Molecular Characterization And Inhibition Of The Myst Acetyltransferase, Hmof

Cheryl Elizabeth Mccullough

University of Pennsylvania, cherylmccullough31@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Biochemistry Commons, and the Chemistry Commons

Recommended Citation


https://repository.upenn.edu/edissertations/2462

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/2462

For more information, please contact repository@pobox.upenn.edu.
Molecular Characterization And Inhibition Of The Myst Acetyltransferase, Hmof

Abstract

Acetylation is one of several post-translational modifications (PTMs) that occurs on histones to regulate chromatin dynamics and function in DNA transcription, replication, repair, and other DNA-templated activities, and is carried out by histone acetyltransferases (HATs). Of the multiple HATs that have been identified to date, the most studied have been classified into five subfamilies based on sequence homology and substrate acetylation properties. This work is focused on the MYST family (named for founding members MOZ, YBF2/SAS3, SAS2, and Tip60) of acetyltransferases, and more specifically, hMOF (human males absent on the first). Despite being the largest, most diverse known family of HATs, the MYST proteins have not been as well studied as other families. The MYST family has been linked with a variety of diseases including Alzheimer’s disease, diabetes, and cancer, and some efforts have been made to develop inhibitors of these proteins. The existing inhibitors for the MYST family, however, are lacking in either potency, selectivity, or favorable pharmacokinetic properties. The first part of this thesis will describe our efforts to identify potent, selective hMOF inhibitors using a high throughput screening campaign. Although this project ultimately uncovered no hMOF inhibitors, it revealed major hurdles that may be encountered when studying the MYST proteins that will be outlined in detail.

The second part of this thesis will discuss autoacetylation, which is a mechanism of regulation for acetyltransferases. The MYST family is autoacetylated at an active site lysine residue to facilitate substrate lysine binding and acetylation, however the mechanism and regulation of this autoacetylation are not perfectly understood. Here, we will describe a molecular investigation of Lys-274 autoacetylation of hMOF. These studies revealed that substitutions of Lys-274 are able to bind cofactor but are destabilized, and are catalytically inactive for histone H4 peptide lysine acetylation, stemming from a disordering of the residue 274-harboring α2-β7 loop. We also provide evidence that a catalytically inactive C316S/E350Q mutant and a K268M mutant can still undergo K274 autoacetylation. Together, these studies point to the critical and specific role of hMOF Lys-274 autoacetylation in hMOF stability and cognate substrate acetylation and argue that binding of AcCoA to hMOF likely drives Lys-274 autoacetylation for subsequent cognate substrate acetylation.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Chemistry

First Advisor
Ronen Marmorstein

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2462
MOLECULAR CHARACTERIZATION AND INHIBITION OF THE MYST ACETYLTRANSFERASE, hMOF

Cheryl E. McCullough

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

___________________________________

Ronen Marmorstein, Ph.D.

Professor of Biochemistry and Biophysics

Graduate Group Chairperson

___________________________________

Gary Molander, Ph.D., Hirschmann-Makineni Professor of Chemistry

Dissertation Committee

E. James Petersson, Ph.D., Associate Professor of Chemistry, Committee Chair
Barry S. Cooperman, Ph.D., Professor of Chemistry
Scott L. Diamond, Ph.D., Arthur E. Humphrey Professor of Chemical and Biomolecular Engineering
Dedicated to Mom, Dad, and Pam
ACKNOWLEDGMENTS

I feel so fortunate to have been surrounded by so many great individuals over the last six years, who have all in one way or another helped me to achieve my goals. First I would like to express my gratitude to Ronen for being a wonderful mentor. His kindness, quick email responses, and dedication to training have been invaluable to my graduate experience. I must also thank my dissertation committee, Dr. Petersson, Dr. Cooperman, and Dr. Diamond, for always pushing me to learn and achieve more, and for thoughtful suggestions on how to progress my science. Furthermore, I would like to thank my undergraduate PI, Dr. Gericke, for encouraging me to pursue a Ph.D, and my high school chemistry teacher, Mrs. Moon-Kutzavich for instructing a great class, leading me to choose chemistry as a college major.

I also need to acknowledge many fellow labmates. I would like to thank: Hua Yuan for being my initial mentor in the lab and for getting me started working with hMOF, Michael Shin for being an enthusiastic student, Adam Olia for helping me with anything and everything in the lab, and for inadvertently teaching me to be more assertive, John Domsic for first teaching me how to solve a crystal structure, Mike Grasso for catbutttface greetings and for sharing with me his enthusiasm for Survivor, Allison Haigney for being a wonderful safety buddy and an even better friend, and her husband Carl Machutta for helpful conversations and assay schematics written on napkins. I also need to thank Glen Liszczak not only for first teaching me enzyme kinetics, but also for his constant encouragement, for being able to make me laugh even on my very worst days, and for vastly improving my overall quality of life. To everyone else in the Marmorstein lab past and present, thank you for many great years filled with laughter, learning, secret grampa frosts, Halloween costumes, jumping contests, words of the day, snacks, and friendship.
I must also thank my wonderful friends. Thank you, Chelsea, for always being there when I need to talk about anything, for helping me to get acquainted with Philadelphia, and for regular pep talks. Katie, thank you for your emotional and medical guidance, for making regular trips to come visit me, and for endless adventures. To Danielle, thanks for trips to center city and the beach, Christmas parties, and sunscreen in the mail. I must also thank my non-human friends: Mr. Bean for being the friendliest neighborhood cat, and for always being eager to be pet, Thunder and Foote for their love and whiskers, and Deloris and Snoopy for their shenanigans.

Lastly I would like to acknowledge my family for their endless love and support. I would like to thank my mom for spending countless hours on the phone with me, and for sending me envelopes filled with whiskers in the mail. I would to thank my dad for his unconditional faith in me. I would like to thank Pam for her constant reassurances, and for keeping me updated on Thunder’s daily activities.
ABSTRACT

MOLECULAR CHARACTERIZATION AND INHIBITION OF THE MYST ACETYLTRANSFERASE, hMOF

Cheryl E. McCullough
Ronen Marmorstein

Acetylation is one of several post-translational modifications (PTMs) that occurs on histones to regulate chromatin dynamics and function in DNA transcription, replication, repair, and other DNA-templated activities, and is carried out by histone acetyltransferases (HATs). Of the multiple HATs that have been identified to date, the most studied have been classified into five subfamilies based on sequence homology and substrate acetylation properties. This work is focused on the MYST family (named for founding members MOZ, YBF2/SAS3, SAS2, and Tip60) of acetyltransferases, and more specifically, hMOF (human males absent on the first). Despite being the largest, most diverse known family of HATs, the MYST proteins have not been as well studied as other families. The MYST family has been linked with a variety of diseases including Alzheimer's disease, diabetes, and cancer, and some efforts have been made to develop inhibitors of these proteins. The existing inhibitors for the MYST family, however, are lacking in either potency, selectivity, or favorable pharmacokinetic properties. The first part of this thesis will describe our efforts to identify potent, selective hMOF inhibitors using a high throughput screening campaign. Although this project ultimately uncovered no hMOF inhibitors, it revealed major hurdles that may be
encountered when studying the MYST proteins that will be outlined in detail.

The second part of this thesis will discuss autoacetylation, which is a mechanism of regulation for acetyltransferases. The MYST family is autoacetylated at an active site lysine residue to facilitate substrate lysine binding and acetylation, however the mechanism and regulation of this autoacetylation are not perfectly understood. Here, we will describe a molecular investigation of Lys-274 autoacetylation of hMOF. These studies revealed that substitutions of Lys-274 are able to bind cofactor but are destabilized, and are catalytically inactive for histone H4 peptide lysine acetylation, stemming from a disordering of the residue 274-harboring α2-β7 loop. We also provide evidence that a catalytically inactive C316S/E350Q mutant and a K268M mutant can still undergo K274 autoacetylation. Together, these studies point to the critical and specific role of hMOF Lys-274 autoacetylation in hMOF stability and cognate substrate acetylation and argue that binding of AcCoA to hMOF likely drives Lys-274 autoacetylation for subsequent cognate substrate acetylation.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ........................................................................................................... iii

ABSTRACT .............................................................................................................................. v

LIST OF TABLES .................................................................................................................... xi

LIST OF FIGURES .................................................................................................................. xii

CHAPTER 1: *Introduction to the Acetyltransferases* ............................................................... 1

1.1 History and Basics of HATs ............................................................................................ 2

1.2 HATs and Disease ............................................................................................................ 4

1.2.1 p300/CBP .................................................................................................................. 4

1.2.2 PCAF/GCN5 .............................................................................................................. 5

1.2.3 HAT1 ........................................................................................................................ 5

1.2.4 Rtt109 ...................................................................................................................... 6

1.3 HAT Inhibitors ............................................................................................................... 7

1.4 Introduction to the MYST Family .................................................................................. 8

1.5 MYST Family Implications in Disease ........................................................................ 11

1.6 MYST Family Inhibitors .............................................................................................. 13

1.6.1 Bisubstrate MYST Inhibitors .................................................................................... 13

1.6.2 Small Molecule MYST Inhibitors ............................................................................. 14

1.7 Autoacetylation ............................................................................................................. 16

1.7.1 Autoacetylation of p300 .......................................................................................... 18

1.7.2 Autoacetylation of Rtt109 ...................................................................................... 20

1.7.3 Autoacetylation of PCAF ........................................................................................ 22

1.7.4 Autoacetylation of MYST Family ......................................................................... 22

1.8 Dissertation Objectives ................................................................................................ 26

CHAPTER 2: *In Vitro Activity Assays for MYST Histone Acetyltransferases and Adaptation for High-Throughput Inhibitor Screening* ............................................. 28

2.1 Preface .......................................................................................................................... 29

2.2 Preparation of Recombinant MYST Proteins ................................................................. 30
2.3 Kinetic Analysis

2.3.1 Primer on Steady State Kinetic Analysis

2.3.2 Specific Kinetic Properties of MYST Acetyltransferases

2.3.3 Controlling for Chemical Acetylation

2.3.4 Kinetic Assay Using Radioactive Acetyl-CoA

2.4 Thermofluor Assay to Investigate Cofactor and Inhibitor Binding

2.5 Inhibitor Screening

2.5.1 General Considerations

2.5.2 Buffer Considerations

2.5.3 Assays

2.5.3.1 Radioactivity Based Assay

2.5.3.2 ELISA Assay

2.5.3.3 Thermofluor Binding Assay

2.5.4 Eliminating Promiscuous Inhibitors

2.5.5 PAINS

2.5.6 Colloidal Aggregators

2.6 Conclusions

CHAPTER 3: High Throughput Screening Results

3.1 Introduction

3.1.1 Preface

3.1.2 Overview

3.1.3 Definitions

3.2 Materials and Methods

3.2.1 Protein Expression, Purification, and Quality Control

3.2.2 High Throughput Assays

3.2.3 Low Throughput Assays

3.2.3.1 Radioactivity Based Assay
3.2.3.2 Thermofluor Assay .........................................................57

3.2.4 Screening Libraries ..........................................................58
3.2.4.1 Spectrum Collection .......................................................58
3.2.4.2 NIH Clinical Collection ..................................................58
3.2.4.3 ChemDiv Library ..........................................................58
3.2.4.4 The Maybridge “Hitfinder” Library .................................59
3.2.4.5 The Natural Product Discovery Institute Library ..............59
3.2.4.6 The Lankenau Chemical Genomics Center Library ..........60

3.2.5 hMOF Crystallization with Merbromin ...............................60

3.2.6 Mechanism of Action (MOA) Assay .................................61

3.3 Results and Discussion ..........................................................62
3.3.1 Initial Validation of Primary Screening Hits .........................62
3.3.4 Cysteine Reactive Compounds ..........................................63
3.3.5 Confirmation of LCGC Active Compounds .........................66
3.3.5.1 Details on Initial Validation of Primary Screening Hits from LCGC .................................................................66
3.3.5.2 Thermofluor Assay Eliminated Two More LCGC Hits ......71
3.3.5.3 LCGC Hits are Irreversible Inhibitors .............................73
3.3.6 Resynthesis and Lack of Validation of LCGC Top Hits ..........76

3.4 Future Directions .................................................................83

CHAPTER 4: Structural and Functional Role of Acetyltransferase hMOF K274 Autoacetylation .................................................................87

4.1 Introduction ...........................................................................88

4.2 Experimental Procedures .....................................................89
4.2.1 Protein Expression and Purification ....................................89
4.2.2 Enzyme Activity Assays ...................................................90
4.2.3 Thermal Denaturation Assay .............................................91
4.2.4 Crystal Structure Determination of MYST HAT Domains ....92
4.2.5 Cell Studies........................................................................................................... 93
4.2.6 Western Blots for Acetyl-Lysine........................................................................ 94
4.3 Results ...................................................................................................................... 95
    4.3.1 No other amino acid substitution at position 274 in hMOF can mimic the
        potentiating role of K274 acetylation on hMOF thermal stability and cognate lysine
        substrate acetylation ..................................................................................... 95
    4.3.2 Crystal structure of hMOF K274P reveals loss of rigidity of the α2-β7 loop
        ...................................................................................................................... 102
    4.3.3 Autoacetylation of K274 does not require active site catalytic residues C316 and E350.
        ...................................................................................................................... 105
    4.3.4 Autoacetylation of K274 increases hMOF half-life in cells.................... 108
4.4 Discussion ............................................................................................................... 110

CHAPTER 5: Conclusions and Future Directions..................................................... 114

5.1 Identification of hMOF Inhibitors ...................................................................... 115
    5.1.1 Conclusions .................................................................................................. 115
    5.1.2 Future Directions: Crystallographic Fragment Screening with hMOF
        ...................................................................................................................... 116
5.2 Autoacetylation of K274 ...................................................................................... 118
    5.2.1 Conclusions .................................................................................................. 118
    5.2.2 Future Directions: Improving hMOF Affinity for H4 Tail....................... 119
    5.2.3 Future Directions: Tracking the K274 Acetyl Group............................... 120
    5.2.4 Attempted Future Directions: Stabilizing the α2-β7 Loop ...................... 123

REFERENCES ........................................................................................................... 125
LIST OF TABLES

Table 1. Autoacetylation of HATs .................................................................17

Table 2: IC$_{50}$ values and Hill slopes calculated from the data presented in Figure 14 compared to values calculated for acetonyl CoA .........................................................70

Table 3: Average melting temperatures for hMOF K274 mutants and MOZ K604 mutants in the presence and absence of 100 µM CoA.................................................................99

Table 4: Data statistics for crystal structures of hMOF-K274P and hMOF-C316S/E350Q.........................................................................................................................103

Table 5: Frequently used abbreviations........................................................124
LIST OF FIGURES

Figure 1. Proposed “ping-pong” catalytic mechanism employed by MYST family acetyltransferases .................................................................................................................. 9

Figure 2. Model for p300 activation by autoacetylation ........................................ 19

Figure 3. Autoacetylation of Rtt109 K290 ............................................................... 21

Figure 4. Close up and schematic of autoacetylation of MYST family acetyltransferases .................................................................................................................. 24

Figure 5. Michaelis-Menten equation, $k_{\text{cat}}$ equation, and graphical representation ...... 33

Figure 6. Progress curve showing the enzymatic activity of the hMOF .......... 36

Figure 7. Michaelis-Menten curves and calculated steady state parameters for hMOF ...................................................................................................................... 38

Figure 8. Thermofluor assay as a readout of hMOF ligand binding .................. 40

Figure 9. Equation and sample data for determining assay Z-factor .................. 42

Figure 10. Sample data for an enzyme recovery assay with a MYST protein .... 50

Figure 11. Chemical structures of the two cysteine-reactive hits ..................... 64

Figure 12. Fo–Fc density for hMOF soaked with merbromin ............................. 65

Figure 13. Data comparing LCGC hits with and without 0.05% tween 20 .... 67

Figure 14. Structures and IC$_{50}$ values for five hits from the LCGC library .... 69

Figure 15. The top three LCGC hits stabilize hMOF in a thermofluor assay .... 72

Figure 16. Attempted Mechanism of Action (MOA) assays using compound 115772 ........................................ 74

Figure 17. Data for enzyme recovery assays ...................................................... 75

Figure 18. Resynthesis of compound 115772 failed to reproduce activity .... 77

Figure 19: Mass spectrometry chromatogram for compound 115772 ................. 78

Figure 20: Mass spectrometry chromatogram for compound 128041 ................. 79

Figure 21: Mass spectrometry chromatogram for compound 14855 ................. 80

Figure 22: Resynthesis Details ........................................................................... 82
Figure 23: Mass Spectrometry chromatogram for another compound from the LCGC Library

Figure 24: Salicylic acid bound to hMOF

Figure 25: Size-exclusion chromatograms of hMOF K274 mutants suggests that they are properly folded

Figure 26: All hMOF K274 mutants are catalytically inactive

Figure 27: hMOF K274 mutants are thermally destabilized, but are still able to bind cofactor

Figure 28: Subset of hMOZ K604 mutations reveals reduced catalytic activity and thermal stability

Figure 29: The hMOF-K274P structure reveals a loss of rigidity in the α2-β7 loop

Figure 30: hMOF-C316S/E350Q is autoacetylated on K274

Figure 31: hMOF WT and hMOF K268M share very similar kinetic profiles

Figure 32: hMOF-WT shows a longer half-life than hMOF-K274R in cells

Figure 33: Overlay of Hat1 bound to H4 peptide structure with hMOF structure reveals that hMOF α2-β7 loop likely makes contacts with H4
CHAPTER 1

Introduction to the Acetyltransferases

Parts of this chapter were originally published in the Journal of Biological Chemistry, (McCullough, Cheryl E.; Song, Shufe; Shin, Michael H.; Johnson, F. Brad; Marmorstein, Ronen. Structural and functional role of acetyltransferase hMOF K274 acetylation. J Biol. Chem. 2016 Aug 26) © the American Society for Biochemistry and Molecular Biology.

