

Regulation of Activity and Selectivity of Histone Deacetylases

by

Kelsey Anne Diffley

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in the University of Michigan
2021

Doctoral Committee:

Professor Carol A. Fierke, Chair
Associate Professor Tomasz Cierpicki
Assistant Professor Yongqing Li
Professor E. Neil G. Marsh
Associate Professor Patrick O'Brien

Kelsey A. Diffley

kdiffley@umich.edu

ORCID iD: [0000-0003-4373-8278](https://orcid.org/0000-0003-4373-8278)

© Kelsey A. Diffley 2021

Acknowledgements

I would like to thank my advisor, Prof. Carol Fierke, for her guidance and mentorship over these past five years. Whether down the hall in the Chemistry building or 1,300 miles away in Texas, I have always felt your support. I thank the past and present members of the Fierke lab: Jan Sheng, Eric Zhang, Hannah Foley, Dr. Jeff Lopez, Dr. Eric Sullivan, Dr. Nancy Wu, Dr. Ben Jennings and Dr. Kip Kaitany for their thoughtful discussions, science or otherwise, coffee breaks, and lunch outings. I particularly thank the last of the Michigan Fierke lab members: Dr. Desireé García-Torres, Andrea Stoddard, Dr. Katherine Welker-Leng for their friendship and support keeping me sane through the end.

I would also like to thank Prof. Neil Marsh, for taking me into his lab and his mentorship and support during my final year. I thank the past and present members of the Marsh lab: Dr. Soumi Ghosh, Dr. Tim Grunkemeyer, Dr. Karl Koebke, April Kaneshiro, Prathamesh Datar, Hannah Chia, Srijoni Majhi, and Ayesha Patel for their help, conversations, and appreciated snacks.

I thank the members of my committee for their advice and expertise.

Lastly, I thank my family for their continued support and love, without which this thesis would not have been possible.

Table of Contents

Acknowledgements	ii
List of Tables	vi
List of Figures	viii
Abstract	xi
Chapter 1 : Introduction	1
Inhibitor Development	5
Importance of Protein-Protein Interactions	8
Substrate Discovery	11
The HDAC Activity “Toolbox”	16
Identity of Active Site Metal	20
Post-translational Modifications	22
Conclusion	25
Chapter 2 : Novel Non-Hydroxamate Inhibition of Histone Deacetylases*	37
Introduction	37
Materials and Methods	39
Reagents	39
CFL1 Library and Expanded Inhibitors	39
Purification of apo-HDAC8	40
Metal-Chelating Fragment Library Screen	41
Dose Response Curves	41
ICP-MS Assay	42
UV-Vis Assay	43
Oxygen Sensitivity of HeLa cell lysates	43
Results	44
Select fragments display potent inhibition	44
HDAC8-Fe(II) selective inhibition	48
Specificity of inhibitor is related to IC_{50}	51
Treatment of HeLa cell lysates	54

Discussion	55
Chapter 3 : The formation of the CoREST complex enhances HDAC1 deacetylase activity and alters selectivity	63
Introduction	63
Materials and Methods	65
Reagents	65
Expression and Purification of HDAC1	66
Expression and Purification of LSD1	68
Expression and Purification of CoREST	69
Expression and Purification of LSD1-CoREST complex	69
6xHis Tag Cleavage	70
Co-Immunoprecipitations	71
Preparation of Singly Acetylated Histone H3	72
Coupled-enzyme Acetate Detection Assay	72
Results	73
Deacetylase activity of recombinant HDAC1 is dependent upon substrate length	73
HDAC1 interacts with LSD1-CoREST complex <i>in vitro</i>	79
Addition of LSD1-CoREST enhances HDAC1 deacetylase activity	81
Discussion	84
HDAC1 substrate selectivity is dependent on post-translational modifications	84
HDAC1 substrate selectivity is dependent on complex formation	85
Interactions significantly enhance HDAC1 deacetylase activity	87
Chapter 4 : Structure-based Prediction of Substrate Selectivity of HDAC6[‡]	95
Introduction	95
Materials and Methods	97
Reagents	97
HDAC6 Expression and Purification	98
Coupled Acetate-Detection Assay	99
Fluor de Lys Assay	100
Calibration of FlexPepBind	101
Running the calibrated protocol on the acetylome	103
Results	103
Measurement of HDAC6 zCD2 substrate selectivity	103
Structure-based computational prediction identifies most substrates of HDAC6	106
Predictions on the human acetylome	110

Comparison of CD2 to CD12 substrate selectivity	114
Discussion	118
Structural differences dictate substrate binding specificity	118
Structure-based model can predict novel HDAC6 substrates	121
Chapter 5 : Conclusions and Future Directions	128
Regulation of HDACs is multi-faceted	128
Identity of <i>in vivo</i> active site metal	128
Post-translational modifications	131
Complex Formation	133
Regulatory mechanisms alter selectivity rather than strictly activity	135
The future of HDAC research	136
Appendix	142

List of Tables

Table 2.1: CFL1 Library Fragments screened against all HDAC8-reconstituted forms.	46
Table 2.2: IC_{50} values of most potent inhibitor fragments	47
Table 2.3: IC_{50} values of parent fragments and select expanded inhibitors	51
Table 3.1: Primers for various construct cloning. Restriction enzymes sites are in bold italics. Mutagenesis sites are in bold/underlined.	69
Table 3.2: Catalytic efficiencies of HDAC1 constructs dependent on expression system	78
Table 3.3: Activation of deacetylase activity upon addition of LSD1-CoREST	84
Table 4.1: Kinetic parameters of peptide substrates used to construct the model (D-TRAINING).....	105
Table 4.2: List of peptides where only a lower limit for k_{cat}/K_M was determined (D-EXT).	106
Table 4.3: Different FlexPepBind protocols evaluated in this study.	107
Table 4.4: Performance metrics of the top 3 protocols on the D-TRAINING dataset (25 peptides)	109
Table 4.5: Kinetic parameters of selected predicted substrates from the human acetylome.....	113

Table 4.6: Kinetic parameters obtained for various HDAC6 constructs with a Lys(Ac)-AMC substrate. 115

Table 4.7: Kinetic parameters of peptide substrates tested with zHDAC6 CD12..... 116

Table A1.1: Top 100 protein acetylome hits from structure-based model of HDAC6.. 142

List of Figures

Figure 1.1: The classes of histone deacetylases and their domain components.....	2
Figure 1.2: Catalytic mechanism of histone deacetylases	3
Figure 1.3: Crystal structures of HDACs with inhibitors bound.....	6
Figure 1.4: Inhibitor structures.....	7
Figure 1.5: Interaction networks of Class I HDACs	10
Figure 1.6: Relative activity of Class I HDACs on various histone acetylation sites	13
Figure 1.7: New substrate discovery methods.....	16
Figure 1.8: HDAC activity assays	19
Figure 2.1: Heat maps illustrating inhibitor potency.....	45
Figure 2.2: Dose response curves of most potent inhibitor fragments against 0.5 μ M HDAC8-Co(II).....	48
Figure 2.3: Metalloform selective inhibitors dose response curves.	49
Figure 2.4: Modification of 4D (left) and 5D (right).....	50
Figure 2.5: Dose response curves of selected expanded inhibitors.	50
Figure 2.6: Bar graphs displaying ICP-MS data of HDAC8 and Fe concentrations	53
Figure 2.7: Inhibitor binding to Fe(II) as measured by UV-Vis spectroscopy,	53

Figure 2.8: Deacetylase activity of HeLa cell lysates with and without treatment with 100 μ M HDAC8-Fe(II) specific inhibitor HPO15.	55
Figure 3.1: Deacetylation rate of recombinant HDAC1 varies with substrate	75
Figure 3.2: Detection of phosphorylated HDAC1	76
Figure 3.3: Coomassie-stained SDS-PAGE gel to illustrate purity of HDAC1 constructs.	79
Figure 3.4: Chromatogram of LSD1-CoREST through size-exclusion chromatography	80
Figure 3.5: Co-immunoprecipitation of HDAC1 using an LSD1 antibody.....	81
Figure 3.6: Comparison of catalytic efficiencies of HDAC1 constructs in catalyzing acetylated Histone H3 substrates in the presence of protein interactors.....	83
Figure 3.7: Substrate selectivity of HDAC1 constructs.	86
Figure 3.8: Proposed formations of HDAC1, LSD1, and CoREST containing complexes	88
Figure 4.1: Dependence of deacetylation rate on substrate concentration for two representative peptides catalyzed by HDAC6.	104
Figure 4.2: Key differences in the substrate binding structure between HDAC6 and HDAC8	108
Figure 4.3: Performance of different protocols on the D-TRAINING set.....	110
Figure 4.4 : Application of the calibrated protocol (RLOOP_R) to the acetylome to detect novel potential HDAC6 substrates.	112

Figure 4.5: Dependence of deacetylation rate on substrate concentration catalyzed by zHDAC6 CD2 for two representative peptides selected from the top acetylome hits. . 113

Figure 4.6: Active site titration of purified zHDAC6 CD12 with Tubastatin A 114

Figure 4.7: Dependence of deacetylation rate on substrate concentration for two representative peptides measured with both CD12 and CD2..... 117

Figure 4.8: Structural differences between zHDAC6 CD2 and CD12..... 120

Figure 5.1: The proposed post-translational modifications of HDAC1. 132

Figure 5.2: Overlapping interactions between HDAC1 and CoREST (RCOR1) or LSD1 (KDMA1). 134

Abstract

Acetylation is an important post-translational modification (PTM). Lysine acetylation is a reversible PTM, where deacetylation is catalyzed by histone deacetylases (HDACs). Histone deacetylase function is crucial for a correctly functioning cell as aberrant acetylation, or deacetylation, has been linked to cancer, diabetes, neurodegeneration, and auto-immune disorders. Yet information about proper regulation of these enzymes is limited. Regulation of HDAC activity and selectivity has been proposed to include: the identity of the divalent active site metal ion, post-translational modifications, and protein interactions to form stable multi-protein complexes. HDAC activity and selectivity is further influenced by substrate amino acid sequence. This thesis explores how these different regulatory measures impact HDAC activity and selectivity.

The biochemically well-characterized HDAC8 was used to investigate novel HDAC inhibitors and it was found that the identity of the active site divalent metal ion plays an important role in determining inhibitor selectivity. The identification and characterization of inhibitors with selective metal-binding groups, particularly Fe(II)-HDAC8 selective inhibitors, demonstrates structural differences between different HDAC8 metalloforms. This work also identified that the tropolone metal binding group potently inhibits HDAC8.

To examine the impact of post-translational modifications and protein interactions on deacetylase activity and selectivity, a simplified CoREST complex including HDAC1 was reconstituted. *In vitro* HDAC1 complex formation significantly increases deacetylase

activity (>10-fold) in comparison to HDAC1 in isolation. The presence of post-translational modifications, specifically phosphorylation, was found to impact substrate selectivity with the identification of a phosphorylation-specific acetylation site, without preventing complex formation.

Finally, to explore the sequence-level substrate selectivity of HDAC6, we successfully constructed a structure-based model of the catalytic domain of HDAC6. This model was used to predict novel substrates that were then validated using peptide mimics. These data demonstrated that the substrate selectivity of HDAC6 is more promiscuous than HDAC8. The comparison of the activity of the single catalytic domain of HDAC6 with HDAC6 containing both catalytic domains demonstrates that the different structural components influence the activity and substrate selectivity profile of the enzyme.

The findings discussed within this thesis illustrate several regulatory factors impart a sizeable contribution to deacetylase activity and selectivity. Such factors include structural components, including cofactors and post-translational modifications, in addition to protein interactions. The contribution of this thesis to the growing knowledge of how HDACs are regulated provides insight into the enzymes' biological function to lead to the development of more effective therapeutic interventions.

Chapter 1 : Introduction

Post-translational modifications (PTMs) of proteins, first identified 60 years ago, have been found to be widespread across the proteome. The first PTM discovered was phosphorylation[1], closely followed by acetylation. Protein acetylation, specifically on lysine residues, was discovered on histone tails in 1963[2, 3]. In the following two decades, numerous other acetylated proteins were found such as tubulin[4] and p53[5]. However, until mass spectrometry was able to be effectively used in a high-throughput manner to identify acetylation, acetylated proteins were thought to be limited and focused to the nucleus. Today, there are thousands of acetylation sites identified throughout the cell[6].

Acetylation is a reversible PTM, thus it can act as a regulatory switch. On histones, acetylation, catalyzed by histone (or lysine) acetyltransferases, generally promotes transcription and deacetylation, catalyzed by histone deacetylases, generally represses transcription[7, 8]. It took 30 years after the discovery of acetylation before the enzymes responsible were isolated and purified. Acetyltransferases and deacetylases were first isolated in yeast and protozoans[9]. Using the knowledge garnered from yeast, an inhibitor capture method was used to isolate the first human deacetylase, histone deacetylase 1 (HDAC1)[10, 11]. The purified HDAC1 demonstrated *in vitro* deacetylase activity against histones, earning its name. Discovery of more deacetylases happened rapidly thereafter[12-19].

The histone deacetylase family is divided into four classes based on homology to yeast deacetylases. Classes I, II, and IV, referred to as HDACs, are metal-dependent deacetylases, requiring a divalent metal for catalysis within a conserved catalytic domain. Class III enzymes, named sirtuins based on their homology to the yeast enzyme, Sir2, utilize a NAD-dependent mechanism. Herein, only the metal-dependent deacetylases, or HDACs, will be discussed in detail (**Figure 1.1**). Class I includes HDAC1, HDAC2, HDAC3, and HDAC8. All are homologous to yeast deacetylase, Rpd3. The Class II HDACs: 4, 5, 6, 7, 9, and 10, are homologous to another yeast deacetylase, Hda1. Class II is further divided into Class IIa and Class IIb, based on sequence homology. Class IIa consists of HDAC4, HDAC5, HDAC7, and HDAC9. Class IIb consists of HDAC6 and HDAC10, where HDAC6 uniquely contains a second deacetylase domain.

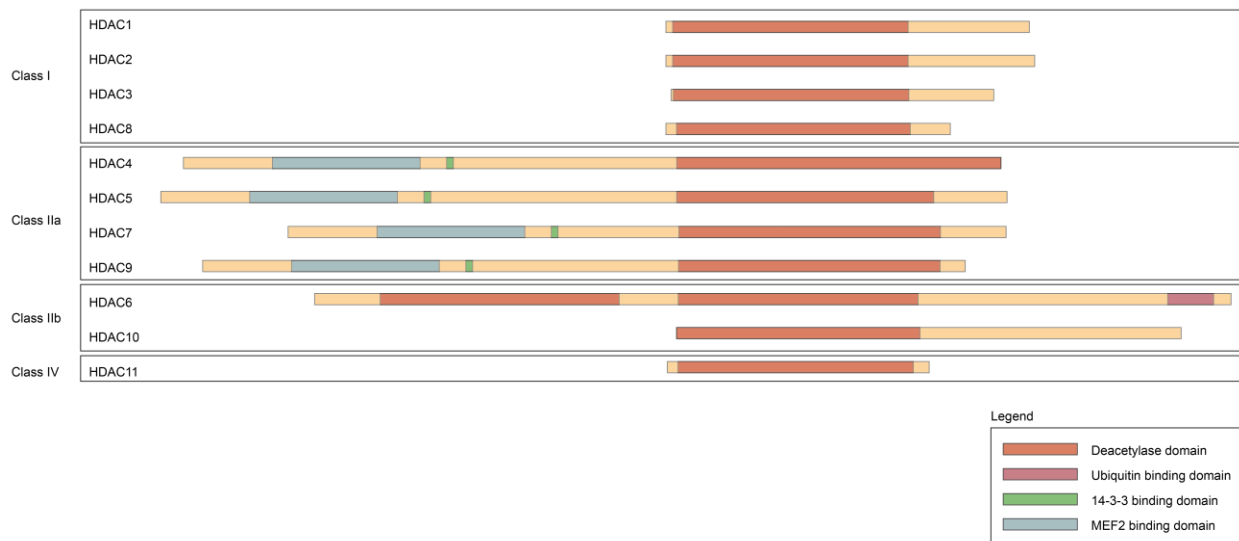


Figure 1.1: The classes of histone deacetylases and their domain components.

All HDACs contain a conserved active site. The active site consists of a His-Asp-Asp metal binding site at the end of an internal cavity that accommodates the lysine side chain. The catalytic mechanism has recently been determined to be most consistent with

an active site conserved histidine serving as a single general base-general acid to activate the metal-water nucleophile and then protonate the lysine leaving group (**Figure 1.2**) [20].

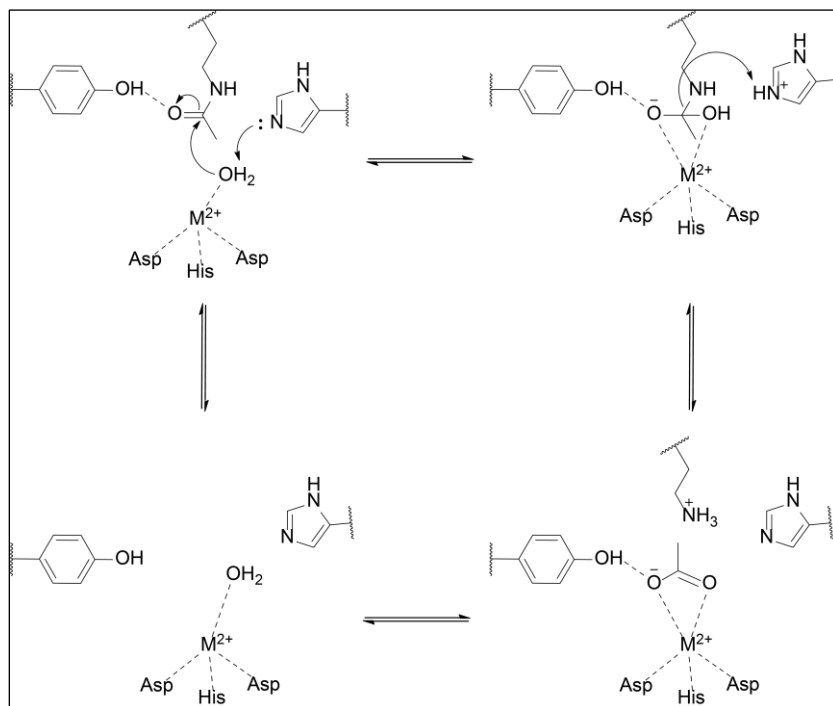


Figure 1.2: Catalytic mechanism of histone deacetylases

HDACs demonstrate varied catalytic efficiency. The Class I enzymes demonstrate the greatest *in vitro* deacetylase activity, apart from HDAC8. HDAC8 has been measured to be ~1000 times less active than HDACs 1-3[21]. HDAC6 demonstrates high catalytic activity similar to HDACs 1-3[21]. Whereas the activity of HDAC10 has been measured to be close to that of HDAC8[21]. Recently, HDAC10 was shown to more efficiently deacetylate polyamines than acetyl-lysine peptides[22]. The Class IIa enzymes are the only members proposed to lack catalytic activity. All other HDACs contain a conserved tyrosine in the active site which forms hydrogen bonds with the substrate's acetyl group (**Figure 1.2**). HDACs 4,5,7,and 9 have an amino acid substitution of the tyrosine to a histidine abolishing deacetylase activity[23]. The Class IIa enzymes have rather been

found to influence deacetylation through protein interactions[24]. Lastly, HDAC11 demonstrates little *in vitro* deacetylase activity. HDAC11 has been shown to more efficiently catalyze deacylation of fatty-acids[25, 26]. The variation of catalytic activity amongst the HDAC enzymes alludes to functional differences.

HDACs are considered to be ubiquitously expressed. Cellular localization of HDACs varies by class. It is widely believed Class I and IV are primarily found in the nucleus, however HDAC8 has been found in the cytoplasm of smooth muscle cells[27-29], indicating potential tissue-specific variability. Class IIa shuttles in and out of the nucleus depending on post-translational modifications and protein interactions[30-32]. Class IIb are primarily cytoplasmic[33], though HDAC6 has also been found in the nucleus and contains a nuclear localization signal and two nuclear export signals[34, 35]. Any regulatory mechanism of HDAC6 localization is unknown.

Though we understand the mechanistic activity of HDACs, the specific biological activity of each enzyme is unclear. All the HDAC enzymes appear to regulate transcription[15, 18, 33, 36-39]. HDACs further regulate the immune response through their involvement in cytokine gene transcription and development of regulatory T-cells[40-45]. Specifically, Class I HDACs are involved in the cell cycle: controlling cell growth[46, 47], proliferation[48-51], differentiation[52, 53], and apoptosis[47, 54, 55]. Class IIa enzymes regulate development and cellular differentiation[56]. Class IIb enzymes play a role in cell proliferation, specifically, HDAC6 activity has been linked to the cellular processes of cell motility[57-59] and autophagy[60-62].

Aberrant HDAC function has been implicated in numerous diseases. Misregulation of acetylation has been connected to cancer[63], neurological diseases

including Alzheimer's[64-66], and immune disorders[67]. The potential role of HDACs in the onset of these disorders has led to a large amount of research into HDAC inhibition. There are currently four FDA-approved inhibitors of Class I HDACs for use in treatment of T-cell lymphoma and multiple myeloma[63] (**Figure 1.4**, denoted with an asterisk). These four inhibitors are not enzyme-selective and have a deleterious effect on the health of the patient[63]. Novel inhibitor research has become a focus of the HDAC field due to the clinical potential but many clinical results of new inhibitors, both selective and non-selective, have been disappointing[63, 68].

Inhibitor Development

The previously developed inhibitors generally have a consistent structure: zinc binding group (ZBG) – linker – cap. The divalent active site metal allows development of inhibitors with a metal-binding group to displace the metal-bound water nucleophile (**Figure 1.3**). The linker region mimics the lysine side chain to fill the internal substrate binding tunnel. The cap group enhances binding to the enzyme by forming additional contacts on the outer protein surface.

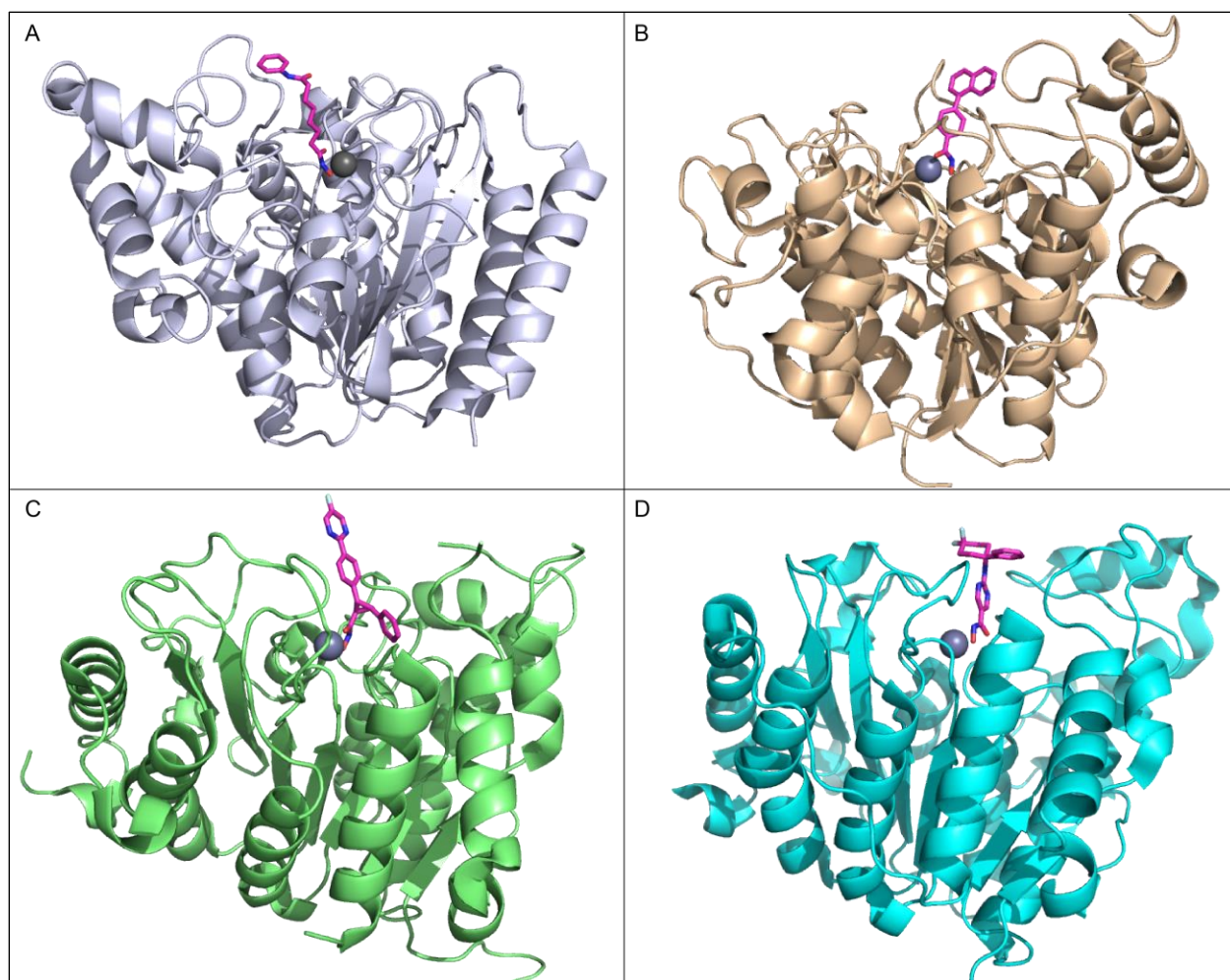


Figure 1.3: Crystal structures of HDACs with inhibitors bound. Inhibitor structures are given in **Figure 1.4**. A) HDAC2 complexed with vorinostat (PDB 4lxz). B) HDAC8 complexed with a selective hydroxamic acid inhibitor 1 (PDB 5fcw). C) HDAC4 complexed with a selective carboxamide inhibitor 2 (PDB 4cbt). D) HDAC6 complexed with selective inhibitor ACY-1083 (PDB 5WGM).

Hydroxamates are the most used ZBG due to their potent metal binding affinity. However, the hydroxamate moiety has poor pharmacokinetics due to its reactivity and is not stable *in vivo* for more than an hour[69]. There are other groups used as ZBGs, including benzamide, carboxylic acids, thiols, and trifluoroketones[70] (**Figure 1.4-A**). However, these moieties have been unable to match the potency of the hydroxamate, with a few exceptions. Research has begun to focus on discovery of novel ZBGs.

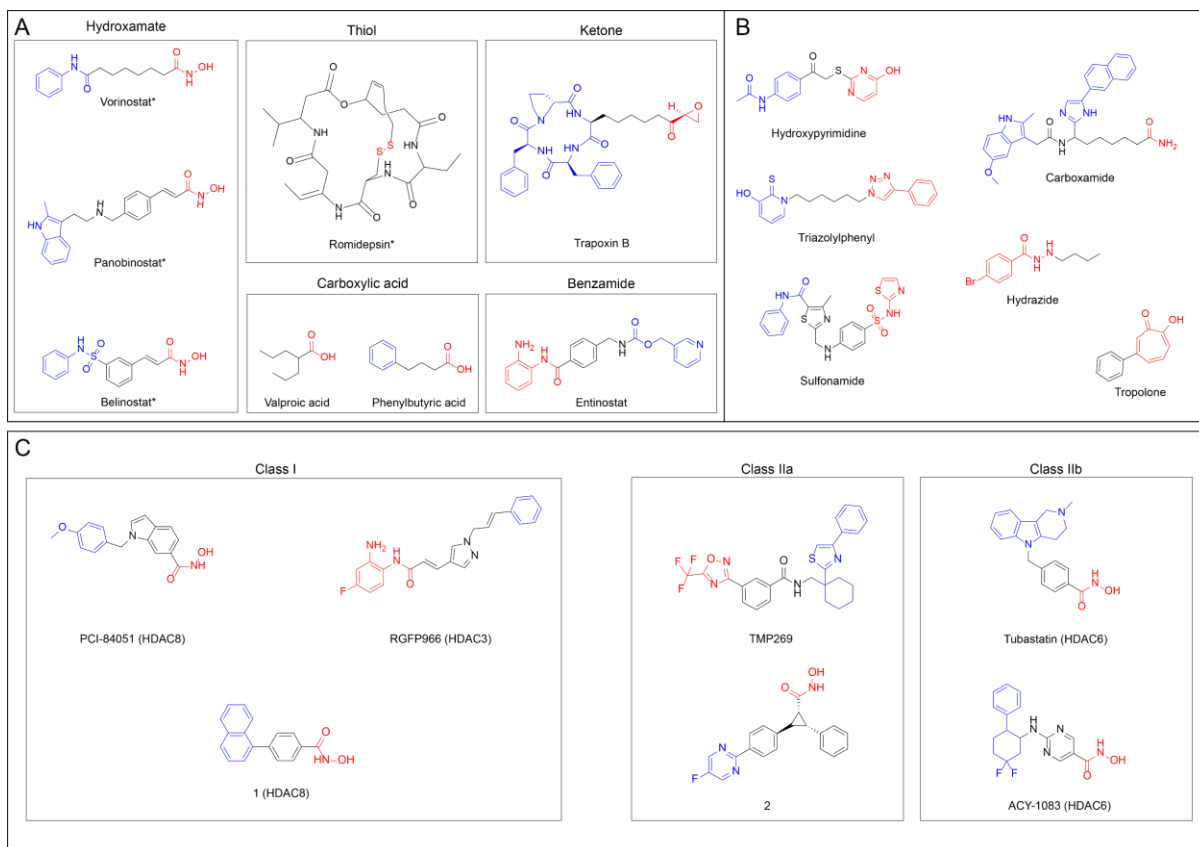


Figure 1.4: Inhibitor structures

The zinc binding group is shown in red, the linker in black, and the cap group in red. (A) Examples of HDAC inhibitors containing traditional zinc binding groups. (B) Examples of HDAC inhibitors containing novel zinc binding groups. (C) Examples of isozyme selective HDAC inhibitors.

There have been numerous attempts to discover novel non-hydroxamate ZBGs. Numerous chemical groups have been used as ZBGs to develop both selective and non-selective HDAC inhibitors. These groups include hydroxypyrimidines, hydrazides, triazolylphenyls, tropolones, sulfonamides, and carboxamides[63, 71-74] (**Figure 1.4-B**). The work into novel ZBGs is encouraging for developing inhibitors with better pharmacokinetics. Though, the overall focus of inhibitor development is still on the traditional ZBGs, with most patents filed from 2012-2017 containing hydroxamates or benzamides, and very few containing a novel ZBG[68].

There is an additional focus on development of isozyme selective HDAC inhibitors with the goal of maintaining biological activity while decreasing toxicity. Crystal structures and molecular docking studies have provided insight into the development of these selective inhibitors. Though all HDACs utilize the same catalytic mechanism, there are enough structure differences between the enzymes to allow for selective binding. HDAC6 has been an exceptional hotspot for inhibitor development with numerous inhibitors developed based on the structure of tubastatin, which contains a hydroxamate ZBG and a large aryl, cap group. Selectivity for HDAC6 seems to be determined by this large hydrophobic, cap group[72]. Selective inhibitors for Class I enzymes utilize structural elements that interact close to the active site and within the internal cavity[72]. PCI-34051, an HDAC8 inhibitor, is an example of a unique linker structure enabling selectivity[75]. RGFP966, a HDAC3 inhibitor, demonstrates the importance of a modified ZBG structure conferring selectivity[76]. Lastly, several selective inhibitors have been developed for Class IIa HDACs. These selective inhibitors contain bulkier aromatic groups around the hydroxamate ZBG instead of a leaner linker[72]. These bulkier groups interact with hydrophobic pockets near the active site to provide for Class IIa inhibitor selectivity[73].

HDAC inhibitor development shows promise for clinical treatment of multiple diseases. However, there is a lack of understanding of the biological pathways such inhibitors would affect. This chapter will summarize the current understanding of biological activity of HDACs and how that activity is regulated.

Importance of Protein-Protein Interactions

Many HDAC do not function alone, rather they form both stable and transient interactions with other proteins. BioGRID, the Biological General Repository for

Interaction Datasets, lists almost 3,000 protein interacting partners for the 11 HDACs[77]. Almost half of those protein interactions are connected to HDAC1 and 2 (**Figure 1.5-A-D**)[77]. In 2013, the most comprehensive study was performed to map the interactome of the HDAC family[78]. This study determined interactions in T-cells by immunoprecipitation of the GFP-tagged HDACs followed by protein identification using mass spectrometry[78]. While the study has some limitations, including only one cell type was analyzed and the use of a bulky GFP tag could prevent some interactions, it does demonstrate the extensive network of HDAC-protein interactions related to numerous biological processes, including many previously unknown interactions[78].

Protein interactions with HDAC1, 2, and 3 are proposed to be necessary for activity. When purified to homogeneity these isozymes have little to no deacetylase activity[79]. HDAC1, 2, and 3 function in stable nuclear complexes (**Figure 1.5-E**). HDAC1 and HDAC2 are found within 3 complexes: Sin3, NuRD, and CoREST[80]. HDAC3 associates with nuclear receptor repressors, SMRT and NCoR[79]. Unlike its Class I counterparts, HDAC8 is catalytically active *in vitro* in a purified, isolated state and does not require any binding partners for deacetylase activity[15]. This is most likely not the case *in vivo* as immunoprecipitated HDAC8 contained several proteins, including proteins involved with the cell cycle and protein transport[78]. The interaction network for HDAC8 is not as extensive as HDAC1, 2, and 3, where there are far more interactions with proteins related to gene expression[78]. The T-cell study did reveal some interesting new HDAC1 interactions involved in protein transport and metabolism[78].

The Class IIa enzymes, as they are catalytically inactive, derive all their function from protein interactions. Interaction with 14-3-3 signaling proteins controls nuclear

localization of the Class Ila enzymes, which in turn controls their ability to interact with MEF2 transcription factors to repress transcription[30, 32, 56, 81]. Class Ila enzymes also interact with the HDAC3/SMRT/NCOR complex[24]. The interaction network found in T-cells between HDAC4, 5, and 7 was found to be rather redundant and predictable between the isozymes as only known interactions were identified[78]. The remaining Class Ila enzyme, HDAC9 demonstrated a distinctly different interaction network with novel interactions identified with proteins related to gene expression, the cell cycle, and RNA processing[78].

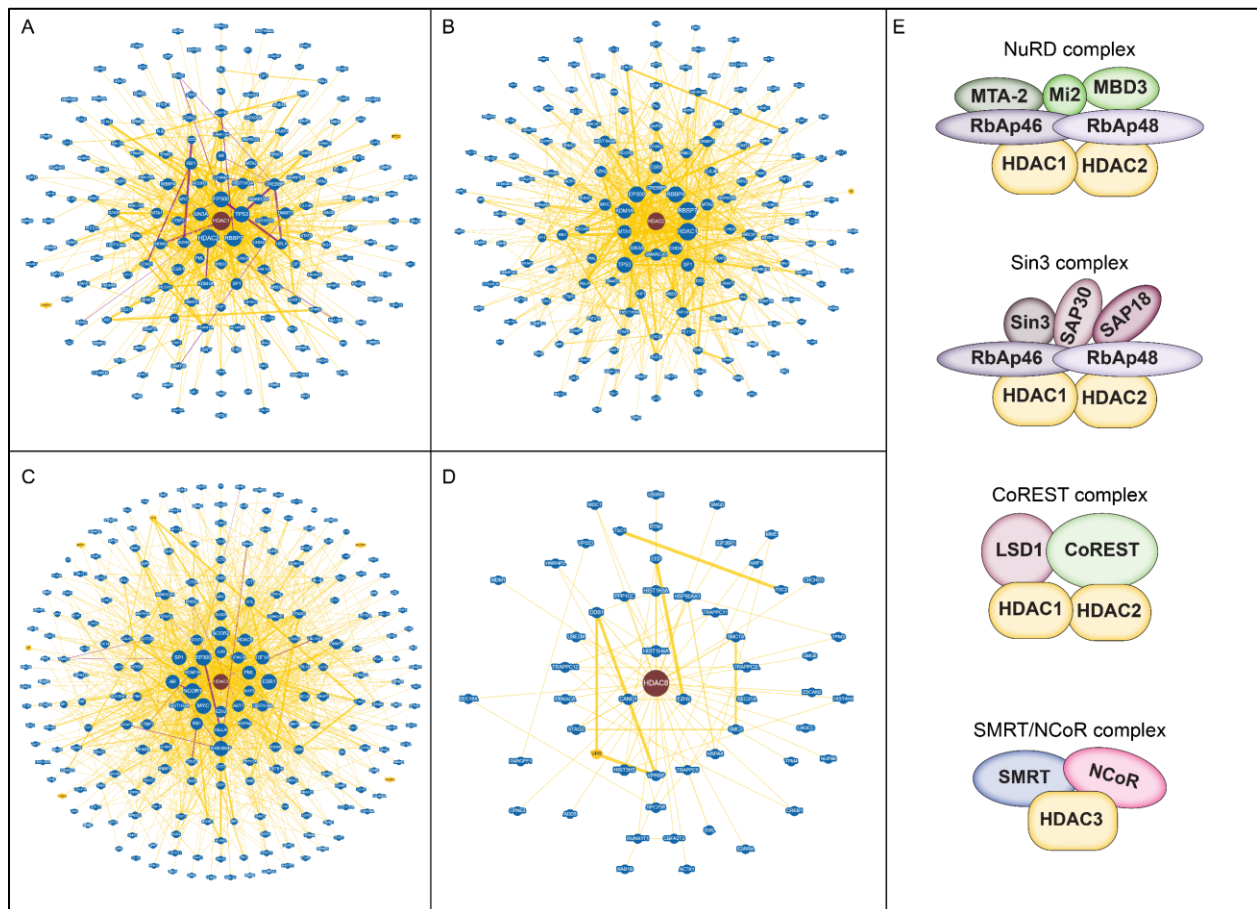


Figure 1.5: Interaction networks of Class I HDACs
 Obtained from BioGRID[77, 82]. (A) HDAC1. (B) HDAC2. (C) HDAC3. (D) HDAC8. (E) Composition of stable nuclear complexes of Class I HDACs

HDAC6 is known to have protein-protein interactions that are more transient in comparison to the stable Class I complexes. HDAC6 uniquely contains a zinc finger, ubiquitin binding domain with high affinity for ubiquitinated proteins[83]. HDAC6 uses this domain to function as a linker between aggregated protein and the motor protein dynein, to facilitate the aggregate's transport to the aggresome[60]. Additionally, a complex containing HDAC6, HDAC11, and HDAC2 has been proposed at the vitamin D receptor involved in the regulation of transcription factor, *MYC*[84]. The HDAC6 interaction network consisted of previously unidentified interactions with proteins involved in ubiquitination, RNA processing, the cell cycle, gene expression, protein transport, and metabolism[78]. A limited interaction network was found for HDAC10[78].

While not much is known about HDAC11, immunoprecipitation of the isozyme revealed an astonishing 124 interacting proteins, a select number of which were validated by further co-immunoprecipitation experiments[78]. The identified proteins are involved in diverse biological processes: gene expression, the cell cycle, RNA processing, ubiquitination, signal transduction, protein transport, protein folding, and metabolism[78]. Additional studies to explore these potential interactions could reveal extensive functionality of HDAC11.

It is apparent HDAC enzymes function within a network of protein interactions. These interactions no doubt affect HDAC catalytic activity and substrate selectivity.

Substrate Discovery

In 2009, immunoaffinity enrichment of acetylated peptides coupled to high-resolution mass spectrometry identified thousands of acetylation sites across the proteome[85]. In the 11 years since that study the number of acetylated proteins and

acetylation sites has only grown. PhosphoSitePlus, a database of post-translational modifications, indicates ~38,000 acetylation sites on ~20,000 proteins have been identified by high-throughput and low-throughput analyses[6]. Many validated acetylated proteins have been linked to sirtuins as the deacetylation enzyme. However, selective sirtuin inhibitors are available to determine enzyme-specific function in contrast to unselective pan-HDAC inhibitors which could contribute to this finding. Despite this, there have been a limited number of acetylation sites where deacetylation is attributed to specific HDACs.

The identification of an acetylated protein as a substrate of a HDAC is traditionally determined by a few common methods. Potential substrates can be identified by *in cellulo* deacetylase inhibition, transient transfection to over-express a HDAC isozyme, and/or knockdown of HDAC expression using siRNA. The effectiveness of these methods can be debated. The lack of isozyme-specific inhibitors limits the detection of isozyme-specific substrates. The functional redundancy of the isozymes could also limit the effectiveness of inhibiting or knocking down expression of one isozyme. Additionally, transient transfection to over-express an isozyme creates an artificial system and assumptions must be made as to whether other pathways are affected that could influence acetylation levels. Furthermore, these methods typically monitor protein acetylation levels by immunoblotting, so substrate discovery is both limited and biased due to the protein of interest needing to be predetermined.

