Abstract
Investigating Immune Pathways Involved with Cancer Drug Metabolism & Viral Evasion
Kirsten M. Knecht
2020

Acting in the first line of defense against viral infections, restriction factors are an important part of the innate immune response. My thesis work focuses on aspects of biology related to two HIV-1 restriction factors: SAMHD1 and APOBEC3G. While the ability of SAMHD1 to inhibit HIV-1 infection by depleting cellular nucleotides required for viral reverse transcription is well-studied, SAMHD1 interference with cancer drugs that closely mimic the chemical structure of nucleotides is still under investigation. In the first part of my thesis, I used X-ray crystallography and biochemical techniques to determine the molecular requirements of cancer drug interactions with the catalytic and allosteric pockets of SAMHD1. These results can be used to improve the efficacy of cancer drugs in patients with high SAMHD1 levels. In the second part of my thesis, I investigate how the virion infectivity factor (Vif) of Maedi-Visna Virus (MVV) recruits ovine APOBEC3 (A3) proteins to an E3 ligase complex to mediate the poly-ubiquitination and degradation of A3 proteins. I used biochemical techniques to map molecular determinants of MVV Vif binding to E3 ligase components (EloB/C and Cul5) and the non-canonical co-factor (CypA). While components of the α domain, including Cul5-, zinc-, and the EloB/C-binding motifs, are similar to other lentiviral Vifs, interactions between the α/β domain of MVV Vif and CypA are unique to the small ruminant lentivirus lineage. I also summarize my work reconstituting E3 ligase sub-complexes containing MVV Vif for structural studies. These studies shed light on how different lentiviral Vif proteins have evolved to target host APOBEC3 proteins for degradation in efforts to escape host restriction.

Investigating Immune Pathways Involved with Cancer Drug Metabolism & Viral Evasion

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

By
Kirsten M. Knecht

Dissertation Director: Yong Xiong

May 2020

Acknowledgements

First, I would like to thank my family for their unconditional love and encouragement throughout my PhD work. My parents and siblings have been my biggest cheer leaders and a source of inspiration to work harder each day. I am also blessed with an extended family of aunts, uncles, and cousins who have provided me with a large network of support.

Next, I must thank my mentors who have provided me with excellent scientific training. I thank my thesis advisor Dr. Yong Xiong for his endless optimism and enthusiasm for our projects. With his guidance, I have become a more independent scientist and improved my communication skills. Also, I thank my thesis committee, Dr. Mark Hochstrasser and Dr. Karen Anderson, for their constructive feedback and valuable insights on my projects. Additionally, I would like to thank my professors from Quinnipiac University who supported my interest in scientific research through my undergraduate studies, Dr. Lise Thomas and Dr. Carol Fenn.

Past and present lab mates have all been an important part of my training at Yale. I especially have to thank Dr. Olga Buzovetsky for the inspirational intensity that she brings to everything she does. Also, thanks to Dr. Samantha Ziegler for providing a reliably calming presence and keeping lab tea time alive. Dr. Katie Digianantonio, Dr. Christopher Lim, and Dr. Sarah Smaga all contributed to the supportive lab environment from which I benefited, as well.

Last but not least, I thank my partner Ruben Atilho for being my most trusted companion. Thank you for the laughter we share and for being my best friend throughout undergrad and grad school. I can't wait to start the next chapter of life together.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Overview of HIV-1 and other lentiviruses

Among the most infamous infectious diseases in modern history, Acquired Immune Deficiency Syndrome (AIDS) was first described by the U.S. Center for Disease Control in 1981 (CDC 1981a, CDC 1981b). Human Immunodeficiency Virus (HIV) is the causative agent. An untreated HIV infection eventually progresses to AIDS, leaving the patient defenseless against opportunistic infections and rare malignancies. Since the start of the global HIV/AIDS epidemic, it is estimated that about 75 million individuals have been infected with HIV and 32 million people have died from HIV/AIDS complications (World Health Organization 2018).

Although significant research progress towards understanding HIV/AIDS has been made in the past four decades, a cure or vaccine remains elusive. Its high mutation rate allows HIV to quickly gain resistance to antiviral drugs, so multi-drug treatments targeting multiple stages of the life cycle are used to increase efficacy. Additionally, latent HIV reservoirs make it difficult to completely rid a patient of HIV. Advances in antiretroviral therapies have increased the life expectancy of people with HIV from about ten years to nearly normal levels (Nakagawa et al. 2013). However, these patients must remain compliant with their multi-drug regimen for the remainder of their lives, and these therapies are not available to all patients. In order to completely eradicate HIV, more fundamental research of HIV biology is necessary.

Viruses are classified by characteristics such as morphology, host species, and nucleic acid type (Lefkowitz et al. 2017). Retroviruses are RNA viruses which use an enzyme called reverse transcriptase to convert their genetic material from single-stranded RNA to double-stranded DNA so it can be integrated it into the host genome (Nisole and Saib 2004). When first isolated from patient samples in 1983, HIV was described as a retrovirus with similarities to the deltaretrovirus Human T-cell Leukemia Virus (HTLV) (Barre-Sinnousi et al. 1983; Gallo et al. 1983). However, later experiments showed that HIV is more closely related to the lentivirus Maedi-Visna Virus (MVV) based on sequence homology and morphology (Gonda et al. 1985; Sonigo et al. 1985). Unlike HTLV, HIV encodes the regulatory proteins Rev and Tat to coordinate a complicated viral RNA transcription and splicing scheme unique to lentiviruses (Petursson et al. 1991).

Lentiviruses are a family of retroviruses clinically characterized by chronic, persistent infections with long incubation periods and fatal outcomes in mammals (Clements and Zink 1996). Paleovirologists estimate that lentiviruses have evolved alongside mammals for millions of years, with the earliest common ancestral lentivirus dating back at least 50 million years ago (Gifford 2012). Lentiviruses can be further categorized into five groups based on their mammalian hosts: (1) primate lentiviruses including HIV and Simian Immunodeficiency Virus (SIV), (2) feline lentiviruses including Feline Immunodeficiency Virus (FIV), (3) bovine lentiviruses including Bovine Immunodeficiency Virus (BIV), (4) small ruminant lentiviruses including MVV and Caprine Arthritis Encephalitis Virus (CAEV), and (5) equine lentiviruses including Equine Infectious Anemia Virus (EIAV) (Figure 1-1). While lentiviruses share similar

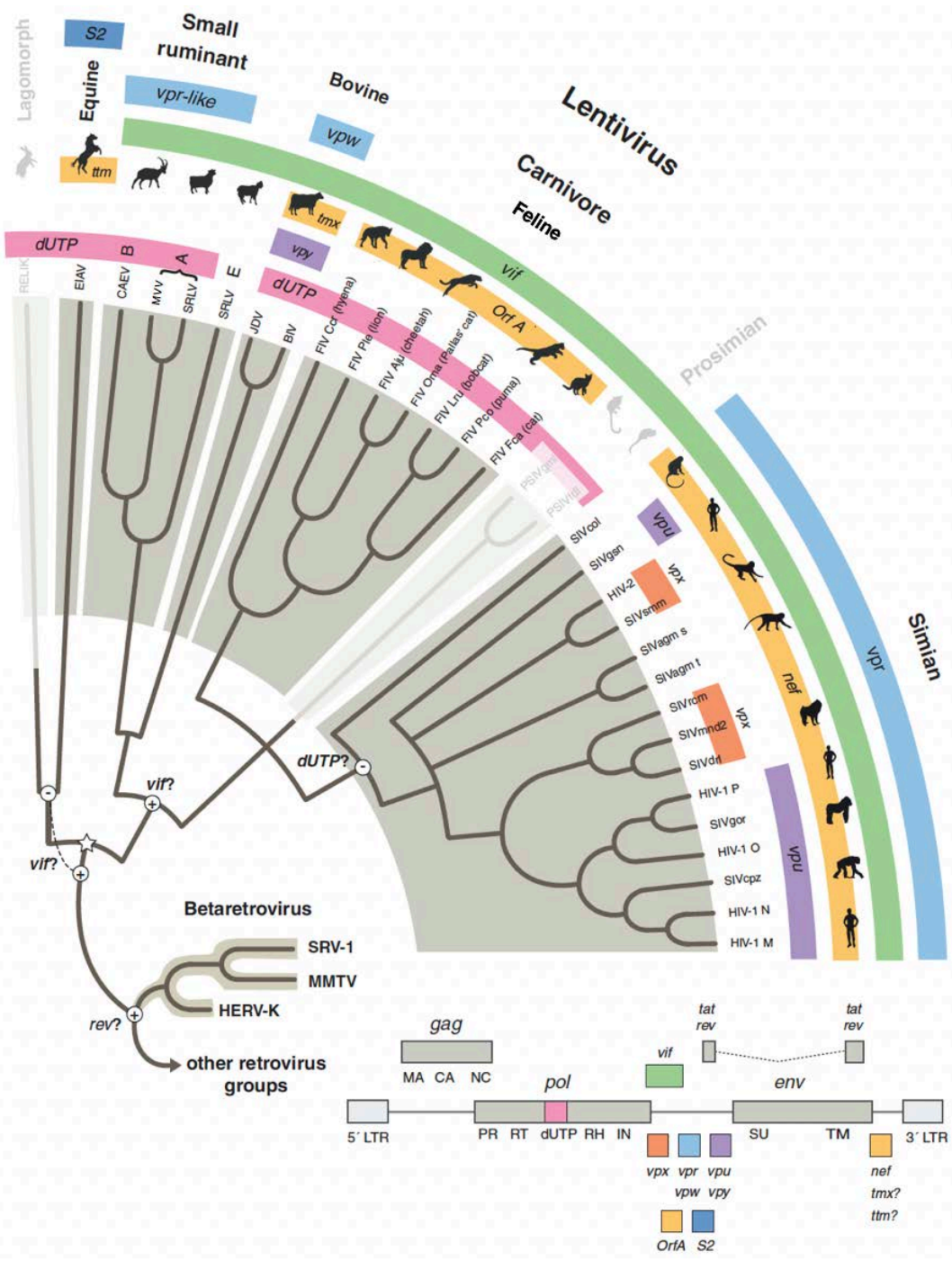

Figure 1-1**.** Phylogeny of lentiviruses with a schematic of genome architecture

All lentiviruses encode structural (*env*, *pol*, and *gag*) and regulatory (*rev* and *tat*) genes, but accessory proteins (*vif*, *vpx*, *vpr*, *nef*, *vpw*, *S2*, and *dUTP*) differ between lineages. Adapted from (Gifford 2012).

overall genetic architecture, accessory proteins which specifically counteract host defense mechanisms can differ between and within lentivirus groups.

Due to a high mutation frequency in lentiviruses, the genetic sequence from two isolates in the same lentivirus group can vary considerably. Also, HIV was acquired through multiple zoonotic transmissions from different primate Simian Immunodeficiency Viruses which adds to genetic variability (Sharp and Hahn 2011). While the main (M) group of HIV type 1 (HIV-1) responsible for the majority of the HIV/AIDS epidemic is most closely related to a SIV which infects chimpanzees, there are at least 9 sub-groups of HIV-1 which have been characterized (Taylor et al. 2008). The genetic diversity and rapid evolution of lentiviruses impose challenges for antiviral therapies and potential vaccine developments.

Small ruminant lentiviruses (SRLV), including Maedi Visna Virus (MVV) and Caprine Arthritis Encephalitis Virus (CAEV), also demonstrate the ability to infect multiple related species of hosts. MVV and CAEV were originally studied in sheep and goats, respectively. However, both MVV and CAEV are capable of infecting both host species, and co-infection has been reported (Minardi da Cruz et al 2013; Ramirez et al. 2013). Furthermore, SRLV can be divided into at least 4 highly divergent groups (A, B, C, and E) with many sub-types in each group (Clotti et al. 2019). More full genomes are needed in order to understand the viral heterogeneity and evolutionary patterns of SRLV. While MVV is distantly related to HIV, studying non-primate lentiviruses helps us to gain a better picture of how lentiviruses have evolved alongside mammals for millions of years.

1.2 HIV life cycle and hijacking of the ubiquitin-proteasome system

Extensive studies have elucidated the steps of the HIV-1 lifecycle in molecular detail (Barre-Sinoussi et al. 2013, Figure 1-2). An infectious HIV virion possesses a lipid membrane decorated with a viral protein, called Env. Inside the virion, a mature capsid protein core holds two copies of the positive sense RNA genome, accessory viral proteins, and host cofactors. Env trimers on the surface of the virion bind to CD4 receptors of a host T cell. This initial interaction is followed by recruitment of CCR5 coreceptors causing a conformational change that allows fusion of the viral membrane with the host plasma membrane.

After fusion, the capsid core and other components of the virion enter the host cytoplasm. The capsid core exploits host microtubule-associated machinery to stimulate viral movement toward the nucleus (Malikov et al. 2015; Rawle and Harrich, 2018). Meanwhile, reverse transcriptase (RT) converts the RNA genome to DNA so that it can enter the nucleus and become incorporated into the host genome by integrase (IN). Then, host machinery transcribes and translates the HIV genes necessary to assemble the next generation of immature virion. Env, Gag, and Pol proteins are trafficked to the membrane where the assembly of nascent immature virions begins. These immature virions, which contain intact Gag polyproteins, then bud off the cellular plasma membrane. Maturation of an infectious virus is mediated by the proteolytic action of protease (PR), which dimerizes and cuts at specific locations of the Gag polyproteins, releasing capsid proteins and allowing the mature capsid core to form (Bell and Lever, 2013).

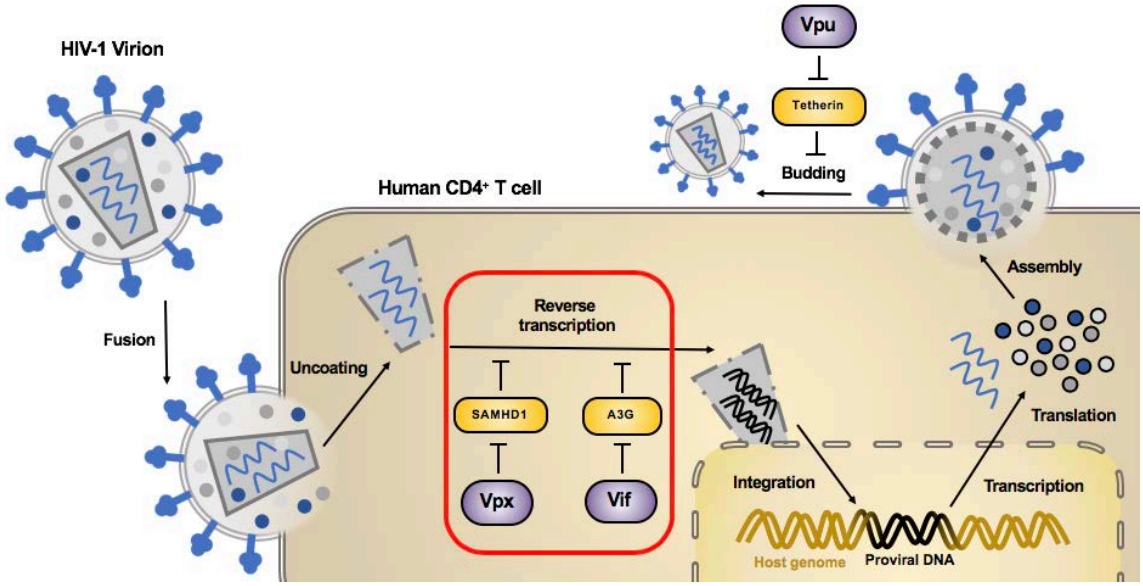

Figure 1-2. HIV lifecycle

Restriction factors are highlighted in gold. Viral accessory proteins that counteract host restriction are highlighted in purple. The two restriction factors of interest to this work are highlighted with a red box.

Throughout the course of millions of years of evolution, viruses and their mammalian host have participated in a molecular arms race to survive. At each step in its lifecycle, HIV-1 must overcome blocks to its replication by the host immune system. Two main branches of the human immune system include the innate and adaptive responses. While the adaptive response specifically neutralizes antigens from past infections, it can take days to weeks to mount a full adaptive response to a new infection. Furthermore, the highly glycosylated Env protein that coats the HIV-1 virion shields the virus from antigen recognition (Ward and Wilson, 2017). This glycan shield remains the main challenge to developing a vaccine for HIV-1.

The innate immune system represents the first line of defense against newly encountered viruses. Innate immune pathways recognize foreign substances, like viral nucleic acid, and use interferons to quickly induce express of proteins called restriction factors that non-specifically inhibit viral replication. Studying the life cycles of lentiviruses like HIV is not only important for novel therapeutic approaches, but can also lead to the discovery of interesting and unexpected biology.

One such discovery is that pathogens commonly target restriction factors for degradation by hijacking the human Ubiquitin-Proteasome System (UPS) (Seissler et al. 2017; Jia et al. 2015; Rojas et al. 2019). Under normal cellular conditions, the UPS functions in combination with the ribosome and molecular chaperones to control protein turnover and maintain proteostasis. Misfolded or unneeded proteins are targeted for destruction by the post-translational, covalent linkage of ubiquitin molecules. Poly-ubiquitination designates proteins to be destroyed by the 26S proteasome complex. The poly-ubiquitination of a target protein is mediated by a series of cellular enzymes labeled E1, E2, and E3. The ubiquitin molecule is first activated by an E1 enzyme, and then it is transferred to an E2. Finally, an E3 ligase complex composed of many protein components bridging the loaded E2 to the target substrate mediates the final ligation of the ubiquitin molecule to the substrate (Figure 1-3A).

Compared to the relatively small pools of two E1 and ~50 E2 enzymes, E3 ligase complexes are extremely diverse with more than 500 distinct complexes. This diversity of E3 ligases imbues the UPS with substrate specificity and regulatory mechanisms important to its function. Viruses target the E3 ligase complex in order to change the specificity of the ubiquitin ligation reaction. For example, HIV-1 Vif changes the

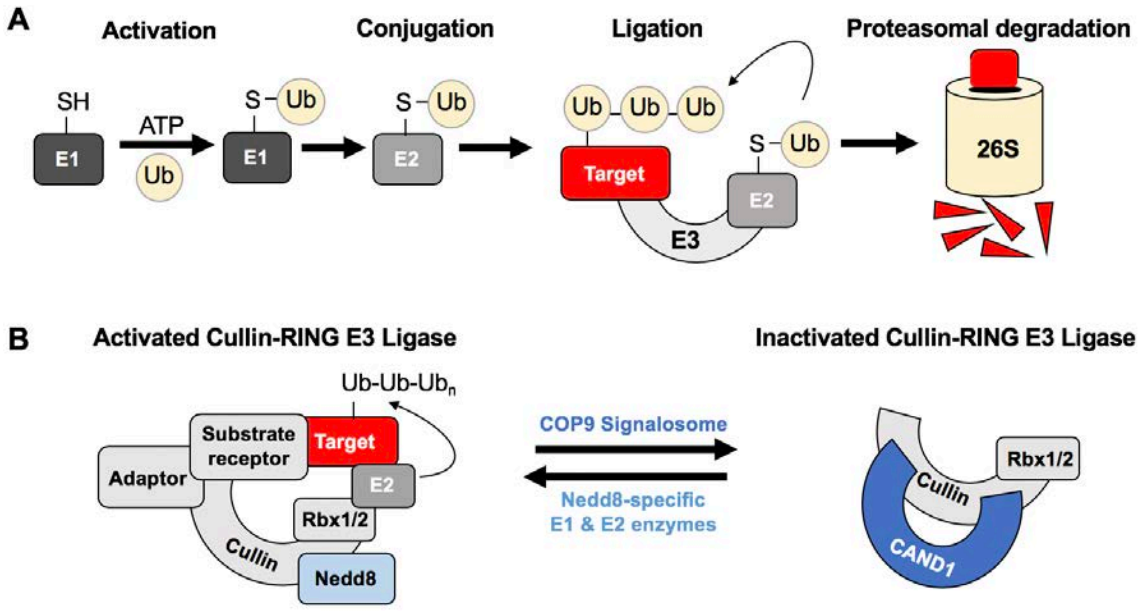

Figure 1-3. The Ubiquitin Proteasome System

(A) Cartoon schematic of the ubiquitin proteasome system, including E1, E2, and E3 enzymes that poly-ubiquitinate a substrate for degradation by the proteasome. (B) Cartoon schematic of Cullin-RING E3 ligase components and regulatory elements.

ubiquitination target of an E3 ligase complex from the family of Cullin-RING ubiquitin ligases. A normal Cullin-RING E3 ligase complex is composed of a scaffold protein (Cul1 to Cul7 or PARC), an adaptor protein (for example, EloB/C), a substrate receptor (for example, VHL), and an E2-activating RING finger protein (Rbx1 or Rbx2) (Figure 1-3B). Additionally, Cullin-RING E3 ligase activity is regulated by the reversible ubiquitin-like modification Nedd8 that is covalently linked to the WHB domain of the Cullin. A recent cryoEM structure of a neddylated Cul1-Rbx1 E3 ligase complex demonstrates that Nedd8 induces a conformational change which promotes ubiquitination of the substrate (Baek et al. 2020). De-neddylation of Cullin-RING E3 ligases by the COP9 signalosome allows an inactivated E3 ligase complex to bind another regulatory protein called CAND1. CAND1 is a helical repeat protein which wraps itself around the

Cullin occluding the neddylation-conjugating lysine residue and driving the exchange of substrate receptors (Goldenberg et al. 2004; Bosu and Kipreos, 2008)

Hijacking the UPS system allows a pathogen to directly destroy restriction factors that would otherwise block infection. Lentiviral accessory proteins such as Vif and Vpx both act as substrate adaptors which recruit APOBEC3G and SAMHD1, respectively, to E3 ligase complexes to be poly-ubiquitinated and subsequently degraded. APOBEC3G and SAMHD1 are two prominent examples of new biology that was discovered while searching for HIV-1 restriction factors. The biological functions of APOBEC3G and SAMHD1 were elucidated only after the discovery that they were each specifically targeted for degradation by a lentiviral accessory protein (Vif and Vpx, respectively). These two restriction factors are the focus of this work and described in more detail in sections below.

1.3 APOBEC3s are potent restriction factors of HIV-1

Early experiments showed that the accessory protein Vif is essential to HIV-1 proliferation only some T cell lines. Since this viral restriction in certain cell lines was inherited when non-permissive and permissive cell lines were fused, the existence of a potent restriction factor with variable expression in T cells was postulated (Strebel et al. 1987; Gabuzda et al. 1992; Simon et al. 1998). Soon after, a member of the previously uncharacterized mammalian apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3 or A3) family of cytidine deaminases was identified as the restriction factor counteracted by Vif (Sheehy et al. 2002). In the absence of Vif, A3G, and subsequently A3F, were shown to potently inhibit HIV-1 infections by

hypermutating single-stranded viral DNA (Sheehy et al. 2003; Harris et al. 2003; Zhang et al. 2003; Wiegand et al. 2004; Lidament et al. 2004; Zheng et al. 2004). A3G multimers become incorporated into an assembling virion in a producer cell by associating with Gag in an RNA-dependent manner (Zennou et al. 2004; Svarovskaia et al. 2004; Khan et al. 2005; Burnett et al. 2007). The cytidine deamination reaction carried out by A3G results in the conversion of cytidines to uridines in the viral genome (Figure 1-4A).

The seven human A3 proteins, labeled A through H, contain either one or two zinc-coordinating domains classified as Z1, Z2, or Z3 domains (LaRue et al. 2008). Crystal structures available for A3 proteins show conservation of the cytidine deaminase fold (Aydin et al. 2014, Figure 1-4B). However, the biological functions are not necessarily conserved amongst different A3s. For instance, A3G, A3F, A3H, and A3D have been shown to inhibit HIV-1, with A3G being the most potent (Hultquist et al. 2011; Chapain et al. 2012). While A3B and A3C do not inhibit HIV-1 infection, they inhibit SIV infection (Yu et al. 2014). Interestingly, several haplotypes of human A3H have been identified, each with varying degrees of HIV-1 inhibition (Wang et al. 2011; Refsland et al. 2014; Naruse et al. 2016). It is possible that each APOBEC gene has evolved to target different viruses or similar viruses in different contexts (Duggal et al. 2012).

Sequence diversity and gene duplication events that expand the repertoire of A3s may be driven by selective pressure imposed by infectious agents. APOBEC genes, especially A3G, demonstrate strong positive selection throughout at least 33 million years of primate evolution indicating selective pressures may have existed long before

lentiviral infections are thought to have originated (Sawyer et al. 2004). This suggests that the family of APOBEC genes may have evolved to defend against a broad range of retrovirus-like genomic attacks. Single-nucleotide polymorphisms in A3G that allow evasion of Vif counteraction have been shown to quickly result in adaption by Vif (Compton et al. 2011; Compton et al. 2013a). Since the human genome mutates at a rate that is orders of magnitude slower than viruses, heterozygosity and genome expansion of restriction factors such as APOBEC3 proteins have been suggested to be a mechanism meant to limit the ability of viruses to counter-adapt (Compton et al. 2013b).

Depending on the species, other mammals have different numbers of A3 genes named simply by their zinc coordination domains and site in the gene locus (LaRue et al. 2009, Figure 1-4C). Non-primate mammalian A3s are less-studied, but presumed to have similar structure and antiviral functions. It has been shown that non-primate A3s (equine, bovine, ovine, and feline) also inhibit lentiviral infections (Zielonka et al. 2009; Munk et al. 2008; Jonsson et al. 2006). Each lentiviral Vif targets at least one host A3 protein for degradation, but usually not A3 proteins from other host species. However, MVV Vif has been shown to degrade multiple host A3s and lacks species specificity (LaRue et al. 2010). This flexibility in A3 recognition may point to similarities between Vif-A3 interactions across species.

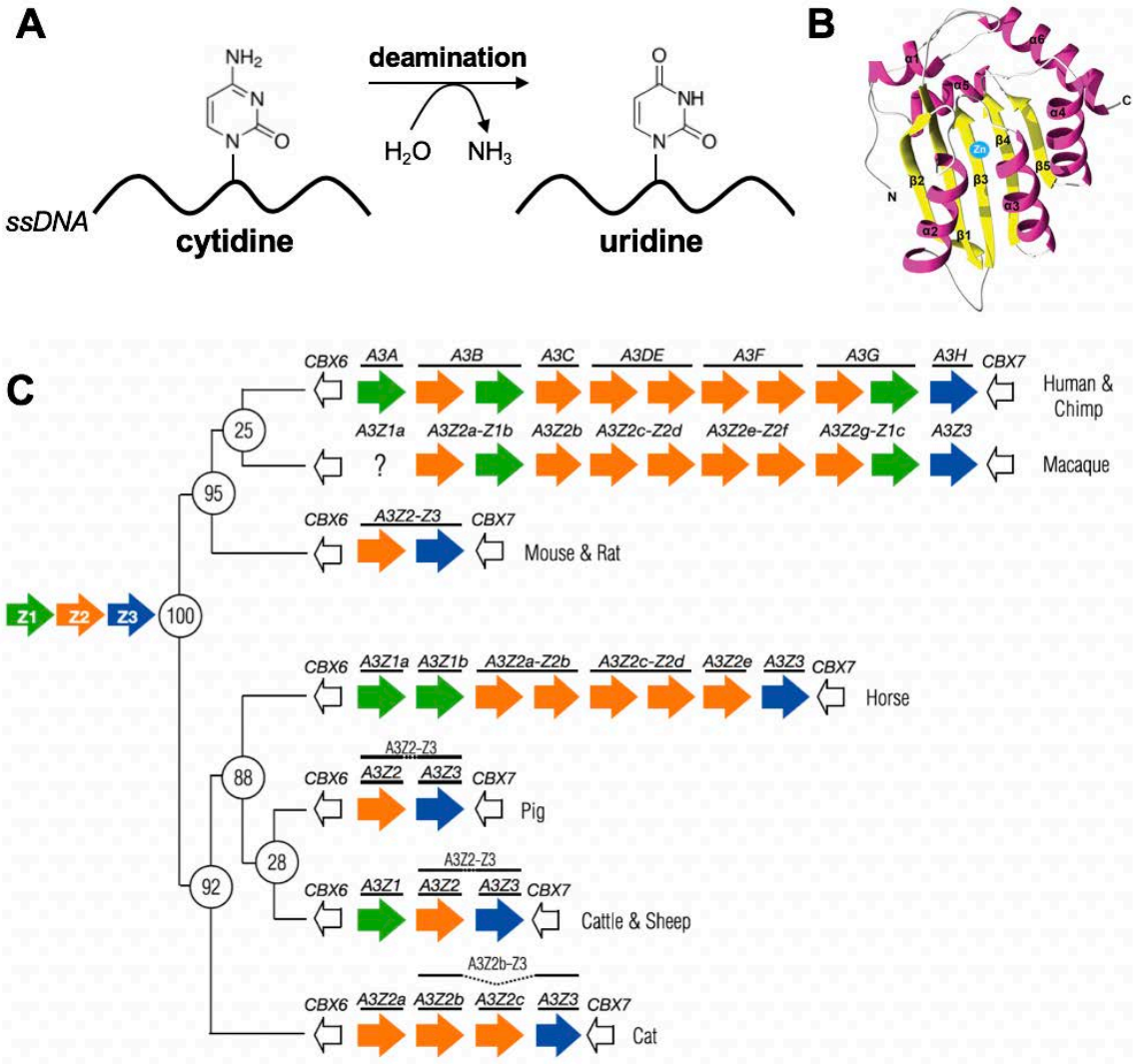

Figure 1-4. A3 family of mammalian cytidine deaminases

(A) Cytidine deamination reaction carried out by A3s. (B) Crystal structure of $A3F_{CTD}$ (PDB ID: 4J4J). Adapted from (Aydin et al. 2014). (C) Schematic of A3 genes in mammals with Z1, Z2, and Z3 domains in green, orange, and blue, respectively. The numbers at tree branch points indicate the estimated time (in millions of years) since divergence of ancestors of present-day species. Adapted from (LaRue et al. 2009).

1.4 Vif targets APOBEC3s for degradation via the proteasome

HIV-1 Vif was first described as a 23 kilodalton protein produced by a short open reading frame overlapping with the *pol* gene (Sodroski et al. 1986; Terwilligier et al. 1986). Although it is not necessary for the production of virions, Vif is essential for HIV-1 infectivity in certain T cell lineages (Strebel et al. 1987; Fisher et al. 1987). These observations led to the hypothesis that Vif counteracts a potent restriction factor of HIV-1 (Madani and Kabat 1998). After identifying A3G as the human restriction factor targeted by Vif (Sheehy et al. 2002), it was further demonstrated that HIV-1 Vif targets A3G for degradation by the proteasome and thereby reduces the amount of A3G incorporated into new virions (Sheehy et al. 2003; Conticello et al. 2003; Kao et al. 2003; Marin et al. 2003;). To accomplish this, HIV-1 Vif recruits A3G to an E3 ligase complex to be ubiquitinated and subsequently degraded by the proteasome (Yu et al. 2003; Yu et al. 2004; Mehle et al. 2004, Figure 1-5A). Its essential role in HIV-1 infectivity makes Vif an attractive target for novel HIV-1 therapies. As a result, elucidating the molecular details allowing HIV-1 Vif to hijack cellular ubiquitin proteasome machinery has become an area of intense research.

HIV-1 Vif hijacks a cellular E3 ligase complex composed a scaffolding protein (Cul5), adaptor proteins (EloB/C), an E2-activating protein (Rbx2), and a non-canonical cofactor (CBFβ). The requirement for CBFβ in this viral E3 ligase complex was surprising because CBFβ is a transcription co-factor not usually involved in the ubiquitin-proteasome system. Nevertheless, CBFβ is required for HIV-1 Vif-mediated degradation of A3G and it is essential for the purification of stable HIV-1 Vif *in vitro* (Jager et al. 2011; Zhang et al. 2012). While Vif proteins from other lentiviruses use similar cellular

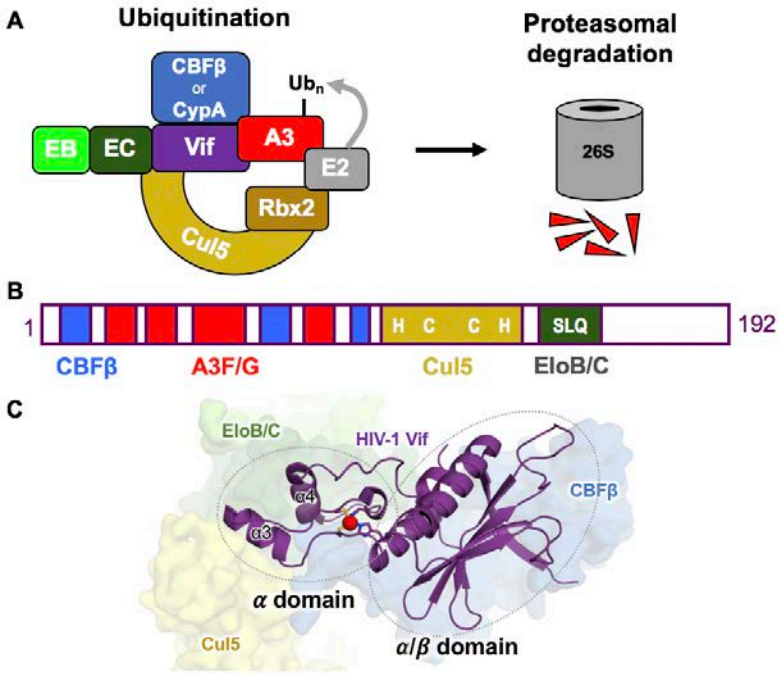

Figure 1-5. HIV-1 Vif degrades A3s via proteasomal degradation

(A) Cartoon schematic of viral E3 ligase mediated by lentiviral Vif that leads to A3 degradation by the proteasome. HIV-1 Vif uses CBFβ as a cofactor, but MVV Vif uses CypA. (B) Bar diagram of HIV-1 Vif binding motifs for cellular factors. (C) Crystal structure of HIV-1 Vif (purple cartoon) with Cul5, EloB/C, and CBFβ (yellow, green, and blue surfaces, respectively). Zinc ion shown as red sphere. Domains of HIV-1 Vif are highlighted. PDB ID: 4N9F (Guo et al. 2014).

E3 ligase components to degrade the cognate A3s in their hosts, only primate lentiviruses use CBFβ as a cofactor. Interestingly, Vif proteins from small ruminant lentiviruses have been shown to use the unrelated CyclophilinA (CypA) instead (Kane et al. 2015; Zhao et al. 2019). CypA is not related to CBFβ in structure, so it is unclear how the two proteins both act as molecular chaperones for Vif proteins. Furthermore, both CypA and CBFβ are highly conserved amongst the mammalian hosts of lentiviruses, so it is unknown why two different lentiviral lineages have evolved the use of two distinct and mutually exclusive cofactors for Vif. Since different lentiviral Vifs are not well conserved, it is possible that

there are significant structural or biochemical differences between Vifs from different lineages.