Parts of this chapter have been reprinted from Methods in Enzymology, Vol. 573, McCullough, Cheryl E.; Marmorstein, Ronen. In vitro activity assays for MYST histone acetyltransferases and adaptation for high-throughput inhibitor screening. pp. 139-160. © Copyright 2016 Elsevier Inc. Academic Press with permission from Elsevier

Parts of this chapter reprinted (adapted) with permission from McCullough, Cheryl E.; Marmorstein, Ronen. Molecular basis for histone acetyltransferase regulation by binding partners, associated domains, and autoacetylation. ACS Chem. Biol. 2016 Mar 18. 11(3):632-642. Copyright 2016 American Chemical Society.”
1.1 History and Basics of HATs

Acetylation is one of several post-translational modifications (PTMs) that occurs on histones to regulate chromatin dynamics and function in DNA transcription, replication, repair, and other DNA-templated activities. The hypothesis that histone acetylation could regulate gene expression was first proposed by Allfrey and coworkers in 1964 (Allfrey, Faulkner et al. 1964). However, it was not for another thirty years that the first histone acetyltransferase, HAT1, was identified by Sternglanz and coworkers (Kleff, Andrulis et al. 1995) and Gottschling and coworkers (Parthun, Widom et al. 1996). In 1996, Allis and coworkers isolated the first acetyltransferase, GCN5, from Tetrahymena, opening the door to the identification of other histone PTMs, the enzymes that mediated the modifications, and their roles in chromatin regulation (Brownell, Zhou et al. 1996).

Of the multiple histone acetyltransferases (HATs) that have been identified to date, the most studied have been classified into five subfamilies based on sequence homology and substrate acetylation properties: HAT1, GCN5/PCAF, MYST (named for founding members MOZ, YBF2/SAS3, SAS2, and Tip60), p300/CBP, and Rtt109. These HATs all share a structurally conserved Acetyl Coenzyme A (AcCoA) binding core, and within these subfamilies proteins share catalytic mechanisms and structurally conserved core-flanking regions. Interestingly, despite the structural homology between many of these proteins, sequence homology is limited for HATs, with the exception of GCN5/PCAF, and HAT1 which belong to the GCN5-related N-acetyltransferase (GNAT) superfamily sharing four regions of sequence conservation spanning ~100 amino acids (Neuwald and Landsman 1997) despite the different substrate specificities between GCN5/PCAF and HAT1 as reviewed in (Vetting, LP et al. 2005).
Histones are not the only proteins that are acetylated. Through acetylmotic studies, we now know that thousands of proteins are acetylated throughout the cell, spanning a wide class spectrum, including transcription factors, kinases, ubiquitin ligases, ribosomal proteins and metabolic enzymes, to mediate a broad range of cellular functions, including cell cycle control, DNA damage check-points, cytoskeleton organization, endocytosis and metabolism (Choudhary, Kumar et al. 2009, Smith and Workman 2009, Spange, Wagner et al. 2009, Zhang, Sprung et al. 2009). While some of the non-histone acetyltransferases have been identified, such as the α-tubulin acetyltransferase, αTAT1 (Akella, Wloga et al. 2010, Shida, Cueva et al. 2010) and the cohesin acetyltransferase, ESCO1 in humans and Eco1 in yeast (Ivanov, Schleiffer et al. 2002, Williams, Garrett-Engele et al. 2003, Hou and Zou 2005), the biologically relevant enzymes, if any, for many of the other acetylation sites are not known. It should be noted that several HATs such as p300/CBP have been shown to also acetylate many non-histone substrates.
1.2 HATs and Disease

Given the pervasiveness of acetylation in the cell, it is not surprising that aberrant protein acetylation or acetyltransferase function has been correlated with multiple human diseases. While this section will not provide a comprehensive review of all the disease associations of the HATs, it will touch on some key examples to make a point of the prevalence of HATs in a number of cancers and other diseases. Specifics regarding the MYST family of HATs and their disease associations will be discussed later in section 1.5.

1.2.1 p300/ CBP

The p300/CBP family of acetyltransferases has been implicated in a variety of diseases. Mutations in p300 or CBP can cause Rubenstein-Taybi syndrome, a disease characterized by facial physical defects and mental retardation (Rubinstein and Taybi 1963, Miller and Rubinstein 1995). The leading diabetes drug, Metformin, was recently shown to act through p300/CBP inhibition, and heterozygous CBP knockout mice are lean and exhibit increased insulin sensitivity (He, Sabet et al. 2009). The integration of HIV-1 virus into the human genome is catalyzed by the viral protein integrase, the activity of which is also increased by p300-mediated acetylation (Cereseto, Manganaro et al. 2005). Furthermore, p300 and CBP have been implicated in a variety of cancers. CBP and p300 null cells formed hematological malignancies (Rebel, Kung et al. 2002). Additionally, the p300 HAT is mutated in a subset of colorectal and gastric cancers making it a bona fide tumor suppressor (Muraoka, Konishi et al. 1996). Furthermore, elevated levels of p300 are often correlated with poor disease prognosis in a variety of
cancers including hepatocellular carcinoma (HCC), prostate cancer, cutaneous squamous cell carcinoma (CSCC), colorectal adenocarcinomas, breast cancer, and non-small cell lung carcinoma (Kaypee, Sudarshan et al. 2016). Another study has found that various types of CBP gene mutations are involved in lung carcinogenesis (Kishimoto, Kohno et al. 2005).

1.2.2 PCAF/GCN5

The PCAF acetyltransferase has been reported to play a role in multiple disease states. PCAF was found to be down regulated in esophageal squamous cell carcinoma (ESCC) (Qin, Fu et al. 2008), in advanced stage ovarian cancers (Sunde, Donninger et al. 2006) and in intestinal type gastric cancers (ITGC) (Ying, Wang et al. 2010). Additionally, drug resistance in breast cancer cell lines has been shown to be correlated with increased PCAF activity (Toth, Boros et al. 2012). Linking aberrant activity to diabetes, PCAF has also been shown to regulate glucose levels and improve insulin sensitivity by repressing gluconeogenesis (Sun, Wang et al. 2014). GCN5 has also been found to be associated with cancer and viral integration as it is over-expressed in non-small cell lung cancer (Chen, Wei et al. 2013), and has also been shown to play a role in HIV-1 viral integration (Terreni, Valentini et al. 2010).

1.2.3 HAT1

Recent studies have linked HAT1 to various diseases, including cancers and hepatitis. One study has found that HAT1 is over expressed in both primary and
metastatic tumors of colorectal cancers (Seiden-Long, Brown et al. 2006). Another study demonstrated that HAT1 plays role in esophageal carcinoma (EC). In this case, HAT1 was found to be upregulated in primary tumors and expression of HAT1 directly coorelated with poor EC tumor differentiation (Xue, Hou et al. 2014). Another report found that a knockdown of HAT1 down regulated hepatitis B DNA in cells, suggesting that HAT1 may play a role in hepatitis B replication (Wang, Liu et al. 2015).

1.2.4 Rtt109

Unlike the other HATs previously described, Rtt109 is a fungal-specific acetyltransferase with no known homolog in humans (Wang, Tang et al. 2008). However, due to its specificity and wide-ranging conservation in fungi, it can be targeted for anti-fungal therapies in humans. For example, it has been found that deletion of Rtt109 in the opportunistic pathogen Candida albicans decreases the mortality rate in infected mice (Lopes da Rosa, Boyartchuk et al. 2010). Rtt109 has also been identified as a drug target for Pneumocystis pneumonia (PcP), an opportunistic infection caused by the fungal species Pneumocystis jirovecii (Dahlin, Kottom et al. 2014).
1.3 HAT inhibitors

In an attempt to better study the acetyltransferases, and to potentially target them therapeutically, there have been various attempts to develop acetyltransferase inhibitors. This section will introduce a few key HAT inhibitors but will not be a comprehensive review. Rather this section acts as a primer for section 1.6 which will discuss MYST specific HAT inhibitors in greater depth.

Lys-CoA is a peptide-CoA conjugated bisubstrate inhibitor of p300 developed by Cole and coworkers. Lys-CoA was both potent with an IC$_{50}$ of approximately 500 nM and selective for p300 over PCAF by approximately 200x. However, due to multiple negatively charged phosphate groups, this inhibitor was not cell permeable (Lau, Kundu et al. 2000). Later, cell-permeable derivatives of Lys-CoA were developed, but these compounds showed poorer potency and selectivity for p300 as compared to the parent compound (Zheng, Balasubramanyam et al. 2005). There are also a number of natural products that have been shown to inhibit HAT activity including anacardic acid which inhibits p300 and PCAF with an IC$_{50}$ of about 7 µM (Balasubramanyam, Swaminathan et al. 2003) and curcumin, which inhibits p300 with an IC$_{50}$ of about 25 µM (Balasubramanyam, Varier et al. 2004). Garcinol, a benzophenone fruit extract, has been shown to inhibit PCAF and p300 activity with an IC$_{50}$ value of 6 µM (Balasubramanyam, Altaf et al. 2004). A monomethylated derivative of garcinol named LTK14 was later developed, which while maintaining the same IC$_{50}$ as garcinol, shows increased specificity towards p300 (Mantelingu, Reddy et al. 2007). Later studies employed virtual ligand screening to identify the compound C646, a p300 specific, cell permeable inhibitor, but it is 20x less potent than Lys-CoA (Bowers, Yan et al. 2010).
1.4 Introduction to the MYST Family

Of the different HAT families that have been characterized to date, the MYST proteins form the largest and most diverse family (Yang 2004), and are conserved from yeast to man (Yuan, Rossetto et al. 2012). MYST family proteins have not been as well studied as some of the other HAT families, yet they play a number of roles in the cell including gene regulation, DNA repair, and cell cycle homeostasis (Sapountzi and Cote 2011). In addition to histones, MYST family proteins also acetylate a variety of other proteins including the p53 tumor suppressor (Mellert and McMahon 2009), ORC1 (Iizuka and Stillman 1999), HIV-1 Tat (Cereseto, Manganaro et al. 2005), and a number of metabolic regulators (Lin, Lu et al. 2009).

MYST proteins share a structurally conserved AcCoA core binding region with other HATs, yet they employ several unique structural and catalytic properties within their conserved HAT domain. The MYST HAT domains contain N- and C-terminal regions that flank the core binding region that are structurally distinct from other HAT families. This includes unique N-terminal C₂HC zinc finger and C-terminal helix-turn-helix motifs required for their chromatin regulatory activities (Yan, Barlev et al. 2000, Lafon, Chang et al. 2007). MYST proteins have been demonstrated to acetylate cognate substrates through a ping-pong catalytic mechanism, whereby AcCoA binds first to transfer the acetyl group to an active site cysteine residue to form an acetyl-enzyme intermediate prior to acetyl-transfer to the cognate lysine substrate. An active site glutamate residue which is buried in a hydrophobic pocket to help raise its pKa for proton extraction is proposed to serve as a key residue for catalysis (Figure 1) (Yan, Harper et al. 2002). However, another study demonstrates that the yeast Esa1 member
Figure 1. Proposed “ping-pong” catalytic mechanism employed by MYST family acetyltransferases. A glutamate in the active site acts to first deprotonate the active site cysteine. Next the acetyl group from acetyl CoA is transferred to the cysteine. The glutamate then protonates the leaving cofactor. The glutamate is then able to deprotonate the substrate lysine so that the cysteine can transfer the acetyl group to the lysine. Finally both the glutamate and cysteine react with water to return to their initial state.
of the MYST family, when assembled within a physiologically relevant piccolo NuA4 complex, shows a strong catalytic dependence on the corresponding active site glutamate residue, but not the active site cysteine residue, suggesting that yeast Esa1 within the piccolo NuA4 complex proceeds through a ternary catalytic mechanism similar to the other HAT families (Berndsen, Albaugh et al. 2007). More recently, MYST acetylation of protein substrates was also found to require active site lysine autoacetylation (Lu, Li et al. 2011, Sun, Guo et al. 2011, Yang, Wu et al. 2012, Yuan, Rossetto et al. 2012), which will be discussed in further detail in section 1.7.4.
1.5 MYST Family Implications in Disease

The altered activities of MYST proteins have also been implicated in human disease. Tip60 is linked to the onset of Alzheimer’s disease (Cao and Sudhof 2001, Baek, Ohgi et al. 2002, Kinoshita, Whelan et al. 2002, Cao and Sudhof 2004). In 2001 it was reported that Tip60 formed a transcriptionally active trimeric complex with nuclear adaptor protein Fe65 and amyloid-β precursor protein (APP), a protein that can be cleaved to release amyloid-β peptide in Alzheimer’s disease (Cao and Sudhof 2001). Later it was found that over expression of APP led to upregulation of Fe65 and Tip60 in transgenic mice (Baek, Ohgi et al. 2002). Furthermore, it was demonstrated that Tip60 could mediate apoptosis induced by the APP cleaved c-terminal fragment (Kinoshita, Whelan et al. 2002).

MYST family dysregulation has also been implicated in a variety of cancers. An RNAi screen identified the Tip60 gene as a modulator of p53-dependent function and senescence avoidance (Berns, Hijmans et al. 2004). Furthermore, Tip60 was demonstrated to be down-regulated in lung and colon cancer tumor tissues where 7 of 20 colon cancer patients and 8 of 20 lung cancer patients had a statistically significant down-regulation of Tip60 (Lleonart, Vidal et al. 2006). Contrastingly, Tip60 was demonstrated to be up-regulated in epithelial tumors of transgenic mice (Hobbs, Wei et al. 2006). One study has reported that loss of acetylation on the MOF (males absent on the first) target lysine 16 of histone H4 is common in many human malignancies. In this study, 25 cancer cell lines including those for leukemia, breast cancer, lung cancer, and colon cancer were studied using LC-ES/MS as compared to the corresponding normal cell lines. Every cancer cell line tested revealed a loss in H4K16 acetylation. These results were then confirmed by isolating histone H4 from the primary tumors of 16
lymphomas and 20 colorectal adenocarcinomas, which showed approximately a 40% decrease in H4K16 acetylation as compared to the normal tissue counterpart cells. This loss of acetylation occurred as an early event, and progressed with time (Fraga, Ballestar et al. 2005). Further studies linked MYST proteins to disease by demonstrating that a recurrent translocation in a subset of acute myeloid leukemias was an in-frame fusion of the HAT domain of the MYST protein MOZ (monocytic leukemia zinc-finger protein) and CBP (CREB-binding protein) HAT with most of the functional domains remaining intact. The proposed mechanism of leukaemogenesis for this fusion protein is aberrant acetylation and transcriptional deregulation of the cAMP and mitogen responsive signaling pathways (Borrow, Stanton et al. 1996). Later it was also found that the MOZ-CBP fusion protein could target the AML1 complex to inhibit AML-1 mediated transcription of hematopoietic genes to induce leukemia (Kitabayashi, Aikawa et al. 2001).
1.6 MYST Family Inhibitors

1.6.1 Bisubstrate MYST Inhibitors

Because of their connections to disease, there have been efforts to develop MYST inhibitors. Existing inhibitors, however, are lacking in potency, selectivity, or favorable pharmacokinetic properties. Wu et al. reported on the development of a series of bisubstrate CoA-peptide MYST inhibitors based on the various substrate specificities of ESA1 and Tip60 (Wu, Xie et al. 2009). The specificity of the bisubstrate inhibitors were tested in comparison to the non-MYST family acetyltransferases, p300 and PCAF, and the potency of the bisubstrates were tested in comparison to the natural product HAT inhibitors anacardic acid and curcumin, which inhibit MYST proteins with approximate IC_{50} values of 300 µM and 200 µM, respectively. The majority of the MYST directed bisubstrate inhibitors had low micromolar IC_{50} values for Tip60 and ESA1.

H4K16CoA was particularly potent against both ESA1 and Tip60, with IC_{50} values of approximately 5.5 µM and 17.6 µM respectively, and was thus approximately 20x more potent than the natural product acetyltransferase inhibitor, anacardic acid. H4K16CoA however lacked specificity for the MYST enzymes and also inhibited p300 with an IC_{50} of approximately 6.6 µM. In fact, most of the bisubstrate inhibitors developed in this study were more potent against p300 than the targeted MYST protein. This could likely be due to the fact that MYST enzymes, unlike other acetyltransferases, do not form ternary complexes for catalysis. Rather, they follow an ordered substrate binding ping-pong mechanism, which is further supported by the fact that the bisubstrate inhibitor H4K16CoA was found to be competitive against AcCoA but noncompetitive against histone substrate when tested with Tip60. In addition to their poor selectivity, the negatively charged CoA moiety on the bisubstrate inhibitors makes them unlikely to be
cell permeable and their use may be limited to strictly *in vitro* studies. However the authors suggest that their use could be expanded by linking a cell permeable motif to the compounds (Wu, Xie et al. 2009).

1.6.2 *Small Molecule MYST inhibitors*

In continuing studies to inhibit MYST family acetyltransferases, virtual screening was performed to identify small molecule inhibitors of Tip60 using the crystal structure of ESA1, the yeast homolog of Tip60, as a model (Wu, Wang et al. 2011). The top 76 hits from the virtual screen were tested for efficacy against Tip60 using *in vitro* biochemical assays and two compounds were shown to inhibit Tip60 with IC$_{50}$ values in the mid micromolar range. Two further compounds were then selected based on structural similarity to the initial two hits for a total of four inhibitors resulting from this screen. While three of the four small molecule inhibitors had better potency than curcumin, the inhibitors still exhibited poor overall potency and the inhibitors did not exhibit selectivity between Tip60 and other HAT families. The authors speculate that this could be because the compounds were found to be competitive with AcCoA and may bind the AcCoA binding pockets shared by the various families of HATs with comparable affinity (Wu, Wang et al. 2011).

