2016: Springer International Publishing.
326. Wilson, a.C., Freiman, R.N., Goto, H., Nishimoto, T., and Herr, W., *VP16 targets an amino-terminal domain of HCF involved in cell cycle progression*. Molecular and cellular biology, 1997. **17**: p. 6139-6146.
327. Mahajan, S.S. and Wilson, A.C., *Mutations in host cell factor 1 separate its role in cell proliferation from recruitment of VP16 and LZIP*. Molecular and cellular biology, 2000. **20**: p. 919-28.
328. Lazarus, M.B., Jiang, J., Kapuria, V., Bhuiyan, T., Janetzko, J., Zandberg, W.F., Vocadlo, D.J., Herr, W., and Walker, S., *HCF-1 is cleaved in the active site of O-GlcNAc transferase.* Science (New York, N.Y.), 2013. **342**: p. 1235-9.
329. Lundberg, S.M., Tu, W.B., Raught, B., Penn, L.Z., Hoffman, M.M., and Lee, S.-I., *Learning the human chromatin network from all ENCODE ChIP-seq data.* bioRxiv, 2015: p. 023911.
330. Itkonen, H.M., Urbanucci, A., Martin, S.E., Khan, A., Mathelier, A., Thiede, B., Walker, S., and Mills, I.G., *High OGT activity is essential for MYC-driven proliferation of prostate cancer cells.* Theranostics, 2019. **9**(8): p. 2183-2197.
331. Kamemura, K. and Hart, G.W., *Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: a new paradigm for metabolic control of signal transduction and transcription*. Prog Nucleic Acid Res Mol Biol, 2003. **73**: p. 107-36.
332. LaBoissière, S., Walker, S., and O'Hare, P., *Concerted activity of host cell factor subregions in promoting stable VP16 complex assembly and preventing interference by the acidic activation domain*. Mol Cell Biol, 1997. **17**(12): p. 7108-18.
333. Wiman, K.G., Clarkson, B., Hayday, A.C., Saito, H., Tonegawa, S., and Hayward, W.S., *Activation of a translocated c-myc gene: role of structural alterations in the upstream region*. Proc Natl Acad Sci U S A, 1984. **81**(21): p. 6798-802.
334. Fresen, K.O. and Hausen, H., *Establishment of EBNA-expressing cell lines by infection of Epstein-Barr virus (EBV)-genome-negative human lymphoma cells with different EBV strains*. Int J Cancer, 1976. **17**(2): p. 161-6.
335. Ambrosio, M.R., Piccaluga, P.P., Ponzoni, M., Rocca, B.J., Malagnino, V., Onorati, M., De Falco, G., Calbi, V., Ogwang, M., Naresh, K.N., Pileri, S.A., Doglioni, C., Leoncini, L., and Lazzi, S.,

The alteration of lipid metabolism in Burkitt lymphoma identifies a novel marker: adipophilin. PLoS One, 2012. **7**(8): p. e44315.
336. Jeong, S.M., Lee, A., Lee, J., and Haigis, M.C., *SIRT4 protein suppresses tumor formation in genetic models of Myc-induced B cell lymphoma*. J Biol Chem, 2014. **289**(7): p. 4135-44.
337. Pajic, A., Spitkovsky, D., Christoph, B., Kempkes, B., Schuhmacher, M., Staege, M.S., Brielmeier, M., Ellwart, J., Kohlhuber, F., Bornkamm, G.W., Polack, A., and Eick, D., *Cell cycle activation by c-myc in a burkitt lymphoma model cell line.* Int J Cancer, 2000. **87**(6): p. 787-93.
338. Kawabata, Y., Hirokawa, M., Kitabayashi, A., Horiuchi, T., Kuroki, J., and Miura, A.B., *Defective apoptotic signal transduction pathway downstream of caspase-3 in human B-lymphoma cells: A novel mechanism of nuclear apoptosis resistance*. Blood, 1999. **94**(10): p. 3523-30.
339. Kawabata, K.C., Ehata, S., Komuro, A., Takeuchi, K., and Miyazono, K., *TGF-beta-induced apoptosis of B-cell lymphoma Ramos cells through reduction of MS4A1/CD20*. Oncogene, 2013. **32**(16): p. 2096-106.
340. McCormick, K.H., Giovanella, B.C., Klein, G., Nilsson, K., and Stehlin, J.S., *Diploid human lymphoblastoid and Burkitt lymphoma cell lines: susceptibility to murine NK cells and heterotransplantation to nude mice*. Int J Cancer, 1981. **28**(4): p. 455-8.
341. Lee, H.J., Kremer, D.M., Sajjakulnukit, P., Zhang, L., and Lyssiotis, C.A., *A large-scale analysis of targeted metabolomics data from heterogeneous biological samples provides insights into metabolite dynamics*. Metabolomics, 2019. **15**(7): p. 103.