Acetylation sites on histone tails have been described as *in vivo* substrates for all Class I HDACs. Histones have been further validated as substrates through *in vitro* deacetylation assays with both peptide analogs and full-length protein. Site selectivity has

additionally been explored for histones using both hyperacetylated full-length histones, where acetylation level was detected with site-specific antibodies, and singly acetylated synthetic peptide analogs[86-88]. HDACs 1, 2 and 3 catalyze deacetylation of all the analyzed acetylated lysines in histone tails but with varying efficiency[86-88] (**Figure 1.6**). It should be noted the activity for HDAC1, 2, and 3 are for their immunopurified, complexed states whereas the activity for HDAC8 is from *E. coli* recombinant enzyme. The different complex components surely infer selectivity, but the substrate selectivity of each complex is unknown.

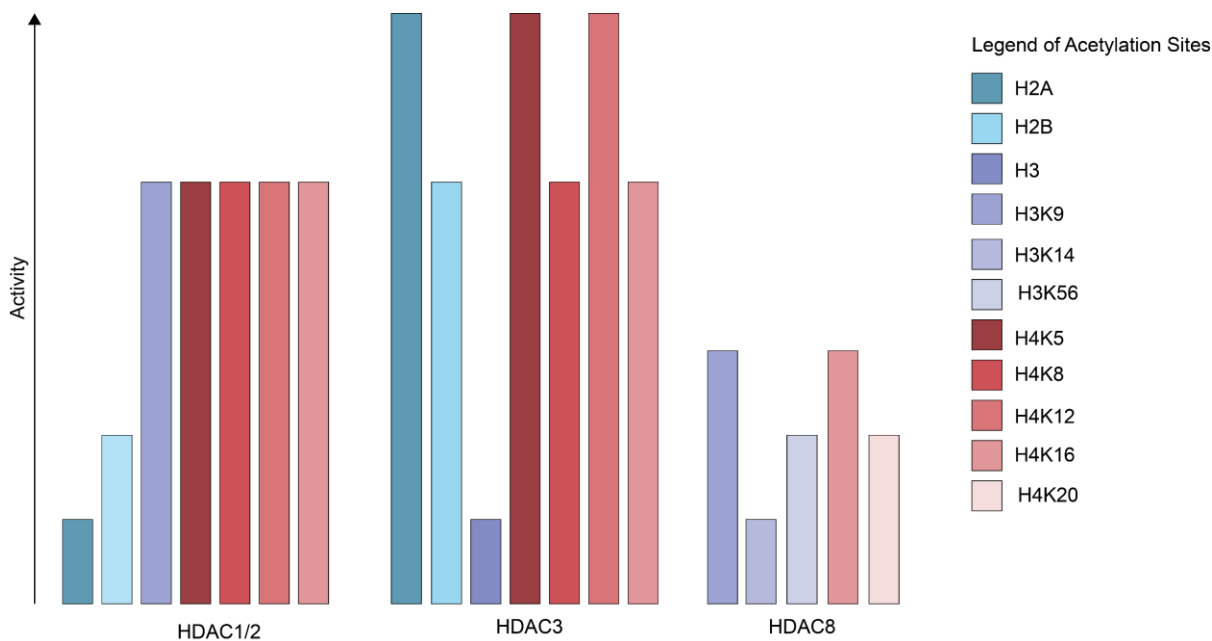


Figure 1.6: Relative activity of Class I HDACs on various histone acetylation sites

Class I HDACs have been linked to several other nuclear non-histone substrates. The transcription factor, p53, was one of the first proposed non-histone protein substrates. Deacetylation of p53 has been linked to multiple isozymes through non-selective *in cellulo* inhibition and transient transfection, as well as *in vitro* peptide analogs[46, 47, 89]. The following protein substrates have only been proposed through

cellular experiments and lack *in vitro* evidence. Both HDAC1 and HDAC2 are proposed to catalyze deacetylation of the CCR4-associated factor 1 (CAF1) that regulates mRNA decay[90]. HDAC1 is additionally proposed to deacetylate DNA methyltransferase 1 (DNMT1) preventing degradation of this protein [52]. HDAC2 is proposed to deacetylate TP53-binding protein 1 (53BP1). Deacetylated 53BP1 binds to damaged chromatin to induce DNA double-strand break repair. HDAC2 has also been proposed to deacetylate phosphatase PTEN down-regulating activity and recruitment to signaling complexes. Lastly, HDAC3 is proposed to deacetylate p65 leading to the nuclear export of the nuclear factor kappa B (NF- κ B) transcription factor complex, repressing transcription[91].

HDAC8-specific protein substrates have been validated both *in vivo* and *in vitro*. The best validated HDAC8 substrate is acetylated structural maintenance of chromosomes protein 3 (SMC3), identified through loss of function mutations in the HDAC8 gene that occur in some patients with the developmental disorder, Cornelia de Lange syndrome[50]. Acetylated SMC3 has been further validated by HDAC8-specific *in cellulo* inhibition and *in vitro* deacetylation by HDAC8[50]. Estrogen related receptor-alpha ($ERR\alpha$) deacetylation is also proposed to be catalyzed by HDAC8, increasing $ERR\alpha$'s DNA binding affinity[92]. Acetylated $ERR\alpha$ has been validated by *in vitro* co-immunoprecipitation with HDAC8 and *in vitro* deacetylation by HDAC8[92].

HDAC6 has been linked to numerous cytosolic substrates. The most extensively validated proposed protein substrates are heat shock protein 90 (HSP90) and tubulin. HDAC6 has been shown to associate with HSP90 and tubulin/microtubules *in vivo*[58, 61]. Inhibition of HDAC6, in addition to over-expression has also been shown to affect the acetylation of HSP90 and α -tubulin *in vivo*[58, 61]. The effect of deacetylation is opposite

for the two substrates as deacetylation enhances HSP90 chaperone activity[61] and deacetylation of tubulin increases microtubule instability[58]. HDAC6 catalytic activity towards HSP90 and α -tubulin, when assembled in microtubules, has also been validated *in vitro*[58, 61]. HDAC6 has additionally been proposed to deacetylate other protein substrates, however they have not been validated *in vitro*. HDAC6 is proposed to catalyze the deacetylation of cortactin, promoting binding of cortactin to F-actin, enhancing cell motility[57], and the protein tau, promoting microtubule binding and preventing pathological aggregation[93-95]. HDAC6 has also been linked to nuclear substrates. HDAC6 is proposed to deacetylate RNA binding protein TDP43 preventing the aggregation of the hyperphosphorylated form of TDP43[62]. The aggregated form of TDP43 is implicated in the development of amyotrophic lateral sclerosis (ALS).

Lastly, HDAC11 associates with and is proposed to deacetylate the replication licensing factor, Cdt1, protecting this protein from proteasomal degradation[96]. HDAC11 additionally associates with and is proposed to deacetylate forkhead box P3 (Foxp3) in T-regulatory cells[43]. Both protein substrates are proposed solely by *in vivo* transient transfection and over-expression of HDAC11[43, 96].

The substrates summarized here likely represent a small number of thousands of potential substrates, as measured by the existence of acetylated proteins in cells. It is apparent traditional experiments are inefficient at identifying the potential multitude of substrates these enzymes act on. Promising high-throughput ways of identifying substrates have been developed in recent years.

The HDAC Activity “Toolbox”

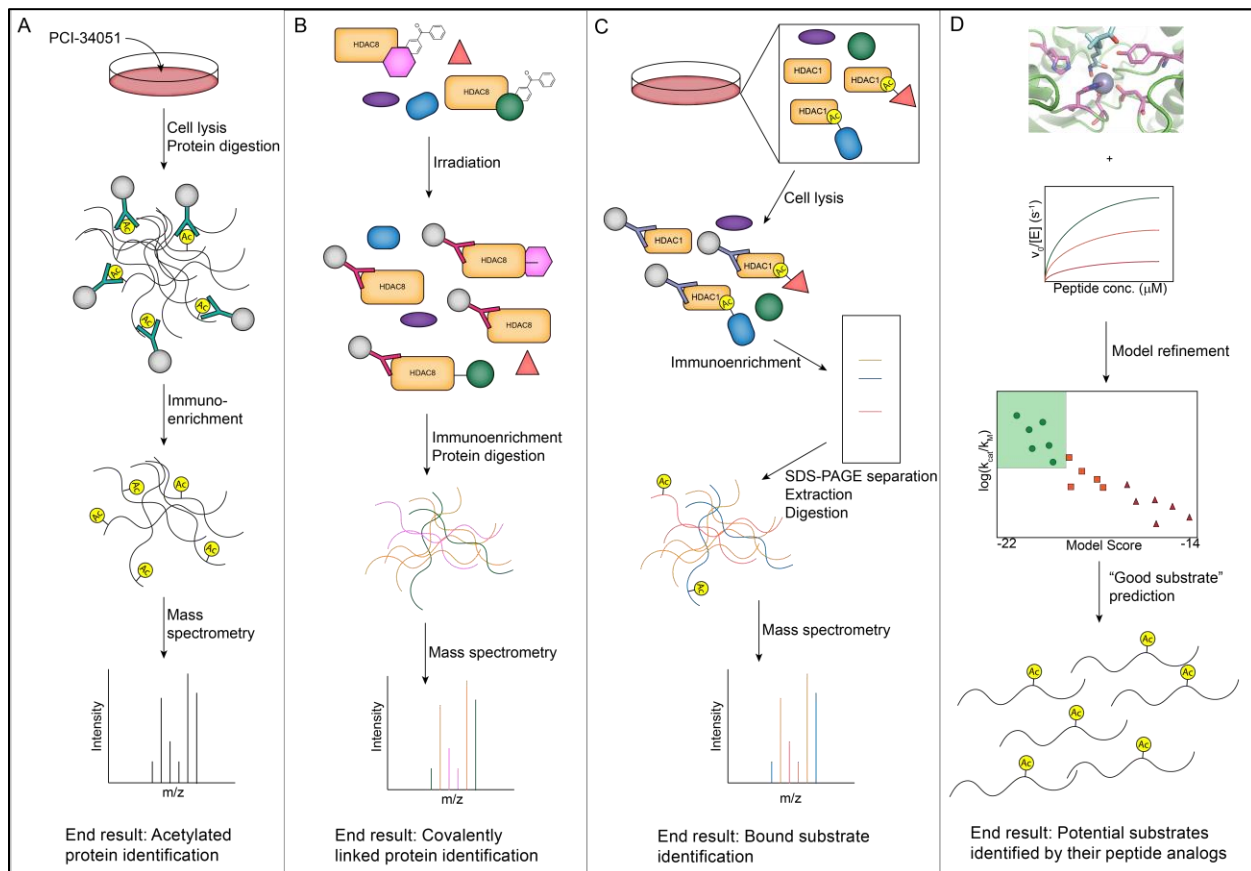


Figure 1.7: New substrate discovery methods.

A) Scheme of *in cellulo* inhibition. B) Scheme of active site photo-crosslinking. C) Scheme of substrate trapping. D) Scheme of active site computational modeling.

One method that has been utilized to determine HDAC-specific substrates is *in cellulo* inhibition. In 2014, Olson et al. used the HDAC8-specific inhibitor, PCI-34051, to increase acetylation of HDAC8 substrates in cell culture[97]. The proteins with substantially increased acetylation due to HDAC8 inhibition were identified using SILAC (stable isotopic labeling of amino acids in cell culture) mass spectrometry (**Figure 1.7-A**). This study was not without its limitations. Only seven proteins were identified as high-confidence HDAC8 substrates. The lack of substrates identified could be due to the

redundant activity of HDACs; once HDAC8 was inhibited another HDAC was able to fill in. The lack of identified substrates could also be due to the low cellular viability of the hydroxamate inhibitor. The experiment was run over 24 hours, much longer than the 1-hour lifespan of hydroxamates in the cell. By the end of the 24 hours a portion of HDAC8 activity would have been restored. Furthermore, this experiment does not demonstrate HDAC8 directly catalyzes deacetylation of these substrates; alteration of the activity of another protein by the inhibitor could lead to decreased acetylation.

The limitations seen with inhibitor use to identify substrates have promoted development of substrate-trapping methods. Two methods that have been developed are active site photo-crosslinking (**Figure 1.7-B**) and inactive trapping mutants (**Figure 1.7-C**). Active site photo-crosslinking has been used to identify over 100 potential HDAC8 substrates[98]. This method incorporates a non-canonical amino acid at an amber stop codon to genetically encode a photo-crosslinker, in this case *p*-benzoyl-L-phenylalanine, proximal to the active site of HDAC8. Transient substrate-enzyme interactions are covalently captured by the photo-crosslinker upon activation with light. In this experiment human tissue culture cell lysates were incubated with recombinant labelled-HDAC8 and potential substrates captured by covalent crosslinking. The identity of the captured proteins was determined by bottom-up mass spectrometry with high stringency controls. The method is high-throughput and effective, identifying the largest number of potential substrates. However, the use of cell lysates and recombinant protein is a limitation due to the loss of cellular compartmentalization and possible disruption of protein complexes.

In contrast, inactive substrate trapping mutants are expressed *in vivo*. The mutants are optimized to display no activity but high substrate binding allowing the isolation of the

enzyme-substrate complex. The method has been successfully used with HDAC1 to identify LSD1 as a novel substrate[99] and identified p53 as a control substrate[100]. The two substrates were identified using two different mutants, indicating screens of multiple mutants must be performed to ensure widespread coverage of substrate interactions. Furthermore, the mutations may alter substrate selectivity and/or complex formation. When this method was coupled to a proteomic, mass spectrometry analysis only six high-stringency hits were found[101].

A last method to help identify substrates is computational modeling. Alam et al. developed a structure-based approach to identifying substrates of HDAC8 (**Figure 1.7-D**) [102]. Based on the Rosetta FlexPepBind framework, the computational model predicts “good” substrates. The model fits peptide analogs of acetylation sites into the active site of HDAC8, determined by crystallography, and calculates if the fit is good enough for efficient binding and catalysis. The HDAC8 model demonstrated a linear correlation between the predicted binding and the $\log k_{cat}/K_M$ for catalysis of peptide deacetylation and predicted the fastest HDAC8 peptide substrate described thus far. This method is more comprehensive than the others discussed. Due to the computational nature of the screen all the acetylation sites identified in the proteome can be rapidly screened as a potential HDAC8 substrate. However, this model format assumes substrate selectivity can be predicted at the peptide level.

In addition to discovering substrates, the deacetylase activity with the proposed substrates must be determined. The kinetic parameter, k_{cat}/K_M , evaluated for each substrate gives insight into substrate selectivity. Deacetylase activity assays have been developed for peptide analogs of substrates and for full-length protein substrates. The

commercially available Fluor de Lys assay measures deacetylase activity with peptide analogs[103]. The peptides contain an acetylated lysine residue within a 5-mer peptide and a C-terminal methyl-coumarin fluorescent tag. Upon deacetylation, the coumarin tag is released by reaction with trypsin which leads to a change in fluorescence (**Figure 1.8-A**). The fluorescent tag dramatically increases ease of use and sensitivity of deacetylation activity. However, it is cost prohibitive to fluorescently tag all peptides one may want to test, outside of those commercially available. The methyl-coumarin tag interacts with HDACs and enhances the catalytic activity, so results may not be directly applicable to selectivity of *in vivo* substrates.

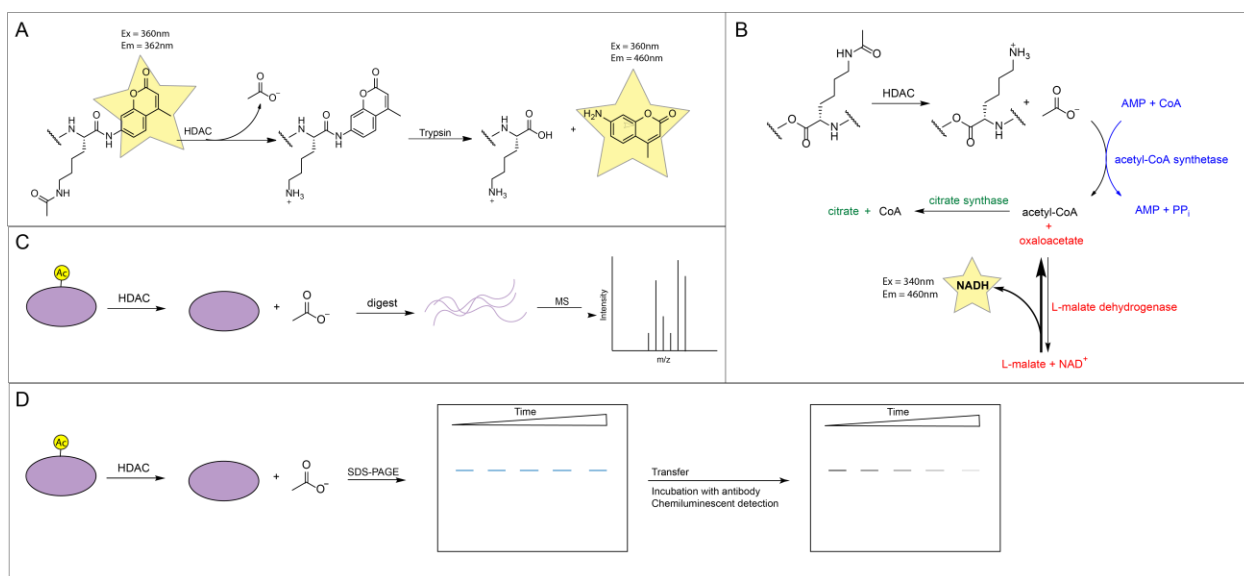


Figure 1.8: HDAC activity assays

(A) Commercially available Fluor de Lys assay. (B) Enzyme-coupled acetate assay. (C) Peptide mass spectrometry assay. (D) Immunoblotting assay.

An enzyme coupled assay measuring acetate production developed by Wolfson et al. allows determination of deacetylase activity on non-labeled peptides[104]. The assay couples the production of acetate to NADH formation, harnessing enzymes from the citric acid cycle, which can be monitored via fluorescence[104] (**Figure 1.8-B**). The

enzyme coupled assay broadens the scope of peptide substrates to be tested and has been used to validate potential substrates identified from some of the previous methods.

The enzyme coupled assay is still limited by its use of peptide analogs of substrates. Castaneda et al. demonstrated that HDAC8 catalytic efficiency increases when full-length acetylated H3/H4 is used versus the peptide analog, thus biologically relevant rates must be determined using full-length substrates[88]. Castaneda et al. used genetically encoded non-canonical amino acid incorporation to express singly acetylated histones. The catalytic rate was then determined through mass spectrometry of the peptide fragments of the histones after they had been subjected to deacetylation by HDAC8 (**Figure 1.8-C**). Deacetylation of acetylated full-length proteins has also been measured by anti-acetyl-lysine immunoblotting (**Figure 1.8-D**). The catalytic rate is determined by the disappearance of acetylated protein over time. This immunoblotting method has been demonstrated using both HDAC8 and HDAC1[105, 106].

Discovery of HDAC-specific substrates using these methods has greatly evolved our understanding of the biological function of these enzymes. HDAC's biological function cannot be fully determined by just substrates, however, as there are many other factors that regulate function and selectivity.

Identity of Active Site Metal

As mentioned, HDACs are metal-dependent deacetylases. It is thought the active site metal ion is zinc(II), based on the first crystal structure of HDAC8 showing a zinc ion in the active site[107]. However, the His-Asp-Asp metal binding site is uncharacteristic for a zinc binding site[108] raising the possibility another metal might be more suited. The effect of different divalent metals in the active site has been extensively studied using

HDAC8. HDAC8 has thus become the most extensively biochemically characterized HDAC and serves as a model for the HDAC catalytic mechanism.

Crystal structures have been solved with multiple divalent metals, including Co(II), Zn(II), Fe(II), and Mn(II)[109]. No significant structural differences in a hydroxamate-bound active site are apparent. However, the deacetylase activity is dependent on the identity of the reconstituted metal as follows: Co(II)>Fe(II)>Zn(II) and minimal activity is observed with Ni(II) and Mn(II)[110]. Due to the intercellular availability of the divalent metals, it is proposed the *in vivo* active site ion is Zn(II) or Fe(II). The K_D values of HDAC8 for Zn(II) and Fe(II) are significantly different, 5pM and 2 μ M respectively[111]. However, the free Zn(II) cellular concentration is estimated in the picomolar range[112, 113] and coincidentally the available Fe(II) concentration is within the micromolar range[114, 115]. Considering the K_D values are comparable to cellular concentrations of the two metals, it cannot be determined based on the affinities and concentrations of Zn(II) and Fe(II) which is the *in vivo* active site metal.

It is possible HDAC8 exists as both a Zn(II) and Fe(II) deacetylase. When expressed in *E.coli*, recombinant HDAC8 contains more iron than zinc[110]. Furthermore, the activity in *E.coli* cell lysates is oxygen-sensitive[110], indicating the presence of oxygen-sensitive Fe(II). It has previously been shown a bacterial deacetylase, UDP-3-O-((R)-3-Hydroxylmyristoyl)-*N*-acetylglucosamine Deacetylase (LpxC), initially labeled as a zinc-dependent deacetylase, is an iron-dependent deacetylase and undergoes a regulatory metal-switching mechanism[116, 117]. The similarities between the two deacetylases, LpxC like HDAC8, is more active with an iron cofactor than zinc, could

indicate HDAC8 additionally has a metal-switching regulatory mechanism. However, there is currently little information about the identity of the metal bound to HDAC8 in cells.

In addition to increased activity with Fe(II), the peptide selectivity of HDAC8 is dependent on the identity of the reconstituted metal. Castaneda et al. demonstrated that HDAC8 activity with acetylated peptide substrates varies between 2-150 fold when the reconstituted metal switched from Zn(II) to Fe(II)[118]. Additionally, the most substantial differences were seen with peptide analogs of potential substrates identified through a proteomic screen. Such a finding suggests that HDAC8-Fe(II) may be an important *in vivo* species and increases the probability that a metal-switching mechanism is involved in biological HDAC8 function.

The exploration of different active site metals with HDAC8 has elucidated a potential important regulatory mechanism of HDACs. It is clear the identity of the divalent metal in HDACs is important for activity and selectivity.

Post-translational Modifications

As with many proteins, HDAC's are regulated by post-translational modifications. HDACs can be acetylated, glycosylated, ubiquitinated, sumoylated, and phosphorylated[6]. The most extensively studied modification is phosphorylation as almost every HDAC is phosphorylated, but the PTM occurs at different sites with different effects.

HDAC1 is doubly phosphorylated at Ser421 and Ser423. HDAC2 is phosphorylated at the corresponding residues, Ser422 and Ser424. The phosphorylation of HDAC1 and HDAC2 seems to be essential for catalytic activity. Pflum et al. found that HDAC1 mutants with the phosphorylated serine changed to alanine reduced deacetylase

activity by 80% and the corresponding phosphomimetic mutants (glutamate) rescued activity[119]. The results could not differentiate whether both phosphorylation sites were necessary for activity, but mass spectrometry analysis of immunoprecipitated HDAC1 demonstrated phosphorylation of both sites[119]. The study additionally found that phosphorylation was important for complex formation. The mutant HDAC1 that removed the phosphorylation sites did not co-immunoprecipitate with known complex partners[119]. It should be noted that the effect of phosphorylation on activity has only been observed with full-length protein substrates, e.g. histones. There is no appreciable difference between the activity of phosphorylated and non-phosphorylated HDAC1 with peptide substrates[120]. This indicates that the changes caused by phosphorylation likely do not directly affect the activity of the metal site. However, it is unknown if such an effect is caused by the lack of interacting proteins with non-phosphorylated HDAC1. HDAC3 is similarly activated by phosphorylation. HDAC3 is phosphorylated on an unconserved residue, Ser424. Like HDAC1 and HDAC2, the modification is critical for enzymatic activity[121].

The last Class I member, HDAC8, is also regulated by phosphorylation. HDAC8 is unique in the Class I enzymes, in that phosphorylation inhibits deacetylase activity. Lee et al. found HDAC8 is phosphorylated at Ser39, an unconserved residue within the catalytic domain. An unphosphorylated mutant (S39A) had no effect on activity, whereas a phosphomimetic mutant (S39E) significantly decreased deacetylase activity with purified histones[122]. Recently, Welker Leng et al. found that the magnitude of the activity decrease of the phosphomimetic mutant (glutamate) was dependent on peptide substrate sequence. Activity decreases varied between 10 and 100-fold dependent on

the peptide substrate[123]. The study demonstrates a potential role of phosphorylation in not only regulating activity but also regulating selectivity of HDACs.

Phosphorylation of Class IIa HDACs affects the cellular localization and in doing so, their regulation of transcription. All members of Class IIa (HDAC4, HDAC5, HDAC7, and HDAC9) are phosphorylated at three, conserved serine residues. Phosphorylation at all three residues induces binding of 14-3-3 proteins, preventing the identification of the nuclear localization signal[30, 32]. The HDACs are then transported out of the nucleus into the cytoplasm. This change in cellular localization prevents Class IIa HDACs from interacting with myocyte enhancing factor-2 proteins (Mef2)[31, 81, 124]. Mef2 proteins are then free to activate gene transcription. In that respect, phosphorylation of Class IIa HDACs inhibits their transcriptional repression activity.

HDAC6 has the most complex regulation by phosphorylation. HDAC6 does not require phosphorylation to be active but HDAC6 is found *in vivo* to be phosphorylated at multiple sites, and the change in activity is dependent on the site. Deribe et al. found phosphorylation by epidermal growth factor receptor kinase (EGFR) at Tyr570 abolishes deacetylase activity as measured by increased acetylation of α -tubulin *in vivo*[125]. In contrast, Williams et al. found extracellular signal-regulated kinase 2 (ERK2) phosphorylates Ser1035 of HDAC6 through the EGFR-Ras-Raf-MEK-ERK signaling cascade, stimulating cell migration[126]. Ser1035 phosphorylation enhances *in vitro* deacetylase activity with α -tubulin as a substrate but does not affect *in vitro* deacetylase activity towards core histones[126]. Additional studies further demonstrate regulation of HDAC6 activity by phosphorylation. Ser458 and Ser22 are also phosphorylation sites that stimulate HDAC6 deacetylase activity. Prevention of phosphorylation at Ser22 decreased

HDAC6 catalytic activity when measured in cell lysates using a peptide fluorometric assay[127]. Phosphorylation at Ser458 increased recombinant HDAC6 catalytic activity when measured *in vitro* using a peptide fluorometric assay. [127, 128]. HDAC6 phosphorylation has also been linked to viral sensing, where phosphorylation by protein kinase c-alpha ($PKC\alpha$) enhances β -catenin deacetylation leading to interferon transcription[129]. Regulation of HDAC6 by phosphorylation demonstrates how deacetylase activity is crucial to biological signaling pathways.

HDAC10 and HDAC11, the least studied HDACs, have no confirmed, regulatory PTM sites. PhosphoSitePlus lists several PTMs found for both enzymes, however they have only been identified through high-throughput proteomic screens[6]. Based on the rest of the HDAC family, HDAC10 and HDAC11 are most likely regulated by phosphorylation or other PTMs but further research is needed to determine the effect.

Modifications of HDACs, particularly phosphorylation, greatly impact the enzyme activity. HDAC8, and potentially HDAC6, serve as models that phosphorylation changes both activity and substrate selectivity. Involvement in specific biological pathways could be dependent upon phosphorylation and the site of phosphorylation.

Conclusion

This dissertation seeks to further explore the factors that affect the activity and selectivity of histone deacetylases. The work focuses on three HDACs: HDAC1, HDAC6, and HDAC8. In Chapter 2, HDAC8 will be used as a model to attempt to elucidate the *in vivo* active site metal by metalloform-specific inhibition with the identification of several metal-binding groups selective for HDAC8-Fe(II) versus HDAC8-Zn(II). Novel non-hydroxamate metal chelator fragments were also analyzed with HDAC8 for inhibitor

development with the tropolone metal binding group identified as the most potent lead molecule for future studies.

In Chapter 3, HDAC1 will serve as a model for elucidating how complexation and protein interactions affect substrate selectivity. This was accomplished by reconstituting the HDAC1-containing CoREST complex *in vitro* and determining activity across multiple potential substrates. These data demonstrated that protein interactions enhance HDAC1 deacetylase activity significantly and activity enhancement can be achieved regardless of the expression system.

Finally, in Chapter 4, the computational substrate discovery method will be tested on HDAC6, to determine if the substrate selectivity for Class IIb HDACs can be determined at the peptide level. A computational model to predict good substrates of the second catalytic domain of HDAC6 was successfully created, though this domain demonstrated low substrate sequence selectivity overall. The addition of the first catalytic domain changed selectivity and activity, indicating that the full structure is important for determining the selectivity profile of HDAC6.

References

1. Fischer, E.H., et al., *Structure of the site phosphorylated in the phosphorylase b to a reaction*. J Biol Chem, 1959. **234**(7): p. 1698-704.
2. Phillips, D.M., *The presence of acetyl groups of histones*. Biochem J, 1963. **87**: p. 258-63.
3. Allfrey, V.G., R. Faulkner, and A.E. Mirsky, *ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS*. Proc Natl Acad Sci U S A, 1964. **51**: p. 786-94.
4. Piperno, G. and M.T. Fuller, *Monoclonal antibodies specific for an acetylated form of alpha-tubulin recognize the antigen in cilia and flagella from a variety of organisms*. J Cell Biol, 1985. **101**(6): p. 2085-94.
5. Gu, W. and R.G. Roeder, *Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain*. Cell, 1997. **90**(4): p. 595-606.
6. Hornbeck, P.V., et al., *PhosphoSitePlus, 2014: mutations, PTMs and recalibrations*. Nucleic Acids Res, 2015. **43**(Database issue): p. D512-20.
7. Hebbes, T.R., A.W. Thorne, and C. Crane-Robinson, *A direct link between core histone acetylation and transcriptionally active chromatin*. The EMBO journal, 1988. **7**(5): p. 1395-1402.
8. Turner, B.M. and G. Fellows, *Specific antibodies reveal ordered and cell-cycle-related use of histone-H4 acetylation sites in mammalian cells*. Eur J Biochem, 1989. **179**(1): p. 131-9.
9. Vidal, M. and R.F. Gaber, *RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in Saccharomyces cerevisiae*. Molecular and Cellular Biology, 1991. **11**(12): p. 6317-6327.
10. Kijima, M., et al., *Trapoxin, an antitumor cyclic tetrapeptide, is an irreversible inhibitor of mammalian histone deacetylase*. Journal of Biological Chemistry, 1993. **268**(30): p. 22429-35.
11. Taunton, J., C.A. Hassig, and S.L. Schreiber, *A Mammalian Histone Deacetylase Related to the Yeast Transcriptional Regulator Rpd3p*. Science, 1996. **272**(5260): p. 408-411.
12. Buggy, J.J., et al., *Cloning and characterization of a novel human histone deacetylase, HDAC8*. Biochem J, 2000. **350 Pt 1**: p. 199-205.

13. Gao, L., et al., *Cloning and functional characterization of HDAC11, a novel member of the human histone deacetylase family*. J Biol Chem, 2002. **277**(28): p. 25748-55.
14. Grozinger, C.M., C.A. Hassig, and S.L. Schreiber, *Three proteins define a class of human histone deacetylases related to yeast Hda1p*. Proc Natl Acad Sci U S A, 1999. **96**(9): p. 4868-73.
15. Hu, E., et al., *Cloning and Characterization of a Novel Human Class I Histone Deacetylase That Functions as a Transcription Repressor*. Journal of Biological Chemistry, 2000. **275**(20): p. 15254-15264.
16. Kao, H.-Y., et al., *Isolation and Characterization of Mammalian HDAC10, a Novel Histone Deacetylase*. Journal of Biological Chemistry, 2002. **277**(1): p. 187-193.
17. Van den Wyngaert, I., et al., *Cloning and characterization of human histone deacetylase 8*. FEBS Lett, 2000. **478**(1-2): p. 77-83.
18. Verdel, A. and S. Khochbin, *Identification of a New Family of Higher Eukaryotic Histone Deacetylases: COORDINATE EXPRESSION OF DIFFERENTIATION-DEPENDENT CHROMATIN MODIFIERS*. Journal of Biological Chemistry, 1999. **274**(4): p. 2440-2445.
19. Voelter-Mahlknecht, S. and U. Mahlknecht, *Cloning and structural characterization of the human histone deacetylase 6 gene*. Int J Mol Med, 2003. **12**(1): p. 87-93.
20. Gantt, S.M., et al., *General Base-General Acid Catalysis in Human Histone Deacetylase 8*. Biochemistry, 2016. **55**(5): p. 820-32.
21. Schultz, B.E., et al., *Kinetics and comparative reactivity of human class I and class IIb histone deacetylases*. Biochemistry, 2004. **43**(34): p. 11083-91.
22. Hai, Y., et al., *Histone deacetylase 10 structure and molecular function as a polyamine deacetylase*. Nat Commun, 2017. **8**: p. 15368.
23. Lahm, A., et al., *Unraveling the hidden catalytic activity of vertebrate class IIa histone deacetylases*. Proceedings of the National Academy of Sciences, 2007. **104**(44): p. 17335-17340.
24. Fischle, W., et al., *Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR*. Molecular cell, 2002. **9**: p. 45-57.
25. Kutil, Z., et al., *Histone Deacetylase 11 Is a Fatty-Acid Deacylase*. ACS Chem Biol, 2018. **13**(3): p. 685-693.
26. Moreno-Yruela, C., et al., *Histone Deacetylase 11 Is an epsilon-N-Myristoyllysine Hydrolase*. Cell Chem Biol, 2018. **25**(7): p. 849-856 e8.

27. Li, J., et al., *Histone deacetylase 8 regulates cortactin deacetylation and contraction in smooth muscle tissues*. Am J Physiol Cell Physiol, 2014. **307**(3): p. C288-95.
28. Waltregny, D., et al., *Expression of histone deacetylase 8, a class I histone deacetylase, is restricted to cells showing smooth muscle differentiation in normal human tissues*. Am J Pathol, 2004. **165**(2): p. 553-64.
29. Waltregny, D., et al., *Histone deacetylase HDAC8 associates with smooth muscle alpha-actin and is essential for smooth muscle cell contractility*. FASEB J., 2005. **19**(8): p. 966-8.
30. Grozinger, C.M. and S.L. Schreiber, *Regulation of histone deacetylase 4 and 5 and transcriptional activity by 14-3-3-dependent cellular localization*. Proceedings of the National Academy of Sciences, 2000. **97**(14): p. 7835-7840.
31. Parra, M. and E. Verdin, *Regulatory signal transduction pathways for class IIa histone deacetylases*. Curr Opin Pharmacol, 2010. **10**(4): p. 454-60.
32. Wang, A., et al., *Regulation of Histone Deacetylase 4 by Binding of 14-3-3 Proteins*. Molecular and cellular biology, 2000. **20**: p. 6904-12.
33. Tong, J.J., et al., *Identification of HDAC10, a novel class II human histone deacetylase containing a leucine-rich domain*. Nucleic acids research, 2002. **30**(5): p. 1114-1123.
34. Bertos, N.R., et al., *Role of the tetradecapeptide repeat domain of human histone deacetylase 6 in cytoplasmic retention*. J Biol Chem, 2004. **279**(46): p. 48246-54.
35. Verdel, A., et al., *Active maintenance of mHDA2/mHDAC6 histone-deacetylase in the cytoplasm*. Curr Biol, 2000. **10**(12): p. 747-9.
36. Gao, L., et al., *Cloning and Functional Characterization of HDAC11, a Novel Member of the Human Histone Deacetylase Family*. Journal of Biological Chemistry, 2002. **277**(28): p. 25748-25755.
37. Fischle, W., et al., *A New Family of Human Histone Deacetylases Related to Saccharomyces cerevisiae HDA1p*. Journal of Biological Chemistry, 1999. **274**(17): p. 11713-11720.
38. Grozinger, C.M., C.A. Hassig, and S.L. Schreiber, *Three proteins define a class of human histone deacetylases related to yeast Hda1p*. Proceedings of the National Academy of Sciences, 1999. **96**(9): p. 4868-4873.
39. Yang, W.-M., et al., *Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3*. Proceedings of the National Academy of Sciences, 1996. **93**(23): p. 12845-12850.

40. Gatla, H.R., et al., *Regulation of Chemokines and Cytokines by Histone Deacetylases and an Update on Histone Decetylase Inhibitors in Human Diseases*. Int J Mol Sci, 2019. **20**(5).
41. Sun, R., M. Hedl, and C. Abraham, *Twist1 and Twist2 Induce Human Macrophage Memory upon Chronic Innate Receptor Treatment by HDAC-Mediated Deacetylation of Cytokine Promoters*. J Immunol, 2019. **202**(11): p. 3297-3308.
42. Wang, L., R. Tao, and W.W. Hancock, *Using histone deacetylase inhibitors to enhance Foxp3(+) regulatory T-cell function and induce allograft tolerance*. Immunol Cell Biol, 2009. **87**(3): p. 195-202.
43. Huang, J., et al., *Histone/protein deacetylase 11 targeting promotes Foxp3+ Treg function*. Sci Rep, 2017. **7**(1): p. 8626.
44. Dahiya, S., et al., *HDAC10 deletion promotes Foxp3(+) T-regulatory cell function*. Sci Rep, 2020. **10**(1): p. 424.
45. de Zoeten, E.F., et al., *Histone deacetylase 6 and heat shock protein 90 control the functions of Foxp3(+) T-regulatory cells*. Mol Cell Biol, 2011. **31**(10): p. 2066-78.
46. Luo, J., et al., *Acetylation of p53 augments its site-specific DNA binding both in vitro and in vivo*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**: p. 2259-64.
47. Luo, J., et al., *Deacetylation of p53 modulates its effect on cell growth and apoptosis*. Nature, 2000. **408**: p. 377-81.
48. Nalawansa, D.A., et al., *HDAC Inhibitor-Induced Mitotic Arrest Is Mediated by Eg5/KIF11 Acetylation*. Cell Chem Biol, 2017. **24**(4): p. 481-492 e5.
49. Ferreira, R., et al., *Cell cycle-dependent recruitment of HDAC-1 correlates with deacetylation of histone H4 on an RB-E2F target promoter*. EMBO reports, 2001. **2**: p. 794-9.
50. Deardorff, M.A., et al., *HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle*. Nature, 2012. **489**(7415): p. 313-7.
51. Yin, F., et al., *LSD1 regulates pluripotency of embryonic stem/carcinoma cells through histone deacetylase 1-mediated deacetylation of histone H4 at lysine 16*. Mol Cell Biol, 2014. **34**(2): p. 158-79.
52. Du, Z., et al., *DNMT1 Stability Is Regulated by Proteins Coordinating Deubiquitination and Acetylation-Driven Ubiquitination*. Science signaling, 2010. **3**: p. ra80.

53. Zhou, Y., R. Santoro, and I. Grummt, *The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription*. EMBO J, 2002. **21**(17): p. 4632-40.
54. Chen, S.H., et al., *HDAC1,2 Knock-Out and HDACi Induced Cell Apoptosis in Imatinib-Resistant K562 Cells*. Int J Mol Sci, 2019. **20**(9).
55. Chen, L.-f., et al., *Duration of Nuclear NF- κ B Action Regulated by Reversible Acetylation*. Science, 2001. **293**(5535): p. 1653-1657.
56. Miska, E.A., et al., *HDAC4 deacetylase associates with and represses the MEF2 transcription factor*. Embo j, 1999. **18**(18): p. 5099-107.
57. Zhang, X., et al., *HDAC6 modulates cell motility by altering the acetylation level of cortactin*. Mol Cell, 2007. **27**(2): p. 197-213.
58. Hubbert, C., et al., *HDAC6 is a microtubule-associated deacetylase*. Nature, 2002. **417**(6887): p. 455-8.
59. Miyake, Y., et al., *Structural insights into HDAC6 tubulin deacetylation and its selective inhibition*. Nat Chem Biol, 2016. **12**(9): p. 748-54.
60. Kawaguchi, Y., et al., *The Deacetylase HDAC6 Regulates Aggresome Formation and Cell Viability in Response to Misfolded Protein Stress*. Cell, 2004. **115**: p. 727-38.
61. Kovacs, J.J., et al., *HDAC6 regulates Hsp90 acetylation and chaperone-dependent activation of glucocorticoid receptor*. Mol Cell, 2005. **18**(5): p. 601-7.
62. Cohen, T.J., et al., *An acetylation switch controls TDP-43 function and aggregation propensity*. Nat Commun, 2015. **6**: p. 5845.
63. Li, Y. and E. Seto, *HDACs and HDAC Inhibitors in Cancer Development and Therapy*. Cold Spring Harb Perspect Med, 2016. **6**(10).
64. Majdzadeh, N., et al., *HDAC4 inhibits cell-cycle progression and protects neurons from cell death*. Dev Neurobiol, 2008. **68**(8): p. 1076-92.
65. Graff, J., et al., *An epigenetic blockade of cognitive functions in the neurodegenerating brain*. Nature, 2012. **483**(7388): p. 222-6.
66. Montgomery, R.L., et al., *Histone deacetylases 1 and 2 control the progression of neural precursors to neurons during brain development*. Proc Natl Acad Sci U S A, 2009. **106**(19): p. 7876-81.
67. Shakespear, M.R., et al., *Histone deacetylases as regulators of inflammation and immunity*. Trends Immunol, 2011. **32**(7): p. 335-43.