Because anacardic acid has been shown to inhibit Tip60, albeit with only mid-micromolar potency, another study designed a series of analogs of anacardic acid to determine if they could be developed into more potent, specific inhibitors against the MYST family. The compounds were evaluated against Tip60, PCAF and p300 in *in vitro* biochemical activity assays to test specificity. Overall, the compounds showed a greater
degree of selectivity towards Tip60 than the two other HATs. The most potent Tip60 anacardic acid analog inhibitors had hydrophobic substitutions at the 6-position in the salicylate ring, whereas the weakest inhibitors had hydrophilic substitutions at the 6-position. Some of the analogs had a phenylethyl substitution at the 6-position, which improved the lipophilicity of those compounds over the parent anacardic acid compound, potentially making them improved candidates for lead development. One analog in particular (referred to as 20 in the referenced manuscript and is now commercially available by the name MG 149) displayed strong potency against Tip60, decreasing enzyme activity by 88% at 200 µM, as well as good selectivity with limited observable inhibition of either p300 or PCAF. This compound was also shown to inhibit hMOF with a similar potency, demonstrating that it may be a specific inhibitor of MYST family HATs. Kinetic studies with MG 149 and Tip 60 revealed that the compound is competitive with AcCoA and noncompetitive with the substrate peptide. The efficacy of the analogs were also tested on the nuclear extracts of HeLa cells and rat brain tissue samples. Confirming what had been observed in vitro, the majority of the compounds displayed a concentration dependent inhibition of histone H4 acetylation. Furthermore, polar substitutions at the 6-position showed less potency and aliphatic substitutions showed increased potency in agreement with the studies on the recombinant Tip60 (Ghizzoni, Wu et al. 2012).
1.7 Autoacetylation

Another mechanism of regulation for acetyltransferases is autoacetylation. Our lab defines autoacetylation as either a chemical or enzymatic event whereby a protein can become acetylated in the presence of no other protein except for itself. Currently, p300, the MYST family of acetyltransferases, Rtt109, and PCAF are all known to be regulated by acetylation to some degree. Interestingly, despite the prevalence of this modification amongst acetyltransferases, each HAT subfamily appears to be regulated by autoacetylation via unique mechanisms (Table 1).
Table 1. Autoacetylation of HATs

<table>
<thead>
<tr>
<th>Protein</th>
<th>Relevant autoacetylated lysine(s)</th>
<th>Functions of acetylation</th>
</tr>
</thead>
<tbody>
<tr>
<td>p300</td>
<td>K1499, K1549, K1554, K1560</td>
<td>• activates enzyme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• regulates autoinhibitory loop</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• decreases $K_m$ for AcCoA and substrate peptide</td>
</tr>
<tr>
<td>MOF</td>
<td>K274</td>
<td>• activates enzyme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• decreases $K_m$ for substrate peptide</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• increases $k_{cat}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• maintains loops conformation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• conserved throughout MYST family</td>
</tr>
<tr>
<td>Rtt109</td>
<td>K290</td>
<td>• activates enzyme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• decreases $K_m$ for AcCoA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• increases $k_{cat}$</td>
</tr>
<tr>
<td>PCAF</td>
<td>K416, K428, K430, K431, K442</td>
<td>• 2-3 fold increase in catalytic activity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• increases structural stability</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• localizes PCAF to nucleus</td>
</tr>
</tbody>
</table>

Only autoacetylated lysines with a known effect on enzyme biochemistry have been listed. MOF has been listed as a representative member of the MYST family.
1.7.1 Autoacetylation of p300

Within the p300 HAT domain there exists a ~40 residue proteolytically sensitive lysine-rich loop that can become heavily acetylated to promote catalytic activity. Studies have found that if this loop of p300 (spanning residues 1520–1560) is deleted the enzyme becomes constitutively active decreasing the $K_m$ for both AcCoA and substrate peptide rather than increasing the catalytic rate (Thompson, Wang et al. 2004). This loop is hypothesized to act as an autoinhibitory loop that regulates p300 activity by blocking substrate binding when unmodified, and exposing the substrate binding region when acetylated (Thompson, Wang et al. 2004). This model for autoinhibition of the hypoacetylated loop is consistent with the p300 HAT domain crystal structure, which shows a highly acidic substrate-binding site that could accommodate a basic hypoacetylated loop more avidly than a hyperacetylated loop (Figure 2). This mechanism of autoinhibition was further supported when it was shown that deletion of this loop in vitro enhanced doxorubicin induced acetylation of the p300 substrate p73 (Thompson, Wang et al. 2004). Along these lines, the effect of this loop deletion on the ability of p300 to bind the substrate ATF-2 was tested. Not only was the loop unnecessary for substrate binding, but also a hypoacetylated loop blocked substrate binding whereas a hyperacetylated loop enhanced substrate binding providing further evidence that this loop harbors autoinhibitory properties (Karanam, Wang et al. 2007).
Figure 2. Model for p300 activation by autoacetylation. When the proteolytically sensitive loop of p300 becomes acetylated, the positive charge is neutralized, and the substrate binding site of the enzyme becomes accessible. Figure adapted from (Yuan and Marmorstein 2013)
1.7.2 Autoacetylation of Rtt109:

Autoacetylation of Rtt109 at lysine 290 is essential for enzyme activation (Albaugh, Arnold et al. 2011). The acetyl group on this lysine allows it to be inserted into a hydrophobic pocket of conserved residues (Figure 3) (Lin and Yuan 2008, Stavropoulos, Nagy et al. 2008, Tang, Meeth et al. 2008). Interestingly, while it seemed that this residue would thus be important for the thermostability of Rtt109, it was found that the melting temperature of acetylated Rtt109 did not differ from that of deacetylated Rtt109, demonstrating that AcK290 likely does not contribute to protein stability (Albaugh, Arnold et al. 2011). Rather, it was found that acetylation of K290 activated Rtt109 by increasing the affinity for AcCoA and also increasing the rate of catalytic turnover (Albaugh, Arnold et al. 2011). The molecular basis by which autoacetylation of Rtt109 facilitates cognate histone lysine acetylation remains unclear.
Figure 3. Autoacetylation of Rtt109 K290. The autoacetylated K290 (shown in green) of Rtt109 in its buried environment. Figure adapted from (Yuan and Marmorstein 2013)
1.7.3 Autoacetylation of PCAF

PCAF (p300/CBP-associated factor) was first identified as a p300/CBP binding protein with histone acetyltransferase activity (Yang, Ogryzko et al. 1996), and further studies of PCAF demonstrated that PCAF becomes autoacetylated in vitro (Herrera, Bergel et al. 1997). Moreover, it was shown that the autoacetylation of PCAF occurs in vivo, and that while PCAF can be acetylated intermolecularly by p300, it also has the capability to acetylate itself intramolecularly at lysines 416, 428, 430, 441, and 442 (Santos-Rosa, Valls et al. 2003). PCAF displayed greater structural stability and a 2–3 fold increase in catalytic activity when autoacetylated (Santos-Rosa, Valls et al. 2003). Noting that the five lysines targeted for autoacetylation in PCAF occur at the nuclear localization signal (NLS) region of PCAF, the effect of autoacetylation on nuclear localization was investigated. By creating a variety of PCAF constructs and mutants to mimic the unacetylated state, it was shown that acetylated PCAF localized to the nucleus whereas unacetylated PCAF localized to the cytoplasm, indicating that autoacetylation likely plays a role in the regulation of nuclear localization for PCAF (Blanco-Garcia, Asensio-Juan et al. 2009).

1.7.4 Autoacetylation of MYST Family

A number of recent reports have revealed that autoacetylation of a conserved lysine residue promotes catalytic activity in the MYST proteins (Lu, Li et al. 2011, Sun, Guo et al. 2011, Yang, Wu et al. 2012, Yuan, Rossetto et al. 2012). In MOF and Tip60 it was shown that acetylation of K274 and K327, respectively, upregulated catalytic activity (Sun, Guo et al. 2011, Yang, Wu et al. 2012). Furthermore, arginine and glutamine...
mutants mimicking lysine and acetylated lysine respectively displayed a >90% loss of catalytic activity for both enzymes (Yang, Wu et al. 2012, Yuan, Rossetto et al. 2012), implying that the specific acetyl-lysine modification was required for catalytic activation. Further biochemical assays also showed that hMOF K274 mutants were defective in H4 binding (Yuan, Rossetto et al. 2012). In a separate study, the autoacetylation of hMOF was correlated with decreased recruitment of the protein to chromatin in vivo, and decreased binding to nucleosomes in vitro (Lu, Li et al. 2011). Furthermore, K274 mutants in drosophila MOF display a decreased overall thermostability profile (Kadlec, Hallaceli et al. 2011). K274 makes extensive contacts in the acetylated form that may contribute to the maintenance of the loop conformation for productive lysine substrate binding. These include hydrogen bonding of the acetyl carbonyl to Tyr301 and Ser303 and hydrophobic interactions of the acetyl methyl with Phe283 and Phe285 (He, Sabet et al. 2009, Sun, Guo et al. 2011, Yuan, Rossetto et al. 2012). The crystal structure of unacetylated hMOF K274, reveals that the lysine is flipped out of the active site by 90° to both block cognate substrate binding and to form a hydrogen bond with the catalytic E350 (Figure 4, right), suggesting that that unacetylated lysine residue inhibits cognate substrate lysine acetylation by both preventing substrate binding and by tying up a key catalytic residue (Figure 4, left) (Yuan, Rossetto et al. 2012). The mode of K274 acetylation has been proposed to proceed through enzymatic autoacetylation intramolecularly (in cis), which is consistent with the proximity of acetylated K274 in the enzyme active site (Sun, Guo et al. 2011, Yuan, Rossetto et al. 2012). However, this acetylation has also been shown to be independent of the E350 catalytic residue and we have recently shown that a 19-residue hMOF peptide (261-279) can undergo chemical, non-enzymatic, acetylation at K274 with elevated AcCoA concentrations
Figure 4. Close up and schematic of autoacetylation of MYST family acetyltransferases (MOF used as a representative member.) When unacetylated K274 forms a hydrogen bond with the catalytic E350. When K274 becomes acetylated (shown in green) it flips 90° out of the active site and forms a hydrogen bond with S303, which may permit substrate binding (shown in pink). Overlaying of the H4K16CoA from the ESA1/H4K16CoA structure (PDB code: 3TO6) reveals that the substrate would clash with unacetylated K274. Figure adapted from (Yuan and Marmorstein 2013)
(Olia, Barker et al. 2015). This finding is consistent with recent studies by us and others that lysine residues near an appropriately positioned basic residue or basic patch within proteins have elevated capacities for chemical acetylation (Baeza, Smallegan et al. 2015, Olia, Barker et al. 2015). For MYST proteins, a conserved lysine residue is found six residues N-terminal to K274, and we showed that an hMOF peptide with the corresponding native K268 or K268R mutation can undergo chemical acetylation while a K268M mutant cannot (Olia, Barker et al. 2015). These findings leave open the possibility that hMOF K274 autoacetylation may happen non-enzymatically to ‘prime’ the enzyme for cognate substrate acetylation.
1.8 Dissertation Objectives

This dissertation will focus on my two main projects: the development of a high throughput assay to screen for small molecule inhibitors of the MYST acetyltransferase, hMOF, and a characterization study on the catalytic and structural significance of the autoacetylated lysine in hMOF, Lys 274.

Chapter 2 will mostly cover information published in a chapter for *Methods in Enzymology: Enzymes of Epigenetics* entitled “In Vitro Activity Assays for MYST Histone Acetyltransferases and Adaptation for High-Throughput Screening.” This chapter will focus more broadly on our high throughput screening campaign, describing methods for the preparation and biochemical characterization of MYST family acetyltransferases, including protocols for the preparation of recombinant protein, enzymatic assays for measuring steady state parameters and binding assays to measure cofactor and inhibitor binding. Details will also be provided on adapting these assays for high throughput screening for small molecule MYST inhibitors. This chapter discusses hurdles that are likely to be encountered, based on our own experiences, when studying the MYST proteins.

Chapter 3 will cover the mostly unpublished details of our screening campaign, discussing the specific results of our screen and the process of hit validation (and lack of compound validation) upon which Chapter 2 was based. This chapter will detail the results from each library that was screened, the biochemical assays used to reveal that compounds were inhibiting hMOF through nuisance behavior rather than specific binding, and the specific series of events that led to the conclusion of this screening campaign without any compounds identified for further development. Chapter 3 will also
provide some reflections on the missteps of the screening campaign and provide future
directions for finding MYST inhibitors.

Chapter 4 will discuss data published in an article in the *Journal of Biological
Chemistry* entitled “Structural and Functional Role of Acetyltransferase hMOF K274
Autoacetylation.” This chapter details a molecular investigation of K274 autoacetylation
of hMOF. These studies point to the critical and specific role of hMOF K274
autoacetylation in hMOF stability and cognate substrate acetylation and argues that
binding of Ac-CoA to hMOF likely drives K274 autoacetylation for subsequent cognate
substrate acetylation.

Finally, this thesis will conclude with Chapter 5 giving conclusions and future
directions of the work described herein.
CHAPTER 2

In Vitro Activity Assays for MYST Histone Acetyltransferases and Adaptation for High-Throughput Inhibitor Screening

2.1 Preface

This chapter will describe methods for biochemical characterization of MYST family acetyltransferases, including protocols for enzyme assays for measuring steady state parameters and binding assays to measure cofactor and inhibitor binding. We also provide details on adapting these assays for high throughput screening of small molecule MYST inhibitors.
2. 2 Preparation of Recombinant MYST Proteins

We generally follow previously described methods for purifying the MYST proteins (hMOF specifically) (Yuan, Rossetto et al. 2012) but with some modifications. To study MYST activity in vitro we purify the HAT domain of hMOF, residues 174-449, with an N-terminal 6xHis tag separated by a TEV protease recognition site. The plasmid is transformed into E. coli BL21(DE3) codon plus RIL (Stratagene) cells and the protein is overexpressed in TB media by induction with 1 mM IPTG and grown at 15°C overnight after the O.D reaches 0.8-1.5. The cells are harvested and lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.5), 0.5 M NaCl, 2 mM β-mercaptoethanol, 5 mM imidazole, 5% glycerol, 0.1% CHAPS and 0.1 mM phenylmethylsulfonyl fluoride. The lysate is cleared by centrifugation at 28,000g for 30 min at 4 °C. The supernatant is then loaded onto Nickel charged NTA resin (Qiagen) that has been equilibrated with lysis buffer. The resin is then washed with 10 column volumes of lysis buffer containing 25 mM imidazole and then eluted with lysis buffer containing 250 mM imidazole buffer in 10 mL fractions. Because we found that the 6xHis tag does not interfere with hMOF activity or crystallization, we do not cleave the tag. Instead we proceed directly from the affinity column to size exclusion chromatography, skipping ion exchange gel chromatography due to hMOF sensitivity to low salt conditions. The hMOF is concentrated and injected onto an FPLC using a HiLoad Superdex 75 16/60 gel filtration column using a gel filtration buffer composed of 20 mM HEPES (pH 7.5), 0.5 M NaCl. We then concentrate the hMOF protein to ~50 µM, aliquot to ~10 µL, flash freeze using liquid nitrogen, and store at -80°C. Once we thaw an aliquot of hMOF we discard the remainder of that aliquot as multiple flash freezings has a negative impact on enzymatic activity. In our experience hMOF expresses and purifies fairly well, and 1L of cells will yield
approximately 5mg of protein. In the case of hMOZ and yESA1, we found that the proteins express better in LB rather than TB, but still express about 10x less protein than hMOF. Both MOZ and ESA1 tolerate ion-exchange well, and achieve a higher level of activity if sizing buffer/storage buffer is at a low pH and contains reducing agent. We use 20 mM sodium citrate pH 5.5, 100 mM NaCl, and 1 mM DTT.
2.3 Kinetic Analysis

2.3.1 Primer on Steady State Kinetic Analysis

The steady state kinetic parameters of a protein include the Michaelis constant ($K_m$), which is the concentration where the reaction rate of the enzyme is at one-half of its maximal velocity ($V_{max}$) and is related to the affinity of an enzyme for its substrate, and the catalytic constant ($k_{cat}$), which is the rate that an enzyme turns over its substrates (Figure 5).

To determine the kinetic parameters for a given enzyme, very carefully conducted activity assays must be performed. The basic steps to follow to determine these parameters for a bisubstrate system such as the MYST family are to hold one substrate at a saturating concentration (~10x $K_m$) while the other substrate is varied. About half of the concentrations of the varied substrate should be below $K_m$ and the other half above $K_m$. Generally a minimum of eight concentrations should be used (Brooks, Geeganage et al. 2004). Results are then graphed as substrate concentrations versus reaction rate to obtain a Michaelis-Menten curve.

A kinetics reaction must be run under steady state conditions, meaning that the reaction is occurring at a linear rate and less than 10% of the substrate is consumed. These parameters ensure that substrate concentration ([S]) can be assumed to remain constant for a given point. To find steady state conditions a reaction progress curve should be obtained before starting kinetic assays. Additionally, enzyme concentration must be significantly lower than substrate concentration (generally at least 100x lower), this allows for enzyme concentration to be equated to zero and exclude from the
Figure 5. Michaelis-Menten equation, $k_{\text{cat}}$ equation and graphical representation of the Michaelis-Menten equation where $v$ = reaction rate, $[S]$ = substrate concentration, $V_{\text{max}}$ = maximum reaction velocity, $K_m$ = Michaelis constant, and $[E_T]$ = total enzyme concentration.
Michaelis-Menten equation. Furthermore, to obtain $k_{cat}$ from a graph of Michaelis-Menten data, assay output signal must be converted to reaction rate. To do so, a standard curve should be generated so that output can be converted to a known amount of substrate conversion. For example, in our radioactivity assays the read out is in CPM (counts per minute) based on how much $^{14}$C-acetyl is transferred from the acetyl CoA to H4. To generate our standard curve we measure a known number of moles of $^{14}$C-acetyl CoA so that CPM can be related to molar transfer of the acetyl group in a linear equation. That value is then divided by reaction volume to obtain a reaction rate with units of Ms$^{-1}$.

### 2.3.2 Specific Kinetic Properties of MYST Acetyltransferases

The MYST family of acetyltransferases exhibit high substrate $K_m$ values. For example, the $K_m$ for the yESA1 HAT domain is $\sim$920 µM for histone H3 (Yan, Barlev et al. 2000). This high $K_m$ for the substrate peptide causes saturating conditions to be extremely high, which can cause complications for kinetic assays. For example in the case of the yESA1, 9.2 mM of H3 would be 10x $K_m$, and thus saturating. To accommodate this MYST family characteristic, we have mainly used radioactivity based assays to measure enzyme activity. Radioactive read outs tend to be quite sensitive, and are not typically limited by binding capacities. For example if we require the attachment of the peptide substrate to a bead for a fluorescence based assay, the binding capacity of the bead might be less than saturating for the peptide and we might then not be able to achieve maximum velocity ($V_{max}$). Furthermore, when using a synthetic peptide as the substrate, using such a high concentration may lead to trifluoroacetic acid (TFA) contamination that could destabilize the enzyme. According to
Genscript’s website, synthetic peptides may contain 10-45% TFA salt. This amount of TFA could be damaging to the enzyme of interest, leading to a much lower $V_{\text{max}}$ in assays where the peptide is the saturating component, or a decrease in activity as peptide concentration increases. To avoid this issue, we have used slightly less than 10$\times K_m$ as an assumed saturating condition.