Many mutational studies have contributed to our understanding of binding motifs and domains required for different cellular factors binding to HIV-1 Vif (Figure 1-5B). CBFβ and A3 interactions have mostly been mapped to the N-terminus of HIV-1 Vif (Mehle et al. 2007; Fribourgh et al. 2014). A PPLP oligomerization motif has also been implicated in HIV-1 Vif binding to A3G (Donahue et al. 2008; Yang et al. 2001). Mutation of the well-conserved BC box (SLQxLA) has been shown to abrogate EloB/C binding to many lentiviral Vifs (Yu et al. 2003; Zhang et al. 2014; Wang et al. 2011). The HCCH zinc binding motif has been shown to be important for HIV-1 Vif (Luo et al. 2005; Xiao et al. 2006; Mehle et al. 2006). While other lentiviral Vif proteins also bind a zinc ion, the exact spacing and identity of the residues coordinating the zinc are not conserved suggesting some flexibility in the structural features of Vifs.

The crystal structure of HIV-1 Vif revealed an entirely novel fold containing a larger domain named the α/β domain that interacts extensively with CBFβ and a smaller helical domain, named the α domain, that interacts with Cul5 and EloB/C (Guo et al. 2014; Figure 1-5C). Consistent with the mutational studies suggesting that the zinc binding HCCH motif was important for Cul5 binding, the structure showed that the zinc coordination helps to stabilize the α domain in a position to interact with Cul5. Supporting CBFβ's role as a structural cofactor, the 4,979 $Å^2$ buried surface area between the α/β domain of HIV-1 and CBFβ is quite extensive (Guo et al. 2014). Besides the vast hydrophobic surface that maps to the N-terminus of HIV-1 Vif, two notable features of the interaction include: (1) an N-terminal β-sheet of HIV-1 Vif which inserts itself into a

β-sheet of CBFβ and (2) C-terminus residues 149-155 of CBFβ which extend into the α domain and zinc finger. Truncations of CBFβ that eliminate its interaction with the zinc finger of HIV-1 Vif did not drastically affect binding to Vif (Zhou et al. 2012). Similarly, mutating the zinc finger of HIV-1 Vif only partially diminished binding to CBFβ (Fribourgh et al. 2014). While the HIV-1 Vif zinc finger may contribute to binding to CBFβ, it is not essential.

While a detailed understanding of how HIV-1 Vif interacts with E3 ligase components is becoming clear, there are still many unanswered questions. For example, it is unclear how HIV-1 Vif uses distinct binding surfaces to interact with different A3 proteins (Aydin et al. 2014). Furthermore, there is no structural information available for Vif proteins from other lentiviruses. A major focus of this work is to understand the molecular determinants involved with recruiting cellular factors to the viral E3 ligase mediated by MVV Vif. While some binding interactions may be well conserved among MVV Vif and HIV-1 Vif, we expect there to be significant structural consequences of the use of CypA as a cofactor instead of CBFβ.

1.5 SAMHD1 restricts HIV-1 by depleting cellular nucleotide pool

The sterile alpha motif (SAM) and histidine-aspartate (HD) domain–containing protein 1 (SAMHD1) was first described as a human homolog of a mouse IFN-gamma induced protein from dendritic cells (Li et al. 2000). The importance of SAMHD1 in human development and innate immunity was highlighted by the discovery that SAMHD1 mutations can cause the rare genetic disorder, Aicardi-Goutieres syndrome (AGS) (Rice et al. 2009). AGS patients suffer from severe neurological defects and

symptoms similar to an auto-immune disease or congenital viral infection thought to be caused by an accumulation of nucleic acids (Aicardi and Goutieres 1984; Crow and Livingston 2008). Interestingly, SAMHD1 was also identified to be an HIV-1 restriction factor counteracted by the lentiviral accessory protein Vpx (Berger et al. 2011; Laguette et al. 2011; Hrecka et al. 2011).

SAMHD1 is a deoxyribonucleotide (dNTP) triphosphohydrolase that cleaves inorganic triphosphate from dNTPs (Goldstone et al. 2011; Powell et al. 2011). The SAMHD1 protein consists of the two domains for which it is named (SAM and HD domains). The smaller SAM domain (residues 1-112) is thought to mediate protein-protein and protein-nucleic acid interactions, whereas the HD domain (residues 113-626) contains catalytic activity (Qiao et al. 2005; Aravind et al. 1998). SAMHD1 can catalyze the hydrolysis of any of the four dNTPs (Figure 1-6A). In addition to the catalytic pocket, SAMHD1 also binds dNTPs at allosteric sites which induce the tetramerization of the protein and is required for dNTPase activity (Figure 1-6B; Yan et al. 2013; Ji et al. 2013; Zhu et al. 2013). To form the tetramer, only dGTP or GTP can bind Allosteric Site 1 and any dNTP can bind the second Allosteric site 2 (Amie et al. 2013; Hansen et al. 2013; Ji et al. 2014).

SAMHD1 inhibits the reverse transcription step of HIV-1 infection by depleting the cellular dNTP levels to concentrations lower than the affinity of HIV-1 reverse transcriptase for dNTPs (Lahoussa et al. 2012; Kim et al. 2012). The HD domain is sufficient for HIV-1 restriction (White et al. 2013). Although it has been suggested that SAMHD1 also degrades incoming HIV-1 genomic RNA, SAMHD1 nuclease activity and

its role in HIV-1 restriction remain controversial (Belaglazova et al. 2013; Ryoo et al. 2014; Seamon et al. 2015; Antonucci et al. 2016).

HIV-2 and its ancestral SIVs, but not HIV-1, contain a gene called *vpx* which targets SAMHD1 for degradation by the proteasome (Goujon et al. 2007; Mahdi et al. 2018). HIV-2 Vpx recruits SAMHD1 to an E3 ligase complex containing Cul4A, Rbx1, DCAF, and DDB1 to be polyubiquitinated (Srivastava et al. 2008; Ahn et al. 2012). Interestingly, Vpx proteins from different SIV isolates have been shown to recognize different regions of the SAMHD1 protein (Schwefel et al. 2014; Schwefel et al. 2015). Without Vpx or another accessory protein dedicated to down-regulating SAMHD1, the mechanism by which HIV-1 overcomes restriction by SAMHD1 remains unclear.

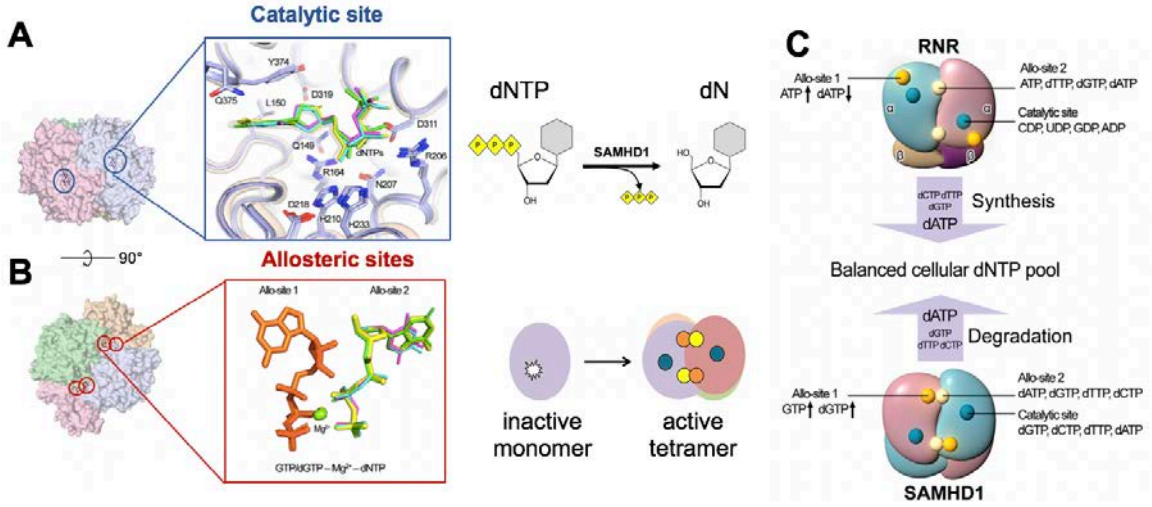

Figure 1-6. SAMHD1 is a dNTPase

(A) dNTPs bound to the catalytic pocket of SAMHD1 and a cartoon schematic of the triphosphohydrolase reaction. Adapted from (Ji et al. 2014). (B) Allosteric pockets of SAMHD1. Adapted from (Ji et al. 2014). (C) Schematic describing biological role of SAMHD1 in balancing the dNTP pool. Adapted from (Ji et al. 2014).

In addition to its role in HIV-1 restriction, SAMHD1 is recognized as an important regulator of nucleotides in the cell (Franzolin et al. 2013). SAMHD1 has been shown to be phosphorylated by CDK2 during the S phase of the cell cycle (Stillman 2013; Welbourn et al. 2013; Cribier et al. 2013; White et al. 2013). It has been proposed that phosphorylation of SAMHD1 may disrupt the tetramer, and impair dNTPase activity (Tang et al. 2015; Arnold et al. 2015). In this way, SAMHD1 activity is regulated to allow high dNTP levels high during S phase, and low dNTP levels during all other phases of the cell cycle. Like SAMHD1, ribonucleotide reductase (RNR) is also regulated by the cell cycle and dNTPs binding to allosteric sites. Together, SAMHD1 and RNR contribute to balancing the dNTP pool and tightly regulating the total level of dNTPs in the cell (Figure 1-6C). This is important for maintaining genomic integrity by influencing the fidelity of DNA synthesis and repair enzymes, which can become low fidelity in the presence of a highly concentrated or biased dNTP pool.

Since dysregulated dNTP pools are linked to genomic instability and tumorigenesis, it has been suggested that SAMHD1 plays a role in tumor suppression by keeping dNTP levels low (Herold et al. 2017; Schaller et al. 2019). SAMHD1 mutations have been found to promote leukemia development in chronic lymphocytic leukemia patients which indicate that SAMHD1 might also play a role in DNA repair (Clifford et al. 2014). Also, low SAMHD1 expression levels have been found in acute myeloid leukemia patient samples (Rassidakis et al. 2018; Jiang et al. 2020). In addition to maintaining the dNTP pool, SAMHD1 has also been implicated in several DNA repair processes that might also contribute to its tumor suppressor activity (Daddacha et al. 2017; Coquel et al. 2018; Cobello-Lobato et al. 2018). While its tumor suppressor

qualities are beneficial, SAMHD1 is also implicated in decreasing the effects of certain chemotherapy treatments, described more below.

1.6 Nucleotide analog drugs and cellular nucleotide metabolism

Nucleotides are important cellular compounds with many essential functions such as the storage of genetic information, signaling cascades, enzyme regulation, and metabolism. Nucleotides have three main chemical components: a nucleobase (adenosine, thymine, cytosine, guanidine, thymine, or uridine), a sugar ring (deoxyribose or ribose), and the presence of up to three phosphate groups.

Nucleotide analogs are a class of drugs that mimic the structure of nucleotides, usually with modifications to one or more of these key chemical components (Figure 1-7A). These drugs are administered to patients as nucleoside prodrugs which enter the cell via specific nucleoside transporters, nucleoside analogs are sequentially phosphorylated to become the active, triphosphorylated form of the drug (Figure 1-7B). This nucleotide analog will then inhibit cell or virus proliferation by interfering with synthesis of genetic material. Many different mechanisms of action have been described for nucleotide analogs, including termination of chain elongation, accumulation of mutations, and direct inhibition of polymerases.

Since cancer cells and viruses require nucleotides to replicate their genomes and actively proliferate, these drugs are attractive chemotherapy and antiviral therapies. However, many nucleotide analog drugs are plagued by issues such as polymerases evolving resistance, poor oral bioavailability, long-term toxicity, and patient-specific variability in efficacy (Jordheim et al. 2013). A deeper understanding of nucleotide

metabolism is necessary to create potent agents with improved properties. In Chapter 2 and the appendix of this thesis, I detail our studies of SAMHD1 interference with nucleotide analog therapies by degrading active triphosphate forms of the drug into non-active nucleoside precursors.

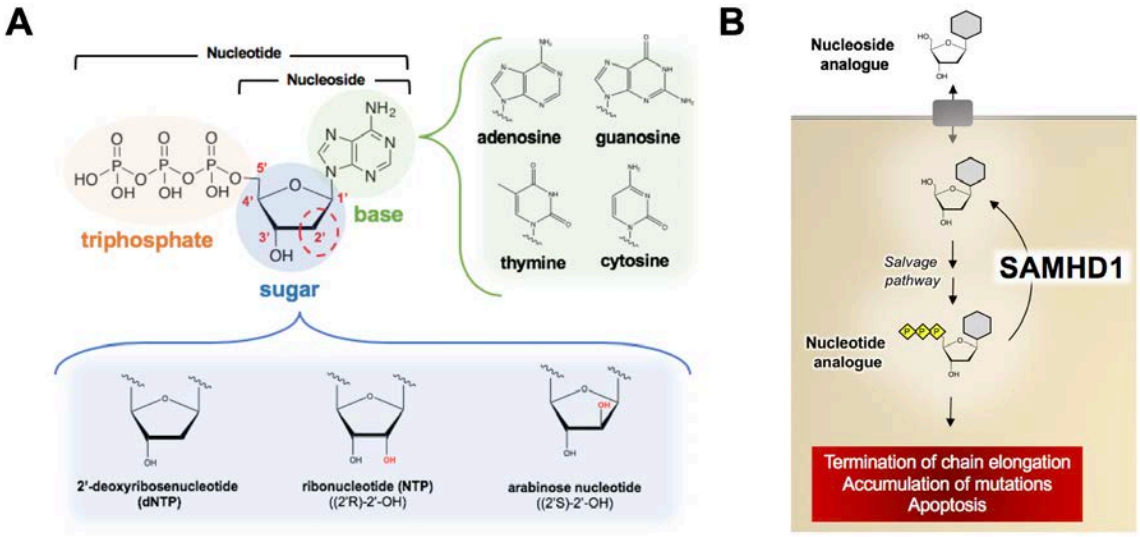

Figure 1-7. Nucleotide analogs

(A) General structure of nucleotide analogs. Canonical nucleobases and chemical modifications to 2' sugar moieties are highlighted. Adapted from (Jordheim et al. 2013). (B) Mechanism of action for nucleotide analogs.

1.7 Aims and Scope

The overall goal of my thesis is to gain a better understanding of host innate immunity in the context of viral infection and cancer. I have used a combination of biochemical and structural approaches to study a viral E3 ligase complex mediated by the MVV Vif and the human enzyme SAMHD1 which harbors the potential to degrade cancer drugs.

In Chapter 2, I probe the interactions of the dNTPase SAMHD1 with nucleotide analog drugs used to treat viral infections and cancers. In this study, I have crystallized each drug with SAMHD1 and characterized their interactions with the catalytic pocket and allosteric sites of the enzyme. These studies provide a framework for understanding how SAMHD1 contributes to challenges associated with nucleotide analog treatments. I also included an appendix which outlines some initial screening for small molecule compounds capable of modulating SAMHD1 activity.

In Chapter 3, I investigate the molecular determinants of MVV Vif binding to cellular cofactors in an E3 ligase complex used to A3s. When compared to previous results with other Vif proteins, results from this section can be used to gain insight into the evolution of lentiviral Vif alongside mammalian host factors. In Chapter 4, I summarize the reconstitution of MVV Vif complexes with CypA, EloBC, and Cul5 used for structural studies. These studies provide insight into the solution behavior of MVV Vif, and will aid in the future determination of the structure of MVV Vif.

1.8 References

Ahn J, Hao C, Yan J, DeLucia M, Mehrens J, Wang C, Gronenborn AM, Skowronski J (2012) HIV/simian immunodeficiency virus (SIV) accessory virulence factor Vpx loads the host cell restriction factor SAMHD1 onto the E3 ubiquitin ligase complex CRL4DCAF1. *J Biol Chem* 287: 12550-8

Aicardi J, Goutieres F (1984) A progressive familial encephalopathy in infancy with calcifications of the basal ganglia and chronic cerebrospinal fluid lymphocytosis. *Ann Neurol* 15: 49-54

Amie SM, Bambara RA, Kim B (2013) GTP is the primary activator of the anti-HIV restriction factor SAMHD1. *J Biol Chem* 288: 25001-6

Aravind L, Koonin EV (1998) The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem Sci* 23: 469-72

Arnold LH, Groom HC, Kunzelmann S, Schwefel D, Caswell SJ, Ordonez P, Mann MC, Rueschenbaum S, Goldstone DC, Pennell S, Howell SA, Stoye JP, Webb M, Taylor IA,

Bishop KN (2015) Phospho-dependent Regulation of SAMHD1 Oligomerisation Couples Catalysis and Restriction. *PLoS Pathog* 11: e1005194

Barre-Sinoussi F, Chermann JC, Rey F, Nugeyre MT, Chamaret S, Gruest J, Dauguet C, Axler-Blin C, Vezinet-Brun F, Rouzioux C, Rozenbaum W, Montagnier L (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science* 220: 868-71

Barre-Sinoussi F, Ross AL, Delfraissy JF (2013) Past, present and future: 30 years of HIV research. *Nat Rev Microbiol* 11: 877-83

Bell NM, Lever AM (2013) HIV Gag polyprotein: processing and early viral particle assembly. *Trends Microbiol* 21: 136-44

Beloglazova N, Flick R, Tchigvintsev A, Brown G, Popovic A, Nocek B, Yakunin AF (2013) Nuclease activity of the human SAMHD1 protein implicated in the Aicardi-Goutieres syndrome and HIV-1 restriction. *J Biol Chem* 288: 8101-10

Berger A, Sommer AF, Zwarg J, Hamdorf M, Welzel K, Esly N, Panitz S, Reuter A, Ramos I, Jatiani A, Mulder LC, Fernandez-Sesma A, Rutsch F, Simon V, Konig R, Flory E (2011) SAMHD1-deficient CD14+ cells from individuals with Aicardi-Goutieres syndrome are highly susceptible to HIV-1 infection. *PLoS Pathog* 7: e1002425

Bosu DR, Kipreos ET (2008) Cullin-RING ubiquitin ligases: global regulation and activation cycles. *Cell Division* 3: 7

Burnett A, Spearman P (2007) APOBEC3G Multimers Are Recruited to the Plasma Membrane for Packaging into Human Immunodeficiency Virus Type 1 Virus-Like Particles in an RNA-Dependent Process Requiring the NC Basic Linker. *Journal of Virology* 81: 5000-5013

Cabello-Lobato MJ, Wang S, Schmidt CK (2017) SAMHD1 Sheds Moonlight on DNA Double-Strand Break Repair. *Trends Genet* 33: 895-897

Chaipan C, Smith JL, Hu WS, Pathak VK (2013) APOBEC3G restricts HIV-1 to a greater extent than APOBEC3F and APOBEC3DE in human primary CD4+ T cells and macrophages. *J Virol* 87: 444-53

Clavel F, Guyader M, Guetard D, Salle M, Montagnier L, Alizon M (1986) Molecular cloning and polymorphism of the human immune deficiency virus type 2. *Nature* 324: 691-5

Clements JE, Zink MC (1996) Molecular biology and pathogenesis of animal lentivirus infections. *Clin Microbiol Rev* 9: 100-17

Clifford R, Louis T, Robbe P, Ackroyd S, Burns A, Timbs AT, Wright Colopy G, Dreau H, Sigaux F, Judde JG, Rotger M, Telenti A, Lin YL, Pasero P, Maelfait J, Titsias M, Cohen DR, Henderson SJ, Ross MT, Bentley D et al. (2014) SAMHD1 is mutated recurrently in chronic lymphocytic leukemia and is involved in response to DNA damage. *Blood* 123: 1021-31

Compton AA, Hirsch VM, Emerman M (2012) The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. *Cell Host Microbe* 11: 91-8

Compton AA, Malik HS, Emerman M (2013) Host gene evolution traces the evolutionary history of ancient primate lentiviruses. *Philos Trans R Soc Lond B Biol Sci* 368: 20120496

Conticello SG, Harris RS, Neuberger MS (2003) The Vif Protein of HIV Triggers Degradation of the Human Antiretroviral DNA Deaminase APOBEC3G. *Current Biology* 13: 2009-2013

Control CfD (1981) Pneumocystis pneumonia--Los Angeles. *MMWR Morb Mortal Wkly Rep* 30: 250-2

Control CfD (1981) Kaposi's sarcoma and Pneumocystis pneumonia among homosexual men--New York City and California. *MMWR Morb Mortal Wkly Rep* 30: 305-8

Coquel F, Silva MJ, Techer H, Zadorozhny K, Sharma S, Nieminuszczy J, Mettling C, Dardillac E, Barthe A, Schmitz AL, Promonet A, Cribier A, Sarrazin A, Niedzwiedz W, Lopez B, Costanzo V, Krejci L, Chabes A, Benkirane M, Lin YL et al. (2018) SAMHD1 acts at stalled replication forks to prevent interferon induction. *Nature* 557: 57-61

Crow YJ, Rehwinkel J (2009) Aicardi-Goutieres syndrome and related phenotypes: linking nucleic acid metabolism with autoimmunity. *Hum Mol Genet* 18: R130-6

Daddacha W, Koyen AE, Bastien AJ, Head PE, Dhere VR, Nabeta GN, Connolly EC, Werner E, Madden MZ, Daly MB, Minten EV, Whelan DR, Schlafstein AJ, Zhang H, Anand R, Doronio C, Withers AE, Shepard C, Sundaram RK, Deng X et al. (2017) SAMHD1 Promotes DNA End Resection to Facilitate DNA Repair by Homologous Recombination. *Cell Rep* 20: 1921-1935

Donahue JP, Vetter ML, Mukhtar NA, D'Aquila RT (2008) The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. *Virology* 377: 49-53

Doyle T, Goujon C, Malim MH (2015) HIV-1 and interferons: who's interfering with whom? *Nat Rev Microbiol* 13: 403-13

Duggal NK, Emerman M (2012) Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat Rev Immunol* 12: 687-95

Fisher AG, Ensoli B, Ivanoff L, Chamberlain M, Petteway S, Ratner L, Gallo RC, Wong-Staal F (1987) The sor gene of HIV-1 is required for efficient virus transmission in vitro. *Science* 237: 888-93

Franzolin E, Pontarin G, Rampazzo C, Miazzi C, Ferraro P, Palumbo E, Reichard P, Bianchi V (2013) The deoxynucleotide triphosphohydrolase SAMHD1 is a major regulator of DNA precursor pools in mammalian cells. *Proc Natl Acad Sci U S A* 110: 14272-7

Fribourgh JL, Nguyen HC, Wolfe LS, Dewitt DC, Zhang W, Yu XF, Rhoades E, Xiong Y (2014) Core binding factor beta plays a critical role by facilitating the assembly of the Vif-cullin 5 E3 ubiquitin ligase. *J Virol* 88: 3309-19

Gabuzda DH, Lawrence K, Langhoff E, Terwilliger E, Dorfman T, Haseltine WA, Sodroski J (1992) Role of vif in replication of human immunodeficiency virus type 1 in CD4+ T lymphocytes. *J Virol* 66: 6489-95

Gallo RC, Sarin PS, Gelmann EP, Robert-Guroff M, Richardson E, Kalyanaraman VS, Mann D, Sidhu GD, Stahl RE, Zolla-Pazner S, Leibowitch J, Popovic M (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science* 220: 865-7

Goldenberg SJ, Cascio TC, Shumway SD, Garbutt KC, Liu J, Xiong Y, Zheng N (2004) Structure of the Cand1-Cul1-Roc1 complex reveals regulatory mechanisms for the assembly of the multisubunit cullin-dependent ubiquitin ligases. *Cell* 119: 517-28

Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, Walker PA, Kelly G, Haire LF, Yap MW, de Carvalho LP, Stoye JP, Crow YJ, Taylor IA, Webb M (2011) HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* 480: 379-82

Gonda MA, Wong-Staal F, Gallo RC, Clements JE, Narayan O, Gilden RV (1985) Sequence homology and morphologic similarity of HTLV-III and visna virus, a pathogenic lentivirus. *Science* 227: 173-7

Goujon C, Riviere L, Jarrosson-Wuilleme L, Bernaud J, Rigal D, Darlix JL, Cimarelli A (2007) SIVSM/HIV-2 Vpx proteins promote retroviral escape from a proteasome-dependent restriction pathway present in human dendritic cells. *Retrovirology* 4: 2

Guo Y, Dong L, Qiu X, Wang Y, Zhang B, Liu H, Yu Y, Zang Y, Yang M, Huang Z (2014) Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* 505: 229-33

Hansen EC, Seamon KJ, Cravens SL, Stivers JT (2014) GTP activator and dNTP substrates of HIV-1 restriction factor SAMHD1 generate a long-lived activated state. *Proc Natl Acad Sci U S A* 111: E1843-51

Harris RS, Petersen-Mahrt SK, Neuberger MS (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* 10: 1247-53

Harris RS, Petersen-Mahrt SK, Neuberger MS (2002) RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell* 10: 1247-53

Herold N, Rudd SG, Sanjiv K, Kutzner J, Myrberg IH, Paulin CBJ, Olsen TK, Helleday T, Henter JI, Schaller T (2017) With me or against me: Tumor suppressor and drug resistance activities of SAMHD1. *Exp Hematol* 52: 32-39

Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, Florens L, Washburn MP, Skowronski J (2011) Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474: 658-61

Hultquist JF, Lengyel JA, Refsland EW, LaRue RS, Lackey L, Brown WL, Harris RS (2011) Human and rhesus APOBEC3D, APOBEC3F, APOBEC3G, and APOBEC3H demonstrate a conserved capacity to restrict Vif-Deficient HIV-1. *Journal of Virology* 85: 11220-11234

Jager S, Kim DY, Hultquist JF, Shindo K, LaRue RS, Kwon E, Li M, Anderson BD, Yen L, Stanley D, Mahon C, Kane J, Franks-Skiba K, Cimermancic P, Burlingame A, Sali A, Craik CS, Harris RS, Gross JD, Krogan NJ (2011) Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* 481: 371-5

Ji X, Tang C, Zhao Q, Wang W, Xiong Y (2014) Structural basis of cellular dNTP regulation by SAMHD1. *Proc Natl Acad Sci U S A* 111: E4305-14

Ji X, Wu Y, Yan J, Mehrens J, Yang H, DeLucia M, Hao C, Gronenborn AM, Skowronski J, Ahn J, Xiong Y (2013) Mechanism of allosteric activation of SAMHD1 by dGTP. *Nat Struct Mol Biol* 20: 1304-9

Jia X, Zhao Q, Xiong Y (2015) HIV suppression by host restriction factors and viral immune evasion. *Curr Opin Struct Biol* 31: 106-14

Jiang H, Li C, Liu Z, Shengjing H (2019) Expression and Relationship of SAMHD1 with Other Apoptotic and Autophagic Genes in Acute Myeloid Leukemia Patients. *Acta Haematologica*

Jonsson SR, Hache G, Stenglein MD, Fahrenkrug SC, Andresdottir V, Harris RS (2006) Evolutionarily conserved and non-conserved retrovirus restriction activities of artiodactyl APOBEC3F proteins. *Nucleic Acids Res* 34: 5683-94

Jordheim LP, Durantel D, Zoulim F, Dumontet C (2013) Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat Rev Drug Discov* 12: 447-64

Kane JR, Stanley DJ, Hultquist JF, Johnson JR, Mietrach N, Binning JM, Jonsson SR, Barelier S, Newton BW, Johnson TL, Franks-Skiba KE, Li M, Brown WL, Gunnarsson HI, Adalbjornsdottir A, Fraser JS, Harris RS, Andresdottir V, Gross JD, Krogan NJ (2015) Lineage-Specific Viral Hijacking of Non-canonical E3 Ubiquitin Ligase Cofactors in the Evolution of Vif Anti-APOBEC3 Activity. *Cell Rep* 11: 1236-50

Kao S, Khan MA, Miyagi E, Plishka R, Buckler-White A, Strebel K (2003) The Human Immunodeficiency Virus Type 1 Vif Protein Reduces Intracellular Expression and Inhibits Packaging of APOBEC3G (CEM15), a Cellular Inhibitor of Virus Infectivity. *Journal of Virology* 77: 11398-11407

Khan MA, Kao S, Miyagi E, Takeuchi H, Goila-Gaur R, Opi S, Gipson CL, Parslow TG, Ly H, Strebel K (2005) Viral RNA is required for the association of APOBEC3G with human immunodeficiency virus type 1 nucleoprotein complexes. *J Virol* 79: 5870-4

Kim B, Nguyen LA, Daddacha W, Hollenbaugh JA (2012) Tight interplay among SAMHD1 protein level, cellular dNTP levels, and HIV-1 proviral DNA synthesis kinetics in human primary monocyte-derived macrophages. *J Biol Chem* 287: 21570-4

Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Segeral E, Yatim A, Emiliani S, Schwartz O, Benkirane M (2011) SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474: 654-7

Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, Bloch N, Maudet C, Bertrand M, Gramberg T, Pancino G, Priet S, Canard B, Laguette N, Benkirane M, Transy C, Landau NR, Kim B, Margottin-Goguet F (2012) SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol* 13: 223-228

Larue RS, Lengyel J, Jonsson SR, Andresdottir V, Harris RS (2010) Lentiviral Vif degrades the APOBEC3Z3/APOBEC3H protein of its mammalian host and is capable of cross-species activity. *J Virol* 84: 8193-201

Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB (2017) Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Research* 46: D708-D717

Li N, Zhang W, Cao X (2000) Identification of human homologue of mouse IFN-gamma induced protein from human dendritic cells. *Immunol Lett* 74: 221-4

Liddament MT, Brown WL, Schumacher AJ, Harris RS (2004) APOBEC3F properties and hypermutation preferences indicate activity against HIV-1 in vivo. *Curr Biol* 14: 1385-91

Luo K, Xiao Z, Ehrlich E, Yu Y, Liu B, Zheng S, Yu XF (2005) Primate lentiviral virion infectivity factors are substrate receptors that assemble with cullin 5-E3 ligase through a HCCH motif to suppress APOBEC3G. *Proc Natl Acad Sci U S A* 102: 11444-9

Madani N, Kabat D (1998) An endogenous inhibitor of human immunodeficiency virus in human lymphocytes is overcome by the viral Vif protein. *J Virol* 72: 10251-5

Mahdi M, Szojka Z, Mótyán JA, Tőzsér J (2018) Inhibitory Effects of HIV-2 Vpx on Replication of HIV-1. *Journal of Virology* 92: e00554-18

Malikov V, da Silva ES, Jovasevic V, Bennett G, de Souza Aranha Vieira DA, Schulte B, Diaz-Griffero F, Walsh D, Naghavi MH (2015) HIV-1 capsids bind and exploit the kinesin-1 adaptor FEZ1 for inward movement to the nucleus. *Nat Commun* 6: 6660

Marin M, Rose KM, Kozak SL, Kabat D (2003) HIV-1 Vif protein binds the editing enzyme APOBEC3G and induces its degradation. *Nat Med* 9: 1398-403

Mehle A, Strack B, Ancuta P, Zhang C, McPike M, Gabuzda D (2004) Vif overcomes the innate antiviral activity of APOBEC3G by promoting its degradation in the ubiquitin-proteasome pathway. *J Biol Chem* 279: 7792-8

Mehle A, Thomas ER, Rajendran KS, Gabuzda D (2006) A zinc-binding region in Vif binds Cul5 and determines cullin selection. *J Biol Chem* 281: 17259-65

Mehle A, Wilson H, Zhang C, Brazier AJ, McPike M, Pery E, Gabuzda D (2007) Identification of an APOBEC3G binding site in human immunodeficiency virus type 1 Vif and inhibitors of Vif-APOBEC3G binding. *J Virol* 81: 13235-41

Minardi da Cruz JC, Singh DK, Lamara A, Chebloune Y (2013) Small ruminant lentiviruses (SRLVs) break the species barrier to acquire new host range. *Viruses* 5: 1867-84

Munk C, Beck T, Zielonka J, Hotz-Wagenblatt A, Chareza S, Battenberg M, Thielebein J, Cichutek K, Bravo IG, O'Brien SJ, Lochelt M, Yuhki N (2008) Functions, structure, and read-through alternative splicing of feline APOBEC3 genes. *Genome Biol* 9: R48

Nakagawa F, May M, Phillips A (2013) Life expectancy living with HIV: recent estimates and future implications. *Curr Opin Infect Dis* 26: 17-25

Naruse TK, Sakurai D, Ohtani H, Sharma G, Sharma SK, Vajpayee M, Mehra NK, Kaur G, Kimura A (2016) APOBEC3H polymorphisms and susceptibility to HIV-1 infection in an Indian population. *J Hum Genet* 61: 263-5

Nisole S, Saib A (2004) Early steps of retrovirus replicative cycle. *Retrovirology* 1: 9

Organization WH (2018) Summary of Global HIV Epidemic. In

Petursson G, Andresdottir V, Andresson O, Torsteinsdottir S, Georgsson G, Palsson PA (1991) Human and ovine lentiviral infections compared. *Comp Immunol Microbiol Infect Dis* 14: 277-87

Powell RD, Holland PJ, Hollis T, Perrino FW (2011) Aicardi-Goutieres syndrome gene and HIV-1 restriction factor SAMHD1 is a dGTP-regulated deoxynucleotide triphosphohydrolase. *J Biol Chem* 286: 43596-600

Qiao F, Bowie JU (2005) The many faces of SAM. *Sci STKE* 2005: re7

Ramírez H, Reina R, Amorena B, de Andrés D, Martínez HA (2013) Small ruminant lentiviruses: genetic variability, tropism and diagnosis. *Viruses* 5: 1175-207

Rassidakis GZ, Herold N, Myrberg IH, Tsesmetzis N, Rudd SG, Henter JI, Schaller T, Ng SB, Chng WJ, Yan B, Ng CH, Ravandi F, Andreeff M, Kantarjian HM, Medeiros LJ, Xagoraris I, Khoury JD (2018) Low-level expression of SAMHD1 in acute myeloid leukemia (AML) blasts correlates with improved outcome upon consolidation chemotherapy with high-dose cytarabine-based regimens. *Blood Cancer J* 8: 98

Rassidakis GZ, Herold N, Myrberg IH, Tsesmetzis N, Rudd SG, Henter JI, Schaller T, Ng SB, Chng WJ, Yan B, Ng CH, Ravandi F, Andreeff M, Kantarjian HM, Medeiros LJ, Xagoraris I, Khoury JD (2018) Low-level expression of SAMHD1 in acute myeloid leukemia (AML) blasts correlates with improved outcome upon consolidation chemotherapy with high-dose cytarabine-based regimens. *Blood Cancer J* 8: 98

Rawle DJ, Harrich D (2018) Toward the "unravelling" of HIV: Host cell factors involved in HIV-1 core uncoating. *PLoS Pathog* 14: e1007270