68. Faria Freitas, M., M. Cuendet, and P. Bertrand, *HDAC inhibitors: a 2013-2017 patent survey*. *Expert Opin Ther Pat*, 2018: p. 1-17.
69. Flipo, M., et al., *Hydroxamates: Relationships between Structure and Plasma Stability*. *Journal of Medicinal Chemistry*, 2009. **52**(21): p. 6790-6802.
70. Zhang, L., et al., *Zinc binding groups for histone deacetylase inhibitors*. *J Enzyme Inhib Med Chem*, 2018. **33**(1): p. 714-721.
71. Patil, V., et al., *3-Hydroxypyridin-2-thione as novel zinc binding group for selective histone deacetylase inhibition*. *J Med Chem*, 2013. **56**(9): p. 3492-506.
72. Roche, J. and P. Bertrand, *Inside HDACs with more selective HDAC inhibitors*. *Eur J Med Chem*, 2016. **121**: p. 451-483.
73. Amin, S.A., N. Adhikari, and T. Jha, *Structure-activity relationships of HDAC8 inhibitors: Non-hydroxamates as anticancer agents*. *Pharmacol Res*, 2018. **131**: p. 128-142.
74. Singh, A.K., A. Bishayee, and A.K. Pandey, *Targeting Histone Deacetylases with Natural and Synthetic Agents: An Emerging Anticancer Strategy*. *Nutrients*, 2018. **10**(6).
75. Balasubramanian, S., et al., *A novel histone deacetylase 8 (HDAC8)-specific inhibitor PCI-34051 induces apoptosis in T-cell lymphomas*. *Leukemia*, 2008. **22**(5): p. 1026-34.
76. Bowers, M.E., et al., *The Class I HDAC inhibitor RGFP963 enhances consolidation of cued fear extinction*. *Learn Mem*, 2015. **22**(4): p. 225-31.
77. Oughtred, R., et al., *The BioGRID interaction database: 2019 update*. *Nucleic Acids Res*, 2019. **47**(D1): p. D529-d541.
78. Joshi, P., et al., *The functional interactome landscape of the human histone deacetylase family*. *Mol Syst Biol*, 2013. **9**: p. 672.
79. Seto, E. and M. Yoshida, *Erasers of histone acetylation: the histone deacetylase enzymes*. *Cold Spring Harbor perspectives in biology*, 2014. **6**(4): p. a018713-a018713.
80. Kelly, R.D. and S.M. Cowley, *The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts*. *Biochem Soc Trans*, 2013. **41**(3): p. 741-9.
81. Dressel, U., et al., *A dynamic role for HDAC7 in MEF2-mediated muscle differentiation*. *J Biol Chem*, 2001. **276**(20): p. 17007-13.

82. Stark, C., et al., *BioGRID: a general repository for interaction datasets*. Nucleic Acids Res, 2006. **34**(Database issue): p. D535-9.
83. Hook, S.S., et al., *Histone deacetylase 6 binds polyubiquitin through its zinc finger (PAZ domain) and copurifies with deubiquitinating enzymes*. Proc Natl Acad Sci U S A, 2002. **99**(21): p. 13425-30.
84. Toropainen, S., et al., *The down-regulation of the human MYC gene by the nuclear hormone 1alpha,25-dihydroxyvitamin D3 is associated with cycling of corepressors and histone deacetylases*. J Mol Biol, 2010. **400**(3): p. 284-94.
85. Choudhary, C., et al., *Lysine acetylation targets protein complexes and co-regulates major cellular functions*. Science, 2009. **325**(5942): p. 834-40.
86. Johnson, C.A., et al., *Human class I histone deacetylase complexes show enhanced catalytic activity in the presence of ATP and co-immunoprecipitate with the ATP-dependent chaperone protein Hsp70*. J Biol Chem, 2002. **277**(11): p. 9590-7.
87. Dose, A., et al., *NMR profiling of histone deacetylase and acetyl-transferase activities in real time*. ACS chemical biology, 2011. **6**(5): p. 419-424.
88. Castaneda, C.A., et al., *HDAC8 substrate selectivity is determined by long- and short-range interactions leading to enhanced reactivity for full-length histone substrates compared with peptides*. J Biol Chem, 2017. **292**(52): p. 21568-21577.
89. Kim, M.M., et al., *Histone deacetylases, HDAC1 and HSIR2, act as a negative regulator of ageing through p53 in human gingival fibroblast*. Mech Ageing Dev, 2004. **125**(5): p. 351-7.
90. Sharma, S., et al., *Acetylation-Dependent Control of Global Poly(A) RNA Degradation by CBP/p300 and HDAC1/2*. Mol Cell, 2016. **63**(6): p. 927-38.
91. Hassa, P.O., et al., *Acetylation of poly(ADP-ribose) polymerase-1 by p300/CREB-binding protein regulates coactivation of NF-kappaB-dependent transcription*. J Biol Chem, 2005. **280**(49): p. 40450-64.
92. Wilson, B.J., et al., *An acetylation switch modulates the transcriptional activity of estrogen-related receptor alpha*. Mol Endocrinol, 2010. **24**(7): p. 1349-58.
93. Cohen, T.J., et al., *The acetylation of tau inhibits its function and promotes pathological tau aggregation*. Nat Commun, 2011. **2**: p. 252.
94. Min, S.W., et al., *Acetylation of tau inhibits its degradation and contributes to tauopathy*. Neuron, 2010. **67**(6): p. 953-66.

95. Noack, M., J. Leyk, and C. Richter-Landsberg, *HDAC6 inhibition results in tau acetylation and modulates tau phosphorylation and degradation in oligodendrocytes*. *Glia*, 2014. **62**(4): p. 535-47.
96. Glozak, M.A. and E. Seto, *Acetylation/Deacetylation Modulates the Stability of DNA Replication Licensing Factor Cdt1*. *Journal of Biological Chemistry*, 2009. **284**(17): p. 11446-11453.
97. Olson, D.E., et al., *An unbiased approach to identify endogenous substrates of "histone" deacetylase 8*. *ACS Chem Biol*, 2014. **9**(10): p. 2210-6.
98. Lopez, J.E., et al., *HDAC8 Substrates Identified by Genetically Encoded Active Site Photocrosslinking*. *J Am Chem Soc*, 2017. **139**(45): p. 16222-16227.
99. Nalawansa, D.A. and M.K. Pflum, *LSD1 Substrate Binding and Gene Expression Are Affected by HDAC1-Mediated Deacetylation*. *ACS Chem Biol*, 2017. **12**(1): p. 254-264.
100. Gomes, I.D. and M.K.H. Pflum, *Optimal Substrate-Trapping Mutants to Discover Substrates of HDAC1*. *Chembiochem*, 2019. **20**(11): p. 1444-1449.
101. Nalawansa, D.A., et al., *HDAC1 Substrate Profiling Using Proteomics-Based Substrate Trapping*. *ACS Chem Biol*, 2018. **13**(12): p. 3315-3324.
102. Alam, N., et al., *Structure-Based Identification of HDAC8 Non-histone Substrates*. *Structure*, 2016. **24**(3): p. 458-68.
103. Wegener, D., et al., *A Fluorogenic Histone Deacetylase Assay Well Suited for High-Throughput Activity Screening*. *Chemistry & Biology*, 2003. **10**(1): p. 61-68.
104. Wolfson, N.A., et al., *An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity*. *Anal Biochem*, 2014. **456**: p. 61-9.
105. Wu, M., et al., *Lysine-14 acetylation of histone H3 in chromatin confers resistance to the deacetylase and demethylase activities of an epigenetic silencing complex*. *Elife*, 2018. **7**.
106. Sullivan, E.D., *Unlocking an HDAC Toolbox: Methods Towards Understanding Isozyme-Specific Activity*, in *Chemical Biology*. 2016, University of Michigan. p. 241.
107. Vannini, A., et al., *Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor*. *Proc Natl Acad Sci U S A*, 2004. **101**(42): p. 15064-9.
108. Sousa, S.F., et al., *The Zinc proteome: a tale of stability and functionality*. *Dalton Trans*, 2009(38): p. 7946-56.

109. Dowling, D.P., et al., *Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function*. *Biochemistry*, 2010. **49**(24): p. 5048-56.
110. Gantt, S.L., S.G. Gattis, and C.A. Fierke, *Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion*. *Biochemistry*, 2006. **45**(19): p. 6170-8.
111. Kim, B., A.S. Pithadia, and C.A. Fierke, *Kinetics and thermodynamics of metal-binding to histone deacetylase 8*. *Protein Sci*, 2015. **24**(3): p. 354-65.
112. Bozym, R.A., et al., *Measuring Picomolar Intracellular Exchangeable Zinc in PC-12 Cells Using a Ratiometric Fluorescence Biosensor*. *ACS Chemical Biology*, 2006. **1**(2): p. 103-111.
113. McCranor, B.J., et al., *Quantitative imaging of mitochondrial and cytosolic free zinc levels in an in vitro model of ischemia/reperfusion*. *J Bioenerg Biomembr*, 2012. **44**(2): p. 253-63.
114. Espósito, B.P., et al., *A Review of Fluorescence Methods for Assessing Labile Iron in Cells and Biological Fluids*. *Analytical Biochemistry*, 2002. **304**(1): p. 1-18.
115. Petrat, F., H. de Groot, and U. Rauen, *Subcellular distribution of chelatable iron: a laser scanning microscopic study in isolated hepatocytes and liver endothelial cells*. *The Biochemical journal*, 2001. **356**(Pt 1): p. 61-69.
116. Gattis, S.G., M. Hernick, and C.A. Fierke, *Active site metal ion in UDP-3-O-((R)-3-hydroxymyristoyl)-N-acetylglucosamine deacetylase (LpxC) switches between Fe(II) and Zn(II) depending on cellular conditions*. *J Biol Chem*, 2010. **285**(44): p. 33788-96.
117. Hernick, M., et al., *Activation of Escherichia coli UDP-3-O-[(R)-3-hydroxymyristoyl]-N-acetylglucosamine deacetylase by Fe²⁺ yields a more efficient enzyme with altered ligand affinity*. *Biochemistry*, 2010. **49**(10): p. 2246-55.
118. Castaneda, C.A., et al., *Active Site Metal Identity Alters Histone Deacetylase 8 Substrate Selectivity: A Potential Novel Regulatory Mechanism*. *Biochemistry*, 2017. **56**(42): p. 5663-5670.
119. Pflum, M.K., et al., *Histone deacetylase 1 phosphorylation promotes enzymatic activity and complex formation*. *J Biol Chem*, 2001. **276**(50): p. 47733-41.
120. Schultz, B.E., et al., *Kinetics and Comparative Reactivity of Human Class I and Class IIb Histone Deacetylases*. *Biochemistry*, 2004. **43**(34): p. 11083-11091.
121. Zhang, X., et al., *Histone deacetylase 3 (HDAC3) activity is regulated by interaction with protein serine/threonine phosphatase 4*. *Genes Dev*, 2005. **19**(7): p. 827-39.

122. Lee, H., N. Rezai-Zadeh, and E. Seto, *Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A*. Mol Cell Biol, 2004. **24**(2): p. 765-73.
123. Welker Leng, K.R., et al., *Phosphorylation of Histone Deacetylase 8: Structural and Mechanistic Analysis of the Phosphomimetic S39E Mutant*. Biochemistry, 2019. **58**(45): p. 4480-4493.
124. Parra, M., et al., *Protein kinase D1 phosphorylates HDAC7 and induces its nuclear export after T-cell receptor activation*. J Biol Chem, 2005. **280**(14): p. 13762-70.
125. Lissanu Deribe, Y., et al., *Regulation of Epidermal Growth Factor Receptor Trafficking by Lysine Deacetylase HDAC6*. Science signaling, 2009. **2**: p. ra84.
126. Williams, K.A., et al., *Extracellular signal-regulated kinase (ERK) phosphorylates histone deacetylase 6 (HDAC6) at serine 1035 to stimulate cell migration*. J Biol Chem, 2013. **288**(46): p. 33156-70.
127. Chen, S., et al., *HDAC6 regulates mitochondrial transport in hippocampal neurons*. PLoS One, 2010. **5**(5): p. e10848.
128. Watabe, M. and T. Nakaki, *Protein kinase CK2 regulates the formation and clearance of aggresomes in response to stress*. Journal of Cell Science, 2011. **124**(9): p. 1519-1532.
129. Zhu, J., C.B. Coyne, and S.N. Sarkar, *PKC alpha regulates Sendai virus-mediated interferon induction through HDAC6 and beta-catenin*. EMBO J, 2011. **30**(23): p. 4838-49.

Chapter 2 : Novel Non-Hydroxamate Inhibition of Histone Deacetylases*

Introduction

Histone deacetylases have been attractive cancer drug targets due to their involvement in the regulation of gene transcription[1-7] and overexpression of HDACs has been implicated in numerous cancers[8]. There are currently four FDA-approved inhibitors of Class I HDACs for use in treatment of T-cell lymphoma and multiple myeloma[8]. Developed HDAC inhibitors have a consistent structure: zinc binding group (ZBG) – linker – cap. The linker and cap groups form contacts within the substrate binding tunnel and protein surface, respectively, and can be varied to confer isozyme selectivity. The ZBG is the primary mode of inhibition as it binds to the divalent active site metal displacing the water nucleophile necessary for catalysis. Hydroxamates have been commonly used as ZBGs, present in 3 of the 4 FDA approved inhibitors, as the group is a potent metal binder. However, hydroxamates have poor pharmacokinetics with an *in vivo* half-life of less than an hour due to the susceptibility of the hydroxamate to hydrolysis under physiological conditions[9]. It can additionally be difficult to achieve good selectivity due to the Zn(II) binding contributing a large portion of target affinity[9]. The issue of selectivity extends to all Zn(II) enzymes and within the HDAC family; all the FDA approved inhibitors are non-selective pan-HDAC inhibitors and exhibit high toxicity

*Previous students George Murphy III and Hannah Foley also contributed to this work. George Murphy III performed the initial screen of CFL1 with HDAC8-Zn(II) and HDAC8-Fe(II). Hannah Foley assisted in performing the dose response curves for the HPO and HPT series.

due to the large effect HDACs 1-3 in particular have on gene regulation[8]. Yet other common ZBGs, benzamides, carboxylic acids etc., have been unable to match the potency of the hydroxamate[10].

Research is currently ongoing to develop novel, non-hydroxamate metal binding groups for HDAC inhibitors[8, 11-14]. The focus has been on developing inhibitors with binding groups specific to Zn(II), due to it being widely accepted that HDACs are Zn(II)-deacetylases. The initial HDAC8 crystal structure showed Zn(II) present in the active site following reconstitution[15]. However, the Asp-Asp-His metal binding site is uncharacteristic of a zinc binding site[16] leading to the HDAC8 active site being extensively studied biochemically and structurally. HDAC8 binds and is activated by additional metal ions, including Fe(II) and Co(II)[16, 17] and it has been shown that the catalytic efficiency and inhibitor affinity varies dependent on reconstituted metal[17]. Yet crystal structures of reconstituted HDAC8 with various metals do not illustrate any structural differences that would account for the different activity[16]. It should be noted that all these crystal structures contain a hydroxamic acid inhibitor suggesting the possibility that the inhibitor alters the geometry of the bound metal ion.

HDAC8 activity in cells has additionally demonstrated oxygen sensitivity suggesting the potential importance of Fe(II) in cellular HDAC8 activity[17, 18]. Alternatively, the oxygen sensitivity of HDACs has been proposed to be due to redox sensitivity. Most Class I HDACs (HDAC1, HDAC2, and HDAC3) have two conserved cysteine residues which can be alkylated by reactive carbonyl species (RCS), a downstream product of reactive oxygen species (ROS), causing decreased deacetylase activity[19]. Class II HDAC4 additionally has been found to contain a cysteine pair

sensitive to oxidation in response to ROS[20]. HDAC8 has been the most recent HDAC reported to have a redox switch[21], where the oxidized form is inactive. The presence of an inactivating redox switch could explain HDAC8 oxygen sensitivity in cell lysates. Nonetheless additional studies have also demonstrated that the substrate selectivity varies with the metal ion bound to HDAC8[22] implicating the identity of the active site metal as a potential regulatory mechanism. The identity of the active site metal in endogenous HDAC8 has been yet to be determined with certainty.

This chapter details the discovery of novel non-hydroxamate inhibitors by use of a metal-chelating fragment library. Some fragments displayed selectivity for inhibition of Fe(II)-compared to Zn(II)-bound HDAC8, leading to these fragments being further structurally modified. These elaborated inhibitors were then used to probe the active site metal bound to HDAC8 in cells but no evidence indicating the presence of HDAC8-Fe(II) was obtained.

Materials and Methods

Reagents

Trace metal certified HEPES, sodium chloride, and potassium chloride were purchased from Sigma. Trace metal certified tips were purchased from Corning (4869, 4694). All other materials were purchased from Fisher at >95% purity unless noted otherwise.

CFL1 Library and Expanded Inhibitors

Prof. Seth Cohen at the University of California San Diego graciously provided the 96-member metal chelator fragment library, CFL1[23], as 50 mM stock solutions in

dimethyl sulfoxide, DMSO. The expanded inhibitors were additionally synthesized in the Cohen lab and provided as 50 mM stock solutions in DMSO.

Purification of apo-HDAC8

Apo-HDAC8 was expressed and purified, followed by reconstitution with Zn(II), Fe(II), or Co(II) as previously described[17] with the following modifications. HDAC8 was reconstituted at a 1:1 molar ratio with Zn(II) using a 1000 ppm Zn ICP standard by incubation on ice for 1 hour. HDAC8 was reconstituted with Co(II) at a 2:1 molar ratio using a 1000 ppm Co ICP standard by incubation on ice for 1 hour. For reconstitution with Fe(II), HDAC8 in storage buffer (25mM MOPS, 1mM TCEP, 5mM KCl) was equilibrated in a Coy labs anaerobic chamber on ice for 1 hour before reconstitution at a 5:1 molar ratio using a solution of Fe(II) sulfate heptahydrate with excess ascorbic acid (50-fold) in reaction buffer (25 mM HEPES pH 7.5, 137 mM NaCl, 2.7 mM KCl) on ice for 1 hour. For activity assays, the HDAC8-Fe(II) was removed from the anaerobic chamber and assayed within 1-2 hours with minimal loss in activity.

The concentration of HDAC8 was determined by titration with fluorescein-SAHA (previously synthesized[24]). Increasing amounts of HDAC8-Co(II) were added to 100 nM fluorescein-SAHA (SAHA $K_i=44\text{nM}$) in individual wells of a Corning 3686 96-well half area microplate (previously soaked in 0.5 mM EDTA, thoroughly rinsed with water and dried). The change in fluorescent polarization (ex. 485 nm, em. 535 nm) was determined after incubation at room temperature for 15 minutes. The data was analyzed using GraphPad Prism to determine the stoichiometric point of saturation and the adjusted concentration of HDAC8.

Metal-Chelating Fragment Library Screen

HDAC8 activity in the presence of the metal-chelating fragments was determined by the Fluor de Lys assay using Enzo Life Sciences Fluor de Lys HDAC8 substrate and Developer II solutions. A control reaction was performed using 0.5 μM HDAC8, 50 μM HDAC8 substrate in 10% DMSO, and reaction buffer. The 50 mM inhibitor stocks in DMSO were diluted to 2 mM in DMSO. The inhibitor, at a final concentration of 200 μM , was added to 50 μM HDAC8 substrate in 1X buffer, preheated to 30°C, and the reaction was initiated with 0.5 μM HDAC8. Various aliquoted timepoints (5 μL) were diluted into 45 μL quench solution (0.73% Developer II, 1.2 μM trichostatin A, 1X buffer) in individual wells of a Corning 3694 96-well half area microplate. After a 10-minute room temperature incubation, the fluorescence of the remaining substrate (ex. 360 nm, em. 362 nm) and resulting product (ex. 360 nm and em. 460 nm) was determined using a PolarStar plate reader. The amount of product in each inhibitor reaction was compared to the control reaction to obtain the % retained activity.

Dose Response Curves

HDAC8 activity was determined using the Fluor de Lys assay. A control reaction was performed containing 0.5 μM HDAC8, 50 μM HDAC8 substrate in 10% DMSO, 1X buffer. The 50 mM inhibitor stocks in DMSO were diluted to various concentrations such that each assay contained a final concentration of 10% DMSO. The diluted inhibitor was added to 50 μM HDAC8 substrate in 1X buffer, preheated to 30°C, and the reaction was initiated by addition of 0.5 μM HDAC8. Various timepoints (n=4) were quenched in individual wells of a Corning 3694 96-well half area microplate and fluorescence determined as previously described. The initial rate of each reaction was compared to the

control reaction to obtain the remaining activity (%). The IC_{50} value was obtained from a fit of **Equation 2.1** to the dependence of the activity on the inhibitor concentration (x) allowing for a variable hill slope (h).

$$y = \frac{100}{1 + \left(\frac{x}{IC_{50}}\right)^h}$$

Equation 2.1

ICP-MS Assay

Reconstituted HDAC8-Fe(II) (9 μ M) was incubated with inhibitor (at 3 times IC_{50}) or DMSO as a control at 30°C for 1 hour under anaerobic conditions in 1X buffer using a Coy anaerobic chamber. The samples were then removed from the anaerobic chamber and processed through Zeba spin 7K MWCO desalting columns (ThermoScientific) to remove unbound inhibitor and Fe(II). A portion of the desalted sample (15 μ L) was diluted to 500 μ L with water for Bradford analysis and the remaining sample (30 μ L) was diluted to 1 mL with 2% metal-free nitric acid for ICP-MS analysis. Coomassie Plus reagent (ThermoScientific) was used to determine the HDAC8 concentration using a Bradford assay. Each desalted, diluted sample was assayed in triplicate and analyzed against an HDAC8 standard curve following the manufacturer (ThermoScientific) protocol. An Agilent 7900 ICP-MS was used to determine the total Fe concentration. The Fe concentration of each desalted, diluted sample was determined from triplicate analysis in the High-Energy Helium mode against a 56 Fe standard curve with 71 Ga as an internal standard. The Fe concentration of the desalted sample was compared to the HDAC8 concentration determined using the Bradford assay.

UV-Vis Assay

Fe(II) was prepared from Fe(II) sulfate heptahydrate in an anaerobic chamber in the presence of excess ascorbic acid (50-fold). Increasing concentrations of inhibitor (10% DMSO in 1X buffer) were added to a constant concentration of Fe(II) in individual wells of a Corning 9017 96-well microplate. The absorbance of each well was measured at the λ_{\max} of the inhibitor-Fe(II) complex using a SpectraMax plate reader. A binding equation (**Equation 2.2**) was fit to these data using GraphPad Prism to determine the K_D for the inhibitor- Fe complex with variable hill slope (h) and y_{\max} = max absorbance change.

$$y = y_{\max} * \frac{x^h}{(K_D^h + x^h)} \quad \text{Equation 2.2}$$

Oxygen Sensitivity of HeLa cell lysates

HeLa cells were purchased from ATCC (CCL-2) and were grown in Gibco™ Dulbecco's modified eagle medium, high glucose (ThermoScientific) supplemented with 10% Hyclone fetal bovine serum in T75 flasks until 80-90% confluent. Cells were harvested using Gibco™ 0.25% Trypsin-EDTA (ThermoScientific), where the sample was stained with trypan blue and the cells counted with a hemocytometer. 1 million cells were lysed in 1X buffer containing 1% Tween-20 and Halt Protease/Phosphatase Inhibitor Cocktail (ThermoScientific) by incubation on ice for 30 minutes under anaerobic conditions. Following lysis, half of the sample was removed from the anaerobic chamber and exposed to oxygen on ice for at least 3 hours. The rest of the sample remained under anaerobic conditions on ice for the same amount of time. The HDAC activity of the aerobic and anaerobic lysate was determined using the Fluor de Lys assay with 100 μ M HDAC8

substrate at 30°C. For each sample, initial and final reaction time points were collected (run in triplicate). Using GraphPad Prism, the reaction rate was determined from the linear correlation between the initial and final timepoints. The different sample means were compared by 2-way ANOVA with Sidak multiple comparisons correction[25].

Results

Select fragments display potent inhibition

The initial screen of the chelator fragment library, CFL1, revealed multiple fragments that were efficient at inhibiting activity of both HDAC8-Zn(II) and HDAC8-Fe(II) (**Figure 2.1-A**). Of the 96 fragments screened, 10 displayed greater than or equal to 50% inhibition at 200 μ M (**Table 2.1**). The fragments that displayed the best inhibition against HDAC8-Zn(II) and HDAC8-Fe(II) were further tested against HDAC8-Co(II) where they additionally displayed good inhibition (**Figure 2.1-B**). To further characterize inhibition, the IC_{50} values of select compounds was determined and the IC_{50} values of three of these compounds are below 30 μ M (**Figure 2.2**). One of those fragments, G12, has an IC_{50} value for each of the metal-reconstituted forms of HDAC8 that is in the nanomolar range (**Table 2.2**).

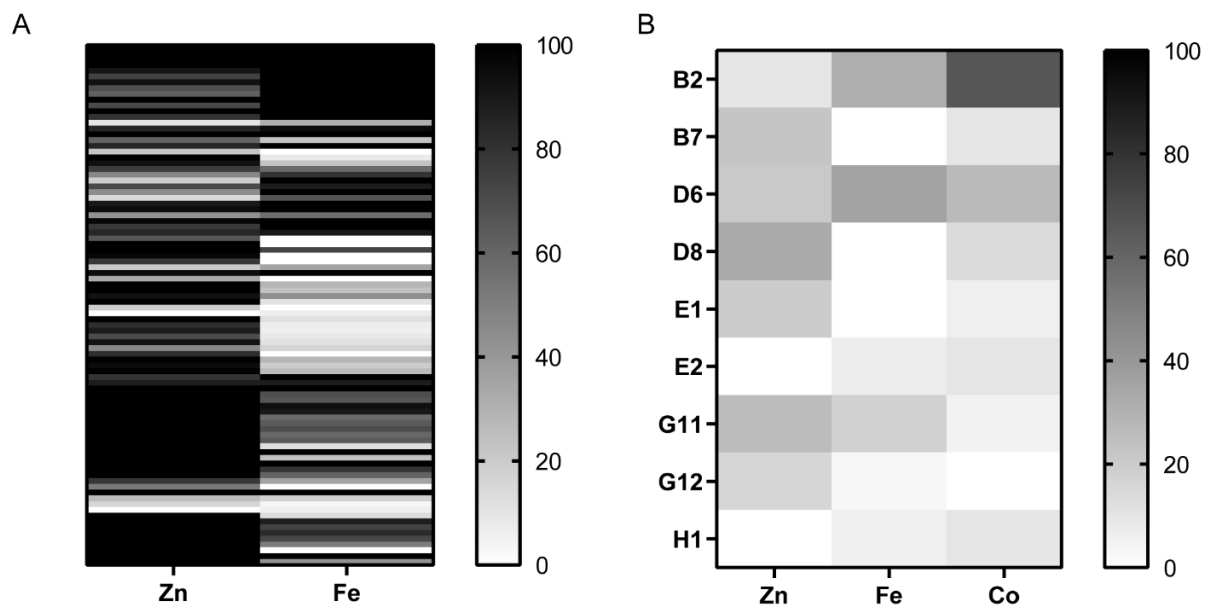
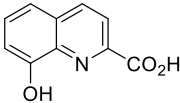
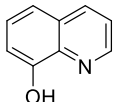
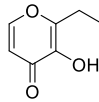
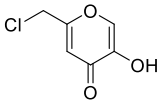
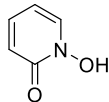
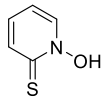
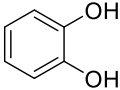
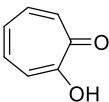
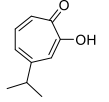
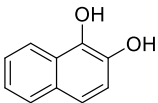


Figure 2.1: Heat maps illustrating inhibitor potency. Grayscale is given for remaining HDAC8 activity (%), more black indicates higher % activity, more white indicates less % activity. A) All CFL1 fragments against HDAC8-Zn(II) and HDAC8-Fe(II). B) Selected hit fragments against all metal-reconstituted forms of HDAC8.

Table 2.1: CFL1 Library Fragments screened against all HDAC8-reconstituted forms

		HDAC8-Zn(II)	HDAC8-Fe(II)	HDAC8-Co(II)
Inhibitor	Structure	Retained activity (%)	Retained activity (%)	Retained activity (%)
B2		10	31	66
B7		40	0	10
D6		21	36	27
D8		33	0	14
E1		20	0	6
E2		0	7	10
G8		57	0	n.d.
G11		26	18	5
G12		16	3	0
H1		0	6	10

^aRetained activity determined from comparison of reconstituted HDAC8 activity with 200 μ M inhibitor to 10% DMSO control (initial velocities from stopped timepoints (n=3-4))

Table 2.2: IC_{50} values of most potent inhibitor fragments

	HDAC8-Zn(II) ^a		HDAC8-Fe(II) ^a		HDAC8-Co(II) ^a	
	IC_{50} (μ M)	Hill Slope	IC_{50} (μ M)	Hill Slope	IC_{50} (μ M)	Hill Slope
D8	90 \pm 30	-1.3 \pm 0.7	26 \pm 2	-1.07 \pm 0.06	40 \pm 20	-1.2 \pm 0.4
H1	8 \pm 2	-1.5 \pm 0.7	n.d. ^b	n/a	20 \pm 10	-0.8 \pm 0.4
G11	19 \pm 8	-0.9 \pm 0.3	11.2 \pm 0.4	-2.0 \pm 0.4	8 \pm 3	-2 \pm 1
G12	0.6 \pm 0.3	-2 \pm 1	0.7 \pm 0.2	-2 \pm 1	0.49 \pm 0.03	-2.0 \pm 0.1

^aThe IC_{50} value (mean \pm standard error) was determined from dose response curves fit with **Equation 2.1** to allow for variable hill slopes. Hill slopes (mean \pm standard error) were determined from GraphPad Prism analysis using Equation 2.1. ^bThe value was not determined.

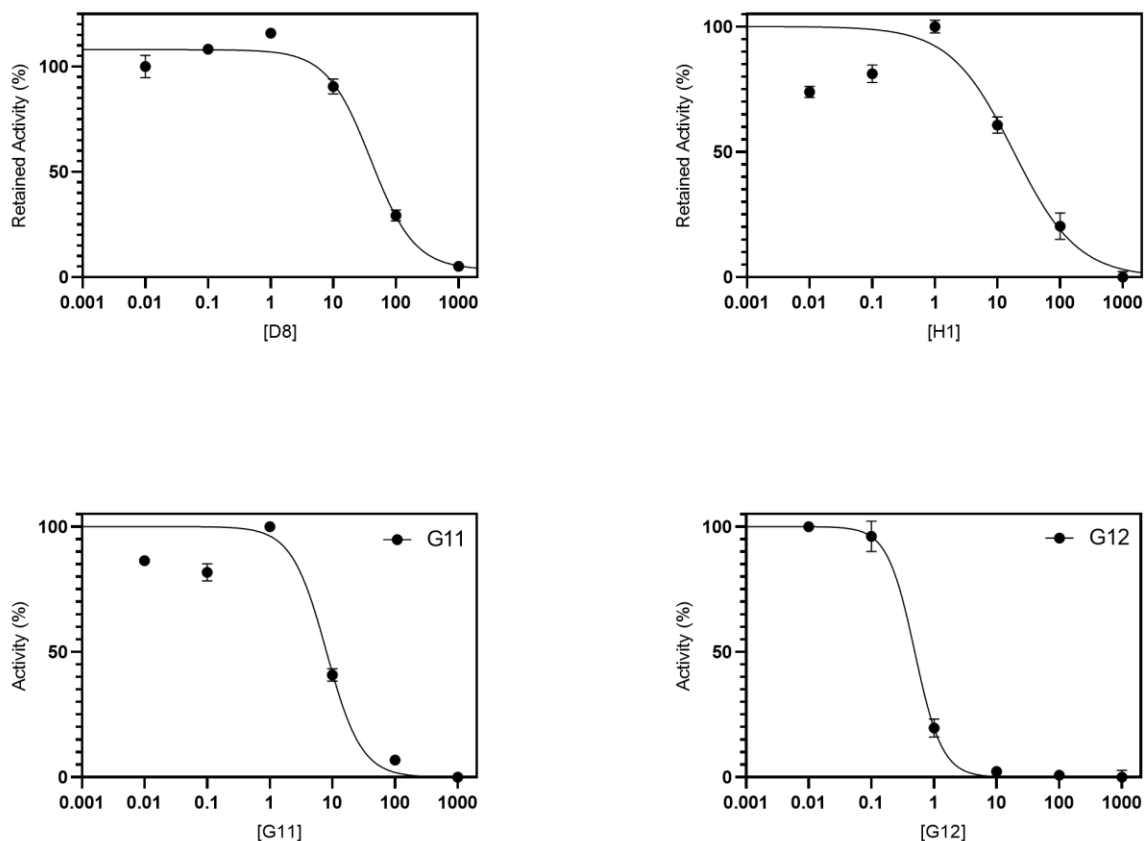


Figure 2.2: Dose response curves of most potent inhibitor fragments against 0.5 μM HDAC8-Co(II). Each point is the initial velocity calculated from a linear fit of stopped timepoints (n=4). Error bars indicate the standard error.

HDAC8-Fe(II) selective inhibition

Interestingly, a greater number of fragments, 27 out of 96, inhibited HDAC8-Fe(II) significantly more than HDAC8-Zn(II), 10 out of 96. No fragments were found that inhibit HDAC8-Zn(II) more than HDAC8-Fe(II). The IC_{50} values for two fragments, D4 and D5, were at least 20 times lower for HDAC8-Fe(II) compared to HDAC8-Zn(II) (**Figure 2.3**). To develop even better HDAC8-Fe(II) specific inhibitors, a small library of compounds were prepared by Dr. Christian Perez from the Cohen group where the structure of these two fragments was elaborated by attaching various linker lengths and cap groups onto

the metal-chelating group (**Figure 2.4**). In general, these added structural features had modest effects on the IC_{50} values and on the selectivity for inhibition of HDAC8-Fe(II) (**Figure 2.5, Table 2.3**). However, more variability in the selectivity for HDAC8-Fe(II) compared to HDAC8-Zn(II) was observed for the D5 based compounds, HPT (**Table 2.3**). The expanded compounds for D4, named HPO, showed some improved inhibition with two compounds having IC_{50} values below 10 μ M, HPO6 and HPO12, for inhibition of HDAC8-Fe(II) (**Table 2.3**). HPO12 also enhanced inhibition of HDAC8-Fe(II) by 5-fold compared to the parent compound (D4).

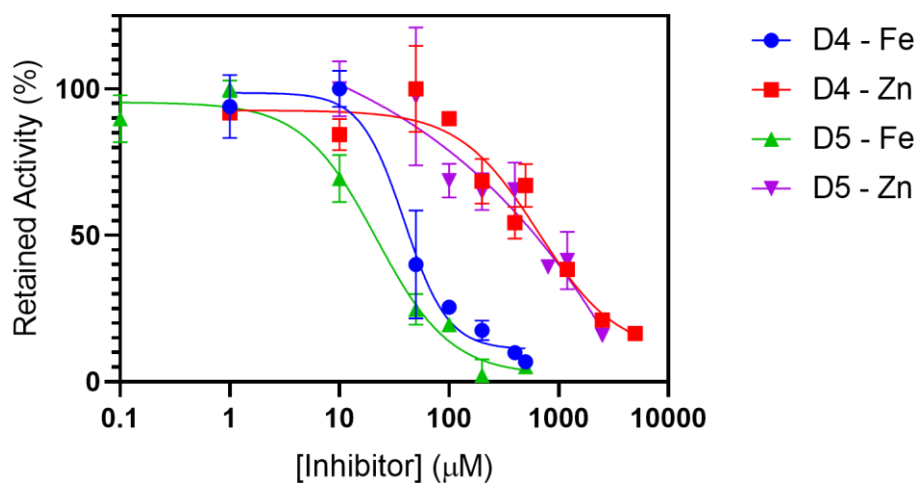


Figure 2.3: Metalloform selective inhibitors dose response curves. The curves illustrate the higher effectiveness against HDAC8-Fe(II) versus HDAC8-Zn(II). Each point is the initial velocity calculated from a linear fit of stopped timepoints ($n=4$). Error bars indicate the standard error.

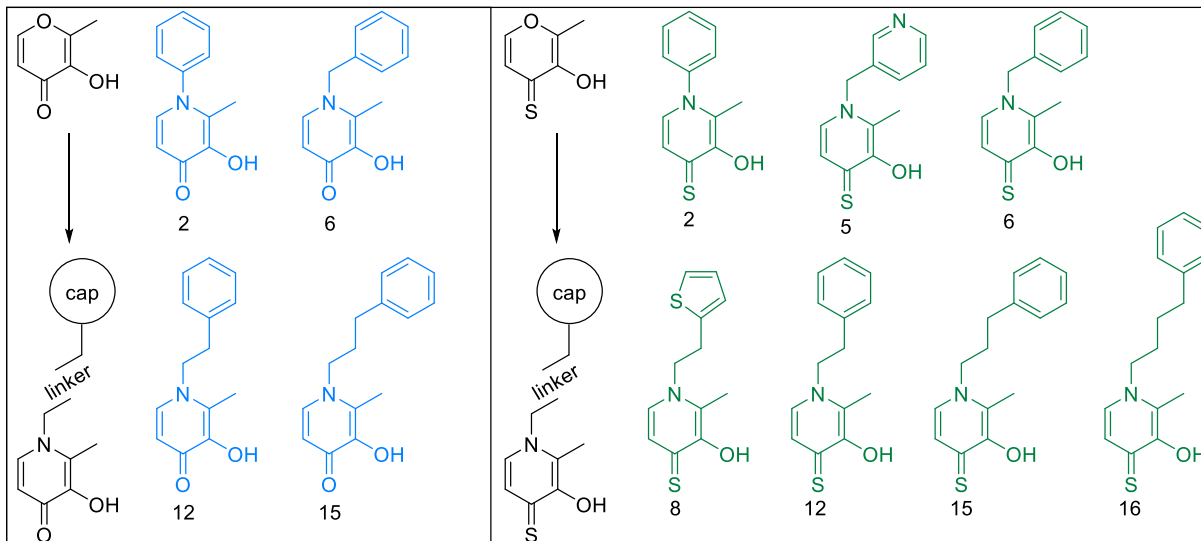


Figure 2.4: Modification of 4D (left) and 5D (right).

The HPO series is shown in blue. The HPT series is shown in green. Naming for both includes the series name followed by the number to represent the structure (shown below in the figure, i.e. HPO-2)

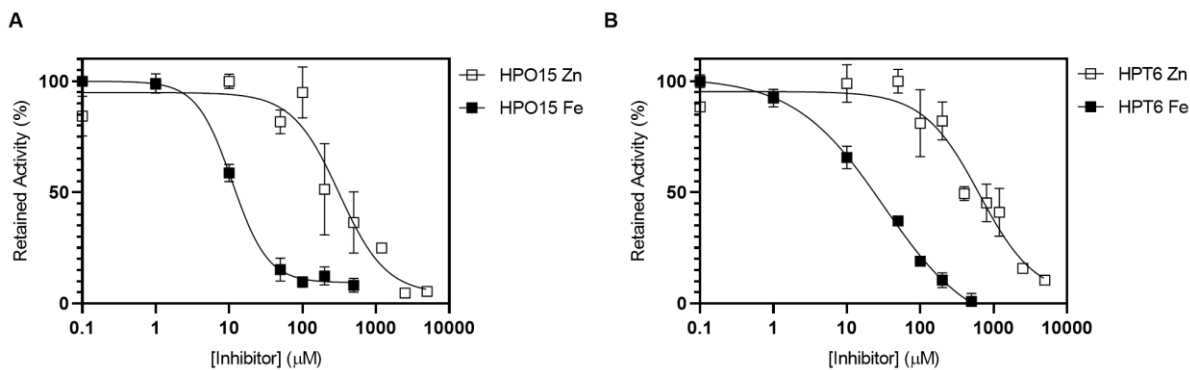


Figure 2.5: Dose response curves of selected expanded inhibitors.