2.3.3 Controlling for Chemical Acetylation

Another hurdle that we have encountered in our lab is chemical acetylation of substrates. Because the acetyl group on acetyl CoA is a fairly labile bond it is possible for the acetyl transfer to occur in the absence of enzyme, especially with high concentrations of acetyl CoA in the presence of basic peptides. To control for this background chemical acetylation, it is imperative to include a control experiment that contains substrate and acetyl CoA but no enzyme. This can be especially important when testing very low activity levels such as when studying catalytic mutants. The chemical acetylation should then be subtracted from assay readouts to obtain a more accurate value for enzyme activity. For an example see (Figure 6).

2.3.4 Kinetic Assay Using Radioactive Acetyl-CoA

To perform kinetic experiments using a radioactivity-based assay we first mix hMOF at 50 nM with H4 peptide in PCR tubes in a buffer containing 40 mM Tris (pH 8.0), 100 mM NaCl, 800 $\mu$M Cys, and 7.5$\mu$M BSA. Depending on which substrate is tested, H4 will either be at varied concentrations or saturating. We begin the
Figure 6. Progress curve showing the enzymatic activity of the hMOF$_{174-449}$ K274M mutant over time as compared to a chemical acetylation control, which contains only the substrate and acetyl-CoA, but no enzyme. The acetylation that occurs in the presence of the mutant enzyme is just slightly greater than the chemical acetylation observed in the absence of enzyme, indicating that this mutant has lost much of its catalytic activity. Note, however, that the activity levels of the mutant enzyme could be misleading without the chemical acetylation control.
assay by adding $^{14}$C-acetyl CoA (which will also be varied or saturating depending on which substrate is being examined) in a final volume of 50 µL. Because these assays are so sensitive, and we want each experiment to run for exactly the same amount of time, we space out each reaction by 30 seconds (that is, begin reaction one at time 0:00, two at 0:30, three at 1:00 etc.). We allow the reaction to proceed for eight minutes, which we have previously determined to be within the linear range for the enzyme. To stop the reaction, we spot 20 µL each of the reaction onto two P81 filter papers, which will bind the positively charged H4 peptide. This step, again, is done in 30-second intervals so that each reaction proceeds for exactly eight minutes (that is, spot reaction one at time 8:00, two at 8:30, three at 9:00 etc.). The filter papers are then washed 3x in 20 mM HEPES (pH 7.5). Because water is insoluble in some scintillation fluids and might interfere with signal readout, we dip each filter paper in acetone and lay each one out to dry on a paper towel. We then put each filter paper in a scintillation vial, add scintillation fluid and read our results on a scintillation counter where CPM will correlate with the amount of $^{14}$C-acetyl that has been transferred from the acetyl CoA to the H4 peptide and thus, enzyme activity. We can then use this data to prepare Michaelis-Menten curves for the enzyme (Figure 7).
Figure 7. Michaelis-Menten curves and calculated steady state parameters for hMOF<sub>174-449</sub>. Note that these values vary somewhat from the aforementioned K<sub>m</sub> values for hMOF of 50 µM and 400 µM for acetyl CoA and H4 peptide, respectively, likely due to the difference in enzyme preparation and buffer conditions. To calculate k<sub>cat</sub> for this data, V<sub>max</sub> was divided by the total enzyme concentration of 50nM.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>K&lt;sub&gt;m&lt;/sub&gt; for AcCoA</td>
<td>27 ± 5 µM</td>
</tr>
<tr>
<td>K&lt;sub&gt;m&lt;/sub&gt; for H4 peptide</td>
<td>650 ± 90 µM</td>
</tr>
<tr>
<td>k&lt;sub&gt;cat&lt;/sub&gt;</td>
<td>160 ± 8 min&lt;sup&gt;-1&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
2.4 Thermofluor Assay to Investigate Cofactor and Inhibitor Binding

A thermofluor assay measures the melting temperature of a protein (Pantoliano, Rhind et al. 2000). If the protein is stabilized by a cofactor or ligand it will have a higher melting temperature. For example, Apo hMOF HAT domain has a $T_m$ of about 37°C, however, upon addition of 100 µM CoA, acetyl CoA, or acetonyl CoA the $T_m$ increases to approximately 47 °C, 50 °C and 52 °C, respectively; interestingly, however, we do not observe a strong thermal shift of hMOF in the presence of H4 peptide (Figure 8). The thermofluor assay uses SYPRO® Orange, a dye that fluoresces in a hydrophobic environment. Thus if SYPRO® Orange is present as a protein is melting it will bind the exposed hydrophobic patches and will fluoresce with increasing intensity as a protein unfolds. This data can then be plotted to produce a melting curve.

There are some caveats to this assay. Firstly, because of the nature of SYPRO® Orange fluorescence in a hydrophobic environment, detergent in the buffer can interfere with signal. In our experience, baseline fluorescence of SYPRO® Orange in the presence of detergent (0.05% Triton-X) is significantly increased, and completely masks any melting data from the protein. Decreasing to 0.01% Triton X shows some improvement but still complicates the melting curve of the protein. It is also possible that the detergent may directly affect protein unfolding properties. If detergent is absolutely required, then other dye options would need to be explored to employ this assay. Additionally, if another protein is present in the sample (a contaminant or BSA, for example) that protein will melt also, producing its own melting curve in tandem with the protein of interest and thus complicating the data.
Figure 8. Thermofluor assay as a readout of hMOF ligand binding. Data for apo hMOF, and hMOF in the presence of 1 mM H4 peptide, and 100 µM CoA, acetyl CoA, and acetonyl CoA, respectively. $T_m$ is the inflection point for each curve: $\sim$37°C for apo hMOF, $\sim$47°C for hMOF with CoA, $\sim$50°C for hMOF with acetyl CoA, $\sim$52°C for hMOF with acetonyl CoA, and $\sim$38°C for hMOF with H4 peptide.
2.5 Inhibitor Screening

2.5.1 General Considerations

When beginning a screening campaign, one should consider what assay is going to be most appropriate for the system being tested and what buffer conditions are appropriate for that assay. Typically, an additive of some kind will be required to terminate all of the reactions in a high throughput screen simultaneously; for our assays we have used 3.7% HCl for this purpose. Also, a control inhibitor is often required to monitor screening quality statistics. Acetonyl-CoA is a non-hydrolyzable acetyl CoA analog that has served us as a control inhibitor against MYST proteins. An orthogonal secondary assay must also be planned to counter screen any hits from the primary assay to eliminate any false positives that may have interfered with the detection system of the primary assay. There are also a few standard statistics for HTS that should be performed before screening compound libraries. One important validation statistic is the Z-factor, a parameter which measures the screening window of an assay to assess assay quality (Figure 9) (Zhang, Chung et al. 1999).

The Z-factor can be any value from 0 to 1 where assay quality increases with increasing Z-factor, which would indicate a large signal window with low signal variability. To obtain a Z-factor a control plate should be run in the assay where half of the reactions are a control for positive signal (enzyme with no inhibitor present/ DMSO only) and the other half of the reactions are controls for background signal (either containing no enzyme, a catalytically dead mutant of the enzyme, or WT enzyme plus a saturating amount of control inhibitor, depending on what is available). The means and standard deviations of the results can then be calculated and put into the equation.
Figure 9. Equation and sample data for determining assay Z-factor. $Z' = Z$-factor, $\sigma_s =$ standard deviation of the positive signal, $\sigma_b =$ standard deviation of the background signal, $\mu_s =$mean of the positive signal, and $\mu_b =$mean of the background signal. Each point in the graph represents the readout for a single reaction within a 384 well plate.
shown in Figure 9 to calculate Z-factor. Generally the minimum acceptable Z-factor value is 0.5, with higher values being more desirable (Zhang, Chung et al. 1999).

In general, when screening for inhibitors each substrate concentration should be set at $K_m$ within a reaction. This allows the assay to be sensitive to competitive inhibitors while maintaining robust signal (Williams and Scott 2009). However, this can pose a problem for MYST family proteins, which tend to have a very high $K_m$ for their substrates. It must be determined if enough of the substrate can be obtained to perform an entire screening campaign, which often involves a minimum of thousands of reactions. Furthermore, if the assay requires immobilization of the substrate, then the binding surface must have a high enough capacity to reach substrate $K_m$. It may be reasonable to use less than $K_m$ of the substrate if the $K_m$ is too high to reasonably achieve in the screen. However in this case, it is important to consider that lowering the substrate concentration will slow the reaction rate. This may mean that the reaction will have to be run for a longer amount of time to obtain a detectable signal to noise ratio. In this case, it is important that the enzyme be stable enough to survive the duration of the assay. Additionally, lowering the concentration of one substrate in a bisubstrate system causes the screen to have a greater bias for finding inhibitors that are competitive with that substrate as it could then be easier to compete off of the enzyme.

2.5.2 Buffer Considerations

It is important to remember that when transitioning from low throughput to high throughput assays there will potentially be a transition to more forms of solution transfer which may translate into a greater variety of contact surfaces for the enzyme/ reaction
mixture (tubing, transfer pins, etc.). In our experience working with the MYST protein hMOF, we have found that while we have had limited difficulty with protein stickiness in low throughput, we lose a significant portion of the protein through the course of transfer steps. To alleviate the effects of a much lower signal than expected when transitioning from low throughput to high throughput we tested the addition of bovine serum albumin (BSA) to our assay buffer, finding that it increased our signal by about 5 times.

Cysteine reactive compounds can lead to a high number of false hits in a screening campaign. In the case of MYST HATs, which have been demonstrated to employ an active site cysteine for acetyltransferase activity (Yan, Harper et al. 2002), this can be a particular problem. It may be appropriate to add some amount of reducing agent to the reaction buffer to limit the amount of inhibition due to compounds reacting with the active site cysteine. This however can lead to more problems as it has been shown that redox cycling, or production of peroxide caused by strong reducing agent interactions with some compounds, can also lead to a high false hit rate (Johnston, Soares et al. 2008). To circumvent this issue, we have included in some of our assays free cysteine in an attempt to outcompete cysteine reactive compounds.

Small amounts of non-ionic detergent can also help to decrease a false hit rate (Feng, Simeonov et al. 2007). This is especially true when screening MYST family acetyltransferases where screening at a higher compound concentration may lead to colloidal aggregation, which will be discussed in more detail in section 2.5.6.
2.5.3 Assays

In the case of the MYST acetyltransferases, we have developed a radioactivity based assay and an ELISA assay to use in high throughput. We have also developed a thermofluor assay to be used as a secondary orthogonal assay in low throughput.

2.5.3.1 Radioactivity Based Assay

Because of the sensitivity and robustness of radioactivity assays, we decided to adapt the radioactivity based P81 filter paper assay discussed in section 2.3.4 to be used in high throughput as a primary assay. The assay is performed by dispensing hMOF (or inactive mutant hMOF E350Q as a negative control) into each well of a 384 well plate followed by the addition of compound and then an H4 peptide/^{14}C-acetyl CoA solution to final concentrations of 50 nM hMOF, 50 µM^{14}C- acetyl CoA, and 400 µM H4 peptide (both substrates are used at ~K_m) in 40 mM Tris (pH 8.0), 100 mM NaCl, 800 µM Cys, and 7.5 µM BSA. The reaction is allowed to proceed for 1 hour (allowing for 40-60% substrate conversion, not performed at steady state) and then 3.7% HCl is added to each well to stop the reaction. The reaction is then transferred from the 384 well plate to four 96 well p81 paper embedded mesh bottom plates. A vacuum manifold is used to pull excess liquid from the wells. Next, each well is washed with 10 mM HEPES (pH 7.5), followed by addition of acetone, which is also vacuumed off. The plate bottoms are then sealed and scintillation fluid is added to each well. The plates are then sealed and read on a TopCount plate reader. The Z-factor for this assay has remained >0.6 for all optimized trials and we have confirmed the ability of the assay to detect inhibition by acetonyl CoA.
2.5.3.2 ELISA Assay

We also developed an ELISA assay, which is performed in automation by first binding H4-GST to a 384-well glutathione coated plate. We have found that hMOF acetylates H4-GST almost as well as H4 peptide, whereas GST-H4 barely becomes acetylated at all. 50 nM enzyme, followed by test compounds are added to the plate followed by 50 µM acetyl CoA. The reaction is allowed to react for 3 hours to obtain a suitable signal window. Primary anti-H4AcK antibody is added and allowed to bind for 30 minutes. Next, the secondary antibody conjugated to horseradish peroxidase (HRP) is added and allowed to bind for 30 minutes. Chemiluminescent substrate is added and is oxidized by HRP to produce light that can be detected on a plate reader. The plate is washed between each step. The Z-factor has remained reproducibly above 0.6. This assay provides a good example of how requiring a binding surface may limit how much substrate can be included in the assay, thus slowing reaction time. In the case of this assay, because screening could take multiple hours, the stability of the assay was tested as a function of time.

Another important aspect of this assay to note is that not all primary antibodies developed to recognize the acetylated substrate perform equally well. In the case of this assay, multiple antibodies were characterized via a dot blot as well as screening within the assay to ensure we could achieve the highest signal to noise ratio with the best performing antibody. In our experience, Histone H4 tetra-acetyl antibody (pAb) (Cat# 39179, 39228), and Histone H4ac (pan-acetyl) antibody (pAb) (Cat# 39243, 39244) from Active Motif have performed relatively well in our assay giving signal windows of approximately 10x and 4x, respectively.
**2.5.3.3 Thermofluor Binding Assay**

We have taken advantage of the thermofluor assay as a secondary orthogonal screen to perform in low throughput to validate compounds identified in the primary assay. We have also begun efforts to translate this assay from low throughput to high throughput for fragment screening. We currently perform the high throughput assay by adding 15 µL of 10 µM hMOF in 20 mM HEPES pH 7.5, 300 mM NaCl to each well, followed by 1 µL of 20 mM fragment compound to a final concentration of 1 mM compound (or 100 µM CoA as a positive control). After allowing the compounds to sit with the protein for 30 minutes, 4 µL of a 1:300 dilution of 5000x SYPRO® Orange dye is added to each well. The plate is then sealed and read on a real-time qPCR instrument. For this assay, we have had some difficulty with deviations in control T<sub>m</sub> and obtaining quality screening statistics. Further optimization will be required to determine if this assay will be suitable for high-throughput screening of MYST proteins.

**2.5.4 Eliminating Promiscuous Inhibitors**

After completing this screening campaign, it was important to focus on the highest quality hits that had the potential to be developed into probes or therapeutics. In general, a good inhibitor should display reproducible dose-dependent inhibition, have a Hill slope of approximately 1.0, bind reversibly to the protein (although covalent inhibitors can be beneficial under certain circumstances they will not be discussed here), and also be selective for the protein of interest. When a compound identified through screening does not meet these criteria, it becomes important to scrutinize it carefully to identify and eliminate compounds that are disrupting enzymatic activity through nonspecific
interactions. For example, through the course of this HTS campaign, hMOF was found to be particularly sensitive towards cysteine reactive compounds, mercury containing compounds, and copper containing compounds. As mentioned earlier, MYST proteins have an active site cysteine responsible for acetyl transfer, thus cysteine reactive compounds may nonspecifically form disulfide bonds to inhibit the protein. Similarly, mercury-containing compounds are also able to interact nonspecifically with the active site cysteine. Additionally, in our experience, copper-containing compounds are especially active against the hMOF. We hypothesize that this occurs because either the copper interferes with the active site cysteine, or otherwise may be displacing the zinc in the zinc finger or both. These three types of compounds are poor candidates for further optimization.

2.5.5 PAINS

Pan-assay interference compounds (PAINS) are a class of compounds that mimic genuine inhibitors by interfering with assay readouts or undergoing nonspecific reactions with proteins. About half of PAINS known today fall into just 16 structural categories (Baell and Walters 2014), and learning these structures can save a lot of time and effort. However, because about half of PAINS do not fit into these categories it is also prudent to check the literature for other reports of a selected compound hitting unrelated proteins as this could be a red flag (Baell and Walters 2014).

In our experience with hMOF screening, it is especially important to verify the identity and purity of the identified hits in the library. Compounds within a library can break down over time into reactive products (Baell and Walters 2014). Through the
course of our assays, we have found that hMOF had numerous hits where mass spec revealed a plethora of compounds were present in the sample when there should have only have been one. This leads us to believe that MYST proteins may be particularly sensitive to these sorts of reactive breakdown products.

Testing for the reversibility of a compound can be a fairly straightforward way to determine if it is interacting with the protein of interest through desirable interactions. If a compound is specifically binding a region of a protein in a dose-dependent manner, then the enzyme should regain activity when the compound is diluted (Copeland 2005). Briefly, to perform this experiment in the context of the MYST protein hMOF, 5 μM enzyme is first incubated with 10 x IC$_{50}$ of a compound for ~30 minutes. The solution is then diluted 100 x such that the enzyme concentration is at 50 nM and the compound is at 0.1 x IC$_{50}$, where little to no activity should be expected from an inhibitor. A reaction is then started by adding H4 peptide and acetyl CoA at K$_m$ and time points are taken every five minutes for an hour. As a positive control for reversibility, we measure enzyme activity with no inhibitor added (DMSO control) as well as enzyme activity with the reversible inhibitor acetyl CoA. As a negative control for reversibility, we used merbromin, a mercury-containing compound containing free mercury that interacts with the active site cysteine. If a compound reversibly binds to the enzyme then the activity of the enzyme should increase with time aligning with the positive controls, and if a compound is irreversible, the enzyme will remain inhibited and little activity will be seen over time, aligning with the negative control (Figure 10). Compounds that are found to be irreversible inhibitors may be nonspecifically modifying the protein of interest and are much less likely to serve as good candidates for optimization.
Figure 10. Sample data for an enzyme recovery assay with a MYST protein.