342. Fernandez-de-Cossio-Diaz, J. and Vazquez, A., *A physical model of cell metabolism.* Sci Rep, 2018. **8**(1): p. 8349.
343. Kim, S., You, S., and Hwang, D., *Aminoacyl-tRNA synthetases and tumorigenesis: more than housekeeping*. Nat Rev Cancer, 2011. **11**(10): p. 708-18.
344. Tameire, F., Verginadis, II, Leli, N.M., Polte, C., Conn, C.S., Ojha, R., Salas Salinas, C., Chinga, F., Monroy, A.M., Fu, W., Wang, P., Kossenkov, A., Ye, J., Amaravadi, R.K., Ignatova, Z., Fuchs, S.Y., Diehl, J.A., Ruggero, D., and Koumenis, C., *ATF4 couples MYC-dependent translational activity to bioenergetic demands during tumour progression.* Nat Cell Biol, 2019. **21**(7): p. 889-899.
345. Zirin, J., Ni, X., Sack, L.M., Yang-Zhou, D., Hu, Y., Brathwaite, R., Bulyk, M.L., Elledge, S.J., and Perrimon, N., *Interspecies analysis of MYC targets identifies tRNA synthetases as mediators of growth and survival in MYC-overexpressing cells*. Proc Natl Acad Sci U S A, 2019. **116**(29): p. 14614-14619.
346. Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J., and Hinnebusch, A.G., *Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain*. Mol Cell, 2000. **6**(2): p. 269-79.
347. Harding, H.P., Zhang, Y., Zeng, H., Novoa, I., Lu, P.D., Calfon, M., Sadri, N., Yun, C., Popko, B., Paules, R., Stojdl, D.F., Bell, J.C., Hettmann, T., Leiden, J.M., and Ron, D., *An integrated stress response regulates amino acid metabolism and resistance to oxidative stress.* Mol Cell, 2003. **11**(3): p. 619-33.
348. Kristie, T.M. and Sharp, P.A., *Purification of the cellular C1 factor required for the stable recognition of the Oct-1 homeodomain by the herpes simplex virus alpha-trans-induction factor (VP16*). J Biol Chem, 1993. **268**(9): p. 6525-34.
349. Parker, J.B., Yin, H., Vinckevicius, A., and Chakravarti, D., *Host Cell Factor-1 Recruitment to E2F-Bound and Cell-Cycle-Control Genes Is Mediated by THAP11 and ZNF143*. Cell reports, 2014. **9**: p. 967-82.
350. Julien, E. and Herr, W., *Proteolytic processing is necessary to separate and ensure proper cell growth and cytokinesis functions of HCF-1*. EMBO Journal, 2003. **22**: p. 2360-2369.
351. Hancock, M.L., Meyer, R.C., Mistry, M., Khetani, R.S., Wagschal, A., Shin, T., Ho Sui, S.J., Naar, A.M., and Flanagan, J.G., *Insulin Receptor Associates with Promoters Genome-wide and Regulates Gene Expression.* Cell, 2019. **177**(3): p. 722-736 e22.
352. Vercauteren, K., Gleyzer, N., and Scarpulla, R.C., *Short hairpin RNA-mediated silencing of PRC (PGC-1-related coactivator) results in a severe respiratory chain deficiency associated with the proliferation of aberrant mitochondria*. J Biol Chem, 2009. **284**(4): p. 2307-19.

353. Wang, H., Mannava, S., Grachtchouk, V., Zhuang, D., Soengas, M.S., Gudkov, A.V., Prochownik, E.V., and Nikiforov, M.A., *c-Myc depletion inhibits proliferation of human tumor cells at various stages of the cell cycle.* Oncogene, 2008. **27**(13): p. 1905-15.
354. Lorenzin, F., Benary, U., Baluapuri, A., Walz, S., Jung, L.A., von Eyss, B., Kisker, C., Wolf, J., and Eilers, M., *Different promoter affinities account for specificity in MYC-dependent gene regulation.* eLife, 2016. **5**: p. 1009-1019.
355. Richart, L., Carrillo-de Santa Pau, E., Rio-Machin, A., de Andres, M.P., Cigudosa, J.C., Lobo, V.J.S.-A., and Real, F.X., *BPTF is required for c-MYC transcriptional activity and in vivo tumorigenesis.* Nat Commun, 2016. **7**: p. 10153.
356. Haggerty, T.J., Zeller, K.I., Osthus, R.C., Wonsey, D.R., and Dang, C.V., *A strategy for identifying transcription factor binding sites reveals two classes of genomic c-Myc target sites*. Proc Natl Acad Sci U S A, 2003. **100**(9): p. 5313-8.
357. Morrish, F., Giedt, C., and Hockenbery, D., c-*MYC apoptotic function is mediated by NRF-1 target genes*. Genes Dev, 2003. **17**(2): p. 240-55.