A) Dose response curves of HPO15 with HDAC8-Zn(II) (open squares) and HDAC8-Fe(II) (filled squares). B) Dose Response curves of HPT6 with HDAC8-Zn(II) (open squares) and HDAC8-Fe(II) (filled squares). The curves illustrate that each inhibitor maintained the metal selectivity seen for the parent fragment. Each point is the initial velocity calculated from a linear fit of stopped timepoints (n=4). Error bars indicate the standard error.

Table 2.3: IC_{50} values of parent fragments and select expanded inhibitors

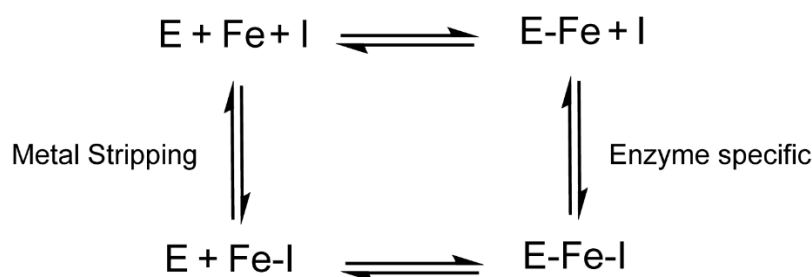
	HDAC8-Zn(II) ^a		HDAC8-Fe(II) ^a		Zn/Fe ratio ^b
	IC_{50} (μ M)	Hill Slope	IC_{50} (μ M)	Hill Slope	
D4	600 \pm 100	-0.8 \pm 0.1	18 \pm 6	-0.8 \pm 0.2	30 \pm 10
HPO2	240 \pm 50	-0.8 \pm 0.1	11.1 \pm 0.8	-1.16 \pm 0.08	22 \pm 4
HPO6	230 \pm 60	-0.45 \pm 0.06	9 \pm 2	-0.7 \pm 0.1	25 \pm 9
HPO12	>5,000 ^c	-0.23 \pm 0.08	4 \pm 2	-0.6 \pm 0.2	>1250
HPO15	320 \pm 70	-1.1 \pm 0.2	14 \pm 2	-1.1 \pm 0.2	23 \pm 6
D5	500 \pm 100	-0.8 \pm 0.2	21 \pm 3	-1.1 \pm 0.2	24 \pm 6
HPT2	390 \pm 60	-1.3 \pm 0.2	22 \pm 5	-0.9 \pm 0.2	17 \pm 5
HPT5	260 \pm 20	-1.5 \pm 0.2	66 \pm 9	-1.3 \pm 0.2	4 \pm 1
HPT6	610 \pm 90	-1.0 \pm 0.2	22 \pm 2	-0.91 \pm 0.07	27 \pm 5
HPT8	360 \pm 40	-1.4 \pm 0.2	40 \pm 10	-0.6 \pm 0.1	9 \pm 2
HPT12	500 \pm 200	-1.1 \pm 0.5	12 \pm 5	-0.8 \pm 0.2	40 \pm 20
HPT15	300 \pm 100	-0.8 \pm 0.2	30 \pm 3	-0.90 \pm 0.07	10 \pm 3
HPT16	800 \pm 200	-0.8 \pm 0.2	21 \pm 4	-0.65 \pm 0.08	40 \pm 10

^aThe IC_{50} value (mean \pm standard error) was determined from dose response curves fit with **Equation 2.1** to allow for variable hill slopes. Hill slopes (mean \pm standard error) were determined from GraphPad Prism analysis using **Equation 2.1**. ^bThe ratio of Zn inhibition over Fe inhibition (\pm propagated error [26]). ^cInhibition did not reach below 50%.

Specificity of inhibitor is related to IC_{50}

Previous measurements determined that the K_D of Fe(II) for HDAC8 is 2 μ M[24], making this metal labile within the active site. Therefore, it is possible that the Fe-specific inhibitors are binding free Fe(II) in solution and sequestering the metal as it dissociates from the enzyme to form inactive, apo-enzyme rather than binding to HDAC8-Fe(II) (**Scheme 2.1**). An ICP-MS assay was developed to test for this metal stripping ability. The ICP-MS was used to determine how much Fe(II) remained bound to HDAC8 after

incubation with the inhibitors followed by rapid separation of the protein from unbound metal and inhibitor using spin desalting columns. These data demonstrated that in both libraries of expanded compounds there were structures that inhibited HDAC8 by removing the metal ion and ones that bound to HDAC8 (**Figure 2.6**). These data also suggest a lower IC_{50} value does not indicate the formation of a HDAC8-Fe(II)-inhibitor complex and the mode of inhibition is likely determined by the relative affinity of the inhibitor for free Fe(II) and HDAC8-Fe(II).



Scheme 2.1: Types of possible inhibition by metal-specific inhibitors.

This hypothesis was confirmed by determining the K_D value of the inhibitor to free Fe(II) using UV-Vis spectroscopy (**Figure 2.7**) where the formation of an Fe(II)-inhibition complex results in a peak in the visible range (absorbance 400-500 nm). The ICP-MS determined specific inhibitor, HPO15, has a K_D for free Fe(II), $89 \pm 5 \mu\text{M}$ ($h = 2.4 \pm 0.2$) that is higher than the measured IC_{50} value, $14 \pm 2 \mu\text{M}$, consistent with an inhibition mechanism of formation of a HDAC8-Fe(II)-HPO15 complex. In contrast, the ICP-MS determined stripping inhibitor, HPO6, has a K_D value for free Fe(II), $<2 \mu\text{M}$ ($h = 2.0 \pm 0.3$) that is lower than the determined IC_{50} value, $9 \pm 2 \mu\text{M}$, consistent with an inhibition mechanism of formation of apo-HDAC8. We were unable to determine the actual Fe(II) K_D for HPO6 as the Fe(II) absorption signal is too low to measure below $2 \mu\text{M}$ (**Figure**

2.7-B). These data confirm the ICP-MS results that HPO15 binds HDAC8-Fe(II) with higher affinity and HPO6 binds free Fe(II) with higher affinity.

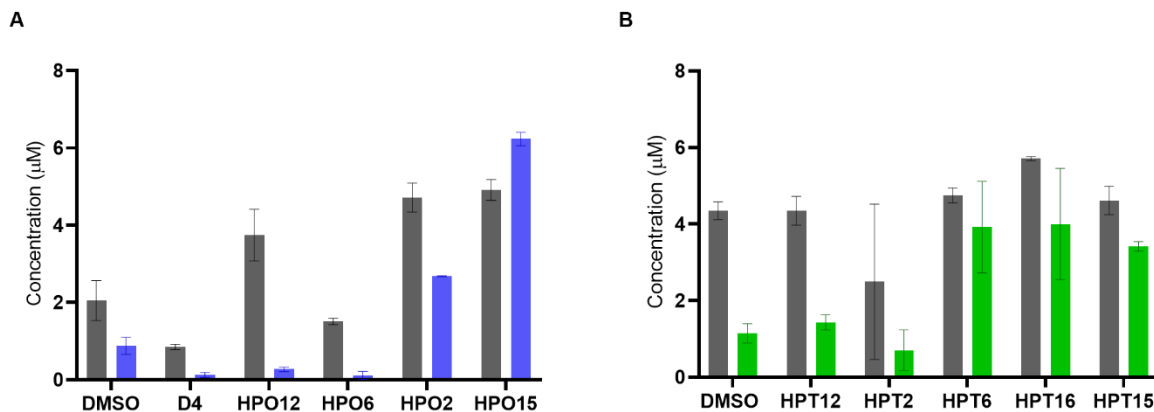


Figure 2.6: Bar graphs displaying ICP-MS data of HDAC8 and Fe concentrations. Concentration of HDAC8 (gray) and the concentration of Fe(II) (A: blue, B: green) remaining after incubation with inhibitors ($[I]=3 \times IC_{50}$). A) HPT series of inhibitors at 1:1 concentration Fe(II) to HDAC8. B) HPO inhibitors at 5:1 concentration Fe(II) to HDAC8. The expanded inhibitors are listed on the x-axis in order of increasing IC_{50} value. Each bar represents the average of duplicated experiments ($N=2$) for concentration of HDAC8 by Bradford assay ($n=3$) and concentration of Fe(II) by ICP-MS ($n=3$).

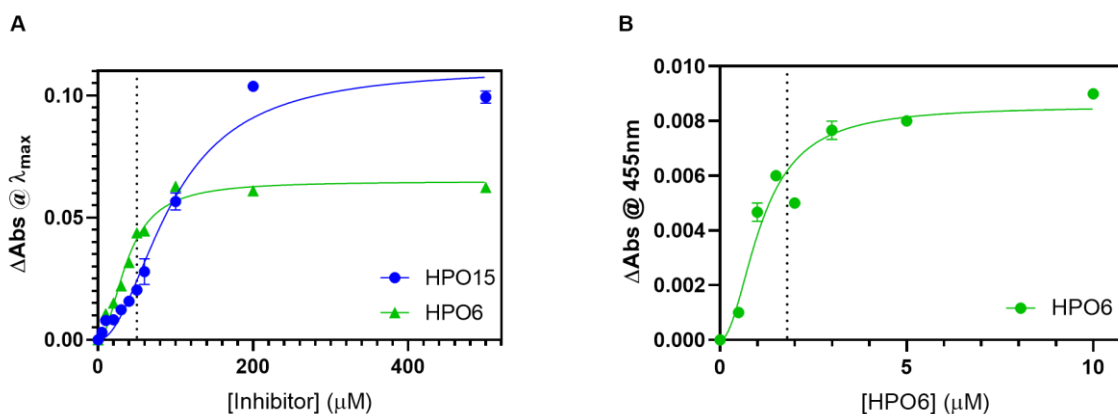


Figure 2.7: Inhibitor binding to Fe(II) as measured by UV-Vis spectroscopy, Data fit using **Equation 2.2**. Each point represents an average ($n=3$) with error bars displaying the standard deviation. A) Titration curve of increasing amounts of inhibitor against $50 \mu\text{M}$ Fe(II). The HDAC8-Fe(II) specific inhibitor, HPO15 (blue), produces a binding curve where the measured K_D value is larger than $50 \mu\text{M}$. The metal stripping inhibitor, HPO6 (green), produces a binding curve where the measured K_D is lower than $50 \mu\text{M}$ and the point of saturation is closer to the total amount of Fe(II), dashed line. B) Titration curve of increasing amounts of HPO6 against $2 \mu\text{M}$ Fe(II). Again, the measured K_D is lower than $2 \mu\text{M}$ and the point of saturation is close to the total amount of Fe(II), dashed line.

Treatment of HeLa cell lysates

After determining that HPO15 was selective for HDAC8-Fe(II), it was then used to treat HeLa cell lysates to try to evaluate if HDAC8-Fe(II) was present. HeLa lysates had been previously determined to display oxygen sensitive HDAC activity[18] and this sensitivity was confirmed in these separate experiments (**Figure 2.8**, left). If the oxygen sensitivity of the deacetylase activity is due to the presence of HDAC8-Fe(II) then addition of HPO15 is predicted to decrease the deacetylase activity of the anaerobic lysates as much or more than the decrease observed by exposure to oxygen. However, incubation of HPO15 with cell lysates did not significantly inhibit the deacetylase activity in anaerobic lysates (**Figure 2.8**, inset). Thus, HPO15 has little effect on the deacetylase activity in HeLa cell lysates and therefore does not provide information about the identity of the active site metal ion.

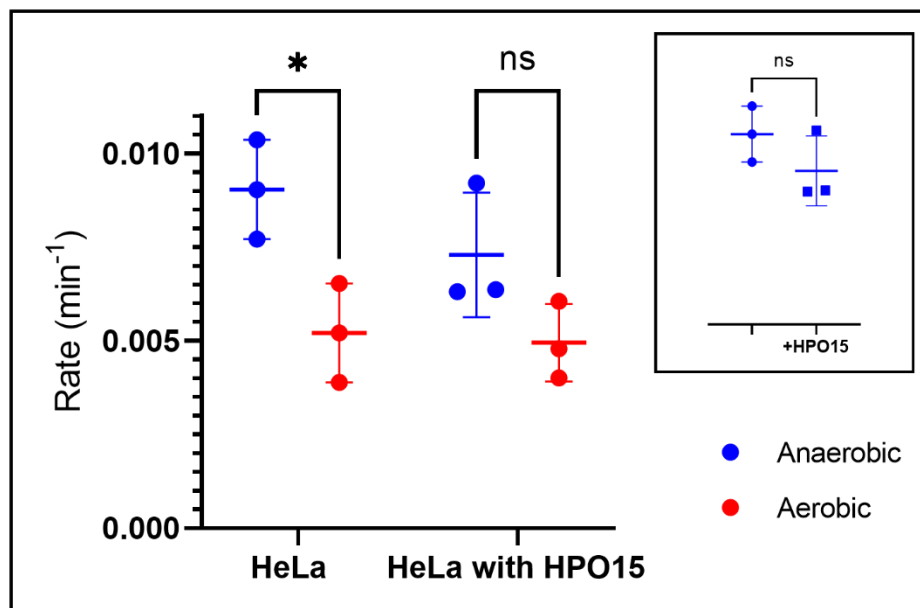


Figure 2.8: Deacetylase activity of HeLa cell lysates with and without treatment with 100µM HDAC8-Fe(II) specific inhibitor HPO15. The untreated lysates display significantly different activity when prepared anaerobically vs aerobically ($p < 0.05$, left). The anaerobic activity of the HPO15 treated lysates is not significantly different than the untreated lysates (inset). Each point represents an average ($N=3$) rate from stopped timepoints ($n=2$).

Discussion

The CFL1 fragment library consisted of chelating groups known to inhibit metalloproteins. Several of the chelating groups, including pyrimidines and carboxylic acids, have been used in described HDAC inhibitors, yet they were not the most potent inhibitors. The most effective hit fragments were tropolone fragments.

The two tropolone fragments, G11 and G12, have very similar structures with G12 having an isopropyl group in the β -position. Due to the increased potency of G12 over G11, it would seem the additional contacts created by this alkyl group are important for binding. Ononye et al. has additionally explored tropolones as potent HDAC inhibitors and found various substituted α - and β -tropolones were selective inhibitors for HDAC2

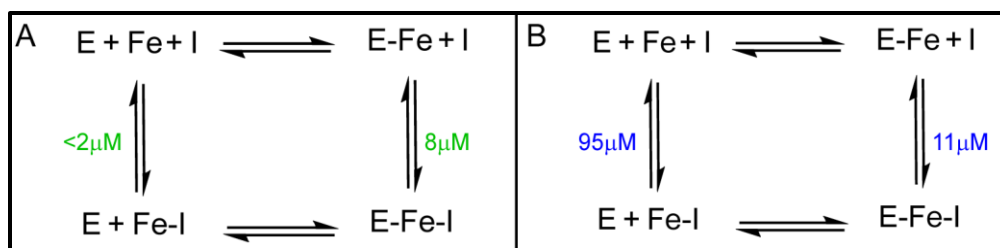
and HDAC8[27]. Tropolones are developed based on a natural product scaffold and have multiple sites of modification to impart enhanced binding and selectivity[27]. The metal binding moiety is additionally much more stable than the hydroxamate counterpart[27]. These positive qualities, combined with the facts that tropolone compounds were the best hits from the fragment library and the only compounds to have sub-micromolar IC_{50} values, point to tropolone compounds as promising future targets for HDAC inhibition.

An interesting outcome of the work presented in this chapter is the identification of metal selective inhibitors. Furthermore, the metal selectivity of the inhibitors was confirmed to not be due to the sequestering the labile Fe(II) by determining the relative affinity to the inhibitors to Fe(II) and HDAC8-Fe(II) through UV-Vis and ICP-MS assays. The selectivity found amongst the inhibitors for solely HDAC8-Fe(II) indicates potential active site structural differences that are dependent on the metal bound with the HDAC8-Fe(II) structure distinct from HDAC8-Zn(II) and HDAC8-Co(II). Structural differences between the reconstituted metalloforms have been proposed previously, as the various metalloforms have differing activity substrate selectivity[17, 22]. Crystallography structural studies have not shown any differences when HDAC8 is bound to a hydroxamate inhibitor, therefore illustrating the need to solve structures of HDAC8 bound to metal selective inhibitors to capture the metalloform structure differences recognized in these other studies.

We modified the two-best metal-selective library fragments, 4D and 5D, in an attempt to enhance potency and selectivity. The series of expanded inhibitors, HPO and HPT, were found to follow the same metal dependence as the initial fragments. For the HPO series, each expanded inhibitor had a lower IC_{50} than the parent fragment, 4D. The

HPT series, however, did not follow the same trend. None of the expanded fragments had a significantly lower IC_{50} than the parent fragment, 5D. Thus, the addition of a linker and cap group did not enhance binding of the HPT inhibitors as it did for the HPO inhibitors meaning that potency of the HPT series was determined by the metal binding group itself. This is surprising considering 4D and 5D only differ by a single atom, oxygen for 4D and sulfur for 5D. These data suggest that the presence of two oxygen atoms as donors for metal binding are more suited for chelation to HDAC8-Fe(II).

While we were able to develop inhibitors with lower IC_{50} values than the parent fragments, there was no correlation between linker length and IC_{50} . However, there was a correlation between linker length and metal stripping propensity. For both the HPO and HPT series, longer linker lengths corresponded to enzyme specific inhibition, whereas shorter linker lengths corresponded to increased metal stripping. This implies the chelator group has preference for binding free Fe(II), as is expected considering these groups were selected due to their metal chelating ability, but the presence of a linker group mimicking the lysine side chain enhanced the affinity for binding HDAC8-Fe(II) (**Scheme 2.2**). We also saw inhibitor binding free Fe(II) exhibits positive cooperativity where $h \sim 2$, indicating a dependence of the mode of inhibition on inhibitor concentration, however as long as the inhibitor exhibits greater affinity for HDAC8-Fe(II) this is not an issue. This highlights the importance of testing for relative affinities when dealing with a labile active site metal ion, as the most potent inhibitors are not always the most effective.



Scheme 2.2: Proposed thermodynamic boxes to illustrate differences in stripping and selective inhibitors. The K_D for Fe(II) is given on the left side and the IC_{50} is given on the right side. A) Box for stripping inhibitor HPO6. The preference for binding free Fe(II) is given from the low K_D . B) Box for selective inhibitor HPO15. The preference for binding HDAC8-Fe(II) is given from the lower IC_{50} than the K_D .

We finally attempted to use the best HDAC8-Fe(II) specific inhibitor, HPO15, to determine if HDAC8 is present as HDAC8-Zn(II) or HDAC8-Fe(II) in the cell. However, the inhibitor treatment could not determine with certainty the presence of HDAC8-Fe(II). Biological replicates were unable to demonstrate a consistent decrease in activity comparable to aerobic activity. The lack of significance could be caused by other interactions that could occur in the more complex environment of cell lysates. It is also unknown whether HPO15 is specific to HDAC8 and most likely is not based on the need for additional modifications for tropolones to achieve HDAC isozyme selectivity[27]. HPO15 may potentially even bind other metalloenzymes, proteins etc. To achieve more confident results, further structural refinement of HPO15 to increase isozyme specificity and increase inhibition efficiency would be necessary.

The work presented in this chapter represents an additional steppingstone in the research of HDAC inhibition and metal identity. We have identified the tropolone metal binding group as a potent HDAC inhibitor and further work refining the structure to increase selectivity will assist in targeted therapeutic development. This work was unable to concretely show whether HDAC8-Fe(II) exists in a significant amount in the cell, but we were able to develop Fe-specific inhibitors and confirm this metal specificity was due to Fe(II) bound to HDAC8 by ICP-MS analysis. The development of HDAC8-Fe(II) specific

inhibitors indicates there are structural differences in the different metal-reconstituted forms of HDAC8. Additional structural studies with inhibitors containing differing metal ligands could help elucidate these structural differences and provide insight into the differing substrate selectivity of the metal-reconstituted forms.

References

1. Hu, E., et al., *Cloning and Characterization of a Novel Human Class I Histone Deacetylase That Functions as a Transcription Repressor*. Journal of Biological Chemistry, 2000. **275**(20): p. 15254-15264.
2. Gao, L., et al., *Cloning and Functional Characterization of HDAC11, a Novel Member of the Human Histone Deacetylase Family*. Journal of Biological Chemistry, 2002. **277**(28): p. 25748-25755.
3. Verdel, A. and S. Khochbin, *Identification of a New Family of Higher Eukaryotic Histone Deacetylases: COORDINATE EXPRESSION OF DIFFERENTIATION-DEPENDENT CHROMATIN MODIFIERS*. Journal of Biological Chemistry, 1999. **274**(4): p. 2440-2445.
4. Tong, J.J., et al., *Identification of HDAC10, a novel class II human histone deacetylase containing a leucine-rich domain*. Nucleic acids research, 2002. **30**(5): p. 1114-1123.
5. Fischle, W., et al., *A New Family of Human Histone Deacetylases Related to Saccharomyces cerevisiae Hda1p*. Journal of Biological Chemistry, 1999. **274**(17): p. 11713-11720.
6. Grozinger, C.M., C.A. Hassig, and S.L. Schreiber, *Three proteins define a class of human histone deacetylases related to yeast Hda1p*. Proceedings of the National Academy of Sciences, 1999. **96**(9): p. 4868-4873.
7. Yang, W.-M., et al., *Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3*. Proceedings of the National Academy of Sciences, 1996. **93**(23): p. 12845-12850.
8. Li, Y. and E. Seto, *HDACs and HDAC Inhibitors in Cancer Development and Therapy*. Cold Spring Harb Perspect Med, 2016. **6**(10).
9. Flipo, M., et al., *Hydroxamates: Relationships between Structure and Plasma Stability*. Journal of Medicinal Chemistry, 2009. **52**(21): p. 6790-6802.
10. Zhang, L., et al., *Zinc binding groups for histone deacetylase inhibitors*. J Enzyme Inhib Med Chem, 2018. **33**(1): p. 714-721.
11. Patil, V., et al., *3-Hydroxypyridin-2-thione as novel zinc binding group for selective histone deacetylase inhibition*. J Med Chem, 2013. **56**(9): p. 3492-506.
12. Roche, J. and P. Bertrand, *Inside HDACs with more selective HDAC inhibitors*. Eur J Med Chem, 2016. **121**: p. 451-483.

13. Amin, S.A., N. Adhikari, and T. Jha, *Structure-activity relationships of HDAC8 inhibitors: Non-hydroxamates as anticancer agents*. *Pharmacol Res*, 2018. **131**: p. 128-142.
14. Singh, A.K., A. Bishayee, and A.K. Pandey, *Targeting Histone Deacetylases with Natural and Synthetic Agents: An Emerging Anticancer Strategy*. *Nutrients*, 2018. **10**(6).
15. Vannini, A., et al., *Crystal structure of a eukaryotic zinc-dependent histone deacetylase, human HDAC8, complexed with a hydroxamic acid inhibitor*. *Proc Natl Acad Sci U S A*, 2004. **101**(42): p. 15064-9.
16. Dowling, D.P., et al., *Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function*. *Biochemistry*, 2010. **49**(24): p. 5048-56.
17. Gantt, S.L., S.G. Gattis, and C.A. Fierke, *Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion*. *Biochemistry*, 2006. **45**(19): p. 6170-8.
18. Gattis, S.G., *Mechanism and metal specificity of zinc-dependent deacetylases*, in *Biological Chemistry*. 2010, University of Michigan.
19. Doyle, K. and F.A. Fitzpatrick, *Redox signaling, alkylation (carbonylation) of conserved cysteines inactivates class I histone deacetylases 1, 2, and 3 and antagonizes their transcriptional repressor function*. *J Biol Chem*, 2010. **285**(23): p. 17417-24.
20. Ago, T., et al., *A redox-dependent pathway for regulating class II HDACs and cardiac hypertrophy*. *Cell*, 2008. **133**(6): p. 978-93.
21. Jansch, N., et al., *The enzyme activity of histone deacetylase 8 is modulated by a redox-switch*. *Redox Biol*, 2019. **20**: p. 60-67.
22. Castaneda, C.A., et al., *Active Site Metal Identity Alters Histone Deacetylase 8 Substrate Selectivity: A Potential Novel Regulatory Mechanism*. *Biochemistry*, 2017. **56**(42): p. 5663-5670.
23. Agrawal, A., et al., *Chelator fragment libraries for targeting metalloproteinases*. *ChemMedChem*, 2010. **5**(2): p. 195-9.
24. Kim, B., A.S. Pithadia, and C.A. Fierke, *Kinetics and thermodynamics of metal-binding to histone deacetylase 8*. *Protein Sci*, 2015. **24**(3): p. 354-65.
25. Abdi, H., *The Bonferroni and Sidak Corrections for Multiple Comparisons*, in *Encyclopedia of Measurement and Statistics*, N. Salkind, Editor. 2007, Sage: Thousand Oaks, CA. p. 103-107.

26. Ku, H.H., *Notes on the Use of Propagation of Error Formulas*. Journal of Research of the National Bureau of Standards - C. Engineering and Instrumentation, 1966. **70C**(4): p. 263-273.
27. Ononye, S.N., et al., *Tropolones as lead-like natural products: the development of potent and selective histone deacetylase inhibitors*. ACS Med Chem Lett, 2013. **4**(8): p. 757-61.

Chapter 3 : The formation of the CoREST complex enhances HDAC1 deacetylase activity and alters selectivity

Introduction

Post-translational modifications (PTMs) of proteins were first identified 60 years ago[1, 2] and are widespread across the proteome[3]. PTMs can regulate a protein's activity, interactions, cellular localization, and degradation. Acetylation, a reversible PTM on lysine residues, was first discovered on histone tails[4]. Acetylated proteins were long thought to be limited to the nucleus, however high-throughput methods using mass spectrometry have identified thousands of acetylation sites throughout the cell. Since lysine acetylation is reversible, this modification can act as a regulatory switch. On histones, acetylation and deacetylation regulates gene transcription making the enzymes that catalyze these reactions essential for proper cell function.

The deacetylase family is divided into four classes based on homology to yeast deacetylases. Classes I, II, and IV, referred to as histone deacetylases or HDACs, are metal-dependent deacetylases, requiring a divalent metal ion for catalysis within a conserved catalytic domain. HDAC1, a Class I deacetylase, is localized to the nucleus functioning to regulate gene transcription[5-11], cell growth[12, 13], proliferation[14-17], differentiation[18, 19], and apoptosis[13, 20, 21]. Mis-regulation of acetylation has been linked to cancer[22], neurological diseases including Alzheimer's[23-25], and immune disorders[26].

There are still significant gaps in our understanding of misregulation of HDAC functions in biological pathways that lead to disease. Many HDACs do not function alone, rather they form both stable and transient interactions with other proteins. BioGRID, the Biological General Repository for Interaction Datasets, lists almost 3,000 total protein interacting partners for all 11 HDACs, with almost half of those protein interactions connected to HDACs 1 and 2[27]. HDAC1 is known to exist in three, stable nuclear complexes, the NuRD[28], Sin3[29], and CoREST[30] complexes, that are proposed to activate the deacetylase activity of HDAC1.

HDAC1 in complex has demonstrated catalytic activity on all four core histones, with varying efficiency[31]. In contrast, when purified to homogeneity HDAC1 has little to no deacetylase activity[32]. The role of specific protein interactions in enhancing HDAC1 activity has not been explored. Recent studies have focused on purifying and measuring the activity of intact HDAC1-containing complexes purified from eukaryotic expression systems[33-35]. The deacetylase activity of these complexes with both single histone and nucleosome substrates has been shown to vary based on acetylation site with each complex containing its own selectivity [34-36]. The CoREST complex, consisting of lysine demethylase 1 (LSD1), repressor element 1 silencing transcription factor corepressor 1 (CoREST), and HDAC1, has demonstrated the greatest deacetylase activity with the least selectivity[36].

Additionally, HDACs are known to be regulated by post-translational modifications, particularly phosphorylation[37-49]. HDAC1 is doubly phosphorylated at Ser421 and Ser423, and this phosphorylation increases the catalytic activity with full-length protein substrates, i.e. histones, as well as enhancing protein-protein interactions[40, 42]. There

is the potential that phosphorylation regulates both activity and selectivity of HDACs. Such a role has recently been shown with HDAC8 where phosphorylation inhibits deacetylase activity, yet the activity decreases vary between 10 and 100-fold, dependent on the peptide substrate[44].

Here, we analyze the deacetylase activity of HDAC1 constructs expressed from both bacterial and eukaryotic expression systems with peptide and protein substrates and reconstitute the core CoREST complex. These studies demonstrate that protein-protein interactions in the CoREST complex enhance deacetylase activity by more than 10-fold. We additionally probed the effect of phosphorylation and phospho-mimic mutations in HDAC1 on the CoREST complex formation demonstrating that phosphorylation is not necessary to form protein interactions but does change substrate selectivity.

Materials and Methods

Reagents

Amintra Maltose Binding Protein affinity resin was purchased from Expedeon (now Abcam). Ni-NTA agarose resin was purchased from Qiagen. Ni-sepharose fast flow prepacked columns were purchased from GE healthcare. Adenosine triphosphate (ATP), coenzyme A (CoA), NAD⁺, NADH, L-malic acid, malate dehydrogenase (MDH), citrate synthase (CS), and mouse monoclonal anti-polyhistidine-alkaline phosphatase antibody were purchased from Sigma. Rabbit monoclonal anti-HDAC1, anti-LSD1, and anti-CoREST antibodies were purchased from Cell Signaling Technologies. N-terminally acetylated and C-terminally carboxylated, singly acetylated lysine peptides were purchased from Peptide 2.0 or Synthetic Biomolecules. 3% (v/v) acetic acid standard was

purchased from RICCA Chemical. All other materials were purchased from Fisher at >95% purity unless noted otherwise.

Expression and Purification of HDAC1

HDAC1 cDNA (BC00301) was purchased from Horizon Discovery. HDAC1 was cloned into a pFastBac vector with a N-terminal His-MBP tag[50], provided by Dr. Clay Brown at the Life Sciences Institute (LSI) Center for Structure Biology, using Gibson Assembly (New England Biolabs). HDAC1 was also cloned into a modified pET12a vector containing a N-terminal 6xHis-SUMO tag by restriction enzyme digestion (KpnI, N-terminus, and XhoI, C-terminus). The 6x-His-SUMO-HDAC1 construct was further modified using site-directed mutagenesis to replace the two serines phosphorylated *in vivo*, S421 and S423, with glutamates to mimic phosphorylation. All primers are given in **Table 3.1**. Proper gene insertion and mutagenesis was confirmed using Sanger sequencing at the University of Michigan sequencing core.

For insect expression, the High Throughput Protein Lab at LSI then expressed the His-MBP-HDAC1 in Tni insect cells grown in 2 L of serum free media at 27°C for 72 hours[50]. The resulting pellet was resuspended in lysis buffer (50 mM Tris pH 8, 200 mM NaCl, 1 mM TCEP, 10% glycerol, 1% Tween-20). The cells were lysed by end over end rotation for 20 minutes at 4°C. The lysate was cleared by centrifugation at 30,000 x g, 45 minutes, at 4°C. The cleared lysate was incubated with 5 mL of Amintra MBP Affinity resin pre-equilibrated with wash buffer (50 mM Tris pH 8, 200 mM NaCl, 1 mM TCEP, 10% glycerol) by end over end rotation for 2 hours. Using a gravity column, the lysate flow-through was collected and the resin was washed with at least 20 column volumes (CVs) of wash buffer. The His-MBP-HDAC1 was eluted with 5 CVs of elution buffer (50 mM Tris

pH 8, 200 mM NaCl, 1 mM TCEP, 10% glycerol, 10 mM maltose). The presence of His-MBP-HDAC1 was confirmed by SDS-PAGE with Coomassie staining. The sample was dialyzed into storage buffer (50 mM Tris pH 8, 200 mM NaCl, 1 mM TCEP, 10% glycerol) at 4°C overnight, followed by concentration using an Amicon 30K MWCO. The estimated purity was 80%. The concentration was determined under native conditions by A_{280} with the calculated extinction coefficient $122,050 \text{ M}^{-1}\text{cm}^{-1}$, aliquoted and flash frozen.

For *E. coli* expression, pET21a containing either WT or the S421E/S423E 6x-His-SUMO-HDAC1 was transformed into BL21(DE3) competent cells and the cells were plated onto LB-agar containing 100 $\mu\text{g}/\text{mL}$ ampicillin. A single colony was selected and used to inoculate a 2xYT starter culture containing 100 $\mu\text{g}/\text{mL}$ ampicillin. The starter culture was grown at 37°C, 250 rpm until cloudy, approximately 4-6 hours. 10 mL starter culture was then used to inoculate 1-L 2xYT containing ampicillin. The cells were grown at 37°C, 150-200 rpm, until the $\text{OD}_{600} = 0.4-0.6$. The cultures were cold-shocked on ice for 10 min and induced with 0.5 mM isopropyl β -D-1-thiogalactopyranoside (IPTG). The cultures were grown for 16-18 hours at 18°C, 150-200 rpm. Harvested cells were resuspended in lysis buffer without detergent (50 mM Tris pH 8, 200 mM NaCl, 1 mM TCEP, 10% glycerol) and lysed using a microfluidizer (Microfluidics). The lysate was cleared as previously described for the insect purification. The cleared lysate was incubated with 5 mL of Ni-NTA resin pre-equilibrated with wash buffer by end over end rotation for 2 hours. Using a gravity column, the lysate flow-through was collected and the resin was washed with at least 20 CVs of wash buffer. The 6xHis-SUMO-HDAC1 was eluted using a stepwise imidazole gradient up to 500 mM. The resulting fractions were analyzed for HDAC1 by SDS-PAGE with Coomassie staining. Fractions containing 6xHis-

SUMO-HDAC1 were dialyzed into storage buffer at 4°C overnight followed by concentration using an Amicon 30K MWCO centrifugal filter. The estimated purity was 50%. The protein concentration was determined using BCA assay (Thermo) with bovine serum albumin (BSA) standards, aliquoted and flash frozen.

Expression and Purification of LSD1

LSD1 cDNA (O60341) was purchased from Sino Biological. LSD1 was cloned into a pET28a vector with a N-terminal thrombin cleavable 6xHis tag by NdeI (N-terminal) and XhoI (C-terminal) restriction enzyme digest. Primers are given in **Table 3.1**. The plasmid was transformed into Rosetta2(DE3)-pLysS competent cells and the cells plated on LB-agar containing 50 µg/mL kanamycin and 34 µg/mL chloramphenicol. A single colony was selected to inoculate a 20 mL 2xYT starter culture containing kanamycin and chloramphenicol. The starter culture was grown at 37°C, 200 rpm until cloudy. 2 L of 2xYT containing kanamycin and chloramphenicol was inoculated with the 20 mL culture and grown at 37°C, 200 rpm until the OD₆₀₀ was 0.6. The flasks were cold-shocked for 10 min on ice, induced by addition of 0.2 mM IPTG and grown 16-18 hours at 16°C, 200 rpm. Harvested pellets were resuspended in lysis buffer without detergent, lysed using a microfluidizer and the lysate was cleared as previously described. Using a gravity column, the cleared lysate was added to 5 mL Ni-NTA resin preequilibrated with wash buffer. The column was washed with 10-20 CVs of wash buffer and the 6xHis-LSD1 was eluted with a stepwise imidazole gradient up to 500 mM. The resulting fractions were analyzed for LSD1 by SDS-PAGE and Coomassie staining. Fractions containing 6xHis-LSD1 were dialyzed into cleavage buffer (storage buffer containing 20 mM CaCl₂) at 4°C overnight

followed by concentration using an Amicon 30K MWCO centrifugal filter, concentration determined by BCA assay, aliquoted and flash frozen. The estimated purity was 90%.

Table 3.1: Primers for various construct cloning. Restriction enzymes sites are in bold italics. Mutagenesis sites are in bold/underlined.

Construct	Forward Primer	Reverse Primer
His-MBP-HDAC1	TACTTCCAATCCAATGCAATGGCGCAGACGCAGGG	TTATCCACTTCCAATGTCAGGCCAAC TTGACCTCCTC
His-SUMO-HDAC1	ACT GGTACC ATGGCGCAGACGCAGGG	ATAG CTCGAGT CAGGC CAACTTGACCTCCTC
His-SUMO-HDAC1 S421E	CTGTGAGGAAGAGTTC GAA GATTCTGAGAGAGGGGAGAG	CTCTCCCTCCTCTTCAGAATC TTCGA ACTCTTCCTCACAG
His-SUMO-HDAC1 S421E/S423E	TGAGGAAGAGTTCGAAGAT GAA GAAGAGGAGG	CCTCCTCTTCT TTC ATCTTCGAACTCTT CCTCA
His-LSD1	GTACAT ATGTTATCTGGGAAGAAGGC	ATATCT CGAGTCACATGCTTGGGGAC TG

Expression and Purification of CoREST

A pET28a vector containing CoREST with a N-terminal thrombin cleavable 6xHis tag was generously provided by Prof. Wenshe Liu at Texas A&M University. 6xHis-thrombin-CoREST plasmid was transformed into BL21(DE3) competent cells and cells plated onto LB-agar containing 50 µg/mL kanamycin. Following transformation, the 6xHis-thrombin-CoREST protein was expressed as previously described for 6xHis-LSD1. Following expression, 6xHis-thrombin-CoREST was purified as previously described for LSD1 but with the following alterations: the purification only used an imidazole gradient up to 200 mM and for concentration a 10K MWCO Amicon centrifugal filter was used. The estimated purity was 60%.

Expression and Purification of LSD1-CoREST complex

An additional pET15 vector containing LSD1 with a N-terminal thrombin cleavable 6xHis tag was generously provided by Prof. Wenshe Liu at Texas A&M University. The pET15 LSD1 vector was co-transformed with the pET28 CoREST vector in BL21(DE3)

competent cells and cells plated onto LB-agar containing 100 µg/mL ampicillin and 50 µg/mL kanamycin and expressed according to the previously described LSD1 protocol but with ampicillin and kanamycin selection. Harvested cells were resuspended and lysed as previously described. The cleared lysate was added onto a pre-packed 5 mL Ni-NTA column at 1 mL/min. The column was washed until the UV signal subsided. LSD1-CoREST was then eluted using a linear gradient up to 500 mM imidazole. The resulting fractions were analyzed for LSD1-CoREST by SDS-PAGE with Coomassie staining. Fractions containing LSD1-CoREST were pooled and dialyzed overnight in cleavage buffer. Following dialysis, pooled fractions were analyzed by size exclusion chromatography, using a Superdex S200 analytical column equilibrated with dialysis buffer, to confirm the presence of the LSD1-CoREST complex. The pooled fractions containing LSD1-CoREST complex, at 80% purity, were concentrated using a 10K MWCO Amicon centrifugal filter.

6xHis Tag Cleavage

The 6xHis tag was cleaved using the Thrombin CleanCleave™ kit by Sigma following the manufacturer's instructions with the following specifications. Briefly, 1-2 mg of protein was incubated with the thrombin agarose beads at room temperature for 3 hours. The cleaved protein was added to a 1 mL prepacked Ni-FF column. The column was washed with 5 mL of lysis buffer. The resulting fractions were analyzed by SDS-PAGE Coomassie. Fractions containing cleaved CoREST were pooled and concentrated. The concentration was determined by BCA assay.

Co-Immunoprecipitations

The various constructs of HDAC1 (6xHis-MBP-HDAC1, 6xHis-SUMO-HDAC1, and 6xHis-SUMO-HDAC1_S421E/S423E) and (6xHis-LSD1)-(6xHis-CoREST) were incubated overnight at a 2:1 molar ratio in phosphate buffered saline (PBS) with rotation at 4°C. LSD1 antibody was added at a 1:100 dilution and allowed to further incubated with rotation at 4°C for one hour. The sample was then added to Dynabeads™ Protein G (Thermo) and incubated with rotation at room temperature for 10 minutes. The beads were washed 3 times with 200 µL PBS. The washed beads were transferred to a separate microtube and eluted with elution buffer (50 mM glycine pH 2.8) and SDS loading buffer (50mM Tris pH 6.8, 1% β-mercaptoethanol, 0.0004% bromophenol blue, 6% glycerol, 2% sodium dodecyl sulfate). The samples were boiled at 95°C for 5 minutes. The eluted sample was run on an SDS-PAGE gel followed by semi-dry transfer onto nitrocellulose membrane. The membrane was blocked with 5% BSA in Tris Buffered Saline with Tween-20 (TBST) for 45 minutes at room temperature with rocking. Primary antibodies corresponding to either HDAC1 (1:2000 dilution in TBST) or LSD1 and CoREST (together, 1:1000 and 1:500 dilution in TBST, respectively) was then incubated with the membrane for 90 minutes at room temperature with rocking. The membrane was washed for 5 minutes, 3 times with TBST. Goat anti-rabbit secondary conjugated to HRP (1:10,000 in TBST) was incubated with membrane for 60 minutes at room temperature. Following an additional washing step, the membrane was monitored for chemiluminescence using a ChemiDoc imager (BioRad) and SuperSignal™ West Pico PLUS chemiluminescent substrate (Thermo).