Refsland EW, Hultquist JF, Luengas EM, Ikeda T, Shaban NM, Law EK, Brown WL, Reilly C, Emerman M, Harris RS (2014) Natural polymorphisms in human APOBEC3H and HIV-1 Vif combine in primary T lymphocytes to affect viral G-to-A mutation levels and infectivity. *PLoS Genet* 10: e1004761

Rice GI, Bond J, Asipu A, Brunette RL, Manfield IW, Carr IM, Fuller JC, Jackson RM, Lamb T, Briggs TA, Ali M, Gornall H, Couthard LR, Aeby A, Attard-Montalto SP, Bertini E, Bodemer C, Brockmann K, Brueton LA, Corry PC et al. (2009) Mutations involved in Aicardi-Goutieres syndrome implicate SAMHD1 as regulator of the innate immune response. *Nat Genet* 41: 829-32

Rojas VK, Park IW (2019) Role of the Ubiquitin Proteasome System (UPS) in the HIV-1 Life Cycle. *Int J Mol Sci* 20

Ryoo J, Choi J, Oh C, Kim S, Seo M, Kim SY, Seo D, Kim J, White TE, Brandariz-Nunez A, Diaz-Griffero F, Yun CH, Hollenbaugh JA, Kim B, Baek D, Ahn K (2014) The ribonuclease activity of SAMHD1 is required for HIV-1 restriction. *Nat Med* 20: 936-41

Sawyer SL, Emerman M, Malik HS (2004) Ancient Adaptive Evolution of the Primate Antiviral DNA-Editing Enzyme APOBEC3G. *PLOS Biology* 2: e275

Schaller T, Herold N (2019) Evidence for SAMHD1 Tumor Suppressor Functions in Acute Myeloid Leukemia. *Acta Haematol*: 1-2

Schwefel D, Boucherit VC, Christodoulou E, Walker PA, Stoye JP, Bishop KN, Taylor IA (2015) Molecular determinants for recognition of divergent SAMHD1 proteins by the lentiviral accessory protein Vpx. *Cell Host Microbe* 17: 489-99

Schwefel D, Groom HC, Boucherit VC, Christodoulou E, Walker PA, Stoye JP, Bishop KN, Taylor IA (2014) Structural basis of lentiviral subversion of a cellular protein degradation pathway. *Nature* 505: 234-8

Seamon KJ, Sun Z, Shlyakhtenko LS, Lyubchenko YL, Stivers JT (2015) SAMHD1 is a single-stranded nucleic acid binding protein with no active site-associated nuclease activity. *Nucleic Acids Res* 43: 6486-99

Seissler T, Marquet R, Paillart JC (2017) Hijacking of the Ubiquitin/Proteasome Pathway by the HIV Auxiliary Proteins. *Viruses* 9

Sharp PM, Hahn BH (2011) Origins of HIV and the AIDS pandemic. *Cold Spring Harb Perspect Med* 1: a006841

Sheehy AM, Gaddis NC, Choi JD, Malim MH (2002) Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. *Nature* 418: 646-50

Sheehy AM, Gaddis NC, Malim MH (2003) The antiretroviral enzyme APOBEC3G is degraded by the proteasome in response to HIV-1 Vif. *Nat Med* 9: 1404-7

Simon F, Mauclere P, Roques P, Loussert-Ajaka I, Muller-Trutwin MC, Saragosti S, Georges-Courbot MC, Barre-Sinoussi F, Brun-Vezinet F (1998) Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nat Med* 4: 1032-7

Simon JH, Gaddis NC, Fouchier RA, Malim MH (1998) Evidence for a newly discovered cellular anti-HIV-1 phenotype. *Nat Med* 4: 1397-400

Sodroski J, Goh W, Rosen C, Tartar A, Portetelle D, Burny A, Haseltine W (1986) Replicative and cytopathic potential of HTLV-III/LAV with sor gene deletions. *Science* 231: 1549-1553

Sonigo P, Alizon M, Staskus K, Klatzmann D, Cole S, Danos O, Retzel E, Tiollais P, Haase A, Wain-Hobson S (1985) Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. *Cell* 42: 369-82

Srivastava S, Swanson SK, Manel N, Florens L, Washburn MP, Skowronski J (2008) Lentiviral Vpx Accessory Factor Targets VprBP/DCAF1 Substrate Adaptor for Cullin 4 E3 Ubiquitin Ligase to Enable Macrophage Infection. *PLOS Pathogens* 4: e1000059

Stillman B (2013) Deoxynucleoside triphosphate (dNTP) synthesis and destruction regulate the replication of both cell and virus genomes. *Proc Natl Acad Sci U S A* 110: 14120-1

Strebel K, Daugherty D, Clouse K, Cohen D, Folks T, Martin MA (1987) The HIV 'A' (sor) gene product is essential for virus infectivity. *Nature* 328: 728-30

Svarovskaia ES, Xu H, Mbisa JL, Barr R, Gorelick RJ, Ono A, Freed EO, Hu WS, Pathak VK (2004) Human apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G (APOBEC3G) is incorporated into HIV-1 virions through interactions with viral and nonviral RNAs. *J Biol Chem* 279: 35822-8

Tang C, Ji X, Wu L, Xiong Y (2015) Impaired dNTPase activity of SAMHD1 by phosphomimetic mutation of Thr-592. *J Biol Chem* 290: 26352-9

Taylor BS, Sobieszczyk ME, McCutchan FE, Hammer SM (2008) The challenge of HIV-1 subtype diversity. *N Engl J Med* 358: 1590-602

Terwilliger E, Sodroski JG, Rosen CA, Haseltine WA (1986) Effects of mutations within the 3' orf open reading frame region of human T-cell lymphotropic virus type III (HTLV-III/LAV) on replication and cytopathogenicity. *J Virol* 60: 754-60

Wang J, Zhang W, Lv M, Zuo T, Kong W, Yu X (2011) Identification of a Cullin5-ElonginB-ElonginC E3 complex in degradation of feline immunodeficiency virus Vif-mediated feline APOBEC3 proteins. *J Virol* 85: 12482-91

Wang X, Abudu A, Son S, Dang Y, Venta PJ, Zheng Y-H (2011) Analysis of Human APOBEC3H Haplotypes and Anti-Human Immunodeficiency Virus Type 1 Activity. *Journal of Virology* 85: 3142-3152

Ward AB, Wilson IA (2017) The HIV-1 envelope glycoprotein structure: nailing down a moving target. *Immunol Rev* 275: 21-32

White TE, Brandariz-Nunez A, Valle-Casuso JC, Amie S, Nguyen L, Kim B, Brojatsch J, Diaz-Griffero F (2013) Contribution of SAM and HD domains to retroviral restriction mediated by human SAMHD1. *Virology* 436: 81-90

White TE, Brandariz-Nunez A, Valle-Casuso JC, Amie S, Nguyen LA, Kim B, Tuzova M, Diaz-Griffero F (2013) The retroviral restriction ability of SAMHD1, but not its deoxynucleotide triphosphohydrolase activity, is regulated by phosphorylation. *Cell Host Microbe* 13: 441-51

Wiegand HL, Doehle BP, Bogerd HP, Cullen BR (2004) A second human antiretroviral factor, APOBEC3F, is suppressed by the HIV-1 and HIV-2 Vif proteins. *Embo j* 23: 2451-8

Xavier Ruiz F, Arnold E (2020) Evolving understanding of HIV-1 reverse transcriptase structure, function, inhibition, and resistance. *Current Opinion in Structural Biology* 61: 113-123

Xiao Z, Ehrlich E, Yu Y, Luo K, Wang T, Tian C, Yu XF (2006) Assembly of HIV-1 Vif-Cul5 E3 ubiquitin ligase through a novel zinc-binding domain-stabilized hydrophobic interface in Vif. *Virology* 349: 290-9

Yan J, Kaur S, DeLucia M, Hao C, Mehrens J, Wang C, Golczak M, Palczewski K, Gronenborn AM, Ahn J, Skowronski J (2013) Tetramerization of SAMHD1 is required for biological activity and inhibition of HIV infection. *J Biol Chem* 288: 10406-17

Yang S, Sun Y, Zhang H (2001) The multimerization of human immunodeficiency virus type I Vif protein: a requirement for Vif function in the viral life cycle. *J Biol Chem* 276: 4889-93

Yu Q, Chen D, Konig R, Mariani R, Unutmaz D, Landau NR (2004) APOBEC3B and APOBEC3C are potent inhibitors of simian immunodeficiency virus replication. *J Biol Chem* 279: 53379-86

Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, Yu XF (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. *Science* 302: 1056-60

Yu Y, Xiao Z, Ehrlich ES, Yu X, Yu XF (2004) Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. *Genes Dev* 18: 2867-72

Zennou V, Perez-Caballero D, Gottlinger H, Bieniasz PD (2004) APOBEC3G incorporation into human immunodeficiency virus type 1 particles. *J Virol* 78: 12058-61

Zhang H, Yang B, Pomerantz RJ, Zhang C, Arunachalam SC, Gao L (2003) The cytidine deaminase CEM15 induces hypermutation in newly synthesized HIV-1 DNA. *Nature* 424: 94-8

Zhang W, Du J, Evans SL, Yu Y, Yu XF (2011) T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. *Nature* 481: 376-9

Zhang W, Wang H, Li Z, Liu X, Liu G, Harris RS, Yu XF (2014) Cellular requirements for bovine immunodeficiency virus Vif-mediated inactivation of bovine APOBEC3 proteins. *J Virol* 88: 12528-40

Zhao Z, Li Z, Huan C, Wang H, Su X, Zhang W (2019) CAEV Vif Hijacks ElonginB/C, CYPA and Cullin5 to Assemble the E3 Ubiquitin Ligase Complex Stepwise to Degrade oaA3Z2-Z3. *Front Microbiol* 10: 565

Zheng YH, Irwin D, Kurosu T, Tokunaga K, Sata T, Peterlin BM (2004) Human APOBEC3F is another host factor that blocks human immunodeficiency virus type 1 replication. *J Virol* 78: 6073-6

Zhou X, Evans SL, Han X, Liu Y, Yu XF (2012) Characterization of the interaction of full-length HIV-1 Vif protein with its key regulator CBFbeta and CRL5 E3 ubiquitin ligase components. *PLoS One* 7: e33495

Zhu C, Gao W, Zhao K, Qin X, Zhang Y, Peng X, Zhang L, Dong Y, Zhang W, Li P, Wei W, Gong Y, Yu XF (2013) Structural insight into dGTP-dependent activation of tetrameric SAMHD1 deoxynucleoside triphosphate triphosphohydrolase. *Nat Commun* 4: 2722

Zielonka J, Bravo IG, Marino D, Conrad E, Perković M, Battenberg M, Cichutek K, Münk C (2009) Restriction of Equine Infectious Anemia Virus by Equine APOBEC3 Cytidine Deaminases. *Journal of Virology* 83: 7547-7559

2 The structural basis for cancer drug interactions with the catalytic and allosteric sites of SAMHD1

2.1 Preface

The work outlined in this chapter has been published in *PNAS* (Knecht* and Buzovetsky* et al. 2018). *Indicates co-first authors. K.M.K., O.B., C.S., D.T., N.F., J.C., O.T.K., and Y.X. designed the research; K.M.K., O.B., C.S., D.T., and V.S. performed the research; K.M.K., O.B., C.S., D.T., F.T., K.R., N.F., G.G., V.B., X.J., O.T.K., J.C., and Y.X. analyzed the data; and K.M.K., O.B., and Y.X. wrote the paper with contributions from O.T.K. and J.C. K.M.K and O.B contributed to the optimization of crystals and structural analysis that resulted in Figures 2-1, 2-5 and 2-6. K.M.K. and O.B. contributed to the oligomerization assays shown in Figures 2-4A and 2-4B. O.B generated data for the HPLC-based activity assay in Figure 2-2C. K.M.K generated malachite-based activity assay data shown in Figures 2-2D, 2-4C, and 2-4D. Our colleagues in the Cinatl Group at Goethe University of Frankfurt and the Keppler group at Ludwig Maximilians University generated the *in vivo* data shown in Figure 2-3.

2.2 Summary

Sterile alpha motif and histidine-aspartate domain–containing protein 1 (SAMHD1) is a deoxynucleoside triphosphate triphosphohydrolase (dNTPase) that depletes cellular dNTPs in non-cycling cells to promote genome stability and to inhibit retroviral and herpes viral replication. In addition to being substrates, cellular nucleotides also allosterically regulate SAMHD1 activity. Recently, it was shown that high expression levels of SAMHD1 are also correlated with significantly worse patient

responses to nucleotide analog drugs that are important for treating a variety of cancers, including acute myeloid leukemia (AML). In this study, we used biochemical, structural, and cellular methods to examine the interactions of various cancer drugs with SAMHD1. We found that both the catalytic and the allosteric sites of SAMHD1 are sensitive to sugar modifications of the nucleotide analogs, with the allosteric site being significantly more restrictive. Nucleotide analogs, including cladribine-TP, clofarabine-TP, fludarabine-TP, vidarabine-TP, cytarabine-TP, and gemcitabine-TP were crystallized in the catalytic pocket of SAMHD1. We find that all of these drugs are substrates of SAMHD1 and that the efficacy of most of these drugs is affected by SAMHD1 activity. Of the nucleotide analogs tested, only cladribine-TP with a deoxyribose sugar efficiently induced the catalytically active SAMHD1 tetramer. Together, these results establish a detailed framework for understanding the substrate specificity and allosteric activation of SAMHD1 with regards to nucleotide analogs, which can be used to improve current cancer and antiviral therapies.

2.3 Introduction

SAMHD1 is a triphosphohydrolase that severs the triphosphate group from deoxynucleoside triphosphates (dNTPs) (Franzolin, Pontarin et al., 2013, Goldstone, Ennis-Adeniran et al., 2011). A major function of SAMHD1 is to reduce the dNTP pool in non-cycling cells, making it an important regulator of dNTP levels in the cell (Franzolin et al., 2013). In addition to its role in regulating genome stability, SAMHD1 is best known for its ability to block infection of a broad range of retroviruses including human immunodeficiency virus type 1 (HIV-1). During viral infection, SAMHD1 depletes the cellular dNTPs needed for reverse transcription of the viral RNA genome

(Berger, Sommer et al., 2011, Bermejo, Lopez-Huertas et al., 2016, Descours, Cribier et al., 2012, Goldstone et al., 2011, Hrecka, Hao et al., 2011, Laguette, Sobhian et al., 2011, Lahouassa, Daddacha et al., 2012). In the cell, SAMHD1 activity is modulated by allosteric activation and phosphorylation (Amie, Bambara et al., 2013, Cribier, Descours et al., 2013, Hansen, Seamon et al., 2014, Ji, Tang et al., 2014, Ji, Wu et al., 2013, Pauls, Badia et al., 2014, Pauls, Ruiz et al., 2014, St Gelais, de Silva et al., 2014, Tang, Ji et al., 2015, Welbourn, Dutta et al., 2013, Yan, Kaur et al., 2013). Nucleotide binding to two allosteric sites of each subunit leads to the assembly of activated SAMHD1 tetramer. Allosteric site (Allo-site) 1 only accommodates guanosine bases (GTP or dGTP), but any canonical dNTP can bind Allo-site 2 (Ji et al., 2013, Koharudin, Wu et al., 2014, Yan et al., 2013, Zhu, Gao et al., 2013, Zhu, Wei et al., 2015). When dNTPs are needed for DNA synthesis, phosphorylation at residue T592 destabilizes the active tetramer of SAMHD1 thereby downregulating SAMHD1 activity (Cribier et al., 2013, Tang et al., 2015, Welbourn et al., 2013).

SAMHD1 is an important general sensor and regulator of the dNTP pools and thus it is crucial to genome maintenance. All canonical dNTPs are both substrates and allosteric activators, and this promiscuity allows SAMHD1 to target therapeutic molecules that resemble nucleotides in structure. Nucleoside analogs are a large class of drugs which are used to treat viral infections and many types of cancers (Balzarini, 2000, Bonate, Arthaud et al., 2006, Ewald, Sampath et al., 2008, Fridle, Medinger et al., 2017, Galmarini, Mackey et al., 2001, Jordheim, Durantel et al., 2013, Liliemark, 1997, Tamamyan, Kadia et al., 2017). These compounds interfere with viral replication or cancer cell proliferation upon incorporation into newly synthesized DNA, resulting in

chain terminations, accumulation of mutations, and often cell apoptosis (Balzarini, 2000, Ewald et al., 2008, Jordheim et al., 2013). Several recent reports have demonstrated that SAMHD1 reduces the efficacy of nucleotide analog drugs by depleting their cellular concentrations (Hollenbaugh, Shelton et al., 2017, Schneider, Oellerich et al., 2017). Strikingly, SAMHD1 expression levels were shown to be highly predictive of patient response to cytarabine, the primary treatment for acute myeloid leukemia (AML) (Schneider et al., 2017). Although the kinetic parameters of SAMHD1 activity has been thoroughly studied for canonical dNTPs (Arnold, Groom et al., 2015, Miazzi, Ferraro et al., 2014, Wang, Bhattacharya et al., 2016), better characterization of SAMHD1 substrate specificity is needed for the optimal administration of current nucleoside analog drugs.

Interestingly, SAMHD1 activity has been shown to increase the efficacy of some nucleotide analogs that are not substrates of SAMHD1 (Amie, Daly et al., 2013, Ordonez, Kunzelmann et al., 2017). In these cases, SAMHD1 activity allows the non-substrate nucleotide analogs to better compete for target active sites by depleting cellular dNTPs. Therefore, depending on how SAMHD1 interacts with a particular nucleotide analog of interest, it might be desirable to either inhibit or increase SAMHD1 activity. To selectively modulate SAMHD1 activity, it is important to fully understand how nucleotide analog drugs either bind to the allosteric site to assemble an active tetramer or bind the catalytic site to be hydrolyzed.

In this study, we characterized SAMHD1 interactions with a panel of nucleotide analogs that are used to treat a variety of cancers and viral infections (Bonate et al., 2006, Ewald et al., 2008, Fridle et al., 2017, Liliemark, 1997, Whitley, Alford et al., 1980, Whitley, Tucker et al., 1980) (Fig. 2-1A). We found that the catalytic site of SAMHD1 is

very promiscuous, allowing SAMHD1 to hydrolyze most of the analogs tested here. On the other hand, Allo-site 2 is more restrictive to modifications of the 2' sugar moiety of the drug. These results are important for the assessment of SAMHD1 as a potential target for cancer therapy, the design of non-hydrolysable derivatives, and the development of modulators of SAMHD1 activity to combine with existing therapies. In addition, this work contributes to a greater understanding of the structural and biochemical principles of SAMHD1 substrate selectivity and allosteric activation.

Nucleoside analog drugs are widely used to treat a variety of cancers and viral infections. With an essential role in regulating the nucleotide pool in the cell by degrading cellular nucleotides, SAMHD1 has the potential to decrease the cellular concentration of frequently prescribed nucleotide analogs and thereby decrease their clinical efficacy in cancer therapy. To improve future nucleotide analog treatments, it is important to understand SAMHD1 interactions with these drugs. Our work thoroughly examines the extent to which nucleotide analogs interact with the catalytic and allosteric sites of SAMHD1. This work contributes to the assessment of SAMHD1 as a potential therapeutic target for cancer therapy and the future design of SAMHD1 modulators that might improve the efficacy of existing therapies.

2.4 Results

2.4.1 Crystal Structures of nucleotide analogs bound to SAMHD1 catalytic pocket

To better understand the structural basis for nucleotide analog binding in the catalytic pocket of SAMHD1, we co-crystallized the inactive catalytic domain of SAMHD1 (residues 113-626 with H206R/D207N mutations to inhibit catalysis) with six

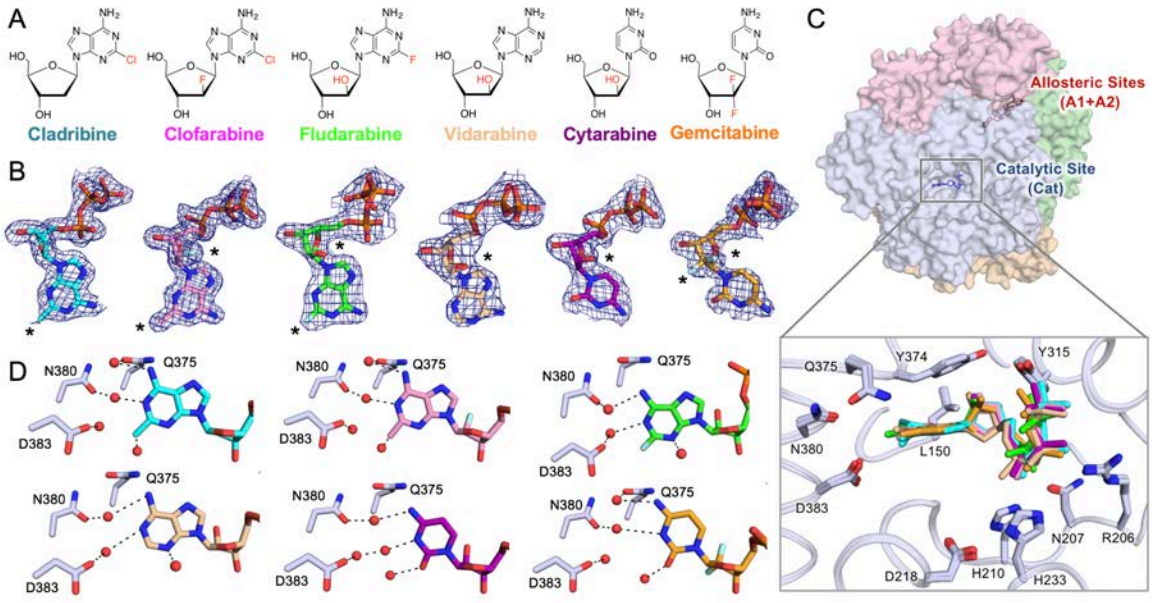

Figure 2-1. Substrate specificity of SAMHD1 is determined by 2' sugar moiety.

(A) Chemical structures of nucleoside analogs used in this study. (B) $2F_o$-F_c electron density ($\sigma = 1.0$) for the nucleotide analog drugs crystallized in the catalytic pocket of SAMHD1. Black asterisks indicate sites of modifications. (C) Top: Transparent surface view of SAMHD1 tetramer with each subunit in a different color. Selected allosteric nucleotides are shown in red sticks and a nucleotide in a catalytic pocket is shown in blue sticks. Bottom: Superposition of all the nucleotide analogs bound to the SAMHD1 catalytic pocket. SAMHD1 backbone is shown as coils with side chains shown as sticks. Cladribine-TP (cyan), clofarabine-TP (magenta), fludarabine-TP (green), vidarabine-TP (wheat), cytarabine-TP (purple), and gemcitabine-TP (orange) are shown as sticks. (D) Water networks (shown as red spheres) observed for each nucleotide analog bound to the SAMHD1 catalytic site. Black dashed lines indicate hydrogen bonds.

selected cancer and antiviral drugs. This SAMHD1 mutant has been previously shown to be identical in conformation and nucleotide-binding properties to the wildtype (WT) enzyme, but it is more amenable to crystallization (Ji et al., 2014, Ji et al., 2013, Koharudin et al., 2014, Zhu et al., 2015). The crystal structures of these SAMHD1-nucleotide analog complexes were determined at resolutions ranging from 1.7 to 2.5 Å

Table 2-1. Data collection and refinement statistics for the seven crystal structures of SAMHD1 HD bound to nucleotide analogs.

Table 1. Data collection and refinement statistics for the seven crystal structures of SAMHD1 HD bound to nucleotide analogues.

	Cladribine-TP	**Clofarabine-TP**	**Fludarabine-TP**	**Vidarabine-TP**	**Cytarabine-TP**	**Gemcitabine-TP**	**((2'R)-2'-F)-dCTP**
Accession Code	6DW4	6DWD	6DWK	6DWJ	6DW3	6DW5	6DW7
Data Collection							
Wavelength (Å)	0.97920	0.97920	0.97920	0.92014	0.97920	0.97910	0.97912
Space Group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions							
a, b, c (Å)	80.3, 142.3, 98.4	80.5, 142.2, 98.8	87.7, 146.7, 99.1	86.9, 146.7, 99.6	87.3, 147.3, 98.5	80.7, 142.9, 98.9	84.5, 146.4, 98.9
α, β, γ (°)	90, 114.1, 90	90, 114.1, 90	90, 114.5, 90	90, 114.5, 90	90, 114.6, 90	90, 113.9, 90	90, 113.7, 90
Molecules/asymmetric unit	4	4	4	4	4	4	4
Resolution (Å)	50-2.00 (2.03-2.00)	50-1.70 (1.73-1.70)	50-2.30 (2.34-2.30)	50-2.50 (2.54-2.50)	50-2.20 (2.24-2.20)	50-1.93 (1.96-1.93)	50-2.5 (2.54-2.50)
Unique reflections	129994 (5987)	209481 (10532)	100587 (5041)	75598 (3796)	110334 (5500)	148608 (7362)	76374 (3868)
R_{merge}	0.080 (0.682)	0.086 (>1)	0.158 (>1)	0.107 (0.737)	0.123 (>1)	0.093 (0.775)	0.126 (0.892)
Mean I / σ I	12.6 (1.6)	15.0 (1.2)	9.3 (1.4)	10.0 (2.2)	7.0 (0.8)	13.2 (1.8)	6.3 (1.4)
Completeness (%)	95.1 (87.7)	94.1 (94.7)	99.5 (99.8)	97.4 (97.1)	98.1 (98.2)	97.0 (96.4)	97.5 (98.7)
Redundancy	3.0 (2.7)	4.2 (4.0)	5.7 (5.6)	3.5 (3.3)	2.9 (2.8)	3.5 (3.3)	2.9 (3.0)
CC½	0.999 (0.518)	0.999 (0.283)	0.983 (0.185)	0.989 (0.688)	0.997 (0.252)	0.998 (0.521)	0.985 (0.427)
Refinement							
Non-hydrogen atoms	17077	17162	16439	16328	16399	16603	16049
R_{work}/R_{free} (%)	16.9/20.2 (26.6/26.5)	17.6/20.2 (30.8/31.4)	17.6/21.2 (26.2/26.0)	17.5/20.8 (26.6/32.7)	19.0/22.8 (37.7/38.1)	17.3/20.5 (26.6/28.2)	21.2/25.0 (39.0/38.6)
Average B factor	32	29	46	74	43	37	67
RMSD							
Bond lengths	0.020	0.019	0.020	0.019	0.019	0.023	0.012
Bond angles	2.0	2.0	2.1	2.0	2.0	2.1	1.6
Ramachandran Analysis							
Preferred regions (%)	98.3	98.3	97.8	98.6	97.5	98.3	98.6
Allowed regions (%)	1.7	1.7	2.1	1.4	2.3	1.7	1.4
Outliers (%)	0	0	0.1	0	0.2	0	0

Statistics in parentheses indicate those for the highest resolution shell.

(Table 2-1), with electron density that allows for unambiguous identification of each substrate in the catalytic pocket (Fig. 2-1B). The structures of SAMHD1 bound to these nucleotide analogs are similar to those obtained with canonical nucleotides (Ji et al., 2014, Koharudin et al., 2014). Nucleotide analog binding does not alter the architecture of the SAMHD1 catalytic pocket (Fig. 2-1C), and these nucleotide analogs adopt similar conformations to the canonical nucleotides in the pocket. This suggests that the nucleotide modifications tested here do not disturb the overall integrity of the catalytic pocket or induce large structural rearrangements. The promiscuous catalytic pocket of SAMHD1 likely accommodates other nucleotide analogs with similar modifications.

2.4.2 Base modification has modest effects on substrate binding

Modifications on the Watson-Crick base edge are well tolerated at the catalytic site of SAMHD1. Consistent with previous observations with dNTPs (Ji et al., 2014, Koharudin et al., 2014), there are no base-specific interactions between SAMHD1 and the bases of the nucleotide analogs. Water networks in the active site of SAMHD1 stabilize the nucleotide analogs in the pocket, similarly to canonical dNTPs. In each case, three to four water molecules bridge interactions from SAMHD1 to the Watson-Crick and sugar edges of the bound analog (Fig. 2-1D). Cytarabine-TP and gemcitabine-TP, each containing an unmodified cytosine base, maintain water network interactions that closely resemble those reported for dCTP (Fig. 2-1D) (Ji et al., 2014). In our previous structural studies, the adenosine base was not resolved in the dATP-SAMHD1 co-crystal structure due to its relatively weak binding affinity to the catalytic site compared to other dNTPs (Ji et al., 2014). However, clear electron density for the bases of clofarabine-TP,

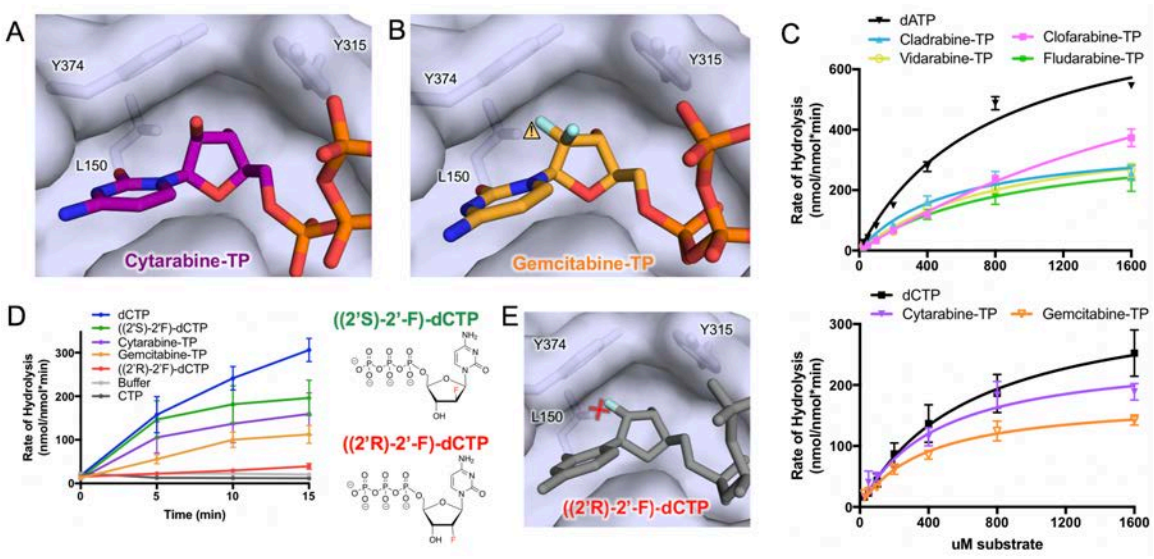

Figure 2-2. Gemcitabine-TP but not ((2'R)-2'-F)-dCTP is hydrolyzed by SAMHD1 *in vitro*.

(A) ((2'S)-2'-OH)) of cytarabine-TP is stabilized by residues Y374 and Y315 through van der Waals interactions. Transparent surface of SAMHD1 is shown with key residues in sticks. (B) 2',2'-difluorine sugar modification of gemcitabine-TP is stabilized by Van der Waals interactions with residues Y374 and Y315 to compensate potential close contact (yellow caution triangle) between ((2'R)-2'-F) atom and residue L150 in the catalytic site. (C) HPLC-based activity assay measuring product produced by pre-assembled SAMHD1 tetramers in the presence of 25-1600 μM nucleotide analog substrates. Error bars represent standard error of mean (SEM) of three independent experiments. (D) Left: dNTPase activity of SAMHD1 over the course of 30 minutes was measured using a malachite green assay. Product is normalized to SAMHD1 concentration (nmol PO_4 / nmol SAMHD1). SAMHD1 tetramers were pre-assembled with 250 μM GTP and dATP and then diluted 100-fold into 125 μM gemcitabine-TP, cytarabine-TP, dCTP, CTP, ((2'S)-2'-F)-dCTP, ((2'R)-2'-F)-dCTP, or buffer. Error bars represent standard error of the mean (SEM) of three independent experiments. Right: chemical structures of ((2'S)-2'-F)-dCTP and ((2'R)-2'-F)-dCTP analogs. (E) A rigid body model of ((2'R)-2'-F)-dCTP (gray sticks) in the catalytic pocket potentially creates a clash (red cross) with residue L150.

cladribine-TP, fludarabine-TP, and vidarabine-TP were observed in the current structures (Fig. 2-1B). The additional chlorine atom (cladribine-TP and clofarabine-TP) or fluorine atom (fludarabine-TP) at the C2 position of the bases contribute to a more extensive water network and stabilize these molecules in the active site (Fig. 2-1D). These results

suggest that small modifications in the base are not likely to disrupt binding to the SAMHD1 catalytic pocket because water networks that contact the base are flexible. Some nucleotide modifications may even enhance the water network surrounding the nucleotide analog, possibly increasing its binding affinity.

2.4.3 2’ sugar moiety regulates SAMHD1 substrate specificity

It is well established that NTPs with ribose sugars ((2’R)-2’-OH) are not substrates of SAMHD1 (Goldstone et al., 2011, Ji et al., 2014). We examined how sugar modifications found in nucleotide analog drugs influence their binding to and hydrolysis by SAMHD1, as it has been shown that arabinose-based nucleotides with 2’S sugar modifications are substrates of SAMHD1 (Hollenbaugh et al., 2017, Schneider et al., 2017). Our crystal structures corroborate the finding that arabinose-based sugars, with either ((2’S)-2’-OH) or ((2’S)-2’-F) modifications, are allowed in the catalytic site of SAMHD1. In addition to the canonical interactions observed with dNTP substrates (Ji et al., 2014), we observe van der Waals and stacking interactions between active site residues Y315 and Y374 and the ((2’S)-2’-OH/F) atoms of cytarabine-TP, clofarabine-TP, fludarabine-TP, vidarabine-TP and gemcitabine-TP in the catalytic pocket (Fig. 2-1C, 2-2A and 2-2B). Since the arabinose-like sugar modifications provide an additional interaction with SAMHD1, these modifications likely stabilize these analogs in the catalytic pocket. In contrast, we crystallized SAMHD1 in the presence of 10 mM ((2’R)-2’-F)-dCTP, which has a ribose-like modification, but electron density for this nucleotide analog was not observed in the catalytic pocket.

Table 2-2. Kinetic constants for nucleotide analog hydrolysis by SAMHD1.

	k_{cat}* (s^{-1})			K_M* (uM)		
dATP	11.4	±	0.8	781	±	118
Cladribine-TP	4.9	±	0.4	537	±	118
Clofarabine-TP	14.2	±	2.7	2912	±	770
Fludarabine-TP	5	±	0.9	867	±	322
Vidarabine-TP	5.6	±	0.6	873	±	180
dCTP	4.7	±	0.7	648	±	230
Cytarabine-TP	3.4	±	0.4	446	±	123
Gemcitabine-TP	2.4	±	0.2	379	±	79

*Mean values from three independent experiments ± S.E.M.