Note that both of the test compounds display irreversible inhibition, as they inhibit enzyme activity after dilution, aligning with the control irreversible inhibitor merbromin, a mercury containing compound that interacts with the active site cysteine. Also note that acetonyl CoA, a reversible inhibitor, aligns very closely to the DMSO control.
2.5.6 Colloidal Aggregators

Colloidal aggregators are compounds that aggregate and sequester a protein to limit its activity (McGovern, Helfand et al. 2003). This can be a particular problem when screening MYST proteins as compounds tend to be screened at higher inhibitor concentrations (low micromolar range) where they can have a higher tendency to aggregate. In general, if a compound is reversibly interacting with an enzyme in a 1:1 stoichiometry then the Hill slope for that compound with the enzyme should be ~1 (Feng, Simeonov et al. 2007), and if the Hill slope is higher than 1, this may indicate that the compound is acting through colloidal aggregate inhibition. Colloidal aggregates can be solubilized by the addition of small amounts of non ionic detergents (~0.01% Triton-X) (McGovern, Helfand et al. 2003). If running an activity assay in the presence of detergent causes the inhibitory activity of a compound to be lost or significantly decreased, this may be a sign that the compound is acting via aggregation rather than by a specific interaction. To avoid initial identification of colloidal aggregators in a primary screen it can be helpful to include small amounts of non-ionic detergent in the assay buffer. However, if detergent interferes with the assay readout of the primary screen, it is important to then include detergent in secondary screens to eliminate compounds acting through colloidal aggregation.
2.6 Conclusions

The MYST family proteins remain one of the more understudied families of the acetyltransferases. This is likely, at least in part, due to their weak affinity for their substrates, which can make them especially challenging to study. Furthermore, the MYST proteins appear to be only weakly inhibited by validated inhibitors, and significantly sensitive to nonspecific inhibitors further complicating efforts to further probe their activity and find new inhibitors.
CHAPTER 3

High Throughput Screening Results
3.1 Introduction

3.1.1 Preface

While the previous chapter sought to provide a general description of procedures and considerations from the HTS campaign, this chapter will delve into the specific results from the HTS including the compounds we investigated, how we discovered their shortcomings, and the hurdles we encountered. This chapter focuses on the results obtained from screening, rather than development or optimization steps, though some results that led to changes of assay conditions will be noted. This chapter provides the context upon which the previous chapter was constructed, and provides the specific data on which the method optimizations and considerations were based.

3.1.2 Overview

This chapter will describe the screening of over 150,000 small molecules and a variety of natural products from six different libraries using two different high throughput screening assays in a search for inhibitors of the MYST family acetyltransferase, hMOF. The first assay employed was an ELISA based assay described in section 2.5.3.2. Briefly this assay was performed in 384-well glutathione coated plates using a H4-GST substrate, a primary antibody that recognized H4 acetyl-lysine, and a luminescent read-out. Chronologically, we began our screening campaign with this assay as our primary assay, and successfully screened three libraries including the Spectrum Collection, the NIH Clinical Collection, and the ChemDiv Library. However, due to a number of technical issues including high false hit rate, and discontinuation of the glutathione plates with which we had optimized the assay, we switched our primary assay from the ELISA
to a radioactivity assay. The high throughput radioactivity assay is described in section 2.5.3.1. Briefly, this assay was performed using $^{14}$C-acetyl-CoA and an H4 peptide substrate. The reaction was carried out in 384-well plates before being transferred using automation to 96 well plates embedded with the H4-binding P81 filter paper.

The results in this chapter will be grouped into the stages of hit validation, even though chronologically, the results were validated following each library screening as a method of quality control. Therefore, the libraries that were screened with the ELISA assay had their initial results rescreened in the ELISA assay, and the libraries that were screened using the radioactive assay, had their initial hits rescreened in the radioactive assay. Because of nuances such as these, the results and discussion in this chapter are grouped into one section to allow for thorough explanations and data clarification.

3.1.3 Definitions

**Hit/Active Compound:** a compound/sample well in the screen that inhibits hMOF by three standard deviations below the mean signal.

**Matrix:** A library that is “matrixed” means that each sample well has a mixture of several compounds (generally 5-10) and each unique compound appears in the screen twice, but shares a well with two completely different sets of compounds. The data in a matrixed library can be deconvoluted by finding which compound is shared between two hit wells. Matrixing a library improves efficiency by increasing the number of compounds that can be screened over a period of time.
3.2 Materials and Methods

3.2.1 Protein Expression, Purification, and Quality Control

To purify hMOF, we followed previously described methods for purifying the MYST proteins (Yuan, Rossetto et al. 2012) but with some modifications, which have been specifically outlined in section 2.2 of this thesis. To ensure reproducibility of results throughout the screening campaign each new preparation of the hMOF protein was tested side by side with previous preparations using low throughput $^{14}$C-acetyl CoA based activity assays prior to use in a screen.

The recombinant histone H4 (residues 1-34) with a C-terminal GST tag in pET28A was expressed and purified as follows. The plasmid was transformed into E. coli BL21(DE3) codon plus RIL (Stratagene) cells and the protein was overexpressed in LB media, induced with 1mM IPTG at OD 0.5-0.8, and grown at 18°C overnight. The cells were harvested and lysed by sonication in lysis buffer (100 mM Tris (pH 8.0), 0.5 M NaCl, 0.2mM EDTA, 1% Triton-X, 5 mM DTT, and 5% glycerol) with 1 mM phenylmethylsulfonyl fluoride followed by centrifugation at 28,000 g for 30 min at 4 °C. The lysate supernatant was then incubated with GST resin (Gold Bio) equilibrated with lysis buffer for 3 hr at 4°C. The resin was washed with 10 column volumes of lysis buffer. The protein was eluted with lysis buffer containing 25 mM reduced glutathione in 10 mL fractions. The eluent was dialyzed overnight at 4°C to eliminate free glutathione. Then the H4-GST protein was concentrated, aliquotted, flash frozen and stored at -80 °C.
3.2.2 High Throughput Assays

Two different high throughput screens (HTS) were employed during the course of this screening campaign. The first assay is the ELISA based HTS, described in detail in section 2.5.3.2. The second assay is the radioactivity based assay, described in detail in section 2.5.3.1.

3.2.3 Low Throughput Assays

3.2.3.1 Radioactivity Based Assay

The low throughput radioactivity based assay was performed using the same parameters as the HTS radioactivity assay, but with slightly different protocols. Briefly, for the low throughput radioactivity assay, hMOF was mixed at 50 nM final concentration with a reaction mixture containing 50 µM \(^{14}\text{C}-\text{acetyl CoA}\), and 400 µM H4 peptide in the presence or absence of a test compound. Reaction components were diluted in a buffer containing 40 mM Tris (pH 8.0), 100 mM NaCl, 800 µM Cys, and 7.5 µM BSA to a final volume of 50 µL. The reaction was allowed to proceed for 1 hour to allow for 40-60% substrate conversion and then stopped by spotting 20 µL each of the reaction onto two P81 filter papers. The filter papers were then washed 3x in 20 mM HEPES (pH 7.5), dipped in acetone and dried. Each filter paper was put into a scintillation vial with scintillation fluid and read on a scintillation counter.

3.2.3.2 Thermofluor Assay

The low throughput thermofluor assay employing the environment sensitive dye, Sypro Orange, is described in sections 2.4 and 4.2.3.
3.2.4 Screening Libraries

3.2.4.1 Spectrum Collection

The Spectrum Collection from MicroSource Discovery Systems, Inc. is a collection of 2560 compounds comprised of 4 smaller screens: the US Drug Collection, the International Drug Collection, the Natural Products Collection, and the Discover Collection. Thus, the Spectrum Collection contains a variety of compounds including those that have reached clinical trials in the United States, or have been marketed abroad in Europe or Asia, as well as purified natural product compounds from animals, plants, microbes, and bioactive compounds identified in peer-review publications that have never been developed into drugs. This library was screened in both the ELISA formatted HTS and the radioactive p81 filter paper based HTS using a final compound concentration of 25 µM.

3.2.4.2 NIH Clinical Collection

The NIH Clinical Collection (NCC) is a compound library assembled by the Molecular Roadmap Initiative at the NIH and is comprised of ~450 small molecules that have been used in clinical trials. Because this library contains so few compounds it was combined with the Spectrum Collection for our screens. This library was screened in both the ELISA formatted HTS and the radioactive HTS using a final compound concentration of 25 µM.

3.2.4.3 ChemDiv Library

The ChemDiv Library is composed of 30,000 compounds from The Discovery Chemistry (DS) set and 20,000 compounds from the Target Specific Chemistry (TS) set.
from ChemDiv, Inc. Together these compounds represent drug-like small molecules, some of which have been shown to have specificity for certain targets including G-coupled protein receptors, kinases, ion channels, proteases, and nuclear hormone receptors. The ChemDiv library was matrixed with five different compounds per sample well with each compound being present in two wells, and was screened with the ELISA HTS assay using a final concentration of 20 µM per compound, and thus 100 µM total compound per well.

3.2.4.4 The Maybridge “Hitfinder” Library

The Hitfinder Library from Maybridge is comprised of 14,400 compounds. These compounds all follow the Lipinski rules for drug-likeness, meaning that they have clogP \( \leq 5 \), hydrogen bond acceptors \( \leq 10 \), hydrogen bond donors \( \leq 5 \), and molecular weight \( \leq 500 \). Furthermore, to minimize false hits, the compounds in the Hitfinder Library have been selected to be non-reactive. This library was screened using the radioactivity p81 filter paper based HTS with a 25 µM final compound concentration.

3.2.4.5 The Natural Product Discovery Institute Library

The Natural Product Discovery Institute (NPDI) collection consists of plant, fungal, and actinomycetous extracts, where multiple rounds of screening and fractionation are performed in collaboration with scientists at NPDI. After each round of screening any extract that shows inhibitory activity against the protein of interest is further fractionated and those fractions are then rescreened and re-fractionated until a
single compound with inhibitory properties is identified. This library was screened using the radioactive HTS.

3.2.4.6- The Lankenau Chemical Genomics Center Library

The Lankenau Chemical Genomics Center (LCGC) Library consists of approximately 120,000 drug-like compounds that follow the Lipinski rules. Any hits obtained from this screen can be obtained in powder form from LCGC for further testing. The entirety of the LCGC library was matrixed with 10 compounds per well, and each compound in two wells, with the final concentration of each compound at 25 µM and was screened with the radioactive HTS.

3.2.5 hMOF Crystallization with Merbromin

hMOF crystals were obtained using the hanging drop method at protein concentrations between 6 and 8 mg/mL, mixed 1:1 with a reservoir buffer composed of 0.2 M ammonium acetate, 28% PEG 3350, and 0.1 M BisTris pH 6.5 and incubated over reservoir buffer. After crystals formed, they were soaked in their well solution plus 10 mM merbromin and incubated over reservoir buffer for 2 days. The crystals were cryoprotected by a short soak in well solution plus 17% ethylene glycol, followed by flash freezing in liquid nitrogen. Data was collected at BNL beamline X-29A. Diffraction data was processed using HKL2000 (Otwinowski and Minor 1997) and the structure was determined using molecular replacement with Phaser (McCoy, Grosse-Kunstleve et al. 2007). Manual model building was carried out using Coot (Emsley and Cowtan 2004)
and refinement was performed using Phenix (Adams, Afonine et al. 2010). Figures were made using PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre Schrödinger, LLC.)

3.2.6 Mechanism of Action (MOA) Assay

Mechanism of action assays were performed using the low throughput radioactivity assay described in section 3.2.3.1. MOA assays were carried out following the same procedures as for Michaelis-Menten curves for determining the $K_m$ and $k_{cat}$ of an enzyme (described in section 2.3.4), the only difference is that they were performed multiple times in the presence of a varying concentrations of compound. Briefly the specific parameters for the data presented in this chapter are given here. We first mixed hMOF at 50 nM with H4 peptide in a buffer containing 40 mM Tris (pH 8.0), 100 mM NaCl, 800 µM Cys, and 7.5 µM BSA. Depending on which substrate was being tested, H4 was either varied from 50 µM – 4 mM or saturating (4 mM). We began the assay by adding $^{14}$C-acetyl CoA which was varied from 15 µM - 500 µM or saturating (500 µM) in a final volume of 50 µL. The reaction proceeded for eight minutes to maintain steady state conditions, and was repeated with 10 different concentrations of compound 115772 ranging from 0 µM to 200 µM for each different substrate concentration.
3.3 Results and Discussion

3.3.1 Initial Validation of Primary Screening Hits

Our initial small molecule screening campaign identified 281 active compounds that appeared to inhibit hMOF. These included 40 hits from the Spectrum/ NCC Library, 17 hits from the ChemDiv Library, and 224 hits from the LCGC library for an overall hit rate of 0.17%. We first retested the hits using the primary screening assay from which they were identified. Approximately 75% of the initial hit compounds showed no inhibitory activity against hMOF upon retest and were eliminated from our pool of potential lead compounds. The eliminated compounds included 24 hits from the Spectrum/ NCC library, all of the compounds from the ChemDiv Library and 172 compounds from the LCGC library. Because the LCGC library was the last library we screened, as well as the largest library, the initial validation steps were more intense than other screens and are described in detail in section 3.3.5.

The NPDI Library screening failed to identify any lead compounds for further testing. After several rounds of screening and fractionation, the only active compounds identified were tannins. Tannins are reviewed in (Khanbabaee and van Ree 2001), and the limited utility of tannins to be developed as inhibitors is discussed in (Beutler 2009). Briefly, tannins are polyphenols found in high concentrations throughout all parts of a plant. The poor prospects for development of tannins is rooted in their mechanism of action, which is to non-covalently bind various proteins (Beutler 2009). Because of this, there is little hope for tannins to be developed to specifically inhibit a protein of interest, leading us to conclude our natural product screening with no pursuable results.
3.3.4 Cysteine Reactive Compounds

Of the compounds that were confirmed from the second round of testing, 16 were from the Spectrum Collection. These 16 compounds were next tested in a concentration series using a secondary low throughput radioactivity assay resulting in only two compounds being confirmed to inhibit hMOF: cystamine dichloride and merbromin (Figure 11). Looking at the structures of these compounds, some chemical details were immediately concerning regarding the specificity of these compounds as inhibitors. As discussed throughout this thesis, hMOF relies on an active site cysteine for activity, and both compounds identified contained moieties that could interfere in some way with that cysteine. In the case of cystamine dichloride, the disulfide bond was particularly disconcerting. Presumably the disulfide bond could cleave and form covalent adducts with the active site cysteine of hMOF to inhibit the enzyme with very few functional groups on the compound that would allow for the interactions with the active site to be specific. With merbromin, the concern was that the mercury atom itself could react with the hMOF active site cysteine. However, because the mercury atom was just a small part of this compound, we continued our study of merbromin, hoping that the overall structure of the compound, and not just the mercury itself, would contribute to specific hMOF inhibition.

To further investigate the mechanism of inhibition of merbromin with hMOF, we grew hMOF crystals and soaked the crystals with merbromin. We were able to collect a data set on an hMOF crystal soaked with 10 mM of merbromin for two days (Figure 12). What we immediately noticed from the Fo−Fc electron density maps from these crystals was large bulbous density around cysteines even at high σ values implying that a heavy metal was present at these locations. What was more important was that at lower σ
Figure 11. Chemical structures of the two cysteine-reactive hits, cystamine dichloride and merbromin, obtained from the high throughput screen of the Spectrum Collection.
Figure 12. Fo-Fc density for hMOF soaked with merbromin. Fo-Fc density (green) contoured at 5σ (inset 2.5σ) for hMOF (magenta) soaked with 10 mM of merbromin for 2 days. The active site cysteine (C316) is colored cyan and shown as sticks. Other cysteines with visible Fo-Fc density at 5σ are shown as magenta sticks. Note in the inset that even at 2.5σ only density for a single metal on the cysteine is visible with no apparent density for the merbromin compound as a whole.
values no further density was revealed in the Fo–Fc maps surrounding the metal, implying that the whole compound was not present, only mercury. We concluded from this data that there were likely mercury contaminants present in our compound sample that could disable the protein with no contribution from the overall structure of the merbromin compound, thus making the merbromin compound unsuitable for further development as an hMOF inhibitor.

This data led us to include 800 µM free cysteine in future assay buffers in an attempt to outcompete hMOF for cysteine reactive compounds, as discussed in the previous chapter section 2.5.4. Furthermore, this data also warned us to be particularly wary of compounds that would react with cysteine, which we would come across repeatedly throughout the screening process, even in the presence of free cysteine in the reaction buffer.

3.3.5 Confirmation of LCGC Active Compounds

3.3.5.1 Details on Initial Validation of Primary Screening Hits from LCGC

Similar to all other libraries, to validate the hits from the LCGC library we retested the compounds in the format of the radioactive high throughput screen which eliminated 172 of the initial hits. Based on our screening compound concentration (25 µM) we suspected that some compounds might inhibit via a mechanism of colloidal aggregation (discussed in section 2.5.6). To eliminate these compounds we rescreened all 224 initial hits from the LCGC screen in the presence and absence of 0.05% tween 20 (Figure 13). A majority of the hits that maintained activity in the presence of a small amount of non-
Figure 13. Data comparing LCGC hits with and without 0.05% tween 20 Data shows the 224 initial hits from the LCGC library as well as positive controls (Columns 2 and 24), negative controls (Columns 1 and 23), and acetonyl CoA dose response control (blue box) in the presence (right) and absence (left) of 0.05% tween 20. The nine compounds selected for further characterization are indicated in red circles. Note that many hits lose activity in the presence of 0.05% tween 20. Compounds that do not lose greater than 10% of their activity and are not selected for further testing are copper salts.
ionic detergent were copper salts, and were eliminated from the list of potential inhibitors due to their ability to disable the protein by binding the active site cysteine. We opted to continue the testing with nine compounds that did not contain copper and varied in their inhibitory activity by less than 10% in the presence or absence of tween 20.

We then further validated that these compounds were not acting through colloidal aggregation by obtaining their powders from LCGC and performing IC$_{50}$ studies using the low throughput radioactivity assay in the presence and absence of tween 20 to confirm a consistent IC$_{50}$ and a Hill slope of $\sim$1. As described in section 2.5.6, one way to ensure that a given compound is not acting through colloidal inhibition is to ensure that addition of a small amount of non-ionic detergent, which would disrupt colloidal aggregates, does not disrupt the inhibitory properties of a compound. Another method which goes hand-in-hand with the IC$_{50}$ test is Hill slope analysis, where a Hill slope of $\sim$1 implies non-cooperative binding of inhibitor to protein. Of the nine compounds that were tested, four compounds failed to show inhibitory activity, and were eliminated. We continued moving forward with the remaining five compounds following this round of testing which showed reproducible IC$_{50}$ values and low Hill slopes (Figure 14, Table 2). It is worth noting that at this point we noticed that most of our Hill slope values were greater than 1, which considering that they were unaffected by tween 20, could be a sign that they were irreversible inhibitors. The reversibility of the top compounds will be addressed in section 3.3.5.3.
Figure 14. Structures and IC\textsubscript{50} values for five hits from the LCGC library in the presence and absence of non-ionic detergent. Column 1: Chemical structures of the five hits from the LCGC library. Column 2: IC\textsubscript{50} curves for these compounds in the absence of 0.05% tween 20. Column 3: IC\textsubscript{50} curves for these compounds in the presence of 0.05% tween 20. Table 2 gives IC\textsubscript{50} calculated values and Hill slopes from these graphs.
Table 2: IC$_{50}$ values and Hill slopes calculated from the data presented in Figure 14 compared to values calculated for acetonyl CoA (data not shown). Error represents the standard deviation calculated from duplicate trials.