358. Ghoshal, K., Majumder, S., Datta, J., Motiwala, T., Bai, S., Sharma, S.M., Frankel, W., and Jacob, S.T., *Role of human ribosomal RNA (rRNA) promoter methylation and of methyl-CpG-binding protein MBD2 in the suppression of rRNA gene expression*. J Biol Chem, 2004. **279**(8): p. 6783-93.
359. Birch, J.L. and Zomerdijk, J.C., *Structure and function of ribosomal RNA gene chromatin*. Biochem Soc Trans, 2008. **36**(Pt 4): p. 619-24.
360. Ptashne, M. and Gann, A., *Transcriptional activation by recruitment*. Nature, 1997. **386**(6625): p. 569-77.
361. Scarpulla, R.C., *Nuclear control of respiratory gene expression in mammalian cells*. J Cell Biochem, 2006. **97**(4): p. 673-83.
362. Oksuz, O., Narendra, V., Lee, C.H., Descostes, N., LeRoy, G., Raviram, R., Blumenberg, L., Karch, K., Rocha, P.P., Garcia, B.A., Skok, J.A., and Reinberg, D., *Capturing the Onset of PRC2-Mediated Repressive Domain Formation.* Mol Cell, 2018. **70**(6): p. 1149-1162 e5.
363. Hershey, J.W., *Translational control in mammalian cells*. Annu Rev Biochem, 1991. **60**: p. 717-55.
364. Laferte, A., Favry, E., Sentenac, A., Riva, M., Carles, C., and Chedin, S., *The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components*. Genes Dev, 2006. **20**(15): p. 2030-40.
365. Felsher, D.W., *Cancer revoked: oncogenes as therapeutic targets*. Nat Rev Cancer, 2003. **3**(5): p. 375-80.
366. Wu, C.H., van Riggelen, J., Yetil, A., Fan, A.C., Bachireddy, P., and Felsher, D.W., *Cellular senescence is an important mechanism of tumor regression upon c-Myc inactivation.* Proc Natl Acad Sci U S A, 2007. **104**(32): p. 13028-33.
367. Wu, C.H., Sahoo, D., Arvanitis, C., Bradon, N., Dill, D.L., and Felsher, D.W., *Combined analysis of murine and human microarrays and ChIP analysis reveals genes associated with the ability of MYC to maintain tumorigenesis*. PLoS Genet, 2008. **4**(6): p. e1000090.
368. Hu, S., Balakrishnan, A., Bok, R.A., Anderton, B., Larson, P.E., Nelson, S.J., Kurhanewicz, J., Vigneron, D.B., and Goga, A., *13C-pyruvate imaging reveals alterations in glycolysis that precede c-Myc-induced tumor formation and regression.* Cell Metab, 2011. **14**(1): p. 131-42.
369. Wolf, E. and Eilers, M., *Targeting MYC Proteins for Tumor Therapy. Annual Review of Cancer Biology*, 2020. **4**(1): p. 61-75.
370. Perez-Garijo, A., Fuchs, Y., and Steller, H., *Apoptotic cells can induce non-autonomous apoptosis through the TNF pathway.* Elife, 2013. **2**: p. e01004.
371. Bremer, E., Samplonius, D., Kroesen, B.J., van Genne, L., de Leij, L., and Helfrich, W., *Exceptionally potent anti-tumor bystander activity of an scFv:sTRAIL fusion protein with specificity for EGP2 toward target antigen-negative tumor cells*. Neoplasia, 2004. **6**(5): p. 636-45.
372. Jain, M., Arvanitis, C., Chu, K., Dewey, W., Leonhardt, E., Trinh, M., Sundberg, C.D., Bishop, J.M., and Felsher, D.W., *Sustained loss of a neoplastic phenotype by brief inactivation of MYC.* Science, 2002. **297**(5578): p. 102-4.

373. Felsher, D.W. and Bishop, J.M., *Reversible tumorigenesis by MYC in hematopoietic lineages*. Mol Cell, 1999. **4**(2): p. 199-207.
374. Giuriato, S., Ryeom, S., Fan, A.C., Bachireddy, P., Lynch, R.C., Rioth, M.J., van Riggelen, J., Kopelman, A.M., Passegue, E., Tang, F., Folkman, J., and Felsher, D.W., *Sustained regression of tumors upon MYC inactivation requires p53 or thrombospondin-1 to reverse the angiogenic switch.* Proc Natl Acad Sci U S A, 2006. **103**(44): p. 16266-71.
375. Albert, B., Kos-Braun, I.C., Henras, A.K., Dez, C., Rueda, M.P., Zhang, X., Gadal, O., Kos, M., and Shore, D., *A ribosome assembly stress response regulates transcription to maintain proteome homeostasis.* Elife, 2019. **8**.