Our activity assays also confirmed that the substrate specificity of SAMHD1 has a general dependency on the stereochemistry of the sugar moiety at the 2' position. We used an HPLC-based assay (Ji et al., 2014) to measure the SAMHD1 enzyme kinetics for these substrates. Larger modifications at the 2'sugar moiety were generally associated with higher K_M and lower k_{cat} values (Fig. 2-2C and Table 2-2). Although most 2' sugar modifications are tolerated and the analogs can bind SAMHD1 in a similar fashion as evidenced by their structures, modest geometry perturbations due to the 2' sugar modifications may lead to less catalytically productive binding. Although gemcitabine-TP is a substrate of SAMHD1, its turnover rate and specific constant are only slightly lower than that of cytarabine-TP (Fig. 2-2C and Table 2-2). A previous report did not observe hydrolysis of gemcitabine-TP, where the concentration of gemcitabine-TP in monocyte-derived macrophages was not dependent on SAMHD1 protein levels (Hollenbaugh et al., 2017). This is perhaps due to the difficulty in detecting the hydrolysis of gemcitabine in the complex cellular environment.

To reconcile the discrepancy regarding gemcitabine-TP hydrolysis, we directly compared the hydrolysis of the doubly modified gemcitabine-TP to the singly modified ((2'S)-2'-F)-dCTP and ((2'R)-2'-F)-dCTP *in vitro* in a time course assay (Fig. 2-2D). We observed that SAMHD1 hydrolyzes ((2'S)-2'-F)-dCTP similar to cytarabine-TP, but gemcitabine-TP is hydrolyzed at a lower rate and ((2'R)-2'-F)-dCTP is unreactive (Fig. 2-2D). In general, these results confirmed other reports that arabinose-based nucleotide analogs are substrates of SAMHD1, whereas the ribose-based ((2'R)-2'-F)-dCTP is not (Hollenbaugh et al., 2017, Schneider et al., 2017). As evidenced by the electron density of gemcitabine-TP in the catalytic pocket, both the 2'S and 2'R fluorine modifications are accommodated in the catalytic pocket of SAMHD1. However, the substructure ((2'R)-2'-F)-dCTP, with a single 2'R modification, is not accommodated. Rigid modeling of this compound into the catalytic pocket predicts that the (2'R)-2'-F atom would clash with L150 (Fig. 2-2D). This suggests that the 2'R modification is permitted only in the context of a 2',2'-difluorine sugar modification, but not alone (Fig. 2-2B and 2-2E). SAMHD1's discrimination between these two substrates may arise from interactions between Y315/Y374 and the 2'S fluorine atom, which might partially compensate for the sterically unfavorable 2'R fluorine modification (Fig. 2-2B and 2-2E). The presence of the 2R' modification may lead to a less suitable positioning of the nucleotide in the catalytic pocket for catalysis as compared to 2'S modifications alone (Fig. 2-2A and Fig. 2-2B). Although we captured gemcitabine-TP in the catalytic pocket at a high concentration (5 mM), the catalytically productive conformations may not be readily formed at lower concentrations used in cellular studies. This is consistent with the low k_{cat} value of gemcitabine-TP in our measurements (Fig. 2-2C). These results support the

notion that the catalytic pocket of SAMHD1 is highly sensitive to the stereochemistry of 2' sugar modification.

2.4.4 SAMHD1 hydrolyzes various nucleotide analogs *in vivo*

Since SAMHD1 is capable of hydrolyzing cancer drugs *in vitro*, it has the potential for decreasing their therapeutic efficacies. We explored the extent to which the triphosphates of each of these nucleotide analogs are degraded by SAMHD1 in the cell following uptake and metabolic activation. A recent report demonstrated a strong inverse correlation between SAMHD1 expression in leukemic blasts and AML patients' clinical response to cytarabine therapy (Schneider et al., 2017). Similarly, our colleagues in the Cinatl Group at Goethe University of Frankfurt and the Keppler group at Ludwig Maximilians University found that the IC_{50} values of fludarabine-TP and clofarabine-TP in AML cell lines were also correlated with SAMHD1 protein expression levels (Fig. 2-3A). However, this effect was not observed for cladribine-TP or gemcitabine-TP. We also tested whether the depletion of SAMHD1 in THP-1 cells via knock-out or targeted proteasomal degradation by Vpx-VLPs affected the IC_{50} values of these drugs. Although a strong effect was observed for cytarabine-TP and a moderate effect was observed for fludarabine-TP, clofarabine-TP, vidarabine-TP and cladribine-TP (Fig. 2-3B), there was no effect on the IC_{50} value for gemcitabine-TP (Fig. 2-3B). To directly measure SAMHD1's effect on each nucleotide analog's concentration in the cell, we used LC-MS/MS to quantify drugs found in cells with or without SAMHD1 expression. Their results show that clofarabine-TP, fludarabine-TP, and cytarabine-TP were strongly depleted by SAMHD1, whereas cladribine-TP was depleted to a low extent and gemcitabine-TP levels remained unaffected (Fig. 2-3C). Although SAMHD1 hydrolyzes

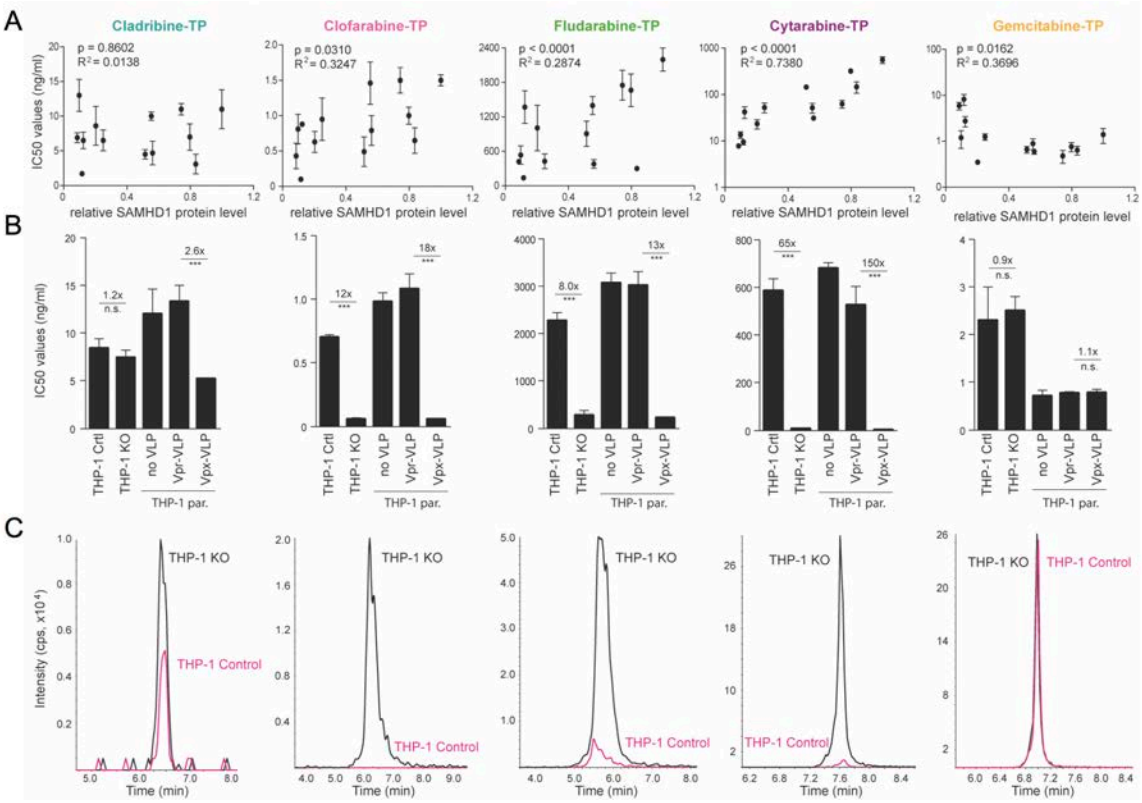

Figure 2-3. SAMHD1 depletes several TPs of nucleoside analogs *in vivo*.

(A) Correlations of cytarabine, clofarabine, fludarabine, cladribine, and gemcitabine concentrations inhibiting 50% of cell viability (IC_{50}) and relative protein expression levels of SAMHD1 in 13 AML cell lines. Relative expression levels (ratios of a SAMHD1/β-actin) are shown as arbitrary units (a.u.). Ratio of SAMHD1/β-actin for parental THP-1 cells is set to 1.0 and ratios of other cell lines are set relative to it. Closed circles represent mean values and error bars indicate SD of three independent experiments. Correlations were analyzed using linear or log-linear regression models. (B) Cytarabine, fludarabine, clofarabine, cladribine or gemcitabine IC_{50} values of THP-1 KO, THP-1 Control cells, or parental THP-1 cells exposed to VSV-G pseudotyped VLPs carrying either lentiviral Vpr (Vpr-VLP, control) or SAMHD1-degrading Vpx proteins (Vpx-VLP). The bars represent mean values and the error bars arethe SD of three independent experiments. The numbers above indicate factor of decrease of the IC_{50} values in the absence of SAMHD1. Statistical analyses were performed using unpaired two-tailed Students' t-test comparing treated samples with untreated control. ***p < 0.001 (C) Representative liquid chromatography tandem mass spectrometry measurements (LC-MS/MS) of cytarabine-TP, fludarabine-TP, clofarabine-TP, cladribine-TP or gemcitabine-TP in THP-1 KO cells (black) and THP-1 control cells (red).

gemcitabine *in vitro* with a low activity, this effect may not be detectable under normal cellular conditions. These cell-based assays indeed corroborate previous reports that gemcitabine-TP is not significantly degraded by SAMHD1 *in vivo* (Hollenbaugh et al., 2017, Schneider et al., 2017). The structural and biochemical framework established in this study also allows for the rational modeling of other known cancer drugs into the catalytic pocket of SAMHD1. For example, the IC_{50} of nelarabine-TP in AML cell lines relies on the expression level of SAMHD1, indicating that it is a substrate of SAMHD1 too. Modeling the metabolite araGTP into the catalytic pocket of SAMHD1 predicts that it would fit into the catalytic pocket like any other arabinose-based nucleotide analog, such as vidarabine-TP.

2.4.5 Allosteric site 2 of SAMHD1 is more restrictive than its catalytic site

In addition to establishing SAMHD1's substrate specificity for nucleotide analogs, we also tested whether these drugs were capable of binding to the allosteric sites to induce the catalytically active SAMHD1 tetramer. None of the nucleotide analogs tested here contained the guanosine base required for Allo-site 1 binding, thus they alone were not sufficient for SAMHD1 activation. To test which analogs bind Allo-site 2, we monitored the oligomerization state of SAMHD1 in the presence of GTP and each of the analogs. Previous studies indicated that clofarabine-TP is an activator of SAMHD1 (Arnold, Kunzelmann et al., 2015), but cytarabine-TP is not (Hollenbaugh et al., 2017, Schneider et al., 2017). Consistent with these reports, our size exclusion chromatography (SEC) assays showed that only clofarabine-TP and cladribine-TP caused a shift in the elution profile of SAMHD1 towards a higher molecular weight species (Fig. 2-4A), with

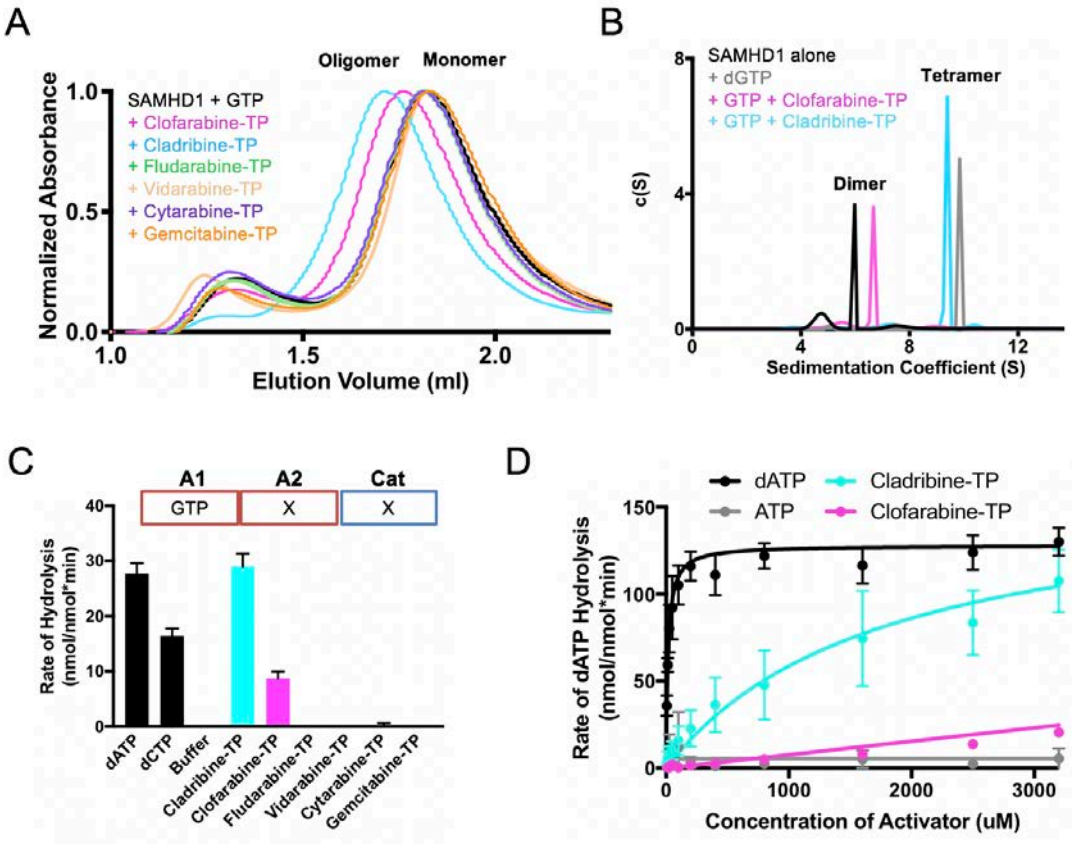

Figure 2-4. The allosteric sites of SAMHD1 are highly restrictive.

(A) Size exclusion chromatography elution profile of SAMHD1 in the presence of 0.5 mM GTP and 4 mM of color-coded nucleotide analog. (B) SV-AUC analysis of SAMHD1 in the absence of nucleotides or the presence of dGTP, GTP with clofarabine-TP, or GTP with cladribine-TP at a final concentration of 150 μM. (C) Malachite green activity assay performed in the presence of 125 μM GTP and 125 μM dCTP, dATP, nucleotide analog, or buffer. Error bars represent SEM of three independent experiments. (D) Malachite green activity assay measuring the hydrolysis of dATP by SAMHD1 tetramers pre-assembled in the presence of 125 μM GTP and 6.3-3200 μM dATP, cladribine-TP, clofarabine-TP, or ATP. Error bars represent SEM of three independent experiments.

cladribine-TP being more effective at inducing SAMHD1 oligomerization.

Sedimentation velocity analytical ultracentrifugation (SV-AUC) also showed that cladribine-TP induced SAMHD1 oligomerization (Fig. 2-4B). However, clofarabine-TP-induced SAMHD1 oligomerization was not detected by SV-AUC, likely due to

experimental constraints requiring about 40-fold less nucleotide analogs compared to SEC. When mixed with GTP, cladribine-TP and to some extent clofarabine-TP both induce SAMHD1 activity (Fig. 2-4C). To compare relative affinities of each drug for Allo-site 2, we measured the activity of SAMHD1 pre-assembled with GTP and increasing concentrations of dATP, cladribine-TP, clofarabine-TP, and ATP activators (Hansen et al., 2014, Ji et al., 2014). Hydrolysis of the dATP substrate was measured as readout for the effective affinity of the activators in the Allo-site 2. While cladribine-TP assembled active SAMHD1 tetramers with a modest efficiency, the apparent affinity of clofarabine-TP for Allo-site 2 is drastically reduced (Fig. 2-4D). Together, these results suggest that only minor sugar modifications are permitted for SAMHD1 allosteric activators.

2.4.6 2' sugar modifications are highly restrictive at the allosteric site 2

To examine how SAMHD1 accommodates nucleotide analogs in the allosteric pocket, we attempted to crystalize SAMHD1 with GTP and each of the nucleotide analogs assayed above. Consistent with our activity assay and oligomerization measurements, only the two allosteric activators, cladribine-TP and clofarabine-TP, resulted in SAMHD1 tetramer crystals. The resulting structures revealed unambiguous electron density for each nucleotide in the Allo-site 2 (Fig. 2-5A). As predicted, the nucleotide analogs bind to Allo-site 2 and GTP binds to the adjacent Allo-site 1. The allosteric sites are not disturbed by cladribine-TP or clofarabine-TP, as the chlorine atom modification at the C2 position of the base does not interfere with the allosteric pocket interactions (Fig. 2-5B). Hydrogen bonds between the adenosine base and residues N119 and N358 (Ji et al., 2014) are preserved in the cladribine-TP and clofarabine-TP

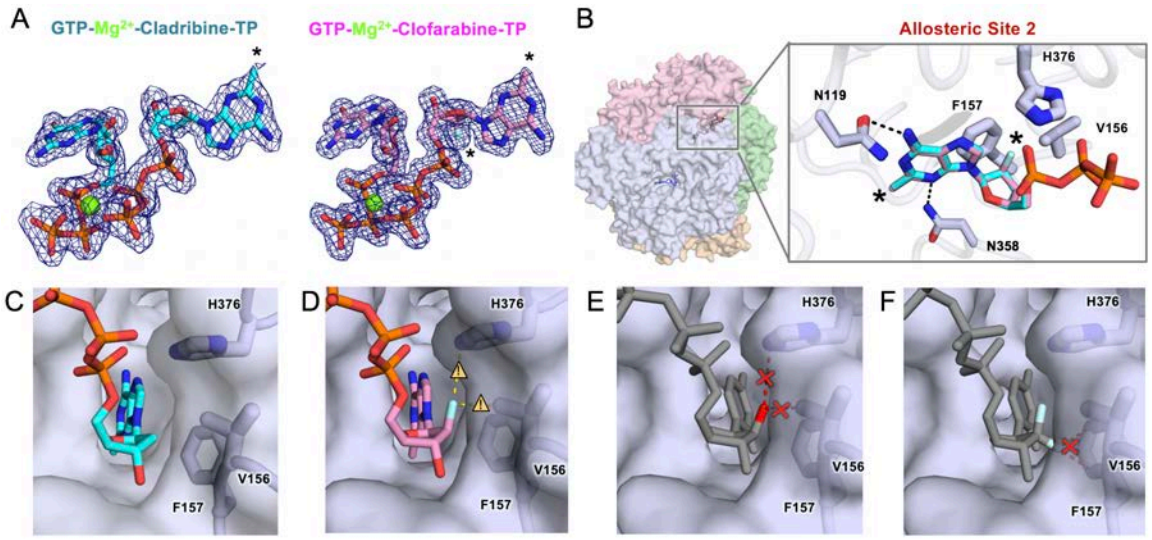

Figure 2-5. Structures of cladribine-TP and clofarabine-TP bound to Allo-site 2 of SAMHD1.

(A) $2F_o$-F_c electron density ($\sigma = 1.0$) for GTP, cladribine-TP, and clofarabine-TP in the allosteric pocket of SAMHD1. Black asterisks indicate sites of modifications, and black dotted lines indicate hydrogen bonds. (B) *Left*: Transparent surface view of the SAMHD1 tetramer. *Right*: Overlay of cladribine-TP (cyan) and clofarabine-TP (pink) in Allo-site 2. Main chain of SAMHD1 is shown as tubes with selected residues and nucleotides represented as sticks. Residues important for gating the 2'-atom are highlighted in thicker sticks. Black asterisks indicate sites of modification. (C) The structure of cladribine-TP (cyan, sticks) in Allo-site 2 with V156, F157, and H376 shown as sticks under the semitransparent surface of SAMHD1. (D) The structure of clofarabine-TP (pink, sticks) in Allo-site 2 with close contacts between the ((2'S)-2'-F) atom and V156, F157, and H376 highlighted with caution triangles and yellow dashed lines. (E) A rigid body model of cytarabine-TP in Allo-site 2 with potential steric clashes between the ((2'S)-2'-OH) group and V156 and H376 highlighted with a red cross and red dashed lines. (F) A rigid body model of gemcitabine-TP in Allo-site 2 with a potential steric clash between the ((2'R)-2'-F) atom and F157 highlighted with a red cross and red dashed lines.

structures, allowing for the correct positioning of each nucleotide in the Allo-site 2 pocket. These results suggest that nucleotide analogs with some similar modifications at the Watson-Crick edge of the base may also be permitted in Allo-site 2.

While modest base modifications do not affect Allo-site 2 binding, modifications to the sugar at the 2' position are restricted. We found that Allo-site 2 excludes all

arabinose-based analogs with the ((2'S)-2'-OH) group, which is consistent with previous observations (Schneider et al., 2017). The ribose-based ((2'R)-2'-F)-dCTP is also excluded from the site, similarly as the NTP molecules reported before (Ji et al., 2014). As expected, Allo-site 2 does not tolerate a nucleotide with 2',2'-difluorine modifications. The exclusion of nucleotides with these sugar modifications is likely due to SAMHD1 residues H376, V156, and F157 which form a tight pocket around the 2' carbon of the dNTPs (Fig. 2-5B and 2-5C). While clofarabine-TP, which contains a ((2'S)-2'-F) modification in this position, is accommodated at this site, the structure shows that the fluorine atom may still cause some steric constraints (Fig. 2-5D). This is consistent with our oligomerization and activity assays, which showed reduced SAMHD1 activation by clofarabine-TP (Fig. 2-4C). Modeling cytarabine-TP in the Allo-site 2 pocket suggests steric clashes arise between the ((2'S)-2'-OH) group and residues H376 and V156 of SAMHD1 (Fig. 2-5E). On the other side of the sugar ring, residue F157 limits the accessibility of nucleotides with ((2'R)-2'-OH) or ((2'R)-2'-F) modifications, such as gemcitabine-TP (Fig. 2-5F). Cladribine-TP is the only deoxyribonucleotide analog tested here, and most likely the lack of a 2' sugar modification allows it to be the strongest allosteric activator of SAMHD1. Overall, Allo-site 2 in SAMHD1 was highly sensitive to modifications of the 2' position of the sugar, and this moiety is a major binding determinant of nucleotide analog binding to the allosteric site.

2.5 Discussion

The study presented herein provides a comprehensive structural and biochemical framework for understanding how a wide range of nucleotide analog drugs interact with SAMHD1. Our biochemical and structural analyses of a panel of nucleotide analogs, with

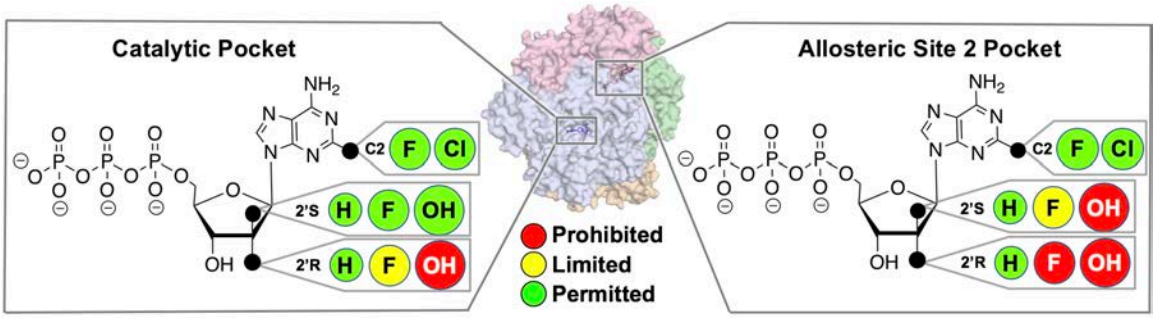

Figure 2-6. Summary of the effects of the 2' sugar moiety on nucleotide analog binding to the catalytic and allosteric sites of SAMHD1.

Middle: Transparent surface view of SAMHD1 tetramer. *Left*: In the catalytic pocket of SAMHD1, while small substitutions such as fluorine atoms at the 2'R and 2'S positions of the sugar are permitted (green circle), their access to the 2'R position is limited (yellow circle). Larger modifications, such as hydroxyl groups, are permitted in the 2'S position, but not the 2'R position (red circle). *Right*: In Allo-site 2 of SAMHD1, hydroxyl groups are prohibited in both 2'R and 2'S positions of the sugar moiety. Fluorine atoms have limited access to the 2'S position, but they are prohibited from the 2'R position. Small base modifications such as fluorine or chlorine atoms are tolerated in both Allo-site 2 and the catalytic site.

a variety of 2' sugar modifications, reveal the detailed binding determinants for the catalytic site of SAMHD1 (Fig. 2-1 and Fig. 2-6). SAMHD1 selects substrates through indirect interactions between water molecules in the catalytic pocket and the base of the analog drugs. The network of water molecules changes to adapt to variations found in the nucleotide analogs. Water-mediated interactions between the substrate and the enzyme allow for binding pocket plasticity and accommodation of different modifications. In addition, the catalytic pocket is accessible to the arabinose-like 2' sugar modifications with the 2'S geometry. This provides a mechanistic understanding of the findings that arabinose-based nucleotides are substrates of SAMHD1 (Hollenbaugh et al., 2017, Schneider et al., 2017). Interestingly, we were able to observe electron density for analogs with modified adenosine bases, which was previously unresolved in our dATP-

SAMHD1 structure (Ji et al., 2014). This indicates that the nucleotide analog modifications may provide additional stabilizing interactions with the catalytic pocket of SAMHD1.

Our results also demonstrate that a nucleotide analog can bind to the allosteric sites of SAMHD1 to activate the enzyme, depending on modifications to its sugar moiety (Fig. 2-5 and Fig. 2-6). We show that modest base modifications, such as a chlorine atom at C2 of adenine, are tolerated at the Allo-site 2. In contrast, Allo-site 2 does not allow arabinose-based or ribose-based nucleotides. Only deoxyribose-based nucleotides, such as cladribine-TP, can efficiently enter Allo-site 2 (Fig. 2-5 and 2-6). The results suggest that any modifications to the 2' carbon of the sugar ring may affect the drug's binding affinity to the allosteric pocket. These findings may help advance the understanding of the effect of current nucleotide analog drugs and guide the design of non-hydrolysable analogs that can activate SAMHD1 for targeted depletion of cellular dNTP pools.

Interestingly, we found that cytarabine-TP is particularly dependent on SAMHD1 cellular expression compared to other drugs tested in this study. Although many of the drugs tested were hydrolyzed by purified SAMHD1 in our biochemical assays, the effect of SAMHD1 expression on drug efficacy varied in cells. This may be due to different interactions between these drugs with other components of the nucleotide metabolism pathway in the cell. For example, even though clofarabine, fludarabine and cladribine are similar in structure, these drugs have been reported to have varying degrees of interactions with cellular factors that affect their ability to be phosphorylated by deoxycytidine kinase and to inhibit ribonucleotide reductase activity and DNA/RNA chain elongation (Bonate et al., 2006). Moreover, nucleoside analogs may differentially

influence activity and expression of cell cycle regulatory proteins such as cyclins and cyclin dependent kinases (Bastin-Coyette, Cardoen et al., 2011, Radosevic, Delmer et al., 2001) which in turn may interact with SAMHD1 and influence its dNTPase activity in leukemic cells (St Gelais, Kim et al., 2018). It remains possible that, although these drugs are hydrolyzed to a similar extent by SAMHD1 *in vitro*, their metabolism in the cell might be influenced by other cellular factors. In the case of cytarabine, our *in vivo* data strongly support that SAMHD1 is one of the main determinants of intracellular cytarabine-TP concentrations. On the other hand, we demonstrated that SAMHD1 strongly hydrolyzed cladribine-TP *in vitro* but had little influence on cladribine-TP concentrations and activity in AML cells. Combining biochemical, biophysical, and cellular studies offers a comprehensive approach to evaluate the response of nucleoside analog drugs to SAMHD1 both at the mechanistic level and in application.

SAMHD1's substrate promiscuity helps it function as a general sensor and regulator of nucleotide pools, but it also allows SAMHD1 to facilitate cancer cells to escape from nucleotide analog treatments. Thus, it is important to consider how different modifications affect a nucleotide analog drug's access to the catalytic and allosteric sites of SAMHD1 (Fig. 2-6) when developing new therapies. Defining SAMHD1 interactions with nucleotide analog drugs may be critical for better predicting patient response to the current and future therapies. More detailed future studies may focus on how each of these analogs competes for relevant active sites. It will shed light on the extent to which nucleotide analogs can be combined to improve current treatments.

2.6 Future Directions

Nucleoside analog drugs are widely used to treat a variety of cancers and viral infections. With an essential role in regulating the nucleotide pool in the cell by degrading cellular nucleotides, SAMHD1 has the potential to decrease the cellular concentration of frequently prescribed nucleotide analogs and thereby decrease their clinical efficacy in cancer therapy. To improve future nucleotide analog treatments, it is important to understand SAMHD1 interactions with these drugs. Our work thoroughly examines the extent to which nucleotide analogs interact with the catalytic and allosteric sites of SAMHD1. This work contributes to the assessment of SAMHD1 as a potential therapeutic target for cancer therapy and the future design of SAMHD1 modulators that might improve the efficacy of existing therapies as outlined in the appendix.

2.7 Experimental Procedures

2.7.1 Nucleotide analog compounds

GTP, dATP, and dCTP were purchased from Thermo Scientific. Cytarabine-TP, cladribine-TP, clofarabine-TP, fludarabine-TP, gemcitabine-TP, ((2'R)-2'-F)-dCTP and ((2'S)-2'-F)-dCTP for *in vitro* assays were purchased from Jena Biosciences. The precursor nucleoside analogs (unphosphorylated) used for all cell culture assays were purchased from the following sources: Tocris (cytarabine, fludarabine, clofarabine, and cladribine), Accord Healthcare GmbH (gemcitabine) and Jena Bioscience (vidarabine and nelarabine). All nucleotide standards, internal standards, and nucleosides for the LC-MS/MS analysis were obtained from Sigma-Aldrich, Silantes or Alsachim (Thomas,

Herold et al., 2015). Cytarabine-$^{13}C_3$ was purchased from Santa Cruz and used for LC-MS/MS analysis.

2.7.2 Protein expression and purification

N-terminal 6×His-tagged SAMHD1 (residues 113-626) was expressed in *Escherichia coli* and purified using Ni-NTA affinity and size-exclusion chromatography as previously described (Ji et al., 2014).

2.7.3 Analytical size exclusion chromatography

Purified samples of SAMHD1 (2 mg/mL, 50 μl) mixed with a final concentration of 500 μM GTP and 4 mM nucleotide analog were applied to a Superdex 200 5/150 GL column (GE Healthcare) pre-equilibrated in 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$ and 0.5 mM tris(2-carboxyethyl)phosphine (TCEP). The UV absorbance at 280 nm was measured as the protein sample eluted from the column.

2.7.4 Analytical ultracentrifugation (AUC)

Sedimentation velocity experiments were performed with a Beckman XL-I analytical ultracentrifuge. Samples were prepared with protein concentration of 0.8-1.3 mg/mL in the buffer containing 50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$ and 0.5 mM TCEP and equilibrated with a final nucleotide concentration of 150 μM. AUC was performed at 35,000 revolutions per minute (r.p.m) and 20°C with an An60-Ti rotor. The experimental parameters including sample partial specific volume, buffer density and viscosity were calculated with SEDNTERP (http://sednterp.unh.edu/). Velocity data were analyzed using the program SEDFIT (Brown & Schuck, 2006).

2.7.5 Crystallization and data collection

Purified SAMHD1 protein in buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$ and 0.5 mM TCEP) was mixed with 1 mM GTP and 10 mM analog nucleotides (5 mM or 0.5 mM for gemcitabine-TP and vidarabine-TP, respectively) with or without 100 μM dATP (or 2.5 mM dATP for vidarabine-TP) and incubated at 4 °C for 15 minutes before crystallization. The small amount of dATP was included to ensure the formation of the SAMHD1 tetramer, as most nucleotide analogs do not bind Allo-site 2. All crystals were grown at 25 °C using the microbatch under-oil method by mixing 1 μL protein (5 mg/mL) with 1 μL crystallization buffer (100 mM SPG (Qiagen) buffer, pH 7.4, 25% PEG 1500). Crystals were cryoprotected by crystallization buffer supplemented with 25% (Vol/Vol) glycerol before frozen in liquid nitrogen. Diffraction data were collected at BNL beamline AMX and the Advanced Photon Source beamline 24-ID. The data statistics are summarized in Table 1.

2.7.6 Structure determination and refinement

The structures were solved by molecular replacement using PHASER (McCoy, Grosse-Kunstleve et al., 2007). We used the previously published SAMHD1 tetramer structure (PDB ID 4BZB), with the bound nucleotides removed, as the search model. The model was refined with iterative rounds of TLS and restrained refinement using *Refmac*5 (Vagin, Steiner et al., 2004) followed by rebuilding the model to the $2F_o$-F_c and the F_o-F_c maps using Coot (Emsley & Cowtan, 2004). Refinement statistics are summarized in Table 1. Coordinates and structure factors have been deposited in the Protein Data Bank, with accession codes listed in Table 1.

2.7.7 Malachite green activity assay

The enzymatic activity assay was adapted from (Seamon & Stivers, 2015). All assays were performed with the purified catalytic domain of SAMHD1 (residues 113-626) at 25°C in a reaction buffer containing 50 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM $MgCl_2$, and 0.5 mM TCEP. Each reaction, containing 10 μM *E. coli* inorganic pyrophosphatase, 0.5 μM SAMHD1, and 125 μM substrate or allosteric activator was quenched with 20 mM EDTA after 5 or 15 minutes. Then, Malachite Green reagent was added to the solution and developed for 15 minutes before the absorbance at 650 nm was measured.

2.7.8 HPLC-based kinetics assay

Reactions were initiated by the addition of pre-assembled SAMHD1 (final concentration of 500 nM) to 12-1600 μM nucleotide analog substrates and incubated at room temperature. Reactions were terminated by a 10× dilution into 20 mM EDTA after 5 minutes. Samples were deproteinized by spinning through an Amicon Ultra 0.5-mL 10-kDa filter (Millipore) for 10 min at 16,000×g. Samples were analyzed by HPLC with 100 μL of sample loaded on the Synergi C18 column 150 × 4.6 mm (Phenomenex). The column was pre-equilibrated in 20 mM ammonium acetate, pH 4.5 (buffer A) and samples were eluted at a flow rate of 1 mL/min with a gradient of methanol (buffer B) over 19 min. UV absorption was recorded at 260 nm.