<table>
<thead>
<tr>
<th>IC$_{50}$ (µM)</th>
<th>Hill Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no tween 20</td>
</tr>
<tr>
<td>14855</td>
<td>14.7 ± 0.9</td>
</tr>
<tr>
<td>115772</td>
<td>20 ± 2</td>
</tr>
<tr>
<td>128041</td>
<td>11.6 ± 0.8</td>
</tr>
<tr>
<td>24727</td>
<td>100 ± 30</td>
</tr>
<tr>
<td>15908</td>
<td>12.1 ± 0.8</td>
</tr>
<tr>
<td>Acetonyl CoA</td>
<td>20 ± 2</td>
</tr>
</tbody>
</table>
3.3.5.2 Thermofluor Assay Eliminated Two More LCGC Hits.

Initially, we had tested all nine of the top LCGC compounds at 100 µM in the thermofluor assay and observed that none of the compounds significantly increased or decreased the thermal stability of hMOF. Now that we confirmed five compounds that had reproducible IC<sub>50</sub> values, we proceeded to retest the effects that these compounds had on the thermal stability of hMOF at higher concentrations using the thermofluor assay. Each compound was tested at 200 µM and 400 µM. Compound 24727 destabilized hMOF by 6°C at 200 µM and compound 15908 had no effect on hMOF at 200 µM and destabilized hMOF by 0.75°C at 400 µM leading to the elimination of these compounds. Compounds 14855, 128041, and 115772 displayed stabilizing effects on hMOF (Figure 15). Compound 14855 had the least effect on hMOF, increasing the apparent melting temperature (T<sub>m</sub>) by only 0.6 °C and 0.9 °C for the respective concentrations. Compound 128041 saw the greatest shift at 200 µM increasing the T<sub>m</sub> by 7.7 °C. However, this result was then complicated by a loss of 1.6 °C in the presence of 400 µM of compound 128041. We suspected that at higher concentrations this compound may start to exhibit poor behavior such as aggregation or poor solubility and did not eliminate it from the study. The best overall behaving compound in the thermofluor assay was 115772 which showed a consistent concentration correlated increase of T<sub>m</sub> with a 1.6 °C increase at 200 µM and a 3.5 °C increase at 400 µM.
Figure 15. The top three LCGC hits stabilize hMOF in a theromfluor assay. Normalized thermal denaturation curves for hMOF presence of 200 µM and 400 µM compound or CoA are shown. SYPRO Orange fluorescence is reported, and the inflection point for each curve is the melting temperature for the protein in the given conditions.
3.3.5.3 LCGC Hits are Irreversible Inhibitors

Because of the stabilization effects observed in our thermofluor data, we proceeded to perform a kinetic mechanism of action (MOA) study starting with compound 115772. However, in the MOA assay, no hMOF inhibition was observed at any concentration of compound 115772 (Figure 16).

A major difference between the IC<sub>50</sub> assays and the MOA assays are that the MOA are performed under steady state conditions, while the IC<sub>50</sub> assays are not. Therefore, while the overall procedures are very similar, the reaction time for the IC<sub>50</sub> assay is 1h (taking the reaction to ~40% completion so that suitable signal can be observed at substrate K<sub>m</sub>), and the MOA reaction time is only 8 min (to assure that reaction is linear and that less than 10% of substrates have been consumed). This data led us to start thinking about on and off rates, and what might contribute to inhibition which occurs in 1 hour but not in 8 minutes. This data also led us decide to test the reversibility of these compounds, which the slightly elevated Hill slope values indicated could be irreversible.

Along these lines, we performed the enzyme recovery assay, which is fully described in section 2.5.5. Briefly, a 100 x enzyme solution is incubated with a compound at 10 x IC<sub>50</sub> concentration. Then, before performing the reaction, the incubated solution is diluted 100x such that the enzyme is at 1 x and the inhibitor is 10 x below the IC<sub>50</sub>. If inhibition is still observed, it indicates that the compound inhibits irreversibly. For our purposes we wanted to further explore how our compound was inhibiting hMOF on longer time scales, if it did not show inhibition in kinetic assays. This experiment revealed that compounds 115772, 128041, and 14855 were all irreversible inhibitors (Figure 17).
Figure 16. Attempted Mechanism of Action (MOA) assays using compound 115772. The legend indicates compound concentration. In this assay we did not observe any inhibition.
Figure 17. **Data for enzyme recovery assays.** Enzyme recovery assays with hMOF and the top 3 LCGC compounds using merbromin as a control for irreversible inhibition and acetonyl CoA as a control inhibitor for reversible inhibition. Note that all three of the LCGC compounds display behaviors of irreversible inhibitors.
While irreversible inhibitors do not tend to be ideal, they are not necessarily useless. Irreversible inhibitors can be useful if they can be made to be specific for a given target, and specificity may arise from chemical interactions between the compound and the target in areas surrounding the covalent interaction (Copeland 2005). Because of this reason we decided to proceed with testing these compounds.

3.3.6 Resynthesis and Lack of Validation of LCGC Top Hits

Because we had a finite amount of our compound stocks from LCGC for testing, we collaborated with Michelle Estrada in the laboratory of Jeff Winkler (Chemistry Department, University of Pennsylvania) to synthesize our remaining compounds so that we could study them further. Compound 115772 was synthesized first, however, the new compound did not reproduce the results of the original. In a dose response trial this newly synthesized compound did not show any inhibition against hMOF, and in a thermofluor assay this compound was destabilizing at both 200 µM and 400 µM (Figure 18).

Similarly, the newly synthesized compound 128041 did not behave the same as the initial compound. In fact, the newly synthesized 128041 could not be tested in any meaningful way because it appeared to be totally insoluble in DMSO. These results led to the obvious question of the integrity of the samples that we had previously been working with, and as such, we sent the original compound samples of 115772, 128041, and 14855 from LCGC for mass spectrometry analysis.

The LCMS chromatograms of compounds 115772, 128041, and 14855 can be seen in Figure 19, Figure 20 and Figure 21, respectively, and reveal not only an overall
Figure 18: Resynthesis of compound 115772 failed to reproduce activity. Data from the resynthesized compound 115772. Top. Compound 115772 failed to inhibit hMOF in a concentration series. Bottom. Thermal denaturation of newly synthesized compound 115772 reveals that it destabilizes hMOF. Addition of varying amounts of CoA shown as a control.
Figure 19: Mass spectrometry chromatogram for compound 115772. Compound structure and molecular weight of compound 115772 is shown in the upper left. Structure and molecular weights of hypothesized break down products shown in left inset.
Figure 20: Mass spectrometry chromatogram for compound 128041. Compound structure and molecular weight of compound 128041 is shown in the upper left. Structure and molecular weight of hypothesized break down product shown in left inset.
**Figure 21**: Mass spectrometry chromatogram for compound 14855. Compound structure and molecular weight of compound 14855 is shown in the upper left.
noisy mixture of compound masses in the sample but also a disappointing lack of a peak for the alleged compounds that they each were supposed to contain. The mass spectrometry chromatogram for compound 115772 (theoretical M.W. 335) was especially noisy, with major peaks that were difficult to assign based on the structure and molecular weight of the parent compound. We considered that the azo bond in compound 115772 may cleave to give two compounds shown in (Figure 22, top) with molecular weights of 166 and 173, however peaks representing these masses were not observed in the chromatogram either. The chromatogram for compound 14855 (theoretical M.W 234) contained all major peaks with molecular weights higher than the molecular weight of the compound itself and could not be assigned. The chromatogram for compound 128041 (theoretical M.W 280,) however had a major peak at MW ~267 which our collaborators hypothesized could possibly be a breakdown product of the parent compound, and also synthesized a related compound with similar functionality (Figure 22, middle). Michelle Estrada synthesized these two compounds, however, they also did not inhibit hMOF when tested in a concentration series using the radioactive low throughput assay (Figure 22, bottom). At this point we were left with compounds of unknown structure for which we did not have enough remaining sample for NMR characterization. As a conclusion, we unfortunately had no choice but to drop the three compounds as chemical structures could not be confirmed, or even identified, even though they may have been potential true inhibitors.
Figure 22: Resynthesis Details. Top: Proposed break down product of compound 128041 (A) based on the mass spectrometry chromatogram and related compound with similar functionality (B). Bottom: Neither compound A nor B inhibited hMOF in a concentration series.
3.4 Future Directions

Many of the lessons learned from this screening campaign are outlined, albeit somewhat indirectly, in Chapter 2. Specifically, not confirming compound structures early was one of the biggest flaws of this study. We confirmed that the compounds were not merely misassigned by both confirming the accuracy of our library keys against those of the scientists at LCGC, as well as mass confirmation of other compounds in the library that were not hits in our screen (Figure 23). As a result of delaying the direct identification of our hits, much time and effort was spent characterizing unknown compounds that could potentially be either specific inhibitors or reactive breakdown products.

Additionally, hMOF, or the MYST family in general, may just be difficult to target. Similar to our own experience a paper reported in 2011 a high throughput screen against the MYST protein, MOZ (Falk, Connor et al. 2011). In this study, ~243,000 lead-like compounds were screened yielding ~1400 initial hits, and only a single scaffold met the criteria of subsequent testing. A shortcoming of this study was that it never revealed the chemical structure of the one successful scaffold, and no follow-up publications have occurred to allow us to fully understand how successful this inhibitor actually was.

New approaches may be needed to successfully target hMOF or the MYST family in general. In future screens in the laboratory, higher quality screening libraries will be obtained, and lead compound structures will be confirmed before delving into characterization. Also, fragment screening is an attractive option. We have already been able to confirm that hMOF binds salicylic acid in the AcCoA binding pocket, through co-
Figure 23: Mass spectrometry chromatogram for another compound from the LCGC Library, showing that other wells contained the compounds that they were assigned to contain with good purity, confirming the problem was not that misassignment of compounds.
crystallization experiments (Figure 24). If we could find other fragment compounds which bind near the salicylate molecule we could design inhibitors based on the combined structures of the fragments. Along these same lines we are currently collaborating with Tyler Higgins of the Winkler Laboratory to design salicylic acid based inhibitors based on the combined structural data of how salicylic acid binds hMOF, along with the specific contacts in the region that are made by acetyl-CoA. Further future directions will be discussed in Chapter 5.
Figure 24: Salicylic acid (pink) bound to hMOF (cyan). hMOF is shown in cyan with C316 and E350 shown in sticks. Salicylic acid binds hMOF in the AcCoA binding pocket.
CHAPTER 4

*Structural and Functional Role of Acetyltransferase hMOF K274 Autoacetylation*

This research was originally published in the *Journal of Biological Chemistry*, (McCullough, Cheryl E.; Song, Shufei; Shin, Michael H.; Johnson, F. Brad; Marmorstein, Ronen. *Structural and functional role of acetyltransferase hMOF K274 acetylation*. J Biol. Chem. 2016 Aug 26) © the American Society for Biochemistry and Molecular Biology.
4.1 Introduction

Because many histone acetyltransferases undergo autoacetylation to potentiate cognate substrate lysine acetylation, and the mode and molecular role of such autoacetylation is poorly understood, we sought to perform an in depth molecular investigation of the mode of hMOF K274 acetylation. A mutational scan of hMOF K274 reveals that all amino acid substitutions of this residue are able to bind cofactor but are significantly destabilized, both in vitro and in cells, and are catalytically inactive for cognate histone H4 peptide lysine acetylation. The X-ray crystal structure of a hMOF K274P mutant suggests that the reduced stability and catalytic activity stems from a disordering of the residue 274-harboring α2-β7 loop. We also provide structural evidence that a C316S/E350Q mutant, which is defective for cognate substrate lysine acetylation; and biochemical evidence that a K268M mutant, that is defective for K274 chemical acetylation in the context of a K274-peptide, can still undergo quantitative K274 autoacetylation. Together, these studies point to the critical and specific role of hMOF K274 autoacetylation in hMOF stability and cognate substrate acetylation and argues that binding of Ac-CoA to hMOF likely drives K274 autoacetylation for subsequent cognate substrate acetylation.
4.2 Experimental Procedures

4.2.1 Protein Expression and Purification

The recombinant hMOF HAT domain (residues 174-449) with an N-terminal 6xHis tag separated by a TEV protease recognition site was expressed and purified essentially as previously described with some modifications (Yuan, Rossetto et al. 2012). Briefly, the plasmid was transformed into E. coli BL21(DE3) codon plus RIL (Stratagene) cells and the protein was overexpressed in TB media, induced with 1mM IPTG at OD 0.8-1.5, and grown at 15°C overnight. The cells were harvested and lysed by sonication in lysis buffer (50 mM HEPES (pH 7.5), 0.5 M NaCl, 2 mM β-mercaptoethanol, 5 mM imidazole, 5% glycerol, 0.1% CHAPS and 1 mM phenylmethylsulfonyl fluoride) followed by centrifugation at 28,000g for 30 min at 4 °C. The lysate supernatant was loaded onto Nickel charged NTA resin (Qiagen) equilibrated with lysis buffer. The resin was washed with 10 column volumes of lysis buffer containing 25 mM imidazole. The protein was eluted with lysis buffer containing 250 mM imidazole buffer in 10 mL fractions. The eluent was concentrated and injected onto an FPLC using a HiLoad Superdex 75 16/60 gel filtration column using sizing buffer (20 mM HEPES (pH 7.5), 0.5 M NaCl). All hMOF mutants were created using the QuikChange protocol from Stratagene. hMOF mutants were expressed and purified using the same protocols as the wild type hMOF protein described above.

The recombinant MOZ HAT domain (residues 505-784) with an N-terminal 6xHis tag separated by a TEV protease recognition site was expressed and purified as follows. The plasmid was transformed into E. coli BL21(DE3) codon plus RIL (Stratagene) cells and the protein was overexpressed in LB media, induced with 1mM IPTG at OD 0.6, and
grown at 15°C overnight. The cells were harvested and lysed by sonication in lysis buffer (50 mM Tris-HCl (pH 8.0), 0.5 M NaCl, 10mM β-mercaptoethanol, 5 mM imidazole, 5% glycerol, 0.01% Triton-X and 1 mM phenylmethylsulfonyl fluoride) followed by centrifugation at 28,000g for 30 min at 4 °C. The lysate supernatant was then loaded onto Nickel charged NTA resin (Qiagen) equilibrated with lysis buffer and the resin was washed with 10 column volumes of lysis buffer containing 25 mM imidazole. The protein was then eluted with lysis buffer containing 250 mM imidazole buffer in 10 mL fractions. The MOZ His tag was cleaved using a 6xHis-tagged TEV protease and the MOZ protein was dialyzed overnight into low salt ion exchange buffer (40mM Tris-HCl, pH 7.0, 50 mM NaCl, 1 mM DTT). The protein solution was then passed over another nickel column to remove TEV protease and uncleaved MOZ. The resin was washed with approximately three column volumes of low salt ion exchange buffer containing 20 mM imidazole. Flow through and washes were pooled and loaded onto a 1 ml HiTrap SP cation exchange column (GE Healthcare). The protein was eluted in ion exchange buffer with a salt gradient (50–1000mM NaCl). Peak fractions were concentrated and run on a HiLoad Superdex 75 16/60 gel filtration column (GE Healthcare) in sizing buffer (20 mM sodium citrate, pH 5.5, 100 mM NaCl, 1 mM DTT). All MOZ mutants were created using the QuikChange protocol from Stratagene. MOZ mutants were expressed and purified using the same protocols as the wild type MOZ protein described above.

4.2.2 Enzyme Activity Assays

For comparing the activity of hMOF mutants (or MOZ mutants) to wild type, a radioactive-based enzymatic assay was employed as previously described (Yan, Harper
et al. 2002). Briefly, 50 nM hMOF or 200 nM MOZ was incubated with 400 µM of substrate peptide (H41-19 peptide for hMOF and H31-19 for MOZ) (GenScript) and 50 µM 14C-acetyl CoA (50-60 mCi/ mmol Moravek) in reaction buffer (100 mM acetate, 50 mM Bis Tris, 50 mM Tris at various pH values 100 mM NaCl, 800 µM cysteine and 0.25mg/mL BSA for hMOF and 40 mM Tris–HCl, pH 8.0, 100 mM NaCl, 1 mM DTT for MOZ ) at a final volume of 50 µL and incubated at room temperature for 1 hour to allow for 40-60% substrate conversion. To quench the reaction, 20 µL of the reaction mixture was added to negatively charged P81 paper (Millipore). For kinetic studies of wild type and K268M hMOF a saturating amount (500 µM) of radiolabeled 14C-acetyl CoA and varying concentrations of the substrate peptide (50–3000 µM) or a saturating amount of substrate peptide (3000 µM) with varying concentrations of 14C-acetyl CoA (15-500 µM) were incubated with 50 nM hMOF (WT or K268M) in 50 µL of reaction volume for 8 min at room temperature. Each reaction was performed in at least octuplicate. The counts per minute (CPM) were converted to molar units using a standard curve of known 14C-acetyl CoA concentrations and values were calculated using GraphPad Prism version 5.0a, GraphPad Software, La Jolla California USA, www.graphpad.com. For both types of radioactivity assays, filter papers were washed three times with 10 mM HEPES (pH 7.5) buffer, dried with acetone, and added to 4 mL of scintillation fluid. Signal was measured with a Packard Tri-Carb 1500 liquid scintillation analyzer.

4.2.3 Thermal Denaturation Assay

Enzyme solutions at 1 µM in 40 mM Tris-HCl, pH 8.0 and 100 mM NaCl were incubated in the presence or absence of 100 µM CoA for 30 minutes followed by

91
addition of a 1:300 dilution of a 5000x concentrate of SYPRO® Orange (Molecular Probes by Life Technologies) at a final volume of 20 µL in a MicroAmp Optical 384-well reaction plate (Applied Biosystems) sealed with Optical Adhesive Covers (Applied Biosystems). Samples were read using a 7900HT Real-Time PCR system (Applied Biosystems) measuring from 20°C to 95°C using a 1% ramp rate. The dye was excited at 490 nm and the emission light was recorded at 575 nm.