376. Golomb, L., Volarevic, S., and Oren, M., *p53 and ribosome biogenesis stress: the essentials*. FEBS Lett, 2014. **588**(16): p. 2571-9.
377. Liu, Y., Deisenroth, C., and Zhang, Y., *RP-MDM2-p53 Pathway: Linking Ribosomal Biogenesis and Tumor Surveillance*. Trends Cancer, 2016. **2**(4): p. 191-204.
378. Gaidano, G., Ballerini, P., Gong, J.Z., Inghirami, G., Neri, A., Newcomb, E.W., Magrath, I.T., Knowles, D.M., and Dalla-Favera, R., *p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia.* Proc Natl Acad Sci U S A, 1991. **88**(12): p. 5413-7.
379. Karpova, M.B., Sanmun, D., Henter, J.I., Smirnov, A.F., and Fadeel, B., *Betulinic acid, a natural cytotoxic agent, fails to trigger apoptosis in human Burkitt's lymphoma-derived B-cell lines*. Int J Cancer, 2006. **118**(1): p. 246-52.
380. O'Malley, J., Kumar, R., Inigo, J., Yadava, N., and Chandra, D., *Mitochondrial Stress Response and Cancer.* Trends Cancer, 2020. **6**(8): p. 688-701.
381. Quiros, P.M., Mottis, A., and Auwerx, J., *Mitonuclear communication in homeostasis and stress*. Nat Rev Mol Cell Biol, 2016. **17**(4): p. 213-26.
382. Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A., and Gorman, A.M., *The integrated stress response.* EMBO Rep, 2016. **17**(10): p. 1374-1395.
383. Dyson, H.J. and Wright, P.E., *Intrinsically unstructured proteins and their functions*. Nat Rev Mol Cell Biol, 2005. **6**(3): p. 197-208.
384. Sharma, R., Raduly, Z., Miskei, M., and Fuxreiter, M., *Fuzzy complexes: Specific binding without complete folding.* FEBS Lett, 2015. **589**(19 Pt A): p. 2533-42.
385. Cordeiro, T.N., Herranz-Trillo, F., Urbanek, A., Estana, A., Cortes, J., Sibille, N., and Bernado, P., *Small-angle scattering studies of intrinsically disordered proteins and their complexes.* Curr Opin Struct Biol, 2017. **42**: p. 15-23.
386. Itkonen, H.M., Minner, S., Guldvik, I.J., Sandmann, M.J., Tsourlakis, M.C., Berge, V., Svindland, A., Schlomm, T., and Mills, I.G., *O-GlcNAc transferase integrates metabolic pathways to regulate the stability of c-MYC in human prostate cancer cells.* Cancer Res, 2013. **73**(16): p. 5277-87.
387. Lee, D.H., Kwon, N.E., Lee, W.J., Lee, M.S., Kim, D.J., Kim, J.H., and Park, S.K., *Increased O-GlcNAcylation of c-Myc Promotes Pre-B Cell Proliferation.* Cells, 2020. **9**(1).
388. Gameiro, P.A. and Struhl, K., *Nutrient Deprivation Elicits a Transcriptional and Translational Inflammatory Response Coupled to Decreased Protein Synthesis*. Cell Rep, 2018. **24**(6): p. 1415-1424.
389. Hannan, K.M., Brandenburger, Y., Jenkins, A., Sharkey, K., Cavanaugh, A., Rothblum, L., Moss, T., Poortinga, G., McArthur, G.A., Pearson, R.B., and Hannan, R.D., *mTOR-dependent regulation of ribosomal gene transcription requires S6K1 and is mediated by phosphorylation of the carboxy-terminal activation domain of the nucleolar transcription factor UBF*. Mol Cell Biol, 2003. **23**(23): p. 8862-77.
390. Campagne, A., Lee, M.K., Zielinski, D., Michaud, A., Le Corre, S., Dingli, F., Chen, H., Shahidian, L.Z., Vassilev, I., Servant, N., Loew, D., Pasmant, E., Postel-Vinay, S., Wassef, M., and Margueron, R., *BAP1 complex promotes transcription by opposing PRC1-mediated H2A ubiquitylation.* Nat Commun, 2019. **10**(1): p. 348.
391. Vo, B.T., Wolf, E., Kawauchi, D., Gebhardt, A., Rehg, J.E., Finkelstein, D., Walz, S., Murphy, B.L., Youn, Y.H., Han, Y.G., Eilers, M., and Roussel, M.F., *The Interaction of Myc with Miz1 Defines Medulloblastoma Subgroup Identity.* Cancer Cell, 2016. **29**(1): p. 5-16.

392. Lamppa, J.W., Tanyos, S.A., and Griswold, K.E., *Engineering Escherichia coli for soluble expression and single step purification of active human lysozyme*. J Biotechnol, 2013. **164**(1): p. 1-8.