Kinetic constants describing nucleotide analog hydrolysis were calculated with the Michaelis-Menten equation (eq. 1) using the Prism software (version 7.0a, GraphPad Software, La Jolla CA),

$$V = E_t * k_{cat} * X/(K_M + X) \quad \text{[eq. 1]}$$

where V is the enzyme velocity, E_t is the total amount of enzyme, and X is the concentration of nucleotide analog. E_t was constrained to a constant value of 0.005 nanomoles. Data shown in Table 2 indicate the mean of three independent experiments with standard error of three independent experiments.

2.7.9 Cells and cell culture

Human AML cell lines including THP-1 (DSMZ no. ACC16; FAB M6), OCI-AML2 (DSMZ No. ACC 99; FAB M4), OC-AML3 (DSMZ No. ACC 582; FAB M4), Molm13 (DSMZ No. ACC 554; FAB M5a), PL-21 (DSMZ No. ACC 536; FAB M3), HL-60 (DSMZ No. ACC 3; FAB M2), MV4-11 (DSMZ No. ACC 102; FAB M5), SIG-M5 (DSMZ No. ACC 468; FAB M5a), ML2 (DSMZ No. ACC 15; FAB M4), NB4 (DSMZ No. ACC 207; FAB M3), KG1 (DSMZ No. ACC 14; FAB not indicated), MonoMac6 (DSMZ No. ACC 124; FAB M5), and HEL (DSMZ No. ACC 11; FAB M6) were obtained from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH). A THP-1 cell clone engineered for SAMHD1 deficiency (THP-1 KO) and a corresponding SAMHD1-positive control cell clone (THP-1 Control) have been reported (Wittmann, Behrendt et al., 2015). All cell lines were cultured in IMDM (Biochrom) supplemented with 10 % FBS (SIG-M5 20% FBS), 4 mM L-Glutamine, 100 IU/mL penicillin, and 100 mg/mL streptomycin at 37°C in a humidified 5 % CO_2 incubator. Cells were routinely tested for mycoplasma contamination (LT07-710, Lonza) and authenticated by short tandem repeat profiling, as reported (Capes-Davis, Reid et al., 2013).

2.7.10 Cell viability assay

Viability of AML cell lines treated with various drug concentrations was determined by the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) dye reduction assay after 96 hours of incubation as described previously (Michaelis, Agha et al., 2015). IC_{50} values were determined using CalcuSyn (Biosoft).

2.7.11 Immunoblotting

Cells were lysed in Triton X-100 sample buffer and proteins separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis. Proteins were blotted onto a nitrocellulose membrane (Thermo Scientific). The membrane was incubated overnight at 4°C with primary antibodies used at the indicated dilutions: SAMHD1 (12586-1-AP, Proteintech, 1:1000), β-actin (3598R-100, BioVision via BioCat, 1:2000). Visualization and quantification were performed using fluorescently labeled secondary antibodies (926-32210 IRDye® 800CW goat anti-mouse and 926-32211 IRDye® 800CW goat anti-rabbit, LI-COR, 1:20000) and Odyssey LICOR.

2.7.12 LC-MS/MS analysis

Cells (1×10^6) were treated with 10 μM of the specific drug and incubated at 37 °C in a humidified 5% CO2 incubator for 6 h. Subsequently, cells were washed twice in 1 mL PBS, pelleted and stored at -20°C until measurement. The concentrations of dNTPs, $^{13}C_3$- cytarabine-TP, fludarabine-TP, clofarabine-TP, cladribine-TP and gemcitabine-TP, in the samples were measured by liquid chromatography-electrospray ionization-tandem mass spectrometry as previously described (Thomas et al., 2015). Briefly, the analytes were extracted by protein precipitation with methanol. An anion exchange HPLC column

(BioBasic AX, 150 x 2.1 mm, Thermo) was used for the chromatographic separation and a 5500 QTrap (Sciex) was used as analyzer, operating as triple quadrupole in positive multiple reaction monitoring (MRM) mode. The analysis of the dNTP was performed as previously described (Thomas et al., 2015). Additionally, $^{13}C_3$- cytarabine-TP, fludarabine-TP, clofarabine-TP, cladribine-TP and gemcitabine-TP were quantified using cytidine-$^{13}C_9$ -$^{15}N_3$-5'-triphosphate as an internal standard (IS). The precursor-to-product ion transitions used as quantifiers were: m/z 487.0 → 115.1 for $^{13}C_3$-cytarabine-TP, m/z 525.7 → 154.1 for fludarabine-TP, m/z 543.7 → 134.0 for clofarabine-TP, m/z 526.0 → 170.0 for cladribine-TP and m/z 504.0 → 326.0 for gemcitabine-TP. Due to the lack of commercially available standards, relative quantification was performed by comparing the peak area ratios (analyte/IS) of the differently treated samples.

2.7.13 Production of virus-like particles (VLPs)

VLPs, carrying either Vpx or Vpr from SIVmac251, were produced by co-transfection of 293T cells with pSIV3+ gag pol expression plasmids and a plasmid encoding VSV-G. The SIVmac251-based gag-pol expression constructs pSIV3+R- (Vpr-deficient) and pSIV3+X- (Vpx-deficient) were previously reported (Gramberg, Sunseri et al., 2010). The SAMHD1 degradation capacity of Vpx-VLPs was determined in THP-1 cells 24 h post transduction by intracellular SAMHD1 staining. AML cell lines were spinoculated with VSV-G pseudotyped VLPs carrying either Vpx or Vpr. Expression of SAMHD1 was monitored by Western blotting.

2.7.14 Statistical information

The average standard errors and standard deviations were calculated from multiple separate experiments as indicated in each figure legends and the results are shown in each graph.

2.7.15 Data deposition

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB), www.pdb.org (PDB ID codes 6DW4, 6DWD, 6DWK, 6DW3, 6DW5, AND 6DW7).

2.8 Acknowledgements

We thank J. Wang, T. Sasaki, V. Duong, and K. Anderson for technical assistance and discussions. This work was supported in part by NIH grants AI136737 (Y.X) and AI120845 (X.J.), by the Yale Cancer Center Pilot Grant (Y.X.), by Hilfe für krebskranke Kinder Frankfurt e.V. and the Frankfurter Stiftung für krebskranke Kinder (J.C and C.S), and by the Deutsche Forschungsgemeinschaft (KE 742/4-1) (O.T.K). O.B. was supported by the predoctoral program in Cellular and Molecular Biology NIH T32 GM007223 and by the National Science Foundation Graduate Research Fellowship DGE1122492. K.M.K. was supported by the predoctoral program in Biophysics NIH T32 GM008283. G.G. was supported by the Hessischen Landesoffensive zur Entwicklung wissenschaftlicher und ökonomischer Exzellenz (LOEWE-Center) Translational Medicine and Pharmacology. This work is in part based upon research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from NIH (P41 GM103403). The Pilatus 6M

detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

2.9 References

Amie SM, Bambara RA, Kim B (2013) GTP is the primary activator of the anti-HIV restriction factor SAMHD1. *J Biol Chem* 288: 25001-6

Amie SM, Daly MB, Noble E, Schinazi RF, Bambara RA, Kim B (2013) Anti-HIV host factor SAMHD1 regulates viral sensitivity to nucleoside reverse transcriptase inhibitors via modulation of cellular deoxyribonucleoside triphosphate (dNTP) levels. *J Biol Chem* 288: 20683-91

Arnold LH, Groom HC, Kunzelmann S, Schwefel D, Caswell SJ, Ordonez P, Mann MC, Rueschenbaum S, Goldstone DC, Pennell S, Howell SA, Stoye JP, Webb M, Taylor IA, Bishop KN (2015) Phospho-dependent Regulation of SAMHD1 Oligomerisation Couples Catalysis and Restriction. *PLoS Pathog* 11: e1005194

Arnold LH, Kunzelmann S, Webb MR, Taylor IA (2015) A continuous enzyme-coupled assay for triphosphohydrolase activity of HIV-1 restriction factor SAMHD1. *Antimicrob Agents Chemother* 59: 186-92

Balzarini J (2000) Effect of antimetabolite drugs of nucleotide metabolism on the anti-human immunodeficiency virus activity of nucleoside reverse transcriptase inhibitors. *Pharmacology & Therapeutics* 87: 175-187

Bastin-Coyette L, Cardoen S, Smal C, de Viron E, Arts A, Amsailale R, Van Den Neste E, Bontemps F (2011) Nucleoside analogs induce proteasomal down-regulation of p21 in chronic lymphocytic leukemia cell lines. *Biochem Pharmacol* 81: 586-93

Berger A, Sommer AF, Zwarg J, Hamdorf M, Welzel K, Esly N, Panitz S, Reuter A, Ramos I, Jatiani A, Mulder LC, Fernandez-Sesma A, Rutsch F, Simon V, Konig R, Flory E (2011) SAMHD1-deficient CD14+ cells from individuals with Aicardi-Goutieres syndrome are highly susceptible to HIV-1 infection. *PLoS Pathog* 7: e1002425

Bermejo M, Lopez-Huertas MR, Garcia-Perez J, Climent N, Descours B, Ambrosioni J, Mateos E, Rodriguez-Mora S, Rus-Bercial L, Benkirane M, Miro JM, Plana M, Alcami J, Coiras M (2016) Dasatinib inhibits HIV-1 replication through the interference of SAMHD1 phosphorylation in CD4+ T cells. *Biochem Pharmacol* 106: 30-45

Bonate PL, Arthaud L, Cantrell WR, Jr., Stephenson K, Secrist JA, 3rd, Weitman S (2006) Discovery and development of clofarabine: a nucleoside analogue for treating cancer. *Nat Rev Drug Discov* 5: 855-63

Brown PH, Schuck P (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophysical journal* 90: 4651-61

Capes-Davis A, Reid YA, Kline MC, Storts DR, Strauss E, Dirks WG, Drexler HG, MacLeod RA, Sykes G, Kohara A, Nakamura Y, Elmore E, Nims RW, Alston-Roberts C, Barallon R, Los GV, Nardone RM, Price PJ, Steuer A, Thomson J et al. (2013) Match criteria for human cell line authentication: where do we draw the line? *Int J Cancer* 132: 2510-9

Cribier A, Descours B, Valadao AL, Laguette N, Benkirane M (2013) Phosphorylation of SAMHD1 by cyclin A2/CDK1 regulates its restriction activity toward HIV-1. *Cell Rep* 3: 1036-43

Descours B, Cribier A, Chable-Bessia C, Ayinde D, Rice G, Crow Y, Yatim A, Schwartz O, Laguette N, Benkirane M (2012) SAMHD1 restricts HIV-1 reverse transcription in quiescent CD4(+) T-cells. *Retrovirology* 9: 87

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica Section D, Biological crystallography* 60: 2126-32

Ewald B, Sampath D, Plunkett W (2008) Nucleoside analogs: molecular mechanisms signaling cell death. *Oncogene* 27: 6522-37

Franzolin E, Pontarin G, Rampazzo C, Miazzi C, Ferraro P, Palumbo E, Reichard P, Bianchi V (2013) The deoxynucleotide triphosphohydrolase SAMHD1 is a major regulator of DNA precursor pools in mammalian cells. *Proc Natl Acad Sci U S A* 110: 14272-7

Fridle C, Medinger M, Wilk MC, Seipel K, Passweg J, Manz MG, Pabst T (2017) Cladribine, cytarabine and idarubicin (CLA-Ida) salvage chemotherapy in relapsed acute myeloid leukemia (AML). *Leuk Lymphoma* 58: 1068-1075

Galmarini CM, Mackey JR, Dumontet C (2001) Nucleoside analogues: mechanisms of drug resistance and reversal strategies. *Leukemia* 15: 875-890

Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, Walker PA, Kelly G, Haire LF, Yap MW, de Carvalho LP, Stoye JP, Crow YJ, Taylor IA, Webb M (2011) HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. *Nature* 480: 379-82

Gramberg T, Sunseri N, Landau NR (2010) Evidence for an activation domain at the amino terminus of simian immunodeficiency virus Vpx. *J Virol* 84: 1387-96

Hansen EC, Seamon KJ, Cravens SL, Stivers JT (2014) GTP activator and dNTP substrates of HIV-1 restriction factor SAMHD1 generate a long-lived activated state. *Proc Natl Acad Sci U S A* 111: E1843-51

Hollenbaugh JA, Shelton J, Tao S, Amiralaei S, Liu P, Lu X, Goetze RW, Zhou L, Nettles JH, Schinazi RF, Kim B (2017) Substrates and Inhibitors of SAMHD1. *PLoS One* 12: e0169052

Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, Florens L, Washburn MP, Skowronski J (2011) Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. *Nature* 474: 658-61

Ji X, Tang C, Zhao Q, Wang W, Xiong Y (2014) Structural basis of cellular dNTP regulation by SAMHD1. *Proc Natl Acad Sci U S A* 111: E4305-14

Ji X, Wu Y, Yan J, Mehrens J, Yang H, DeLucia M, Hao C, Gronenborn AM, Skowronski J, Ahn J, Xiong Y (2013) Mechanism of allosteric activation of SAMHD1 by dGTP. *Nat Struct Mol Biol* 20: 1304-9

Jordheim LP, Durantel D, Zoulim F, Dumontet C (2013) Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat Rev Drug Discov* 12: 447-64

Koharudin LM, Wu Y, DeLucia M, Mehrens J, Gronenborn AM, Ahn J (2014) Structural basis of allosteric activation of sterile alpha motif and histidine-aspartate domain-containing protein 1 (SAMHD1) by nucleoside triphosphates. *J Biol Chem* 289: 32617-27

Laguette N, Sobhian B, Casartelli N, Ringeard M, Chable-Bessia C, Segeral E, Yatim A, Emiliani S, Schwartz O, Benkirane M (2011) SAMHD1 is the dendritic- and myeloid-cell-specific HIV-1 restriction factor counteracted by Vpx. *Nature* 474: 654-7

Lahouassa H, Daddacha W, Hofmann H, Ayinde D, Logue EC, Dragin L, Bloch N, Maudet C, Bertrand M, Gramberg T, Pancino G, Priet S, Canard B, Laguette N, Benkirane M, Transy C, Landau NR, Kim B, Margottin-Goguet F (2012) SAMHD1 restricts the replication of human immunodeficiency virus type 1 by depleting the intracellular pool of deoxynucleoside triphosphates. *Nat Immunol* 13: 223-228

Liliemark J (1997) The clinical pharmacokinetics of cladribine. *Clin Pharmacokinet* 32: 120-31

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658-674

Miazzi C, Ferraro P, Pontarin G, Rampazzo C, Reichard P, Bianchi V (2014) Allosteric regulation of the human and mouse deoxyribonucleotide triphosphohydrolase sterile alpha-motif/histidine-aspartate domain-containing protein 1 (SAMHD1). *The Journal of biological chemistry* 289: 18339-46

Michaelis M, Agha B, Rothweiler F, Loschmann N, Voges Y, Mittelbronn M, Starzetz T, Harter PN, Abhari BA, Fulda S, Westermann F, Riecken K, Spek S, Langer K, Wiese M, Dirks WG, Zehner R, Cinatl J, Wass MN, Cinatl J, Jr. (2015) Identification of flubendazole as potential anti-neuroblastoma compound in a large cell line screen. *Sci Rep* 5: 8202

Ordonez P, Kunzelmann S, Groom HC, Yap MW, Weising S, Meier C, Bishop KN, Taylor IA, Stoye JP (2017) SAMHD1 enhances nucleoside-analogue efficacy against HIV-1 in myeloid cells. *Sci Rep* 7: 42824

Pauls E, Badia R, Torres-Torronteras J, Ruiz A, Permanyer M, Riveira-Munoz E, Clotet B, Marti R, Ballana E, Este JA (2014) Palbociclib, a selective inhibitor of cyclin-dependent kinase4/6, blocks HIV-1 reverse transcription through the control of sterile alpha motif and HD domain-containing protein-1 (SAMHD1) activity. *AIDS* 28: 2213-22

Pauls E, Ruiz A, Badia R, Permanyer M, Gubern A, Riveira-Munoz E, Torres-Torronteras J, Alvarez M, Mothe B, Brander C, Crespo M, Menendez-Arias L, Clotet B, Keppler OT, Marti R, Posas F, Ballana E, Este JA (2014) Cell cycle control and HIV-1 susceptibility are linked by CDK6-dependent CDK2 phosphorylation of SAMHD1 in myeloid and lymphoid cells. *J Immunol* 193: 1988-97

Radosevic N, Delmer A, Tang R, Marie JP, Ajchenbaum-Cymbalista F (2001) Cell cycle regulatory protein expression in fresh acute myeloid leukemia cells and after drug exposure. *Leukemia* 15: 559-566

Schneider C, Oellerich T, Baldauf HM, Schwarz SM, Thomas D, Flick R, Bohnenberger H, Kaderali L, Stegmann L, Cremer A, Martin M, Lohmeyer J, Michaelis M, Hornung V, Schliemann C, Berdel WE, Hartmann W, Wardelmann E, Comoglio F, Hansmann ML et al. (2017) SAMHD1 is a biomarker for cytarabine response and a therapeutic target in acute myeloid leukemia. *Nat Med* 23: 250-255

Seamon KJ, Stivers JT (2015) A High-Throughput Enzyme-Coupled Assay for SAMHD1 dNTPase. *J Biomol Screen* 20: 801-9

St Gelais C, de Silva S, Hach JC, White TE, Diaz-Griffero F, Yount JS, Wu L (2014) Identification of cellular proteins interacting with the retroviral restriction factor SAMHD1. *J Virol* 88: 5834-44

St Gelais C, Kim SH, Maksimova VV, Buzovetsky O, Knecht K, Shepard C, Kim B, Xiong Y, Wu L (2018) A Cyclin-binding Motif in Human SAMHD1 Is Required for Its HIV-1 Restriction, dNTPase Activity, Tetramer Formation, and Efficient Phosphorylation. *Journal of virology*

Tamamyan G, Kadia T, Ravandi F, Borthakur G, Cortes J, Jabbour E, Daver N, Ohanian M, Kantarjian H, Konopleva M (2017) Frontline treatment of acute myeloid leukemia in adults. *Crit Rev Oncol Hematol* 110: 20-34

Tang C, Ji X, Wu L, Xiong Y (2015) Impaired dNTPase activity of SAMHD1 by phosphomimetic mutation of Thr-592. *J Biol Chem* 290: 26352-9

Thomas D, Herold N, Keppler OT, Geisslinger G, Ferreiros N (2015) Quantitation of endogenous nucleoside triphosphates and nucleosides in human cells by liquid chromatography tandem mass spectrometry. *Anal Bioanal Chem* 407: 3693-704

Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta crystallographica Section D, Biological crystallography* 60: 2184-95

Wang Z, Bhattacharya A, Villacorta J, Diaz-Griffero F, Ivanov DN (2016) Allosteric Activation of SAMHD1 Protein by Deoxynucleotide Triphosphate (dNTP)-dependent Tetramerization Requires dNTP Concentrations That Are Similar to dNTP Concentrations Observed in Cycling T Cells. *The Journal of biological chemistry* 291: 21407-21413

Welbourn S, Dutta SM, Semmes OJ, Strebel K (2013) Restriction of virus infection but not catalytic dNTPase activity is regulated by phosphorylation of SAMHD1. *J Virol* 87: 11516-24

Whitley R, Alford C, Hess F, Buchanan R (1980) Vidarabine. *Drugs* 20: 267-282

Whitley RJ, Tucker BC, Kinkel AW, Barton NH, Pass RF, Whelchel JD, Cobbs CG, Diethelm AG, Buchanan RA (1980) Pharmacology, Tolerance, and Antiviral Activity of Vidarabine Monophosphate in Humans. *Antimicrobial Agents and Chemotherapy* 18: 709-715

Wittmann S, Behrendt R, Eissmann K, Volkmann B, Thomas D, Ebert T, Cribier A, Benkirane M, Hornung V, Bouzas NF, Gramberg T (2015) Phosphorylation of murine SAMHD1 regulates its antiretroviral activity. *Retrovirology* 12: 103

Yan J, Kaur S, DeLucia M, Hao C, Mehrens J, Wang C, Golczak M, Palczewski K, Gronenborn AM, Ahn J, Skowronski J (2013) Tetramerization of SAMHD1 is required for biological activity and inhibition of HIV infection. *J Biol Chem* 288: 10406-17

Zhu C, Gao W, Zhao K, Qin X, Zhang Y, Peng X, Zhang L, Dong Y, Zhang W, Li P, Wei W, Gong Y, Yu XF (2013) Structural insight into dGTP-dependent activation of tetrameric SAMHD1 deoxynucleoside triphosphate triphosphohydrolase. *Nat Commun* 4: 2722

Zhu CF, Wei W, Peng X, Dong YH, Gong Y, Yu XF (2015) The mechanism of substrate-controlled allosteric regulation of SAMHD1 activated by GTP. *Acta crystallographica Section D, Biological crystallography* 71: 516-24

3 Mapping interactions of maedi-visna virus Vif with cellular factors

3.1 Summary

The mammalian apolipoprotein B mRNA-editing enzyme catalytic polypeptide-like 3 (APOBEC3 or A3) family of cytidine deaminases restrict viral infections by mutating viral DNA and impeding reverse transcription. To overcome this antiviral activity, most lentiviruses express a viral accessory protein called Vif, which recruits A3 proteins to Cullin-RING E3 ubiquitin ligases for ubiquitylation and subsequent proteasomal degradation. While primate lentiviral Vif proteins utilize the transcription factor CBFβ as a non-canonical cofactor to stabilize the complex, maedi-visna virus (MVV) Vif hijacks cyclophilin A (CypA) instead. Since CBFβ and CypA are both highly conserved among mammals, the requirement for two different cellular cofactors suggests that these two Vif proteins have different biochemical and structural properties. Our results demonstrate that while some common motifs between HIV-1 Vif and MVV Vif are involved in recruiting Cul5, different determinants in MVV Vif are required in cofactor binding and stabilization of the E3 ligase complex. Results from this study advance our understanding of the mechanism of MVV Vif recruitment of cellular factors and the evolution of lentiviral Vif proteins.

3.2 Introduction

Lentiviruses have engaged in an evolutionary molecular arms race with mammalian host immune systems for millions of years (Gifford 2011; Harris et al. 2016). A common mechanism that viruses have developed to overcome restriction is hijacking

the host cellular degradation machinery pathways to destroy host antiviral restriction factors (Barry and Fruh 2006). For example, most lentiviruses encode the accessory protein known as virion infectivity factor (Vif) to target the APOBEC3 (apolipoprotein B mRNA-editing catalytic polypeptide-like 3 or A3) family of restriction factors for destruction by the proteasome (Yu et al. 2003; Conticello et al. 2003; Sheehy et al. 2003). In the absence of Vif, A3s potently block lentiviral infection by hypermutating viral DNA and blocking reverse transcription (Albin and Harris 2010).

In contrast to the highly conserved E3 ligase components that Vif hijacks, the A3 family of proteins in mammals are much more diverse (LaRue et al. 2010; Yoshikawa et al. 2016). Different mammals have varying numbers of A3 genes, each containing one or two domains with a zinc-coordinating DNA cytosine deaminase motif that can be clustered into three distinct phylogenetic groups called Z1, Z2, or Z3 (LaRue et al. 2008). While the seven human A3s are named A3A-A3H, the three sheep (and all other non-primate) A3s are simply named for their Z domain (LaRue et al. 2009). While there is only 49% amino acid sequence identity between human A3H and the corresponding sheep oaA3Z3 protein, MVV Vif can degrade both of them while HIV-1 selectively degrades A3H and not oaA3Z3 (Larue et al. 2010). Variations between the target A3 proteins from different hosts could potentially lead to major differences in the structural and biochemical requirements for HIV-1 and MVV Vif proteins. Notably, the same HIV-1 Vif has developed different binding modes for different human A3 proteins (Aydin et al. 2014; Hu et al. 2019).

One striking difference between HIV-1 and MVV Vif is the requirement for different non-canonical cofactors to stabilize the E3 ligase complex. HIV-1 Vif requires

the transcription factor CBFβ to form stable E3 ligase complexes and degrade A3s (Jager et al. 2012; Zhang et al. 2012). Outside the context of binding HIV-1 Vif, CBFβ is known to form a heterodimer with the RUNX1 transcription factor and control the gene expression of cellular processes, including T cell development (Collins et al. 2009). It has been suggested that limiting HIV-1 competes with RUNX1 for CBFβ binding and alters RUNX1 target gene expression (Kim et al. 2013). Instead of CBFβ, MVV Vif requires the prolyl isomerase cyclophilin A (CypA) as its non-canonical cofactor (Kane et al. 2015). CypA is an immunophilin molecule which regulates immune responses and is also a cofactor for HIV-1 infection through interactions with the viral capsid (Watashi et al. 2007). Analysis of CypA active site mutants suggest that prolyl isomerase activity is not necessary for association of CypA in a reconstituted MVV Vif complex, but is important for A3 degradation in vivo (Kane et al. 2015). Although CBFβ and CypA are not related to each other, they are each highly conserved from humans and sheep (100% and 99% amino acid sequence identity between human and sheep for CBFβ and CypA, respectively). This divergence in cofactor requirement could point to a difference in the structure between these different lentiviral Vifs. It is also possible that the different biological functions of these CBFβ and CypA apply constraints to the viruses.

Extensive studies have characterized the interactions between HIV-1 Vif and human E3 ligase components. A crystal structure of HIV-1 Vif with CBFβ, EloB/C, and the N-terminus of Cul5 revealed that the HIV-1 Vif has a novel protein fold composed of two domains: a small α domain (helices α3 and α4) and a larger α/β domain (Guo et al. 2014; Figure 1A). Within the α domain, a zinc-binding HCCH motif is important for positioning helix α3 to interact with Cul5 and a modified SOCS box motif is important

for helix α4 for to interact with EloC (Xiao et al. 2006; Mehle et al. 2006; Yu et al. 2004). The α/β domain of HIV-1 Vif includes extensive interactions with CBFβ and is known to be the site of binding for APOBEC3s (Aydin et al. 2014; Hu et al. 2019). Despite this detailed understanding of HIV-1 interactions with host factors, it is unknown whether other lentiviral Vif proteins share similar structural architecture.

In this study, we set out to understand the molecular determinants for MVV Vif binding to cellular components. We use biochemical assays to analyze the effect of mutations and truncations on purified Vif protein binding to CypA and Cul5. We identified a metal-binding motif of MVV Vif that is important for CypA binding. We have also explored the effect of an IR motif on MVV Vif binding to Cul5. Finally, our results suggest that a putative C-terminal helix is required for CypA binding. Notably, none of the mutations or truncations tested here had any negative effects on MVV Vif binding to EloB/C. Together, these results highlight important similarities and differences between MVV Vif and other lentiviral Vif proteins and allow a better understanding of their evolution.

3.3 Results

3.3.1 A zinc coordination motif is important for MVV Vif binding to CypA

Vif proteins originating from different lentiviruses have been shown to bind a zinc ion with different motifs. For instance, HIV-1 Vif contains a HCCH zinc binding motif which is conserved among primate lentiviruses (Luo et al. 2005; Paul et al. 2005; Mehle et al. 2006; Xiao et al. 2006). The HCCH motif of HIV-1 Vif is structurally important for positioning helix α3 to interact with Cul5 (Guo et al. 2014; Figure 3-1A). Likewise,

KCCC, CCCC, and HCCC motifs have been proposed to be structurally important for zinc binding in FIV, CAEV, and BIV Vif proteins, respectively (Figure 3-1B, Stern et al. 2010; Zhao et al. 2019; Zhang et al. 2014). This diversity among zinc binding motifs highlights biochemical flexibility of this structural element while maintaining a conserved function.

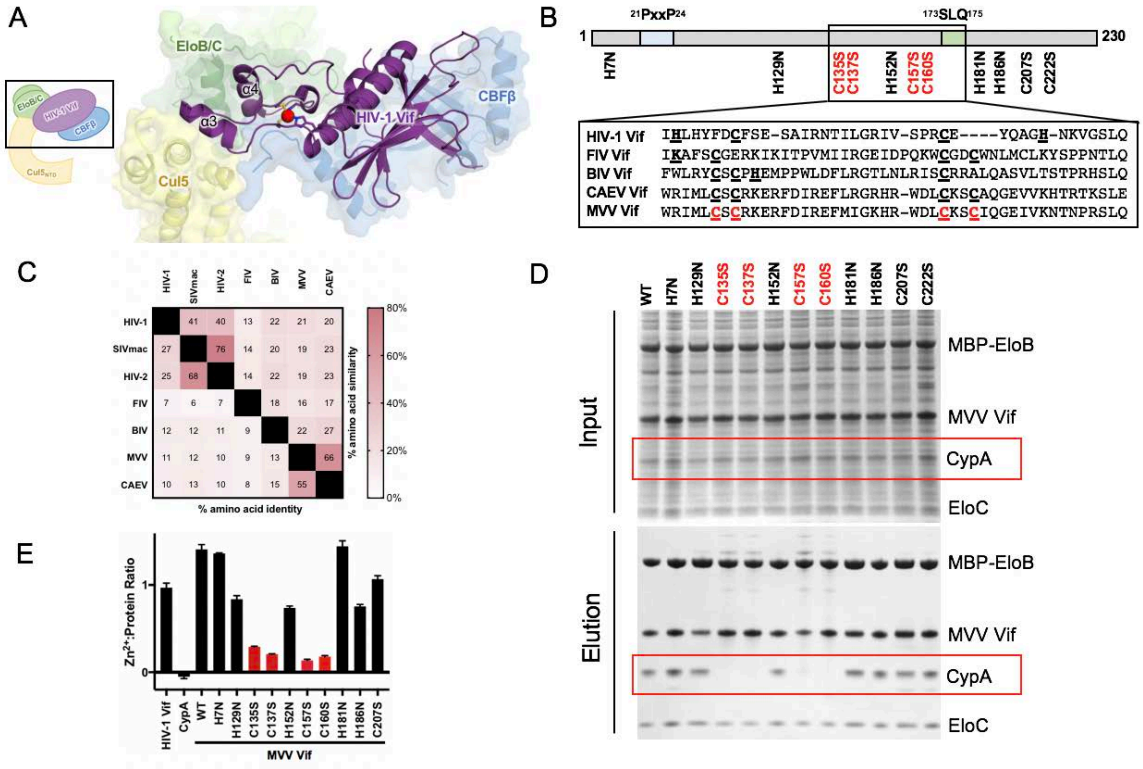

Figure 3-1. Zinc coordination motif is important for MVV Vif binding to CypA

(A) Crystal structure of HIV-1 Vif (purple cartoon) shows zinc (red sphere) and Cul5 interface (yellow cartoon) (Guo et al. 2014, PDB ID: 49NF). (B) Bar diagram representation of MVV Vif protein highlight previously identified motifs important for CypA and EloC binding (blue and green, respectively) and mutants created for this assay. Inset show alignment of lentiviral Vif proteins in region of protein where zinc coordination motifs have been identified (bold and underlined residues). (C) Amino acid sequence similarity and identity comparisons of representative lentiviral Vif proteins. (D) Coomassie-stained SDS-PAGE analysis of purified protein complexes. L = cell lysate; E = final purified protein elution after sequential NiNTA and maltose affinity purifications. (E) A colorimetric assay for metal released from purified protein after denaturation. Error bars represent standard deviation from the mean.

Interestingly, while the HCCH motif of HIV-1 Vif is important for Cul5 binding, it is dispensable for interaction with its cellular cofactor CBFβ (Fribourgh et al. 2014; Hu et al. 2019; Mehle et al. 2006; Xiao et al. 2006). In contrast, a recent study showed that the zinc-binding CCCC motif of CAEV Vif was essential for binding its cofactor CypA (Zhao et al. 2019). This difference in the function of the zinc-binding motifs may suggest significant structural differences in the overall fold of HIV-1 Vif compared to MVV Vif.

As members of the larger Small Ruminant Lentivirus continuum, MVV Vif and CAEV Vif are highly conserved with each other, especially compared to other Vif proteins outside this group (Figure 3-1C; Ramirez et al. 2013; Minardi da Cruz et al. 2013; Highland 2017). The highest amino acid sequence identity between the MVV and CAEV Vif sequences is found within the α domain region that contains this CCCC motif. This suggests that CAEV and MVV most likely use very similar mechanisms to bind EloB/C and Cul5.

To test whether MVV Vif also contains a zinc-binding motif, we individually mutated all histidines and cysteines of the MVV Vif protein to residues that do not coordinate metals, purified protein sub-complexes, and used a colorimetric assay to test for the presence of zinc ions. The mutant MVV Vif proteins were tagged with 6xHistidine and co-expressed with CypA, MBP-tagged EloB, and EloC in *E. coli.* Sequential affinity purification steps were used to isolate sub-complexes with both the 6xHistidine-tagged Vif mutant and the MBP-tagged EloB. As observed by SDS-PAGE analysis, all four components of the sub-complex were found in similar levels in the cell lysate, but mutations to the CCCC motif previously characterized in CAEV Vif resulted in a dramatic reduction of binding to CypA (Figure 3-1D). This result shows that, like

CAEV Vif (Zhao et al. 2019), the CCCC motif of MVV Vif is important for co-factor CypA binding and therefore the overall structural integrity of Vif itself.

To show that the CCCC motif of MVV Vif coordinates a zinc ion, we used a colorimetric Zincon assay to measure the amount of zinc co-purified with the mutant Vif complexes. Purified Vif protein complexes were denatured, deproteinated, and the resulting supernatant was assessed for metal ions, as previously described (Sabel et al. 2010). A zinc-to-protein ratio was calculated by dividing the concentration of purified protein by the concentration of metal ions detected using Zincon (Figure 3-1E). As controls, a purified HIV Vif complex was shown to contain one zinc ion per protein molecule and purified CypA was shown not to contain any metal ions, as expected. While wildtype and most mutant MVV Vif proteins bind a single zinc ion, mutations to the CCCC motif dramatically reduced the amount of zinc extracted from the Vif complex (Figure 3-1E). Notably, disrupting this motif and CypA binding did not affect binding to EloB/C. This indicates that the CCCC zinc-binding motif of MVV Vif is conserved both in sequence and in function with CAEV Vif's CCCC motif in recruiting the cofactor CypA.