Preparation of Singly Acetylated Histone H3

Singly acetylated *Xenopus leavis* histone H3 had been previously purified (as detailed in Castaneda et al.[51]) and stored as lyophilized protein at -20°C. Lyophilized H3 was resuspended in assay buffer (25 mM HEPES pH8, 137 mM NaCl, 5 mM KCl) at room temperature over 20 minutes with gentle agitation. Any insoluble particulate was removed by centrifugation at 10,000 x g for 5 minutes. Resuspended H3 was dialyzed overnight in assay buffer at 4°C using 3500 MWCO Slide-a-Lyzer™ dialysis cassettes (Thermo). The concentration of H3 sample following dialysis was determined by Bradford assay (Thermo) with bovine serum albumin (BSA) standards. The sample was then aliquoted and flash frozen.

Coupled-enzyme Acetate Detection Assay

Acetyl CoA Synthetase (ACS) was expressed and purified as previously described[52, 53].

The coupled acetate-detection assay or simply the 'acetate assay' was performed as previously described with a few modifications[53]. Briefly, lyophilized peptides were re-suspended in water. Peptide concentration was determined by the BCA assay. Reactions containing 100-500 μM peptide or 10-30 μM protein in assay buffer were initiated with 0.1-1 μM HDAC1 at 37°C. Timepoints, 60 μL, were quenched with 5 μL of 10% hydrochloric acid and kept on ice. Timepoints were flash frozen with liquid nitrogen and stored at -80°C until work-up.

Coupled solution (50 mM HEPES, pH 8, 400 μM ATP, 10 μM NAD⁺, 30 μM CoA, 0.07 U/μL CS, 0.04 U/μL MDH, 50 μM ACS, 100 mM NaCl, 3 mM KCl, 50 mM MgCl₂, and

2.5 mM L-malic acid) was prepared the day of the work-up and incubated at room temperature away from light for at least 25 minutes. Timepoints were thawed and neutralized with 15 μ L of freshly prepared and filtered 6% sodium bicarbonate. Neutralized timepoints and each acetate standard, 60 μ L, were added to 10 μ L coupled solution in a black, flat-bottomed, half-area, non-binding, 96-well plate (Corning No. 3686). The resulting NADH fluorescence (Ex 340 nm, Em 460 nm) was measured on a PolarStar fluorescence plate reader until the signal reached equilibrium. Using the acetate standard curve, the final fluorescence of each timepoint was converted to μ M acetate product and the initial velocity of the reaction was determined from the time-dependence of the appearance of acetate. The catalytic efficiency of the reaction (k_{cat}/K_M) was determined using **Equation 3.1**.

$$\frac{v_o}{[E]} = \frac{k_{cat}[S]}{(K_M + [S])} \quad \text{Equation 3.1}$$

Results

Deacetylase activity of recombinant HDAC1 is dependent upon substrate length

Full-length human HDAC1 was initially expressed in *Escherichia coli* (*E. coli*), rEchHDAC1, however this enzyme was inefficient ($k_{cat}/K_M < 0.5 \text{ M}^{-1}\text{s}^{-1}$) at catalyzing deacetylation of acetylated peptide substrates. Additionally, deacetylase activity against peptide substrates was not increased by preparing a mutant that contains two glutamate side chains mimicking proposed *in vivo* phosphorylation sites (Ser421Glu/Ser423Glu; phosphomimic mutant) shown to be important for activation of deacetylase activity[40]. Full-length human HDAC1 expressed in Tni insect cells, rTniHDAC1, and purified to

homogeneity contained increased deacetylase activity against peptide substrates ($k_{cat}/K_M = 2.3 \pm 0.3 \text{ M}^{-1}\text{s}^{-1}$; **Figure 3.1-A**). The HDAC1 expressed in insect cells was additionally shown to be phosphorylated using the Pro-Q™ Diamond phosphoprotein stain (**Figure 3.2**). However, the increased deacetylase activity was still minimal, particularly compared to HDAC1 purchased from BPS Biosciences. To explore whether HDAC1 catalyzes deacetylation of protein substrates more readily due to additional interactions, as has previously been shown with HDAC8[51], we measured deacetylase activity using singly acetylated *Xenopus leavis* Histone H3 prepared using non-canonical amino acid incorporation[51].

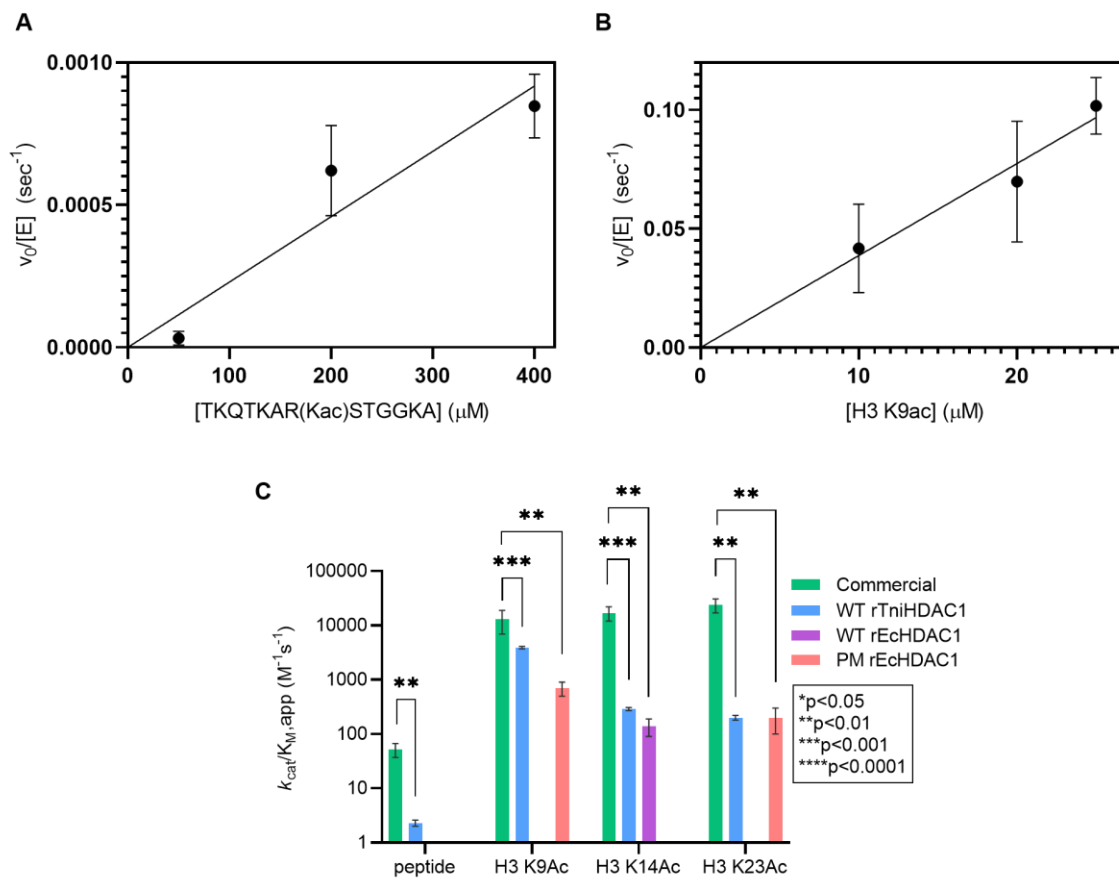


Figure 3.1: Deacetylation rate of recombinant HDAC1 varies with substrate

A) Dependence of WT rTniHDAC1 reaction rate on the substrate concentration of a H3K9ac peptide analog (TKQTARK(ac)STGGKA) measured using the acetate assay. B) Dependence of WT rTniHDAC1 reaction rate on substrate concentration of singly acetylated Histone H3 at Lys9 (H3 K9Ac). C) Comparison of catalytic efficiencies of HDAC1 constructs catalyzing deacetylation of Histone H3 substrates. Exact values are given in Table 2. Bars represent the average value (N=2-4) with error bars representing S.D. Denoted p-values are calculated for comparison to the value of the commercially available HDAC1 (green) by t-test with Sidak's correction for multiple comparisons[54]. #PM: phosphomimic, S421E/S423E.

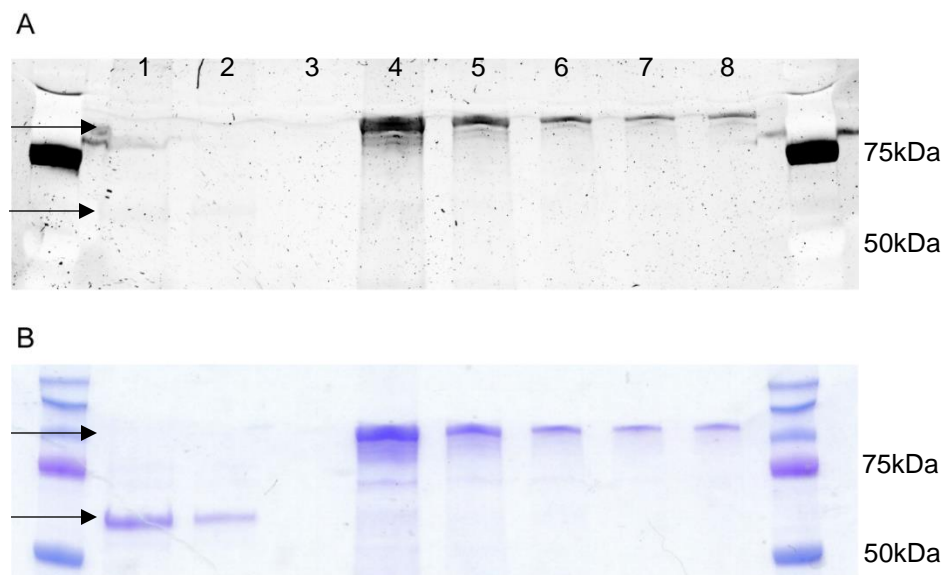


Figure 3.2: Detection of phosphorylated HDAC1

A) Pro-Q Diamond Phosphoprotein stained SDS-PAGE gel. Lanes 1-2: Serial dilutions of WT rEchHDAC1 (~55 kDa). Lane 3: empty. Lanes 4-8: Serial dilutions of MBP-rTniHDAC1 (~97 kDa). B) The same gel from A) stained with Coomassie following the fluorescent stain. Arrows indicate where to expect the bands based on the molecular weight.

To measure deacetylase activity with an acetylated protein substrate, we used the acetate assay previously developed for peptide substrates[53] to measure the production of free acetate. We measured acetate production in reactions of HDAC1 with H3 acetylated at lysine-9 (K9ac) with and without the potent inhibitor SAHA, where the presence of SAHA decreased the measured activity ($<100 \text{ M}^{-1}\text{s}^{-1}$) demonstrating that the observed activity was catalyzed by HDAC1. Both the *E. coli* phosphomimic (PM) mutant and recombinant HDAC1 expressed in insect cells catalyze deacetylation of H3 K9ac more rapidly than the peptide analog. For rTniHDAC1, the increase is about 1000-fold (**Figure 3.1-A vs. Figure 3.1-B**), and the increase for the PM rEchHDAC1 was >5000-fold. Interestingly, the wild-type (WT) rEchHDAC1 did not display detectable deacetylase activity ($k_{\text{cat}}/K_{\text{M}} < 0.05 \text{ mM}^{-1}\text{s}^{-1}$) using histone H3 K9ac. When the substrate scope was expanded to additional histone H3 acetylation sites, lysine-14 (K14ac) and lysine-23

(K23ac), WT rEcHDAC1 catalyzed deacetylation of H3 K14ac at an observable rate ($k_{cat}/K_M = 0.14 \text{ mM}^{-1}\text{s}^{-1}$). In comparison, the PM rEcHDAC1 had broader site selectivity and displayed measurable deacetylase activity towards H3 K9ac and H3 K23ac but not H3 K14ac indicating that the phosphomimic mutations altered substrate selectivity. WT rTniHDAC1 catalyzed deacetylation of all three H3 acetylation sites tested (**Table 3.2**).

When the activity towards both peptide and protein substrates was compared to commercially available HDAC1, purchased from BPS Biosciences, all constructs were significantly less active (**Figure 3.1-C**). However, HDAC1 exists in multiple stable protein complexes *in vivo*[28-30], and these protein-protein interactions are proposed to enhance the activity of HDAC1[32]. The commercially available HDAC1 contains additional proteins that might enhance activity. In contrast, HDAC1 expressed in either insect cells or *E. coli* has fewer and/or different protein contaminants than the commercially available material (**Figure 3.3**). Previously HDAC1 has been purified in the respective *in vivo* complexes and demonstrated to have significant catalytic activity[33-35]. To test the function of the protein-protein interactions we were interested in reconstituting the minimal CoREST complex *in vitro* and analyze the effect of individual constituents of the complex on HDAC1 activity and selectivity.

Table 3.2: Catalytic efficiencies of HDAC1 constructs dependent on expression system

k_{cat}/K_M (mM ⁻¹ s ⁻¹) ^a							
WT (Insect)		WT (<i>E. coli</i>)		S421E/S423E (<i>E. coli</i>)		Comm. ^f	
HDAC1 ^b	HDAC1- LSD1- RCOR1 ^c	HDAC1 ^b	HDAC1- LSD1- RCOR1	HDAC1 ^b	HDAC1- LSD1- RCOR1 ^c	HDAC1 ^c	
Peptide ^d	0.0023±0.0003	0.034±0.005	<0.0001 ^e	n.d ^g	<0.0001 ^e	n.d ^g	0.052±.015
H3 K9ac	3.9±0.2	30±10	<0.05 ^e	29±9	0.7±0.2	30±20	13±6
H3 K14ac	0.29±0.02	18±6	0.14±0.05	15±7	<0.5 ^e	20±10	17±5
H3 K23ac	0.20±.02	30±20	<0.05 ^e	<5 ^e	0.2±0.1	20±10	24±7

^aEach value is calculated as an average of initial velocities (N=2-4) at [substrate]= 15 μM. The initial velocity was calculated from a stopped time course (n=3-4). Error represents S.E.M. calculated using GraphPad Prism analysis.

^b[HDAC1] = 1 μM. ^c[HDAC1] = 0.1 μM. ^d13-mer peptide based on H3 K9ac: TKQTAR(Kac)STGGKA ^eActivity could not be measured, so assay detection limit is given. ^fCommercially available HDAC1. ^gThe value was not determined.

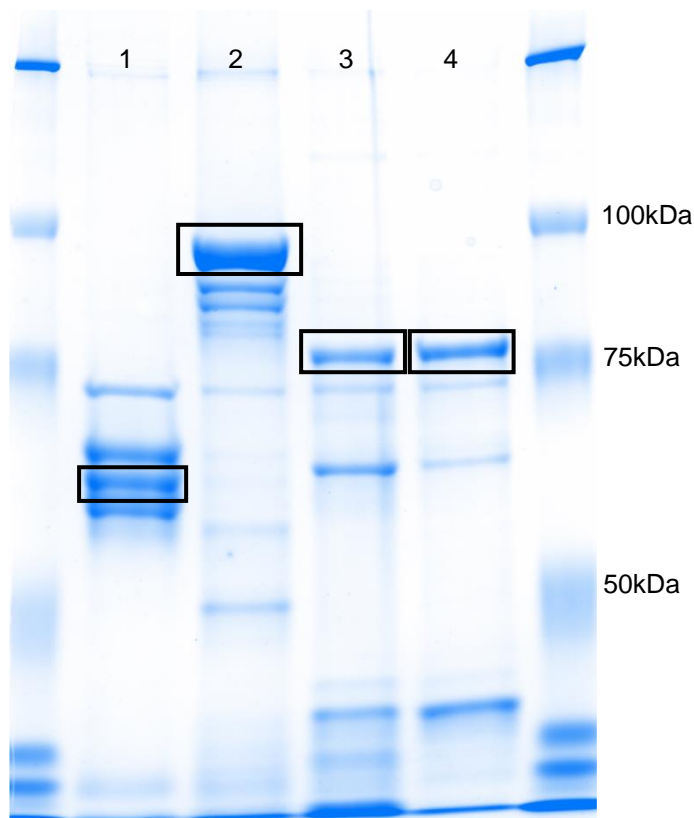


Figure 3.3: Coomassie-stained SDS-PAGE gel to illustrate purity of HDAC1 constructs. 4 μ g of protein was loaded per lane. The HDAC1 band is outlined with a black box. Lane 1 – commercially available HDAC1-FLAG (~56 kDa). Lane 2 – MBP-rTniHDAC1 (~96 kDa). Lane 3 – WT SUMO-rEchHDAC1 (~67 kDa). Lane 4 – PM SUMO-rEchHDAC1 (~67kDa). The two outside lanes are molecular weight standards.

HDAC1 interacts with LSD1-CoREST complex *in vitro*

Full-length human LSD1 and full-length human CoREST were recombinantly expressed and purified from *E. coli*, both individually and together. 6xHis-LSD1 and 6xHis-CoREST were confirmed to co-purify in a complex, hereby referred to as LSD1-CoREST, by size-exclusion chromatography (SEC) (**Figure 3.4**). Only WT rTniHDAC1 and PM rEchHDAC1 co-immunoprecipitated with LSD1 alone using a LSD1 antibody, though all constructs co-immunoprecipitated with LSD1-CoREST, demonstrating the formation of a complex (**Figure 3.5**). Attempts to identify a direct interaction between HDAC1 and CoREST by immunoprecipitation using the 6xHis-tag on HDAC1 were

unsuccessful. Though phosphorylation has been proposed to be necessary for HDAC1 complex formation[40], LSD1-CoREST was shown to co-immunoprecipitate with WT rTniHDAC1, WT rEcHDAC1, and PM rEcHDAC1. While the WT rTniHDAC1 was confirmed to be phosphorylated, the WT rEcHDAC1 does not possess this post-translational modification (**Figure 3.2**). The importance of phosphorylation was apparent for affinity and/or stability of a HDAC1-LSD1 interaction.

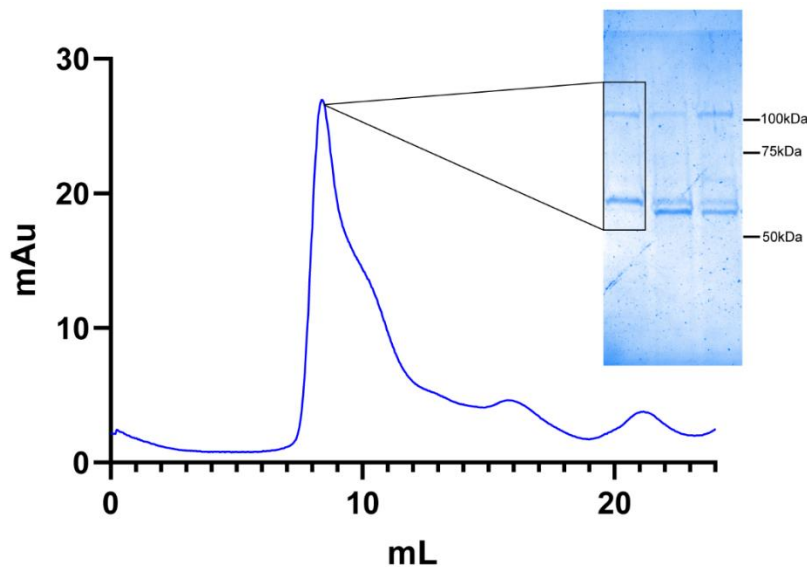


Figure 3.4: Chromatogram of LSD1-CoREST through size-exclusion chromatography
Inset is a Coomassie stained SDS-PAGE gel with the fraction corresponding to the major peak highlighted with a box. The gel indicates there are only two bands present, the top band corresponds to LSD1 (~92 kDa) and the bottom band corresponds to CoREST (~55 kDa).

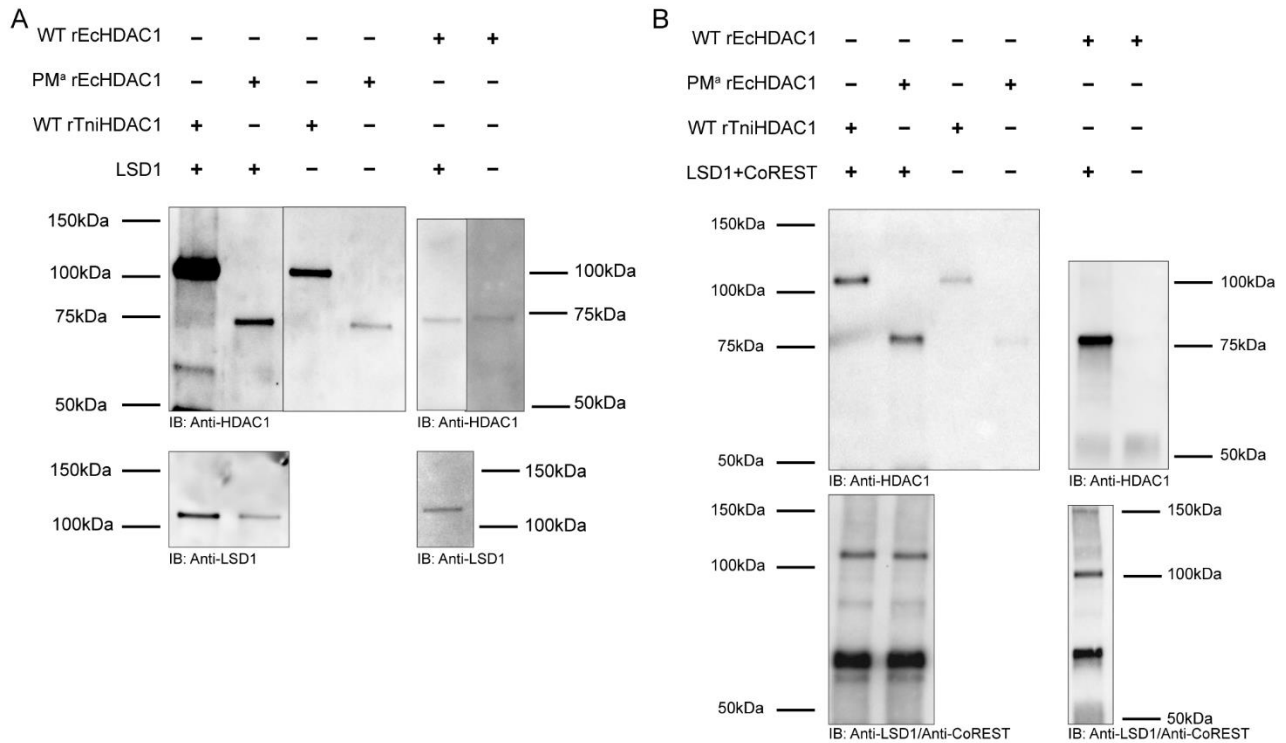


Figure 3.5: Co-immunoprecipitation of HDAC1 using an LSD1 antibody
 Expected molecular weights: rTniHDAC1 96kDa, rEcHDAC1: 67kDa, LSD1: 92kDa, CoREST: 53kDa.
 A) Co-immunoprecipitation of HDAC1 incubated with LSD1. B) Co-immunoprecipitation of HDAC1 incubated with LSD1-CoREST. ^aPM: phosphomimic, S421E/S423E.

Addition of LSD1-CoREST enhances HDAC1 deacetylase activity

The deacetylase activity of an HDAC1-LSD1-CoREST complex reconstituted by incubation of the subunits *in vitro* was determined with both peptide and protein substrates. WT rTniHDAC1 incubated with LSD1 or CoREST individually displayed no enhancement of deacetylase activity with peptide substrate (**Figure 3.6-A**). Similarly, no enhancement of activity was observed when both LSD1 and CoREST, purified individually, were incubated with HDAC1 (**Figure 3.6-A**). However, when WT rTniHDAC1 was incubated with the pre-formed, co-purified LSD1-CoREST complex significant enhancement of deacetylase activity of peptide substrate was observed, comparable with the activity of commercially available HDAC1 (**Figure 3.6-A**). In contrast, the deacetylase activity of HDAC1 purified from *E. coli* with peptide substrates was not enhanced by

incubation with LSD1-CoREST. The catalytic activity of HDAC1-LSD1-CoREST was increased for deacetylation of acetylated H3 (**Figure 3.6-B**). For both insect and PM rEchHDAC1, the enhancement of activity applied to all tested acetylation sites. Interestingly, for the WT rEchHDAC1, activity was enhanced for only H3 K9ac and H3 K14ac while activity for H3 K23ac not detectable (**Table 3.2**). There was no detected deacetylase activity of the LSD1-CoREST complex alone when tested with peptide or protein substrates.

The extent of activation observed upon incubation with the LSD1-CoREST subcomplex varied with the acetylation site; the fold increase in activity ranged from approximately 10 to over 300-fold (**Table 3.3**). WT rTniHDAC1 and PM rEchHDAC1 showed similar activation (within 2-fold) upon addition of the LSD1-CoREST subcomplex for the substrates H3 K9ac and H3 K23ac, although addition of LSC1-CoREST activated deacetylation of H3 K23ac 10-fold more than H3 K9ac. This trend was reversed for WT rEchHDAC1, with LSD1-CoREST activating deacetylase activity for H3 K9ac, but not H3 K23ac. Finally, the H3 K14ac substrate showed the most diverse activation across the constructs.

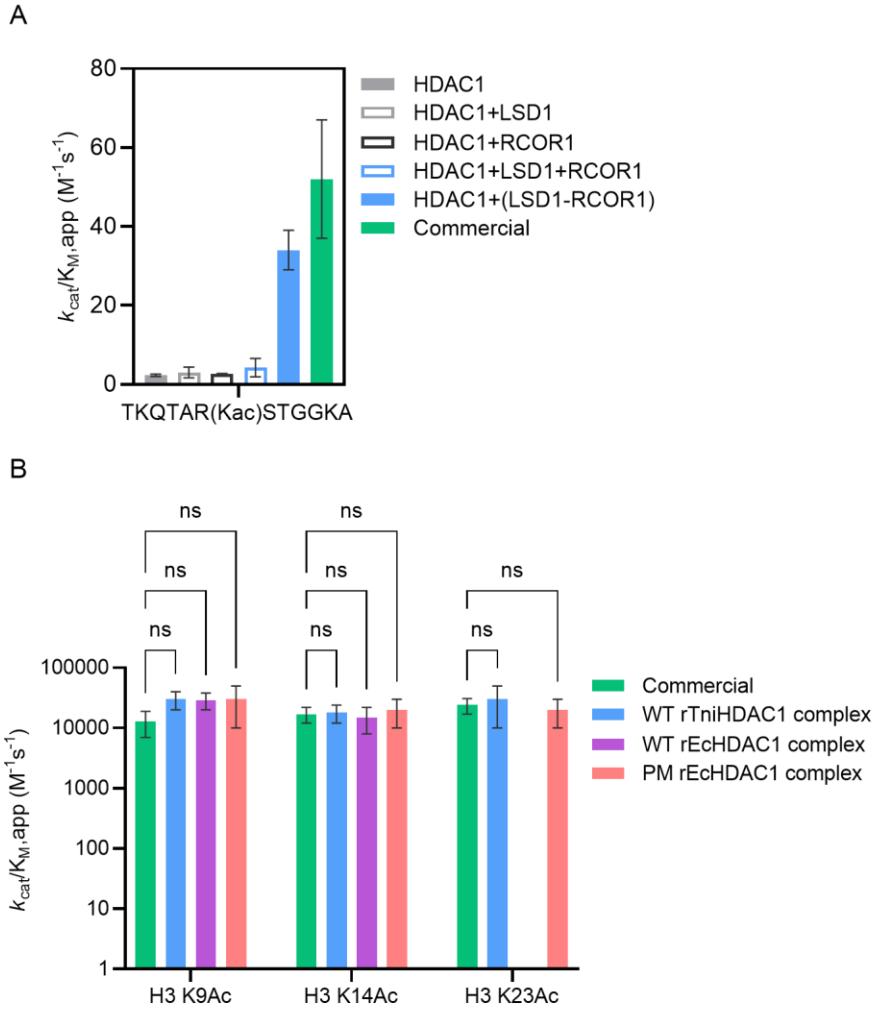


Figure 3.6: Comparison of catalytic efficiencies of HDAC1 constructs in catalyzing acetylated Histone H3 substrates in the presence of protein interactors

A) Comparison of the catalytic efficiency of WT rTniHDAC1 upon addition of LSD1 and CoREST, both individually (gray and black outline, respectively) and combined (blue outline: LSD1 and CoREST purified individually, blue: LSD1-CoREST co-purified in complex), and commercially available HDAC1 (green). Each bar is calculated from an initial velocity using stopped timepoints (n=4) at $[S] = 200\mu M$, error bars represent the standard error calculated by GraphPad Prism analysis. B) Catalytic efficiencies calculated for HDAC1 incubated with the LSD1-CoREST subcomplex. Exact values are given in **Table 3.2**. Bars represent the average value (N=2-4) with error bars representing S.D. Denoted p-values (ns = not significant) are calculated for the comparison to the value of the commercially available HDAC1 (green) by t-test with Sidak's correction for multiple comparisons[54]. *PM: phosphomimic, S421E/S423E.

Table 3.3: Activation of deacetylase activity upon addition of LSD1-CoREST

	Ratio of Activity Activation ^a		
	H3K9Ac	H3K14Ac	H3K23Ac
Insect WT	8±3	60±20	200±100
<i>E. coli</i> WT	>500	110±60	n/a
<i>E. coli</i> S421E/S423E	40±30	>400	100±70

^aCalculated using values from **Table 3.2**. The error of the ratios was calculated using propagation of uncertainty[55].

Discussion

We have performed kinetic characterization of recombinant HDAC1 with and without the known interacting proteins LSD1 and CoREST. To compare HDAC1 activity when alone versus present in complex, we aimed to reconstitute the CoREST complex *in vitro*. Using LSD1 and CoREST expressed separately and as a complex, we saw *in vitro* interaction with all HDAC1 constructs. The *in vitro* HDAC1-LSD1-CoREST complex displayed significantly enhanced deacetylation efficiency. The measured deacetylase activity was then comparable to commercially available HDAC1 indicating additional co-purifying proteins present are important for enhancing HDAC1 deacetylase efficiency.

HDAC1 substrate selectivity is dependent on post-translational modifications

The analyzed substrates included both histone peptide analogs, and histone proteins selected due to being known substrates of HDAC1. Initially, HDAC1 expressed and purified out of the insect cells was the most active and was the only construct able to catalyze deacetylation of acetylated-peptide substrates. The WT rTniHDAC1 could have

additional post-translational modifications present that impact deacetylation efficiency, and we demonstrated that recombinant WT rTniHDAC1 is phosphorylated (**Figure 3.2**). We further explored the impact of phosphorylation on deacetylase activity by creating a phosphorylation mimic in *E. coli* using the negatively charged glutamate substitute for phosphorylated serine residues. The *E. coli* phosphorylation mimic still did not demonstrate significant deacetylase activity towards peptide substrates. Yet, we were able to detect deacetylase activity of all constructs towards histone protein substrates. The insect WT HDAC1 deacetylated all histone acetylation sites. The phosphorylation mimic HDAC1 expressed in *E. coli* demonstrated broader substrate selectivity by deacetylating two acetylation sites (H3 K9ac and H3 K23ac) compared to wild type HDAC1 expressed in *E. coli* that catalyzed deacetylation of a different acetylation site (H3 K14ac) (**Table 3.2**). When compared to the HDAC1 expressed in insect cells, the proteins expressed in *E. coli* each had similar activity (~2-fold) for one of the acetylation sites (H3 K23ac for the PM and H3 K14ac for the WT) yet were much less active for other acetylation sites (~4-60 fold).

HDAC1 substrate selectivity is dependent on complex formation

When alone, WT rTniHDAC1 and PM rEchHDAC1 displayed similar substrate selectivity with greatest activity towards H3 K9ac and less activity towards H3 K14ac and H3 K23ac (**Figure 3.7-A**). Upon addition of LSD1-CoREST, both constructs then displayed similar activity for all H3 acetylation sites (**Figure 3.7-B**). Interestingly, WT rEchHDAC1 in the presence of LSD1-CoREST had similar activity to WT rTniHDAC1 towards H3 K9ac and H3 K14ac, but still had no detectable deacetylase activity towards H3 K23ac (**Figure 3.7-B**). Similar activity for H3 K23ac for WT rTniHDAC1 and PM

rEcHDAC1 seems to indicate the K23ac site is specific to phosphorylated HDAC1. The overall similarity in substrate selectivity for the phosphomimic expressed in *E. coli* and the phosphorylated HDAC1 expressed in insect cells is encouraging, as it suggests the phosphomimic can be successfully used to probe HDAC1 activity and selectivity.

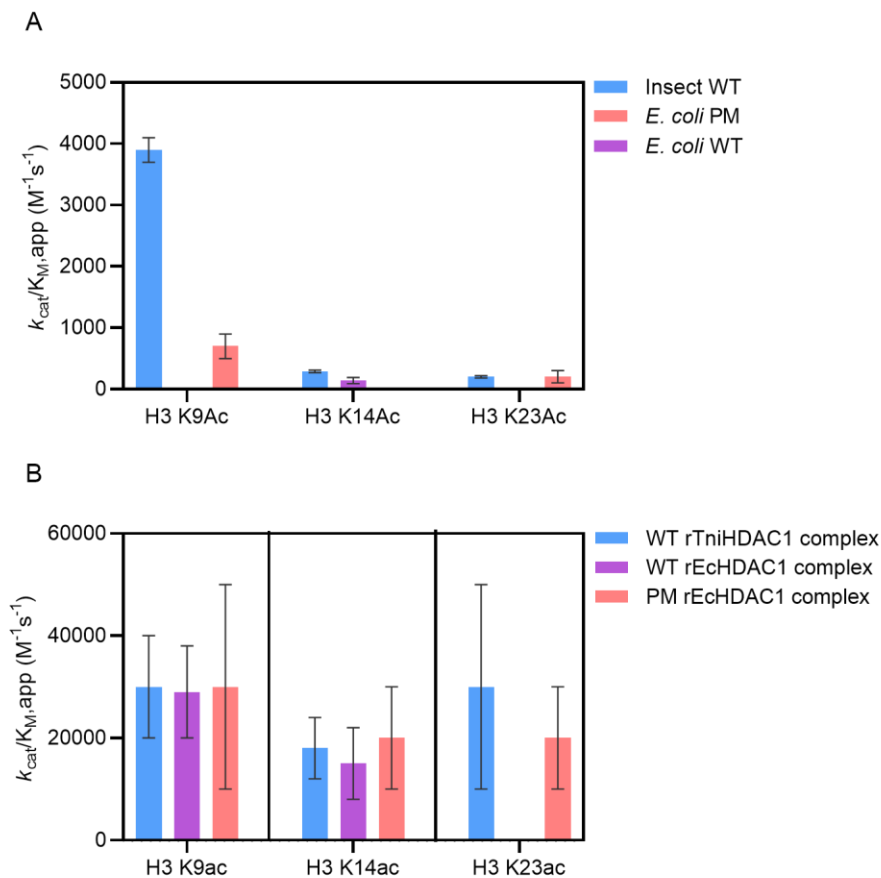


Figure 3.7: Substrate selectivity of HDAC1 constructs. A) Deacetylase activity of HDAC1 alone against acetylated histone H3. B) Deacetylase activity of HDAC1 in the presence of LSD1-CoREST against acetylated histone H3.

These results suggest that HDAC1 depends on the primary amino acid sequence for recognition of substrates and catalytic activity. Histone H3 in this study was alone in solution, which does not occur biologically as Histone H3 is found in more structured tetramers with Histone H4[56]. Further studies of deacetylation of the histone H3/H4

tetramer, the histone H3/H4/H2A/H2B octamer, and the nucleosome (octamer with DNA) would be useful to elaborate on the observed activation differences. The nucleosome is especially a key interest as there have been conflicting studies of which component of the complex interacts with the nucleosome allowing deacetylation, LSD1[34] or CoREST[57].

Interactions significantly enhance HDAC1 deacetylase activity

This work concretely shows that the presence of the LSD1-CoREST subcomplex significantly increases HDAC1 deacetylase activity and, possibly, alters selectivity. An important distinction is that both LSD1 and CoREST are necessary to enhance HDAC1 deacetylase activity; this activation was achieved by co-purifying LSD1 and CoREST as a complex from recombinant expression in *E. coli*. CoREST consists of two distinct SANT domains, SANT1 on the N-terminus and SANT2 on the C-terminus. The ELM2/SANT1 domain of CoREST has been established as necessary for HDAC1-CoREST interaction[30]. And the SANT2/Linker domain of CoREST has an established interaction with the tower domain of LSD1 as characterized by X-ray crystallography[58]. Recent structural studies indicate that the full CoREST complex forms via the LSD1 tower and CoREST linker domains, and CoREST has an additional contact with HDAC1 via its ELM2/SANT1 domain[34]. Our data indicate that the tower domain of LSD1 interacts directly and independently with phosphorylated HDAC1, as shown by the co-immunoprecipitation of HDAC1 and LSD1 (**Figure 3.8**). However, the formation of the CoREST-LSD1 subcomplex is required for the interaction of CoREST with HDAC1 and to enhance HDAC1 deacetylation (**Figure 3.8**). This is shown by the absence of co-immunoprecipitation of HDAC1 and CoREST and the lack of activity enhancement of HDAC1 upon addition of CoREST. Interestingly, the interaction of the LSD1-CoREST

subcomplex is not dependent on HDAC1 phosphorylation. We are unable to determine if these interactions differences are dependent on different binding affinities of phosphorylated and unphosphorylated HDAC1 to LSD1 versus LSD1-CoREST. Crosslinking studies indicate the region of HDAC1 containing the phosphorylation sites interacts with both LSD1 and CoREST [34]. The additional interaction points enabled by CoREST could be enough to overcome weaker interaction points with LSD1. Further exploration into the binding affinities and formation constants for the different HDAC1 constructs would help elucidate the importance of HDAC1 phosphorylation *in vivo*.

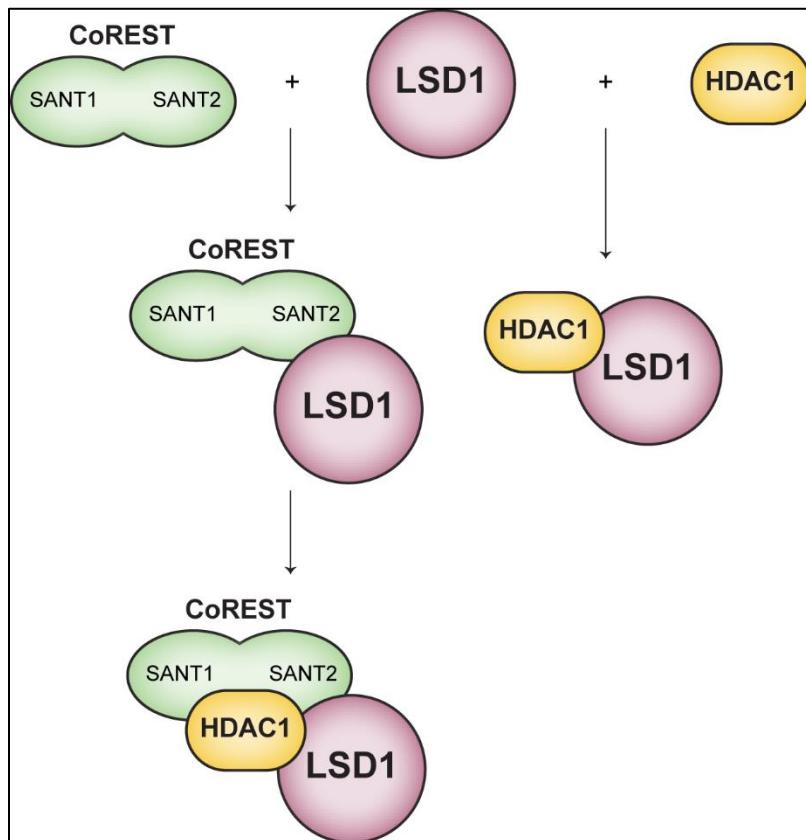


Figure 3.8: Proposed formations of HDAC1, LSD1, and CoREST containing complexes

Successful *in vitro* reconstitution of a HDAC1-containing complex will allow further exploration of how these complexes are assembled. The CoREST complex is known to

be involved in gene transcription regulation due to its interactions with histones, and the core components are known to be important for cell differentiation and cell cycle regulation, making these proteins valuable cancer therapeutic targets[59]. The core components of the complex are an HDAC (either HDAC1 or HDAC2), LSD1, and CoREST[59]. The core CoREST complex explored here could be further expanded to include other known protein interactors, such as HDAC2. The addition of other interactions may further alter the activity and selectivity of the CoREST complex. It is established histone tails are known substrates for HDAC1-containing complexes and it is worthwhile to continue establishing the activity and selectivity for histone substrates, however the NuRD HDAC1 complex has been found to deacetylate non-histone proteins, such as p53[13]. The exploration of how complexes affect selectivity towards non-histone substrates is valuable to further understand HDAC1 involvement in biological processes. Furthermore, the use of bacterially expressed enzyme allows analysis of unmodified enzyme in addition to incorporating specific post-translational modifications, which is harder to control in eukaryotic expression systems. Understanding the formation of these complexes and how they are regulated in their cellular environment will ultimately increase our understanding of how HDACs are implicated in cancer and other diseases.