4.2.4 Crystal Structure Determination of MYST HAT Domains

All protein crystals were obtained using the hanging drop method at protein concentrations between 6 and 8 mg/mL. Crystals of hMOF-K274P were obtained by first concentrating the protein in the presence of 100 µM MG 149 (Selleck Chemicals), then incubated with a 5 x molar excess of MG 149 before mixing 1:1 with a reservoir buffer composed of 0.28 M Ammonium Chloride, 26%PEG 3350, and 0.1M BisTris pH 6.5 and incubated over reservoir buffer. Crystals of hMOF-C316S/E350Q were obtained by incubating protein with a 5 x molar excess of MG 149 before mixing 1:1 with a reservoir buffer composed of 0.2 M Ammonium Chloride, 28%PEG 3350, and 0.1M BisTris pH 6.5. Despite the presence of MG 149 in the crystal preparatory conditions, no density for this molecule was observed in the final electron density maps.

All native data sets were collected at APS 23 ID-B. Diffraction data were processed using HKL2000 (Otwinowski and Minor 1997) for hMOF-K274P and XDS (Kabsch 2010) for hMOF-C316S/E350Q and the structures were determined using molecular replacement with Phaser (McCoy, Grosse-Kunstleve et al. 2007). Iterative cycles of manual model building were carried out using Coot (Emsley and Cowtan 2004)
and refinement was performed using Phenix (Adams, Afonine et al. 2010). Figures were made using PyMOL (The PyMOL Molecular Graphics System, Version 1.2r3pre Schrödinger, LLC.)

4.2.5 Cell Studies

DNA fragments containing V5/His-hMOF and V5/His-hMOF(K274R) were obtained from pcDNA4-based plasmids (Yuan, Rossetto et al. 2012) and were cloned into vector pLU-TREmin-MCS-pPURO using restriction sites Sall and Agel. The rtTA3 gene was cloned into the Sall and BamHI sites of pLU-EF1-MCS-BLAST. Sequential transfections of pLU-EF1-rtTA3-BLAST plasmids and V5/His-hMOF plasmids into 293T cells were performed using Lipofectamine 2000 (Invitrogen), according to the manufacturer's guidelines. 10 µg/mL blasticidin and 1.5 µg/mL puromycin were used for selection of successfully transfected 293T cells. 293T cells were maintained in DMEM (with pyruvate, Mediatech) and 10% FBS at 37°C, under 6% CO2.

V5/His-hMOF and V5/His-hMOF (K274R) levels were assessed by western blotting of lysates. Western detection of actin in a separate region of the same blot was used as a loading control. The antibodies used were as follows: Anti-V5, P/N 46-0705, Lot: 1707842 from Invitrogen, (1:2500), anti-actin (I-19), sc-1616, Lot: L1108, (1:2500), and donkey anti-goat IgG HRP: sc-2020, Lot: G1615, (1:3000) from Santa Cruz Biotech, and rabbit anti-mouse IgG HRP: ab97046, Lot: GR220045-5, (1:3000) from Abcam.
4.2.6 Western Blots for Acetyl-Lysine

K274 acetylation levels of hMOF C316S/E350Q, hMOF K274A, hMOF WT and hMOF K268M were assessed by western blotting of purified proteins. The antibodies used were as follows: Acetylated-Lysine (Ac-K-103) Mouse mAb, 9681S, Lot: 11 from Cell Signaling Technology, (1:1000), and Goat Anti-Mouse IgG (H+L)-HRP Conjugate: L1721011, Lot 350000839, (1:4000) from Bio-Rad.
4.3 Results

4.3.1 No other amino acid substitution at position 274 in hMOF can mimic the potentiating role of K274 acetylation on hMOF thermal stability and cognate lysine substrate acetylation

Because lysine 274 has been shown to have roles in both protein stability and substrate binding and acetylation in vitro, we sought to determine if other amino acids could mimic the effect of K274 autoacetylation. To do so, we prepared recombinant 6x-His-tagged wild-type and mutated hMOF HAT domains containing K274 mutated to every other natural amino acid to assess protein thermostability and catalytic activity. Following affinity purification using the 6x-His-tag, the wild-type and mutant proteins eluted off of a size exclusion column at the same position, suggesting that the mutant proteins were properly folded (Figure 25). Each mutant was tested for catalytic activity, and to ensure that the K274 mutation was not merely affecting catalysis by changing the pKₐ of the active site of the enzyme, activity was tested over a range of pH values (Figure 26). Strikingly, these studies revealed that every K274 mutant protein was significantly defective for cognate lysine acetylation, retaining less than 5% of wild-type activity.
Figure 25: Size-exclusion chromatograms of hMOF K274 mutants suggests that they are properly folded. UV chromatograms of a representative subset of hMOF K274 mutants applied to HiLoad Superdex 75 16/60 gel filtration column. Each mutant displayed a similar profile and the major peak at approximately 75 mL was collected.
Figure 26: All hMOF K274 mutants are catalytically inactive. Relative enzymatic activity of hMOF K274 mutants in comparison to the wild type hMOF over a pH range from 5 to 9 measured as a function of H41-19 peptide acetylation is shown in the chart. hMOF-WT activity at pH 8.0 is set to 1. All mutants retain less than 5% of wild-type enzymatic activity. Error bars represent standard deviations determined from quadruplicate measurements.
To test the thermal stability of the K274 hMOF mutants, each one was tested in a thermal denaturation assay. All mutants produced a melting curve revealing an average loss of 9.3°C in thermal stability (with the exception of K274L, which did not produce a curve from which an inflection point could be calculated) relative to wild-type hMOF (Table 3 and Figure 27A). Interestingly, however, upon addition of 100 µM Co-enzyme A (CoA) cofactor, the stability of the mutants increased an average of 7.9°C and wild-type hMOF stability increased by 6°C (Figure 27B and Table 3). These results demonstrate that the hMOF K274 mutants have compromised thermal stability relative to the wild-type protein but still retain the capacity for binding cofactor.

To assess whether the effect of the hMOF K274 mutations are general to other MYST proteins, we prepared a subset of the hMOF mutants in the recombinant MYST hMOZ HAT domain for enzymatic and thermal stability studies. These studies demonstrate that of the subset of corresponding K604 mutations (to A, R, M and Q) tested, all of them had significantly compromised catalytic activity for cognate lysine acetylation (Figure 28A) and thermal stability relative to wild-type hMOZ (Figure 28B and Table 3). This result suggests that the importance of K274 acetylation on hMOF activity and thermal stability is likely general to other MYST proteins. It should be noted however, that the degree of protein stabilization mediated by autoacetylation at this position may vary from protein to protein within the MYST family.
Table 3: Average melting temperatures for hMOF K274 mutants and MOZ K604 mutants in the presence and absence of 100 µM CoA.

<table>
<thead>
<tr>
<th></th>
<th>hMOF K274X apo $T_m$ (°C)</th>
<th>hMOF K274X+100µM CoA $T_m$ (°C)</th>
<th>hMOZ K604X apo $T_m$ (°C)</th>
<th>hMOZ K604X +100µM CoA $T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K (WT)</td>
<td>40.6 ± 0.4</td>
<td>46.9 ± 0.1</td>
<td>40.3 ± 0.5</td>
<td>45.2 ± 0.4</td>
</tr>
<tr>
<td>D</td>
<td>30.9 ± 0.6</td>
<td>42.8 ± 0.1</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>Q</td>
<td>30.9 ± 1.8</td>
<td>37.2 ± 0.2</td>
<td>36.0 ± 0.3</td>
<td>41.2 ± 0.1</td>
</tr>
<tr>
<td>T</td>
<td>29.6 ± 1.0</td>
<td>38.5 ± 0.0</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>G</td>
<td>31.7 ± 0.7</td>
<td>40.5 ± 0.4</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>M</td>
<td>30.5 ± 0.9</td>
<td>40.2 ± 0.4</td>
<td>34.5 ± 0.3</td>
<td>40.6 ± 0.2</td>
</tr>
<tr>
<td>H</td>
<td>32.3 ± 1.5</td>
<td>41.0 ± 0.4</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>E</td>
<td>29.3 ± 1.7</td>
<td>40.4 ± 0.5</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>R</td>
<td>31.8 ± 1.5</td>
<td>40.3 ± 0.2</td>
<td>37.6 ± 0.9</td>
<td>43.1 ± 0.2</td>
</tr>
<tr>
<td>F</td>
<td>30.1 ± 1.4</td>
<td>38.9 ± 0.2</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>S</td>
<td>31.5 ± 0.9</td>
<td>41.2 ± 0.7</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>C</td>
<td>30.0 ± 1.2</td>
<td>37.8 ± 0.5</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>P</td>
<td>32.7 ± 0.3</td>
<td>40.1 ± 0.4</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>Y</td>
<td>31.5 ± 2.0</td>
<td>38.8 ± 0.6</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>L</td>
<td>-</td>
<td>-</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>A</td>
<td>31.0 ± 0.5</td>
<td>40.1 ± 0.7</td>
<td>34.8 ± 0.6</td>
<td>39.9 ± 0.3</td>
</tr>
<tr>
<td>N</td>
<td>32.7 ± 0.3</td>
<td>40.0 ± 0.5</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>V</td>
<td>32.1 ± 1.0</td>
<td>38.5 ± 0.8</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>I</td>
<td>31.7 ± 0.6</td>
<td>38.3 ± 0.5</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
<tr>
<td>W</td>
<td>32.3 ± 0.4</td>
<td>34.1 ± 0.5</td>
<td>N.R.</td>
<td>N.R.</td>
</tr>
</tbody>
</table>

N.R.=Data not recorded. Note that data could not be fit for K274L. Error is determined from a minimum of triplicate measurements.
Figure 27: hMOF K274 mutants are thermally destabilized, but are still able to bind cofactor. Normalized thermal denaturation curves for hMOF K274 mutants in the absence (A) and presence (B) of 100 µM CoA are shown. SYPRO Orange fluorescence is shown, and the inflection point for each curve is the melting temperature for the protein.
Figure 28: Subset of hMOZ K604 mutations reveals reduced catalytic activity and thermal stability. (A) Enzymatic activity of MOZ K604 mutants as compared to WT at pH 8.0. (B) Thermal denaturation experiment with MOZ-WT and K604 mutants in the presence or absence of 100 μM CoA.
4.3.2 Crystal structure of hMOF K274P reveals loss of rigidity of the α2-β7 loop

In order to understand the molecular basis for the compromised catalytic activity and thermal stability of the hMOF K274 mutants relative to the wild-type protein, we attempted to obtain crystals of each of the hMOF K274 HAT domain mutants for X-ray structure determination. We were only able to obtain crystals of the hMOF-K274P HAT domain mutant, which formed in space group P2₁2₁2₁ with one molecule in the asymmetric unit and diffracted to 2.6 Å resolution. The structure was refined to an R_work/R_free of 24.0/27.6% with good geometry (Table 4). The structure overlays very well with the wild type hMOF HAT domain (PDB code 3QAH) with a RMSD of 0.85 Å² (excluding the α2-β7 loop). Strikingly, in contrast to the wild-type hMOF structure containing the acetylated K274, which shows ordered electron density corresponding to the α2-β7 loop (Figure 29A), the hMOF-K274P mutant shows no clear density for the α2-β7 loop (Figure 29B). These studies suggest that hMOF K274 acetylation contributes to the integrity of the α2-β7 loop of the activated protein. Notably, the active site C316 and E350 residues of hMOF-K274P are in the same position as in the wild type structure (Figure 29C), demonstrating that the α2-β7 loop and acetylated K274 do not contribute to orienting the key catalytic active site residues. Although it is also possible that the introduction of a proline residue, rather than loss of acetyllysine could be responsible for the disorder of the α2-β7 loop, the observation that the hMOF K274P has a similar thermal stability to all other hMOF K274 mutants, suggests that the absence of acetyllysine at least contributes to the observed disorder of the α2-β7 loop.
Table 4: Data statistics for crystal structures of hMOF-K274P and hMOF-C316S/E350Q.

<table>
<thead>
<tr>
<th>Crystal parameters</th>
<th>hMOF-K274P</th>
<th>hMOF-C316S/E350Q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2;2;2_1</td>
<td>P2;2;2_1</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td>a=46.094, b=55.846, c=122.405</td>
<td>a=46.148, b=55.676, c=122.311</td>
</tr>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>36.82 - 2.598</td>
<td>27.84 - 2.166</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>10235 (994)</td>
<td>17365 (1682)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>5.9 (5.9)</td>
<td>5.5 (5.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (99.8)</td>
<td>99.6 (99.6)</td>
</tr>
<tr>
<td>R_{merge} (%)</td>
<td>11.2 (64.9)</td>
<td>5.9 (61.5)</td>
</tr>
<tr>
<td>I/σ</td>
<td>18.05 (2.5)</td>
<td>10.3 (1.2)</td>
</tr>
<tr>
<td>Refinement statistics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R_{work}/R_{free} (%)</td>
<td>23.96/ 27.57</td>
<td>19.05/ 23.37</td>
</tr>
<tr>
<td>Root-mean-square deviation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.006</td>
<td>0.008</td>
</tr>
<tr>
<td>Angles</td>
<td>0.91</td>
<td>0.90</td>
</tr>
<tr>
<td>Average B-factor (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>41.16</td>
<td>44.17</td>
</tr>
<tr>
<td>Water</td>
<td>34.95</td>
<td>42.93</td>
</tr>
<tr>
<td>Ions</td>
<td>37.28</td>
<td>47.4</td>
</tr>
<tr>
<td>Ramachandran statistics (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>Allowed</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Outliers</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Values in parenthesis are for the highest resolution shell.
Figure 29: The hMOF-K274P structure reveals a loss of rigidity in the $\alpha_2$-$\beta_7$ loop. Omit Fo-Fc density (green) contoured at 3$\sigma$ for the $\alpha_2$-$\beta_7$ loop of hMOF WT (PDB ID 3QAH)(A) and hMOF K274P (B). Note the significant lack of density for the $\alpha_2$-$\beta_7$ loop caused by introduction of the hMOF K274P mutation. (C) Comparison of the wild-type and K274P hMOF active sites.
4.3.3 Autoacetylation of K274 does not require active site catalytic residues C316 and E350

To investigate the mode of K274 autoacetylation, we prepared a hMOF C316S/E350Q HAT domain double mutant and asked whether this mutant is capable of K274 autoacetylation. We first prepared hMOF C316S/E350Q in bacteria and confirmed that this mutant had no detectable activity against a cognate substrate lysine in the form of a 19-residue histone H4 N-terminal peptide, while the wild-type hMOF HAT domain showed robust activity (Figure 30A). To determine the K274 autoacetylation state of hMOF C316S/E350Q, we performed a western blot using a pan-acetyl lysine antibody, which indicated that the hMOF C316S/E350Q mutant was acetylated (Figure 30B). We also crystalized this mutant and determined its structure to 2.17 Å resolution. The structure was refined to an $R_{\text{work}}/R_{\text{free}}$ of 19.1/23.4% with good geometry (Table 4). Strikingly, the refined structure revealed that K274 was acetylated with full occupancy (Figure 30C). This observation demonstrates that K274 autoacetylation does not require the active site catalytic residues C316 and E350.

We have previously demonstrated that an isolated peptide of the α2-β7 loop can be chemically acetylated at residue K274 and that the basic nature of native residue K268 or a mutant residue R268 supports K274 autoacetylation, and an uncharged K268M mutant does not (Olia, Barker et al. 2015). To test the importance of K268 for K274 autoacetylation and subsequent hMOF cognate lysine substrate acetylation in the context of the folded hMOF HAT domain, we prepared the recombinant hMOF K268M HAT domain mutant for enzymatic analysis against a histone H4 cognate lysine substrate for comparison with the wild-type hMOF HAT domain. Surprisingly, the kinetic
Figure 30: hMOF-C316S/E350Q is autoacetylated on K274. (A) Comparison of histone H4 acetylation activity of hMOF WT and hMOF C316S/E350Q. (B) Western blot using a pan-acetyl lysine antibody comparing the amount of protein present on the membrane to the signal present on exposed film. (C) Structure of the hMOF-C316S/E350Q mutant showing simulated annealing electron density at 1.5σ around the acetylated K274. S303 is shown as a point of reference.
Michaelis-Menten curves are shown for hMOF-WT and hMOF-K268M against acetyl CoA (A) and histone H4 peptide (B). Different $k_{cat}$ values were calculated for data in each graph to demonstrate similarities in each individual case without bias from averaging. To calculate $k_{cat}$ for this data, $V_{max}$ was divided by the total enzyme concentration of 50nM. Error bars represent standard deviations determined from a minimum of octuplicate measurements.

**Figure 31.** hMOF WT and hMOF K268M share very similar kinetic profiles.
parameters for the wild type and K268M mutant enzyme are remarkably similar (Figure 31). A western blot using a pan-acetyl lysine antibody confirmed that the hMOF K268M mutant is autoacetylated at wild-type levels (Figure 30B). These results suggest that in the context of the folded hMOF HAT domain, K268 does not contribute significantly to K274 autoacetylation and that autoacetylation may be promoted by Ac-CoA binding and positioning for effective K274 autoacetylation.