3.3.2 IR Motif is important for MVV Vif binding to Cul5

Hydrophobic residues in helix α3 of HIV-1 Vif are important for Cul5 binding and recruitment of the E3 ligase complex (Xiao et al. 2006; Guo et al. 2014; Figure 3-2A). Recently, an IR motif found in a similar region of FIV Vif and CAEV Vif has been shown to be important for Cul5 binding (Gu et al. 2018; Zhang et al. 2019; Figure 3-2B). A consensus sequence of HIV-1 Vif shows that a hydrophobic residue is frequently found

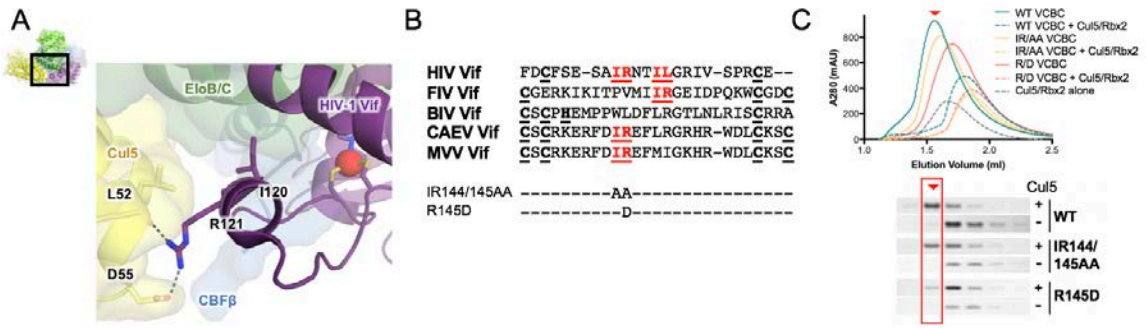

Figure 3-2. IR motif is important for MVV Vif binding to Cul5

(A) Crystal structure of HIV-1 Vif (purple) binding to Cul5 (yellow) (Guo et al. 2014, PDB ID: 4N9F). (B) *Top*: Alignment of lentiviral Vifs highlighting IR motifs shown to be important for Cul5 binding in red. *Bottom*: Mutants of MVV Vif tested in this assay. (C) *Top*: SEC analysis of MVV Vif complexes with or without Cul5. *Bottom*: Coomassie-stained SDS-PAGE analysis of fractions from SEC assay. Only band for MVV Vif is shown for clarity.

in the position corresponding to this isoleucine residue (Xiao et al. 2006). The HIV-1 Vif E3 ligase structure maps these residues to the α3 helix and shows the corresponding arginine in HIV-1 Vif interacts with Cul5 directly (Guo et al. 2014). While the isoleucine contributes to a hydrophobic surface between HIV-1 Vif and EloC, the arginine participates in hydrogen bonding with the carbonyl of L52 and side chain of D55 of Cul5. Secondary structure prediction software also predicts MVV Vif to have an α helix in this region which also contains an IR motif (residues I144 and R145). However, the molecular determinants of MVV Vif binding to Cul5 have yet to be elucidated.

To test whether this IR motif is important for MVV Vif binding to Cul5, we performed site-directed mutagenesis in this region (Figure 3-2B). Purified mutant Vif sub-complexes were analyzed by size exclusion chromatography (SEC) alone and mixed with Cul5 and peak fractions were analyzed by SDS-PAGE (Figure 3-2C). While each mutant sub-complex elutes similarly in the absence of Cul5, the addition of Cul5 results

in varying degrees of larger complex formation compared to wildtype. The most significant shifts in the SEC elution profile and SDS-PAGE gel band intensity were observed for wild-type MVV Vif upon addition of Cul5/Rbx2, indicating strongest binding affinity (red arrow in Figure 3-2C). The IR144/145AA Vif complex modestly affects the elution shift indicating that binding has not been severely disrupted. It is possible that these residues do contribute to the overall interface between MVV Vif and Cul5, but that alanine mutations are not deleterious enough to abrogate the larger interaction. We found that a more radical mutation (R145D) was necessary to abrogate the interaction between MVV Vif and Cul5. Based on the HIV-1 Vif E3 ligase structure, this substitution of a negatively charged aspartate with the positively charged arginine likely results in a charge repulsion with residue D55 of Cul5, not allowing Cul5 to bind efficiently. This result suggests that interface of MVV Vif with Cul5 contains a region similar to other lentiviral Vif interactions with Cul5.

3.3.3 A putative C-terminal helix is important for MVV Vif binding to CypA

To map the regions necessary for binding to CypA, we performed a series of MVV Vif truncations to assess the requirement of the N- or C-terminus for binding. A putative helix lies at both the N- and C-termini of the protein, so we set out to test whether either of these predicted secondary structural elements are required for cofactor binding (Figure 3-3A). The putative N-terminal helix (residues 4-14) lies just upstream of the ^{21}PxxP^{24} motif which has been shown to bind the catalytic pocket of CypA (Kane et al. 2015). While CAEV and MVV Vifs share this motif, other lentiviral Vifs do not possess this N-terminal extension. Similarly, the C-terminal helix (residues 218-226) is conserved in CAEV Vif sequence, but not in other Vifs.

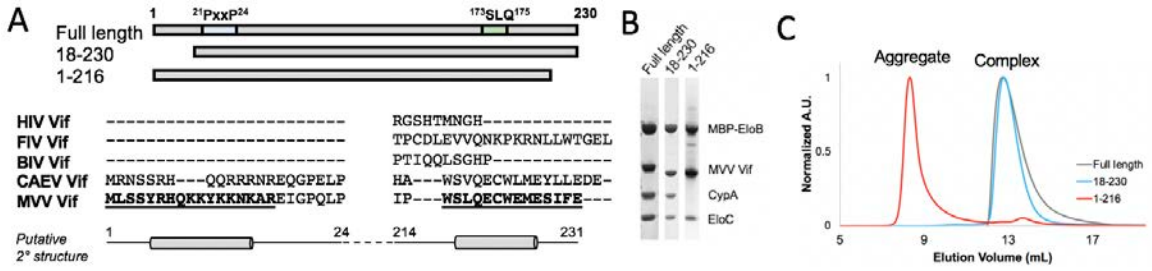

Figure 3-3. Putative C-terminal helix is important for MVV Vif binding to CypA

(A) *Top*: Bar diagram for N- and C-terminal truncation constructs. *Bottom*: Sequence alignment of lentiviral Vif proteins at the N-terminus and C-terminus with putative secondary structure prediction of MVV Vif below. (B) Coomassie-stained SDS-PAGE and (C) size exclusion chromatography analysis of truncated MVV Vif complexes compare to wildtype MVV Vif.

MVV Vif truncations were co-expressed with CypA, EloB, and EloC in *E. coli*, co-purified using affinity purification, and purity was analyzed by SDS-PAGE (Figure 3-3B). While deleting the N-terminus of MVV Vif did not seem to affect overall complex behavior, deleting the C-terminal helix abrogated MVV Vif binding to CypA and caused the complex to aggregate (Figure 3-3C). This demonstrates that this putative C-terminal helix is important for the overall structural integrity of MVV Vif. It is possible that residues within this putative helix directly participate in important CypA interactions or indirectly contribute to the overall structural integrity and folding of the Vif protein. Neither truncation affected the ability of MVV Vif to bind EloB/C, supporting a model where destabilization of the α/β domain does not affect the ability of the α domain to bind EloB/C.

3.4 Discussion

Despite little sequence conservation among Vif proteins from different lentiviruses, they are functionally conserved to target the host restriction factor APOBEC3 proteins by hijacking the same cellular ubiquitin-proteasome degradation pathways. To achieve this common function, some level of structural conservation also exists among the lentiviral Vifs. Interestingly, many of the known lentiviral Vifs need an additional cofactor for stability and function, in which cases divergence emerges. In particular, it is rather intriguing that the unrelated cofactors CBFB and CypA exploited by HIV-1 and MVV Vifs, respectively, are each highly conserved (>99% identical) and simultaneously exist in both host cells, and yet each virus apparently independently evolved to depend on one of the two. The results presented herein advance our understanding of the similarities and differences of these two Vif-cellular cofactor interactions (Figure 3-4), which provides insights into the cellular immunity mechanisms of each host that the different Vif proteins have evolved alongside.

Our results support that lentiviral Vif proteins hijack the canonical component of the Cul5-based E3 ligase machinery of the host through a common but slightly variable set of interactions. Since Cul5 and other E3 ligase components are highly conserved among mammals, it is expected that each Vif would use similar binding interfaces with the ligase. The importance of the conserved Vif BC-box in binding EloB/EloC is well established (Yu et al. 2003; Larue et al. 2010). While the Cul5-binding interface seems to be moderately conserved in the amino acid sequence of Vif, it is unclear how much the molecular details of the interactions are maintained among different Vif proteins. For instance, the Cul5-binding region of BIV Vif has been modeled to have a slightly

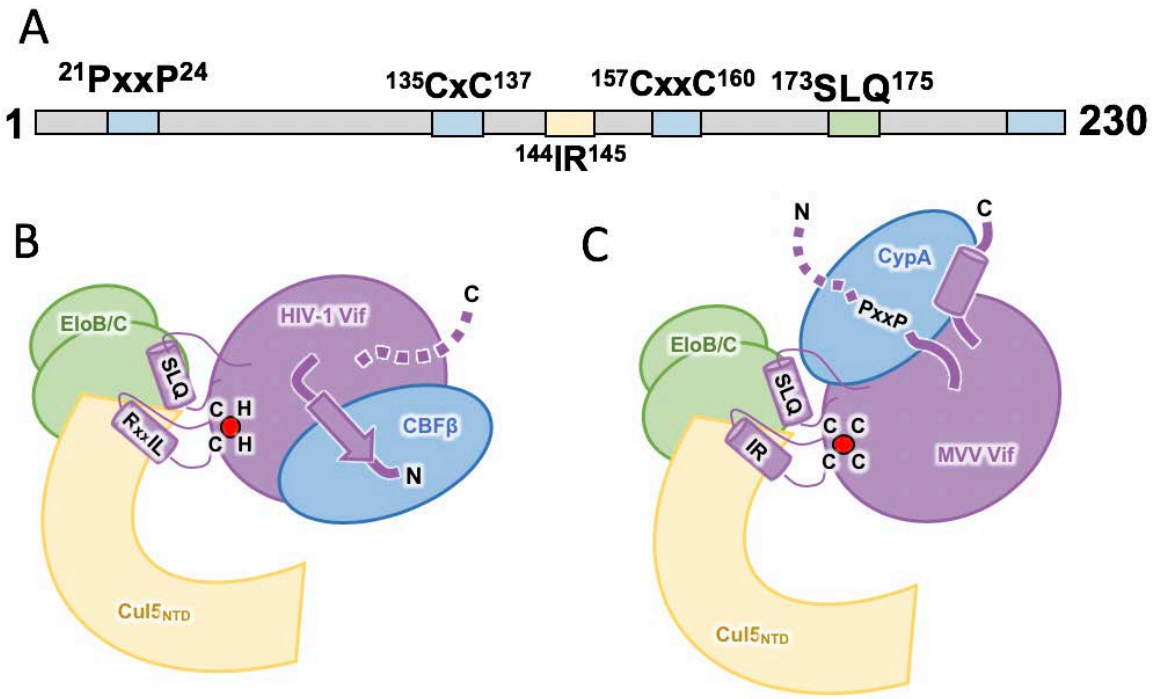

Figure 3-4. Summary of MVV Vif binding motifs

(A) Schematic of MVV Vif motifs important for binding cellular factors. Regions important for binding CypA are shown in blue, Cul5 in yellow, and EloB/C in green. (B) Cartoon schematic of HIV-1 Vif interactions with CBFβ, Cul5, and EloB/C. (C) Cartoon schematic of MVV Vif interactions with CypA, Cul5, and EloB/C. Important residues including the C135/C137/C157/C160 zinc-binding motif, I144/R145 Cul5-binding motif, and the P21/P24 CypA-binding motif are highlighted. N- and C-termini are labeled.

different structural fold than that of HIV-1 Vif (Zhang et al. 2014). Similarly, the IR motif of FIV Vif is predicted to have a Cul5-binding mode distinct from that observed in the crystal structure of the HIV-1 Vif E3 ligase complex (Gu et al. 2018). Even the closely related CAEV Vif and MVV Vif may have slightly different interactions with Cul5. While a single alanine mutation of either IR motif residue of CAEV Vif abolished binding to Cul5 (Zhang et al. 2019), our results show that a more severe charge-swapped mutation to this region is required to abrogate MVV Vif binding to Cul5. Nevertheless,

the involvement of the IR motif in different lentiviral Vifs points to a similar interaction at a common site.

Our data further provide a better understanding of the different modes of interactions employed by lentiviral Vif proteins to recruit divergent non-canonical cofactors. Our mapping results highlight the unique components of MVV Vif which contribute to CypA binding. Combining this information with the HIV-1 Vif structure allows us to start building a conceptual model for MVV Vif binding to CypA (Figure 4B). Previously, it was shown that the ^{21}PxxP^{24} motif found at the N-terminal region of MVV Vif likely interacts with the active site of CypA (Kane et al. 2015). We found that a putative α–helix preceding the ^{21}PxxP^{24} motif is not required for the interaction. It is possible that inherent flexibility in this region of MVV Vif is important for the ability of the ^{21}PxxP^{24} motif to reach into CypA's narrow active site binding groove. This is in contrast to the critical N-terminal β-strand of HIV-1 Vif that inserts into the CBFβ core, forming an integral part of the structure. The relative importance of a structural element at the C-terminus of MVV Vif is another striking divergence from the HIV-1 Vif structure. While the C-terminal region of HIV-1 Vif is thought to be flexible and dispensable for CBFβ interaction, our study shows that the putative C-terminal helix of MVV Vif is essential for binding to CypA.

Our results also reveal a unique cofactor-recruiting function in the CCCC motif of MVV Vif. This motif corresponds to the HCCH motif in HIV-1 Vif, whose integrity is required for positioning the Vif α-domain for interaction with Cul5 but is dispensable for recruiting the cofactor CBFβ (Mehle et al. 2006; Xiao et al. 2006; Fribourgh et al. 2014, Hu et al 2019). In contrast, the CCCC motif of MVV Vif is essential for interaction with

the cofactor CypA in our experiments. While it is likely that the CCCC motif of MVV Vif is important for Cul5 binding as well, it cannot be directly tested with our aggregated MVV Vif variant complexes which purified without CypA. This motif is thought to be a conserved zinc-binding element in lentiviral Vif molecules, as demonstrated for HIV-1 Vif and the highly similar CAEV Vif. Consistently, our biochemical data demonstrated the metal binding capability of the CCCC motif of MVV Vif. However, a previous report found that adding the Zinc-chelating TEPN to tissue culture samples does not affect MVV Vif's ability to degrade A3 protein *in vivo* (Zhang *et al.* 2014). This discrepancy may be contributed by the complexity of tissue culture samples. It is possible that the zinc was not sufficiently chelated from MVV Vif, which might have a very high affinity for the metal ion once it is bound. It is also technically possible that there are other metal-independent mechanisms that allow MVV Vif to degrade A3s in cells.

The work described herein show that although key characteristics of lentiviral Vif proteins are preserved to recruit host E3 ubiquitin ligase machinery, very different mechanisms have evolved for the binding of non-canonical cofactors. The α-domain of Vif is likely highly conserved in structure and function in recognition for host Cul5 and EloB/C. The metal-binding motif in the conserved α-domain can nonetheless adopt additional functions in recruiting the non-canonical cofactors. However, the larger α/β domain of Vif proteins which has evolved to interact with a variety of target A3 proteins and highly divergent non-canonical cofactors (such as CBFβ or CypA) are likely very different in both sequence and structure. Future studies will determine the extent to which different Vif proteins have evolved different molecular mechanisms for interacting with these targets and cofactors. Also, more work is needed to understand the effect of Vif and

its non-canonical cofactor on the catalysis of ubiquitin ligation. It is unclear whether Vif influences the normal regulation of these E3 ligase complexes. For instance, interactions with CAND1 or neddylation of Cul5 might be influenced by Vif and therefore change the ability of these complexes to recognize and ubiquitinate substrates.

3.5 Experimental Procedures

3.5.1 Protein expression and purification

Co-expression of MVV Vif (N-terminal 6xHis tag) with CypA, EloB (N-terminal MBP tag) and EloC was induced in BL21DE3 cells with the addition of 300 μM IPTG. After an overnight induction at 16 °C, cells were harvested via centrifugation, resuspended in Buffer A (50 mM Tris pH 8, 400 mM NaCl, 10 mM imidazole, 0.1 mM TCEP). Cells were lysed with a microfluidizer, clarified via centrifugation, and applied to a NiNTA resin. After washing the column with Buffer A, the protein complex was eluted with Buffer B (50 mM Tris-HCl pH 8, 400 mM NaCl, 400 mM imidazole, 0.1 mM TCEP), and applied to amylose resin equilibrated with Buffer C (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.1 mM TCEP). Protein complex was eluted with Buffer D (50 mM Tris-HCl pH 8, 150 mM NaCl, 10 mM maltose, 0.1 mM TCEP). The purified protein complex was applied to a SEC column equilibrated in Buffer C, and the subsequent peak was pooled, concentrated and stored at -80 °C. HIV-1 Vif was co-expressed with

3.5.2 Sequence alignment and identity matrix

ClustalW (Madeira et al. 2019) and Sequence Manipulation Suite (Stothard 2000) were used to create sequence alignments and generate sequence identities between different representative Vif amino acid sequences. The Vif sequences were obtained from

GenBank, and the accession numbers are: P12504.1 for HIV-1, BAM76138.1 for HIV-2, ANT85793.1 for SIVmac, BAX00793.1 for FIVpetaluma, NP_040564.1 for BIV, ALP75978.1 for CAEV, and CAA01215.1 for MVV.

3.5.3 Zincon assay

Zincon, borate, and Zn^{2+} standard solutions were prepared as previously described (Sabel *et al.* 2010). Purified protein samples (300 μM) were acidified and denatured by the addition of concentrated HCl and trichloroacetic acid to final concentrations of 0.2 M and 10% (w/v), respectively. Denatured protein precipitate was removed by centrifugation at 16,000 rcf for 5 minutes. The resulting clarified supernatant was transferred to a fresh tube and neutralized with NaOH. After this neutralized sample was diluted 40-fold into 50 mM borate buffer at pH 9, 1.6 mM Zincon stock solution was added for a final concentration of 40 μM Zincon. Mixtures were mixed immediately, and absorbance at 620 nm was measured after 30 minutes using a plater reader. Values plotted represent the average of 3 repeats with error bars representing standard deviation.

3.5.4 SEC binding assay

Purified protein samples (50 μL, 40 μM) were applied to a Superdex 200 5/150 GL column (GE Healthcare) pre-equilibrated in Buffer C. Protein elution was monitored via measuring absorbance at 280 nm and SDS-PAGE analysis of fractions.

3.6 References

Albin JS, Harris RS (2010) Interactions of host APOBEC3 restriction factors with HIV-1 in vivo: implications for therapeutics. Expert Rev Mol Med 12: e4

Aydin H, Taylor MW, Lee JE (2014) Structure-guided analysis of the human APOBEC3-HIV restrictome. Structure 22: 668-84

Barry M, Fruh K (2006) Viral modulators of cullin RING ubiquitin ligases: culling the host defense. Sci STKE 2006: pe21

Conticello SG, Harris RS, Neuberger MS (2003) The Vif Protein of HIV Triggers Degradation of the Human Antiretroviral DNA Deaminase APOBEC3G. Current Biology 13: 2009-2013

Fribourgh JL, Nguyen HC, Wolfe LS, Dewitt DC, Zhang W, Yu XF, Rhoades E, Xiong Y (2014) Core binding factor beta plays a critical role by facilitating the assembly of the Vif-cullin 5 E3 ubiquitin ligase. J Virol 88: 3309-19

Gifford RJ (2012) Viral evolution in deep time: lentiviruses and mammals. Trends Genet 28: 89-100

Gu Q, Zhang Z, Gertzen CGW, Haussinger D, Gohlke H, Munk C (2018) Identification of a Conserved Interface of Human Immunodeficiency Virus Type 1 and Feline Immunodeficiency Virus Vifs with Cullin 5. J Virol 92

Guo Y, Dong L, Qiu X, Wang Y, Zhang B, Liu H, Yu Y, Zang Y, Yang M, Huang Z (2014) Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. Nature 505: 229-33

Harris RS, Anderson BD (2016) Evolutionary Paradigms from Ancient and Ongoing Conflicts between the Lentiviral Vif Protein and Mammalian APOBEC3 Enzymes. PLoS Pathog 12: e1005958

Highland MA (2017) Small Ruminant Lentiviruses: Strain Variation, Viral Tropism, and Host Genetics Influence Pathogenesis. Vet Pathol 54: 353-354

Hu Y, Desimmie BA, Nguyen HC, Ziegler SJ, Cheng TC, Chen J, Wang J, Wang H, Zhang K, Pathak VK, Xiong Y (2019) Structural basis of antagonism of human APOBEC3F by HIV-1 Vif. Nat Struct Mol Biol 26: 1176-1183

Jager S, Kim DY, Hultquist JF, Shindo K, LaRue RS, Kwon E, Li M, Anderson BD, Yen L, Stanley D, Mahon C, Kane J, Franks-Skiba K, Cimermancic P, Burlingame A, Sali A, Craik CS, Harris RS, Gross JD, Krogan NJ (2011) Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. Nature 481: 371-5

Kane JR, Stanley DJ, Hultquist JF, Johnson JR, Mietrach N, Binning JM, Jonsson SR, Barelier S, Newton BW, Johnson TL, Franks-Skiba KE, Li M, Brown WL, Gunnarsson HI, Adalbjornsdottir A, Fraser JS, Harris RS, Andresdottir V, Gross JD, Krogan NJ (2015) Lineage-Specific Viral Hijacking of Non-canonical E3 Ubiquitin Ligase Cofactors in the Evolution of Vif Anti-APOBEC3 Activity. Cell Rep 11: 1236-50

Kim DY, Kwon E, Hartley PD, Crosby DC, Mann S, Krogan NJ, Gross JD (2013) CBFbeta stabilizes HIV Vif to counteract APOBEC3 at the expense of RUNX1 target gene expression. Mol Cell 49: 632-44

Klase Z, Yedavalli VS, Houzet L, Perkins M, Maldarelli F, Brenchley J, Strebel K, Liu P, Jeang KT (2014) Activation of HIV-1 from latent infection via synergy of RUNX1 inhibitor Ro5-3335 and SAHA. PLoS Pathog 10: e1003997

LaRue RS, Andresdottir V, Blanchard Y, Conticello SG, Derse D, Emerman M, Greene WC, Jonsson SR, Landau NR, Lochelt M, Malik HS, Malim MH, Munk C, O'Brien SJ, Pathak VK, Strebel K, Wain-Hobson S, Yu XF, Yuhki N, Harris RS (2009) Guidelines for naming nonprimate APOBEC3 genes and proteins. J Virol 83: 494-7

LaRue RS, Jonsson SR, Silverstein KA, Lajoie M, Bertrand D, El-Mabrouk N, Hotzel I, Andresdottir V, Smith TP, Harris RS (2008) The artiodactyl APOBEC3 innate immune

repertoire shows evidence for a multi-functional domain organization that existed in the ancestor of placental mammals. BMC Mol Biol 9: 104

Larue RS, Lengyel J, Jonsson SR, Andresdottir V, Harris RS (2010) Lentiviral Vif degrades the APOBEC3Z3/APOBEC3H protein of its mammalian host and is capable of cross-species activity. J Virol 84: 8193-201

Madeira F, Park Ym, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research 47: W636-W641

Mehle A, Thomas ER, Rajendran KS, Gabuzda D (2006) A zinc-binding region in Vif binds Cul5 and determines cullin selection. J Biol Chem 281: 17259-65

Minardi da Cruz JC, Singh DK, Lamara A, Chebloune Y (2013) Small ruminant lentiviruses (SRLVs) break the species barrier to acquire new host range. Viruses 5: 1867-84

Ramírez H, Reina R, Amorena B, de Andrés D, Martínez HA (2013) Small ruminant lentiviruses: genetic variability, tropism and diagnosis. Viruses 5: 1175-207

Sabel CE, Neureuther JM, Siemann S (2010) A spectrophotometric method for the determination of zinc, copper, and cobalt ions in metalloproteins using Zincon. Anal Biochem 397: 218-26

Stern MA, Hu C, Saenz DT, Fadel HJ, Sims O, Peretz M, Poeschla EM (2010) Productive replication of Vif-chimeric HIV-1 in feline cells. J Virol 84: 7378-95

Stothard P (2000) The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28: 1102, 1104

Watashi K, Shimotohno K (2007) Cyclophilin and viruses: cyclophilin as a cofactor for viral infection and possible anti-viral target. Drug Target Insights 2: 9-18

Xiao Z, Ehrlich E, Yu Y, Luo K, Wang T, Tian C, Yu XF (2006) Assembly of HIV-1 Vif-Cul5 E3 ubiquitin ligase through a novel zinc-binding domain-stabilized hydrophobic interface in Vif. Virology 349: 290-9

Yoshikawa R, Izumi T, Nakano Y, Yamada E, Moriwaki M, Misawa N, Ren F, Kobayashi T, Koyanagi Y, Sato K (2016) Small ruminant lentiviral Vif proteins commonly utilize cyclophilin A, an evolutionarily and structurally conserved protein, to degrade ovine and caprine APOBEC3 proteins. Microbiol Immunol 60: 427-36

Yu X, Yu Y, Liu B, Luo K, Kong W, Mao P, Yu XF (2003) Induction of APOBEC3G ubiquitination and degradation by an HIV-1 Vif-Cul5-SCF complex. Science 302: 1056-60

Yu Y, Xiao Z, Ehrlich ES, Yu X, Yu XF (2004) Selective assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 ubiquitin ligase complex through a novel SOCS box and upstream cysteines. Genes Dev 18: 2867-72

Zhang W, Du J, Evans SL, Yu Y, Yu XF (2011) T-cell differentiation factor CBF-beta regulates HIV-1 Vif-mediated evasion of host restriction. Nature 481: 376-9

Zhang W, Wang H, Li Z, Liu X, Liu G, Harris RS, Yu XF (2014) Cellular requirements for bovine immunodeficiency virus Vif-mediated inactivation of bovine APOBEC3 proteins. J Virol 88: 12528-40

Zhao Z, Li Z, Huan C, Wang H, Su X, Zhang W (2019) CAEV Vif Hijacks ElonginB/C, CYPA and Cullin5 to Assemble the E3 Ubiquitin Ligase Complex Stepwise to Degrade oaA3Z2-Z3. Front Microbiol 10: 565

4 Reconstituting MVV Vif E3 ligase complexes for structural studies

4.1 Preface

In this chapter, I outline some interesting observations made in pursuit of the still elusive high-resolution structure of MVV Vif complexes. This work was done in close collaboration with colleagues who assisted with preparing electron microscopy (EM) grids, EM imaging, and EM data analysis. Dr. Katie Digianantonio and Kaifeng Zhou assisted with screening samples, and the talented post doctoral fellow Dr. Yingxia Hu created all of the EM images, 2D class averages, and 3D reconstructions presented in this chapter. Additionally, Dr. Bill Eliason and Dr. Swapnil Devarkar performed the light scattering experiment in Section 4.4.4.

4.2 Introduction

The Cullin-RING ubiquitin ligases are found in all eukaryotic cells, and many viruses have evolved mechanisms to hijack this machinery to target host proteins for degradation by the proteasome (Barry and Fruh 2006). By studying these virus-host interactions, we stand to gain more information about biological processes related to both viral and cellular proliferation. Although the structure of HIV-1 Vif with cellular E3 ligase components was solved (Guo et al. 2014), it is unknown whether other lentiviral Vif proteins utilize similar structural features to assemble their respective E3 ligase complexes.

One striking difference between HIV-1 Vif and MVV Vif is their requirement for different non-canonical cofactors. HIV-1 Vif requires CBFβ, while MVV Vif requires CypA (Jager et al. 2012; Kane et al. 2015). CBFβ and CypA are unrelated in structure or biological function, and they cannot substitute one another in HIV-1 or MVV Vif complexes. Since lentiviral Vif proteins are not well conserved, it is hard to predict how MVV Vif would interact with CypA. Interactions with CBFβ dominate the structural integrity of HIV-1 Vif, so it is unknown whether other lentiviral Vif proteins will have similar structural features.

In this chapter, I summarize the reconstitution of MVV Vif complexes for structural studies, including crystallography and cryo-electron microscopy (cryoEM). First, I outline construct designs that helped to stabilize the 4-component MVV Vif sub-complexes. Then, I summarize work that has been assembling the 6-component complex, including Cul5/Rbx2. In the last section, I detail surprising results that we observed when CypA was fused to MVV Vif, and how it might help us to solve the structure in the future.

4.3 Results and Discussion

4.3.1 MBP tagged EloB improves expression and purification of 4-component MVV Vif complex from *E. coli*

In our pursuit of the MVV Vif structure, we first sought to purify the minimal stable complex. Early attempts to co-express large amounts of MVV Vif, CypA, and EloB/C in *E. coli* were unsuccessful. I found that the 6xHis-Thioredoxin (TRX)-CypA fusion protein used by Kane et al. (2015) over-expressed as a single component, but low

expression of the other three components made purification of 4-component complexes (MVV Vif, CypA, EloB, and EloC or VCBC) challenging. After removing the 6xHis-TRX tag from CypA and adding a 6xHis tag to MVV Vif, I found that expression of all four components were equally low.

After many expression and purification attempts without solubility tags, I found that the addition of a maltose binding protein (MBP) to the N-terminus of EloB greatly improved expression of the entire VCBC complex in *E. coli* and allowed for the purification of large amounts of VCBC complex (Figure 4-1A). I found that the sequence of Ni-NTA resin, amylose resin, anion exchange, and size exclusion chromatography (SEC) yielded a VCBC complex of high purity and monodispersity. A typical preparation of this complex can yield about 40 mg of purified protein per 2 L culture. Although the anion exchange step can be omitted, I found this results in a large void volume peak with a high 260/280 ratio appears by gel filtration. Since I do not see this when I purify CypA or EloB/C alone, this suggests that MVV Vif forms higher order oligomers mediated by co-purified cellular nucleic acids. It has been shown that HIV-1 Vif binds nucleic acids in a manner that is important for viral packaging and infectivity (Dettenhoffer et al. 2000; Zhang et al. 2000; Khan et al. 2001). It is possible that RNA binding is a conserved functional feature between HIV-1 and MVV Vif proteins.

The MBP tag likely aids in expression by increasing the solubility of the VCBC complex. The MBP tag is rigidly fused to EloB, meaning that there is no flexible linker between the C-terminus of MBP and N-terminus of EloB. Since EloB/C and CypA are both easily expressed in *E. coli* and purified without solubility tags, it was surprising that a solubility tag on EloB/C could help the expression and behavior of MVV Vif. While the

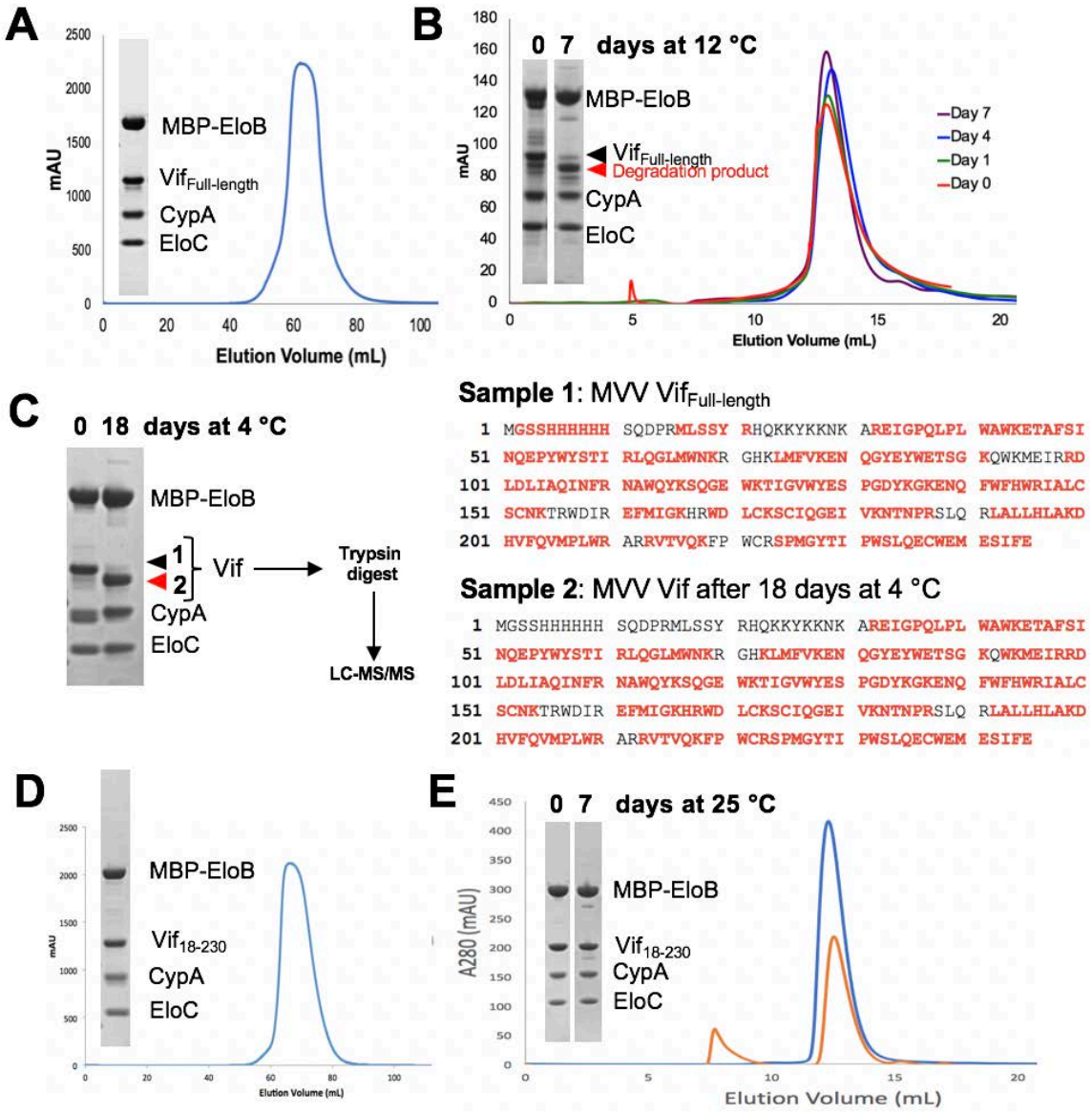

Figure 4-1. N-terminal residues of MVV Vif is not necessary for complex formation.

(A) SEC and SDS-PAGE analysis of VCBC complex. (B) Stability assay showing SEC and SDS-PAGE analysis of VCBC complex after incubation at 12 °C. (C) Summary of mass spectrometry experiment and results. (D) SEC and SDS-PAGE analysis of VCBC complex with truncated N-terminus of MVV Vif. (E) Stability assay showing SEC and SDS-PAGE analysis of N-terminally truncated MVV Vif VCBC complex after incubation at 25 °C.

sequence of cofactor binding or a detailed mechanism of MVV Vif protein folding is unknown, it has been suggested that the BC box of HIV-1 Vif begins to fold upon binding of EloB/C (Stanley et al. 2008; Bergeron et al. 2010). This together with the data shown in Chapter 3 showing that MVV Vif binds EloB/C in the absence of CypA suggests that binding to EloB/C may be an important first step in the formation of an VCBC complex or MVV Vif folding.