References

1. Phillips, D.M., *The presence of acetyl groups of histones*. Biochem J, 1963. **87**: p. 258-63.
2. Fischer, E.H., et al., *Structure of the site phosphorylated in the phosphorylase b to a reaction*. J Biol Chem, 1959. **234**(7): p. 1698-704.
3. Hornbeck, P.V., et al., *PhosphoSitePlus, 2014: mutations, PTMs and recalibrations*. Nucleic Acids Res, 2015. **43**(Database issue): p. D512-20.
4. Allfrey, V.G., R. Faulkner, and A.E. Mirsky, *ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS*. Proc Natl Acad Sci U S A, 1964. **51**: p. 786-94.
5. Hu, E., et al., *Cloning and Characterization of a Novel Human Class I Histone Deacetylase That Functions as a Transcription Repressor*. Journal of Biological Chemistry, 2000. **275**(20): p. 15254-15264.
6. Gao, L., et al., *Cloning and Functional Characterization of HDAC11, a Novel Member of the Human Histone Deacetylase Family*. Journal of Biological Chemistry, 2002. **277**(28): p. 25748-25755.
7. Verdel, A. and S. Khochbin, *Identification of a New Family of Higher Eukaryotic Histone Deacetylases: COORDINATE EXPRESSION OF DIFFERENTIATION-DEPENDENT CHROMATIN MODIFIERS*. Journal of Biological Chemistry, 1999. **274**(4): p. 2440-2445.
8. Tong, J.J., et al., *Identification of HDAC10, a novel class II human histone deacetylase containing a leucine-rich domain*. Nucleic acids research, 2002. **30**(5): p. 1114-1123.
9. Fischle, W., et al., *A New Family of Human Histone Deacetylases Related to Saccharomyces cerevisiae HDA1p*. Journal of Biological Chemistry, 1999. **274**(17): p. 11713-11720.
10. Grozinger, C.M., C.A. Hassig, and S.L. Schreiber, *Three proteins define a class of human histone deacetylases related to yeast Hda1p*. Proceedings of the National Academy of Sciences, 1999. **96**(9): p. 4868-4873.
11. Yang, W.-M., et al., *Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3*. Proceedings of the National Academy of Sciences, 1996. **93**(23): p. 12845-12850.
12. Luo, J., et al., *Acetylation of p53 augments its site-specific DNA binding both in vitro and in vivo*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**: p. 2259-64.

13. Luo, J., et al., *Deacetylation of p53 modulates its effect on cell growth and apoptosis*. Nature, 2000. **408**: p. 377-81.
14. Nalawansa, D.A., et al., *HDAC Inhibitor-Induced Mitotic Arrest Is Mediated by Eg5/KIF11 Acetylation*. Cell Chem Biol, 2017. **24**(4): p. 481-492 e5.
15. Ferreira, R., et al., *Cell cycle-dependent recruitment of HDAC-1 correlates with deacetylation of histone H4 on an RB-E2F target promoter*. EMBO reports, 2001. **2**: p. 794-9.
16. Deardorff, M.A., et al., *HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle*. Nature, 2012. **489**(7415): p. 313-7.
17. Yin, F., et al., *LSD1 regulates pluripotency of embryonic stem/carcinoma cells through histone deacetylase 1-mediated deacetylation of histone H4 at lysine 16*. Mol Cell Biol, 2014. **34**(2): p. 158-79.
18. Du, Z., et al., *DNMT1 Stability Is Regulated by Proteins Coordinating Deubiquitination and Acetylation-Driven Ubiquitination*. Science signaling, 2010. **3**: p. ra80.
19. Zhou, Y., R. Santoro, and I. Grummt, *The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription*. EMBO J, 2002. **21**(17): p. 4632-40.
20. Chen, S.H., et al., *HDAC1,2 Knock-Out and HDACi Induced Cell Apoptosis in Imatinib-Resistant K562 Cells*. Int J Mol Sci, 2019. **20**(9).
21. Chen, L.-f., et al., *Duration of Nuclear NF- κ B Action Regulated by Reversible Acetylation*. Science, 2001. **293**(5535): p. 1653-1657.
22. Li, Y. and E. Seto, *HDACs and HDAC Inhibitors in Cancer Development and Therapy*. Cold Spring Harb Perspect Med, 2016. **6**(10).
23. Majdzadeh, N., et al., *HDAC4 inhibits cell-cycle progression and protects neurons from cell death*. Dev Neurobiol, 2008. **68**(8): p. 1076-92.
24. Graff, J., et al., *An epigenetic blockade of cognitive functions in the neurodegenerating brain*. Nature, 2012. **483**(7388): p. 222-6.
25. Montgomery, R.L., et al., *Histone deacetylases 1 and 2 control the progression of neural precursors to neurons during brain development*. Proc Natl Acad Sci U S A, 2009. **106**(19): p. 7876-81.
26. Shakespear, M.R., et al., *Histone deacetylases as regulators of inflammation and immunity*. Trends Immunol, 2011. **32**(7): p. 335-43.

27. Oughtred, R., et al., *The BioGRID interaction database: 2019 update*. Nucleic Acids Res, 2019. **47**(D1): p. D529-d541.
28. Xue, Y., et al., *NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities*. Mol Cell, 1998. **2**(6): p. 851-61.
29. Fleischer, T.C., U.J. Yun, and D.E. Ayer, *Identification and characterization of three new components of the mSin3A corepressor complex*. Mol Cell Biol, 2003. **23**(10): p. 3456-67.
30. You, A., et al., *CoREST is an integral component of the CoREST- human histone deacetylase complex*. Proceedings of the National Academy of Sciences, 2001. **98**(4): p. 1454-1458.
31. Johnson, C.A., et al., *Human class I histone deacetylase complexes show enhanced catalytic activity in the presence of ATP and co-immunoprecipitate with the ATP-dependent chaperone protein Hsp70*. J Biol Chem, 2002. **277**(11): p. 9590-7.
32. Seto, E. and M. Yoshida, *Erasers of histone acetylation: the histone deacetylase enzymes*. Cold Spring Harbor perspectives in biology, 2014. **6**(4): p. a018713-a018713.
33. Kalin, J.H., et al., *Targeting the CoREST complex with dual histone deacetylase and demethylase inhibitors*. Nat Commun, 2018. **9**(1): p. 53.
34. Song, Y., et al., *Mechanism of Crosstalk between the LSD1 Demethylase and HDAC1 Deacetylase in the CoREST Complex*. Cell reports, 2020. **30**(8): p. 2699-2711.e8.
35. Wu, M., et al., *Lysine-14 acetylation of histone H3 in chromatin confers resistance to the deacetylase and demethylase activities of an epigenetic silencing complex*. Elife, 2018. **7**.
36. Wang, Z.A., et al., *Diverse nucleosome Site-Selectivity among histone deacetylase complexes*. eLife, 2020. **9**: p. e57663.
37. Du, Y., et al., *aPKC phosphorylation of HDAC6 results in increased deacetylation activity*. PLoS One, 2015. **10**(4): p. e0123191.
38. Dressel, U., et al., *A dynamic role for HDAC7 in MEF2-mediated muscle differentiation*. J Biol Chem, 2001. **276**(20): p. 17007-13.
39. Williams, K.A., et al., *Extracellular signal-regulated kinase (ERK) phosphorylates histone deacetylase 6 (HDAC6) at serine 1035 to stimulate cell migration*. J Biol Chem, 2013. **288**(46): p. 33156-70.

40. Pflum, M.K., et al., *Histone deacetylase 1 phosphorylation promotes enzymatic activity and complex formation*. J Biol Chem, 2001. **276**(50): p. 47733-41.
41. Zhang, X., et al., *Histone deacetylase 3 (HDAC3) activity is regulated by interaction with protein serine/threonine phosphatase 4*. Genes Dev, 2005. **19**(7): p. 827-39.
42. Cai, R., et al., *Mammalian histone deacetylase 1 protein is posttranslationally modified by phosphorylation*. Biochem Biophys Res Commun, 2001. **283**(2): p. 445-53.
43. Lee, H., N. Rezai-Zadeh, and E. Seto, *Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A*. Mol Cell Biol, 2004. **24**(2): p. 765-73.
44. Welker Leng, K.R., et al., *Phosphorylation of Histone Deacetylase 8: Structural and Mechanistic Analysis of the Phosphomimetic S39E Mutant*. Biochemistry, 2019. **58**(45): p. 4480-4493.
45. Watabe, M. and T. Nakaki, *Protein kinase CK2 regulates the formation and clearance of aggresomes in response to stress*. Journal of Cell Science, 2011. **124**(9): p. 1519-1532.
46. Parra, M., et al., *Protein kinase D1 phosphorylates HDAC7 and induces its nuclear export after T-cell receptor activation*. J Biol Chem, 2005. **280**(14): p. 13762-70.
47. Tsai, S.C. and E. Seto, *Regulation of histone deacetylase 2 by protein kinase CK2*. J Biol Chem, 2002. **277**(35): p. 31826-33.
48. Grozinger, C.M. and S.L. Schreiber, *Regulation of histone deacetylase 4 and 5 and transcriptional activity by 14-3-3-dependent cellular localization*. Proceedings of the National Academy of Sciences, 2000. **97**(14): p. 7835-7840.
49. Wang, A., et al., *Regulation of Histone Deacetylase 4 by Binding of 14-3-3 Proteins*. Molecular and cellular biology, 2000. **20**: p. 6904-12.
50. Brown, W.C., et al., *New ligation-independent cloning vectors compatible with a high-throughput platform for parallel construct expression evaluation using baculovirus-infected insect cells*. Protein Expr Purif, 2011. **77**(1): p. 34-45.
51. Castaneda, C.A., et al., *HDAC8 substrate selectivity is determined by long- and short-range interactions leading to enhanced reactivity for full-length histone substrates compared with peptides*. J Biol Chem, 2017. **292**(52): p. 21568-21577.
52. Reger, A.S., J.M. Carney, and A.M. Gulick, *Biochemical and Crystallographic Analysis of Substrate Binding and Conformational Changes in Acetyl-CoA Synthetase*. Biochemistry, 2007. **46**(22): p. 6536-6546.

53. Wolfson, N.A., et al., *An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity*. *Anal Biochem*, 2014. **456**: p. 61-9.
54. Abdi, H., *The Bonferroni and Sidak Corrections for Multiple Comparisons*, in *Encyclopedia of Measurement and Statistics*, N. Salkind, Editor. 2007, Sage: Thousand Oaks, CA. p. 103-107.
55. Ku, H.H., *Notes on the Use of Propagation of Error Formulas*. *Journal of Research of the National Bureau of Standards - C. Engineering and Instrumentation*, 1966. **70C**(4): p. 263-273.
56. Arents, G., et al., *The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix*. *Proceedings of the National Academy of Sciences*, 1991. **88**(22): p. 10148-10152.
57. Lee, M.G., et al., *Functional Interplay between Histone Demethylase and Deacetylase Enzymes*. *Molecular and Cellular Biology*, 2006. **26**(17): p. 6395-6402.
58. Yang, M., et al., *Structural Basis for CoREST-Dependent Demethylation of Nucleosomes by the Human LSD1 Histone Demethylase*. *Molecular Cell*, 2006. **23**(3): p. 377-387.
59. Kelly, R.D. and S.M. Cowley, *The physiological roles of histone deacetylase (HDAC) 1 and 2: complex co-stars with multiple leading parts*. *Biochem Soc Trans*, 2013. **41**(3): p. 741-9.

Chapter 4 : Structure-based Prediction of Substrate Selectivity of HDAC6[‡]

Introduction

For cells and organisms to survive and adapt to different conditions, complex and fine-tuned, context-dependent regulation is crucial. Much of this regulation is achieved by post-translational modifications (PTM) that can change the behavior of a protein. One of the major regulatory modifications is acetylation and subsequently deacetylation. Histone deacetylases, HDACs, catalyze the removal of an acetyl group from the post-translational modification of acetyl-lysine in proteins. They are divided into two major groups, zinc-dependent HDACs and NAD-dependent HDACs called sirtuins. Zinc-dependent HDACs are further divided into subclasses based on their homology to yeast enzymes[1].

Histone deacetylase 6 (HDAC6) is a class IIB deacetylase and is the only HDAC to contain two deacetylase domains with distinct specificities. The first domain specifically catalyzes deacetylation of acetylated C-terminal lysine residues[2], while the second domain shows broad substrate selectivity[3]. HDAC6 is primarily localized to the cytoplasm and HDAC6 deacetylase activity has been linked to several cytoplasmic

[‡]Reproduced, in part, from a manuscript in preparation: [Structure-based Assessment of KDAC6 Specificity](#). Varga, J., Diffley, K., Welker-Leng, K.R., Alam, N., Fierke, C.A., Schueler-Furman, O. Kelsey Diffley and Katherine R. Welker-Leng performed the initial zHDAC6 CD2 deacetylase activity assays for training the model. Julia Varga performed all computational modeling and optimization. Kelsey Diffley performed the additional deacetylase activity assays for acetylome prediction and zHDAC6 CD12. Julia Varga, Kelsey Diffley, Katherine R. Welker-Leng, Carol A. Fierke, and Ora Schueler-Furman analyzed the data. Julia Varga, Kelsey Diffley, Carol A. Fierke, and Ora Schueler-Furman wrote the manuscript. The text was revised by Kelsey Diffley.

cellular processes: microtubule stability and cell motility through deacetylation of α -tubulin[4]; lysosome and autophagosome fusion by cortactin deacetylation[5]; protein folding by regulation of Hsp90 activity[6, 7]; innate immunity via deacetylation of retinoic acid inducible gene-I protein; and aggresome formation using HDAC6's ubiquitin binding domain[8]. While the broad specificity of HDAC6 has been reported, an understanding of the selectivity determinants is still lacking, as is a detailed understanding of the underlying structural basis that makes this particular HDAC more promiscuous than others, such as HDAC8[9].

Many of the enzymes that install or remove PTMs act on short linear motifs (SLIMs) that are often exposed. Therefore, their substrate selectivity may be approximated by short peptides that mimic the region[10]. Different types of prediction methods for finding putative substrates have been developed. Many sequence-based predictions find modification sites based on position-specific scoring matrices (PSSMs)[11], that are derived from a large set of substrates. However, these do not account for possible interdependencies between amino acids at different positions in the substrate, nor do they include secondary structure elements that might be important for recognition. Machine learning-based approaches can be used for these aims (e.g. HMMs[12] and naive Bayes[13]), but such approaches depend on considerable amounts of data.

Structure-based methods can complement sequence-based methods, in cases of non-canonical motifs, as we have previously shown in the case of the PTM enzymes protein farnesyltransferase (FTase)[14] and HDAC8[9]. This approach assumes that the local peptide sequence of the substrate is a main determinant of selectivity. Our framework to assess substrate selectivity, Rosetta FlexPepBind[15], consists of

calibrating a protocol by generating structural models of the interaction of candidate substrates with the enzyme[16] and using this model to evaluate their respective binding affinity. The underlying assumption is that by properly defining and reinforcing the catalytically competent substrate binding conformation, the estimated binding ability can be taken as a proxy for substrate reactivity. Accuracy of the calibrated protocol can be estimated by applying it to an independent test set and then using it on candidate peptides with unknown activity to identify new substrates.

In this study we utilized an accurate biochemical assay that measures acetate production following deacetylation[17] to quantify the catalytic activity of HDAC6 for specific peptides and to establish a gold standard set of peptide substrates. Based on these activities, we calibrated the structure-based approach FlexPepBind to evaluate activity of potential substrates. Calibration revealed important structural differences between HDAC6 and HDAC8 that form the basis of the considerable difference in selectivity of these two deacetylases. Finally, application of this method to screen the acetylome identified novel potential regulatory mechanisms based on HDAC6-dependent regulation.

Materials and Methods

Reagents

High flow amylose resin was purchased from New England Biolabs and Ni-NTA agarose was purchased from Qiagen. Adenosine triphosphate (ATP), coenzyme A (CoA), NAD⁺, NADH, L-malic acid, malate dehydrogenase (MDH), citrate synthase (CS), and mouse monoclonal anti-polyhistidine-alkaline phosphatase antibody were purchased

from Sigma. Boc-Lys-AMC was purchased from Bachem. N-terminally acetylated and C-terminally carboxylated, singly acetylated peptides were purchased from Peptide 2.0 or Synthetic Biomolecules. 3% (v/v) acetic acid standard was purchased from RICCA Chemical. All other materials were purchased from Fisher at >95% purity unless noted otherwise.

HDAC6 Expression and Purification

The plasmids and protocols for the expression and purification of zebrafish HDAC6 catalytic domain 2 (zCD2, residues 440-798) and both catalytic domains (zCD12, residues 60-798) were generously provided by David Christianson (University of Pennsylvania). HDAC6 zCD2 was expressed and purified as described previously[18]. HDAC6 zCD12 was expressed and purified as previously described with several alterations for expression optimization[3]. Briefly, zCD12 was expressed in *E. coli* BL21(DE3) cells grown in 2xYT in the presence of 50 mg/L kanamycin. Cells were grown at 37°C until the OD₆₀₀=1.0, then cooled for 1 h to 16°C. The media was supplemented with 200 μM zinc sulfate 30 min before induction with 75 μM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested after 18 h and purified according to the zCD2 protocol[18]. The concentration of zCD12 was confirmed by an active site titration with tubastatin A using the Fluor de Lys assay (see below). Various concentrations of tubastatin A (0.025-2 μM) were added to 50 μM Boc-Lys-AMC peptide and the reaction was initiated by addition of ~0.1 μM HDAC6 (as determined by A₂₈₀, ext. coeff. = 85,260 M⁻¹cm⁻¹) at 30°C. A control reaction was also performed without tubastatin A. The rate of each inhibitor reaction was compared to the control reaction to obtain the quantity of remaining activity (%). Using GraphPad prism, the % remaining activity was

plotted versus inhibitor concentration to determine the stoichiometric point of saturation where [inhibitor]=[enzyme].

Coupled Acetate-Detection Assay

Acetyl CoA Synthetase (ACS) was expressed and purified as previously described[17, 19].

The coupled acetate-detection assay or simply the 'acetate assay' was performed as previously described with a few modifications[17]. Briefly, lyophilized peptides were re-suspended in water when possible or with minimal quantities of acid, base, or organic solvent required to improve solubility. Peptide concentration was determined by one or more of the following methods: measuring A_{280} using the amino acid extinction coefficients if the peptide contained a tryptophan or tyrosine; using the fluorescamine assay if the peptide contained a free lysine[20]; performing the bicinchoninic (BCA) assay using bovine serum albumin (BSA) as a standard; or determining the concentration of acetate produced by complete deacetylation by HDAC6.

Reactions containing 10-2000 μ M singly acetylated peptides in 1X HDAC6 assay buffer (50 mM HEPES, pH 8.0, 137 mM NaCl, 2.7 mM KCl, 1 mM $MgCl_2$) were initiated by addition of 0.1-1 μ M HDAC6 at 30°C. At various time points a 60 μ L aliquots were quenched with 5 μ L of 10% hydrochloric acid and kept on ice until assay completion (no more than 90 minutes). Timepoints were flash frozen with liquid nitrogen and stored at -80°C until work-up.

Coupled solution (50 mM HEPES, pH 8, 400 μ M ATP, 10 μ M NAD^+ , 30 μ M CoA, 0.07 U/ μ L CS, 0.04 U/ μ L MDH, 50 μ M ACS, 100 mM NaCl, 3 mM KCl, 50 mM $MgCl_2$, and 2.5 mM L-malic acid) was prepared the day of the assay work-up and incubated at room

temperature away from light for at least 25 minutes. Assay aliquots were quickly thawed and neutralized with 15 μL of freshly prepared and filtered 6% sodium bicarbonate. Neutralized timepoints or controls (acetate or NADH standards), 60 μL , were added to 10 μL coupled solution (or 1X assay buffer for NADH standards) in a black, flat-bottomed, half-area, non-binding, 96-well plate (Corning No. 3686). The resulting NADH fluorescence (Ex 340 nm, Em 460 nm) of standards and timepoints was read on a PolarStar fluorescence plate reader at 1-3-minute increments until the signal reached equilibrium. The standard curve for the acetate controls was compared to the NADH standards to verify assay function. When possible, a positive control reaction for enzyme activity was included. Using the acetate standard curve, the fluorescence of each timepoint was converted to μM product, and the dependence of the initial rate of the reaction (<10%) on the substrate concentration was measured. Using GraphPad Prism, the Michaelis-Menten equation (**Equation 4.1**) was fit to the resulting dependence of the initial velocity on substrate concentration to determine the kinetic parameters k_{cat}/K_M , k_{cat} , and K_M . Standard error was calculated using GraphPad Prism analysis. 25 peptides were used in the training set and another 16 peptides were used for validation of the protocol.

$$\frac{v_o}{[E]} = \frac{k_{\text{cat}}[S]}{(K_M + [S])} \quad \text{Equation 4.1}$$

Fluor de Lys Assay

Reactions containing 25-200 μM Boc-Lys-AMC peptide in 1X HDAC6 buffer were initiated with 0.01-0.1 μM HDAC6. Various aliquoted timepoints (5 μL) were quenched in individual wells of a Corning 3694 96-well half area microplate containing 45 μL quench solution (0.73% Developer II, 1.2 μM trichostatin A, 1X HDAC6 buffer). After a 10-minute

room temperature incubation, the fluorescence of the remaining substrate (Ex 360 nm, Em 362 nm) and the resulting product (Ex 360 nm, Em 460 nm) was determined using a PolarStar plate reader. Standard curves were made for each substrate concentration to reflect the change in product/substrate fluorescence ratio. Using these standard curves, each reaction timepoint was converted to μM product, and initial rate of the reaction (<10%) was measured at various substrate concentrations. Using GraphPad Prism, the Michaelis-Menten equation (**Equation 4.1**) was fit to the resulting dependence of the initial velocity on substrate concentration to determine the kinetic parameter $k_{\text{cat}}/K_{\text{M}}$.

Calibration of FlexPepBind

Running FlexPepBind requires the creation of a starting structure to generate a template (or a set of templates, as in the present study) for threading peptides. We used the structure of HDAC6 catalytic domain 2 from PDB (Protein Data Bank[21]) (CD2; PDB ID 5EFN) which was crystallized in a complex with a coumarin-linked trimer peptide substrate. In every described Rosetta protocol, we used Rosetta v2020.28. The protocol implemented in this study is like the one used in our previous study on HDAC8 specificity[9], and in the following we mainly highlight the differences.

To enforce a catalytic-competent binding conformation we defined constraints that characterize substrate binding as given by the solved structures of HDAC6 bound to ligands. Constraints were defined for Rosetta runs as done with HDAC8[9]. These include (1) interactions coordinating the proper binding of the Zn^{2+} ion required for enzymatic activity, (2) interactions between the acetylated lysine side chain and the binding pocket, and (3) a dihedral angle constraint in the peptide between residues 3 and 4 (*i.e.*, adjacent

to the acetylated lysine in the modeled hexamers) to enforce a cis-peptide bond. All the distances between interacting residues were measured on structure PDB ID 5EFN.

The best substrate peptide (sequence: *EGK_{Ac}FVR*) was built in the binding pocket using the corresponding atoms of trichostatin A. The rest of the peptide was added in an extended conformation. FlexPepDock was run on this structure with the constraints added, generating *nstruct*=1000 decoys with different setups. In contrast to the HDAC8 study, here we selected not only the top-scoring structure, but rather the top 5 structures, according to the measures *l_sc* or *reweighted_sc*, depending on the protocol (Protocols Pxa and Pxb, respectively). Every peptide of the training dataset was threaded onto these starting structures using the Rosetta *fixbb* protocol and running FlexPepDock with minimization only. For each peptide sequence, the best score (*reweighted_sc*) among the 5 templates was used to reflect its substrate strength (final scoring according to *l_sc* resulted in inferior predictions and was therefore not followed up). For protocols P1x & P2x, the receptor backbone was not moved, or only minimized, respectively. Since these protocols did not yield satisfactory predictions, for protocols P3x and P4x, the template was optimized using loop modeling prior to the peptide docking step. Loop modeling was performed with kinematic closure with fragments protocol [22] on the loop of residues 455-467 (**Figure 4.2**). Fragments were generated as described in the fragment picker application manual and 1000 decoys were generated with the loop modeling protocol. The top (top2) scoring decoys were selected for protocol P3 (P4) to accommodate different possible loop and peptide conformations.

Running the calibrated protocol on the acetylome

The dataset of the human acetylome was extracted from the PhosphoSitePlus database[23] (downloaded on 23/09/2020). Hexamer peptides around acetylated sites on human proteins with at least one low-throughput experiment to support were derived with two leading and three trailing residues around the modification site. Only peptides spanning a full hexamer (*i.e.*, the modification is not at the termini) were selected. Interaction data for HDAC6 was downloaded from BioGRID[24] on 11/09/2020 from database version 4.1.190. For pathway analysis, we used Reactome with Pathway Browser version 3.7 and database release 74.

Results

Measurement of HDAC6 zCD2 substrate selectivity

We used an enzyme-coupled acetate-detection assay, or simply the ‘acetate assay’ (see Methods), developed and used previously[9, 17], to measure the catalytic efficiency (k_{cat}/K_M) of HDAC6 catalyzing deacetylation of acetylated peptides. These peptides included sequences from known HDAC6 substrate proteins, as well as a set of selected peptides with reported acetylation sites that were used in previous studies[9, 25, 26]. We analyzed the substrate selectivity of the second deacetylase domain (CD2) of *Danio rerio* HDAC6 for these experiments which is more stable and has been shown to be a valid substitute for human HDAC6[3]. To ensure an accurate determination of k_{cat}/K_M we measured HDAC6-catalyzed deacetylation at a minimum of four peptide concentrations, with at least two concentrations below K_M (**Figure 4.1**, squares). A total of 25 peptides which met these criteria were used in the training set (**Table 4.1**).

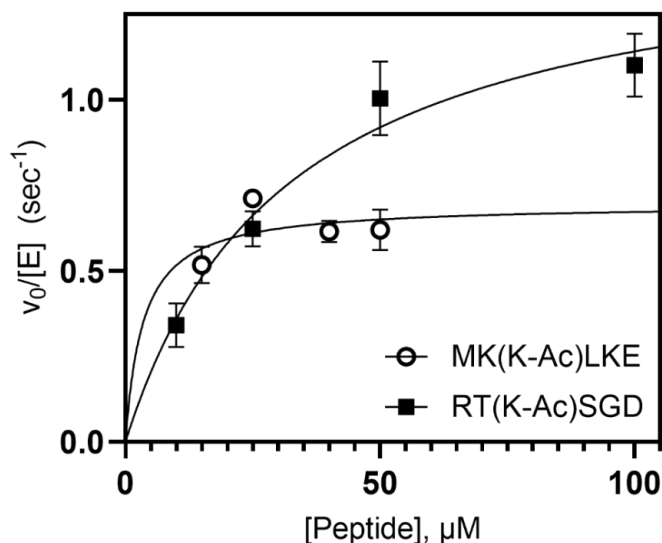


Figure 4.1: Dependence of deacetylation rate on substrate concentration for two representative peptides catalyzed by HDAC6.

The initial velocity for each substrate concentration was determined by taking a linear regression of a time course consisting of a minimum of three timepoints, error bars reflect the standard error. The kinetic parameters are determined from a nonlinear least square fit of the Michaelis-Menten equation to the data and are listed in Table 1. Black squares: example of data which met the criteria to produce accurate k_{cat}/K_M values. Open circles: example of overfit data resulting in calculation of a lower limit for k_{cat}/K_M , due to the K_M being lower than the detection limit of the assay.

Additionally, we included 15 peptides (set D-EXT, **Table 4.2** above line) where the value of k_{cat} was measured accurately but the K_M value was lower than the limit of detection for the acetate assay ($\sim 10\text{-}20 \mu\text{M}$ substrate, **Figure 4.1**) allowing determination of only a lower limit for the value of k_{cat}/K_M . The measured k_{cat}/K_M values span the range of three orders of magnitude. We defined a cutoff of $k_{\text{cat}}/K_M = 10^4 \text{ M}^{-1}\text{s}^{-1}$ to distinguish 7 non-substrates from the rest. Two additional peptides were tested but were not included in further analysis (**Table 4.2**, below line).

Table 4.1: Kinetic parameters of peptide substrates used to construct the model (D-TRAINING).

Peptide ^a	Protein (site of modification)	k_{cat}/K_M ($M^{-1}s^{-1}$) ^b	k_{cat} (s^{-1}) ^b	K_M (μM) ^b
EG(K-Ac)FVR	LMNA (K450ac)	220,000 ± 40,000	4.2 ± 0.7	19 ± 4
SD(K-Ac)TIG	TUBA1A, TUBA4A, TUBA3E, TUBA1C, TUBA1B, TUBA4A (K40ac)	70,000 ± 30,000	2.1 ± 0.3	29 ± 15
AM(K-Ac)HRS	MYO1G (K90ac)	63,000 ± 9,000	1.9 ± 0.2	30 ± 7
LA(K-Ac)EMK	CFAP157 (K35ac)	60,000 ± 20,000	1.4 ± 0.2	20 ± 10
TG(K-Ac)TVA	TCOF1 (K146ac)	50,000 ± 10,000	1.37 ± 0.09	27 ± 7
RT(K-Ac)SGD	ARHGEF1 (K234ac)	47,000 ± 9,000	1.5 ± 0.2	32 ± 9
SQ(K-Ac)KTF	GRP94 (K682ac)	44,000 ± 9,000	1.6 ± 0.1	40 ± 10
AG(K-Ac)RIA	DIP2A (K50ac)	40,000 ± 10,000	1.9 ± 0.5	50 ± 30
YK(K-Ac)FYE	HSP90A (K436ac)	40,000 ± 10,000	1.8 ± 0.6	21 ± 9
YG(K-Ac)LRK	ACTN1 (K195ac)	31,000 ± 4,000	6 ± 0.7	200 ± 80
EG(K-Ac)TNY	ZNF587 (K209ac)	29,000 ± 7,000	2 ± 0.4	70 ± 30
QK(K-Ac)VKE	ZNF280D (K209ac)	28,000 ± 6,000	1.8 ± 0.2	60 ± 20
YE(K-Ac)EKE	KIF5A, KIF5C (K348ac); KIF5B (K346ac)	27,000 ± 5,000	1.8 ± 0.1	60 ± 10
EG(K-Ac)TGE	ZKSCAN1 (K310ac)	25,000 ± 4,000	1.07 ± 0.08	43 ± 9
LS(K-Ac)KSK	PARP1 (K209ac)	24,000 ± 3,000	1.7 ± 0.2	70 ± 20
PA(K-Ac)ESP	ATBF1 (K3416ac); treacle (K904Ac)	22,000 ± 4,000	1.08 ± 0.05	49 ± 9
AM(K-Ac)KIR	CDK1 (K33ac)	16,000 ± 2,000	0.98 ± 0.09	60 ± 10
QY(K-Ac)KEL	LMNA (K260ac)	13,200 ± 700	1.58 ± 0.08	120 ± 10
KT(K-Ac)PIW	HSP90A (K294ac)	12,000 ± 2,000	0.8 ± 0.1	60 ± 20
AH(K-Ac)RGS	DDP3 (K294ac)	9,000 ± 3,000	>1	>150
KL(K-Ac)KKE	MYH1 (K1085ac)	6,000 ± 1,000	0.50 ± 0.09	90 ± 30
EV(K-Ac)KMT	MAP4 (K847ac)	5,600 ± 800	0.19 ± 0.02	35 ± 7
GY(K-Ac)KTK	RPL4 (K162ac)	5,000 ± 2,000	2.3 ± 0.5	400 ± 200
AH(K-Ac)KSH	S100AB (K84ac)	4,700 ± 400	>0.5	>100
PL(K-Ac)KDR	RPL3 (K393ac)	2,100 ± 200	>0.5	>200
SW(K-Ac)DGL	ACTN2 (K181ac)	1,200 ± 300	>0.3	>200

^aPeptides above the line were labeled as substrates, below the line were labeled as non-substrates (cutoff = 10,000 $M^{-1}s^{-1}$). ^bValues (mean ± standard error) were calculated using **Equation 4.1** from initial velocities (n=3-4) from 4 substrate concentrations with 0.1-0.5 μM HDAC6.

Table 4.2: List of peptides where only a lower limit for k_{cat}/K_M was determined (D-EXT).

Peptide ^a	Protein (site of modification)	k_{cat}/K_M ($M^{-1} s^{-1}$) ^b	k_{cat} (s^{-1}) ^b	K_M (μM) ^b
YD(K-Ac)DEV	SIRT1 (K430ac)	>110,000	0.28 ± 0.04	<25
ME(K-Ac)FKI	XPOT(K627ac)	>80,000	2.0 ± 0.5	<25
ET(K-Ac)YRW	VDAC1 (K61ac)	>70,000	4 ± 1	<50
GG(K-Ac)RVM	VPS35 (K35ac)	>60,000	1.4 ± 0.2	<25
YD(K-Ac)LRK	ACTN4 (K214ac); TUBG1 (K397ac)	>50,000	1.2 ± 0.2	<25
HG(K-Ac)EVG	PCPB1 (K23ac, K57ac)	>40,000	2.2 ± 0.5	<50
QG(K-Ac)SGN	MYO18B (K401ac)	>40,000	1.1 ± 0.2	<25
ID(K-Ac)RTI	EEF1A1 (K31ac), EEF1A2, EEF1A1P5 (K36ac)	>40,000	1.0 ± 0.2	<25
GA(K-Ac)DEP	TCOF1 (K1414ac)	>40,000	0.9 ± 0.2	<25
VS(K-Ac)RKL	UBA1 (K1024ac), ARHGAP32 (K674ac)	>40,000	1.0 ± 0.1	<25
YE(K-Ac)FRN	USP32 (K209ac)	>40,000	1.1 ± 0.2	<25
MK(K-Ac)LKE	PASD1 (K379ac)	>30,000	0.7 ± 0.1	<25
SQ(K-Ac)YKR	MART3 (K473ac)	>30,000	1.3 ± 0.4	<50
CG(K-Ac)GLE	CSRP1 (K151ac)	>20,000	1.1 ± 0.4	<50
SE(K-Ac)ILQ	TCOF1 (K294ac)	>20,000	1.2 ± 0.2	<50
EG(K-Ac)GNG	EIF5 (K28Ac)	>17,000	0.84 ± 0.09	<50
ME(K-Ac)KKE	GBP7 (K389Ac)	$3,300 \pm 200^c$	$>0.3^c$	$>150^c$
QD(K-Ac)PLR	CCDC86 (K261Ac)	>2,000	0.11 ± 0.04	<50

^aThe peptides listed below the line were not used for further analysis for the substrate model. ^bValues (mean \pm standard error) were calculated from initial velocities ($n=3-4$) from 4 substrate concentrations. ^cValues were calculated from initial velocities ($n=3-4$) from 3 substrate concentrations.

Structure-based computational prediction identifies most substrates of HDAC6

The calibration of FlexPepBind was performed using the k_{cat}/K_M values for HDAC6-catalyzed deacetylation of hexamer peptides (**Table 4.1**, hereinafter referred to as D-TRAINING), using an approach similar to the one previously applied to HDAC8 and FTase enzymes[9, 14]. First, the best measured substrate of HDAC6 (EGK_{Ac}FVR, derived from prelamin) was docked using Rosetta FlexPepDock[9] into the binding pocket of a

solved CD2 HDAC6 structure. Peptides from the D-TRAINING set were then threaded onto this template set and minimized (see Methods). The top-scoring model (according to reweighted score) was used as an estimate for their binding ability (Protocols BASIC_I and BASIC_R in **Table 4.3**). The performance of the protocol was evaluated based on the calculated binary distinction (AUC values) and correlation between experimental values and Rosetta scores (Pearson correlation).

Table 4.3: Different FlexPepBind protocols evaluated in this study. RLOOP_R, the final protocol, is highlighted in bold.

Protocol name	Receptor backbone minimization	Loop modeling	Starting structure selected by
BASIC_I	-	-	interface score
BASIC_R	-	-	reweighted score
RMIN_I	✓	-	interface score
RMIN_R	✓	-	reweighted score
RLOOP_I	-	✓	interface score
RLOOP_R	-	✓	reweighted score
RLOOPS_I	-	✓✓	interface score
RLOOPS_R	-	✓✓	reweighted score

Using the standard protocols, BASIC_I and BASIC_R, resulted in low AUC values and no correlation between the experimental values of our dataset and the Rosetta scores (reweighted score). We note that using a similar protocol we were able to achieve excellent distinction in our previous study of HDAC8 substrate selectivity[9]. Examination of possible differences between the structures of HDAC6 and HDAC8 highlighted a significantly longer loop in HDAC6 (455-467) near the binding site that might rearrange for substrate binding (**Figure 4.2**). We therefore first ran the same protocol with receptor backbone minimization in the refinement step for template generation (corresponding

protocols RMIN_I and RMIN_R). This allowed the loop to move away from the binding site, opening it up. This slight modification in the protocol resulted however only in a slight improvement of correlation to experimental values, while the binary distinction did not change much.

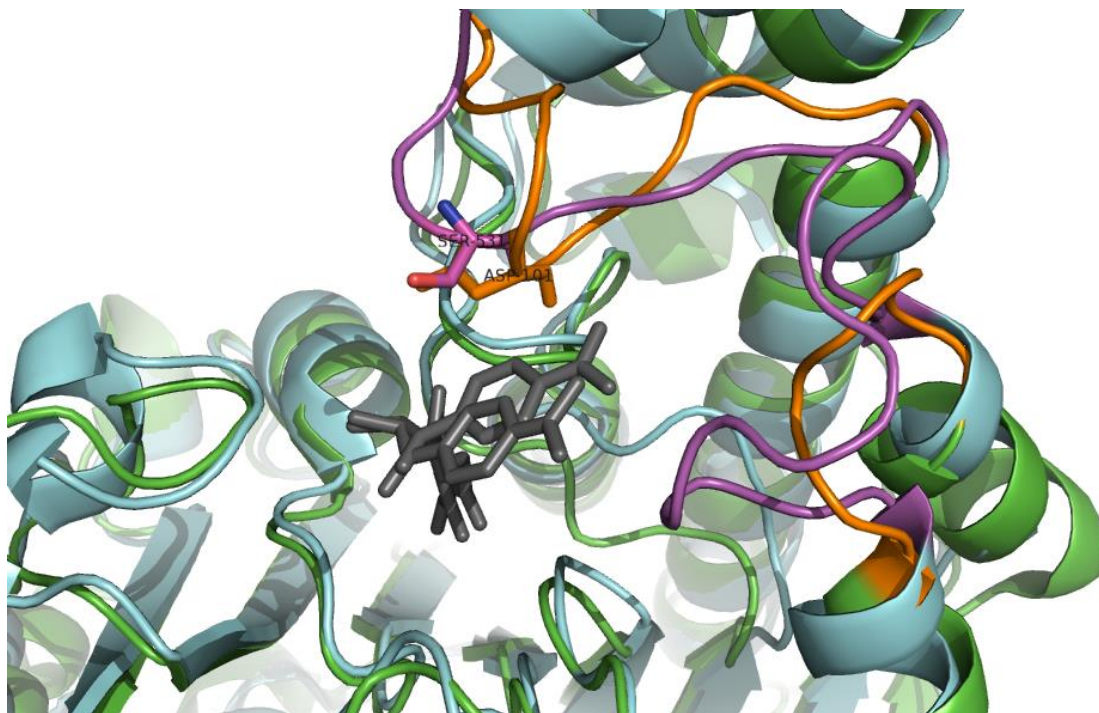


Figure 4.2: Key differences in the substrate binding structure between HDAC6 and HDAC8. Note the different loop structures (orange/magenta) and the difference in a key residue which forms a hydrogen bonding to the substrate and is shown in stick structure (Ser531 magenta, Asp101 orange). Both structures contain trichostatin A bound to the metal ion (dark gray). Green/orange: HDAC8 (PDB ID 1t64). Cyan/magenta: HDAC6 CD2 (PDB ID 5eek).

We proposed that additional receptor loop flexibility might improve the performance of our protocol. To test this, we modelled this loop (455-467) using the Rosetta kinematic loop closure with fragments protocol ([22], generating 1000 models; see Methods). We picked the top-scoring model and repeated the above protocol involving docking the best substrate peptide into the binding pocket, and threading different candidate substrates onto top-scoring models (protocols RLOOP_I and RLOOP_R, respectively). By using starting structures from this setup, we were able to

obtain good performance for the D-TRAINING set (**Table 4.4** and **Figure 4.3**). When selecting among a set of loop structures obtained from loop modeling (protocol RLOOPS_I), the results did not improve further. Loop flexibility is crucial for selectivity prediction as these were the only models to provide good correlations.