4.3.4 Autoacetylation of K274 increases hMOF half-life in cells

Our findings that K274 autoacetylation plays a specific and precise role in thermal stability of the hMOF HAT domain in vitro, led us to ask whether K274 acetylation is required for hMOF protein stability in cells. To measure the stability of hMOF and hMOF (K274R) within cells, we expressed V5-tagged versions of the proteins under the control of a Tet-On promoter. Doxycycline was added to induce steady-state levels of the proteins, and then the rate of protein decay was followed after removal of doxycycline and the addition of cyclohexamide to block protein synthesis. hMOF-K274R decayed more rapidly than the wild type protein (Figure 32A). The mutant protein also decayed more rapidly than wild type when the experiment was performed without cyclohexamide, thus ruling-out potential contributions from other effects of the drug (e.g. inhibition of autophagy) (Figure 32B). These findings demonstrate that the K274R mutation destabilizes hMOF in vivo, and indicate that hMOF K274 autoacetylation contributes to protein stability in cells.
Figure 32: hMOF-WT shows a longer half-life than hMOF-K274R in cells. (A)
The levels of V5-tagged hMOF-WT and hMOF-K274R were followed by
immunoblotting at the indicated time points after doxycycline withdrawal and
treatment with cyclohexamide. (B) Same as (A), except without cyclohexamide.
Immunoblotting for actin on the same blot provides a loading control.
4.4 Discussion

The results described here reveal that an autoacetylated lysine residue at position 274 of hMOF contains the essential specific chemical properties to promote hMOF stability and activity. Strikingly, we find that hMOF K274 autoacetylation cannot be mimicked by other residues, arguing for the unique chemical properties of an autoacetylated K274 residue in hMOF catalysis. While previous reports have proposed that hMOF K274 autoacetylation may be required to ‘unlock’ the protein off state by making the catalytic E350 available for the catalytic reaction and making the cognate lysine substrate binding site more accessible (Sun, Guo et al. 2011, Yuan, Rossetto et al. 2012), we have found that mutations of K274 to very small amino acids such as glycine or alanine do not activate the protein. These data would suggest that hMOF K274 autoacetylation also plays a more direct, positive role in the active state of the enzyme, rather than merely counteracting reactivity in the off state. All but one of the hMOF K274 mutants were demonstrated to bind CoA, and the K274P crystal structure reveals that the key C316 and E350 catalytic residues are unperturbed. The α2-β7 loop, however, was observed to be disordered in the hMOF-K274P structure, consistent with the loss of thermal stability for the K274 mutants in vitro and the reduced half-life to hMOF-K274R in cells. These observations demonstrate that in addition to facilitating cognate lysine substrate binding (Yuan, Rossetto et al. 2012), hMOF K274 autoacetylation promotes protein stability to also indirectly contribute to cognate lysine substrate catalysis. We previously reported on a structure of hMOF in which the α2-β7 loop could be modeled in two alternate conformations, one in which K274 was acetylated and another where K274 was unacetylated (Yuan, Rossetto et al. 2012). Interestingly, the α2-β7 loop could be traced in both cases, however K274 refined to a
higher B-factor in the unacetylated relative to acetylated form, 36.6 Å² vs. 30.9 Å², respectively, consistent with the greater rigidity of α2-β7 loop when K274 is acetylated.

While autoacetylation of hMOF K274 has been demonstrated to be required for cognate histone substrate binding (Yuan, Rossetto et al. 2012), no MYST family proteins have been crystallized with their cognate lysine peptide substrates. To gain insights into how hMOF might recognize its cognate substrate and how the K274-containing α2-β7 loop might participate in this recognition, we aligned the structurally similar Hat1 HAT of the Hat1/Hat2/H4 peptide crystal structure (PDB ID 4PSW) with hMOF (PDB ID 3QAH). This superposition demonstrates that a Hat1 loop containing E162-D168 overlays well with the α2-β7 loop of hMOF (Figure 33). In the Hat1 structure, four of six residues in this region are reported to make interactions with the H4 peptide (E162, A163, N165, and I167) (Li, Zhang et al. 2014). We hypothesize that the same trend may extrapolate to the α2-β7 loop of hMOF such that the hMOF α2-β7 loop might be used for cognate histone H4 peptide recognition in a similar way. Similar ideas have been proposed comparing hMOF to Gcn5 bound to H3 substrate (Kadlec, Hallacli et al. 2011, Sun, Guo et al. 2011). Such a model would explain how disruption of the α2-β7 loop may compromise histone H4 binding for cognate lysine acetylation.

We provide crystallographic evidence that hMOF K274 becomes acetylated in the absence of both known catalytic residues, C316 and E350, which makes a strong argument for the autoacetylation at this position to not be mediated by acid/base catalysis as proposed for cognate lysine substrate acetylation. We have also found that in contrast to the trends seen in hMOF peptides (Olia, Barker et al. 2015), mutation of the preceding basic residue, K268 in context of the HAT domain does not appear to
Figure 33: Overlay of Hat1 bound to H4 peptide structure with hMOF structure reveals that hMOF α2-β7 loop likely makes contacts with H4. hMOF is shown in cyan with the α2-β7 loop in navy, Hat1 is shown in light pink with the E162-D168 region (corresponding to the α2-β7 loop in hMOF) shown in red. A region of the H4 peptide from the Hat1 structure is shown in magenta.
affect enzyme activity. Thus, the mechanism for the regulation of hMOF K274 appears likely to be driven more by Ac-CoA binding and positioning, potentially to prime hMOF for cognate lysine acetylation.
CHAPTER 5

Conclusions and Future Directions
5.1 Identification of hMOF Inhibitors

5.1.1 Conclusions

We were unable to identify any inhibitors from our high throughput screening campaign of hMOF. While we did identify three compounds that irreversibly inhibited hMOF with stabilizing effects, our inability to confirm the identity of these compounds made it impossible for us to characterize them further. We also encountered numerous false hits throughout our screening campaign, with one major culprit being colloidal aggregates. One study has reported that in a screening campaign testing compounds in a concentration series from 3 nM to 30 µM, 95% of hits were shown to act through aggregate-based inhibition. Furthermore, 60% of the compounds that formed colloidal aggregates at 5 µM did not form aggregates at 1 µM and aggregate formation increased with increasing compound concentration (Feng, Simeonov et al. 2007). Therefore, it is not quite so surprising that many of our hits were colloidal aggregates at 20-25 µM. It is also important to remember that even with these high screening concentrations, our overall hit rate was fairly low at 0.17%. We had initially tried screening some libraries at much lower compound concentrations (5 µM) which yielded no hits at all. Additionally, even the best known MYST acetyltransferase inhibitors have IC$_{50}$ values greater than 5 µM (discussed in section 1.6). Taken together, these data suggest that it may be very hard to target the available chemical space using small molecules in the MYST enzymes and that perhaps small molecule HTS may not be the best approach for finding hMOF inhibitors. For this reason and others, I propose that crystallographic fragment screening would be a reasonable route to finding hMOF inhibitors.
Fragments differ from small molecule drug-like compounds in a variety of ways. Fragments have lower molecular weight, increased polarity and solubility, and less chemical complexity. The goal of fragment screening is to either find multiple fragments which bind in adjacent regions of a target protein and then design an inhibitor by linking those fragments or to find a single fragment and expand it based on known structural information (Joseph-McCarthy, Campbell et al. 2014). Because of the limited number of contacts made by a fragment, enzymatic activity readouts are generally insufficient for identifying fragment binders. Rather, assays that detect binding, such as SPR, NMR, ITC, and crystallography as well as some other are employed (Joseph-McCarthy, Campbell et al. 2014).

There are three major criteria that should be met to successfully perform crystallographic fragment screening. Firstly, X-ray diffraction resolution will generally need to be better than 2.8 Å to identify the bound fragment. Secondly, crystals must be resilient to soaking conditions, and thirdly the crystal conditions/ crystal packing must allow for fragment binding (Patel, Bauman et al. 2014). hMOF has already been proven to meet all of these criteria. hMOF crystals, in my hands, frequently diffract to better than 2.8 Å at synchrotron beam sources. I have also subjected multiple hMOF crystals to multiple-day long soaks with various compounds without fracturing crystals or losing diffraction. Additionally, not only can CoA be easily soaked into pre-formed crystals, demonstrating the availability of the binding pocket in the crystallographic form, but I have also observed the fragment-sized compound salicylic acid bound to hMOF. A hurdle that is often reported in crystallographic fragment screening is when crystal conditions bind the crystal, causing fragments to need to compete with crystallization
conditions (Patel, Bauman et al. 2014). hMOF crystals in my hands, however, have been void of visible crystal condition molecules in the structure, with mostly just water and ions occupying the protein pockets. For these reasons, I think hMOF would be an ideal candidate for crystallographic fragment screening.
5.2 Autoacetylation of K274

5.2.1 Conclusions

A mutational scan of hMOF K274 revealed that all amino acid substitutions of this residue were able to bind cofactor but were significantly destabilized, both in vitro and in cells, and are catalytically inactive for cognate histone H4 peptide lysine acetylation. The X-ray crystal structure of a hMOF K274P mutant suggested that the reduced stability and catalytic activity stemmed from a disordering of the residue 274-harboring α2-β7 loop. While we hypothesized that the disordering of this loop results in a loss of binding to the H4 peptide, we were unable to confirm this hypothesis using a variety of biochemical techniques due to the relatively weak affinity of hMOF for the substrate histone tail. Furthermore, there are no crystal structures of any of the MYST family proteins bound to the histone tail. As future directions for this project, I would like to explore ways to improve H4 tail binding/affinity to hMOF so that it can be crystallized and further studied. Additionally, the loss of catalytic activity for cognate substrate correlated with the mutation of K274 raises the question of whether this residue is involved in the catalysis for the MYST family. In this respect I also propose a future direction of tracking the K274 acetyl group to see if it is transferred from the enzyme to the substrate. Lastly, throughout the course of this project I have attempted a number of other experiments to learn more about the role of K274 acetylation that were overall, unsuccessful. Specifically, I attempted to make a series of hMOF double mutants that would recapture the stability of the stability of the α2-β7 loop. For the sake of record and for a fuller understanding of the direction of my future studies, I will detail these attempted experiments at the conclusion of this section.
5.2.2 Future Directions: Improving hMOF Affinity for H4 Tail

One hypothesis as to why none of the MYST family acetyltransferases have been crystallized with their cognate substrate peptides is because of the sequential nature of their ping-pong catalytic mechanism. Before the H4 peptide binds, AcCoA must first bind the active site, transfer an acetyl group to the active site cysteine, and then exit the active site. It is reasonable to speculate that the hMOF crystals have never revealed H4 binding is because the active site is not primed to bind H4. As such I would like to design an hMOF mutant where the active site mimics the active site following acetyl transfer from AcCoA to the active site cysteine, perhaps by mutating the cysteine to a glutamine in an attempt to mimic acetyl-cysteine, and then use this mutant to attempt co-crystallization with the H4 peptide.

Another problem that could be causing the difficulty in co-crystallizing the H4 peptide with hMOF is the low affinity of the protein towards the peptide. Perhaps to overcome this I could create an hMOF construct that has a section of the H4 tail cloned into the N or C terminal of the protein with a linker. This would allow for a higher local concentration of H4 tail in constant proximity with its binding site. Current crystallization protocols for hMOF leave the 6xHis tag and TEV cleavage site intact; implying that excess peptide on the N-terminus of the protein will not abrogate crystallization. If this region were replaced with the relevant sequence of the H4 peptide it could ensure a 1:1 mixture of hMOF:H4 peptide in our crystals which may induce the H4 binding due to the increased local concentration.

Crystallographic information on how hMOF binds the substrate peptide would not only allow us to confirm our hypothesis on the role of the autoacetylation of hMOF in
catalysis, but would also provide a key piece of information on the MYST family that has been thus far missing.

5.2.3 Future Directions: Tracking the K274 Acetyl Group

Every K274 mutant enzyme was significantly defective for cognate lysine acetylation, despite being properly folded, which raises the question of whether this amino acid is a catalytic residue for the MYST family. To investigate this issue, I propose an experiment to track the acetyl group on K274 to determine if it can be transferred to a cognate substrate.

To perform this experiment, I would use a 6xHis tagged WT hMOF as well as a 6xHis tagged hMOF C316S mutant to control for acetyl transfer occurring through the active site cysteine. Each enzyme would first be incubated with the promiscuous yeast deacetylase HST2 and NAD⁺ to remove any existing acetylation. Following deacetylation, each hMOF enzyme would be isolated by binding to nickel resin in a column and washing away HST2. Next the hMOF would be incubated with 14C-AcCoA to allow transfer of 14C acetyl to K274. Following the incubation, the enzymes would be washed to remove free 14C-AcCoA, followed by a wash with a saturating amount of cold CoA to displace any 14C-AcCoA that may remain bound to the AcCoA binding pocket of the enzymes. A sample of both WT hMOF and hMOF C316S would be collected at this point to measure the degree of acetylation of the samples. Next, to measure acetyl transfer, both the WT and mutant hMOF would be incubated with H4 peptide. Following incubation with the H4 peptide, the H4 peptide would be eluted off the columns and collected. The nickel-bound hMOF would be washed again to remove any remaining H4
peptide, and another sample would be collected to measure the degree of acetylation of the hMOF enzymes following acetyl transfer. The amount of acetylation on the collected H4 peptide or the difference in acetylation between pre- and post- H4 incubated enzyme would indicate the amount of acetyl transfer that occurred for either of the two enzymes. Then comparing the difference between the amount of acetyl transferred for the WT hMOF and the hMOF C316S mutant to decipher to what extent, if any, the acetyl transfer is occurring through K274.

One caveat of this proposed experiment is that trying to isolate the possible catalytic contribution of K274 from the active site cysteine could be very difficult to achieve. A previous study has demonstrated that when incubated with AcCoA, the MYST active site cysteine will become acetylated (Yan, Harper et al. 2002) While I propose using a hMOF C316S mutant as a control, this mutant has been shown to be catalytically inactive against cognate substrate and likely would not transfer acetyl group to the H4 peptide. Similarly, I could include a control of a hMOF K274Q in an attempt to isolate the catalytic contribution of the active site cysteine, but again, hMOF K274 mutants have been demonstrated to be catalytically inactive. Thus, the major difficulty in assessing the catalytic potential of lysine 274 will be isolating the catalytic contribution of the active site cysteine.

Alternatively, another method we could use to help differentiate between acetylation originating from the cysteine versus the lysine would be mass spectrometry. For this two-part experiment I would first want to determine if the acetyl group on the active site lysine 274 were sufficiently labile to be transferred in the absence of a deacetylase, for instance. To test this I would once again deacetylate the hMOF with HST2 to remove any preexisting acetylation. Then I would incubate the enzyme with $^{13}$C-
AcCoA, and take a sample for mass spectrometry to ensure that the K274 is acetylated with $^{13}$C-acetyl. Next I would wash away the $^{13}$C-AcCoA, and add an excess of $^{12}$C-AcCoA (cold) and H4 to allow the enzyme to go through several acetylation cycles, then wash the hMOF again and take another sample for mass spectrometry. To avoid different peptide lengths in the mass spectrometry data for acetylated versus deacetylated lysines, the hMOF would be denatured and treated with $^{12}$C-acetic anhydride prior to trypsinization such that cut sites are uniform and only occur at arginines. If the lysine 274 remains acetylated with $^{13}$C-acetyl, this would indicate that this residue is not participating in catalysis and the experiment is concluded. If the lysine is not acetylated with $^{13}$C-acetyl, but rather $^{12}$C-acetyl, then I move on to the second experiment to confirm that the acetyl group from lysine 274 is transferred to the substrate. For this experiment I would incubate a large amount (low micromolar) of hMOF with $^{13}$C-AcCoA once again, and wash. Then I would add an equimolar amount of H4 as well as an excess of cold $^{12}$C-AcCoA. I would then isolate the H4 peptide following the reaction and use mass spectrometry to determine if $^{13}$C-acetyl is transferred to the H4 peptide. This experiment would still have the caveat that I could not ensure that the $^{13}$C-acetyl was not traveling through the active site cysteine, especially as an acetyl cysteine is generally not stable enough to be detected in mass spectrometry experiments. However, if I were able to show that the acetyl-lysine were labile enough to transfer the acetyl group from the first experiment, then perhaps I could, in this case, use a hMOF C316S mutant to show some percentage of acetyl transfer through K274. Of course the caveat here becomes that if the catalytic mechanism requires both the cysteine and the lysine, then once again teasing out the individual contributions of either amino acid becomes complicated.
5.2.4 Attempted Future Directions: Stabilizing the α2-β7 Loop

Because our crystallographic data from the K274P structure revealed a loss of rigidity in the α2-β7 loop of hMOF, and because we hypothesized that the positioning of this loop was likely important for cognate substrate binding, we sought to make additional mutations in hMOF that would stabilize this loop in the presence of a K274 mutation. Specifically, because acetylated K274 hydrogen bonds to S303, I sought to make mutations at position 303 that could potentially make hydrogen bonds to various K274 mutants. First I tried to mimic the acK274/S303 interaction by making a hMOF K274S/S303Q mutant that would, in essence, swap the two residues, using the glutamine as an acetylated lysine mimic. This mutant however was insoluble, and I hypothesized that perhaps the glutamine, being shorter in length than an acetylated lysine, was not long enough to be in close enough proximity to S274 for hydrogen bonding. Following this logic, I made two more mutants, which I hypothesized would have the length required for hydrogen bonding: hMOF K274N/S303Q and hMOF K274Q/S303Q. Despite the fact that both of these double mutants were soluble, which was an improvement on the prior hMOF K274S/S303Q mutant, they both thermally destabilized the hMOF enzyme by an even greater degree than their K274 single mutant counterparts. Ultimately, I was unable to reestablish the stability of the α2-β7 loop.
Table 5: Frequently used abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcCoA</td>
<td>Acetyl Coenzyme A</td>
</tr>
<tr>
<td>acK</td>
<td>acetyl lysine</td>
</tr>
<tr>
<td>CBP</td>
<td>CREB-binding protein</td>
</tr>
<tr>
<td>CoA</td>
<td>Coenzyme A</td>
</tr>
<tr>
<td>CPM</td>
<td>counts per minute</td>
</tr>
<tr>
<td>GNAT</td>
<td>GCN5-related N-acetyltransferase</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathione S-transferase</td>
</tr>
<tr>
<td>H4</td>
<td>Histone H4</td>
</tr>
<tr>
<td>H4K16</td>
<td>Histone H4 lysine 16</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyltransferase</td>
</tr>
<tr>
<td>HAT1</td>
<td>Histone acetyltransferase 1</td>
</tr>
<tr>
<td>hMOF</td>
<td>human males absent on the first</td>
</tr>
<tr>
<td>HTS</td>
<td>High Throughput Screen</td>
</tr>
<tr>
<td>LCGC</td>
<td>Lankenau Center for Chemical Genomics</td>
</tr>
<tr>
<td>MOF</td>
<td>males absent on the first</td>
</tr>
<tr>
<td>MOZ</td>
<td>monocytic leukemia zinc-finger protein</td>
</tr>
<tr>
<td>MYST</td>
<td>(named for founding members) MOZ, YBF2/SAS3, SAS2, and Tip60</td>
</tr>
<tr>
<td>NPDI</td>
<td>Natural Product Discovery Institute</td>
</tr>
<tr>
<td>PAINS</td>
<td>Pan-assay interference compounds</td>
</tr>
<tr>
<td>PCAF</td>
<td>P300/CBP-associated factor</td>
</tr>
<tr>
<td>PTM</td>
<td>Post translational modification</td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco etch virus</td>
</tr>
<tr>
<td>Tip60</td>
<td>60 kDa Tat-interactive protein</td>
</tr>
</tbody>
</table>
REFERENCES