4.3.2 N-terminal residues of MVV Vif are not required for CypA binding

Since large amounts of the initial VCBC complex with full-length Vif (residues 1-230) was easy to obtain and exhibited good solution behavior, it became our first target for crystallization. Many crystallization trays were set up with this initial complex at different concentrations (~0.5-5 mg/mL), incubated at different temperatures (4, 12, and 25 °C) with or without additives (LMNG and DDM) using the microbatch under oil method. However, no protein crystal hits were obtained for this complex. In a test for the stability of the VCBC complex, I incubated an aliquot of VCBC complex at 12 °C for several days and observed that the complex remained stable as monitored by SEC (Figure 4-1B). However, SDS-PAGE analysis of the aged sample revealed that individual protein components of the complex had begun to degrade. Most notably, the MVV Vif band had almost entirely shifted down about 4 kDa in size to a stable degradation product. This degradation is likely due to contaminating proteases that presumably cut flexible or unstructured regions of the MVV Vif protein.

We hypothesized that this stable degradation product of MVV Vif could become a minimal construct target for crystallography, so we set out to identify the region of MVV

Vif that had been degraded. I incubated another aliquot of the VCBC complex at 4 °C for 18 days, analyzed the resulting sample by SDS-PAGE and excised bands from the gel that corresponded to the full-length and degraded MVV Vif (Figure 4-1C). With assistance from Yale's Keck Mass Spectrometry services group, these protein bands were trypsin-digested and analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS). Results from this experiment show that peptides could be confidently mapped to most of the protein sequence including the C-terminus for both samples. However, N-terminal residues were missing from the degraded MVV Vif sample. Specifically, the N-terminal 6xHis tag and residues 1-16 of MVV Vif were not captured by LC-MS/MS. Since trypsin cuts peptides at lysine and arginine residues which are present in this region, the exact N-terminus of the degraded MVV Vif cannot be unambiguously determined. Nonetheless, we created a new construct of MVV Vif starting at residue 18. The resulting protein was easily expressed in *E. coli* and purified similarly as full-length MVV Vif (Figure 4-1D). Furthermore, this construct was more stable than full-length MVV Vif at 25 °C (Figure 4-1E). Again, I placed a lot of effort into setting up crystallization trays with this construct at different temperatures and concentrations, but no crystals were obtained. Crystallization efforts might be hindered by the low solubility of this complex.

The sequence of the N-terminus of MVV Vif is interesting for a few reasons. For one, it is the site of CypA binding to the ^{21}PxxP^{24} motif. Also, it has an unusually long stretch of positively charged residues. Secondary structure prediction server predicts that there may be a helix in this region. However, it is possible that this charged stretch is unstructured. Perhaps flexibility in this region is actually require to help the ^{21}PxxP^{24}

motif reach into the narrow binding groove of CypA. While the N-terminal residues are not necessary for the complex formation shown here, it is very possible that these residues are important for other biological functions *in vivo*. More studies will be needed to test the effect of deleting these residues from MVV Vif on viral infectivity in cells.

4.3.3 6-component MVV Vif complex is a difficult EM target

After successful purification of the 4-component MVV Vif complex, I sought to reconstitute larger E3 ligase complexes including Cullin proteins. It is currently unclear whether MVV Vif selective hijacks Cul2 or Cul5 in the cell and whether the Cul2/Cul5 binding for MVV Vif is similar to HIV-1 Vif. One study showed that Cul5, but not Cul2, bound MVV Vif overexpressed in 293T cells (Zhang et al. 2014). Another study showed that N-terminally tagged MVV Vif could bind Cul5 and Cul2, but C-terminally tagged MVV Vif could only bind Cul2 (Kane et al. 2015). In the same assay, both Cul2 and Cul5 bound HIV-1 Vif regardless of whether it was N- or C-terminally tagged. This is consistent with the structure of HIV-1 Vif bound to $Cul5_{NTD}$ which shows that neither the N- or C- terminus is close to the Cul5 binding interface (Guo et al. 2014). Together, these results suggest that the overall structure of MVV Vif binding to Cul2/Cul5 will differ from that of HIV-1 Vif.

To learn more about how MVV Vif recruits Cullins to its E3 ligase complex, we sought to obtain a structure with either Cul2 or Cul5. I found that MVV Vif bound both Cul2 or Cul5 using SEC (Figures 4-2A and 4-2B). When mixed with equimolar amounts of VCBC complex and Cul2 or Cul5, a large shift in the elution profile indicates that a larger complex has been formed. The SDS-PAGE analysis also confirms that all

components of the complex are present. These results confirm that MVV Vif is capable of binding both Cul2 and Cul5 *in vitro*, but more detailed studies are needed to determine the physiological role of each during MVV infection.

We ultimately chose to pursue a 6-component MVV Vif complex with Cul5/Rbx2 (abbreviated as VCBC-CR) for structural studies because it yielded reproducibly monodisperse peaks by SEC. I proceeded to set up crystallization trays, but found that solubility of this complex is similarly low as the VCBC complex. Next, we turned to electron microscopy (EM). CryoEM images showed particles which resembled the C-shaped architecture that Cullin-RING E3 ligase complexes are known to adopt (Figure 4-2C). Subsequent 2D class averages were promising showing detailed secondary structural features, especially the helical repeats of the Cul5 (Figure 4-2D). However, it was clear that the distribution of VCBC-CR complex particles was dominated by a few preferred orientations. While it is easy to pick out C-shaped side views of the particles, other views of the particles from different angels were rare. Also, some particles were dissociated into smaller species due to damage from the air-water interface (Glaeser and Han 2017). Attempts to shield the protein with detergents, such as DDM or LMNG, have not been successful at improving sample quality. Moreover, the 2D class averages showed that the density of the VCBC portion of the complex appeared more blurred than the rest of the complex (Figure 4-2D), indicating an intrinsic flexibility within the VCBC subcomplex or between the VCBC subcomplex and Cul5/Rbx2.

It was possible to create a low-resolution 3D reconstruction of the VCBC-CR (Figure 4-2E). The resulting electron density resembles the C-shaped architecture that we expect from this complex. However, the resolution is limited by the preferred orientation

of the molecule buried in ice and its intrinsic flexibility in solution. Recent studies suggest that the HIV-1 Vif complex may have inherent flexibility between its domains (Ball et al. 2019). If MVV Vif is also inherently dynamic in nature, this would hinder our ability to capture a single dominant structure. It may be possible that further optimization of this six-component VCBC-CR complex could yield a higher resolution reconstruction, but significant challenges, such as the preferred orientation of particles and complex flexibility, would need to be overcome.

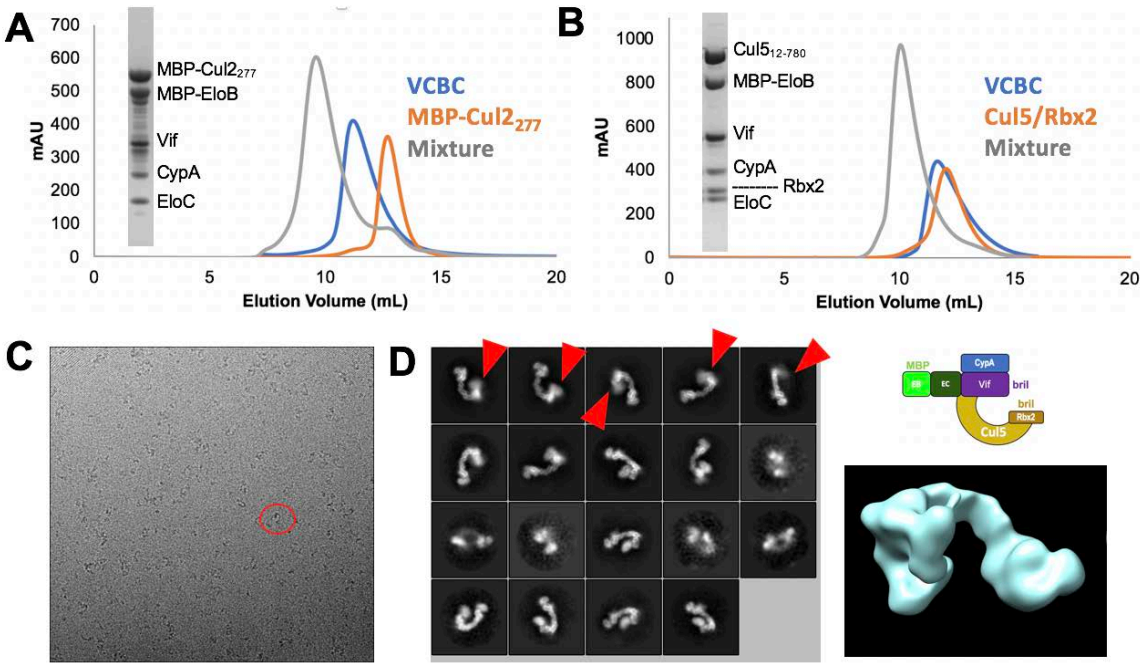

Figure 4-2. 6-component MVV Vif complex is heterogeneous

(A) SEC and SDS-PAGE analysis of MVV Vif complex with Cul2. (B) SEC and SDS-PAGE analysis of MVV if complex with Cul5/Rbx2. (C) Representative cryoEM image of VCBC-CR complex with a C-shaped particle circled in red. (D) Representative 2D class averages of VCBC-CR complex with blurred VCBC region highlighted with red arrows in top row. (E) 3D reconstruction and cartoon schematic of VCBC-CR complex.

4.3.4 Fusing CypA to N-terminus of MVV Vif results in higher order oligomers

In some stability assays, I observed that a small aggregate peak forms over time, which was composed of Vif/EloB/EloC without CypA. I wondered whether the VCBC complex would be more stable if the MVV Vif-CypA interaction was strengthened by fusing them together. In order to increase the probability that Vif and CypA stay bound throughout the course of our structural optimization trials, I fused CypA to the N-terminus of the MVV Vif (residues 18-230) by engineering a flexible amino acid linker between the two proteins. When designing this fusion construct, I wanted to balance two considerations. The first consideration was that excess linker could contribute to unstructured and flexible regions that are not conducive to crystallization and may be easily degraded by contaminating proteases. Secondly, I did not know the distance between the C-terminus of CypA and the N-terminus of MVV Vif in the context of the three-dimensional VCBC complex. The linker length I chose must be sufficiently long enough to allow the native conformation of the binding interaction to exist.

With this in mind, I created three constructs with three different linker lengths (0, 10, and 30 amino acids) and tested their expression and behavior empirically. While CypA-0-Vif and CypA-10-Vif constructs did not express well in *E. coli*, the CypA-30-Vif was expressed and purified similarly as the original VCBC complex (Figure 4-3A). It is likely that the CypA-0-Vif and CypA-10-Vif constructs did not express well because the shorter linkers did not give Vif and CypA enough freedom to bind in the correct orientation. For reference, the 30 amino acid linker is approximately 105 angstroms long (assuming an average length of 3.5 Angstroms for each amino acid), so the MVV Vif and CypA proteins have ample flexibility to bind. For reference, the longest dimension of the

HIV-1 Vif E3 ligase complex containing HIV-1 Vif, CBFβ, EloB/C, and $Cul5_{NTD}$ is about 120 angstroms (Guo et al. 2014).

One surprising characteristic of the CypA-30-Vif complex with MBP-EloB/C (abbreviated C30VBC) is that it appears to form oligomers. In a SEC coupled multi-angle light scattering (MALS) experiment comparing the original VCBC complex to the C30VBC complex, the VCBC complex appears close to its expected "monomer" size while the C30VBC is approximately a "dimer" (Figure 4-3B). When the C30VBC sample is concentrated, I have even observed higher order oligomers of C30VBC by SEC. This unexpected dimer species may aid the structural determination by Cryo-EM, which would otherwise be limited by the small size of the monomer VCBC particles.

An initial negative stain grid was extremely promising, showing homogeneous particles that seemed to be approximately the correct size for a C30VBC dimer (Figure 4-3C). Next, Alicia optimized grid freezing conditions for cryoEM, and collected data on the Krios Titan at Yale's West Campus microscopy facilities. After picking particles, 2D classifications revealed a molecule with apparent C2 symmetry (Figure 4-3D). Encouragingly, the 2D classifications show detailed secondary structure features of the central core of the molecule, but remain fuzzy at the furthest tips of the molecule. An initial low-resolution 3D reconstruction indicates that the central portion of the molecule is the dimer C30V portion of the complex, and the extended arms of the molecule represent the MBP-EloB/C (Figure 4-3E). We plan to repeat these grid conditions, in hopes that these C30VBC particles will bring a high-resolution 3D reconstruction if we collect more data.

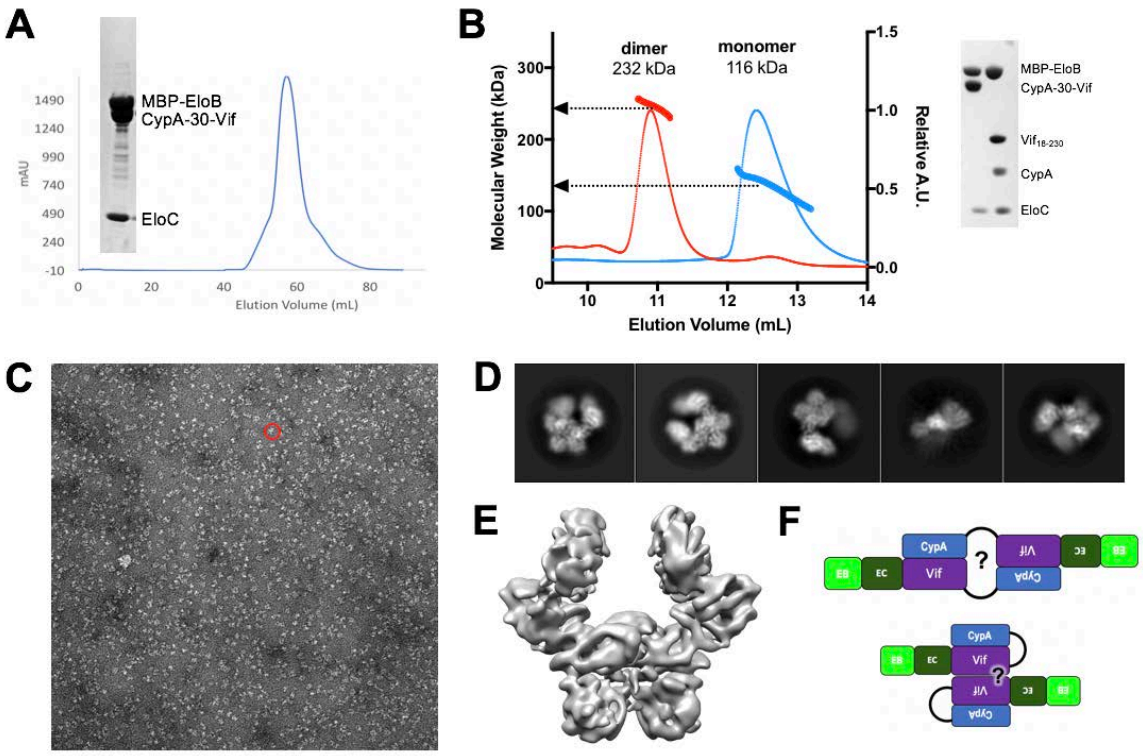

Figure 4-3. C30VBC complex forms oligomers

(A) SEC and SDS-PAGE analysis of C30VBC complex. (B) SEC-MALS and SDS-PAGE analysis comparing VCBC to C30VBC samples. (C) Representative negative stain EM image of C30VBC complex. (D) 2D class averages calculated from a cryo-EM dataset of C30VBC. (E) 3D reconstruction of C30VBC complex. (F) Cartoon schematics showing two possible dimerization mechanisms of C30VBC.

4.4 Future Directions

Solving the structure of MVV Vif will allow us to understand the molecular details of the MVV Vif-CypA interaction and assess the structural conservation between MVV Vif and HIV-1 Vif. From an evolutionary standpoint, it will be interesting to see how two lentiviral Vifs have evolved to recruit different non-canonical cofactors to their respective E3 ligases to perform similar functions. Although we have not solved a high-resolution structure of MVV Vif yet, my experience with purifying these complexes and

optimizing the samples for crystallization and EM sheds light on potential avenues to get there in the future.

While working with the MVV Vif complexes, I have found that it is essential to have a solubility tag on the complex to keep MVV Vif behaving well. I have focused mostly on the MBP-tagged EloB/C construct in this chapter, but I have also had some success purifying samples with the smaller b_{562}RIL solubility tag (Chun et al. 2012). In the future, complexes with this b_{562}RIL tag should be screened more extensively for crystallization and EM. It is possible that the MBP tag had inhibited crystal formation in previous constructs and this more compact solubility tag could stabilize the complex and allow for different crystal packing alternatives.

Additionally, while the full-length Cul5/Rbx2 constructs have been the main target of our structural studies, it is possible that smaller truncations of N-terminal truncations of Cul5 or Cul2 could make a better crystallization construct. For instance, the $Cul5_{NTD}$ crystallized with HIV-1 Vif only contained the first three helical repeats of Cul5 (Guo et al. 2014). One way that we could potentially stabilize and improve the solution behavior of these protein constructs would be to fuse the b_{562}RIL tag to the C-terminus of these truncated proteins in an attempt to shield hydrophobic residues left exposed by the truncation. In addition to pursuing the structures of MVV Vif with Cul5 of Cul2, it will also be interesting to interrogate the binding determinants for MVV Vif to bind these proteins.

The C30VBC complex is perhaps the most intriguing construct for structural studies because of its interesting oligomerization potential. Although the C30VBC

complex alone would most likely be too small for EM, the possibility of a stable dimer can push it into a range that is feasible with current technology. If we can succeed in solving the structure of this MVV Vif dimer, more studies should be done to determine the biological relevance of the dimer and understand the molecular details of dimerization. It is still not clear whether the dimer formed by C30VBC is indicative of a true physiological property of MVV Vif or an artifact of the fusion construct (Figure 4-3F). While it has been shown that dimerization of HIV-1 Vif through its PPLP motif is important for infectivity (Yang et al. 2001; Yang et al. 2003; Auclair et al. 2007; Miller et al. 2007; Donahue et al. 2008), dimerization has not yet been suggested to be important for MVV Vif activity *in vivo*. If the subsequent mutational analysis proves that this dimer is in fact important for MVV infectivity *in vivo*, this could shed light onto yet another similarity between HIV-1 Vif and MVV Vif.

A limitation to these experiments is that not all components of the catalytically active E3 ligase complex are present in our study. First of all, we have not yet purified the substrate, ovine A3 proteins, to test binding to our reconstituted E3 ligase complexes. It would be interesting to map how they interact with MVV Vif and whether there are any direct interactions with the non-canonical cofactor, CypA. Secondly, we have not included the ubiquitin-conjugating E2 enzyme in any of our complexes. From these studies, we will not be able to understand how the overall complex architecture allows for catalysis. Finally, even when Cul5-Rbx2 are present in our reconstituted E3 ligase complex, we have never attempted to neddylate Cul5. Since neddylation causes a conformational change to the overall architecture of the complex. In the future, it might

be beneficial to neddylate Cul5 before analyzing the complex by cryoEM, as it might improve overall behavior or reduce troublesome flexibility of the complex.

4.5 Experimental Procedures

4.5.1 Constructs and cloning

MVV Vif and CypA constructs described by Kane et al. (2015) were gifted from John Gross at University of California San Francisco. MVV Vif was cloned into the first multiple cloning site (MCS) with an N-terminal 6xHis tag, and CypA was cloned into the second MCS of a pET-Duet vector. To create the C30V construct, CypA and a thirty amino acid linker were cloned between the 6xHis tag and MVV Vif in the first MCS of a pET-Duet vector. EloB (residues 1–118) with an N-terminal MBP tag and EloC (residues 17–112) were cloned into the pCDF-Duet vector. Cul5 (residues 12-780) and Rbx2 were cloned into a pRSF-Duet vector.

4.5.2 Expression and Purification

Plasmids encoding 6xHis tagged MVV Vif, CypA, MBP-tagged EloB, and EloC were transformed into *E. coli* BL-21(DE3) cells and protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) at 16°C in Terrific Broth overnight. Plasmids encoding 6xHis tagged Cul5 and Rbx2 were transformed into *E.coli* BL-21(DE3) cells and protein expression induced with 0.5 mM IPTG at 16°C in Luria Broth overnight.

Cells were harvested and lysed by a microfluidizer. For MVV Vif complexes, the lysate was clarified by centrifugation and the supernatant was applied first to a Ni-NTA

resin, then maltose resin, and followed by anion exchange and size exclusion chromatography. Cul5 and Rbx2 were purified similarly, omitting the maltose resin step. Proteins were analyzed after each step by SDS-PAGE.

4.5.3 LC-MS/MS

MVV Vif complexes were analyzed by SDS-PAGE before and after a long 18 day incubation at 4 °C. MVV Vif gel bands were excised from the gel using a clean razor blade, and samples were analyzed by Jean Kanyo in the Mass Spectrometry & Proteomics department of the W. M. Keck Foundation Biotechnology Resource Laboratory. After trypsin digestion, liquid chromatography tandem mass spectrometry was performed. The spectra were analyzed using the Mascot algorithm using the MVV Vif sequence as a custom database (Hirosawa et al. 1993).

4.5.4 Negative stain sample preparation

After size exclusion chromatography, a peak fraction of purified protein complex was chosen for analysis by negative stain EM. Sample was applied to a carbon coated copper grid (Electron Microscopy) pre-treated by glow-discharging at 25 mA for 30 seconds, and then stained with 2% uranyl acetate for 45 seconds. Images were collected by Alicia Hu on a Jeol JEM-1230 electron microscope in Kline Biology Tower or a FEI Talos L120C electron microscope in Bass Center at Yale University.

4.5.5 CryoEM sample preparation

After size exclusion chromatography, a peak fraction of purified protein complex was chosen for analysis by cryoEM. Sample was applied to a C-Flat2/1-3C copper grid

(Electron Microscopy Sciences) pretreated by glow-discharging at 8 mA for 20 seconds. The grid was blotted at 10 °C with 100% humidity and plunge-frozen in liquid ethane using a FEI Vitrobot Mark IV (Thermo Fisher). Images were acquired by Alicia Hu on the FEI Talos L120C microscope and the FEI Titan Krios G2 at Yale University. EM data was processed by Alicia Hu using RELION (Scheres et al. 2012).

4.6 References

Ball KA, Chan LM, Stanley DJ, Tierney E, Thapa S, Ta HM, Burton L, Binning JM, Jacobson MP, Gross JD (2019) Conformational Dynamics of the HIV-Vif Protein Complex. *Biophys J* 116: 1432-1445

Bergeron JR, Huthoff H, Veselkov DA, Beavil RL, Simpson PJ, Matthews SJ, Malim MH, Sanderson MR (2010) The SOCS-box of HIV-1 Vif interacts with ElonginBC by induced-folding to recruit its Cul5-containing ubiquitin ligase complex. *PLoS Pathog* 6: e1000925

Chun E, Thompson AA, Liu W, Roth CB, Griffith MT, Katritch V, Kunken J, Xu F, Cherezov V, Hanson MA, Stevens RC (2012) Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* 20: 967-76

Dettenhofer M, Cen S, Carlson BA, Kleiman L, Yu XF (2000) Association of human immunodeficiency virus type 1 Vif with RNA and its role in reverse transcription. *J Virol* 74: 8938-45

Donahue JP, Vetter ML, Mukhtar NA, D'Aquila RT (2008) The HIV-1 Vif PPLP motif is necessary for human APOBEC3G binding and degradation. *Virology* 377: 49-53

Glaeser RM, Han BG (2017) Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophys Rep* 3: 1-7

Guo Y, Dong L, Qiu X, Wang Y, Zhang B, Liu H, Yu Y, Zang Y, Yang M, Huang Z (2014) Structural basis for hijacking CBF-beta and CUL5 E3 ligase complex by HIV-1 Vif. *Nature* 505: 229-33

Hirosawa M, Hoshida M, Ishikawa M, Toya T (1993) MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. *Comput Appl Biosci* 9: 161-7

Jager S, Kim DY, Hultquist JF, Shindo K, LaRue RS, Kwon E, Li M, Anderson BD, Yen L, Stanley D, Mahon C, Kane J, Franks-Skiba K, Cimermancic P, Burlingame A, Sali A, Craik CS, Harris RS, Gross JD, Krogan NJ (2011) Vif hijacks CBF-beta to degrade APOBEC3G and promote HIV-1 infection. *Nature* 481: 371-5

Kane JR, Stanley DJ, Hultquist JF, Johnson JR, Mietrach N, Binning JM, Jonsson SR, Barelier S, Newton BW, Johnson TL, Franks-Skiba KE, Li M, Brown WL, Gunnarsson HI, Adalbjornsdottir A, Fraser JS, Harris RS, Andresdottir V, Gross JD, Krogan NJ (2015) Lineage-Specific Viral Hijacking of Non-canonical E3 Ubiquitin Ligase Cofactors in the Evolution of Vif Anti-APOBEC3 Activity. *Cell Rep* 11: 1236-50

Khan MA, Aberham C, Kao S, Akari H, Gorelick R, Bour S, Strebel K (2001) Human immunodeficiency virus type 1 Vif protein is packaged into the nucleoprotein complex through an interaction with viral genomic RNA. *J Virol* 75: 7252-65

Miller JH, Presnyak V, Smith HC (2007) The dimerization domain of HIV-1 viral infectivity factor Vif is required to block virion incorporation of APOBEC3G. *Retrovirology* 4: 81

Scheres SHW (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology* 180: 519-530

Yang B, Gao L, Li L, Lu Z, Fan X, Patel CA, Pomerantz RJ, DuBois GC, Zhang H (2003) Potent Suppression of Viral Infectivity by the Peptides That Inhibit Multimerization of Human Immunodeficiency Virus Type 1 (HIV-1) Vif Proteins. *Journal of Biological Chemistry* 278: 6596-6602

Yang S, Sun Y, Zhang H (2001) The multimerization of human immunodeficiency virus type I Vif protein: a requirement for Vif function in the viral life cycle. *J Biol Chem* 276: 4889-93

Zhang H, Pomerantz RJ, Dornadula G, Sun Y (2000) Human immunodeficiency virus type 1 Vif protein is an integral component of an mRNP complex of viral RNA and could be involved in the viral RNA folding and packaging process. *J Virol* 74: 8252-61

Zhang W, Wang H, Li Z, Liu X, Liu G, Harris RS, Yu XF (2014) Cellular requirements for bovine immunodeficiency virus Vif-mediated inactivation of bovine APOBEC3 proteins. *J Virol* 88: 12528-40

5 Appendix I: Modulating SAMHD1 activity to improve nucleotide analog therapies

5.1 Preface

Parts of Section 5.3.1 have been published in *Nature Communications* (Oellerich et al. 2019). K.M.K. and O.B. designed experiments and analyzed data. K.M.K. generated malachite-based activity assay data shown in Figures 5-2B, 5-3D, and 5-4D. K.M.K. and O.B. contributed to the structural analysis shown in Figure 5-2G and the high-throughput screen shown in Figures 5-4A and 5-4B.

5.2 Introduction

Leukemia, lymphoma, and myeloma are diagnosed in all ages and accounts for about 9.5% of deaths in 2010 in the US (Chhikara and Parang 2010). Acute myeloid leukemia (AML) is identified by the rapid increase of immature white blood cells that crowd and accumulate around the bone marrow (Hasserjian 2013). Currently the most popular treatment is a combination of chemotherapy and infusion of the drug cytarabine (Tamamyan et al. 2017). Cytarabine was isolated from a Caribbean sponge in 1959 and since then has been widely used for treatment of AML. Despite the wide use, it is not a very potent and effective treatment; it only cures about 30% of adult patients (Jordheim et al. 2013). To improve the effectiveness of chemotherapy for the treatment of AML, it is essential to understand why cytarabine does not eradicate most of the leukemic stem cells.

Cytarabine represents the first FDA-approved nucleoside analog drug, but this class of compounds has since grown to include many other related molecules that are utilized to treat multiple types of cancers and viral infections (Jordheim et al. 2013). These drugs are antimetabolites that interfere with intracellular enzymes, like DNA polymerases. When a drug, such as cytarabine, is incorporated into newly synthesized DNA, it can result in chain terminations or accumulation of mutations that activate DNA damage checkpoints and arrest the cell cycle (Ewald et al. 2008).

Knowledge of nucleoside analog metabolism is important to predict patient outcomes during chemotherapy. As a regulator of the cellular dNTP pool, SAMHD1 may also play an important role in nucleoside analog metabolism and efficacy. SAMHD1 is a homotetrameric enzyme that hydrolyzes cellular dNTPs into nucleosides and triphosphates. The activity of SAMHD1 is regulated by nucleotide binding at two allosteric sites (allo-site), GTP or dGTP at Allo-site 1 and any of the four cellular dNTPs at Allo-site 2 (Ji et al. 2014). Because of SAMHD1's role in nucleotide metabolism, it is important to understand how it interacts with nucleoside analogs. It was previously shown that SAMHD1 increases the efficacy of AZT, a nucleoside inhibitor for HIV reverse transcriptase, by depleting natural dNTPs that compete with AZT binding to reverse transcriptase (Amie et al. 2013).

Strikingly, the triphosphohydrolase activity of SAMHD1 reduces cytarabine cytotoxicity in patient leukemic blasts (Schneider et al. 2107; Herold et al. 2017). Patients with low SAMHD1 expression levels, determined by immunohistochemistry, were more likely to achieve complete remission (CR) with cytarabine treatment than patients with high SAMHD1 expression levels (Figure 5-1A). Furthermore, the event free survival rate

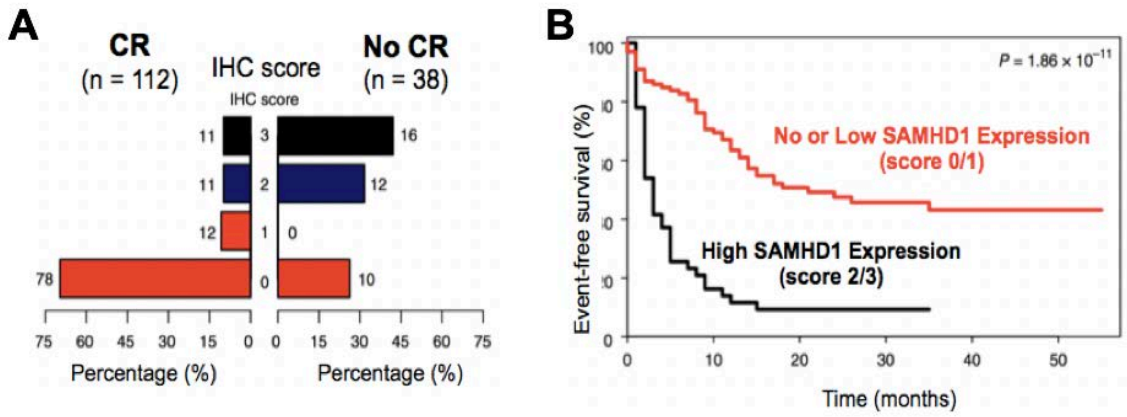

Figure 5-1. SAMHD1 interferes with cytarabine efficacy in AML patients

(A) Comparison of SAMHD1 expression levels in patients treated with cytarabine who attained complete remission (CR) and no CR. Adapted from (Schneider et al. 2017). (B) Kaplan-Meier analyses for event-free survival of patients with AML treated with cytarabine and segmented by SAMHD1 expression. Adapted from (Schneider et al. 2017).

of patients with high SAMHD1 expression was significantly lower than patients with no or low SAMHD1 expression (Figure 5-1B). These data strongly suggest that SAMHD1 activity is affecting cytarabine treatment, likely by hydrolyzing the drug and decreasing its concentration in the cell. This also raises the question of whether other nucleoside analogs used to treat AML are substrates of SAMHD1 and whether their potency is dependent on SAMHD1 triphosphohydrolase activity.

Our overall goal is to improve the effectiveness of the current AML treatments by modulating the activity of SAMHD1. In this chapter, I describe our investigation of SAMHD1 interactions with two more nucleotide analogs used to treat AML. Then, I report the results of a high-throughput screen used to discover find inhibitors of SAMHD1. Information obtained by these studies might help to improve the prediction of patient responses to nucleoside analog drugs and eventually lead to SAMHD1 inhibitors or enhancers that can be combined with existing therapies.

5.3 Results

5.3.1 Characterization of azacytidine-TP and decitabine-TP with the catalytic and allosteric sites of SAMHD1

In addition to the studies performed in Chapter 2, we have also investigated the interactions between SAMHD1 and the hypomethylating agents, decitabine triphosphate (DAC-TP), and azacytidine triphosphate (AZA-TP) (Oellerich et al. 2019). Both DAC and AZA are used interchangeably as a frontline therapy for older acute myeloid leukemia (AML) patients, but mechanisms underlying varied patient responses are unknown (Tawfik et al. 2014; Stahl et al. 2018). A recent study suggested that SAMHD1 protects cancer cells from DAC (Herold et al. 2017), so we set out to investigate the biochemical and structural nature of SAMHD1 interactions with these DAC-TP and AZA-TP.