Table 4.4: Performance metrics of the top 3 protocols on the D-TRAINING dataset (25 peptides)

Protocol	Specificity ^a	Sensitivity ^b	MCC ^c	AUC ^d	Pearson Correlation [p-value]	Spearman Correlation
RLOOP_I	1 (1)	0.56 (0.62)	0.51	0.79 (0.85)	-0.58 [0.002]	-0.61 [0.001]
RLOOP_R	0.71 (0.71)	0.67 (0.74)	0.34	0.75 (0.85)	-0.53 [0.005]	-0.55 [0.003]
RLOOPS_I	0.71 (0.71)	0.77 (0.89)	0.46	0.77 (0.80)	-0.41 [0.035]	-0.57 [0.003]

^aSpecificity measures the proportion of negatives correctly identified. ^bSensitivity measures the proportion of positives correctly identified. ^cMatthew correlation coefficient. ^dArea under the ROC curve.

We first tested our protocols on dataset D-EXT, since our measurements indicated that these peptides are substrates, although accurate k_{cat}/K_M values could not be determined due to low K_M values (**Table 4.2**). We found that even though the RLOOP_I protocol performed best for D-TRAINING in terms of AUC and correlation values, it predicted that most of these peptides were non-substrates (**Figure 4.3-A**, depicted with triangles). Therefore, to reduce the number of false negatives, we moved forward with protocol RLOOP_R, which showed similar performance on the D-TRAINING set (**Figure 4.3-B,-E** and **Table 4.4**). To define substrates, we defined two cutoffs: -890 gives the best value by the Youden index ($Y = \text{specificity} + \text{sensitivity} - 1$), while the stricter cutoff of -900 safely eliminates all false positive peptides in the D-TRAINING set.

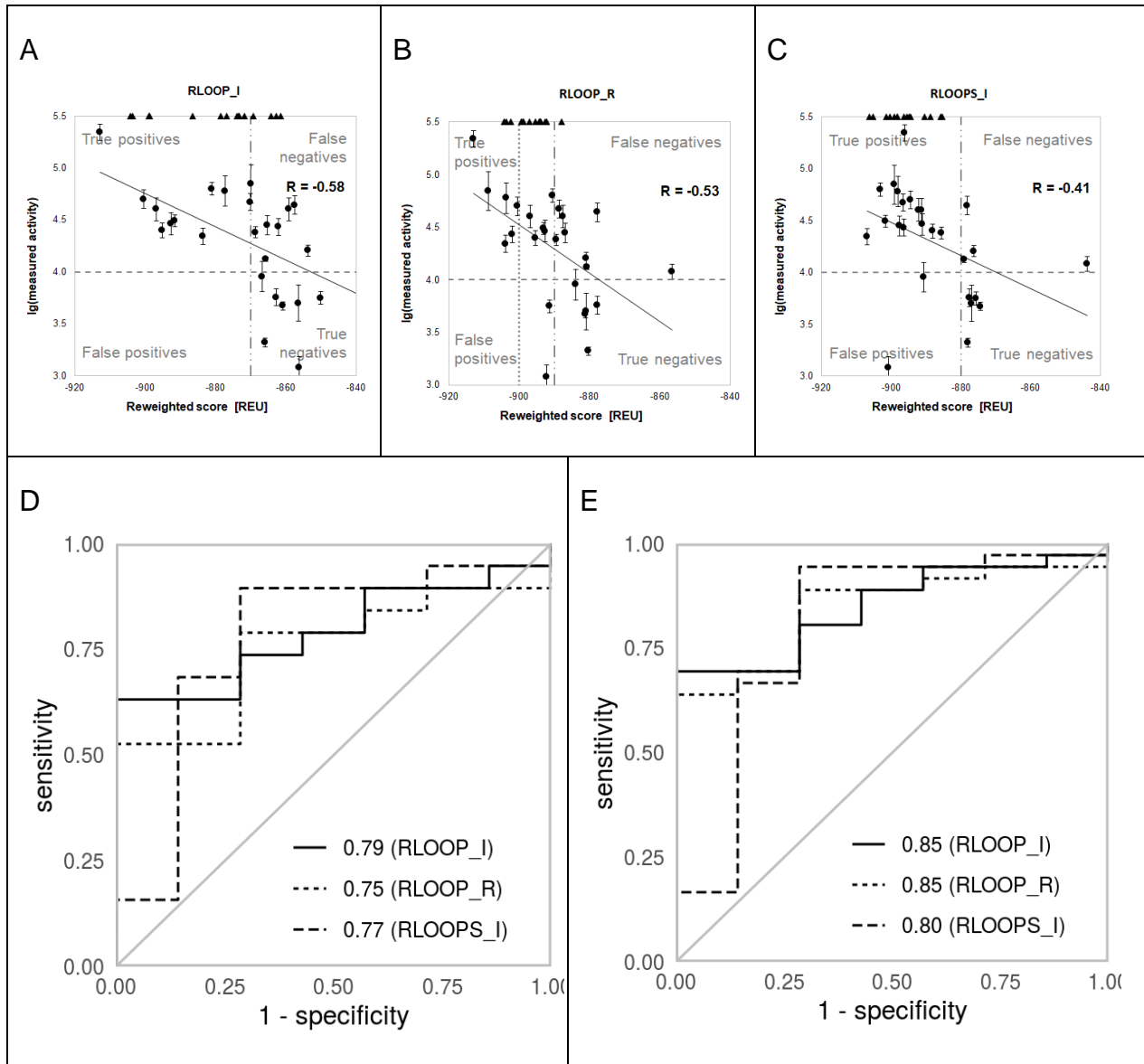


Figure 4.3: Performance of different protocols on the D-TRAINING set.

(A-C) Correlation plots between experimental (y-axis, logarithmic scale) and predicted substrate activities (x-axis). (A) Protocol RLOOP_I (cutoff: -870), (B) Protocol RLOOP_R (cutoff: -890, dotted dashed line), (C) Protocol RLOOPS_I (cutoff: -880, dotted dashed line. Strict cutoff: -900, dashed line). Circles: D-TRAINING dataset; triangles: D-EXT (presented on edge due to only having lower limits). (D-E) Area Under the Curve (AUC) plots for the three protocols (RLOOP_I, RLOOP_R and RLOOPS_I; solid, dashed and dotted line, respectively), reflecting the ability to distinguish between substrates and non-substrates. (D) D-TRAINING set (25 peptides) (E) D-TRAINING + D-EXT (25 + 16 peptides).

Predictions on the human acetylome

To detect new potential HDAC6 substrates, we used our calibrated protocol (RLOOP_R) to screen the human acetylome (from PhosphoSitePlus[23]; we used only peptides annotated from low-throughput experiments) (Figure 4.4-A). This screen

detected around 242 and 600 peptides classified as substrates by our strict and non-strict cutoff (belonging to 141 and 257 proteins, respectively). Few (3) of these peptides score better than the top-scoring peptide in the D-TRAINING set. In comparison to our previous study on HDAC8 selectivity[9], many more potential substrates are suggested, which is in line with the reported significantly lower selectivity observed for HDAC6.

We also compared the sequence logo created from these peptides to the one created from the training dataset (D-TRAINING), to ensure that we did not simply recreate the sequence specificity of that dataset (**Figure 4.4-B,-C**). Although there are some similarities, such as the preference for glycine at P₋₁ and for valine, tyrosine, and methionine at P₊₂, the sequence logo is different than the original database and has more variability.

We selected 11 peptide analogs to measure activity with HDAC6 to test the robustness of our prediction model. 10 of the 11 selected peptides gave k_{cat}/K_M values over 10,000 M⁻¹s⁻¹, characterizing them as substrates by our criteria (**Table 4.5, Figure 4.5**). One peptide did not meet the substrate cutoff, however, only a lower limit for the value of k_{cat}/K_M was determined so this peptide could be within substrate range as well. Since this was the lowest scored of the selected peptides, a stricter substrate cutoff of -900 (which eliminated all false positives in the training set (**Figure 4.3**)) could be more robust. Overall, the structure-based prediction model for zHDAC6 CD2 was able to predict good peptide substrate analogs for this enzyme with high accuracy.

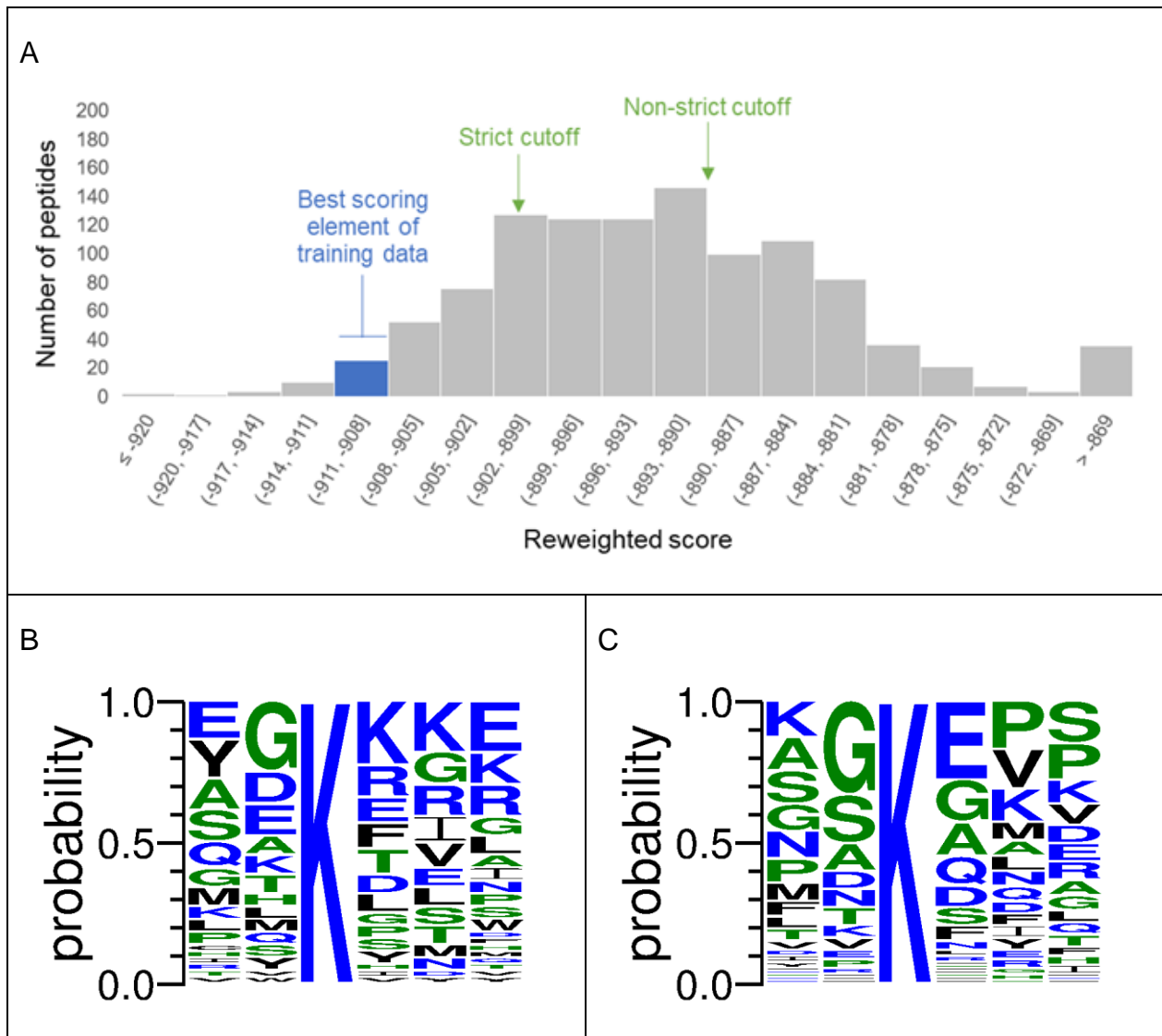


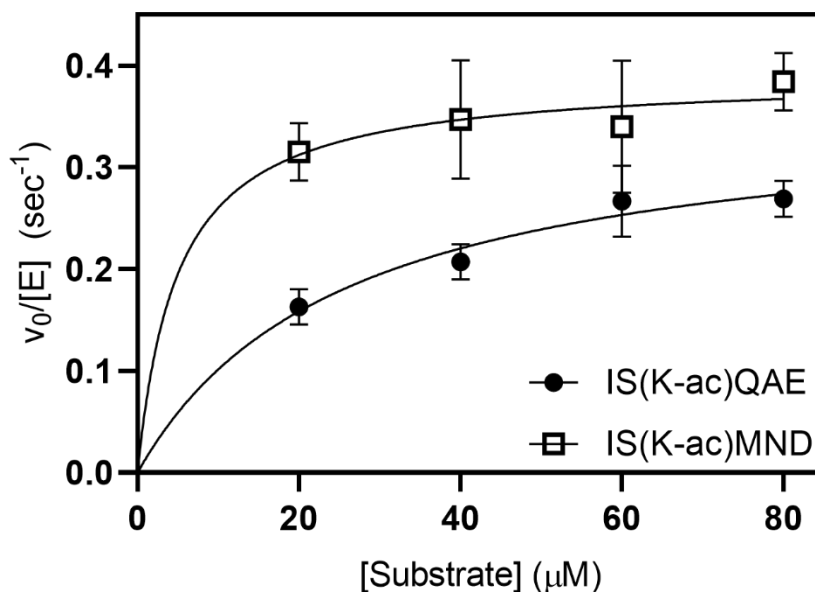
Figure 4.4 : Application of the calibrated protocol (RLOOP_R) to the acetylome to detect novel potential HDAC6 substrates.

(A) Distribution of scores obtained for acetylated peptides (as annotated in the Phosphosite database). The bin for the best scoring peptide in the training set is highlighted in blue, and the green arrows mark the threshold for discriminating substrates defined for RLOOP_R (-890 and -900 for non-strict and strict cutoff, respectively) (B-C) Sequence logos of (B) measured peptides (D-TRAINING set) and (C) top 100 peptides predicted by our protocol.

Table 4.5: Kinetic parameters of selected predicted substrates from the human acetylome.

Peptide	Protein (site of modification)	Score	k_{cat}/K_M ($M^{-1} s^{-1}$) ^a	k_{cat} (s^{-1}) ^a	K_M (μM) ^a
PC(K-Ac)EVD	NFAT5 (K282ac)	-923.56	>10,000	0.22	≤ 20
PG(K-Ac)EEK	FOXO1 (K440ac)	-914.77	>16,000	0.32	≤ 20
FP(K-Ac)EAK	EGFR (K1179ac)	-913.74	>26,000	0.52 ± 0.02	<20
IS(K-Ac)MND	IFI16 (K451ac)	-911.69	>20,000	0.39 ± 0.02	<20
KG(K-Ac)QAE	HMG1 (K61ac)	-909.3	$14,000 \pm 3,000$	0.36 ± 0.04	26 ± 8
HS(K-Ac)GFG	TARDBP (K145ac)	-909.19	>45,000	0.9 ± 0.1	<20
SG(K-Ac)GKK	GATA1 (K312ac)	-908.68	>14,000	0.27 ± 0.03	<20
AG(K-Ac)FGP	CGAS (K50ac)	-905.59	>10,000	0.2 ± 0.1	<20
QA(K-Ac)SPP	MEF2C (K239ac)	-905.57	>35,000	0.7 ± 0.2	<20
MG(K-Ac)GVS	ENO1 (K60ac)	-905.4	>17,000	0.34 ± 0.09	<20
NG(K-Ac)LTG	GAPDH (K227ac)	-897.7	>8,000	0.16 ± 0.06	<20

^aValues (mean \pm standard error) were calculated using **Equation 4.1** from initial velocities ($n=4$) from 4 substrate concentrations.

**Figure 4.5:** Dependence of deacetylation rate on substrate concentration catalyzed by zHDAC6 CD2 for two representative peptides selected from the top acetylome hits.

Comparison of CD2 to CD12 substrate selectivity

While we were able to create a model to predict substrates of HDAC6, it is based on a truncated form of the enzyme. The truncated form contains the catalytic domain, CD2, proposed to contain the deacetylase activity, but the other catalytic domain, CD1, likely serves some purpose in activity and potentially selectivity of HDAC6. We therefore explored the activity of a tandem construct, CD12, with a subset of the same peptides tested with CD2.

When initially testing the CD12 construct, we found it was less active than CD2. The difference in activity was surprising as previous studies with both constructs found them to have similar activity[3]. An active site titration using the HDAC6-specific inhibitor Tubastatin A[27] was performed with CD12 to ensure accurate measurement of the concentration of active enzyme. Assuming a 1:1 binding of inhibitor to HDAC6, as inhibitors have been developed primarily for the CD2 domain[27], the titration indicated an active enzyme concentration comparable to the total concentration determined by absorbance at 280 nm (**Figure 4.6**).

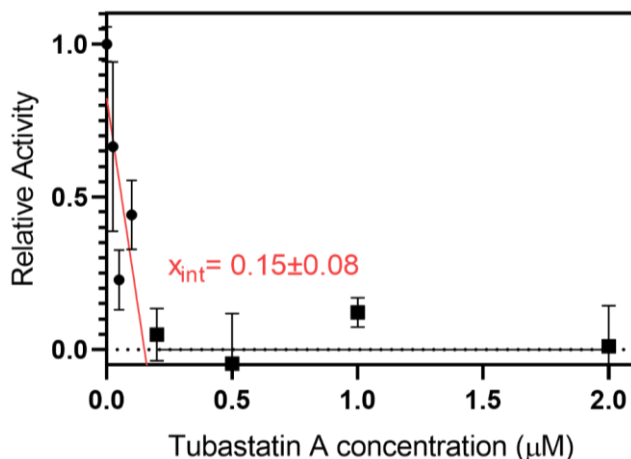


Figure 4.6: Active site titration of purified zHDAC6 CD12 with Tubastatin A. The calculated x-intercept of when activity was completely inhibited was within error of the A_{280} calculated concentration (0.1 μM).

We additionally measured the activity with a substrate that has been previously tested with full-length zebrafish HDAC6 and full kinetic parameters have been published[28]. When tested with this substrate, Boc-Lys(Ac)-AMC, our value for k_{cat}/K_M was higher than the previously published number due to both increased in the k_{cat} value and a decrease in the K_m values (**Table 4.6**). These comparisons suggest that our protein constructs have good activity and that the additional residues present in zHDAC6 CD12 and full-length zHDAC6 lower the reactivity with peptide substrates.

Table 4.6: Kinetic parameters obtained for various HDAC6 constructs with a Lys(Ac)-AMC substrate.

Construct	k_{cat}/K_M ($M^{-1}s^{-1}$)	$k_{cat}(s^{-1})$	K_M (μM)
zHDAC6 CD2 ^a	>150,000	3.7±0.6	<25
zHDAC6 CD12 ^a	30,000±10,000	0.49±0.07	20±10
zHDAC6 ^b	6,000±1,000	0.26±0.01	45±8

^a k_{cat}/k_M values are presented as mean ± standard error, calculated from a saturation curve containing 4 substrate concentrations where the initial velocity was calculated from 4 timepoints. ^bThe values given for zHDAC6, a full-length construct, were taken from Miyake et al[28] for comparison to zHDAC6 CD12.

We then proceeded to measure the activity of CD12 with peptide substrates previously tested with CD2 (**Table 4.7**). We measured the kinetic parameters for deacetylation of a total of 13 peptides catalyzed by CD12 and found that a majority of the peptides contained significantly different kinetic parameters than CD2 (**Table 4.7**). The k_{cat} values for CD12 were lower than CD2 for 11 of the 13 peptides by 10-fold, on average. Furthermore, the K_M values for 8 of the 13 peptides were lower for CD12 compared to CD2 and in many cases below the assay detection limit making it possible to only calculate lower limits for the value of k_{cat}/K_M . There were no peptides tested which had greater activity with CD12 than CD2.

Table 4.7: Kinetic parameters of peptide substrates tested with zHDAC6 CD12

Peptide	Protein (site of modification)	k_{cat}/K_M ($M^{-1} s^{-1}$) ^a	k_{cat} (s^{-1}) ^a	K_M (μM) ^a
KT(K-Ac)PIW	HSP90A (K294ac)	6,000±5,000	0.10±0.02	20±20
ME(K-Ac)KKE	GBP7 (K389Ac)	>2,000	0.06±0.03	<25
MK(K-Ac)LKE	PASD1 (K379ac)	>2,000	0.06±0.04	<35
KL(K-Ac)KKE	MYH1 (K1085ac)	>2,000	0.06±0.02	<25
AM(K-Ac)HRS	MYO1G (K90ac)	>5,000	0.14±0.05	<25
ME(K-Ac)FKI	XPOT(K627ac)	350±50	>.15	>0.07
YD(K-Ac)LRK	ACTN4 (K214ac); TUBG1 (K397ac)	5,000±4,000	0.16±0.08	30±40
ID(K-Ac)RTI	EEF1A1 (K31ac), EEF1A2, EEF1A1P5 (K36ac)	2,000±1,000	0.15±0.07	70±60
EV(K-Ac)KMT	MAP4 (K847ac)	600±400	0.1±0.1	200±300
QK(K-Ac)VKE	ZNF280D (K209ac)	>2,000	0.047±0.004	<25
AG(K-Ac)RIA	DIP2A (K50ac)	>6,000	0.09±0.06	<15
QD(K-Ac)PLR	CCDC86 (K261Ac)	>2,000	0.1±0.1	<25
QG(K-Ac)SGN	MYO18B (K401ac)	10,000±4,000	0.15±0.02	15±8

^aValues (mean ± standard error) were calculated from saturation curves constructed from initial velocities (n=4) of at least 3 substrate concentrations.

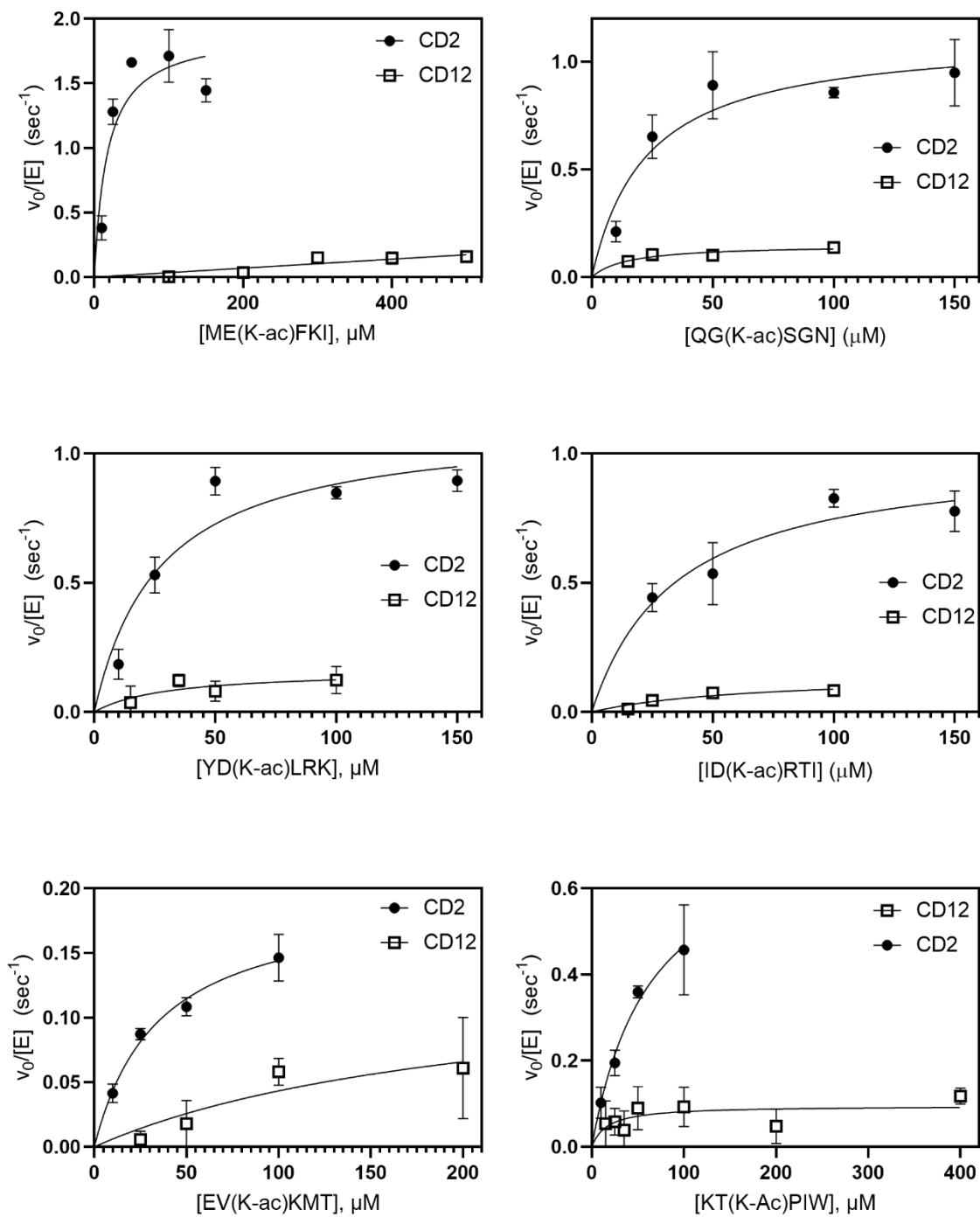


Figure 4.7: Dependence of deacetylation rate on substrate concentration for two representative peptides measured with both CD12 and CD2.

Discussion

We have successfully produced a structure-based model to predict HDAC6 substrates. However, the increased activity and promiscuity of HDAC6 made this process more difficult than for the previous HDAC isozyme studied, HDAC8. We randomly selected the peptides in the training set with high sequence variability, yet only a small fraction (7 out of 50) was characterized as non-substrates. Furthermore, the difference in activity between a substrate and non-substrate was smaller than what was previously seen with HDAC8, where the difference was over four orders of magnitude compared to the two orders of magnitude seen for HDAC6. A consequence of the condensed area of peptide activity is that good correlations are more difficult to obtain. By recognizing and utilizing structural differences between the isozymes, we were able to build a model with a correlation above 0.5.

Structural differences dictate substrate binding specificity

For HDAC8 we were able to obtain good predictions without introducing any backbone receptor flexibility[9]. Comparison of the HDAC6 and HDAC8 structures highlighted two main differences in the loops forming the binding pocket that could lead to differences in binding selectivity. First, the loop with the residue that forms a hydrogen bond and positions the substrate's acetylated lysine, D101 in HDAC8 and S531 in HDAC6, contact the substrate at a similar position, but stem from a very different loop structure (**Figure 4.2**). This loop participates in the formation of the pocket accommodating the residue preceding the acetylated lysine (P-1). In HDAC6 this pocket is considerably smaller, explaining the significant enrichment for glycine at P-1. Second, the loop located near the pocket that accommodates the residues C-terminal to the

acetylated lysine (P₊₁) is significantly longer in HDAC6 (12 residues compared to 7 in HDAC8) and more hydrophobic. The larger size of the loop, together with its more nonpolar character, suggest that this HDAC6 loop may be more flexible, allowing for adaptations of the binding groove resulting in a more promiscuous binding pattern. We showed here that this loop needs to move for our protocol to succeed, suggesting that this pocket opens for substrate binding.

We additionally demonstrated that the second catalytic domain (CD1), present in a more full-length construct of HDAC6, effects catalytic activity and selectivity of CD12. When evaluating the small number of sequences of peptides which had similar activities for CD2 and CD12, there is a preference for lysine residues at P₋₂, P₊₁, and P₊₂ and for glutamate at P₊₃ (**Figure 4.8-B**). Sequences of peptides where CD2 activity was greater than CD12 do not follow any striking residue preferences, with only a slight preference for lysine at P₊₂ (**Figure 4.8-A**). These differences suggest structural differences between CD12 versus CD2 in binding substrates. However, the crystal structures of full-length zHDAC6 and CD2 do not show any significant structural differences (**Figure 4.8-C**). There is a slight perturbation of the loop which accommodates the C-terminal peptide residues, which perhaps limits the amount of loop opening seen as necessary for modeling the promiscuity of CD2. Sampling of yet more loop structures, prior or together with the peptide structure optimization, can in the future pave the way for application of this approach to more flexible receptors, provided proper calibration is applied to focus on relevant, but dynamic sets of conformers.

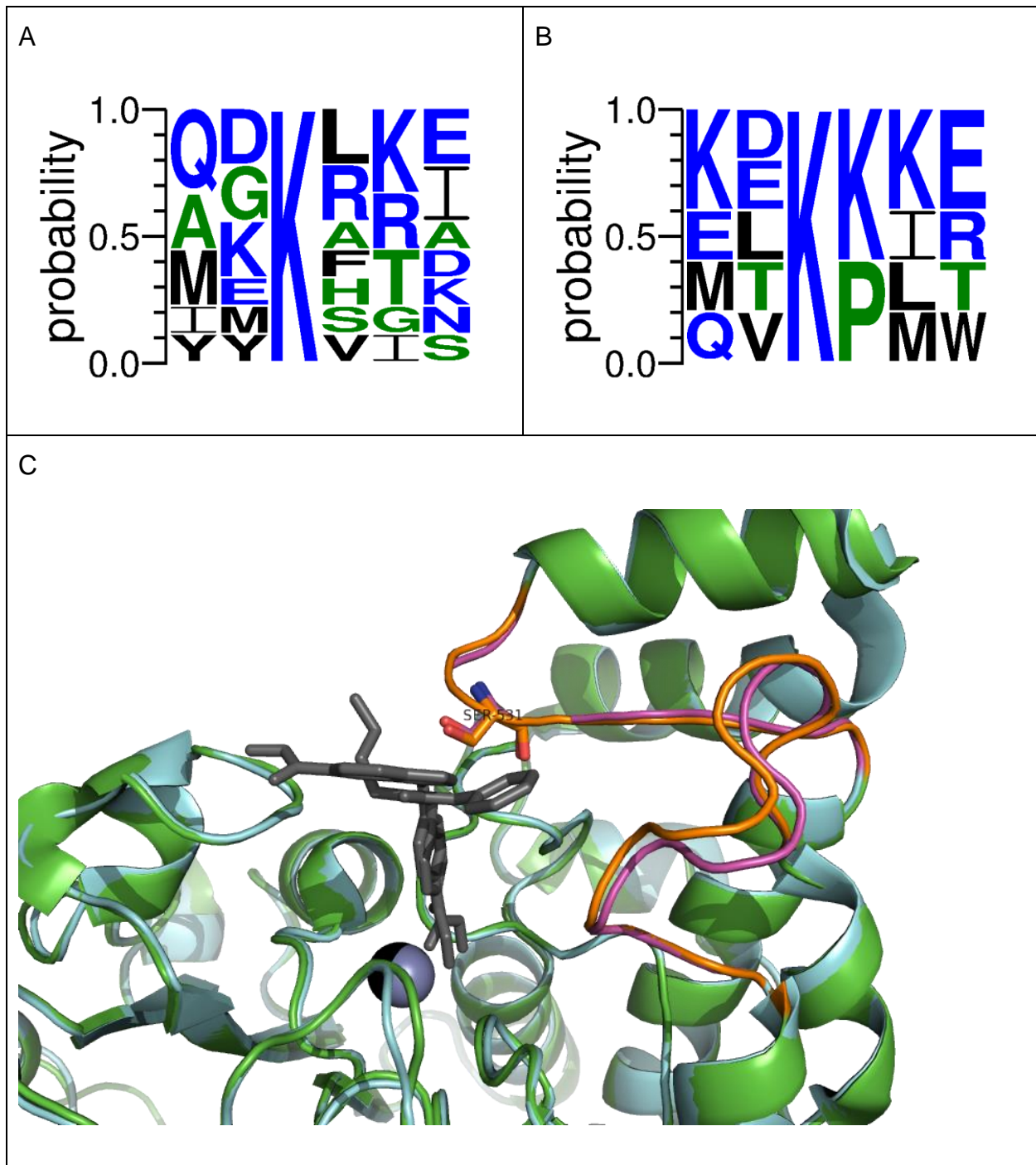


Figure 4.8: Structural differences between zHDAC6 CD2 and CD12.

A) Sequence logo of peptides which display greater activity with CD2 than CD12. B) Sequence logo of peptides which display similar activity with CD2 and CD12. C) Overlay of CD2 (green/orange) and CD12(cyan/magenta) structures.

Structure-based model can predict novel HDAC6 substrates

The proposed model seems to be especially suited to confidently predicting good substrates of HDAC6. When the model was applied to the known acetylome, over 300 sites were identified below the substrate cutoff of -890. The proteins belonging to the top 100 best scoring peptides (see Appendix) were submitted to Reactome Pathway analysis[29] to identify the likely pathways they are involved in. Using this analysis we identified several pathways which have been previously reported to be linked to HDAC6, such as autophagy and aggrephagy[30], transcriptional regulation by TP53 [31], NOTCH1 signaling[32, 33], EGFR signaling[34, 35], MAPK signaling[36], and cytosolic sensors of pathogen-associated DNA[37, 38]. It is encouraging that our protocol can confirm HDAC6 involvement in these pathways.

To further test our protocol's reliability, we determined HDAC6 activity with 11 predicted substrate peptides. HDAC6 demonstrated catalytic efficiency above the substrate cutoff for almost all the peptides (10 out of 11). Only two of these substrates, epidermal growth factor receptor (EGFR) and TAR DNA-binding protein 43 (TDP43), have previously been explored as direct substrates of HDAC6[34, 35, 39, 40]. The other proteins for whom peptide analogs were tested represent novel potential substrates of HDAC6. A few of these proteins have identified interactions with HDAC6: nuclear factor of activated T-cells 5 (NFAT5 [41]), alpha enolase (ENO1 [42]), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH [42]). The misregulation of several predicted substrates, ENO1, a tumor associated antigen[43], GAPDH, regulator of cell death[44], and forkhead box protein M1 (FOXO1), a transcription factor responsible for cell proliferation[45], has been implicated in cancer phenotypes. HDAC6 activity is further

connected to the immune response with the potential substrates of: non-histone chromosomal protein HMG-14 (HMGN1), an alarmin which triggers the innate and adaptive immune response[46]; gamma-interferon-inducible protein 16 (IFI16), which triggers host response to viral DNA[47]; and cyclic GMP-AMP synthase (CGAS), a cytosolic DNA sensor[48]. New roles for HDAC6 could potentially be found in the acetylation of NFAT5[41] and MEF2C[49], transcription factors important for muscle cell development, and GATA1, an important transcription factor in blood cells whose acetylation promotes bromodomain binding to chromatin[50]. The predicted substrates are diverse, yet most fit within functions previously linked to HDAC6, immunity[37, 51, 52] and cancer[53, 54], but this analysis also identified some potential novel roles of HDAC6 in muscle and blood cell development.

The successful construction of a structure-based model for another HDAC isozyme indicates the positive ability of utilizing this method to predict novel substrates. There have been several methods used to predict HDAC substrates including *in vivo* inhibition[55] and substrate-trapping pulldown methods including photo-crosslinking[25] and mutant capture[56]. While these methods take advantage of the enzymes in their natural, biologically relevant state, data analysis is more cumbersome, and the number of substrates identified is limited (≤ 100). Using a computational model allows the screening of all known acetylated proteins upfront with less effort, where identified substrates can then be followed up on in more detailed, specific analysis. Structure-based modelling is not without its limitations, as it is dependent on substrate selectivity being predicted at the local sequence level, which was challenging and somewhat lacking for HDAC6 in comparison to HDAC8. Even with these limitations, using the prediction model was

illuminating for discovering novel substrates and novel avenues of substrate exploration for HDAC6.

References

1. Yang, X.J. and E. Seto, *The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men*. Nat Rev Mol Cell Biol, 2008. **9**(3): p. 206-18.
2. Kutil, Z., et al., *The unraveling of substrate specificity of histone deacetylase 6 domains using acetylome peptide microarrays and peptide libraries*. Faseb j, 2019. **33**(3): p. 4035-4045.
3. Hai, Y. and D.W. Christianson, *Histone deacetylase 6 structure and molecular basis of catalysis and inhibition*. Nat Chem Biol, 2016. **12**(9): p. 741-7.
4. Hubbert, C., et al., *HDAC6 is a microtubule-associated deacetylase*. Nature, 2002. **417**(6887): p. 455-8.
5. Zhang, X., et al., *HDAC6 modulates cell motility by altering the acetylation level of cortactin*. Mol Cell, 2007. **27**(2): p. 197-213.
6. Kovacs, J.J., et al., *HDAC6 regulates Hsp90 acetylation and chaperone-dependent activation of glucocorticoid receptor*. Mol Cell, 2005. **18**(5): p. 601-7.
7. de Zoeten, E.F., et al., *Histone deacetylase 6 and heat shock protein 90 control the functions of Foxp3(+) T-regulatory cells*. Mol Cell Biol, 2011. **31**(10): p. 2066-78.
8. Kawaguchi, Y., et al., *The Deacetylase HDAC6 Regulates Aggresome Formation and Cell Viability in Response to Misfolded Protein Stress*. Cell, 2004. **115**: p. 727-38.
9. Alam, N., et al., *Structure-Based Identification of HDAC8 Non-histone Substrates*. Structure, 2016. **24**(3): p. 458-68.
10. Gibson, T.J., et al., *Experimental detection of short regulatory motifs in eukaryotic proteins: tips for good practice as well as for bad*. Cell Commun Signal, 2015. **13**: p. 42.
11. Kumar, M., et al., *ELM-the eukaryotic linear motif resource in 2020*. Nucleic Acids Res, 2020. **48**(D1): p. D296-d306.
12. Weatheritt, R.J., et al., *The identification of short linear motif-mediated interfaces within the human interactome*. Bioinformatics, 2012. **28**(7): p. 976-82.
13. Tallorin, L., et al., *Discovering de novo peptide substrates for enzymes using machine learning*. Nat Commun, 2018. **9**(1): p. 5253.

14. London, N., et al., *Identification of a novel class of farnesylation targets by structure-based modeling of binding specificity*. PLoS Comput Biol, 2011. **7**(10): p. e1002170.
15. Alam, N. and O. Schueler-Furman, *Modeling Peptide-Protein Structure and Binding Using Monte Carlo Sampling Approaches: Rosetta FlexPepDock and FlexPepBind*. Methods Mol Biol, 2017. **1561**: p. 139-169.
16. London, N., et al., *In silico and in vitro elucidation of BH3 binding specificity toward Bcl-2*. Biochemistry, 2012. **51**(29): p. 5841-50.
17. Wolfson, N.A., et al., *An enzyme-coupled assay measuring acetate production for profiling histone deacetylase specificity*. Anal Biochem, 2014. **456**: p. 61-9.
18. Welker-Leng, K., *Investigating the Substrate Selectivity and Regulation of Histone Deacetylases*, in *Chemistry*. 2019, University of Michigan.
19. Reger, A.S., J.M. Carney, and A.M. Gulick, *Biochemical and Crystallographic Analysis of Substrate Binding and Conformational Changes in Acetyl-CoA Synthetase*. Biochemistry, 2007. **46**(22): p. 6536-6546.
20. Huang, X. and M. Hernick, *A fluorescence-based assay for measuring N-acetyl-1-D-myo-inositol-2-amino-2-deoxy- α -D-glucopyranoside deacetylase activity*. Anal Biochem, 2011. **414**(2): p. 278-81.
21. Berman, H.M., et al., *The Protein Data Bank*. Nucleic Acids Res, 2000. **28**(1): p. 235-42.
22. Stein, A. and T. Kortemme, *Improvements to robotics-inspired conformational sampling in rosetta*. PLoS One, 2013. **8**(5): p. e63090.
23. Hornbeck, P.V., et al., *PhosphoSitePlus, 2014: mutations, PTMs and recalibrations*. Nucleic Acids Res, 2015. **43**(Database issue): p. D512-20.
24. Stark, C., et al., *BioGRID: a general repository for interaction datasets*. Nucleic Acids Res, 2006. **34**(Database issue): p. D535-9.
25. Lopez, J.E., et al., *HDAC8 Substrates Identified by Genetically Encoded Active Site Photocrosslinking*. J Am Chem Soc, 2017. **139**(45): p. 16222-16227.
26. Sullivan, E.D., *Unlocking an HDAC Toolbox: Methods Towards Understanding Isozyme-Specific Activity*, in *Chemical Biology*. 2016, University of Michigan. p. 241.
27. Butler, K.V., et al., *Rational design and simple chemistry yield a superior, neuroprotective HDAC6 inhibitor, tubastatin A*. J Am Chem Soc, 2010. **132**(31): p. 10842-6.

28. Miyake, Y., et al., *Structural insights into HDAC6 tubulin deacetylation and its selective inhibition*. Nature Chemical Biology, 2016. **12**(9): p. 748-754.
29. Fabregat, A., et al., *Reactome pathway analysis: a high-performance in-memory approach*. BMC bioinformatics, 2017. **18**(1): p. 142.
30. Lamark, T. and T. Johansen, *Aggrephagy: selective disposal of protein aggregates by macroautophagy*. Int J Cell Biol, 2012. **2012**: p. 736905.
31. Ryu, H.W., et al., *HDAC6 deacetylates p53 at lysines 381/382 and differentially coordinates p53-induced apoptosis*. Cancer Lett, 2017. **391**: p. 162-171.
32. Deskin, B., et al., *Inhibition of HDAC6 Attenuates Tumor Growth of Non-Small Cell Lung Cancer*. Transl Oncol, 2020. **13**(2): p. 135-145.
33. Deskin, B., et al., *Requirement of HDAC6 for activation of Notch1 by TGF- β 1*. Sci Rep, 2016. **6**: p. 31086.
34. Wang, Z., et al., *HDAC6-mediated EGFR stabilization and activation restrict cell response to sorafenib in non-small cell lung cancer cells*. Med Oncol, 2016. **33**(5): p. 50.
35. Liu, W., et al., *HDAC6 regulates epidermal growth factor receptor (EGFR) endocytic trafficking and degradation in renal epithelial cells*. PLoS One, 2012. **7**(11): p. e49418.
36. Haakenson, J., et al., *HDAC6-Dependent Functions in Tumor Cells: Crossroad with the MAPK Pathways*. Crit Rev Oncog, 2015. **20**(1-2): p. 65-81.
37. Moreno-Gonzalo, O., et al., *HDAC6 controls innate immune and autophagy responses to TLR-mediated signalling by the intracellular bacteria *Listeria monocytogenes**. PLoS Pathog, 2017. **13**(12): p. e1006799.
38. Moreno-Gonzalo, O., F. Mayor, Jr., and F. Sánchez-Madrid, *HDAC6 at Crossroads of Infection and Innate Immunity*. Trends Immunol, 2018. **39**(8): p. 591-595.
39. Deribe, Y.L., et al., *Regulation of epidermal growth factor receptor trafficking by lysine deacetylase HDAC6*. Sci Signal, 2009. **2**(102): p. ra84.
40. Cohen, T.J., et al., *An acetylation switch controls TDP-43 function and aggregation propensity*. Nat Commun, 2015. **6**: p. 5845.
41. Herbelet, S., et al., *Localization and Expression of Nuclear Factor of Activated T-Cells 5 in Myoblasts Exposed to Pro-inflammatory Cytokines or Hyperosmolar Stress and in Biopsies from Myositis Patients*. Frontiers in Physiology, 2018. **9**(126).

42. Watson, G.W., et al., *HDAC6 activity is not required for basal autophagic flux in metastatic prostate cancer cells*. *Exp Biol Med* (Maywood), 2016. **241**(11): p. 1177-85.
43. Cappello, P., et al., *Alpha-Enolase (ENO1), a potential target in novel immunotherapies*. *Front Biosci* (Landmark Ed), 2017. **22**: p. 944-959.
44. Zhang, J.Y., et al., *Critical protein GAPDH and its regulatory mechanisms in cancer cells*. *Cancer Biol Med*, 2015. **12**(1): p. 10-22.
45. Gartel, A.L., *FOXM1 in Cancer: Interactions and Vulnerabilities*. *Cancer Res*, 2017. **77**(12): p. 3135-3139.
46. Yang, D., et al., *High-mobility group nucleosome binding domain 1 (HMGN1) functions as a Th1-polarizing alarmin*. *Semin Immunol*, 2018. **38**: p. 49-53.
47. Guise, A.J., et al., *Histone deacetylases in herpesvirus replication and virus-stimulated host defense*. *Viruses*, 2013. **5**(7): p. 1607-32.
48. Sun, L., et al., *Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway*. *Science*, 2013. **339**(6121): p. 786-91.
49. Dong, C., et al., *Myocyte enhancer factor 2C and its directly-interacting proteins: A review*. *Progress in Biophysics and Molecular Biology*, 2017. **126**: p. 22-30.
50. Katsumura, K.R. and E.H. Bresnick, *The GATA factor revolution in hematology*. *Blood*, 2017. **129**(15): p. 2092-2102.
51. Choi, S.J., et al., *HDAC6 regulates cellular viral RNA sensing by deacetylation of RIG-I*. *Embo j*, 2016. **35**(4): p. 429-42.
52. Liu, H.M., et al., *Regulation of Retinoic Acid Inducible Gene-I (RIG-I) Activation by the Histone Deacetylase 6*. *EBioMedicine*, 2016. **9**: p. 195-206.
53. Aldana-Masangkay, G.I. and K.M. Sakamoto, *The role of HDAC6 in cancer*. *J Biomed Biotechnol*, 2011. **2011**: p. 875824.
54. Li, T., et al., *Histone deacetylase 6 in cancer*. *J Hematol Oncol*, 2018. **11**(1): p. 111.
55. Olson, D.E., et al., *An unbiased approach to identify endogenous substrates of "histone" deacetylase 8*. *ACS Chem Biol*, 2014. **9**(10): p. 2210-6.
56. Nalawansa, D.A., et al., *HDAC1 Substrate Profiling Using Proteomics-Based Substrate Trapping*. *ACS Chem Biol*, 2018. **13**(12): p. 3315-3324.

Chapter 5 : Conclusions and Future Directions

Regulation of HDACs is multi-faceted

My thesis has explored regulation of HDACs by probing several variables, namely: identity of the active site metal, post-translational modifications, and complex formation. It is known in the field that these variables play key roles in regulation, but there is limited information about how they affect HDAC deacetylase activity. My research has contributed to the growing knowledge of these regulatory measures and will provide insights into new directions of research into HDACs.

Identity of *in vivo* active site metal

The field at large has not explored the identity of the *in vivo* active site metal of HDACs. The prevailing view is that HDACs are Zn(II)-dependent deacetylases. No significant structural differences between HDAC8 with reconstituted metals have been observed[1]. However, there have been continued examples of varied deacetylase activity and selectivity between the different reconstituted forms[2, 3], which indicates metal identity is changing some aspect of enzyme catalysis. My exploration of metal specific inhibitors detailed in Chapter 2 illustrates that there are significant differences in ligand affinity depending on the active site metal ion, corroborating the previous studies demonstrating the differing selectivity of the different reconstituted forms of HDAC8.

With evidence of these differences, it would be interesting to explore what aspects of activity are most affected. The k_{cat} and K_M values for a coumarin labeled peptide have

been determined for HDAC8-Zn(II) and HDAC8-Fe(II), where HDAC8-Zn(II) has a higher k_{cat} (about 2-fold) but also has a larger K_M (about 5-fold)[3]. A smaller K_M for HDAC8-Fe(II) could indicate a structure that is more conducive for substrate binding. It has previously been proposed Fe(II) and Co(II) form a stronger interaction with the acetylated substrate than Zn(II) leading to the observed activity differences and the lower K_M values for both HDAC8-Fe(II) and HDAC8-Co(II)[3]. This would seem to indicate similar structure between HDAC8-Fe(II) and HDAC8-Co(II), though that contradicts what I observed with the metal-specific inhibitors in Chapter 2 as the metal-specific inhibitors were unable to inhibit HDAC8-Co(II). However, only having this detailed kinetic information for one peptide limits our ability to make conclusions. It would be useful to know if this trend also depends on substrate sequence, but only overall catalytic efficiency (k_{cat}/K_M) has been determined for unlabeled peptides[2] due to lower activity and higher K_M values, so it is unknown if the trend is generalizable to HDAC8 kinetics. Ideally, a crystal structure of HDAC8-Fe(II) bound to a Fe(II)-specific inhibitor or structures of unliganded different metalloforms would visualize the structural differences and how these changes relate to substrate affinity and reactivity.

To validate the importance of HDAC8-Fe(II), an *in vivo* connection needs to be demonstrated. The cellular concentrations of the two metal ions are comparable to their binding affinities to HDAC8[4-7], so switching between the two metals could be used as a regulatory mechanism that alters both activity and selectivity. This has been a difficult area to research as the endogenous expression levels of HDAC8 are too low to provide an adequate amount of metal-bound enzyme to accurately measure the metal identity by ICP-MS. When HDAC8 is overexpressed in tissue culture cells the bound metal ion is

mainly the tight-binding Zn(II) (unpublished data from Fierke lab). It is possible that the preference for Zn(II) when HDAC8 is overexpressed could be due to Fe(II) being highly regulated in cells with insufficient Fe(II) for the additional protein[6].

A different way to examine if HDAC8 is regulated by Fe(II) and Zn(II) is by exploring the regulation of HDACs in cells. HDAC8-Fe(II) would be sensitive to reactive oxygen species, so HDAC activity and selectivity would change based on cellular environment. I demonstrated in Chapter 2 that HDAC activity in cell lysates is oxygen sensitive, and HDAC activity has been connected to hypoxic conditions through regulation of hypoxia-inducible factor 1 α [8-10]. Though the oxygen sensitivity of HDACs has also been connected to redox sensitive residues in several members of the HDAC family where oxidation of these residues inhibits deacetylase activity, it has not been explored if this oxidation changes both activity and selectivity.

The lack of clarity indicates further studies need to be conducted to determine if the identity of the active site metal is a component of *in vivo* regulation of HDACs. Thus far, the primary studies have been performed with HDAC8 but perhaps further studies with other HDAC isozymes such as HDAC6 and HDAC1 could help elucidate if this possibility of metal switching is specific to HDAC8 or pertains to the whole enzyme family. One way to start exploring this area is to prepare apo-HDAC6 and apo-HDAC1 and use the identified metal-specific inhibitors with different reconstituted forms of HDAC6 and HDAC1 to see if the ligand affinity and catalytic activity are also dependent on the active site metal ion.

Post-translational modifications

HDACs themselves have been found to be subject to modifications as knowledge of acetylation and other post-translational modifications has expanded across the proteome. Among the HDACs I have discussed within my thesis, phosphorylation has been the only PTM studied as important for regulating the activity of HDAC6 and HDAC8 (confirmed by low-throughput experiments). I demonstrated that HDAC1 phosphorylation is important for both activity and selectivity. No doubt other PTMs regulate selectivity of HDAC1, as several other PTMs, such as acetylation and sumoylation, have been observed.

In Chapter 3 I explored two phosphorylation sites, Ser421 and Ser423, which are known to be important for activating deacetylase activity[11]. Two additional, less studied, phosphorylation sites, Ser406 and Ser393, are located outside of the deacetylase domain of HDAC1 (**Figure 5.1**). Phosphorylation of Ser393 has shown to increase deacetylase activity[12]. Phosphorylation of Ser406 occurs in early mitotic cells from prophase to metaphase, but the effect of this PTM is not known[13]. All of these phosphorylation sites exist within a 30-residue section of HDAC1 potentially making this region an important regulatory component of HDAC1 activity.

Acetylation has additionally been recently explored on HDAC1. Contrary to phosphorylation, acetylation inhibits deacetylase activity and induces association with the glucocorticoid receptor in response to hormone stimulation[14]. There are numerous potential acetylation sites: K218, K220, K432, K438, K439, and K441[14]. Two of these sites, K218 and K220, are located within the conserved deacetylase domain. The

remaining sites are located not far from the proposed phosphorylation sites, further indicating that this region of HDAC1 is an important for cellular regulation via PTMs.

The final PTM sites identified in HDAC1 are K444 and K476 which are sumoylated. Sumoylation at these sites is known to regulate HDAC1's biological functions, but unknown to what degree[15], and has been found to be important in reducing amyloid plaques[16].

The effects of these different PTMs on activity can be determined by generating the *in vivo* PTMs *in vitro*. Non-canonical amino acid incorporation can be utilized to study the effects of these post-translational modifications. Incorporation of non-canonical amino acids (NCAA) in recombinant proteins has advanced in the last 10 years to a point where high fidelity and yield can be accomplished[17-23]. *E.coli* is still the easiest organism to achieve NCAA incorporation[24] and I have shown in Chapter 3 HDAC1 expressed in *E.coli* can recapitulate the activity of mammalian and insect expressed enzyme. Analogs of phosphoserine and acetylated lysine have been successfully incorporated into expressed proteins[25, 26]. NCAA incorporation cannot be used for sumoylation, as it is a protein, but due to the similarity to ubiquitination, sumoylation has been achieved by *in vitro* assays[27].

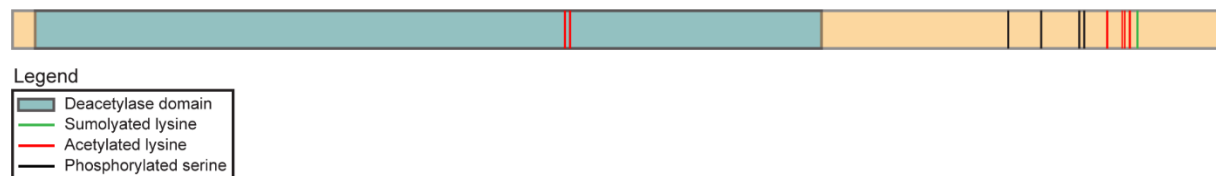


Figure 5.1: The proposed post-translational modifications of HDAC1. The sites are denoted by lines, colored by modifications (see Legend). The conserved deacetylase domain is shown in light blue.

It is curious that regulatory post-translational modifications are concentrated within the C-terminal region of HDAC1 rather than in the conserved deacetylase domain (**Figure 5.1**). Therefore, HDAC1 can be separated into a deacetylase domain and a regulatory domain; other members of the deacetylase family may also follow this division. As I have shown in Chapter 3, phosphorylation of HDAC1 on Ser421 and Ser423 affects the activity and selectivity of HDAC1 but does not clearly stimulate protein-protein interactions. An exploration of the effects of potential PTM sites on activity and complex formation as well as whether these PTMs co-exist or have any synergistic effects are important future questions.

Complex Formation

My work described in Chapter 3 demonstrates that complex formation between HDAC1 and other proteins has an enormous effect on activity. This work only explored interactions with two other proteins, a small fraction of the protein interactions that have been identified with HDAC1 (~600 proteins[28, 29]). Furthermore, this work was on a known stable complex. There have only been 3 such complexes identified for HDAC1, leaving hundreds of other interactions that could have similar effects on activity.

These interactions are typically explored by *in vivo* experimentation or co-purification of the entire complex. My work has illustrated that *in vitro* reconstitution and activity assays can also be effective in evaluating the effects of protein-protein interactions. Using *in vitro* reconstitution additionally allows for more flexibility in building up these interactions to study their individual and combined effects. For HDAC1 the method of building up interactions is particularly useful because the known stable complexes, CoREST, NuRD, and Sin3, have base protein components although there

are many additional interacting proteins that could affect activity and selectivity. For the CoREST complex, LSD1 and CoREST both share an additional 8 interacting proteins with HDAC1 (determined from BioGRID[28, 29]). The interaction networks of LSD1 and CoREST with HDAC1 show even more potential interactors (**Figure 5.2**), which opens numerous questions as to how this complex is regulated within cells.

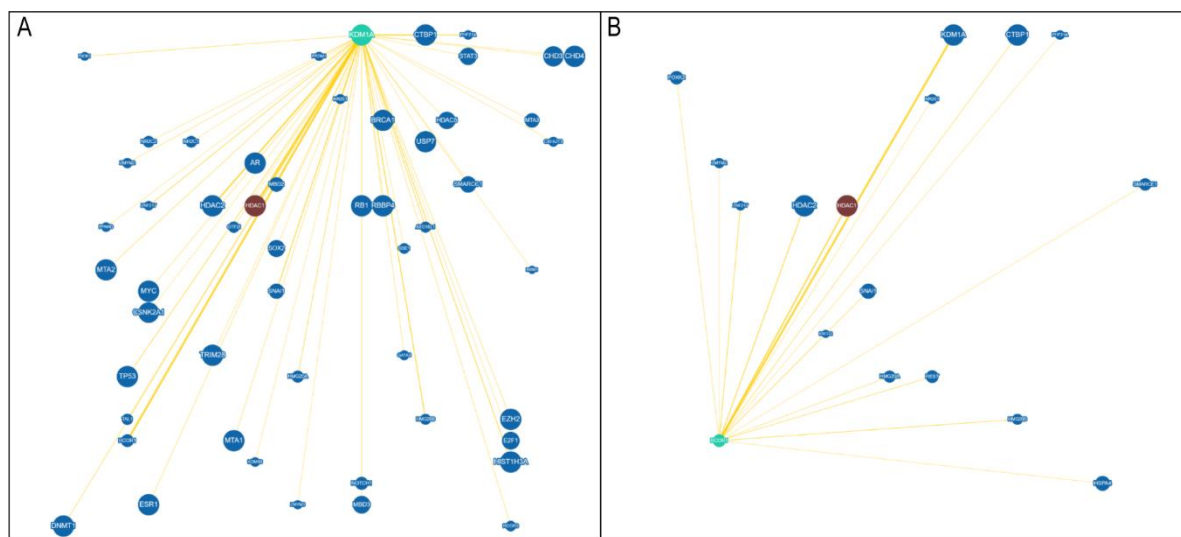


Figure 5.2: Overlapping interactions between HDAC1 and CoREST (RCOR1) or LSD1 (KDMA1). A) The blue circles indicate interactors of both LSD1 (KDMA1, in green) and HDAC1 (in maroon). B) The blue circles indicate interactors of both CoREST (RCOR1, in green) and HDAC1 (in maroon). Obtained from BioGRID[28, 29].

In future studies it would be very interesting to explore potential protein interactors with HDAC6 and HDAC8. Both of these HDACs have more limited identified potential interactions, as indicated by co-immunoprecipitation experiments[30]. For HDAC8 the most concretely identified interactor with HDAC8 is structural maintenance of chromosomes protein 3 (SMC3)[30] which is also a substrate[31-33]. In addition to SMC3, HDAC8 has been connected to other proteins involved in the cell cycle and cell division[30]. Interactions with HDAC8 are also proposed with cytoskeleton associated proteins, primarily those expressed in muscle cells[30]. As with all of these identified

interactions, it is not clear whether they are HDAC8 substrates or if they regulate HDAC8 activity.

HDAC6 uniquely contains a ubiquitin binding domain, so it is no surprise that there are numerous interactions with ubiquitin-containing proteins[30]. HDAC6 has also been shown to interact with dynein to shuttle proteins to aggresomes[34]. It is unknown how or whether the deacetylase activity of HDAC6 is involved in these interactions, though catalytically inactive mutants are not as effective as wild-type enzyme in regulating aggresome formation[34].

Regulatory mechanisms alter selectivity rather than just activity

Commonly regulatory measures either activate or inactive enzymatic activity. Recent studies have demonstrated that regulation of HDACs is more complicated since these regulatory modifications may also impact selectivity, not just activity. The impact on selectivity highlights the need to identify the most likely cellular substrates of these enzymes. The knowledge of isozyme-specific substrates will be crucial for understanding how changes in selectivity due to regulatory modifications will affect the implicated biological pathways.

The Fierke lab had previously collaborated with Prof. Ora Furman-Schueler's group at Hebrew University of Jerusalem to create a very successful model of HDAC8 selectivity[35]. In Chapter 4, I discussed the continued use of structure-based modeling to assist in discovering novel HDAC substrates as we applied this method to a second HDAC, HDAC6. It became apparent that HDAC6, at least the single CD2 domain of HDAC6, does not contain as much substrate selectivity at the sequence level as HDAC8. It would be interesting to develop a structure-based model of a full-length HDAC6

construct to determine if the lack of selectivity is mediated by structural changes induced by the presence of the other catalytic domain. I began exploring these selectivity differences in Chapter 4 and saw that the presence of both domains changes the selectivity of HDAC6. However, with a limited sample size it is difficult to ascertain how selectivity changed. Additional attempts to elucidate the substrate selectivity of HDAC6 reinforce the promiscuity previously seen; when HDAC1, HDAC3, HDAC6, and HDAC8 were screened against a diverse 3-mer coumarin labeled peptide library, it was found that HDAC6 has the broadest selectivity of all the isozymes[36]. A limitation of these studies was that peptide substrate analogs are used. The peptides are unable to mimic any contacts the full structure of the substrate might form outside the active site. These additional contacts were shown to be important with HDAC1 in Chapter 3.

The lack of sequence level selectivity suggests that other factors play a role in determining the substrate selectivity of HDAC6, and potentially other HDAC isozymes. The external signaling factors that could dictate HDAC activity include active site metal identity, post-translational modifications, and protein interactions. Measures to determine how each of these factors affect certain HDACs can be achieved by utilizing the methods of site-directed mutagenesis and NCAA incorporation, apo-enzyme purification with metal reconstitution, and complex co-purification or immunoprecipitation.

The future of HDAC research

Based on the body of work presented in this thesis I believe the future of HDAC research should focus in the following areas: substrate discovery, *in vitro* analysis followed by *in vivo* analysis, and inhibitor development. All these categories in unison

build a complete picture of the biological impacts of HDAC deacetylase activity and targeting this activity to treat and prevent disease.

One of the most important foci of HDAC research should be on substrate discovery. As illustrated within this thesis, we are slowly building an understanding of how substrate selectivity is determined by HDAC but there is still a long way to go before we truly understand what is happening within the cell. There are very few confirmed substrates for any of the HDAC isozymes within the tens of thousands of acetylation sites identified within the human proteome. There have been multiple substrate discovery methods utilized, each with their own positives and limitations. Further refinement of these substrate discovery methods to incorporate important factors that determine selectivity will be essential for a complete isozyme-specific substrate list to be determined.

In vitro analysis of the activity and selectivity of HDAC enzymes can be used to identify the most important selectivity factors. HDACs will need to be screened against peptide substrates with and without potentially important selectivity factors (i.e. post-translational modifications, other interacting proteins). It is additionally necessary that the *in vitro* and *in vivo* analyses work together to test and confirm results to provide the best analysis of substrate selectivity.

Finally, with all the information gained from substrate selectivity and regulatory factors I believe the field will be able to identify inhibitors that lead to better patient outcomes. There are only currently 4 FDA-approved HDAC inhibitors, and the majority of HDAC inhibitors that enter clinical trials fail. In a study of the clinical trials between 2013 and 2017 only 1% reached Phase 4[37]. The failure rate could be due to not understanding enough of the biological function of the HDAC enzymes, so the effect of

the inhibition of the enzymes cannot be well predicted. The continued development of isozyme-specific inhibitors along with non-hydroxamate inhibitors will lead to more targeted and more efficacious, long-lasting therapeutics.

References

1. Dowling, D.P., et al., *Structures of metal-substituted human histone deacetylase 8 provide mechanistic inferences on biological function*. *Biochemistry*, 2010. **49**(24): p. 5048-56.
2. Castaneda, C.A., et al., *Active Site Metal Identity Alters Histone Deacetylase 8 Substrate Selectivity: A Potential Novel Regulatory Mechanism*. *Biochemistry*, 2017. **56**(42): p. 5663-5670.
3. Gantt, S.L., S.G. Gattis, and C.A. Fierke, *Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion*. *Biochemistry*, 2006. **45**(19): p. 6170-8.
4. Bozym, R.A., et al., *Measuring Picomolar Intracellular Exchangeable Zinc in PC-12 Cells Using a Ratiometric Fluorescence Biosensor*. *ACS Chemical Biology*, 2006. **1**(2): p. 103-111.
5. McCranor, B.J., et al., *Quantitative imaging of mitochondrial and cytosolic free zinc levels in an in vitro model of ischemia/reperfusion*. *J Bioenerg Biomembr*, 2012. **44**(2): p. 253-63.
6. Espósito, B.P., et al., *A Review of Fluorescence Methods for Assessing Labile Iron in Cells and Biological Fluids*. *Analytical Biochemistry*, 2002. **304**(1): p. 1-18.
7. Petrat, F., H. de Groot, and U. Rauen, *Subcellular distribution of chelatable iron: a laser scanning microscopic study in isolated hepatocytes and liver endothelial cells*. *The Biochemical journal*, 2001. **356**(Pt 1): p. 61-69.
8. Qian, D.Z., et al., *Class II histone deacetylases are associated with VHL-independent regulation of hypoxia-inducible factor 1 alpha*. *Cancer Res*, 2006. **66**(17): p. 8814-21.
9. Fath, D.M., et al., *Histone deacetylase inhibitors repress the transactivation potential of hypoxia-inducible factors independently of direct acetylation of HIF-alpha*. *J Biol Chem*, 2006. **281**(19): p. 13612-9.
10. Ye, M., et al., *Histone deacetylase 5 promotes the migration and invasion of hepatocellular carcinoma via increasing the transcription of hypoxia-inducible factor-1 α under hypoxia condition*. *Tumour Biol*, 2017. **39**(6): p. 1010428317705034.
11. Pflum, M.K., et al., *Histone deacetylase 1 phosphorylation promotes enzymatic activity and complex formation*. *J Biol Chem*, 2001. **276**(50): p. 47733-41.
12. Feng, B., et al., *Mitogen-activated protein kinase phosphatase 3 (MKP-3)-deficient mice are resistant to diet-induced obesity*. *Diabetes*, 2014. **63**(9): p. 2924-34.

13. Segré, C.V., et al., *A monoclonal antibody specific for prophase phosphorylation of histone deacetylase 1: a readout for early mitotic cells*. *MAbs*, 2016. **8**(1): p. 37-42.
14. Qiu, Y., et al., *HDAC1 Acetylation Is Linked to Progressive Modulation of Steroid Receptor-Induced Gene Transcription*. *Molecular Cell*, 2006. **22**(5): p. 669-679.
15. David, G., M.A. Neptune, and R.A. DePinho, *SUMO-1 Modification of Histone Deacetylase 1 (HDAC1) Modulates Its Biological Activities*. *Journal of Biological Chemistry*, 2002. **277**(26): p. 23658-23663.
16. Tao, C.C., et al., *Epigenetic regulation of HDAC1 SUMOylation as an endogenous neuroprotection against A β toxicity in a mouse model of Alzheimer's disease*. *Cell Death Differ*, 2017. **24**(4): p. 597-614.
17. Neumann-Staubitz, P. and H. Neumann, *The use of unnatural amino acids to study and engineer protein function*. *Curr Opin Struct Biol*, 2016. **38**: p. 119-28.
18. Wan, W., J.M. Tharp, and W.R. Liu, *Pyrrolysyl-tRNA synthetase: an ordinary enzyme but an outstanding genetic code expansion tool*. *Biochim Biophys Acta*, 2014. **1844**(6): p. 1059-70.
19. Wang, L., *Engineering the Genetic Code in Cells and Animals: Biological Considerations and Impacts*. *Acc Chem Res*, 2017. **50**(11): p. 2767-2775.
20. Tuley, A., et al., *The genetic incorporation of thirteen novel non-canonical amino acids*. *Chem Commun (Camb)*, 2014. **50**(20): p. 2673-5.
21. Wang, Y.S., et al., *Genetic incorporation of twelve meta-substituted phenylalanine derivatives using a single pyrrolysyl-tRNA synthetase mutant*. *ACS Chem Biol*, 2013. **8**(2): p. 405-15.
22. Tharp, J.M., et al., *Genetic incorporation of seven ortho-substituted phenylalanine derivatives*. *ACS Chem Biol*, 2014. **9**(4): p. 884-90.
23. Wang, Y.S., et al., *A rationally designed pyrrolysyl-tRNA synthetase mutant with a broad substrate spectrum*. *J Am Chem Soc*, 2012. **134**(6): p. 2950-3.
24. Leisle, L., et al., *Incorporation of Non-Canonical Amino Acids*. *Advances in experimental medicine and biology*, 2015. **869**: p. 119-151.
25. Neumann, H., S.Y. Peak-Chew, and J.W. Chin, *Genetically encoding N(epsilon)-acetylysine in recombinant proteins*. *Nat Chem Biol*, 2008. **4**(4): p. 232-4.
26. Park, H.-S., et al., *Expanding the Genetic Code of Escherichia coli with Phosphoserine*. *Science*, 2011. **333**(6046): p. 1151.

27. Vethantham, V. and J.L. Manley, *In vitro* sumoylation of recombinant proteins and subsequent purification for use in enzymatic assays. Cold Spring Harbor protocols, 2009. **2009**(1): p. pdb.prot5121-pdb.prot5121.
28. Oughtred, R., et al., *The BioGRID interaction database: 2019 update*. Nucleic Acids Res, 2019. **47**(D1): p. D529-d541.
29. Stark, C., et al., *BioGRID: a general repository for interaction datasets*. Nucleic Acids Res, 2006. **34**(Database issue): p. D535-9.
30. Joshi, P., et al., *The functional interactome landscape of the human histone deacetylase family*. Mol Syst Biol, 2013. **9**: p. 672.
31. Deardorff, M.A., et al., *HDAC8 mutations in Cornelia de Lange syndrome affect the cohesin acetylation cycle*. Nature, 2012. **489**(7415): p. 313-7.
32. Decroos, C., et al., *Compromised structure and function of HDAC8 mutants identified in Cornelia de Lange Syndrome spectrum disorders*. ACS Chem Biol, 2014. **9**(9): p. 2157-64.
33. Decroos, C., et al., *Biochemical and structural characterization of HDAC8 mutants associated with Cornelia de Lange syndrome spectrum disorders*. Biochemistry, 2015. **54**(42): p. 6501-13.
34. Kawaguchi, Y., et al., *The Deacetylase HDAC6 Regulates Aggresome Formation and Cell Viability in Response to Misfolded Protein Stress*. Cell, 2004. **115**: p. 727-38.
35. Alam, N., et al., *Structure-Based Identification of HDAC8 Non-histone Substrates*. Structure, 2016. **24**(3): p. 458-68.
36. Riester, D., et al., *Factors affecting the substrate specificity of histone deacetylases*. Biochemical and Biophysical Research Communications, 2007. **357**(2): p. 439-445.
37. Faria Freitas, M., M. Cuendet, and P. Bertrand, *HDAC inhibitors: a 2013-2017 patent survey*. Expert Opin Ther Pat, 2018: p. 1-17.

Appendix

Table A1.1: Top 100 protein acetylome hits from structure-based model of HDAC6

Peptide	Reweightedscore	# proteins	protein	Gene	Protein name	modsite
PCKEVD	-923.56	1	O94916-2	NFAT5	Nuclear factor of activated T-cells 5	K282-ac
STKEVD	-923.014	1	Q9H4B7	TUBB1	Tubulin beta-1 chain	K324-ac
NSKQPA	-917.96	1	P20248	CCNA2	Cyclin-A2	K95-ac
MSKEQF	-915.87	1	Q16236	NFE2L2	Nuclear factor erythroid 2-related factor 2	K487-ac
PGKEEK	-914.77	1	Q08050	FOXM1	Forkhead box protein M1	K440-ac
AGKQLR	-914.17	1	Q8NC51	SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein	K68-ac
FPKEAK	-913.741	1	P00533	EGFR	Epidermal growth factor receptor	K1179-ac
KVKEVL	-913.316	1	Q15831	STK11	Serine/threonine-protein kinase STK11	K64-ac
AGKFKR	-913.02	1	P39748	FEN1	Flap endonuclease 1	K375-ac
ASKESH	-912.884	1	P40763	STAT3	Signal transducer and activator of transcription 3	K49-ac
NDKEAA	-912.30	1	P27695	APEX1	DNA-(apurinic or apyrimidinic site) endonuclease	K35-ac
ISKMND	-911.69	1	Q16666	IFI16	Gamma-interferon-inducible protein 16	K451-ac
KGKGKP	-911.38	1	P26358	DNMT1	DNA (cytosine-5)-methyltransferase 1	K1117-ac
SKKDPE	-911.311	1	P19838	NFKB1	Nuclear factor NF-kappa-B p105 subunit	K441-ac

KNKEQH	-910.88	1	O60934	NBN	Nibrin	K504-ac
YSKQMQ	-910.82	1	Q92769	HDAC2	Histone deacetylase 2	K90-ac
LNKEMV	-910.753	1	Q8NEB9	PIK3C3	Phosphatidylinositol 3-kinase catalytic subunit type 3	K781-ac
SGKVMR	-910.65	1	Q9NUB1	ACSS1	Acetyl-coenzyme A synthetase 2-like, mitochondrial	K642-ac
GPKGIG	-910.59	1	P50461	CSRP3	Cysteine and glycine-rich protein 3	K69-ac
TSKA VS	-910.47	1	P29590	PML	Protein PML	K515-ac
GSKNVD	-910.36	1	O60934	NBN	Nibrin	K208-ac
FGKFER	-910.05	1	O00571	DDX3X	ATP-dependent RNA helicase DDX3X	K118-ac
AEKQPS	-909.33	1	Q09472	EP300	Histone acetyltransferase p300	K970-ac
KGKQAE	-909.30	1	P05114	HMGN1	Non-histone chromosomal protein HMG-14	K61-ac
KKKEFE	-909.28	1	O15392	BIRC5	Baculoviral IAP repeat-containing protein 5	K122-ac
KAKDPT	-909.232	1	Q00613	HSF1	Heat shock factor protein 1	K524-ac
HSKGFG	-909.19	1	Q13148	TARDBP	TAR DNA-binding protein 43	K145-ac
PAKFPS	-909.17	1	P17482	HOXB9	Homeobox protein Hox-B9	K27-ac
GVKSPG	-909.03	1	Q01094	E2F1	Transcription factor E2F1	K120-ac
SGKSAK	-908.95	1	Q13569	TDG	G/T mismatch-specific thymine DNA glycosylase	K87-ac
GSKQNS	-908.80	1	P29084	GTF2E2	Transcription initiation factor IIE subunit beta	K52-ac
KEKEMN	-908.76	1	Q9NR30	DDX21	Nucleolar RNA helicase 2	K137-ac
SGKGKK	-908.68	1	P15976	GATA1	Erythroid transcription factor	K312-ac
MTKDLA	-908.27	1	P48735	IDH2	Isocitrate dehydrogenase [NADP], mitochondrial	K413-ac
PGKGVK	-908.25	1	Q01094	E2F1	Transcription factor E2F1	K117-ac

NSKSVP	-908.18	1	O60563	CCNT1	Cyclin-T1	K380-ac
LNKSPP	-908.15	1	Q01543	FLI1	Friend leukemia integration 1 transcription factor	K240-ac
PGKALV	-907.94	1	P19338	NCL	Nucleolin	K116-ac
PIKEDS	-907.751	1	P00533	EGFR	Epidermal growth factor receptor	K1061-ac
NGKEQL	-907.70	1	Q8NHS0	DNAJB8	DnaJ homolog subfamily B member 8	K223-ac
TAKAVD	-907.675	1	P51692	STAT5B	Signal transducer and activator of transcription 5B	K694-ac
VGKEPS	-907.66	1	O94925	GLS	Glutaminase kidney isoform, mitochondrial	K311-ac
KDKRNQ	-907.54	1	P41235	HNF4A	Hepatocyte nuclear factor 4-alpha	K108-ac
KGKGKG	-907.30	3		COL	Collagen	
PSKLDS	-907.145	1	P07900	HSP90AA1	Heat shock protein HSP 90-alpha	K69-ac
SGKAKT	-907.13	1	P0C0S5	H2AZ1	Histone H2A.Z	K11-ac
SGKGNP	-907.10	1	P35222	CTNNB1	Catenin beta-1	K49-ac
ENKCPV	-907.077	1	Q09472	EP300	Histone acetyltransferase p300	K1800-ac
YGKLPP	-907.05	1	Q9Y2D1	ATF5	Cyclic AMP-dependent transcription factor ATF-5	K29-ac
MKKEMI	-907.03	1	O75469	NR1I2	Nuclear receptor subfamily 1 group I member 2	K109-ac
NVKAKI	-906.88	1	P62988	UBB	Polyubiquitin-B [Cleaved into: Ubiquitin]	K27-ac
NTKNHD	-906.867	1	Q09472	EP300	Histone acetyltransferase p300	K1699-ac
TSKNKS	-906.847	1	Q09472	EP300	Histone acetyltransferase p300	K1558-ac
GAKFPI	-906.799	1	P12931	SRC	Proto-oncogene tyrosine-protein kinase Src	K426-ac

DGKDVM	-906.76	1	P06744	GPI	Glucose-6-phosphate isomerase	K116-ac
LNKEKK	-906.75	1	P14316	IRF2	Interferon regulatory factor 2	K29-ac
ADKDYS	-906.56	1	P07195	LDHB	L-lactate dehydrogenase B chain	K82-ac
NGKGRP	-906.51	1	P15923	TCF3	Transcription factor E2-alpha	K34-ac
TRKEME	-906.47	1	P42858	HTT	Huntingtin	K343-ac
AEKQRP	-906.42	1	Q96EP5	DAZAP1	DAZ-associated protein 1	K150-ac
INKEQF	-906.35	1	Q09472	EP300	Histone acetyltransferase p300	K1228-ac
ALKAPS	-906.23	1	O60566	BUB1B	Mitotic checkpoint serine/threonine-protein kinase BUB1 beta	K250-ac
SAKGFG	-906.10	1	Q99814	EPAS1	Endothelial PAS domain-containing protein 1	K685-ac
FTKDKH	-906.04	1	Q16236	NFE2L2	Nuclear factor erythroid 2-related factor 2	K443-ac
VTKGDS	-906.00	1	Q09472	EP300	Histone acetyltransferase p300	K1542-ac
CDKEYV	-905.94	1	O43623	SNAI2	Zinc finger protein SNAI2	K166-ac
GDKRNQ	-905.94	1	Q09472	EP300	Histone acetyltransferase p300	K423-ac
FSKEAS	-905.86	1	Q9P0J1	PDP1	[Pyruvate dehydrogenase [acetyl-transferring]]-phosphatase 1, mitochondrial	K202-ac
RSKEIT	-905.84	1	P17844	DDX5	Probable ATP-dependent RNA helicase DDX5	K80-ac
KPKAPL	-905.84	1	O95600	KLF8	Krueppel-like factor 8	K95-ac
GGKAGK	-905.83	1	P0C0S5	H2AZ1	Histone H2A.Z	K4-ac
GTKAVT	-905.78	1	P33778	H2BC3	Histone H2B type 1-B	K116-ac
LVKEFF	-905.74	1	P11021	HSPA5	Endoplasmic reticulum chaperone BiP	K376-ac
SGKAPR	-905.72	2		FOXO	Forkhead box protein	

SDKTVE	-905.67	1	Q09472	EP300	Histone acetyltransferase p300	K1331-ac
AAKDKS	-905.67	1	P05114	HMG1	Non-histone chromosomal protein HMG-14	K42-ac
MGKGDP	-905.66	2		HGB	High mobility group protein	
FGKIIR	-905.60	1	P49773	HINT1	Histidine triad nucleotide-binding protein 1	K21-ac
AGKFGP	-905.59	1	Q8N884	CGAS	Cyclic GMP-AMP synthase	K50-ac
QAKSPP	-905.565	1	Q06413	MEF2C	Myocyte-specific enhancer factor 2C	K239-ac
MGKGVV	-905.40	1	P06733	ENO1	Alpha-enolase	K60-ac
AGKGGY	-905.37	1	Q13547	HDAC1	Histone deacetylase 1	K218-ac
NGKSYP	-905.33	1	Q13330	MTA1	Metastasis-associated protein MTA1	K626-ac
SAKELP	-905.21	1	Q09472	EP300	Histone acetyltransferase p300	K1499-ac
DGKHVV	-905.10	1	P62937	PPIA	Peptidyl-prolyl isomerase A cis-trans	K125-ac
MSKQEL	-905.00	1	O94992	HEXIM1	Protein HEXIM1	K284-ac
KGKYYA	-904.97	1	Q13547	HDAC1	Histone deacetylase 1	K220-ac
NAKQHK	-904.96	1	O60934	NBN	Nibrin	K233-ac
VKKEIQ	-904.88	1	Q15831	STK11	Serine/threonine-protein kinase STK11	K97-ac
KSKGQV	-904.85	1	P09874	PARP1	Poly [ADP-ribose] polymerase 1	K508-ac
AGKQLE	-904.84	1	P62988	UBB	Polyubiquitin-B [Cleaved into: Ubiquitin]	K48-ac
GGKARA	-904.77	2		H	Histone	
PRKEPV	-904.68	1	P26358	DNMT1	DNA	K961-ac
LDKAVS	-904.63	1	Q09472	EP300	Histone acetyltransferase p300	K1473-ac
PAKAVT	-904.63	1	P19338	NCL	Nucleolin	K102-ac

NGKDSK	-904.60	1	P06748	NPM1	Nucleophosmin	K212-ac
ANSFVG	-904.539	1	P36507	MAP2K2	Dual specificity mitogen-activated protein kinase 2	S226-ac
PSKSAP	-904.45	1	P33778	H2BC3	Histone H2B type 1-B	K5-ac
PSKSVL	-904.33	1	Q08050	FOXM1	Forkhead box protein M1	K614-ac
LNKNMQ	-904.327	1	Q06413	MEF2C	Myocyte-specific enhancer factor 2C	K234-ac