Structurally, DAC-TP and AZA-TP are very similar (Figure 5-2A). Both nucleotide analogs contain the same modified nucleobase that closely resembles cytosine, except for the carbon at position 5 of the base that is replaced with a nitrogen atom. The only difference between these two nucleotide analogs is that DAC-TP contains a 2'-deoxyribose sugar, while AZA-TP possesses a ribose sugar with a 2'-hydroxyl group. Neither AZA-TP nor DAC-TP were hydrolyzed by SAMHD1 in the absence of additional activators in an *in vitro* enzymatic assay, suggesting that they

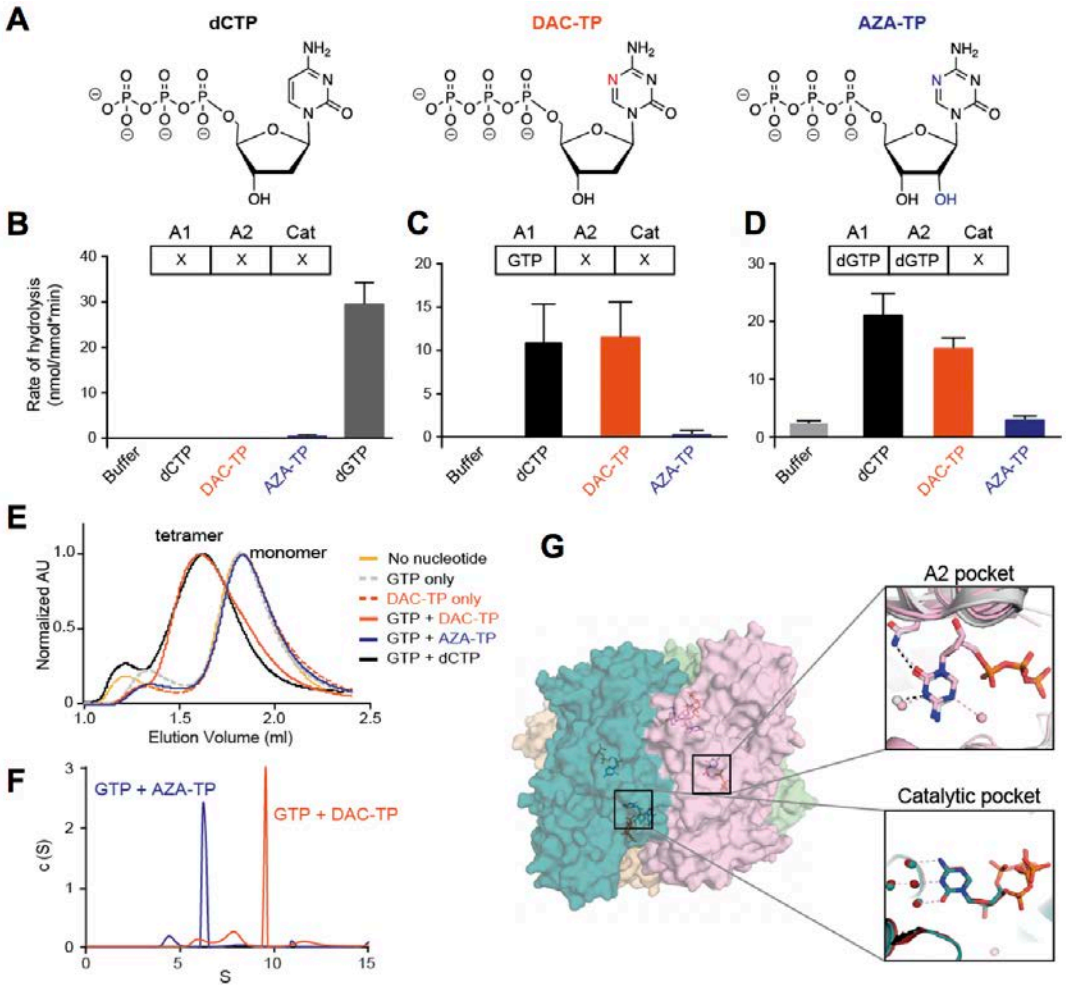

Figure 5-2. Interactions of SAMHD1 with hypomethylating agents, DAC-TP and AZA-TP.

(A) Chemical structures of dCTP, decitabine triphosphate (DAC-TP), and azacytidine triphosphate (AZA-TP). (B-D) dNTPase activity assay for recombinant SAMHD1. The rate of hydrolysis for the indicated nucleotide triphosphates was quantified in the presence of (B) buffer alone, (C) GTP, which can occupy allosteric site 1 (A1), or (D) dGTP, which can occupy A1 and allosteric site 2 (A2). Cat: Catalytic site. (E) Size-exclusion chromatograms of SAMHD1 in absence of nucleotide (yellow line) or in the presence of GTP alone (gray dashed line), DAC-TP alone (orange dashed line), GTP + DAC-TP (orange line), GTP + AZA-TP (blue line), or GTP + dCTP (black line). (F) Sedimentation velocity of SAMHD1 incubated in the presence of GTP in combination with either AZA-TP or DAC-TP. (G) Structure of the SAMHD1 tetramer in complex with DAC-TP. Each SAMHD1 subunit is shown as a surface representation, colored in light pink, deep teal, light green and wheat. In the insets, nucleotides are shown as sticks. (Top, right panel) An overlay of SAMHD1 structures in complex with dCTP (PDB ID: 4TNP; gray) or DAC-TP (light pink) in allosteric site 2. (Bottom, right panel) An overlay of SAMHD1 structures in complex with dCTP (PDB ID: 4TNP; gray) or DAC-TP (deep teal) in the catalytic pocket. This figure was adapted from (Oellerich et al. 2019).

cannot bind both allosteric sites to activate the enzyme (Figure 5-2B). However, when GTP was added to the mixture, DAC-TP (but not AZA-TP) was hydrolyzed by SAMHD1 just as efficiently as canonical dCTP (Figure 5-2C). As GTP can only bind the A1 site (Ji et al. 2014), this implies that DAC-TP can occupy the A2 site to activate SAMHD1-mediated hydrolysis. In an established SAMHD1 pre-assembly assay (Ji et al. 2014), SAMHD1 tetramers were pre-assembled in the presence of dGTP and subsequently diluted into either reaction buffer alone or buffer containing dCTP, DAC-TP, or AZA-TP (Figure 5-2D). dCTP and DAC-TP, but not AZA-TP, were hydrolyzed, confirming that DAC-TP, but not AZA-TP, is a substrate of SAMHD1 enzyme. This is consistent with our previous observation that modest modifications to the Watson-Crick edge of a nucleobase do not perturb binding to the catalytic pocket of SAMHD1 (Knecht and Buzovetsky et al. 2018).

To confirm our finding that DAC-TP can promote SAMHD1 tetramerization by binding to the A2 site, we monitored the oligomerization state of SAMHD1 in the presence of the A1-site activator GTP following addition of either DAC-TP or AZA-TP. Size-exclusion chromatography revealed that DAC-TP, but not AZA-TP, induced the formation of higher-order SAMHD1 oligomers (Figure 5-2E). In line with these results, the sedimentation velocity measured by an analytical ultracentrifugation assay confirmed that DAC-TP, but not AZA-TP, induces the formation of SAMHD1 tetramers (Figure 5-2F).

To elucidate how DAC-TP binds the catalytic pocket and A2 site of SAMHD1, DAC-TP was co-crystallized with the inactivated catalytic domain of SAMHD1 (residues

Table 5-1. Data collection and refinement statistics for the crystal structures of SAMHD1 HD bound to DAC-TP

Data Collection	
Wavelength (Å)	0.97920
Space Group	P 1 21 1
Cell dimensions	
a, b, c (Å)	87.5, 146.5, 98.6
α, β, γ (°)	90.0, 114.6, 90.0
Molecules/asymmetric unit	4
Resolution (Å)	50-2.14 (2.18-2.14)
Unique reflections	125603 (6000)
R_{merge}	0.132 (>1)
I/σ	9.3 (0.88)
Completeness (%)	98.9 (95.0)
Redundancy	4.1 (3.3)
CC½	0.993 (0.175)
Refinement	
Number of nonhydrogen atoms	16496
R_{work}/R_{free} (%)	17.4/22.5 (33.7/34.7)
Average B factor	41.73
RMSD	
Bond lengths	0.016
Bond angles	1.86
Ramachandran Analysis	
Preferred regions (%)	97.1
Allowed regions (%)	2.8
Outliers (%)	0.1

Statistics in parentheses indicate those for highest resolution shell.

113-626 with H206R and D207N mutations), which has been extensively used to study nucleotide binding to SAMHD1 (Ji et al. 2013; Ji et al. 2014). The crystal structure of the complex, determined at a 2 Å resolution (Table 5-1), demonstrated that DAC-TP molecules can occupy both the A2 pocket and catalytic site of the SAMHD1 tetramer (Figure 5-2G). Comparison of the DAC-TP/SAMHD1 structure with a previously

determined dCTP/SAMHD1 structure (PDB ID 4TNP; Ji et al. 2014) showed that the two structures align very well with an overall root mean square deviation of ~0.4 Å. The hydrogen bonding interactions between the SAMHD1 catalytic pocket and DAC-TP are the same as those involved in dCTP binding. Although an additional water molecule forms a hydrogen bond with the 5 nitrogen moiety that is only present in the base of DAC-TP, this extra water molecule, interestingly, did not engage in any interaction with SAMHD1. Moreover, SAMHD1's catalytic pocket was not altered by the interaction with DAC-TP, suggesting that DAC-TP adopts a binding mode similar to that of canonical dCTP (Figure 5-2G, lower inset). In line with our biochemical analyses, AZA-TP did not co-crystallize with SAMHD1. Taken together, DAC-TP is an allosteric activator and substrate of SAMHD1, while AZA-TP does not interact with SAMHD1.

Our collaborators went on to confirm that SAMHD1 triphosphohydrolase activity does indeed have an effect on DAC-TP, but not AZA-TP, levels and cytotoxicity *in vivo* (Oellerich et al. 2019). This indicates substantial differences in the mode of action of these HMAs that are often regarded to be interchangeable in AML therapy. More research is needed to identify biomarkers that reliably guide the decision on whether patients should be treated with DAC or AZA. However, SAMHD1 has the potential to become a relevant biomarker for the stratification of HMAs in AML patients and a therapeutic target for the improvement of DAC-based therapies. Combined with the results of Chapter 2, these findings contribute to our overall understanding of how SAMHD1 can affect the metabolism of nucleotide analog drugs.

5.3.2 High-throughput screening for SAMHD1 inhibitors

As a regulator of the nucleotide pool in the cell, modulation of SAMHD1 could potentially improve patient response to nucleotide analog drugs. In cases where SAMHD1 is capable of degrading a nucleotide analog, SAMHD1 activity will decrease the concentrations of nucleotide analog in the cell (Figure 5-3A). For example, SAMHD1 expression in AML patients was correlated with worse rates of remission when treated with cytarabine (Schneider et al. 2017). The 2' arabinose sugar of cytarabine allows for entry into the catalytic pocket of SAMHD1, making it a good substrate of SAMHD1.

However, SAMHD1 activity also has the potential to increase the efficacy of certain nucleotide analog drugs (Figure 5-3B). If the nucleotide analog is not a substrate of SAMHD1, then SAMHD1 activity will leave the concentration of the nucleotide analog high relative to depleted dNTP pool. Since nucleotide analogs compete with cellular dNTPs for active sites of polymerases, SAMHD1 activity has the potential to enhance the efficacy of nucleotide analog drugs. For example, reducing SAMHD1 levels via Vpx-mediated degradation resulted in lower dNTP pool concentrations and increased sensitivity to the HIV-1 nucleoside reverse transcriptase inhibitor zidovudine (Amie et al. 2013). The 3' cyano group of zidovudine is not accommodated by the catalytic pocket of SAMHD1, making it a poor substrate of SAMHD1.

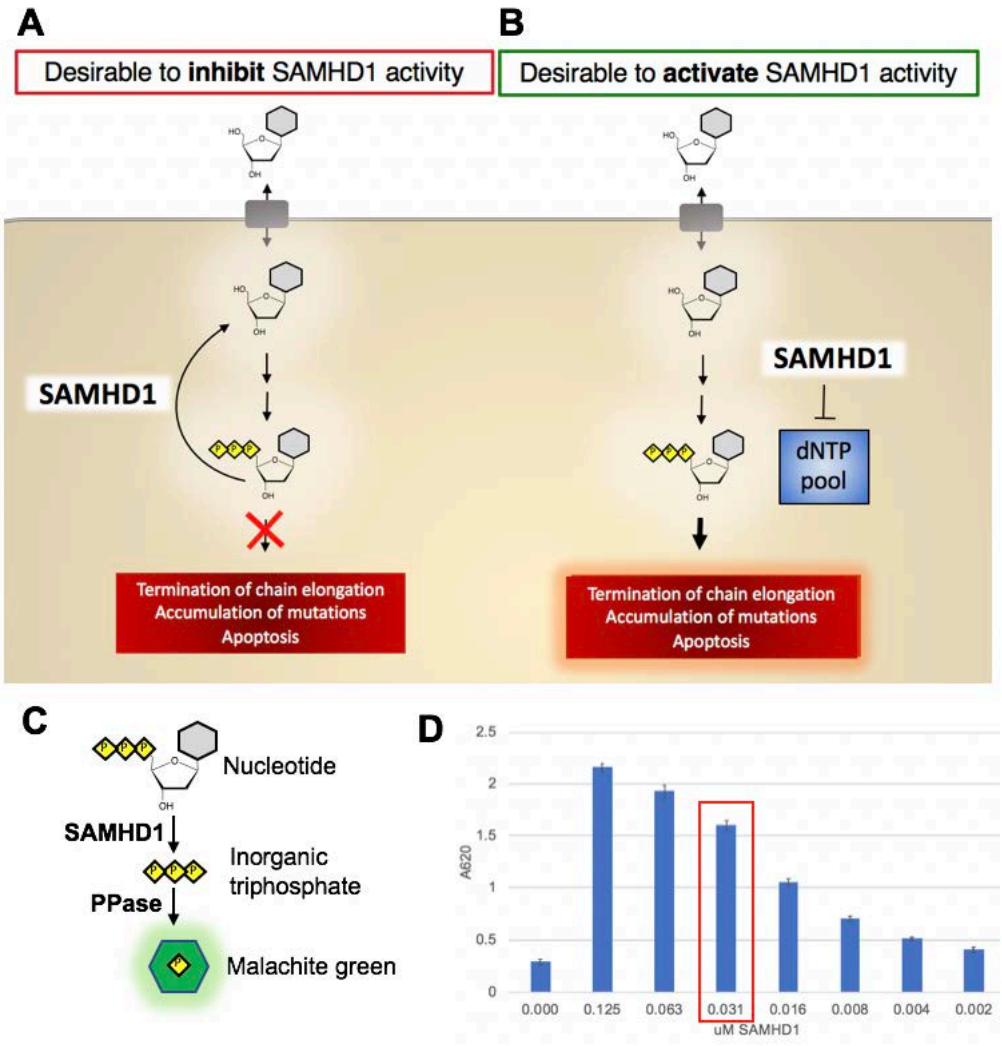

Figure 5-3. Optimizing SAMHD1 activity assay for high-throughput screening

(A) Cartoon schematic showing a situation where SAMHD1 activity is harmful to nucleotide analog therapy. (B) Cartoon schematic showing a situation where SAMHD1 activity is helpful to nucleotide analog therapy. (C) Cartoon schematic showing the steps of the malachite green colorimetric assay for SAMHD1 activity: (1) SAMHD1 cleaves the inorganic triphosphate from dNTPs, (2) inorganic pyrophosphatase (PPase) hydrolyzes triphosphate into inorganic phosphates, and (3) malachite green changes color when bound to inorganic phosphate.

Since SAMHD1 has emerged as an important therapeutic target, we sought to find small molecules that could either activate or inhibit SAMHD1 activity. To accomplish this, we optimized the colorimetric malachite green assay for a high-throughput screen for modulators of SAMHD1 activity (Figure 5-3C; Seamon and Stivers 2015). In this assay, a secondary enzyme (PPase) is used to cleave inorganic phosphates from the triphosphate produced by SAMHD1 catalysis. Many assay parameters were optimized including reaction time, malachite development time, dNTP substrate concentration, and SAMHD1 concentration. We sought to find conditions that would yield the best signal to noise ratio and most reproducible results using a 384-well plate and high-throughput pipetting machines. Figure 5-2D shows a representative dilution series where all variables were kept constant except SAMHD1 concentration. While there is about 0.3 absorbance units (AU) of background noise in the absence of SAMHD1, adequate signal above 1 AU is observed for SAMHD1 concentrations above 15 nM. Ultimately, we chose the 30 nM SAMHD1 concentration for this assay because at this concentration the signal to noise is above 3, and only about 70% of the maximum amount of product had been created. This means that we can potentially capture an increase or decrease in SAMHD1 activity recorded in this assay.

After the optimal assay conditions were determined, we screened 5,763 compounds from libraries in the Yale Center for Molecular Discovery for modulation of SAMHD1 activity. In order to increase the likelihood that a potential hit would be bioactive with favorable pharmacological properties, we chose libraries enriched for FDA-approved drugs. Another consideration we had while choosing the small molecule libraries to screen was the three-dimensional architecture of the catalytic and allosteric

pockets of SAMHD1. As much as possible, we tried to find libraries with structural features that resemble nucleotides. However, none of the molecule libraries included compounds with triphosphate groups due to their highly labile nature.

The assay performance was assessed by comparing control reactions to reactions with test compounds. The test compounds were spread across 20 different 384-well plates (Figure 5-4A). Negative controls inhibition containing SAMHD1 but no test compound were normalized and set to 0% effect (black arrows in Figure 5-4A). Positive controls for inhibition containing reactants except for SAMHD1 were set to 100% effect (red arrows in Figure 5-4A). While most compounds imposed little to no effect on SAMHD1 activity with the 3 sigma cutoff being about 10.2% effect, some compounds did seem to produce a significant inhibitory effect on SAMHD1 (examples circled in red in Figure 5-4A). None of the compounds significantly increased SAMHD1 activity. Overall, 326 compounds had greater than a 10.2 % inhibitory effect (5.6% hit rate) and 5 compounds had greater than a 40% inhibitory effect (0.1% hit rate). The assay itself performed well, with a steady signal to background ratio and z' score across the 20 plates of reactions (Figure 5-4B).

Interestingly, nucleotide analogs were not common in the hit list. The highest-ranking nucleotide was AMP which showed 17% effect. This may be due to the fact that none of the nucleotide analogs tested here were triphosphorylated, which is likely important for binding to the catalytic pocket. Most compounds with high inhibitory effects were compounds that are known commonly found to cause artificial interference in assays (Gilbert et al. 2016). After removing compounds with known aggregating or metal chelating properties known to interfere in assays such as this, we chose to follow

up on three hits from the screen: benzobromane, octinoxate, and clonidine (Figure 5-4C). These compounds all share aromatic rings, but are otherwise structurally diverse. Each compound ranged in 20-30% effect on SAMHD1 activity during the high-throughput screen. However, when a similar activity assay was repeated, neither octinozate nor clonidine affected SAMHD1 activity significantly more than the DMSO only control (Figure 5-4D). Benzobromane did decrease activity by approximately 30% in this second assay, but more validation will be needed in order to confirm that it is directly inhibiting SAMHD1.

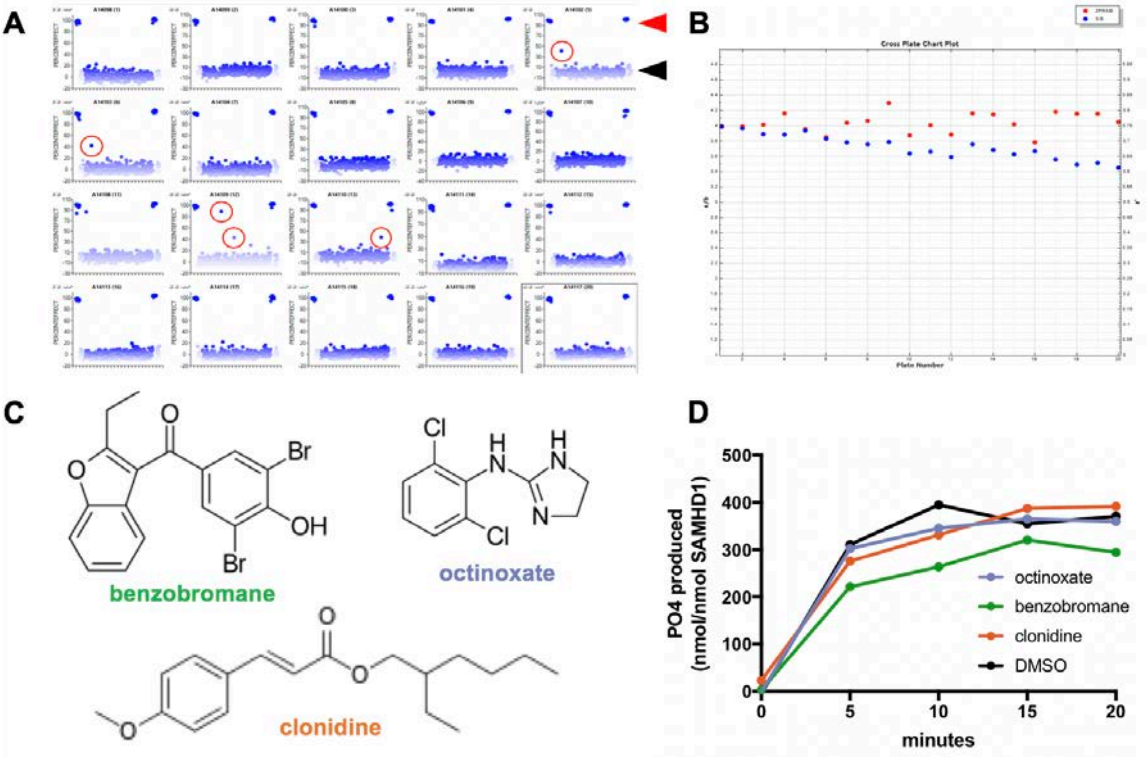

Figure 5-4. Results of small molecule screen for SAMHD1 modulators

(A) Scatterplot of inhibitory effect of small molecules on SAMHD1 activity across 20 plates in high-throughput screen. Negative control with no compound added is indicated with black arrow and positive control with no SAMHD1 added is indicated with red arrow. Top inhibitory hits are highlighted with red circles. (B) Signal to background ratios and z' scores for each plate. (C) Three representative hits from the screen. (D) Validation assay for inhibition of SAMHD1 activity.

5.4 Future Directions

With importance in the metabolism of both cellular nucleotide pools and nucleotide analogs, SAMHD1 is an attractive therapeutic target. Since SAMHD1 expression and activities vary depending on the patient, a potential personalized medicine strategy could be to screen patients for SAMHD1 mutations and expression levels before the administration of nucleotide analog drugs. More clinical research should be performed to stratify the response of patients to nucleotide drugs by SAMHD1 expression. By understanding the clinical relationship between different nucleotide analogs and SAMHD1 expression, physicians may be able to prescribe the most effective drug for a particular patient based on their personal SAMHD1 expression levels. For instance, if an AML patient is known to have higher SAMHD1 expression levels, the physician might be advised to administer AZA over DAC, instead of choosing one arbitrarily to start.

Depending on the nucleotide analog, it might be useful to either inhibit or enhance SAMHD1 activity. In the future, more avenues of modulating SAMHD1 activity should be explored for therapeutic use. For instance, small molecule inhibitors could potentially bind the catalytic pocket and competitively inhibit SAMHD1 catalysis. However, another mechanism of inhibition could include targeting the allosteric pocket required for assembling the active tetramer form of the protein. It is possible that a small molecule could directly bind to either of the allosteric sites or bind in an entirely different area that might block the dimer or tetramer interfaces. Furthermore, it might be possible to target

SAMHD1 expression through CRISPR-based or RNAi-based techniques. However, the off-target effects of SAMHD1 down regulation must be explored.

5.5 Experimental Procedures

5.5.1 Protein expression and purification

N-terminal 6×His-tagged SAMHD1 constructs were expressed in *Escherichia coli* BL21(DE3). Cells were harvested, resuspended in lysis buffer (50 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole, 5 mM $MgCl_2$, 0.5 mM TCEP), and lysed using a microfluidizer. Lysate was clarified by centrifugation (26,892×g for 25 minutes). SAMHD1 protein was purified using Ni-NTA affinity and size-exclusion chromatography in SAMHD1 Buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, and 0.5 mM TCEP).

5.5.2 Analytical size exclusion chromatography

Purified samples of SAMHD1 (2 mg per mL, 50 μL) mixed with a final concentration of 500 μM GTP and 2-4 mM nucleotide analog were applied to a Superdex 200 5/150 GL column (GE Healthcare) pre-equilibrated with SAMHD1 Buffer. The UV absorbance at 280 nm was measured as the protein sample eluted from the column, and values were normalized to their respective peak heights.

5.5.3 Analytical ultracentrifugation

Sedimentation velocity experiments were performed with a Beckman XL-I analytical ultracentrifuge. Samples were prepared with protein concentration of 0.8-1.3 mg/ml in SAMHD1 Buffer and equilibrated with a final concentration of 150 μM

nucleotides. AUC was performed at 169,167×g and 20°C with an An60-Ti rotor. The experimental parameters including sample partial specific volume, buffer density and viscosity were calculated with SEDNTERP (http://sednterp.unh.edu/). Velocity data were analyzed with the program SEDFIT (Brown and Schuck 2006).

5.5.4 Crystallization and data collection

SAMHD1 protein in buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 5 mM $MgCl_2$ and 0.5 mM TCEP) was mixed with 1 mM GTP and 10 mM nucleotide analogs and incubated at 4 °C for 15 minutes before crystallization. All crystals were grown at 25 °C using the microbatch under-oil method by mixing 1 μL protein (5 mg per mL) with 1 μL crystallization buffer (100 mM SPG (Qiagen) buffer, pH 7.4, 25% PEG 1500). Crystals were cryoprotected by crystallization buffer supplemented with 25% (Vol/Vol) glycerol before being frozen in liquid nitrogen. Diffraction data were collected at the Advanced Photon Source beamline 24-IDC. The data statistics are summarized in Table 5-1.

5.5.5 Structure determination and refinement

The structures were solved by molecular replacement using PHASER (McCoy et al. 2007). The previously published SAMHD1 tetramer structure, with all bound nucleotides removed was used as the search model (PDB ID 4BZB). The model was refined with iterative rounds of TLS and restrained refinement using *Refmac*5 (Vagin et al. 2004), followed by rebuilding the model to the $2F_o$-F_c and the F_o-F_c maps using Coot (Emsley et al. 2004). Refinement statistics are summarized in Table 5-1. Coordinates and

structure factors have been deposited in the Protein Data Bank (PDB accession number: 6CM2).

5.5.6 Malachite assay

The enzymatic activity assay was modified from Seamon and Stivers (2015). All assays were performed with purified wild-type SAMHD1 (residues 113–626) at 25°C in a reaction buffer containing 50 mM Tris-HCl pH 8, 150 mM NaCl, 5 mM $MgCl_2$, and 0.5 mM TCEP. Each 40 μL reaction, containing 10 μM pyrophosphatase, 0.5 μM SAMHD1, and 125 μM substrate or allosteric activator was quenched with 40 μL 20 mM EDTA after 15 minutes. Then, 20 μL Malachite Green reagent was added to the solution and developed for 15 minutes before the absorbance at 650 nm was measured.

5.5.7 High-throughput screen for SAMHD1 inhibitors

Peter Gareiss assisted with assay design and implementation at Yale's Center for Molecular Discovery. The malachite assay described above was modified for high the high-throughput assay format. Similar buffers and protein were prepared as above. A library of 5763 compounds were added to Corning 3702 clear 384-well plates. Each 20 μL reaction, containing 10 μM pyrophosphatase, 0.03 μM SAMHD1, and 125 μM dGTP activator, was quenched with 20 μL 20 mM EDTA after 60 minutes. Then, 20 μL Malachite Green reagent was added to the solution and developed for 60 minutes before the absorbance at 620 nm was measured using and Envision Self Service plate reader.

5.6 References

Amie SM, Daly MB, Noble E, Schinazi RF, Bambara RA, Kim B (2013) Anti-HIV host factor SAMHD1 regulates viral sensitivity to nucleoside reverse transcriptase inhibitors via

modulation of cellular deoxyribonucleoside triphosphate (dNTP) levels. *J Biol Chem* 288: 20683-91

Bohl SR, Bullinger L, Rucker FG (2018) Epigenetic therapy: azacytidine and decitabine in acute myeloid leukemia. *Expert Rev Hematol* 11: 361-371

Brown PH, Schuck P (2006) Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation. *Biophys J* 90: 4651-61

Chhikara BS, Parang K (2010) Development of cytarabine prodrugs and delivery systems for leukemia treatment. *Expert Opin Drug Deliv* 7: 1399-414

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126-32

Ewald B, Sampath D, Plunkett W (2008) Nucleoside analogs: molecular mechanisms signaling cell death. *Oncogene* 27: 6522-37

Gilberg E, Jasial S, Stumpfe D, Dimova D, Bajorath J (2016) Highly Promiscuous Small Molecules from Biological Screening Assays Include Many Pan-Assay Interference Compounds but Also Candidates for Polypharmacology. *Journal of Medicinal Chemistry* 59: 10285-10290

Hasserjian RP (2013) Acute myeloid leukemia: advances in diagnosis and classification. *Int J Lab Hematol* 35: 358-66

Herold N, Rudd SG, Ljungblad L, Sanjiv K, Myrberg IH, Paulin CB, Heshmati Y, Hagenkort A, Kutzner J, Page BD, Calderon-Montano JM, Loseva O, Jemth AS, Bulli L, Axelsson H, Tesi B, Valerie NC, Hoglund A, Bladh J, Wiita E et al. (2017) Targeting SAMHD1 with the Vpx protein to improve cytarabine therapy for hematological malignancies. *Nat Med* 23: 256-263

Ji X, Tang C, Zhao Q, Wang W, Xiong Y (2014) Structural basis of cellular dNTP regulation by SAMHD1. *Proc Natl Acad Sci U S A* 111: E4305-14

Ji X, Wu Y, Yan J, Mehrens J, Yang H, DeLucia M, Hao C, Gronenborn AM, Skowronski J, Ahn J, Xiong Y (2013) Mechanism of allosteric activation of SAMHD1 by dGTP. *Nat Struct Mol Biol* 20: 1304-9

Jordheim LP, Durantel D, Zoulim F, Dumontet C (2013) Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. *Nat Rev Drug Discov* 12: 447-64

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658-674

Oellerich T, Schneider C, Thomas D, Knecht KM, Buzovetsky O, Kaderali L, Schliemann C, Bohnenberger H, Angenendt L, Hartmann W, Wardelmann E, Rothenburger T, Mohr S, Scheich S, Comoglio F, Wilke A, Strobel P, Serve H, Michaelis M, Ferreiros N et al. (2019) Selective inactivation of hypomethylating agents by SAMHD1 provides a rationale for therapeutic stratification in AML. *Nat Commun* 10: 3475

Schneider C, Oellerich T, Baldauf HM, Schwarz SM, Thomas D, Flick R, Bohnenberger H, Kaderali L, Stegmann L, Cremer A, Martin M, Lohmeyer J, Michaelis M, Hornung V, Schliemann C, Berdel WE, Hartmann W, Wardelmann E, Comoglio F, Hansmann ML et al. (2017) SAMHD1 is a biomarker for cytarabine response and a therapeutic target in acute myeloid leukemia. *Nat Med* 23: 250-255

Seamon KJ, Stivers JT (2015) A High-Throughput Enzyme-Coupled Assay for SAMHD1 dNTPase. *J Biomol Screen* 20: 801-9

Stahl M, DeVeaux M, Montesinos P, Itzykson R, Ritchie EK, Sekeres MA, Barnard JD, Podoltsev NA, Brunner AM, Komrokji RS, Bhatt VR, Al-Kali A, Cluzeau T, Santini V, Fathi AT, Roboz GJ, Fenaux P, Litzow MR, Perreault S, Kim TK et al. (2018) Hypomethylating agents in relapsed and refractory AML: outcomes and their predictors in a large international patient cohort. *Blood Adv* 2: 923-932

Tamamyan G, Kadia T, Ravandi F, Borthakur G, Cortes J, Jabbour E, Daver N, Ohanian M, Kantarjian H, Konopleva M (2017) Frontline treatment of acute myeloid leukemia in adults. *Crit Rev Oncol Hematol* 110: 20-34

Tang C, Ji X, Wu L, Xiong Y (2015) Impaired dNTPase activity of SAMHD1 by phosphomimetic mutation of Thr-592. *J Biol Chem* 290: 26352-9

Tawfik B, Sliesoraitis S, Lyerly S, Klepin HD, Lawrence J, Isom S, Ellis LR, Manuel M, Dralle S, Berenzon D, Powell BL, Pardee T (2014) Efficacy of the hypomethylating agents as frontline, salvage, or consolidation therapy in adults with acute myeloid leukemia (AML). *Ann Hematol* 93: 47-55

Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, Murshudov GN (2004) REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. *Acta Crystallogr D Biol Crystallogr* 60: 2184-95

6 Appendix II: List of Publications

*Indicates equal author contributions

Oellerich T*, Schneider C*, Thomas D, **Knecht KM**, Buzovetsky O, Kaderali L, Schliemann C, Bohnenberger H, Angenendt L, Hartmann W, Wardelmann E, Rothenburger T, Mohr S, Scheich S, Comoglio F, Wilke, A, Stroebel P, Serve H, Ferreirós N, Geisslinger G, Xiong Y, Keppler OT, Cinatl J. 2019. Selective inactivation of hypomethylating agents by SAMHD1 provides a rationale for therapeutic stratification in AML. *Nat Commun.* **10**: 3475.

Weber E*, Buzovetsky O*, Heston L*, Yu KP, **Knecht KM**, El-Guindy A, Miller G, Xiong Y. A Noncanonical Basic Motif of Epstein-Barr Virus ZEBRA Protein Facilitates Recognition of Methylated DNA, High-Affinity DNA Binding, and Lytic Activation. *J Virol.* **93**: e00724-19.

Atilho RM, Arachchilage GM, Greenlee EB, **Knecht KM**, Breaker RR. 2019. A bacterial riboswitch class for the thiamin precursor HMP-PP employs a terminator-embedded aptamer. 2019. *Elife* **8**: e45210.

Oi C, Treado JD, Levine ZA, Lim CS, **Knecht KM**, Xiong Y, O'Hern CS, Regan L. 2018. A threonine zipper that mediates protein-protein interactions: Structure and prediction. *Protein Sci*. **27**: 1969-1977.

Knecht KM*, Buzovetsky O*, Schneider C, Reiss K, Keppler OT, Batista VS, Ji X, Xiong Y. 2018. The structural basis for cancer drug interactions with the catalytic and allosteric sites of SAMHD1. *Proc Natl Acad Sci USA*. **115**: E10022-E10031.

Chen S, Bonifati S, Qin Z, St. Gelais C, Kodigepalli KM, Barrett BS, Kim SH, Antonucci JM, Ladnerd KJ, Buzovetsky O, **Knecht KM**, Xiong Y, Yount JS, Guttridge DC, Santiago ML, Wu L. 2018. SAMHD1 Suppresses Innate Immune Responses to Viral Infections and Inflammatory Stimuli by Inhibiting the NF-kB and Interferon Pathways. *Proc Natl Acad Sci U S A*. **115**: E3798-E3807.

Antonucci JM, Kim SH, St. Gelais C, Bonifati S, Buzovetsky O, **Knecht KM**, Duchon AA, Xiong Y, Musier-Forsyth K, Wu L. 2018. SAMHD1 Impairs HIV-1 Gene Expression and Reactivation of Viral Latency in CD4+ T-Cells. *J Virol*. **92:** e00292-18.

St. Gelais C, Kim SH, Maksimova VV, Buzovetsky O, **Knecht KM**, Shepard C, Kim B, Xiong Y, Wu L. 2018. A cyclin-binding motif in human SAMHD1 is required for its HIV-1 restriction, dNTPase activity, tetramer formation, and efficient phosphorylation. *J Virol*. **92**:e01787-17.

Buzovetsky O*, Tang C*, **Knecht KM**, Antonucci JM, Wu L, Ji X, Xiong Y. 2018. The SAM domain of mouse SAMHD1 is critical for its activation and regulation. *Nat Commun*. **9**:411.